Stochastic Characterization, Microstructure Reconstruction and Transport Property Prediction for Porous Media

Swansea University
Prifysgol Abertawe

Jinlong Fu

College of Engineering
Swansea University

Submitted to Swansea University in fulfilment of the requirements for the Degree of
Doctor of Philosophy

January 2021

To my beloved family...

"Learning without thought is labor lost; thought without learning is perilous."
—— **Confucius**

"Men love to wonder, and that is the seed of our science."
—— **Emerson**

"All models are wrong, but some are useful."
—— **George Box**

"To ordain conscience for heaven and earth; to secure life and fortune for the people; to continue lost teachings for past sages; to establish peace for all future generations."
—— **Zai Zhang**

Declaration

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed .....[redacted]..... (candidate)

Date .....14/01/2021.....

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated.
Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed .....[redacted]..... (candidate)

Date .....14/01/2021.....

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access approved by the Swansea University**.

Signed .....[redacted]..... (candidate)

Date .....14/01/2021.....

Publications

The following publications are based on the research presented in this thesis:

(1) **Jinlong Fu**, Jiabin Dong, Yongliang Wang, Yang Ju, D. Roger J. Owen and Chenfeng Li. Resolution effect: An error correction model for intrinsic permeability of porous media estimated from lattice Boltzmann method. *Transport in Porous Media*, 2020, 132:627–656.

(2) **Jinlong Fu**, Shaoqing Cui, Song Cen and Chenfeng Li. Statistical characterization and reconstruction of heterogeneous microstructures using deep neural network. *Computer Methods in Applied Mechanics and Engineering*, 2020, 373:113516.

(3) **Jinlong Fu**, Hywel R. Thomas and Chenfeng Li. Tortuosity of porous media: Image analysis and physical simulation. *Earth-Science Reviews*, 2020, 103439.

(4) **Jinlong Fu**, Hywel R. Thomas and Chenfeng Li. Permeability prediction for natural porous rocks through feature selection and machine learning. *International Journal of Engineering Science*. (Under Review)

(5) **Jinlong Fu**, Shaoqing Cui, Song Cen and Chenfeng Li. Stochastic reconstruction of 3D pore microstructures from 2D cross-sectional images via supervised machine learning. *Computers & Structures*. (Under Review)

(6) **Jinlong Fu**, Hywel R. Thomas and Chenfeng Li. Statistical characterization and analysis of pore microstructures: Morphology and topology. *Earth-Science Reviews*. (Under Review)

(7) Shaoqing Cui, **Jinlong Fu**, Song Cen, Hywel R. Thomas and Chenfeng Li. The correlation between statistical descriptors of heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering*. (Under Review)

(8) **Jinlong Fu**, Hywel R. Thomas and Chenfeng Li. Determination of the optimal relaxation parameter in lattice Boltzmann model to reduce uncertainty in permeability estimation for complex porous media. (In Preparation)

Acknowledgements

Thanks to angles in white who are fighting the coronavirus COVID-19, this world still has a paradise side and life is beautiful like summer flowers as before. Without such a protective umbrella, this thesis would never have been possible.

Foremost, I would like to express my deep and sincere gratitude to my supervisors: Prof. Chenfeng Li, Prof. D. Roger J. Owen, and Prof. Hywel R. Thomas, for their continuous support to my PhD study. Over the past four years, Prof. Chenfeng Li has created a free, nurturing and rigorous academic environment for me, in which I can shape myself to be a qualified researcher under his invaluable guidance. He always kept a high degree of scientific enthusiasm to motivate me to overcome research difficulties, and constructive suggestions always came just in time due to his profound knowledge in my research area. Besides, Prof. Owen and Prof. Thomas provided me insightful comments and encouraged me to explore challenging problems, which kept me confident throughout my PhD study.

In addition to my supervisors, I would like to thank Prof. Quentin Fisher (Leeds University), Dr. Bjornar Sandnes, Prof. Yang Ju (China University of Mining and Technology), Prof. Yuntian Feng, Prof. Zuo Zhuang (Tsinghua University), Prof. Xianghua Xie, Dr. Ramin Bostanabad (Northwestern University) and Dr. Minglu Wu (China University Of Petroleum), for their helpful advice and discussions.

Next, I am very grateful to my colleagues, also friends at Zienkiewicz Center for Computational Engineering. I acquired many scientific research skills from Dr. Dunhui Xiao, Dr. Min Luo, Dr. Yongliang Wang and Dr. Yan Shang, which greatly improves my research quality. Besides, Shaoqing Cui, Bin Chen, Jinsheng Wang, Beatriz Ramos Barboza, Ashutosh Bhokare, Si Li, Sizeng You, Jiabin Dong, Yanan Sun, Zhi Li, Tongming Qu, Han Wang, Shan Zhong and others are my fellow travelers on the research journey, and we strove forward with each other's company, which helped me to keep passion and interest in my research.

Special thanks go to Zienkiewicz PhD scholarship (Swansea University), China Scholarship Council and Great Britain China Center, for their financial supports. And I also greatly appreciate Imperial College London, University of Stuttgart, University College London, and Digital Rocks Portal, for their publicly shared micro-CT data of various porous rocks.

Words can hardly express my gratitude to my dear parents, beloved sisters and brothers, beautiful wife and lovely daughter, who always make me feel happy even when I was frustrated by research difficulties, and they are the driving force to motivate me to set sail again all the time.

Abstract

Porous media are ubiquitous in the natural environment and engineering, and typical examples include rocks, soils, and concretes. Their transport properties (i.e. permeability, effective diffusivity, formation resistivity factor and thermal conductivity) usually exhibit strong uncertainty, due to the intricacy, stochasticity and heterogeneity inherent in pore microstructures. Although high-definition visualizations of opaque porous media can be derived from advanced microscopy imaging techniques (such as micro-CT), it remains a critical challenge to effectively link microstructural characteristics to transport properties. This thesis is devoted to improving the understanding of structure-property relationships of porous media, thereby predicting macroscopic transport properties from observable microstructure informatics. The research objective has been achieved through three interrelated modules: stochastic characterization, microstructure reconstruction and predictive model construction.

To bridge the gaps between microstructural characteristics and transport properties, quantitative characterization of porous media in explicit expression is the essential prerequisite, through which the stochastic morphology of complicated microstructures is reduced to a small set of descriptors related to macroscopic transport properties. A comprehensive review of statistical characterization of pore microstructures is presented, where a wide variety of morphological descriptors are collected through an extensive literature survey, to provide microstructural informatics from the global, local, geometrical and topological perspectives. Tortuosity of porous media is one of the key parameters to model transport properties, and it is systematically examined from the viewpoints of concept and evaluation method in this study. The correlations between geometrical and physical tortuosities are further analyzed, based on which phenomenological relations between them are established.

With the limited availability of digital microstructures, the inherent stochasticity of porous media can rarely be captured by using a small number of samples. The complete computational dataset should cover the representative samples with all possible configurations. Stochastic reconstruction of 3D pore microstructure is an effective way to provide large numbers of samples of arbitrary size for analyzing transport properties of porous media. A novel method is presented to statistically characterize and reconstruct heterogeneous microstructures through a deep neural network model, which can generate 3D pore microstructure samples by well preserving statistical equivalence, long-distance connectivity and transport properties. Besides, another new approach is developed to stochastically reconstruct 3D pore microstructures from 2D cross-sectional images through supervised machine learning, which can rapidly produce more realistic and accurate 3D microstructure samples compared with other three classical approaches.

The digital microstructure of high-quality provides a high-fidelity framework for pore-scale simulations of fluid flows, permitting one to evaluate transport properties or explore specific physical phenomena. However, the results obtained from low-resolution images of pore microstructures are often compromised with significant errors, known as the resolution effect. The resolution effect on permeability evaluation from the lattice Boltzmann method is quantitatively investigated, and an error correction model is constructed to reduce/eliminate this resolution effect by identifying the primary error causes. The model uses correlated morphological descriptors to quantify the resolution effect and achieve error correction.

What's more, the dependence of permeability on microstructural characteristics of porous media is fundamentally studied through feature selection and machine learning. The morphological descriptors significant to permeability are highlighted and selected through the performance-driven feature selection. In essence, the selected morphological descriptors provide a deep and interpretable insight into the underlying microstructure-permeability linkage. The machine learning-based permeability model is thus built by using the optimal subset of morphological descriptors as the feature data, and the prediction model exhibits excellent performances in predictive accuracy and general applicability.

Table of contents

List of figures

List of tables

Chapter 1

Introduction

1.1 Background and motivation

Porous media are ubiquitous in nature, science and engineering, typical examples include rocks, soils, concretes, cement, porous electrodes, heterogeneous catalysts and bone tissues, which consist of solid skeletal matrix and pore space, allowing fluid to pass through or surround. It is well known that the pore structure fundamentally determines the macroscopic transport properties of a porous medium [2]. In geoscience, the transport behaviors of fluid permeation, mass diffusion, current transmission, and heat conduction through porous rocks are highly dependent on the pore geometry and connectivity, which determine the corresponding petrophysical properties: permeability [34], effective diffusivity [352], formation resistivity factor [37], and thermal conductivity [357] respectively.

Structure-property linkages of porous media are embodied in many geologic applications. For radioactive waste management [264, 299, 398], the internal specific surface area of the back-fill clay (such as bentonite) is decisive to its sorption affinities for organic and inorganic ions, which plays a vital role in preventing radionuclide to migrate from a deep geological repository to the surface environment over a long period. As to geological carbon dioxide sequestration [337, 192, 40], CO_2 is injected into the underground storage reservoir and isolated from the atmosphere by the geological barrier with low permeability, where pore microstructures of rocks in the reservoir formation and the geological barrier are strongly correlated to the storage capacity and the sealing performance respectively. In the case of groundwater contaminant transport [31, 170], the microstructural characteristics of aquifer materials (such as permeable or fractured rocks) determine a variety of transport mechanisms including permeation, diffusion and dispersion.

In addition to transport properties, the macroscopic elastic properties of porous rocks and concrete are also acutely sensitive to pore structures [256, 15]. The interaction between

the cementitious material and environment is controlled by the pore structure as well, which leads to structural degradation and determines the material durability [27, 360, 292]. The electrode microstructure has a significant influence on the performance of solid oxide fuel cell in terms of energy storage and conversion [456, 24]. For the heterogeneous catalyst, the pore architecture is directly linked to the reaction efficiency [361, 245]. Besides, the processes of cell migration, nutrient diffusion, waste removal, and bone remodeling are greatly affected by the structure of bone tissue [138].

Currently, the relationship between pore microstructures and macroscopic physical properties is one of the most elementary questions in numerous porous media research disciplines (e.g., geoscience, material science, civil engineering, energy storage, chemical engineering, and bioengineering), which is still without being conclusive [445, 460, 125, 195, 438, 18]. The structure-property linkage is of great value to model/predict macroscopic properties of a given porous medium just from its measurable microstructural features, without conducting expensive experiments or and complex numerical simulations. To bridge the gaps between the microstructural characteristics and the effective macroscopic properties, direct measurement on the pore space and quantitative characterization of geometric complexity in an explicit expression are the essential prerequisites [12].

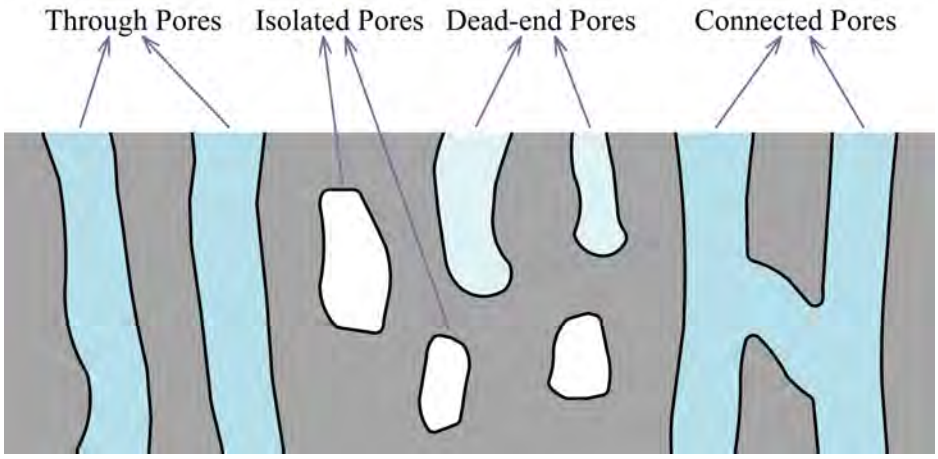


Fig. 1.1 Schematic representation of typical pores in porous media

However, it is difficult to completely measure and characterize the intricacy of pore microstructures due to various reasons: (1) Only a few microstructural properties, such as porosity, pore surface area and pore size distribution, can be measured from laboratory experiments [486, 259];(2) The internal structures of the opaque porous media are often at the microscale, even nanoscale, and thus advanced microscopy imaging techniques are the only option for noninvasive and nondestructive measurements [474, 12]; (3) Porous media usually exhibit structure on hierarchical scales, which means the structure element itself has a

structure at lower scale [492]; (4) Structural elements of porous media, irrespective the scale, show a huge variety of complex geometry in terms of shape and size, and exhibit strong randomness in the spatial distribution of different phases [460]. In general, a porous medium involves a large number of connected, isolated, dead-end, geometry-irregular, randomly distributed and hierarchical pores, constituting an anfractuous pore microstructure that is extremely complicated to achieve a completely quantitative representation (as illustrated in Figure 1.1).

1.2 Pore microstructure measurement techniques

Various measuring techniques have been applied to acquire the geometry and topology information of pore structures [12], which can be broadly grouped into two categories: experimental measurement and microscopy imaging. Although the scopes of pore size covered by these measurement techniques are overlapped, they are still complementary in general, as illustrated in Figure 1.2.

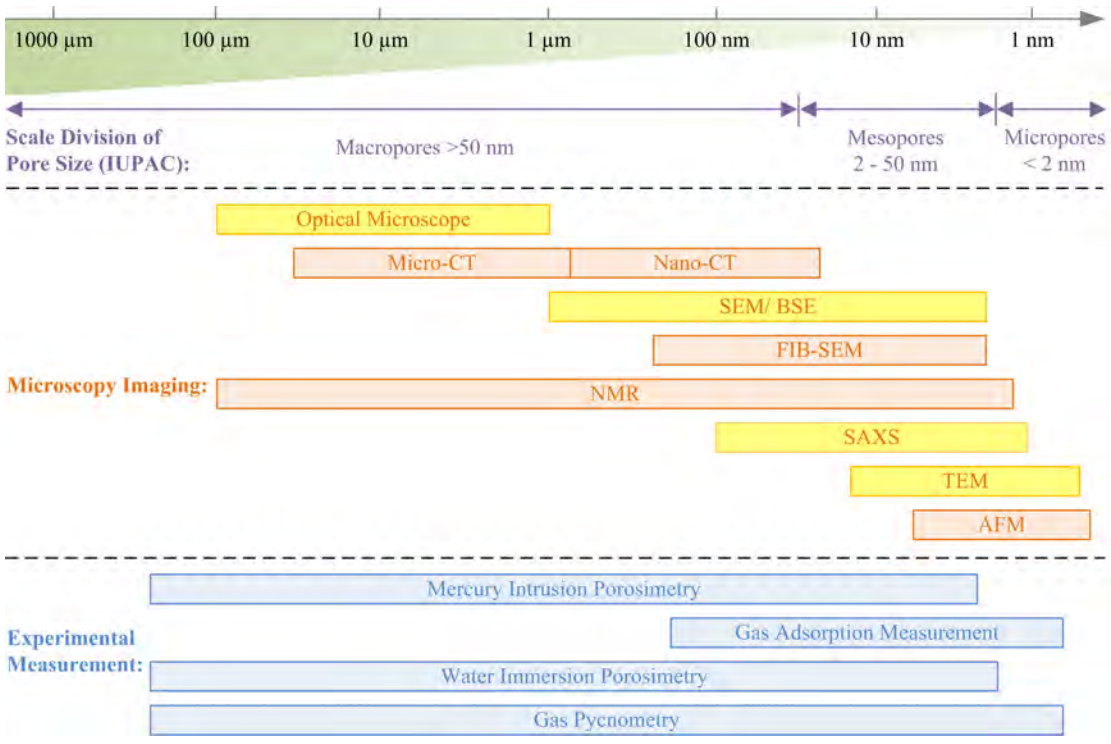


Fig. 1.2 Experimental measurement approaches and microscopy imaging techniques (2D and 3D imaging techniques are denoted by texts with yellow and pink backgrounds respectively) to characterize pore microstructures

1.2.1 Experimental measurement approaches

There are several experimental measurement approaches to characterizing microstructures of porous media in non-destructive manners, such as gas adsorption measurement, water immersion porosimetry (WIP), and gas pycnometer. From these experimental measurements, information on effective pore space can be obtained, including porosity, pore size distribution (PSD) and specific surface area, but not the full connectivity of pore space [486].

Mercury intrusion porosimetry (MIP) [155, 145] could be the most popular experimental measurement to investigate pore structures in an indirect manner, due to its simple physio-chemical principle and fast operation process, and it is capable of covering the pore size ranging from about 3.5 nm to 500 μm. However, the determination of PSD from MIP is based on the assumption of cylindrical pores, which is untrue for porous media with irregular pore geometry. In practice, MIP systematically misallocates almost all pore sizes of complicated porous media, and it tends to underestimate the actual pore sizes, because of the "throat" and "ink bottle" effects.

Gas adsorption [89, 88] is widely used to quantify the pore microstructures of tight sandstones and shales, and the measurable pore diameter range is from 0.3 to 300 nm. During the testing process, the gas adsorption volume is measured together with the relative equilibrium adsorption pressure, and isotherm interpretation is then required to achieve microstructural characteristics. The classical Brunauer-Emmett-Teller (BET) method is popular to evaluate specific surface area, but it is not applicable for pore media with narrow pores. Pore volume and PSD can be derived from the well-known Barrett-Joyner-Halenda (BJH) method, but the BJH method simply describes a porous medium as a collection of cylindrical or slit pores.

WIP [254, 202] is a standard porosity measurement, where the testing sample is saturated with a liquid with known density and then the pore volume can be determined from the weight difference between the completely saturated and dry states. The bulk volume of the testing sample can be determined from Archimedes' Principle by immersing it in the liquid. Thus, the porosity can be calculated as the ratio of the pore volume to the bulk volume. Water is an ideal choice of the saturating and immersing fluid, because it is able to penetrate small pores and capillaries.

Gas pycnometry [432, 12], also referred to as helium pycnometry, is to measure the volume of solid particles for low permeable porous rocks. Combined with the measurement of bulk or pore volume, the effective porosity of a rock sample can be estimated. It is based on Boyle-Mariotte's law that describes the volume-pressure relationship of an experimental gas. Helium (He) is usually used as the experimental gas, due to its ideal properties. There are at least three types of gas pycnometry reported in literature [432]: constant-volume,

variable-volume and comparative, and constant-volume gas pycnometry is the most popular one.

1.2.2 Advanced microscopy imaging techniques

How pores are internally connected is of primary interest, because connectivity and topology of pore networks are crucial to fluid flow and mass transport behaviors. However, conventional experimental measurements can only provide very limited information about internal pore structures, and the interpretation of experimental data is not straightforward, which requires careful modeling.

Recently, modern microscopy imaging techniques [12, 259] have been able to digitize the geometry of a porous medium, usually called a microstructure, into a 2D or 3D image at different resolution levels and various length scales. The digital microstructure can be used for a variety of simulation or diagnostic purposes, which radically boosts microstructural analyses. Scanning electron microscopy (SEM), backscattered electron (BSE) imaging, atomic force microscopy (AFM), optical microscopy and transmission electron microscopy (TEM) are commonly used techniques to provide 2D non-invasive visualizations of microstructures. Besides, 3D digital microstructures can be acquired through X-ray micro-computed tomography (micro-CT or μCT), Nuclear magnetic resonance (NMR) imaging, and focused ion beam scanning electron microscopy (FIB-SEM). Although the imaging mechanisms of various microscopy techniques are different, the acquired digital images at the same resolution level are comparable [328]. As shown in Figure 1.3, 1.4 and 1.5, the 2D visualizations of microstructures inside reservoir rocks are obtained from modern microscopy imaging techniques. Various types of pores, such as inter-granular pores, intra-granular pores and fractures, constitute the complicated pore systems inside a sandstone, a carbonate rock and a shale.

Optical microscope [308] uses visible light and a system of lenses to create magnified visual of small objects, and the magnified visual can be captured by a light-sensitive camera to produce a 2D photomicrograph, from which the pore system and the solid minerals on the surface of a porous rock sample can be evaluated. The pore system can be highlighted by impregnating epoxy to the sample for better identification [12]. The resolution range of an optical microscope is essentially limited by the length of the visible light wave, which is usually lower than the half-wavelength (0.4-0.7 μm). Optical petrology is a simple and piratical way to investigate pore structures, but the real pore system (including inter-granular pores) in 3D space can not be completely reflected by the surface information.

The micro-CT scanner [238, 388] is able to obtain internal microstructures of porous rocks in a non-destructive and non-invasive manner. An X-ray beam is sent out from the

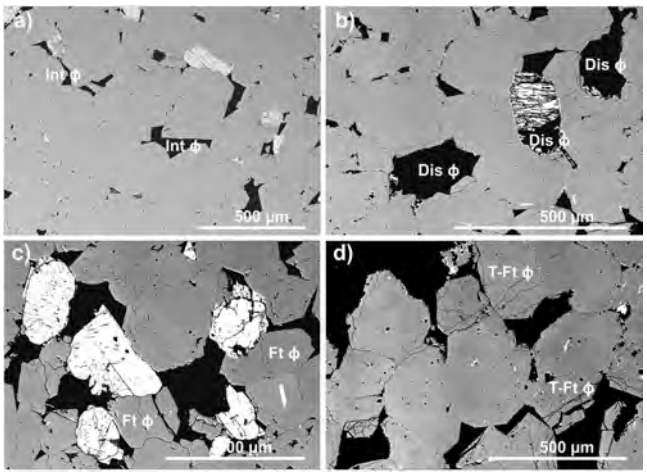


Fig. 1.3 SEM-BSE images showing different types of pores inside Clashach sandstone, where pore space is shown in dark: (a) intergranular pores, (b) secondary dissolution pores, (c) fractures, (d) transgranular fractures, and micropores on each image (redraw after [133])

radioactive source, and it is attenuated when passing through materials, due to scattering and absorption. The attenuated X-ray beam is detected by a detector, and then the attenuation degree can be converted to a projection image according to Beer-Lambert law. A continuous series of cross-sectional projection images are taken from different angles by rotating the sample 360°, from which the 3D model is reconstructed to provide an internal microstructure. The typical spatial resolution of micro-CT is from about 1 to 50 μm, and rock cores are normally constrained to a few millimeters to acquire representative images at the micro-scale. Currently, the nano-CT facility is able to provide scanning images with the spatial resolution better than 50 nm, allowing researchers to observe smaller pores at the sub-micro scale. Besides, SAXS [499] is capable of quantifying the nano-scale density difference in a rock sample by analyzing the elastic scattering behavior, to provide structural information of dimensions ranging from 1 to 100 nm.

SEM [423, 397] is a popular imaging technique to acquire 2D visualizations of pore microstructures. It generates images by scanning the sample surface with a focused beam of electrons, where the electrons interact with the atoms in the sample to produce various signals, including secondary electrons and back-scattered electrons (BSE). In SEM mode, secondary electrons emitted by atoms are detected, and the signal intensity yields topography information. The common SEM instrument can achieve resolution around 5-20 nm. BSE can also be detected to form images, but it usually produces lower resolution than secondary electrons. In addition, SEM can also be used in transmission mode, where the focused beam of electrons is transmitted through a thin specimen (less than 100 nm) to form an image. TEM is capable of returning resolution at the atomic level, which has been used to characterize

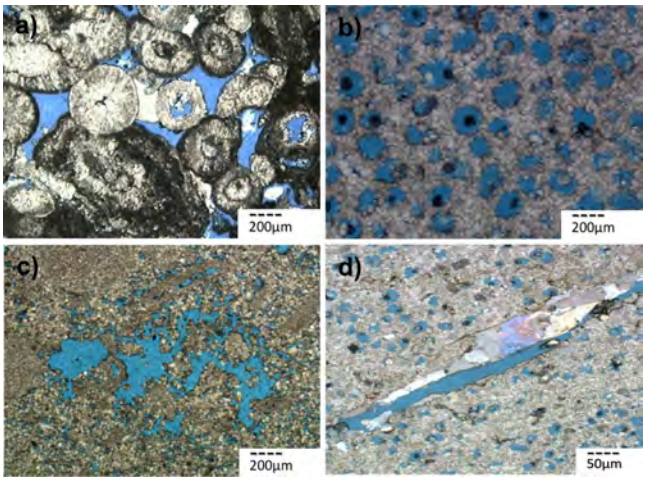


Fig. 1.4 Casting thin section images showing different types of pores inside Carboniferous carbonate rock, where pore space is shown in blue: (a) intergranular dissolution pores, (b) intragranular dissolution pores, (c) intercrystalline dissolution pores, and (d) dissolution fractures (redraw after [400])

the metal-Zeolite catalyst systems [248]. However, the connectivity of a 3D pore space cannot be accurately learned from the surface information provided by 2D SEM images. The combination of focused ion beams (FIB) and SEM (FIB-SEM) provides a new route to acquire the 3D pore structure with very high-resolution [323, 48], where ion beams make very fine slices through etching the sample, enabling sequential SEM images to be obtained. The 3D model can be built by stacking 2D sequential images together. The FIB-SEM is a destructive technique, but it can typically achieve voxel dimensions of tens of nanometers.

NMR [259] is another non-destructive technique to quantify pore microstructure, which is able to cover almost the full range of pores inside reservoir rocks. NMR extracts pore structure information from the relaxation process of hydrogen nuclei in the pore fluids, where the spin axis relaxation times of protons are measured in the presence of static and pulsed magnetic field [12, 486]. There are two types of relaxation times in NMR: the longitudinal and transverse. The pore size distribution of a porous media sample is commonly extracted from the transverse relaxation time, because the transverse relaxation time can be obtained much faster and contains sufficient information on pore size. Typically, NMR is suitable for quantifying pore diameter ranging from 2 nm to 1 μm, depending on the absorbate.

AFM [212] is the latest scanning probe microscope techniques, which is capable of characterizing pore features down to the atomic scale. AFM has been applied to quantify pore structures and gas flow behaviors inside the ultra-tight shales [212, 107]. In AFM, a probe at the tip of a cantilever is used to scan over the sample surface, and interaction (force) between the probe and the sample is measured simultaneously. The local attractive and

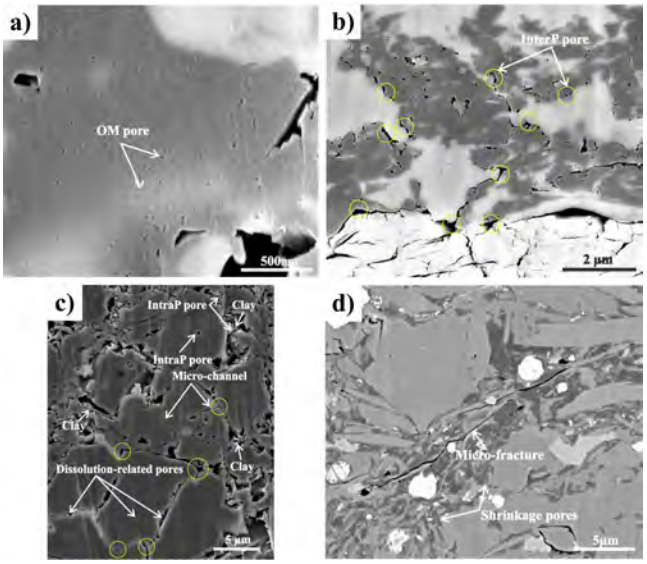


Fig. 1.5 SEM images showing different types of pores inside shale, where pore space is shown in dark: (a) organic matter pores, (b) intergranular pores, (c) intragranular pores and micro-channel, and (d) micro-fracture and shrinkage pores (redraw after [500])

repulsive force between the probe and sample surface bend the cantilever beam, and the vertical and lateral deflections of the cantilever are measured through an optical lever. The optical lever operates by reflecting a laser beam off the back of the cantilever, in order to amplify the deflection to be a measurable signal that can be captured by the photo-detector. Generally, the lateral resolution that AFM can achieve is relatively low (about 30 nm), but the vertical resolution can reach up to 0.1 nm.

Moreover, chemical characterization and composition/crystallographic data can be obtained accompanying with microstructural data, by combining energy-dispersive X-ray spectroscopy (EDXS) [331], secondary-ion mass spectrometry (SIMS) [84] or electron backscatter diffraction (EBSD) [203] with the above mentioned microscopy imaging technologies. The internal variations of microstructures responding to the changing ambient environment can also be tracked, allowing 4D analysis of microstructural evolution in the time scale [67, 380]. The noninvasive microscope imaging techniques leave the physical samples intact and unaltered, allowing complimentary analysis such as experimental measurements performed on the samples.

1.3 Digital rock physics

Benefiting from the tremendous advances in microscopy imaging techniques, digital rock physics (DRP) [11, 48] has evolved into a new discipline to investigates and calculates the

transport/physical properties of porous rocks in a non-destructive manner, as illustrated in Figure 1.6. DRP uses microscopy imaging techniques such as micro-CT to obtain digital rock images with varying resolutions and sizes. The representations of complex pore microstructure can be obtained by separating the pore space from mineral matrix phases through the segmentation processing on raw images, and then physical simulations can be performed on the pore-scale microstructures to estimate the effective properties of rocks (e.g. permeability, elastic moduli, effective diffusivity, thermal conductivity and formation resistivity factor).

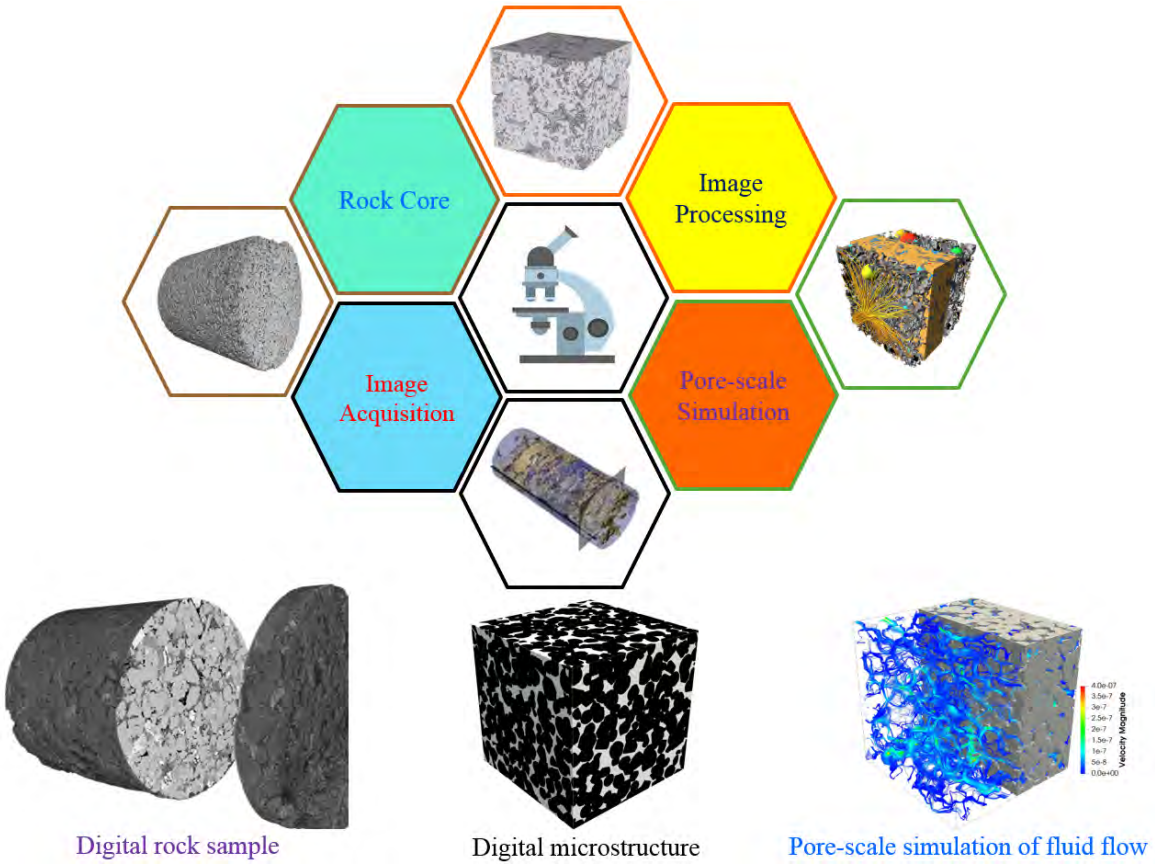


Fig. 1.6 Schematic illustration of the digital rock physics technique to study transport properties of porous rocks

1.3.1 Image processing and segmentation

After the image acquisition, the raw scanning image should be processed to remove artifacts and noise, and then image segmentation is conducted to separate the pore space from the solid matrix for pore space analysis [388]. Due to the wide use of lab-based micro-CT scanners, we take the micro-CT image as the example to illustrate image processing and segmentation.

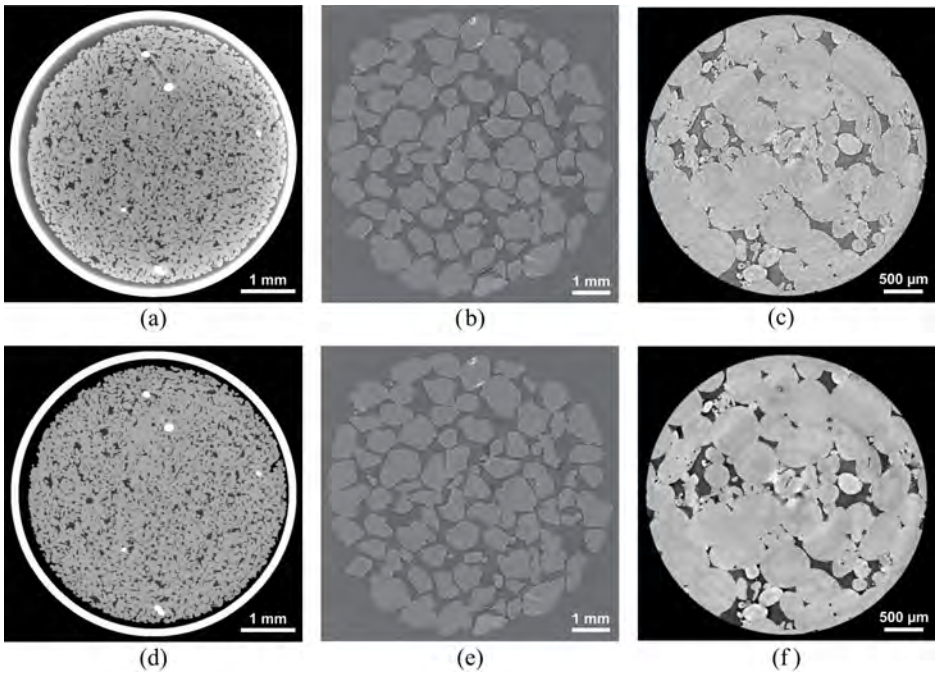


Fig. 1.7 Illustration of image processing to remove artifacts and noise in micro-CT images: (a) A 2D slice of a simulated silicon sample with beam hardening artifacts (after [193]); (b) A cross-section of a sand pack sample (uniformly graded (0.60–1.18 mm) Leighton Buzzard sand) with ring artifact (after [81]); (c) A cross-sectional slice of a Ketton limestone sample with noise (after [412]); (d) Correction of the beam hardening artifacts; (e) Alleviation of the ring artifact; (f) Removal of image noise

Artifacts and noise removal

There are two types of commonly encountered scanning artifacts [238] in micro-CT images: beam hardening and ring artifacts, which have dramatic effects on intensity-based image segmentation. When an X-ray beam passes through a scanned object, its lower energy components are preferentially absorbed by the object surface, leading to an increase in the average energy of the X-ray beam. This effect is referred to as beam hardening, and it can cause cupping artifact (brighter object edges compared to the image center) and streaking artifact (dark shadows surrounding the bright areas in the image), as shown in Figure 1.7a. Various remedies can be used to correct beam hardening artifacts, such as scanning preparation, outer covering filter and data processing, and more details can be found in relevant references [238, 193]. By comparing Figure 1.7a and d, one can observe that the beam hardening artifacts are effectively corrected.

As to the ring artifact [388], it appears as full or partial circles in the center of the micro-CT scanning image, as shown in Figure 1.7b. The ring artifact is caused by the shifts in output from the detector array of a micro-CT scanner, leading to anomalous values of voxel

intensity in the photograph [238]. Due to its linkage to beam hardening, the ring artifact can be eliminated at the scanning stage by using the same methods, such as scanning preparation and outer covering filter. Besides, the ring artifact is more amenable to software methods. For example, it can be alleviated by renormalizing the reconstructed image at the sinogram stage through Fourier filtering [322].

Noise is almost inevitable during the image acquisition process, which brings great difficulty to image segmentation [206]. Noise can be significantly reduced from raw scanning images through various filtering methods, so as to enhance the contrast between different phases, as illustrated in Figure 1.7c and f. The mean, Gaussian, median, shock, anisotropic diffusion, total variation and nonlocal means filters are frequently used to smooth raw scanning images, and more details can be found in relevant references [226, 388].

Image segmentation

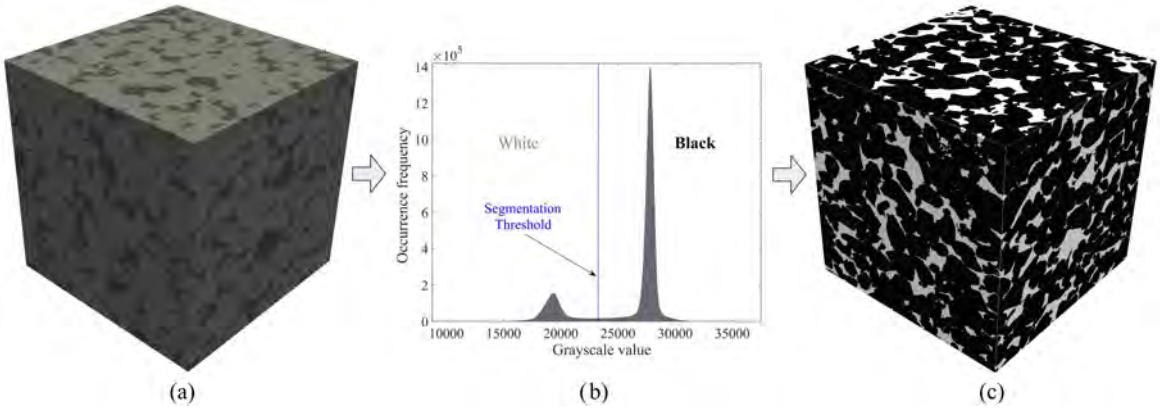


Fig. 1.8 Illustration of image segmentation for a Gildehauser sandstone sample via global thresholding: (a) The 3D micro-CT image in 16-bit grayscale mode, where the pore space is shown in dark; (b) The histogram of the grayscale values, based on which the segmentation threshold is determined; (c) The 3D segmented image with two distinct phases, where white region corresponds to the pore space and the solid matrix is shown in black

The raw micro-CT image is in a grayscale or color mode with continuous pixel/voxel intensities, as shown in Figure 1.8a and b. Therefore, it is important to convert the raw scanning image from the continuum model to a discrete form that permits quantitative characterization of the porous microstructure and subsequent pore-scale modeling of fluid flow [206]. An image segmentation process is commonly implemented to separate the pore space from the solid matrix. A broad range of image segmentation methods have been developed and applied in porous media research, such as global thresholding, locally adaptive thresholding, region growing methods, deformable surface methods, probabilistic fuzzy

clustering, Bayesian methods, hybrid method and manual segmentation, and more details about image segmentation can be found in related references [206, 388, 268].

After image segmentation, the raw scanning image is discretized into a binary form with two distinct phases: pore space and solid matrix, as shown in Figure 1.8c. The pore space P is represented as the solid matrix body M embedded in Euclidean space Ω, and a particular configuration of a pore structure X can be described by using the characteristic (or indicator) function as follows [186]:

$$X(\mathbf{r}) = \begin{cases} 1 & \text{for } \mathbf{r} \in P \\ 0 & \text{for } \mathbf{r} \in M \end{cases} \tag{1.1}$$

where $\mathbf{r} \in \Omega$ is the position vector of the pixel/voxel on the digitized grid, and $X(\mathbf{r})$ is the corresponding pixel/voxel intensity value.

1.3.2 Pore-scale modeling of fluid flow

The digital microstructure of high-quality provides a high fidelity framework for numerical simulations of different physical processes at the pore scale, permitting one to evaluate effective physical properties or explore specific physical phenomena, from which a profound understanding of the impacts of the pore microstructure on physical processes can be gained directly [369, 11, 48, 427]. Pore-scale simulations allow for systematic variations of pore geometries, fluid properties, operating conditions and boundary conditions to evaluate their influences, which is almost impossible to achieve with physical experiments [305, 486]. For the effective transport properties of porous media, pore-network modeling (PNM) and direct numerical simulation (DNS) are the two main pore-scale modeling approaches to mimicking transport processes [48], from which permeability, effective diffusivity, electrical and thermal conductivity can be evaluated according to Darcy's, Fick's, Ohm's and Fourier's law, receptively.

Pore network modeling

PNM [486] simplifies the complicated pore space into a topologically representative network of pore bodies interconnected by pore throats with ideal shapes (such as sphere and cylinder), where the size and connectivity of each network element are determined from the geometry and topology of the pore microstructure according to specific criteria. The transport behaviors within each network element are described by semi-analytical laws (such as Hagen-Poiseuille law), which makes the computation considerably fast and enables researches to incorporate strong heterogeneity in modeling large volumes [48]. PNM is naturally appropriate for

capillary-controlled transport processes. However, the transport properties obtained from PNM could be less credible, due to the simplification of the complicated pore space. It is still a great challenge to obtain reliable results from PNM, by identifying the features and phenomena relevant to network modeling and ignoring the remaining parts to simplify the computational complexity.

Direct numerical simulation

On the contrary, DNS directly discretizes the pore space into computing elements by preserving pore geometry (voxels can be used as the computing elements), and transport equations (such as the Navier-Stokes or Laplace equations) are numerically solved or approximated on the computational meshes [305, 48]. Lattice Boltzmann method (LBM), finite element method (FEM) and finite volume method (FVM) are commonly used to approximate or solve the transport equations at the pore scale. Generally, DNS can provide direct insight into the impact of pore microstructure on transport properties.

However, this direct approach has severe limitations: (1) Intensive computation. The 3D digital microstructure with large representative size and high resolution usually contains millions, even billions of computational elements (or voxels), therefore, massive parallel programming, long computing time, high-performance computation platform and large data storage are usually required to run numerical simulations (such as the lattice Boltzmann simulation of multiphase flow) [283, 384]. (2) Lack of representativeness. The results obtained from physical simulations performed on single or several digital microstructure sample(s) with limited sizes may not represent the real property of the porous medium, because properties of natural porous media can be heterogeneous and show strong randomness at all length scales [445, 103]. (3) Multi-scale problem. As mentioned above, a pore medium usually involves pores with various characteristic length scales, and these pores can not be completely captured by a single image, because all microscopy imaging techniques have a limited ratio between image size and image resolution. It is possible to built digital models covering heterogeneous properties at several length scales, but these models can be too detailed to solve the fluid flow equations via standard discretization methods [132, 283].

1.4 Research objective and layout of the thesis

1.4.1 Research objective

As discussed above, modern microscopy imaging techniques can provide digital microstructures of porous media, and effective transport properties can be estimated by performing

pore-scale simulations on them. However, the crucial point is how to deeply understand the dependence of transport properties on the microstructural characteristics of porous media.

The objective of this thesis is to gain deep insight into the microstructure-property linkages of porous media, thereby to construct reliable predictive models of transport properties. Such microstructure-property models are of great value, based on which transport properties can be directly estimated from the observable microstructure informatics, avoiding high-cost and time-consuming experiments or computation-intensive simulations.

The research objective is achieved through three interrelated modules, including stochastic characterization, microstructure reconstruction and predictive model construction:

- **Stochastic characterization** of pore microstructures in explicit expression is the essential prerequisite to model transport properties, where the microstructural complexity of porous media is reduced to a small set of morphological descriptors related to transport properties.

- **Microstructure reconstruction** is an effective and economic way to statistically generate large numbers of samples for transport property modeling, so as to capture the inherent stochasticity of porous media. Because the complete computational dataset should cover the representative samples with all possible configurations.

- **Predictive model construction** is not only to establish reliable prediction models to directly evaluate transport properties from relevant morphological descriptors, but also to fundamentally study and interpret the underlying microstructure-property relationships of porous media.

1.4.2 Layout of the thesis

The research works corresponding to the three interrelated modules are distributed in eight chapters of this thesis, as illustrated in Figure 1.9.

In **Chapter 2**, a comprehensive review of statistical characterization of pore microstructures is presented, where a broad variety of morphological descriptors are collected through an extensive literature survey. And the applications of morphological descriptors to characterize pore microstructures and model transport properties are systematically summarized as well. Basically, these morphological descriptors quantify pore microstructures from different perspectives, from which suitable subsets of descriptors can be selected to model transport properties or explore specific problems. This chapter lays a solid foundation for all subsequent studies in this thesis, including stochastic microstructure reconstruction, error

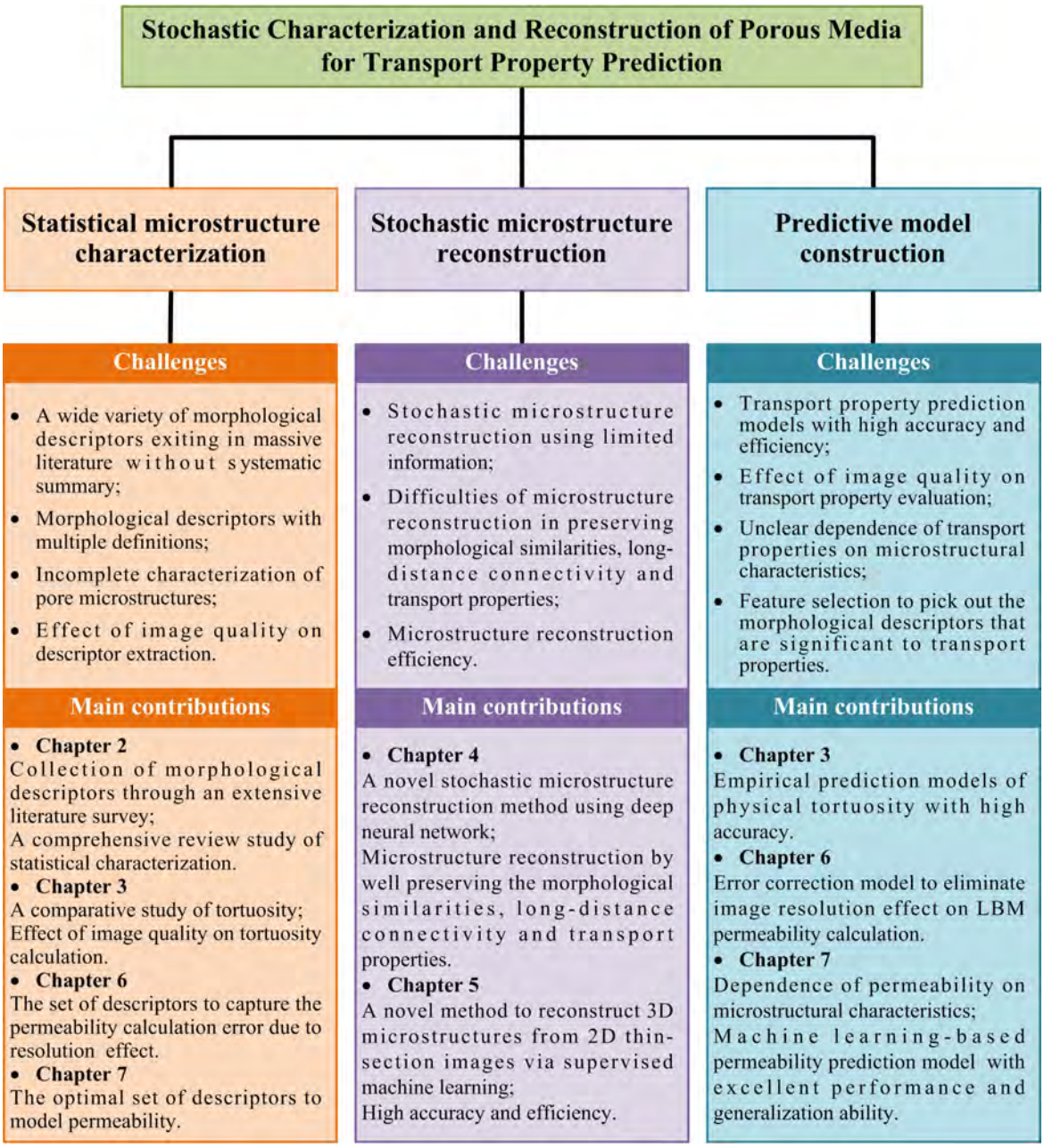


Fig. 1.9 The structure arrangement of the thesis

correction modeling due to the resolution effect, and the fundamental investigation of the microstructure-permeability relationship.

In **Chapter 3**, tortuosity of porous media, as a key parameter to model transport properties, is systematically examined from the viewpoints of definition and evaluation method. Various types of tortuosities are broadly divided into two groups: geometrical and physical tortuosity, depending on whether the analysis object is the pore space itself or the transport flow through

the pore space. Phenomenological relations between geometrical and physical tortuosities are established to predict physical tortuosities from geometrical ones, avoiding computation-intensive simulations. Tortuosity plays an important role in subsequent studies of this thesis, including error correction modeling due to the resolution effect, and establishment of the microstructure-permeability linkage.

In **Chapter 4**, a novel method is presented to statistically characterize and reconstruct random microstructures via deep neural networks, from which large numbers of pore microstructure samples can be rapidly generated by preserving statistical equivalence, morphological similarities, long-distance connectivity and transport properties. The reconstructed microstructure samples in this chapter are used to further explore the relationships between microstructural characteristics and transport properties.

In **Chapter 5**, a novel method is developed to statistically reconstruct 3D porous microstructures from 2D cross-sectional images using supervised machine learning. The proposed method is able to reconstruct both 3D isotropic and anisotropic microstructures, and it exhibits great advantages over the commonly used reconstruction methods, in terms of accuracy and efficiency.

In **Chapter 6**, the resolution effect on permeability evaluated from LBM is quantitatively investigated through identifying the primary causes of error, based on which an error correction model is proposed. The model uses such microstructure characteristics as connected porosity, specific surface area and diffusion tortuosity to quantify the resolution effect and achieve error correction. The error correction model makes image resolution reduction more meaningful and creditable for LBM permeability evaluation of porous media, thereby supporting its adoption in practical applications. The study in this chapter provides credible permeability results for the subsequent study to deeply investigate the microstructure-permeability linkage.

In **Chapter 7**, the dependence of permeability on microstructural characteristics of natural porous rocks is fundamentally investigated through feature selection and machine learning. The optimal subset of morphological descriptors is picked out from the feature pool containing dozens of descriptors through the performance-driven feature selection. The feature selection result provides a deep and interpretable insight into the microstructure-permeability relationship. The machine learning-based permeability prediction model, corresponding to the optimal subset morphological descriptors, possesses an excellent performance in predicting permeability values spanning 4 orders of magnitude.

Finally, **Chapter 8** summarizes the main attributions and innovations of this thesis and points out directions for future research.

Chapter 2

Statistical characterization of porous media: Morphology and topology

2.1 Introduction

Quantitative description of porous media is an essential prerequisite to build microstructure-property relationships or to reveal physical phenomena related to microstructural characteristics. The pore space inside natural porous rock usually exhibits great disorder and strong randomness, which should be quantitatively characterized in statistical terms to involve the spatial distribution of microstructural features. Statistical characterization [441, 186, 321, 460] of the pore microstructure is an attractive way to model/predict transport or physical properties, where the microstructural complexity is reduced to a limited number of meaningful descriptive parameters related to the properties of interest.

Massive efforts have been made to analyze pore microstructures, and a broad range of statistical characterization approaches have been developed in the literature, including geometrical measurements [92, 186, 12], topological quantification [90, 460, 18], pore network interpretation [28] and others. However, there are several crucial problems in microstructure characterization through image analysis, which have not been properly solved yet, and they have significant impacts on the development of reliable structure-property relationships: (1) Ambiguous definitions. Some morphological descriptors are not strictly defined, even with multiple definitions, whose concepts are strongly coupled with specific operations that could cause inconsistent results. (2) Incomplete characterization. There is no such a basic set of computable/measurable parameters that are capable of completely characterizing the microstructures of natural porous media. Usually, physical properties depend on unknown sets of microstructural variables and parameters, and the internal correlations between mor-

phological descriptors are also not fully understand. (3) Image analysis errors. Microstructure information extracted from digitized images is subject to a number of errors, which could come from image binarization and preparation, finite image size and boundary effect, finite image resolution and discretization effect, as well as key parameter selection related to image analysis algorithms. (4) Improper usage. Some microstructure information can be extracted both from 2D and 3D images, and equivalent transformation between them can be made through stereology (stereology provides meaningful quantitative descriptions of the geometry of 3D structures from measurements that are made on 2D images). Accurate descriptions of pore connectivity and topology always require 3D microstructures, but the relevant measures, such as tortuosity, constriction factor and pore/throat size distribution, are often extracted from 2D slices by researchers, leading to mistakes in microstructure characterization and transport properties modeling.

Considering the above problems, it is essential to systematically investigate microstructural characteristics, so as to lay a solid foundation to explore how pore microstructures affect the macroscopic properties of porous media by integrating the important microstructural informatics. A comprehensive review of statistical characterization of pore microstructures is thus conducted, where the theoretical bases and extraction methods of a wide variety of morphological descriptors are involved, aiming to provide a profound understanding of limitations and opportunities in microstructure characterization and analysis.

2.2 Microstructure characterization via image analysis

Broad diversity of morphological descriptors have been proposed to characterize porous media in the literature. The commonly used descriptors are collected through an extensive literature study, and they are summarized and grouped in Table 2.1. To achieve a better understanding of microstructure characterization, the morphological descriptors are applied to a group of Fontainebleau sandstone samples with a wide porosity range. Fontainebleau sandstone [371, 7] only possesses inter-granular porosity covering a range from 3.0% to 30.0% without noticeable grain-size variation, which makes it the most commonly used natural porous rock to investigate the structure-property correlations independently of other influencing factors [122, 92, 34, 371, 7, 142]. In this study, seven 3D digital microstructure samples of Fontainebleau sandstone are used [35], whose porosities are 8.61%, 10.15%, 12.57%, 15.36%, 17.62%, 20.64% and 24.50%, respectively. The image sizes of these digital samples are of $480 \times 480 \times 480$ voxels with a voxel resolution of 5.7 μm, representing equivalent physical samples of 2.736 mm^3 cube, as shown in Figure 2.1.

Table 2.1 Summary of morphological descriptors to characterize pore microstructures

Morphological descriptors	Porous media	Imaging techniques	Image dimension	Applications	Representative references
Porosity (Pore volume fraction)	Rocks, soils, concretes, and others.	Micro-CT, FIB-SEM, SEM and others.	2D and 3D	Microstructure characterization; Modeling/prediction of transport/physical properties; Stochastic microstructure reconstruction.	[185], [91], [125], [2], [47], [178], and [12].
Specific surface area	Rocks, soils, concretes, and others.	Micro-CT, FIB-SEM, SEM and others.	2D and 3D	Microstructure characterization; Modeling/prediction of transport/physical properties; Stochastic microstructure reconstruction.	[50], [246], [185], [221], [2], [255], [367] and [12].
Euler characteristic (Pore connectivity function)	Soils and sintered glass.	Serial-section imaging and Micro-CT	3D	Quantification of pore connectivity; Measurement of pore size; Prediction of hydraulic properties; Interpretation of capillary trapping; Measurement of fluid topology.	[457], [459], [458], [17], [181], [286] and [387].
Minkowski functionals	Sandstones, limestones and Boolean models.	Micro-CT	2D and 3D	Microstructure characterization; Stochastic microstructure reconstruction; Prediction of physical properties; Digital image segmentation.	[306], [262], [394], [315] and [388].
Minkowski functions	Soils, sand, silica and Boolean models.	Micro-CT and SEM	3D	Microstructure characterization; Prediction of permeability; Measurement of pore size; Linking capillary trapping.	[310],[460], [390], [267] and [150].
Integral of mean curvature	Glass packing beads.	Micro-CT	3D	Estimation of capillary pressure.	[273], [304] and [19].
Two-point correlation function	Sandstones, glass packing beads and other heterogeneous materials.	Micro-CT, SEM and BSE	2D and 3D	Prediction of permeability; Microstructure characterization; Stochastic microstructure reconstruction; Evaluation of effective diffusivity; Description of fluid velocity field.	[92], [46], [441], [215],[262], [487], [191], [218], and [52].
Normalized autocovariance function	Sandstones and dolomite rocks.	BSE, SEM and Micro-CT	2D and 3D	Microstructure characterization; Prediction of permeability; Stochastic microstructure reconstruction.	[493], [33], [208], [277], [442], [217], and [262].
Two-point cluster correlation function	Sandstone and other heterogeneous materials.	Micro-CT	3D	Microstructure characterization; Stochastic microstructure reconstruction.	[444], [441], [213], [77] and [52].

Morphological descriptors	Porous media	Imaging techniques	Image dimension	Applications	Representative references
Pair connectivity function	Sandstones, soils, Boolean models, and other heterogeneous materials.	SEM and Micro-CT.	2D and 3D	Microstructure characterization; Evaluation of pore connectivity; Stochastic microstructure reconstruction.	[9], [471], [242], [13] and [389].
Field-field correlation function	Porous rocks, and composite materials.	Micro-CT and SEM	2D and 3D	Microstructural characterization; Stochastic microstructure reconstruction.	[374], [373], [213], [406] and [14].
Surface correlation functions	Concretes, disk packing, sphere packing, and other heterogeneous materials.	– –	2D and 3D	Estimation of permeability; Microstructure characterization; Stochastic microstructure reconstruction.	[378], [441], [216], [379]and [296].
Spherical contact distribution function	Boolean models, nanocomposites, solid oxide fuel cell, and paper.	Micro-CT, FIB-SEM and TEM	2D and 3D	Microstructure characterization; Measurement of pore size.	[319], [267], [287], [144] and [297].
Lineal path function	Sandstones and heterogeneous materials.	Micro-CT and SEM-BSE	2D and 3D	Microstructure characterization; Stochastic microstructure reconstruction; Prediction of permeability.	[291], [92], [105], [441], [70] and [175].
Chord length distribution function	Sandstones and heterogeneous media, fuel cell materials and composite.	Micro-CT and SEM	2D and 3D	Microstructural characterization; Stochastic microstructure reconstruction; Prediction of permeability; Estimation of pore size.	[446], [441], [267], [92], [182], [74] and [416].
Local porosity distribution function	Sandstones, sedimentary rocks and mudstone.	Micro-CT and BSE	2D and 3D	Microstructure characterization; Stochastic microstructure reconstruction; Prediction of permeability; Estimation of effective electrical conductivity.	[183], [184]; [44], [186], [341], [262] and [101].
Coarseness	Sandstones, fiber material and other heterogeneous materials.	Micro-CT and SEM	2D and 3D	Microstructural characterization.	[290], [92], [364] and [408].

Morphological descriptors	Porous media	Imaging techniques	Image dimension	Applications	Representative references
Local percolation probabilities	Sandstones, sedimentary rocks, cement paste and mudstone.	Micro-CT and BSE	2D and 3D	Microstructure characterization; Stochastic microstructure reconstruction; Prediction of permeability; Estimation of effective electrical conductivity.	[187], [45], [262], [200], [341], [389] and [101].
Total fraction of percolating cells	Sandstones, carbonate rocks and mudstone.	Micro-CT and BSE	2D and 3D	Microstructure characterization; Stochastic microstructure reconstruction; Prediction of permeability; Estimation of effective electrical conductivity.	[187], [44], [45], [341], [262], [101] and [171].
Pore size distribution	Sandstones, soils, limestones, clay, cement, concretes, and battery electrodes.	SEM-FIB, Micro-CT, and SEM	3D	Microstructure characterization; Prediction of permeability; Prediction of durability; Prediction of effective diffusivity.	[340], [115], [281], [292], [321], [257], [267], [237], [145], [431], [121] and [174].
Pore throat size distribution	Sandstones, soils, limestones, clay, cement, concretes, and battery electrodes.	SEM-FIB, Micro-CT, and SEM	3D	Microstructure characterization; Prediction of permeability; Prediction of durability; Prediction of effective diffusivity.	[146], [281], [407], [321], [239], [195], [121] and [174].
Geometrical tortuosity (Tortuosity factor)	Sandstones, soils, limestones, clay, cement, concretes, and battery electrodes.	SEM-FIB, Micro-CT, and SEM	2D and 3D	Microstructure characterization; Prediction of permeability; Prediction of durability; Prediction of effective diffusivity; Prediction of durability.	[325], [280], [159], [428], [223], [124], [75], [422], [96], [353] and [142]
Geometrical tortuosity distribution	Sandstones, battery electrodes, and fuel cell.	SEM-FIB, and Micro-CT	3D	Microstructure characterization; Prediction of permeability; Prediction of durability; Prediction of effective diffusivity.	[435], [280], [223], [75] and [80].
Constriction factor	Sandstones, porous ceramic, clay, fule cell electrodes, and collagen fibril.	Micro-CT, and FIB	2D and 3D	Microstructural characterization; Prediction of permeability; Prediction of conductivity; Prediction of effective diffusivity.	[195], [472], [34], [236], [422], [330], [329] and [41].

Morphological descriptors	Porous media	Imaging techniques	Image dimension	Applications	Representative references
Pore coordination number	Sandstones, sedimentary rocks, fuel cell and silica monolith.	Optical microscope, Micro-CT and SEM	2D and 3D	Microstructure characterization; Prediction of permeability; Estimation of effective electrical conductivity.	[122], [281], [32], [32], [435], [197], [365] and [7].
Fractal characterization	Sandstones, carbonate rocks, biomaterials, Sierpinski carpet and coal.	Micro-CT, FBI-SEM, and SEM	2D and 3D	Microstructure characterization; Prediction of permeability; Prediction of effective diffusivity; Stochastic reconstruction of microstructures.	[1], [496], [350], [179], [327], [113], [224], [12], [229] and [485].

Fig. 2.1 The 3D digital microstructures (binary images) of a group of Fontainebleau sandstone samples with different porosities ϕ (pore space is shown in white, and solid matrix is shown in black)

2.2.1 Minkowski functionals

The idea of using Minkowski functionals to characterize structural complexity emerged from the field of statistical physics, where it was evident that numerous physical processes are linked with the geometry and spatial pattern of stochastic structures [391, 307]. Minkowski functionals show the potential to achieve geometric characterization of pore structures and fluid phase distribution inside porous media.

As early as 1997, Vogel started to use Euler characteristic as an index of connectivity of soil structures, where Euler characteristic was quantified as a function of minimum pore diameter, called connectivity function [457]. The pore size distribution and connectivity function developed from Minkowski functionals were used to predict the hydraulic properties of soils [459]. Latter, Vogel (2002) [458] used Euler characteristic as a function of porosity threshold to characterize the topology (connectivity) of porous media. Mecke and Arns (2005) [306] used Minkowski functionals as the morphological measure to quantitatively characterize porous media with complex structure, and also to predict the transport and thermodynamic properties in porous media. The effects of Minkowski functionals on the permeability of porous media were analyzed by Lehmann et al. (2008) [267], and they found permeability is much more sensitive to volume and surface than to curvature and connectivity. Vogel et al. (2010) [460] further developed the Minkowski quantities to be functions of pore size or aggregate size through morphological filter processing, in order to characterize the soil structures, which have the potential to be linked to physical properties. Scholz et al. (2012) [394] found that the permeability of 2D porous structures can be determined from Euler characteristic. Recently, Minkowski functionals were also used to reconstruct statistically equivalent microstructure samples for porous media [262, 389, 315].

In addition to microstructure characterization and reconstruction, Minkowski functionals also have been applied to quantitatively represent the morphological characteristics of two-phase flow in different researches, to understand two-phase fluid distribution [267], capillary pressure [19], pore-scale displacement mechanisms [387], flow regimes during immiscible displacement [17] and relative permeability [286].

Basic Minkowski densities

Minkowski functionals are basic morphological measures defined for binary spatial structures [460]. Given a n-dimensional binary structure, a set of $n+1$ Minkowski functionals can be obtained through integral geometry, where area, perimeter and Euler characteristic are for 2D structures, and volume, surface, integrated mean curvature (related to mean breath) and integrated total curvature or Euler characteristic are for 3D structures [307, 267, 18].

The theoretical basis of Minkowski functionals necessary for structural characterization is briefly introduced here, and more in-depth analyses can be found in related references [418, 393, 18].

After image segmentation, the digital microstructure of a porous medium is represented as a binary structure. Take a 3D digital microstructure X as the case, the first functional M_0 is total volume of pore space:

$$M_0(X) = \int_{\Omega} X(\mathbf{r})\, \mathrm{d}v \tag{2.1}$$

where dv is the volume element. The second functional M_1 measures the total area of the pore-solid interface ∇X through integral geometry:

$$M_1(X) = \int_{\nabla X} \mathrm{d}s \tag{2.2}$$

where ds is the surface element (as illustrated in Figure 2.2). The third functional M_2 is the integral of mean curvature over the pore-solid interface ∇X:

$$M_2(X) = \frac{1}{2} \int_{\nabla X} \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \mathrm{d}s \tag{2.3}$$

where R_1 and R_2 are the principal radii of curvature of surface element ds (as illustrated in Figure 2.2). Generally, the radius of convex curvature is positive, and it is negative for concave curvature. The fourth functional M_3 is the integral of Gaussian curvature over the pore-solid interface ∇X:

$$M_3(X) = \int_{\nabla X} \left(\frac{1}{R_1 R_2} \right) \mathrm{d}s = 4\pi\, \chi(X) \tag{2.4}$$

where $\chi(X)$ is the Euler characteristic of pore space. The measurement units of M_0, M_1 and M_2 are $[L^3]$, $[L^2]$ and $[L]$, respectively, while M_3 is dimensionless.

To eliminate the effect of sample size, the four Minkowski functionals are normalized with respect to the bulk volume $V(X)$ of the porous media, yielding four Minkowski densities m_k [460]:

$$m_k(X) = \frac{M_k(X)}{V(X)} \tag{2.5}$$

Obviously, m_0 is the absolute porosity ϕ of porous media, defined as the fraction of pore space over the bulk volume. It can be simply calculated from the ratio of pore voxel number to the total voxel number via image analysis [136, 454]. Porosity [204] is one of the most fundamental properties of porous media such as reservoir rocks, which is a measure of fluid storage capability determining the amount of valuable resource fluid that

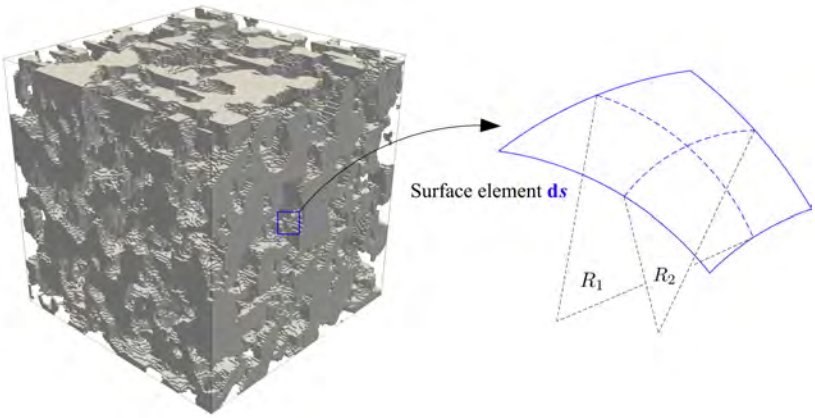
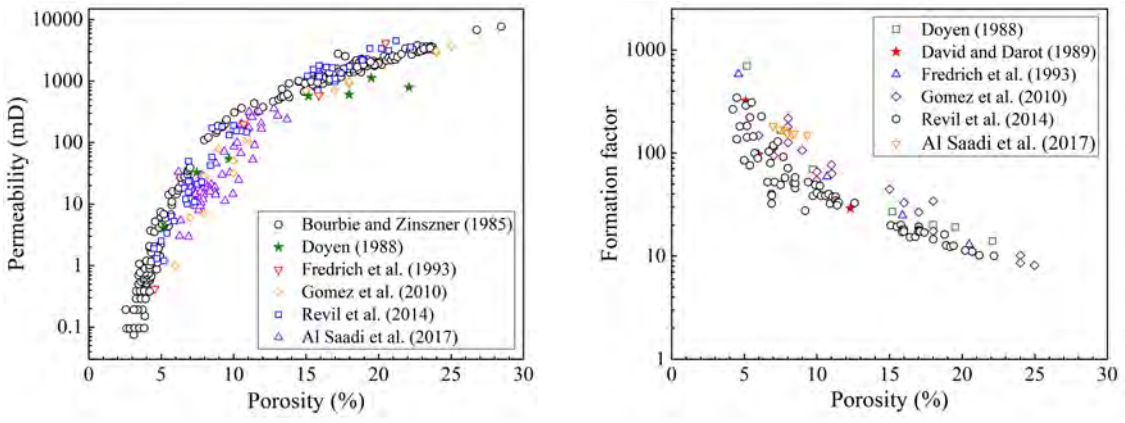


Fig. 2.2 Illustration of a surface element ds on the pore-solid interface ∇X of a 3D pore microstructure X and locally defined radii R_1 and R_2 of curvature

may have accumulated. Connected porosity ϕ_c, defined as the percentage of interconnected pore space with respect to the bulk volume, is commonly used to model/predict transport properties [5, 454]. As shown in Figure 2.3, transport properties of Fontainebleau sandstone are functions of porosity. Besides, effective porosity ϕ_e is also proposed though excluding the dead-end pores or determining the effective pore space according to fluid streamlines [246, 34].



(a) (b)

Fig. 2.3 Transport properties of Fontainebleau sandstone as functions of porosity: (a) Intrinsic permeability (experimental data is collected from [56], [122], [140], [158] and [7]); (b) Formation factor (experimental data is collected from [122], [109], [140], [158], [371] and [7])

The second Minkowski density m_1 corresponds to specific surface area S, which is defined as the total interface area between pore space and solid matrix per unit bulk volume of porous media [2]. Specific surface area is another critical geometric characteristic to understand

transport behaviors inside porous media [50, 246, 221, 367]. In addition, mean curvature m_2 is related to pore shape, and total curvature m_3 characterizes the topology of pore space, which is an important index of connectivity. An isolated pore provides a positive Euler characteristic, so a negative value of total curvature with a large magnitude usually means good connectivity. The Minkowski functionals of 2D/3D pore structures can be determined by considering local binary patterns, and more details about the computation can be found in [333] and [266]. Here, we calculate the Minkowski densities for the group of Fontainebleau sandstone samples in Figure 2.1 through image analysis, and the results are shown in Figure 2.4.

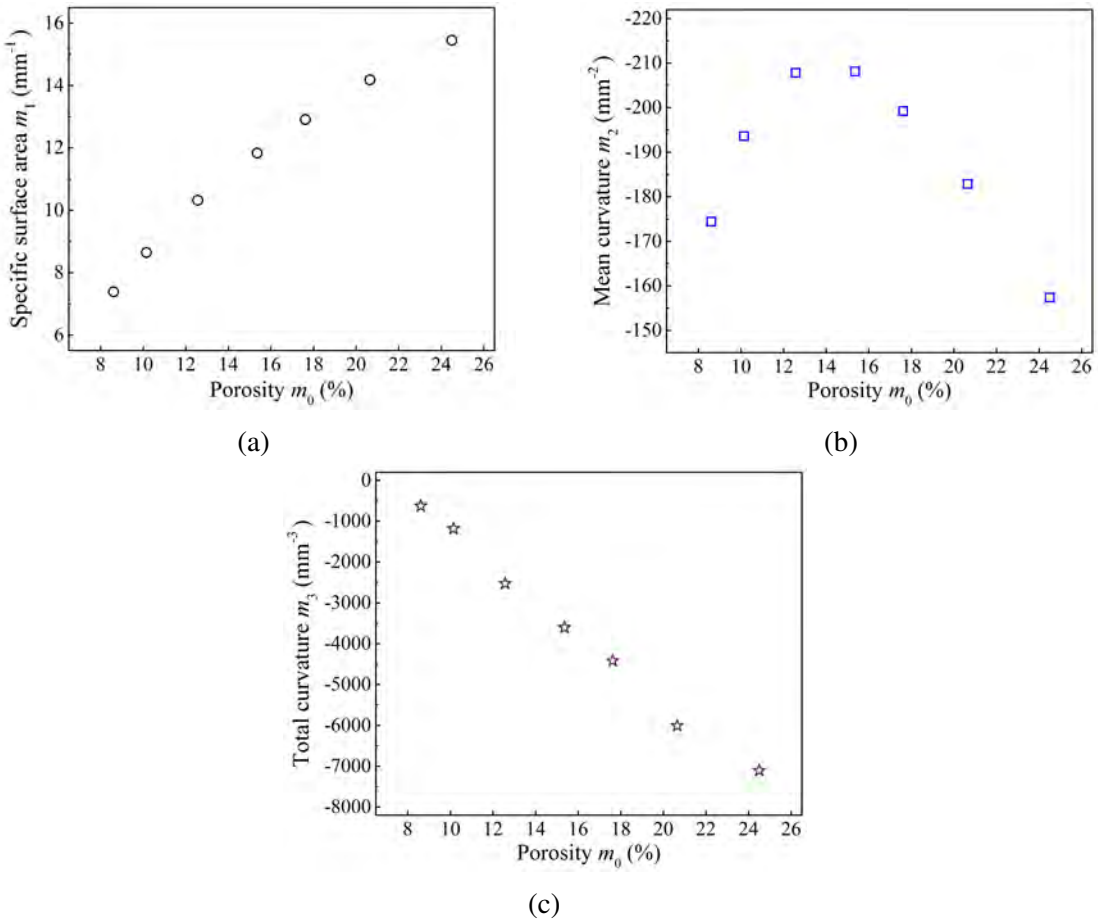


Fig. 2.4 The Minkowski densities m_k of the group of Fontainebleau sandstone samples: (a) Specific surface area m_1; (b) Mean curvature m_2; (c) Total curvature m_3

Minkowski functions

As discussed above, each Minkowski functional/density provides a single value by integrating over the pore space or the pore-solid surface, representing the average property of the

overall porous structure. However, Minkowski functionals/densities may not be sufficient to characterize the geometric features that are relevant to flow processes inside porous media [267], because the structural informatics such as pore size distribution and anisotropy are not involved. Therefore, Minkowski densities are also determined as functions of pore size by researchers [457, 459, 460, 390], called Minkowski functions, so as to provide geometric information on pore surface area and pore connectivity, as well as on pore/aggregate size distribution.

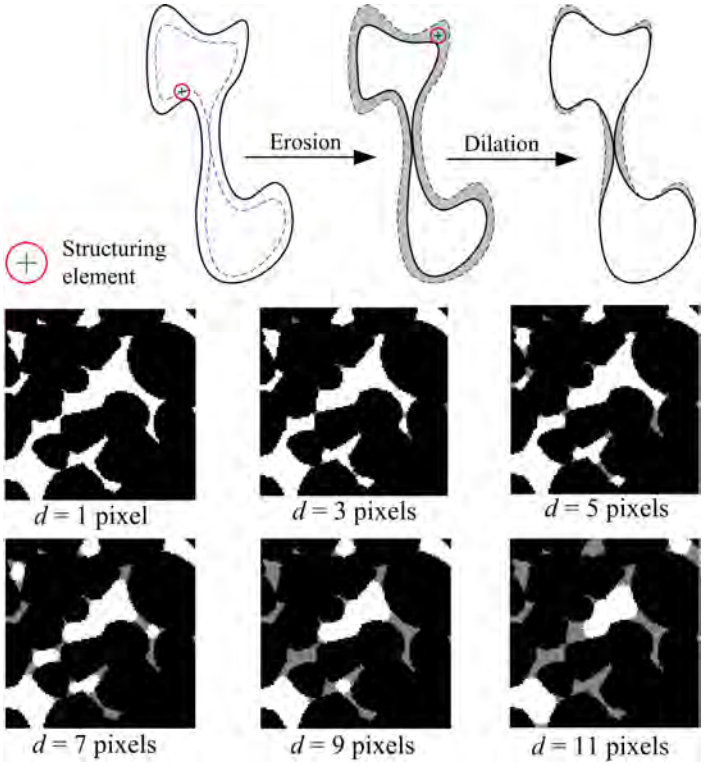


Fig. 2.5 Schematic illustration of morphological opening on a 2D image using a circular structuring element with increasing diameter d (pore space is in white, solid matrix is in black, and gray areas are the pores smaller than the structuring element after morphological opening)

Through a filter procedure based on mathematical morphology, the pores that are larger than a specified size can be extracted. This procedure, known as morphological opening, is an erosion operation followed by a dilation operation by using the same structuring element, as illustrated in Figure 2.5. By using a spherical structuring element with a prescribe diameter d, the pores smaller this diameter d are filtered out after morphological opening, and then Minkowski densities can be computed as functions of pore diameter by changing the diameter of the spherical structuring element, which are called Minkowski functions $m_k(d)$ [460, 390]. As shown in Figure 2.6, the Minkowski functions $m_0(d)$ of the group of

Fontainebleau sandstone are plotted. It should be noted that the first Minkowski function $m_0(d)$ is cumulative pore-size distribution representing the porosity larger than a given pore diameter, and its derivative is an alternative of pore-size distribution [459], as shown in Figure 2.29b. The fourth Minkowski function $m_4(d)$, also called pore connectivity function, is defined as the Euler characteristic in dependence of minimum pore size considered, which provides quantitative information on the connectivity within and between different classes of pore size [459].

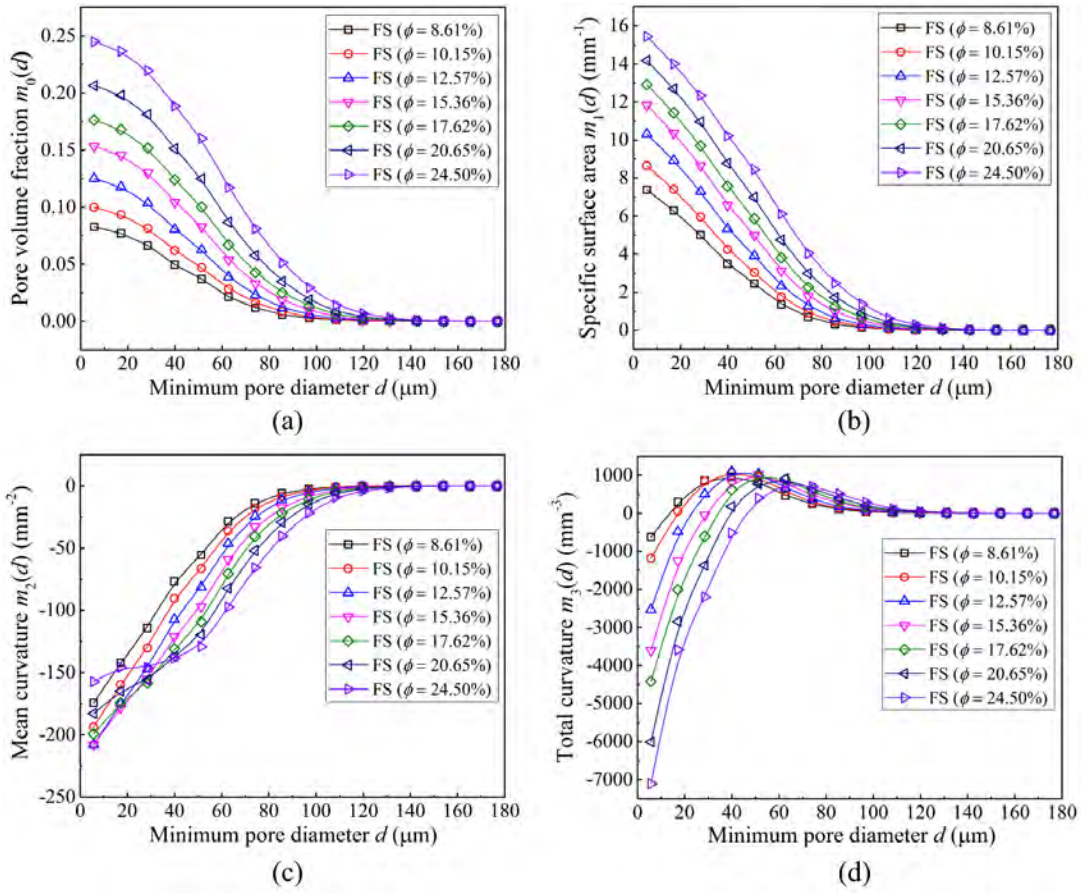


Fig. 2.6 The Minkowski functions $m_k(d)$ of the group of Fontainebleau sandstone samples: (a) Cumulative pore size distribution function $m_0(d)$; (b) $m_1(d)$; (c) $m_2(d)$; (d) Pore connectivity function $m_3(d)$

2.2.2 Spatial correlation functions

As early as in 1996, spatial correlation functions had been used to characterize the microstructures of rocks and other porous media by researchers, providing quantitative measures to understand the dependence of transport properties on the pore microstructure [92, 46].

The macroscopic physical properties of heterogeneous materials have been explicitly expressed as series expansions containing certain integrals of spatial correlation functions [441, 443]. Additionally, spatial correlation functions are also frequently used for stochastic reconstruction to generate statistically equivalent microstructure samples for porous media [493, 495, 217, 214].

Two/n-point correlation function

The n-point correlation function $S_n(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n)$ is commonly used to statistically characterize microstructures, which provides the probability of finding n points in the same material phase located at $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n$ [92, 441].

$$S_n(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n) = \langle X(\mathbf{r}_1)X(\mathbf{r}_2)...X(\mathbf{r}_n)\rangle \tag{2.6}$$

where $\langle\cdot\rangle$ denotes ensemble averaging, which is equivalent to volume averaging under ergodic hypothesis. For statistically homogeneous and isotropic media, the n-point correlation function only depends on the relative positions of the n points. The infinite family of n-point correlation functions exclusively and exhaustively describes the random nature of pore microstructures.

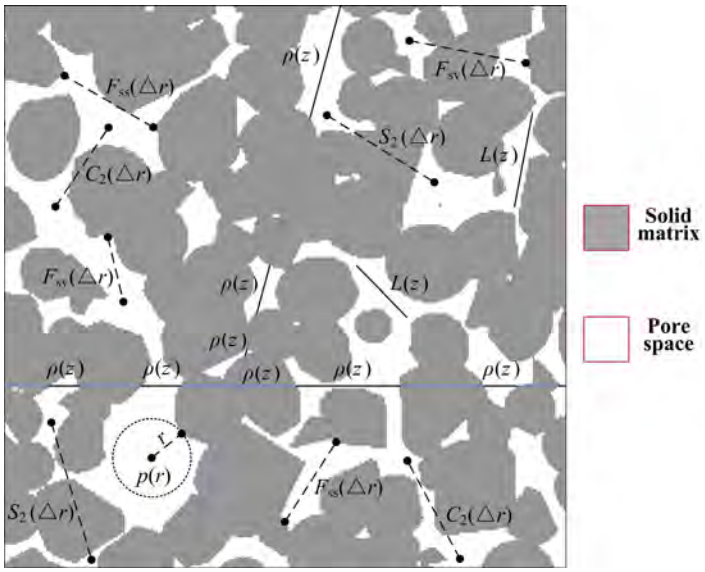


Fig. 2.7 Schematic illustration of statistical descriptors of 2D pore microstructures

The porosity (volume fraction) ϕ can also be considered as the one-point correlation function, which can be simply interpreted as the probability of finding a point in the pore space [92, 441]. As to the two-point correlation function $S_2(\mathbf{r}_1, \mathbf{r}_2)$, it depicts the probability

of finding any two points separated by distance $\Delta\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ who are located in the pore space (as illustrated in Figure 2.7). For the statistically homogeneous and isotropic porous media, $S_2(\mathbf{r}_1, \mathbf{r}_2)$ only depends on the magnitude of separation distance Δr, given by

$$S_2(\mathbf{r}_1, \mathbf{r}_2) = S_2(\mathbf{r}_2 - \mathbf{r}_1) = S_2(\Delta r) \tag{2.7}$$

where $\Delta r = |\mathbf{r}_2 - \mathbf{r}_1|$. In particular, the two-point correlation function has the significant properties as follows, which can be seen in Figure 2.8:

$$\begin{aligned} S_2(0) &= \phi \\ \lim_{\Delta r \to \infty} S_2(\Delta r) &= \phi^2 \end{aligned} \tag{2.8}$$

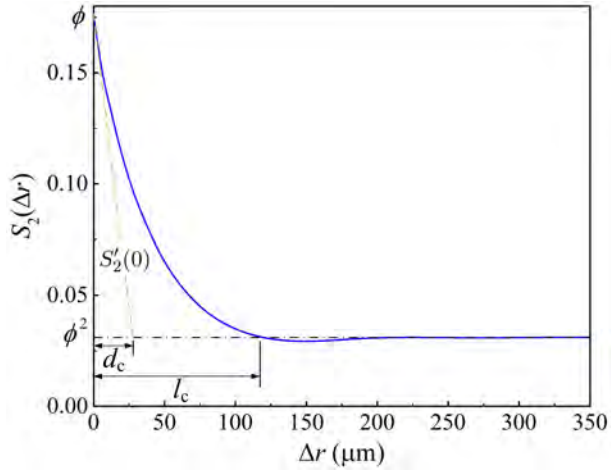


Fig. 2.8 An example of two-point correlation function $S_2(\Delta r)$

Two-point correlation function also carries the information of specific surface area S. When Δr is close to 0, the slop of $S_2(\Delta r)$ is proportional to S [185, 495, 443], given by:

$$\frac{\mathrm{d}(S_2(\Delta r))}{\mathrm{d}(\Delta r)}|_{\Delta r=0} = -S_2'(0) = \begin{cases} -S/2 & \text{for 1D space;} \\ -S/\pi & \text{for 2D space;} \\ -S/4 & \text{for 3D space.} \end{cases} \tag{2.9}$$

The effective pore size d_c can also be estimated from $S_2(\Delta r)$ [46, 12], given by

$$d_c = \frac{\phi - \phi^2}{-S_2'(0)} \tag{2.10}$$

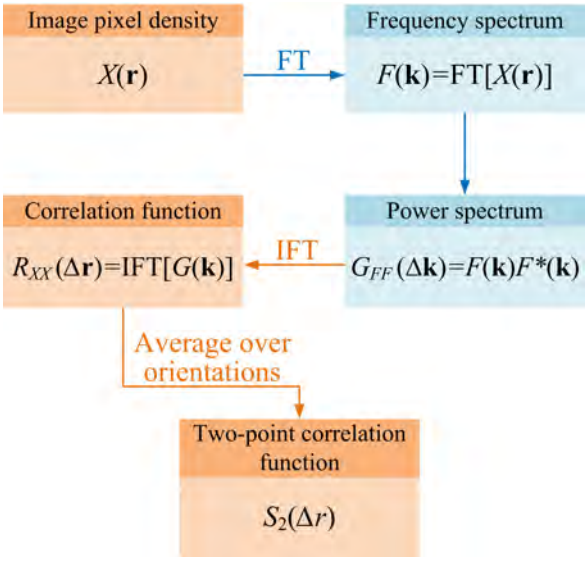


Fig. 2.9 The flow diagram of computing two-point correlation function via Fourier transform

The value of d_c is directly measured from the curve of $S_2(\Delta r)$, as illustrated in Figure 2.8, and it is related to the hydraulic diameter $d_h = 2\phi/S$ [2, 34, 344]. Additionally, several characteristic length scales can be extracted from $S_2(\Delta r)$ as follows [92]:

$$l_a = \int_0^\infty \left(S_2(\Delta r) - \phi^2\right) \mathrm{d}(\Delta r) \tag{2.11}$$

$$l_b = \sqrt{\int_0^\infty \left(S_2(\Delta r) - \phi^2\right) \Delta r \mathrm{d}(\Delta r)} \tag{2.12}$$

Another characteristic length scale is the distance l_c where $S_2(\Delta r)$ dwindles to its asymptotic value of ϕ^2 (as shown in Figure 2.8). All these length scales mentioned above have been used as probes to understand the influence of pore structure on transport/physical properties of porous media.

Generally, there are two common ways to extract two-point correlation function from a digital microstructure. One is the Monte Carlo method to randomly caste line segments with length Δr in a specified orientation and then to count the fraction of times that both of the endpoints fall in pore space [441]. The problem of this method is that it is computationally slow. The other way is using the radial integration of the power spectrum of the digital image via the Fourier transform [12]. This method is based on the famous Wiener-Khinchin theorem [135], which is much more efficient to compute two-point correlation function, compared to the former way [217]. Two-point correlation function is computed through the Fourier transform (FT) of the image, multiplying by the complex conjugate, inverse

Fourier transform (IFT) of the power spectrum and averaging the correlation function over all orientations, as illustrated in Figure 2.9.

It should be noted that the two-point correlation function discussed above is the void-void correlation function of porous media. Actually, grain-grain and void-grain correlation function also have been used to characterize pore microstructures [46, 214]. The two-point correlation functions and the corresponding high-level microstructural parameters of the group of Fontainebleau sandstone samples are provided in Figure 2.10 and Table 2.2, respectively.

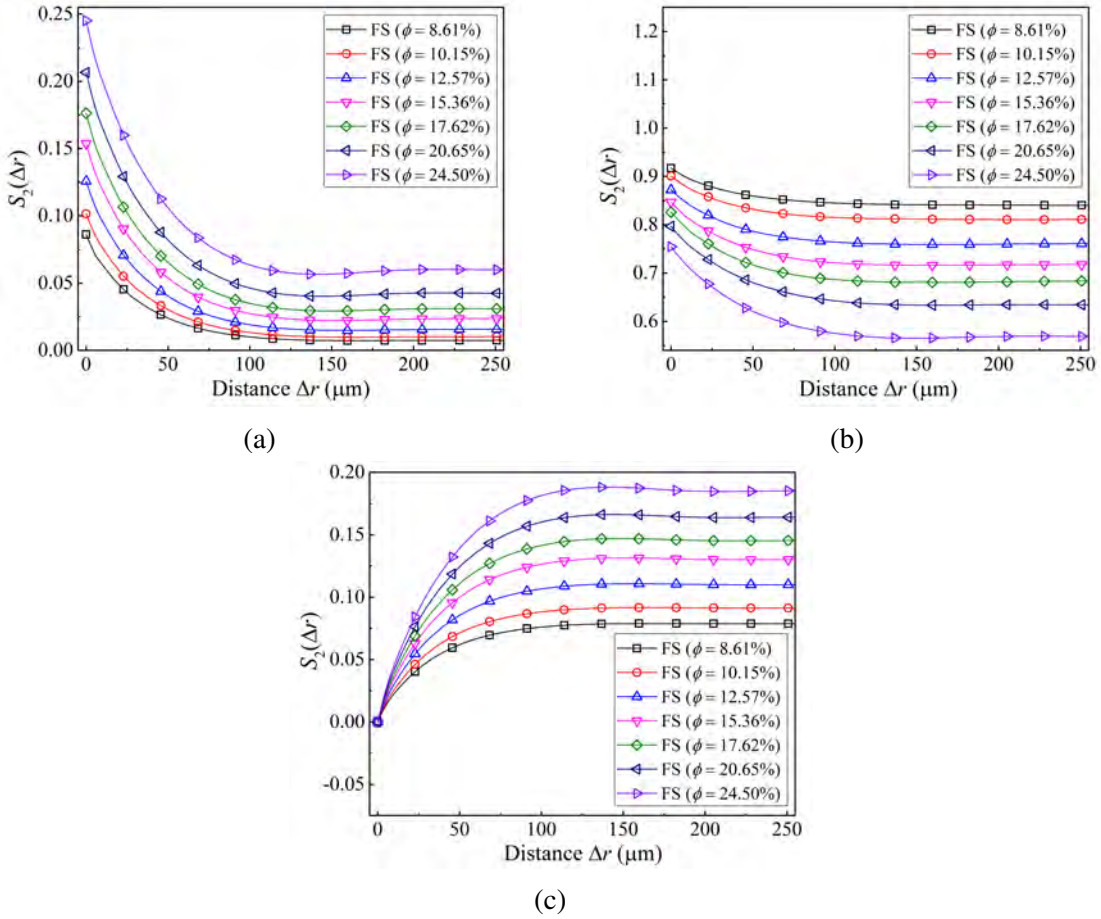


Fig. 2.10 The two-points correlation functions $S_2(\Delta r)$ of the group of Fontainebleau sandstone samples: (a) Void-void correlation; (b) Grain-grain correlation; (c) Void-grain correlation

Besides, three-point or other higher-order correlation functions, involving multiple reference points, are also valuable statistics for describing porous microstructures, which always encompass more information than lower-order correlation functions [213]. However, correlation functions with order higher than three encounter both analytical and numerical difficulties to compute [217]. Therefore, two-point correlation function, carrying informa-

tion of porosity and specific surface area, is the most commonly used correlation function, which has been arisen in rigorous expressions of effective transport or mechanical properties [46, 441].

Table 2.2 Microstructural parameters extracted from $S_2(\Delta r)$, $R(\Delta r)$, $\rho(z)$ and $\mu(\widetilde{\phi},L)$ of the group of Fontainebleau sandstone samples

Microstructural parameters	Fontainebleau sandstone samples with different porosities						
	8.61%	10.15%	12.57%	15.36%	17.62%	20.65%	24.50%
S (mm^{-1})	11.08	12.50	14.71	16.72	18.23	19.98	21.75
d_c (μm)	28.42	29.17	29.89	31.10	31.86	32.80	34.02
d_h (μm)	23.28	23.46	24.36	25.96	27.30	29.12	31.72
l_a (μm)	2.63	3.06	3.66	4.48	4.95	5.65	6.43
l_b (μm)	7.21	8.29	9.59	10.76	11.32	11.62	11.88
l_c (μm)	143.08	133.53	126.76	122.63	119.13	114.77	110.92
l_d (μm)	33.42	33.55	33.30	34.46	34.10	34.48	34.76
$\bar{z}$ (μm)	41.74	43.33	46.04	49.69	52.79	56.11	61.50
$\bar{r}$ (μm)	16.49	17.01	17.66	18.53	19.23	20.11	21.40
$L^\star$ (μm)	239.40	222.30	222.30	222.30	188.10	171.00	171.00

Normalized/scaled auto-covariance function

Normalized/scaled auto-covariance function $R(\Delta r)$, a variation of two-point correlation function $S_2(\Delta r)$, is widely used to statistically characterize and reconstruct porous microstructures [33, 277, 442, 215]. For statistically homogeneous and isotropic porous media, $R(\Delta r)$ is defined as

$$\begin{aligned} R(\Delta r) &= \frac{\langle (X(\mathbf{r}_1)-\phi)(X(\mathbf{r}_2)-\phi)\rangle}{\phi-\phi^2} \\ &= \frac{\langle X(\mathbf{r}_1)X(\mathbf{r}_2)\rangle-\phi^2}{\phi-\phi^2} \\ &= \frac{S_2(\Delta r)-\phi^2}{\phi-\phi^2} \end{aligned} \tag{2.13}$$

where $\Delta r = |\mathbf{r}_2-\mathbf{r}_1|$ again. Obviously, $R(\Delta r)$ is equivalent to $S_2(\Delta r)$ in terms of the carried morphology information of pore microstructures, and it has properties [186] as follows,

$$\begin{aligned} R(0) &= 1 \\ \lim_{\Delta r\to\infty} R(\Delta r) &= 0 \end{aligned} \tag{2.14}$$

In addition, the integral of $R(\Delta r)$ was used as a length scale to predict permeability of porous rocks [208]:

$$l_\mathrm{d} = \int_0^\infty R(\Delta r)\mathrm{d}(\Delta r) = \frac{l_\mathrm{a}}{\phi-\phi^2} \tag{2.15}$$

Normalized/scaled auto-covariance functions and the corresponding high-level microstructural parameter of the group of Fontainebleau sandstone samples are provided in Figure 2.11 and Table 2.2, respectively.

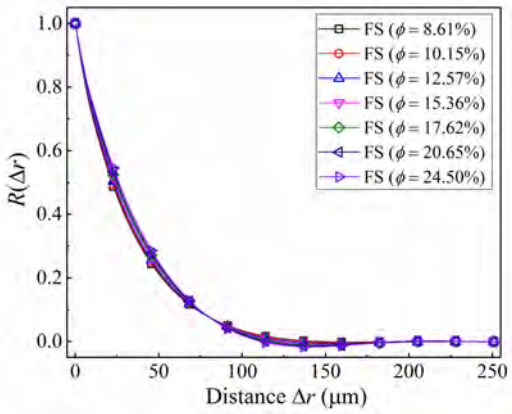


Fig. 2.11 Normalized auto-covariance functions $R(\Delta r)$ of the group of Fontainebleau sandstone samples

Two-point cluster correlation function

Although two-point correlation function $S_2(\mathbf{r}_1,\mathbf{r}_2)$ carries fundamental and indispensable information of pore microstructures, it cannot provide sufficient topological characteristics (clustering or connectedness) of pore space [441, 216]. As an analogue of $S_2(\mathbf{r}_1,\mathbf{r}_2)$, two-point cluster correlation function $C_2(\mathbf{r}_1,\mathbf{r}_2)$ is defined as the probability of finding both points $\mathbf{r}_1$ and $\mathbf{r}_2$ in the same cluster of void phase, which contains topological connectivity information [444, 441, 231].

The relationship between $S_2(\mathbf{r}_1,\mathbf{r}_2)$ and $C_2(\mathbf{r}_1,\mathbf{r}_2)$ is interpreted by decomposing $S_2(\mathbf{r}_1,\mathbf{r}_2)$ into a connected part and a disconnected part:

$$S_2(\mathbf{r}_1,\mathbf{r}_2) = C_2(\mathbf{r}_1,\mathbf{r}_2) + E_2(\mathbf{r}_1,\mathbf{r}_2) \tag{2.16}$$

where $E_2(\mathbf{r}_1,\mathbf{r}_2)$ denotes the probability of finding both points $\mathbf{r}_1$ and $\mathbf{r}_2$ in different clusters of void phase.

Similarly, $C_2(\mathbf{r}_1,\mathbf{r}_2)$ only depends on the magnitude of separation distance Δr for the statistically homogeneous and isotropic porous media, given by

$$C_2(\mathbf{r}_1,\mathbf{r}_2) = C_2(\mathbf{r}_2-\mathbf{r}_1) = C_2(\Delta r) \tag{2.17}$$

where $\Delta r = |\mathbf{r}_2-\mathbf{r}_1|$ again.

And $C_2(\Delta r)$ also has important properties [213] as follows:

$$\begin{aligned} C_2(0) &= \phi \\ \lim_{\Delta r \to \infty} C_2(\Delta r) &= 0 \end{aligned} \tag{2.18}$$

Finally, it should be noted that two-point cluster correlation function is an intrinsic measure for 3D microstructures, which can not be simply measured from 2D cross-section of porous media [441]. The two-point cluster correlation functions $C_2(\Delta r)$ of the group of Fontainebleau sandstone samples are given in Figure 2.12, which are very close to the two-point correlation functions $S_2(\Delta r)$ in Figure 2.10a. This is due to the good connectivity of pore space inside the Fontainebleau sandstone samples, which means most of the pores are interconnected to be one pore cluster. For pore media with many isolated pores, the $C_2(\Delta r)$ exhibits great difference from the $S_2(\Delta r)$.

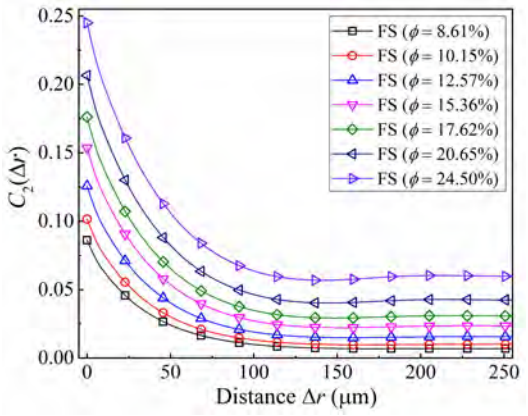


Fig. 2.12 Two-point cluster correlation functions $C_2(\Delta r)$ of the group of Fontainebleau sandstone samples

Pair connectivity function

Pair connectivity function $H(\mathbf{r}_1, \mathbf{r}_2)$ is defined as the probability of two points $\mathbf{r}_1$ and $\mathbf{r}_2$ are linked by a continuous path within the pore space P [9, 13, 389], given by

$$H(\mathbf{r}_1, \mathbf{r}_2) = \text{Prob}\{\mathbf{r}_1 \leftrightarrow \mathbf{r}_2,\ \mathbf{r}_2 \in P | \mathbf{r}_1 \in P\} \tag{2.19}$$

where $\mathbf{r}_1 \leftrightarrow \mathbf{r}_2$ means the points $\mathbf{r}_1$ and $\mathbf{r}_2$ are connected through a pore path. This connectivity measurement combines the concept of connected path used in percolation theory with the well-established lag-dependence [471, 389].

For the statistically homogeneous and isotropic porous media, $H(\mathbf{r}_1, \mathbf{r}_2)$ only depends on the magnitude of separation distance Δr. Apparently, pair connectivity function is in close accordance to two-point cluster correlation function $C_2(\Delta r)$ discussed above.

$$H(\Delta r) = \frac{C_2(\Delta r)}{\phi} \tag{2.20}$$

where $\Delta r = |\mathbf{r}_2 - \mathbf{r}_1|$ again. An average connectivity distance l_e is defined as the integral of $H(\Delta r)$ over all distance [242, 13], given by:

$$l_e = \int_0^\infty H(\Delta r) \mathrm{d}(\Delta r) \tag{2.21}$$

The pair connectivity functions of the group of Fontainebleau sandstone samples are provided in Figure 2.13, and the average connectivity distances are given in Table 2.2 .

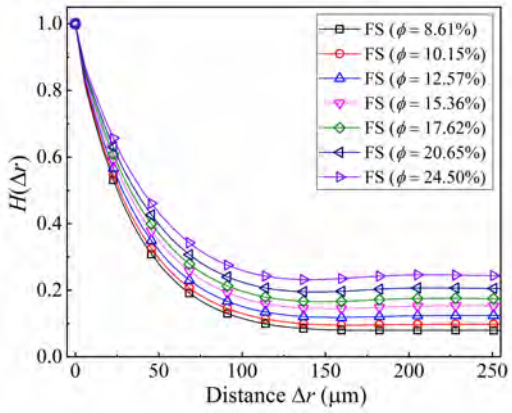


Fig. 2.13 Pair connectivity functions $H(\Delta r)$ of the group of Fontainebleau sandstone samples

Field-field correlation function

A pore microstructure is usually modeled as a truncation of a certain realization of random field [277, 372, 135, 14], and field-field correlation function is used to characterize the underlying random field $Y(\mathbf{r})$ instead of the porous microstructure itself $X(\mathbf{r})$ [213, 406]. The random field is often assumed to be statistically homogeneous up to the second order, so that its mean and variance are invariant when shifted in space. Gaussian random field (GRF) is commonly used to model microstructures due to its explicit and simple probability distribution.

The realization of a standard GRF $Y(\mathbf{r})$ at each point $\mathbf{r}$ marginally follows a standard normal distribution, and its marginal probability density function is as follows:

$$f\left(Y(\mathbf{r})=y_{\mathrm{r}}\right)=\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-y_{\mathrm{r}}^{2}}{2}\right) \tag{2.22}$$

where y_{r} is a realization of the GRF at point $\mathbf{r}$. And the cumulative distribution function is then given by

$$F\left(Y(\mathbf{r})=y_{\mathrm{r}}\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y_{\mathrm{r}}}\exp\left(\frac{-\zeta^{2}}{2}\right)\mathrm{d}\zeta \tag{2.23}$$

where ζ is a dummy variable. The porous microstructure $X(\mathbf{r})$ in binary form can be considered as the result from the truncation of a certain relation of the corresponding GRF $Y(\mathbf{r})$, with the thresholding value ϕ:

$$X(\mathbf{r})=\begin{cases}1 & \text{if } F\left(Y(\mathbf{r}=y_{\mathrm{r}})\right)\leq\phi\\ 0 & \text{otherwise}\end{cases} \tag{2.24}$$

where ϕ is the porosity.

For a statistically homogeneous and isotropic porous microstructure, field-field correlation function $G(\Delta r)$ is a direct way to characterize the underlying GRF $Y(\mathbf{r})$, given by

$$G(\mathbf{r}_1,\mathbf{r}_2)=\langle Y(\mathbf{r}_1)Y(\mathbf{r}_2)\rangle=G(\Delta r) \tag{2.25}$$

where $\Delta r=|\mathbf{r}_2-\mathbf{r}_1|$ again.

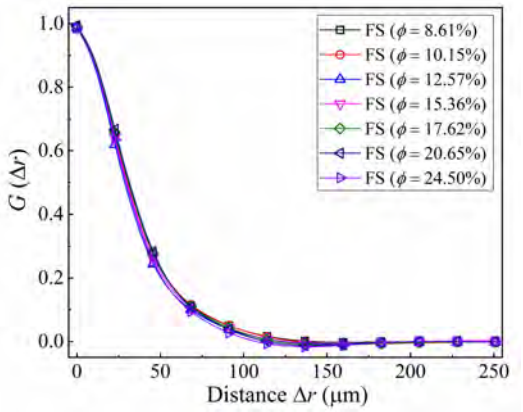


Fig. 2.14 Field-field correlation functions $G(\Delta r)$ of the group of Fontainebleau sandstone samples

Basically, the field-field correlation function $G(\Delta r)$ is the correlation of two points separated by distance Δr in the GRF $Y(\mathbf{r})$. The transformation of the underlying GRF $Y(\mathbf{r})$ from the porous microstructure $X(\mathbf{r})$ can be implemented through iterative optimization methods or explicit transformation, and more details can be found in relevant references [374, 373, 405, 213, 135]. As shown in Figure 2.14, the field-field correlation functions of the group of Fontainebleau sandstone samples are given.

Surface correlation functions

Surface correlation functions [441, 216, 296] contain essential information about the pore-solid interface ∇X, which are of fundamental importance to transport properties of fluid-saturated porous media, including permeability, effective diffusivity and trapping constant [377, 378, 443]. The specific surface area S mentioned above is the simplest surface correlation function, which is a constant everywhere for homogeneous media:

$$S = \langle \nabla X \rangle \tag{2.26}$$

Specific surface area S can also be considered as a one-point correlation function, but it cannot be interpreted as probability, because the likelihood of finding a point on the pore-solid interface is zero.

As to two-point surface correlation functions, surface-surface correlation function F_{ss} and surface-void correlation function F_{sv} are often used to characterize pore microstructures. For statistically homogeneous and isotropic porous media, they are defined as follows:

$$F_{\mathrm{ss}}(\mathbf{r}_1, \mathbf{r}_2) = \langle \nabla X(\mathbf{r}) \nabla X(\mathbf{r} + \Delta \mathbf{r}) \rangle = F_{\mathrm{ss}}(\Delta r) \tag{2.27}$$

$$F_{\mathrm{sv}}(\mathbf{r}_1, \mathbf{r}_2) = \langle \nabla X(\mathbf{r}) X(\mathbf{r} + \Delta \mathbf{r}) \rangle = F_{\mathrm{sv}}(\Delta r) \tag{2.28}$$

where $\Delta r = |\mathbf{r}_2 - \mathbf{r}_1|$ again. The asymptotic behaviors of these two surface correlation functions are given by:

$$\lim_{\Delta r \to \infty} F_{\mathrm{ss}}(\Delta r) = S^2 \tag{2.29}$$

$$\lim_{\Delta r \to \infty} F_{\mathrm{sv}}(\Delta r) = S\phi \tag{2.30}$$

As shown in Figure 2.15, the surface-surface correlation functions and surface-void correlation functions of the group of Fontainebleau sandstone samples are provided.

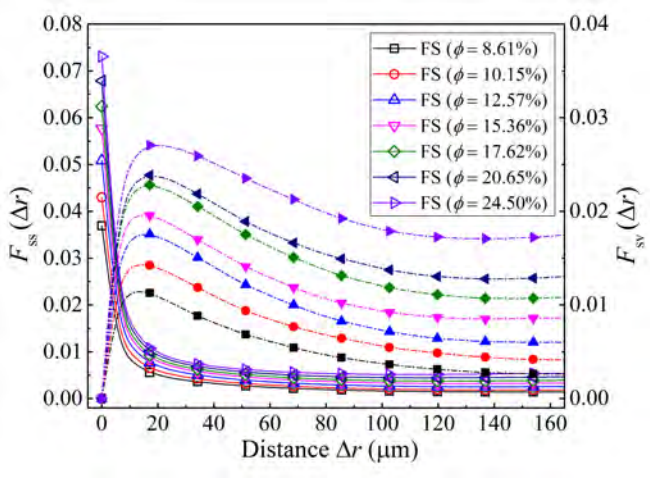


Fig. 2.15 Surface correlation functions $F_{ss}(\Delta r)$ (solid line + empty label) and $F_{sv}(\Delta r)$ (dashed line + solid label) of the group of Fontainebleau sandstone samples

2.2.3 Contact distribution function

Contact distribution functions are used to characterize the stochastic geometry [83], and some special cases have been applied in porous media research, such as spherical and linear contact distribution function [319, 186].

Spherical contact distribution function

For a given porous medium X and a convex compact set B containing the internal origin o, contact distribution function $H_B(d)$ is defined as

$$H_B(d) = 1 - \frac{\text{Prob}\{M \cap (d \times B) = \varnothing\}}{\phi} \quad \text{for } d \geq 0 \tag{2.31}$$

where M denotes the solid matrix, $\varnothing$ denotes empty set, and ϕ is the porosity. A particular important case is that the convex compact set B is a unit sphere $S(o, d = 1)$ centered at the point o, yielding the spherical contact distribution function $H_S(d)$:

$$H_S(d) = 1 - \frac{\text{Prob}\{M \cap S(o,d) = \varnothing\}}{\phi} \quad \text{for } d \geq 0 \tag{2.32}$$

where $S(o,d)$ denotes a sphere centered at point o with diameter d. More specifically, $H_S(d)$ is defined as the probability that the diameter of a randomly centered sphere completely falling into the pore space is not larger than d.

Spherical contact distribution function $H_S(d)$ not only accounts for pore shapes but also pore size distribution. Its probability density is an alternative to characterize pore size

distribution function $p(d)$ of porous media [186, 267], given by:

$$p(d) = \frac{\mathrm{d}H_S(d)}{\mathrm{d}d} \tag{2.33}$$

where d here is equivalent to pore diameter. The $p(d)$ can be considered as the fraction of spheres centered at pore space with diameter d which have no intersection with solid matrix. The spherical contact distributions and pore size distributions of the group of Fontainebleau sandstone samples are given in Figure 2.16 and 2.29b, respectively.

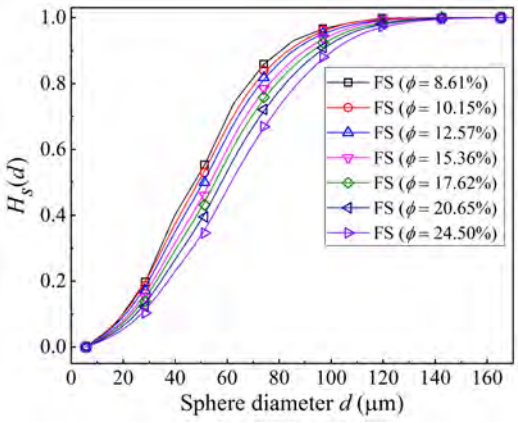


Fig. 2.16 Spherical contact distributions of the group of Fontainebleau sandstone samples

Linear contact distribution function

Analogous to spherical contact distribution function, linear contact distribution function $H_l(z)$ [319] corresponds to the case where the convex compact set B is a unit line segment $l(o,z=1)$. $H_l(z)$ is often redefined as lineal path function $L(z)$ to characterize porous media [186]. For statistically isotropic porous media, lineal path function $L(z)$ [291, 92, 333, 186] provides the probability that a line segment with length z wholly falls inside the pore space P when randomly casting it into the porous medium sample X, given by:

$$L(z) = \phi(1 - H_l(z)) = \mathrm{Prob}\{M \cap l(o,z) = \varnothing\} \tag{2.34}$$

The schematic illustration of $L(z)$ can been in Figure 2.7.

For the line segment with length $z \to 0$, $L(z)$ reduces to the probability of finding a point in the pore space. As to the line segment with large length, $L(z)$ decays to zero rapidly.

$$\begin{aligned} L(0) &= \phi \\ \lim_{z\to\infty} L(z) &= 0 \end{aligned} \tag{2.35}$$

Generally, $L(z)$ is considered to contain partial topological connectedness information of pore space. The lineal path functions of the group of Fontainebleau sandstone samples are given in Figure 2.17.

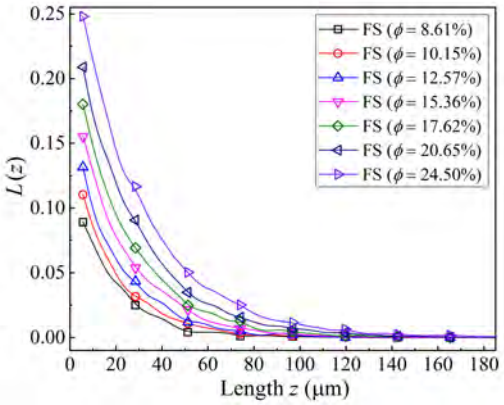


Fig. 2.17 Lineal path functions $L(z)$ of the group of Fontainebleau sandstone samples

Chord length distribution function

Throwing an infinitely long line into a porous medium, it will be intersected into line segments by the pore-solid interface, and the line segments lying in pore or solid phase are called chords, as illustrated in Figure 2.7. Chord length is thus defined as the distance between two intersections of a line with the pore-solid boundary. For statistically isotropic porous media, chord length distribution function $\rho(z)$ is defined as the probability of finding a chord with the length between z and $z+\mathrm{d}z$ completely lying inside pore space [441].

Chord length distribution function $\rho(z)$ [319, 92, 333] is highly related to lineal path function $L(z)$ though

$$\rho(z) = \frac{\bar{z}}{\phi}\frac{\mathrm{d}^2(L(z))}{\mathrm{d}z^2} \tag{2.36}$$

where $\bar{z}$ is the mean chord length and it is defined as

$$\bar{z} = \int_0^\infty z\rho(\zeta)\mathrm{d}\zeta \tag{2.37}$$

where ζ is a dummy variable.

Chord length distribution function $\rho(z)$ only possesses limited information on connectivity of pore network, because connectedness is evaluated along lineal test lines. The chord length distribution functions and the mean chord lengths of the group of Fontainebleau sandstone samples are provided in Figure 2.18 and Table 2.2, respectively.

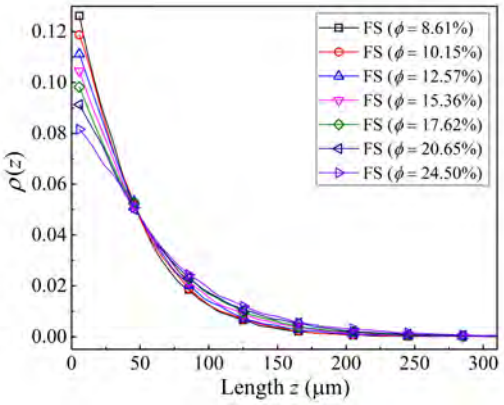


Fig. 2.18 Chord length distribution functions $\rho(z)$ of the group of Fontainebleau sandstone samples

2.2.4 Local porosity characterization

Local porosity characterization [290, 184, 186, 408] is a class of scale-dependent measurements to quantitatively represent the porosity fluctuation and connectivity variation for porous media. As illustrated in Figure 2.19, variations in local porosity and connectivity of porous media are presented. Two geometric characteristics, including local porosity distribution and coarseness [290, 92], have been successively created to measure the local volume fraction fluctuation in porous media. Local percolation probabilities and total fraction of percolation cells are used to characterize the size-dependent connectivity variation inside porous media [183, 44]. The local porosity characteristics are highly related to transport properties of porous media, including permeability and conductivity [45, 186].

Local porosity distribution

Local porosity distribution [183, 44, 186] is to measure local porosities from compact subsets of a porous medium and collect the measurements in a histogram form to represent the empirical probability density. A measurement cell $K(\mathbf{r}, L)$ is defined as a cube (or a square) with side-length L centered at the lattice vector $\mathbf{r}$ inside the 3D (or 2D) pore microstructure

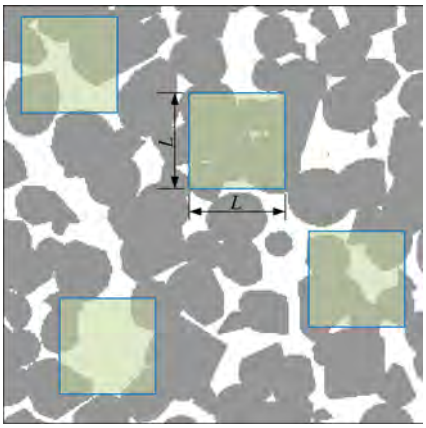


Fig. 2.19 Schematic illustration of variations in local porosity and connectivity of porous media

(as shown in Figure 2.19), and local porosity $\phi_{\text{local}}(\mathbf{r},L)$ is defined as the porosity of this measurement cell $K(\mathbf{r},L)$:

$$\phi_{\text{local}}(\mathbf{r},L) = \frac{V(P \cap K(\mathbf{r},L))}{V(K(\mathbf{r},L))} \tag{2.38}$$

where $V(\cdot)$ denotes the volume.

Local porosity distribution $\mu(\widetilde{\phi},L)$ provides the probability density of finding a local porosity equal to $\widetilde{\phi}$ when the side-length of measurement cell is set to L, and it is defined as

$$\mu(\widetilde{\phi},L) = \frac{1}{N\Delta\widetilde{\phi}} \sum_{\mathbf{r}} \delta(\widetilde{\phi} - \phi_{\text{local}}(\mathbf{r},L)) \tag{2.39}$$

where N is the number of placements of the measurement cells $K(\mathbf{r},L)$, and $\delta(\widetilde{\phi} - \phi_{\text{local}}(\mathbf{r},L))$ is the Dirac δ function which is defined as follows:

$$\delta(\widetilde{\phi} - \phi_{\text{local}}(\mathbf{r},L)) = \begin{cases} 1, \text{ if } |\widetilde{\phi} - \phi_{\text{local}}(\mathbf{r},L)| \leq \mathrm{d}\widetilde{\phi}; \\ 0, \text{ otherwise.} \end{cases} \tag{2.40}$$

where $\mathrm{d}\widetilde{\phi}$ denotes the interval to build the histogram associated with local porosity distribution. Essentially, $\mu(\widetilde{\phi},L)\mathrm{d}\widetilde{\phi}$ measures the probability of finding a measurement cell with side-length L and local porosity between $\widetilde{\phi}$ and $\widetilde{\phi} + \mathrm{d}\widetilde{\phi}$.

Obviously, $\mu(\widetilde{\phi},L)$ depends on the side-length L of the measurement cell, but there is no common rule to determine an optimal value of L for a specific porous medium. A special length scale $L^{\star}$ is defined to measure the pore size, given by:

$$L^{\star} = \min\{L : \mu(0,L) = \mu(1,L) = 0\} \tag{2.41}$$

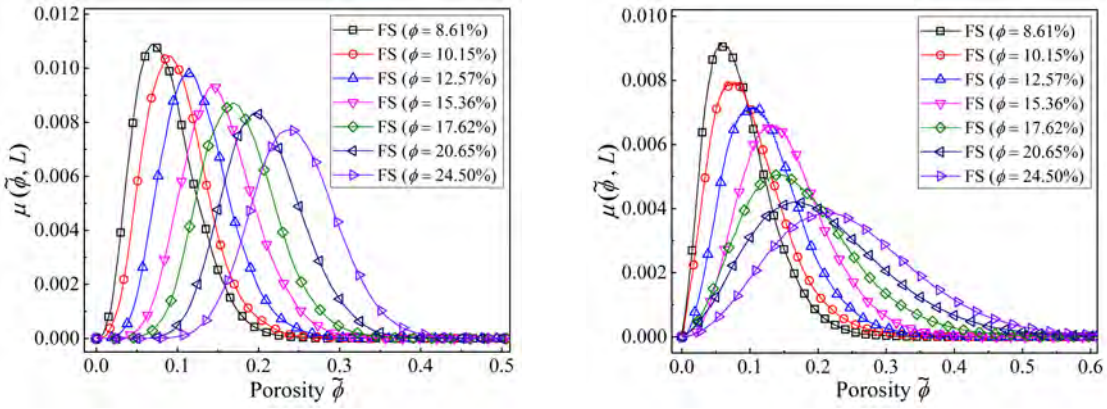
where the Dirac δ function at $\phi_{\text{local}} = 0$ and $\phi_{\text{local}} = 1$ vanish for the first time. And the length scale $L^\star$ is usually used as the characteristic length to extract the local porosity distribution $\mu(\widetilde{\phi}, L^\star)$ and the local percolation probability $\lambda_\alpha(\widetilde{\phi}, L^\star)$ from porous media [44, 186].

If the porous medium is macroscopically homogeneous, its local porosity distribution $\mu(\widetilde{\phi}, L)$ has following properties:

$$\begin{aligned} \lim_{L\to 0} \mu(\widetilde{\phi}, L) &= \phi\delta(\widetilde{\phi} - 1) + (1-\phi)\delta(\widetilde{\phi}) \\ \lim_{L\to\infty} \mu(\widetilde{\phi}, L) &= \delta(\widetilde{\phi} - \phi) \end{aligned} \tag{2.42}$$

where ϕ is bulk porosity of the porous medium.

Besides, local densities of surface area, mean curvature and total curvature are also defined in the way analogous to local porosity distribution, but they are not used frequently in microstructure characterization (more details can be found in [185]). The local porosity distributions and the corresponding length scales $L^\star$ of the group of Fontainebleau sandstone samples are given in Figure 2.20 and 2.2, receptively.



(a) $L = 50$ voxels (b) $L = L^\star$

Fig. 2.20 Local porosity distributions $\mu(\widetilde{\phi}, L)$ of the group of Fontainebleau sandstone samples (the value of $L^\star$ decreases with porosity, which is equal to 42, 39, 39, 39, 33, 30 and 30 voxels respectively for these sandstone samples)

Coarseness

Coarseness $C(L)$ provides a quantitative measure of size-dependent fluctuation of local porosity $\phi_{\text{local}}(\mathbf{r}, L)$, and it has extensive use to characterize stochastic microstructures [290, 92, 364, 408]. The standard deviation associated with the local porosity $\phi_{\text{local}}(\mathbf{r}, L)$ of measurement cells $K(\mathbf{r}, L)$ can be used to represent the fluctuation of spatially local porosity

for statistically homogeneous media. Coarseness $C(L)$ is defined as the standard deviation divided by the mean of $\phi_{\text{local}}(\mathbf{r},L)$, given by

$$\begin{aligned} C(L) &= \frac{\sqrt{\left\langle \phi^2_{\text{local}}(\mathbf{r},L)\right\rangle - \left\langle \phi_{\text{local}}(\mathbf{r},L)\right\rangle^2}}{\left\langle \phi_{\text{local}}(\mathbf{r},L)\right\rangle} \\ &= \frac{\sqrt{\left\langle \phi^2_{\text{local}}(\mathbf{r},L)\right\rangle - \phi^2}}{\phi} \end{aligned} \tag{2.43}$$

where ϕ is the bulk porosity again. Coarseness $C(L)$ has following properties, when the measurement cell $K(\mathbf{r},L)$ has infinitely small or large size L:

$$\begin{aligned} &\lim_{L\to 0} C(L) = \frac{\sqrt{\phi(1-\phi)}}{\phi} \\ &\lim_{L\to \infty} C(L) = 0 \end{aligned} \tag{2.44}$$

It should be noted that the measurement cell $K(\mathbf{r},L)$ can be an arbitrary-shaped observation region, not limited to a cube or square. The coarseness results of the group of Fontainebleau sandstones are given in Figure 2.21.

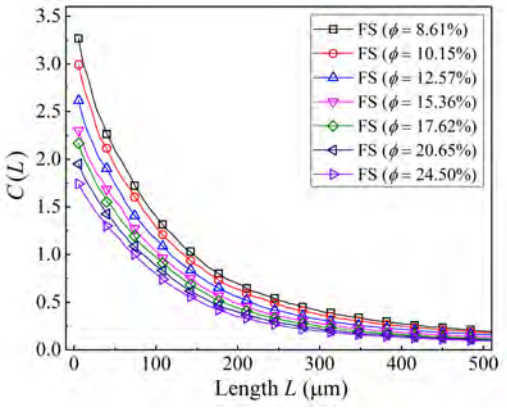


Fig. 2.21 Coarseness $C(L)$ results of the group of Fontainebleau sandstone samples

Local percolation probabilities

Local percolation probabilities [183, 44, 186, 101] are used to characterize the connectivity of measurement cells $K(\mathbf{r},L)$ with the specific local porosity. A measurement cell $K(\mathbf{r},L)$ is termed percolating in direction α, if there exists at least one pore path lying inside this measurement cell that connects two points on the opposite surfaces perpendicular to this direction.

The indicator function of the measurement cell $K(\mathbf{r},L)$ percolating in direction α is given by

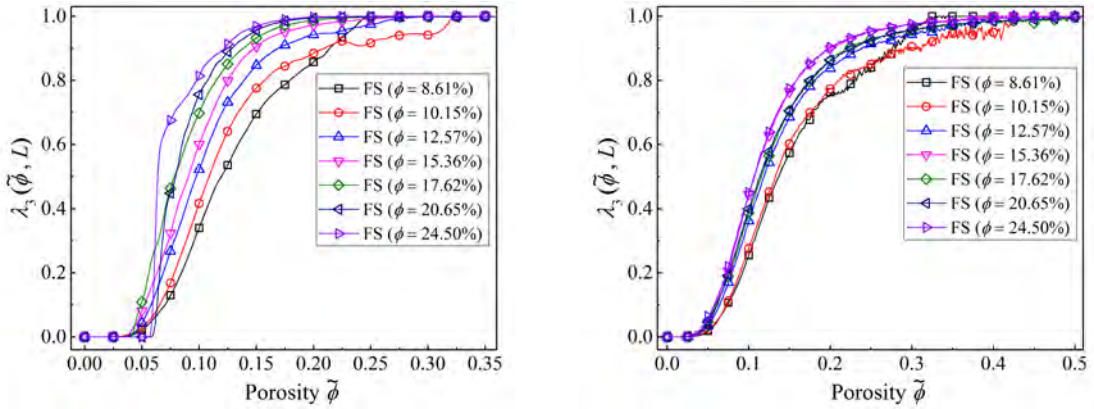
$$\Lambda_\alpha(\mathbf{r},L) = \begin{cases} 1, \text{ if } K(\mathbf{r},L) \text{ percolates in the direction } \alpha; \\ 0, \text{ otherwise.} \end{cases} \tag{2.45}$$

Thus, $\Lambda_x(\mathbf{r},L)=1$, $\Lambda_y(\mathbf{r},L)=1$ or $\Lambda_z(\mathbf{r},L)=1$ indicates the measurement cell is percolating in the x-, y- or z-direction. It should be noted that $\Lambda_1(\mathbf{r},L)=1$ and $\Lambda_3(\mathbf{r},L)=1$ indicate the measurement cell is percolating in at least one of the principal directions and in all the three principal directions, respectively.

The local percolating probability $\lambda_\alpha(\widetilde{\phi},L)$ in the direction α is defined as follows:

$$\lambda_\alpha(\widetilde{\phi},L) = \frac{\sum_{\mathbf{r}} \Lambda_\alpha(\mathbf{r},L)\delta(\widetilde{\phi}-\phi_{\text{local}}(\mathbf{r},L))}{\sum_{\mathbf{r}} \delta(\widetilde{\phi}-\phi_{\text{local}}(\mathbf{r},L))} \tag{2.46}$$

Basically, $\lambda_\alpha(\widetilde{\phi},L)$ provides the fraction of measurement cells with side-length L and local porosity $\widetilde{\phi}$ that are percolating in direction α. Local percolating probabilities of the group of Fontainebleau sandstone samples are provided in Figure 2.22.



(a) $L = 50$ voxels (b) $L = L^\star$

Fig. 2.22 Local porosity probabilities $\lambda_3(\widetilde{\phi},L)$ of the group of Fontainebleau sandstone samples

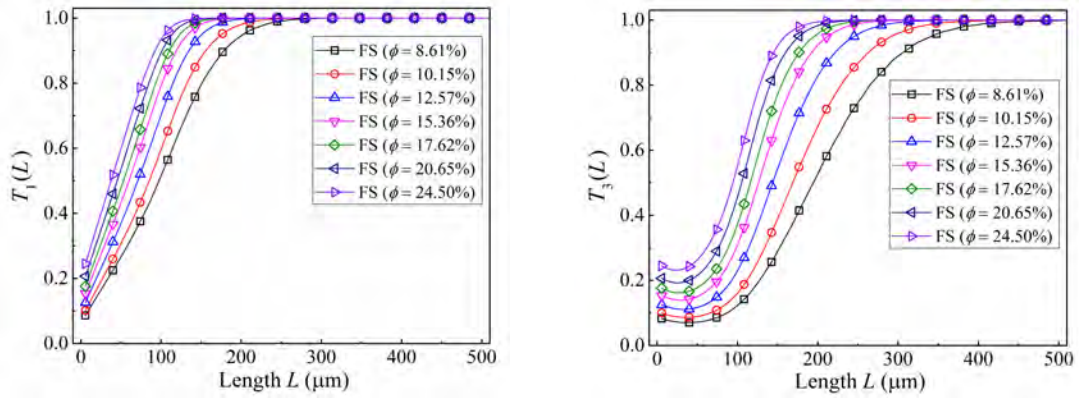
Total fraction of percolating cells

Total fraction of percolating cells $T_\alpha(L)$ [44, 186, 101] provides the fraction of all measurement cells $K(\mathbf{r},L)$ with side-length L that are percolating in the direction α, and it is

calculated from the integration over all local porosities:

$$T_\alpha(L) = \int_0^1 \mu(\widetilde{\phi}, L)\lambda_\alpha(\widetilde{\phi}, L)\,\mathrm{d}\widetilde{\phi} \tag{2.47}$$

Total fraction of percolating cells $T_\alpha(L)$ quantifies the connectivity characteristic of the pore network, which is important for modeling transport properties [262]. The Total fractions of percolating cells fo the group of Fontainebleau sandstone samples are given in Figure 2.23.



(a) Percolation in one of three principal directions (b) Percolation in all three principal directions

Fig. 2.23 Total fractions of percolation cells $T_1(L)$ and $T_3(L)$of the group of Fontainebleau sandstone samples

2.2.5 Pore size distribution

The size of pores within porous media could affect the transport and mechanical properties, so pore size distribution (PSD) is one of the most important descriptors to quantitatively represent pore structure, and then to understand the structure-property linkages. Generally, pores are classified into three categories according to their sizes: micropores (pore diameter $d \leq 2$ nm), mesopores (2 nm $< d \leq 50$ nm), and macropores ($d > 50$ nm), which was recommended by the International Union of Pure and Applied Chemistry [410]. A wide pore size spectrum associated with ultrafine pore sizes may be exhibited in porous media such as shale, which sometimes makes it difficult to obtain the full PSD information.

It is difficult to define the "pore body" inside natural porous rocks, and there is no uniformity of PSD [185, 281]. Due to the mathematically unambiguous definitions of PSD [321], several image analysis methods have been proposed to extract PSD from images. These methods can be generally classified into the intercept method, the random point

method, the pore skeleton method, the discrete method, the continuous method and the morphological opening method. It should be noted that PSD can only be extracted from digital microstructures in 3D form, because it contains connectedness information of pore space.

Intercept method

The PSD obtained from the intercept method [140, 194, 74] reflects the probability of intercepted lengths of randomly oriented lines by the pore-solid interface, as illustrated in Figure 2.24. Both lineal path function and chord length distribution function can be used as the intercept PSDs to estimate characterize the pore structure, and more details have been given in Section 2.2.3. However, intercept PSD may poorly represent the pore size of porous media with complicated morphologies. For example, very different interception lengths are observed at the location of a bottle neck (as shown in Figure 2.24), which may lead to relatively wider pore size distribution.

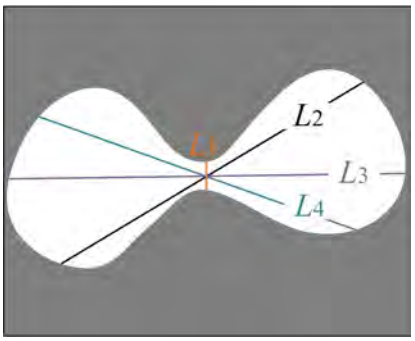


Fig. 2.24 Schematic diagram of PSD determination using the intercept method

Random point method

The random point method [291, 185, 92, 441] is one of the simplest approach to obtaining PSD $p(r)$, where $p(r)\mathrm{d}r$ is defined as the probability of randomly choosing a point in the pore space that lies at the distance between r and $r+\mathrm{d}r$ to the pore-solid interfaces. Essentially, this distance r is equal to the radius of the sphere centered at the randomly selected point and inscribed to the pore-solid surface, as illustrated in Figure 2.7. The sum of probability is equal to 1, given by:

$$\int_0^\infty p(r)\mathrm{d}r = 1 \tag{2.48}$$

PSD $p(r)$ satisfies properties as follows:

$$\begin{aligned} \lim_{r\to 0} p(r) &= \frac{S}{\phi} \\ \lim_{r\to \infty} p(r) &= 0 \end{aligned} \tag{2.49}$$

Besides, the mean pore size $\bar{r}$ is defined as the first moment of $p(r)$, given by:

$$\bar{r} = \int_0^\infty r p(r) \mathrm{d}r \tag{2.50}$$

In this study, the distances of all pore pixels/voxels to the pore-solid interfaces are calculated, to form a distance map, as shown in Figure 2.25a. The distance data is then collected in a histogram form to represent the empirical probability density of pore radius. The PSDs and mean pore sizes of the group of Fontainebleau sandstone samples are provided in Figure 2.25b and Table 2.2, respectively. Generally, the PSD obtained from this random point method provides less information of the geometrical complexity of pore space.

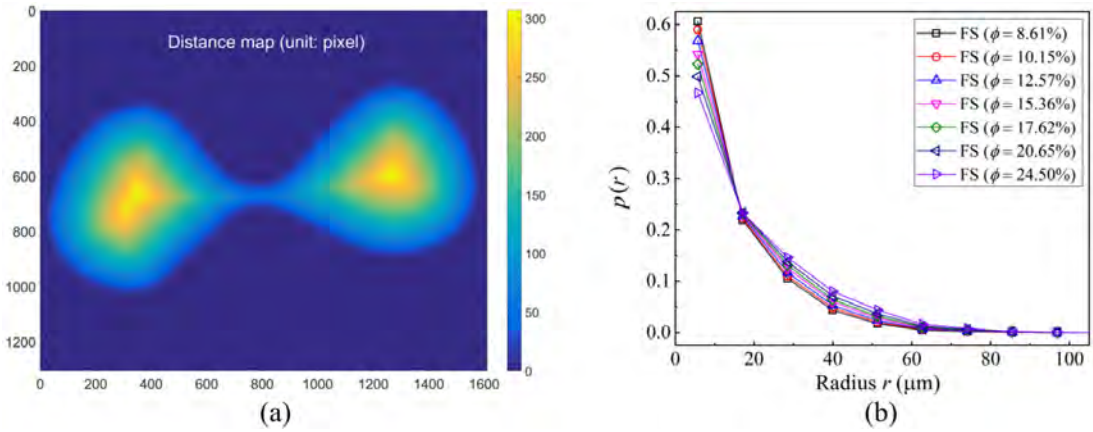


Fig. 2.25 (a) Schematic diagram of distance map, and (b) PSD $p(r)$ of the group of Fontainebleau sandstone samples

Pore skeleton method

The pore skeleton method [280, 115, 281, 431] is analogous to the random point method, but it selects the points on the medial axis (pore skeleton) of the pore space as the centers of circles/spheres inscribed to the pore-solid interfaces, in order to capture geometric characteristics related to transport flow. The pore skeleton can be extracted from a pore space through a skeletonization process, such as the homotopic thinning algorithm [280] and the fast marching algorithm [431]. The distances from the points on the pore skeleton to the

pore-solid interfaces are equal to the radii of circles/spheres that are centered on the pore skeleton and inscribed to the pore-solid interfaces. The occurrence frequencies of radii are used to estimate the PSD. Here, the homotopic thinning algorithm is adopted to extract pore skeleton, and then the inscribed circles centered on the pore skeleton are determined, as shown in Figure 2.26a. The pore-skeleton-based PSDs of the group of Fontainebleau sandstone samples are shown in Figure 2.26b.

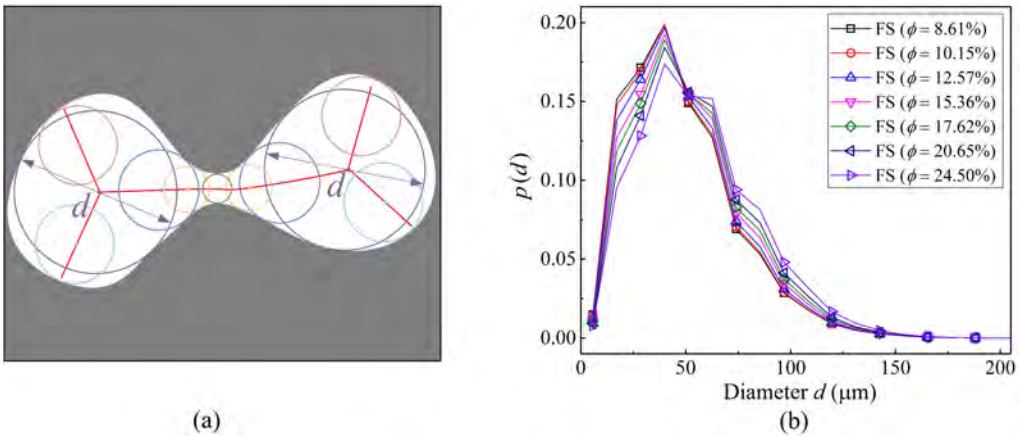


Fig. 2.26 Determination of PSD by using the continuous method: (a) schematic diagram of circles centered at the pore skeleton and inscribed to pore-solid interface, and (b) the continuous PSDs of the group of Fontainebleau sandstone samples

Discrete method

The entire pore network can be split into discrete objects based on a specific criterion, and then equivalent pore diameters d_{eq} are derived from the areas (2D) or the volumes (3D) of the discrete pore bodies [321, 194], as illustrated in Figure 2.27a. However, there is a lack of reliable methods to recognize individual pore objects, and the splitting process may introduce significant uncertainties to the determination of PSD [320].

Here, the watershed algorithm [309] is adopted to partition the pore network, and the discrete PSD is then calculated in a histogram form to represent the empirical probability density of pore diameter. The discrete PSDs of the group of Fontainebleau sandstone samples are given in Figure 2.27b.

Continuous method

In reality, the pore space inside a natural porous rock does not consist of isolated cavities with simple shapes, which is a continuous 3D network with irregular geometry, so the conventional

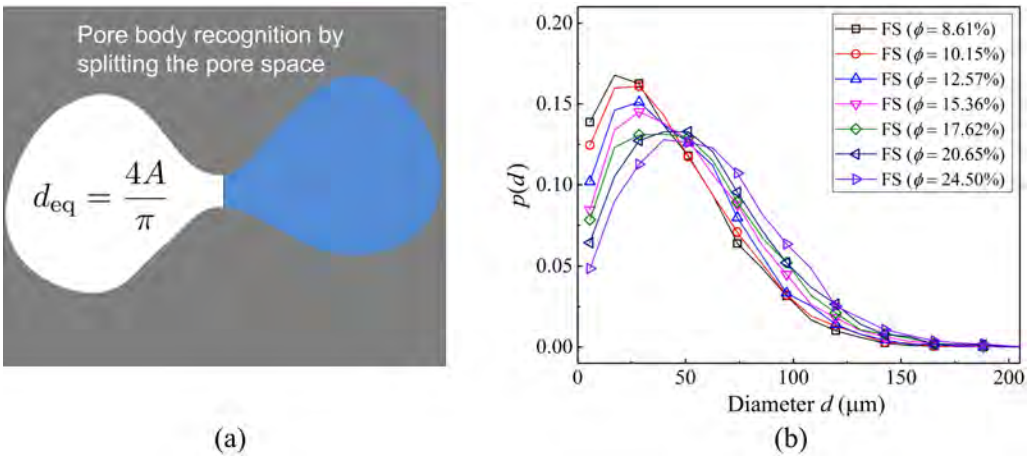


Fig. 2.27 PSD determination by using the discrete method: (a) schematic diagram of pore splitting, and (b) the discrete PSDs of the group of Fontainebleau sandstone samples

concept of "discrete PSD" breaks down. The continuous method treats the pore space as a single continuum, where the PSD at a specific pore diameter d is determined from the amount of pore area (2D) or volume (3D) that can be potentially occupied by the circle or sphere with diameter d [321, 194].

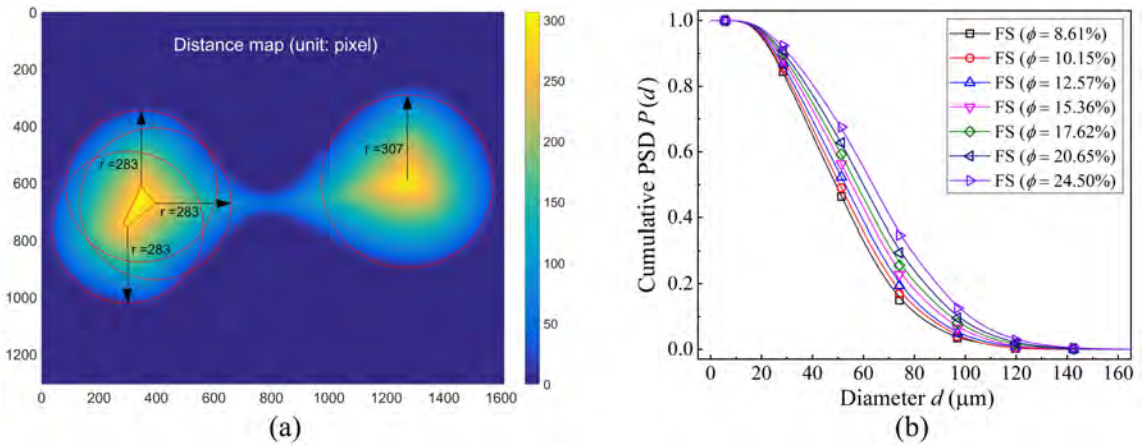


Fig. 2.28 Determination of PSD by using the continuous method: (a) schematic diagram of distance map, and (b) the continuous PSDs of the group of Fontainebleau sandstone samples

In this study, the distance map is firstly calculated from the binary pore mask, which contains the distances between all void pixels/voxels to the pore-solid interfaces (as can be seen in Figure 2.28a). The distance map provides information where the centers of circles or spheres with radius $r = d/2$ should be located without overlapping with the solid matrix. The location set $LS(r)$, where void pixels/voxels are farther than r away from the pore-solid interfaces, can then be determined from the distance map. Subsequently, $LS(r)$ is dilated by

a circular or spherical structuring element with radius r, and the obtained area or volume is the part of pore space that can be completely experienced by the circle or sphere with radius r. By progressively reducing the diameter $d = 2r$, more constricted area or volume such as bottle necks and narrow corners can be occupied by the circle or sphere. The cumulative PSD $P(d)$ can be calculated from the ratio of occupied area or volume to the entire pore space. The cumulative PSDs of the group fo Fontainebleau sandstone samples are provided in Figure 2.28b.

Morphological opening method

PSD can also be determined through a morphological opening process, where successive morphological opening is performed on the pore space with a circular or spherical structuring element (SE) of increasing size [122, 431, 121]. After each opening operation, the resulting pore space only preserves the pores larger than the SE with a specific radius. The opening operation is repeated with increasing SE size until the entire pore space vanishes. The pore volume fraction with a specific radius can be calculated from the difference of pore volumes between two successive opening steps, so as to estimate the PSD. The obtained pore radius is equivalent to the radius of the maximum inscribed circle or sphere inside the pore, as shown in Figure 2.29a.

Actually, this method has been discussed in Section 2.2.1 and 2.2.3 to determine the Minkowski functions and the spherical contact distribution function, respectively. The morphological opening process is illustrated in Figure 2.5. The morphological-opening PSDs of the group of Fontainebleau sandstone samples are shown in Figure 2.29b.

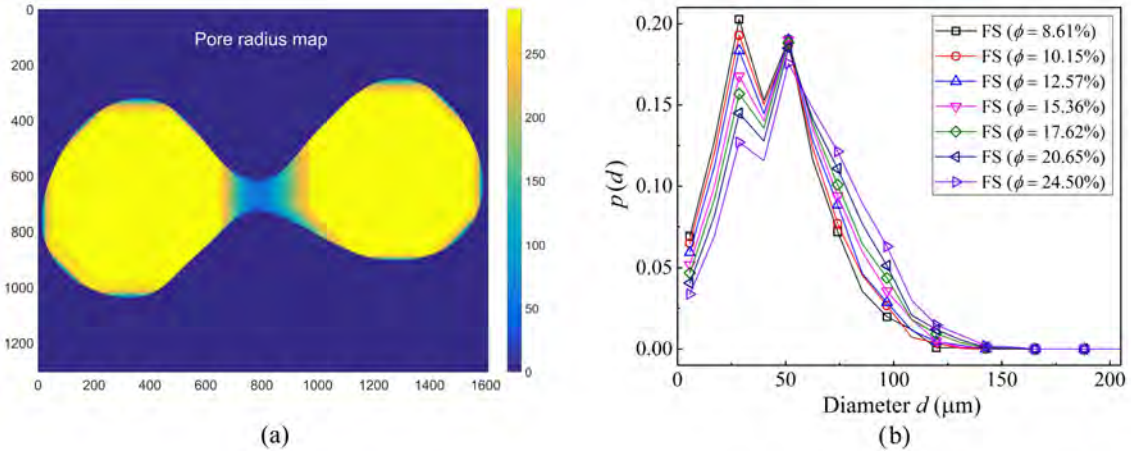
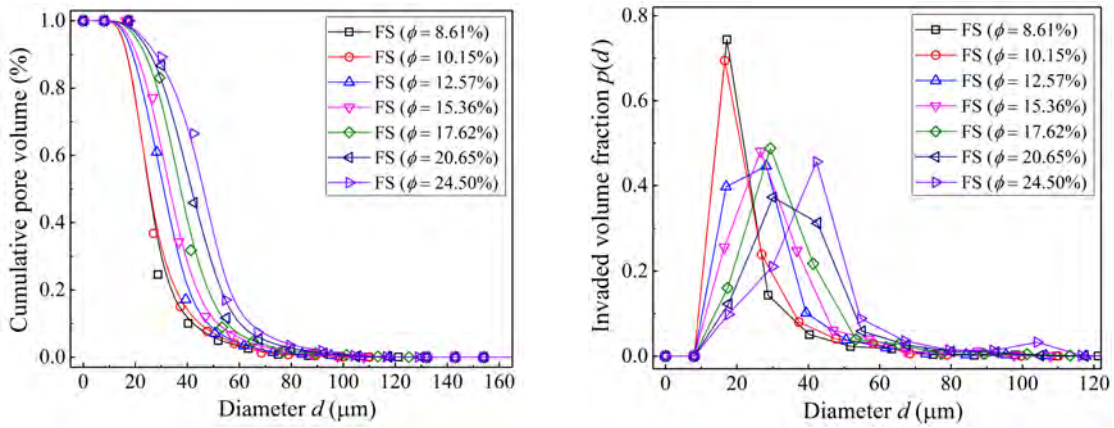


Fig. 2.29 Determination of PSD by using the morphological opening method: (a) schematic diagram of pore radius map, and (b) the morphological-opening PSDs $p(d)$ of the group of Fontainebleau sandstone samples

2.2.6 Pore throat size distribution

A pore throat [281, 239] is defined as a local minimum of the cross-sectional area along the pore channel. Throats locally control the movement of flow and play an important role in determining the bulk transport properties of porous media. The MIP simulation [321, 121, 174] is a popular approach to extracting throat size information from porous media, because the mercury intrusion process is dominated by the effect of constriction/bottleneck [195]. Besides, there are several throat-finding algorithms that have been developed to locate throats in the pore network and estimate the throat size distribution, such as the wedge-based algorithm [407], the Dijkstra-based shortest length algorithm [281], and the planar dilation algorithm [239]. Due to the wide use in characterizing pore structures [121] and modeling transport properties [195], the MIP simulation is emphatically introduced in this work. With the availability of 3D digital microstructures, the mercury intrusion process can be mimicked through a specific region-growing algorithm [146, 321].



(a) Cumulative throat size distributions (b) Throat size distributions

Fig. 2.30 Pore throat size distributions of the group of Fontainebleau sandstone samples evaluated from MIP simulations

The cubic digital microstructure sample is virtually surrounded with mercury by setting active mercury layers separating from the cube surfaces with a fixed distance. The contact angle θ between mercury and solid matrix is assumed to be 180°, so that the intrusion flow from each active mercury location can be represented as a sphere with diameter d. The exact contact angle value is not important here, because it just acts as an overall scale factor in the Washburn equation [465], which does not affect the dependence of intrusion pressure on equivalent pore size. If a sphere centered at an actuation point with diameter d can be fitted in the pore space without overlapping the solid matrix, all void voxels within the spherical space that are still empty will be recognized as the new mercury actuation regions for the

next flow propagation step. The iterative cycle continues until there is no new intrusion space to explore. By executing this simulation scheme with a continuous rang of spherical diameter d, a complete MIP curve can be obtained to represent the relationship between intruded pore volume and spherical diameter d. In essence, this simulation scheme can be considered as a region-growing algorithm, where a new void voxel meeting the growing criterion is permitted to join the intrusion region. Throat size distributions evaluated from MIP simulations for the group of Fontainebleau sandstone samples are shown in Figure 2.30.

Recently, morphological opening has been used for MIP simulation, and the procedure is similar to the determination of morphological-opening PSD, as described in section 2.2.5. The MIP simulation needs to check the pore connectivity after each image erosion step and to remove the eroded pore space that is inaccessible to the active mercury surface [121]. MIP simulation results of the group of Fontainebleau sandstone samples are given in Figure 2.31.

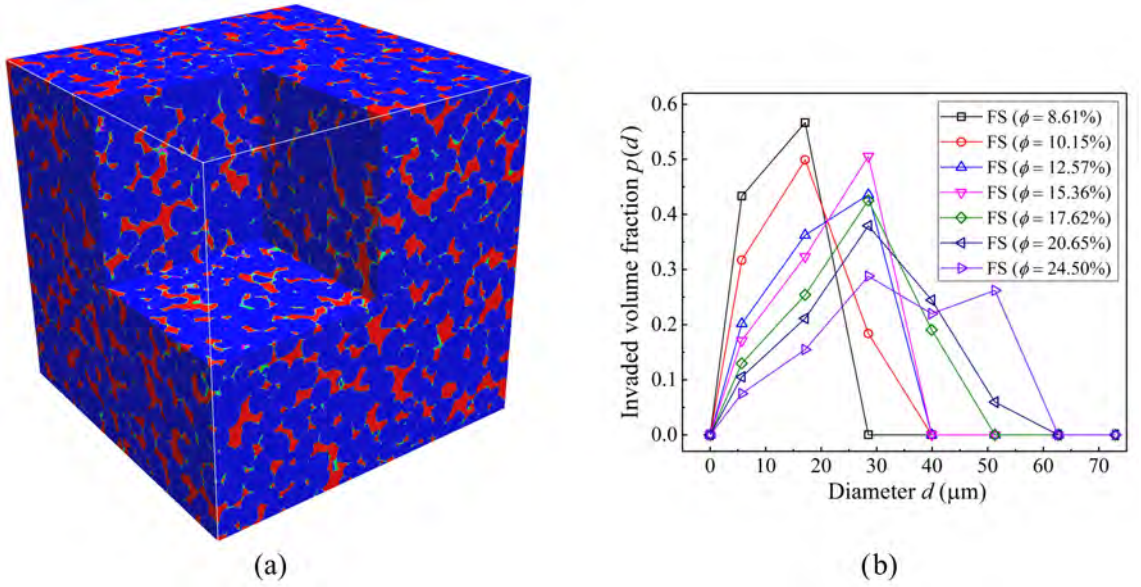


Fig. 2.31 The MIP simulations (morphological opening method) performed on the group of Fontainebleau sandstone samples: (a) Visualization of differentiated pores after mercury intrusion, when the diameter of spherical structuring element is set to be 28.5 μm (the mercury-intruded pores in red, the empty pore in green and the solid matrix is in blue); (b) Pore throat size distributions

2.2.7 Tortuosity

Porous media, whether natural or artificial, are with highly chaotic pore space and especially complex microstructures. The internal pore channels are usually tortuous and meandering, along which the cross-sectional areas are also uneven and varying. The flow paths of fluid

permeation, molecular diffusion, current transmission, and heat conduction are shaped by the pore channels, where the presence of solid matrix in porous media greatly hinders these transport processes. Tortuosity is a dominant microstructural measurement to characterize the sinuosity and complexity of internal percolation paths, and it is also used to quantify retarding effects of pore microstructures on different transport processes [90, 26, 153].

The tortuosity concept varies depending on the specific applications, but in general, it can be understood from two interrelated perspectives: (1) a geometrical measure to characterize the sinuosity of pore channels (flow paths) inside the porous medium; (2) a physical characteristic to quantify the retarding effect (resistance) of the porous microstructure on transport processes through the porous medium. The nature of tortuosity has not been completely understood yet, and it is very confusing whether tortuosity is an intrinsic microstructural property of the porous medium, a physical characteristic depending on the type of transport process taking place inside the porous medium, or just a 'fudge factor' to match predictions with experimental results [153].

Since the concept of tortuosity can be understood from physical and geometrical perspectives, we thus broadly divide tortuosity into two major categories: physical tortuosity and geometrical tortuosity, depending on whether the analysis object is the transport flow through the pore space or the pore space itself. The physical tortuosity can be further divided into hydraulic, electrical, diffusional, and thermal tortuosity, according to the specific transport process occurring inside porous media [90].

Physical tortuosity

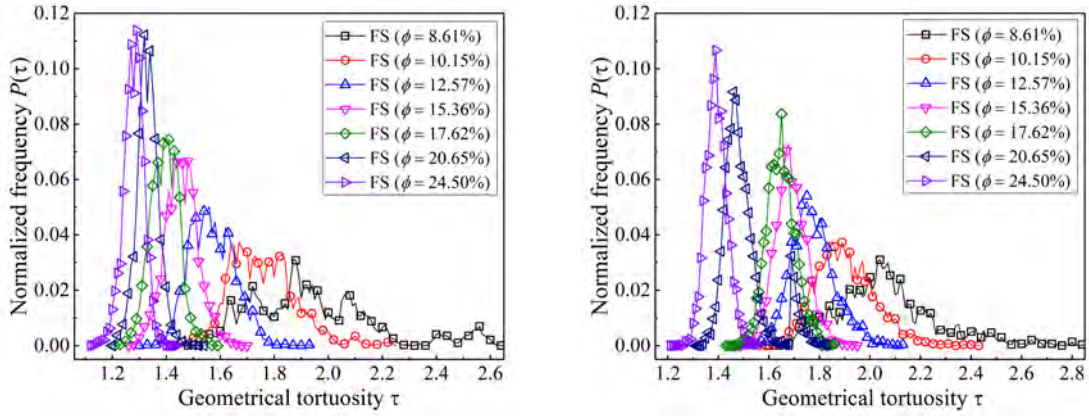
Given the digital microstructures of porous media, the transport processes including fluid permeation, molecular diffusion, electrical conduction and heat transfer can be simulated using various numerical methods to compute the corresponding physical tortuosities. More details will be provided in Chapter 3.

Geometrical tortuosity

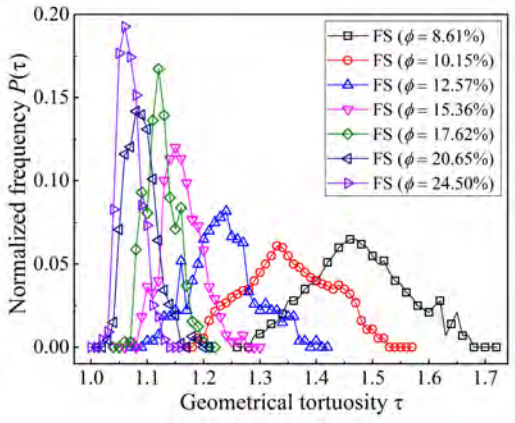
Geometrical tortuosity is defined as the ratio of the effective length of flow paths to the straight-line distance in the macroscopic flow direction [90, 153], to describe the degree of streamlines deviating from the straight line, which can be evaluated through image analysis. More details will be given in Chapter 3.

Geometrical tortuosity distribution

Geometrical tortuosity is a single-valued factor greater than 1, which is calculated as the average ratio of the shortest pore channel length $L_{\min}$ to the corresponding straight-line distance L. The identified shortest pore channel connecting two opposite faces can be different, when the starting or ending point changes. Therefore, it is natural to define the geometrical tortuosity distribution to characterize the length variation of the connected pore channel inside porous media [280, 435, 223, 74, 80]. Compared to one single tortuosity value, the geometrical tortuosity distribution contains more information relevant to microstructural properties, which can be used to correlate microstructural characteristics to transport properties.



(a) DSPSM (b) SSPSM



(c) FMM

Fig. 2.32 Geometrical tortuosity distributions of the group of Fontainebleau sandstone samples

The direct shortest path searching method (DSPSM) [74], skeleton shortest searching method (SSPSM) [280] and fast marching method (FMM) [223] are used to extract the geometrical tortuosity distribution, and the results of the group of Fontainebleau sandstones

are shown in Figure 2.32. The geometrical tortuosity distribution results calculated by DSPSM, SSPSM and FMM are inconsistent, although they show similar trends varying with porosity. More details about these three methods can be found in Chapter 3.

2.2.8 Constrictivity

The constriction and expansion along pore channels (as shown in Figure 2.33) converge and diverge the fluid streamlines, which will lead to variation in fluid velocity, thereby hindering the transport process [34]. As discussed above, tortuosity does not account for the variation in cross-sectional area along the pore channels [153]. To quantitatively characterize the transport resistance called 'bottleneck effect', constrictivity δ was therefore introduced by researchers [195, 34, 422]:

$$\frac{\sigma_{\text{eff}}}{\sigma_0} = \frac{\phi\delta}{\tau} \tag{2.51}$$

where σ_0 is the electrical conductivity of a fluid, and σ_{eff} is the effective electrical conductivity measured on the test porous media sample. The influences of constrictivity on transport processes are usually considered together with the effects of tortuosity τ.

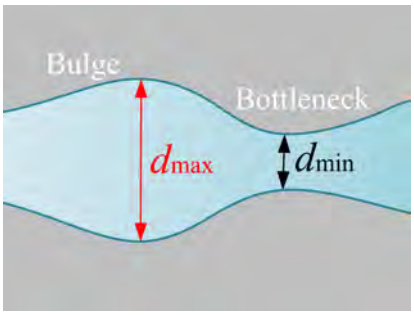


Fig. 2.33 Sketch of the non-constricted bulge and bottlenecks along the pore channel

Constrictivity δ is a dimensionless parameter whose value lies in between 0 and 1, and it characterizes the transport resistance that is inverse proportional to the width of bottleneck [195, 34]. However, there is no rigorous and unique definition for constrictivity.

Petersen (1958) [354] is one of the pioneers to study the effect of constriction on diffusion, and he proposed a simple mathematical expression of constriction factor β_{Petersen} for a single cylindrical pore channel with varying cross-section areas, given by:

$$\beta_{\text{Petersen}} = \frac{A_{\max}}{A_{\min}} \tag{2.52}$$

where $A_{\max}$ and $A_{\min}$ are the maximum and minimum cross-section area respectively along the pore channel.

Latter, Currie (1960) [106] introduced a factor f to represent the contrition effect on diffusion through a tube with non-uniform cross section:

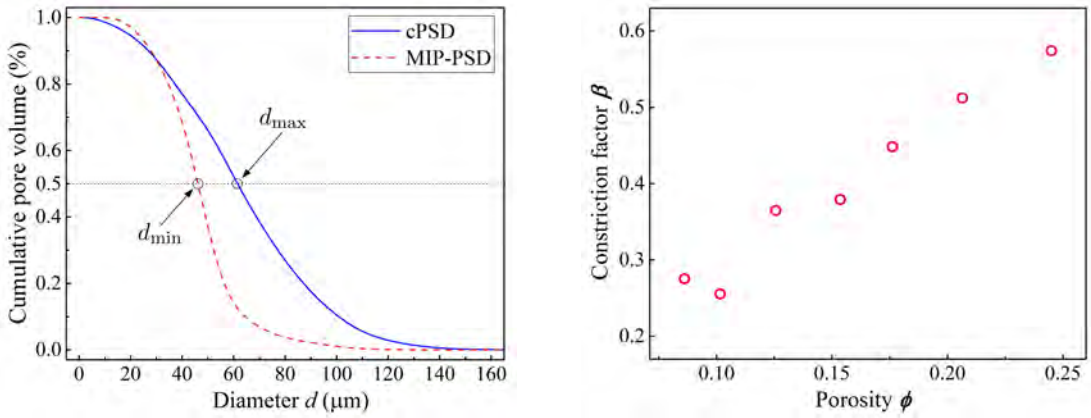
$$f = \frac{\sqrt{A_{\max} A_{\min}}}{A_{\text{mean}}} \tag{2.53}$$

where A_{mean} is the average cross-section area along the cube. Although β_{Petersen} and f are well defined in geometric perspective, they are limited to ideal geometry of single pore capillary.

To quantify the constriction factor of real porous media with complicated geometry, a proper methodology is required to measure the cross-sectional areas at the constrictions and the bulges, as shown in Figure 2.33. Holzer et al. (2013) [195] used the inverse of Petersen's constriction factor to characterize the bottleneck effect, whose value is in between 0 and 1, given by:

$$\beta = \frac{A_{\min}}{A_{\max}} = \left(\frac{d_{\min}}{d_{\max}}\right)^2 \tag{2.54}$$

where $d_{\max}$ and $d_{\min}$ are the diameters of pore cross sections at the non-constricted 'bulges' and the bottlenecks respectively along pore channels.



(a) (b)

Fig. 2.34 (a) Schematic illustration of constriction factor β determined from cumulative pore size distributions; (b) Constriction factors β of the Fontainebleau sandstone samples

The values of $d_{\max}$ and $d_{\min}$ can be determined from PSD of pore media with complicated geometries. As discussed in Section 2.2.5, the PSD results obtained from diverse methods are inconsistent, due to the conceptual differences across them. Basically, the PSD obtained from MIP simulation (MIP-PSD) is controlled the pore constriction, while continuous PSD (cPSD) is insensitive to the bottleneck effect [321]. Therefore, cPSD and MIP-PSD were adopted by [195], from which $d_{\max}$ and $d_{\min}$ can be extracted respectively, to quantitatively represent the

dimensions of bottlenecks and non-constricted bulges for complicated microstructures. The 50% volume fractiles (d_{50}) of the cPSD and MIP-PSD curves are considered as the measures of $d_{\max}$ and $d_{\min}$ respectively, as illustrated in Figure 2.34a. And then, constriction factor β can be determined according to Eqn. (2.54).

As a microstructural parameter, constriction factor is quantified from measurable geometrical attributes (i.e. $d_{\max}$ and $d_{\min}$). Following the above procedure, constriction factor results of the group of Fontainebleau sandstones are calculated and plotted in Figure 2.34b, and the corresponding cPSD and MIP-PSD curves can be seen in Figure 2.28a and 2.30b, respectively.

2.2.9 Pore coordination number

Pore coordination number η [379, 365] represents the number of adjacent pore bodies connected to a specific pore, which is used to characterize the topology of porous media. As a fundamental characteristic of pore networks, the coordinate number has been extensively invoked for studying flow and transport in porous rocks [32, 79, 451].

Generally, there are two different approaches to determining the pore coordination number of a porous medium. In the first approach, the skeleton (medial axis) is extracted from the pore space firstly, and then coordinate number η is calculated as the number of branches of the medial axis meeting at a junction of the pore skeleton [122, 281, 197], as shown in Figure 2.35. The average coordination number $\bar{\eta}$ is an important topological parameter, defined as the average number of branches meeting at one node in the skeleton. In this study, the homotopic thinning algorithm [280] is applied to reduce the pore space to a skeleton of one voxel thickness with the preservation of topological properties. The average coordination numbers of the group of Fontainebleau sandstone samples are given in Figure 2.36a.

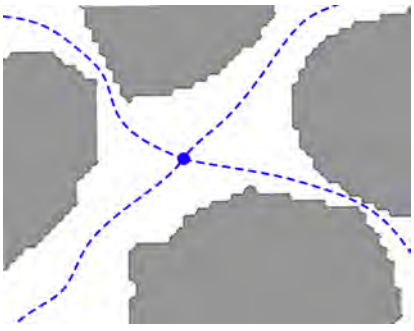
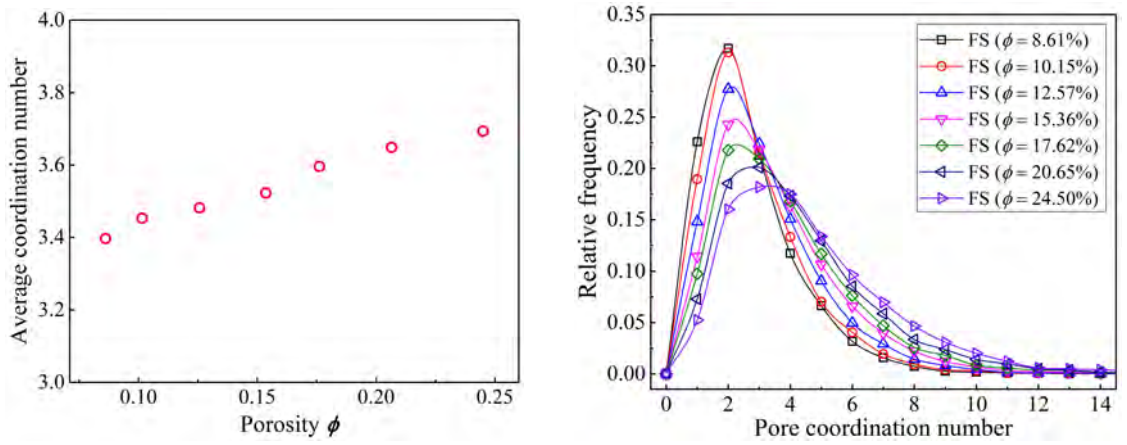


Fig. 2.35 The 2D schematic illustration of a pore with coordination number $\bar{\eta}$=4 (the blue dash lines denote the pore skeleton)



(a) (b)

Fig. 2.36 (a) The average coordinate numbers calculated from the first approach, and (b) the coordination number distributions calculated from the second approach for the group of Fontainebleau sandstone samples.

The second approach compartmentalizes the pore space into individual pore bodies and calculates the coordinate number of a pore as the number of adjacent pore bodies connected to it [197, 365], as shown in Figure 2.37b. In this study, the watershed segmentation algorithm [309] is adopted to partition the pore space, and the pore coordination number distribution (PCND) is then calculated in a histogram form to represent the empirical probability density of pore coordination number. The PCNDs of the group of Fontainebleau sandstone samples are provided in Figure 2.36b.

(a) Original pore space (b) Segmentation of pore space

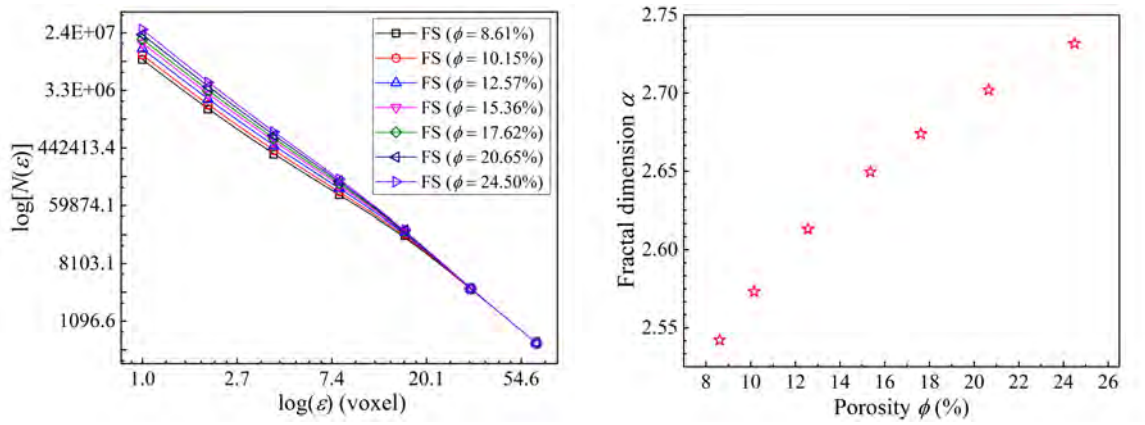
Fig. 2.37 The 2D schematic illustration of determination of pore coordination number through pore space segmentation

2.2.10 Fractal characterization

Fractal geometry theory [131] has been broadly applied in the microstructural analysis of natural porous media. There are several fractal parameters such as fractal dimension, lacunarity and succolarity, and these fractal parameters are complementary to characterize complicated pore microstructures [327, 12, 485]. These fractal parameters of pore geometry extracted from digital images have also been employed to relate microstructural characteristics to transport properties[2, 485], and to stochastically reconstruct microstructure samples [224].

Fractal dimension

Fractal dimension [232, 12, 229, 485], as the most conventional fractal parameter, is a measurement of scaling irregularity and complexity of pore microstructures, which has been widely used in microstructure characterization and macroscopic property modeling for reservoir rocks. Given the digital microstructure of a porous medium, the fractal dimension can be efficiently estimated from the image through the popular box-counting method [279, 271].



(a) (b)

Fig. 2.38 Determination of fractal dimension for the Fontainebleau sandstone samples via the boxing-counting method: (a) the $\log[n(\varepsilon)]$-$\log(\varepsilon)$ graphs and (b) fractal dimension results α

In the box-counting method, the fraction dimension is measured from a binary image, and the image is subdivided into a series of boxes of side-length ε. The number $n(\varepsilon)$ of required boxes to completely cover the void phase in the image is counted. Reducing or increasing the box size ε and recounting the number $n(\varepsilon)$ of required boxes by repeating the above procedures, the relationship between $\log[n(\varepsilon)]$ and $\log(\varepsilon)$ is created. The $\log[n(\varepsilon)]$-$\log(\varepsilon)$ plot is used to determine the fractal dimension from the slope of the regression line, as shown

in Figure 2.38a. The box-counting fractal dimension α [131] is defined as:

$$\alpha = \lim_{\varepsilon \to 0} \frac{\log[n(\varepsilon)]}{\log(1/\varepsilon)} \tag{2.55}$$

The principles of the box-counting method to extract the fractal dimension from 2D and 3D microstructures are almost identical. Here, a cube-counting method is used to estimate the fractal dimensions of the group of Fontainebleau sandstone samples, and the results are given in Figure 2.38b.

Lacunarity

Lacunarity is a relatively new fractal parameter to quantitatively measure how clustered the pore microstructure is, which reflects the degree of homogeneity or heterogeneity or translational/rotational invariance of the porous medium [327, 12]. Generally, a low lacunarity value means homogeneity of pore microstructures, while a high lacunarity value indicates heterogeneity. The porous microstructures with the same fractal dimension can be distinguished by the lacunarity. It has been applied to represent the pore structures and evaluate the transport properties of reservoirs rocks [148, 485], soils [85] and other porous media [240, 104].

The box counting method is the simplest method to calculate lacunarity [8, 12], where a box of side-length ε is glided over the binary image in a non-overlapping or an overlapping manner (also called sliding box counting method). The number M of pore pixels in each box of size ε is recorded during the box sliding process, and the number of boxes contain M pore pixels is counted as well, marked as $n(M,\varepsilon)$. The occurrence frequency of the boxes of side-length ε containing M pore pixels is calculated to represent probability density function $Q(M,\varepsilon)$, which is calculated as follows:

$$Q(M,\varepsilon) = \frac{n(M,\varepsilon)}{N(\varepsilon)} \tag{2.56}$$

where $N(\varepsilon)$ is the total number of boxes of size ε.

The lacunarity $\delta(\varepsilon)$ at a specific box size ε is defined as the ratio between the second moment and the square of the first moment of $Q(M,\varepsilon)$, given by:

$$\delta(\varepsilon) = \frac{\sum M^2 Q(M,\varepsilon)}{[\sum M Q(M,\varepsilon)]^2} \tag{2.57}$$

Gradually increasing or reducing the box size, lacunarity $\delta(\varepsilon)$ as a function of ε can be obtained. Besides, lacunarity $\delta(\varepsilon)$ satisfies properties as follows:

$$\begin{aligned} \lim_{\varepsilon \to 0} \delta(\varepsilon) &= \frac{1}{\phi} \\ \lim_{\varepsilon \to \infty} \delta(\varepsilon) &= 1 \end{aligned} \tag{2.58}$$

where ϕ is the bulk porosity the porous medium.

Similar to the sliding box-counting method described above, the lacunarity of a 3D image can be calculated by using a cubic box of side-length ε through the same procedure. The lacunarity results of the group of Fontainebleau sandstone samples are provided in Figure 2.39.

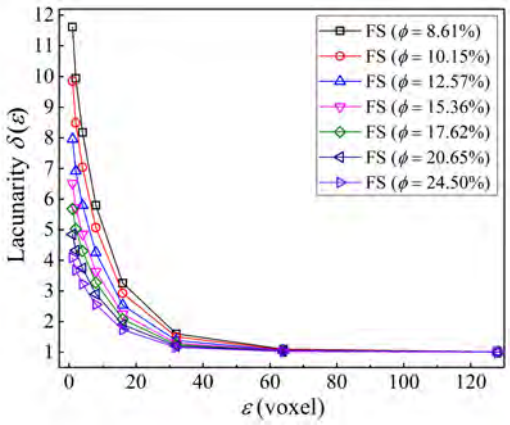


Fig. 2.39 The lacunarity $\delta(\varepsilon)$ results of the group of Fontainebleau sandstone samples

Succolarity

Succolarity is a fractal parameter containing connectivity and intercommunication information, which can be used to characterize the percolation degree that allows fluid to flow through the pore microstructure in a direction [112, 113, 327, 12, 485]. The calculation approach of succolarity was put forward by De Melo and Conci (2008, 2013) [112, 113], which is an ideal box-counting method contains four steps. Succolarity of a 2D image is calculated in different directions: from left to right (x^+), from right to left (x^-), from bottom to top (y^+) and from top to bottom (y^-).

In the first step, a virtual pressure field is applied to the 2D microstructure in x^+ direction (or other directions), and a virtual fluid starting from the most-left column of pore pixels is driven by the pressure field to flow into the microstructure. All of the pore pixels connected

to the most-left column are detected according to the 4-connectivity criterion, and these selected pixels form the penetrable area for the next set of succolarity calculation.

In the second step, the 2D image that only contains pore pixels belonging to the penetrable area is divided into boxes of size ε. The occupation percentage (OP) of pore pixels is measured for each box, given by:

$$OP(i,j) = \frac{NP(i,j)}{\varepsilon^2} \tag{2.59}$$

where i and j are the coordinates of the centroid of a box, and $NP(i,j)$ is the number of pore pixels in the box. The average pressure (AP) applied to each box in x^+ direction is related to the location of its centroid, given by:

$$AP(i,j) = j - 0.5 \tag{2.60}$$

Finally, the box-counting succolarity $\psi(\varepsilon, x^+)$ at specific box size ε and in the x^+ direction is defined as:

$$\psi(\varepsilon, x^+) = \frac{\sum[OP(i,j) \times AP(i,j)]}{\sum AP(i,j)} \tag{2.61}$$

Gradually increasing or reducing the box size, the succolarity $\psi(\varepsilon)$ as function ε of can be acquired. And the succolarities in other directions can be calculated in the same procedure, and they have been used to measure the connectivities of the microstructure in different directions.

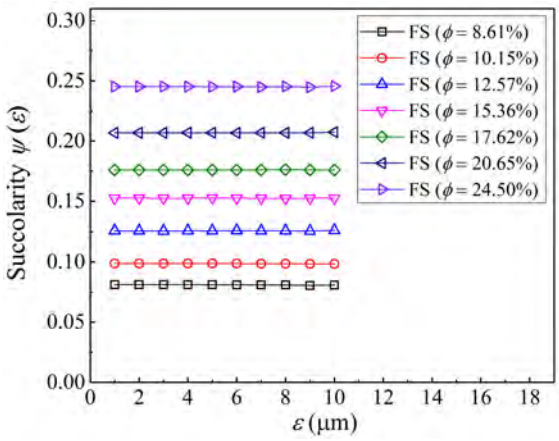


Fig. 2.40 The succolarity $\psi(\varepsilon)$ results of the group of Fontainebleau sandstone samples

The calculation procedure of 3D succolarity is similar to that of 2D succolarity, where the 3D microstructure is applied with the virtual pressure filed in 6 directions and divided into cubic boxes [113]. The average succolarities of the group of Fontainebleau sandstone

samples over 6 directions are given in Figure 2.40, and it seems that succolarity is invariant to the cubic box size for these samples.

2.3 Concluding remarks

The main contribution of this chapter is to fundamentally investigate the statistical characterization of pore microstructures, where a broad variety of morphological descriptors have been collected from an extensive literature study. The morphological descriptors are statistical descriptions of microstructural complexity within porous media, which are essential to deeply understand microstructure-property relationships, to reveal specific physical phenomena, or to explore the transport behaviors of fluid flow relevant to microstructural characteristics. The comprehensive review of microstructure characterization in this chapter provides a sophisticated understanding of the limitations and opportunities of morphological descriptors in porous media research.

The commonly used morphological descriptors are introduced in terms of theoretical principles and determination methods, and they are roughly categorized into four levels to provide microstructural informatics from global, local, geometrical and topological perspectives. Porosity and specific surface area are the typical descriptors at the first level, which simply represent the global microstructural properties of porous media via single numbers. However, the detailed morphology of pore systems is not captured by them. When it comes to the second level, local microstructural characteristics are measured by descriptors such as local porosity distribution, coarseness and local percolation probability. As to the third level, geometric attributes of porous media are quantified in different aspects such as shape and size, and the frequently-used descriptors include pore/throat size distribution function, mean curvature, spatial correlations functions, surface roughness, contact distribution functions and others. The fourth level focuses on the topology of porous media, which is related to long-distance connectivity and percolation of the pore network. Integrate of total curvature (Euler characteristic), pore connectivity function and total fraction of percolation cells are commonly-used indexes of connectivity. Geometrical tortuosity reflects the sinuosity degree of percolation paths inside porous media, while constriction factor quantitatively represents the cross-section variation along pore channels.

It should be noted that this chapter lays a solid foundation for subsequent studies in this thesis, including stochastic microstructure reconstruction (Chapter 4 and 5), error correction modeling due to the resolution effect (Chapter 6), and the fundamental investigation of the microstructure-permeability relationship (Chapter 7).

Chapter 3

Tortuosity of porous media: Image analysis and physical simulation

3.1 Introduction

Porous media, whether natural or artificial, comprise chaotic microstructures so that the internal flow path for fluid permeation, molecular diffusion, electrical conduction, and heat transfer are tortuous and meandering. In order to quantify the transport behavior of flow passing through porous media, an accurate understanding of the heterogeneous microstructure is required [48, 474, 64]. Such microstructural descriptors as porosity [246], specific surface area [355], correlation functions [92, 186], pore size distribution [321], Minkowski functionals [460, 18], constrictivity [195, 236] and tortuosity [34, 490] play vital roles in characterizing the morphology of porous media. Among them, tortuosity is a special parameter, which is not only used to explicitly describe the sinuosity and complexity of internal percolation paths but also to quantify transport properties of porous media [90, 26, 153].

Tortuosity is an important concept in many scientific and engineering fields. In geoscience, tortuosity is closely related to the transport behavior of fluid permeation, molecular diffusion, electrical conduction, and heat conduction, thereby affecting such petrophysical properties as permeability, effective diffusivity, formation resistivity factor, and thermal conductivity [300, 23, 420]. In energy storage and conversion, tortuosity is a key factor for electrodes design to maximize battery capacity [235, 24, 497]. For chemical catalysts, tortuosity has a great influence on reaction efficiency [361, 245]. In acoustic applications, tortuosity is closely linked to sound propagation and velocity variation in pore space [99, 501]. For water treatment, the tortuosity of the membrane microstructure significantly influences the

distillation performance [228]. In bone tissue engineering, tortuosity affects cell migration through the scaffold, nutrient diffusion, waste removal, and bone remodeling [54, 147, 41].

Despite this broad presence in diverse fields, tortuosity is not consistently defined by geologists, engineers, chemists and biologists. Extensive efforts have been made to determine the tortuosity of porous media, but the evaluation methods and results differ between various types of tortuosity. Also, different types of tortuosity are often interchangeably used in literature [404, 153, 438]. In the following subsections, different approaches to determining the tortuosity of porous media are briefly reviewed, including experimental measurement, theoretical/phenomenological modeling, image analysis and physical simulation.

3.1.1 Experimental measurement

To model the permeability of porous media, the concept of tortuosity was first introduced by Kozeny (1927) [250] based on a capillary tube experiment and then further refined by Carman (1937) [71]. Tortuosity was initially only used as a tuning parameter to account for the effect of complicated pore structures on fluid flows in porous media [26, 51]. The presence of the solid phase in porous media causes the flow path to deviate from the straight line; and to approximate the effective flow path, the length scale of the pore channel is adjusted by tortuosity. Besides fluid permeation, the tortuosity concept has been applied to other transport behaviors of porous media as well, including molecular diffusion, electrical conduction and heat transfer.

There are two widely-used experimental approaches to measuring the tortuosity of porous media. The first approach is based on the electrical conductivity experiment, where the porous test sample is saturated with an electrolyte and the tortuosity is evaluated using the effective conductivity measured on the porous sample [149, 467, 26]. The second approach measures the diffusion coefficient of a nonreactive species diffusing through the porous media sample, and the tortuosity is estimated using the measured diffusivity [26, 467, 176]. Compared to the electrical conductivity experiment, the diffusion process in porous media is much slower, and the corresponding experiment is sensitive to the selected chemical species and temperature. Typically, the tortuosity values measured from electrical conductivity and diffusion experiments are different [149]. Therefore, different tortuosities are defined depending on the underlying physical process and the measurement protocol [490].

3.1.2 Theoretical/phenomenological models

A large number of theoretical or phenomenological models have been proposed to estimate the tortuosity of porous media, and most of them simply express tortuosity as functions

of porosity [316, 261, 124, 4]. The Bruggeman correlation [61, 119] is one of the most commonly used porosity-tortuosity relationships, which is expressed as an exponential equation, given by

$$\tau^2 = \phi^{1-\alpha} \tag{3.1}$$

where τ denotes the tortuosity, ϕ is the porosity and α is an empirical constant called Bruggeman exponent. Considering the relationship between porosity ϕ and formation resistivity factor F, i.e. $\tau^2 = \phi F$ [483, 93], the Bruggeman relation is formally consistent with the well-known Archie's law [16]: $F = \phi^{-\alpha}$. Another famous model was presented by [94], which treats tortuosity as a logarithmic function of porosity:

$$\tau^2 = 1 - p\ln\phi \tag{3.2}$$

where p is an empirical constant related to the grain shape. The porosity-tortuosity model put forward by Iversen and Jørgensen [210] is also popular, where the linear relationship between porosity and tortuosity was observed as follows:

$$\tau^2 = 1 + q(1-\phi) \tag{3.3}$$

where q is an empirical constant depending on the type of porous media. Du Plessis and Masliyah [123] derived an analytical model from an idealized granular pore microstructure without involving any undetermined parameter, which is mathematically expressed as

$$\tau = \frac{\phi}{1-(1-\phi)^{2/3}} \tag{3.4}$$

Besides, another representative theoretical model was derived from a fixed bed of randomly packed identical particles by Lanfrey et al. [261], given by

$$\tau = 1.23\frac{(1-\phi)^{4/3}}{\xi^2\phi} \tag{3.5}$$

where ξ is the shape factor (sphericity) of particle, $\xi = 1$ for sphere and $\xi < 1$ for non-spherical particles.

The above five porosity-tortuosity relations are representatives of various models for tortuosity estimations. Despite the simple formulations, phenomenological models usually contain empirical parameters or uncertain coefficients such as α, p, q and ξ in above equations, and accurate determinations of them are difficult. Even for a small class of porous media, these adjustable parameters can vary significantly causing great errors for the

tortuosity estimation [153, 438]. To improve the prediction accuracy, tortuosity models with more sophisticated formulation and more parameters have been proposed, but typically they are only suitable for a narrow range of microstructures [358, 301, 261, 86]. Ghanbarian et al. [153] presented a critical review of tortuosity models including geometrical, hydraulic, electrical and diffusional tortuosities, and it was found that various porosity-tortuosity models differ greatly and may not be used interchangeably. Moreover, although many actual porous media are anisotropic in nature, most theoretical and phenomenological models assume the medium to be isotropic.

Microstructural features such as the shape, size, orientation and spatial distribution of grains and pores all affect the tortuosity of porous media [453], but most phenomenological models rely mainly on the porosity [404, 153]. A porous medium contains a large number of geometrically irregular pore bodies, some are connected, and others are isolated or dead-end. The length scales of randomly distributed pores can range over several orders of magnitude. Besides, the solid-pore interfaces are typically rough with a significant resistance effect on transport flow. These complications make it extremely challenging, if not impossible, to develop an explicit mathematical model that can accurately predict the tortuosity for general porous media. Indeed, the tortuosity values calculated from various phenomenological models often differ greatly from experimental measurements or numerical simulation results and are even incorrect in some cases [404, 301, 261, 438].

3.1.3 Image analysis

As illustrated in Figure 3.1, modern microscopy imaging techniques are capable of providing non-destructive visualization and characterization of 2D/3D porous microstructures at scales of nanometers to millimeters [259]. Raw microscopic images are typically processed and stored as pixels or voxels in grey scale (as shown in Figure 3.19a), and through image segmentation they can be converted into binary forms, where the pore space is separated from the solid matrix (as shown Figure 3.19c). The segmented digital microstructure with sufficient size and adequate resolution can be directly used as a representative elementary volume (REV) for further analyses of geometry measurement and physical simulation.

With the increasing availability of digital microstructures, the effects of complex microstructures on the macroscopic transport properties have attracted significant attention in recent years [48, 474, 64]. Owing to the high efficiency, it is desirable to directly evaluate the tortuosity through image analysis. Thus, to describe the degree of flow paths deviating from straight lines in porous media, geometrical tortuosity is defined as the ratio of the effective flow-path length to the straight-line distance in the macroscopic flow direction [153]. Geometrical tortuosity is usually computed by considering the shortest pore channels as the

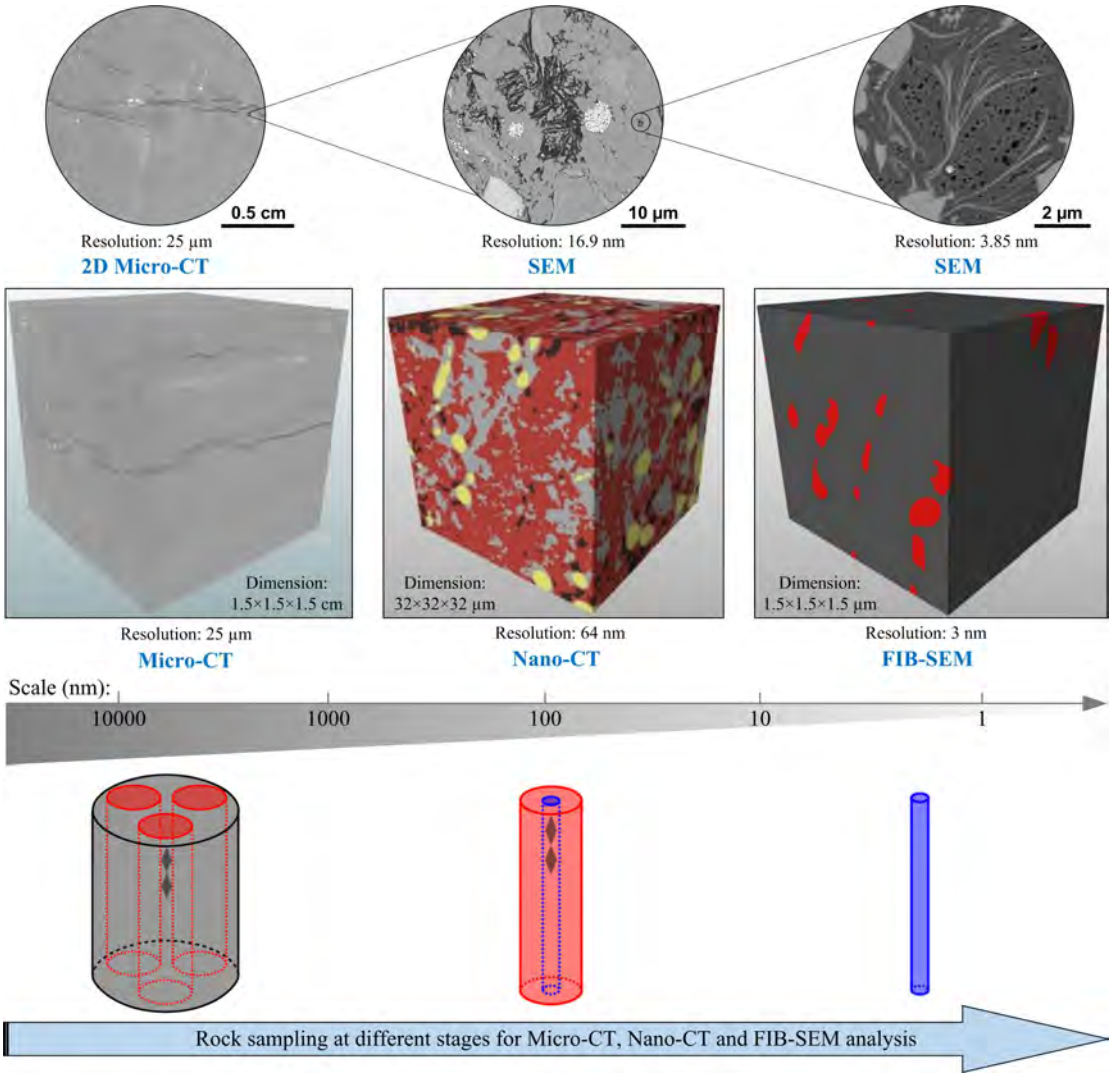


Fig. 3.1 Multiscale microstructure characterization of a shale sample by using different microscopy imaging techniques: The 2D images on the first line provide microstructural information at three length scales, where micro-fractures, mineralogical components and organic pores are identified; The 3D images in the second row visualize internal structure at three length scales, where micro-fractures, mineralogical components and macro/meso-pores are distinguished; The images at the bottom are graphic illustrations about rock sample preparation for microscopy imaging at different length scales (the raw picture materials are collected from [332] and [117])

actual flow paths through porous media [428, 75]. Various algorithms, such as Dijkstra's algorithm and fast marching algorithm, have been used to search for the shortest pore channels that can be equivalent to the effective flow paths. Some popular methods to calculate geometrical tortuosity are listed in Table 3.1, including the direct shortest path searching

method (DSPSM), the skeleton shortest path searching method (SSPSM), the fast marching method (FMM) and the pore centroid method (PCM), etc.

Table 3.1 Image analysis methods to calculate geometrical tortuosity of porous media

Evaluation method	Porous media	Image dimension	Representative reference
DSPSM (Dijkstra's algorithm)	Porous rocks, nanocomposites, and fuel cell electrodes.	3D	[159], [75], [464], [282] and [422]
DSPSM (Wavefront propagation algorithm)	Structures of sodium chloride compacts.	2D	[381]
DSPSM (A-star algorithm)	Structures of sintered compacts.	2D	[313]
SSPSM (Dijkstra's algorithm)	Porous rocks, fuel cell electrodes, alumina, and paper microstructures.	3D	[280], [21], [428], [402] and [6]
FMM	Porous rocks, fuel cell electrodes, infiltrated aluminium, cellular materials, and energy storage materials.	3D	[62], [223], [347], [80] and [431]
PCM	Porous rocks, soil and, fuel cell electrodes.	3D	[162], [116], [403] and [225]
Path tracking method	Artificial models.	3D	[415]

It should be noted that geometrical tortuosity is a microstructural characteristic, independent from the specific transport process in porous media [90]. Moreover, although constrictions and bottlenecks of pore channels can greatly affect the transport behaviors [195, 34], geometrical tortuosity only considers the longitudinal distance of possible flow paths and neglects the flow variation in the cross-section of pore channels. Also, the flow paths identified by image analysis can differ greatly from the real flow paths in porous media [153]. These limitations may explain why geometrical tortuosity often fails to accurately model/predict such transport properties of porous media as permeability, effective diffusivity and conductivity.

3.1.4 Physical simulation

High-quality digital microstructures also make it possible to mimic experimental measurements with high fidelity numerical simulations, so that tortuosities can be evaluated from physical simulations of fluid permeation, electrical conduction, molecular diffusion and heat transfer in porous media [124, 98, 438]. These flux-based tortuosities (namely the hydraulic,

Table 3.2 Numerical simulation methods for the prediction of physical tortuosities

Tortuosity type	Numerical method	Porous media	Image dimension	Representative reference
Hydraulic	Lattice gas automata (The precursor to LBM)	Artificial models.	3D	[498]
Hydraulic	LBM	Porous rocks and artificial models.	2D and 3D	[247], [301], [124], [302] and [114]
Hydraulic	FEM	Artificial models.	2D	[383]
Hydraulic	FVM	Artificial models.	2D	[295]
Electrical	Lattice gas automata (The precursor to LBM)	Artificial models.	3D	[498]
Electrical	FEM	Artificial models.	2D	[383]
Thermal	FVM	Fuel cell electrodes.	3D	[98], [440] and [60]
Thermal	FEM	Fuel cell electrodes.	3D	[426]
Diffusional	FVM	Porous rocks and battery cathode.	3D	[205] and [353]
Diffusional	FEM	Fuel cell electrodes.	3D	[476] and [80]
Diffusional	FDM	Porous rocks and fuel cell electrodes.	3D	[358], [96], [137] and [23]
Diffusional	LBM	Soils, and fuel cell electrodes.	3D	[467], [211] and [342]
Diffusional	Gauss–Seidel method	Porous alumina.	3D	[402]
Diffusional	RWM	Porous rocks, fuel cell electrodes, concretes and bead packs.	2D and 3D	[326], [325], [211], [241] and [447]

electrical, diffusional and thermal tortuosity) are collectively referred to as physical tortuosities in this paper. Table 3.2 summarizes the popular numerical methods for pore-scale simulations of various transport processes in porous media to evaluate physical tortuosities, including the lattice Boltzmann method (LBM), the finite element method (FEM), the finite volume method (FVM), the finite difference method (FDM), the random walk method (RWM) and others.

In contrast to geometrical tortuosity, which is exclusively determined by the porous microstructure, physical tortuosities are related to both the porous microstructure and the transport flow inside it. For a specific transport process in a porous medium, the associated physical tortuosity describes the resistance effect from the solid microstructure. Although transport processes in porous media can be fundamentally different, physical tortuosities

are often interchangeably used in the literature [90]. For example, the electrical tortuosity was considered to be identical to the hydraulic tortuosity [72, 318, 108, 466], and many researchers took the former as an approximation of the latter to model permeability of porous media [111, 110, 26, 383]. The diffusional tortuosity obtained from molecular diffusion simulation or random walk simulation was also used to substitute the hydraulic tortuosity in hydraulic conductivity studies of porous rocks [325, 300, 353, 23]. Numerical simulation of thermal conduction was adopted to evaluate the diffusional tortuosity of porous media [98, 440, 60, 426]. Besides, geometrical tortuosity is commonly used as an alternative of physical tortuosities to model transport properties of porous media [313, 428, 75, 195, 34, 225, 274], such as permeability, effective diffusivity, formation resistivity factor and thermal conductivity.

Many open-source or commercial software packages have been developed for microstructural analysis, as summarized in Table 3.3. These software tools can all provide morphological characteristics or physical properties of porous media, but they adopt different methods (as listed in Table 3.1 and 3.2) to calculate tortuosity, which often leads to inconsistent results.

3.1.5 Motivation and significance

A summary of the origin, evolution and variation of the tortuosity concept was given by Clennell (1997) [90], where the relationships between different types of tortuosity were analyzed. The tortuosity concept varies depending on the specific application, but in general, it can be understood from two interrelated perspectives: (1) a geometrical measure to characterize the sinuosity of pore channels (flow paths) inside the porous medium; (2) a physical characteristic to quantify the retarding effect (resistance) of the heterogeneous microstructure on a specific transport process inside the porous medium. The nature of tortuosity has not been fully understood, so that it remains confusing whether tortuosity is an intrinsic microstructural property of the porous medium, a physical characteristic depending on the type of transport process, or just a "fudge factor" to tune predictions to match with experimental results [153].

A diverse range of approaches have been developed and applied to evaluate tortuosities of porous media, including experimental measurement, theoretical/phenomenological models, image analysis, and physical simulation, but inconsistent results are often obtained from different evaluation methods [149, 404, 153, 347, 438]. There is no rigorous and unique definition for tortuosity as its physical essence has not been completely understood. Other factors, including experimental facility and condition, modeling assumption and data, numerical simulation method and boundary condition, and image preparation and analysis methods, also have significant influence on the tortuosity results. Further research is required to tackle

Table 3.3 Open-source and commercial software packages for the prediction of tortuosities

Name	Tortuosity type	Calculation method	Nature	Reference
iMorph	Geometrical	FMM	Open-source	[62]
Tort3D	Geometrical	DSPSM	Open-source	[75]
TORT3D	Geometrical	SSPSM	Open-source	[6]
Palabos	Hydraulic Diffusional	LBM simulation of laminar flow LBM simulation of molecular diffusion	Open-source	[345]
OpenLB	Hydraulic Diffusional	LBM simulation of laminar flow LBM simulation of molecular diffusion	Open-source	[339]
TauFactor	Diffusional	FDM	Open-source	[96]
PuMA	Diffusional	FDM or RWM	Open-source	[136]
pytrax	Diffusional	RWM	Open-source	[447]
Avizo	Geometrical Hydraulic Diffusional Electrical Thermal	PCM FVM simulation of laminar flow FVM simulation of molecular diffusion FVM simulation of electrical conduction FVM simulation of heat conduction	Commercial	[20]
Simpleware	Geometrical Hydraulic Diffusional Electrical Thermal	SSPSM FEM simulation of laminar flow FEM simulation of molecular diffusion FEM simulation of electrical conduction FEM simulation of heat conduction	Commercial	[409]
GeoDict	Geometrical Hydraulic Diffusional Electrical Thermal	DSPSM or SSPSM FVM simulation of laminar flow FVM simulation of molecular diffusion FVM simulation of electrical conduction FVM simulation of heat conduction	Commercial	[151]

the nature of tortuosity and the relationship between different evaluation approaches, which are critical for understanding the transport behaviors inside porous media.

In this study, we focus on modern approaches to evaluating the tortuosity of porous media, i.e. image analysis and physical simulation. It is noted that the tortuosity τ is a scalar-valued factor greater than 1. Neither the tortuosity factor τ^2 (i.e. the square of tortuosity) [127] nor the tortuosity coefficient τ^{-1} (i.e. the inverse of the tortuosity) [189] is considered in this work. Geometrical and physical tortuosities are systematically reviewed in Section 3.2 and Section 3.3, respectively, to clarify the definition and evaluation method of each type of tortuosity. A comparison study is conducted for the geometrical tortuosity evaluations in Section 3.4.1, where the geometrical tortuosity values obtained from different image analysis methods are found to be inconsistent. Although geometrical tortuosity has a relatively clear definition, several uncertainties exist in its evaluation methods via image analysis. Another comparison study is presented in Section 3.4.2 for numerical simulations

of different transport flows in porous media, to examine the relationship between physical tortuosities. The underlying physics is explored to explain the differences and similarities. Large discrepancies are observed between geometrical and physical tortuosities, which questions strongly the common practice of using the former as the substitute for the latter. To address this problem, empirical correlations between geometrical and physical tortuosities are established in Section 3.5 for a wide range of porous media samples. Finally, the influences from REV size and image resolution on tortuosity evaluation are discussed in Section 3.6, and the correlation length is used to determine the optimal image parameters.

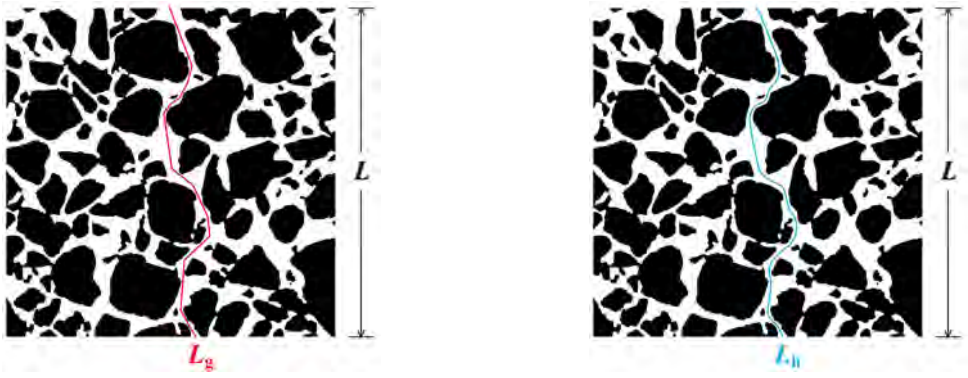
3.2 Geometrical tortuosity

3.2.1 Definition of geometrical tortuosity

To describe the degree of streamlines in porous media deviating from the straight line, geometrical tortuosity is defined as the ratio of the effective length of flow paths to the straight-line distance L in the macroscopic flow direction [90, 153]. As illustrated in Figure 3.2, the transport flow inside porous media is shaped by the pore network, so the shortest pore channel L_g is usually adopted as the effective flow path L_h to compute geometrical tortuosity τ_g [2], given by

$$\tau_g = \frac{L_g}{L} \tag{3.6}$$

Compared with the zigzag line of the shortest pore channel, the effective flow path is much smoother.



(a) (b)

Fig. 3.2 (a) The shortest pore-channel (geodesic) length L_g, (b) the effective flow-path length L_h and the straight-line (Euclidean) distance L in a 2D digital microstructure

Geometrical tortuosity can be considered as a microstructural characteristic as it is purely determined by the geometrical and morphological features of porous media [90]. Various image analysis methods have been developed to evaluate geometrical tortuosity from the digital microstructures of porous media, including the direct shortest-path search method (DSPSM), the skeleton shortest-path search method (SSPSM), the fast marching method (FMM) and the pore centroid method (PCM), etc. These image-based methods directly operate on the pixel/voxel data, which are usually easy to implement and computationally efficient. It is however worth noting that geometrical tortuosity does not account for the variation of cross-section along the pore channels, although the constriction and bottleneck of porous microstructures do play a vital role in transport properties [438].

3.2.2 Direct shortest-path search method

The DSPSM searches for the shortest pore channel directly on the pixel/voxel data by using the shortest-path search algorithms [82, 422], such as Dijkstra's algorithm [428, 75], the wavefront propagation algorithm [482] and the A-star algorithm [313]. In this subsection, a representative DSPSM based on Dijkstra's algorithm is briefly recapped.

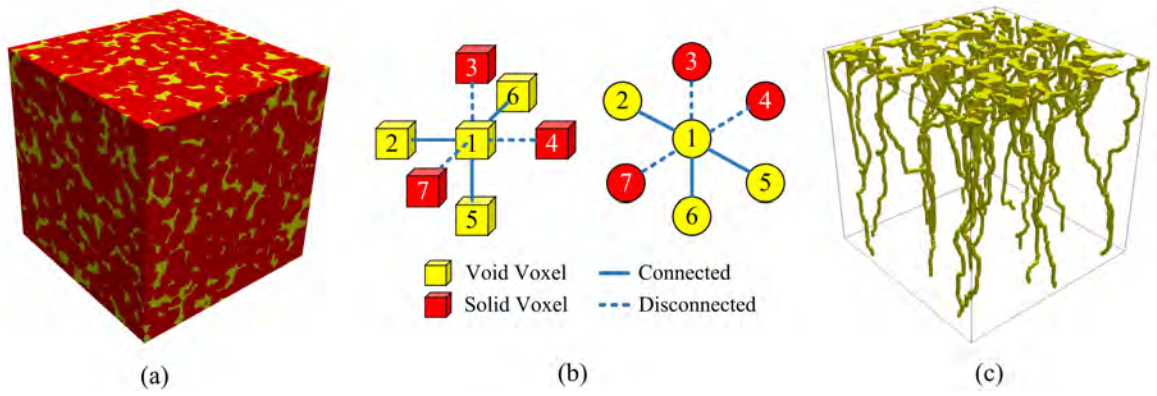


Fig. 3.3 (a) A 3D digital microstructure (yellow denotes the pore space and red is the solid matrix); (b) Conversion of 3D dataset to 2D graph for path search, based on 6-neighbour-connectivity criterion; (c) The shortest pore channels between top and bottom surface determined by using Dijkstra's algorithm

The shortest pore channel is defined as the shortest connected route starting from a surface voxel on one side of the digital microstructure and ending at another surface voxel on the opposite side, and it must always remain in the pore space [75]. Changing the starting or ending voxel will lead to a different shortest path, so geometrical tortuosity τ_g is usually calculated as the average ratio of the shortest pore channel length L_{min} to the corresponding

straight-line distance L:

$$\tau_g = \left\langle \frac{L_{min}}{L} \right\rangle \tag{3.7}$$

The algorithm first converts the 3D microstructure dataset into a 2D adjacency matrix, so that the digital microstructure in Figure 3.3a is modeled by the graph in Figure 3.3b. Each node in the graph represents a voxel in the microstructure, and the connections between nodes represent the voxel connectivities, where the 6-neighbour-connectivity criterion is usually adopted. The Euclidean distance between two adjacent void voxels in the digital microstructure is used as the weight of linkage between the corresponding nodes in the graph, and all weights are equal to 1 for the 6-neighbour-connectivity case. Dijkstra's algorithm is then used to search for the shortest path from a designate starting node to an ending node that represents a pore voxel on the opposite surface. The search is performed for each pore voxel on the start surface to identify all pore channels of interest, as shown in Figure 3.3c. Finally, geometrical tortuosity τ_g of the porous medium in Figure 3.3a can be calculated according to Eq. (3.7).

3.2.3 Skeleton shortest-path search method

The shortest pore channels identified by DSPSM can deviate greatly from the medial axis of pore network, especially for porous media with large porosities or high-resolution microstructures. As shown in Figure 3.4, the shortest path identified by Dijkstra's algorithm in the wide channel shown in Figure 3.4b deviates greatly from the central axis and such deviation is much smaller for the narrow channel shown in Figure 3.4a. As a result, the DSPSM may underestimate geometrical tortuosity. Besides, to manage computational cost, the simple 6-neighbour-connectivity criterion as shown in Figure 3.3b is usually adopted by the DSPSM, leading to bumpy results as shown in Figure 3.3c, which could introduce extra errors into geometrical tortuosity evaluation.

One way to reduce the above error is to perform the path search on the pore skeleton instead of the pore space [280, 428, 402, 6]. Formed by the medial axis of a pore space, the skeleton preserves to a large extent of the original topological and geometrical properties. The homotopic thinning algorithm [265] is often used to extract the 3D medial axis skeleton from the 3D microstructure, and then Dijkstra's algorithm is applied on the skeleton to identify the shortest pore channels, as illustrated in Figure 3.5. Geometrical tortuosity is then calculated according to Eq. (3.7). The medial axis skeleton has a much lower dimensionality compared to the original pore space, which greatly simplifies the path search task. Therefore, Dijkstra's algorithm can adopt 18- or 26-neighbor connectivity. Generally, the tortuosity value obtained from the pore skeleton is greater than that from the pore space [422], especially for porous

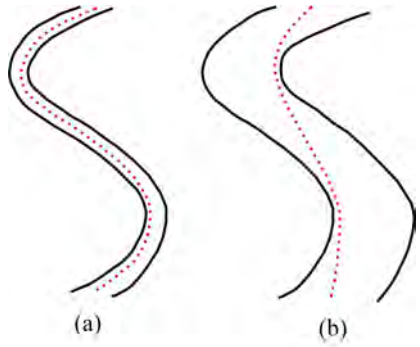


Fig. 3.4 The shortest paths (red dashed lines) identified by Dijkstra's algorithm in a narrow pore channel (left) and a wide channel (right)

media with high porosities. More details about the skeleton extraction can be found in relevant literatures [265, 362].

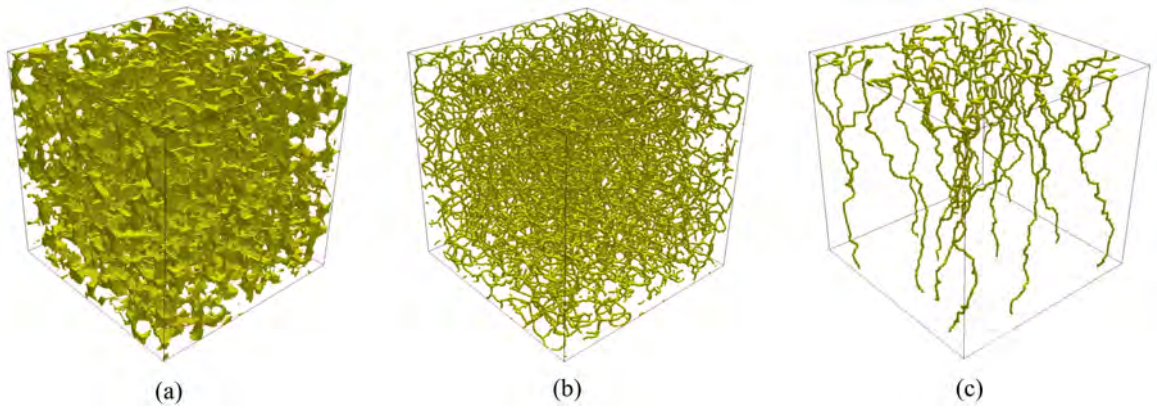


Fig. 3.5 (a) The pore space of the 3D porous medium in Figure 3.3a; (b) The medial axis skeleton extracted from the pore space; (c) The shortest pore channels identified between top and bottom surface

3.2.4 Fast marching method

The FMM is another frequently-used algorithm to find the shortest pore channel inside porous media [399, 223, 173, 347, 222, 431]. It is an efficient computational algorithm for tracking and modeling the propagation of a wavefront moving in its normal direction with a known speed. At a given position $\mathbf{x}$, the motion of the wavefront is described by the Eikonal equation [399, 25, 173] as follows:

$$\|\nabla t(\mathbf{x})\| F(\mathbf{x}) = 1,\ F(\mathbf{x}) > 0 \tag{3.8}$$

where $\|\nabla t(\mathbf{x})\|$ is the magnitude of the gradient vector of the arrival function $t(\mathbf{x})$, and $F(\mathbf{x})$ is the propagation speed of the wavefront at the location $\mathbf{x}$. By efficiently approximating the solution to the Eq. (3.8), the FMM calculates the time $t(\mathbf{x})$ required by the wavefront originating from the source to arrive at the other points in the computational domain. As $F(\mathbf{x}) > 0 \; \forall \mathbf{x}$, the wavefront only expands outward, and hence the points further away from the source have greater $t(\mathbf{x})$ values. If $F(\mathbf{x}) = 1$ is set throughout the investigation domain, the arrival time map $t(\mathbf{x})$ is equivalent to the distance map $d(\mathbf{x})$, which provides the distances to all locations from the source, as shown in Figure 3.6c.

The FMM can be directly applied to a regular grid such as a 3D digital image, to evaluate geometrical tortuosity τ_g of a porous medium. The pore voxels on the starting surface are selected as the sources of wavefront expansion, and the opposite surface is set as the destination plane. The lengths of shortest pore channels L_{min} connecting the source voxels and destination voxels can be obtained from the distance map $d(\mathbf{x})$, by setting $F(\mathbf{x}) = 1$ for the pore space and $F(\mathbf{x}) = 10^{-6}$ (a small value) for the solid matrix, where the wavefront moves much faster in the pore space than in the solid matrix. Similarly, the straight-line distances L between the starting and ending voxels can also be obtained from the distance map $d(\mathbf{x})$, by setting $F(\mathbf{x}) = 1$ for both pore space and solid matrix, where the porous media sample is considered as a homogeneous domain allowing the wavefront to propagate freely. Once the distance data L_{min} and L are determined by the FMM, geometrical tortuosity τ_g can be calculated according to Eq. (3.7).

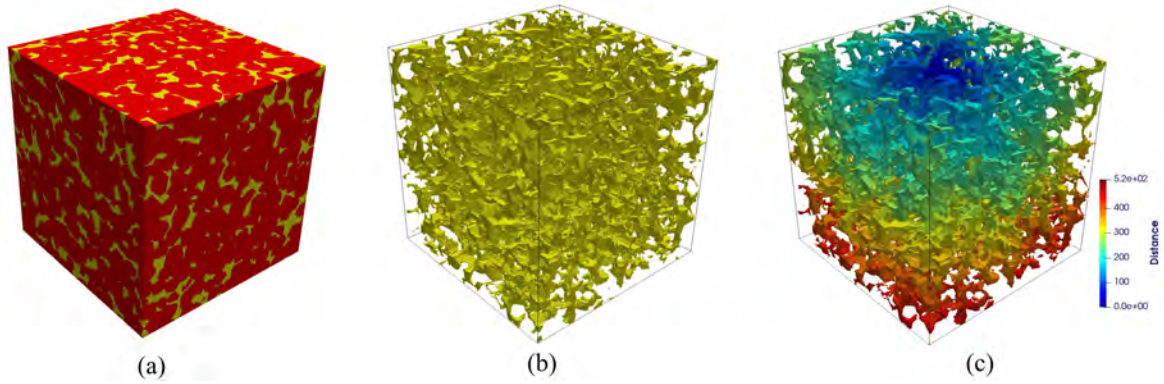


Fig. 3.6 (a) The 3D digital microstructure; (b) The pore space of the 3D porous medium, which is the investigation domain of FMM; (c) The distance map calculated from FMM by setting the pore voxel on the top center as the source of wavefront expansion

3.2.5 Pore centroid method

The PCM estimates geometrical tortuosity τ_g by calculating the average change of pore centroid locations between adjacent 2D slices in the 3D microstructure [162, 414, 98, 438]. As shown in Figure 3.7, the effective pore channel is determined by calculating the coordinate (x_i, y_i, z_i) of pore centroid on each layer and then connecting these pore centroids into a path. The length L_g of the effective pore path can be calculated as:

$$L_g = \sum_{i=1}^{N-1} L_i = \sum_{i=1}^{N-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2} \tag{3.9}$$

where i enumerates the 2D slices, and N is the total number of 2D slices in the 3D microstructure. Geometrical tortuosity τ_g is then calculated as the ratio of L_g to L, according to Eq. (3.6).

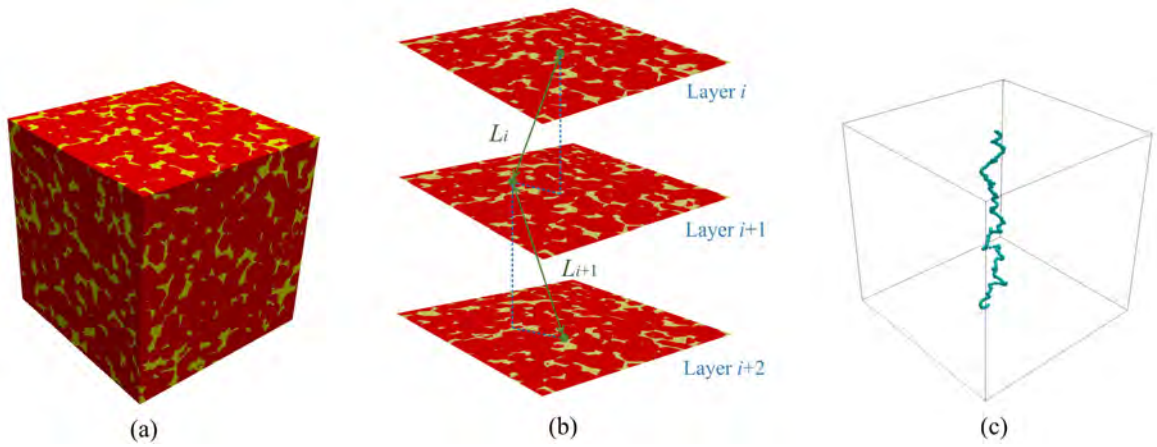


Fig. 3.7 (a) The 3D digital microstructure; (b) The distances L_i between the pore centroids on adjacent 2D slices; (c) The effective pore path from top to bottom surface determined by PCM

3.3 Physical tortuosities

In contrast to geometrical tortuosity characterizing the pore microstructure itself, physical tortuosities focus on the transport processes occurring inside the porous medium, including transfers of volume, charge, mass and energy. Numerical simulations can be performed on the digital microstructures of porous media to mimic various transport phenomena at the pore scale, thereby to compute corresponding physical tortuosities.

3.3.1 Analogy among different transport phenomena

The study of transport phenomena in porous media [2, 87, 30] concerns fluid permeation, electrical conduction, molecular diffusion, heat transfer and momentum exchange, and different transport phenomena share notable commonalities in terms of macroscopic mechanisms. Generally, these transport phenomena are governed by two primary principles at the steady state [42, 59, 359]: (1) the conservation law, which demands the quantity of transport matter (i.e. mass, charge, energy, momentum or others) must be conserved; (2) the constitutive relation, which describes the behavior of transport matter responding to stimuli.

Table 3.4 The analogy among various transport phenomena in porous media

Transport phenomenon (Transport matter/quantity)	**Driving force**	**Constitutive relation**	**Phenomenological property**	**Definitive law**
Fluid permeation (Volume)	Hydraulic gradient ∇H (Pressure gradient)	$\mathbf{J}_{\mathrm{h}} = -K\nabla H$	Hydraulic conductivity K (Intrinsic permeability)	Darcy's law
Molecular diffusion (Mass)	Concentrate gradient ∇C	$\mathbf{J}_{\mathrm{d}} = -D\nabla C$	Diffusion coefficient D	Fick's law
Electrical conduction (Charge)	Voltage gradient ∇V	$\mathbf{J}_{\mathrm{e}} = -\sigma\nabla V$	Electrical conductivity σ	Ohm's law
Heat transfer (Energy)	Temperature gradient ∇T	$\mathbf{J}_{\mathrm{t}} = -\lambda\nabla T$	Thermal conductance λ	Fourier's law
Newtonian viscous flow (Momentum)	Velocity gradient ∇U	$\mathbf{J}_{\mathrm{m}} = -\mu\nabla U$	Dynamic viscosity μ	Newton's law of viscosity

Note: $\mathbf{J}_{\mathrm{h}}$, $\mathbf{J}_{\mathrm{d}}$, $\mathbf{J}_{\mathrm{e}}$, $\mathbf{J}_{\mathrm{t}}$ and $\mathbf{J}_{\mathrm{m}}$ denote fluid flux, diffusion flux, charge flux, heat flux and momentum flux (shear stress) respectively.

For different transport phenomena, if the density of transport matter is constant and independent of time and space, the continuity equation can be simply expressed as follows [42, 59, 359]:

$$\nabla \cdot \mathbf{J} = 0 \tag{3.10}$$

where $\mathbf{J}$ is the steady-state flux of transport matter. As to the constructive relations, transfers of fluid volume, mass, charge, heat and momentum are almost described identically, which can be seen in Table 3.4. Different transport processes are governed by similar macroscopic laws (i.e. Darcy's, Fick's, Ohm's, Fourier's and Newton's law), which have a general expression as follows [42, 59, 359]:

$$\mathbf{J} = -\delta\nabla F_{\mathrm{drive}} \tag{3.11}$$

where the steady-state flux $\mathbf{J}$ of transport matter is proportional to the applied driving force $\nabla F_{\mathrm{drive}}$, and the proportionality constant δ is the phenomenological coefficient corresponding

to transport property (i.e. intrinsic permeability, diffusion coefficient, electrical conductivity, thermal conductance or dynamic viscosity).

Massive efforts have been made to study the 'analogy' among various transport phenomena, aiming to achieve the utility of them [42, 59, 359]. Partially due to this mathematical analogy, the physical tortuosities associated with four different transport phenomena (i.e. fluid permeation, electric conduction, molecular diffusion and heat transfer) are often interchangeably used in the literature, without rigorous justification. However, this mathematical analogy in no way means that the physical mechanisms underlying different transport processes are completely equivalent in any aspect. For example, there are discrepancies between the transport streamlines of different transport phenomena happening inside a specific porous medium, which is the primary cause of the inconsistencies between different types of physical tortuosity (more detailed explanation is provided in Section 3.4). To address this gap, we systematically investigate hydraulic, electrical, diffusional and thermal tortuosity in the following subsections to clarify their relationships.

3.3.2 Hydraulic tortuosity

Definition of hydraulic tortuosity

Hydraulic tortuosity τ_{h} was first proposed by Kozeny (1927) [250] and then further refined by Carman (1937) [71] to model permeability κ of porous media through the well-known Kozeny-Carman equation [90]:

$$\kappa = \frac{\phi^3}{\beta \tau_{\mathrm{h}}^2 S^2} \tag{3.12}$$

where ϕ is the porosity defined as the fraction of pore space in the porous medium, S is the specific surface area equal to the ratio of interstitial surface area to bulk volume, and β is the shape factor (a constant depending on the type of granular material).

Hydraulic tortuosity τ_{h} is defined as [90]:

$$\tau_{\mathrm{h}} = \frac{L_{\mathrm{h}}}{L} \tag{3.13}$$

where L_{h} denotes the length of effective hydraulic flow path and L the straight-line distance through the porous medium in the macroscopic flow direction.

The length L_{h} is usually measured from the streamlines of steady state pore-scale flow. Hydraulic tortuosity τ_{h} can be considered as a dimensionless parameter to describe the average elongation of streamlines in a porous medium with respect to the homogeneous flow [124]. However, the flow flux in a random porous medium varies continuously along the

flow path affected by its cross section, shape, orientation, branching and rejoining, making it difficult to identify the streamlines of pore-scale flow.

Evaluation via fluid simulation

The fluid flow in a digital microstructure can be simulated using such numerical schemes as LBM [219] and FVM [356], after which hydraulic tortuosity τ_{h} can be calculated by analyzing the streamlines determined by the pore-scale simulation. Several methods have been proposed to calculate τ_{h} based on the streamlines. The effective length of hydraulic flow path L_{h} is often evaluated as a weighted average of streamline lengths [498, 124]:

$$\tau_{\mathrm{h}} = \frac{1}{L}\left(\frac{\sum_i w_i l_i}{\sum_i w_i}\right) \tag{3.14}$$

where i enumerates the separated streamlines, l_i is the length of the ith streamline, and w_i is the weight of the ith streamline contributing to the overall flux.

Different methods have been proposed to determine the weight w_i. Zhang (1995) [498] used the travel time t_i of a fluid particle moving along the ith streamline to determine the weight w_i, given by: $w_i = 1/t_i$. Considering the relationship between the travel time and the fluid particle speed along the streamline, Koponen (1996) [247] modified the calculation as follows:

$$\tau_{\mathrm{h}} = \frac{1}{L}\frac{\int_\Omega l(\mathbf{r})v(\mathbf{r})\,\mathrm{d}\Omega}{\int_\Omega v(\mathbf{r})\,\mathrm{d}\Omega} \approx \frac{1}{L}\frac{\sum_j l(\mathbf{r}_j)v(\mathbf{r}_j)}{\sum_j v(\mathbf{r}_j)} \tag{3.15}$$

where Ω denotes the volume domain of pore space, $l(\mathbf{r})$ is the length of the streamline passing through a point $\mathbf{r}$, and $v(\mathbf{r}) = |\mathrm{v}(\mathbf{r})|$ is the tangential velocity of the fluid at point $\mathbf{r}$. It was later proved by [124] and [302] that Eq. (3.15) is equivalent to:

$$\tau_{\mathrm{h}} = \frac{\int_\Omega v(\mathbf{r})\,\mathrm{d}\Omega}{\int_\Omega v_x(\mathbf{r})\,\mathrm{d}\Omega} = \frac{\langle v\rangle}{\langle v_x\rangle} \tag{3.16}$$

where $v_x(\mathbf{r})$ is the velocity component along the macroscopic flow direction at point $\mathbf{r}$, and $\langle\cdot\rangle$ denotes the spatial average over the pore space Ω.

Eq. (3.16) directly calculates hydraulic tortuosity from the steady-state fluid velocity filed, avoiding the calculation of streamlines. This method is widely used as it significantly simplifies the calculation of hydraulic tortuosity for porous media with complicated geometries. It should be noted that Eq. (3.16) is based on two assumptions: the fluid is incompressible and the fluid flow is not reentrant. For the fluid flow passing through porous media with a low Reynolds number, the effect of reentrant is negligible.

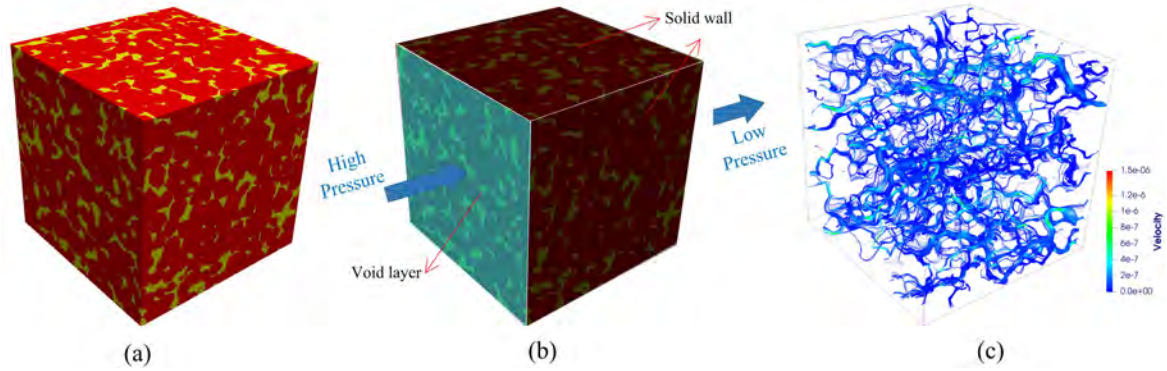


Fig. 3.8 (a) A 3D porous media REV; (b) Boundary conditions for the LBM simulation of fluid flow passing through a porous media REV; (c) The steady-state fluid velocity field inside the porous media REV

LBM is adopted in this work to simulate the fluid flow in digital microstructures, as shown in Figure 3.8c. The LBM models the fluid particles through a time-dependent distribution propagating on a regular lattice [252]. The pore voxels of digital microstructures serve as the regular lattice, and each lattice node is located in the center of corresponding pore voxel. Starting with an initial state, the particle distribution function moves from one lattice node to its neighboring nodes at each time step, and evolves itself locally subject to both mass and momentum conservation. The local equilibrium particle distribution function corresponds to an ideal state, where the particle distribution tends to a specific macroscopic state, and it is specifically selected to recover the macroscopic Navier-stokes equations. For stationary, creeping, and incompressible Newtonian fluid flow in porous media, the Navier-Stokes equations can be simplified to Darcy's law (as expressed in Table 3.4).

Pore-scale simulations using LBM are well established and comprehensively documented in many literatures [219, 258, 139]. In this work, the conventional D3Q19 lattice arrangement and Bhatnagar-Gross-Krook (BGK) collision operator are adopted. Since the steady-state flow in porous media is insensitive to the initial state, we simply set the flow velocity to 0 and the fluid density to 1 for the whole domain as the initial conditions. As shown in Figure 3.8b, two types of boundary condition are adopted: the no-slip boundary at the pore-solid interface and the constant pressure boundary at inlet and outlet. To apply the constant pressure gradient to the cubic microstructure sample, two void layers are added to the outlet and inlet faces. As to the other four faces that are parallel to the global flow direction, solid walls are added to isolate the cubic REV from the outside. Driven by a constant pressure gradient, the fluid flows from the inlet face towards the outlet face steadily, as shown in Figure 3.8b. The LBM simulation runs iteratively until the steady state is reached, and then hydraulic tortuosity can be calculated following Eq. (3.16).

3.3.3 Electrical tortuosity

Definition of electrical tortuosity

Electrical tortuosity is measured through the electrical conduction experiment, where the porous media sample is saturated with an electrolyte of electrical conductivity σ_0 and driven by a voltage difference, the electric charges are transmitted through the pore space filled by the electrolyte liquid [149, 467, 26]. The effective conductivity σ_{eff} measured on the test sample is smaller than the electrolyte conductivity σ_0, due to the retarding effect of solid microstructure on the electric charge transmission [477]. The difference between σ_0 and σ_{eff} is represented by electrical tortuosity τ_{e} as follows [448, 348, 260]:

$$\sigma_{\text{eff}} = \sigma_0 \frac{\phi}{\tau_{\text{e}}^2} \tag{3.17}$$

where ϕ is the porosity, and $1/\tau_{\text{e}}^2$ can be understood as the fraction of pore volumes having the same efficiency for electrical conductivity as the electrolyte.

To be consistent with hydraulic tortuosity τ_{h} in Eq. (3.13)), electrical tortuosity τ_{e} was redefined as [477, 90, 68]:

$$\tau_{\text{e}} = \frac{L_{\text{e}}}{L} \tag{3.18}$$

where L_{e} is the effective length of the electrical conduction path. Electrical tortuosity of a porous medium is defined as a retardation factor of the electrical conduction in Eq. (3.17), and also defined as an elongation factor of electrical conduction path in Eq. (3.18). However, the equivalence between these two definitions is derived from a simple capillary tube model [26], while it remains unclear whether or not a similar correlation holds for general porous media [124].

Evaluation via electrical conduction simulation

Electrical conduction in porous media can be simulated by various numerical methods such as LBM [427], FEM [383], FDM [396] and RWM [209]. In this work, we use the Avizo package [20] to simulate electrical charges transmitting through 3D digital microstructures and electrical tortuosity is calculated according to Eq. (3.17). The Avizo solver is based FVM [317], where the image voxels are used as the volume elements. The solid phase is assumed to be homogeneous and insulating, while the pore space is saturated with an electrolyte of electrical conductivity σ_0. As shown in Figure 3.9b, a constant electrical potential difference is applied on the inlet and outlet faces of the cubic digital microstructure, while the other four faces are set as isolating. The boundary condition for the pore-solid interface is set as

follows:

$$\nabla V \cdot \mathbf{n} = 0 \tag{3.19}$$

where $\mathbf{n}$ is the unit normal of the pore-solid surface, pointing to the solid phase.

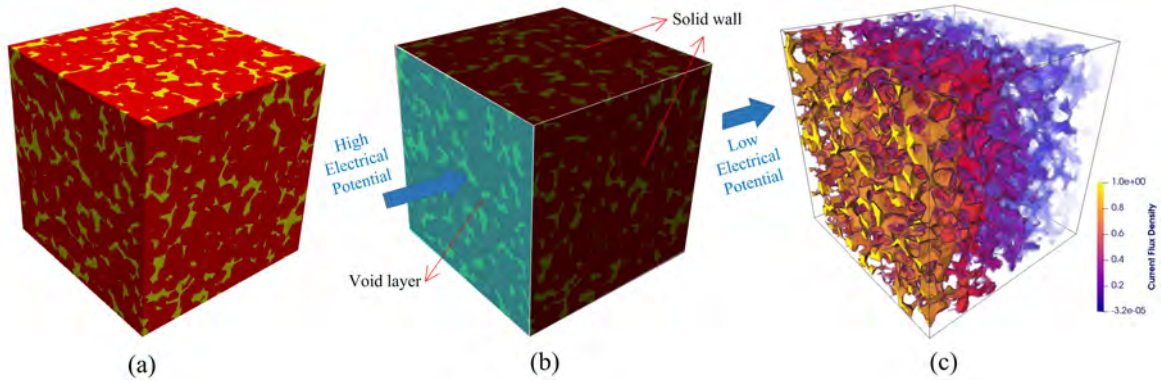


Fig. 3.9 (a) A 3D porous media REV; (b) Boundary conditions for the FVM simulation of electrical conduction in the porous media REV; (c) The steady-state electrical flux density inside the porous media REV

Driven by the constant electrical potential gradient, electrical charges are transmitted through the porous media sample filled with electrolyte. The electric conduction at the steady state is governed by Ohm's law and the conservation of charges, as expressed in Table 3.4 and Eq. (3.10) respectively. The effective (apparent) electrical conductivity σ_{eff} can be estimated by:

$$\frac{J_{\text{e}}^{(\text{REV})}}{A} = \sigma_{\text{eff}} \frac{V_{\text{outlet}} - V_{\text{inlet}}}{L} \tag{3.20}$$

where $J_{\text{e}}^{(\text{REV})}$ is the electrical flux passing through the outlet face of the porous media sample, A is the cross section area perpendicular to the flow direction, V_{inlet} and V_{outlet} are the applied electrical potentials at the inlet and outlet faces, respectively, L is the distance between inlet and outlet faces. The total electrical flux $J_{\text{e}}^{(\text{REV})}$ can be obtained through an integration:

$$J_{\text{e}}^{(\text{REV})} = \int_A -\sigma_0 \nabla V \, \mathrm{d}A \tag{3.21}$$

Substituting Eq. (3.20) and Eq. (3.21) into Eq. (3.17), electrical tortuosity τ_{e} can be computed as:

$$\tau_{\text{e}} = \sqrt{\frac{\sigma_0}{\sigma_{\text{eff}}}\phi} = \sqrt{\frac{A\sigma_0 \frac{V_{\text{outlet}} - V_{\text{inlet}}}{L}}{J_{\text{e}}^{(\text{REV})}}\phi} = \sqrt{\frac{J_{\text{e}}^{(\text{Free})}}{J_{\text{e}}^{(\text{REV})}}\phi} \tag{3.22}$$

where $J_{\text{e}}^{(\text{Free})}$ is the electrical flux passing through a homogeneous medium with the same dimensions, of the electrical conductivity σ_0, and driven by the same electrical potential

gradient. As discussed earlier, electrical tortuosity τ_e is often used to substitute hydraulic tortuosity τ_h whose direct evaluation is more time-consuming and costly.

3.3.4 Diffusional tortuosity

Definition of diffusional tortuosity

The tortuosity of porous media can also be measured by using a diffusion experiment, where a diffusing gases are typically used as the probe due to their high diffusion rates [449, 251]. The effective diffusivity D_{eff} in porous media is lower than the bulk diffusivity D_0 of the diffusing gas [149]. Diffusional tortuosity τ_d is determined by the ratio of D_{eff} to D_0 as [449, 127, 55, 312, 440]:

$$D_{\text{eff}} = D_0 \frac{\phi \delta}{\tau_d^2} \tag{3.23}$$

where δ is the constriction coefficient due to the variation of pore diameter along the diffusion pathway, and it depends on the ratio of the solute diameter (probe molecule dimension) to the pore diameter. The kinetic diameter of probe gas molecule is usually smaller than 1 nm [411], and when the pore diameter is larger than 1 μm, δ is approximately 1 [163].

By modeling the porous medium with cylindrical capillaries, diffusional tortuosity τ_d is also defined as the ratio between the effective length of diffusive path L_d and the straight-line distance L [127]:

$$\tau_d = \frac{L_d}{L} \tag{3.24}$$

Generally speaking, mass diffusion inside porous media may involve ordinary molecular (Fickian) diffusion [199] and Knudsen diffusion [298], depending on the ratio of the mean free path of the probe molecules to the pore diameter. If the mean free path is relatively short compared to the pore size, only the molecular diffusion occurs, and the diffusion process can be described by Fick's law as expressed in Table 3.4. Knudsen diffusion takes place when the mean free path is comparable to the pore size, where the molecules collide with the pore walls more frequently than with each other, impeding the molecule transport. The diffusion transport usually decreases as the probe molecule size increases, because more collisions occur between the molecules and the pore walls. Therefore, the tortuosity obtained from different diffusion measurements may differ, depending on the diffusion regime and the probe molecule. The thresholds of pore size are 1 μm and 10 nm for Fickian diffusion and Knudsen diffusion, respectively [484, 149]. That is, the Knudsen effect is negligible when pore size is larger than 1 μm, and in this case the diffusion tortuosity is independent of the probe molecule. In this work, only the ordinary molecular diffusion is considered

and diffusional tortuosity τ_{d} is evaluated according to Eq. (3.23), where the constriction coefficient δ is set as 1.

Evaluation via molecular diffusion simulation

The molecular diffusion process in porous media can be simulated by such numerical methods as FVM [438], LBM [342] and FDM [95]. In this work, we adopt the open-source solver TauFactor [96], which is based on FDM and directly uses image voxels as the discretization mesh for simulation. As shown in Figure 3.10, diffusional tortuosity is calculated by comparing the steady-state diffusive flux passing through the cubic porous media REV and the cubic homogeneous REV with the same dimensions. A constant concentration difference is applied between inlet and outlet faces of both REVs, while the other four faces parallel to the diffusional flow direction are set as isolating solid walls. The boundary condition at the pore-solid interface is set as:

$$\nabla C \cdot \mathbf{n} = 0 \tag{3.25}$$

Driven by the constant concentration gradient, the diffusive specie moves from the inlet face to the outlet face. When the steady state is reached, the diffusional flux passing through the porous media REV and the homogeneous REV can be obtained from Fick's law:

$$J_{\mathrm{d}}^{(\mathrm{REV})} = -AD_{\mathrm{eff}}\frac{\Delta C}{L} \tag{3.26}$$

$$J_{\mathrm{d}}^{(\mathrm{Free})} = -AD_0\frac{\Delta C}{L} \tag{3.27}$$

where $J_{\mathrm{d}}^{(\mathrm{REV})}$ and $J_{\mathrm{d}}^{(\mathrm{Free})}$ are the diffusion fluxes passing through the porous media REV and the homogeneous REV, respectively. Taking the ratio of $J_{\mathrm{d}}^{(\mathrm{REV})}$ to $J_{\mathrm{d}}^{(\mathrm{Free})}$ and then rearranging the equation, diffusional tortuosity τ_{d} can be expressed as:

$$\tau_{\mathrm{d}} = \sqrt{\frac{D_0}{D_{\mathrm{eff}}}\phi} = \sqrt{\frac{J_{\mathrm{d}}^{(\mathrm{Free})}}{J_{\mathrm{d}}^{(\mathrm{REV})}}\phi} \tag{3.28}$$

Evaluation via random walk simulation

Diffusional tortuosity τ_{d} of a porous medium can also be statistically estimated from the random walk process of non-absorbing particles, which can be considered as a simplified diffusion simulation [326, 325, 211, 201, 438]. The random walk simulation starts by randomly distributing a number of walkers into the pore space at $t = 0$. Then, each walker

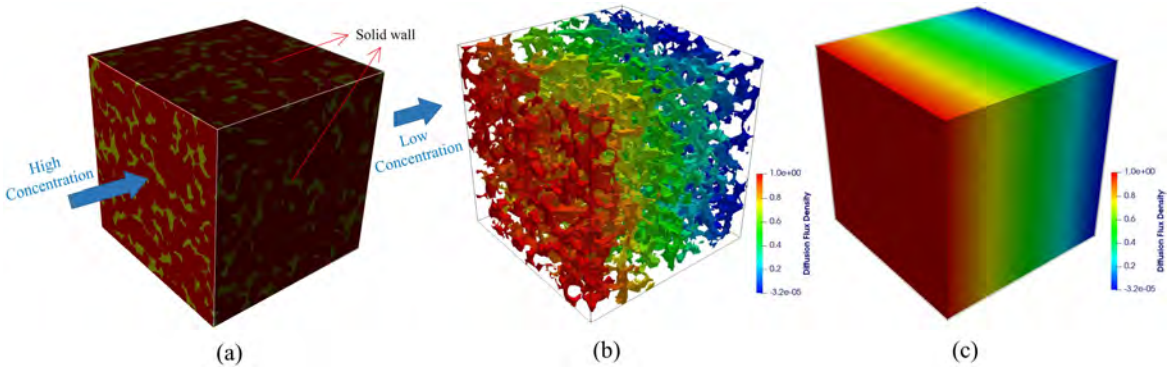


Fig. 3.10 (a) The schematic diagram of boundary conditions for the FDM simulation of molecular diffusion in a porous media REV; (b) The diffusional flux density at the steady state inside the porous media REV; (c) The diffusional flux density at the steady state inside the homogeneous REV

randomly selects one of its neighboring voxels (e.g. following the 6-neighbor-connectivity criterion) as its possible position at the next time step. If the selected voxel belongs to the pore phase, the walker moves to the new position; otherwise the walker stays at the current location. With the random walking procedure repeated, the mean square displacement $\langle r^2(t)\rangle$ of all random walkers is a function of time step t:

$$\langle r^2(t)\rangle = \frac{1}{N}\sum_{i=1}^{N}\left\{[x_i(t)-x_i(0)]^2+[y_i(t)-y_i(0)]^2+[z_i(t)-z_i(0)]^2\right\} \tag{3.29}$$

where N is the number of random walkers, $x_i(t)$, $y_i(t)$ and $z_i(t)$ are the coordinates of the ith walker at time t, and $\langle\cdot\rangle$ denotes the ensemble average.

For random walkers walking inside a cubic homogeneous REV, the mean square displacement $\langle r^2(t)\rangle_{\text{Free}}$ of them is related to the diffusion coefficient D_{Free} as follows:

$$D_{\text{Free}} = \frac{1}{6}\frac{\mathrm{d}\langle r^2(t)\rangle_{\text{Free}}}{\mathrm{d}t} \tag{3.30}$$

where D_{Free} is the diffusion coefficient in the cubic homogeneous REV (e.g. the diffusivity of gas or water). While for random walkers in a porous medium REV, the mean square displacement $\langle r^2(t)\rangle_{\text{Pore}}$ of them is reduced due to the obstruction effect from solid phase, which is related to the effective diffusion coefficient D_{Pore} as follows:

$$D_{\text{Pore}} = \phi\frac{1}{6}\frac{\mathrm{d}\langle r^2(t)\rangle_{\text{Pore}}}{\mathrm{d}t} \tag{3.31}$$

D_{Free} is a constant because of the unrestricted diffusion in the cubic homogeneous REV, but D_{pore} is time-dependent due to local heterogeneity present in porous media. The random walk trajectory is restricted by the solid phase, thereby reducing the diffusivity.

Diffusional tortuosity τ_{d} is calculated from the ratio of D_{Free} to D_{Pore} as follows:

$$\tau_{\text{d}} = \sqrt{\frac{D_{\text{Free}}}{D_{\text{Pore}}}\phi} = \sqrt{\frac{\mathrm{d}\left\langle r^2(t)\right\rangle_{\text{Free}}/\mathrm{d}t}{\mathrm{d}\left\langle r^2(t)\right\rangle_{\text{Pore}}/\mathrm{d}t}} \tag{3.32}$$

Long-time data of the mean square displacement $\left\langle r^2(t)\right\rangle$ from large numbers of random walkers can fully experience the porous microstructure, and the slope $\mathrm{d}\left\langle r^2(t)\right\rangle_{\text{Pore}}/\mathrm{d}t$ approximates to a constant value, as shown in Figure 3.11c.

A natural porous medium such as rock often possesses an anisotropic pore structure. If the porous medium is significantly anisotropic, the effective diffusivity D_{Pore} is tensor instead of a scalar, and Eqs. (3.29) to (3.32) break down [325]. Directional mean square displacements $\left\langle x^2(t)\right\rangle$, $\left\langle y^2(t)\right\rangle$ and $\left\langle z^2(t)\right\rangle$ are required to evaluate the directional tortuosities of the anisotropic porous medium [325, 211], given by

$$\left\langle x^2(t)\right\rangle_{\text{Pore}} = \frac{1}{N}\sum_{i=1}^{N}[x_i(t) - x_i(0)]^2 \tag{3.33}$$

$$\left\langle y^2(t)\right\rangle_{\text{Pore}} = \frac{1}{N}\sum_{i=1}^{N}[y_i(t) - y_i(0)]^2 \tag{3.34}$$

$$\left\langle z^2(t)\right\rangle_{\text{Pore}} = \frac{1}{N}\sum_{i=1}^{N}[z_i(t) - z_i(0)]^2 \tag{3.35}$$

$$\left\langle x^2(t)\right\rangle_{\text{Free}} = \left\langle y^2(t)\right\rangle_{\text{Free}} = \left\langle z^2(t)\right\rangle_{\text{Free}} = \frac{1}{3}\left\langle r^2(t)\right\rangle_{\text{Free}} \tag{3.36}$$

Then, the directional tortuosities τ_x, τ_y and τ_z can be calculated as follows:

$$\tau_x = \sqrt{\frac{\mathrm{d}\left\langle x^2(t)\right\rangle_{\text{Free}}/\mathrm{d}t}{\mathrm{d}\left\langle x^2(t)\right\rangle_{\text{Pore}}/\mathrm{d}t}} \tag{3.37}$$

$$\tau_y = \sqrt{\frac{\mathrm{d}\left\langle y^2(t)\right\rangle_{\text{Free}}/\mathrm{d}t}{\mathrm{d}\left\langle y^2(t)\right\rangle_{\text{Pore}}/\mathrm{d}t}} \tag{3.38}$$

$$\tau_z = \sqrt{\frac{\mathrm{d}\left\langle z^2(t)\right\rangle_{\text{Free}}/\mathrm{d}t}{\mathrm{d}\left\langle z^2(t)\right\rangle_{\text{Pore}}/\mathrm{d}t}} \tag{3.39}$$

In our implementation, the random walk simulation lasts for 1,000,000 time steps with 10,000 random walkers, and the periodic boundary condition is set to avoid the random walkers moving out the cubic REV sample with a finite size. Data obtained from the first 500,000 time steps are dismissed, because the random walkers have not fully experienced the porous microstructure at the early stage of the simulation.

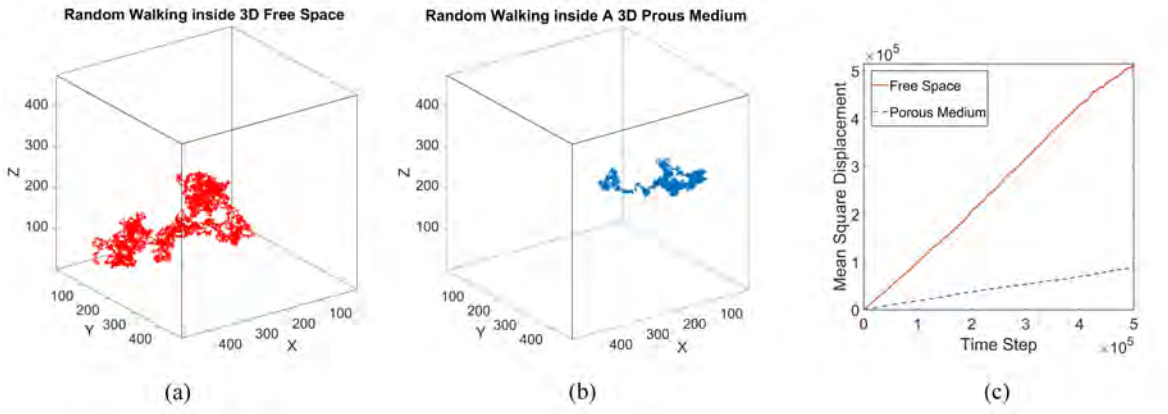


Fig. 3.11 (a) The trajectory of a single random walker moving inside the free space for 1,000,000 time steps; (b) The trajectory of a single random walker moving inside the pore space for 1,000,000 time steps; (c) Mean-square displacements of the random walking averaged over 10,000 walkers

3.3.5 Thermal tortuosity

Definition of thermal tortuosity

Heat transfer in porous media is mainly determined by the solid phase conduction. The convective and radiative effects can be largely ignored [338], because the thermal conductivity of solid matrix is usually much higher than the void phase of air. Therefore, the concept of thermal tortuosity is sometimes related to the winding paths of heat flow through the solid phase of a porous medium [338, 49]. However, this work focuses on the tortuosity of the pore space, and as such we simulate the heat conduction passing through the void phase of the porous medium, by assuming the solid phase to be thermal insulating. The effective thermal conductivity λ_{eff} can be evaluated from the heat flux simulation, based on which thermal tortuosity τ_{t} of pore space can be obtained as follows [98, 97]:

$$\lambda_{\text{eff}} = \lambda_0 \frac{\phi}{\tau_{\text{t}}^2} \tag{3.40}$$

where λ_0 is the bulk thermal conductivity of air.

Evaluation via heat transfer simulation

Heat transfer through the pore space can be numerically simulated by using FVM [97], FDM [473] and LBM [166]. The FVM-based Avizo package [20] is adopted in this work to simulate the heat flux passing through porous media, and the thermal tortuosity is calculated according to Eq. (3.40). The pore space is fully filled with air of thermal conductivity λ_0, while the solid phase is assumed to be homogeneous and thermal insulating. A constant temperature difference is applied between input and output faces of the cubic porous media REV, and the other faces are set as thermal isolating. The boundary condition on the pore-solid interface is set as follows:

$$\nabla T \cdot \mathbf{n} = 0 \tag{3.41}$$

Driven by the temperature difference, heat transfers inside the pore space from the higher temperature area to the lower temperature area. When the steady-state is reached, the heat flow can be described by the conservation of energy in Eq. (3.10) and Fourier's law (as expressed in Table 3.4). The effective (apparent) thermal conductivity λ_{eff} can be calculated by:

$$\frac{J_{\text{t}}^{(\text{REV})}}{A} = \lambda_{\text{eff}} \frac{T_{\text{outlet}} - T_{\text{inlet}}}{L} \tag{3.42}$$

where $J_{\text{t}}^{(\text{REV})}$ is the heat flux passing through the output face of the cubic REV, and T_{intlet} and T_{outlet} are the applied temperature at the input and output faces, respectively. The total heat flux $J_{\text{t}}^{(\text{REV})}$ can be obtained by integrating the local heat flux computed from Fourier's law:

$$J_{\text{t}}^{(\text{REV})} = \int_A -\lambda_0 \nabla T \, \mathrm{d}A \tag{3.43}$$

Substituting Eq. (3.42) and Eq. (3.43) into Eq. (3.40), thermal tortuosity τ_{t} can be computed as:

$$\tau_{\text{t}} = \sqrt{\frac{\lambda_0}{\lambda_{\text{eff}}}\phi} = \sqrt{\frac{A\lambda_0 \frac{T_{\text{outlet}} - T_{\text{inlet}}}{L}}{J_{\text{t}}^{(\text{REV})}}\phi} = \sqrt{\frac{J_{\text{t}}^{(\text{Free})}}{J_{\text{t}}^{(\text{REV})}}\phi} \tag{3.44}$$

where $J_{\text{t}}^{(\text{Free})}$ is the heat flux passing through a cubic homogeneous REV with the same dimensions and filled with air of thermal conductivity λ_0.

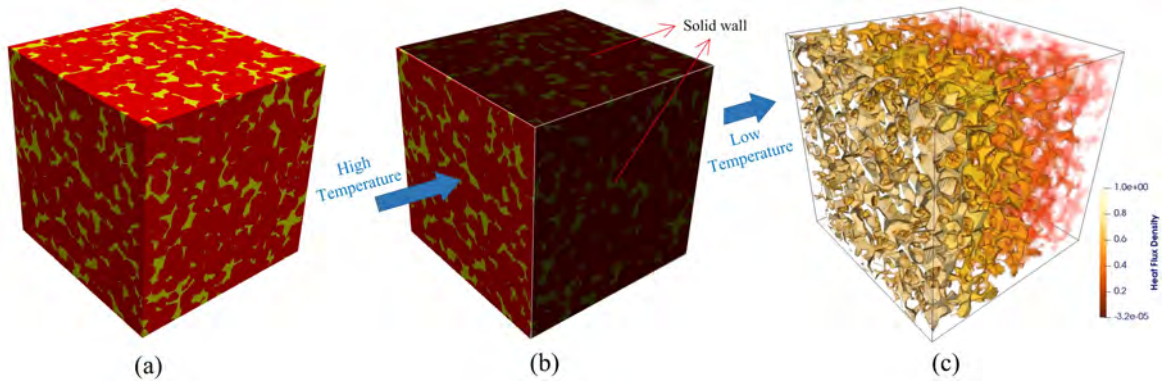


Fig. 3.12 (a) A 3D porous media REV; (b) Boundary conditions of the FVM simulation of heat flux passing through the porous media REV; (c) The heat flux density inside the porous media REV at the steady state

3.4 Comparison and analysis

A set of Fontainebleau sandstones with porosity ranging from 8.61% to 24.50% are used to compare the tortuosity results evaluated by different methods. As shown in Figure 2.1, the 3D REV models (in binary forms) are of $480 \times 480 \times 480$ voxels with a voxel resolution of 5.700 μm, representing an equivalent physical sample of 2.736 mm^3 cube [35]. For each REV, both geometrical and physical tortuosities are computed for three principal directions using the methods reviewed in Section 3.2 and Section 3.3, respectively. Fontainebleau sandstone is considered isotropic and to minimize the sampling error, a characteristic tortuosity τ is obtained by averaging three directional tortuosities [97]:

$$\frac{1}{\tau^2} = \frac{1}{3}\left(\frac{1}{\tau_x^2} + \frac{1}{\tau_y^2} + \frac{1}{\tau_z^2}\right) \tag{3.45}$$

where τ_x, τ_y and τ_z are tortuosities in the x, y and z direction respectively.

The tortuosity results evaluated from different methods for the set of Fontainebleau sandstone are summarized in Table 3.5 and plotted in Figure 3.13. Although the overall trends of tortuosity in relation to porosity are similar, remarkable discrepancies can be observed across different methods, especially for the samples with low porosity.

3.4.1 Comparison between geometrical tortuosity results

The geometrical tortuosities of Fontainebleau sandstone samples are computed using four image-based methods, as reviewed in Section 3.2. It can be observed in Table 3.5 and Figure 3.13 that remarkable differences are present between the tortuosity results, although the

Table 3.5 Geometrical and physical tortuosities of the Fontainebleau sandstone samples

Tortuosity type	Calculation method/model	Tortuosity values of samples with different porosities						
		8.61%	10.15%	12.57%	15.36%	17.62%	20.65%	24.50%
Geometrical	DSPSM	1.91	1.76	1.57	1.46	1.40	1.34	1.28
Geometrical	SSPSM	2.07	1.98	1.83	1.71	1.65	1.63	1.61
Geometrical	FMM	1.54	1.32	1.20	1.13	1.11	1.07	1.05
Geometrical	PCM	2.97	2.79	2.41	2.39	2.15	2.03	1.84
Hydraulic	LBM simulation	2.02	1.95	1.81	1.67	1.59	1.52	1.45
Electrical	FVM simulation	4.32	3.50	2.90	2.29	2.04	1.82	1.65
Diffusional	FDM simulation	4.33	3.51	2.85	2.29	2.04	1.82	1.65
Diffusional	RWM simulation	4.30	3.49	2.90	2.36	2.08	1.85	1.73
Thermal	FVM simulation	4.33	3.50	2.91	2.29	2.04	1.82	1.65
Diffusional	Bruggeman relation ($\alpha = 1.5$)	1.85	1.77	1.68	1.60	1.54	1.48	1.42
Hydraulic	Bruggeman relation ($\alpha = 1.8$)	2.67	2.50	2.29	2.12	2.00	1.88	1.76
Electrical	Archie's law ($\alpha = 2.5$)	6.29	5.56	4.74	4.08	3.68	3.26	2.87
Hydraulic	Comiti's model ($p = 0.49$)	1.48	1.46	1.42	1.38	1.36	1.33	1.30
Diffusional	Iversen's model ($q = 2$)	1.68	1.67	1.66	1.64	1.63	1.61	1.58
Hydraulic	Du Plessis's model	1.48	1.47	1.47	1.46	1.45	1.45	1.43
Geometrical	Lanfrey's model ($\xi = 0.9$)	15.64	12.97	10.10	7.92	6.66	5.40	4.26

trends in relation to porosity are similar. The largest and smallest geometrical tortuosity results are produced by the PCM and the FMM, respectively, and the difference between them exceeds 75%. The DSPSM and the SSPSM provide similar geometrical tortuosities, due to the same searching algorithm (Dijkstra's algorithm) being used. Because of the difference in image preparation, the tortuosity results from the DSPSM are usually smaller than that calculated from the SSPSM, as discussed in Section 3.2.3. Tortuosity results obtained from the FMM are always smaller than that calculated from the DSPSM and the SSPSM, which means the shortest pore channels identified by the FMM are 'shorter' than that searched by Dijkstra's algorithm. A stable difference between the geometrical tortuosity results computed from the FMM and Dijkstra's algorithm can be observed. As to the PCM, it is more suitable for the single pore channel case, where it recognizes the central axis of pore channel as the effective flow path.

Several factors could affect image analysis results, thereby bringing uncertainties to the computation of geometrical tortuosity: (1) Image preparation. For instance, the DSPSM directly searches for the shortest pore channel within the raw pore space, while the SSPSM operates on the pore skeleton after image skeletonization, leading to inconsistent results of geometrical tortuosity between them. The adopted skeletonization algorithm can also affect the extracted topology of the 3D graph [195] and then influence the tortuosity calculation. (2) The definition of the effective flow path L_g. The shortest pore channel L_{min} is often taken as

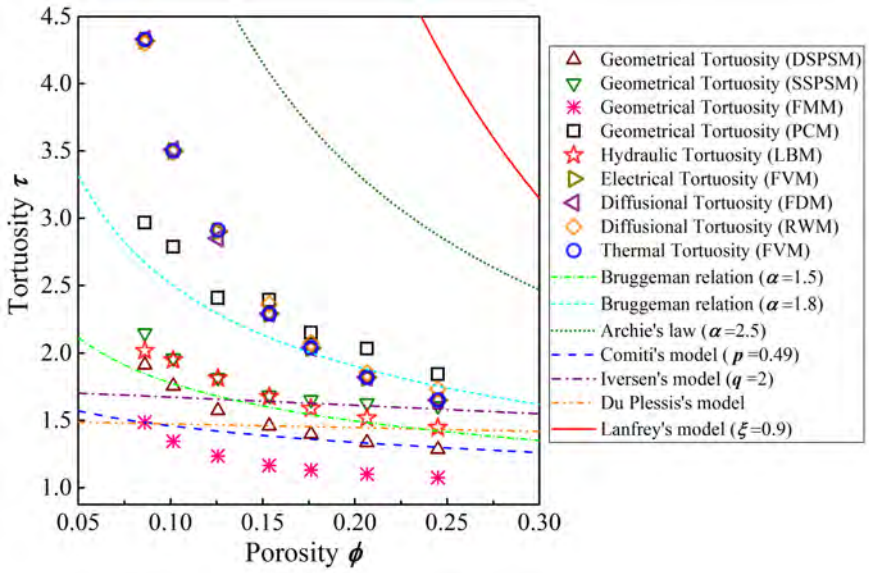


Fig. 3.13 The tortuosity results calculated from different image analysis methods, physical simulations and porosity-tortuosity models for the set of Fontainebleau sandstone samples

the effective flow path through the complex pore microstructure [280], but different criteria have also been used by other methods such as the PCM. (3) The path searching algorithm and the definition of pixel/voxel connectivity. For examples, the shortest pore channel searched by the fast marching method is different from the one obtained by Dijkstra's algorithm, and the shortest path based on the 6-neighbor-connectivity criterion (as shown in Figure 3.14) is usually longer than that based on the 18- or 26-neighbor-connectivity criterion. (4) The distance metric. A number of distance metrics exist for regular image grids, among which Manhattan distance, Euclidean distance and Chebyshev distance are widely used [80]. These distance metrics yield different distances between voxels, as shown in Figure 3.15. The choice of distance metric directly affects the identification of the shortest/effective pore channel, thereby affecting the geometrical tortuosity result.

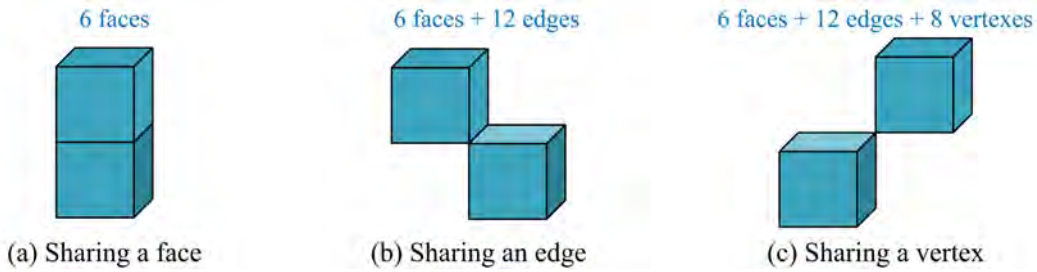


Fig. 3.14 Different connectivity criteria between adjacent image voxels: (a) 6-neighbor-connectivity; (b) 18-neighbor-connectivity; (c) 26-neighbor-connectivity

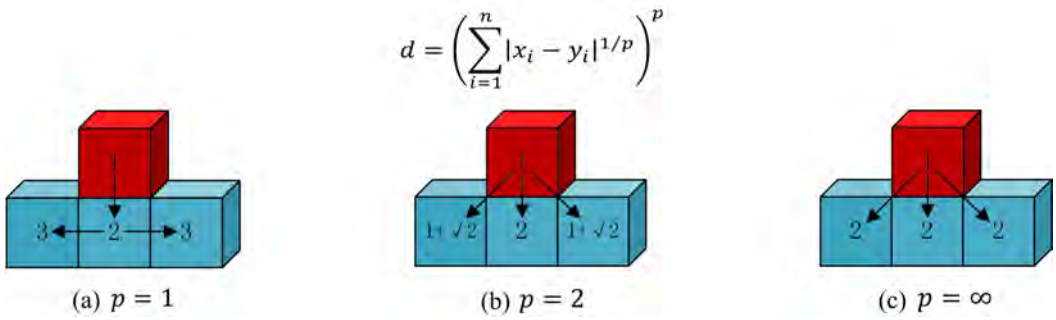


Fig. 3.15 Different distance metrics between image voxels: (a) Manhattan distance; (b) Euclidean distance; (c) Chebyshev distance

What's more, the effects coming from image resolution and image size on tortuosity computation cannot be ignored, but they can be reduced/avoided by using the digital REV samples with sufficient high resolutions and large sizes, which will be further discussed in Section 3.6. With the above intrinsic uncertainties present in various image-based calculation methods, proper use of geometrical tortuosity requires a standard procedure, clear definition and reasonable parameter selection, to ensure reproducible and portable results.

3.4.2 Comparison between physical tortuosity results

The physical tortuosities of the set of Fontainebleau sandstone samples are computed via different physical simulations as discussed in Section 3.3, and the results are recorded in Table 3.5 and plotted in Figure 3.13. It is observed that the diffusional, electrical and thermal tortuosity of each sample are almost identical. This is because the underlying transport physics of molecular diffusion, electrical conduction and heat transfer at steady state are described by similar constitutive relations, namely Fick's, Ohm's and Fourier's laws. Diffusivity and conductivity are directly related to each other through Nernst-Einstein equation as well [293]. Numerical simulations of these three transport processes are governed by Laplace equation with exactly the same mathematical expression, and the boundary conditions including the external driving force, outside surface and pore-solid interface are identical as well. Recognizing the consistency among diffusional, electrical and thermal tortuosities, a general definition of physical tortuosity τ can be expressed as [26]:

$$\left(\frac{J}{\nabla F_{\text{drive}}}\right)_{\text{eff}} = \frac{\phi}{\tau^2}\left(\frac{J}{\nabla F_{\text{drive}}}\right)_0 \tag{3.46}$$

where J denotes the transport flux, ∇F_{drive} denotes the driving force gradient, and the subscripts 'eff' and '0' denote the effective conductivity/diffusivity in porous media and the bulk conductivity/diffusivity in homogeneous media, respectively.

However, the general definition in Eq. (3.46) is not applicable to hydraulic tortuosity. As shown in Table 3.5 and Figure 3.13, the hydraulic tortuosity differs significantly from the other three physical tortuosities, especially for the porous media samples with lower porosity. Fluid permeation inside porous media not only contains the bulk movement of flow volume driven by pressure gradient, but also involves the momentum transfer from the pore wall towards the central axis of flow, so the underlying physics of hydrodynamic flow is greatly different from that of molecular diffusion, electrical conduction or heat transfer. Essentially, pore-scale fluid flow is governed by Navier–Stokes equations, which are fundamentally different from the Laplace equation that governs the other three transport phenomena. Further, due to the viscous resistance between the pore wall and fluid, the flow velocity reduces to zero at pore walls (no-slip boundary condition) and gradually increases towards the medial axes of pore channels, forming a velocity distribution as shown in Figure 3.16a. For the transport processes of molecular diffusion, electrical conduction and heat transfer, the flow intensity suddenly declines to zero at the pore walls, forming a uniform intensity distribution as shown in Figure 3.16b. Therefore, the pressure-driven fluid flow is much more hindered by the small pores inside porous media, which can make the bulk moving streamlines of fluid permeation significantly different from the transfer paths of the other three transport phenomena.

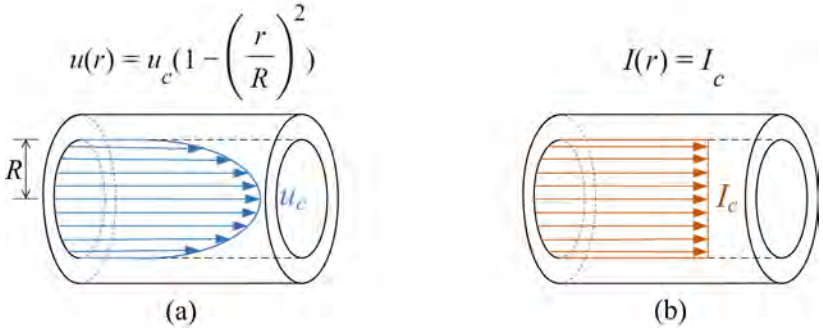


Fig. 3.16 (a) The velocity profile of Newtonian fluid flow inside a pipe with 'no-slip' boundary condition; (b) The intensity profile of molecular diffusion, electrical conduction and heat transfer inside a pipe

Another useful observation from Table 3.5 and Figure 3.13 is that the random walk simulation can provide tortuosity results close to the diffusional tortuosity values obtained from molecular diffusion simulations. However, the results from random walk simulation rely heavily on the numbers of random walkers as well as the simulation time step, especially for

porous media with complicated microstructures. The difference between molecular diffusion and random walk simulation increases with porosity. This is because the pore space is not fully experienced by the random walkers, leading to the overestimation of tortuosity. The accuracy of random walk simulation can be improved by increasing the numbers of random walkers and time steps.

It should be noted that four different numerical methods (i.e. LBM, RWM, FDM and FVM) are used here to simulate various transport phenomena at the pore scale, so numerical errors on the computations of physical tortuosities can be different. For example, all numerical simulations directly use image voxels as the basic computing elements, but the connectivity between adjacent pore voxels follow different criteria during the discretization of governing equations. Basically, the 18-neighbor-connectivity criterion is used in the LBM simulation because of the adoption of the D3Q19 lattice arrangement, while the RWM, FDM and FVM simulations follow the 6-neighbor-connectivity criterion which only permits direct exchange of dependent variables occurring between adjacent pore voxels with common faces. However, according to the comparison results in previous studies [211, 98, 433, 96, 439], as well as our computational experiences (as shown in Figure 3.13, 3.26 and 3.29), numerical errors are very tiny and negligible for the high-fidelity simulations performed on 3D high-resolution images of porous media samples. Therefore, the above analysis results on physical tortuosities are reliable, although different numerical simulation methods are used for this comparison study. Besides, image resolution and image size are important influencing factors for physical tortuosity calculation as well, which will be further discussed in Section 3.6.

3.4.3 Performances of theoretical/phenomenological models

It is evident from Table 3.5 and Figure 3.13 that there are close inverse correlations between porosity and tortuosity results computed from different image analysis methods or physical simulations. In this subsection, five different porosity-tortuosity models are applied to estimate the tortuosity results of the set of Fontainebleau sandstone samples, as expressed in Eqs. (3.1) to (3.5). As discussed in Section 3.1.2, these porosity-tortuosity relations usually contain empirical parameters or uncertain coefficients that are difficult to accurately determine, and irrational values of them may lead to great errors in tortuosity estimation. Recommended values of these adjustable parameters are found through a literature survey, and they are adopted here to assess the performances of theoretical/phenomenological models on tortuosity evaluation.

For Bruggeman relation or Archie's law (as expressed in Eq. (3.1)), its empirical exponent α usually ranges from 1.3 to 2.5 [311], and there are various methods to determine α for a particular porous medium [153]. However, reliable values of α are often determined from

experimental measurements or numerical simulations [438]. The value of α is assigned to be 1.5 in the standard form of Bruggeman relation to estimate diffusional tortuosity of porous media [438]. Besides, Mota et al. (2001) [316] found α was equal to 1.8 for granular beds of spherical particles, from which the estimated hydraulic tortuosity well agreed with the experimental data. As to Comiti's model (Eq. (3.2)), its empirical constant p varies greatly for plates with different height-side ratios [94]. Mauret and Renaud (1997) [303] found p=0.49 was appropriate for a capillary model of granular and fibrous beds to estimate hydraulic tortuosity, which was later confirmed by an experimental study. For the empirical model (Eq. (3.3)) proposed by Iversen and Jørgensen (1993) [210], its coefficient q=2 was found to be appropriate for diffusional tortuosity of sandy sediments through diffusion experiments. Considering the non-spherical grains, the sphericity factor ξ involved in Lanfrey's model (Eq. (3.5)) should be smaller than 1 [261] to estimate geometrical tortuosity for natural porous media. Here, α= 1.5, 1.8 and 2.5 (the upper limit), $p = 0.49$, $q = 2$ and $\xi = 0.9$ are adopted for the above porosity-tortuosity relationships to evaluate the tortuosity values of the set of Fontainebleau sandstone samples, and the results are recorded in Table 3.5 and plotted Figure 3.13.

In general, none of these five models can provide tortuosity values that closely match the results computed from any image analysis methods or physical simulations, which is in accordance with the remarks on theoretical/phenomenological models in Section 3.1.2. Comiti's model, Iversen's model and Du Plessis's model tend to underestimate physical tortuosities (compared with the physical simulation results), especially for the porous media samples with low porosity, while Lanfrey's model greatly overestimates the geometrical tortuosity values of the set of Fontainebleau sandstones. The primary reason of such model failures is that both model constructions and corresponding parameter calibrations are based on porous media with high porosity (such as sandy sediments, granular beds and fibrous beds). Therefore, it is no wonder that the above models can rarely provide reliable tortuosity estimations for natural porous rocks with relatively low porosity. As shown in Figure 3.17, a large group of pore media samples with a wide range of porosity are used to test these porosity-tortuosity models, and the results further confirm the above viewpoint.

As can be seen in Table 3.5, Figure 3.13 and 3.17, three different values are assigned to the exponent α in Bruggeman relation or Archie's law to estimate tortuosity values of Fontainebleau sandstones. It seems that Bruggeman relation or Archie's law that treats tortuosity as an exponential function of tortuosity is more reasonable than the other four models, and it is possible to accurately evaluate physical tortuosities by selecting an appropriate value for the Bruggeman exponent α. To better correlate Bruggeman model with tortuosity, an additional scaling factor β was introduced by researchers, and Bruggeman

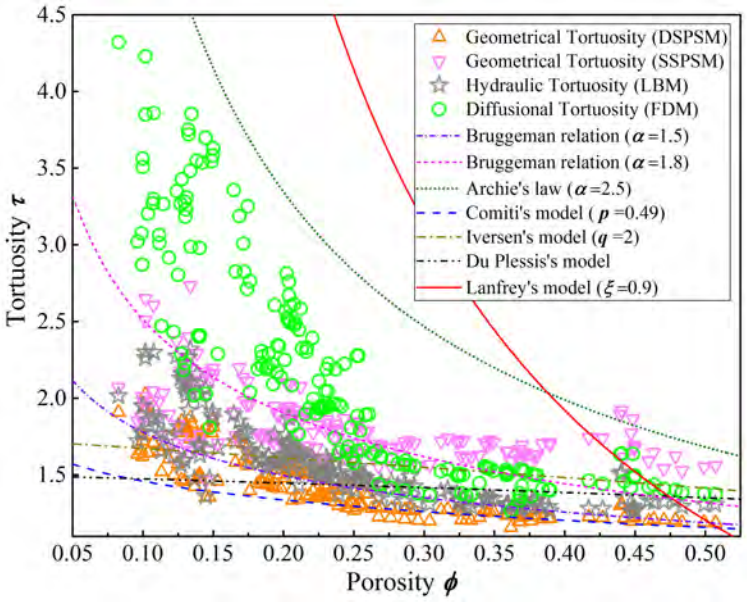


Fig. 3.17 Tortuosity results calculated from different image analysis methods, physical simulations and porosity-tortuosity models for a group of porous media samples with a broad range of porosity (detailed information about the porous media samples are given in Section 3.5)

model was extended to be $\tau = \beta\phi^{\alpha}$ [438]. The crucial point is how to determine the values of adjustable parameters α and β that directly affect the estimation accuracy of tortuosity. It has been proved that both α and β depend on the analyzed pore microstructures, and they vary significantly even for a small class of porous media, which leads to conflicting results on the validity of Bruggeman relation in a broad range of studies [153, 438]. Therefore, the tortuosity predictions from Bruggeman relation are not always consistent with experimental data or numerical simulation results, and researchers usually alter the adjustable parameters to fit the experimental or numerical values.

It is understandable that different porous media with the same porosity may have different transport path lengths, so tortuosity cannot be a function of porosity only, especially for low-porosity rocks that possess extremely complicated microstructures with a large number of isolated and dead-end pores. Natural porous rocks are usually very different from and much more complex than the spherical structures that closely follow the initial hypotheses of porosity-tortuosity relationship derivation. The aforementioned porosity-tortuosity models do not take account of other microstructural characteristics that have considerable effects on tortuosity evaluation, such as pore size, shape, surface, distribution and connectivity. In summary, it can be stated that the porosity-tortuosity models are appropriate and reliable for porous media that are similar to the pore microstructures used to construct corresponding

relationships and calibrate adjustable coefficients, but they are less suitable for complicated pore networks without adequate characterization.

3.5 Relationships between geometrical and physical tortuosities

It is usually compute-intensive to evaluate physical tortuosities via pore-scale simulations of transport processes, but geometrical tortuosity can be obtained much cheaply from image analysis. Therefore, geometrical tortuosity is commonly used as the substitute for physical tortuosities to model and predict transport properties. However, due to conceptual differences, great discrepancies exist between geometrical and physical tortuosities. As shown in Figure 3.13, none of the image-based methods is able to provide a geometrical tortuosity result that is close to any one of the physical tortuosities. In order to bridge such gaps, we established quantitative relationships between geometrical and physical tortuosities in this section, so that tortuosity can be better used in porous media research.

As the porosity varies, a strong correlation exists between geometrical and physical tortuosities, which can be observed in Figure 3.13. To confirm this trend, a large set of porous media (mainly geological materials) with diverse morphologies are used to explore the relationships between geometrical and physical tortuosities, including sandstones, sand packs, carbonate rocks (limestone), synthetic silica, bead packs, and others. The 3D micro-CT images of these porous media samples are collected from several publicly shared data sources, and due to page limit, only a part of them are shown in Figure 3.18 in 2D form. The raw micro-CT images are in gray-scale to prepare for image analysis and physical simulation. First they are processed and segmented into binary forms by using ImageJ [207], a popular image processing tool in the digital rock physics community. As illustrated in Figure 3.19, 3D micro-CT images of four representative porous media samples are segmented using Otsu's method [343].

The micro-CT databases employed in this work include 120 digital microstructure samples. These raw images are representative covering a diverse set of porous media, but more samples are still required in order to develop phenomenological models to link geometrical and physical tortuosities. A stochastic microstructure reconstruction method has been newly developed using deep neural networks, which will be detailedly explained in Chapter 4. This method is capable of quickly generating 3D digital microstructure samples by preserving the statistical equivalence, long-distance connectivity and transport properties of the training images. Here, the segmentations of the above micro-CT models are used as the guides to

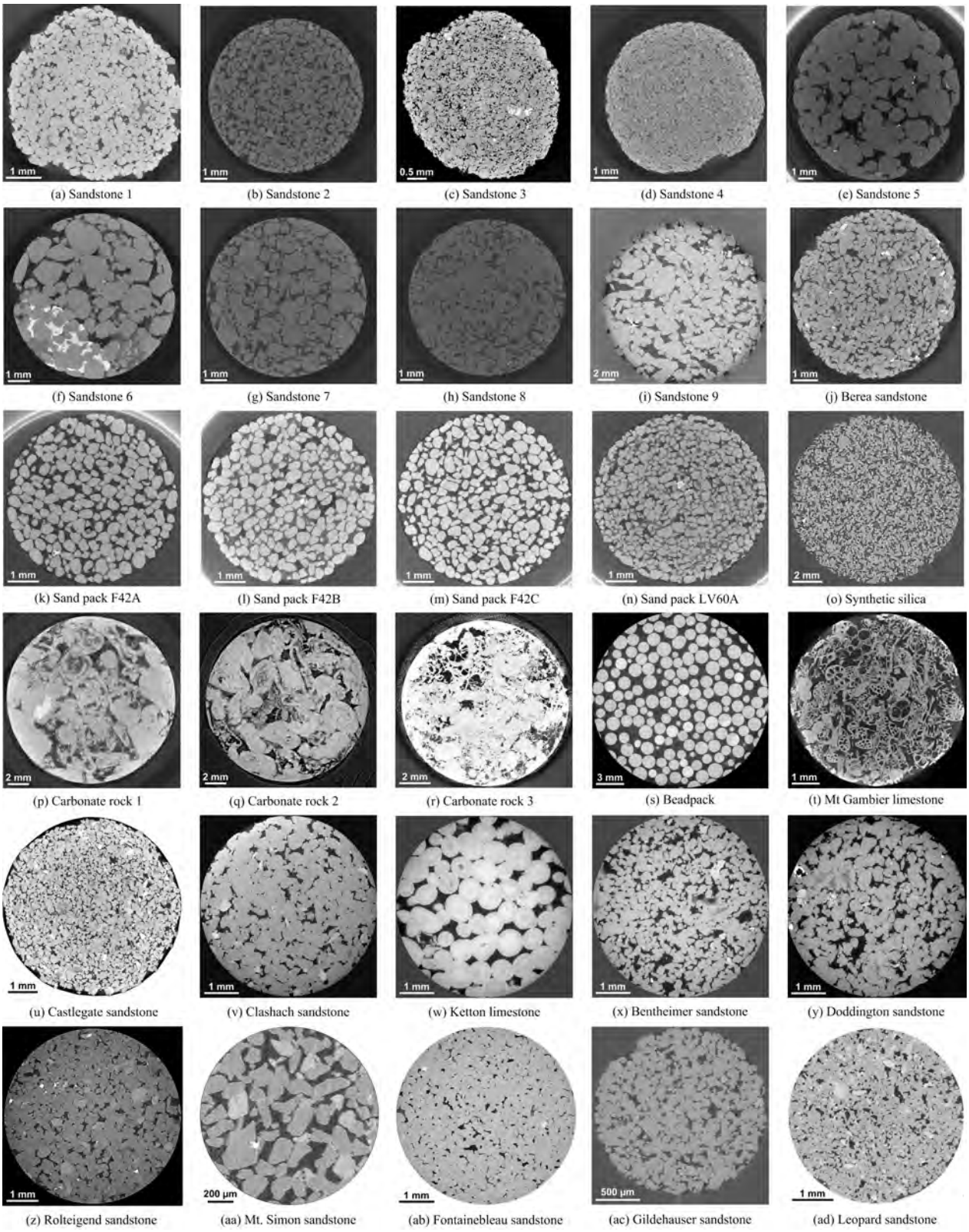


Fig. 3.18 The 2D cross-sections of 3D micro-CT images of different porous media samples (pore is shown in dark): data in (a)-(r) are collected from [120]; data in (s)-(u) are collected from [455]; data in (v)-(y) are collected from [401]; data in (z) is collected from [386]; data in (aa) is collected from [180]; data in (ab) is collected from [262]; data in (ac) is collected from [63]; data in (ad) is collected from [36]

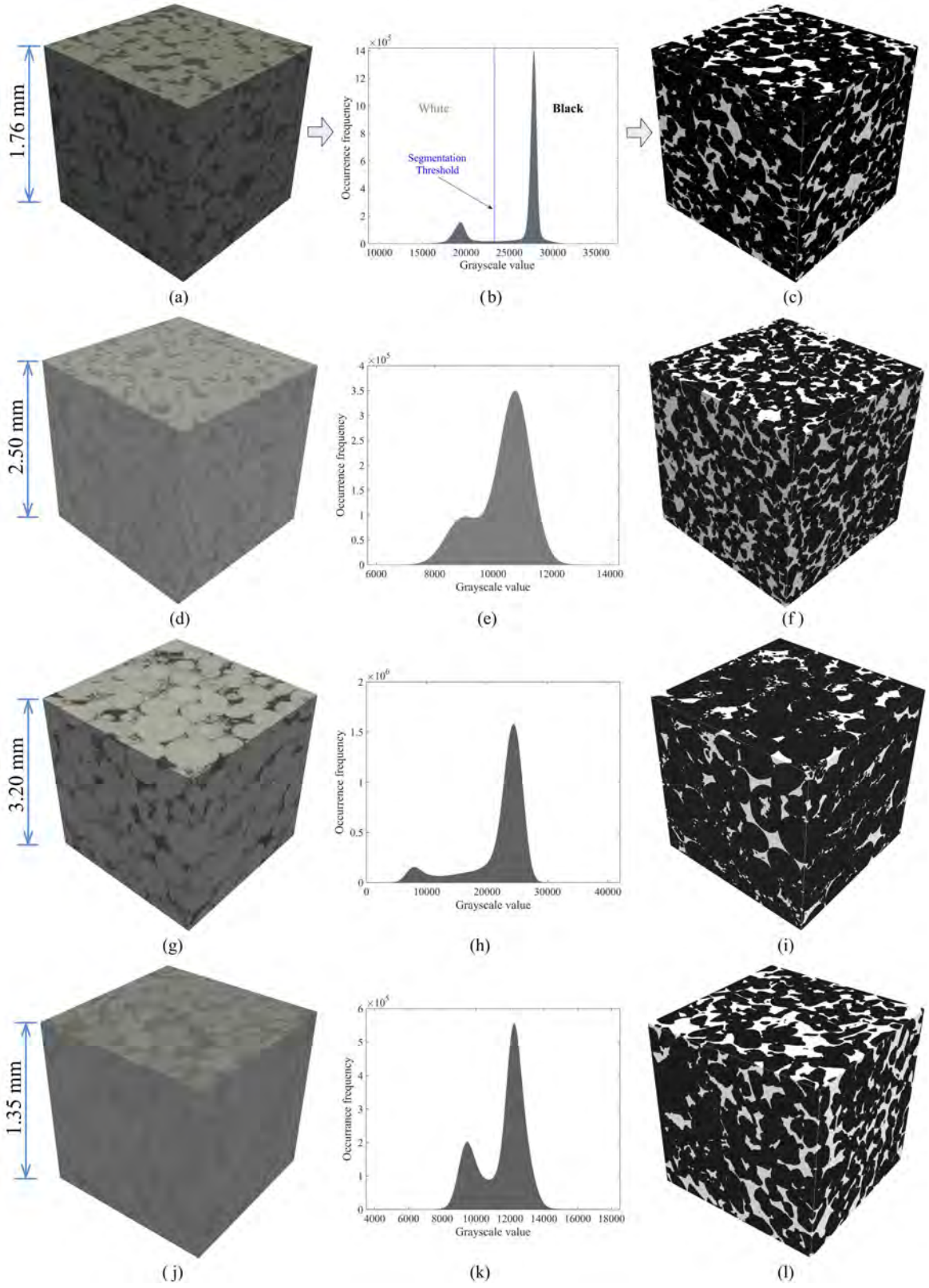
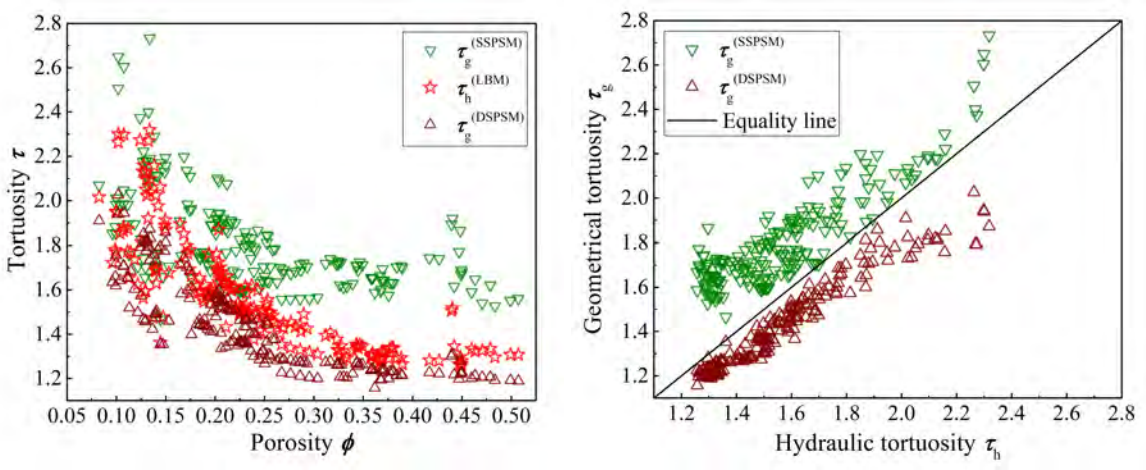


Fig. 3.19 Segmentation of micro-CT images through Otsu's method: (a) Gildehauser sandstone, (d) Bentheimer sandstone, (c) Ketton limestone and (d) Mt. Simon sandstone; (b), (e), (h) and (k) are the histograms of grayscale values of corresponding micro-CT images; (c), (f), (i) and (l) are the segmented images in binary form



(a) (b)

Fig. 3.20 (a) The hydraulic tortuosity obtained with LBM simulation and the geometrical tortuosities obtained from SSPSM and DSPSM for porous media at different porosity; (b) Comparison between geometrical and hydraulic tortuosities

reconstruct new porous media samples. In total, 200 microstructure samples (120 samples from raw micro-CT images plus 80 reconstructed samples) are used in this work to investigate the relationships between geometrical and physical tortuosities, with the porosity varying from 8.61% to 50.72%. Specifically, the relationship between geometrical tortuosity and hydraulic tortuosity is described in Section 3.5.1, and the relationship between geometrical tortuosity and diffusional/electrical/thermal tortuosity is described in Section 3.5.2.

3.5.1 Prediction model for hydraulic tortuosity

As shown in Figure 3.13, the hydraulic tortuosity lies in between the geometrical tortuosity results calculated from the DSPSM and SSPSM. This phenomenon can be easily understood by comparing the shortest pore channels with the effective flow path in Figure 3.2 and 3.4. Thus, it is natural to use the DSPSM and the SSPSM results to develop a better approximation to hydraulic tortuosity, thereby avoiding the costly fluid flow simulations. Indeed, the same trend is also observed for the whole set of 200 porous media samples, as shown in Figure 3.20, where the geometrical tortuosity results calculated from the DSPSM and the SSPSM form nicely the upper and lower bounds of hydraulic tortuosity determined from LBM simulations.

The above observation suggests that the average length of fluid streamlines in a porous medium is bounded by the shortest paths identified by the DSPSM and SSPSM. Based on the above data and taking into account the variation with respect to porosity, we propose the

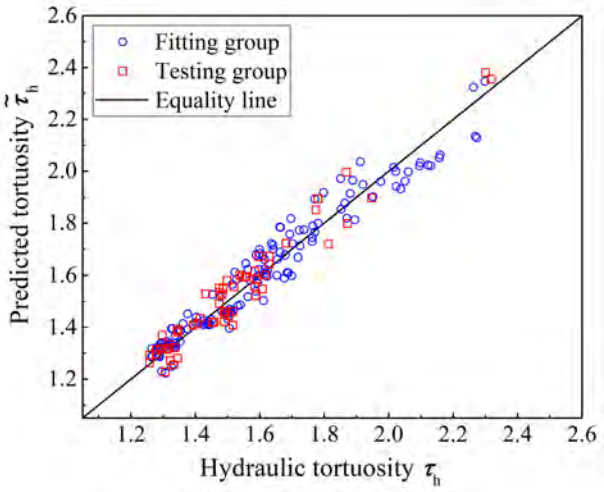


Fig. 3.21 Comparison between the hydraulic tortuosity predicted using the proposed model and the reference value obtained from LBM simulations

following phenomenological model to predict hydraulic tortuosity:

$$\widetilde{\tau}_h = a\phi^b \tau_g^{(\mathrm{DSPSM})} + (1 - a\phi^b)\tau_g^{(\mathrm{SSPSM})} \tag{3.47}$$

where $\widetilde{\tau}_h$ denotes the predicted hydraulic tortuosity, a and b are empirical coefficients, and $\tau_g^{(\mathrm{DSPSM})}$ and $\tau_g^{(\mathrm{SSPSM})}$ are the geometrical tortuosity results calculated from the DSPSM and SSPSM, respectively.

The set of porous media with 200 samples are randomly divided into the fitting group (70%) and the testing group (30%). The coefficients a and b are identified as 1.3100 and 0.5435 respectively, through fitting the tortuosity and porosity data extracted from the porous media samples in the fitting group. The mean error and the largest error of the fitted model are 2.99% and 7.41%, respectively. The fitted model is applied to the porous media samples in the testing group, to predict the hydraulic tortuosity values based the geometrical tortuosity and porosity data. As shown in Figure 3.21, the predictions $\widetilde{\tau}_h$ from the above model are in good agreement with the results τ_h obtained from LBM simulations, and the mean error and the largest error of prediction are 3.11% and 7.19%, respectively.

3.5.2 Prediction model for diffusional tortuosity

As discussed in Section 3.4.2, diffusional, electrical and thermal tortuosity are almost identical, and they are unified into a general form (as illustrated in Eq. 3.46). Here, we take diffusional tortuosity as the representative to establish the prediction model. The geometrical tortuosity $\tau_g^{(\mathrm{DSPSM})}$ calculated by the DSPSM is selected for the task, because it shows a

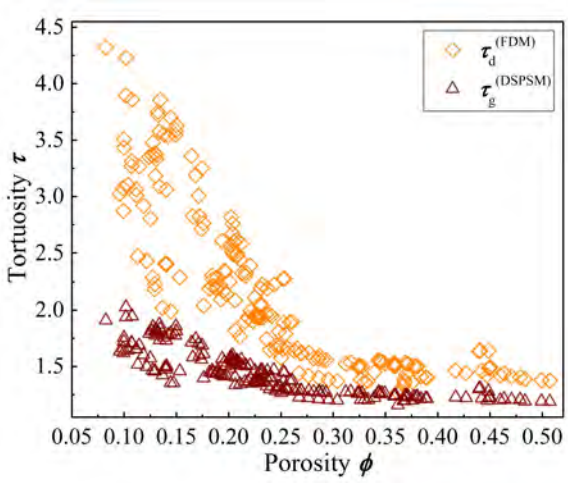


Fig. 3.22 The geometrical and diffusional tortuosities of porous media at varying porosity

stronger correlation to diffusional tortuosity and does not involve any extra image preparation that may cause errors. As shown in Figure 3.13, the diffusional tortuosity τ_{d} values are always larger than the geometrical tortuosity values $\tau_{\mathrm{g}}^{(\mathrm{DSPSM})}$, and their difference decreases with the growth of porosity ϕ. This observation also holds for other porous media samples with wider range of porosity, as shown in Figure 3.22. Here, an empirical model is going to be established to link geometrical tortuosity to diffusional tortuosity. Although the gap between $\tau_{\mathrm{g}}^{(\mathrm{DSPSM})}$ and τ_{d} is related to porosity, it is found that τ_{d} can be accurately estimated from $\tau_{\mathrm{g}}^{(\mathrm{DSPSM})}$ without considering porosity ϕ through the following empirical model:

$$\widetilde{\tau}_{\mathrm{d}} = m(\tau_{\mathrm{g}}^{(\mathrm{DSPSM})})^{n} \tag{3.48}$$

where m and n are empirical constants.

The same fitting and testing group of porous media samples are used here. The values of m and n are evaluated to be 0.9505 and 2.2310 respectively, though fitting the tortuosity data extracted from the porous media samples in the fitting group. The mean error and the largest error of the fitted model are 3.58% and 8.50%, respectively. Applying the fitted model to the testing group, the predicted diffusional tortuosity $\widetilde{\tau}_{\mathrm{d}}$ are in good agreement with the numerical simulation result τ_{d}, as demonstrated in Figure 3.23. The mean error and the largest error of the predictions are 3.07% and 7.71%, respectively.

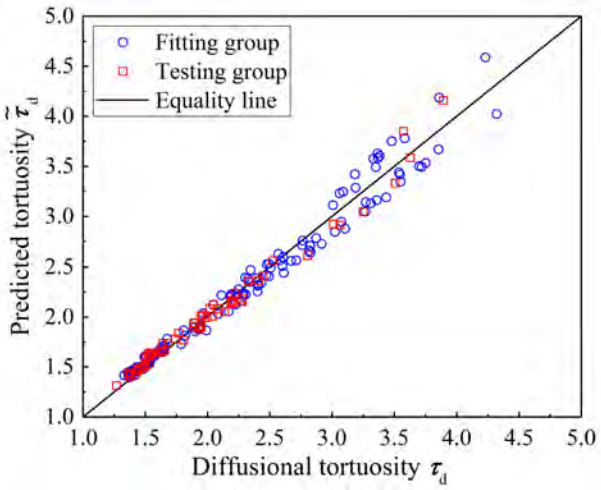


Fig. 3.23 Comparison between the predicted diffusional tortuosity obtained from the proposed model and the reference value obtained from molecular diffusion simulation

3.6 Issues on REV size and resolution

REV size and resolution (voxel size) are two critical issues that need to be carefully considered when computing tortuosity values of digital microstructures through image analysis or physical simulation. Generally, reliable pore-scale analysis requires large digital microstructure with high resolution. To serve as a good statistical representative for the specific porous medium, the REV sample needs to be sufficiently large [103], and to provide accurate pore space geometry close to the reality, the digital microstructure must be at a sufficiently high-resolution level [142].

Through image analysis or physical simulation, accurate tortuosity results can be obtained from large REV samples with high resolution [51, 353, 165]. However, it is still a challenging task to deal with 3D REV samples with super-high resolutions, because they can contain hundreds of millions (or even billions) of voxels. High-performance computing platforms are often required to perform pore-scale analyses on these digital microstructures. In practice, the raw micro-CT images are often rescaled to a lower resolution, in order to reduce the computational cost. The resolution reduction is particularly common when using GPU platforms, as a finer REV model often exceeds the memory capacity of a single graphics card [220]. Therefore, the choice of image resolution of the REV sample is often a compromise between accuracy requirement and computational complexity.

Naturally, the critical REV size and resolution for porous media depend on the characteristic length of morphological features. In this work, the correlation length is selected as the characteristic length to study the effects of image size and resolution on tortuosity evaluation, because it is related to both mean pore radius and hydraulic radius [234]. Specifically, the

correlation length l is defined as the distance where the two-point correlation function [441] dwindles to the asymptotic value of ϕ^2 [92], as illustrated in Figure 3.25.

3.6.1 REV size

A Fontainebleau sandstone with porosity about 14.00% is selected to study the effect of sample size on tortuosity evaluation. Its 3D micro-CT image with a resolution of 7.324 μm [188] is segmented using ImageJ [207], as shown in Figure 3.24. The two-point correlation function can be extracted from the binary segmentation, and the correlation length l is measured as 17.50 voxels, as demonstrated in Figure 3.25. To study the influence of sample size on tortuosity calculation, the 3D digital microstructure is cut into cubic subsamples with side lengths L of 5, 10, 15, 20 and 25 times of the correlation length l, as shown in Figure 3.24b.

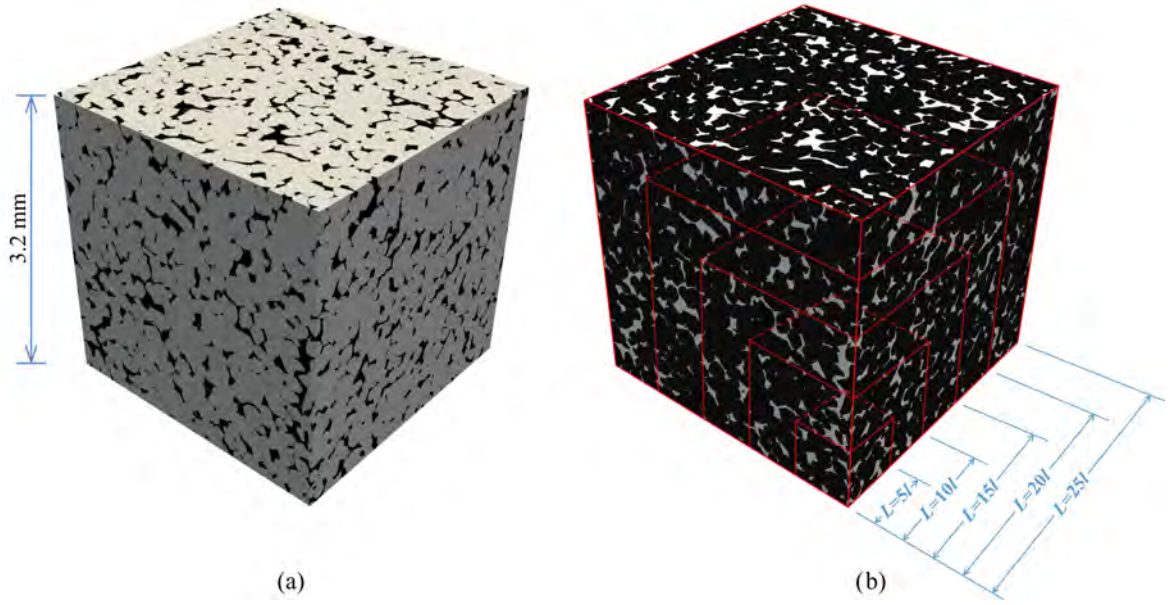


Fig. 3.24 (a) The micro-CT image of a Fontainebleau sandstone sample with resolution of 7.324 μm; (b) The binary segmentation of the micro-CT image, and it is partitioned to subsamples of different sizes (L is the side length of cubic sample, and l is the correlation length)

Image analysis and physical simulation are performed on these subsamples to compute geometrical and physical tortuosity, and the results are plotted in Figure 3.26. It can be observed that both geometrical and physical tortuosity values fluctuate greatly for small samples, especially the diffusional tortuosity computed by the RWM where the results fluctuate from 2.37 to 3.84. As the sample size increases, both geometrical and physical

tortuosities converge gradually. The larger fluctuation observed in smaller samples can be attributed to two causes: (1) the morphological characteristics of smaller samples are not statistically representative and (2) the boundary effects of various computing algorithms or numerical methods on tortuosity calculations are more visible for smaller samples. It is also observed that the tortuosity results computed from all methods become stable when the sample size L increases to $20l$, indicating the critical REV size L_c for tortuosity calculation:

$$L_c = 20l \tag{3.49}$$

This critical REV size $L_c = 20l$ can also be a valid reference for other types of porous media. But more significantly, the correlation length extracted from the two-point correlation function is demonstrated to be an effective characteristic length to determine the REV size for tortuosity evaluation.

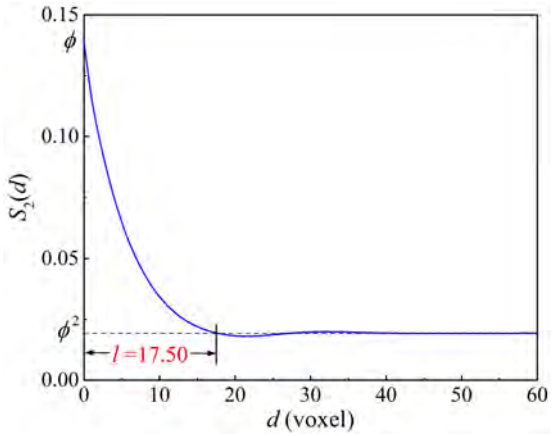
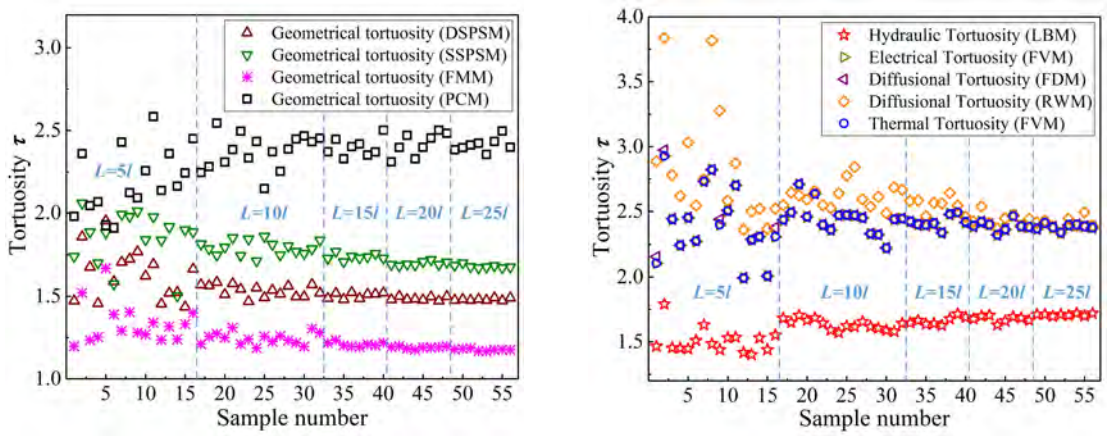


Fig. 3.25 The two-point correlation function $S_2(d)$ and correlation length l of the Fontainebleau sandstone sample with resolution of 7.324 μm

3.6.2 REV resolution

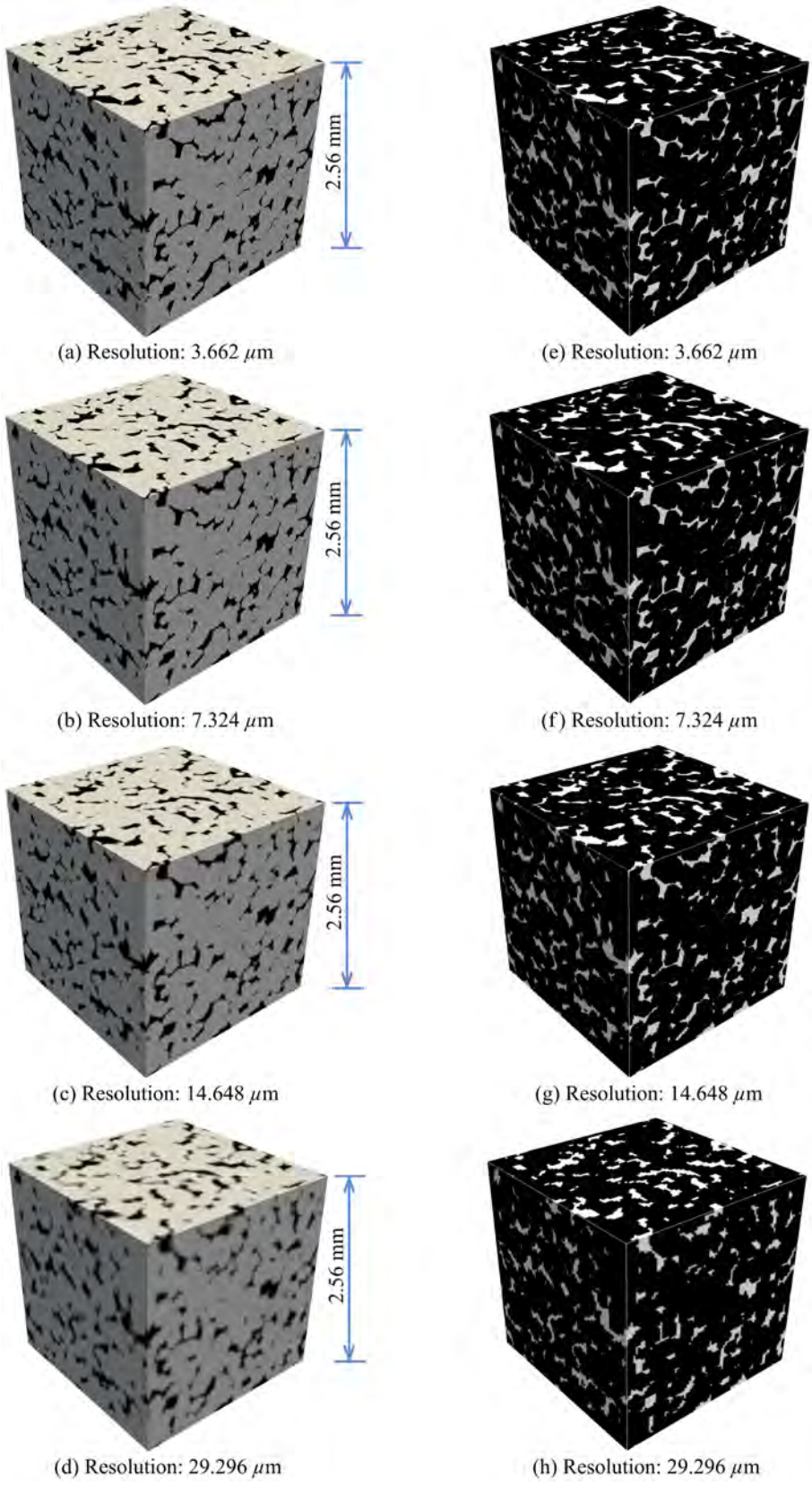
The 3D micro-CT images at four resolution levels (e.g. 3.662, 7.324, 14.648 and 29.296 μm) of the Fontainebleau sandstone [188] are used to study the resolution effect on tortuosity evaluation, as shown in Figure 3.27. To avoid the sample size effect, these micro-CT images are segmented and then partitioned into subsamples with side length $L = 20l$. The correlation length l of the Fontainebleau sandstone is about 128.170 μm, which is equal to 35.00, 17.50, 8.75 and 4.38 voxels respectively for the digital microstructures from fine to coarse. The correlation length l can be directly measured from the two-point correlation function curves, as illustrated in Figure 3.28.



(a) Geometrical tortuosity (b) Physical tortuosity

Fig. 3.26 The effects of sample size on the calculation results of different tortuosities

Performing image analysis and pore-scale simulation on the digital microstructures with different resolutions, the geometrical and physical tortuosity values are calculated, and the results are plotted in Figure 3.29. The REV resolution is found to have different effects on tortuosity evaluations of the four image-based methods. The geometrical tortuosity results calculated by the DSPSM and FMM are sensitive to image resolution, and they show similar increase pattern as voxel size becomes larger. There are two primary sources of error caused by the resolution effect: geometrical error and image analysis error. The difference between the acquired digital microstructure and the real microstructure is termed as the geometrical error, and this error increases as the digital microstructure becomes coarse, due to losing geometrical information during the pixel/voxel binning. Bad pore-connectivity inside the digital microstructure with low resolution usually results in a larger value of geometrical tortuosity. Image analysis error refers to the difference between the shortest pore channel identified by the algorithm and the actual shortest pore channel inside a porous microstructure. Usually, the shortest pore channel identified from a fine image is 'shorter' than that obtained from a coarse image, as illustrated in Figure 3.4. On the contrary, the SSPSM and PCM are less influenced by the REV resolution, and the geometrical tortuosity only slightly declines as the image resolution becomes lower. As described in Section 3.2.3, the SSPSM searches the shortest pore channel from the medial axis skeleton of the pore space, which greatly reduces the image analysis error discussed above. The PCM is very simple, and it does not consider pore-connectivity and does not involve a path searching algorithm either. As the PCM only takes into account the pore centroid position on each layer, it is usually less affected by the image resolution. Generally, the resolution effect on the four image-based methods is very much limited when the correlation length l is larger than 17.5 voxels, as shown in Figure



(a) Resolution: 3.662 μm (e) Resolution: 3.662 μm

(b) Resolution: 7.324 μm (f) Resolution: 7.324 μm

(c) Resolution: 14.648 μm (g) Resolution: 14.648 μm

(d) Resolution: 29.296 μm (h) Resolution: 29.296 μm

Fig. 3.27 The micro-CT images and corresponding binary segmentations of a Fontainebleau sandstone sample at four different resolution levels

3.29a. Therefore, for the calculation of geometrical tortuosity, it is recommended to use digital microstructures with the voxel size smaller than $l/17.5$.

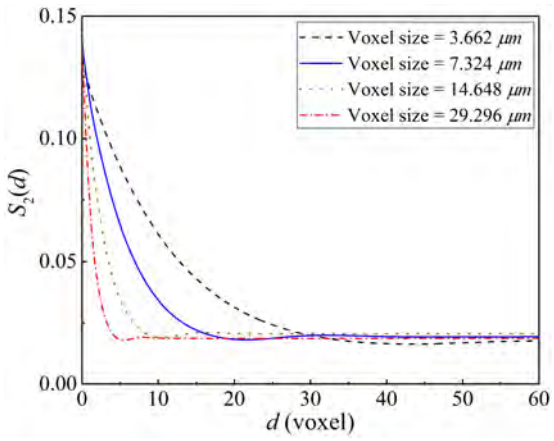
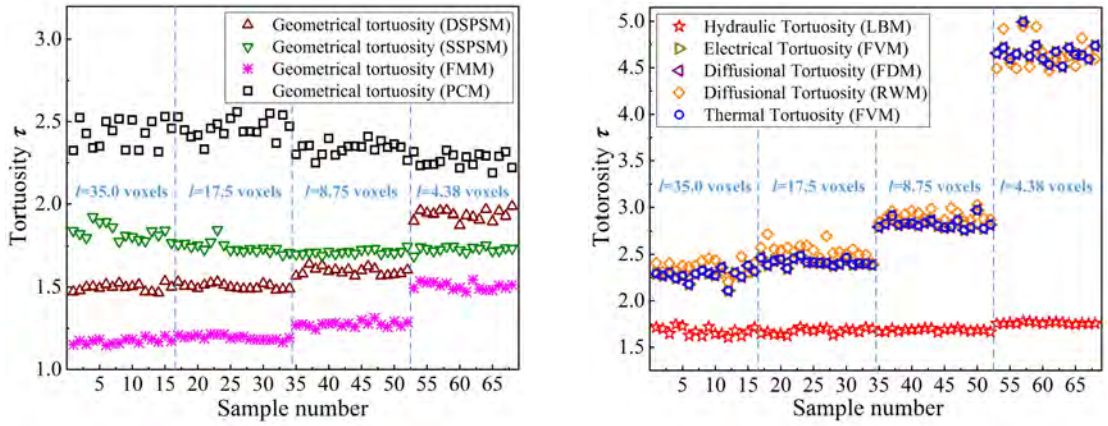


Fig. 3.28 The two-point correlation functions $S_2(d)$ of digital microstructure samples at four different resolution levels



(a) Geometrical tortuosity (b) Physical tortuosity

Fig. 3.29 The effects of image resolution on calculation results of different tortuosities

As observed in Figure 3.29b, the hydraulic tortuosity evaluation is quite stable when the image resolution changes, and reasonably accurate results can be obtained on coarse images. Although hydraulic tortuosity is less affected by REV resolution, it has been proven that the resolution effect on the hydraulic conductivity (or permeability) is remarkable [142]. For the evaluation of diffusional, electrical and thermal tortuosity, the resolution effect is significant. As can be seen in Figure 3.29b, the diffusional tortuosity results increase with the voxel size. The poor pore-connectivity inside the coarse image leads to greater retardation effects on the transport processes of molecular diffusion, electrical conduction and heat transfer, thereby

overestimating the tortuosity values. Physical tortuosities are estimated from numerical simulations where image voxels are directly used as the discretization elements. Therefore, the increase of voxel size may introduce extra numerical errors. For example, when LBM is used to simulate fluid flow passing through a porous medium, the fluid-solid interface will be 'misidentified' if a suboptimal value is selected for the single relaxation time, and this error will be greatly amplified on a coarse grid [142]. As shown in Figure 3.29b, the resolution effect on the physical tortuosity calculation is generally negligible when the REV resolution is higher than R_c:

$$R_c = l/17.5 \tag{3.50}$$

This critical resolution level $R_c = l/17.5$ can play as an effective reference for other types of porous media as well, but the key here is that the correlation length can be used as the measure to select a proper image resolution level for tortuosity evaluation.

3.7 Discussion and concluding remarks

3.7.1 Discussion

The concept of tortuosity is used to model and predict transport properties of porous media, such as permeability, effective diffusivity, electrical conductivity and thermal conductance. However, it seems that the physical tortuosities evaluated through numerical simulations of various transport phenomena are redundant, because corresponding transport properties can be directly estimated from physical simulations without the requirement of such intermediate variables. Inverse computation of physical tortuosity from transport property (as defined in Eq. (3.17), (3.23), (3.32) and (3.40)) just runs counter to the original intention of introducing tortuosity. Massive efforts have been made to directly relate tortuosity to porosity, so as to derive explicit models for tortuosity evaluation, as discussed in Section 3.1.2. The broad variety of porosity-tortuosity models usually lack general applicability for natural porous media with complicated pore networks, due to the absence of consideration of other important microstructural characteristics, which has been deeply discussed in Section 3.4.3. Besides, reliable calibrations of the empirical parameters or uncertainty coefficient involved in these porosity-tortuosity relations often require experimental measurement or numerical simulation, which heavily discounts their values in practical applications.

Geometrical tortuosity can be an effective alternative to physical tortuosities to fulfill the intended purpose in modeling and predicting transport properties, because it is a purely microstructural feature of porous media that can be quickly determined through image analysis. However, no general standard exists in extracting geometrical tortuosity from the

digital microstructures of porous media, and thus uncertainties could be brought into image analysis, due to the different options in terms of image preparation, definition of the effective flow path, path searching algorithm, pixel/voxel connectivity criterion, distance metric, image resolution and other factors. Although geometrical tortuosity is conceptually sound, inconsistent results are often obtained from different image-based calculation methods, as discussed in Section 3.4.1. A lack of reference makes it have no way to assess the performances of different image-based methods and judge the rationality of geometrical tortuosity results. According to our own computational experience, the DSPSM and SSPSM using Dijkstra's algorithm are recommended to extract geometrical tortuosity from the digital porous media sample, because they can efficiently and stably provide geometrical tortuosity results that are the lower and upper approximations of hydraulic tortuosity respectively, as discussed in Section 3.5.1.

Tortuosity cannot simply be a characterization of pore space itself just from the geometric perspective, but it should also be related to the actual streamlines of transport flows. For electrical, diffusional and thermal flow, the transfers of charge, mass and energy depend on the available cross-sectional area of the transport path, but geometrical tortuosity doesn't take consideration of the constrictions or bottlenecks along the pore channel. As to hydrodynamic flow, the movement of fluid volume relies on both the pore size and cross-sectional shape along the permeation route, and momentum exchange is also involved into it due to the viscous effect, but these important factors are never taken into account in geometrical tortuosity. In view of this, we developed two phenomenological models based on a large group of porous media samples with a broad range of porosity in Section 3.5, aiming to bridge the gaps between geometrical and physical tortuosities. The physical tortuosity values estimated from the geometrical ones through the proposed models can be more appropriate for transport property prediction.

Along the lines of the above discussion, two issues on tortuosity may need to be further resolved: (1) If geometrical tortuosity is an intrinsic microstructural property, it should have a unique value for a particular porous medium. How to standardize the determination procedure of geometrical tortuosity to yield a unique value is the first issue. (2) Geometrical tortuosity is not exactly equivalent to physical tortuosities, so how to better understand the gaps between them is the second issue. Only after that, geometrical tortuosity together with other supplementary microstructural characteristics can completely take the role of physical tortuosities in modeling and predicting transport properties.

3.7.2 Concluding remarks

From concepts to evaluation methods, this study systematically investigates the tortuosity of porous media. The study focuses on modern approaches to evaluating the tortuosity of porous media based on digital microstructures, covering a wide range of image analysis and numerical simulation methods. The various types of tortuosity are broadly categorized into two groups: geometrical and physical tortuosity, depending on whether the analysis object is the pore space itself or the transport flow through the pore space. Physical tortuosity is further classified into four types: hydraulic, electrical, diffusional, and thermal tortuosity, corresponding to the specific transport process.

Different physical tortuosities are often used interchangeably in the literature, ignoring their distinct differences in definition and evaluation method. To clarify the situation, hydraulic, electrical, diffusional, and thermal tortuosities are examined by using numerical simulations of the corresponding transport processes in porous media. It is found that electrical, diffusional and thermal tortuosity are similar in value, but they visibly differ from hydraulic tortuosity. This is because the underlying physics of molecular diffusion, electrical conduction, and heat transfer are equivalent, and the steady-states of these transport phenomena are described by Laplace's equation. This is distinctly different from Navier-Stokes equations that govern the hydrodynamic fluid flow in porous media. Therefore, electrical, diffusional, and thermal tortuosity can be unified to be one retarding factor, but it cannot be interchangeably used with hydraulic tortuosity.

The increasing availability of high-quality digital microstructures makes it possible to perform high-fidelity numerical simulations for transport flows in porous media. However, the computational cost of pore-scale simulation is high and it increases rapidly with the size of digital REVs. In contrast, geometrical tortuosity can be cheaply evaluated from image analysis, making it desirable to explore an alternative approach to evaluating physical tortuosities without involving costly physical simulation. It is found that the commonly-used image analysis methods do not provide consistent geometrical tortuosity results and they all differ greatly from the physical tortuosity values. To overcome this problem, two phenomenological models are proposed to evaluate physical tortuosities using the geometrical tortuosity, one for the hydraulic tortuosity prediction, and the other for the diffusional/electrical/thermal tortuosity prediction. Independent checks are performed to confirm the accuracy of these predictions.

Finally, the effects of image size and resolution on the tortuosity evaluation are analyzed as well. The correlation length extracted from two-point correlation function is found to be an effective measure to determine the critical thresholds for REV size and resolution.

It's worth reminding that tortuosity plays an important role in subsequent studies of this thesis, including stochastic microstructure reconstruction (Chapter 4 and 5), error correction modeling due to the resolution effect (Chapter 6), and the fundamental investigation of the microstructure-permeability relationship (Chapter 7).

Chapter 4

Statistical characterization and reconstruction of heterogeneous microstructures using deep neural network

4.1 Introduction

Random heterogeneous materials comprised of at least two distinct phases (including void) are ubiquitous in engineering and natural environment, and typical examples include composites, alloys, concretes, soils and rocks. Their physical properties (i.e. transport, elastic, and conductive properties) usually exhibit strong uncertainty, due to the random distributions of different phases and the phase discontinuity on the interfaces [443]. Microstructural characteristics of heterogeneous media are considered to be the key to understanding their macroscopic behaviors [34, 167]. Modern microscopy imaging techniques have been able to digitize the geometry of a heterogeneous medium, usually called a microstructure, into a 2D or 3D image at different resolution levels and various length scales. The digital microstructure can be used for a variety of simulation or diagnostic purposes.

Although useful insight into heterogeneous media can be derived from modern microscopy imaging techniques, it remains a critical challenge to effectively link microstructural features to material properties. Such knowledge is of great value for predicting material properties from the observable microstructures [443]. Additionally, microstructural characteristics can also be linked to material formation processes, which can help to design and manufacture materials to possess desirable properties.

In practice, material properties are usually measured by experiments on physical samples or estimated from numerical simulations on digital microstructure samples. However, physical sample acquisition, such as rock sample drilling, is usually expensive, and microscopy imaging is also of high cost [480]. Without sufficient physical (or digital) samples, direct experiments (or numerical simulations) are unable to correctly assess heterogeneous media due to their stochastic nature. The complete computational model of a heterogeneous medium is far beyond a small number of microstructure samples, and it should cover representative samples with all possible configurations. In other words, the complete computational model is an ensemble of representative/statistical volume elements that share the same averaged characteristics. This situation promotes researchers to find an alternative approach to understanding the microstructure-property relationships with limited availability of physical (or digital) raw samples [224].

A reliable way to quantitatively characterize the complex microstructure is an essential prerequisite to establishing effective microstructure-property relationships. Stochastic characterization provides a statistical description of microstructural features, where the stochastic morphology of complex microstructures is reduced to a small set of descriptors related to material properties [443]. As mentioned above, the number of raw microstructures is usually limited, due to the high cost of advanced microscopy imaging techniques, which brings a practical difficulty for accurate characterization. Therefore, stochastic microstructure reconstruction was introduced to generate microstructural samples with statistical equivalence by using limited microstructural information [495], and it has been proved to be an effective and economical way to tackle this problem.

Various methods have been developed to reconstruct heterogeneous microstructures over the past few decades. Among them, the stochastic optimization method [495] and Gaussian random field method [363, 135, 134] are the most commonly-used approaches. The stochastic optimization method generates a microstructure sample through an optimization procedure by setting up the optimization objective with microstructural descriptors [495]. The reconstructed trial sample is iteratively adjusted to minimize the statistical difference (measured by the selected descriptors) between it and the original microstructure, to produce a microstructural sample that is statistically equivalent to the original one. The simulated annealing algorithm is often used to drive the optimization procedure. This reconstruction method is flexible and robust, and the accuracy of realizations can be controlled by setting the appropriate termination criterion for the optimization process. However, this method is very computation-intensive, due to its iterative nature of the optimization procedure and the repetitive calculation of statistical descriptors. Gaussian random field method [363, 135, 134] models a microstructure as a Gaussian random field, and it produces a microstructure sample

by truncating a manually generated realization of random filed. Generally, this types of method are much faster than the stochastic optimization method to generate microstructure samples, but it only captures the first- and second-order statistical information, which is inadequate to preserve the morphological complexity of heterogeneous microstructures. Besides, some other reconstruction methods, including process-based method [341], Markov Chain Monte Carlo simulation [480], multiple-point statistics method [336], patch-based method [430] and texture synthesis method [285], are also used frequently.

Effective reconstruction methods should allow accurate and efficient reconstruction of realistic microstructures, such that geometrical measurement and numerical simulation can be performed on the reconstructed samples to investigate the impact of microstructures on material properties. Recently, machine learning [368] and deep learning [3] have been successfully applied to characterize and reconstruct complex heterogeneous media, because of their powerful capacities to analyze complex data and explore hidden connections.

Caers (2001) [66] first used artificial neural networks (ANN) to statistically model connected reservoirs. Latter, Sundararaghavan and Zabaras (2005) [429] considered microstructure reconstruction as a pattern recognition problem, and they adopted a support vector machine (SVM) to reconstruct 3D microstructures by using statistical information extracted from planar images. The SVM-based reconstruction method can generate microstructure samples with multiple features closely matching the original microstructure, but *a prior* database of 3D microstructures is required for feature extraction and SVM model training, which is often unavailable in practice. In 2016, Bostanabad et al. (2016) [52, 53] adopted the decision tree (supervised machine learning) model to learn the stochastic feature of heterogeneous microstructures, and then statistically equivalent microstructure samples are generated based on this characterization. As a simple supervised machine learning model, the decision tree may not fully capture the complex morphology of heterogeneous microstructures. As Markov random field assumption was used for microstructure characterization, this method may have limitations in capturing long-distance morphologies. Cang et al. (2017) [69] developed a method to characterize and reconstruct heterogeneous microstructures via convolution deep belief networks (CDBN). This method can generate samples that statistically preserve the long-distance morphological features as well as the critical fracture strength values. Currently, this method can only be used for 2D microstructure reconstruction, and it cannot guarantee the accuracy, because the key parameters such as threshold are determined empirically. Mosser et al. (2017) [315] proposed a method to reconstruct 3D porous media by using the generative adversarial network (GAN), where the GAN serves as an implicit description of the probability distribution of the 3D image dataset. The reconstructed samples closely match the reference porous media in terms of pore morphology and single-phase

permeability, but this method requires *a prior* dataset of 3D digital images to train the GAN model. More recently, the convolutional neural network (CNN) [276, 462] has been used to characterize and reconstruct heterogeneous microstructures. GAN and CNN are powerful tools to characterize microstructures, but the efficiency to reconstruct new samples still needs to improve, especially for large 3D microstructures.

Based on deep neural network (DNN), this study presents a new method for statistical characterization and reconstruction of heterogeneous microstructures in both 2D and 3D formats. To examine its performance, a series of reconstruction tests and comparisons involving many different types of microstructures are conducted, where the reconstruction quality is assessed by comparing both morphological and physical properties of reconstructed and target microstructures. The main technical advantage includes: (1) using sparse stacked autoencoder (SSAE) and SoftMax classifier, the proposed method is accurate, efficient, and applicable to a wide range of heterogeneous materials with various morphological features; (2) using a multi-level approach to statistically characterize and reconstruct microstructures in a hierarchical manner, the new method can accurately capture long-distance morphological features, making it particularly suitable for the application in porous media with long-range connectivity.

The remainder of this chapter is organized as follows: In Section 4.2, statistical microstructure characterization using deep neural network is presented in detail through a 2D microstructure, including basic assumption, training data collection and process, and 'SSAE+Softmax' model training; Section 4.3 explains the procedure of stochastic reconstruction based on the microstructure characterization from the trained 'SSAE+Softmax' model; Section 4.4 examines the performance of reconstructing 2D microstructures with distinct morphologies, where the proposed method is systematically compared with the recent machine learning-based method proposed by Bostanabad (2016) [52, 53]; In Section 4.5, the method is extended for 3D microstructures, and the multi-level approach is developed for microstructures with long-distance morphological features; Finally, conclusions are drawn in Section 4.6.

4.2 Microstructure characterization

The raw SEM or micro-CT images of heterogeneous microstructures are in grayscale, so they are usually segmented to represent different phases of the microstructures with distinct intensity values [388]. Statistically characterizing the distributions of different phases within the microstructures is critical, which is the first step to reconstruct new microstructure samples with statistical equivalence. In this section, we take the 2D heterogeneous microstructure with

two phases as the example to illustrate the proposed microstructure characterization method using deep neural networks, and this method can be readily extended to 3D microstructures with multiple phases.

4.2.1 Markov random field

An example of 2D digital microstructure with two phases is shown in Figure 4.2a, where its two phases are represented by white and black regions respectively. In computer vision, a digital image is a dot matrix data structure which represents a typically rectangular grid of pixels. Let the matrix $\boldsymbol{X}_{(s_1 \times s_2)}$ denote the set of pixels in a 2D digital microstructure with size $s_1 \times s_2$:

$$\boldsymbol{X} = \{X_{ij} \mid 1 \leq i \leq s_1, 1 \leq j \leq s_2\} \tag{4.1}$$

where X_{ij} denotes the element in the matrix $\boldsymbol{X}$, representing the intensity value of the pixel at row i and column j. For the 2-phase microstructure, X_{ij} is a binary variable described by the phase function as follows:

$$X_{ij} = \begin{cases} 1 & \text{if } (i,j) \text{ is located at white phase} \\ 0 & \text{if } (i,j) \text{ is located at black phase} \end{cases} \tag{4.2}$$

Markov random field assumption

Due to the stochastic nature of heterogeneous microstructure, the matrix $\boldsymbol{X}$ corresponding to a microstructure sample can be described by the joint probability distribution function (PDF) $p(X_{11}, X_{12}, \ldots, X_{ij}, \ldots, X_{s_1,s_2})$ or $p(\boldsymbol{X})$, which is the perfect characterization of the microstructure $\boldsymbol{X}$. If $p(\boldsymbol{X})$ could be learned, it can be directly used to reconstruct statistically equivalent microstructure samples. However, $p(\boldsymbol{X})$ usually has an extremely high dimension equal to the number of pixels in the image, making it impracticable to estimate $p(\boldsymbol{X})$ from limited number of digital images $\boldsymbol{X}$. To simplify the microstructure characterization problem, the digital image of a heterogeneous medium is assumed to be a Markov random filed, which can greatly reduce the computing complexity of $p(\boldsymbol{X})$.

MRF models a texture as a local and stationary random process, and it has been proven to be an effective method for a wide range of textures in computer vision applications [272]. Specifically, MRF has two assumptions [469]:

- **Locality**: each pixel X_{ij} in the digital image is only related to its neighboring pixels $\boldsymbol{N}_{ij}$ within a sufficiently large area:

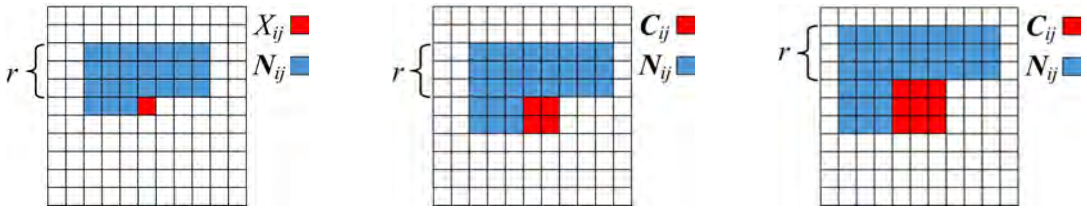
$$p(X_{ij} | \boldsymbol{X}^{(-ij)}) = p(X_{ij} | \boldsymbol{N}_{ij}) \tag{4.3}$$

where $\boldsymbol{X}^{(-ij)}$ denotes the rest of pixels in $\boldsymbol{X}$ excluding pixel X_{ij}.

- **Stationarity**: $p(X_{ij}|\boldsymbol{N}_{ij})$ is the same for all pixels regardless of the pixel location (i, j), which signifies that different regions in the image always share the same statistical characteristics.

Data template and extension of MRF assumption

The geometry and size of spatial region covering the neighboring pixels $\boldsymbol{N}_{ij}$ are important for the application of MRF. In this study, the neighboring pixels $\boldsymbol{N}_{ij}$ within the L-shaped region are adopted to characterize the central pixel X_{ij}, and the whole area covering both the central pixel and its neighborhood is called a 'data template', as shown in Figure 4.1. The reason to choose the L-shaped data template is that the neighboring pixels $\boldsymbol{N}_{ij}$ are prior to the central pixel X_{ij} in the raster scan order, which meets the requirement of microstructure reconstruction in our method (more details will be discussed in Section 4.3). The size of data template is measured by the radius r, and it should be large enough to capture the fundamental morphology features in the image.



(a) 1 central pixel (b) 4 central pixels (c) 9 central pixels

Fig. 4.1 Data templates with central pixel(s) and neighboring pixels (neighborhood radius $r = 3$ pixels)

The classical MRF model only focuses on the relationship between one central pixel X_{ij} and its surrounding neighboring pixels $\boldsymbol{N}_{ij}$ [272]. In this study, we extend the classical MRF assumption into more general forms, where the correlation between multiple central pixels $\boldsymbol{C}_{ij}$ and their surrounding neighboring pixels $\boldsymbol{N}_{ij}$ can also be considered, and the corresponding data templates are shown in Figures 4.1b and 4.1c. The locality and stationarity assumptions are both applied equally to the cases with multiple central pixels, and Eq. (4.3) is rewritten as:

$$p(\boldsymbol{C}_{ij}|\boldsymbol{X}^{(-ij)}) = p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij}) \tag{4.4}$$

where $\boldsymbol{X}^{(-ij)}$ denotes the set of pixels in $\boldsymbol{X}$ excluding the central pixels $\boldsymbol{C}_{ij}$.

The possible configurations of multiple central pixels are much more diverse than the case of a single central pixel. Take the 4-central-pixel case as an example, there are 2^4

possible patterns for these 4 central pixels in a 2-phase microstructure. The purposes of MRF extension are two-fold: (1) to capture the complexity (long-distance morphology) of microstructures by using the proposed multi-level approach (as described in Section 4.5); and (2) to speed up the microstructure reconstruction process by updating multiple pixels at each iteration step (as discussed in Section 4.3). One can choose the most suitable template for microstructure characterization and reconstruction to balance efficiency and accuracy.

4.2.2 Data event collection

The data template is used to collect primary local patterns from the microstructure, so it has significant influence on microstructure characterization, thereby determining the quality of microstructure reconstruction. The selected data template scans over the entire digital image $\boldsymbol{X}$, termed training image here, to collect data events $(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})$ in a raster scan order. Each data event $(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})$ is a paired observation, and its configuration varies with the location (i, j) to represent the variation of local spatial pattern in the training image $\boldsymbol{X}$. The scanning process starts from the top left pixels with enough neighboring pixels to fit the data template. The border pixels are not scanned as the central pixel, because they do not have sufficient neighborhoods to accommodate the data template. Figure 4.2 illustrates the data event collection by using the selected data template to scan over the entire training image.

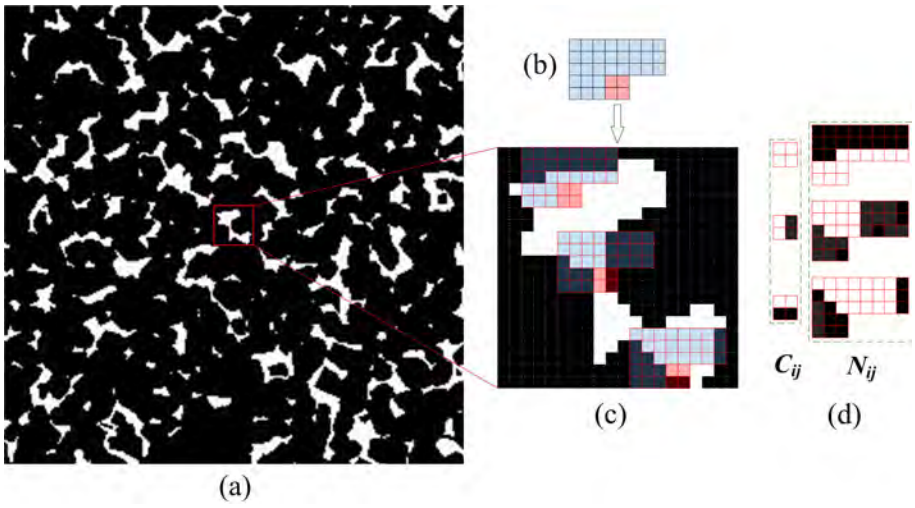


Fig. 4.2 Illustration of data event collection from the training image: (a) The training image of a 2D microstructure; (b) Data template with 4 central pixels ($r = 3$ pixels); (c) Partial enlarge detail of the training image with data template scanning on it; (d) Three examples of data event $(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})$

The conditional probability distribution function (CPDF) $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$ associated with the data template can be estimated from the occurrence frequencies of data events collected from

the training image:

$$p(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij}) = \frac{O(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})}{\sum O(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})} = \frac{O(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})}{O_{\text{all}}} \tag{4.5}$$

$$p(\boldsymbol{N}_{ij}) = \frac{O(\boldsymbol{N}_{ij})}{\sum O(\boldsymbol{N}_{ij})} = \frac{O(\boldsymbol{N}_{ij})}{O_{\text{all}}} \tag{4.6}$$

$$p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij}) = \frac{p(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})}{p(\boldsymbol{N}_{ij})} = \frac{O(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})}{O(\boldsymbol{N}_{ij})} \tag{4.7}$$

where $O(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})$ and $O(\boldsymbol{N}_{ij})$ are the occurrence numbers of data event $(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})$ and $(\boldsymbol{N}_{ij})$ respectively, and O_{all} is the total number of all data events collected from the training image.

4.2.3 Microstructure characterization via deep neural network

The core of microstructure characterization is to determine the CPDF $p(X_{ij}|\boldsymbol{N}_{ij})$ or $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$ of the training image $\boldsymbol{X}$. However, explicit construction of the CPDF is complicated and computational intensive [469], especially for heterogeneous microstructures with complicated geometries, sometimes even impossible. As mentioned above, the phase value of each pixel X_{ij} can only be 1 (white) or 0 (black) in the binary representation of a microstructure. Therefore, the conditional probability $p(X_{ij}|\boldsymbol{N}_{ij})$ associated with 1-central-pixel data template can be considered as a class probability, which provides the probability of the central pixel X_{ij} to be white or black when given its neighboring pixels $\boldsymbol{N}_{ij}$. It is a binary classification problem, where the phase value of X_{ij} is the classification category and the configuration of $\boldsymbol{N}_{ij}$ is the classification feature. In addition, the data events $(X_{ij}, \boldsymbol{N}_{ij})$ collected from the training image are paired observations, which are perfect training dataset to train a machine learning model for classification [43]. Taking the collected data events $(X_{ij}, \boldsymbol{N}_{ij})$ as the training dataset, the machine learning classifier can be built to minimize the probability of misclassification for any input features $\boldsymbol{N}_{ij}$. The class probability stored in the fitted machine learning model will be an accurate estimation of the CPDF $p(X_{ij}|\boldsymbol{N}_{ij})$. As to the CPDF $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$ associated with data template having multiple central pixels, it can also be estimated in the same way. But data conversion is required to assign a label L_{ij} to each configuration of the multiple central pixels $\boldsymbol{C}_{ij}$, and then the label L_{ij} together with the neighbouring pixels $\boldsymbol{N}_{ij}$ will serve as a pair of observation to train the machine learning model. In other words, the binary classification problem discussed above becomes a multi-class classification problem here, but the basic principles remain. More details about data conversion will be explained in Section 4.2.3.

An appropriate machine learning model is critical for the above classification (pattern recognition) task. Considering the high dimension of feature data (feature data $\boldsymbol{N}_{ij}$ usually

contain hundreds of pixels, and the number of voxels will exceed 1000 for 3D microstructure cases) and the large number of classification categories (the number of possible classification categories for the n-central-pixel case is 2^n, and it becomes even larger for multi-phase microstructures), deep neural network appears to be the most attractive option. DNN models [392] with multiple layers are effective for solving classification problems with complex data, and each layer can learn/capture features at a different level.

Data conversion for the central pixels in data template

For the 1-central-pixel data template, the central pixel has two possible configurations, and each configuration can be labeled by the phase value. However, for the 4- or 9-central-pixel data template, there are 2^4 or 2^9 possible configurations for the central pixels. An index label is required for each possible configuration, so that the machine learning model can be trained to map each label to the corresponding feature data.

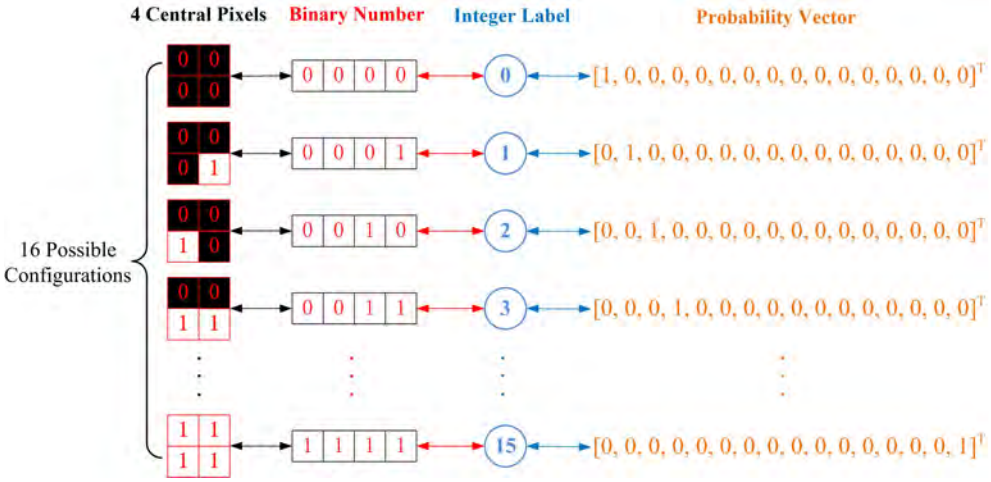


Fig. 4.3 Flow chart of data conversion (taking 4-central-pixel data template as the example)

As shown in Figure 4.3, a 4-central-pixel data template is used to explain the data conversion procedure. Given 4 central pixels $\boldsymbol{C}_{ij}$, there are 16 possible pattern configurations in total, and for each configuration, the phase values are first recorded as a binary number, which is then translated into an integer label, denoted by L_{ij}. Finally, the integer label is translated into a probability vector so that the class probabilities can be directly outputted from the trained DNN model upon new input data. The data conversion procedure remains the same for a n-central-pixel data template with two phases. As to multi-phase microstructures, such as 3-phase microstructures, the data conversion still works by changing the binary number to a ternary number. It should be noted that the data conversion is reversible, which

is important for the stochastic reconstruction of microstructure samples through probability sampling.

Data compression for the neighboring pixels in data template

Although MRF assumption has been adopted to reduce dimensionality for microstructure characterization, the number of neighboring pixels in the data template is usually large, especially for 3D microstructures with long-distance morphological characteristics. This section explains how to extract features from the raw pixel inputs using stacked sparse autoencoder (SSAE) [190, 488, 463].

Autoencoder is an unsupervised machine learning algorithm that learns the representation of high-dimension input data by finding the correlation between the data, and is especially suitable for dimensionality reduction. Consisting of an encoder and a decoder, an autoencoder is a multi-layer feed-forward neural network trained to represent the input through back-propagation. It compresses the input data into lower dimensional code and then reconstructs the output from this latent-space representation. During the training process, the discrepancy between input and reconstruction is minimized by applying back-propagation, and a set of parameters including weights W and biases b are optimized, as shown in Figure 4.4.

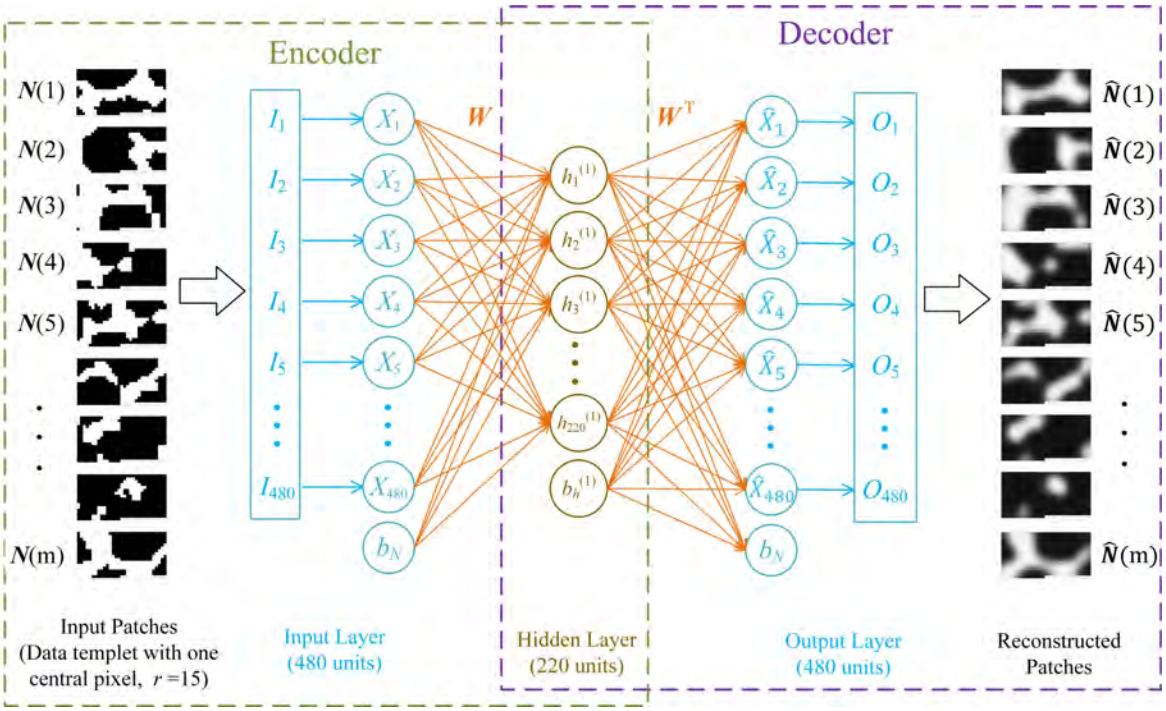


Fig. 4.4 The architecture of basic sparse autoencoder to compress local patterns of heterogeneous microstructures (each input patch contains 480 pixels)

A set of neighboring pixels $\boldsymbol{N}_{ij}$ is called a 'training patch' here. Let $\boldsymbol{Pa} = (\boldsymbol{N}(1), \boldsymbol{N}(2), \ldots, \boldsymbol{N}(m))$ denote the entire training patches collected from the training image, where $\boldsymbol{N}(k) \in \mathcal{R}^{d_\mathrm{N}}$, m and d_N are the number of training patches and the number of pixels in each patch, respectively. Let $\boldsymbol{h}^{(l)}(k) = (h_1^{(l)}(k), h_2^{(l)}(k), \ldots, h_{d_\mathrm{h}}^{(l)}(k))^T$ denote the learned high-level feature at layer l for the k-th patch, where d_h is the number of hidden units in the current layer l. Throughout this section, the superscript and subscript on a notation are used to define the hidden layer and unit in this layer, respectively. For example, the $h_i^{(1)}$ in Figure 4.4 represents the i-th unit in the 1st hidden layer. For simplicity, $\boldsymbol{N}$ and $\boldsymbol{h}^{(l)}$ are used to denote an input patch and its representation at hidden layer l, respectively.

Figure 4.4 shows the architecture of a sparse autoencoder. Generally, the input layer of the autoencoder consists in an encoder, where input patches $\boldsymbol{N}$ are compressed into lower dimensional representations $\boldsymbol{h}$. The output layer is a decoder which is trained to reconstruct approximations $\widehat{\boldsymbol{N}}$ of the input patches $\boldsymbol{N}$ from the hidden representation $\boldsymbol{h}$. In essence, training an autoencoder is to optimize the parameters by setting the minimization of the discrepancy between input $\boldsymbol{N}$ and output $\widehat{\boldsymbol{N}}$ as the objective. This discrepancy is represented by a cost function with three terms as follows:

$$\mathcal{L}(\theta) = \frac{1}{m}\sum_{k=1}^{m} L(\boldsymbol{N}(k), d_{\hat{\theta}}(e_{\check{\theta}}(\boldsymbol{N}(k)))) + \alpha \sum_{j=1}^{n} KL(\rho||\hat{\rho}_j) + \beta\|W\|_2^2 \tag{4.8}$$

The first term in Eq. (4.8) is the mean sum-of-squared error that accounts for the discrepancy between input $\boldsymbol{N}$ and output $\widehat{\boldsymbol{N}}$ over the entire set of data. Encoder $e_{\check{\theta}}(\cdot)$ compresses the input $\boldsymbol{N} \in \mathcal{R}^{d_\mathrm{N}}$ into the hidden representation $\boldsymbol{h} \in \mathcal{R}^{d_\mathrm{h}}$ through the following equation:

$$\boldsymbol{h} = e_{\check{\theta}}(\boldsymbol{N}) = f(\boldsymbol{W}\boldsymbol{N} + \boldsymbol{b}_\mathrm{h}) \tag{4.9}$$

where $\boldsymbol{W}$ is a $d_\mathrm{h} \times d_\mathrm{N}$ weight matrix, and $\boldsymbol{b}_\mathrm{h} \in \mathcal{R}^{d_\mathrm{h}}$ is a bias vector. The encoder is parametrized by $\check{\theta} = (W, b_\mathrm{h})$. Decoder $d_{\hat{\theta}}(\cdot)$ transforms the hidden representation $\boldsymbol{h}$ into the input space $\widehat{\boldsymbol{N}}$ by:

$$\widehat{\boldsymbol{N}} = d_{\hat{\theta}}(\boldsymbol{h}) = f(\boldsymbol{W}^T\boldsymbol{h} + \boldsymbol{b}_\mathrm{N}) \tag{4.10}$$

where $\boldsymbol{W}^\mathrm{T}$ is a $d_\mathrm{N} \times d_\mathrm{h}$ weight matrix, and $\boldsymbol{b}_\mathrm{N} \in \mathcal{R}^{d_\mathrm{N}}$ is a bias vector. Hence, the decoder is parameterized by $\hat{\theta} = (W^\mathrm{T}, b_\mathrm{N})$. The activation function $f(\cdot)$ is a sigmoid logistic function:

$$f(z) = \frac{1}{1+e^{-z}} \tag{4.11}$$

where z is the pre-activation of a neuron. The cost function $L(\boldsymbol{N}, \widehat{\boldsymbol{N}})$ quantifies the discrepancy between input $\boldsymbol{N}$ and the reconstruction $\widehat{\boldsymbol{N}}$, given by:

$$L(\boldsymbol{N}, \widehat{\boldsymbol{N}}) = (\boldsymbol{N} - \widehat{\boldsymbol{N}})^2 \tag{4.12}$$

The second term in Eq. (4.8) describes the sparsity cost, where n is the number of units in the hidden layer, and the index j scans across all hidden units in the network. $\hat{\rho}_j$ is the average activation value of hidden unite j over the entire data set, and ρ is the desired activation value. $KL(\rho||\hat{\rho}_j)$ denotes the Kullback-Leibler (KL) divergence between ρ and $\hat{\rho}_j$, given by:

$$KL(\rho||\hat{\rho}_j) = \rho \log(\frac{\rho}{\hat{\rho}_j}) + (1-\rho) \log \frac{1-\rho}{1-\hat{\rho}_j} \tag{4.13}$$

The third term in Eq. (4.8) defines the regularization cost (also called weight decay term), which tends to decrease the magnitude of the weight to prevent overfitting.

$$\|W\|_2^2 = tr(\boldsymbol{W}^T \boldsymbol{W}) \tag{4.14}$$

The basic sparse autoencoders can also be stacked to form a hierarchical sparse deep model, where high-order features are compressed, organized and extracted. A stacked sparse autoencoder (SSAE) is a deep neural network consisting of multiple sparse autoencoders, where the output of each autoencoder layer is fed into the inputs of the successive autoencoder layer. The architecture of an SSAE that consists of two autoencoder layers is shown in Figure 4.5. The figure illustrates only the encoder part of each basic autoencoder, because the decoder part is not involved in the feature dimensionality reduction when estimating the class probability for an input patch (further explanation is provided in the next subsection). The SSAE yields a function S: $\mathcal{R}^{d_{\mathrm{N}}} \rightarrow \mathcal{R}^{d_{\mathrm{h}^{(2)}}}$, which depresses an input raw pixels of a patch to a new feature representation $\boldsymbol{h}^{(2)} = S(\boldsymbol{N})$. Due to large numbers of initial parameters (weights and bias) in SSAE, it is challenging to effectively optimize the parameters of these autoencoders. Pretraining the network of a deep autoencoder to make the initial parameters close to the good solutions is a practical strategy to overcome this problem. The greedy layer-wise learning algorithm is an effective method to pretrain a deep network by training each layer in turn.

SSAE+Softmax classifier: An implicit form of CPDF

Softmax classifier [43] is a multi-class neural network model using supervised learning algorithm, which is used to classify the final output from the SSAE into multiple classes and calculates a probability for each possible class. The architecture of SSAE with two sparse

autoencoders followed by a Softmax classifier is shown in Figure 4.5. The first autoencoder layer of the SSAE model learns primary features $\boldsymbol{h}^{(1)}$ from the raw pixel input $\boldsymbol{N}$. Then, these primary features are used as the input for the second autoencoder layer of the SSAE model to learn secondary features $\boldsymbol{h}^{(2)}$ from the primary features. Next, these secondary features are served as the input for a Softmax classifier layer, which is trained to map the secondary features $\boldsymbol{h}^{(2)}$ to the corresponding labels $\boldsymbol{L}$. These two sparse autoencoders and the Softmax classifier are combined together to form an 'SSAE+Softmax' model with two hidden layers and a classification layer, as illustrated in Figure 4.5. After the 'SSAE+Softmax' classifier is properly trained, it can provide a probability vector for a new testing patch. This probability vector contains all probabilities of the testing patch belonging to corresponding classes.

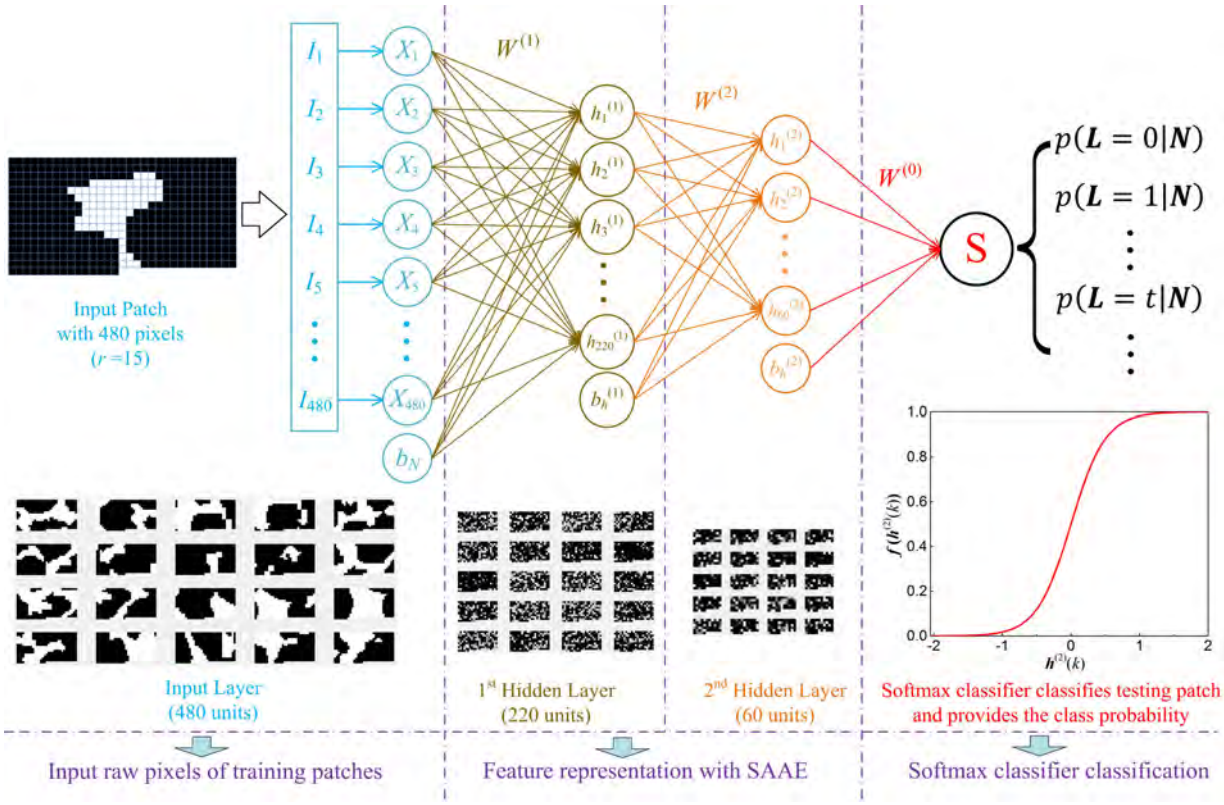


Fig. 4.5 Schematic diagram of the 'SSAE+Softmax' model consisting of two autoencoder layers and a Softmax classifier layer to classify the testing patches (local morphology patterns) of a microstructure (Only the encoder part of each sparse autoencoder is drawn here, because the decoder part is not involved in the feature dimensionality reduction for estimating the class probability)

Mathematically, the class probability of an input patch $\boldsymbol{N}$ belonging to class t can be written as:

$$p(\boldsymbol{L}=t|\boldsymbol{N}) = p(\boldsymbol{L}=t|\boldsymbol{h}^{(2)}) = Softmax(\boldsymbol{W}^{(0)}\boldsymbol{h}^{(2)}+\boldsymbol{b}^{(0)}) = \frac{e^{W_t^{(0)}h^{(2)}+b_t^{(0)}}}{\sum_k^K e^{W_k^{(0)}h^{(2)}+b_k^{(0)}}} \tag{4.15}$$

where $\boldsymbol{W}^{(0)}$ and $\boldsymbol{b}^{(0)}$ are the weight matrix and bias vector of the softmax classifier, respectively, and K is the number of classification categories. Considering the mapping relationship between the label $\boldsymbol{L}$ and the configuration of central pixels $\boldsymbol{C}$, the class probability $p(\boldsymbol{L}|\boldsymbol{N})$ from the 'SSAE+Softmax' classifier is equivalent to the conditional probability $p(\boldsymbol{C}|\boldsymbol{N})$. From this perspective, the 'SSAE+Softmax' classifier model trained by the training data $(\boldsymbol{X}_{ij},\boldsymbol{N}_{ij})$ or $(\boldsymbol{C}_{ij},\boldsymbol{N}_{ij})$ is an implicit form of the conditional probability distribution (CPDF) $p(\boldsymbol{X}_{ij}|\boldsymbol{N}_{ij})$ or $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$. In practice, cross-validation should be conducted to prevent overfitting the 'SSAE+Softmax' classifier model, and thereby to improve its predictive performance on new datasets.

4.3 Microstructure reconstruction

In the previous section, statistical microstructure characterization is obtained by training an 'SSAE+Softmax' classifier model to learn the CPDF $p(X_{ij}|\boldsymbol{N}_{ij})$ or $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$ of the training image $\boldsymbol{X}$. This section describes how to generate statistically equivalent microstructure samples based on the obtained characterization.

4.3.1 Microstructure reconstruction procedure

The joint PDF $p(\boldsymbol{Y})$ of the reconstructed microstructure sample $\boldsymbol{Y}$ can be expressed as:

$$\begin{aligned} p(\boldsymbol{Y}) &= p(Y_{11})p(Y_{12}|Y_{11})p(Y_{13}|Y_{11},Y_{12})\cdots p(Y_{ij}|Y_{11},Y_{12},\cdots Y_{i(j-1)})\cdots \\ &= p(Y_{11})p(Y_{12}|\boldsymbol{Y}_{(<12)})p(Y_{13}|\boldsymbol{Y}_{(<13)})\cdots p(Y_{ij}|\boldsymbol{Y}_{(<ij)})\cdots \end{aligned} \tag{4.16}$$

where $\boldsymbol{Y}_{(<ij)}$ represents the pixel sequence ahead of the pixel Y_{ij}. If all above conditional probabilities $p(Y_{ij}|\boldsymbol{Y}_{(<ij)})$ are known, it is feasible to sequentially generate pixel Y_{ij} by sampling from the $p(Y_{ij}|\boldsymbol{Y}_{(<ij)})$. As shown in Figure 4.1a, the neighboring pixels $\boldsymbol{N}_{ij}$ in the data template are all in front of the central pixel X_{ij} in the raster scan order. With the MRF assumption, the following relation holds:

$$p(Y_{ij}|\boldsymbol{Y}_{(<ij)}) = p(Y_{ij}|\boldsymbol{N}_{ij}^{(Y)}) \tag{4.17}$$

where $\boldsymbol{N}_{ij}^{(Y)}$ is the set of neighboring pixels surrounding the central pixel Y_{ij}. Substituting Eq. (4.17) into Eq. (4.16), the joint PDF $p(\boldsymbol{Y})$ is rewritten as:

$$p(\boldsymbol{Y}) = p(Y_{11})p(Y_{11}|\boldsymbol{N}_{11}^{(Y)})p(Y_{12}|\boldsymbol{N}_{12}^{(Y)})\cdots p(Y_{ij}|\boldsymbol{N}_{ij}^{(Y)})\cdots \tag{4.18}$$

where each conditional probability $p(Y_{ij}|\boldsymbol{N}_{ij}^{(Y)})$ can be obtained by inputting $\boldsymbol{N}_{ij}^{(Y)}$ into the trained 'SSAE+Softmax' classifier. The microstructure sample $\boldsymbol{Y}$ can be reconstructed through sequential probability sampling by giving an initial guess of $p(Y_{11})$. The effect of initialization can be ignored when the pixel sequence is long enough, therefore, the reconstructed microstructure sample $\boldsymbol{Y}$ is statistically equivalent to the training image $\boldsymbol{X}$ in terms of the joint PDF.

Basically, the above procedure is to reconstruct a microstructure sample through generating pixels one by one. By using the data template with multiple central pixels (as shown in Figures 4.1b and 4.1c), multiple pixels can be generated at each step, which can greatly speed up the reconstruction process. In these cases, the joint PDF $p(\boldsymbol{Y})$ can be calculated as:

$$p(\boldsymbol{Y}) = p(\boldsymbol{C}_{11}^{(Y)})p(\boldsymbol{C}_{11}^{(Y)}|\boldsymbol{N}_{11}^{(Y)})p(\boldsymbol{C}_{12}^{(Y)}|\boldsymbol{N}_{12}^{(Y)})\cdots p(\boldsymbol{C}_{ij}^{(Y)}|\boldsymbol{N}_{ij}^{(Y)})\cdots \tag{4.19}$$

where $\boldsymbol{C}_{ij}^{(Y)}$ denotes the multiple central pixels in the image $\boldsymbol{Y}$, and $p(\boldsymbol{C}_{ij}^{(Y)}|\boldsymbol{N}_{ij}^{(Y)})$ can be obtained from the trained 'SSAE+Softmax' classifier by inputting $\boldsymbol{N}_{ij}^{(Y)}$.

The algorithmic workflow of the proposed characterization and reconstruction for 2D microstructures is summarized in Algorithm 1. The 4-central-pixel data template is used as the example to describe the procedure of microstructure characterization and reconstruction, and a schematic illustration is provided in Figure 4.6.

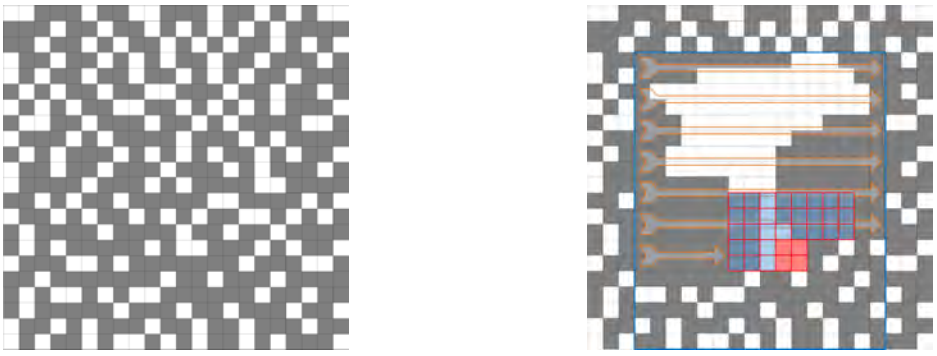
4.3.2 Boundary effect

As shown in Figure 4.6b, the boundary pixels of the initial image $\boldsymbol{Y}_0$ do not have enough neighboring pixels as the inputs for the trained 'SSAE+Softmax' classifier model, so their phase values will not be updated during the reconstruction process. However, these boundary pixels still play as neighboring pixels for the corresponding inner pixels, which will cause a negative effect on microstructure reconstruction, as shown in Figure 4.7b. To tackle the boundary effect problem, a simple method is to remove the boundary and affected area from the reconstructed image and only pick the central part as the final result, as shown in Figure 4.7b. In this study, a more rational approach using periodical/reflected boundary conditions is proposed to overcome the boundary effect problem. Specifically, the boundary of the initial reconstruction is first removed to obtain a defect-free reconstruction (as shown in

Algorithm 1: Stochastic characterization and reconstruction of 2D heterogeneous microstructures

Microstructure Characterization:
Data: The training image $\boldsymbol{X}$ with size $s_1 \times s_2$;
Data: The selected 4-central-pixel data template with neighboring radius equal to r.
while $r+1 \leq i \leq s_1-1$ and $r+1 \leq j \leq s_2-r-1$ **do**
  Collect the data event $(\boldsymbol{C}_{ij}, \boldsymbol{N}_{ij})$ from the training image $\boldsymbol{X}$ by using the selected data template;
  Convert the multiple central pixels $\boldsymbol{C}_{ij}$ to corresponding label data L_{ij};
  Move the data template to the next step in the raster scan order.
end
Train the 'SSAE+Softmax' classifier model $\boldsymbol{M}$ by using the training data $(\boldsymbol{L}, \boldsymbol{N})$;
Result: The trained 'SSAE+Softmax' classifier model $\boldsymbol{M}$.

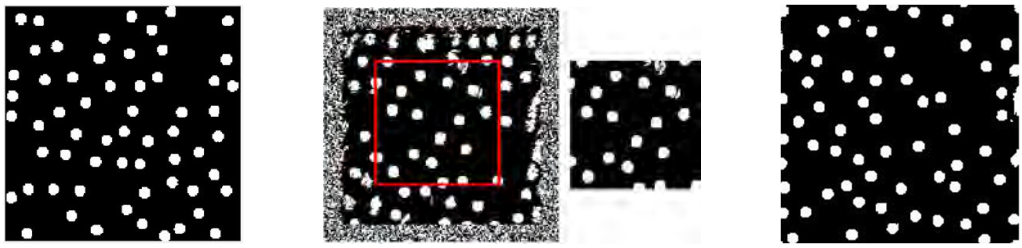
Microstructure Reconstruction:
Data: The trained 'SSAE+Softmax' classifier model;
Data: The selected 4-central-pixel data template with neighboring radius equal to r;
Data: A 2D lattice grid with size $s_3 \times s_4$.
Initialization: Assign binary white noise to the 2D lattice grid as the initial image $\boldsymbol{Y}_0$;
while $r+1 \leq i \leq s_3-1$ *and* $r+1 \leq j \leq s_4-r-1$ **do**
  Extract the data event $(\boldsymbol{C}_{ij}^{(Y_0)}, \boldsymbol{N}_{ij}^{(Y_0)})$ from the initial image by using the selected data template;
  Input the $\boldsymbol{N}_{ij}^{(Y_0)}$ to the trained 'SSAE+Softmax' classifier model $\boldsymbol{M}$ to obtained the class probability vector $p(\boldsymbol{L}|\boldsymbol{N}_{ij}^{(Y_0)})$;
  Generate a label data L_{ij} by probability sampling from the class probability vector $p(\boldsymbol{L}|\boldsymbol{N}_{ij}^{(Y_0)})$;
  Convert the label data L_{ij} to the corresponding configuration of multiple pixels $\boldsymbol{C}_{ij}^{(Y)}$;
  Update corresponding pixels $\boldsymbol{C}_{ij}^{(Y_0)}$ by using $\boldsymbol{C}_{ij}^{(Y)}$;
  Move the data template to the next step in the raster scan order.
end
Result: A statistically equivalent reconstructed microstructure sample $\boldsymbol{Y}$.



(a) (b)

Fig. 4.6 Schematic illustration of microstructure reconstruction procedure: (a) A binary white noise image as the initial state; (b) Pixels are updated by using the 4-central-pixel data template to scan the initial image in the raster scan order (It should be noted that the boundary pixels outside the blue square will not be updated, because they do not have enough neighboring pixels)

Figure 4.7b); then the defect-free sample is used as a tile to build an enlarged copy with periodic/reflected boundaries; and finally the intermediate sample with periodic/reflected boundaries is used as the initialization to reconstruct the final sample as shown in 4.7c. This periodical/reflected boundary approach can effectively eliminate the boundary effect problem in the reconstruction.



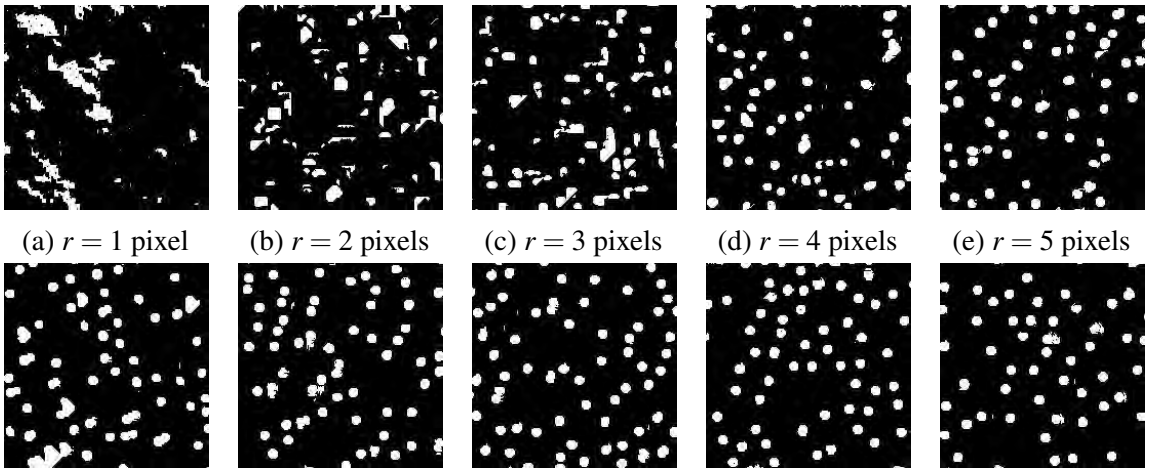
(a) (b) (c)

Fig. 4.7 (a) The training image; (b) The reconstructed image without setting periodical boundary; and (c) The reconstructed image with setting periodical boundary

4.3.3 Size of data template

The radius r of the neighboring pixel zone in the data template is an important parameter affecting the reconstruction quality. A digital image of heterogeneous microstructure can be considered as a spatial signal, so the size of the data template should be large enough to cover the local morphological features. Using the image in Figure 4.7a as the training image,

new samples are reconstructed by using 1-central-pixel data templates with different radius r, and the reconstructed results are shown in Figure 4.8. The microstructure reconstruction quality is continuously improved before r increasing to a threshold value, and no significant improvement is observed after that. This test confirms that the selection of r has a significant influence on microstructure reconstruction quality, and it needs to be greater than a minimum value to achieve an accurate reconstruction.



(a) $r = 1$ pixel (b) $r = 2$ pixels (c) $r = 3$ pixels (d) $r = 4$ pixels (e) $r = 5$ pixels

(f) $r = 6$ pixels (g) $r = 7$ pixels (h) $r = 8$ pixels (i) $r = 10$ pixels (j) $r = 15$ pixles

Fig. 4.8 Microstructure reconstruction by using data templates with different radii r

In our experience, the minimum value of r can be estimated from the two-point correlation function (TPCF) [443] of the training image. The correlation length a provides a reference value for r, and it can be directly measured from TPCF, as shown in Figure 4.9. The correlation length a can also be computed by fitting the TPCF with an exponential function as defined below [100]:

$$S_2(d) = (\phi - \phi^2)\exp(-\frac{3d}{a}) + \phi^2 \tag{4.20}$$

where $S_2(d)$ denotes TPCF, ϕ is volume fraction of calculated phase, and d is the distance between two points. The correlation length a of the training image in Figure 4.7a is 9.83 pixels, and the reconstructed result by setting $r = 10$ is visually consistent with the training image, as shown in Figure 4.8i. As to the 4-central-pixel and 9-central-pixel data template, the minimum sizes of r can also be determined from the correlation length a, which are equal to $a - 0.5$ and $a - 1$, respectively.

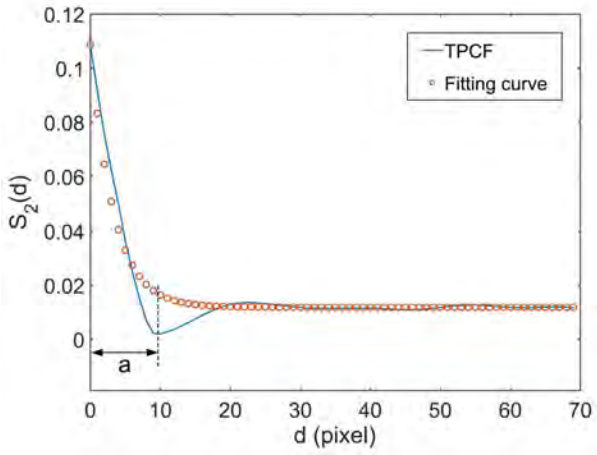


Fig. 4.9 Two-point correlation function and its fitting curve

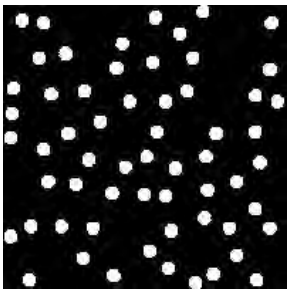
4.4 Results and discussion

In this section, two groups of microstructures with distinct morphologies are used to examine the proposed statistical characterization and reconstruction method. The new method is also compared with a recent machine-learning based approach [52, 53] that uses the decision tree model to learn the stochastic features of heterogeneous media. Different statistical descriptors [443], including two-point correlation function (TPCF), two-point cluster correlation function (TPCCF) and lineal path function (LPF), are used to assess statistical equivalence between reconstructed and target microstructures. In addition, the time costs for microstructure characterization and reconstruction by using different data templates are also compared. Microstructure characterization and reconstruction methods used in this study are all implemented with MATLAB, and all tests are performed on a PC with an Intel Xeon 3.5GHZ CPU and 32GB memory.

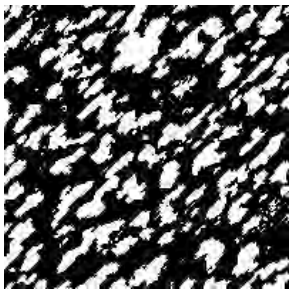
4.4.1 Microstructures with short-distance correlations

Original microstructures

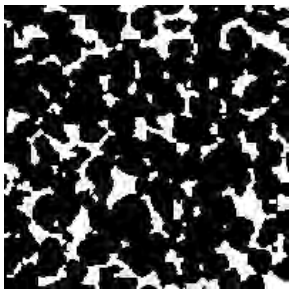
Representative images of three different microstructures are chosen as the training images, as shown in Figure 4.10. Each of these microstructures has unique morphology with relative small size, and the correlation lengths of them are 10, 10, 12 pixels respectively, which can be measured from the TPCFs in Figure 4.12. The first one is an isotropic microstructure with circular inclusions embedded in the matrix; the second one is an anisotropic microstructure with irregular-shaped inclusions; and the last microstructure is a porous medium with strong randomness in pore shapes and sizes.



(a) Microstructure I



(b) Microstructure II



(c) Microstructure III

Fig. 4.10 The 2D training images of three microstructures with short-distance correlations (image size: 200 $\times$ 200 pixels)

For each microstructure, we use all three data templates (as shown in Figure 4.1) to collect data events from the training image, and then to train 'SSAE+Softmax' classifier models in order to learn the CPDFs of the training image. Finally, each trained 'SSAE+Softmax' classifier model is used to generate a set of microstructures with 50 samples, and these samples are compared with the target microstructures by measuring the statistical descriptors. The sizes of data template is determined from the correlation length of the original microstructure, as described in Section 4.3.3. The parameters of 'SSAE+Softmax' classifier model training are summarized in Table 4.1. For comparison, Bostanabad's method [52, 53] is also used to reconstruct these three microstructures by using data templates with the same sizes.

Table 4.1 The parameters for 'SSAE+Softmax' classifier model training

Microstructure	Data template	Template size r	Unite number of each layer		
			Input layer	Hidden layer 1	Hidden layer 2
I	1-central-pixel	10 pixels	220	80	40
I	4-central-pixel	10 pixels	240	80	40
I	9-central-pixel	10 pixels	260	120	40
II	1-central-pixel	10 pixels	220	80	40
II	4-central-pixel	10 pixels	240	120	40
II	9-central-pixel	10 pixels	260	140	60
III	1-central-pixel	12 pixels	312	120	50
III	4-central-pixel	12 pixels	336	140	60
III	9-central-pixel	11 pixels	308	140	60

Reconstruction results

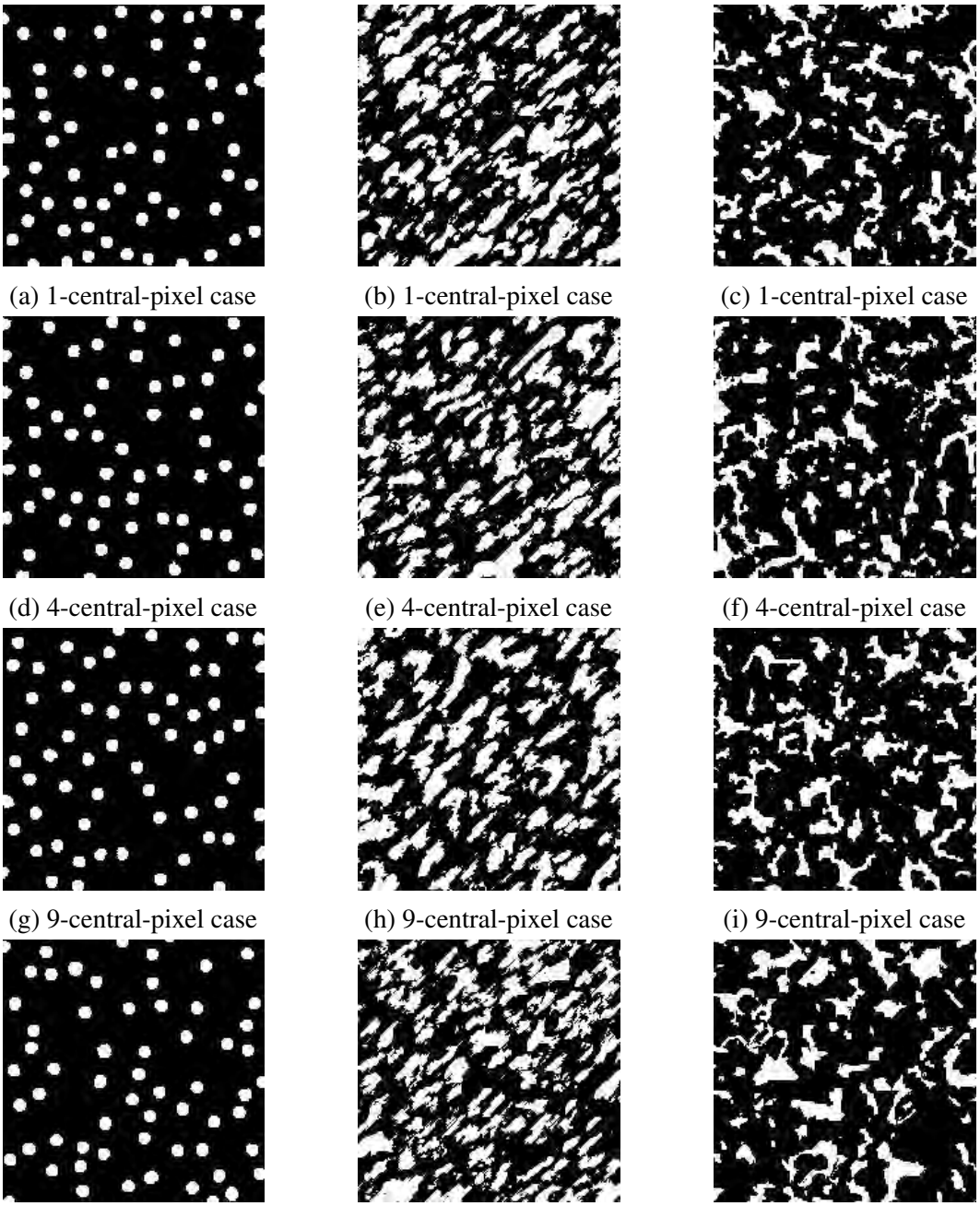
As shown in Figure 4.11, the three columns of images are the representative reconstruction results corresponding to the three target microstructures in Figure 4.10, by using the proposed method and Bostanabad's method. The first three rows (from top to bottom) of images are reconstructed from the proposed method by using the three data templates in Figure

4.1, respectively. The last row provides reconstruction results of Bostanabad's method. Visual inspection of these reconstructed samples implies that the isotropy, anisotropy, and randomness of target microstructures are well preserved, and no obvious discrepancies are observed between the samples reconstructed from the proposed method and Bostanabad's method.

In order to assess whether these reconstructed samples are statistically equivalent to the original microstructures, we compared the TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and the LPF ($L(d)$) between them. All three morphological descriptors are computed for the white phase in the images, and the results are shown in Figure 4.12, 4.13 and 4.14. Generally, the curves of $S_2(d)$, $C_2(d)$ and $L(d)$ measured from the original microstructures are located inside the red intervals which are the ranges of the statistical descriptors extracted from each set of reconstructed microstructures with 50 samples. The relative errors (quantified by L_2 norm error) between the averages of morphological descriptors from reconstructed samples and that of the original microstructure are listed in Table 4.2. Each averaged curve calculated from the 50 reconstructed samples are very closed to the curve measured from the original microstructures, with a relative error within 5.00%, and the greatest variation of the statistical descriptor extracted from a particular reconstructed sample is usually less than 10.00%.

More specifically, the reconstructed samples of the first microstructure not only preserve the shape of circular inclusions, but also maintain the spatial distribution of the circular inclusions as evidenced by the good agreements of statistical descriptors. As to the second microstructure, the reconstructed samples preserve well the anisotropic property and the irregularity of inclusion shapes, which indicates the raster scan order of pixel generation during the reconstruction process does not affect the reconstruction quality. The third microstructure shows the greatest stochasticity in terms of pore shape and pore size, and the trained 'SSAE+Softmax' classifier models is again able to capture stochasticity and complexity of pore geometry and reconstruct statistically equivalent samples.

The reconstructed samples from the proposed method by using different data templates do not differ much visually, and the statistical equivalence shows a similar accuracy level (as shown in Figure 4.11 and Table 4.2). However, the efficiency of microstructure characterization and reconstruction by using different data templates varies greatly, as illustrated in Table 4.2. The time cost of 'SSAE+Softmax' classifier model training (microstructure characterization) increases significantly from using 1-central-pixel template to using 9-central-pixel template, while the corresponding time consumption of reconstruction declines notably. Besides the selection of data template, the complexity of the target microstructure also influence the computational cost, because large neighborhood and complicated classifier model are required for such microstructures. Compared with the 1-central-pixel case, the reconstruction



(a) 1-central-pixel case (b) 1-central-pixel case (c) 1-central-pixel case

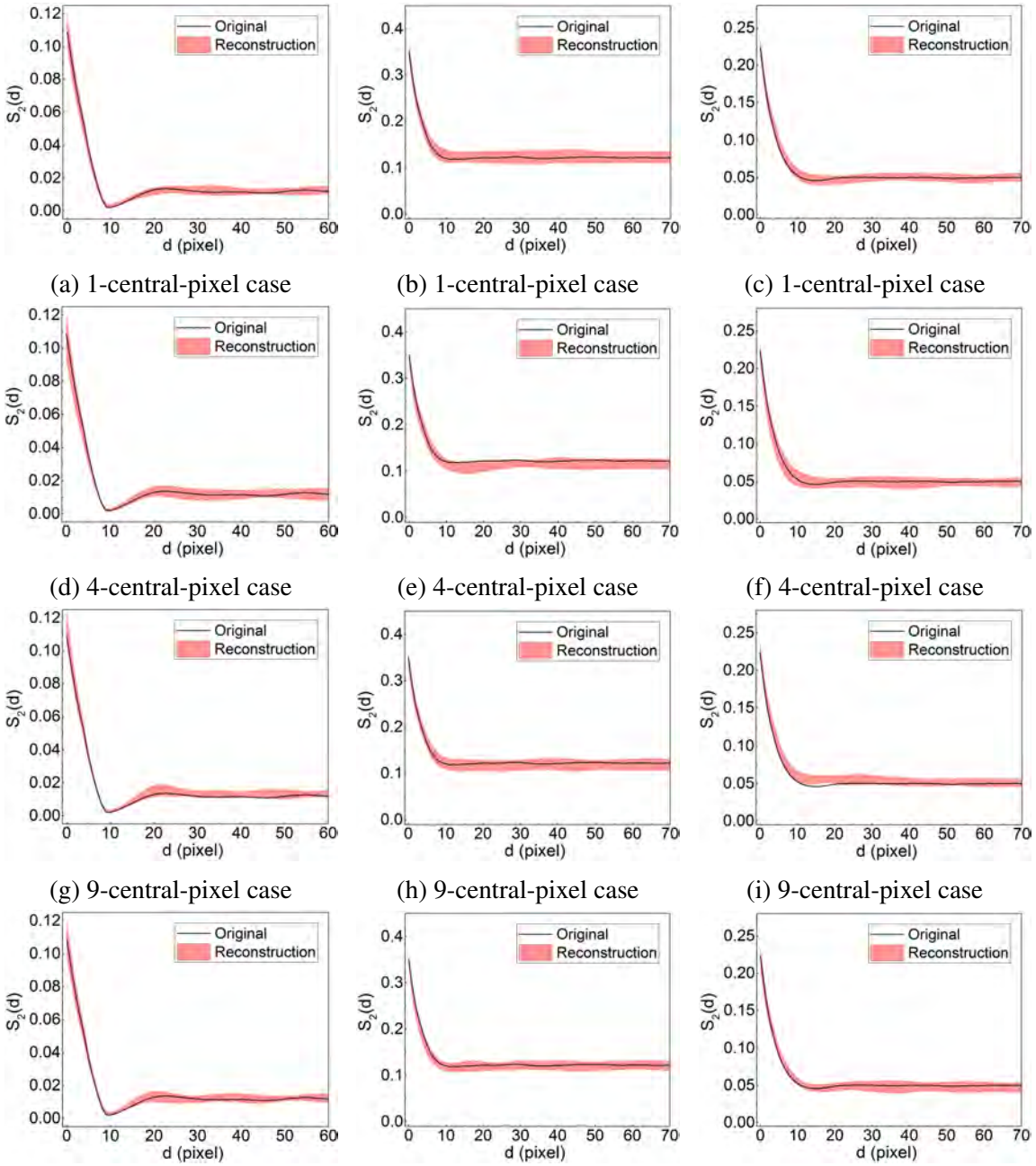
(d) 4-central-pixel case (e) 4-central-pixel case (f) 4-central-pixel case

(g) 9-central-pixel case (h) 9-central-pixel case (i) 9-central-pixel case

(j) Bostanabad's method (k) Bostanabad's method (l) Bostanabad's method

Fig. 4.11 Representative reconstruction results (image size: 200×200 pixels) by using the proposed method and Bostanabad's method (the three columns of microstructures from left to right are reconstruction results corresponding to microstructure I, II and III respectively)

using 4- or 9-central-pixel template requires more computer memory, but the reconstruction speed can be enhanced remarkably. In addition, the trained 'SSAE+Softmax' classifier model can be stored and used repeatedly to reconstruct microstructure samples with any size.



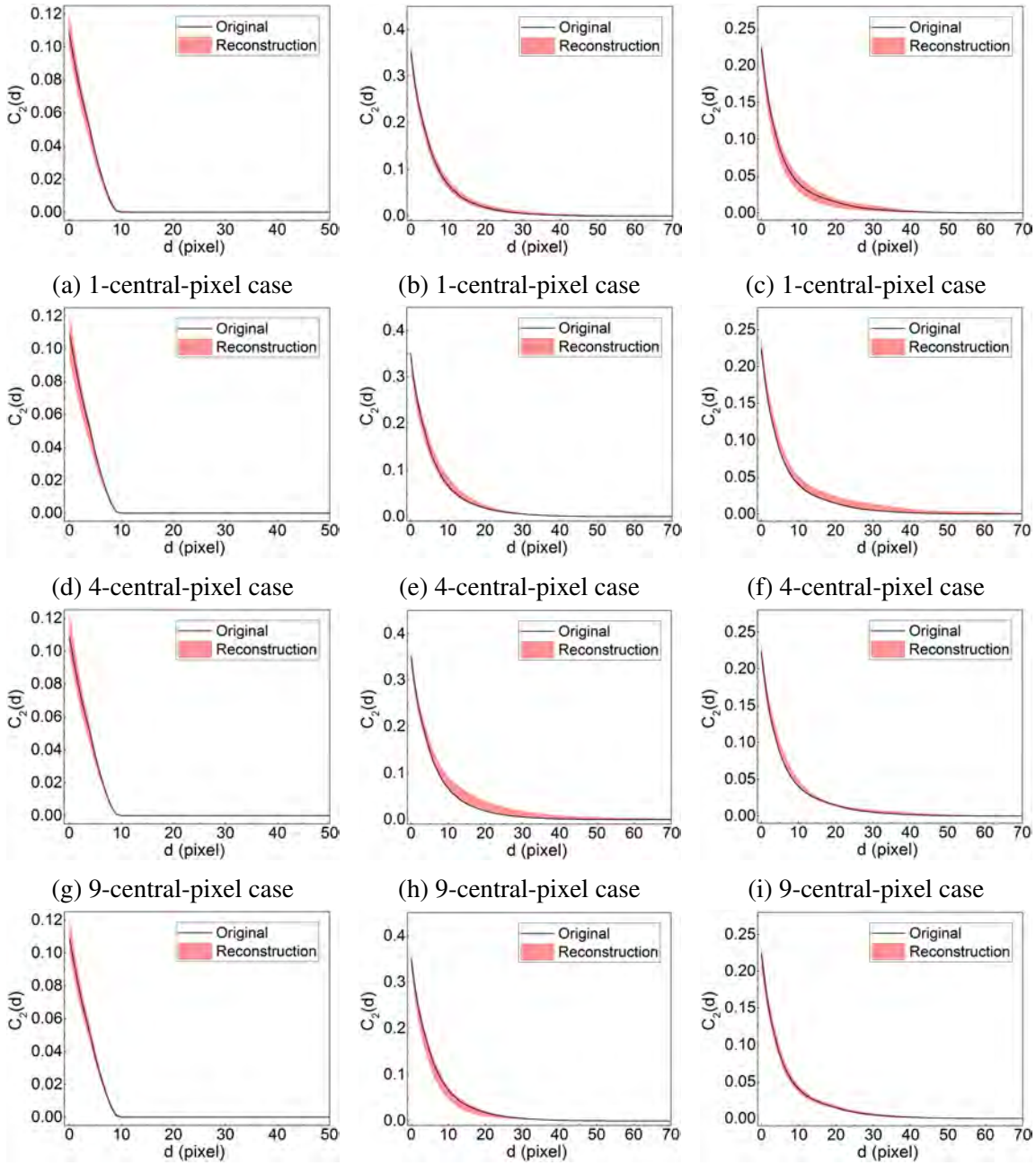
(a) 1-central-pixel case (b) 1-central-pixel case (c) 1-central-pixel case

(d) 4-central-pixel case (e) 4-central-pixel case (f) 4-central-pixel case

(g) 9-central-pixel case (h) 9-central-pixel case (i) 9-central-pixel case

(j) Bostanabad's method (k) Bostanabad's method (l) Bostanabad's method

Fig. 4.12 Comparison of TPCF $S_2(d)$ between the original and reconstructed microstructures from the proposed method and Bostanabad's method (the three columns of figures from left to right are TPCFs corresponding to the microstructure I, II and III respectively)



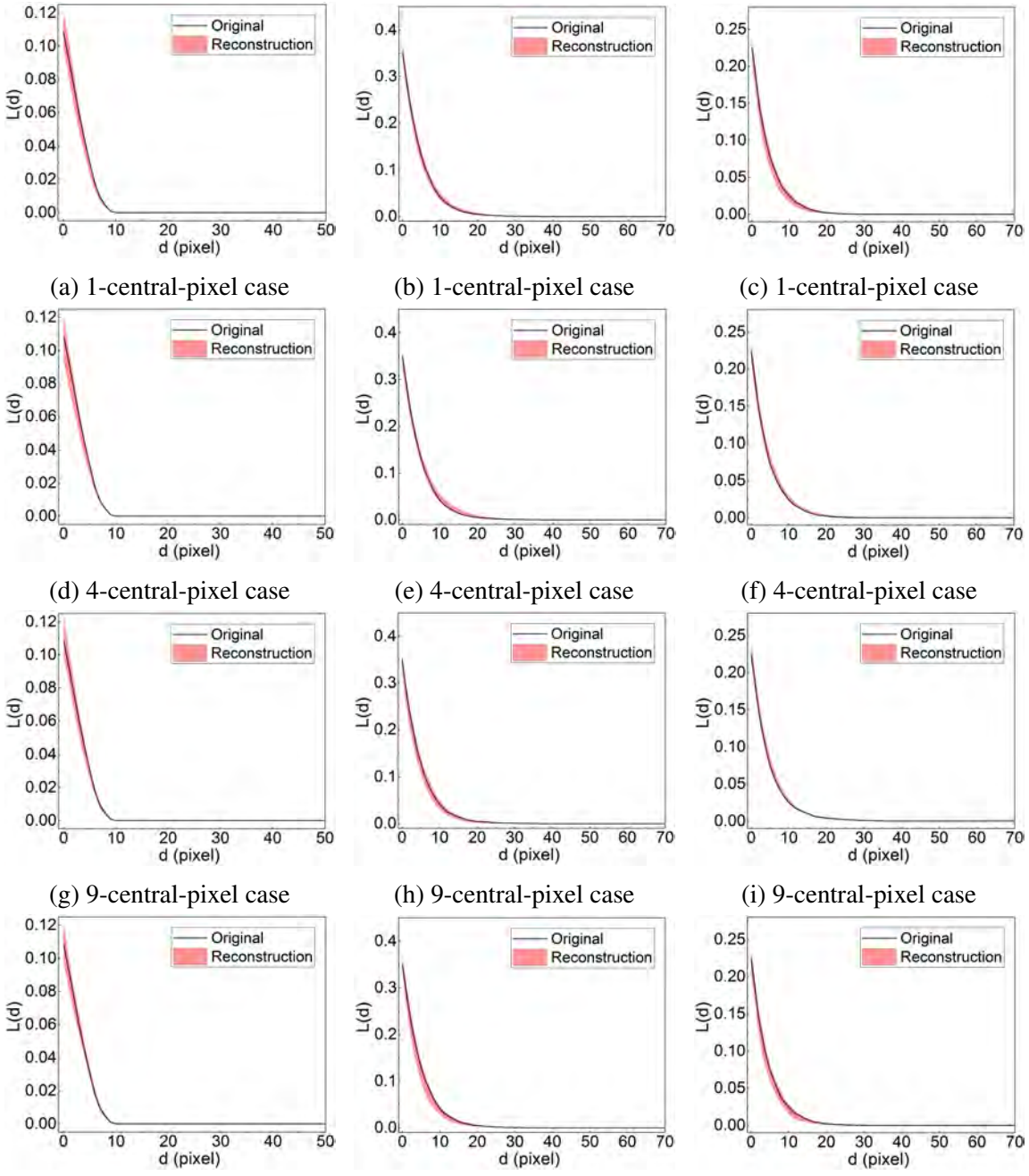
(a) 1-central-pixel case (b) 1-central-pixel case (c) 1-central-pixel case

(d) 4-central-pixel case (e) 4-central-pixel case (f) 4-central-pixel case

(g) 9-central-pixel case (h) 9-central-pixel case (i) 9-central-pixel case

(j) Bostanabad's method (k) Bostanabad's method (l) Bostanabad's method

Fig. 4.13 Comparison of TPCCF $C_2(d)$ between the original and reconstructed microstructures from the proposed method and Bostanabad's method (the three columns of figures from left to right are TPCCFs corresponding to the microstructure I, II and III respectively)



(a) 1-central-pixel case (b) 1-central-pixel case (c) 1-central-pixel case

(d) 4-central-pixel case (e) 4-central-pixel case (f) 4-central-pixel case

(g) 9-central-pixel case (h) 9-central-pixel case (i) 9-central-pixel case

(j) Bostanabad's method (k) Bostanabad's method (l) Bostanabad's method

Fig. 4.14 Comparison of LPF $L(d)$ between the original and reconstructed microstructures from the proposed method and Bostanabad's method (the three columns of figures from left to right are LPFs corresponding to the microstructure I, II and III respectively)

Table 4.2 Average reconstruction errors in volume fraction (ϕ), TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and LPF ($L(d)$), as well as the average computational costs of microstructure characterization and reconstruction

Microstructure	Data template/ Method	$\lvert\Delta\phi\rvert$	$\lvert\Delta S_2(d)\rvert$	$\lvert\Delta C_2(d)\rvert$	$\lvert\Delta L(d)\rvert$	Average CPU time (s)	
						Characterization	Reconstruction
I	1-central-pixel	2.55%	2.83%	1.34%	1.17%	161.5	198.9
I	4-central-pixel	3.15%	2.53%	1.25%	0.90%	336.1	52.6
I	9-central-pixel	2.76%	3.21%	1.51%	1.52%	529.2	26.9
I	Bostanabad's method	2.69%	2.74%	1.36%	1.28%	15.2	41.4
II	1-central-pixel	3.90%	2.38%	1.85%	1.21%	139.7	166.4
II	4-central-pixel	3.58%	3.10%	2.92%	1.73%	290.6	53.7
II	9-central-pixel	3.67%	3.45%	2.31%	1.82%	1293.5	24.7
II	Bostanabad's method	3.49%	2.46%	2.11%	1.32%	19.1	62.8
III	1-central-pixel	3.43%	2.90%	2.84%	1.41%	116.2	169.1
III	4-central-pixel	3.07%	2.38%	1.67%	1.56%	270.2	55.4
III	9-central-pixel	2.58%	3.41%	1.07%	1.77%	848.8	29.2
III	Bostanabad's method	2.97%	3.06%	1.19%	1.84%	18.7	54.1

For the above three microstructures with short-distance morphologies, the proposed method and Bostanabad's method have similar reconstruction accuracy, as recorded in Table 4.2. Bostanabad's method is more efficient in microstructure characterization, because the decision tree model is much simpler than the 'SSAE+Softmax' classifier model. The proposed method can reconstruct samples much faster than Bandstand's method when the 9-central-pixel template is adapted. Therefore, for reconstruction of multiple samples, the proposed method will achieve a better overall efficiency than Bostanabad's method, and the more samples reconstructed the more time saved. However, the key advantage of the new method is to accurately generate microstructures with long-distance morphological features, which will be discussed in the following section.

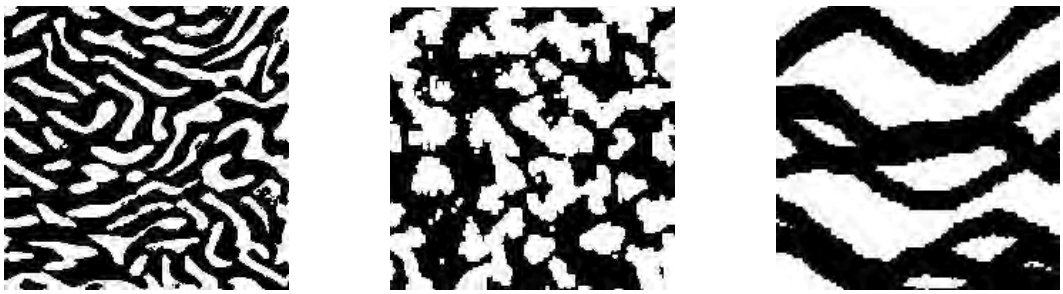
4.4.2 Microstructures with long-distance correlations

This section considers microstructures with long-distance correlations. Again, both the proposed reconstruction method and Bostanabad's method are tested on three microstructures with distinct morphological features.

Original microstructures

As shown in Figure 4.15, three distinct microstructures with long-distance features are tested. The largest inclusions in microstructure IV and microstructure V exceed the half size of training images, and the wavy stripes in microstructure VI even penetrate through the whole training image. For microstructure IV and microstructure V, both the proposed method with three different data templates and Bostanabad's method are used for statistical

characterization reconstruction. The parameters of 'SSAE+Softmax' classifier model training are listed in Table 4.3. The data template sizes are determined by the correlation length, and for fair comparison the data templates of same sizes are used by Bostanabad's method.



(a) Microstructure IV (b) Microstructure V (c) Microstructure VI

Fig. 4.15 The 2D training images of three microstructures with long-distance correlations (image size: 200 × 200 pixels)

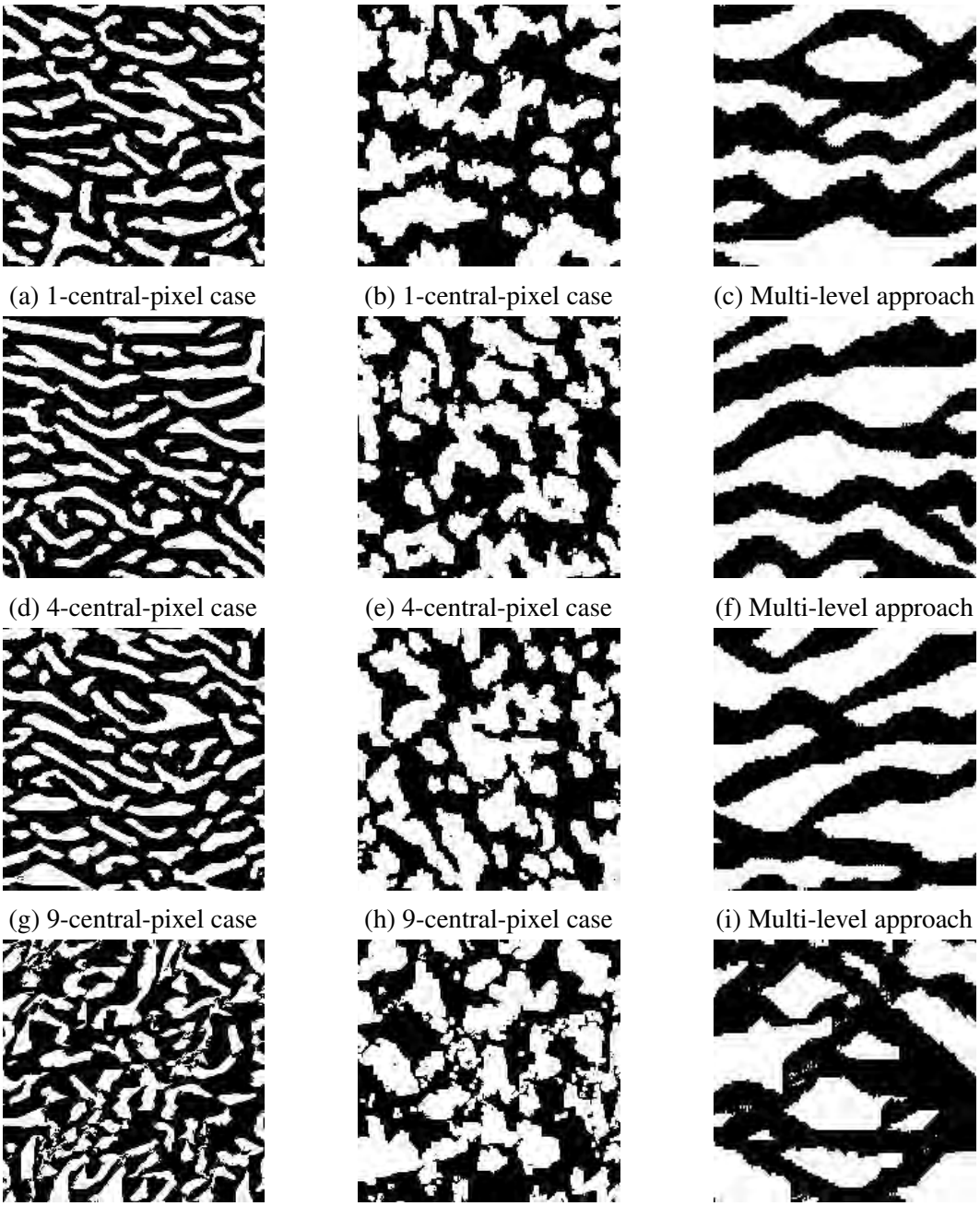
Table 4.3 The parameters for 'SSAE+Softmax' classifier model training

Microstructure	Data template	Template size r	Unite number of each layer		
			Input layer	Hidden layer 1	Hidden layer 2
IV	1-central-pixel	10 pixels	220	80	40
IV	4-central-pixel	10 pixels	240	80	40
IV	9-central-pixel	10 pixels	260	120	60
V	1-central-pixel	20 pixels	840	200	80
V	4-central-pixel	19 pixels	798	200	80
V	9-central-pixel	19 pixels	836	240	80

As to microstructure VI, the multi-level version of the proposed method is adopted to capture the super large morphological features. The multi-level approach in this study is specifically developed to generate 3D porous media samples that preserve long-distance connectivity of pore networks, and its reconstruction procedures are described in details in Section 4.5.2. Microstructure samples are reconstructed in three levels, and 1-, 4- and 4-central-pixel templates are used in the first, second and third level of reconstruction respectively, where the template sizes are 4, 8, and 14 pixels respectively.

Reconstruction results

Figure 4.16 shows the representative reconstruction results corresponding to the three target microstructures in Figure 4.15. Visually, both the local morphologies and long-distance connectivities of the original microstructures are well preserved by the reconstructed samples from the proposed method, while the performance of Bostanabad's method is poorer, especially for microstructure VI.



(a) 1-central-pixel case (b) 1-central-pixel case (c) Multi-level approach

(d) 4-central-pixel case (e) 4-central-pixel case (f) Multi-level approach

(g) 9-central-pixel case (h) 9-central-pixel case (i) Multi-level approach

(j) Bostanabad's method (k) Bostanabad's method (l) Bostanabad's method

Fig. 4.16 Representative reconstruction results (image size: 200×200 pixels) by using the proposed method and Bostanabad's method (the three columns of microstructures from left to right are reconstruction results corresponding to the microstructure IV, V and VI respectively)

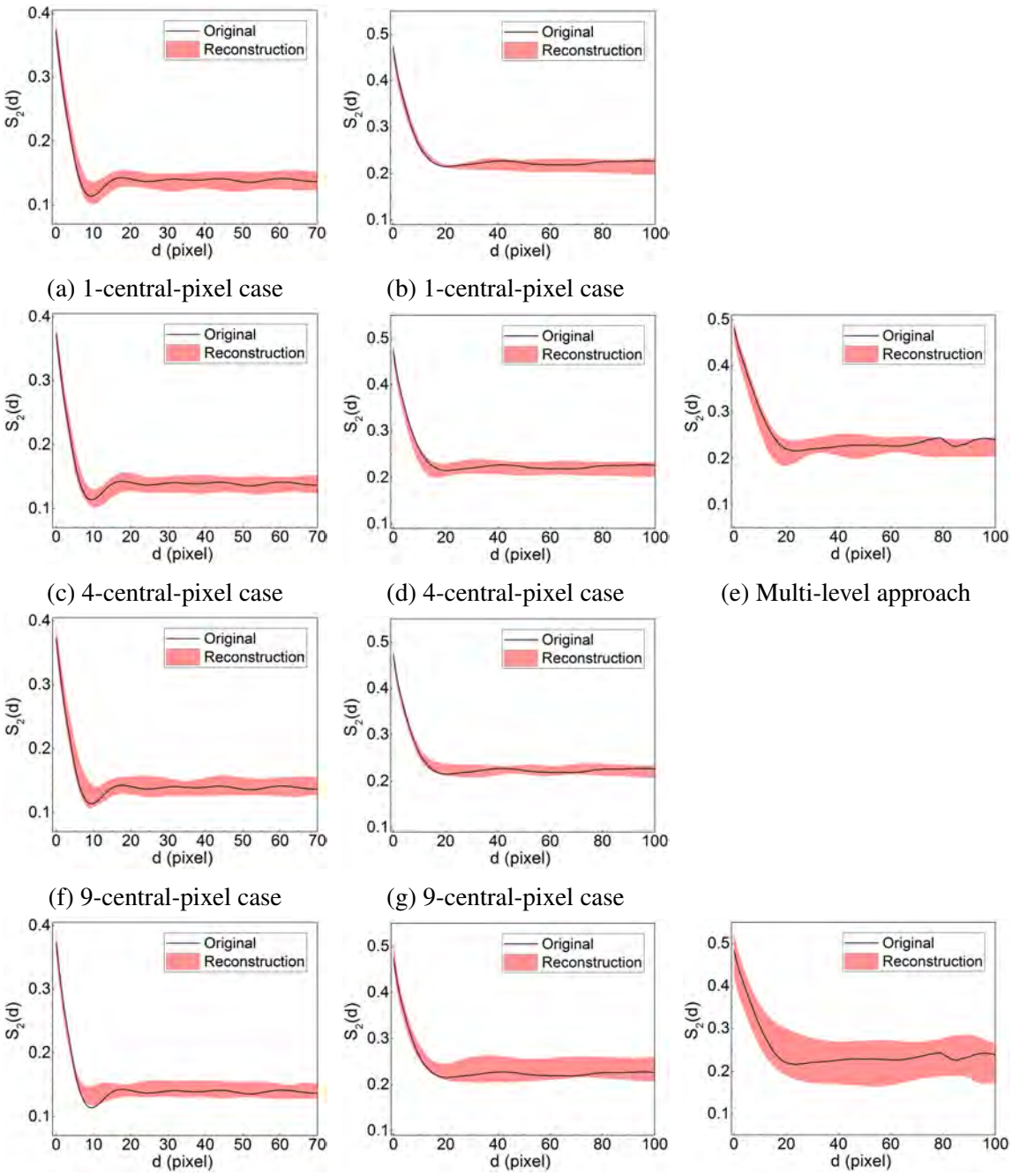
The proposed method well captures the inclusion pattern in microstructure IV as well as its long-distance correlation. The reconstructed samples are also statistically equivalent to the original microstructure measured by TPCF, TPCCF and LPF, as illustrated in Figure

4.17, Figure 4.18 and Figure 4.19, respectively. As to Bostanabad's method, it approximately captures the long-distance morphology of microstructure IV by adopting a sufficiently large data template, but it loses local morphology information of inclusions. As shown in Figure 4.16, the inclusions in the reconstructed sample are disordered and very different from the layered pattern in the original sample. This is because the simple decision tree model used by Bostanabad's method cannot capture both the local details and the long-distance morphological features at the same time. The 'SSAE+Softmax' classifier model adopted by the proposed method is capable of preserving more comprehensively morphological features of the microstructure.

The irregularly shaped inclusions in microstructure V are of different sizes, while the morphology exhibits long-distance features. Both the stochastic nature and long-distance morphology of microstructure V are well preserved in the samples reconstructed from the proposed method, as shown in Figure 4.16, Figure 4.17, Figure 4.18 and Figure 4.19. By contrast, the reconstructed samples from Bostanabad's method cannot accurately preserve the long-distance morphological features and as shown in Figure 4.18 and Figure 4.19, the TPCCF and LPF of reconstructed samples are below that of the original microstructure. Compared to the 'SSAE+Softmax' classifier model, the decision tree model adopted by Bostanabad's method is less effective to learn and express the complex morphology of microstructure V.

Microstructure VI contains long wavy stripes penetrating through whole medium, so the multi-level version of the proposed method is used to characterize and reconstruct the microstructure in a hierarchical way. The low-frequency (long-distance) morphology is maintained at the low level of reconstruction, and high-frequency details are added at the following higher levels of reconstruction. Three representative reconstructed samples are provided in Figure 4.16, from which one can observe that the long-distance morphological features of microstructure VI are well preserved. The good agreements of TPCF, TPCCF and LPF between the original and reconstructed microstructures also confirm the statistical equivalence between them, as shown in Figure 4.17, Figure 4.18 and Figure 4.19, respectively. As to Bostanabad's method, it fails to capture the long-distance morphology that is far beyond the size of selected data template, although date templates of different sizes were tried in this study. The disagreements of TPCCF and LPF between the original and reconstructed microstructures also demonstrate the limitation of Bostanabad's method in reconstructing microstructures with long-distance correlations.

The average reconstruction errors measured by statistical descriptors are recorded in Table 4.4, as well as the average time costs of microstructure characterization and reconstruction by using different methods. The training of 'SSAE+Softmax' classifier model is generally more



(a) 1-central-pixel case (b) 1-central-pixel case

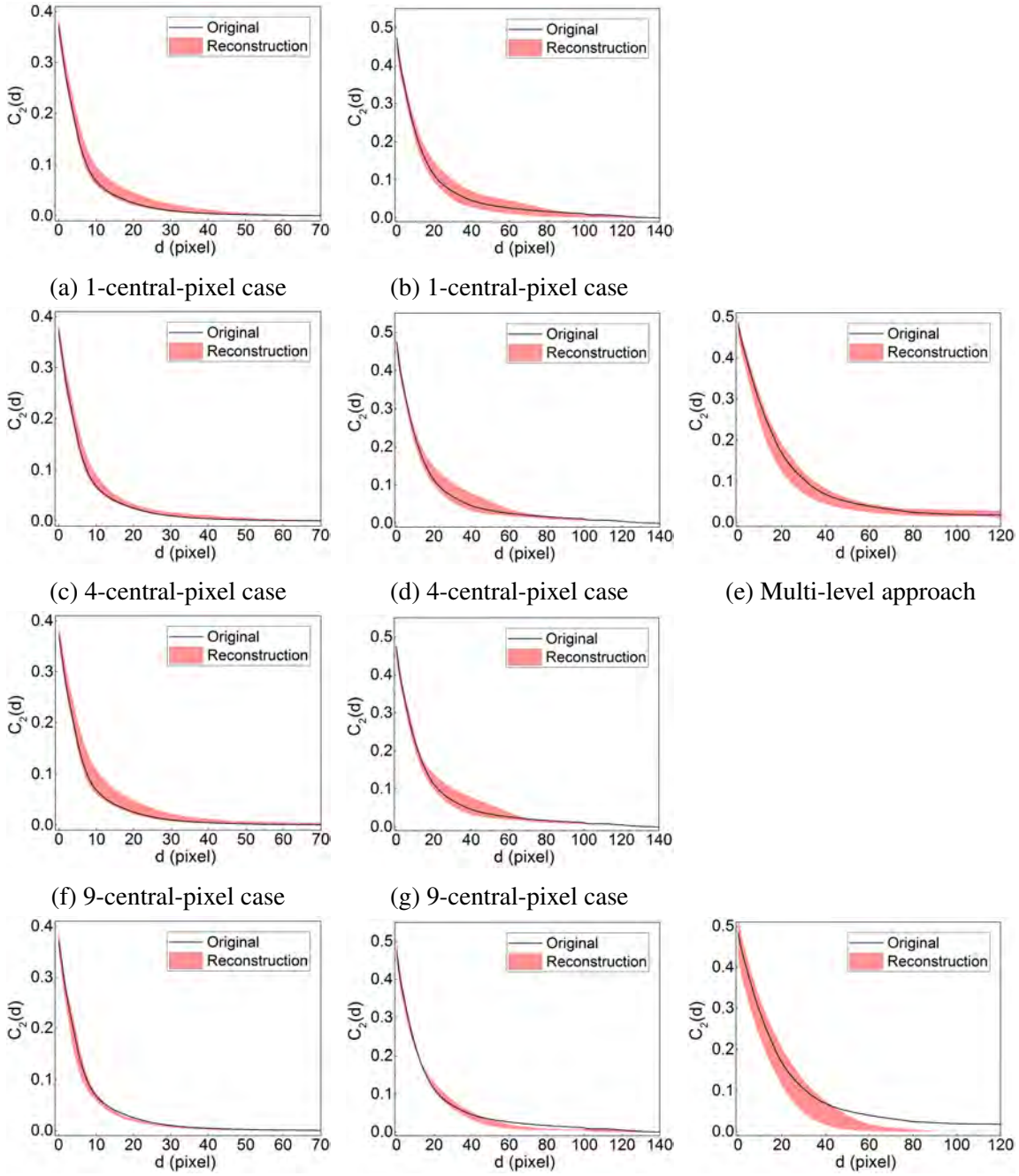
(c) 4-central-pixel case (d) 4-central-pixel case (e) Multi-level approach

(f) 9-central-pixel case (g) 9-central-pixel case

(h) Bostanabad's method (i) Bostanabad's method (j) Bostanabad's method

Fig. 4.17 Comparison of TPCF $S_2(d)$ between the original and reconstructed microstructures from the proposed method and Bostanabad's method (the three columns of figures from left to right are TPCFs corresponding to the microstructure IV, V and VI respectively)

time consuming than the training of the decision tree model, but the former model adopted by the proposed methods can better learn and express the complexity of microstructures



(a) 1-central-pixel case (b) 1-central-pixel case

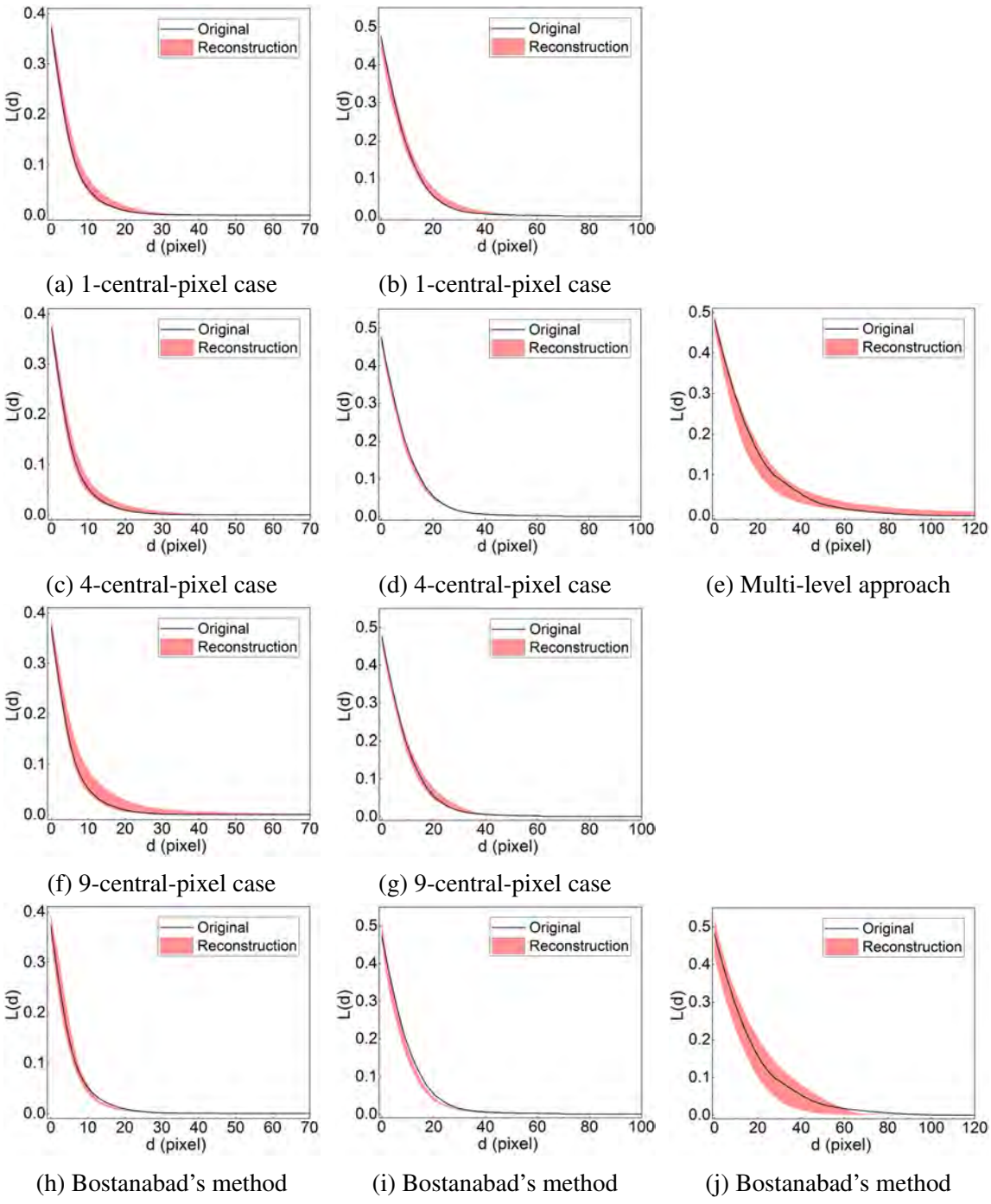
(c) 4-central-pixel case (d) 4-central-pixel case (e) Multi-level approach

(f) 9-central-pixel case (g) 9-central-pixel case

(h) Bostanabad's method (i) Bostanabad's method (j) Bostanabad's method

Fig. 4.18 Comparison of TPCCF $C_2(d)$ between the original and reconstructed microstructures from the proposed method and Bostanabad's method (the three columns of figures from left to right are TPCCFs corresponding to the microstructure IV, V and VI respectively)

with long-distance morphologies. Once the 'SSAE+Softmax' classifier model corresponding to a microstructure is properly trained, it can be saved for reuse. Using the 4-central-pixel



(a) 1-central-pixel case (b) 1-central-pixel case

(c) 4-central-pixel case (d) 4-central-pixel case (e) Multi-level approach

(f) 9-central-pixel case (g) 9-central-pixel case

(h) Bostanabad's method (i) Bostanabad's method (j) Bostanabad's method

Fig. 4.19 Comparison of LPF $L(d)$ between the original and reconstructed microstructures from the proposed method and Bostanabad's method (the three columns of figures from left to right are LPFs corresponding to the microstructure IV, V and VI respectively)

data template, the proposed method has a similar reconstruction efficiency as Bostanabad's method, while using the 9-central-pixel data template, the reconstruction efficiency of the

Table 4.4 Average reconstruction errors in volume fraction (ϕ), TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and LPF ($L(d)$), as well as the average computational costs of microstructure characterization and reconstruction

Microstructure	Data template (Method)	$\lvert\Delta\phi\rvert$	$\lvert\Delta S_2(d)\rvert$	$\lvert\Delta C_2(d)\rvert$	$\lvert\Delta L(d)\rvert$	Average CPU time (s)	
						Characterization	Reconstruction
IV	1-central-pixel	3.68%	3.15%	3.56%	2.40%	127.5	162.8
IV	4-central-pixel	3.35%	3.23%	2.27%	1.73%	284.0	49.6
IV	9-central-pixel	3.72%	4.19%	4.21%	3.62%	930.2	25.1
IV	Bostanabad's method	3.59%	5.12%	3.94%	3.48%	17.9	46.5
V	1-central-pixel	4.28%	4.35%	5.56%	2.67%	145.7	205.4
V	4-central-pixel	4.09%	4.43%	4.69%	1.39%	255.3	52.8
V	9-central-pixel	3.84%	3.59%	4.14%	2.36%	1032.7	27.5
V	Bostanabad's method	7.71%	8.04%	8.15%	4.25%	18.3	49.3
VI	Multi-level approach	4.82%	5.17%	3.71%	4.51%	465.1	90.2
VI	Bostanabad's method	10.31%	12.65%	32.46%	8.74%	19.2	51.4

proposed method gets doubled. In summary, the proposed method has unique advantages in accurately characterizing and efficiently reconstructing heterogeneous microstructures with long-distance correlations.

4.5 Extension for 3D Microstructures

The methodology presented in Section 4.2 and Section 4.3 can be readily extended to characterize and reconstruct 3D microstructures. The procedures of microstructure characterization and reconstruction are similar to that of 2D microstructures, while the main difference is on the data template. We introduce four data templates for 3D microstructures, as shown in Figure 4.20. The first two data templates have full neighboring voxels in cubic space, which cover the neighboring voxels that are prior to the central voxel(s) in the raster scan order, and the numbers of central voxels of these two templates are 1 and 8, respectively. Considering memory constraint and computational complexity, the number of neighboring voxels in the 3D data template can be reduced. Another two data templates with partial neighboring voxels are therefore created, and they only cover the neighboring voxels in three orthogonal planes, as shown in Figure 4.20c and 4.20d.

Although the partial templates are used for 3D microstructure characterization and reconstruction to reduce the memory requirement, sometimes the size of data template has to be very large in order to capture the largest morphological features in the training image, such as the long-distance connectivity inside the porous microstructures, which may exceed the computational capacity of ordinary computers. To overcome this limitation and efficiently capture long-distance features, we propose a multiple-level reconstruction strategy, which is explained in Section 4.5.2.

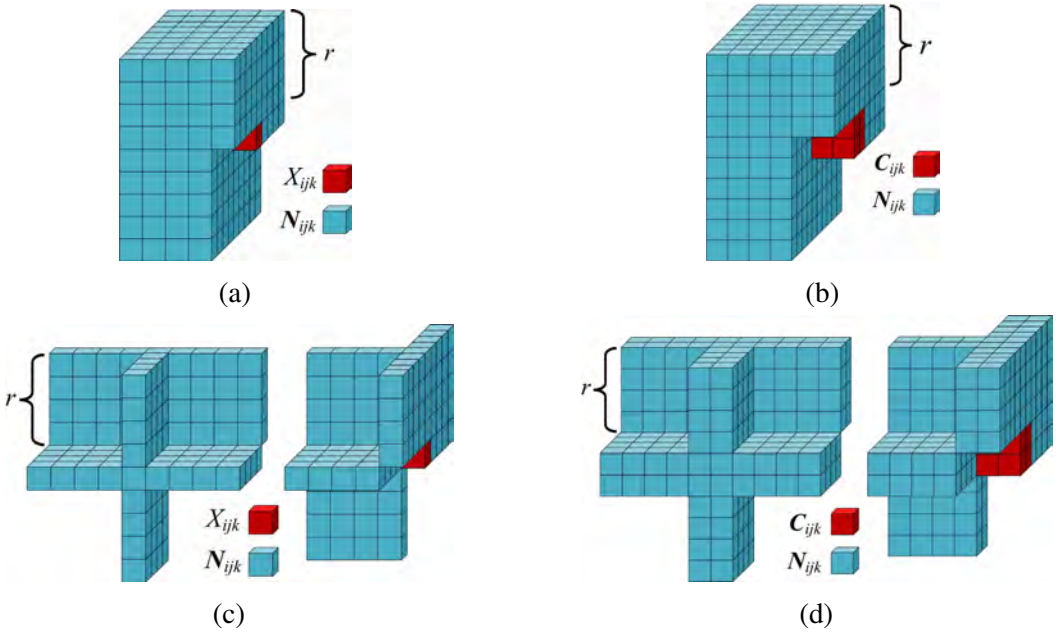


Fig. 4.20 The 3D data templates (neighborhood radius $r = 4$ voxels): (a) 1 central voxel with full neighboring voxels, (b) 8 central voxels with full neighboring voxels, (c) 1 central pixel with partial neighboring voxels, and (d) 8 central voxels with partial neighboring voxels

4.5.1 Single-level approach

The single-level approach characterizes and reconstructs 3D microstructures in the same procedure as that of 2D microstructures, as described in Algorithm 1. Firstly, the selected 3D data template is used to scan the entire 3D training image $\boldsymbol{X}$ to collect data events $(X_{ijk}, \boldsymbol{N}_{ijk})$ or $(\boldsymbol{C}_{ijk}, \boldsymbol{N}_{ijk})$, and then the 'SSAE+Softmax' classifier model is trained to estimate the CPDF $p(X_{ijk}|\boldsymbol{N}_{ijk})$ or $p(\boldsymbol{C}_{ijk}|\boldsymbol{N}_{ijk})$ of the 3D training image. Finally, 3D statistically equivalent samples can be generated by using the trained 'SSAE+Softmax' classifier model which is considered as an implicit form of CPDF $p(X_{ijk}|\boldsymbol{N}_{ijk})$ or $p(\boldsymbol{C}_{ijk}|\boldsymbol{N}_{ijk})$. A 3D microstructure samples is reconstructed slice by slice, and the voxels on each slice are generated in the raster scan order.

Original microstructures

In this section, three 3D microstructures with different morphological features are selected to demonstrate the single-level approach, as shown in Figure 4.21. The first microstructure is a clustered isotropic nanocomposite with silica inclusions randomly embedded in the rubber matrix. The volume fraction of silica inclusion is 9.55%, and the size of the training image is $150 \times 150 \times 150$ voxels. The second microstructure is anisotropic, and its

(a) Microstructure I (b) Microstructure II (c) Microstructure III

Fig. 4.21 The 3D training images of three microstructures with different morphological features

morphological features in orthogonal directions are visibly different in terms of size and shape of the secondary phase. The volume fraction of the secondary phase is 39.16%, and the image size is $100 \times 100 \times 100$ voxels. The last image is a porous microstructure of synthesis silica with pore space (white phase) occupying 42.73% of the bulk volume, which has long-distance morphological features and the image size is $150 \times 150 \times 150$ voxels. Different data templates are used to characterize and reconstruct these microstructures, and a set of 50 samples are generated for each case. The statistical equivalence between the original microstructures and the reconstructed samples is evaluated by comparing different morphological descriptors, including TPCF $S_2(d)$, TPCCF $C_2(d)$, and LPF $L(d)$, which are all computed from the white phase in each microstructure. The neighborhood radius of data template is determined by the correlation length measured from the TPCF, which can be seen in Figure 18. The parameters of the 'SSAE+Softmax' classifier model are listed in Table 4.5.

Table 4.5 The parameters for 'SSAE+Softmax' classifier model

Microstructure	Data template	Template size r	Unite number of each layer		
			Input layer	Hidden layer 1	Hidden layer 2
I	1, full	7 voxels	1687	220	80
I	1, partial	7 voxels	315	100	50
II	1, partial	11 voxels	759	160	70
II	8, partial	10 voxels	1320	220	100
III	1, partial	13 voxels	1053	200	85
III	8, partial	12 voxels	1872	300	120

Reconstruction results

As shown in Figure 4.22, the three columns of images are the representative reconstructions for the three original microstructures in Figure 4.21, respectively. These reconstructed samples have the same sizes as the corresponding training images, and they are reconstructed by using different data templates in order to study the influence of data template on microstructure reconstruction. Visual comparisons between the original microstructures and reconstructed samples imply that the morphological features of the original microstructures, e.g. stochasticity, anisotropy and connectivity, can be well preserved by the reconstructed samples. To evaluate the statistical equivalence between the original microstructure and corresponding reconstructed samples, as well as to quantify the variation of reconstructed samples, morphological descriptors are compared and the results are shown in Figures 4.23, 4.24 and 4.25. The red intervals in Figures 4.23, 4.24 and 4.25 are $S_2(d)$, $C_2(d)$ and $L(d)$ computed from the reconstructed microstructure samples. The black solid curves are the descriptors extracted from original microstructures, and they are all located inside the red intervals. The L_2 norm error between the averages of morphological descriptors extracted from reconstructed samples and that of the original microstructure are usually less than 5.00% (as listed in Table 4.6), which confirms statistical equivalence is well preserved.

Two data templates with full and partial neighborhoods are used to reconstruct samples for the first microstructure, whose morphological features are relatively simple and in small size. Although the reconstructed sample by using these two data templates are visually similar, the data template covering full neighborhood generates samples with smaller errors and less variation, which can be observed from Figures 4.23, 4.24 and 4.25. The data template covering partial neighborhoods can also produce samples with acceptable error and variation, and it significantly reduces the memory requirement and time cost of microstructure characterization and reconstruction. As to the second and third microstructures, they have long-distance morphologies, so the sizes of data templates need to be large enough to capture the main morphological features. The data templates covering partial neighborhoods are adopted to reconstruct samples for them, because these data templates require less memory to operate. No significant difference is observed between the reconstructed samples using partial data templates with 1 and 8 central voxel(s). The reconstructed samples are all visually similar, and the statistical evaluations show similar accuracy level, as shown in Figures 4.23, 4.24, 4.25. According to Table 4.6, the efficiencies of characterization and reconstruction by using different data templates show great difference. The time cost of microstructure characterization by using 8-central-voxel template is higher than using 1-central-voxel template, while the cost comparison for microstructure reconstruction is opposite. The trained

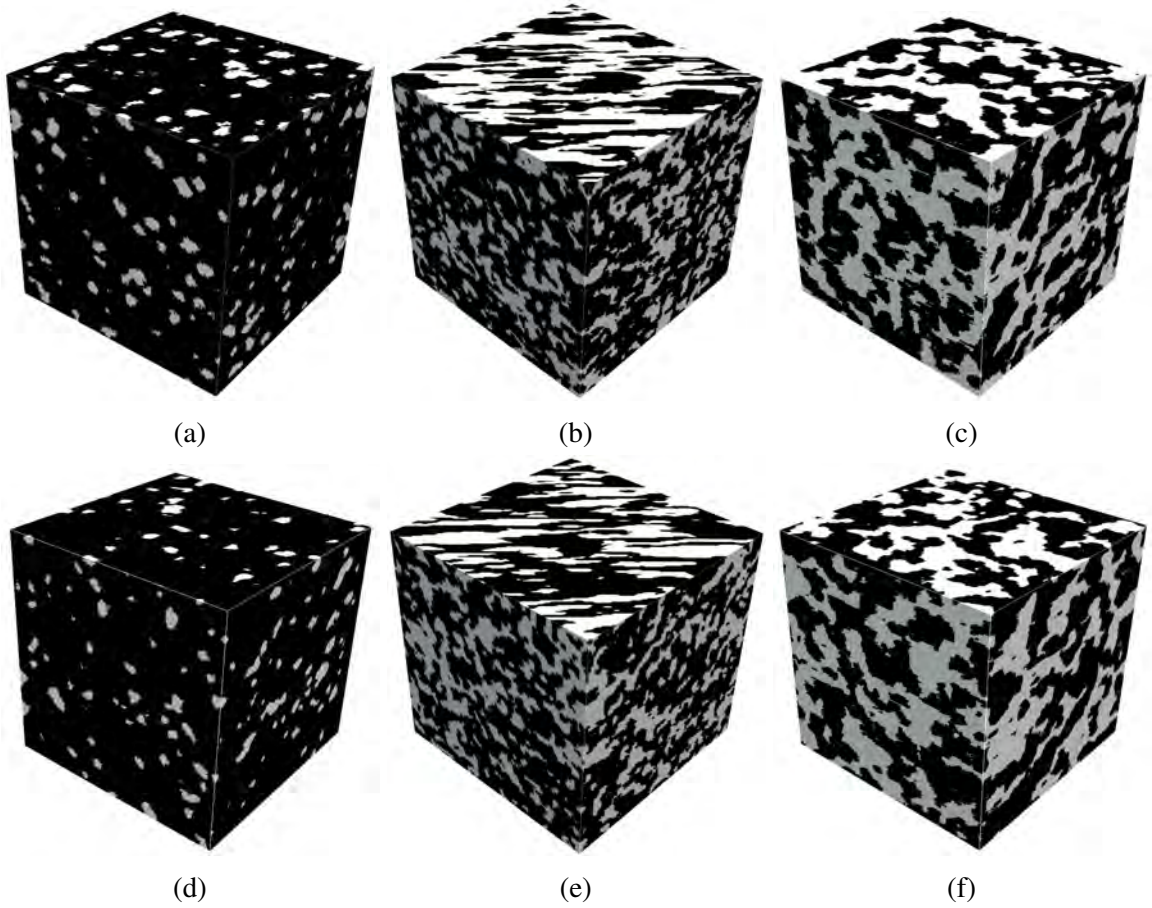


Fig. 4.22 Microstructure reconstruction results by using different data templates: (a) 1 central voxel with full neighborhoods, (b) 1 central voxel with partial neighborhoods, (c) 1 central voxel with partial neighborhoods, (d) 1 central voxel with partial neighborhoods, (e) 8 central voxel with partial neighborhoods, and (f) 8 central voxels with partial neighborhoods

'SSAE+Softmax' classifier model can be stored for reuse, so microstructure reconstruction by using 8-central-voxel template is time-saving in the long run.

Table 4.6 Average reconstruction errors of volume fraction (ϕ), TPCF ($S_2(d)$), TPCCF ($C_2(d)$), and LPF ($L(d)$), as well as the average computational costs of microstructure characterization and reconstruction

Microstructure (Data template)	$\lvert\Delta\phi\rvert$	$\lvert\Delta S_2(d)\rvert$	$\lvert\Delta C_2(r)\rvert$	$\lvert\Delta L(d)\rvert$	Average CPU time (s)	
					Characterization	Reconstruction
I (1, full)	2.25%	2.02%	1.85%	1.21%	10142.2	8029.4
I (1, partial)	2.83%	2.89%	2.49%	2.31%	3669.0	7783.9
II (1, partial)	4.35%	4.77%	4.77%	4.59%	4202.7	3038.4
II (8, partial)	4.07%	4.07%	4.06%	3.88%	6329.4	819.8
III (1, partial)	4.65%	4.58%	4.58%	4.68%	5068.8	8077.4
III (8, partial)	4.88%	4.45%	4.46%	4.23%	12142.5	1809.4

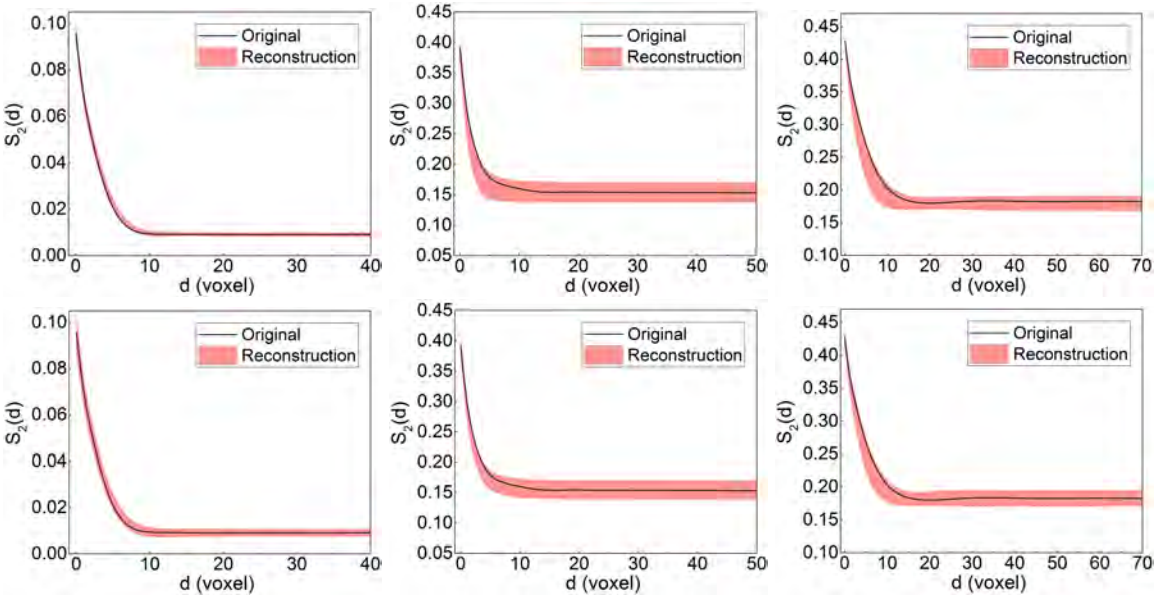


Fig. 4.23 Comparison of TPCF $S_2(d)$ between the original and reconstructed microstructures

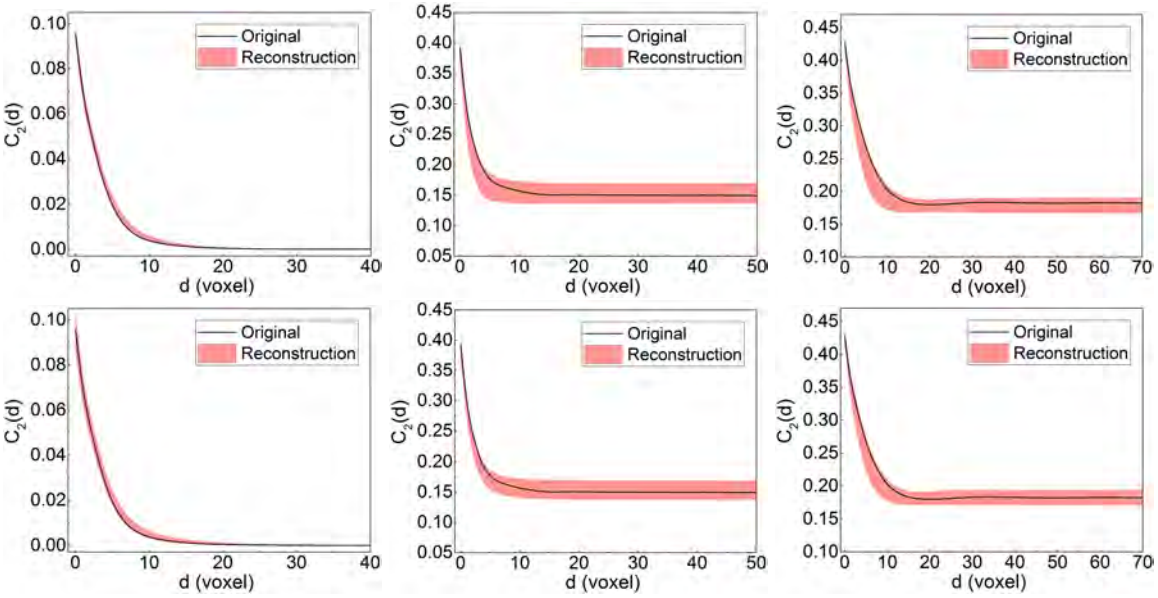


Fig. 4.24 Comparison of TPCCF $C_2(d)$ between the original and reconstructed microstructures

4.5.2 Multi-level approach

The single-level approach captures the morphological features of microstructures by using data templates of a sufficiently large size. However, due to limitations in computer memory and computing power, it may not be adequate to capture the long-distance connectivity of complicated pore network within such porous microstructure as sandstone. In this section, we

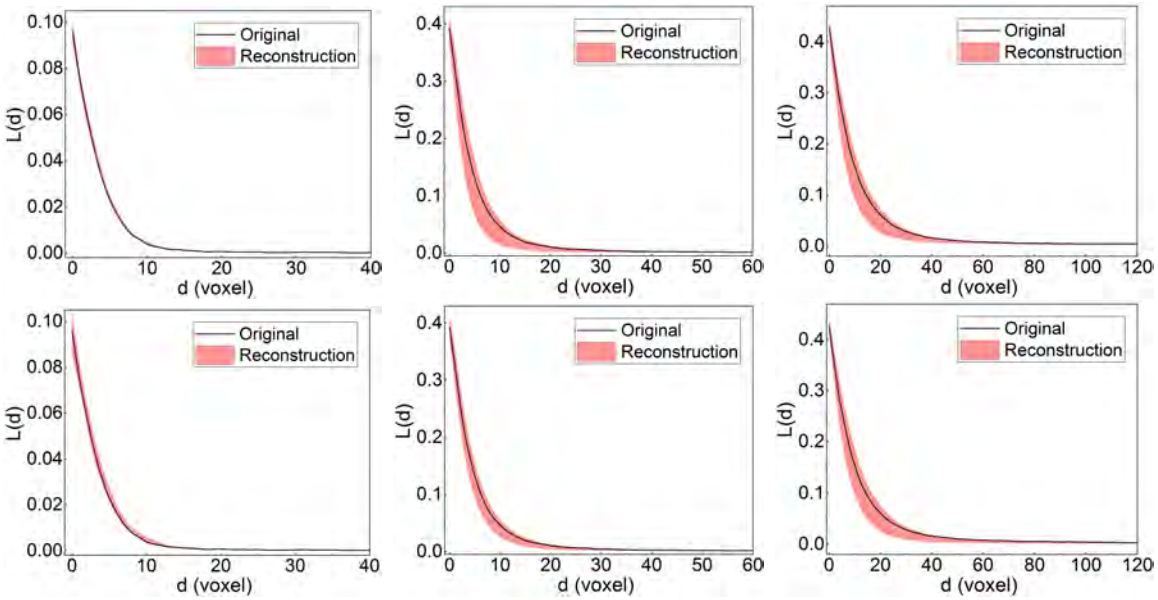


Fig. 4.25 Comparison of LPF $L(d)$ between the original and reconstructed microstructures

solve this problem by using a multi-level image pyramid, where the large-scale morphological features can be represented more compactly by using a data template with a smaller size at the lower image pyramid level. The multi-level approach is particularly developed to characterize and reconstruct porous microstructures with long-distance connectivity.

The three-level approach is taken as an example to explain the procedures to characterize and reconstruct 3D microstructures in a multi-level manner. First, a 3D Gaussian pyramid [65] is built in three levels for the original microstructure. The three-level image pyramid plays as the training images for microstructure characterization and reconstruction at different levels, which can be denoted as $\boldsymbol{X}_{\text{high}}$, $\boldsymbol{X}_{\text{middle}}$, and $\boldsymbol{X}_{\text{low}}$, respectively. Then, for the first level, $\boldsymbol{X}_{\text{low}}$ is used as the training image, and an 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{low}}$ is trained to learn the morphological features of $\boldsymbol{X}_{\text{low}}$ by using a selected data template. Based on the fitted 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{low}}$, a microstructure sample $\boldsymbol{Y}_{\text{low}}$ can be generated at the lowest level. Basically, the microstructure characterization and reconstruction at the first level is the same as the single-level approach. Next, for the second level, $\boldsymbol{X}_{\text{middle}}$ and $\boldsymbol{X}_{\text{low}}$ are used together to serve as the training images. The only modification is that each set of neighboring voxels $\boldsymbol{N}$ not only contains the neighboring voxels in current level but also covers the neighboring voxels and central voxel(s) in the lower level, in order to maintain the morphological consistency between $\boldsymbol{X}_{\text{middle}}$ and $\boldsymbol{X}_{\text{low}}$. The second 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{middle}}$ is trained by using the collected training data, based on which a microstructure sample $\boldsymbol{Y}_{\text{middle}}$ can be produced conditional to the reconstructed sample $\boldsymbol{Y}_{\text{low}}$ at the lower level. Finally, as to the third level, the procedures are the same as that of

the second level, where an 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{high}}$ at the highest level is trained based on $\boldsymbol{X}_{\text{high}}$ and $\boldsymbol{X}_{\text{middle}}$, after which the final microstructure sample $\boldsymbol{Y}_{\text{high}}$ can be reconstructed conditional to $\boldsymbol{Y}_{\text{middle}}$.

In summary, the multi-level approach characterizes and reconstructs microstructures in a hierarchical way, where the low-frequency (long-distance) morphological features are captured at a lower level, conditional to which the high-frequency details are added at the higher levels. The same procedures can also be used to characterize and reconstruct 2D microstructures with long-distance morphologies, as discussed in Section 4.4.2. The algorithmic workflow of the multi-level approach is summarized in Algorithm 2, and a schematic illustration is provided in Figure 4.26.

Algorithm 2: Stochastic characterization and reconstruction of 3D heterogeneous microstructures using the three-level approach

Microstructure Characterization:

Data: The Gaussian pyramid with three levels playing as the training images $\boldsymbol{X}_{\text{high}}$, $\boldsymbol{X}_{\text{middle}}$, and $\boldsymbol{X}_{\text{low}}$;

Data: The selected data templates $\boldsymbol{DT}_{\text{high}}$, $\boldsymbol{DT}_{\text{middle}}$, and $\boldsymbol{DT}_{\text{low}}$ for different levels.

Collect the data events $(\boldsymbol{C}_{\text{low}}, \boldsymbol{N}_{\text{low}})$ from the training image $\boldsymbol{X}_{\text{low}}$ by using the data template $\boldsymbol{DT}_{\text{low}}$, and then train the 'SSAE+Softmax' classifier models $\boldsymbol{M}_{\text{low}}$;

Collect the data events $(\boldsymbol{C}_{\text{middle}}, \boldsymbol{N}_{\text{middle}} + \boldsymbol{N}_{\text{low}} + \boldsymbol{C}_{\text{low}})$ from the training image $\boldsymbol{X}_{\text{middle}}$ and $\boldsymbol{X}_{\text{low}}$ by using the data templates $\boldsymbol{DT}_{\text{middle}}$ and $\boldsymbol{DT}_{\text{low}}$, and then train the 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{middle}}$;

Collect the data events $(\boldsymbol{C}_{\text{high}}, \boldsymbol{N}_{\text{high}} + \boldsymbol{N}_{\text{middle}} + \boldsymbol{C}_{\text{middle}})$ from the training image $\boldsymbol{X}_{\text{high}}$ and $\boldsymbol{X}_{\text{middle}}$ by using the data templates $\boldsymbol{DT}_{\text{high}}$ and $\boldsymbol{DT}_{\text{middle}}$, and then train the 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{high}}$;

Result: The trained 'SSAE+Softmax' classifier models $\boldsymbol{M}_{\text{high}}$, $\boldsymbol{M}_{\text{middle}}$ and $\boldsymbol{M}_{\text{low}}$.

Microstructure Reconstruction:

Data: The trained 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{high}}$, $\boldsymbol{M}_{\text{middle}}$ and $\boldsymbol{M}_{\text{low}}$;

Data: The selected data templates $\boldsymbol{DT}_{\text{high}}$, $\boldsymbol{DT}_{\text{middle}}$, and $\boldsymbol{DT}_{\text{low}}$ for different levels.

Reconstruct the first-level microstructure $\boldsymbol{Y}_{\text{low}}$ based on the trained 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{low}}$;

Conditional to $\boldsymbol{Y}_{\text{low}}$, reconstruct the second-level microstructure $\boldsymbol{Y}_{\text{middle}}$ based on the trained 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{middle}}$;

Conditional to $\boldsymbol{Y}_{\text{middle}}$, reconstruct the third-level microstructure $\boldsymbol{Y}_{\text{high}}$ based on the trained 'SSAE+Softmax' classifier model $\boldsymbol{M}_{\text{high}}$;

Result: A statistically equivalent reconstructed microstructure sample $\boldsymbol{Y}_{\text{high}}$.

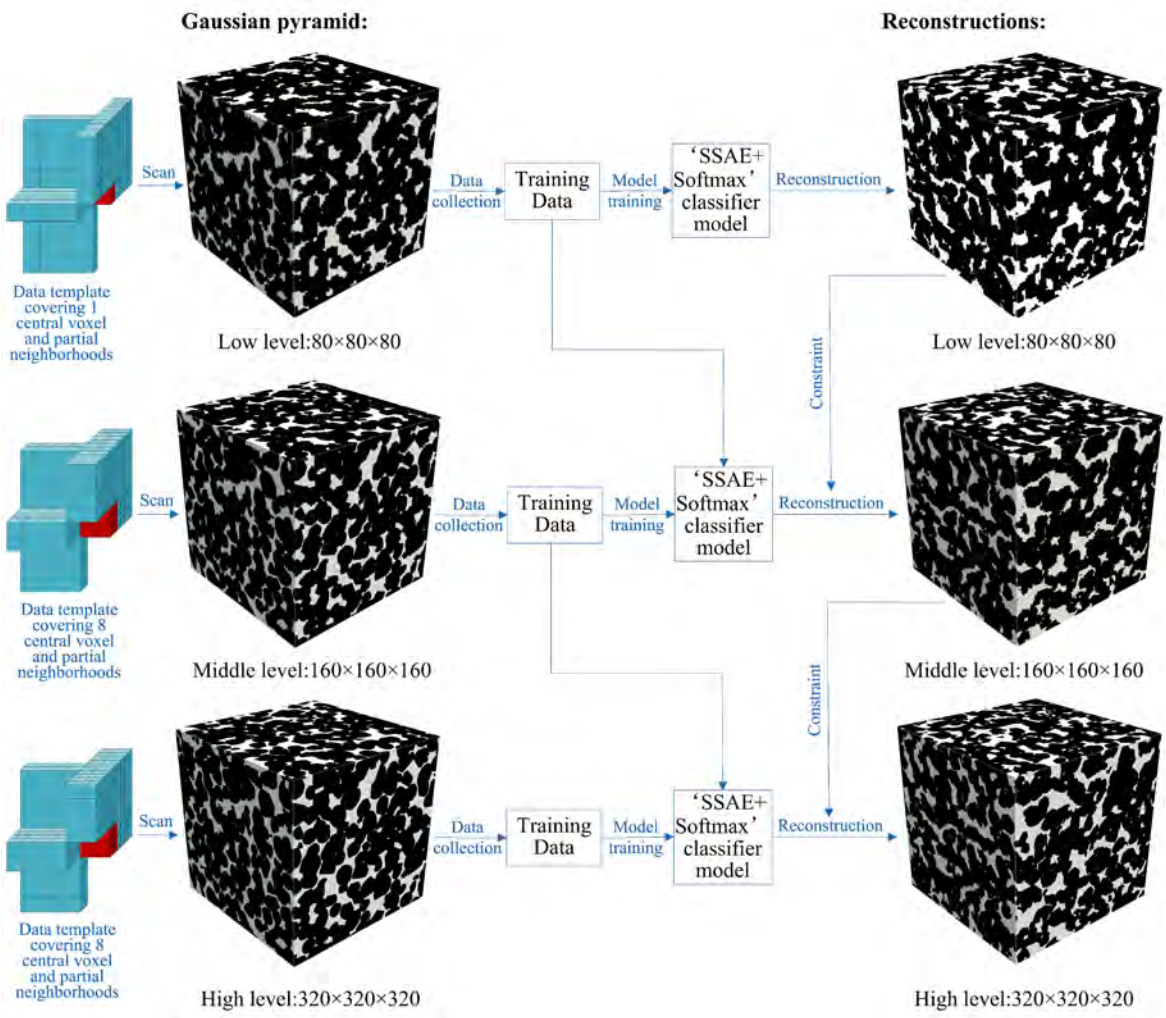


Fig. 4.26 Schematic diagram of microstructure characterization and reconstruction using the three-level approach

Original microstructures

Fontainebleau sandstone is chosen to test the performance of the proposed multi-level approach to characterize and reconstruct porous microstructures, and its reconstruction results are also compared with that of Bostanabad's method [52, 53]. Fontainebleau sandstone [341] is an isotropic porous medium, which is often used as the reference standard for validating models of porous media due to its special properties. It consists of monocrystalline quartz grains with an average size of about 200 μm and it does not contain clay, inside which only inter-granular porosity exists. A group of 3D digital microstructures (resolution: 5.7μm) of three Fontainebleau sandstone samples are used, whose porosities are 15.6%, 20.9%, and 24.5%, respectively [34]. The digital microstructures are shown in Figure 4.27, where

the white region denotes the pore space and the black region denotes the solid matrix. These digital microstructures are used as training images for stochastic characterization and reconstruction.

(a) ϕ =15.6% (b) ϕ =20.9% (c) ϕ =24.5%

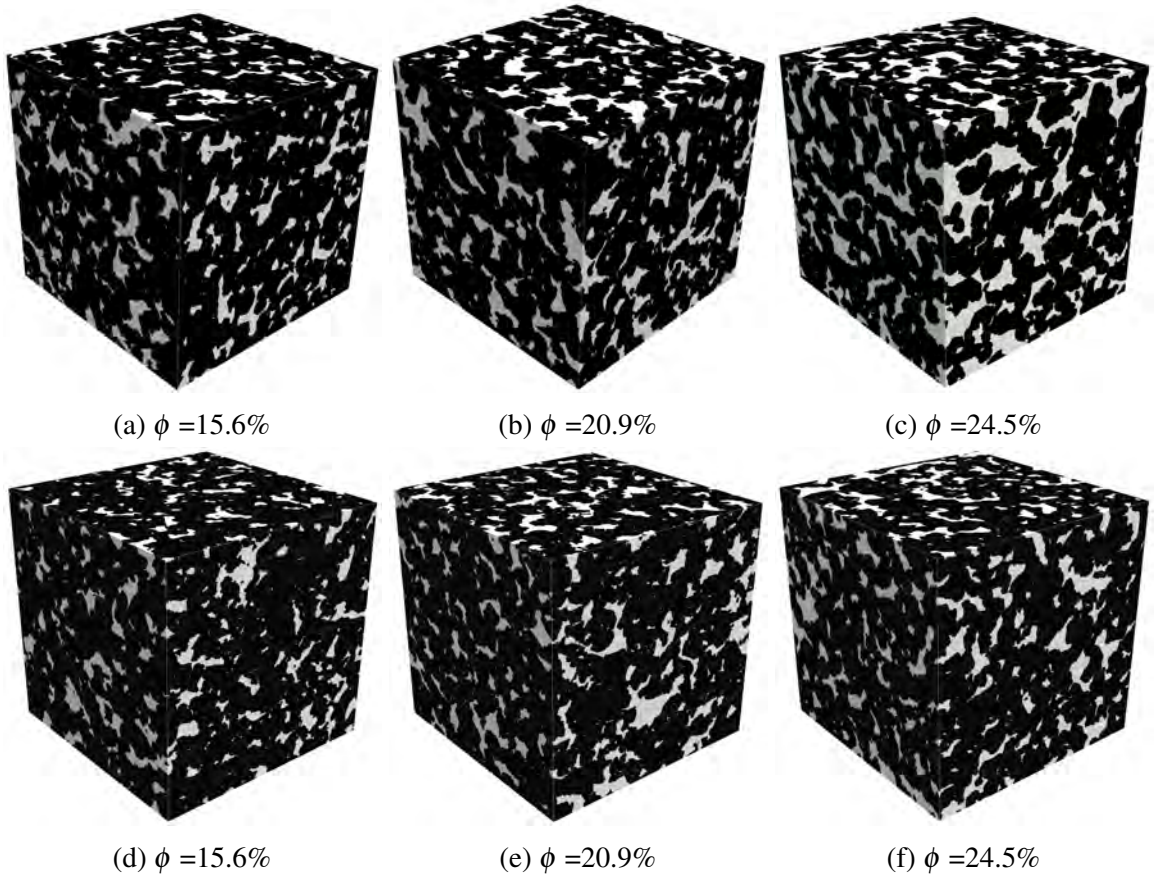
Fig. 4.27 The 3D training images (320×320×320) of Fontainebleau sandstone microstructures with different porosities ϕ

The three-level approach is used to characterize and reconstruct the microstructures of Fontainebleau sandstones, so as to capture the long-distance connectivity as well as the geometrical complexity of the pore networks. For the first level, the data template $\boldsymbol{DT}_{\text{low}}$ covering 1 central voxel and partial neighboring voxels is selected, and the radius of $\boldsymbol{DT}_{\text{low}}$ is 7 voxels which can be measured from the TPCFs $S_2(d)$ of the training images $\boldsymbol{X}_{\text{low}}$. As to the second level, the training images $\boldsymbol{X}_{\text{middle}}$ become finner than $\boldsymbol{X}_{\text{low}}$, and the data template $\boldsymbol{DT}_{\text{middle}}$ covering 8 central voxels and partial neighboring voxels is chosen, because 8 voxels in the 3D image of current level covers the same domain of 1 voxel in the 3D image of the lower level. Limited by the computing capacity, the radius of $\boldsymbol{DT}_{\text{middle}}$ is also 7 voxels in this study. For the third level, the data template $\boldsymbol{DT}_{\text{high}}$ covering 8 central voxels and partial neighboring voxels is used, and the radius of $\boldsymbol{DT}_{\text{high}}$ is set to be 6 voxels. In additional, the data template with partial neighborhoods is use for Bostanabad's method, and the size of data template is determined by correlation length.

Reconstruction results

For each Fontainebleau sandstone sample, a group of 30 microstructure samples are reconstructed by the proposed multi-level approach and Bostanabad's method individually, and representative reconstruction results are shown in Figure 4.28. The Fontainebleau sandstone sample with the porosity of 24.5% is taken as the example to illustrate the characterization and reconstruction procedure in Figure 4.26. The Gaussian pyramid with three levels is used

as the training images, where the training image is half sized from a high level to a lower level with the main morphology maintained. At the low level, the long-distance morphology is captured compactly by using a data template of a relatively small size. Conditional to the reconstruction result at the low level, morphological details are added to the reconstruction at the middle level, which makes the reconstruction result much closer to the original microstructure, especially in terms of the local connectivity between isolated pores. After adding more high-frequency morphology features at the highest level, the final reconstruction result becomes visually indistinguishable compared to the original microstructure. By contrast, the reconstructed samples from Bostanabad's method have smaller pores and less pore connectivity, as shown in Figure 4.28.

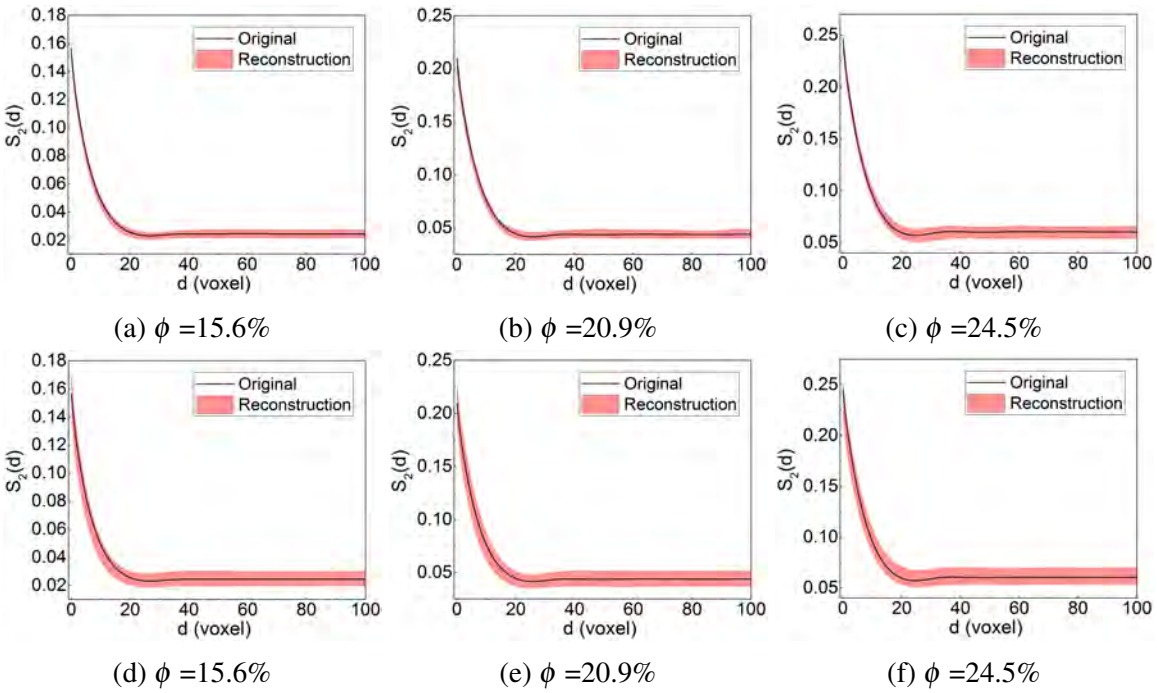


(a) ϕ =15.6% (b) ϕ =20.9% (c) ϕ =24.5%

(d) ϕ =15.6% (e) ϕ =20.9% (f) ϕ =24.5%

Fig. 4.28 The representative reconstruction results (320×320×320) for Fontainebleau sandstone samples by using different methods: the proposed multi-level approach (a)∼(c) and Bostanabad's method (d)∼(f)

Statistical equivalence between the original and reconstructed microstructures is accessed by TPCF, as shown in Figure 4.29. Both the proposed multi-level approach and Bostanabad's method show good performance to maintain the stochastic features of original microstructures,



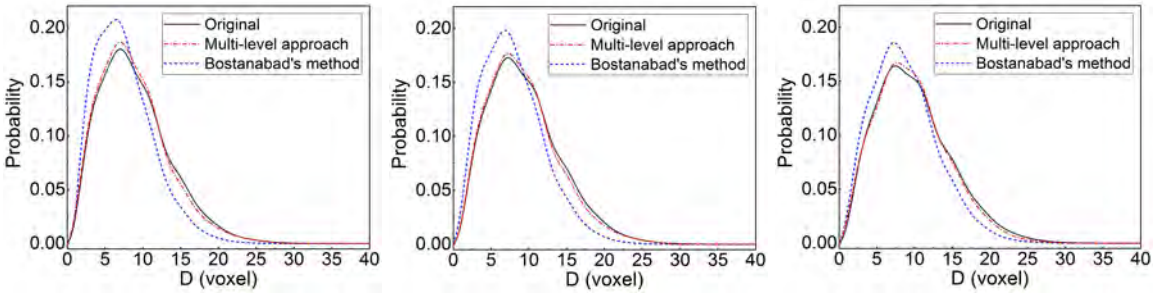
(a) ϕ =15.6% (b) ϕ =20.9% (c) ϕ =24.5%

(d) ϕ =15.6% (e) ϕ =20.9% (f) ϕ =24.5%

Fig. 4.29 Comparison of TPCF $S_2(d)$ between the original and reconstructed microstructures from different methods: the proposed multi-level approach (a)∼(c) and Bostanabad's method (d)∼(f)

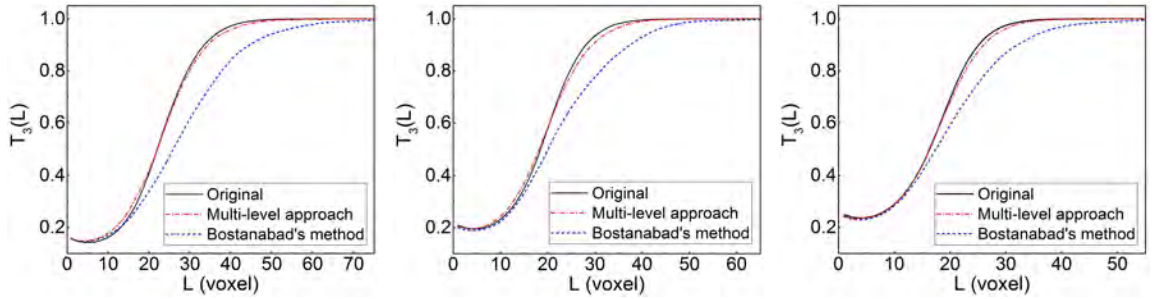
but the multi-level approach has higher accuracy in reconstructing samples in terms of TPCF. Pore size distribution (PSD) $P(D)$ [281] is used to evaluate the reconstruction results from the angle of pore geometry, and it measures the distribution of equivalent diameters D of pore bodies. In Figure 4.30, the black solid curve is the PSD extracted from the original microstructure, while the red and blue dashed curves are the average results of PSDs calculated from 30 reconstructed samples. Bostanabad's method tends to generate smaller pores inside the reconstructed samples compared to the pores inside the original microstructures, and pore geometry information at a large scale is not fully captured. The proposed multi-level approach is however able to accurately capture the geometry information of both small and large pores, which is confirmed by the good agreements of PSDs extracted from the reconstructed and original microstructures. Total fraction of percolating cells (TFPC) $T_3(L)$ [341] measures topology information of pore network, which reflects the connectivity characteristics of porous media. As shown in Figure 4.31, the average TFPCs extracted from the 30 samples reconstructed by Bostanabad's method are just below the reference TFPCs, especially at long distance, which means the reconstructed samples have worse pore connectivity compared to the original microstructure. However, the long-distance connectivity of the pore network is well preserved in the samples reconstructed by the

proposed multi-level approach, which is confirmed by the good matches of TFPCs between the original and reconstructed microstructures. In general, the proposed multi-level approach is very powerful to capture the morphological features of porous microstructures with long-distance connectivity.



(a) ϕ =15.6% (b) ϕ =20.9% (c) ϕ =24.5%

Fig. 4.30 Comparison of PSD $P(D)$ between the original and reconstructed microstructures



(a) ϕ =15.6% (b) ϕ =20.9% (c) ϕ =24.5%

Fig. 4.31 Comparison of TFPC $T_3(L)$ between the original and reconstructed microstructures

Transport properties

In this section, transport properties, including intrinsic permeability and diffusional tortuosity factor, are compared between the original and reconstructed microstructures to assess the reconstruction quality. Both intrinsic permeability and diffusional tortuosity factor are directly related to geometrical attributes of the porous microstructure, such as porosity, specific surface, and pore/throat size distribution [34, 96]. Lattice Boltzmann Method (LBM) [252, 142] and Finite Difference Method (FDM) [269] are adopted to evaluate the permeability and tortuosity factor through open-source solvers: *Palabos* [114] and *TauFactor* [96], respectively. These two open-source solvers are widely used by the research community of porous media, where pore-scale simulations of fluid flow and molecular

diffusion can be directly performed on the digital microstructures by using the voxels as the mesh elements, and they have good flexibility to work with complex geometries.

The LBM is used to simulate a single-phase fluid flow with low Reynolds number ($Re \ll 1$) passing through a porous medium under a constant pressure gradient. When the fluid flow reaches a steady state, it can be described by Darcy's law, and the intrinsic permeability κ of the porous medium is quantified by the following equation:

$$\kappa = -\frac{\mu}{\nabla p}\langle \mathbf{u} \rangle \tag{4.21}$$

where ∇p is the pressure gradient along the direction of fluid flow, μ is the dynamic viscosity of the fluid, and $\langle \mathbf{u} \rangle$ denotes the average fluid velocity across the simulation domain.

The FDM is adopted to simulate the steady diffusive flow through a porous medium under a constant concentration gradient. Compared to the free diffusion without geometry constraint, the diffusive transport through the porous medium is remarkably decreased because of the convolutions of flow paths inside the porous medium. The steady-state diffusive flow is described by the Fick's first law, and the decrease of diffusive transport in a porous medium is quantified by tortuosity factor τ [163]:

$$\mathbf{J} = -D_{\mathrm{e}} \nabla C \tag{4.22}$$

$$D_{\mathrm{e}} = D\frac{\phi}{\tau} \tag{4.23}$$

where $\mathbf{J}$ denotes the diffusion flux, D_{e} is the effective diffusivity, ∇C is the concentration gradient, D is the intrinsic diffusivity of the conductive phase (gas or liquid) filling the pores, and ϕ is porosity.

Performing LBM simulations on both the original and reconstructed microstructures of Fontainebleau sandstones, we obtain the velocity fields of fluid flow at steady states, as shown in Figure 4.32. Driven by the same pressure gradient, the flow velocity magnitude range inside the microstructure reconstructed from the multi-level approach is close to that inside the original microstructure. As to the microstructure reconstructed from Bostanabad's method, it has smaller velocity magnitude and fewer flow streamlines. The permeability results evaluated from LBM simulations are recorded in Figure 4.34a. It can be observed that the permeabilities of the samples reconstructed from the multi-level approach fluctuate around the reference values (red stars). The average permeability value of each group of 30 reconstructed samples is close to the corresponding reference permeability, with all errors less than 5.00%. The presence of permeability variation in reconstructed samples is expected, because these reconstructed samples are generated from probability sampling and they

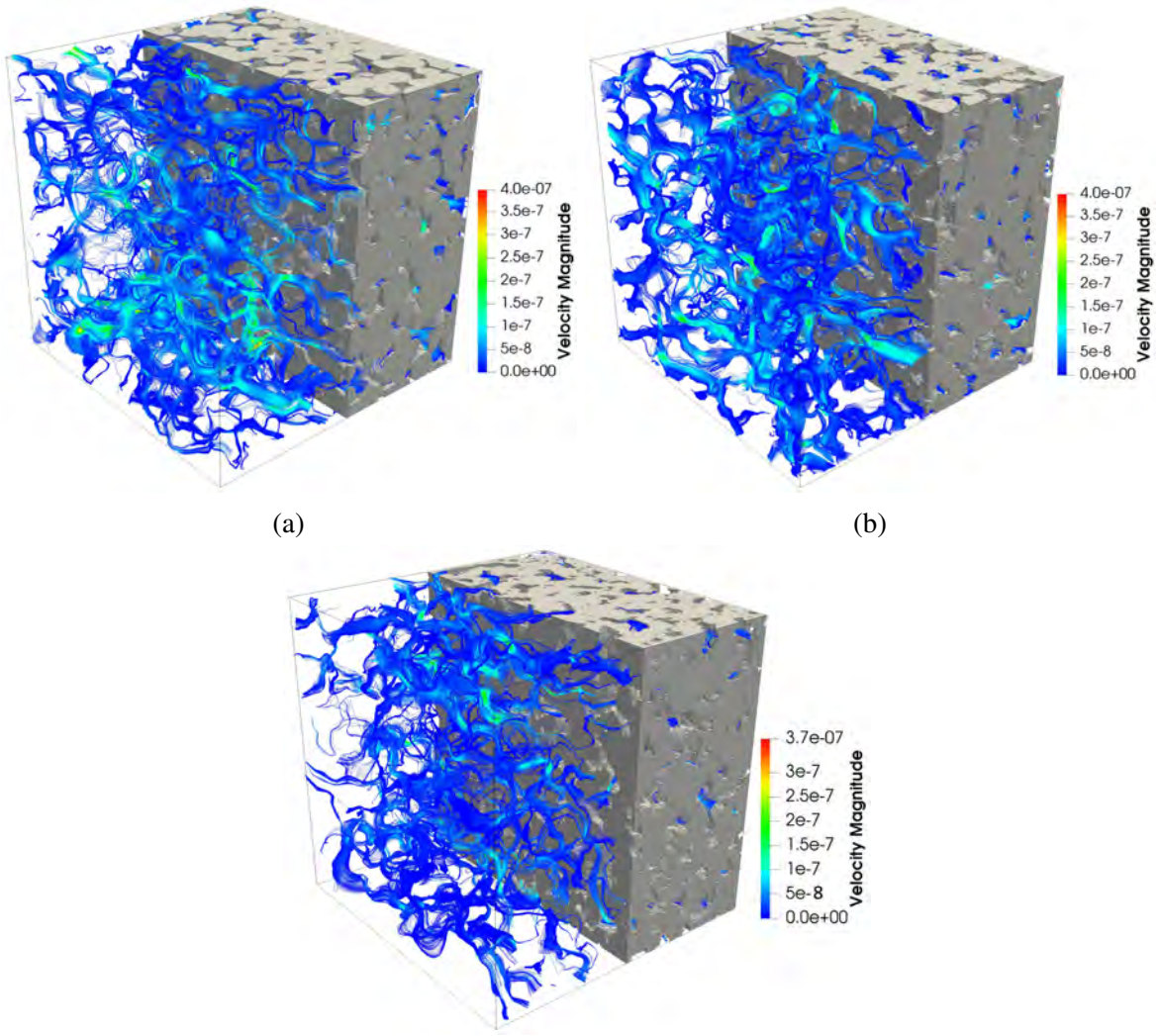


Fig. 4.32 The fluid velocity fields at steady state obtained from LBM simulations: (a) Original microstructure ($\phi = 24.5\%$); (b) The representative microstructure ($\phi = 24.5\%$) reconstructed from the multi-level approach; (c) The representative microstructure ($\phi = 24.5\%$) reconstructed from Bostanabad's method

statistically represent the heterogeneous nature of the concerned microstructures. However, Bostanabad's method fails to preserve the intrinsic permeabilities of original microstructures, and the reconstructed samples generally have significantly smaller permeabilities compared to the reference values, with errors ranging from -79.70% to -36.10%.

FDM simulations are applied to both the original and reconstructed microstructures, from which the steady-state flux density fields of diffusion are acquired. As shown in Figure 4.33, the output flux from the sample reconstructed by the multi-level method is close to that

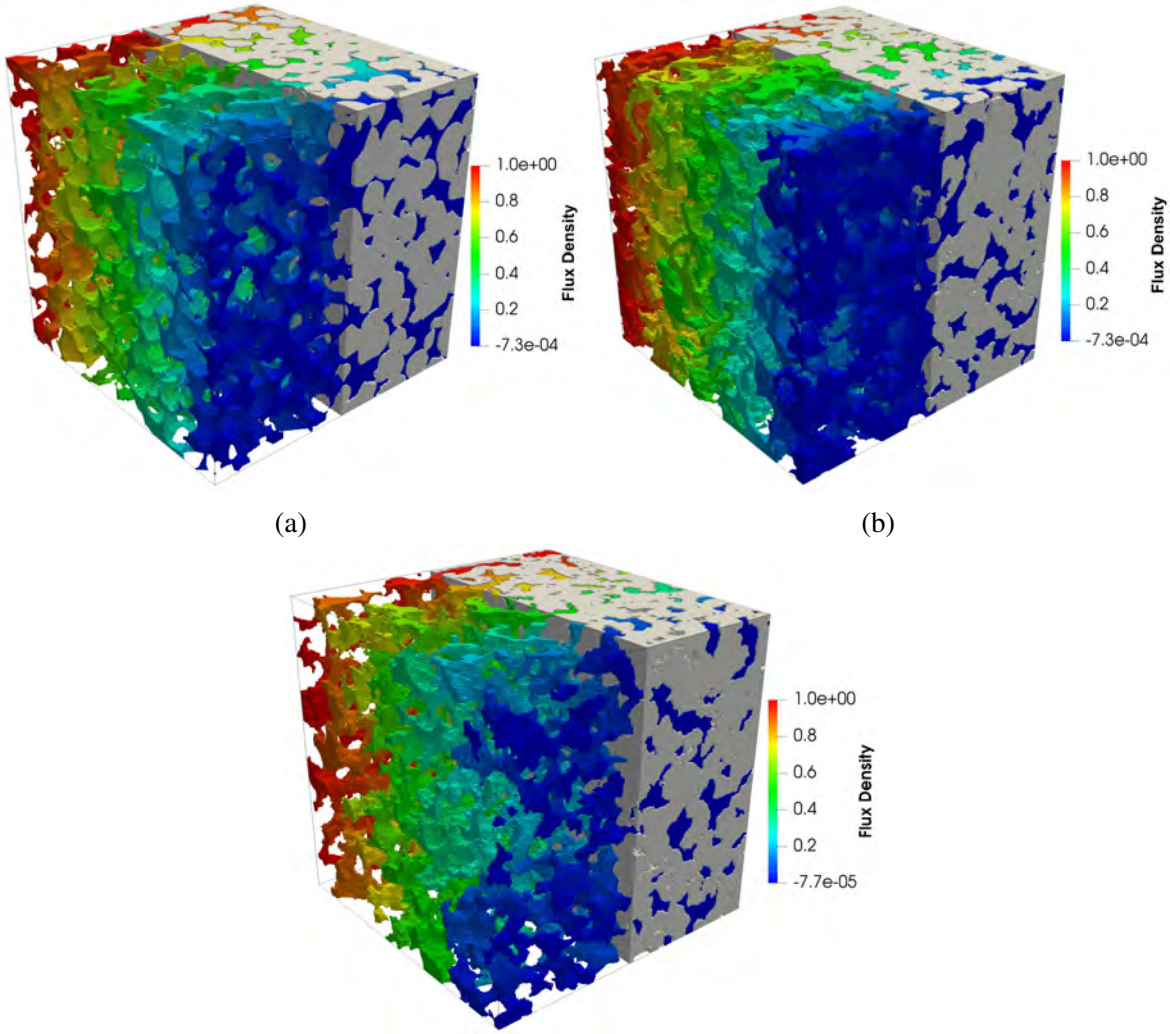
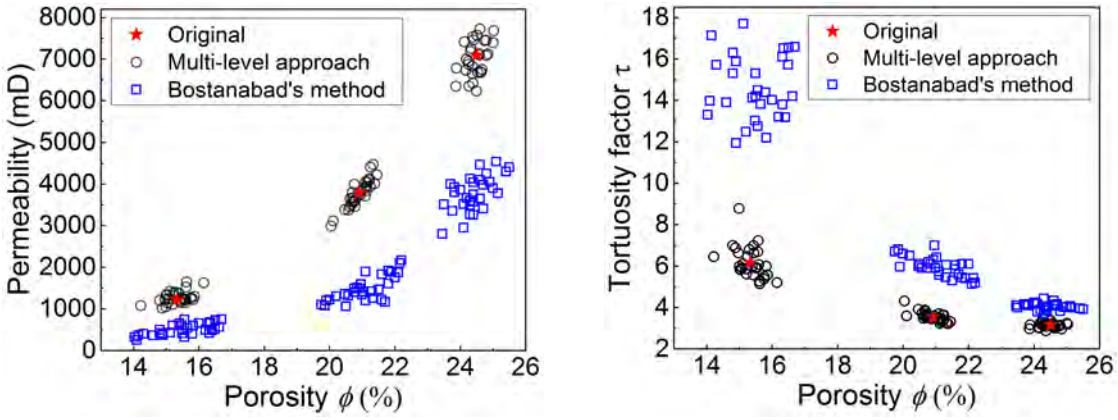


Fig. 4.33 The flux density fields at steady state obtained from FDM simulations: (a) Original microstructure ($\phi = 24.5\%$); (b) The representative microstructure ($\phi = 24.5\%$) reconstructed from the multi-level approach; (c) The representative microstructure ($\phi = 24.5\%$) reconstructed from Bostanabad's method

from the original microstructure, under the same concentrate gradient. But the reconstructed microstructure from Bostanabad's method outputs much less diffusional flux, compared to the original one. Tortuosity factors are evaluated from the steady-state flux density fields (as plotted in Figure 4.34b), which quantify the retarding effects of microstructures on diffusion. Again, the tortuosity factors of reconstructed samples from the multi-level method are located around the reference values, with the average errors smaller than 5.00%. However, Bostanabad's method is not able to maintain the tortuosity factors of original microstructures, and the tortuosity factors of its reconstructed sample are 25.80% to 186.70%



(a) Intrisinc permeability (b) Diffusional tortuosity factor

Fig. 4.34 The comparison of transport properties between the original and reconstructed microstructures

larger than the reference values. In summary, the proposed multi-level approach shows excellent performance to statistically preserve transport properties of porous media, and the reconstructed samples can be used for relevant researches.

4.6 Conclusions

The main contribution of this chapter is to present a versatile method that can accurately characterize and efficiently reconstruct heterogeneous microstructures using deep neural network (DNN), and it only requires one (or a few) real microstructure(s) to serve as the training image(s). The most prominent advantage of this method is its powerful capacity to capture long-distance morphological features, which overcomes the short-distance limitation of many MRF-based methods.

Specifically, the digital microstructure image is treated as an extended Markov random field (MRF), in order to improve microstructure characterization quality and reconstruction speed. Different data templates are designed to collect data events (local morphology patterns) from the target image, and these data events are used as training data to fit an 'SSAE+Softmax' model for the microstructure characterization purpose. In essence, the fitted 'SSAE+Softmax' model is an implicit representation of the conditional probability distribution function of pixels/voxels in the target image, from which statistically equivalent microstructure samples can be generated through a specific reconstruction procedure. Moreover, a multi-level approach is developed for microstructures with long-distance correlations, where microstructure characterization and reconstruction are carried out in a hierarchical manner, to avoid the short-distance limitation inherent in the MRF assumption.

The new method is tested on both 2D and 3D microstructures with distinct morphologies, and reconstruction qualities are evaluated through comparisons between target and reconstructed microstructures, in terms of statistical information, geometrical attribute, topological features and transport properties. The comparisons show excellent consistency, and the statistical equivalence, morphological similarity and transport properties are all well preserved by the reconstructed microstructure samples. In general, the proposed method is applicable to a variety of microstructures, especially suitable for microstructures with long-distance morphological features. It is also flexible to balance the computation memory requirement and reconstruction speed by choosing different data templates or switching between single-level and multi-level approach. It is noted that the proposed method can be readily extended for heterogeneous media with multiple phases, although only two-phase microstructures are involved in this work.

It should be noted that large numbers of 3D microstructure samples generated by the proposed reconstruction method are used to further explore the relationships between microstructural characteristics and transport properties in Chapter 3 and 7.

Chapter 5

Stochastic reconstruction of 3D pore microstructures from 2D cross-sectional images via supervised machine learning

5.1 Introduction

As mentioned in Chapter 1, various modern microscopy imaging techniques have been developed to obtain the 3D or 2D digital images of porous media at different scales. Micro-CT [388] and FIB-SEM [48] are often used to scan the physical samples of porous media and construct 3D digital visualizations at the microscopic level. However, high-quality 3D datasets of pore microstructures are still difficult and expensive to acquire, and 3D datasets are usually limited to small volume measurement which may bring uncertainty to statistical significance. In practice, high-fidelity 2D thin-sections are much easier and cheaper to obtain through SEM or other microscopy facilities, which are often the only supplied graphic information of porous rocks in engineering. The 3D representations of pore space are of great necessity to quantitatively characterize the porous media and to perform numerical simulations of transport processes [336, 48], thereby to establish reliable structure-property relationships. In this case, 3D microstructures need to be created by preserving the spatial dependences inferred from the available 2D thin-sections [120]. Stochastic microstructure reconstruction is an effective way to provide large numbers of samples of arbitrary sizes for microstructural analyses. The focus of this study is to develop a stochastic microstructure reconstruction method, where statistically equivalent 3D microstructures can be accurately and efficiently reconstructed by using the limited information of 2D thin-sections.

Generally, stochastic optimization-based method [495] and Gaussian random field transformation [363, 135] are the most popular stochastic reconstruction methods for heterogeneous media, and other methods such as Markov Chain Monte Carlo (MCMC) simulation [480], multiple-point statistics method [336], patch-based method [430], process-based method [341] and texture synthesis method [285] are also used frequently.

For the reconstruction of porous media, the connectivity and geometry of pore space must be accurately preserved, as well as the flow and transport properties. However, the stochastic optimization-based method and Gaussian random field-based method often fail to reproduce samples by maintaining the long-distance characteristics, because these methods only use the low-order statistical information that is inadequate to fully capture the microstructural complexity of porous media. Effective reconstruction methods allow accurate and efficient generation of realistic microstructures, and subsequent analyses can be performed on the reconstructed samples to predict macroscopic properties. Due to the above issues, various techniques have been developed into relatively new algorithms aiming to characterize and reconstruct the microstructures of heterogeneous porous media, among which machine learning has been proved to be a promising tool to solve these problems.

Recently, Bostanabad et al. (2016) [52, 53] proposed a novel method to statistically reconstruct 2D and 3D microstructures using supervised machine learning. This method follows a similar reconstruction concept as texture synthesis, but it adopts a supervised learning algorithm to determine the central pixel value from the phase values of its neighboring pixels. Because of the involvement of supervised learning, the method performs more efficiently and flexibly than conventional texture synthesis, and the stochastic signature and microstructural complexity of heterogeneous media can also be well captured by it.

Currently, Bostanabad's method [52, 53] can only reconstruct 2D microstructures from 2D reference samples, or generate 3D microstructures from 3D reference samples. Inspired by this method, a new multiple-voxel interaction scheme is further developed to reconstruct 3D heterogeneous pore microstructures from 2D thin-section images. It should be noted that this study is by no means a simple extension of Bostanabad's method. Due to information asymmetry between 2D and 3D microstructures, the supervised learning models trained by 2D thin section images cannot be directly used to generate 3D microstructures through a mapping mode. Therefore, the dimensional extension of 2D morphological information needs to be implemented in a proper way to generate 3D microstructures. Our strategy for the dimensional extension is as follow: (1) For isotropic reconstruction cases, 2D multiple-points statistics on one plane, which is captured by the fitted supervised machine learning models, are rotated 90° around each principle axis; (2) For anisotropic reconstruction cases, supervised machine learning models are trained based on 2D images on all three principal

planes to capture the anisotropic features. The probabilities of the voxel phase conditional to its neighboring voxels on each 2D plane are combined to provide an approximation of the conditional probability of the voxel phase in 3D space through an interacting scheme. Based on this conditional probability, statistically equivalent 3D microstructure samples can be reconstructed through sequential probability sampling.

5.2 Statistical microstructure characterization

As discussed in Chapter 4, the underlying full joint probability distribution function (PDF) $p(\boldsymbol{X})$ is an ideal characterization of the pore microstructure $\boldsymbol{X}$. Due to the extremely high dimension of $p(\boldsymbol{X})$, the Markov random filed (MRF) [469] assumption is adopted to reduce the computing complexity of $p(\boldsymbol{X})$, thereby to simplify the microstructure characterization issue. With the locality and stationarity assumptions of MRF, the core of microstructure characterization is to learn the relationship between a pixel and its neighboring pixels, aiming to capture local morphology patterns. The conditional probability distribution function (CPDF) $p(X_{ij}|\boldsymbol{N}_{ij})$ is used to take on the role of statistical microstructure characterization, where X_{ij} denotes a pixel in a image, and $\boldsymbol{N}_{ij}$ denotes the neighboring pixels surrounding it.

5.2.1 Data event collection

The L-shaped data template that contains 1 central pixel and its neighbors is adopted to collect data events from the 2D digital microstructure (also called training image), as shown in Figure 5.1. The process of data collection is illustrated in Figure 5.2. As described in Chapter 4, the CPDF $p(\boldsymbol{C}_{ij}|\boldsymbol{N}_{ij})$ associated with the data template can be estimated from the occurrence frequencies of data events collected from the training image:

$$p(X_{ij},\boldsymbol{N}_{ij}) = \frac{O(X_{ij},\boldsymbol{N}_{ij})}{\sum O(X_{ij},\boldsymbol{N}_{ij})} = \frac{O(X_{ij},\boldsymbol{N}_{ij})}{O_{\text{all}}} \tag{5.1}$$

$$p(\boldsymbol{N}_{ij}) = \frac{O(\boldsymbol{N}_{ij})}{\sum O(\boldsymbol{N}_{ij})} = \frac{O(\boldsymbol{N}_{ij})}{O_{\text{all}}} \tag{5.2}$$

$$p(X_{ij}|\boldsymbol{N}_{ij}) = \frac{p(X_{ij},\boldsymbol{N}_{ij})}{p(\boldsymbol{N}_{ij})} = \frac{O(X_{ij},\boldsymbol{N}_{ij})}{O(\boldsymbol{N}_{ij})} \tag{5.3}$$

where $O(X_{ij},\boldsymbol{N}_{ij})$ and $O(\boldsymbol{N}_{ij})$ are the occurrence numbers of data event $(X_{ij},\boldsymbol{N}_{ij})$ and $(\boldsymbol{N}_{ij})$ respectively, and O_{all} is the total number of all data events collected from the training image.

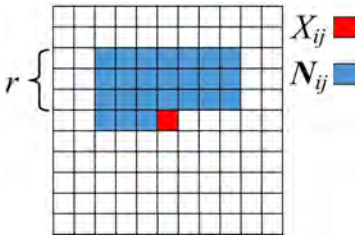


Fig. 5.1 Data template with neighborhood radius r=2 pixels

5.2.2 Implicit modeling of CPDF via supervised machine learning

Explicit construction of CPDF $p(X_{ij}|\mathbf{N}_{ij})$ should be avoided due to its high computational expense [469], especially for highly complicated microstructures, and the explicit probability sampling may also make the subsequent microstructure reconstruction procedure more complex. In this study, the CPDF $p(X_{ij}|\mathbf{N}_{ij})$ is estimated by using the collected data event to fit a supervised learning model, and the fitted model can be served as an implicit representation of $p(X_{ij}|\mathbf{N}_{ij})$ to reconstruct new statistically equivalent microstructure samples.

As the phase value of each pixel can only be 1 (pore) or 0 (solid) for the binary representation of a pore microstructure, so $p(X_{ij}|\mathbf{N}_{ij})$ can be considered as a class probability where X_{ij} is the classification category and $\mathbf{N}_{ij}$ is the classification features. Data events $(X_{ij}, \mathbf{N}_{ij})$ are paired observations, which are perfect training data to train a supervised machine learning model for classification [324]. Taking the collected data events $(X_{ij}, \mathbf{N}_{ij})$ as the training data, a classifier can be built through minimizing the probability of misclassification for any input features $\mathbf{N}_{ij}$. The class probability stored in the fitted model can be an accurate estimation of CPDF $p(X_{ij}|\mathbf{N}_{ij})$.

Commonly-used supervised machine learning algorithms [249], such as decision tree, artificial neural network, Bayesian network and support vector machine, are able to fulfill this classification purpose. Among them, the decision tree is found particularly suitable to capture the spatial dependencies of pore microstructure via implicitly modeling CPDF $p(X_{ij}|\mathbf{N}_{ij})$, based on which microstructure samples can quickly be reconstructed due to its simplicity.

A decision tree is a hierarchical structure consisting of nodes and directed edges, as shown in Figure 5.2e, where the root node is located at the top of the tree model, leaf nodes are located at the bottom, and internal nodes are between the root and leaves. Each non-leaf node contains attribute test conditions to separate input feature data that have different characteristics. Each leaf node is assigned a class label after computing all attributes, and the paths from the root to leaves represent the classification rules. Essentially, the decision tree recursively partitions the variable space into regions, and each region corresponding to

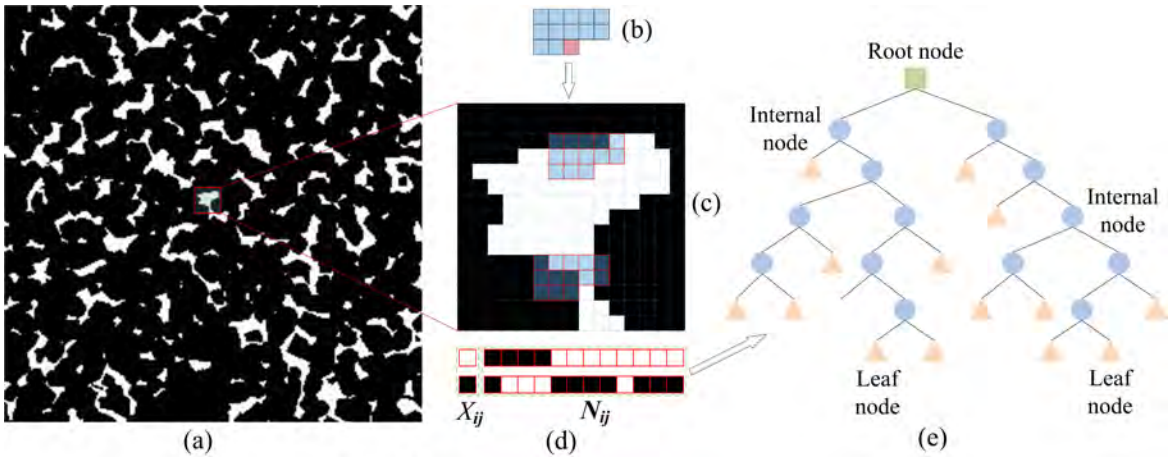


Fig. 5.2 Illustration of data event collection and decision tree training: (a) The training image $\boldsymbol{X}$ of a porous medium; (b) A data template with $r = 2$; (c) Partial enlarged detail of training image with the data template scanning over it; (d) Two examples of data events $(X_{ij}, \boldsymbol{N}_{ij})$; (e) Training of decision tree model

one leaf node, which is defined by the split sequences from the root to leaf. In this study, X_{ij} and $\boldsymbol{N}_{ij}$ play the roles of label and feature vector respectively to fit the decision tree model. It should be noted that the format of $\boldsymbol{N}_{ij}$ needs to be rearranged to an array for the model training, as shown in Figure 5.2d and e.

After the decision tree model is properly trained based on the train data, the model is able to predict the response for a new observation by bringing it to traverse from the root node to the correct leaf node. Each leaf node stores a predicted class probability associated with the response variable. This class probability is estimated as the ratio between the number of training observations with response value belonging to that class and the number of training observations with predictor values falling into that leaf node, which is equivalent to the conditional probability in Eq. (5.3). In our case, X_{ij} and $\boldsymbol{N}_{ij}$ of the training observations $(X_{ij}, \boldsymbol{N}_{ij})$ corresponding to pixel location (i, j) are the response variable and predictor variables respectively. Therefore, the fitted decision tree model can be considered as an implicit representation of CPDF $p(X_{ij}|\boldsymbol{N}_{ij})$ corresponding to the training image $\boldsymbol{X}$.

Decision tree algorithm, as well as other supervised machine learning algorithms, contain tuning parameters that can be adjusted to enhance its predictive performance on new observations. Cross-validation, as a widely-used nonparametric method, is adopted to optimize the tuning parameters for the decision tree model. Commonly, the decision tree model is firstly over-fitted to ensure the $p(X_{ij}|\boldsymbol{N}_{ij})$ of the training image $\boldsymbol{X}$ is fully preserved, and cross-validation is then applied to prune the over-fitted tree model and eliminate the redun-

dant noise, aiming to achieve an optimal model with the best approximate of $p(X_{ij}|\mathbf{N}_{ij})$. As to the basic theory of decision tree, one can refer to relevant literature [58, 249].

5.3 Stochastic microstructure reconstruction

After implicitly modeling of the CPDF $p(X_{ij}|\mathbf{N}_{ij})$ of the 2D target image via decision tree, the next key issue is how to generate (reconstruct) statistically equivalent 3D microstructures based on this characterization. As discussed in Section 5.1, Bostanabad's method has been well developed to reconstruct 2D (3D) microstructures from 2D (3D) training images [52, 53], and its reconstruction procedure can be interpreted as stochastic recombination of local morphology patterns captured by the pretrained decision tree model, where only the size expansion occurs during the microstructure reconstruction. While for the 3D microstructure reconstruction from 2D training images, both size expansion and dimensional extension of local morphology patterns are involved in the reconstruction process. Here, an interacting scheme is developed to reconstruct 3D microstructures from 2D thin-section images.

5.3.1 Basic idea of morphology integration

Generally, a 2D thin-section image cannot completely represent the corresponding 3D microstructure that contains a stack of 2D images, but it can be considered as a realization from a probability sampling of the morphological information of the 3D microstructure in 2D space [480]. If the 3D microstructure is stationary, it is reasonable to use a single 2D thin-section image of proper size as an effective representative for all layers in the same direction [430]. Therefore, it is possible and feasible to reconstruct 3D microstructures from 2D thin-sections.

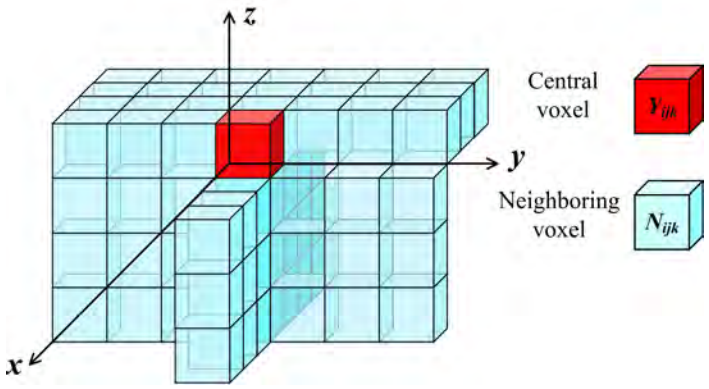


Fig. 5.3 A 3D data template involving 2D data templates on three perpendicular planes (r=3)

The MRF assumption is still applicable for 3D microstructures, which means the phase value of a central voxel Y_{ijk} at any site of the 3D image only depends on its neighboring voxels $\mathbf{N}_{ijk}$ within a large enough space. Here, we further simplify the MRF assumption, where the value of a central voxel only depends on its neighboring voxels on three perpendicular planes: *xy*-plane, *yz*-plane and *zx*-plane within a sufficiently long range. A 3D data template is thus created by combining three 2D data templates, as shown in Figure 5.3.

Data events are collected by using the 2D data template to scan 2D thin-section image(s), and then they are used to fit the decision tree model(s). For the 3D isotropic reconstruction, the statistics of local morphology pattern on the *xy*-plane captured by the decision tree model can be directly transformed to the *yz*-plane and *zx*-plane. While for the 3D anisotropic microstructure, 2D training images in the three perpendicular directions are all required, and three decision models are trained separately to capture the anisotropic properties. The point is how to integrate the 2D statistics on three separate planes into the 3D statistics of local morphology patterns. As illustrated in Figure 5.3, where 2D data templates on three principal planes are combined to be a 3D data template. And the relationship between the central voxel Y_{ijk} and its neighboring voxels $\mathbf{N}_{ijk}$ is mathematically expressed as follows:

$$\begin{aligned} p(Y_{ijk}|\mathbf{N}_{ijk}) &= p(Y_{ijk}|\mathbf{N}_{ijk}^{(zx)}, \mathbf{N}_{ijk}^{(xy)}, \mathbf{N}_{ijk}^{(yz)}) \\ &\approx \frac{1}{3} \cdot \left[p(Y_{ijk}|\mathbf{N}_{ijk}^{(xy)}) + p(Y_{ijk}|\mathbf{N}_{ijk}^{(yz)}) + p(Y_{ijk}|\mathbf{N}_{ijk}^{(zx)}) \right] \end{aligned} \tag{5.4}$$

where $\mathbf{N}_{ijk}^{(xy)}$, $\mathbf{N}_{ijk}^{(yz)}$ and $\mathbf{N}_{ijk}^{(zx)}$ are the neighboring voxels on *xy*-plane, *yz*-plane and *zx*-plane respectively.

As discussed in Section 2.3, the properly fitted decision tree model is a good implicit approximation of CPDF, therefore, $p(Y_{ijk}|\mathbf{N}_{ijk})$ can be calculated as:

$$p(Y_{ijk}|\mathbf{N}_{ijk}) \approx \frac{1}{3} \cdot (p_{ijk}^{(xy)} + p_{ijk}^{(yz)} + p_{ijk}^{(zx)}) \tag{5.5}$$

where $p_{ijk}^{(xy)}$, $p_{ijk}^{(yz)}$ and $p_{ijk}^{(zx)}$ are class probabilities stored in the fitted decision tree models. For 3D isotropic reconstruction, $p_{ijk}^{(xy)}$, $p_{ijk}^{(yz)}$ and $p_{ijk}^{(zx)}$ can be extracted from the same decision tree model. For 3D anisotropic reconstruction, $p_{ijk}^{(xy)}$, $p_{ijk}^{(yz)}$ and $p_{ijk}^{(zx)}$ should be extracted from three different decision tree models that are trained by using the data events collected from *xy*-plane, *yz*-plane and *zx*-plane respectively.

5.3.2 Microstructure reconstruction procedure

The 3D CPDF $p(Y_{ijk}|\boldsymbol{N}_{ijk})$ is estimated through a special interacting scheme, based on which statistically equivalent 3D microstructure samples can be generated. Suppose there is a underlying full joint PDF $p(\boldsymbol{Y})$ of the 3D reconstructed samples $\boldsymbol{Y}$, and $p(\boldsymbol{Y})$ can be expressed in factorial from as follows:

$$\begin{aligned} p(\boldsymbol{Y}) &= p(Y_{111})p(Y_{112}|Y_{111})p(Y_{113}|Y_{111},Y_{112})\cdots p(Y_{ijk}|Y_{111},Y_{112},\cdots Y_{ij(k-1)})\cdots \\ &= p(Y_{111})p(Y_{112}|\boldsymbol{Y}^{(<112)})p(Y_{113}|\boldsymbol{Y}^{(<113)})\cdots p(Y_{ijk}|\boldsymbol{Y}^{(<ijk)})\cdots \end{aligned} \tag{5.6}$$

where $\boldsymbol{Y}^{(<ijk)}$ represents the voxel sequence ahead of the voxel Y_{ijk}.

If all CPDFs $p(Y_{ijk}|\boldsymbol{Y}^{(<ijk)})$ are known to us, it is feasible to sequentially generate voxels by probability sampling from the full joint PDF $p(\boldsymbol{Y})$, there by to reconstruct 3D microstructure samples. To be specific, voxel Y_{111} is generated from $p(Y_{111}|\boldsymbol{Y}^{(<111)})$ firstly, with random initialization of $\boldsymbol{Y}^{(<111)}$, and then voxel Y_{112} is generated from $p(Y_{112}|\boldsymbol{Y}^{(<112)})$, conditional to Y_{111}. The remaining voxels can all be generated by using same procedure. When the voxel sequence is long enough, the effect of initialization can be ignored.

As shown in Figure 5.3, the neighboring voxels $\boldsymbol{N}_{ijk}$ in the 3D data template are all in front of the central voxels Y_{ijk} in a raster scan order. With the MRF assumption, the following relationship is established:

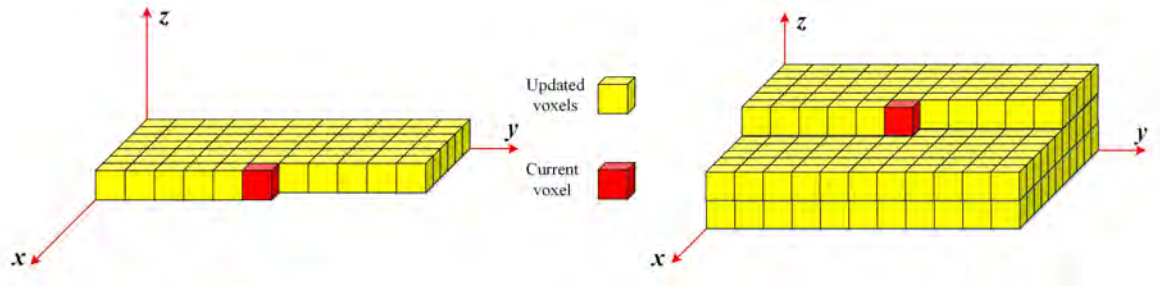
$$p(Y_{ijk}|\boldsymbol{Y}^{(<ijk)}) = p(Y_{ijk}|\boldsymbol{N}_{ijk}) \tag{5.7}$$

Substituting Eq. (5.7) to Eq. (5.6), the full joint PDF $p(\boldsymbol{Y})$ is rewritten as follows:

$$p(\boldsymbol{Y}) = p(Y_{111})p(Y_{112}|\boldsymbol{N}_{112})\cdots p(Y_{ijk}|\boldsymbol{N}_{ijk})\cdots \tag{5.8}$$

The CPDF $p(Y_{ijk}|\boldsymbol{N}_{ijk})$ can be estimated by using Eq. (5.5), where class probabilities are extracted from the pretrained decision tree models.

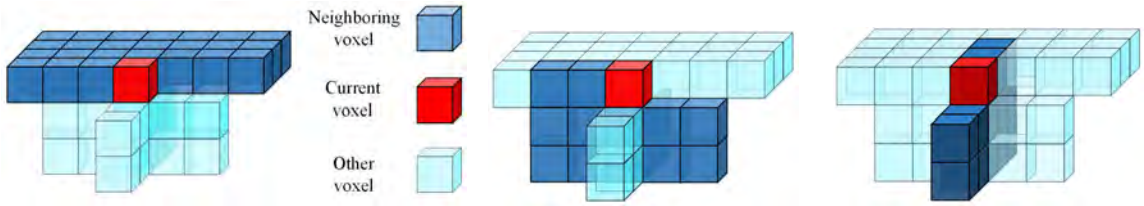
In this method, the 3D microstructure sample is reconstructed layer by layer, and voxels on each layer are generated one by one in a raster scan order, as illustrated in Figure 5.4. The phase value of each voxel Y_{ijk} is generated by probability sampling from the CPDF $p(Y_{ijk}|\boldsymbol{N}_{ijk})$ that is estimated through Eq. (5.5). The class probabilities $p_{ijk}^{(xy)}$, $p_{ijk}^{(yz)}$ and $p_{ijk}^{(zx)}$ are obtained by inputing the neighboring voxels $\boldsymbol{N}_{ijk}^{(xy)}$, $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$ into the fitted decision tree models $\boldsymbol{M}^{(xy)}$, $\boldsymbol{M}^{(yz)}$, and $\boldsymbol{M}^{(zx)}$ respectively.



(a) (b)

Fig. 5.4 Schematic illustration of a 3D microstructure reconstruction: (a) reconstruction of the first layer and (b) reconstruction of the third layer

It should be noted that the number of voxels in $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$ are less than that of $\boldsymbol{N}_{ijk}^{(xy)}$ for the reconstruction of the first few layers. For example, the number of voxels in $\boldsymbol{N}_{ij1}^{(yz)}$ and $\boldsymbol{N}_{ij1}^{(zx)}$ are 0 for the reconstruction of 1st layer, which means only the $\boldsymbol{N}_{ij1}^{(xy)}$ with neighboring radius r is required to reconstruct the 1st layer. The neighboring radii of $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$ gradually increase as more layers are generated. As shown in Figure 5.5, the neighboring radii of $\boldsymbol{N}_{ij3}^{(xy)}$ $\boldsymbol{N}_{ij3}^{(yz)}$ and $\boldsymbol{N}_{ij3}^{(zx)}$ are equal to 3, 2 and 2 voxels respectively, for the reconstruction of the 3rd layer.



(a) (b) (c)

Fig. 5.5 Neighboring voxels on (a) xy-plane with $r_{xy} = 3$ voxels, (b) on yz-plane with $r_{yz} = 2$ voxels and (c) zx-plane with $r_{zx} = 2$ voxels

During the reconstruction process of the first r layers, multiple decision tree models are required because of the variation of neighboring voxel numbers on yz-plane and zx-plane. After that, the neighboring radii of $\boldsymbol{N}_{ijk}^{(xy)}$ $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$ are fixed to be a specific value r. More details about boundary condition and determination of neighboring radius can be found in Chapter 4. The following pseudocode summarizes the procedure to reconstruct new 3D microstructure samples from 2D thin-section images, as described in Algorithm 3.

Algorithm 3: Stochastic reconstruction of 3D microstructures from 2D cross-sectional image(s) using supervised machine learning

Data: The radius r of neighboring voxel (pixel) area.
Data: The pretrained supervised learning models $\boldsymbol{M}$.
Data: A 3D lattice grid with size $l \times w \times h$.
Initialization: Assign binary white noise to the 3D lattice grid as the initial state $\boldsymbol{Y}_0$;
while $k \leq r$ **do**
  Select the fitted supervised machine learning models $\boldsymbol{M}_{(r)}^{(xy)}$, $\boldsymbol{M}_{(k-1)}^{(yz)}$, and $\boldsymbol{M}_{(k-1)}^{(zx)}$;
  Reconstruct the k^{th} layer $\boldsymbol{L}_k$ of $\boldsymbol{Y}_0$, based on $\boldsymbol{M}_{(r)}^{(xy)}$, $\boldsymbol{M}_{(k-1)}^{(yz)}$, and $\boldsymbol{M}_{(k-1)}^{(zx)}$;
  while $r+1 \leq i \leq l$ *and* $r+1 \leq j \leq w-r$ **do**
    Extract the neighboring voxels $\boldsymbol{N}_{ijk}$ from $\boldsymbol{Y}_0$, including $\boldsymbol{N}_{ijk}^{(xy)}$, $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$;
    Input $\boldsymbol{N}_{ijk}^{(xy)}$, $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$ to the $\boldsymbol{M}_{(r)}^{(xy)}$, $\boldsymbol{M}_{(k-1)}^{(yz)}$, and $\boldsymbol{M}_{(k-1)}^{(zx)}$ respectively, to obtain class probability $p_{ijk}^{(xy)}$, $p_{ijk}^{(yz)}$ and $p_{ijk}^{(zx)}$;
    Compute $p(Y_{ijk}|\boldsymbol{N}_{ijk})$ according to Eq. (5.5);
    Generate the phase value of voxel Y_{ijk} by probability sampling from $p(Y_{ijk}|\boldsymbol{N}_{ijk})$, and update the corresponding voxel in $\boldsymbol{Y}_0$.
  end
end
while $r+1 \leq k \leq h$ **do**
  Select the fitted supervised machine learning models $\boldsymbol{M}_{(r)}^{(xy)}$, $\boldsymbol{M}_{(r)}^{(yz)}$, and $\boldsymbol{M}_{(r)}^{(zx)}$;
  Reconstruct the k^{th} layer $\boldsymbol{L}_k$ of $\boldsymbol{Y}_0$, based on $\boldsymbol{M}_{(r)}^{(xy)}$, $\boldsymbol{M}_{(r)}^{(yz)}$, and $\boldsymbol{M}_{(r)}^{(zx)}$;
  while $r+1 \leq i \leq l$ *and* $r+1 \leq j \leq w-r$ **do**
    Extract the neighboring voxels $\boldsymbol{N}_{ijk}$ from $\boldsymbol{Y}_0$, including $\boldsymbol{N}_{ijk}^{(xy)}$, $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$;
    Input $\boldsymbol{N}_{ijk}^{(xy)}$, $\boldsymbol{N}_{ijk}^{(yz)}$ and $\boldsymbol{N}_{ijk}^{(zx)}$ to the $\boldsymbol{M}_{(r)}^{(xy)}$, $\boldsymbol{M}_{(r)}^{(yz)}$, and $\boldsymbol{M}_{(r)}^{(zx)}$ respectively, to obtain class probability $p_{ijk}^{(xy)}$, $p_{ijk}^{(yz)}$ and $p_{ijk}^{(zx)}$;
    Compute $p(Y_{ijk}|\boldsymbol{N}_{ijk})$ according to Eq. (5.5);
    Generate the phase value of voxel Y_{ijk} by probability sampling from $p(Y_{ijk}|\boldsymbol{N}_{ijk})$, and update the corresponding voxel in $\boldsymbol{Y}_0$.
  end
end
Reture: A statistically equivalent 3D reconstructed microstructure sample $\boldsymbol{Y}$

5.3.3 Noise reduction and image smoothing

As discussed above, 2D thin-section images cannot completely represent corresponding 3D microstructures, so the decision tree model fitted by 2D data cannot fully capture the 3D morphological information. Due to this incompleteness, the reconstructed 3D microstructures are inevitable mingled with noise [336], as shown in Figure 5.6a.

To improve the microstructure reconstruction quality, image noise-reduction processing can be carried out. Besides, non-percolating solid and pore components may exist in the raw stochastic realizations of porous microstructures. The non-percolating solid components are isolated islands suspended inside the pore space, which is impossible in the real porous media. Therefore, they are directly replaced by pore space, and the effect of this replacement has been proved to be negligible [277]. Okabe (2004) [334] adopted dilation-erosion processes to smooth the sold-pore boundaries, and additional operation was carried out to adjust the porosity of the reconstructed porous microstructure, through which noise is significantly reduced.

(a) (b)

Fig. 5.6 Noise reduction and image smoothing for 2D layers in the 3D reconstructed microstructure: (a) before image processing and (b) after image processing

Inspired by previous studies described above, the reconstructed microstructures in this study are processed through the following procedures:

(1) 2D image process after each layer is reconstructed:
   i. Remove the small isolated noise;
   ii. Apply the opening process (combination of dilation and erosion) to smooth pore-solid interfaces.

(2) 3D image process after the whole microstructure is reconstructed:
   i. Replace the non-percolating solid phase with pore space;
   ii. Adjust the pore-solid ratio to preserve the bulk porosity.

5.4 Results and comparison

Here, the proposed reconstruction method based on supervised learning (SL) is applied to two different random media with isotropic and anisotropic properties. To highlight the superiority of the proposed method, it is compared with other three commonly used reconstruction methods, including the simulated annealing (SA) method [495], Gaussian random field-based

(GRF) method [363, 135], and Markov Chain Monte Carlo (MCMC) method [480], to reconstruct 3D microstructures from 2D thin-section images. Generally, the effectiveness of microstructure reconstruction depends on how well the inherent stochasticity of the original microstructure is preserved. Therefore, a comprehensive comparison is conducted to evaluate the statistical equivalence between the reconstructed and original microstructures.

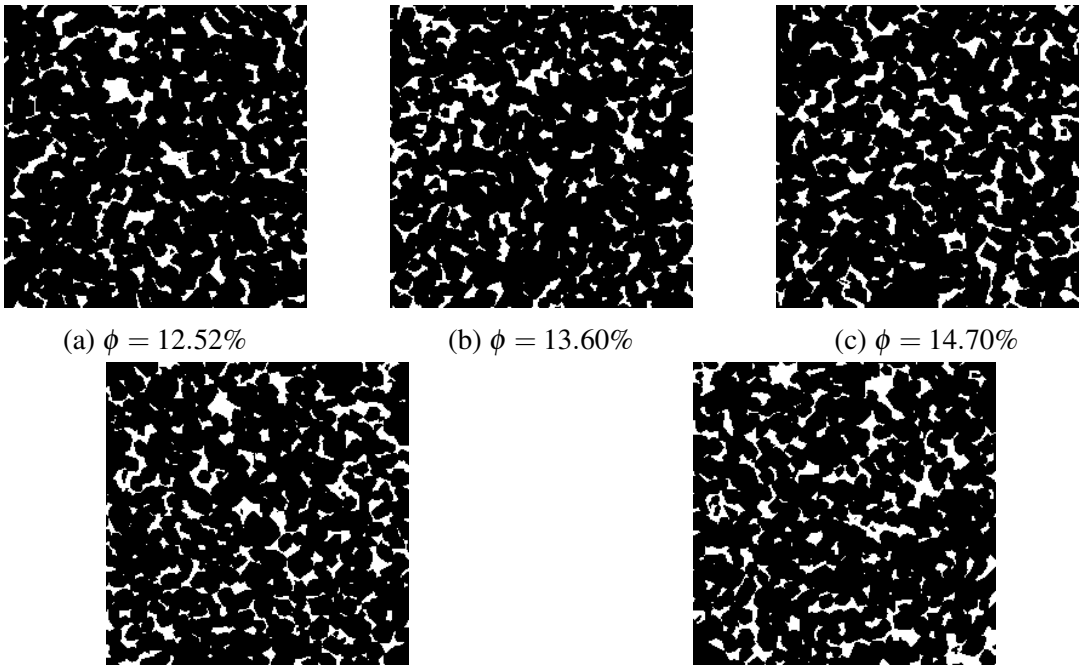
5.4.1 Example 1: Isotropic porous medium

Fontainebleau sandstone is chosen as the example of isotropic porous medium, which is often used as the reference standard for validation, due to its special properties [341, 262]. A 3D micro-CT scanning image of Fontainebleau sandstone discretized at a resolution of 14.65 μm is obtained from Institute for Computation Physics, University Stuttgart, Germany. A sub-model containing $150 \times 150 \times 150$ voxels is cut from it, which is used as the 3D reference model for microstructure reconstruction (as shown in Figure 5.9). Representative 2D slices are selected from the 3D reference image, and they are used as the 2D training images to fit the decision tree model. Considering the porosity variation of different layers in the 3D image (the bulk porosity is 14.7%), we select five representative 2D slices with different porosities as the training images, and their porosities are 12.52%, 13.60%, 14.70%, 15.80%, and 16.89% respectively, as shown in Figure 5.7.

Based on the selected 2D images, 3D microstructure samples are generated from the SA method, GRF method, MCMC method and the proposed SL method. From each reconstruction method, a group of 75 microstructure samples are reproduced, and statistical/morphological descriptors are extracted from these samples to access the reconstruction performance in terms of statistical equivalence and morphological similarity. For the proposed SL method, the size r of data template is 10 pixels/voxels, and it can be determined from the two-point correlation functions of the 2D slices, as shown in Figure 5.8. For the SA method, two-point correlation function [495] is selected as the optimization objective during the microstructure reconstruction process.

Visual comparison

The experimental micro-CT model and the representative microstructure samples reconstructed from different methods are illustrated in Figure 5.9, and the image size of them are $150 \times 150 \times 150$ voxels. Visual inspection implies that the morphology of the 3D reconstructed samples using different methods are significantly different from each other, and only the SL sample closely resembles the micro-CT model in terms of pore shape and pore distribution. For the SA samples, they are reconstructed by preserving the porosity and



(a) $\phi = 12.52\%$ (b) $\phi = 13.60\%$ (c) $\phi = 14.70\%$

(d) $\phi = 15.80\%$ (e) $\phi = 16.89\%$

Fig. 5.7 The selected slices with difference porosities ϕ (the pore space is shown in white, and solid matrix is shown in black)

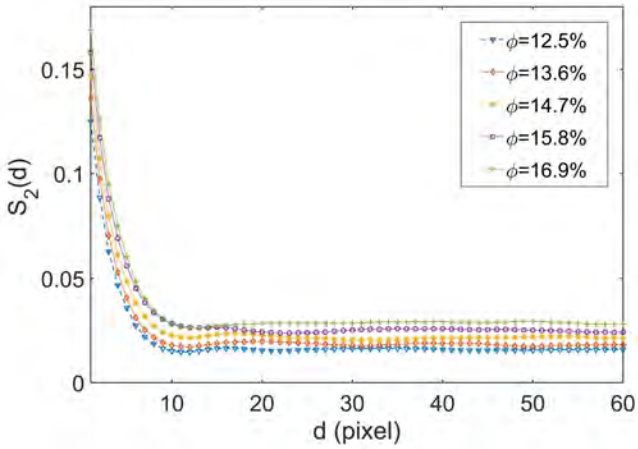
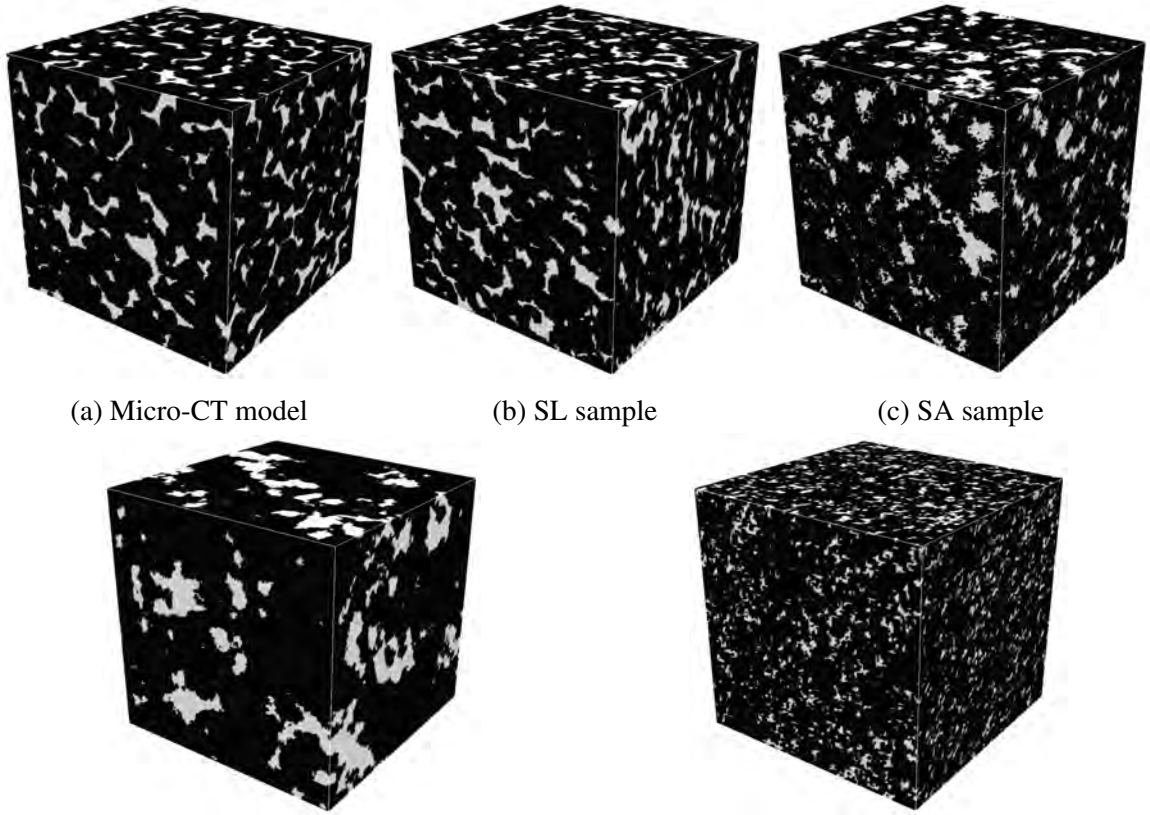


Fig. 5.8 Two-point correlation functions of the selected 2D slices

two-point correlation function of the 2D reference images. Visual observation of the SA model reveals that it contains more isolated small pores compared with the micro-CT model, and isolated solid components exist inside the pore space, which is impossible in the real porous rocks. The GRF samples preserve the porosity and autocorrelation function of the 2D slice selected from the micro-CT model, but the pore shapes are more round, which is distinctly different from the strip shape of pores within the micro-CT model. Visually, the

MCMC method provides the worst reconstruction result, because it rarely maintains the morphology of the micro-CT model.



(a) Micro-CT model (b) SL sample (c) SA sample

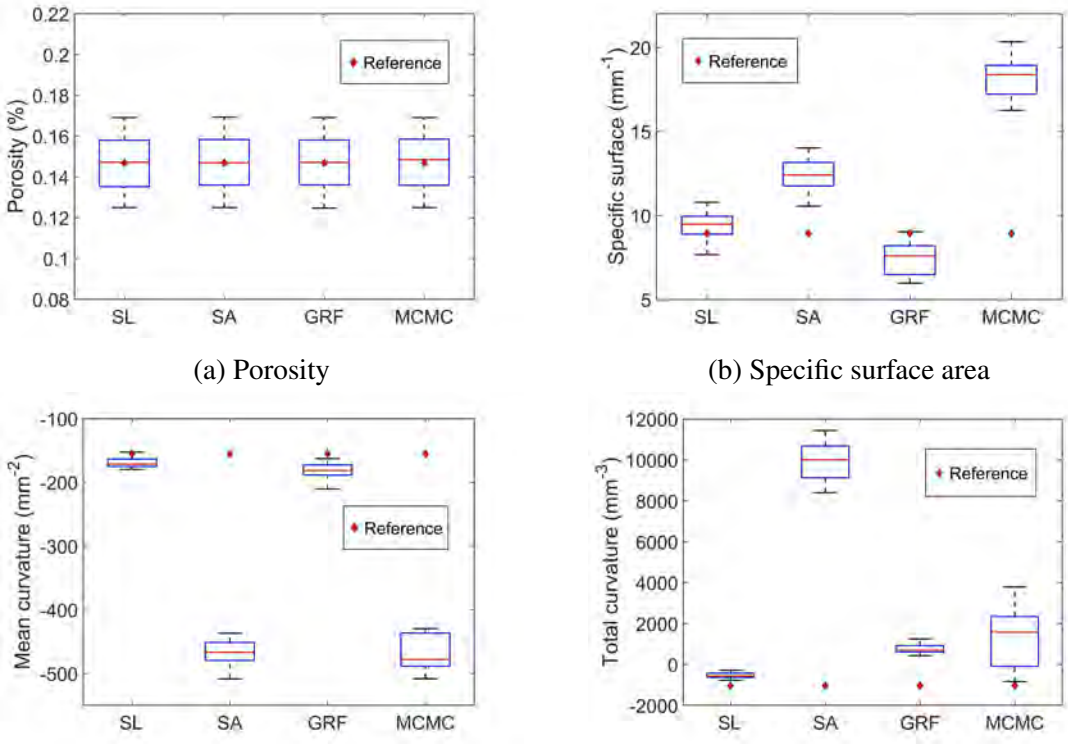
(d) GRF sample (e) MCMC sample

Fig. 5.9 Visual comparison between the 3D micro-CT model and the representative microstructure samples reconstructed from different methods

Comparison based on Minkowski functionals

The four boxplots in Figure 5.10 illustrate the comparisons of Minkowski functionals between the micro-CT model and the 3D reconstructed models. For each type of Minkowski functionals, the statistics are collected from the four groups of 3D samples reconstructed by different methods, based on which the corresponding boxplots are drawn.

As can be seen from Figure 5.10a, 3D microstructure samples reconstructed by all the four methods can well preserve the reference porosity. However, when it comes to the specific surface area, mean curvature, and total curvature, only the SL samples can accurately maintain the corresponding reference values (as shown in Figure 5.10b-c), which means the proposed SL method can capture more morphological information than the other three



(a) Porosity

(b) Specific surface area

(c) Mean curvature

(d) Total curvature

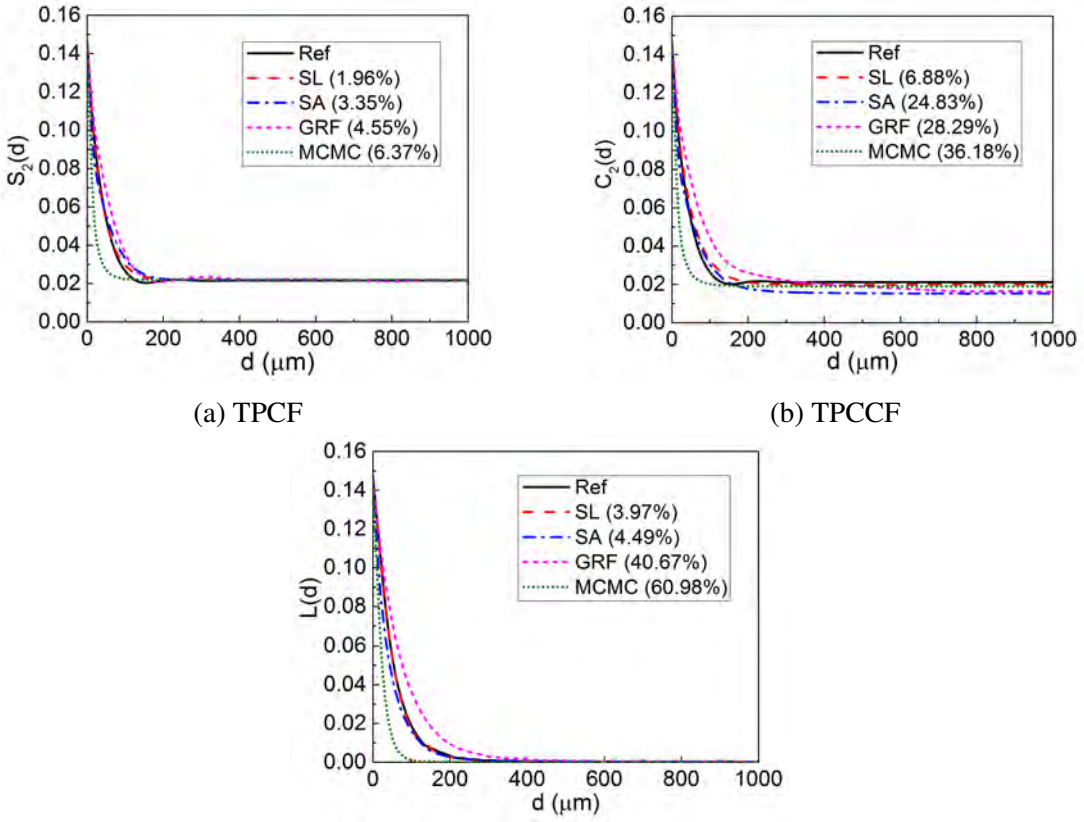
Fig. 5.10 Comparisons of Minkowski functionals between the micro-CT model (the reference values) and the reconstructed microstructure samples

methods. Compared to the SA method and the MCMC method, the GRF-based method exhibits better performance, because the ranges of specific surface area and mean curvature extracted from the GRF samples cover the reference values, and the total curvature range is not far away from its reference value. The SA samples differ greatly from the micro-CT model in terms of specific surface area, mean curvature and total curvature, but the performance of the SA method could be improved by involving more descriptors as the optimization objectives. As to the MCMC method [480], it is filed to capture high-order morphological information, because it only uses multiple-point statistics within a very short range to reconstruct microstructure samples.

Comparison based on statistical correlation functions

Three statistical descriptors are extracted from both the 3D micro-CT model and the reconstructed microstructure samples to check the statical equivalence, including two-point correlation function (TPCF), two-point cluster correlation function (TPCCF) and lineal path function (LPF). All the functions are calculated for the pore space. As shown in Figure

5.11a-c, the dashed curves correspond to the averages of statistical descriptors over 75 reconstructed microstructure samples, and the solid curve is the functions extracted from the micro-CT model. The dashed curves are compared with the solid ones in terms of the Euclidean norm, and the relative errors are provided in the brackets of each figure (the same below).



(a) TPCF

(b) TPCCF

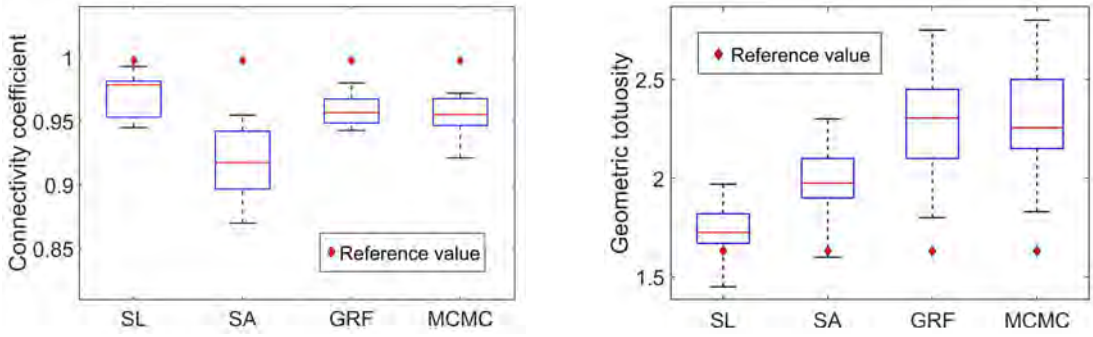
(c) LPF

Fig. 5.11 Comparisons of statistical correlation functions between the micro-CT model and the reconstructed microstructure samples

The proposed SL method possesses the highest level of accuracy in preserving statistical equivalence. The SA samples can well maintain the information of TPCF and LPF, but not TPCCF, which means the SA method is failed to completely capture the long-distance connectivity of the pore network. As to the GRF method and MCMC method, they exhibit poor performance in capturing the information of TPCCF and LPF, which means they can only generate microstructure samples by preserving low-order statistics.

Comparison based on connectivity coefficient and geometric tortuosity

Connectivity coefficient [481] is defined as the proportion of the connected pore volume to the bulk pore volume, which can be used to represent the connectivity characteristic of porous media. As can be seen from Figure 5.12a, the red diamond mark denotes the connectivity coefficient of the micro-CT model, and its value is 0.99, which means almost the entire pore space is connected and very few isolated pores exist. The boxplots describe the distributions of connectivity coefficients of the reconstructed microstructure samples, and the average values of them are all below the reference value, which means the reconstructed samples contain more isolated pores than the micro-CT model. By comparison, the proposed SL method shows the best performance in maintaining the connectivity coefficient.



(a) Connectivity coefficient (b) Geometric tortuosity

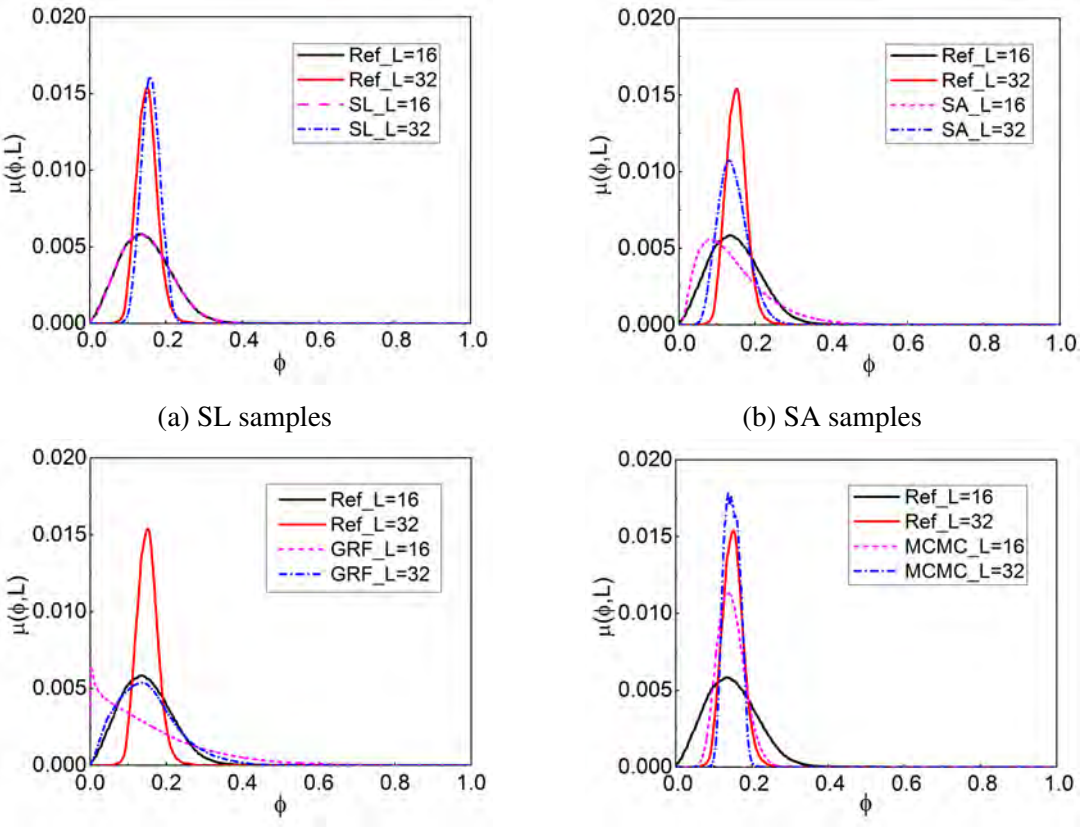
Fig. 5.12 Comparisons of (a) connectivity coefficient and (b) geometric tortuosity extracted from the 3D micro-CT model and the reconstructed microstructure samples

Geometrical tortuosity [90, 153] is defined as the ratio of the effective length of flow paths to the straight-line distance in the macroscopic flow direction, to describe the degree of streamlines deviating from the straight line, whose value is usually greater than 1. As shown in Figure 5.12b, the average tortuosity values of the SL samples, SA samples, GRF samples and MCMC samples are 1.725, 1.98 2.325 and 2.289 respectively, and they are all lager than the reference value (1.631) from the micro-CT model, but the SL samples are still the closest to the reference model in terms of the geometrical tortuosity.

Comparison based on local porosity characterization

Figure 5.13 demonstrates the comparisons between the micro-CT model and the reconstructed microstructure samples in terms of local porosity distribution (LPD) $\mu(\phi, L)$ on two different scales. The dashed curves are average results of LPDs over 75 reconstructed samples, and the solid curves correspond to the LPDs extracted from the micro-CT model. By comparison,

the SL samples are the best match to the reference model in terms of LPDs on two different scales, and the SA samples show small differences from the reference model. As to the GRF samples and the MCMC samples, the LPD curves of them significantly deviate from the reference ones.



(a) SL samples

(b) SA samples

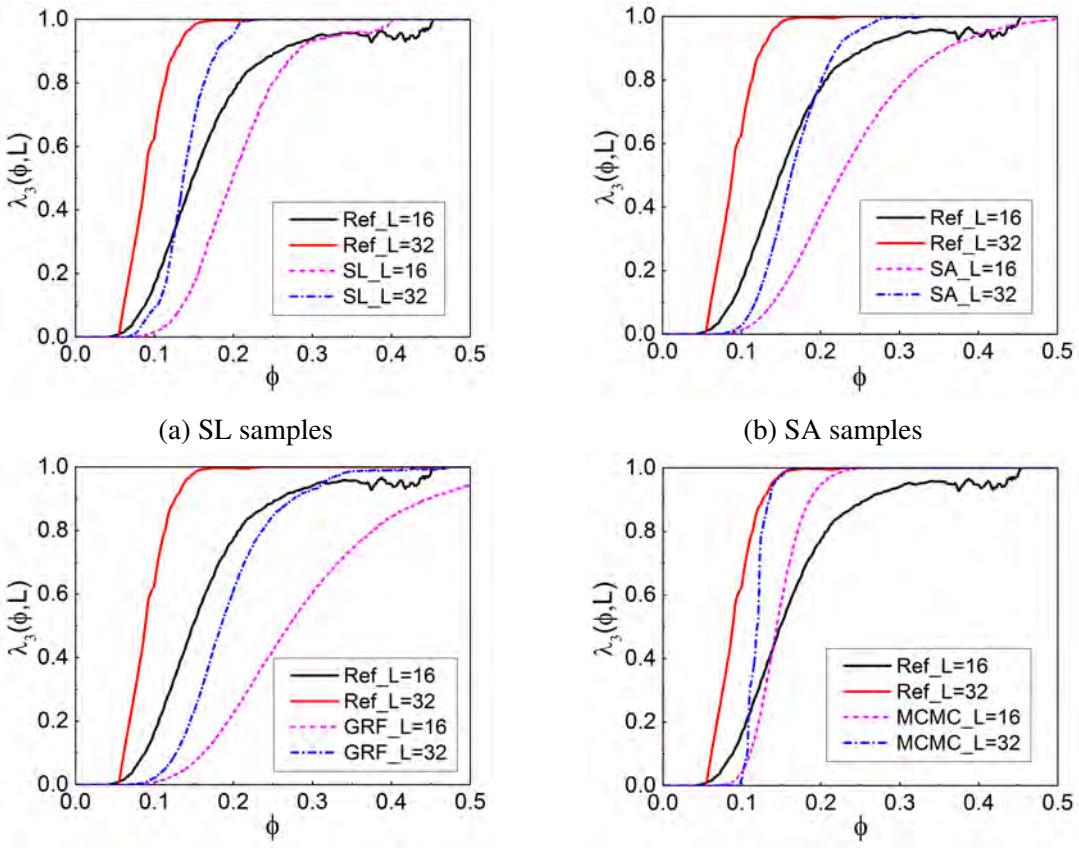
(c) GRF samples

(d) MCMC samples

Fig. 5.13 Comparisons of LPDs extracted from the 3D micro-CT model and the reconstructed microstructure samples on two different scales (L=16 and L=32 voxels)

In Figure 5.14, local percolation probabilities (LPP) $\lambda_3(\phi, L)$ are extracted from both the micro-CT model and the reconstructed microstructure samples on two different scales. Local percolation probability represents the fluctuations of pore connectivity, but none of the four reconstruction methods is able to completely capture this connectivity characteristics. The primary reason is that 2D thin-sections limitedly reflect the topological information of the 3D pore network, so the 3D microstructure samples reconstructed from 2D thin-sections cannot fully preserve the connectivity characteristics. Even so, the proposed SL method and the MCMC method show comparatively good performance in maintaining pore connectivity.

Figure 5.15 provides the total fractions of percolating cells (TFPC) of both the micro-CT model and the reconstructed microstructure samples. TFPC directly reflects the connectivity



(a) SL samples (b) SA samples

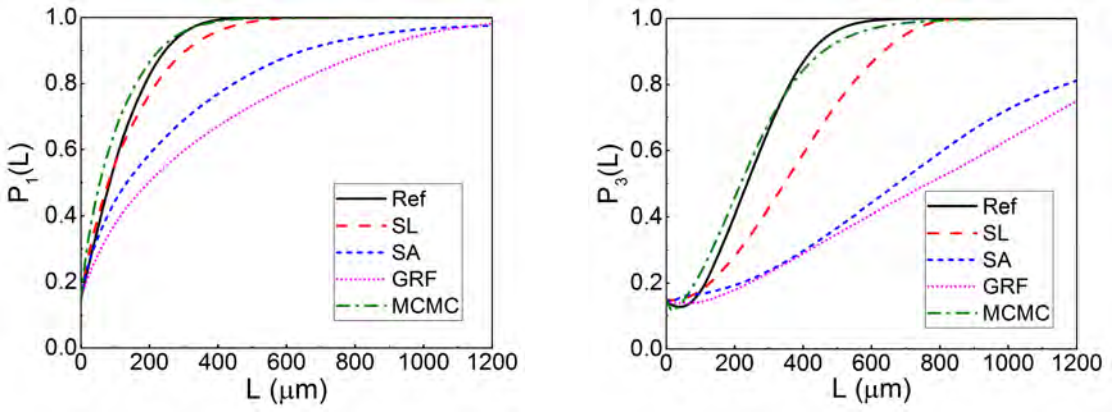
(c) GRF samples (d) MCMC samples

Fig. 5.14 Comparisons of LPPs extracted from the 3D micro-CT model and the reconstructed microstructure samples on two different scales (L=16 and L=32 voxels)

characteristics of porous media, but only the TFPC curves of MCMC samples well match the reference curves. As to the SL samples, their TFPC curves reasonably agree with the reference TFPC curve in one direction, but not the TFPC curve in three directions. The SA samples and GRF samples show a huge discrepancy with the reference model in aspects of TFPC, which reveals that the connectivity characteristics of them are much worse than that of the reference model.

5.4.2 Example 2: Anisotropic microstructure

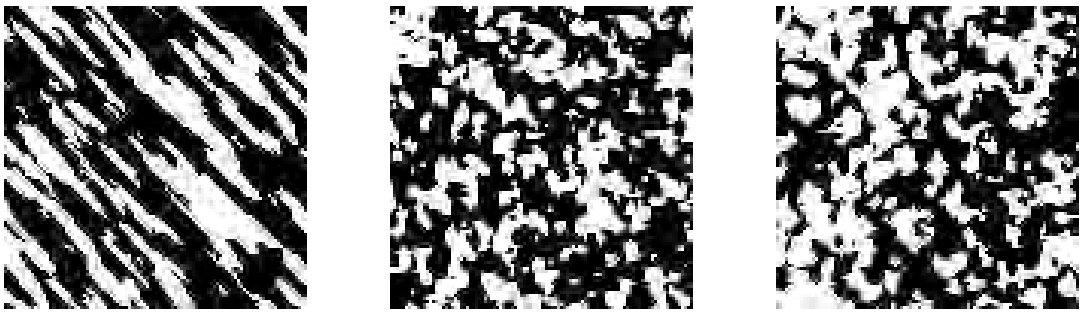
Here, the proposed reconstruction method based on supervised learning (SL) is used to reconstruct a 3D anisotropic microstructure from 2D thin-sections. The 3D reference model with $100 \times 100 \times 100$ voxels is shown in Figure 5.17a, and representative 2D slices in three perpendicular planes are selected from the 3D reference model to play as the 2D training images for 3D reconstruction, as shown in Figure 5.16. Based on the set of 2D training



(a) Percolation in one principal direction (b) Percolation in three principal directions

Fig. 5.15 Comparisons of TFPCs extracted from the 3D micro-CT model and the reconstructed microstructure samples

images, the proposed SL method and the SA method, MCMC method are used to reconstruct 3D microstructures (The GRF method is not used here, because it is generally for isotropic microstructure reconstruction). A group of 50 samples are reconstructed by using each reconstruction method.



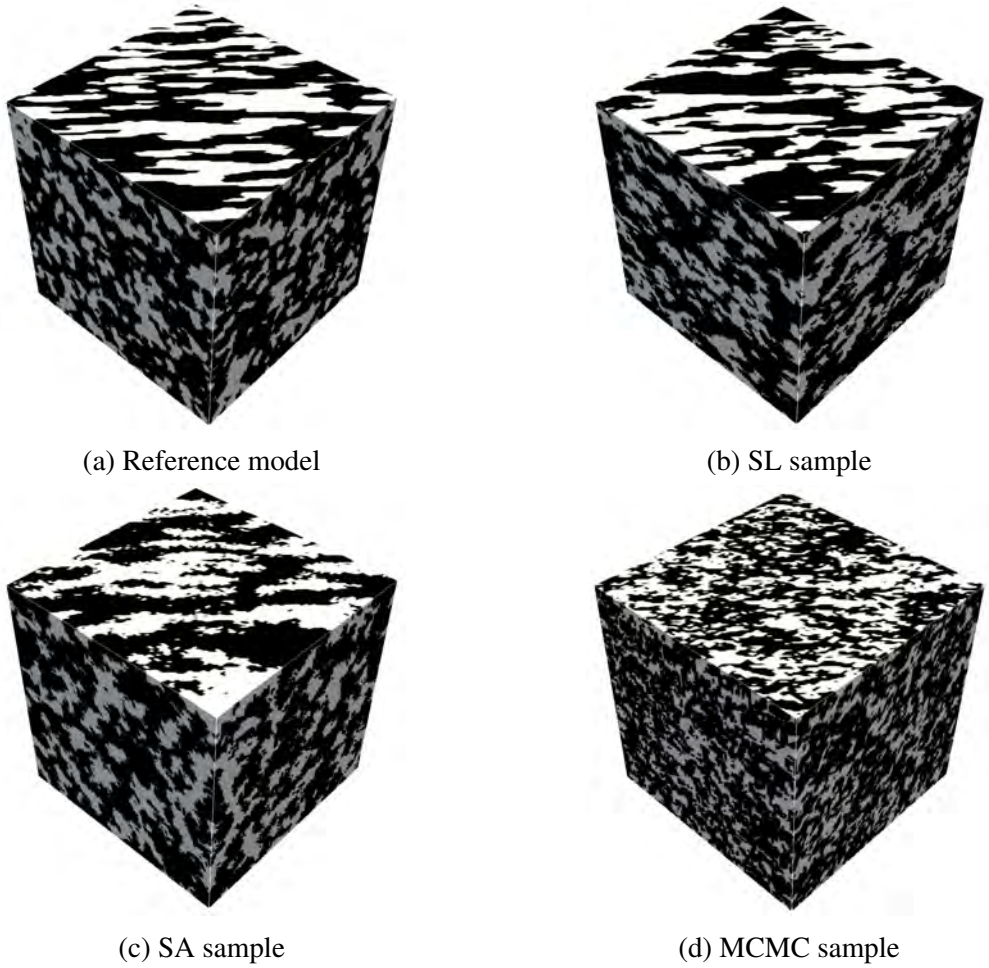
(a) *xy*-plane (b) *yz*-plane (c) *zx*-plane

Fig. 5.16 A set of 2D training images are selected from the 3D reference image in three perpendicular planes

For the proposed SL method, the data template size is determined from the TPCF of each 2D training image. According to the TPCF curves in Figure 5.19a-c, the radii of data template on three perpendicular planes are determined to be $r^{(xy)} = 16$, $r^{(yz)} = 5$ and $r^{(zx)} = 5$ pixels, respectively.

Visual comparison

As shown in Figure 5.17, 3D visualizations (image size: $100 \times 100 \times 100$ voxels) of the reference model and representative reconstructed microstructure samples are provided. Visual inspection implies that only the SL samples and SA samples can well preserve the morphology of the reference model, especially the anisotropic feature. The SL samples also closely resemble the reference model in terms of grain shapes and particle sizes, while the SA samples contain more isolated grains of small size. It should be noted that the SA samples are reconstructed by setting the TPCFs in different directions as the multiple optimization objectives, which makes the reconstruction more accurate but computationally intensive. As to the MCMC samples, they are failed to preserve the basic morphology of the reference model.



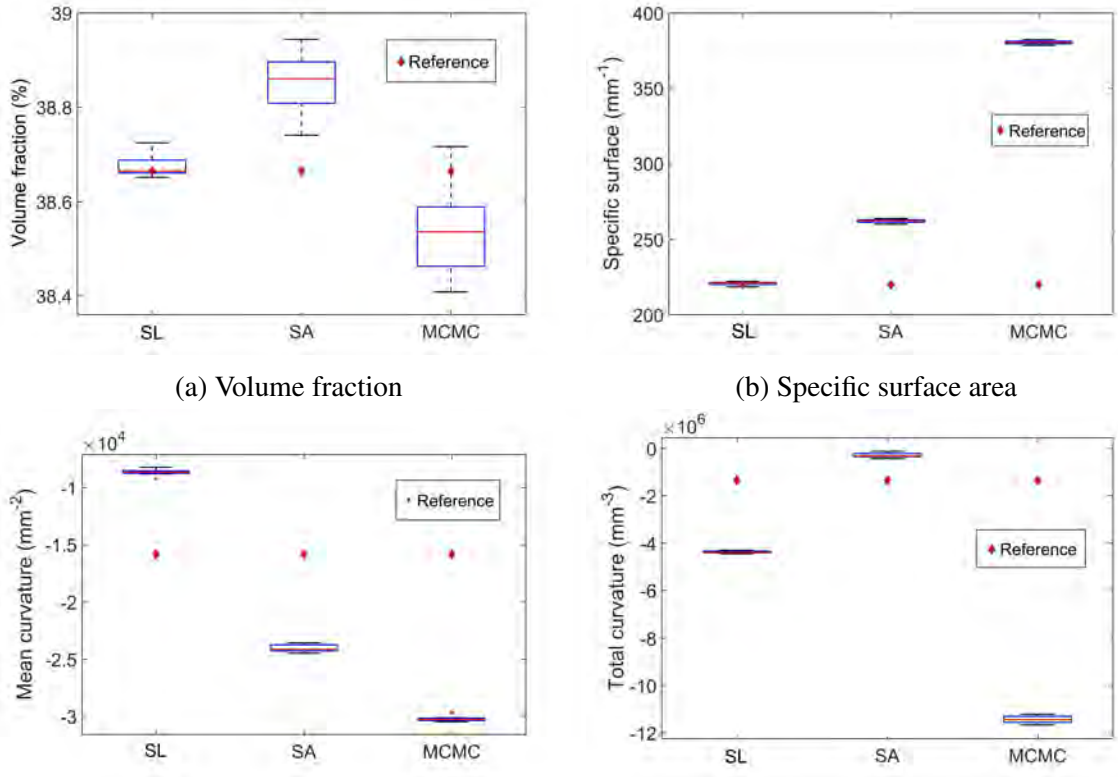
(a) Reference model (b) SL sample

(c) SA sample (d) MCMC sample

Fig. 5.17 Visual comparison between the 3D reference model and the representative microstructure samples reconstructed from different methods

Comparison based on Minkowski functionals

The four boxplots in Figure 5.18 illustrate the comparisons of Minkowski functionals between the 3D reference model and reconstructed microstructure samples. For each type of Minkowski functionals, the statistics collected from 50 reconstructed microstructure samples, based on which the boxplots are drawn. All of the 3D reconstructed models agree well with the reference model in terms of volume fraction. However, only the SL samples can maintain the specific surface area of the reference model. As to the mean curvature and total curvature, all of the reconstructed samples are failed to preserve them.



(a) Volume fraction

(b) Specific surface area

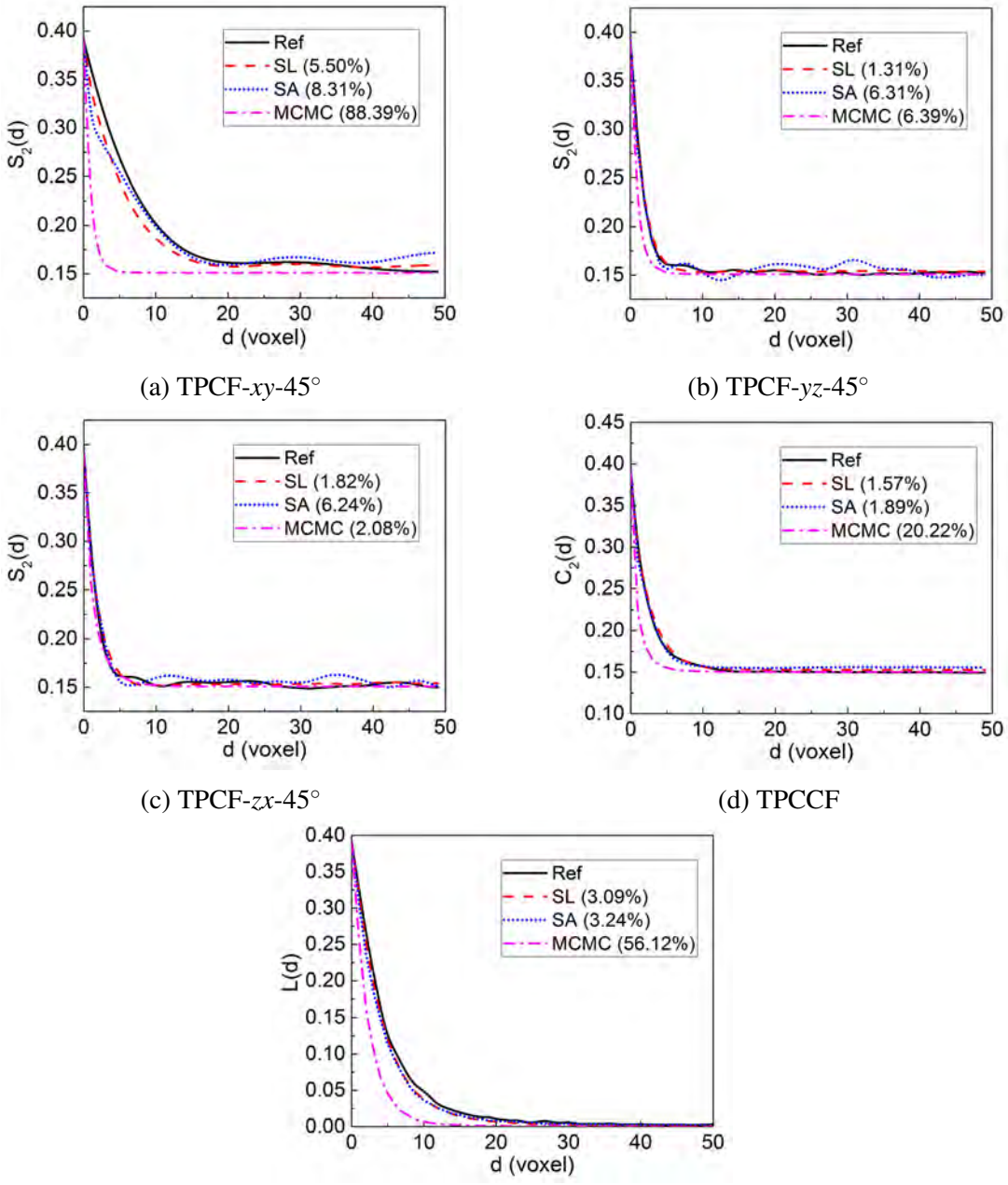
(c) Mean curvature

(d) Total curvature

Fig. 5.18 Comparisons of Minkowski functionals between the reference model and the reconstructed microstructure samples

Comparison based on statistical correlation functions

To inspect how well the reconstructed samples preserve the anisotropic feature of the 3D reference model, we compare the TPCFs in three different directions. Figure 5.19a-c display the TPCFs in the directions of 45-degree on *xy*-plane, *yz*-plane, and *zx*-plane, respectively.



(a) TPCF-*xy*-45°

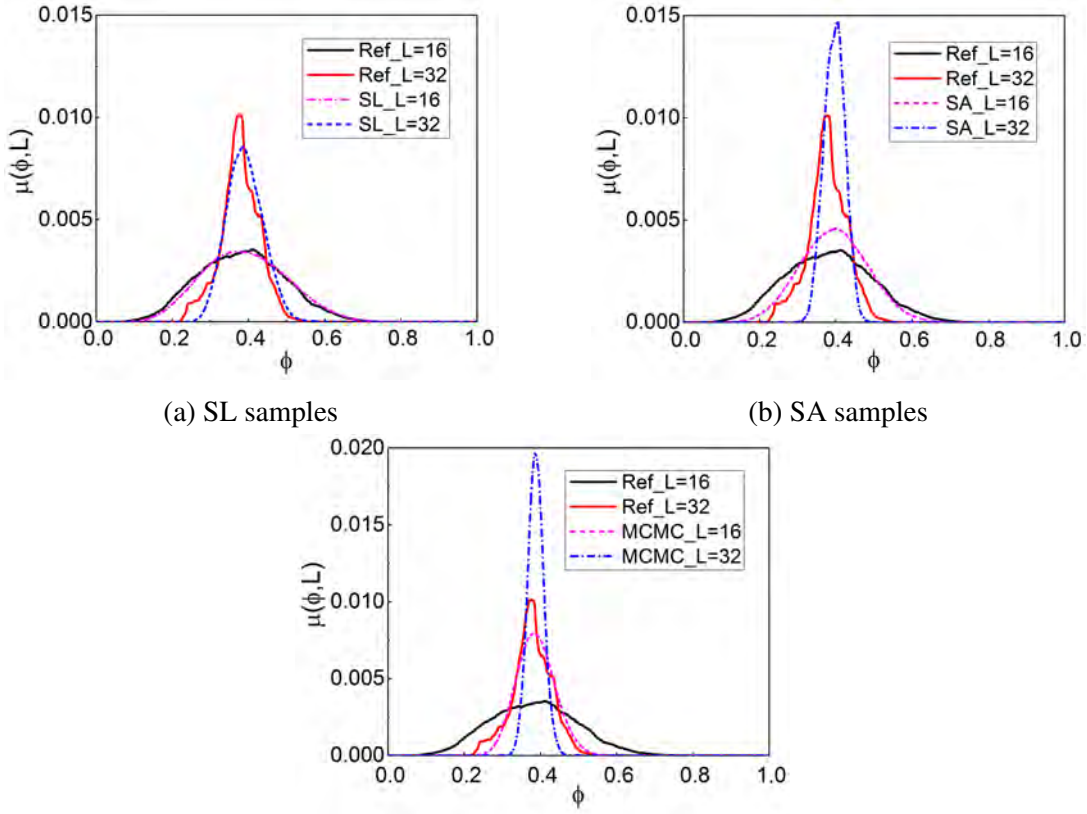
(b) TPCF-*yz*-45°

(c) TPCF-*zx*-45°

(d) TPCCF

(e) LPF

Fig. 5.19 Comparisons of statistical correlation functions between the 3D reference model and the reconstructed microstructure samples



(a) SL samples

(b) SA samples

(c) MCMC samples

Fig. 5.20 Comparisons of LPDs extracted from the 3D reference model and the reconstructed microstructure samples on two different scales (L=16 and L=32 voxels)

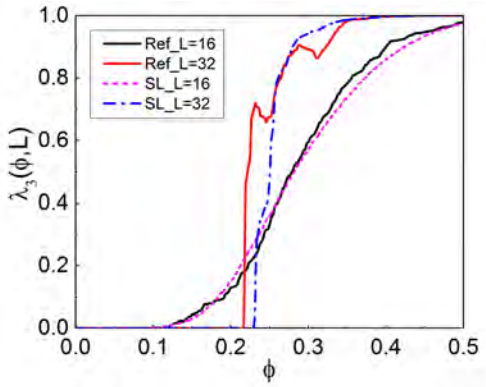
The TPCF curves extracted from the SL samples and SA samples well match the reference curves, while the accuracy of SL samples is much higher than that of SA samples. As to MCMC samples, the relative error of the TPCF in the direction of 45-degree on xy-plane is up to 88.39%, which means MCMC samples are failed to maintain the anisotropic feature of the reference model. Besides, TPCCF and LPF are also extracted from the reference model and reconstructed samples. The SL samples show the best performance in preserving the TPCCF and LPF of the reference model, and the SA samples can also reasonably preserve TPCCF and LPF. As to the MCMC samples, they are failed to capture both TPCCF and LPF, especially in the short distance.

Comparison based on local volume fraction characterization

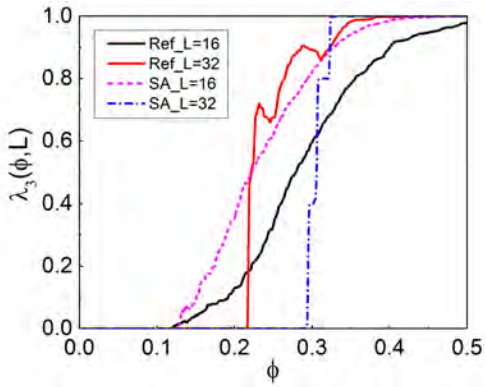
Figure 5.20 demonstrates the comparisons of local volume fraction distributions (LVFDs) $\mu(\phi, L)$ between the 3D reference model and reconstructed microstructure samples, and the LVFDs are calculated on two different scales, with the side lengths of measurement cells

L=16 and L=32 voxels respectively. Apparently, the SL samples show the best performance in preserving the information of LVFD, but the SA method and MCMC method are failed to capture LVFD.

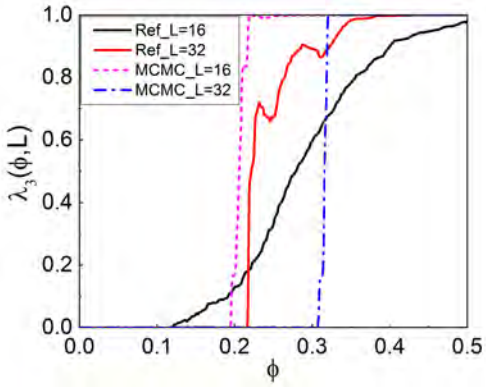
In Figure 5.21, the local percolation probabilities (LPPs) $\lambda_3(\phi, L)$ extracted from the 3D reference model and reconstructed microstructure samples are compared on two different scales ($L = 16$ and $L = 32$ voxels). Since LPP represents the connectivity fluctuation of a specific material phase, so the connectivity characteristics of SL samples are closest to that of the reference model. As the SA and MCMC samples, their connectivity characteristics are significantly different from that of the reference model.



(a) SL samples



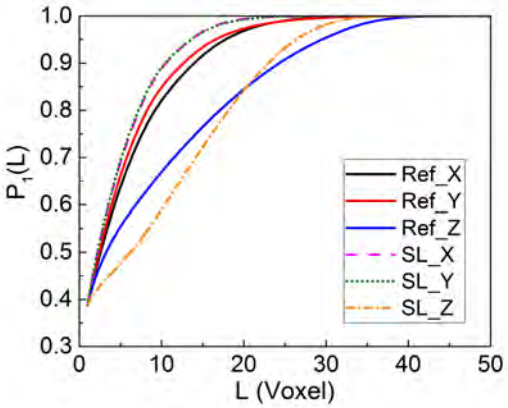
(b) SA samples



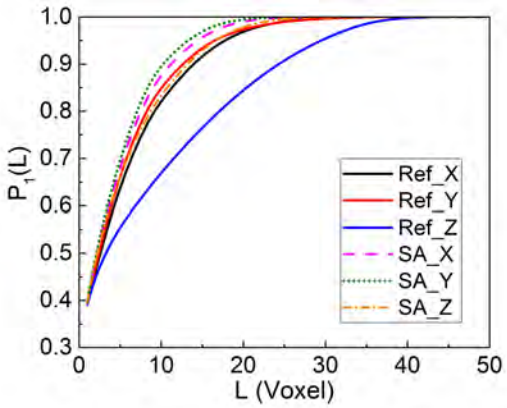
(c) MCMC samples

Fig. 5.21 Comparisons of LPPs extracted from the 3D reference model and the reconstructed microstructure samples on two different scales (L=16 and L=32 voxels)

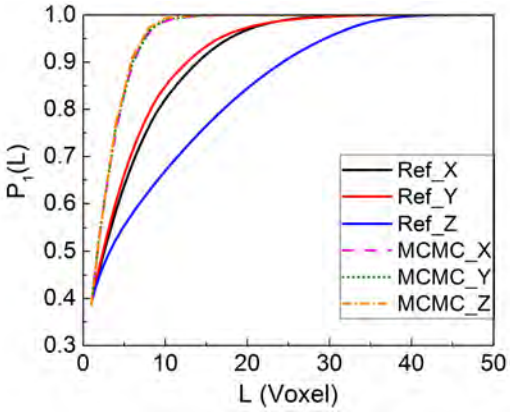
Figure 5.22 provides the total fractions of percolating cells (TFPCs) $P_1(L)$ extracted from the 3D reference model and reconstructed microstructure samples. Due to the anisotropic property of the reference model, the TFPC in z-directions is remarkably different from that in x-direction and y-direction. It is clear that only the SL samples reasonably preserve the anisotropic connectivity of the reference model.



(a) SL samples



(b) SA samples

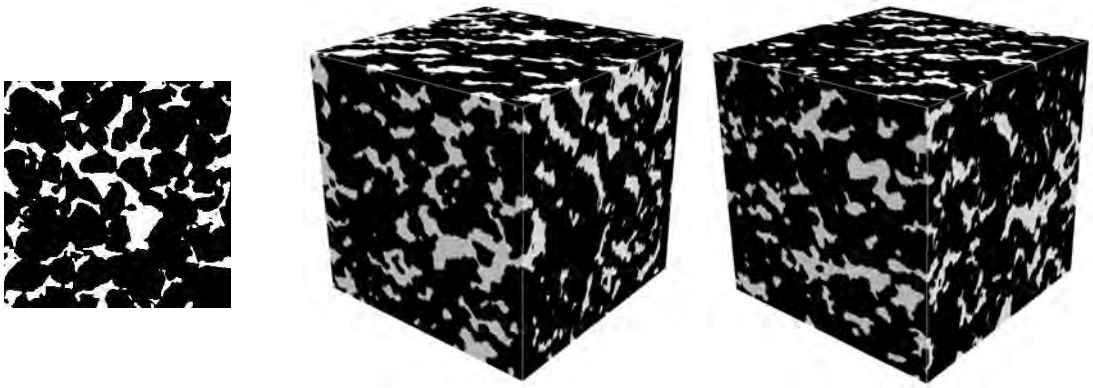


(c) MCMC samples

Fig. 5.22 Comparisons of TFPCs extracted from the 3D reference model and the reconstructed microstructure samples

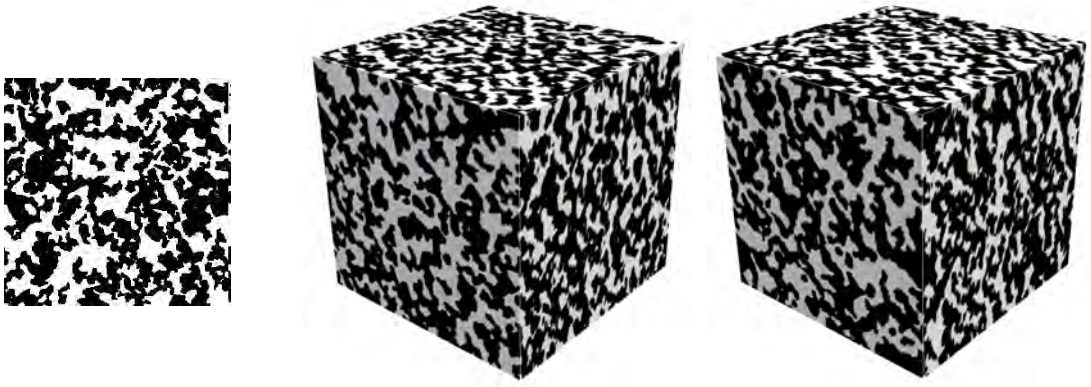
5.4.3 Other examples

To demonstrate the powerful capacity of the proposed microstructure reconstruction method, it is applied to reconstruct various 3D microstructures with distinct morphological features from corresponding 2D thin-sections, as illustrated in Figure 5.23, 5.24, 5.25, 5.27 and 5.26.



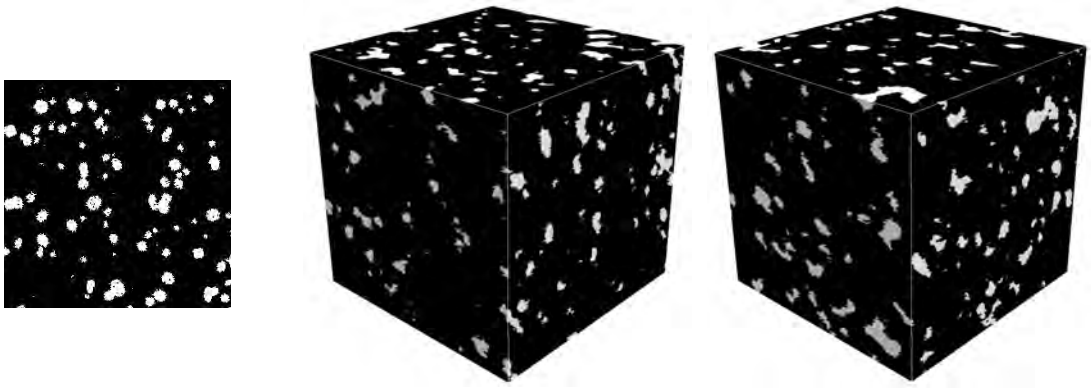
(a) 2D training image (b) 3D reconstructed sample I (c) 3D reconstructed sample II

Fig. 5.23 Berea sandstone



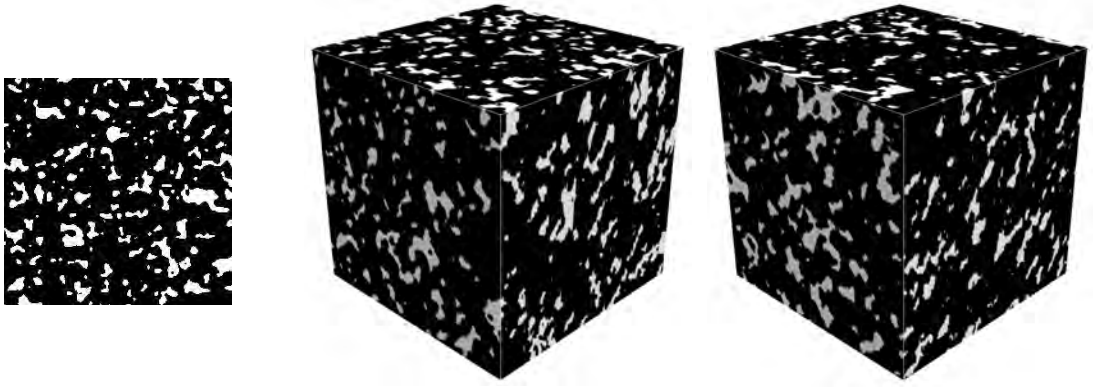
(a) 2D training image (b) 3D reconstructed sample I (c) 3D reconstructed sample II

Fig. 5.24 Synthetic silica



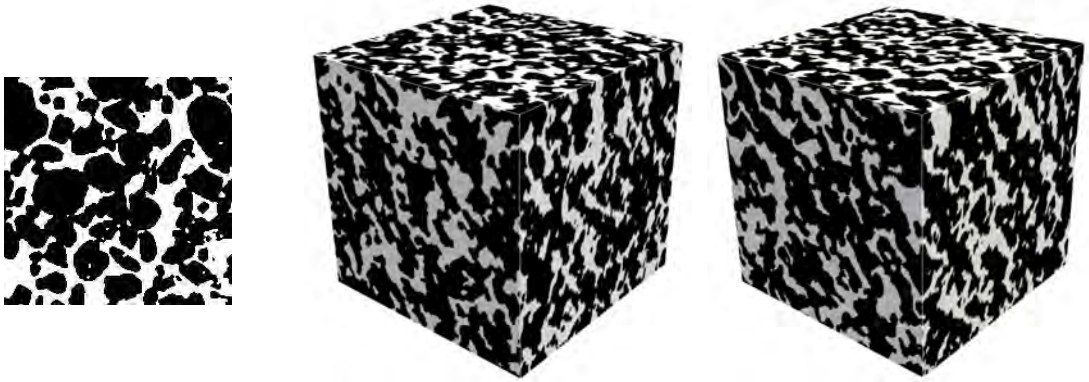
(a) 2D training image (b) 3D reconstructed sample I (c) 3D reconstructed sample II

Fig. 5.25 A nanocomposite with silica embedded in rubber matrix



(a) 2D training image (b) 3D reconstructed sample I (c) 3D reconstructed sample II

Fig. 5.26 Sandstone IV



(a) 2D training image (b) 3D reconstructed sample I (c) 3D reconstructed sample II

Fig. 5.27 Sandstone VIII

5.5 Discussion and conclusions

The main contribution of this chapter is to develop a novel method to statistically reconstruct 3D heterogeneous porous microstructures from 2D thin-section images via supervised machine learning. The method starts by collecting data events from the 2D training images, and then the collected data events are fed to the supervised learning models to learn the local morphological patterns of the microstructure. Based on the statistical information stored in the fitted supervised learning models, 3D microstructures can be constructed through a specific procedure to integrate morphological information.

Essentially, the fitted supervised machine learning model is an implicit representation of multiple-points statistics associated with the data template. For isotropic microstructures, they can be statistically characterized by learning the multiple-points statistics from the representative 2D planes in one direction. While for anisotropic microstructures, the multiple-

points statistics of 2D planes in the three principal directions are required to capture the anisotropy.

The class probabilities stored in the fitted supervised learning model(s) are good approximations of the conditional probabilities of each 2D plane, and 2D statistical information on three perpendicular planes can be integrated to provide an estimation of the conditional probability of voxel in 3D space. In this way, the morphological information hidden in the 2D training images is dug out by supervised machine learning and then extended to 3D spatial morphology through information integration, based on which statistically equivalent 3D microstructures can be generated through probability sampling. The 3D microstructure is constructed through a layer-by-layer scheme, where voxels on each layer are generated one by one. The operation of information integration here is to maintain the continuity and connectivity between adjacent layers.

Table 5.1 Computational cost of microstructure reconstruction using different methods

Reconstruction method	Average CPU time cost(s)	
	Example 1: Isotropic microstructure (size: $150 \times 150 \times 150$ voxels)	Example 2: Anisotropic microstructure (size: $100 \times 100 \times 100$ voxels)
The SL method (the proposed method)	3415	1080
The SA method	498841	87940
The GRF method	85	-
The MCMC method	1475	345

The proposed method is compared with three commonly used reconstruction methods, to reconstruct both 3D isotropic and anisotropic microstructures by using the limited information from 2D thin-sections, and their performances are assessed in terms of reconstruction accuracy and efficiency. The four reconstruction methods are performed on an ordinary computer (Intel Xeon CPU 3.50GHz), and the computational costs are listed in Table 5.1. Generally, the proposed method can well preserve the spatial dependencies of the reference microstructure, and to generate more realistic and accurate microstructures with higher efficiency, compared with the other three classical methods.

However, the 3D reconstructed microstructure samples are not completely consistent with the 3D reference models, especially in terms of long-range connectivity. To summarize, the reasons are threefold: Firstly, 2D thin-section images are not the complete representatives of the corresponding 3D microstructure, based on which the 3D reconstructed sample cannot fully preserve the morphological information. Proper selection of 2D slices as the training images can improve the qualities of 3D reconstruction samples. Secondly, with the MRF assumption, local morphology patterns are collected from the training images to fit the supervised learning models, so the reconstructed samples may lose some global

morphological features such as long-range connectivity. Finally, the image post-processing of reconstructed samples may bring a certain degree of distortion, but currently, it is an indispensable process for microstructure reconstruction based on incomplete information. More reasonable and flexible image post-processing techniques can be used to improve reconstruction results.

Another important advantage of the proposed method is that various open-source machine learning toolboxes are available, which makes the proposed method easy to be implemented based on them. Moreover, the microstructure examples involved in this study are with two phases, but the proposed method can be easily extended to reconstruct multi-phase microstructures.

Chapter 6

Resolution effect: An error correction model for permeability of porous media estimated from lattice Boltzmann method

6.1 Introduction

Digital rock physics (DRP) studies, in a non-destructive manner, how pore-scale processes govern macroscopic rock properties, such as transport, elastic, thermal and electrical properties [11]. DRP uses microscopic imaging techniques such as X-ray micro-computed tomography (micro-CT) to obtain 3D digital rock images with varying resolutions and sizes. The 3D representations of complex pore geometry can be obtained by separating the pore space from mineral matrix phases through the segmentation processing on raw images, and then physical simulations can be performed on the pore-scale microstructures to estimate the effective properties of rocks (e.g. permeability, elastic moduli, effective diffusivity, thermal conductivity and formation resistivity factor). In this study, we focus on intrinsic permeability, one of the most important transport properties of porous media that measure the ability of a structure to transmit fluid. The intrinsic permeability of porous media is of fundamental importance for many geological and geotechnical applications, such as oil and gas production [154], geothermal energy extraction, contaminant hydrogeology and carbon geological sequestration.

When core samples are unavailable for experimental measurements, microscopy imaging and numerical simulation can be performed on cutting samples to determine the permeability

of a porous rock. Lattice Boltzmann method (LBM) is one of the most widely-used numerical solvers for pore-scale simulation of fluid flow through porous media [417], because it can be directly applied to the digital pore-scale microstructures without simplification or meshing and it has good flexibility to work with complex flow geometries. The LBM is developed from the microscopic kinetic equation (Boltzmann equation) for gas, where the mathematical model is expressed as a particle distribution equation, and such macroscopic properties as density, pressure and velocity can be calculated from the particle distribution function [252]. Compared to laboratory measurements, the LBM has the advantage to directly link the flow field to the micro-scale geometry of porous media. The method is well-suited for evaluating single-phase flow properties of porous rocks, and can deliver reliable results for permeability, dispersion coefficients and effective reaction rates, etc. [48].

Out of several variations, the most popular LBM formulation is arguably the single relaxation Bhatnagar-Gross-Krook (BGK) collision model [78] with a standard bounce-back (SBB) scheme to represent the no-slip condition on solid walls [252]. The success of LBM is, to a great degree, founded on the BGK model to solve the incompressible Navier-Stocks equations. Although mass and momentum conserving, the single relaxation time LBM (SRT-LBM) using the SBB scheme has a major drawback that the exact physical position of fluid-solid boundary depends on the viscosity value due to the associated numerical approximation [177]. The viscosity-dependent boundary condition becomes more problematic in pore-scale simulations of fluid flow through porous media, because it makes the predicted macroscopic properties (e.g. permeability) dependent on the simulation parameter viscosity. In LBM, the kinematic viscosity ν is controlled by the relaxation time τ through the equation: $\nu = c_s^2 \Delta t(\tau - 0.5)$. However, there is no common rule to determine the optimal value of τ when performing SRT-LBM on a specific porous medium. It is generally recommended to choose a relaxation time around 1.0 as a trade-off between the accuracy, stability and efficiency of LBM simulation [252]. In addition, a higher resolution grid is often suggested to reduce the dependency of computed permeability on viscosity.

Accurate pore-scale analysis using LBM requires high-resolution and large-size representative volume elements (RVEs), which can take a long time to simulate, even on high-performance computing platforms [401]. The general resolutions of micro-CT images used for DRP analysis range from a few microns to dozens of microns, depending on the rock types and the physical phenomena [48]. Current micro-CT facilities are capable to provide digital images with more than $2000 \times 2000 \times 2000$ voxels at the submicron resolution, and the LBM solvers have to be massively parallelized to deal with such large models. In practice, the raw micro-CT images are often rescaled to a lower resolution, in order to reduce the computational cost of LBM simulation to an acceptable level. The resolution reduction is

particularly common when using GPU platforms, as a finer RVE model often exceeds the memory capacity of a single graphics card [220]. Therefore, the choice of image resolution is usually a compromise between accuracy requirement and computation cost.

The resolution effect on permeability evaluation is well known to the DRP community, and it occurs not only with LBM simulation but also with other numerical methods. As early as 2000, Edie et al. [126] studied the effect of image resolution on fluid flow simulation in porous media. They took a 3D digital model of a Fontainebleau sandstone sample constructed from micro-CT at a resolution of 7.5 μm, and coarsened it to five different resolution levels ranging from 15 μm to 45 μm. Lattice gas simulation was performed on these models, and the permeability results were found to vary by a factor of 5 over the resolution range. Later, Pan et al. (2001) [346] studied the effect of voxel size on permeability calculation by performing LBM on random sphere packings, and they found the permeability result decreased when the image became finer before a threshold level of resolution was reached. In 2004, Keehm and Mukerji [234] researched the influence of grid resolution on LBM permeability evaluation, and they found the permeability was consistently overestimated with the increase of grid space. Many other researchers also reported similar observations that the permeability evaluation from LBM generally increases as the image resolution reduces (voxel size increases) [335, 51, 161, 385]. These previous studies only focused on the resolution effect coming from the variations of pore geometry and pore structure due to resolution reduction, without considering the influence of relaxation time. This limits the practical applicability of these studies. As shown in Figure 6.7b, the computed permeability with the relaxation time of 0.625 decreases as the image resolution reduces, which is opposite to some of the former observations in the literature.

Recently, Peng et al. (2012) [351] investigated the resolution effect on pore structure characterization of a Berea sandstone, and they found relatively low-resolution images (voxel size up to 5.29 μm) can be used to calculate permeability, because small pores do not influence permeability significantly. However, for micro-CT images with even lower resolutions, they tended to overestimate the pore size and thereby the permeability [353]. Alyafei et al. (2015) [10] used Pore Network Modeling (PNM) to study the resolution effect on permeability prediction for sandstones and limestones, but they did not observe a clear trend in permeability prediction with respect to the voxel size. More recently, Shah et al. (2016) [401] studied the influence of voxel resolution on petrophysical properties of sandstones and carbonates by using both LBM and PNM, and they showed that petrophysical properties are consistent over the resolution range after using a specific numerical coarsening scheme to rescale images. But Latief et al. (2017) [263] observed that the image resolution reduction performed by hardware-based and software-based approaches both have significant

influences on permeability results calculated by LBM. Liu et al. (2018) [284] adopted the Finite Volume Method (FVM) to estimate the permeabilities from micro-CT images with different resolutions, and they found the permeability predictions are stable when the image resolution is higher than a specific threshold resolution. Similarly, Guan et al. (2018) [164] also used FVM to calculate the permeability of a Bentheimer sandstone, and they found permeability changes rapidly when the voxel size is larger than a certain value, which implies an optimal resolution may exist to balance efficiency and accuracy of permeability evaluation. In addition, Borujeni et al. (2013) [51] found grid coarsening on the Finite Element Method (FEM) mesh leads to a reduction of computed permeability, and they attributed it to the tightening of pore space rather than a numerical error.

In summary, the resolution effect on permeability evaluation for porous media is widely recognized for virtually all numerical methods including LBM, PNM, FVM and FEM, but they are rarely addressed in a quantitative manner in previous studies, making the resolution reduction questionable in practical use. In this study, we focus on the resolution effect on permeability evaluation from the LBM simulation, aiming to establish an error correction model that can effectively reduce, even eliminate the permeability error caused by image resolution reduction.

The rest of this chapter is organized as follows. Section 6.2 briefly summarizes the basic theory of LBM, our implementation of LBM to evaluate permeabilities of porous media, and the verification of our implementation is also presented in this section. In Section 6.3, we investigate the resolution effect on permeability calculation by performing LBM on a group of digital rock samples, and the geometry error and the boundary error are identified as the primary error sources due to the resolution effect. In Section 6.4, two models are established to capture the geometry error and the boundary error respectively, which can be used to compensate for the resolution effect on LBM permeability evaluation. Section 6.5 examines the performance of the proposed error correction model on different types of porous media. Finally, conclusions and discussions of future research are summarized in Section 6.6.

6.2 Permeability evaluation using LBM

6.2.1 Basic theory of LBM

LBM [252] models the fluid flow through a time-dependent distribution of fluid particles propagating on a regular lattice. In DRP research, pore voxels of segmented rock images serve as the regular lattice for LBM to simulate the fluid flow, and each lattice node is located in the center of the corresponding pore voxel. The numerical grid of lattice Boltzmann

simulation is coincident with the voxel grid in this study. The particle distribution function $f_i(\mathbf{x},t)$ represents the probability of finding a fluid particle with the lattice velocity $\mathbf{c}_i$ in the location $\mathbf{x}$ and at the time t. Beginning with an initial state, $f_i(\mathbf{x},t)$ moves from one lattice node to its neighboring nodes at each time step, and evolves itself locally subject to both mass and momentum conservation.

In this work, we adopt the conventional LBM scheme with D3Q19 lattice arrangement (as shown in Figure 6.1) and the Bhatnagar-Gross-Krook (BGK) collision operator [78]. The evolution of $f_i(\mathbf{x},t)$ along the direction of $\mathbf{c}_i$ from the time t to $t+\Delta t$ can be expressed as:

$$f_i(\mathbf{x}+\mathbf{c}_i\Delta t, t+\Delta t) - f_i(\mathbf{x},t) = -\frac{1}{\tau}\left[f_i(\mathbf{x},t) - f_i^{\mathrm{eq}}(\mathbf{x},t)\right] \tag{6.1}$$

where τ is the single-relaxation time, $f_i^{\mathrm{eq}}(\mathbf{x},t)$ is the equilibrium distribution function, and the subscript i indicates the direction of lattice velocity around the lattice node. The relaxation time τ is a function of kinematic lattice viscosity ν of simulated fluid, i.e. $\nu = c_s^2\Delta t(\tau - 0.5)$, where c_s is the lattice speed of sound and it is assigned with the dimensionless value of $\sqrt{1/3}$.

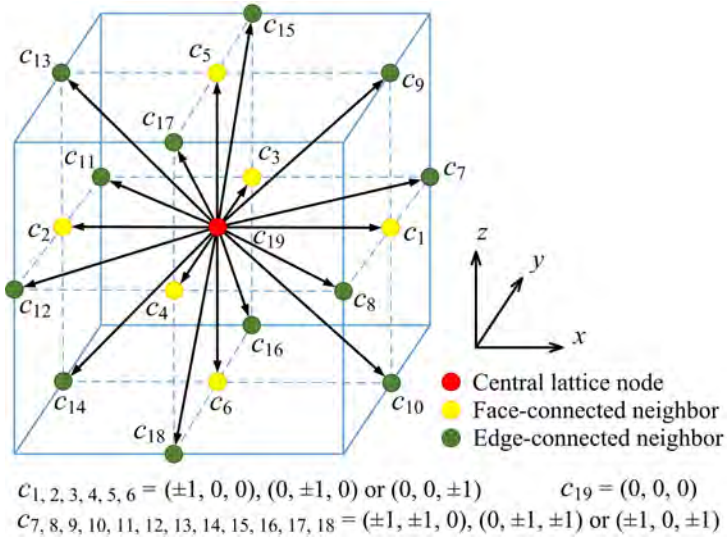


Fig. 6.1 Lattice structure of D3Q19 model with 19 velocity directions per lattice node

The equilibrium distribution function $f_i^{\mathrm{eq}}(\mathbf{x},t)$ corresponds to an ideal state where the particle distributions tend to a specific macroscopic state, in order to recover the macroscopic Navier-stokes equations. For the D3Q19 lattice arrangement with BGK collision operator, $f_i^{\mathrm{eq}}(\mathbf{x},t)$ is given by the following equation [78]:

$$f_i^{\mathrm{eq}}(\mathbf{x},t) = w_i\rho\left[1 + 3(\mathbf{c}_i\cdot\mathbf{u}) + \frac{9(\mathbf{c}_i\cdot\mathbf{u})^2}{2} - \frac{3(\mathbf{u}\cdot\mathbf{u})}{2}\right] \tag{6.2}$$

where w_i is the weight factor of D3Q19 lattice structure, ρ is the fluid density, and $\mathbf{u}$ is the macroscopic fluid velocity. For the D3Q19 lattice model, the weight factors equal to $\frac{12}{36}$, $\frac{2}{36}$ and $\frac{1}{36}$ for the velocity directions of the central lattice node, face-connected neighbors and edge-connected neighbors, respectively.

At the end of each time step, the macroscopic properties of fluid flow, including density ρ and velocity $\mathbf{u}$, can be approximated from $f_i(\mathbf{x},t)$ through the following equations, and these macroscopic properties will be used for the LBM computation at the next time step.

$$\rho = \sum_{i=1}^{n} f_i \tag{6.3}$$

$$\mathbf{u} = \frac{\sum_{i=1}^{n} f_i \mathbf{c}_i}{\rho} \tag{6.4}$$

where n is the number of lattice directions (n=19 in D3Q19 lattice structure).

6.2.2 Implementation of lattice Boltzmann simulation

To evaluate the intrinsic permeability of a porous medium, we consider a cubic digital rock sample as shown in Figure 6.2, and use SRT-LBM to simulate a single-phase fluid flow (incompressible Newtonian fluid) with low Reynolds number ($Re \ll 1$) passing through it. Under a constant pressure difference between the inlet and outlet faces, the fluid is driven through the pore microstructure. When the fluid flow reaches a steady state, it can be described by Darcy's law, and the intrinsic permeability κ of the porous medium is quantified by the following equation:

$$\kappa = -\frac{\mu}{\nabla p} \langle \mathbf{u} \rangle \tag{6.5}$$

where ∇p is the pressure gradient along the direction of macroscopic fluid flow, μ is the dynamic viscosity of fluid, and $\langle \mathbf{u} \rangle$ denotes the average fluid velocity across the simulation domain.

Since the initial condition of fluid flow in LBM is not sensitive for studying the steady-state flow and the associated long-term behavior [413], we simply set the initial flow velocity $\mathbf{u} = 0$ and fluid density $\rho = 1$ throughout the domain in the LBM simulation. Three types of boundary conditions are adopted in the LBM simulation for fluid flow inside porous media: the no-slip boundary condition (SBB scheme) on the pore-solid surface, fixed pressure boundary condition at the inlet and outlet faces, and periodic boundary condition added on the boundary faces that are parallel to the macroscopic flow direction. The fluid flows in the porous media only pass through the connected pore space, while the isolated pore space

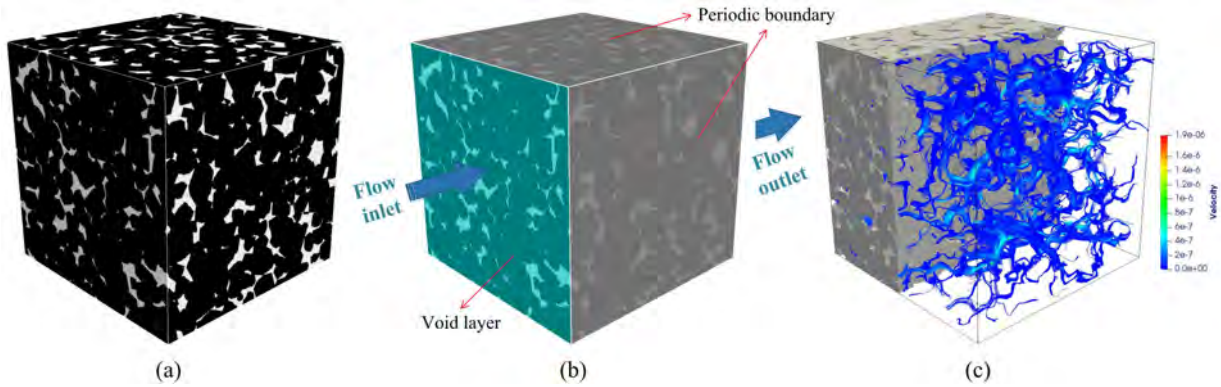


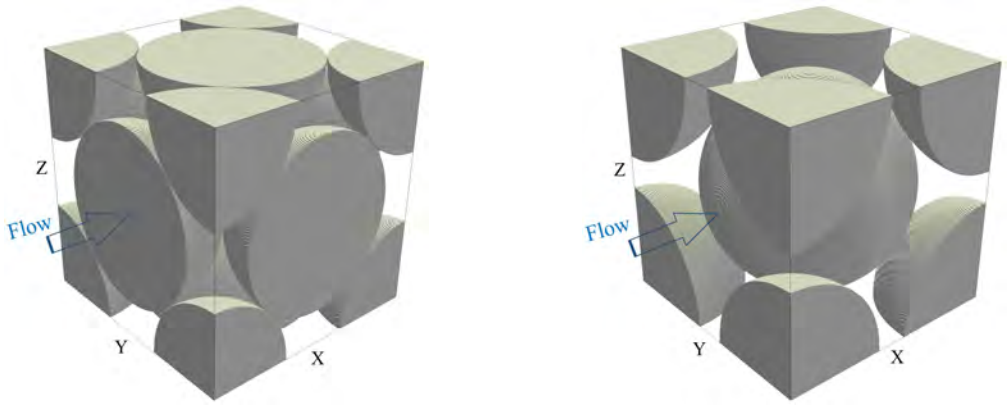
Fig. 6.2 LBM evaluation of intrinsic permeability for a porous medium: (a) The 3D digital image of the porous microstructure; (b) The schematic diagram of fluid flow through the porous medium; (c) The steady-state fluid velocity field inside the porous medium obtained from LBM simulation

does not contribute to the permeability. The 18-connectivity criterion is used to identify the connected pore space, which is consistent with the 19-velocity lattice arrangement (D3Q19). The LBM simulation runs continuously until the steady state of fluid flow is reached. The fluid flow is assumed to be stable when the standard deviation of average kinetic energy falls below 10^{-6} (the maximum number of iteration steps is 60,000).

6.2.3 Verification of LBM implementation

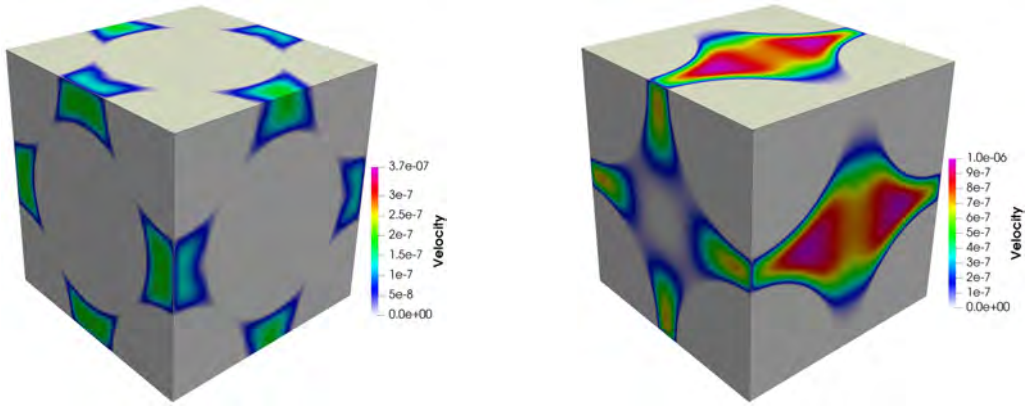
The intrinsic permeability defined by Darcy's law is valid for laminar flow with a low Reynolds number *Re*. Further, the permeability of porous media is independent of fluid properties including viscosity, density and pressure gradient. If the *Re* increases beyond approximately 1, Darcy's law starts to fail [29], because the inertial force plays an important role and it can not be neglected. In this section, we apply our implementation of SRT-LBM to calculate the permeabilities for two simple structures with known analytical permeability values. Firstly, the permeability results are checked with the pressure gradient varied over a wide range, in order to ensure the porous media flow is in the laminar regime. Secondly, the calculated permeability results are compared with the analytical values to verify our LBM implementation.

The 3D structures of the face-centered cube (FCC) and body-centered cube (BCC) are created as shown in Figure 6.3. The porosities of FCC and BCC are 25.95% and 31.98%, respectively. The permeabilities of FCC and BCC can be analytically determined by solving the associated Stocks equations [76], and the dimensionless permeability values (κ/D^2, where D is the sphere diameter) are 1.7360×10^{-4} (for FCC) and 5.0230×10^{-4} (for BCC).



(a) FCC (domain:$\sqrt{2}D \times \sqrt{2}D \times \sqrt{2}D$) (b) BCC (domain:$2/\sqrt{3}D \times 2/\sqrt{3}D \times 2/\sqrt{3}D$)

Fig. 6.3 The representations of FCC structure with the sphere diameter D = 340.12 voxels and BCC structure with the sphere diameter D = 347.28 voxels



(a) Velocity field inside FCC (b) Velocity field inside BCC

Fig. 6.4 The steady-state fluid velocity fields obtained from LBM simulations

Pressure gradients varying over several orders of magnitude are applied along the X direction of FCC and BCC models. With the relaxation time $\tau = 1.0$, the steady-state fluid velocity fields obtained from LBM simulations are shown in Figure 6.4. The calculated permeabilities of FCC and BCC models under different pressure gradients are listed in Table 6.1. It can be observed that the calculated dimensionless permeability (κ/D^2) varies with the applied pressure gradient, until the pressure gradient is smaller than a specific value. To examine the cause of permeability fluctuation, the Reynolds number Re of steady-state flow is calculated for each simulation case. The predicted permeabilities in the cases of high-pressure gradients have significant errors because the corresponding Reynolds numbers are much greater than 1.0 invalidating the assumption of Darcy's law. In the cases of low-pressure gradients, the calculated permeabilities are consistent with small relative errors at -1.092%

and -0.884% compared to the analytical values of FCC and BCC, respectively. As discussed in Section 6.3 later, the permeability estimation error mainly comes from the geometry error and the boundary error, both of which are negligible for the FCC and BCC structures considered here because of the fine-resolution grid adopted in the calculation. The verification confirms our LBM implementation is able to accurately estimate the permeability of porous media when the boundary conditions are set correctly. In the following sections, the resolution effect on permeability evaluation for porous media will be investigated quantitatively to develop an error correction model, for which the flow regime will be controlled with a low Reynolds number ($Re \ll 1$).

Table 6.1 Fluid flow simulations under different pressure gradients by using SRT-LBM

Structure	∇p (lattice unit)	ν (lattice unit)	$\langle \mathbf{u} \rangle$ (lattice unit)	κ/D^2	Error (%)	$Re = \frac{\langle \mathbf{u} \rangle D}{\nu}$
FCC	2.0661×10^{-4}	0.1667	1.9647×10^{-2}	1.3763×10^{-4}	-20.718	40.00
	2.0661×10^{-5}	0.1667	2.4595×10^{-3}	1.7229×10^{-4}	-0.753	5.01
	2.0661×10^{-6}	0.1667	2.4545×10^{-4}	1.7195×10^{-4}	-0.952	0.50
	2.0661×10^{-7}	0.1667	2.4514×10^{-5}	1.7173×10^{-4}	-1.079	4.99×10^{-2}
	2.0661×10^{-8}	0.1667	2.4511×10^{-6}	1.7170×10^{-4}	-1.092	4.99×10^{-3}
	2.0661×10^{-9}	0.1667	2.4510×10^{-7}	1.7170×10^{-4}	-1.092	4.99×10^{-4}
BCC	7.4257×10^{-5}	0.1667	3.3397×10^{-2}	6.2441×10^{-4}	24.311	69.43
	2.4753×10^{-5}	0.1667	8.6362×10^{-3}	4.8440×10^{-4}	-3.563	17.95
	2.4753×10^{-6}	0.1667	8.8652×10^{-4}	4.9725×10^{-4}	-1.006	1.84
	2.4753×10^{-7}	0.1667	8.8766×10^{-5}	4.9789×10^{-4}	-0.878	1.84×10^{-1}
	2.4753×10^{-8}	0.1667	8.8762×10^{-6}	4.9786×10^{-4}	-0.884	1.84×10^{-2}
	2.4753×10^{-9}	0.1667	8.8761×10^{-7}	4.9786×10^{-4}	-0.884	1.84×10^{-3}

Note: ∇p is the pressure gradient, ν is the kinetic viscosity, $\langle \mathbf{u} \rangle$ is the average flow velocity and D is the sphere diameter.

6.3 Resolution effect on permeability calculation though pore-scale simulation

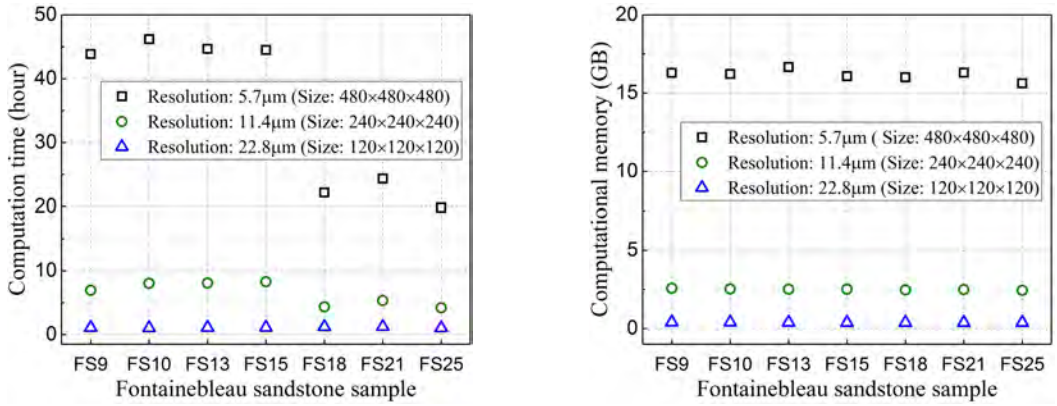
6.3.1 Permeability evaluation via SRT-LBM simulation

By using our implementation of SRT-LBM, pore-scale simulations are performed on a group of porous media to investigate the effects of image resolution and relaxation time on permeability results. Seven Fontainebleau rock samples are analyzed, whose porosities are 8.6%, 10.1%, 12.5%, 15.3%, 17.6%, 20.6%, and 24.5%, respectively [34]. The digital images of these samples are of 480×480×480 voxels with a voxel resolution of 5.7 μm, which represents an equivalent physical sample of 2.7 mm^3 cube. To check the resolution effect, these raw images are rescaled to 240×240×240 voxels and 120×120×120 voxels, representing the voxel resolution at around 11.4 μm and 22.8 μm, respectively. Figure 6.5 shows an example of the digital rock models at different resolution levels. The LBM

simulations are run on the ASTUTE HPC cluster in Swansea University (72 nodes, 2×14 core CPUs per node, 64 GB RAM per node).

(a) Resolution= 5.7 μm (b) Resolution= 11.4 μm (c) Resolution= 22.8 μm

Fig. 6.5 The binary representations of Fontainebleau sandstone with porosity 15.3% at three different resolution levels (white phase denotes pore space and black phase denotes solid matrix)

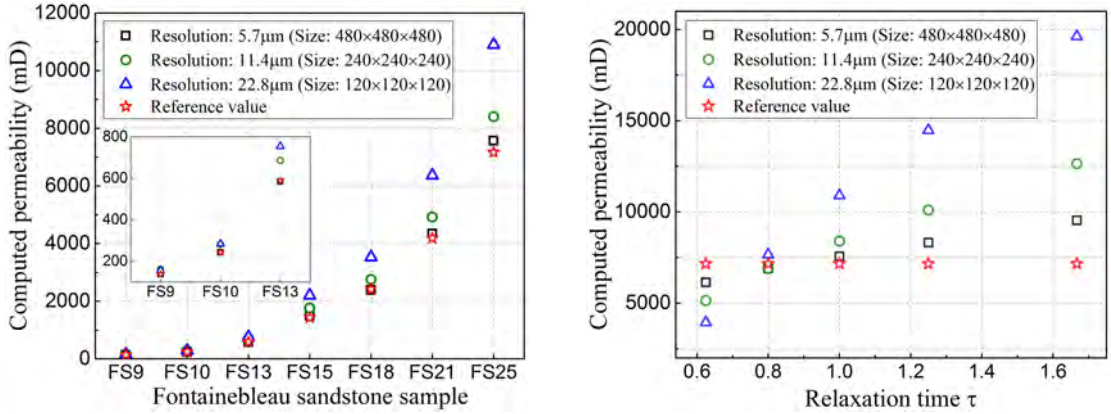


(a) Computation time (b) Computational memory

Fig. 6.6 The computation time and memory requirement of SRT-LBM simulations to evaluate intrinsic permeabilities (seven Fontainebleau sandstone samples FS9, FS10, FS13, FS14, FS15, FS18, FS21 and FS25 are studied, whose porosities are 8.6%, 10.1%, 12.5%, 15.3%, 17.6%, 20.6%, and 24.5%, respectively)

By setting the relaxation time $\tau = 1.0$, the permeabilities of these digital rocks with different resolutions are calculated through SRT-LBM simulations. As shown in Figure 6.6a and 6.6b, reducing the image resolution not only improves the efficiency of LBM simulation significantly, but it also greatly reduces the memory requirement. However, image resolution reduction also causes a negative impact on permeability evaluation, due

to the loss of geometric information related to the finer pore-scale microstructure. It is found that the permeability results evaluated from SRT-LBM simulations are different before and after image coarsening. As shown in Figure 6.7a, the computed permeabilities from the highest resolution images are closed to the reference values obtained from laboratory measurement [56], but the coarsened images tend to overestimate the permeability values. The evaluation errors caused by resolution reduction are between 12.48% and 47.37% for these Fontainebleau sandstone samples. For some other porous media, the evaluation errors caused by resolution reduction even exceed 100%, as illustrated in Figure 6.19a, 6.20a and 6.21a.



(a) Influence of image resolution

(b) Influence of relaxation time

Fig. 6.7 The influences of image resolution and relaxation time on permeability estimation from SRT-LBM simulations

By choosing different values of relaxation time τ, the SRT-LBM simulations are performed on a digital Fontainebleau sandstone sample with 24.50% porosity, and the permeability results are shown in Figure 6.7b. Apparently, the computed permeability is dependent on the relaxation time (viscosity), and they appear to be positively correlated. The extent to which the computed permeability depends on the relaxation time differs for digital rock images at different resolutions. Low-resolution images are very sensitive to the relaxation time, which has a significant influence on the accuracy of computed permeability. As shown in Figure 6.7b, the computed permeability at $\tau = 1.667$ is four times larger than the computed permeability at $\tau = 0.625$ for the image with the lowest resolution. For high-resolution images, the dependence of computed permeability on the relaxation time is much weaker, so the choice of τ value is less worrying.

From the above parametric investigations, it is clear that LBM simulations performed on higher resolution images provide more accurate permeability predictions. The reasons are: (1) the digital rock image at higher resolution contains more details so that the pore

geometry and structure are closer to the morphological reality; (2) the higher resolution grid makes LBM simulations less sensitive to the relaxation time, in which the physical location of the fluid-solid boundary is less misplaced when a sub-optimal relaxation time is chosen. However, practical limitations in computing power and waiting time often force the resolution of digital rock image to be reduced so that the associated pore-scale LBM simulation is computationally feasible. This resolution reduction usually has a remarkable effect on the accuracy of LBM simulation results, which is referred to as the resolution effect. Two primary sources of permeability evaluation error are identified in relation to the resolution effect: (1) the geometry error, which comes from the changes of pore geometry, including pore-solid interface and pore connectivity, as a result of voxel reassignment from image rescaling; (2) the boundary error, which comes from the aggravation of mismatch between fluid-solid boundary and pore-solid interface for the low-resolution image, when a sub-optimal relaxation time is chosen. Both the geometry error and the boundary error are related to the pore geometry, therefore, it is possible to use suitable geometric attributes to represent and quantify them, and then to correct the permeability evaluation error due to the resolution effect.

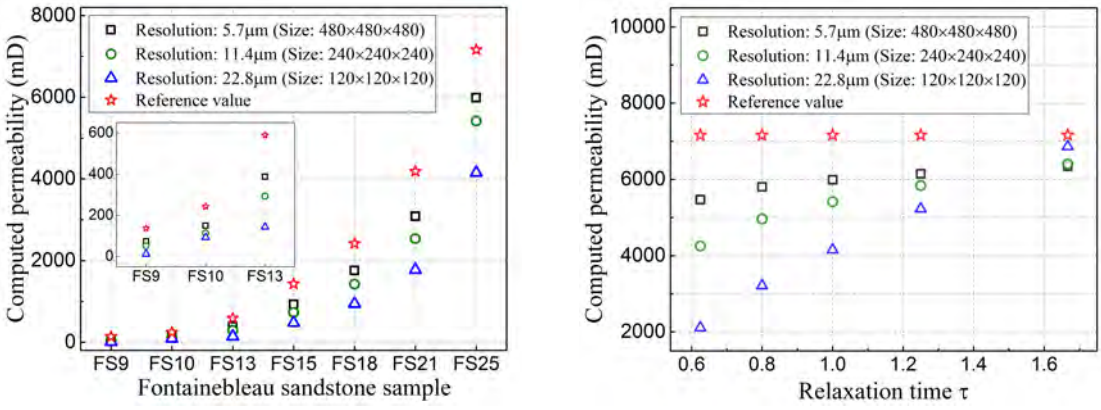
6.3.2 Comparison with other numerical schemes

The resolution effect on permeability calculation via SRT-LBM is studied in the above section, where the permeability results are compared with the reference values obtained from laboratory measurement. In this section, the examination of the resolution effect on permeability calculation is extended to other numerical schemes.

As discussed above, the permeability result calculated from SRT-LBM strongly depends on the relaxation time τ. Multiple-relaxation time LBM (MRT-LBM) was thus introduced by d'Humieres (2002) [118] to overcome some of the limitations of SRT-LBM, particularly to maintain the stability at low relaxation time. MRT-LBM allows different moments of the distribution function to relax at different rates, while for SRT-LBM all moments relax at the same rate [252]. A collision matrix is constructed to define individual relaxation times for all variables, which provides the maximum flexibility to optimize LBM stability. Further details about MRT-LBM can be found in relevant literature [118, 252]. The procedures to compute the intrinsic permeabilities of porous media via MRT-LBM are also well documented in the literature [129, 370]. We implement the MRT-LBM program to calculate intrinsic permeability for 3D porous media samples under the same initial state and boundary conditions as the SRT-LBM implementation (explained in Section 6.2).

By setting the relaxation time $\tau = 1.0$, the permeabilities of the digital Fontainebleau rocks with different resolutions are computed from MRT-LBM simulations, which are plotted

in Figure 6.8a. The image resolution is observed to have a great influence on the permeability results, and the computed permeabilities from the finest images are close to the reference values. Compared with the SRT-LBM simulation, the MRT-LBM simulation does not provide more accurate permeability evaluations when the relaxation time is $\tau = 1.0$. By choosing different values of relaxation time τ, the permeability results of a digital Fontainebleau rock with 24.50% porosity are computed via MRT-LBM simulations. It is observed that the viscosity-dependence (or relaxation time dependency) of MRT-LBM is much weaker than that of SRT-LBM, which is a well-known advantage of MRT-LBM. Therefore, if a sub-optimal relaxation time τ is chosen (for instance, when $\tau \geq 1.4$ is taken for the Fontainebleau sandstone samples), MRT-LBM usually provides more accurate permeability results compared to SRT-LBM. It is possible to adjust the MRT collision parameters so as to greatly reduce the viscosity-dependence, but such parameter adjustment mainly relies on experience [129, 370]. In addition, compared with SRT-LBM, MRT-LBM requires longer simulation time (1.5 to 2 times) and more computing memory (more than 1.5 times).



(a) Influence of image resolution

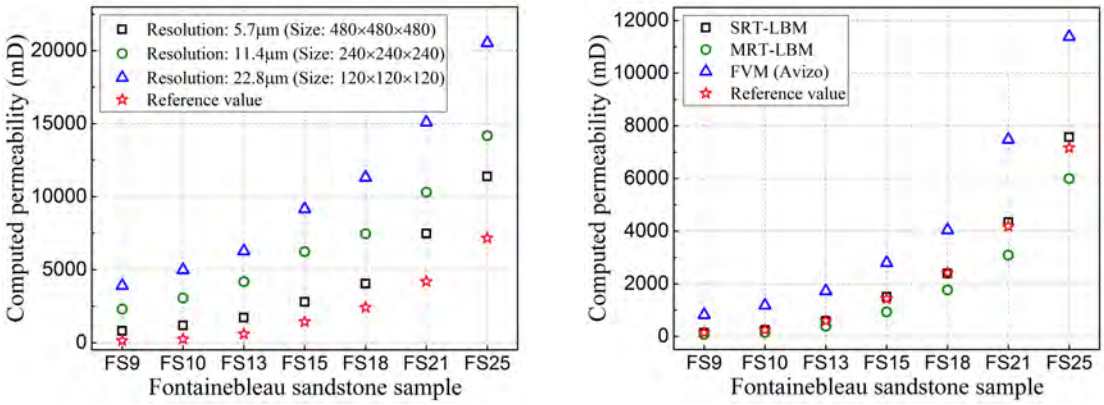
(b) Influence of relaxation time

Fig. 6.8 The influences of image resolution and relaxation time on permeability estimation from MRT-LBM simulations

Finite volume method (FVM) is another commonly-used numerical scheme to simulate pore-scale fluid flow [475, 356]. To examine the resolution in relation to the FVM scheme, a commercial FVM fluid dynamic solver, Avizo (FEI, Hillsboro, Oregon, USA), is used to simulate incompressible laminar flow through porous media to estimate intrinsic permeability. Avizo solves Navier–Stokes equations through voxel-based FVM [172], where the image voxels are used as the volume elements, and no additional meshing step is required. The no-slip boundary condition is set at the fluid-solid interface, and periodical boundary conditions are imposed on the surfaces of the cubic sample. Driven by the pressure difference between inlet and outlet, laminar flow with low Reynolds number is simulated to pass through the

porous medium sample. When the steady state is reached, the intrinsic permeability can be estimated based on Darcy's law.

As shown in Figure 6.9a, the permeability results are obtained by performing FVM simulations on digital Fontainebleau rocks with different resolutions. It is observed that the resolution effect on permeability calculation through FVM simulation is greater than the LBM simulation. Overall, FVM simulation tends to overestimate the permeability of porous media, especially when the image resolution is low. Besides, the permeability results of the digital Fontainebleau rocks with the highest resolutions are compared between LBM (including SRT-LBM and MRT-LBM with relaxation time $\tau = 1$) and FVM simulations, as illustrated in Figure 6.9b. For the case of $\tau = 1$, SRT-LBM provides the most accurate permeability results among three different numerical schemes, which at least demonstrates that SRT-LBM is capable of providing comparable permeability results when appropriate relaxation time is selected. Taking into account the simplicity, efficiency and popularity of SRT-LBM, further studies are desired to reduce or eliminate the resolution effect on its permeability calculation.



(a) Influence of image resolution (b) Comparison of permeability results

Fig. 6.9 (a) The influence of image resolution on permeability estimation from FVM simulations; (b) Comparison between permeability results estimated from SRT-LBM, MRT-LBM and FVM simulations

6.4 Error correction model for resolution effect

Accurate intrinsic permeability is expected to be obtained from the digital rock image with high resolution, but due to the high cost of LBM simulation, the image is usually rescaled into a lower resolution. As illustrated in Section 6.3, resolution reduction can significantly reduce the computational complexity of LBM simulation, but it also causes the geometry

and boundary errors that affect the pore-scale simulation results. Both types of error are associated to the pore geometry: the geometry error mainly comes from the changes of pore-solid interface and pore connectivity resulting from voxel reassignment; and the boundary error is primarily caused by the aggravation of mismatch between fluid-solid boundary and pore-solid interface when a sub-optimal relaxation time is chosen. In this section, we look for suitable geometric attributes to represent and quantify the geometry and boundary errors in permeability evaluation, and propose an error correction model to recover the accurate permeability from the results evaluated using coarsened images.

6.4.1 Digital rocks at different resolutions

A set of cubic micro-CT models of Fontainebleau sandstone are considered in this study. Fontainebleau sandstone is an isotropic porous medium, which is often used as the reference standard for validating models of porous media. It consists of monocrystalline quartz grains with an average size of about 200 μm, inside which only inter-granular porosity exists. Digital rock samples from the open database at Stuttgart University's Institute of Computational Physics [188] are used in this study.

The Fontainebleau sandstone sample with physical size of 15 mm^3 is imaged by micro-CT scanner at four different resolution levels: 3.662 μm, 7.324 μm, 14.648 μm and 29.296 μm, as shown in Figure 6.10. The micro-CT images are segmented into binary images by using *ImageJ*, a popular image processing tool in the DRP community. The default thresholding method in *ImageJ* is adopted here, which is a variation of the IsoData algorithm [452], and the segmentation results are also given in Figure 6.10, where the white phase indicates pore space and the black phase denotes solid grains. After image segmentation, the 3D binary images are partitioned into small RVEs for pore-scale simulation using LBM. The image sizes of RVEs at the four resolution levels (from high to low) are $640\times640\times640$, $320\times320\times320$, $160\times160\times160$ and $80\times80\times80$, respectively, and representative examples are shown in Figure 6.10. For each resolution level, around 15 RVEs are selected to perform SRT-LBM simulation of pore-scale fluid flow and evaluate their intrinsic permeabilities.

6.4.2 The geometry error coming from voxel assignment

Image acquisition and segmentation are essentially discretization processes, which assign the voxel value in the digital image to represent the material phases at the microscopic scale. Discrepancies in terms of pore geometry exist between digital rock images obtained at different resolutions, as highlighted by the red circles in Figure 6.12b (which are binary segmentations of micro-CT images in Figure 6.12a). Higher resolution images generally

(a) Micro-CT image (3.662 μm)

(b) Binary segmentation (3.662 μm)

(c) Micro-CT image (7.324 μm)

(d) Binary segmentation (7.324 μm)

(e) Micro-CT image (14.648 μm)

(f) Binary segmentation (14.648 μm)

(g) Micro-CT image (29.296 μm)

(h) Binary segmentation (29.296 μm)

Fig. 6.10 The 3D micro-CT images and corresponding binary segmentations (white = pore, black = solid) of the Fontainebleau sandstone at four different resolution levels

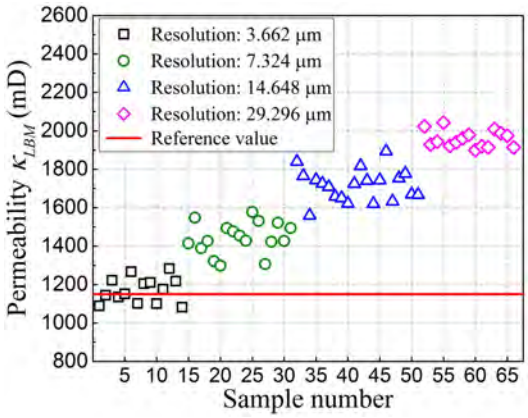


Fig. 6.11 Permeability results evaluated by LBM simulations for the RVEs at different resolutions

provide a closer representation for the reality inside rocks, while lower resolution images may lose fine-scale details of pore space geometry. The permeability evaluation error purely caused by the discrepancies of pore geometry between digital rock images at different resolutions is referred to as the geometry error. In practice, it is however difficult to separate the geometry error from the boundary error caused by relaxation time, unless the optimal relaxation times τ_0 is selected, where the exact pore-solid interface is used as the fluid-solid boundary during the LBM simulation. In this section, we set $\tau = 1.0$ to calculate the permeabilities for all RVEs of Fontainebleau sandstone sample at different resolutions, in order to avoid the influence coming from the relaxation time. The computed permeability results are recorded in Figure 6.11, and the reference permeability of the Fontainebleau sandstone sample is about 1150 mD [56]. The permeabilities evaluated from the RVEs at the highest resolution are close to the reference value, while the permeability results from the RVEs at lower resolutions are overestimated with errors ranging from 23.9% to 68.0%. It is clear that the resolution effect does have a significant impact on the LBM permeability evaluation.

The primary sources of error are the discrepancies of the pore-solid interface and pore connectivity between RVEs at different resolution levels. Therefore, it is possible, at least in principle, to characterize the evaluation error by using suitable geometric attributes of porous media. A well known empirical predication for the intrinsic permeability of porous media is Kozeny-Carman equation [90]:

$$\kappa = \frac{\phi^3}{\beta \tau_{\mathrm{H}}^2 S^2} \tag{6.6}$$

where ϕ is the porosity, S is the specific surface area equal to the ratio of the total pore-solid surface area to the bulk volume, β is a shape factor, and τ_{H} is the hydraulic tortuosity of the

porous medium. The Kozeny-Carman equation implies that porosity, specific surface area and tortuosity of the digital rock image could be useful geometric attributes to account for the permeability evaluation error due to the resolution effect.

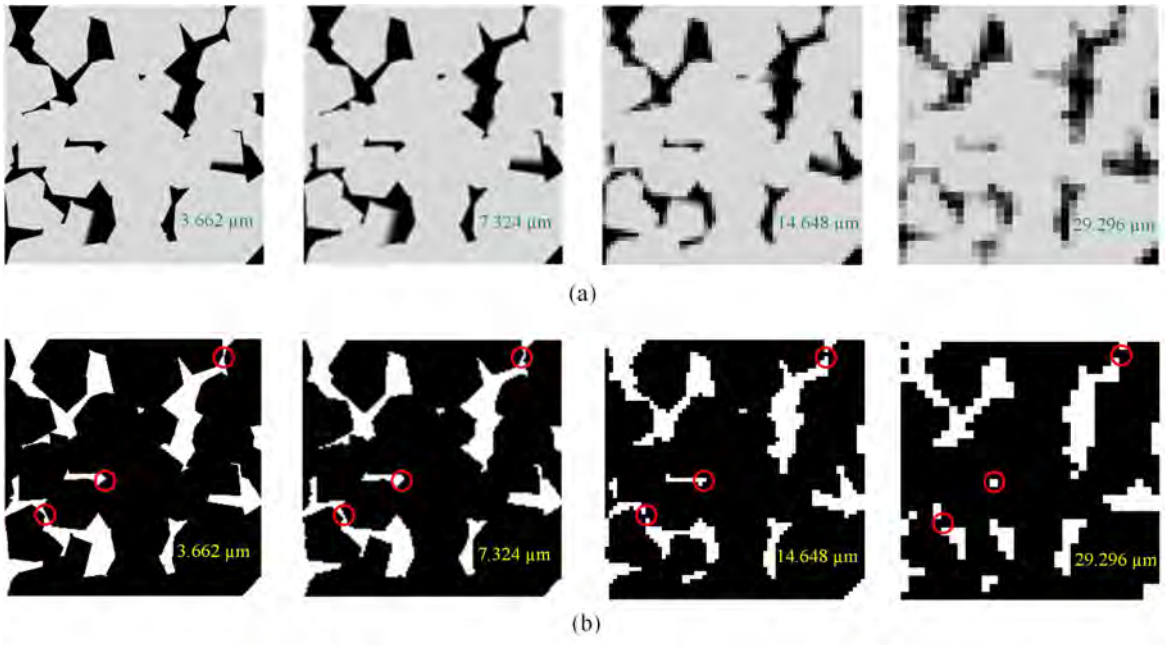


Fig. 6.12 (a) 2D slices of the 3D micro-CT images and (b) 2D slices of the corresponding binary segmentations at four different resolution levels

Note that fluid flows in the porous media only pass through the connected pore space in the LBM simulation. Therefore, we extract connected pore space from each RVE sample using the 18-connectivity criterion, and then calculate the connected porosity ϕ_c and the connected specific surface area S_c to investigate the geometric difference between RVEs at different resolutions. Figures 6.13a-b show the ϕ_c and S_c values extracted from these RVEs. The average connected porosities at four different resolutions are 13.674%, 13.981%, 14.143% and 13.687%, respectively. The average connected specific surface area at four different resolutions are 10.333, 9.659, 8.668 and 6.823 mm^{-1}, respectively.

The hydraulic tortuosity τ_H is used in the Kozeny-Carman equation to account for the retardation effect of tortuous flow paths, and it is defined as the ratio of flow path length to the Euclidean distance between inlet and outlet [90]. Although the definition of hydraulic tortuosity is conceptually sound, it is extremely difficult, if not impossible, to accurately estimate it from a digital image without performing a pore-scale flow simulation. Therefore, the geometrical tortuosity [90] and the diffusional tortuosity [163] are examined instead in this work. Through comparisons, the diffusional tortuosity is adopted to characterize the complexity of flow paths inside porous media. The diffusional tortuosity τ_d is obtained from

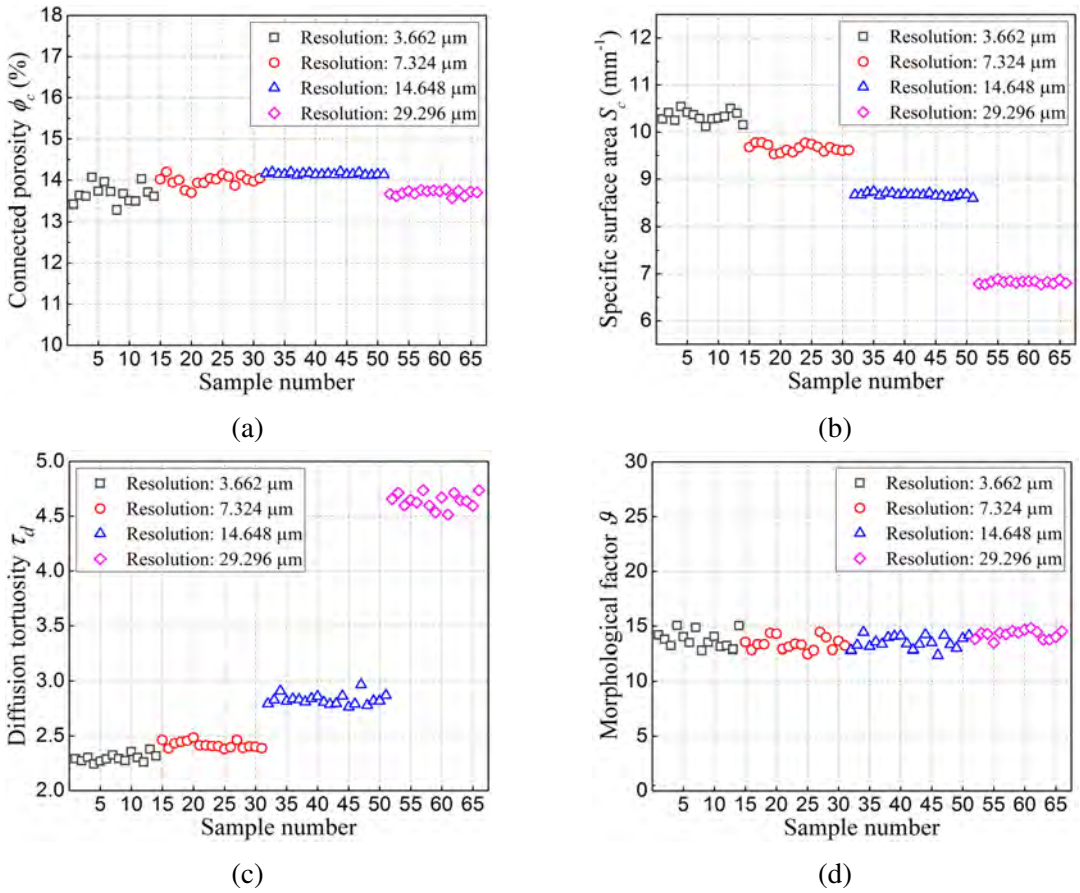


Fig. 6.13 The geometric attributes of Fontainebleau sandstone RVEs at different resolution levels: (a) Connected porosity ϕ_c; (b) Connected specific surface area S_c; (c) Diffusion tortuosity τ_d; (d) Morphological factor ϑ

diffusion simulation by the following equation:

$$D_p = \frac{\phi \sigma D_b}{\tau_d^2} \tag{6.7}$$

where D_p is the effective diffusion coefficient in the porous medium, σ is the constrictivity factor, and D_b is the diffusion coefficient of air or water filled inside the pore space. The constrictivity factor $\sigma \approx 1.0$ when the pore size is lager than 1 μm [163]. In this study, the diffusion tortuosity is obtained using the open-source code *TauFactor* [96]. Figure 6.13c shows the diffusion tortuosities calculated from these RVEs at four different resolutions, whose average values are 2.298, 2.419, 2.829 and 4.641, respectively. The diffusion tortuosity monotonically increases as the resolution of RVEs drops.

Following the above analysis, an improved model is proposed to link the intrinsic permeability κ_{LBM} evaluated using LBM simulation and the geometric attributes of digital rock

images:

$$\kappa_{\mathrm{LBM}} = \frac{\phi_{\mathrm{c}}^3}{\vartheta\sqrt{\tau_{\mathrm{d}}}S_{\mathrm{c}}^2} \tag{6.8}$$

where ϕ_{c} denotes connected porosity, S_{c} connected specific surface area, τ_{d} diffusion tortuosity, and ϑ a morphological factor independent from the resolution of digital images. By substituting the values of LBM permeability κ_{LBM}, connected porosity ϕ_{c}, connected specific surface area S_{c} and diffusion tortuosity τ_{d} of these RVEs into Eq. (6.8), the morphological factor ϑ can be obtained, as shown in Figure 6.13d. The numerical results confirm that the morphological factor ϑ is indeed a constant, independent from the image resolution.

As the geometric parameters ϕ_{c}, S_{c} and τ_{d} can all be easily extracted from 3D digital images, Eq. (6.8) can be used to correct the error of LBM permeability evaluation obtained at lower resolution. Specifically, as the morphological factor ϑ is a constant independent from image resolution, the following relation holds:

$$\vartheta = \frac{\phi_{\mathrm{c,h}}^3}{\kappa_{\mathrm{h}}\sqrt{\tau_{\mathrm{d,h}}}S_{\mathrm{c,h}}^2} = \frac{\phi_{\mathrm{c,l}}^3}{\kappa_{\mathrm{l}}\sqrt{\tau_{\mathrm{d,l}}}S_{\mathrm{c,l}}^2} \tag{6.9}$$

where the subscripts h and l denote the quantities corresponding to the high-resolution RVE_{h} and the low-resolution RVE_{l}, respectively. Rearranging the above equation yields:

$$\kappa_{\mathrm{h}} = \kappa_{\mathrm{l}}\frac{\phi_{\mathrm{c,h}}^3\sqrt{\tau_{\mathrm{d,l}}}S_{\mathrm{c,l}}^2}{\phi_{\mathrm{c,l}}^3\sqrt{\tau_{\mathrm{d,h}}}S_{\mathrm{c,h}}^2} = \kappa_{\mathrm{l}}C_{\mathrm{resolution}} \tag{6.10}$$

where $C_{\mathrm{resolution}} = \dfrac{\phi_{\mathrm{c,h}}^3\sqrt{\tau_{\mathrm{d,l}}}S_{\mathrm{c,l}}^2}{\phi_{\mathrm{c,l}}^3\sqrt{\tau_{\mathrm{d,h}}}S_{\mathrm{c,h}}^2}$ is the up-scaling factor fully determined by the geometric attributes of the 3D digital models RVE_{h} and RVE_{l}. It should be noted that Eq. (6.10) is established based on the computed permeabilities from LBM simulations by setting $\tau = 1.0$.

6.4.3 The boundary error coming from mismatch between fluid-solid boundary and pore-solid interface

The standard bounce-back (SBB) scheme is most popular in the literature to represent no-slip solid walls. In this scheme, the incoming fluid particles that hit the solid walls at a certain time step are bounced back to the nodes in the pore space where they came from. The SBB scheme assumes the fluid-solid boundary is located approximately half-way between the bounce-back solid node and its neighboring fluid node [252], as illustrated in Figure 6.14. However, it does not always locate the fluid-solid boundary at the half-way between the bounce-back solid node and its neighboring fluid node [252].

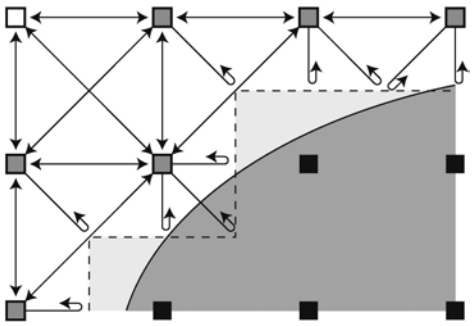


Fig. 6.14 Fluid-solid boundary handling in LBM simulations by using the SBB scheme [288] (Solid matrix is shown in dark gray, pore space (fluid phase) in white. The black squares are the lattice nodes in the solid phase, the gray-filled squares are the fluid nodes with solid node neighbor(s), and the white squares are lattice nodes completely within the fluid phase. The stair-shaped dash line represents the fluid-solid boundary in LBM simulations)

The classical lattice Boltzmann model with BGK collision operator and SBB scheme has a well-know defect that the actual physical position of fluid-solid boundary depends on the relaxation time τ. The spatial error is related to the grid spacing and proportional to $(\tau - 1/2)^2$ [156]. This can be easily observed from the analytic solution for Poiseuille-Hagen flow between two parallel plates [157]:

$$H_{\text{eff}}^2 = H^2 + 48\nu^2 - 1 = H^2 + \frac{16}{3}(\tau - 1/2)^2 - 1 \tag{6.11}$$

where H is the exact physical width between the parallel plates, and H_{eff} is the effective width. In this example, the solid wall is exactly located at the half-way between the bounce-back solid node and its neighboring fluid node if and only if $\tau = 0.933$.

However, there is no common rule to determine the optimal relaxation time for porous media with complex geometry [198]. The fluid-solid boundary of LBM simulation is not exactly located at the pore-solid interface of the porous medium, when a sub-optimal relaxation time is chosen. This mismatch between the fluid-solid boundary and pore-solid interface causes a numerical error to the LBM simulation, which is referred to as the boundary error. It tends to be aggravated after resolution reduction, and the permeability computed from the low-resolution image is more sensitive to the relaxation time than that computed from the high-resolution image (see Section 6.3).

The permeabilities of the Fontainebleau sandstone RVEs are computed through LBM simulations at different relaxation times ranging from 0.625 to 1.667, and the results are recorded in Figure 6.17a. The permeability results of four representative RVEs at different resolutions are given in Figure 6.15, from which it can be observed the computed permeability almost increases linearly with the relaxation time. Similar observations are also recorded by

other researches [129]. Hence, we assume a positive linear correlation between the computed permeability and the relaxation time. The computed permeability data at different relaxation time are linearly fitted as:

$$M = \frac{\kappa(\tau) - \kappa(1)}{\tau - 1} \tag{6.12}$$

where $\kappa(\tau)$ is the computed permeability at relaxation time τ, and M is the slope of the fitted line. The dashed lines in Figure 6.15 show the linear fitting results, which agree well with the original permeability data.

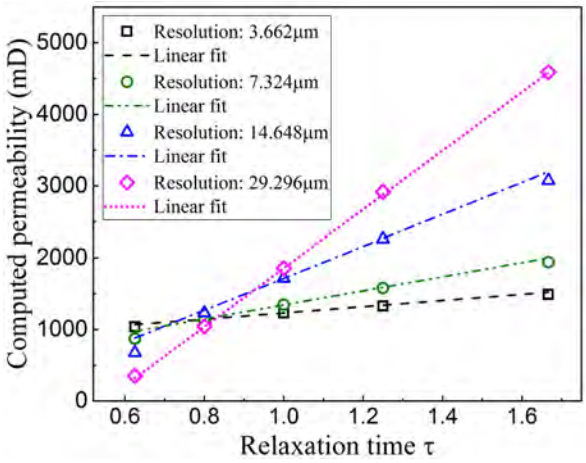


Fig. 6.15 The dependence of computed permeability on relaxation time for images at different resolution levels

If the slope M for a specific porous medium can be obtained without performing more LBM simulations, then the computed permeability $\kappa(\tau)$ at different relaxation time can be readily predicted from $\kappa(1)$ based on Eq. (6.12). As discussed above, the relaxation time controls the actual location of the fluid-solid boundary. Based on our experiments, it is assumed the effective connected porosity $\phi_c(\tau)$ and the effective specific surface area $S_c(\tau)$ of a porous medium are functions of the relaxation time τ, while other parameters such as morphological factor ϑ (or shape factor β) and tortuosity τ_d (or τ_H) are independent from the relaxation time τ. According to the Kozeny-Carman equation (6.6), the following relationship can be derived:

$$\begin{aligned} M = \frac{\mathrm{d}\kappa(\tau)}{\mathrm{d}\tau} &= \frac{\mathrm{d}\left(\dfrac{\phi^3(\tau)}{\beta\tau_d^2 S^2(\tau)}\right)}{\mathrm{d}\tau} \\ &= \frac{1}{\beta\tau_d^2}\left(3\frac{\phi^2(\tau)}{S^2(\tau)}\frac{\mathrm{d}\phi(\tau)}{\mathrm{d}\tau} - 2\frac{\phi^3(\tau)}{S^3(\tau)}\frac{\mathrm{d}S(\tau)}{\mathrm{d}\tau}\right) \end{aligned} \tag{6.13}$$

The computed permeability $\kappa(\tau)$ is assumed to be proportional to relaxation time τ, so the slope M is independent of τ. Suppose an optimal relaxation time τ_0 exists such that the fluid-solid boundary of the LBM simulation is exactly located at the pore-solid interface of the porous medium, the above equation can then be written as:

$$M = \frac{1}{\beta \tau_{\mathrm{d}}^2} \left(3 \frac{\phi_{\mathrm{c}}^2(\tau_0)}{S_{\mathrm{c}}^2(\tau_0)} \frac{\mathrm{d}\phi_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau} - 2 \frac{\phi_{\mathrm{c}}^3(\tau_0)}{S_{\mathrm{c}}^3(\tau_0)} \frac{\mathrm{d}S_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau} \right) \tag{6.14}$$

where $\phi_{\mathrm{c}}(\tau_0)$ denotes the connected porosity and $S_{\mathrm{c}}(\tau_0)$ the connected specific surface area. Both $\phi_{\mathrm{c}}(\tau_0)$ and $S_{\mathrm{c}}(\tau_0)$ can be directly extracted from the pore geometry, as shown in Figure 6.13a and 6.13b, respectively.

The remaining difficulty is to characterize $\frac{\mathrm{d}\phi_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau}$ and $\frac{\mathrm{d}S_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau}$. Note that the effective pore space expands as the relaxation time increases. As shown in Figure 6.16, the pore space expands by one voxel (pixel) thickness as a result of the relaxation time increasing from τ_0 to $\tau_0 + T$, where T is a small positive increment. Using this artificial case, $\frac{\mathrm{d}\phi_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau}$ and $\frac{\mathrm{d}S_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau}$ can be estimated as:

$$\frac{\mathrm{d}\phi_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau} \approx \frac{\phi_{\mathrm{c}}(\tau_0 + T) - \phi_{\mathrm{c}}(\tau_0)}{T} \tag{6.15}$$

$$\frac{\mathrm{d}S_{\mathrm{c}}(\tau_0)}{\mathrm{d}\tau} \approx \frac{S_{\mathrm{c}}(\tau_0 + T) - S_{\mathrm{c}}(\tau_0)}{T} \tag{6.16}$$

where $\phi_{\mathrm{c}}(\tau_0 + T)$ and $S_{\mathrm{c}}(\tau_0 + T)$ are the connected porosity and connected specific surface area of the porous medium with pore space expanded by one voxel (pixel) thickness, and they can be directly extracted from the expanded pore geometry, as shown in Figure 6.17c and 6.17d.

(a) Original pore space (b) The solid border (in yellow) (c) Expansion of pore space

Fig. 6.16 Illustration of pore space expanding for one pixel (voxel) thickness

Substituting Eq. (6.15), (6.16) and (6.6) into Eq. (6.14) yields:

$$\begin{aligned} M &= \frac{1}{\beta\tau_{\mathrm{d}}^2 T}\left(3\frac{\phi_{\mathrm{c}}^2(\tau_0)}{S_{\mathrm{c}}^2(\tau_0)}(\phi_{\mathrm{c}}(\tau_0+T)-\phi_{\mathrm{c}}(\tau_0))-2\frac{\phi_{\mathrm{c}}^3(\tau_0)}{S_{\mathrm{c}}^3(\tau_0)}(S_{\mathrm{c}}(\tau_0+T)-S_{\mathrm{c}}(\tau_0))\right) \\ &= \frac{1}{\beta\tau_{\mathrm{d}}^2 T}\left(3\frac{\phi_{\mathrm{c}}^2(\tau_0)}{S_{\mathrm{c}}^2(\tau_0)}(\phi_{\mathrm{c}}(\tau_0+T)-\phi_{\mathrm{c}}(\tau_0))-2\frac{\phi_{\mathrm{c}}^3(\tau_0)}{S_{\mathrm{c}}^3(\tau_0)}(S_{\mathrm{c}}(\tau_0+T)-S_{\mathrm{c}}(\tau_0))\right)\frac{\kappa(\tau_0)}{\kappa(\tau_0)} \\ &= \frac{\kappa(\tau_0)}{T}\left(3\frac{\phi_{\mathrm{c}}(\tau_0+T)-\phi_{\mathrm{c}}(\tau_0)}{\phi_{\mathrm{c}}(\tau_0)}-2\frac{S_{\mathrm{c}}(\tau_0+T)-S_{\mathrm{c}}(\tau_0)}{S_{\mathrm{c}}(\tau_0)}\right) \end{aligned} \tag{6.17}$$

where $\kappa(\tau_0)$ is the computed permeability at the optimal relaxation time τ_0 from the LBM simulation. The values of $\phi_{\mathrm{c}}(\tau_0)$, $\phi_{\mathrm{c}}(\tau_0+T)$, $S_{\mathrm{c}}(\tau_0)$ and $S_{\mathrm{c}}(\tau_0+T)$ are geometric attributes which can be quickly extracted from the pore geometries before and after the pore space expansion. However, $\kappa(\tau_0)$ and T are unknown, and are difficult to determine for a specific porous medium.

To overcome the above difficulty, the dimensionless permeability $\kappa(\tau)/\kappa(1)$ is considered, where the permeability calculated with the relaxation time $\tau = 1.0$ is used to normalize the permeability computed with an arbitrary relaxation time τ. As $\kappa(1)$ is a constant for a given digital rock image, the slope of the linear fitting to the dimensionless permeabilities is given by:

$$\begin{aligned} m = \frac{M}{\kappa(1)} &= \frac{\kappa(\tau_0)}{\kappa(1)T}\left(3\frac{\phi_{\mathrm{c}}(\tau_0+T)-\phi_{\mathrm{c}}(\tau_0)}{\phi_{\mathrm{c}}(\tau_0)}-2\frac{S_{\mathrm{c}}(\tau_0+T)-S_{\mathrm{c}}(\tau_0)}{S_{\mathrm{c}}(\tau_0)}\right) \\ &= C_\kappa E \end{aligned} \tag{6.18}$$

where the coefficient $C_\kappa = \dfrac{\kappa(\tau_0)}{\kappa(1)T}$ and the expansion ratio $E = 3\frac{\phi_{\mathrm{c}}(\tau_0+T)-\phi_{\mathrm{c}}(\tau_0)}{\phi_{\mathrm{c}}(\tau_0)} - 2\frac{S_{\mathrm{c}}(\tau_0+T)-S_{\mathrm{c}}(\tau_0)}{S_{\mathrm{c}}(\tau_0)}$.

As shown in Figure 6.17a, the RVE group of Fontainebleau sandstone are simulated by using LBM with different relaxation times, ranging from 0.625 to 1.667. For each RVE sample, a strong linear correlation exists between the computed permeability and the relaxation time, as illustrated in Figure 6.15. The normalized slope m for each RVE sample is plotted in Figure 6.17b. With the decrease of image resolution, the normalized slope increases, which implies the permeabilities computed from lower resolution images are more sensitive to the relaxation time. The geometric attributes including the connected porosity and the connected specific surface area are extracted from the original and expanded porous media, and the results are given in Figures 6.13a, 6.13b, 6.17c and 6.17d, respectively. According to Eq. (6.18), the expansion ratio E is obtained from the geometric attributes for each RVE sample, which are plotted in Figure 6.17e. By substituting the normalized slope m and the expansion ratio E to Eq. (6.18), the coefficient $C_\kappa = \dfrac{\kappa(\tau_0)}{\kappa(1)T}$ can be obtained, as plotted in

Figure 6.17f. It is observed that the coefficient C_κ remains constant as the resolution varies. Hence, C_κ is independent of image resolution for a specific porous medium:

$$C_\kappa = \frac{m_\mathrm{h}}{E_\mathrm{h}} = \frac{m_\mathrm{l}}{E_\mathrm{l}} \tag{6.19}$$

where the subscripts h and l denote the quantities corresponding to the high resolution RVE_h and the low resolution RVE_l, respectively. The slope m_h can be estimated from m_l as follows:

$$m_\mathrm{h} = m_\mathrm{l}\frac{E_\mathrm{h}}{E_\mathrm{l}} = m_\mathrm{l} C_\mathrm{relaxation} \tag{6.20}$$

where $C_\mathrm{relxation} = \dfrac{E_\mathrm{h}}{E_\mathrm{l}}$, E_h and E_l are expansion ratios directly extracted from RVE_h and RVE_l, respectively, and m_l can be quickly estimated from LBM simulations performed on RVE_l.

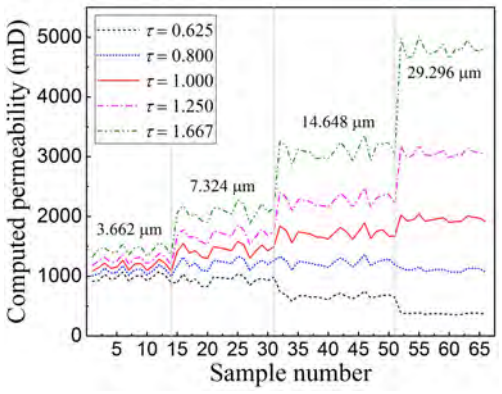
By combining Eq. (6.10) and Eq. (6.20), the computed permeabilities of RVE_h can be estimated from the computed permeabilities of RVE_l, given by:

$$\kappa_\mathrm{h}(\tau) = C_\mathrm{resolution} C_\mathrm{relaxation}\left(\kappa_\mathrm{l}(\tau) - \kappa_\mathrm{l}(1)\right) + C_\mathrm{resolution}\kappa_\mathrm{l}(1) \tag{6.21}$$

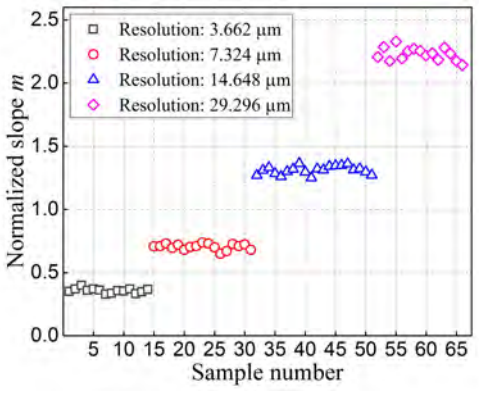
where $C_\mathrm{resolution}$ is the up-scaling factor. Both $C_\mathrm{resolution}$ and $C_\mathrm{relaxation}$ are determined by the geometric attributes of the 3D digital models RVE_h and RVE_l. It should be noted that Eq. (6.21) can be used to correct the permeability error due to the resolution effect for LBM simulations with any relaxation time τ.

6.4.4 Remarks on image resizing methods

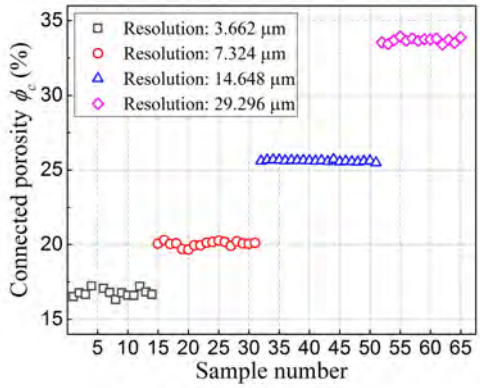
The choice of image resizing methods may affect the image rescaling results, because different voxel reassignment schemes can potentially introduce different geometry errors. There are two categories of image resizing methods for adjusting the resolution of micro-CT images of porous rocks: the device-based method and the software-based method. The device-based method is implemented by the specific settings on the micro-CT facility during image acquisition, to achieve a trade-off between the image resolution and the size of physical samples. The spatial resolution of a micro-CT image depends not only on scanning parameters such as focal-spot size and detector element dimensions, but also the resolving power linked to the image reconstruction. After image acquisition, the image resolution can also be artificially adjusted (reduced) by using software-based methods such as image interpolation [130] and image pyramid reduction [65], to make the associated LBM simulation more computationally affordable. As shown in Figure 6.18, different software-based methods are performed on a binary segmentation of micro-CT image, and



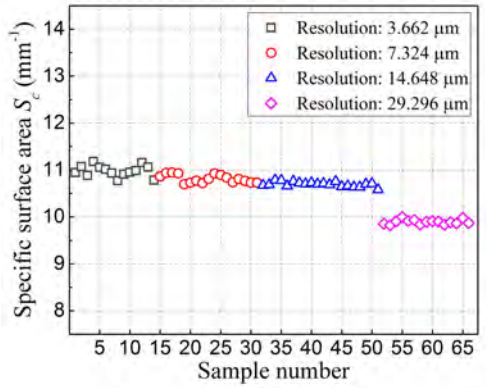
(a) Computed permeability results



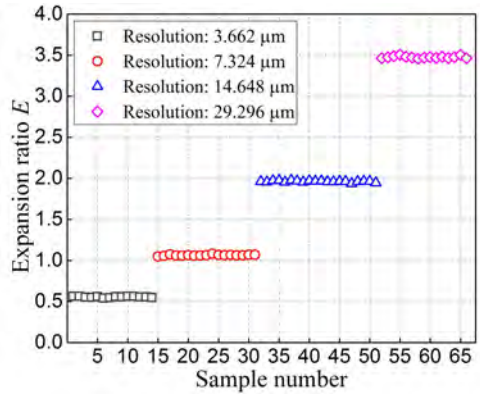
(b) Normalized slope values



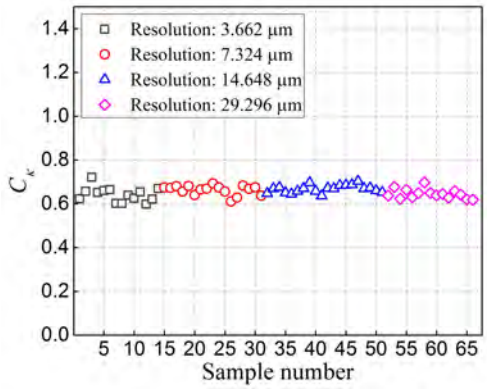
(c) Connected porosity



(d) Specific surface area



(e) Expansion ratio values



(f) The values of C_κ

Fig. 6.17 The computed results of the Fontainebleau sandstone RVEs at different resolution levels: (a) Computed permeability $\kappa(\tau)$; (b) Normalized slope m; (c) Connected porosity $\phi_c(\tau_0+T)$ of expanded pore geometry; (d) Connected specific surface area $S_c(\tau_0+T)$ of expanded pore geometry; (e) Expansion ratio E; (f) C_κ

very slight differences are observed between these rescaling results. LBM simulations are applied to these 3D digital models (corresponding to the 2D slices in Figure 6.18), and

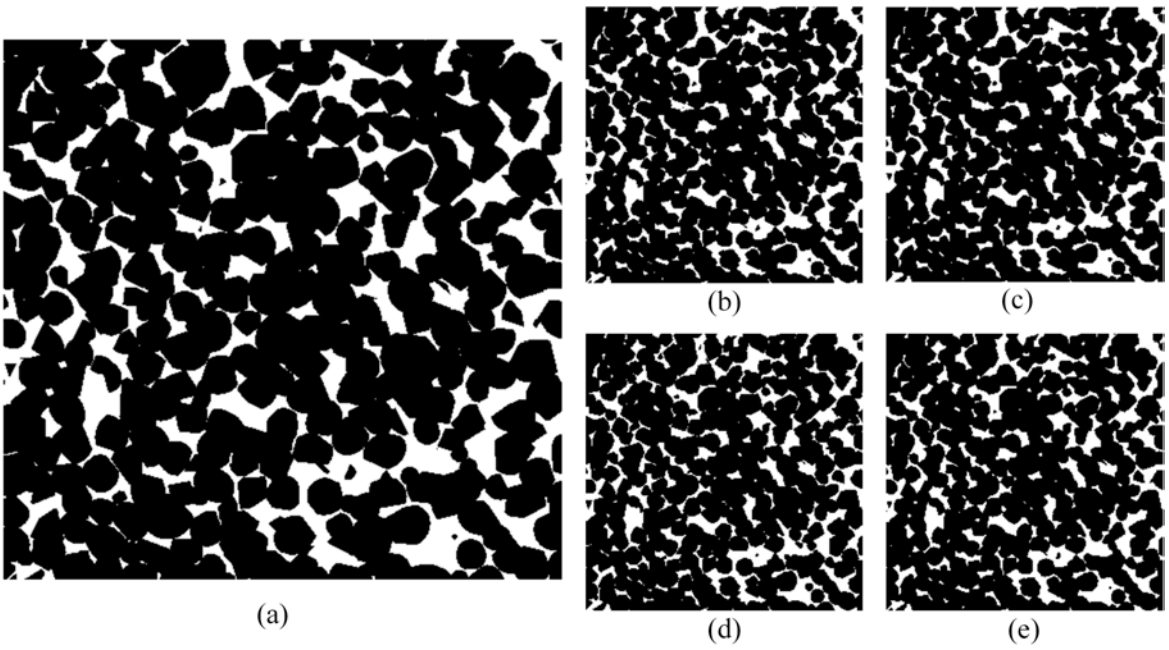


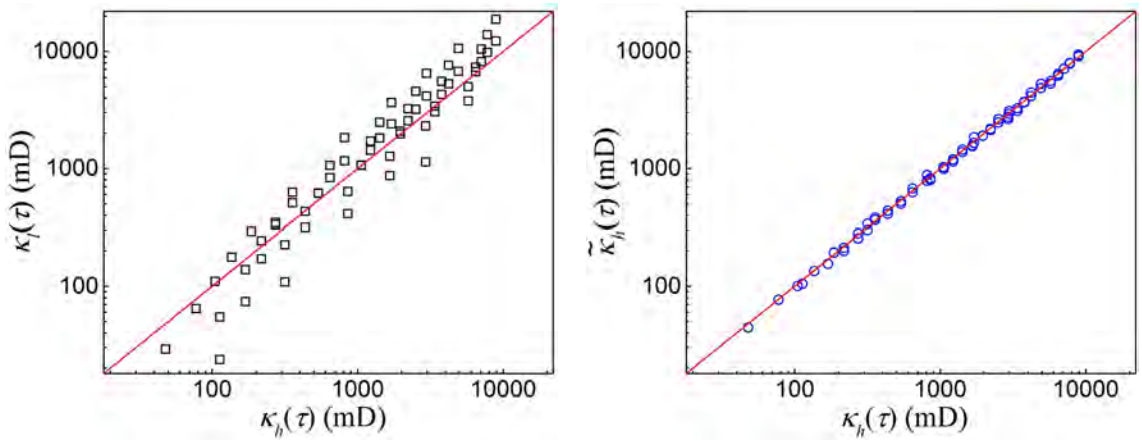
Fig. 6.18 Comparison between different image resizing methods: (a) A 2D slice in the original 3D image at resolution 5.7 μm and it is rescaled to 11.4 μm by using four different software-based resizing methods; (b) The rescaled image by the cubic interpolation method; (c) The rescaled image by the linear interpolation; (d) The rescaled image by the nearest interpolation method; (d) The rescaled image by the Gaussian pyramid reduction method

the permeability results at $\tau = 1$ are 4343.5 (original), 4919.1 (cubic interpolation), 4928.7 (linear interpolation), 4947.1 (nearest interpolation), and 4890.5 mD (Gaussian pyramid reduction), respectively. The permeability differences between the four rescaled images are negligible compared with the evaluation error caused by resolution reduction and bad choice of relaxation time.

It is possible, at least in principle, to develop a specialized image resizing method which is able to preserve comprehensively the pore geometry details and avoid the geometry error. However, such methods are content-dependent and it is extremely difficult, if not impossible, to design a resizing scheme for porous media due to their complicated pore geometries. In this study, the cubic interpolation method is adopted to rescale the image, after which the result is binarized using the global threshold scheme [343] to keep the porosity consistent. The proposed error correction model (as illustrated in Eq. (6.21) is established based on micro-CT images with different resolutions resulted from the device-based method. In the next section, the error correction model is demonstrated for the resolution effect arising from the software-based resizing method.

6.5 Results and discussion

The error correction model in Eq. (6.21) is established following the analysis of a Fontainebleau sandstone with a porosity of about 14.00%. It is necessary to examine the applicability of the model on other Fontainebleau sandstones with different porosities. The Fontainebleau sandstone samples discussed in Section 6.1 are employed for this purpose, whose porosities range from 8.61% to 24.50%. The resolution of raw images is 5.7 μm, and they are rescaled to 11.4 μm and 22.8 μm. The permeabilities of these digital rock images with different resolutions are evaluated through LBM simulations, and different relaxation times ranging from 0.625 to 1.667 are selected in the simulation.



(a) Permeability errors (b) Permeability correction

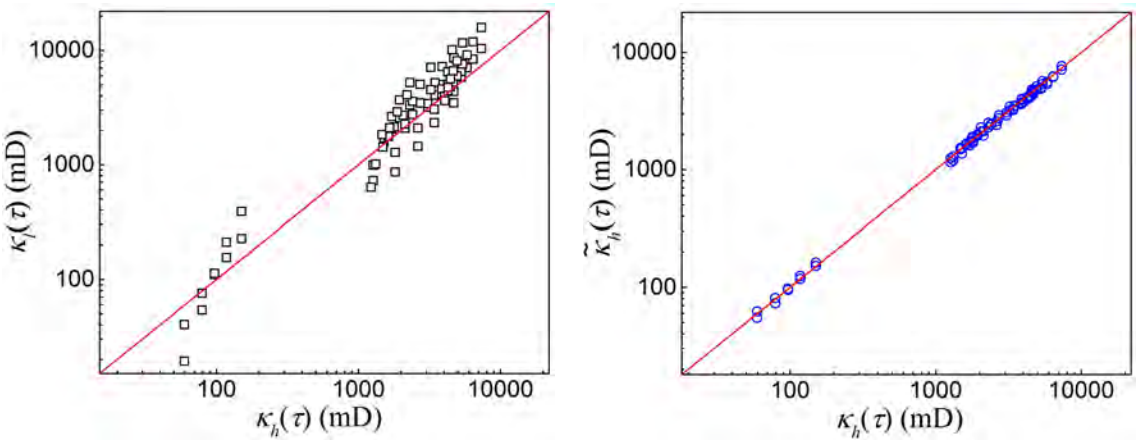
Fig. 6.19 Tests on Fontainebleau sandstone of different porosities: (a) Errors of computed permeabilities between high and low-resolution images; (b) Error correction for the permeabilities computed from low-resolution images

As shown in Figure 6.19a, the permeabilities $\kappa_l(\tau)$ computed from low resolution images are different from the permeabilities $\kappa_h(\tau)$ computed from high resolution images. The relative errors between $\kappa_l(\tau)$ and $\kappa_h(\tau)$ are significant, varying between -79.02% to 126.06%. After applying the error correction, the permeability results $\tilde{\kappa}_h(\tau)$ are recovered from $\kappa_l(\tau)$, and as shown in Figure 6.19b they are in good agreement with $\kappa_h(\tau)$. The errors between $\tilde{\kappa}_h(\tau)$ and $\kappa_h(\tau)$ are less than 5.00% for all rescaled images with 11.4 μm resolution. As to the rescaled images with 22.8 μm resolution, the errors between $\tilde{\kappa}_h(\tau)$ and $\kappa_h(\tau)$ are slightly larger, but all within 6.50%. These testes demonstrate the error correction model is not limited by the porosity range of Fontainebleau sandstones. The image information and geometric attributes of the Fontainebleau sandstone samples are listed in Table 6.2.

Table 6.2 The intermediate variables of permeability error correction for Fontainebleau sandstones of different porosities

Sandstone	Resolution (μm)	Size (voxel)	$\phi_c(\tau_0)$	$S_c(\tau_0)$ (mm^{-1})	τ_d	$\phi_c(\tau_0+T)$	$S_c(\tau_0+T)$ (mm^{-1})
Fontainebleau 1	5.700	480^3	8.26%	7.384	4.950	11.96%	8.503
	11.400	240^3	7.88%	5.966	6.882	14.04%	7.654
Fontainebleau 2	5.700	480^3	9.99%	8.645	3.658	14.29%	9.875
	11.400	240^3	9.65%	7.190	4.684	17.09%	9.104
Fontainebleau 3	5.700	480^3	12.51%	10.320	2.915	17.59%	11.593
	11.400	240^3	12.41%	8.956	3.573	21.56%	11.066
	22.800	120^3	11.77%	6.862	6.162	27.07%	9.919
Fontainebleau 4	5.700	480^3	15.33%	11.833	2.295	21.10%	13.039
	11.400	240^3	15.28%	10.437	2.616	25.77%	12.487
	22.800	120^3	15.07%	8.435	3.630	33.40%	11.589
Fontainebleau 5	5.700	480^3	17.61%	12.911	2.041	23.85%	14.003
	11.400	240^3	17.58%	11.527	2.278	28.99%	13.423
	22.800	120^3	17.48%	9.485	2.954	37.66%	12.491
Fontainebleau 6	5.700	480^3	20.64%	14.181	1.8153	27.43%	15.082
	11.400	240^3	20.63%	12.800	1.996	33.09%	14.411
	22.800	120^3	20.57%	10.671	2.438	42.70%	13.287
Fontainebleau 7	5.700	480^3	24.50%	15.447	1.652	31.80%	16.090
	11.400	240^3	24.49%	14.092	1.784	37.95%	15.270
	22.800	120^3	24.46%	11.977	2.102	48.40%	13.864

Note: $\phi_c(\tau_0)$ and $\phi_c(\tau_0+T)$ are connected porosities of the porous medium and the expanded porous medium, respectively. $S_c(\tau_0)$ and $S_c(\tau_0+T)$ are connected specific surface areas of the porous medium and the expanded porous medium, respectively. τ_d is the diffusional tortuosity of the porous medium (same to the tables below).



(a) Permeability errors (b) Permeability correction

Fig. 6.20 Tests on different types of sandstones: (a) Errors of computed permeabilities between high and low resolution images; (b) Error correction for the permeabilities computed from low resolution images

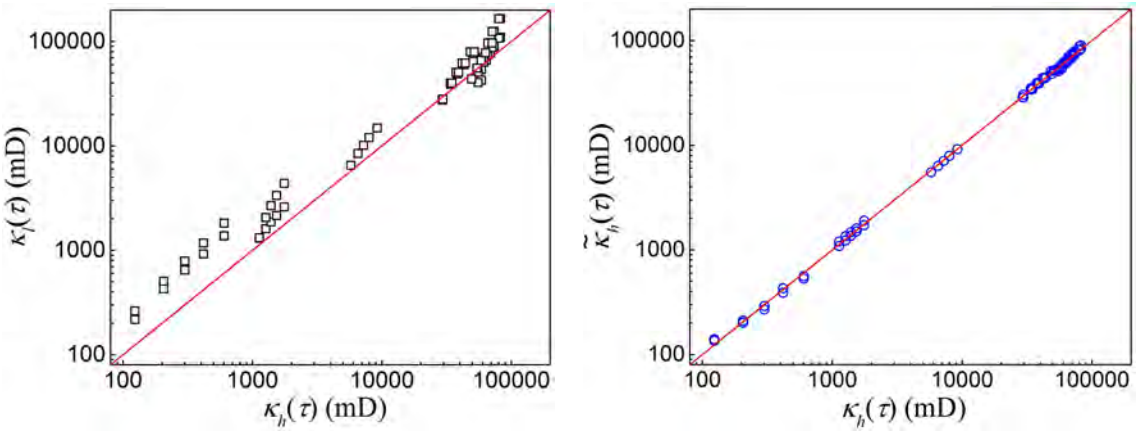
Next, the proposed error correction model is applied to different types of sandstones for further applicability test. Six types of sandstones are examined in addition to Fontainebleau sandstone, and the image information and geometric attributes of the sandstone samples are

given in Table 6.3. The raw images of these sandstone samples are rescaled into two lower resolutions. Still the relative errors between the computed permeabilities $\kappa_l(\tau)$ and $\kappa_h(\tau)$ are remarkable, ranging from -67.58% to 159.76% as shown in Figure 6.20a. By applying the error correction model, the permeability errors are dramatically reduced to less than 5.00% in most cases (the largest error is 7.15%), as shown in Figure 6.20b. These results demonstrate the error correction model works well for different types of sandstones.

Table 6.3 The intermediate variables of permeability error correction for different types of sandstones

Sandstone	Resolution (μm)	Size (voxel)	$\phi_c(\tau_0)$	$S_c(\tau_0)$ (mm^{-1})	τ_d	$\phi_c(\tau_0+T)$	$S_c(\tau_0+T)$ (mm^{-1})
Berea	5.345	400^3	19.53%	15.815	2.133	26.60%	17.420
	10.690	200^3	19.42%	13.953	2.411	32.50%	16.407
	21.380	100^3	19.06%	10.593	3.421	40.28%	13.885
Doddington	2.775	480^3	20.30%	14.196	1.680	23.46%	14.556
	5.549	240^3	20.26%	12.722	2.006	25.96%	13.525
	11.098	120^3	20.15%	11.120	2.265	30.34%	12.543
Bentheimer 1	6.007	400^3	21.56%	13.954	1.935	28.54%	14.999
	12.014	200^3	21.55%	12.661	2.106	34.57%	14.521
	24.028	100^3	21.51%	10.535	2.661	44.70%	13.303
Bentheimer 2	5.000	480^3	25.43%	16.744	1.794	32.63%	17.412
	10.000	240^3	25.36%	14.626	1.896	37.83%	16.021
	20.000	120^3	25.21%	11.974	2.203	46.41%	14.155
Bentheimer 3	5.000	480^3	22.30%	15.037	1.908	28.80%	15.779
	10.000	240^3	22.26%	13.723	2.033	33.73%	14.923
	20.000	120^3	22.14%	11.578	2.393	41.83%	13.462
Gildehauser	4.400	400^3	19.26%	14.320	2.107	24.55%	15.507
	8.800	200^3	19.19%	13.530	2.279	29.32%	15.587
	17.600	100^3	19.09%	11.621	2.799	37.60%	14.900
Leopard	3.475	480^3	13.37%	15.722	4.396	18.20%	18.100
	6.950	240^3	13.18%	13.861	5.216	21.95%	17.555
	13.901	120^3	12.48%	10.360	10.849	26.62%	14.831
Castlegate	5.600	480^3	20.62%	16.417	2.155	28.44%	17.984
	11.200	240^3	20.60%	15.325	2.208	35.64%	17.914
	22.400	120^3	20.50%	12.588	2.759	47.33%	15.954

Finally, the error correction model is further tested on other porous media including carbonate rocks, sand packs and synthetic silica, etc. The detailed information about these porous media samples is given in Table 6.4. The permeability errors caused by resolution reduction show similar patterns as that of sandstones. As shown in Figure 6.21a, the errors between $\kappa_l(\tau)$ and $\kappa_h(\tau)$ is significant, ranging from -27.61% to 179.40%. By applying the proposed error correction model, the permeability errors are successfully reduced to less than 5.00% in most of the cases (the largest error is 7.80%), as can be seen in Figure 6.21b. These tests demonstrate the proposed error model is applicable for many more types of porous media.



(a) Permeability errors (b) Permeability correction

Fig. 6.21 Tests on other types of porous media. (a) Errors of computed permeabilities between high and low resolution images; (b) Error correction for the permeabilities computed from low resolution images

Table 6.4 The intermediate variables of permeability error correction for different types of porous media

Porous media	Resolution (μm)	Size (voxel)	$\phi_c(\tau_0)$	$S_c(\tau_0)$ (mm^{-1})	τ_d	$\phi_c(\tau_0+T)$	$S_c(\tau_0+T)$ (mm^{-1})
Carbonate 1	2.850	400^3	21.24%	20.677	2.255	26.06%	20.276
	5.700	200^3	21.06%	16.565	2.750	28.76%	16.759
	11.400	100^3	21.03%	12.831	3.211	33.06%	13.570
Carbonate 2	5.345	400^3	14.02%	10.007	5.484	18.54%	10.362
	10.690	200^3	13.82%	7.878	5.819	20.98%	8.243
	21.380	100^3	13.77%	5.540	5.798	23.94%	6.062
Sand pack F42A	9.996	300^3	32.65%	10.070	1.304	40.56%	9.944
	19.992	150^3	32.55%	8.930	1.555	46.59%	8.939
	39.984	75^3	32.29%	7.589	1.754	56.69%	7.765
Sand pack F42B	10.002	300^3	33.35%	10.611	1.334	41.66%	10.373
	20.004	150^3	33.22%	9.518	1.535	48.15%	9.381
	40.008	75^3	32.95%	7.997	1.735	58.63%	8.033
Sand pack F42C	10.002	300^3	32.86%	10.689	1.339	41.23%	10.500
	20.004	150^3	32.77%	9.768	1.454	48.04%	9.563
	40.008	75^3	32.55%	8.211	1.622	58.93%	8.120
Sand pack LV60A	10.002	300^3	36.78%	15.054	1.382	48.70%	14.197
	20.004	150^3	36.67%	13.027	1.509	57.21%	12.029
Sand pack LV60B	8.851	300^3	36.85%	15.642	1.363	41.66%	11.722
	17.702	150^3	36.75%	13.637	1.467	48.15%	10.601
Sand pack LV60C	10.002	300^3	37.59%	15.963	1.390	50.24%	15.033
	20.004	150^3	37.48%	13.708	1.483	59.11%	12.600
Synthetic silica	3.850	150^3	44.73%	44.577	1.452	55.29%	40.903
	7.700	75^3	44.78%	36.431	1.575	61.94%	32.454

In summary, the error correction model has been tested on 24 different porous media samples for LBM permeability evaluation, and it succeeds in all cases and significantly

reduces the errors caused by the resolution effect. Combining the resolution reduction and the proposed error correction model, one can evaluate the intrinsic permeability of porous media both accurately and efficiently.

6.6 Conclusions and future work

The resolution effect related to the permeability evaluation of porous media using LBM simulation is systematically investigated in this chapter, and an error correction model is proposed to compensate for the accuracy loss due to the use of low-resolution images. Two primary sources of error are identified for the resolution effect: (1) the geometry error, which is caused by the variation of pore geometry between digital images at different resolution levels; (2) the boundary error, which is caused by misidentification of the fluid-solid boundary for LBM simulations. A specific set of geometric attributes are defined to quantify these two errors, from which a novel error correction model is proposed. A large number of numerical tests have been performed to examine the performance of the new error correction model, covering the differences in rock types, resizing methods, image resolutions and relaxation times. In all our tests, the error correction model reduces dramatically the permeability evaluation error and the remaining error is less than 5.00% in most cases. By ensuring simultaneously accuracy and efficiency for LBM permeability evaluation, the error correction model makes the practice use of image resolution reduction more meaningful and creditable in digital rock physics applications.

The new error correction model is not without limitation. If the image resolution is overly reduced, the morphological difference between the original and coarsened images may become too large to be workable for error correction. It is especially true for the micro-CT images scanned at low resolution and for the porous media with very low porosity. As for future research, it is interesting to look for the optimal relaxation time of LBM simulation for a specific porous medium such that the fluid-solid boundary is exactly located at the pore-solid interface. Moreover, the resolution effect on multi-phase fluid flow simulations through LBM or other numerical methods can also be studied in a similar way.

It's worth reminding that the error correction model provides more credible permeability values of porous media samples for the subsequent study to deeply investigate the microstructure-permeability linkage in Chapter 7.

Chapter 7

Dependence of permeability on microstructural characteristics of porous rocks: Investigation and modeling through feature selection and machine learning

7.1 Introduction

Permeability quantifies the ability of a porous medium to transmit fluid [29], which is a fundamental parameter to understanding the transport behaviors of fluid flows. Permeability plays a critical role in many geological applications, including oil and gas recovery, geothermal energy exploitation, CO_2 underground storage, radioactive waste disposal and contaminant hydrogeology. The permeable pore spaces within natural geologic materials are often highly distinct from each other, which leads to an extremely broad range of permeability values, varying up to 13 orders of magnitude [294]. Essentially, the transport properties of a porous medium strongly depend on its microstructural characteristics. The structure-property relationship is one of the most elementary issues in porous media research, from which the macroscopic physical property is expected to be predicted from the measurable microstructure informatics. However, the intricacy, stochasticity and heterogeneity inherent in the natural porous rocks make it difficult to accurately and rapidly evaluate permeability, especially for tight rocks with low porosity. Therefore, a deep insight into the dependence of

permeability on microstructural characteristics is always being pursued, aiming to achieve a reliable and efficient method for permeability prediction [46, 92, 489, 34, 142].

Laboratory measurement is the routine way to determine permeability, where fluid flow is driven by a constant pressure difference to pass through a rock core, and permeability is evaluated according to Darcy's law when the fluid flow reaches the steady state [29]. One the other hand, great efforts have been made to incorporate permeability using analytical or empirical models (as listed in Table 7.1), such as the famous semi-empirical Kozeny-Carman relation [143] and many variants deriving from it [38, 102, 34, 470]. Generally, these models rely on specific microstructural characteristics of porous media, such as porosity, specific surface area, tortuosity, characteristic length, pore size, constriction factor, fractal dimension and others. Despite the simplicity and convenience in actual practice, analytical models are often highly idealized and empirical models usually contain adjustable parameters of uncertainty, which make them only appropriate and reliable for specific or ideal pore microstructures and often break down for natural rocks with complicated pore networks.

Recently, the digital rock physics (DRP) technique has been rapidly developed to be an alternative to laboratory measurement, which is convenient and reliable to characterize microstructural attributes and compute petrophysical properties [11, 142]. The DRP technique applies advanced microscopy imaging facility [48, 12], such as X-ray micro-computed tomography (micro-CT) and focused ion beam scanning electron microscopy (FIB-SEM), to obtain 3D visualization of pore microstructures at the microscale, which provides a new routine to explore microscopic physics of transport phenomena. After the image acquisition, high-fidelity numerical simulations of transport processes can be performed on the digital microstructures to evaluate corresponding transport properties or explore specific physical phenomena [11, 47].

Pore-network modeling (PNM) and direct numerical simulation (DNS) are the two primary pore-scale computing approaches to mimicking transport processes happening inside porous rocks. PNM simplifies the complicated pore space into a topologically representative network of pore bodies interconnected by pore throats with ideal shapes (such as sphere and cylinder), according to specific criteria [486, 47]. The transport behaviors within each network element are described by semi-analytical laws (such as Hagen-Poiseuille law), which makes the computation considerably fast and enables multi-scale modeling to incorporate strong heterogeneity in large volumes. PNM is naturally appropriate for capillary-controlled transport processes. However, the transport property results estimated from PNM are often less accurate, due to the simplification of the complicated pore space. It is still of great challenge to gain reliable results from PNM by identifying the microstructural features

relevant to network modeling and ignoring the remaining parts to simplify the computational complexity [486].

On the contrary, DNS directly discretizes the raw pore space into computing elements by preserving pore geometry (voxels can be used as the computing elements), and then transport equations (such as Navier-Stokes or Laplace equations) are numerically solved or approximated on the computational meshes [11, 48]. Lattice Boltzmann method (LBM), finite element method (FEM) and finite volume method (FVM) are commonly used to approximate or solve transport equations at the pore scale. Generally, DNS can provide direct insight into the impact of pore microstructure on transport properties, but it has severe limitations in computational intensity. The 3D digital microstructure with a large representative size and high resolution usually contains hundreds of millions, even billions of computational elements (or voxels). As a consequence, massively parallel programming, long computing time, high-performance computing (HPC) platform and large data storage are usually required to run such huge numerical simulations [283, 384]. The compute-intensive nature of DNS makes it difficult to accommodate all the details of pore microstructures and involve all the relevant transport physics.

Generally, it is a 'brute-force' way to explore the relationships between microstructural characteristics and transport properties by solely using pore-scale simulations. Currently, many attempts have been made to build the surrogate microstructure-property models through artificial intelligence, in order to rapidly and accurately predict macroscopic properties from the measurable microstructure informatics. Due to the powerful capacities in massive data analysis and hidden rule exploration, machine/deep learning algorithms are becoming more and more popular in this field, especially the convolutional neural network (CNN). Typically, CNN [3] is capable of automatically extracting task-related features from spatial data such as images through its convolution layers, avoiding the manual feature selection procedure, and it has achieved tremendous success in the computer vision field. Therefore, many similar studies have been conducted to construct CNN-based surrogate models for permeability prediction, where the 2D or 3D digital images of porous microstructures are directly used as the input data [419, 479, 424, 227, 436, 196, 434]. Besides, CNN has also been applied to establish the linkages between microstructures and other macroscopic properties/behaviors for various heterogeneous materials, including effective thermal conductivity [468], effective elastic moduli [73, 275], effective diffusivity [478], P/S-wave velocity [230], formation factor [366] and fluid velocity filed [382].

However, several critical issues and challenges have not been properly solved for the CNN-based surrogate modeling of physical properties in porous media research. (1) Dimensional homogeneity problem. In some studies [275, 227, 436], the 'linkages' between digital

images and physical-property variables are directly built through CNN without nondimensionalization process, which is contrary to the principle of dimensional homogeneity. (2) Computational problem. High computational complexity and excessive memory requirement are the inherent defects of the 3D CNN, which strictly limit both the quality and quantity of 3D training images. However, a reliable prediction model needs a large amount of representative elementary volumes (REVs) with high quality to feed it, and such high-intensive computation usually demands an HPC platform. (3) Feature extraction problem. Kernels (convolving windows) are applied across the input image to extract local features, but the internal connection of components as well as the relative spatial relationships are not captured by the convolution layers of CNN. It means that the global features of porous media (such as long-distance connectivity and topological information) that are crucial to transport properties are rarely used for permeability prediction by CNN. (4) Lack of spatial invariance. The internal representation of a pore microstructure in CNN is not independent of view angle, which means the rotation of the input image can affect the prediction result. This issue can be solved through data augmentation, but the computational expense of DNN model training will be dramatically increased. (5) Overfitting problem. The CNN model is easily over-fitted due to a large database for training. (6) Low-level interpretability. The complicated CNN architecture, formed by a deep stack of distinct layers, is often referred to as a 'black box', because it is difficult to understand the underlying mechanics and no inherent way exists to interpret how features influence a particular prediction. (7) Lack of flexibility. Once a CNN model is fitted for physical property prediction, both the size and resolution of input images are fixed, which is very inflexible for the common cases where adjustments of image size or resolution have to be made without losing information.

According to the discussion above, it is clear that the usage of CNN doesn't contribute to any better understandings of transport/mechanical behaviors within pore microstructures because of its poor explainability. In contrast, simple regression algorithms, including linear regression, decision tree, random forest, support vector machine and shallow neural network, are much easier to be interpreted than CNN, which can be used to reveal the underlying mechanisms of the microstructure-property linkages. Additionally, the manual feature-extraction free of CNN doesn't constitute a comparative advantage over the other machine learning algorithms that require predefined feature variables, especially in porous media research. Because various morphological descriptors have already been properly designed in the literature to statistically characterize pore microstructures (as listed in Table 7.1), and they can be used as the input features to construct microstructure-property relationships through machine learning. Compared with the unreadable features extracted by CNN, morphological descriptors characterize porous microstructures from multiple perspectives

with clear physical meanings, and these descriptors can compose the feature pool to construct machine learning-based surrogate permeability models. Feature selection can be conducted to keep the morphological descriptors that are significant to permeability and remove the abundant/irrelevant ones, through which the microstructural complexity is reduced to a limited number of descriptive parameters related to permeability, and then a permeability prediction model with excellent performance can be achieved, as illustrated in Figure 7.1. What's more important is that the dependence of permeability on microstructural characteristics can also be well interpreted through the feature selection process, providing a deep insight into the implicit microstructure-permeability relationship.

Although simple supervised machine learning algorithms were adopted to model effective properties of porous media in previous studies [450, 461, 128, 437, 376], they didn't achieve the purpose to yield insightful understandings of the linkages between macroscopic properties and microstructural characteristics, and the corresponding pore-scale behaviors are still poorly understood. This study distinguishes itself from previous efforts in five aspects: (1) Plenty of digital rocks with diverse morphologies were acquired from the advanced microscopy imaging technique (micro-CT) at high-resolution levels, which are used to construct the predictive model to ensure its generalization ability. (2) A large number of 3D microstructure samples are stochastically reconstructed by preserving statistical equivalence, morphological similarity and transport properties, which are used as the raw data to capture the stochasticity in permeability modeling. (3) A wide variety of morphological descriptors are collected through an extensive literature survey for the purpose of comprehensive characterization of pore microstructures. (4) High-fidelity simulations of pore-scale fluid flow passing through the representative elementary volumes (REV) of digital microstructures are performed on the HPC platform, to calculate permeability values. (5) Feature selection is conducted to gain the optimal feature subset for predictive model construction and enhance the model interpretability.

Basically, a data-driven framework is developed in this chapter to study the dependence of macroscopic property on microstructural characteristics, which is generally applicable to analyze different physical (hydraulic, thermal, electrical, diffusional and mechanical) properties of porous media, although only permeability is studied here. The remaining parts of this chapter is organized as follows: In Section 7.2, the methodology of surrogate permeability modeling through feature selection and machine learning are systematically presented, and the raw datasets are well prepared and organized, including microstructure sample generation, morphological descriptor extraction and permeability evaluation; In Section 7.3, three types of feature selection are briefly introduced, and various feature selection methods are tried to pick out the microstructure characteristics that are significant

to permeability; In Section 7.4, two permeability prediction models are constructed through wrapper type feature selection and shallow neural networks, where the dependence of permeability on microstructure characteristics is deeply explained, and their predictive performances are also compared with that of two analytical/empirical models; The main contributions and key findings are summarized in Section 7.5.

7.2 Methodology and data preparation

In this study, multiple techniques are fused into the data-driven framework to investigate the microstructure-permeability relationship and to construct the predictive model for natural porous rocks, as illustrated in Figure 7.1. There are six components involved in this framework, including digital rock acquirement, stochastic microstructure reconstruction, statistical characterization of pore microstructures, pore-scale simulation, feature selection and machine learning-based surrogate modeling.

7.2.1 Digital rock acquirement

Various types of porous media (mainly porous rocks) are used in this work, to explore the relationship between microstructural characteristics and intrinsic permeability, including sandstones, carbonate rocks, sand packs, synthetic silicas and others. The pore microstructures within them are distinct in terms of pore size, shape, geometry and topology, with a wide porosity range from 6.85% to 50.73%, based on which a generalized prediction model of permeability is expected to achieve with high reliability.

Advanced microscopy imaging facilities, such as the micro-CT scanner, are capable of obtaining deep insight into the opaque porous rocks, to provide 3D digital visualization of internal structures at the micro-scale. The raw (micro-CT) scanning image is usually in a grayscale or color mode with continuous pixel/voxel intensities, as can be seen in Figure 7.2a. It is necessary to convert the raw scanning image from the continuous model to a discrete form that permits quantitative characterization of the porous microstructures and subsequent pore-scale simulation of fluid flow. As shown in Figure 7.2, the raw micro-CT scanning image of a Mt. Simon sandstone sample [243] is taken as the example to illustrate image processing and segmentation, and a segmented image in binary form is finally obtained by separating the pore space from the solid matrix. The segmented image, also called digital microstructure, allows statistical quantification of the pore network, and it also provides high-fidelity meshes for numerical simulations of fluid flow at the pore scale, permitting one

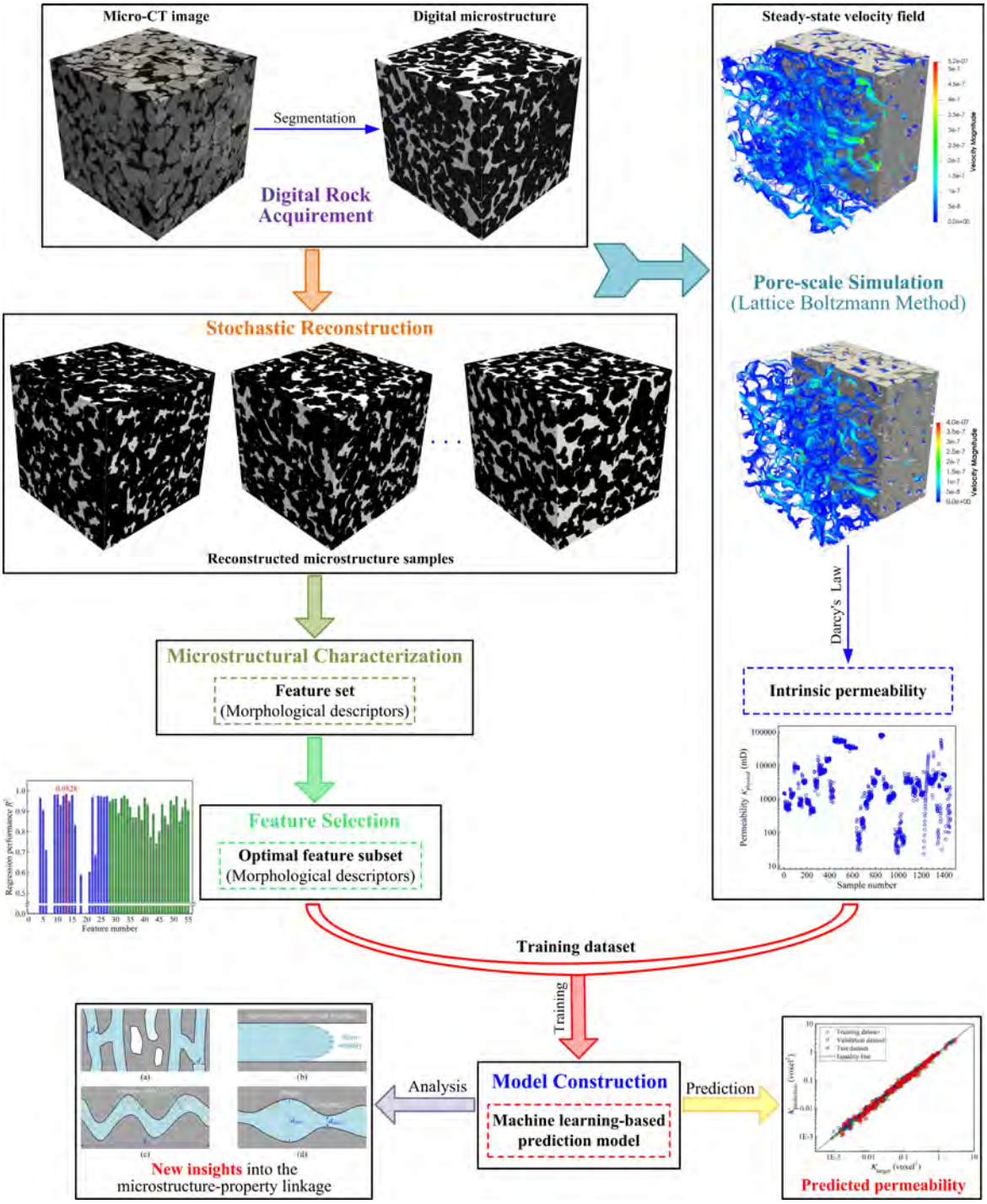


Fig. 7.1 The data-driven framework to investigate the microstructure-property relationship and to construct the permeability prediction model for porous rocks, which primarily contains six modules: (1) Digital rock acquirement, (2) Stochastic microstructure reconstruction, (3) Microstructural characterization, (4) Feature selection, (5) Pore-scale simulation (LBM), and (6) Predictive model construction

to evaluate transport properties. More details on image processing and segmentation can be found in relevant references [388, 206].

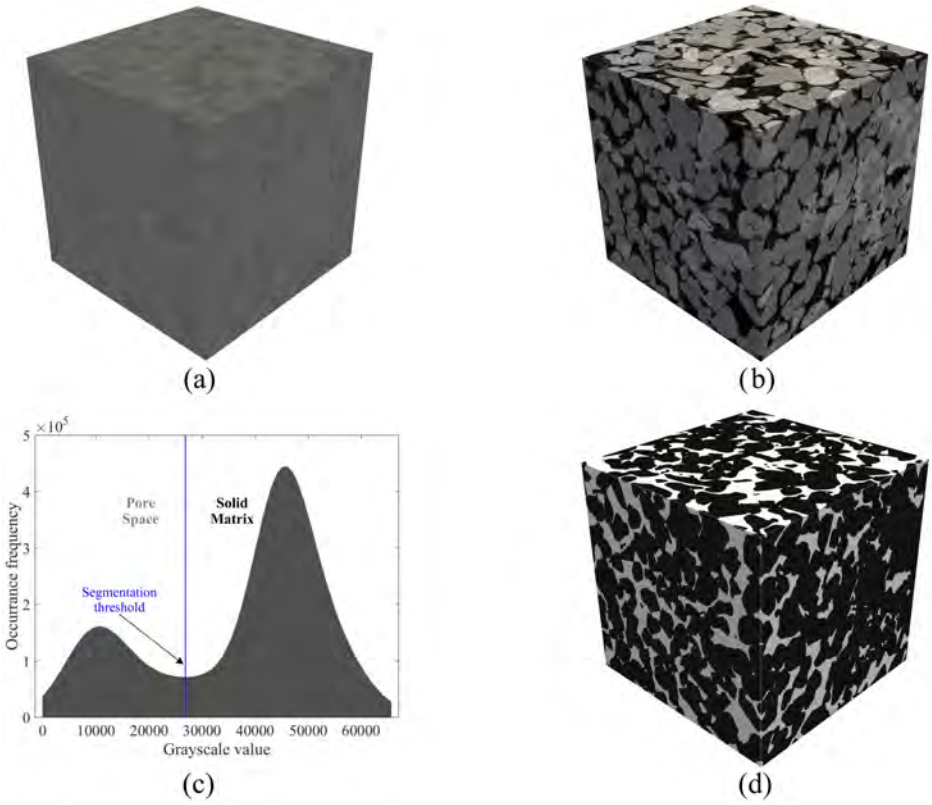
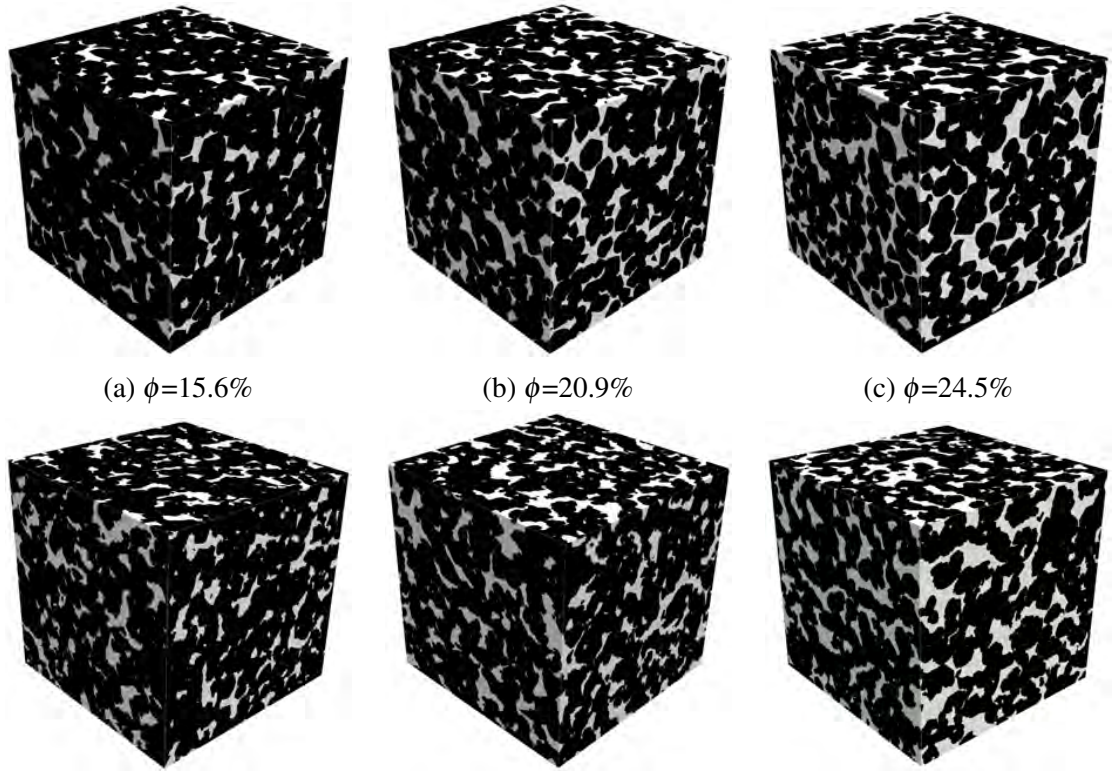


Fig. 7.2 Illustration of image processing and segmentation: (a) The raw micro-CT scanning image of a Mt. Simon sandstone sample (resolution is 2.80 μm, and image size is 480^3 voxels); (b) The grayscale image after denoising and enhancement; (c) The histogram of voxel grayscale value; and (d) The binary image segmented by a global thresholding method (pore space is shown in white, and solid matrix is shown in black)

It should be noted that the micro-CT scanning images used here are collected from several open-access databases, and these images are processed and segmented by using *ImageJ* [207], a popular image processing tool in the DRP community. Due to the high costs of rock core drilling and microscopy imaging, there are only a limited number of 3D digital rock samples available. The number of micro-CT scanning images is 185, covering 37 types of porous media with distinct morphological features, and parts of them are shown in Figure 7.8 and 7.9.

7.2.2 Stochastic microstructure reconstruction

The transport properties of porous media usually exhibit strong uncertainty, due to the random distribution of pore bodies. The limited number of digital rocks obtained from micro-CT scanners are far from adequate to involve all possible morphology configurations of pore microstructures, much less to reveal the stochastic nature of porous rocks. In general, the complete computational dataset [141] is an ensemble of representative/statistical volume elements that cover all morphological possibilities and share the same averaged characteristics, based on which a generalized prediction model can be achieved with high reliability.



(a) ϕ=15.6% (b) ϕ=20.9% (c) ϕ=24.5%

(d) ϕ=15.6% (e) ϕ=20.9% (f) ϕ=24.5%

Fig. 7.3 Stochastic microstructure reconstruction (image size: 320^3 voxels): (a), (b) and (c) are the original microstructures of Fontainebleau sandstones with different porosities ϕ; (d), (e) and (f) are representative reconstruction results

Stochastic microstructure reconstruction [495, 335] is an effective and economical approach to generating statistically equivalent samples of porous media, and the numerous reconstructed samples can be used to investigate the microstructure-property relationships when the availability of real porous media samples is limited. Here, a novel reconstruction method developed in Chapter 4 [141] is applied to generate 3D pore microstructure samples,

and the reconstructed samples have been proved to be able to preserve statistical equivalence, morphological similarity, long-distance connectivity and transport properties. As can be seen in Figure 7.3, Fontainebleau sandstone samples with different porosities are taken as examples to illustrate stochastic microstructure reconstruction using the novel method. This method is used to generate 1270 pore microstructure samples (with image size varying from 200^3 to 320^3 voxels), guided by the morphological information extracted from the real digital microstructures. The real digital rocks together with the reconstructed microstructure samples compose the raw dataset for subsequent analyses, which contains 1455 samples in total.

7.2.3 Pore microstructure characterization

The intrinsic permeability of a porous medium only depends on its microstructural characteristics, therefore, it is an essential prerequisite to quantitatively characterizing the pore microstructure in an explicit expression, so as to derive an interpretable prediction model. The pore space inside natural porous rock usually exhibits great disorder and strong randomness, which needs to be quantitatively characterized in statistical terms. Through statistical characterization [443], the microstructural complexity of a porous medium is reduced to a small set of morphological descriptors related to macroscopic physical properties. A broad range of microstructure characterization approaches have been developed for pore media, including geometrical measurement, topological representation, fractal analysis and others. As listed in Table 7.1, a wide variety of morphological descriptors have been collected through an extensive literature study in Chapter 2, and they will be used as the microstructural features for permeability prediction. About half of these descriptors were used to model the transport properties of porous media in previous studies, and corresponding permeability prediction models are also provided in Table 7.1.

Various types of morphological descriptors characterize pore microstructures from different perspectives, and they can be roughly categorized into four levels. Porosity and specific surface area are the typical descriptors at the first level to simply represent the global/mean properties of porous microstructures via single numbers, but they ignore the detailed morphological features of pore networks that may have significant effects on transport processes. When it comes to the second level, local or size-dependent features are measured by morphological descriptors such as local porosity distribution, coarseness, local percolation probabilities, and lacunarity. As to the third level, geometric attributes of porous media are quantified from various aspects such as pore shape, pore size and surface roughness. The frequently-used descriptors include pore/throat size distribution, mean curvature, chord length distribution, lineal path function, and spatial correlations functions. The fourth level

Table 7.1 The commonly used morphological descriptors for statistical microstructure characterization and permeability (κ) modeling

Morphological descriptor (Determination method)	Data dimension	Empirical or analytical model	Representative reference
D1: Absolute porosity ϕ	1	$\kappa = \frac{\xi^2 D_p^2 \phi^3}{180(1-\phi)^2}$	[71]
D2: Connected porosity ϕ_c	1	$\kappa \approx \phi_c d_t^2$	[152]
D3: Specific surface area S	1	$\kappa = \phi^2/(cS^2F)$	[278]
D4: Integral of mean curvature m_2	1	–	[267]
D5: Integral of total curvature m_3	1	–	[460]
D6: Geometrical tortuosity τ_g (Direct shortest path search method)	1	$\kappa = \phi^3/(cS^2\tau_g^2)$	[246] and [74]
D7: Geometrical tortuosity τ_g (Skeleton shortest path search method)	1	$\kappa = \phi^3/(cS^2\tau_g^2)$	[142] and [428]
D8: Constriction factor β (Mercury intrusion porosimetry simulation)	1	$\kappa = L_h^2\phi_c\beta/(8\tau_g^2)$	[195] and [34]
D9: Constriction factor β (Morphological opening method)	1	$\kappa = L_h^2\phi_c\beta/(8\tau_g^2)$	[34] and [121]
D10: Mean chord length $\langle z \rangle$	1	$\kappa = \frac{\phi\langle z\rangle^2}{32\tau_g}\left(\frac{\langle z^2\rangle}{2\langle z\rangle^2} + \sum\langle\cos\vartheta_m\rangle\right)$	[92] and [39]
D11: Average pore coordination number η	1	$\kappa = 2\pi w d_h^4(\eta - 1.5)^b/l^2$	[197]
D12: Average pore size d (Continuous method)	1	$\kappa = cd^2/F$	[321]
D13: Average pore size d (Discrete method)	1	$\kappa = c\phi d^2$	[194]
D14: Average pore size d (Morphological opening method)	1	$\kappa = cd^2/F$	[349] and [121]
D15: Average pore size d (Random point method)	1	$\kappa = 0.8391G^{-1.3334}\phi^2 d^2$	[441]
D16: Average pore size d (Skeleton method)	1	$\kappa = \phi^{5.5}d^2/5.6$	[115]
D17: Average pore throat size d_t	1	$\kappa = d_t^2/(226F)$	[278]
D18: Effective pore size d_c	1	$\kappa = d_c^2\phi/((1-\phi)^2 c\tau_g^2)$	[46]
D19: Hydraulic pore diameter d_h	1	$\kappa = d_h^2\phi/(c\tau_g^2)$	[29]
D20: Characteristic length I l_a	1	–	[92]
D21: Characteristic length II l_b	1	$\kappa \leq 2l_b^2/(3(1-\phi)^2)$	[92]
D22: Characteristic length III l_c	1	–	[92]
D23: Characteristic length IV l_d	1	$\ln(\kappa) = a + b\ln(\phi) + c\ln(l_d)$	[208]
D24: Average connectivity distance l_e	1	–	[242]
D25: Characteristic length V $L^\star$	1	–	[186]
D26: Fractal dimension α	1	$\kappa = \pi\alpha d_{max}^4/(128A(4-\alpha))$	[496]
D27: Succolarity ψ	1	$\kappa = 0.0239e^{23.98\psi}$	[485]
D28: Lacunarity $\delta(\varepsilon)$	10	–	[327]
D29: Chord length distribution $\rho(z)$	70	–	[319]
D30: Lineal path function $L(z)$	50	–	[186]
D31: Spherical contact distribution function $H_S(d)$	16	–	[267]
D32: 1st Minkowski function $m_0(d)$	16	–	[460]
D33: 2nd Minkowski function $m_1(d)$	16	–	[18]
D34: 3rd Minkowski function $m_2(d)$	16	–	[460]
D35: 4th Minkowski function $m_3(d)$	16	–	[18]
D36: Two-point correlation function $S_2(r)$	50	–	[46]
D37: Two-point cluster correlation function $C_2(r)$	50	–	[216]
D38: Normalized auto-covariance function $R(r)$	50	–	[33]
D39: Pair connectivity function $H(r)$	50	–	[242]
D40: Surface-surface correlation function $F_{SS}(r)$	20	–	[377]
D41: Surface-void correlation function $F_{SV}(r)$	50	–	[378]
D42: Local porosity distribution $\mu(\widetilde{\phi}, L = 50)$	100	–	[44]
D43: Local porosity distribution $\mu(\widetilde{\phi}, L = L^\star)$	100	–	[186]
D44: Local percolation probabilities $\lambda(\widetilde{\phi}, L = 50)$	50	–	[101]
D45: Local percolation probabilities $\lambda(\widetilde{\phi}, L = L^\star)$	50	–	[101]
D46: Total fraction of percolating cells $T_1(L)$	75	–	[262]
D47: Total fraction of percolating cells $T_3(L)$	75	–	[186]
D48: Pore coordination number distribution $O(\eta)$	20	–	[197]
D49: Pore size distribution $p(d)$ (Continuous method)	50	–	[321]
D50: Pore size distribution $p(d)$ (Discrete method)	30	–	[194]
D51: Pore size distribution $p(d)$ (Morphological opening method)	20	–	[121]
D52: Pore size distribution $p(d)$ (Random point method)	30	–	[441]
D53: Pore size distribution $p(d)$ (Skeleton method)	30	–	[115]
D54: Pore throat size distribution $p(d_t)$	15	–	[281]
D55: Coarseness $C(L)$	100	–	[364]

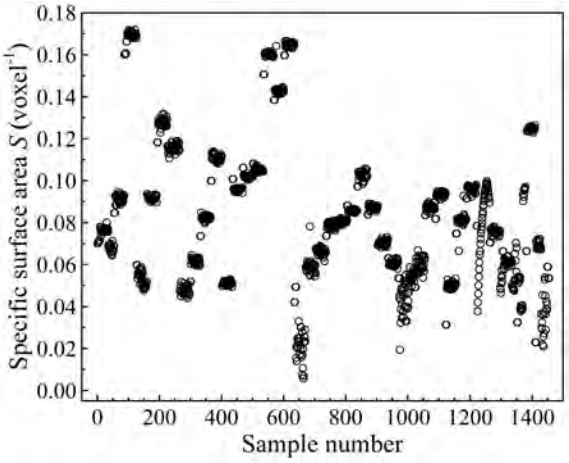
focuses on the topological characteristics of pore microstructures, which is related to long-distance connectivity and percolation of pore networks. Total curvature (Euler characteristic), two-point cluster correlation function, pair connectivity function, total fraction of percolating cells and succolarity are commonly-used indexes of connectivity. Pore coordination number represents the number of adjacent pore bodies connected to a specific pore. Besides, geometrical tortuosity characterizes the sinuosity and complexity of percolation paths inside porous media, while constriction factor quantitatively represents cross-sectional variation along pore channels.

All of the morphological descriptors listed in Table 7.1 are extracted from the digital microstructure dataset with 1455 samples, and they are going to be the possible predictor variables to feed the machine learning model for permeability evaluation, as illustrated in Figure 7.1. The results of 12 representative descriptors are given in Figure 7.4 and 7.5. Some morphological descriptors have multiple definitions, leading to various determination methods, so different determination methods are used here to achieve a microstructure characterization as comprehensively as possible.

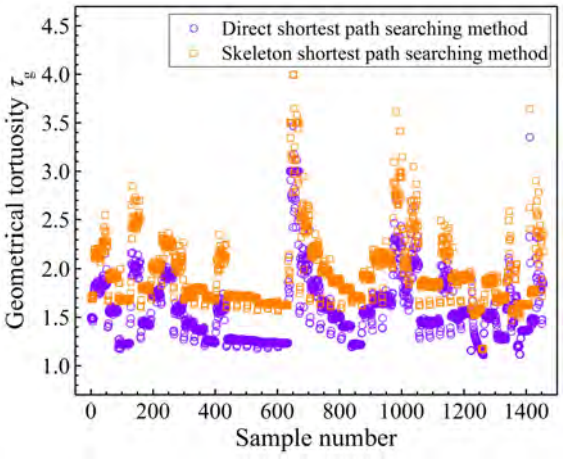
It should be noted that image resolutions of the digital microstructure samples are all around 5 μm, which means the microstructural analyses here are conducted in voxel domains with different length scales. For the dimensionless descriptors, such as porosity, geometrical tortuosity, constriction factor, and poor coordination number, no additional data process is required. As to the descriptors with length dimension, such as specific surface area, mean curvature, average pore size and characteristic length, they are all quantified by using voxel as the basic length unit here, instead of converting them into physical length scale. Because it is easy to achieve the seamless combination between the morphological descriptors in the voxel length unit and the LBM permeability in the lattice length unit, just by setting the lattice length equal to the voxel size for each porous media sample.

7.2.4 Permeability evaluation via lattice Boltzmann simulation

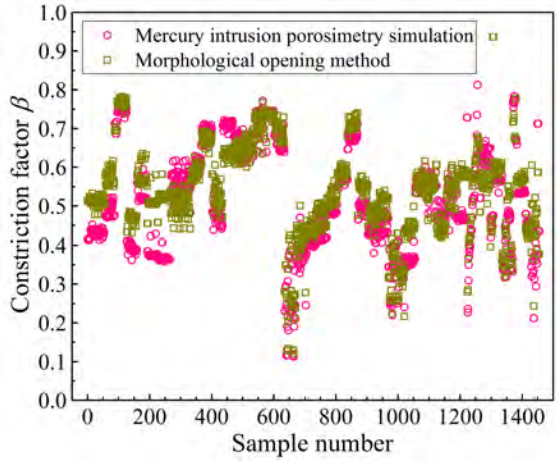
Fore pore-scale simulation of fluid flow, lattice Boltzmann method (LBM) is more rigorous than pore network modeling (PNM), and the former can provide more reliable permeability evaluations for porous media samples [486]. Besides, the lattice Boltzmann simulation is directly performed on voxel domain of the digital microstructure without simplification, and the computed permeability values in the lattice unit can be directly linked to the morphological descriptors in the voxel unit, avoiding additional data conversion/processing. Therefore, LBM is adopted in this work to evaluate the intrinsic permeabilities of the set of digital microstructure samples. Further details about the numerical methodology of LBM can be



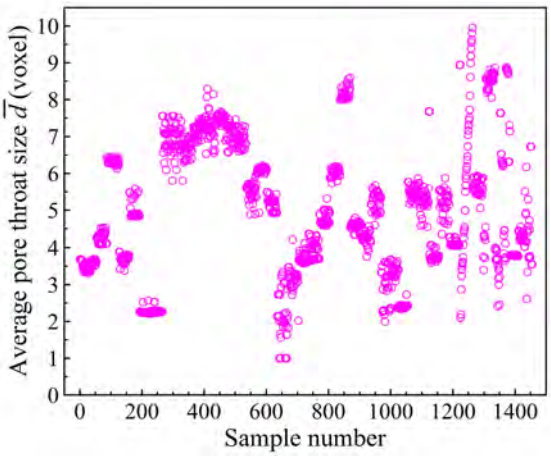
(a) Specific surface area



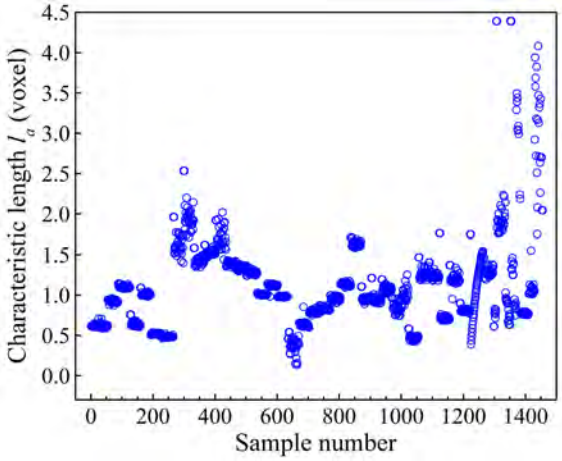
(b) Geometrical tortuosity



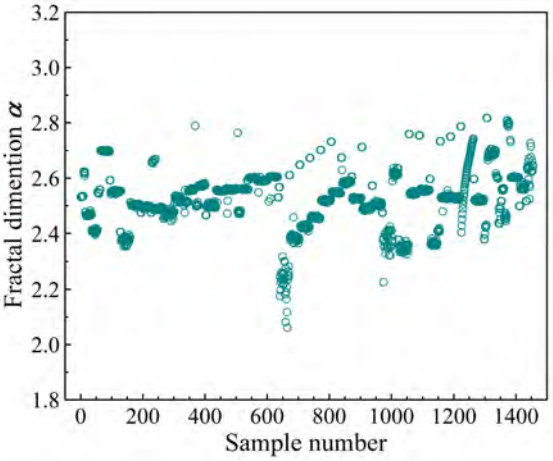
(c) Constriction factor



(d) Average pore throat size

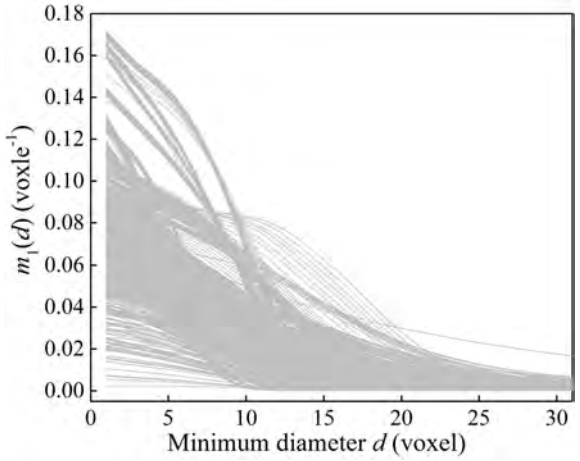


(e) Characteristic length

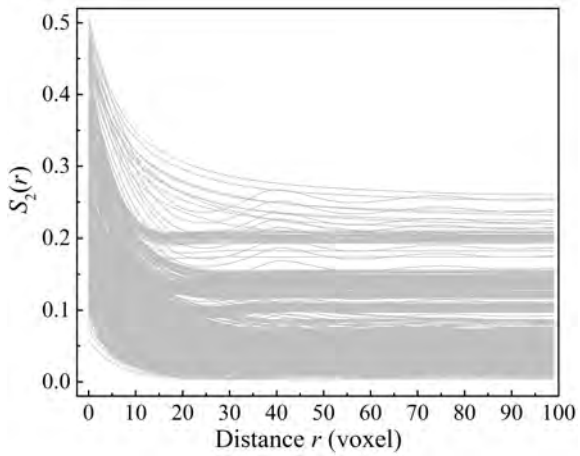


(f) Fractal dimension

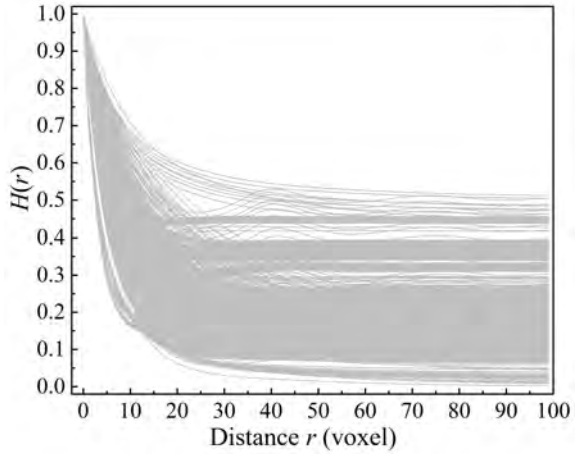
Fig. 7.4 Representative morphological descriptors extracted from the dataset with 1455 microstructure samples



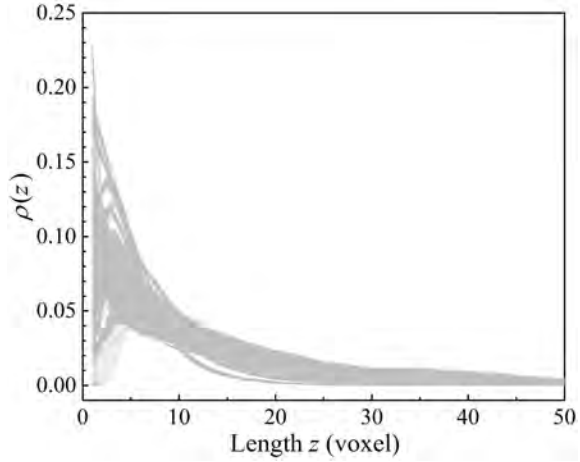
(a) 2nd Minkowski function



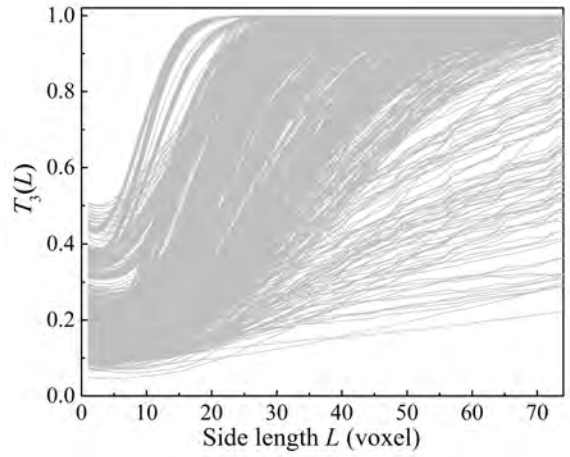
(b) Two-point correlation function



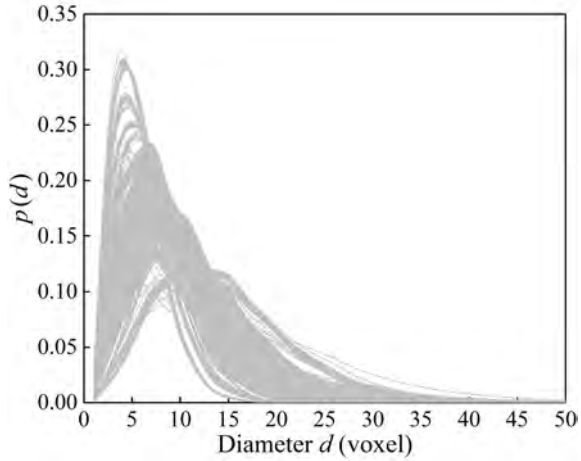
(c) Pair connectivity function



(d) Chord length distribution



(e) Total fraction of percolating cells



(f) Pore size distribution

Fig. 7.5 Representative morphological descriptors extracted from the dataset with 1455 microstructure samples

found in the relevant literature [252], and the procedures to evaluate intrinsic permeabilities of porous media through lattice Boltzmann method are well explained in Chapter 6.

Driven by a constant pressure difference between the inlet and outlet faces, LBM is performed on the cubic digital rock sample, to simulate a single-phase fluid flow with a low Reynolds number ($Re \ll 1$) passing through it. When the fluid flow reaches a steady state, it can be described by Darcy's law, and the intrinsic permeability κ of the porous medium is quantified by the following equation:

$$\kappa = -\frac{\mu}{\nabla p}\langle \mathbf{u} \rangle \tag{7.1}$$

where ∇p is the pressure gradient along the direction of macroscopic fluid flow, μ is the dynamic viscosity of the fluid, and $\langle \mathbf{u} \rangle$ denotes the volume-averaged fluid velocity across the simulation domain. The Mt. Simon sandstone sample in Figure 7.2 is taken as the example to illustrate the lattice Boltzmann simulation, as shown in Figure 7.6.

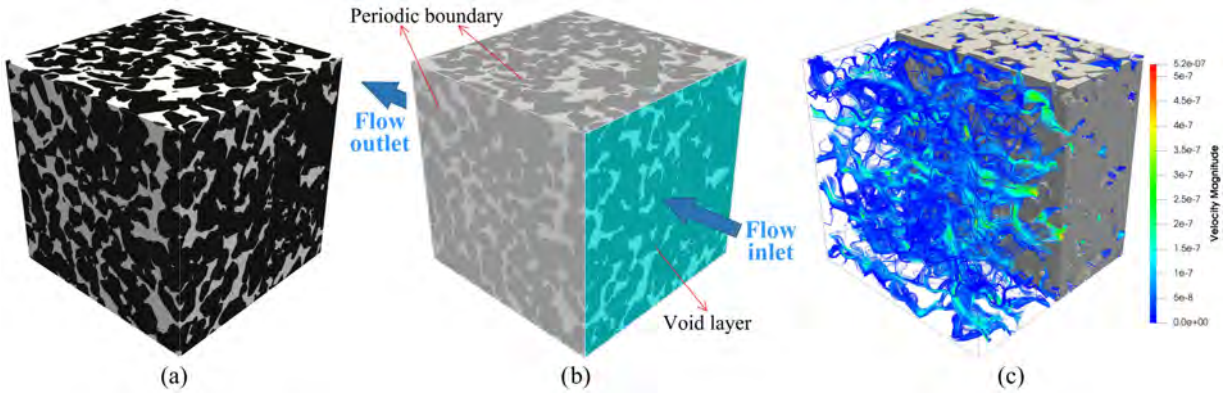


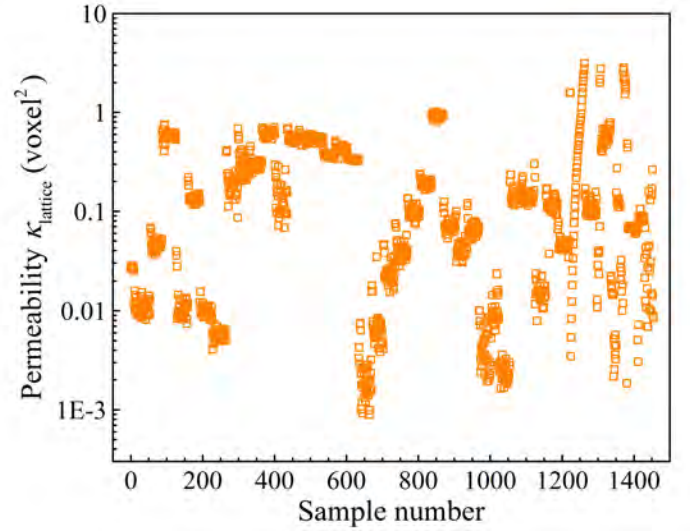
Fig. 7.6 Evaluation of intrinsic permeability through lattice Boltzmann simulation: (a) The 3D digital microstructure of a Mt. Simon sandstone sample; (b) The boundary conditions; (c) The steady-state fluid velocity field inside the porous medium

The permeability value computed from lattice Boltzmann simulation is in dimensionless lattice unit, which can be converted to the physical unit via the following equation [425]:

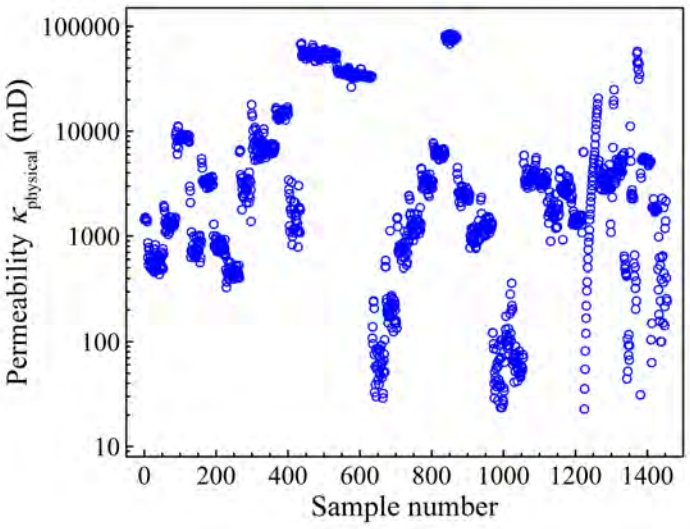
$$\kappa_{\text{physical}} = \kappa_{\text{lattice}} \left(\frac{L_{\text{physical}}}{L_{\text{lattice}}} \right)^2 \tag{7.2}$$

where κ_{physical} and κ_{lattice} are the permeability values in physical and lattice unit, respectively; and L_{physical} and L_{lattice} are the lengths of any identical feature in the physical sample and the LBM domain, respectively. Since the numerical grid of LBM is coincident with the voxel grid in this study, so the value of $\frac{L_{\text{physical}}}{L_{\text{lattice}}}$ is equal to the image resolution.

For each microstructure sample, lattice Boltzmann simulations of fluid flow are conducted along with three-axial directions, and the average value of three directional permeabilities is used to investigate the microstructure-permeability relationship. As shown in Figure 7.7, the permeability values of the digital microstructure dataset with 1455 samples are given both in the lattice and physical unit, and the permeability values span in a broad range over 4 orders of magnitude. As mentioned above, the permeability values in the lattice unit will be used as the response variable to fit the machine learning-based prediction model (as illustrated in Figure 7.1).



(a) Permeability values in the lattice unit



(b) Permeability values in the physical unit

Fig. 7.7 The intrinsic permeability results of the set of 1455 microstructure samples evaluated from lattice Boltzmann simulations

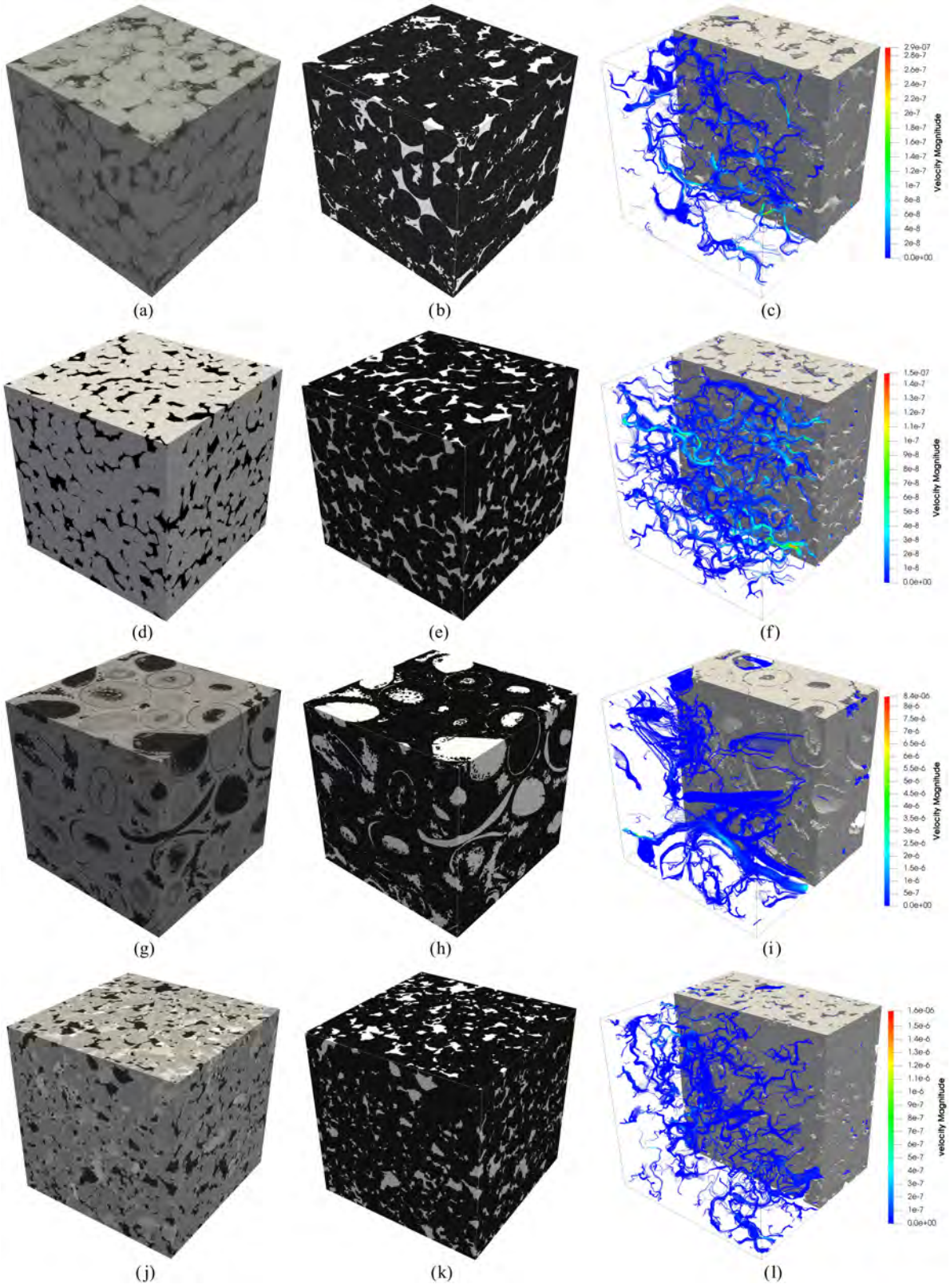


Fig. 7.8 Digital rock samples and lattice Boltzmann simulations: The micro-CT scanning images of (a) Ketton carbonate, (d) Fontainebleau sandstone, (g) Savonnières carbonate, and (j) Leopard sandstone; (b), (e), (h) and (k) are the segmented images (digital pore microstructures); (c), (f), (i) and (l) are the flow velocity fields at steady states

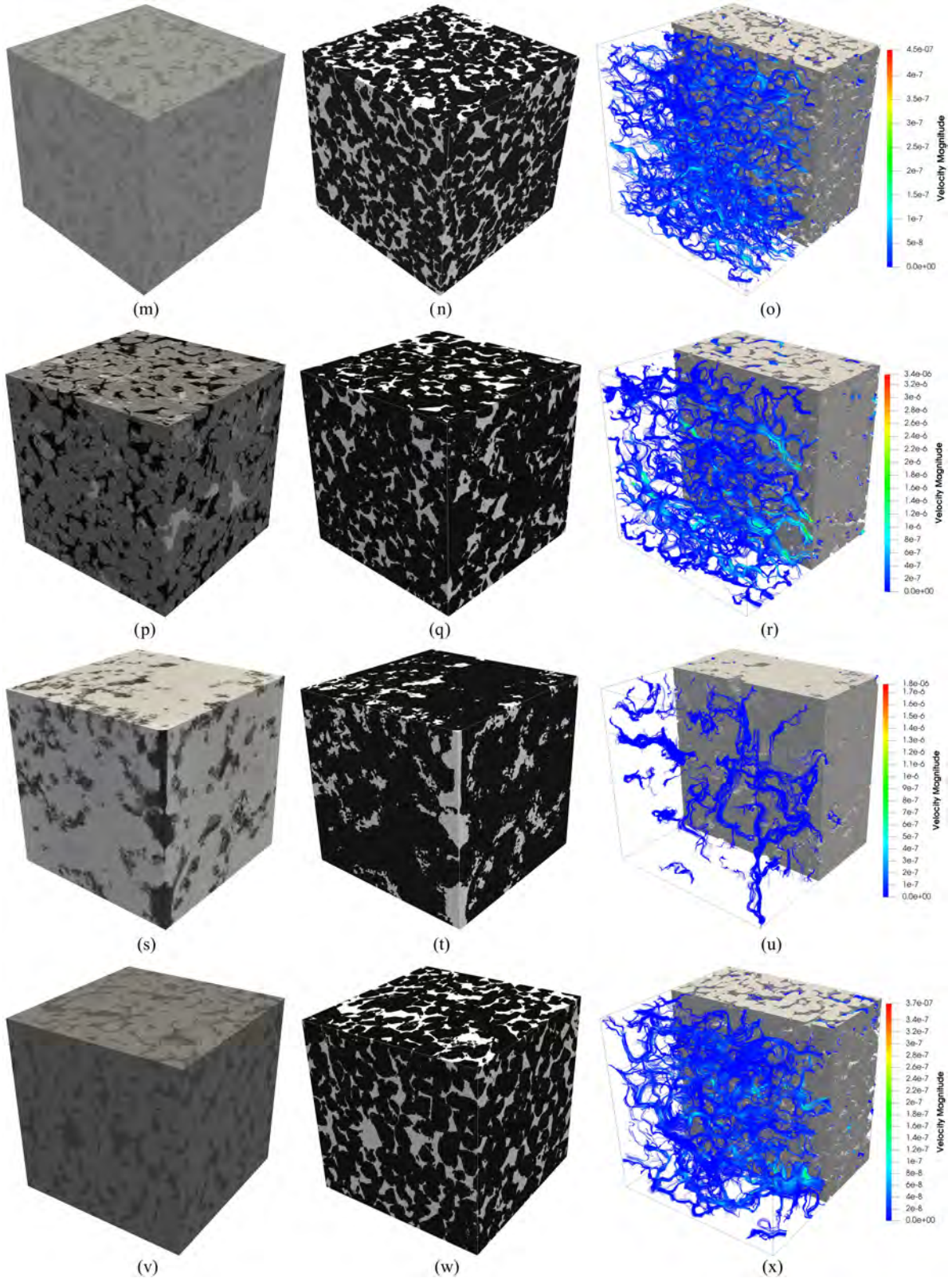


Fig. 7.9 Digital rock samples and lattice Boltzmann simulations: The micro-CT scanning images of (m) Bentheimer sandstone, (p) Berea sandstone, (s) Massangis Jaune carbonate, and (v) Gildehauser sandstone; (n), (q), (t) and (w) are the segmented images; (o), (r), (u) and (x) are the flow velocity fields at steady states

7.3 Feature selection

It is well-known that the intrinsic permeability of a porous medium only depends on its microstructural characteristics [29, 142], but some morphological descriptors (in Table 7.1) statistically characterizing microstructural complexity may be irrelevant or less significant to permeability. The unnecessary involvement of irrelevant and abundant features can increase the model complexity and make it harder to interpret. On the other hand, a dumb regression model can consequently be constructed with lower accuracy and reliability in prediction. Therefore, feature selection is an indispensable data preprocessing for predictive model construction, where the most relevant and significant features are picked out from a large set of morphological descriptors in Table 7.1.

For a dataset of m observations $\{\mathbf{x}_i, \kappa_i\}$ $(i = 1, \ldots m)$ consisting of n input feature variables $x_{i,j}$ $(j = 1, \ldots n)$ and a output permeability value κ_i, various methods can be applied to pick out the feature variables $x_{i,j}$ that are important to the response κ_i. Generally, feature selection techniques [168, 270] can be divided into three categories from the perspective of selection strategy, which are the filter, wrapper and embedded methods. Among them, the filter type feature selection is independent of the training algorithm, while the wrapper and embedded methods interact with a particular learning process. All of these methods are tried here to undertake the feature selection task due to their wide applicability.

The objective of feature selection in this work is threefold: (1) Enhancing interpretability of the implicit regression model to obtain deep insights into the underlying dependence of permeability on microstructural characteristics; (2) Reducing the computational complexity and avoiding overfitting to built a cost-effective predictor by using the selected subset of features; (3) Achieving a generalized and sensible model with the optimal performance in permeability prediction.

7.3.1 Filter type feature selection

Filter type feature selection [168, 270] assesses feature importance according to certain data characteristics, so it is unrelated to any learning algorithms. Typically, a filter method consists of two steps: feature importance ranking and feature filtering. Different feature evaluation criteria have been proposed to rank feature importance, such as feature correlation, mutual information, the feature discriminative ability, the feature ability to maintain the data manifold, and the feature capacity to reconstruct the raw data. Four representative criteria of feature importance evaluation are briefly introduced as follows:

(i) Correlation coefficient

$$R\left[\mathbf{x}_j\right] = \frac{\sum_{i=1}^{m}(x_{i,j}-\overline{x_j})(\kappa_i-\overline{\kappa})}{\sqrt{\sum_{i=1}^{m}(x_{i,j}-\overline{x_j})^2\sum_{i=1}^{m}(\kappa_i-\overline{\kappa})^2}} \tag{7.3}$$

where the bar notation stands for an average over the index j. The value of $R\left[\mathbf{x}_j\right]$ is between -1 and 1, and a value close to -1 or 1 indicates strongly linear dependency between a feature variable and the target. In additional to Pearson's correlation, other similar criteria such as Kendall's and Spearman's correlation can also be used.

(ii) RReliefF importance weight
The RReliefF algorithm [375] evaluates the feature importance weight $W\left[\mathbf{x}_j\right]$ according to the capacity of the feature vector $\mathbf{x}_j$ to distinguish the response values. It punishes the feature for not distinguishing different response values and rewards the features for not separating similar response values, providing the following probabilities:

$$W\left[\mathbf{x}_j\right] = P(\text{similar } \mathbf{x}_j \mid \text{dissimilar } \kappa) - P(\text{similar } \mathbf{x}_j \mid \text{dissimilar } \kappa) \tag{7.4}$$

The RReliefF algorithm is capable of revealing the conditional dependencies between features, from which a unified evaluation of feature importance to response values can be obtained.

(iii) F-test importance score
F-test [22] is a type of statistical test using the F-statics to test the null hypotheses H, which can be used to examine the significance of each feature. An F-statistic F is the ratio between the variance of two models, given by

$$F = -\frac{\text{Between-group varaince}}{\text{Within-group varaince}} \tag{7.5}$$

which tests the null hypothesis (that the responses value grouped by feature variables are extracted from populations sharing the same mean) with respected to the alternative hypothesis (that these populations do not share the same mean).

The p-value of the F-statistics is used to indicate feature importance, providing the largest probability under the null hypothesis H:

$$p = \Pr(F \geq t \mid H) \tag{7.6}$$

where t is the observed value. The feature importance score is given in the form of $-\log(p)$, where a lager score means the corresponding feature is significant to the response.

(iv) Nearest neighbor-based feature weight
Neighborhood component analysis (NCA) [491] selects features through optimizing the prediction accuracy of nearest neighbor regression, and the corresponding objective function for minimization is expressed as follows:

$$f(\mathbf{w}) = \frac{1}{m}\sum_{i=1}^{m} l_i + \lambda \sum_{j=1}^{n} w_j^2 \tag{7.7}$$

where $\mathbf{w}$ is the weighting vector indicating feature importance, l_i is a loss function measuring the disagreement between the target κ and prediction $\hat{\kappa}$, the regularization term $\lambda \sum_{j=1}^{n} w_j^2$ is to alleviate overfitting, and the regularization parameter λ can be tuned via cross validation.

7.3.2 Embedded type feature selection

Embedded methods [168, 270] conduct feature selection during the process of model fitting, which are deeply embedded in specific learning algorithms such as decision tree, random forest and Gaussian process regression.

(i) Decision tree
During the training process of a decision tree model [289], feature importance is evaluated from the sum of changes in the mean squared error (MSE) due to splits on each feature and the number of branch nodes. This sum is taken over all splits at each branch node, where the MSE is estimated as node error weighted by node probability. And then, feature importance associated with the corresponding split is calculated as the difference between the MSE for the parent node and the total MSE for children nodes.

(ii) Random forest
Random forest [57] is a predictive model composed of a weighted combination of multiple decision trees, where feature importance can be evaluated by permutation to measure the influence degree of a feature variable in predicting the response. If a feature is influential to the response, then permuting its values will affect the prediction error; while if a feature is uninfluential, then permuting its values has an ignorable

effect on model prediction. Feature importance is computed as the difference of model prediction errors before and after variable permutation.

(iii) Gaussian process regression
Gaussian process regression (GPR) [395] is a non-parametric Bayesian approach towards regression problems, where feature importance can be evaluated from corresponding separate length scales of the kernel function. In the case of the squared exponential kernel function, a large value of the length scale indicates the feature variable is not influential on the model response. The feature weights are computed as the exponential of the negative length scales. The values of normalized feature weights are constrained between 0 and 1, and the irrelevant features have weights close to 0.

7.3.3 Wrapper type feature selection

Compared with the above-introduced filter and embedded methods, wrapper methods are applicable to heterogeneous feature selection. Considering the dimension difference between morphological descriptors listed in Table 7.1, wrapper type feature selection is worth trying. Wrapper methods [244, 168, 169] assess the quality of feature selection according to the predictive performance of the predefined learning algorithm. It searches the optimal feature subset through evaluating the possible combinations of features greedily based on a certain evaluation criterion. For regression problems, the coefficient of determination R^2, the root-mean-squared error (RMSE) and the mean-absolute error (MAE) can be used as the metrics to evaluate the model performance, which are mathematically expressed as following equations respectively:

$$R^2 = 1 - \frac{\sum_{i=1}^{m}(\kappa_i - \hat{\kappa}_i)^2}{\sum_{i=1}^{m}(\kappa_i - \overline{\kappa})^2} \tag{7.8}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\kappa_i - \hat{\kappa}_i)^2} \tag{7.9}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}|\kappa_i - \hat{\kappa}_i| \tag{7.10}$$

where κ_i and $\hat{\kappa}_i$ are the target and predicted permeability respectively corresponding to the i-th sample, and $\overline{\kappa}$ is the average value of the target permeability. R^2 quantifies the degree to which the feature variables explain the variation of the response, and its value ranges from 0 to 1, where a larger value indicates a better model performance.

Exhaustive search is a 'brute-force' strategy in wrapper type feature selection, which usually requires enormous amounts of computation, especially when the number of feature

variables is large. By contrast, greedy search strategies are of lower computation cost, which can be further divided into two categories: forward selection and backward elimination. Here, the wrapper method with sequential forward adding strategy [160] is adopted, as explained in Figure 7.10. Starting with a null model, each morphological descriptor in Table 7.1 is used individually to fit the machine-learning model, and the descriptor that achieves the best predictive performance (the maximum R^2 value) is picked out as the first selected feature. A new machine-learning model with two features is then fitted by sequentially combining the previously selected feature with one of the remaining descriptors, and the descriptor resulting in the largest R^2 is selected as the second feature. The above procedure is repeated iteratively until no improvement of predictive performance or reaching the desired number of included features, and a subset of features is consequently selected through this performance-driven process.

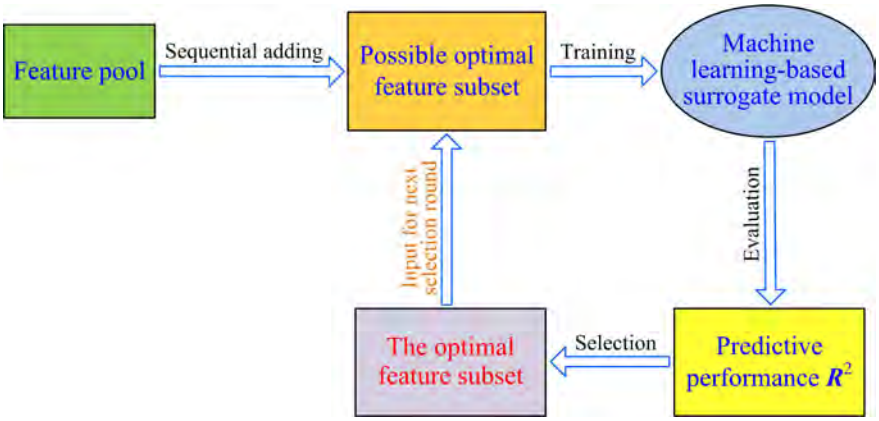
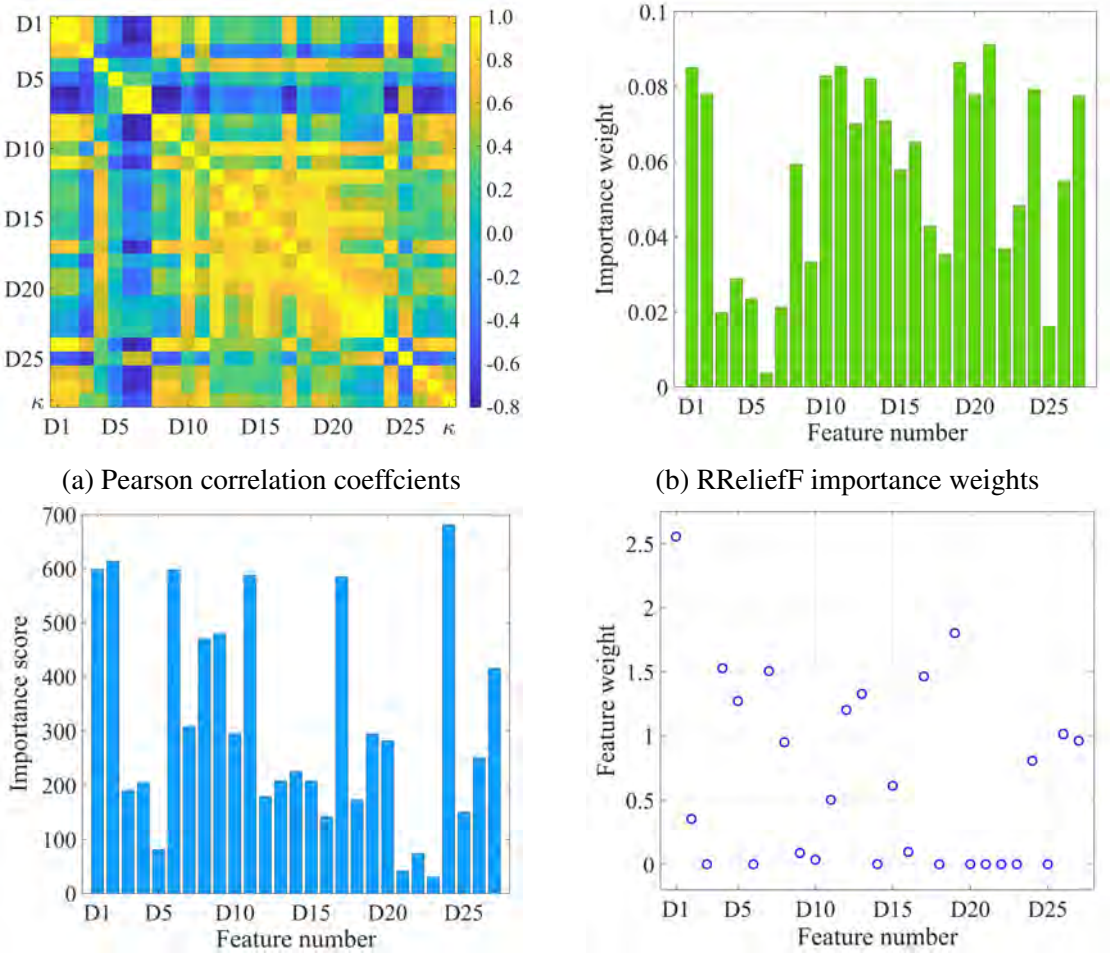


Fig. 7.10 The flow chart of the wrapper method to select features through a sequential forward adding strategy

7.3.4 Feature selection results

As listed in Table 7.1, the first 27 morphological descriptors are in the format of a single number, while the remaining 28 descriptors are in the form of distributions with different dimensions. Generally, the filter and embedded methods are not applicable to feature selection with heterogeneous data, and the wrapper type feature selection possesses strong versatility. Considering the above situation, feature selection is firstly performed on the first 27 morphological descriptors by using different methods, and later the wrapper method is applied to select features from the entire feature pool with 55 descriptors in Section 7.4.

The results of feature importance ranking are estimated for the first 27 descriptors by using four different filter methods, as shown in Figure 7.11. Due to the difference of feature



(a) Pearson correlation coeffcients

(b) RReliefF importance weights

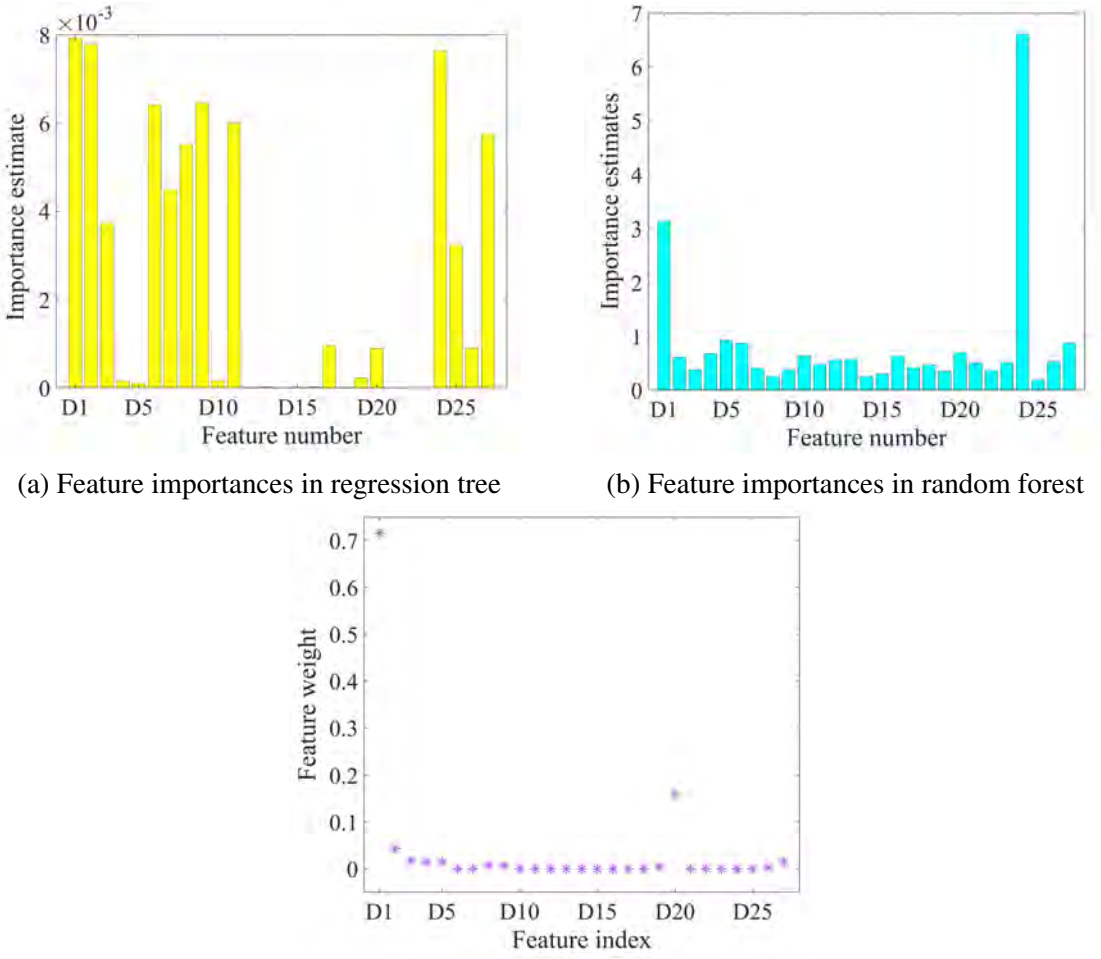
(c) F-test importance scores

(d) NCA feature weights

Fig. 7.11 The feature importance ranking results estimated from four filter methods

evaluation criteria, the importance ranking results estimated from these methods are not entirely consistent with each other, but the overall assessment results are similar. In all of the four filter methods, porosity (D1) is identified as one of the most important microstructure characteristics to permeability (κ), which is in accord with common understandings of porous media. However, specific surface area (D3) is evaluated to be an insignificant/irrelevant feature, which is contrary to the basic view that specific surface area is critical to permeability of porous media. Filter methods evaluate the importance of feature variables individually, but a feature variable that is recognized to be unimportant by itself can be significant to the response when used with other features [168]. Filter methods are unable to detect the joint importance of multivariable features, which is one of the main drawbacks.

Three embedded methods are used to assess the importance weights of the first 27 morphological descriptors to permeability, and corresponding results are plotted in Figure



(a) Feature importances in regression tree

(b) Feature importances in random forest

(c) Feature weights in GPR

Fig. 7.12 The feature importance ranking results estimated from three embedded methods

7.12. Because different learning algorithms are embedded in these three feature selection processes, the importance ranking orders of morphological descriptors are inconsistent. Similar to the results of filter methods, porosity (D1) is also selected as an important microstructure feature by embedded methods, but specific surface area (D3) is assigned with low importance scores, especially in random forest and GPR models. Besides, embedded type feature selection is associated with specific learning algorithms, which is inflexible for prediction model construction.

The intended purpose of feature selection has not been achieved by using the filter or embedded methods. On the one hand, not all of the morphological descriptors in Table 7.1 are covered; On the other hand, feature importance is not reasonably assigned to some particular descriptors. This task will be continuously undertaken through wrapper type feature selection. Because the wrapper method and the choice of learning algorithms are interrelated to each

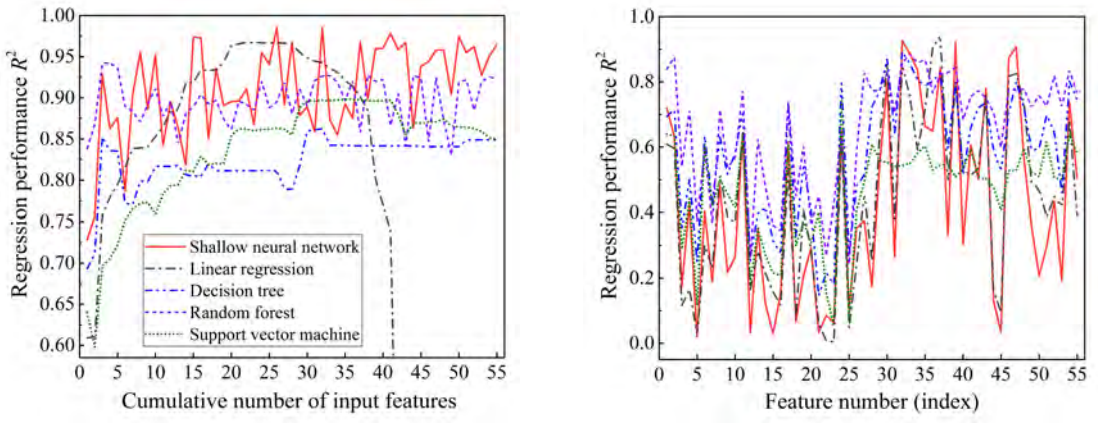
other, the feature selection results are put together with prediction model construction in the next section.

7.4 Results and discussion

7.4.1 Choice of learning algorithm

The methodology of permeability prediction has been illustrated in Figure 7.1, where the microstructure-permeability relationship is implicitly built in a supervised learning manner. Although deep learning algorithms such as CNN have an acceptable performance on surrogate modeling, they are usually too complicated to provide a clear interpretation of the underlying linkage between microstructure characteristics and permeability. Generally, prediction accuracy increases with model complexity, but at the expense of explainability [253]. Besides, blind pursuit of model accuracy also tends to overfit the training data, lowering the prediction performance on unseen data.

There are a number of alternative regression algorithms which have better explainability, such as linear regression [314], decision tree [289], random forest [57], support vector machine [421] and shallow neural network [494]. To identify the important features from the 55 morphological descriptors via wrapper type feature selection, the chosen learning algorithm should be able to deal with high-dimensional data, thereby avoiding the influence of model incapability on large data processing. Besides, the model responses of the chosen learning algorithm should also be sensitive to influential feature variables, so as to capture feature significance or insignificance.



(a) Cumulative features (b) Individual feature

Fig. 7.13 Comparison of regression performances based on different learning algorithms

By using the entire data with 1455 observations, the model regression performances (measured by R^2) corresponding to 5 learning algorithms are preliminarily investigated, and the results are plotted in Figure 7.13. Firstly, regression models are fitted by gradually increasing the number of input features, where all of the morphological descriptors in Table 7.1 are sequentially added to the feature set. Secondly, each descriptor is used as an individual feature to feed the regression model.

As can be seen in Figure 7.13a, the shallow neural network model is not only able to cope with high-dimensional data, but it can also achieve the best regression performance when suitable features are given. The other four models are inferior to the shallow neural network in terms of regression performance, among which linear regression is failed to model high-dimensional data and support vector machine requires many input features to reach an acceptable performance. In Figure 7.13a and b, the curves of regression performance corresponding to shallow neural network fluctuate strongly, which indicates the model response is sensitive to input feature variables. According to the above comparative analysis, the shallow neural network is chosen as the supervised learning algorithm to undertake the feature selection task through the wrapper method. The selected features are then used to implicitly establish the microstructure-permeability linkage for prediction purpose. What's more, a good explanation of how microstructural characteristics determine the intrinsic permeability of a porous medium can be obtained as well.

7.4.2 Prediction model construction

A shallow neuron network model with two hidden layers is used to explore the dependency of permeability on microstructural characteristics, where and the numbers of neurons in the 1st and 2nd hidden layers are 40 and 20 respectively. More detailed information on the shallow neuron network model is provided in the Appendices of this chapter.

After data normalization, the entire dataset is randomly split into three subsets: training (60%), validation (20%) and testing (20%). The training dataset is used to fit the neuron network, where network parameters are adjusted to minimize regression error. The validation dataset is used to measure network generalization, and the training process is halted when generalization stops improving, so as to avoid overfitting. The testing dataset is used to provide an independent measure of the model performance on new observations. Due to the variation of initial conditions (such as initial network parameters and data divisions), training the neural network model multiple times will generate different results. Therefore, in each case of inputted features, the neural network model is are fitted for 50 times, and the average output is used to evaluate the predictive performance during the process of wrapper type feature selection.

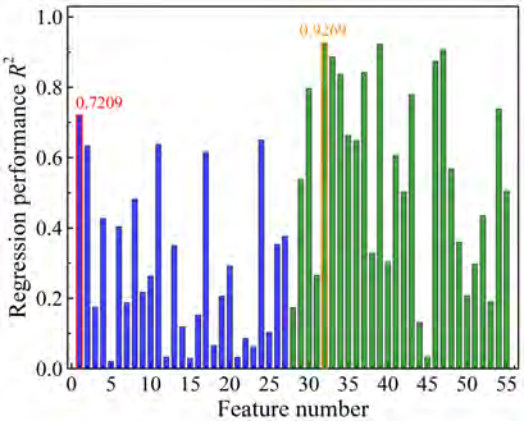


Fig. 7.14 Regression performances of the shallow neuron network models trained by using individual features

Figure 7.13b is redrawn into a bar graph, as shown in Figure 7.14. The first 27 morphological descriptors in Table 7.1 are in the format of a single number, which are shown as blue bars, and we call them blue descriptors for convenience. As to the remaining 28 morphological descriptors, they are in the format of distribution, which are called green descriptors here. Roughly, individual utilization of a green descriptor can yield better predictive performance than a blue descriptor, because the former contains more predictor variables than the latter. However, most of the morphological descriptors yield poor permeability predictions ($R^2 < 0.8$) when individually using them, which may imply morphological descriptors should be used together for better characterization of porous microstructures.

As can be observed in Figure 7.13a, the regression performance improving with the number of used features is the main trend, but some morphological descriptors play negative roles in permeability prediction, leading to local fluctuations of the performance curve. Therefore, feature selection is a necessary operation to obtain the optimal subset of morphological descriptors that are relevant and significant to permeability. Besides, the maximum value of R^2 is 0.9812, and it measures the predictive performance of the neural network model trained by using the first 26 descriptors in Table 7.1, which may suggest an excellent prediction model can be achieved by only using the blue descriptors.

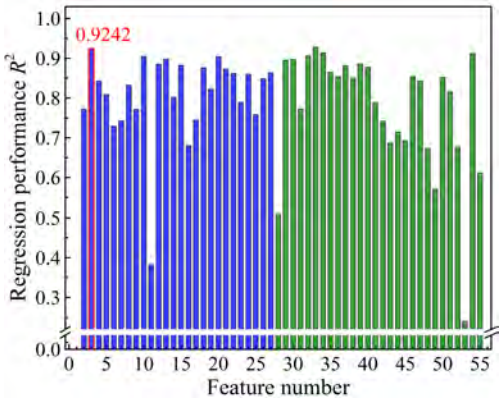
Permeability prediction model I

In this subsection, feature selection is restricted to the blue descriptors by using the wrapper method, aiming to build a cost-effective prediction model with fewer input variables. According to Figure 7.14a, absolute porosity (D1) is selected in the first step, because it yields a neural network model with the best predictive performance ($R^2 = 0.7209$) among

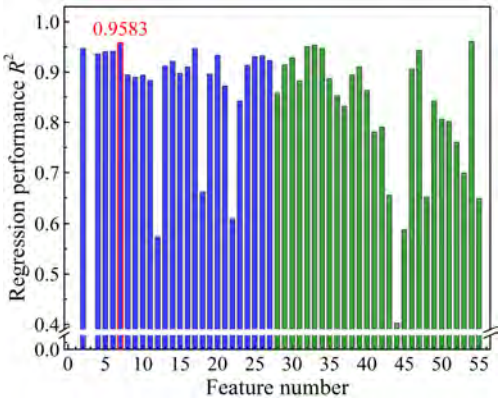
the blue descriptors. The importance of D1 to permeability is identified by all types of feature selection, including filter, embedded and wrapper methods. In the second step, D1 is combined with the remaining descriptors to fit neural network models, and the predictive performances of these models are remarkably improved, as shown in Figure 7.15a. The maximum value of R^2 is 0.9242, and the corresponding descriptor D3 (specific surface area) is selected as the second feature. In contrast to filter and embedded methods, the significance of specific surface area to permeability can be well recognized through wrapper type feature selection. Repeating the above procedures, morphological descriptors D7, D20, D2, D8, D17, D19, D13 and D26 are successively picked out, as exhibited in Figure 7.15b–i.

The effect of feature selection on performance improvement is significant, where predictive performance increases continuously when more selected morphological descriptors are fed to the neural network model. When the first 6 selected descriptors are used for model training, its predictive performance is comparable to another model fitted by using 26 descriptors, which can be seen by comparing Figure 7.15e with Figure 7.13a. When 8 selected descriptors are used to train the neural network model, its performance reaches the peak, where $R^2 = 0.9937$. Continuing to add more descriptors for model training, the performance of the fitted neural network model starts to decline. Therefore, the optimal feature subset, obtained from wrapper type feature selection, contains 8 morphological descriptors, which are absolute porosity, geometrical tortuosity, characteristic length I, connected porosity, constriction factor, average pore throat size and hydraulic pore diameter.

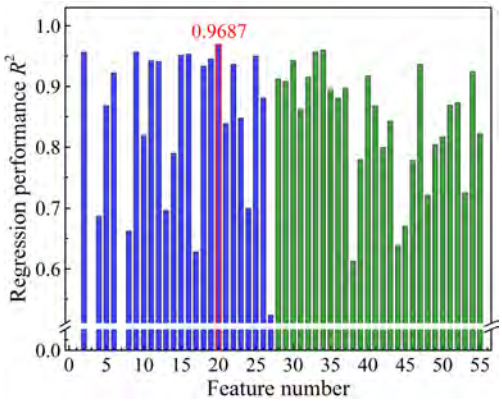
The above feature selection result is thought to be reasonable, because the selected morphological descriptors had been directly used to build analytical/empirical models for permeability prediction in previous researches, as demonstrated in Table 7.1. For example, the first 3 selected descriptors (porosity, specific surface area and geometrical tortuosity) are involved in the well-known Kozeny-Carmon relation [90, 142]. Different from the filter or embedded methods that one-sidedly analyze the relationship between an individual descriptor and permeability, the wrapper method selects features through a multi-variable analytical methodology.



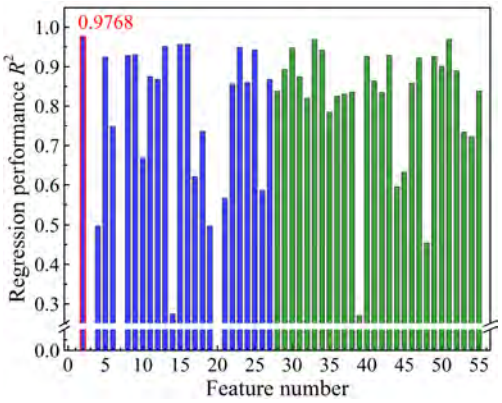
(a) D3: Specific surface area S



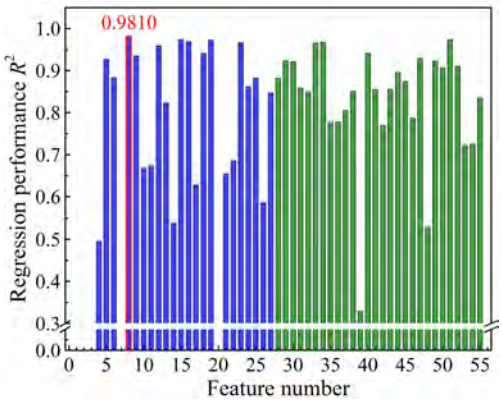
(b) D7: Geometrical tortuosity τ_g



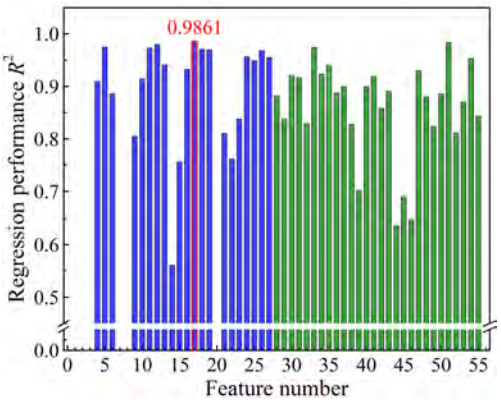
(c) D20: Characteristic length I l_a



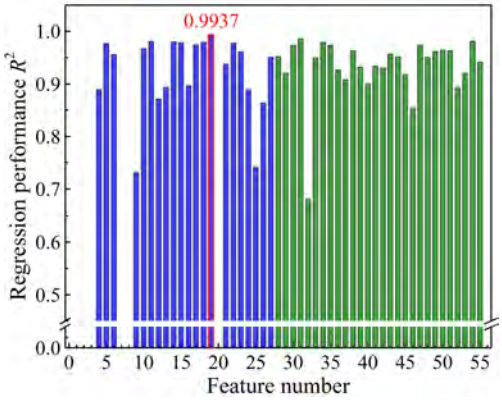
(d) D2: Connected porosity ϕ_c



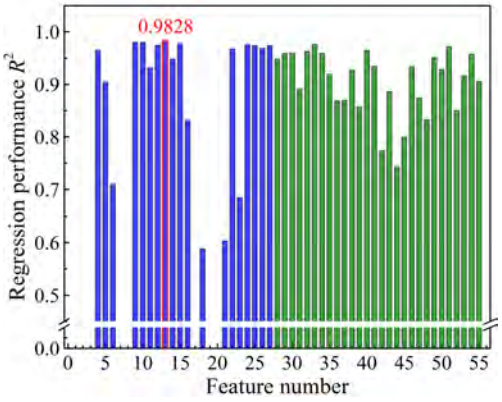
(e) D8: Constriction factor β



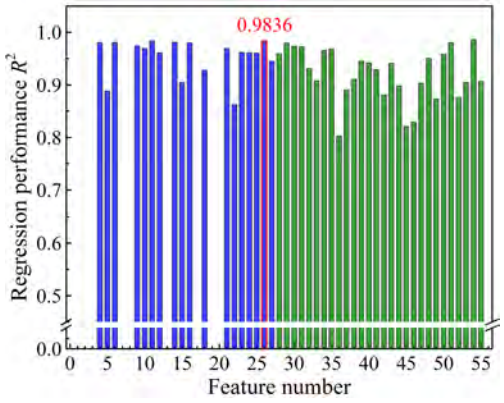
(f) D17: Average pore throat size d_t



(g) D19: Hydraulic pore diameter d_h



(h) D13: Average pore size d



(i) D26: Fractal dimension α

Fig. 7.15 The results of feature selection in each step via the wrapper method

Essentially, the 8 selected morphological descriptors quantitatively characterize pore microstructures from five perspectives, as illustrated in Figure 7.16, and the dependence of permeability on microstructural characteristics can be interpreted as follows : (1) Absolute and connected porosity represent the entire pore space and the permeable portion permitting fluid to flow through respectively; (2) Specific surface area approximately reflects the area of the fluid-solid interface that provides adhesive friction to fluid flow; (3) Characteristic length, hydraulic pore diameter and pore throat size quantify the distances between opposite walls of pore channels allowing fluid to pass through; (4) Geometrical tortuosity measures the sinuosity of percolating pore paths that extends the average length of flow streamlines; (5) Constriction factor characterizes the cross-sectional variation along pore channels which leads to convergence and divergence of streamlines.

Based on the optimal set of selected features, the neural network model is fitted for prediction purpose, and the results are shown in Figure 7.17. The model exhibits high generalization capability to predict permeability spanning 4 orders of magnitude with an average error of 6.92% for natural reservoir rocks. Compared to the popular PNM who has an average error ranging from 33.00% to 40.00% in permeability evaluation [28, 486], this machine learning-based prediction model has much better performances in terms of accuracy and efficiency.

Specifically, the relative errors of the fitted neural network model in permeability prediction are within $\pm$10.00% for 81.17% of the porous microstructure samples, and the error magnitude is less than 20.00% for 95.05% of the used samples. Less than 5.00% of the samples have relatively larger prediction errors between $\pm$20.00% and $\pm$40.00%, due to the extreme complexities of these porous microstructures. For some special porous media, their complicated microstructures exhibit strong randomness and heterogeneity, which cause great

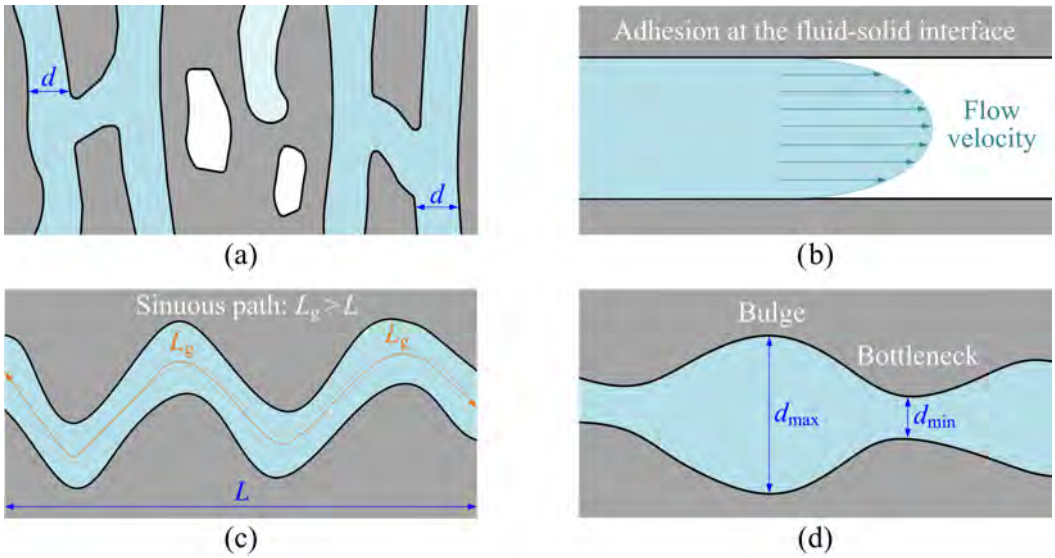


Fig. 7.16 Graphic illustrations of microstructural characteristics: (a) Interconnected, dead-end and isolated pores, and the distances between opposite solid walls; (b) Adhesion at the fluid-solid interface; (c) Sinuous fluid flow path; and (d) Cross-section variation along pore channel

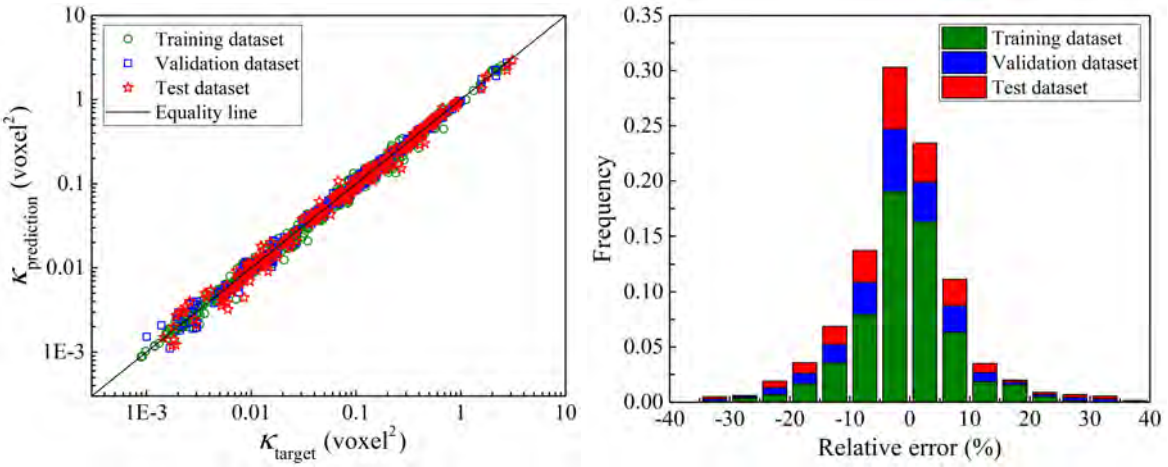


Fig. 7.17 Performance of prediction model I: (1) Comparison between the target and predicted permeability values; (a) The relative error distribution of permeability prediction

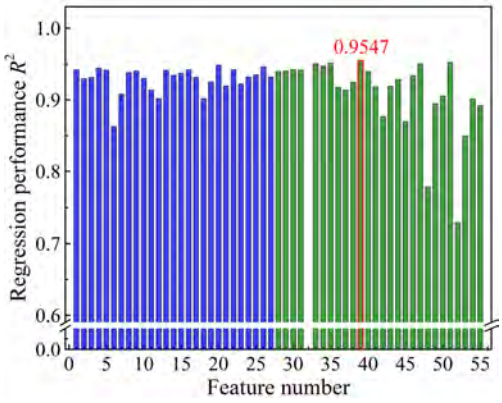
difficulties to completely capture the microstructural characteristics related to permeability. Considering the diversity of porous media samples used in this work that possesses a super large variation of permeability, it is understandable that a generalized prediction model shows a certain of incompleteness for some individual samples but has acceptable accuracy.

Permeability prediction model II

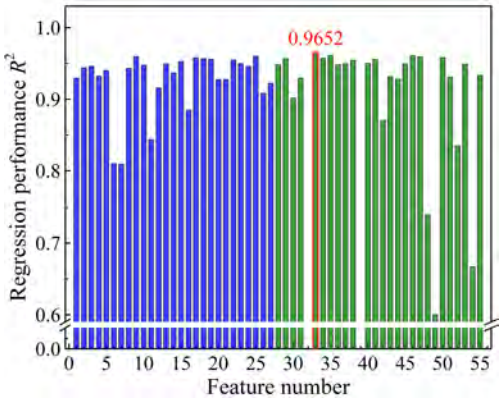
In this subsection, microstructural characteristics are selected from the feature pool containing both blue and green descriptors, to built another neural network model for permeability prediction. According to Figure 7.14, D32 (1st Minkowski function) is firstly selected because it yields the prediction model with the best performance ($R^2 = 0.9269$). Subsequently, D39, D33, D19, D8, D7, D17, D18, D35 and D54 are picked out one by one through the wrapper method of feature selection, as demonstrated in Figure 7.18. The peak of predictive performance ($R^2 = 0.9881$) is achieved by using the first 9 selected morphological descriptors to train the neural network model, and the selected feature set contains 1st Minkowski function, pair connectivity function, 2nd Minkowski function, hydraulic pore diameter, constriction factor, geometrical tortuosity, average pore throat size, effective pore size, and 4th Minkowski function.

Comparing the 1st and 2nd feature selection results, 4 morphological descriptors (hydraulic pore diameter, constriction factor, geometrical tortuosity, average pore throat size) are found in common, as can be seen in Figure 7.15 and 7.18. As to the remaining selected descriptors, some of them are also highly related. For example, porosity and specific surface area are fully contained in the 1st and 2nd Minkowski function respectively, which can be considered as a single point on a curve. Basically, the two sets of selected features share similar morphological information, but the main difference between them lies in the pair connectivity function and 4th Minkowski function. Both of these two descriptors are important indicators of pore connectivity that affect the spatial distribution of fluid flow inside porous media.

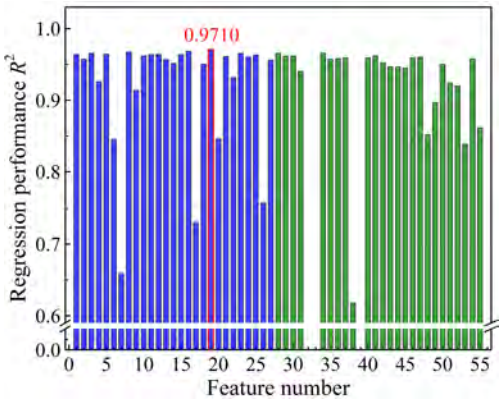
During the second process of feature selection, the predictive performance starts from a higher point ($R^2 = 0.9269$), but it increases slowly as more selected descriptors are added, and R^2 reaches a lower peak value of 0.9881, as can be seen in Figure 7.18. The permeability prediction results of the neural network model trained by the optimal set of descriptors are given in Figure 7.19. The average error of permeability prediction from this model is 7.76%, which is slightly larger than that of the first prediction model who has an average error of 6.92%. To be specific, the second permeability prediction model has errors within $\pm 10.00\%$ for 75.26% of the porous media samples used in this study, and the ratio increases to 93.13% for samples whose permeability values can be reliably evaluated with errors between -20.00% and 20.00%. As to the remaining samples (less than 7.00%), the predicted permeabilities of them have relatively large errors ranging from $\pm 20.00\%$ and $\pm 40.00\%$, but still acceptable.



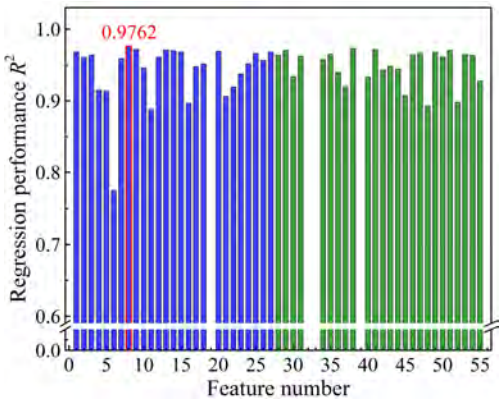
(a) D39: Pair connectivity function $H(r)$



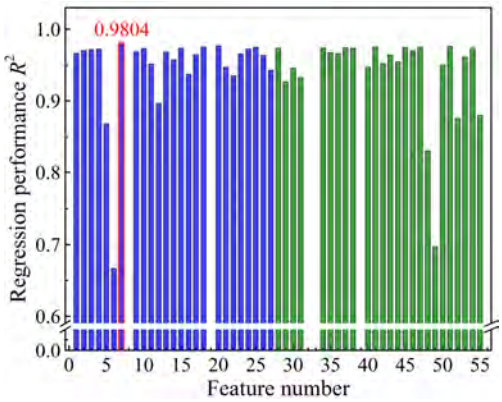
(b) D33: 2nd Minkowski function $m_1(d)$



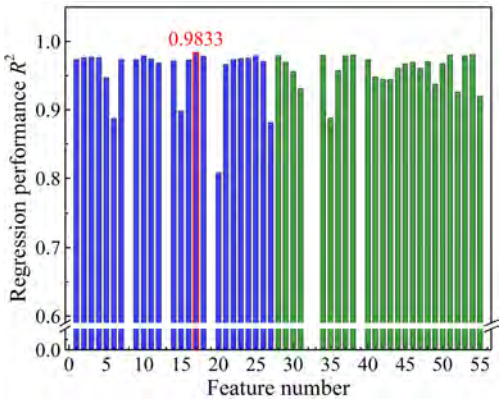
(c) D19: Hydraulic pore diameter d_{h}



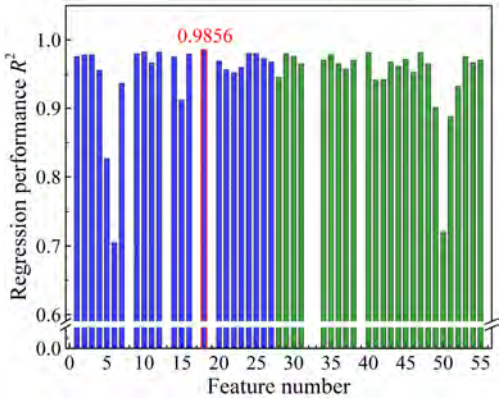
(d) D8: Constriction factor β



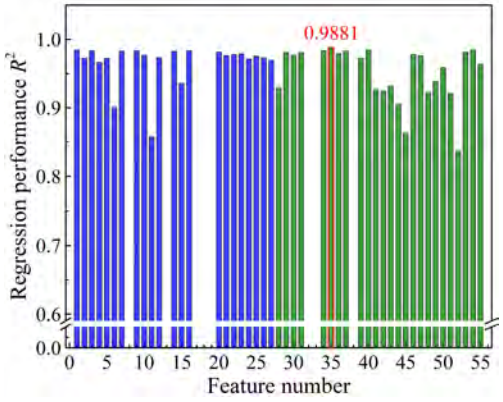
(e) D7: Geometrical tortuosity τ_{g}



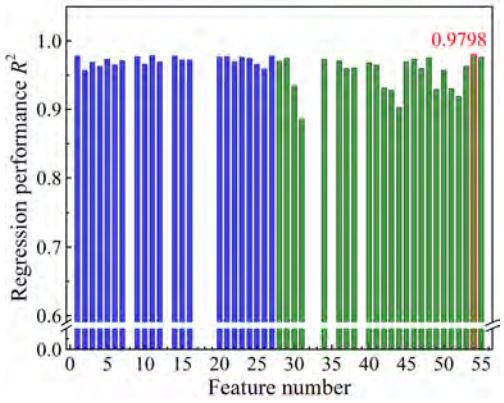
(f) D17: Average pore throat size d_{t}



(g) D18: Effective pore size d_{c}

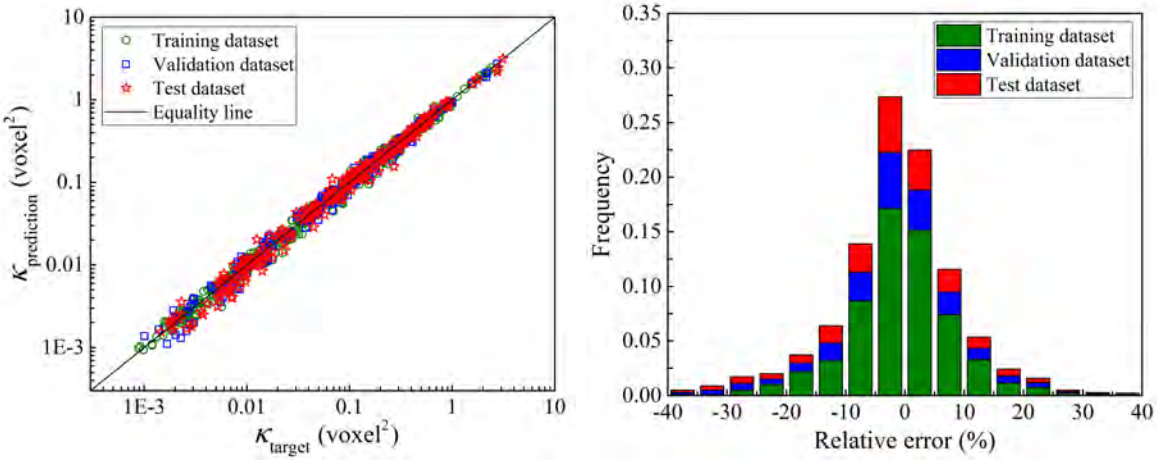


(h) D35: 4th Minkowski function $m_3(d)$



(i) D54: Pore throat size distribution $p(d_t)$

Fig. 7.18 The results of feature selection in each step via wrapper method



(a) (b)

Fig. 7.19 Performance of the prediction model II: (1) Comparison between the target and predicted permeability values; (a) The relative error distribution of permeability prediction

The second prediction model is slightly inferior to the first model in terms of prediction accuracy, although the feature set with many more variables (103 variables) are fed to the second prediction model, compared with the feature set with 8 variable input to the first model. It may imply that the blue descriptors are more suitable to model permeability compared with the green descriptors, because the formers contain more morphological information that is significant to permeability. It should be noted that the objective of this study is not just to provide reliable models for permeability prediction, but also to highlight the morphological descriptors that are most relevant to permeability through the feature selection processing performed on a fairly simple neural network model. Without doubt, the predictive accuracy

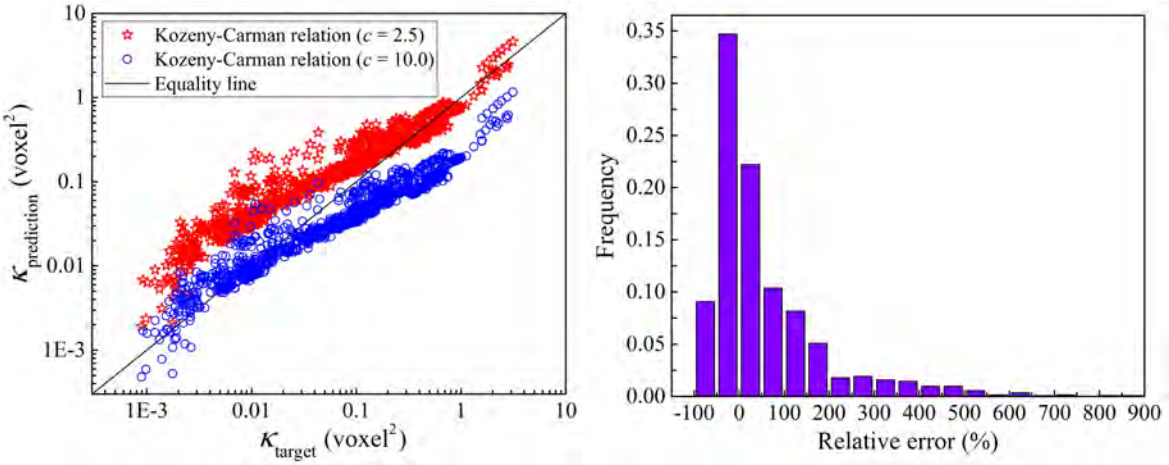
can be improved by fitting a more complicated and powerful model, but the risk of overfitting will be increased as well, and the generalization performance can thus be declined.

7.4.3 Comparisons

In this subsection, the permeability prediction models established through feature selection and machine learning are compared with two popular analytical/empirical models, to highlight the advantage of the proposed data-driven framework in transport property analysis. As discussed in the introduction section, the semi-empirical Kozeny-Carman relation [90] is one of the best-known models to estimate permeability, given by:

$$\kappa = \frac{\phi^3}{cS^2\tau_g^2} \tag{7.11}$$

where c is a dimensionless parameter called shape factor. Generally, shape factor c is considered as a constant, varying with the types of porous media, but it is determined empirically [489]. Its value is around 2.5 for beds packed with spherical particles [233]. Due to microstructural complexities, the shape factors should be larger than 2.5 for natural porous rocks. Here, 2.5 and 10.0 are taken as the upper and lower limits of shape factor receptively, to estimate the permeability values of the set of porous media samples, and the results are shown in Figure 7.20.



(a) Permeability estimation results (b) The relative error distribution

Fig. 7.20 Performance of Kozeny-Carman relation in permeability estimation

There are three common morphological descriptors involved in both Kozeny-Carman relation and the machine learning-based prediction model I, which are porosity ϕ, specific surface area S and geometrical tortuosity τ_g. However, the machine learning-based prediction

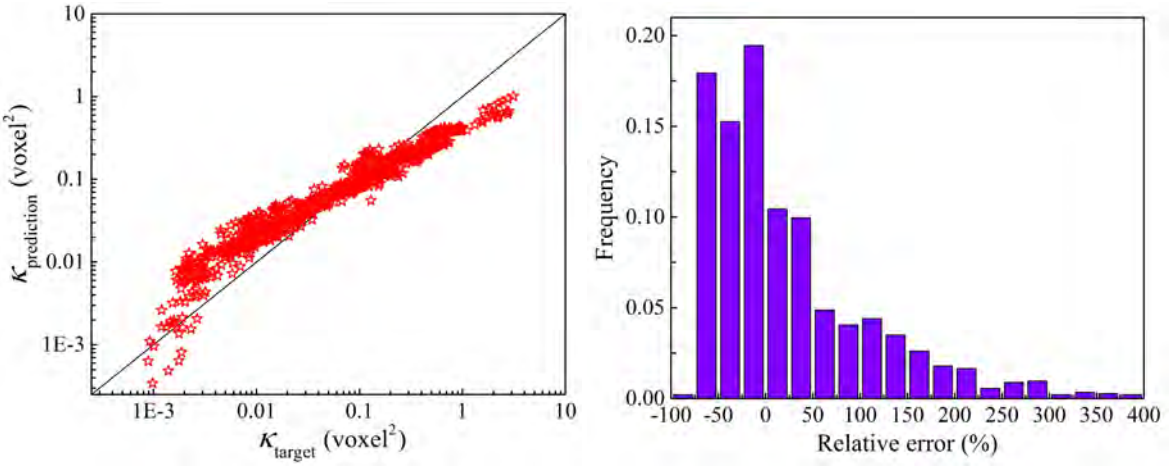
model I possesses a much better performance than Kozeny-Carman equation in terms of accuracy, which can be easily seen by comparing Figure 7.17 with Figure 7.20. The estimation error of Kozeny-Carman relation mainly comes from the misunderstanding of shape factor c, because it is a parameter related to microstructural characteristics such as constriction factor, hydraulic pore diameter and pore throat size, rather than a constant. Even though the shape factor can be used as a constant for simplicity, empirical determination of its value also has a great influence on permeability estimation results. As shown in Figure 7.20a, permeability values are generally overestimated and underestimated for the cases of $c = 2.5$ and $c = 10.0$ respectively. The average values of the permeability estimation results at these two cases are used to assess the performance of Kozeny-Carman relation, and the relative error distribution is given in Figure 7.20b. The permeability estimation error ranges widely from -75.00% to 875.00%, and the average error magnitude is 86.69%. In general, Kozeny-Carman relation shown hug uncertainty and less reliability in permeability estimation for natural porous rocks, mainly due to empirical selection of the shape factor.

Inspired by Kozeny-Carman relation, Berg et al. [34] proposed an analytical model to describe permeability through pore structure parameters, without introducing free constant. Although Berg's model reproduces Darcy's law for idealized pipe flow, it is also used to evaluate permeability for natural rocks. Berg's model is mathematically expressed as follows:

$$\kappa = \frac{L_{\mathrm{h}}^2 \phi_c \beta}{8\tau_{\mathrm{g}}^2} \tag{7.12}$$

where L_{h} is called hydraulic length, representing the effective pore size inside a porous medium. It can be easily found that the remaining three structure parameters are also involved in the machine learning-based permeability prediction model I. Here, the value of L_{h} is estimated from the average pore diameter d by using the skeleton method [115], and then the permeability values of the set of porous media samples are calculated through Berg's model, as shown in Figure 7.21.

Compared with Kozeny-Carman relation, Berg's model exhibits a better performance in permeability estimation, and its greatest success is the exclusion of empirical parameter by introducing constriction factor β. But permeability is still not completely described by Berg's model for natural porous rocks, because the permeability estimations have remarkable errors ranging from -87.25% to 387.25%, as shown in Figure 7.21b. To construct explicit permeability models by introducing more microstructural characteristics may not be a feasible practice, because the nonlinear relationship between microstructure and permeability is extremely complicated for natural porous rocks. In view of the above-mentioned facts, the rationality of implicitly modeling permeability through machine learning to involve



(a) Permeability estimation results (b) The relative error distribution

Fig. 7.21 Performance of Berg's model in permeability estimation

further microstructural informatics can be understood, and these machine learning-based permeability models show great advantages over the empirical/analytical models in terms of predictive accuracy. In the meanwhile, the interpretations of such implicit microstructure-permeability models are maintained as well through the feature selection process.

7.5 Conclusions

In this chapter, the dependence of permeability on microstructural characteristics of natural porous rocks is fundamentally investigated through feature selection and machine learning (shallow neural network). Various morphological descriptors are collected from an extensive literature survey, and they quantitatively characterize pore microstructures from different perspectives. These morphological descriptors compose the raw feature pool, from which microstructural characteristics that are significant to permeability are identified by using wrapper type feature selection.

The optimal set of morphological descriptors contains absolute and connected porosity, specific surface area, geometrical tortuosity, characteristic length, constriction factor, average pore throat size and hydraulic pore diameter. The feature selection result provides a deep insight into the microstructure-permeability relationship, which can be interpreted as follows: (1) Absolute and connected porosity represent the entire void fraction and the permeable portion conducting fluid flow respectively; (2) Specific surface area approximately reflects the fluid-solid interface where adhesive friction is applied to fluid flow; (3) Characteristic length, pore throat size and hydraulic pore diameter are indicators with different emphases to

represent the average distance between opposite solid walls, where fluid flow is constricted between them; (4) Geometrical tortuosity quantifies the prolongation degree of streamlines due to the sinuosity of percolating pore path; (5) Constriction factor measures the cross-section fluctuation along pore channels that leads to divergence and convergence of streamlines at the expansion and shrinkage positions respectively.

Based on the optimal set morphological descriptors, a machine learning-based model is constructed for permeability prediction purpose, and it exhibits an excellent generalization performance in predicting permeability values spanning 4 orders of magnitude for natural porous rocks, with the average error less than 6.92%. Compared with commonly used empirical/analytical permeability relations, the machine learning-based prediction model possesses great superiorities in terms of predictive accuracy and general applicability. Basically, this study provides a new routine to fundamentally understand structure-property linkages for porous media, which can be straightforwardly applied to model different macroscopic properties, such as effective diffusivity, thermal conductivity, formation factor and effective elastic moduli, by involving significant microstructural informatics.

Appendices

Artificial neural networks [494] are function approximators to map the inputs to the output through many interconnected computation elements called neurons. Each elementary neuron possesses a certain degree of approximation capacity, and a powerful learning performance can be achieved by cohesively combining many neurons. In a basic neuron, as illustrated in Figure 7.22, the weighted sum of inputs x_i plussing the bias b is calculated through a transfer function, and the sum is then fed to a non-linear activation function to obtain an output y:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{7.13}$$

where the activation function f can be a sigmoid function, hyperbolic tangent function or rectified linear unit.

Many basic neurons can be connected in various ways to be a multilayer architecture, called a neural network, where the output from a prior neuron is used as the input of the following neuron. It has been proved that a fairly simple neural network is capable of fitting any practical functions [253]. The shallow neural networks only have one or two hidden layers compares to the deep ones, and each hidden layer comprises of many basic neurons. Considering model complexity, prediction accuracy and interpretability, a shallow neural network is adopted to construct the implicit microstructure-permeability model in this study.

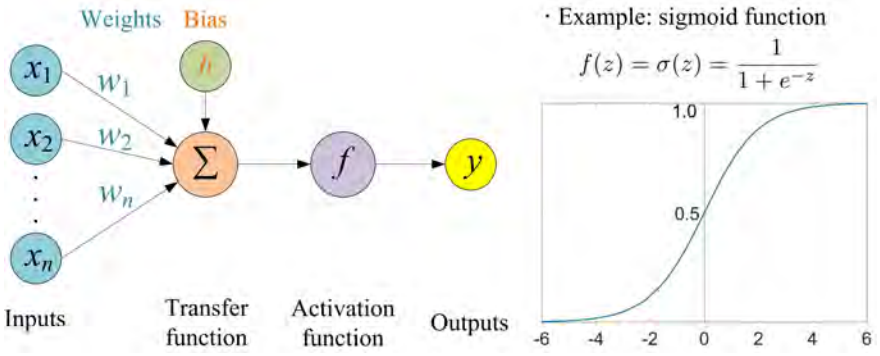


Fig. 7.22 The graphic model of a basic neuron

Morphological descriptors D are used as the microstructural features to feed a shallow neural network model with 2 hidden layers, and the final output is a permeability prediction $\hat{\kappa}$ of the porous media sample, as illustrated in Figure 7.23.

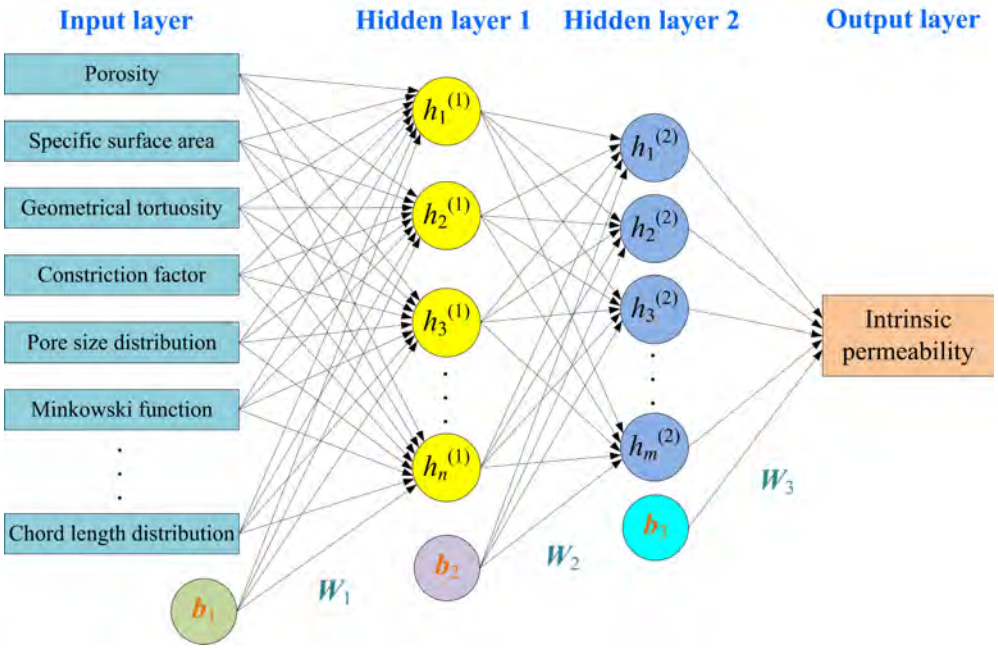


Fig. 7.23 The graphic illustration of a shallow neural network model with 2 hidden layers

The permeability prediction $\hat{\kappa}$ is computed through a series of forward-propagation equations that occur at particular layers, given by:

$$H^{(1)} = f_1\left(W_1^{\mathrm{T}} D + b_1\right) \tag{7.14}$$

$$H^{(2)} = f_2\left(W_2^{\mathrm{T}} H^{(1)} + b_2\right) \tag{7.15}$$

$$\hat{\kappa} = f_3\left(W_3^{\mathrm{T}} H^{(2)} + b_3\right) \tag{7.16}$$

where $H^{(1)}$ and $H^{(2)}$ are the outputs of the 1st and 2nd hidden layers. Training data are used to fit the shallow neural network model by adjusting the weight matrices and bias vectors, where the cost function used to quantify the errors between observations and predictions is minimized through an optimization algorithm such as gradient descent. Cross-validation is usually performed to avoid overfitting, thereby to improve the generalized predictive ability for new observations. More details about parameter optimization of neural networks can be found in relevant references [494, 253].

Chapter 8

Conclusions and outlook

8.1 Conclusions

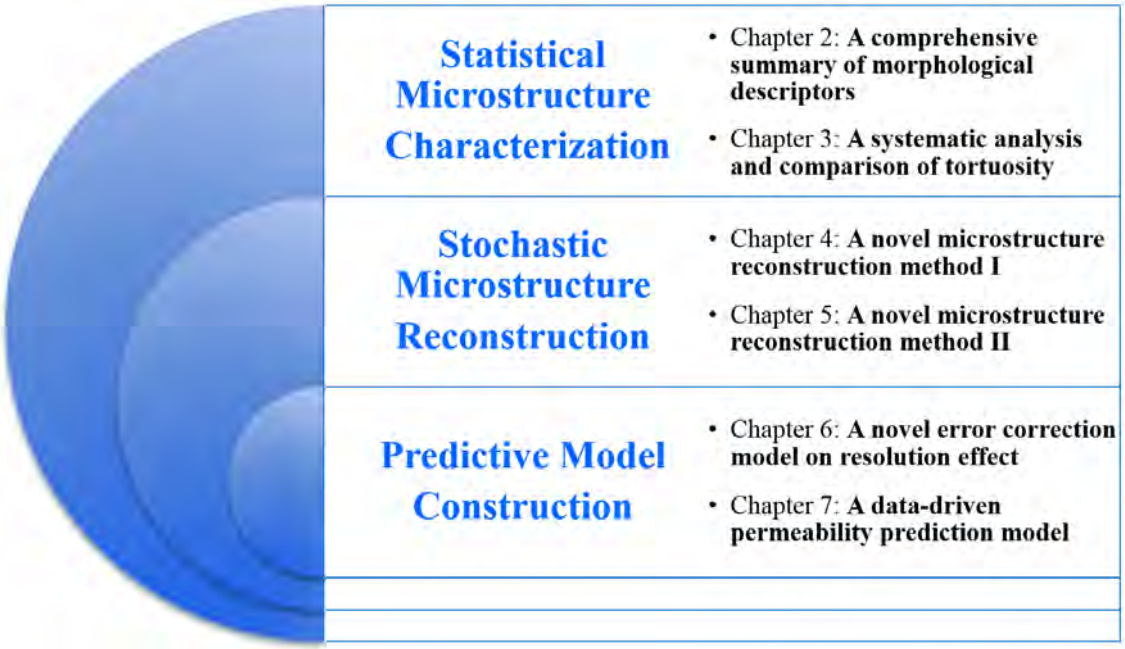


Fig. 8.1 The main contributions and innovations of this thesis

The present thesis is devoted to anatomizing the dependence of transport properties on microstructural characteristics of porous media through statistical microstructure characterization, stochastic microstructure reconstruction and predictive model construction, as illustrated in Figure 8.1. Overall, a comprehensive review of statistical characterization of pore microstructures has been provided; a systematic study of tortuosity evaluation methods, including image analysis and physical simulation, has been conducted; a novel stochastic

microstructure reconstruction method has been developed through deep neural network to generate 2D/3D microstructure samples with high quality; a new method has been proposed to statistically reconstruct 3D microstructures from the limited information provided by 2D thin-sections via supervised machine learning; an error correction model has been established to reduce/eliminate the resolution effect on permeability estimated from LBM; the microstructure-permeability relationship has been fundamentally investigated through the performance-driven process of feature selection, and the machine learning-based permeability prediction model has been constructed by using the optimal subset of microstructural features.

The main contributions and innovations made in each chapter are summarized as follows:

- **Chapter 2** fundamentally investigates the statistical characterization of pore microstructures by presenting a comprehensive review, where a broad variety of morphological descriptors are gathered through an extensive literature survey. The morphological descriptors are quantitative expressions of porous media, which are essential to obtain deep insight into the structure-property linkages, to explore the relevant transport behaviors of fluid flow, or to reveal underlying mechanisms of specific physical phenomena. The commonly used morphological descriptors are introduced in terms of theoretical bases and extraction methods, and they are roughly categorized into four levels to provide microstructural characteristics from global, local, geometrical and topological perspectives. This chapter provides a profound understanding of the limitations and opportunities in microstructure characterization and analysis.

- **Chapter 3** systematically examines geometrical, hydraulic, electrical, diffusional, and thermal tortuosities from the viewpoints of concept and evaluation method. Physical tortuosities are defined to describe different transport processes occurring inside porous media, while geometrical tortuosity is considered as a microstructural characteristic, and they can be evaluated through various image analysis and numerical simulation methods. Different types of tortuosity can differ greatly in value, but they are used interchangeably in many works of literature. Thus, the correlations between geometrical and physical tortuosities are further analyzed, which also takes into account the influences of both image size and resolution. From the correlation analysis, phenomenological relations between geometrical and physical tortuosities are established, so that the latter can be accurately predicted by using the former, avoiding computationally intensive simulations of transport processes.

- **Chapter 4** presents a novel method to statistically characterize and reconstruct random microstructures through a deep neural network (DNN) model, which can be used to study the microstructure-property relationships. In this method, the digital microstructure is assumed to be a stationary Markov random field (MRF), and local patterns covering the basic morphological features are collected to train a DNN model, after which statistically equivalent samples can be generated through a DNN-guided reconstruction procedure. To overcome the short-distance limitation associated with the MRF assumption, a multi-level approach is developed to preserve the long-distance morphological features of heterogeneous microstructures. The proposed microstructure reconstruction method can capture both morphological characteristics and physical properties, which is efficient, accurate and versatile to generate 2D/3D microstructures with long-distance morphology, such as porous media with long-range connectivity.

- **Chapter 5** proposes a new approach to stochastically reconstructing 3D heterogeneous porous media from the 2D cross-sectional images through supervised machine learning. Data events collected from the 2D cross-sectional images are fed to the supervised learning model to learn the multiple-point statistics of the original microstructure. For the isotropic microstructure, the multiple-points statistics on one (or several) representative 2D plane are sufficient to reflect the 3D morphological information. While for the anisotropic microstructure, the multiple-points statistics of 2D planes in the three principal directions are required to capture the anisotropy. The class probabilities stored in the fitted supervised learning model(s) are good approximations of the multiple-points statistics of each 2D plane, and 2D statistical information on three perpendicular planes can be integrated to provide an estimation of the conditional probability of voxel in 3D space, based on which statistically equivalent 3D microstructures can be generated through probability sampling. Compared with the other three classical reconstruction methods, the proposed method can rapidly generate more realistic and accurate 3D microstructures.

- **Chapter 6** quantitatively investigates the resolution effect on permeability evaluation from the lattice Boltzmann method (LBM), and an error correction model is established to reduce/eliminate this resolution effect by identifying the primary error causes. The error correction model uses such geometric attributes as connected porosity, specific surface area and diffusional tortuosity to quantify the permeability error due to resolution effect and achieve error correction. Demonstrated on various types of porous media including sandstone, carbonate rock, sand pack and synthesis silica, etc., the proposed error correction model can effectively correct the errors in LBM permeability

evaluation caused by the resolution effect. The error correction model makes image resolution reduction more meaningful and creditable for LBM permeability evaluation of porous media, thereby supporting its adoption in practical applications.

- **Chapter 7** deeply explores the dependence of permeability on microstructural characteristics of natural porous rocks through performance-driven feature selection, and the machine learning-based permeability model is constructed by using the optimal subset of morphological descriptors as the feature data. The optimal feature set contains absolute and connected porosity, specific surface area, geometrical tortuosity, characteristic length, constriction factor, average pore throat size and hydraulic pore diameter, which provides a deep insight into the microstructure-permeability linkage that can be interpreted as follows: Absolute and connected porosity represent the entire void fraction and the permeable portion conducting fluid flow respectively; Specific surface area approximately reflects the fluid-solid interface where adhesive friction is applied to fluid flow; Characteristic length, pore throat size and hydraulic pore diameter are indicators with different emphases to represent the average distance between opposite solid walls, where fluid flow is constricted between them; Geometrical tortuosity quantifies the prolongation degree of streamlines due to the sinuosity of percolating pore path; Constriction factor measures the cross-section fluctuation along pore channels that leads to divergence and convergence of streamlines at the expansion and shrinkage positions respectively. Additionally, the machine learning-based permeability model possesses an excellent generalization performance to estimate permeability values spanning 4 orders of magnitude with the average error less than 6.92%, and it exhibits great superiorities over empirical/analytical permeability relations in terms of predictive accuracy and generalization ability.

8.2 Outlook

The research work presented in this thesis has improved the understanding of the relationships between microstructural characteristics and transport properties of porous media, but some problems still require further research. Future studies that may be beneficial are suggested as follows:

- Although extensive efforts have been made to quantitatively characterize pore microstructures, the morphological descriptors determined from digitized images are subject to a number of errors and uncertainties, including finite image size, finite resolution, binarization error, boundary effect, and errors related to parameter selection (such

as connectivity criterion and distance metrics) of determination algorithms. These errors should be constrained within acceptable ranges for microstructural analysis.

- Stochastic reconstruction of 3D heterogeneous pore microstructure is an effective way to provide large numbers of samples to analyze the transport properties of porous media. However, it is still a challenge to rapidly generate 3D digital samples of pore microstructures with arbitrary desired morphologies, especially for 3D microstructure reconstruction from 2D cross-sectional images. A more powerful method remains to be developed to promote the microstructure-property study in porous media research.

- For permeability evaluation from lattice Boltzmann simulation, viscosity dependence is one of its major drawbacks. According to our preliminary findings, it is possible to determine the optimization relaxation time of LBM simulation for a specific porous medium, where the fluid-solid boundary is exactly located at the pore-solid interface, thereby avoiding the viscosity dependence.

- This thesis provides a new routine to fundamentally investigate the dependence of permeability on microstructural characteristics of porous media, which can also be applied to study other macroscopic properties by involving significant microstructural informatics, such as effective diffusivity, thermal conductivity, formation factor and elastic moduli. Besides, it is well worth establishing microstructure-property linkages for sandstones containing finer-grained materials (i.e. clay).

- It is interesting to apply the morphological descriptors to quantitatively represent the characteristics of multi-phase flow to understand the fluid distribution, capillary pressure, relative permeability, fluid displacement mechanisms and other pore-scale transport behaviors.

References

[1] Adler, P. (1996). Transports in fractal porous media. *Journal of Hydrology*, 187(1-2):195–213.

[2] Adler, P. (2013). *Porous Media: Geometry and Transports*. Elsevier.

[3] Agrawal, A. and Choudhary, A. (2019). Deep materials informatics: Applications of deep learning in materials science. *MRS Communications*, 9(3):779–792.

[4] Ahmadi, M. M., Mohammadi, S., and Hayati, A. N. (2011). Analytical derivation of tortuosity and permeability of monosized spheres: A volume averaging approach. *Physical Review E*, 83(2):026312.

[5] Ahmed, T. (2018). *Reservoir Engineering Handbook*. Gulf Professional Publishing.

[6] Al-Raoush, R. I. and Madhoun, I. T. (2017). Tort3d: A matlab code to compute geometric tortuosity from 3d images of unconsolidated porous media. *Powder Technology*, 320:99–107.

[7] Al Saadi, F., Wolf, K., and Kruijsdijk, C. (2017). Characterization of fontainebleau sandstone: Quartz overgrowth and its impact on pore-throat framework. *Journal of Petroleum & Environmental Biotechnology*, 7(328):2.

[8] Allain, C. and Cloitre, M. (1991). Characterizing the lacunarity of random and deterministic fractal sets. *Physical Review A*, 44(6):3552.

[9] Allard, D. et al. (1993). On the connectivity of two random set models: the truncated gaussian and the boolean. In *Geostatistics Tróia'92*, pages 467–478. Springer.

[10] Alyafei, N., Raeini, A. Q., Paluszny, A., and Blunt, M. J. (2015). A sensitivity study of the effect of image resolution on predicted petrophysical properties. *Transport in Porous Media*, 110(1):157–169.

[11] Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., Keehm, Y., Krzikalla, F., Lee, M., Madonna, C., et al. (2013). Digital rock physics benchmarks—part ii: Computing effective properties. *Computers & Geosciences*, 50:33–43.

[12] Anovitz, L. M. and Cole, D. R. (2015). Characterization and analysis of porosity and pore structures. *Reviews in Mineralogy and Geochemistry*, 80(1):61–164.

[13] Antoine, M., Javaux, M., and Bielders, C. (2009). What indicators can capture runoff-relevant connectivity properties of the micro-topography at the plot scale? *Advances in Water Resources*, 32(8):1297–1310.

[14] Apourvari, S. N. and Arns, C. H. (2016). Image-based relative permeability upscaling from the pore scale. *Advances in Water Resources*, 95:161–175.

[15] Arandigoyen, M. and Alvarez, J. (2007). Pore structure and mechanical properties of cement–lime mortars. *Cement and Concrete Research*, 37(5):767–775.

[16] Archie, G. E. et al. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the AIME*, 146(01):54–62.

[17] Armstrong, R. T., McClure, J. E., Berrill, M. A., Rücker, M., Schlüter, S., and Berg, S. (2016). Beyond darcy's law: The role of phase topology and ganglion dynamics for two-fluid flow. *Physical Review E*, 94(4):043113.

[18] Armstrong, R. T., McClure, J. E., Robins, V., Liu, Z., Arns, C. H., Schlüter, S., and Berg, S. (2019). Porous media characterization using minkowski functionals: theories, applications and future directions. *Transport in Porous Media*, 130(1):305–335.

[19] Armstrong, R. T., Porter, M. L., and Wildenschild, D. (2012). Linking pore-scale interfacial curvature to column-scale capillary pressure. *Advances in Water Resources*, 46:55–62.

[20] Avizo (2017). ThermoFisher Scientific. Oregon, USA. Website. https://www.thermofisher.com.

[21] Axelsson, M. and Svensson, S. (2010). 3d pore structure characterisation of paper. *Pattern Analysis and Applications*, 13(2):159–172.

[22] Bache, K. and Lichman, M. (2013). UCI machine learning repository.

[23] Backeberg, N. R., Iacoviello, F., Rittner, M., Mitchell, T. M., Jones, A. P., Day, R., Wheeler, J., Shearing, P. R., Vermeesch, P., and Striolo, A. (2017). Quantifying the anisotropy and tortuosity of permeable pathways in clay-rich mudstones using models based on x-ray tomography. *Scientific Reports*, 7(1):14838.

[24] Bae, C.-J., Erdonmez, C. K., Halloran, J. W., and Chiang, Y.-M. (2013). Design of battery electrodes with dual-scale porosity to minimize tortuosity and maximize performance. *Advanced Materials*, 25(9):1254–1258.

[25] Bærentzen, J. A. (2001). On the implementation of fast marching methods for 3d lattices.

[26] Barrande, M., Bouchet, R., and Denoyel, R. (2007). Tortuosity of porous particles. *Analytical Chemistry*, 79(23):9115–9121.

[27] Basheer, L., Kropp, J., and Cleland, D. J. (2001). Assessment of the durability of concrete from its permeation properties: A review. *Construction and Building Materials*, 15(2-3):93–103.

[28] Baychev, T. G., Jivkov, A. P., Rabbani, A., Raeini, A. Q., Xiong, Q., Lowe, T., and Withers, P. J. (2019). Reliability of algorithms interpreting topological and geometric properties of porous media for pore network modelling. *Transport in Porous Media*, 128(1):271–301.

[29] Bear, J. (2013). *Dynamics of Fluids in Porous Media*. Courier Corporation.

[30] Bear, J. and Bachmat, Y. (2012). *Introduction to Modeling of Transport Phenomena in Porous Media*, volume 4. Springer Science & Business Media.

[31] Bear, J. and Cheng, A. H.-D. (2010). *Modeling Groundwater Flow and Contaminant Transport*, volume 23. Springer Science & Business Media.

[32] Bekri, S., Xu, K., Yousefian, F., Adler, P., and Thovert, J.-F. (2000). Pore geometry and transport properties in north sea chalk. *Journal of Petroleum Science & Engineering*, 25(3-4):107–134.

[33] Bentz, D. P. and Martys, N. S. (1994). Hydraulic radius and transport in reconstructed model three-dimensional porous media. *Transport in Porous Media*, 17(3):221–238.

[34] Berg, C. F. (2014). Permeability description by characteristic length, tortuosity, constriction and porosity. *Transport in Porous Media*, 103(3):381–400.

[35] Berg, C. F. (2016). Fontainebleau 3d models. http://www.digitalrocksportal.org/projects/57.

[36] Berg, S., Armstrong, R., and Wiegmann, A. (2018). Gildehauser sandstone. http://www.digitalrocksportal.org/projects/134.

[37] Bernabé, Y., Zamora, M., Li, M., Maineult, A., and Tang, Y. (2011). Pore connectivity, permeability, and electrical formation factor: a new model and comparison to experimental data. *Journal of Geophysical Research: Solid Earth*, 116(B11).

[38] Berryman, J. G. and Blair, S. C. (1987). Kozeny–carman relations and image processing methods for estimating darcy's constant. *Journal of Applied Physics*, 62(6):2221–2228.

[39] Bertei, A., Nucci, B., and Nicolella, C. (2013). Effective transport properties in random packings of spheres and agglomerates.

[40] Bickle, M. J. (2009). Geological carbon storage. *Nature Geoscience*, 2(12):815–818.

[41] Bini, F., Pica, A., Marinozzi, A., and Marinozzi, F. (2019). A 3d model of the effect of tortuosity and constrictivity on the diffusion in mineralized collagen fibril. *Scientific Reports*, 9(1):1–14.

[42] Bird, R. B. (2002). Transport phenomena. *Applied Mechanics Reviews*, 55(1):R1–R4.

[43] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

[44] Biswal, B., Manwart, C., and Hilfer, R. (1998). Three-dimensional local porosity analysis of porous media. *Physica A: Statistical Mechanics and its Applications*, 255(3-4):221–241.

[45] Biswal, B., Manwart, C., Hilfer, R., Bakke, S., and Øren, P. (1999). Quantitative analysis of experimental and synthetic microstructures for sedimentary rock. *Physica A: Statistical Mechanics and its Applications*, 273(3-4):452–475.

[46] Blair, S. C., Berge, P. A., and Berryman, J. G. (1996). Using two-point correlation functions to characterize microgeometry and estimate permeabilities of sandstones and porous glass. *Journal of Geophysical Research: Solid Earth*, 101(B9):20359–20375.

[47] Blunt, M. J. (2017). *Multiphase Flow in Permeable Media: A pore-scale Perspective*. Cambridge University Press.

[48] Blunt, M. J., Bijeljic, B., Dong, H., Gharbi, O., Iglauer, S., Mostaghimi, P., Paluszny, A., and Pentland, C. (2013). Pore-scale imaging and modelling. *Advances in Water Resources*, 51:197–216.

[49] Bodla, K. K., Murthy, J. Y., and Garimella, S. V. (2010). Microtomography-based simulation of transport through open-cell metal foams. *Numerical Heat Transfer, Part A: Applications*, 58(7):527–544.

[50] Borner, F. D., Schon, J. H., et al. (1991). A relation between the quadrature component of electrical conductivity and the specific surface area of sedimentary rocks. *The Log Analyst*, 32(05).

[51] Borujeni, A. T., Lane, N., Thompson, K., and Tyagi, M. (2013). Effects of image resolution and numerical resolution on computed permeability of consolidated packing using lb and fem pore-scale simulations. *Computers & Fluids*, 88:753–763.

[52] Bostanabad, R., Bui, A. T., Xie, W., Apley, D. W., and Chen, W. (2016a). Stochastic microstructure characterization and reconstruction via supervised learning. *Acta Materialia*, 103:89–102.

[53] Bostanabad, R., Chen, W., and Apley, D. (2016b). Characterization and reconstruction of 3d stochastic microstructures via supervised learning. *Journal of Microscopy*, 264(3):282–297.

[54] Botchwey, E. A., Dupree, M. A., Pollack, S. R., Levine, E. M., and Laurencin, C. T. (2003). Tissue engineered bone: Measurement of nutrient transport in three-dimensional matrices. *Journal of Biomedical Materials Research Part A: An Official Journal of The Society for Biomaterials, The Japanese Society for Biomaterials, and The Australian Society for Biomaterials and the Korean Society for Biomaterials*, 67(1):357–367.

[55] Boudreau, B. P. (1996). The diffusive tortuosity of fine-grained unlithified sediments. *Geochimica et Cosmochimica Acta*, 60(16):3139–3142.

[56] Bourbie, T. and Zinszner, B. (1985). Hydraulic and acoustic properties as a function of porosity in fontainebleau sandstone. *Journal of Geophysical Research: Solid Earth*, 90(B13):11524–11532.

[57] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

[58] Breiman, L. (2017). *Classification and Regression Trees*. Routledge.

[59] Brodkey, R. S. and Hershey, H. C. (2003). *Transport Phenomena: A Unified Aprroach*. Brodkey publishing.

[60] Brown, L., Abdulaziz, R., Tjaden, B., Inman, D., Brett, D., and Shearing, P. (2016). Investigating microstructural evolution during the electroreduction of uo2 to u in licl-kcl eutectic using focused ion beam tomography. *Journal of Nuclear Materials*, 480:355–361.

[61] Bruggeman, V. D. (1935). Berechnung verschiedener physikalischer konstanten von heterogenen substanzen. i. dielektrizitätskonstanten und leitfähigkeiten der mischkörper aus isotropen substanzen. *Annalen Der Physik*, 416(7):636–664.

[62] Brun, E., Vicente, J., Topin, F., and Occelli, R. (2008). Imorph: A 3d morphological tool to fully analyse all kind of cellular materials. *Cellular Metals for Structural and Functional Applications*.

[63] Bultreys, T. (2016). Massangis jaune carbonate. http://www.digitalrocksportal.org/projects/73.

[64] Bultreys, T., De Boever, W., and Cnudde, V. (2016). Imaging and image-based fluid transport modeling at the pore scale in geological materials: A practical introduction to the current state-of-the-art. *Earth-Science Reviews*, 155:93–128.

[65] Burt, P. and Adelson, E. (1983). The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4):532–540.

[66] Caers, J. (2001). Geostatistical reservoir modelling using statistical pattern recognition. *Journal of Petroleum Science and Engineering*, 29(3-4):177–188.

[67] Cai, B., Wang, J., Kao, A., Pericleous, K., Phillion, A., Atwood, R., and Lee, P. (2016). 4d synchrotron x-ray tomographic quantification of the transition from cellular to dendrite growth during directional solidification. *Acta Materialia*, 117:160–169.

[68] Cai, J., Wei, W., Hu, X., and Wood, D. A. (2017). Electrical conductivity models in saturated porous media: A review. *Earth-Science Reviews*, 171:419–433.

[69] Cang, R., Xu, Y., Chen, S., Liu, Y., Jiao, Y., and Yi Ren, M. (2017). Microstructure representation and reconstruction of heterogeneous materials via deep belief network for computational material design. *Journal of Mechanical Design*, 139(7).

[70] Čapek, P., Hejtmánek, V., Brabec, L., Zikánová, A., and Kočiřík, M. (2009). Stochastic reconstruction of particulate media using simulated annealing: improving pore connectivity. *Transport in Porous Media*, 76(2):179–198.

[71] Carman, P. C. (1937). Fluid flow through granular beds. *Transactions of the Institution of Chemical Engineers*, 50:150–166.

[72] Carman, P. C. (1956). *Flow of Gases through Porous Media*. Butterworths Scientific Publications: London.

[73] Cecen, A., Dai, H., Yabansu, Y. C., Kalidindi, S. R., and Song, L. (2018). Material structure-property linkages using three-dimensional convolutional neural networks. *Acta Materialia*, 146:76–84.

[74] Cecen, A., Wargo, E., Hanna, A., Turner, D., Kalidindi, S., and Kumbur, E. (2012a). 3-d microstructure analysis of fuel cell materials: spatial distributions of tortuosity, void size and diffusivity. *Journal of The Electrochemical Society*, 159(3):B299.

[75] Cecen, A., Wargo, E., Hanna, A., Turner, D., Kalidindi, S., and Kumbur, E. (2012b). 3-d microstructure analysis of fuel cell materials: spatial distributions of tortuosity, void size and diffusivity. *Journal of The Electrochemical Society*, 159(3):B299–B307.

[76] Chapman, A. and Higdon, J. J. (1992). Oscillatory stokes flow in periodic porous media. *Physics of Fluids A: Fluid Dynamics*, 4(10):2099–2116.

[77] Chen, D., He, X., Teng, Q., Xu, Z., and Li, Z. (2014). Reconstruction of multiphase microstructure based on statistical descriptors. *Physica A: Statistical Mechanics and Its Applications*, 415:240–250.

[78] Chen, H., Chen, S., and Matthaeus, W. H. (1992). Recovery of the navier-stokes equations using a lattice-gas boltzmann method. *Physical Review A*, 45(8):R5339.

[79] Chen, S.-C., Lee, E. K., and Chang, Y.-I. (2003). Effect of the coordination number of the pore-network on the transport and deposition of particles in porous media. *Separation and Purification Technology*, 30(1):11–26.

[80] Chen-Wiegart, Y.-c. K., DeMike, R., Erdonmez, C., Thornton, K., Barnett, S. A., and Wang, J. (2014). Tortuosity characterization of 3d microstructure at nano-scale for energy storage and conversion materials. *Journal of Power Sources*, 249:349–356.

[81] Cheng, Z. and Wang, J. (2018). A particle-tracking method for experimental investigation of kinematics of sand particles under triaxial compression. *Powder Technology*, 328:436–451.

[82] Cherkassky, B. V., Goldberg, A. V., and Radzik, T. (1996). Shortest paths algorithms: Theory and experimental evaluation. *Mathematical Programming*, 73(2):129–174.

[83] Chiu, S. N., Stoyan, D., Kendall, W. S., and Mecke, J. (2013). *Stochastic Geometry and Its Applications*. John Wiley & Sons.

[84] Choi, K. S., Zhu, Z., Sun, X., De Moor, E., Taylor, M. D., Speer, J. G., and Matlock, D. K. (2015). Determination of carbon distributions in quenched and partitioned microstructures using nanoscale secondary ion mass spectroscopy. *Scripta Materialia*, 104:79–82.

[85] Chun, H. C., Giménez, D., and Yoon, S. W. (2008). Morphology, lacunarity and entropy of intra-aggregate pores: Aggregate size and soil management effects. *Geoderma*, 146(1-2):83–93.

[86] Chung, D.-W., Ebner, M., Ely, D. R., Wood, V., and García, R. E. (2013). Validity of the bruggeman relation for porous electrodes. *Modelling and Simulation in Materials Science and Engineering*, 21(7):074009.

[87] Civan, F. (2011). *Porous Media Transport Phenomena*. John Wiley & Sons.

[88] Clarkson, C. R., Solano, N., Bustin, R. M., Bustin, A., Chalmers, G., He, L., Melnichenko, Y. B., Radliński, A., and Blach, T. P. (2013). Pore structure characterization of north american shale gas reservoirs using usans/sans, gas adsorption, and mercury intrusion. *Fuel*, 103:606–616.

[89] Clarkson, C. R., Wood, J., Burgis, S., Aquino, S., Freeman, M., et al. (2012). Nanopore-structure analysis and permeability predictions for a tight gas siltstone reservoir by use of low-pressure adsorption and mercury-intrusion techniques. *SPE Reservoir Evaluation & Engineering*, 15(06):648–661.

[90] Clennell, M. B. (1997). Tortuosity: a guide through the maze. *Geological Society, London, Special Publications*, 122(1):299–344.

[91] Cnudde, V., Cwirzen, A., Masschaele, B., and Jacobs, P. (2009). Porosity and microstructure characterization of building stones and concretes. *Engineering Geology*, 103(3-4):76–83.

[92] Coker, D. A., Torquato, S., and Dunsmuir, J. H. (1996). Morphology and physical properties of fontainebleau sandstone via a tomographic analysis. *Journal of Geophysical Research: Solid Earth*, 101(B8):17497–17506.

[93] Coleman, S. and Vassilicos, J. (2008). Transport properties of saturated and unsaturated porous fractal materials. *Physical Review Letters*, 100(3):035504.

[94] Comiti, J. and Renaud, M. (1989). A new model for determining mean structure parameters of fixed beds from pressure drop measurements: application to beds packed with parallelepipedal particles. *Chemical Engineering Science*, 44(7):1539–1545.

[95] Cooper, S. (2015). Quantifying the transport properties of solid oxide fuel cell electrodes.

[96] Cooper, S., Bertei, A., Shearing, P., Kilner, J., and Brandon, N. (2016). Taufactor: An open-source application for calculating tortuosity factors from tomographic data. *SoftwareX*, 5:203–210.

[97] Cooper, S., Eastwood, D., Gelb, J., Damblanc, G., Brett, D., Bradley, R., Withers, P., Lee, P., Marquis, A., Brandon, N., et al. (2014). Image based modelling of microstructural heterogeneity in lifepo4 electrodes for li-ion batteries. *Journal of Power Sources*, 247:1033–1039.

[98] Cooper, S. J., Kishimoto, M., Tariq, F., Bradley, R. S., Marquis, A. J., Brandon, N. P., Kilner, J. A., and Shearing, P. R. (2013). Microstructural analysis of an lscf cathode using in situ tomography and simulation. *ECS Transactions*, 57(1):2671–2678.

[99] Corapcioglu, M. Y. (1996). *Advances in Porous Media*, volume 3. Elsevier.

[100] Corson, P. B. (1974). Correlation functions for predicting properties of heterogeneous materials. ii. empirical construction of spatial correlation functions for two-phase solids. *Journal of Applied Physics*, 45(7):3165–3170.

[101] Cosenza, P., Prêt, D., and Zamora, M. (2015). Effect of the local clay distribution on the effective electrical conductivity of clay rocks. *Journal of Geophysical Research: Solid Earth*, 120(1):145–168.

[102] Costa, A. (2006). Permeability-porosity relationship: A reexamination of the kozeny-carman equation based on a fractal pore-space geometry assumption. *Geophysical Research Letters*, 33(2).

[103] Costanza-Robinson, M. S., Estabrook, B. D., and Fouhey, D. F. (2011). Representative elementary volume estimation for porosity, moisture saturation, and air-water interfacial areas in unsaturated porous media: Data quality implications. *Water Resources Research*, 47(7).

[104] Cousins, T. A., Ghanbarian, B., and Daigle, H. (2018). Three-dimensional lattice boltzmann simulations of single-phase permeability in random fractal porous media with rough pore–solid interface. *Transport in Porous Media*, 122(3):527–546.

[105] Cule, D. and Torquato, S. (1999). Generating random media from limited microstructural information via stochastic optimization. *Journal of Applied Physics*, 86(6):3428–3437.

[106] Currie, J. (1960). Gaseous diffusion in porous media. part 2.-dry granular materials. *British Journal of Applied Physics*, 11(8):318.

[107] Darabi, H., Ettehad, A., Javadpour, F., and Sepehrnoori, K. (2012). Gas flow in ultra-tight shale strata. *Journal of Fluid Mechanics*, 710:641–658.

[108] David, C. (1993). Geometry of flow paths for fluid transport in rocks. *Journal of Geophysical Research: Solid Earth*, 98(B7):12267–12278.

[109] David, C., Darot, M., et al. (1989). Permeability and conductivity of sandstones. In *ISRM International Symposium*. International Society for Rock Mechanics and Rock Engineering.

[110] De Lima, O. and Niwas, S. (2000). Estimation of hydraulic parameters of shaly sandstone aquifers from geoelectrical measurements. *Journal of Hydrology*, 235(1-2):12–26.

[111] de Lima, O. A. (1995). Water saturation and permeability from resistivity, dielectric, and porosity logs. *Geophysics*, 60(6):1756–1764.

[112] De Melo, R. and Conci, A. (2008). Succolarity: Defining a method to calculate this fractal measure. In *2008 15th International Conference on Systems, Signals and Image Processing*, pages 291–294. IEEE.

[113] De Melo, R. H. and Conci, A. (2013). How succolarity could be used as another fractal measure in image analysis. *Telecommunication Systems*, 52(3):1643–1655.

[114] Degruyter, W., Burgisser, A., Bachmann, O., and Malaspinas, O. (2010). Synchrotron x-ray microtomography and lattice boltzmann simulations of gas flow through volcanic pumices. *Geosphere*, 6(5):470–481.

[115] Delerue, J., Perrier, E., Yu, Z., and Velde, B. (1999). New algorithms in 3d image analysis and their application to the measurement of a spatialized pore size distribution in soils. *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*, 24(7):639–644.

[116] Desrues, J., Viggiani, G., and Besuelle, P. (2010). *Advances in X-ray Tomography for Geomaterials*, volume 118. John Wiley & Sons.

[117] Devarapalli, R. S., Islam, A., Faisal, T. F., Sassi, M., and Jouiad, M. (2017). Micro-ct and fib–sem imaging and pore structure characterization of dolomite rock at multiple scales. *Arabian Journal of Geosciences*, 10(16):361.

[118] d'Humieres, D. (2002). Multiple–relaxation–time lattice boltzmann models in three dimensions. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 360(1792):437–451.

[119] Dias, R., Teixeira, J. A., Mota, M., and Yelshin, A. (2006). Tortuosity variation in a low density binary particulate bed. *Separation and Purification Technology*, 51(2):180–184.

[120] Dong, H. (2008). *Micro-CT imaging and pore network extraction*. PhD thesis, Department of Earth Science and Engineering, Imperial College London.

[121] Dong, H., Gao, P., and Ye, G. (2017). Characterization and comparison of capillary pore structures of digital cement pastes. *Materials and Structures*, 50(2):154.

[122] Doyen, P. M. (1988). Permeability, conductivity, and pore geometry of sandstone. *Journal of Geophysical Research: Solid Earth*, 93(B7):7729–7740.

[123] Du Plessis, J. P. and Masliyah, J. H. (1991). Flow through isotropic granular porous media. *Transport in Porous Media*, 6(3):207–221.

[124] Duda, A., Koza, Z., and Matyka, M. (2011). Hydraulic tortuosity in arbitrary porous media flow. *Physical Review E*, 84(3):036319.

[125] Dullien, F. A. (2012). *Porous Media: Fluid Transport and Pore Structure*. Academic press.

[126] Edie, M. S., Olson, J. F., Burns, D. R., and Toksoz, M. N. (2000). The effect of image resolution on fluid flow simulations in porous media. Technical report, Massachusetts Institute of Technology. Earth Resources Laboratory.

[127] Epstein, N. (1989). On tortuosity and the tortuosity factor in flow and diffusion through porous media. *Chemical Engineering Science*, 44(3):777–779.

[128] Erofeev, A., Orlov, D., Ryzhov, A., and Koroteev, D. (2019). Prediction of porosity and permeability alteration based on machine learning algorithms. *Transport in Porous Media*, 128(2):677–700.

[129] Eshghinejadfard, A., Daróczy, L., Janiga, G., and Thévenin, D. (2016). Calculation of the permeability in porous media using the lattice boltzmann method. *International Journal of Heat and Fluid Flow*, 62:93–103.

[130] Fadnavis, S. (2014). Image interpolation techniques in digital image processing: an overview. *International Journal of Engineering Research and Applications*, 4(10):70–73.

[131] Falconer, K. (2004). *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons.

[132] Farmer, C. (2002). Upscaling: a review. *International Journal for Numerical Methods in Fluids*, 40(1-2):63–78.

[133] Farrell, N. and Healy, D. (2017). Anisotropic pore fabrics in faulted porous sandstones. *Journal of Structural Geology*, 104:125–141.

[134] Feng, J., Cen, S., Li, C., and Owen, D. (2016). Statistical reconstruction and karhunen–loève expansion for multiphase random media. *International Journal for Numerical Methods in Engineering*, 105(1):3–32.

[135] Feng, J., Li, C., Cen, S., and Owen, D. (2014). Statistical reconstruction of two-phase random media. *Computers & Structures*, 137:78–92.

[136] Ferguson, J. C., Panerai, F., Borner, A., and Mansour, N. N. (2018). Puma: The porous microstructure analysis software. *SoftwareX*, 7:81–87.

[137] Finegan, D. P., Cooper, S. J., Tjaden, B., Taiwo, O. O., Gelb, J., Hinds, G., Brett, D. J., and Shearing, P. R. (2016). Characterising the structural properties of polymer separators for lithium-ion batteries in 3d using phase contrast x-ray microscopy. *Journal of Power Sources*, 333:184–192.

[138] Florencio-Silva, R., Sasso, G. R. d. S., Sasso-Cerri, E., Simões, M. J., and Cerri, P. S. (2015). Biology of bone tissue: structure, function, and factors that influence bone cells. *BioMed Research International*, 2015.

[139] Fredrich, J., DiGiovanni, A., and Noble, D. (2006). Predicting macroscopic transport properties using microscopic image data. *Journal of Geophysical Research: Solid Earth*, 111(B3).

[140] Fredrich, J., Greaves, K., and Martin, J. (1993). Pore geometry and transport properties of fontainebleau sandstone. In *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, volume 30, pages 691–697. Elsevier.

[141] Fu, J., Cui, S., Cen, S., and Li, C. (2021). Statistical characterization and reconstruction of heterogeneous microstructures using deep neural network. *Computer Methods in Applied Mechanics and Engineering*, 373:113516.

[142] Fu, J., Dong, J., Wang, Y., Ju, Y., Owen, D. R. J., and Li, C. (2020a). Resolution effect: An error correction model for intrinsic permeability of porous media estimated from lattice boltzmann method. *Transport in Porous Media*, 132(3):627–656.

[143] Fu, J., Thomas, H. R., and Li, C. (2020b). Tortuosity of porous media: Image analysis and physical simulation. *Earth-Science Reviews*, page 103439.

[144] Gaiselmann, G., Neumann, M., Holzer, L., Hocker, T., Prestat, M. R., and Schmidt, V. (2013). Stochastic 3d modeling of la0. 6sr0. 4coo3- δ cathodes based on structural segmentation of fib–sem images. *Computational Materials Science*, 67:48–62.

[145] Gao, Z. and Hu, Q. (2013). Estimating permeability using median pore-throat radius obtained from mercury intrusion porosimetry. *Journal of Geophysics and Engineering*, 10(2):025014.

[146] Garboczi, E. J. and Bentz, D. P. (1990). Digitised simulation of mercury intrusion porosimetry. *Ceram. Trans.*, 16:365–380.

[147] Gariboldi, M. I. and Best, S. M. (2015). Effect of ceramic scaffold architectural parameters on biological response. *Frontiers in Bioengineering and Biotechnology*, 3:151.

[148] Garrison Jr, J. R., Pearn, W. C., and von Rosenberg, D. (1992). The fractal menger sponge and sierpinski carpet as models for reservoir rock/pore systems: I.; theory and image analysis of sierpinski carpets. *In Situ;(United States)*, 16(4).

[149] Garrouch, A. A., Ali, L., and Qasem, F. (2001). Using diffusion and electrical measurements to assess tortuosity of porous media. *Industrial & Engineering Chemistry Research*, 40(20):4363–4369.

[150] Geistlinger, H., Ataei-Dadavi, I., Mohammadian, S., and Vogel, H.-J. (2015). The impact of pore structure and surface roughness on capillary trapping for 2-d and 3-d porous media: Comparison with percolation theory. *Water Resources Research*, 51(11):9094–9111.

[151] GeoDict (2020). Math2Market. Kaiserslautern, Germany. Website. http://www.geodict.com.

[152] Géraud, Y. (1994). Variations of connected porosity and inferred permeability in a thermally cracked granite. *Geophysical Research Letters*, 21(11):979–982.

[153] Ghanbarian, B., Hunt, A. G., Ewing, R. P., and Sahimi, M. (2013). Tortuosity in porous media: a critical review. *Soil science Society of America Journal*, 77(5):1461–1477.

[154] Ghanizadeh, A., Clarkson, C., Aquino, S., Ardakani, O., and Sanei, H. (2015). Petrophysical and geomechanical characteristics of canadian tight oil and liquid-rich gas reservoirs: I. pore network and permeability characterization. *Fuel*, 153:664–681.

[155] Giesche, H. (2006). Mercury porosimetry: a general (practical) overview. *Particle & Particle Systems Characterization*, 23(1):9–19.

[156] Ginzbourg, I. and Adler, P. (1994). Boundary flow condition analysis for the three-dimensional lattice boltzmann model. *Journal de Physique II*, 4(2):191–214.

[157] Ginzburg, I. and d'Humieres, D. (2003). Multireflection boundary conditions for lattice boltzmann models. *Physical Review E*, 68(6):066614.

[158] Gomez, C. T., Dvorkin, J., and Vanorio, T. (2010). Laboratory measurements of porosity, permeability, resistivity, and velocity on fontainebleau sandstoneslaboratory measurements on sandstones. *Geophysics*, 75(6):E191–E204.

[159] Gommes, C. J., Bons, A.-J., Blacher, S., Dunsmuir, J. H., and Tsou, A. H. (2009). Practical methods for measuring the tortuosity of porous materials from binary or gray-tone tomographic reconstructions. *AIChE Journal*, 55(8):2000–2012.

[160] Goodarzi, M., Dejaegher, B., and Heyden, Y. V. (2012). Feature selection methods in qsar studies. *Journal of AOAC International*, 95(3):636–651.

[161] Gooya, R., Bruns, S., Müter, D., Moaddel, A., Harti, R. P., Stipp, S. L. S., and Sørensen, H. O. (2016). Effect of tomography resolution on the calculated microscopic properties of porous materials: Comparison of sandstone and carbonate rocks. *Applied Physics Letters*, 109(10):104102.

[162] Gostovic, D., Smith, z. J., Kundinger, D., Jones, K., and Wachsman, E. (2007). Three-dimensional reconstruction of porous lscf cathodes. *Electrochemical and Solid-state Letters*, 10(12):B214–B217.

[163] Grathwohl, P. (2012). *Diffusion in Natural Porous Media: Contaminant Transport, Sorption/Desorption and Dissolution Kinetics*, volume 1. Springer Science & Business Media.

[164] Guan, K. M., Nazarova, M., Guo, B., Tchelepi, H., Kovscek, A. R., and Creux, P. (2018). Effects of image resolution on sandstone porosity and permeability as obtained from x-ray microscopy. *Transport in Porous Media*, pages 1–13.

[165] Guan, K. M., Nazarova, M., Guo, B., Tchelepi, H., Kovscek, A. R., and Creux, P. (2019). Effects of image resolution on sandstone porosity and permeability as obtained from x-ray microscopy. *Transport in Porous Media*, 127(1):233–245.

[166] Guo, Z. and Zhao, T. (2005). A lattice boltzmann model for convection heat transfer in porous media. *Numerical Heat Transfer, Part B*, 47(2):157–177.

[167] Gupta, A., Cecen, A., Goyal, S., Singh, A. K., and Kalidindi, S. R. (2015). Structure–property linkages using a data science approach: application to a non-metallic inclusion/steel composite system. *Acta Materialia*, 91:239–254.

[168] Guyon, I. and Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3(Mar):1157–1182.

[169] Guyon, I., Gunn, S., Nikravesh, M., and Zadeh, L. A. (2008). *Feature Extraction: Foundations and Applications*, volume 207. Springer.

[170] Hadley, P. W. and Newell, C. (2014). The new potential for understanding groundwater contaminant transport. *Groundwater*, 52(2):174–186.

[171] Hajizadeh, A. and Farhadpour, Z. (2012). An algorithm for 3d pore space reconstruction from a 2d image using sequential simulation and gradual deformation with the probability perturbation sampler. *Transport in Porous Media*, 94(3):859–881.

[172] Harlow, F. H. and Welch, J. E. (1965). Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *The Physics of Fluids*, 8(12):2182–2189.

[173] Hassouna, M. S. and Farag, A. A. (2007). Multistencils fast marching methods: A highly accurate solution to the eikonal equation on cartesian domains. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(9):1563–1574.

[174] Hauswirth, S. C., Abou Najm, M. R., and Miller, C. T. (2019). Characterization of the pore structure of porous media using non-newtonian fluids. *Water Resources Research*, 55(8):7182–7195.

[175] Havelka, J., Kučerová, A., and Sỳkora, J. (2016). Compression and reconstruction of random microstructures using accelerated lineal path function. *Computational Materials Science*, 122:102–117.

[176] He, W., Zou, J., Wang, B., Vilayurganapathy, S., Zhou, M., Lin, X., Zhang, K. H., Lin, J., Xu, P., and Dickerson, J. H. (2013). Gas transport in porous electrodes of solid oxide fuel cells: a review on diffusion and diffusivity measurement. *Journal of Power Sources*, 237:64–73.

[177] He, X., Zou, Q., Luo, L.-S., and Dembo, M. (1997). Analytic solutions of simple flows and analysis of nonslip boundary conditions for the lattice boltzmann bgk model. *Journal of Statistical Physics*, 87(1-2):115–136.

[178] Hemes, S., Desbois, G., Urai, J. L., Schröppel, B., and Schwarz, J.-O. (2015). Multi-scale characterization of porosity in boom clay (hades-level, mol, belgium) using a combination of x-ray μ-ct, 2d bib-sem and fib-sem tomography. *Microporous and Mesoporous Materials*, 208:1–20.

[179] Henderson, N., Brêttas, J. C., and Sacco, W. F. (2010). A three-parameter kozeny–carman generalized equation for fractal porous media. *Chemical Engineering Science*, 65(15):4432–4442.

[180] Herring, A., Sheppard, A., Turner, M., and Beeching, L. (2018). Multiphase flows in sandstones. http://www.digitalrocksportal.org/projects/135.

[181] Herring, A. L., Andersson, L., Schlüter, S., Sheppard, A., and Wildenschild, D. (2015). Efficiently engineering pore-scale processes: The role of force dominance and topology during nonwetting phase trapping in porous media. *Advances in Water Resources*, 79:91–102.

[182] Hidajat, I., Rastogi, A., Singh, M., Mohanty, K., et al. (2002). Transport properties of porous media reconstructed from thin-sections. *SPE Journal*, 7(01):40–48.

[183] Hilfer, R. (1991). Geometric and dielectric characterization of porous media. *Physical Review B*, 44(1):60.

[184] Hilfer, R. (1992). Local-porosity theory for flow in porous media. *Physical Review B*, 45(13):7115.

[185] Hilfer, R. (1996). Transport and relaxation phenomena in porous media. *Advances in Chemical Physics*, 92:299–424.

[186] Hilfer, R. (2002). Review on scale dependent characterization of the microstructure of porous media. *Transport in Porous Media*, 46(2-3):373–390.

[187] Hilfer, R., Rage, T., and Virgin, B. (1997). Local percolation probabilities for a natural sandstone. *Physica A: Statistical Mechanics and its Applications*, 241(1-2):105–110.

[188] Hilfer, R. and Zauner, T. (2011). High-precision synthetic computed tomography of reconstructed porous media. *Physical Review E*, 84(6):062301.

[189] Hillel, D. (2003). *Introduction to Environmental Soil Physics*. Elsevier.

[190] Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507.

[191] Hlushkou, D., Liasneuski, H., Tallarek, U., and Torquato, S. (2015). Effective diffusion coefficients in random packings of polydisperse hard spheres from two-point and three-point correlation functions. *Journal of Applied Physics*, 118(12):124901.

[192] Holloway, S. (2007). Carbon dioxide capture and geological storage. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1853):1095–1107.

[193] Holt, J., Paziresh, M., Kingston, A., and Sheppard, A. (2014). Correction of beam hardening artefacts in microtomography for samples imaged in containers. In *Developments in X-Ray Tomography IX*, volume 9212, page 92120A. International Society for Optics and Photonics.

[194] Holzer, L., Iwanschitz, B., Hocker, T., Münch, B., Prestat, M., Wiedenmann, D., Vogt, U., Holtappels, P., Sfeir, J., Mai, A., et al. (2011). Microstructure degradation of cermet anodes for solid oxide fuel cells: Quantification of nickel grain growth in dry and in humid atmospheres. *Journal of Power Sources*, 196(3):1279–1294.

[195] Holzer, L., Wiedenmann, D., Münch, B., Keller, L., Prestat, M., Gasser, P., Robertson, I., and Grobéty, B. (2013). The influence of constrictivity on the effective transport properties of porous layers in electrolysis and fuel cells. *Journal of Materials Science*, 48(7):2934–2952.

[196] Hong, J. and Liu, J. (2020). Rapid estimation of permeability from digital rock using 3d convolutional neural network. *Computational Geosciences*, 24:1523–1539.

[197] Hormann, K., Baranau, V., Hlushkou, D., Höltzel, A., and Tallarek, U. (2016). Topological analysis of non-granular, disordered porous media: determination of pore connectivity, pore coordination, and geometric tortuosity in physically reconstructed silica monoliths. *New Journal of Chemistry*, 40(5):4187–4199.

[198] Hosa, A., Curtis, A., and Wood, R. (2016). Calibrating lattice boltzmann flow simulations and estimating uncertainty in the permeability of complex porous media. *Advances in Water Resources*, 94:60–74.

[199] Hosticka, B., Norris, P., Brenizer, J., and Daitch, C. (1998). Gas flow through aerogels. *Journal of Non-crystalline Solids*, 225:293–297.

[200] Hu, J. and Stroeven, P. (2005). Local porosity analysis of pore structure in cement paste. *Cement and Concrete Research*, 35(2):233–242.

[201] Hu, L. B., Savidge, C., Rizzo, D. M., Hayden, N., Hagadorn, J. W., and Dewoolkar, M. (2012). Commonly used porous building materials: Geomorphic pore structure and fluid transport. *Journal of Materials in Civil Engineering*, 25(12):1803–1812.

[202] Huang, Z., Zhang, Y., Xie, L., Zhao, P., He, B., and Ren, L. (2020). Comparative study of porosity test methods for shale. *Arabian Journal of Geosciences*, 13(2):94.

[203] Humphreys, F. (2004). Characterisation of fine-scale microstructures by electron backscatter diffraction (ebsd). *Scripta Materialia*, 51(8):771–776.

[204] Hunt, A., Ewing, R., and Ghanbarian, B. (2014). *Percolation Theory for Flow in Porous Media*, volume 880. Springer.

[205] Hutzenlaub, T., Asthana, A., Becker, J., Wheeler, D., Zengerle, R., and Thiele, S. (2013). Fib/sem-based calculation of tortuosity in a porous licoo2 cathode for a li-ion battery. *Electrochemistry Communications*, 27:77–80.

[206] Iassonov, P., Gebrenegus, T., and Tuller, M. (2009). Segmentation of x-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures. *Water Resources Research*, 45(9).

[207] ImageJ (2016). Website. https://imagej.net/Welcome.

[208] Ioannidis, M., Kwiecien, M., and Chatzis, I. (1996). Statistical analysis of the porous microstructure as a method for estimating reservoir permeability. *Journal of Petroleum Science and Engineering*, 16(4):251–261.

[209] Ioannidis, M., Kwiecien, M., and Chatzis, I. (1997). Electrical conductivity and percolation aspects of statistically homogeneous porous media. *Transport in Porous Media*, 29(1):61–83.

[210] Iversen, N. and Jørgensen, B. B. (1993). Diffusion coefficients of sulfate and methane in marine sediments: Influence of porosity. *Geochimica et Cosmochimica Acta*, 57(3):571–578.

[211] Iwai, H., Shikazono, N., Matsui, T., Teshima, H., Kishimoto, M., Kishida, R., Hayashi, D., Matsuzaki, K., Kanno, D., Saito, M., et al. (2010). Quantification of sofc anode microstructure based on dual beam fib-sem technique. *Journal of Power Sources*, 195(4):955–961.

[212] Javadpour, F. et al. (2009). Nanopores and apparent permeability of gas flow in mudrocks (shales and siltstone). *Journal of Canadian Petroleum Technology*, 48(08):16–21.

[213] Jiang, Z., Chen, W., and Burkhart, C. (2013). Efficient 3d porous microstructure reconstruction via gaussian random field and hybrid optimization. *Journal of Microscopy*, 252(2):135–148.

[214] Jiao, Y. and Chawla, N. (2014). Modeling and characterizing anisotropic inclusion orientation in heterogeneous material via directional cluster functions and stochastic microstructure reconstruction. *Journal of Applied Physics*, 115(9):093511.

[215] Jiao, Y., Stillinger, F., and Torquato, S. (2008). Modeling heterogeneous materials via two-point correlation functions. ii. algorithmic details and applications. *Physical Review E*, 77(3):031135.

[216] Jiao, Y., Stillinger, F., and Torquato, S. (2009). A superior descriptor of random textures and its predictive capacity. *Proceedings of the National Academy of Sciences*, 106(42):17634–17639.

[217] Jiao, Y., Stillinger, F. H., and Torquato, S. (2007). Modeling heterogeneous materials via two-point correlation functions: Basic principles. *Physical Review E*, 76(3):031110.

[218] Jin, C., Langston, P. A., Pavlovskaya, G. E., Hall, M. R., and Rigby, S. P. (2016). Statistics of highly heterogeneous flow fields confined to three-dimensional random porous media. *Physical Review E*, 93(1):013122.

[219] Jin, G., Patzek, T., and Silin, D. (2004). Direct prediction of the absolute permeability of unconsolidated and consolidated reservoir rock. spe 90084. In *2003 SPE Annual Technical Conference and Exhibition (Houston, Texas, USA), SPE*.

[220] Jones, B. and Feng, Y. (2016). Effect of image scaling and segmentation in digital rock characterisation. *Computational Particle Mechanics*, 3(2):201–213.

[221] Jones, S. B. and Or, D. (2002). Surface area, geometrical and configurational effects on permittivity of porous media. *Journal of Non-Crystalline Solids*, 305(1-3):247–254.

[222] Jørgensen, P. S., Ebbehøj, S. L., and Hauch, A. (2015). Triple phase boundary specific pathway analysis for quantitative characterization of solid oxide cell electrode microstructure. *Journal of Power Sources*, 279:686–693.

[223] Jørgensen, P. S., Hansen, K. V., Larsen, R., and Bowen, J. R. (2011). Geometrical characterization of interconnected phase networks in three dimensions. *Journal of Microscopy*, 244(1):45–58.

[224] Ju, Y., Zheng, J., Epstein, M., Sudak, L., Wang, J., and Zhao, X. (2014). 3d numerical reconstruction of well-connected porous structure of rock using fractal algorithms. *Computer Methods in Applied Mechanics and Engineering*, 279:212–226.

[225] Kaczmarek, Ł. D., Zhao, Y., Konietzky, H., Wejrzanowski, T., and Maksimczuk, M. (2017). Numerical approach in recognition of selected features of rock structure from hybrid hydrocarbon reservoir samples based on microtomography. *Studia Geotechnica et Mechanica*, 39(1):13–26.

[226] Kaestner, A., Lehmann, E., and Stampanoni, M. (2008). Imaging and image processing in porous media research. *Advances in Water Resources*, 31(9):1174–1187.

[227] Kamrava, S., Tahmasebi, P., and Sahimi, M. (2020). Linking morphology of porous media to their macroscopic permeability by deep learning. *Transport in Porous Media*, 131(2):427–448.

[228] Karanikola, V., Corral, A. F., Jiang, H., Sáez, A. E., Ela, W. P., and Arnold, R. G. (2017). Effects of membrane structure and operational variables on membrane distillation performance. *Journal of Membrane Science*, 524:87–96.

[229] Karimpouli, S. and Tahmasebi, P. (2017). A hierarchical sampling for capturing permeability trend in rock physics. *Transport in Porous Media*, 116(3):1057–1072.

[230] Karimpouli, S. and Tahmasebi, P. (2019). Image-based velocity estimation of rock using convolutional neural networks. *Neural Networks*, 111:89–97.

[231] Karsanina, M. V., Gerke, K. M., Skvortsova, E. B., and Mallants, D. (2015). Universal spatial correlation functions for describing and reconstructing soil microstructure. *PloS One*, 10(5):e0126515.

[232] Katz, A. and Thompson, A. (1985). Fractal sandstone pores: implications for conductivity and pore formation. *Physical Review Letters*, 54(12):1325.

[233] Kaviany, M. (2012). *Principles of Heat Transfer in Porous Media*. Springer Science & Business Media.

[234] Keehm, Y. and Mukerji, T. (2004). Permeability and relative permeability from digital rocks: Issues on grid resolution and representative elementary volume. In *SEG Technical Program Expanded Abstracts 2004*, pages 1654–1657. Society of Exploration Geophysicists.

[235] Kehrwald, D., Shearing, P. R., Brandon, N. P., Sinha, P. K., and Harris, S. J. (2011). Local tortuosity inhomogeneities in a lithium battery composite electrode. *Journal of The Electrochemical Society*, 158(12):A1393–A1399.

[236] Keller, L. M., Hilger, A., and Manke, I. (2015). Impact of sand content on solute diffusion in opalinus clay. *Applied Clay Science*, 112:134–142.

[237] Keller, L. M., Holzer, L., Wepf, R., and Gasser, P. (2011). 3d geometry and topology of pore pathways in opalinus clay: Implications for mass transport. *Applied Clay Science*, 52(1-2):85–95.

[238] Ketcham, R. A. and Carlson, W. D. (2001). Acquisition, optimization and interpretation of x-ray computed tomographic imagery: applications to the geosciences. *Computers & Geosciences*, 27(4):381–400.

[239] Kim, J.-W., Kim, D., and Lindquist, W. (2013). A re-examination of throats. *Water Resources Research*, 49(11):7615–7626.

[240] Kim, J.-W., Perfect, E., and Choi, H. (2007). Anomalous diffusion in two-dimensional euclidean and prefractal geometrical models of heterogeneous porous media. *Water resources research*, 43(1).

[241] Kishimoto, M., Iwai, H., Miyawaki, K., Saito, M., and Yoshida, H. (2013). Improvement of the sub-grid-scale model designed for 3d numerical simulation of solid oxide fuel cell electrodes using an adaptive power index. *Journal of Power Sources*, 223:268–276.

[242] Knudby, C. and Carrera, J. (2005). On the relationship between indicators of geostatistical, flow and transport connectivity. *Advances in Water Resources*, 28(4):405–421.

[243] Kohanpur, A. H., Valocchi, A., and Crandall, D. (2019). Micro-ct images of a heterogeneous mt. simon sandstone sample. http://www.digitalrocksportal.org/projects/247.

[244] Kohavi, R., John, G. H., et al. (1997). Wrappers for feature subset selection. *Artificial Intelligence*, 97(1-2):273–324.

[245] Kolitcheff, S., Jolimaitre, E., Hugon, A., Verstraete, J., Carrette, P.-L., and Tayakout-Fayolle, M. (2017). Tortuosity of mesoporous alumina catalyst supports: Influence of the pore network organization. *Microporous and Mesoporous Materials*, 248:91–98.

[246] Koponen, A., Kataja, M., and Timonen, J. (1997). Permeability and effective porosity of porous media. *Physical Review E*, 56(3):3319.

[247] Koponen, A., Kataja, M., and Timonen, J. v. (1996). Tortuous flow in porous media. *Physical Review E*, 54(1):406.

[248] Koster, A., Ziese, U., Verkleij, A., Janssen, A., Graaf, J. d., Geus, J. W., De Jong, K., et al. (2000). Development and application of 3-dimensional transmission electron microscopy (3d-tem) for the characterization of metal-zeolite catalyst systems.

[249] Kotsiantis, S. B., Zaharakis, I., and Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, 160:3–24.

[250] Kozeny, J. (1927). Über kapillare leitung des wassers im boden, vol. 136. *Sitzungsber Akad Wiss, Wien*.

[251] Kreamer, D. K., Weeks, E. P., and Thompson, G. M. (1988). A field technique to measure the tortuosity and sorption-affected porosity for gaseous diffusion of materials in the unsaturated zone with experimental results from near barnwell, south carolina. *Water Resources Research*, 24(3):331–341.

[252] Krüger, T., Kusumaatmaja, H., Kuzmin, A., Shardt, O., Silva, G., and Viggen, E. M. (2017). The lattice boltzmann method. *Springer International Publishing*, 10:978–3.

[253] Kuhn, M., Johnson, K., et al. (2013). *Applied Predictive Modeling*, volume 26. Springer.

[254] Kuila, U., McCarty, D. K., Derkowski, A., Fischer, T. B., and Prasad, M. (2014). Total porosity measurement in gas shales by the water immersion porosimetry (wip) method. *Fuel*, 117:1115–1129.

[255] Kuila, U. and Prasad, M. (2013). Specific surface area and pore-size distribution in clays and shales. *Geophysical Prospecting*, 61(2):341–362.

[256] Kumar, M. and Han, D.-h. (2005). Pore shape effect on elastic properties of carbonate rocks. In *SEG Technical Program Expanded Abstracts 2005*, pages 1477–1480. Society of Exploration Geophysicists.

[257] Kumar, R. and Bhattacharjee, B. (2004). Assessment of permeation quality of concrete through mercury intrusion porosimetry. *Cement and Concrete Research*, 34(2):321–328.

[258] Kutay, M. E., Aydilek, A. H., and Masad, E. (2006). Laboratory validation of lattice boltzmann method for modeling pore-scale flow in granular materials. *Computers and Geotechnics*, 33(8):381–395.

[259] Lai, J., Wang, G., Wang, Z., Chen, J., Pang, X., Wang, S., Zhou, Z., He, Z., Qin, Z., and Fan, X. (2018). A review on pore structure characterization in tight sandstones. *Earth-Science Reviews*, 177:436–457.

[260] Landesfeind, J., Hattendorff, J., Ehrl, A., Wall, W. A., and Gasteiger, H. A. (2016). Tortuosity determination of battery electrodes and separators by impedance spectroscopy. *Journal of The Electrochemical Society*, 163(7):A1373–A1387.

[261] Lanfrey, P.-Y., Kuzeljevic, Z., and Dudukovic, M. (2010). Tortuosity model for fixed beds randomly packed with identical particles. *Chemical Engineering Science*, 65(5):1891–1896.

[262] Latief, F., Biswal, B., Fauzi, U., and Hilfer, R. (2010). Continuum reconstruction of the pore scale microstructure for fontainebleau sandstone. *Physica A: Statistical Mechanics and its Applications*, 389(8):1607–1618.

[263] Latief, F., Fauzi, U., Irayani, Z., and Dougherty, G. (2017). The effect of x-ray micro computed tomography image resolution on flow properties of porous rocks. *Journal of Microscopy*, 266(1):69–88.

[264] Lee, S. and Tank, R. (1985). Role of clays in the disposal of nuclear waste: a review. *Applied Clay Science*, 1(1-2):145–162.

[265] Lee, T.-C., Kashyap, R. L., and Chu, C.-N. (1994). Building skeleton models via 3-d medial surface axis thinning algorithms. *CVGIP: Graphical Models and Image Processing*, 56(6):462–478.

[266] Legland, D., Kiêu, K., and Devaux, M.-F. (2007). Computation of minkowski measures on 2d and 3d binary images. *Image Analysis &Stereology*, 26(2):83–92.

[267] Lehmann, P., Berchtold, M., Ahrenholz, B., Tölke, J., Kaestner, A., Krafczyk, M., Flühler, H., and Künsch, H. (2008). Impact of geometrical properties on permeability and fluid phase distribution in porous media. *Advances in Water Resources*, 31(9):1188–1204.

[268] Leu, L., Berg, S., Enzmann, F., Armstrong, R. T., and Kersten, M. (2014). Fast x-ray micro-tomography of multiphase flow in berea sandstone: A sensitivity study on image processing. *Transport in Porous Media*, 105(2):451–469.

[269] LeVeque, R. J. (2007). *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*, volume 98. Siam.

[270] Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., and Liu, H. (2017a). Feature selection: A data perspective. *ACM Computing Surveys (CSUR)*, 50(6):1–45.

[271] Li, J., Du, Q., and Sun, C. (2009). An improved box-counting method for image fractal dimension estimation. *Pattern Recognition*, 42(11):2460–2469.

[272] Li, S. Z. (2009). *Markov random field modeling in image analysis*. Springer Science & Business Media.

[273] Li, T., Schlüter, S., Dragila, M. I., and Wildenschild, D. (2018a). An improved method for estimating capillary pressure from 3d microtomography images and its application to the study of disconnected nonwetting phase. *Advances in Water Resources*, 114:249–260.

[274] Li, W., Zou, C.-c., Wang, H., and Peng, C. (2017b). A model for calculating the formation resistivity factor in low and middle porosity sandstone formations considering the effect of pore geometry. *Journal of Petroleum Science and Engineering*, 152:193–203.

[275] Li, X., Liu, Z., Cui, S., Luo, C., Li, C., and Zhuang, Z. (2019). Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning. *Computer Methods in Applied Mechanics and Engineering*, 347:735–753.

[276] Li, X., Zhang, Y., Zhao, H., Burkhart, C., Brinson, L. C., and Chen, W. (2018b). A transfer learning approach for microstructure reconstruction and structure-property predictions. *Scientific Reports*, 8.

[277] Liang, Z., Fernandes, C., Magnani, F., and Philippi, P. (1998). A reconstruction technique for three-dimensional porous media using image analysis and fourier transforms. *Journal of Petroleum Science and Engineering*, 21(3-4):273–283.

[278] Liang, Z., Ioannidis, M., and Chatzis, I. (2000). Permeability and electrical conductivity of porous media from 3d stochastic replicas of the microstructure. *Chemical Engineering Science*, 55(22):5247–5262.

[279] Liebovitch, L. S. and Toth, T. (1989). A fast algorithm to determine fractal dimensions by box counting. *Physics Letters A*, 141(8-9):386–390.

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

[281] Lindquist, W. B., Venkatarangan, A., Dunsmuir, J., and Wong, T.-f. (2000). Pore and throat size distributions measured from synchrotron x-ray tomographic images of fontainebleau sandstones. *Journal of Geophysical Research: Solid Earth*, 105(B9):21509–21527.

[282] Litster, S., Epting, W., Wargo, E., Kalidindi, S., and Kumbur, E. (2013). Morphological analyses of polymer electrolyte fuel cell electrodes with nano-scale computed tomography imaging. *Fuel Cells*, 13(5):935–945.

[283] Liu, J., Pereira, G. G., Liu, Q., and Regenauer-Lieb, K. (2016). Computational challenges in the analyses of petrophysics using microtomography and upscaling: a review. *Computers & Geosciences*, 89:107–117.

[284] Liu, T., Jin, X., and Wang, M. (2018). Critical resolution and sample size of digital rock analysis for unconventional reservoirs. *Energies*, 11(7):1798.

[285] Liu, X. and Shapiro, V. (2015). Random heterogeneous materials via texture synthesis. *Computational Materials Science*, 99:177–189.

[286] Liu, Z., Herring, A., Arns, C., Berg, S., and Armstrong, R. T. (2017). Pore-scale characterization of two-phase flow using integral geometry. *Transport in Porous Media*, 118(1):99–117.

[287] Lively, B., Bizga, J., Li, B., Smith, M., and Zhong, W.-H. (2014). Quantified non-destructive dispersion assessment via macrodispersion and ac/dc ratio analyses for hybrid polycarbonate/graphite nanoplatelet/carbon nanotube composites. *Composites Part B: Engineering*, 56:567–574.

[288] Llewellin, E. (2010). Lbflow: An extensible lattice boltzmann framework for the simulation of geophysical flows. part i: theory and implementation. *Computers & Geosciences*, 36(2):115–122.

[289] Loh, W.-Y. (2002). Regression tress with unbiased variable selection and interaction detection. *Statistica sinica*, pages 361–386.

[290] Lu, B. and Torquato, S. (1990). Local volume fraction fluctuations in heterogeneous media. *The Journal of Chemical physics*, 93(5):3452–3459.

[291] Lu, B. and Torquato, S. (1992). Lineal-path function for random heterogeneous materials. *Physical Review A*, 45(2):922.

[292] Lu, S., Landis, E., and Keane, D. (2006). X-ray microtomographic studies of pore structure and permeability in portland cement concrete. *Materials and Structures*, 39(6):611–620.

[293] Lu, X. (1997). Application of the nernst-einstein equation to concrete. *Cement and Concrete Research*, 27(2):293–302.

[294] Luhmann, A. J., Tutolo, B. M., Bagley, B. C., Mildner, D. F., Seyfried Jr, W. E., and Saar, M. O. (2017). Permeability, porosity, and mineral surface area changes in basalt cores induced by reactive transport of co 2-rich brine. *Water Resources Research*, 53(3):1908–1927.

[295] Luo, L., Yu, B., Cai, J., and Zeng, X. (2014). Numerical simulation of tortuosity for fluid flow in two-dimensional pore fractal models of porous media. *Fractals*, 22(04):1450015.

[296] Ma, Z. and Torquato, S. (2018). Precise algorithms to compute surface correlation functions of two-phase heterogeneous media and their applications. *Physical Review E*, 98(1):013307.

[297] Machado Charry, E., Neumann, M., Lahti, J., Schennach, R., Schmidt, V., and Zojer, K. (2018). Pore space extraction and characterization of sack paper using μ-ct. *Journal of Microscopy*, 272(1):35–46.

[298] Malek, K. and Coppens, M.-O. (2003). Knudsen self-and fickian diffusion in rough nanoporous media. *The Journal of Chemical Physics*, 119(5):2801–2811.

[299] Marty, N. C., Fritz, B., Clément, A., and Michau, N. (2010). Modelling the long term alteration of the engineered bentonite barrier in an underground radioactive waste repository. *Applied Clay Science*, 47(1-2):82–90.

[300] Masís-Meléndez, F., Deepagoda, T. C., de Jonge, L. W., Tuller, M., and Moldrup, P. (2014). Gas diffusion-derived tortuosity governs saturated hydraulic conductivity in sandy soils. *Journal of Hydrology*, 512:388–396.

[301] Matyka, M., Khalili, A., and Koza, Z. (2008). Tortuosity-porosity relation in porous media flow. *Physical Review E*, 78(2):026306.

[302] Matyka, M. and Koza, Z. (2012). How to calculate tortuosity easily? In *AIP Conference Proceedings 4*, volume 1453, pages 17–22. AIP.

[303] Mauret, E. and Renaud, M. (1997). Transport phenomena in multi-particle systems—i. limits of applicability of capillary model in high voidage beds-application to fixed beds of fibers and fluidized beds of spheres. *Chemical Engineering Science*, 52(11):1807–1817.

[304] McClure, J. E., Berrill, M. A., Gray, W. G., and Miller, C. T. (2016). Influence of phase connectivity on the relationship among capillary pressure, fluid saturation, and interfacial area in two-fluid-phase porous medium systems. *Physical Review E*, 94(3):033102.

[305] Meakin, P. and Tartakovsky, A. M. (2009). Modeling and simulation of pore-scale multiphase fluid flow and reactive transport in fractured and porous media. *Reviews of Geophysics*, 47(3).

[306] Mecke, K. and Arns, C. (2005). Fluids in porous media: a morphometric approach. *Journal of Physics: Condensed Matter*, 17(9):S503.

[307] Mecke, K. R. (2000). Additivity, convexity, and beyond: applications of minkowski functionals in statistical physics. In *Statistical Physics and Spatial Statistics*, pages 111–184. Springer.

[308] Mertz, J. (2019). *Introduction to Optical Microscopy*. Cambridge University Press.

[309] Meyer, F. (1994). Topographic distance and watershed lines. *Signal Processing*, 38(1):113–125.

[310] Michielsen, K. and De Raedt, H. (2001). Integral-geometry morphological image analysis. *Physics Reports*, 347(6):461–538.

[311] Moctezuma-Berthier, A., Vizika, O., and Adler, P. M. (2002). Macroscopic conductivity of vugular porous media. *Transport in Porous Media*, 49(3):313–332.

[312] Moldrup, P., Olesen, T., Komatsu, T., Schjønning, P., and Rolston, D. (2001). Tortuosity, diffusivity, and permeability in the soil liquid and gaseous phases. *Soil Science Society of America Journal*, 65(3):613–623.

[313] Montes, J., Cuevas, F., and Cintas, J. (2007). Electrical and thermal tortuosity in powder compacts. *Granular Matter*, 9(6):401–406.

[314] Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to Linear Regression Analysis*, volume 821. John Wiley & Sons.

[315] Mosser, L., Dubrule, O., and Blunt, M. J. (2017). Reconstruction of three-dimensional porous media using generative adversarial neural networks. *Physical Review E*, 96(4):043309.

[316] Mota, M., Teixeira, J. A., Bowen, W. R., and Yelshin, A. (2001). Binary spherical particle mixed beds: porosity and permeability relationship measurement.

[317] Moukalled, F., Mangani, L., Darwish, M., et al. (2016). The finite volume method in computational fluid dynamics. *An advanced introduction with OpenFoam® and Matlab®. Nueva York: Springer. Recuperado de http://www. gidropraktikum. narod. ru/Moukalled-et-al-FVM-OpenFOAM-Matlab. pdf*.

[318] Mualem, Y. and Friedman, S. (1991). Theoretical prediction of electrical conductivity in saturated and unsaturated soil. *Water Resources Research*, 27(10):2771–2777.

[319] Muche, L. and Stoyan, D. (1992). Contact and chord length distributions of the poisson voronoi tessellation. *Journal of Applied Probability*, pages 467–471.

[320] Münch, B., Gasser, P., Holzer, L., and Flatt, R. (2006). Fib-nanotomography of particulate systems—part ii: Particle recognition and effect of boundary truncation. *Journal of the American Ceramic Society*, 89(8):2586–2595.

[321] Münch, B. and Holzer, L. (2008). Contradicting geometrical concepts in pore size analysis attained with electron microscopy and mercury intrusion. *Journal of the American Ceramic Society*, 91(12):4059–4067.

[322] Münch, B., Trtik, P., Marone, F., and Stampanoni, M. (2009). Stripe and ring artifact removal with combined wavelet—fourier filtering. *Optics Express*, 17(10):8567–8591.

[323] Munroe, P. R. (2009). The application of focused ion beam microscopy in the material sciences. *Materials Characterization*, 60(1):2–13.

[324] Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

[325] Nakashima, Y. and Kamiya, S. (2007). Mathematica programs for the analysis of three-dimensional pore connectivity and anisotropic tortuosity of porous rocks using x-ray computed tomography image data. *Journal of Nuclear Science and Technology*, 44(9):1233–1247.

[326] Nakashima, Y. and Watanabe, Y. (2002). Estimate of transport properties of porous media by microfocus x-ray computed tomography and random walk simulation. *Water Resources Research*, 38(12).

[327] N'Diaye, M., Degeratu, C., Bouler, J.-M., and Chappard, D. (2013). Biomaterial porosity determined by fractal dimensions, succolarity and lacunarity on microcomputed tomographic images. *Materials Science and Engineering: C*, 33(4):2025–2030.

[328] Nelson, G., Harris, W., Lombardo, J., Izzo, J., Chiu, W. K., Tanasini, P., Cantoni, M., Van Herle, J., Comninellis, C., Andrews, J., et al. (2011). Comparison of x-ray nanotomography and fib-sem in quantifying the composite lsm/ysz sofc cathode microstructure. *ECS Transactions*, 35(1):2417–2421.

[329] Neumann, M., Hirsch, C., Staněk, J., Beneš, V., and Schmidt, V. (2019). Estimation of geodesic tortuosity and constrictivity in stationary random closed sets. *Scandinavian Journal of Statistics*, 46(3):848–884.

[330] Neumann, M., Staněk, J., Pecho, O. M., Holzer, L., Beneš, V., and Schmidt, V. (2016). Stochastic 3d modeling of complex three-phase microstructures in sofc-electrodes with completely connected phases. *Computational Materials Science*, 118:353–364.

[331] Newbury, D. E. and Ritchie, N. W. (2013). Elemental mapping of microstructures by scanning electron microscopy-energy dispersive x-ray spectrometry (sem-eds): extraordinary advances with the silicon drift detector (sdd). *Journal of Analytical Atomic Spectrometry*, 28(7):973–988.

[332] Ning, Y., He, S., Liu, H., Wang, H., Qin, G., et al. (2016). A rigorous upscaling procedure to predict macro-scale transport properties of natural gas in shales by coupling molecular dynamics with lattice boltzmann method. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

[333] Ohser, J. and Mücklich, F. (2000). *Statistical Analysis of Microstructures in Materials Science*. Wiley.

[334] Okabe, H. (2004). *Pore-scale modelling of carbonates*. PhD thesis, University of London.

[335] Okabe, H. and Blunt, M. J. (2004). Prediction of permeability for porous media reconstructed using multiple-point statistics. *Physical Review E*, 70(6):066135.

[336] Okabe, H. and Blunt, M. J. (2005). Pore space reconstruction using multiple-point statistics. *Journal of Petroleum Science and Engineering*, 46(1-2):121–137.

[337] Okamoto, I., Li, X., and Ohsumi, T. (2005). Effect of supercritical co2 as the organic solvent on cap rock sealing performance for underground storage. *Energy*, 30(11-12):2344–2351.

[338] Olives, R. and Mauran, S. (2001). A highly conductive porous medium for solid–gas reactions: effect of the dispersed phase on the thermal tortuosity. *Transport in Porous Media*, 43(2):377–394.

[339] OpenLB (2019). Website. https://www.openlb.net/.

[340] Ordóñez, S., Fort, R., and Del Cura, M. G. (1997). Pore size distribution and the durability of a porous limestone. *Quarterly Journal of Engineering Geology and Hydrogeology*, 30(3):221–230.

[341] Øren, P.-E. and Bakke, S. (2002). Process based reconstruction of sandstones and prediction of transport properties. *Transport in Porous Media*, 46(2-3):311–343.

[342] Ostadi, H., Rama, P., Liu, Y., Chen, R., Zhang, X., and Jiang, K. (2010). 3d reconstruction of a gas diffusion layer and a microporous layer. *Journal of Membrane Science*, 351(1-2):69–74.

[343] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.

[344] Ozgumus, T., Mobedi, M., and Ozkol, U. (2014). Determination of kozeny constant based on porosity and pore to throat size ratio in porous medium with rectangular rods. *Engineering Applications of Computational Fluid Mechanics*, 8(2):308–318.

[345] Palabos (2019). Univeristy of Geneva. Geneva, Switzerland. Website. https://palabos.unige.ch/.

[346] Pan, C., Hilpert, M., and Miller, C. T. (2001). Pore-scale modeling of saturated permeabilities in random sphere packings. *Physical Review E*, 64(6):066702.

[347] Pardo-Alonso, S., Vicente, J., Solórzano, E., Rodriguez-Perez, M. Á., and Lehmhus, D. (2014). Geometrical tortuosity 3d calculations in infiltrated aluminium cellular materials. *Procedia Materials Science*, 4:145–150.

[348] Patel, K. K., Paulsen, J. M., and Desilvestro, J. (2003). Numerical simulation of porous networks in relation to battery electrodes and separators. *Journal of Power Sources*, 122(2):144–152.

[349] Paterson, M. (1983). The equivalent channel model for permeability and resistivity in fluid-saturated rock—a re-appraisal. *Mechanics of Materials*, 2(4):345–352.

[350] Pendleton, D., Dathe, A., and Baveye, P. (2005). Influence of image resolution and evaluation algorithm on estimates of the lacunarity of porous media. *Physical Review E*, 72(4):041306.

[351] Peng, S., Hu, Q., Dultz, S., and Zhang, M. (2012a). Using x-ray computed tomography in pore structure characterization for a berea sandstone: Resolution effect. *Journal of Hydrology*, 472:254–261.

[352] Peng, S., Hu, Q., and Hamamoto, S. (2012b). Diffusivity of rocks: Gas diffusion measurements and correlation to porosity and pore size distribution. *Water Resources Research*, 48(2).

[353] Peng, S., Marone, F., and Dultz, S. (2014). Resolution effect in x-ray microcomputed tomography imaging and small pore's contribution to permeability for a berea sandstone. *Journal of Hydrology*, 510:403–411.

[354] Petersen, E. (1958). Diffusion in a pore of varying cross section. *AIChE Journal*, 4(3):343–345.

[355] Petersen, L. W., Moldrup, P., Jacobsen, O. H., and Rolston, D. (1996). Relations between specific surface area and soil physical and chemical properties. *Soil Science*, 161(1):9–21.

[356] Piller, M., Schena, G., Nolich, M., Favretto, S., Radaelli, F., and Rossi, E. (2009). Analysis of hydraulic permeability in porous media: from high resolution x-ray tomography to direct numerical simulation. *Transport in Porous Media*, 80(1):57.

[357] Pimienta, L., Sarout, J., Esteban, L., and Piane, C. D. (2014). Prediction of rocks thermal conductivity from elastic wave velocities, mineralogy and microstructure. *Geophysical Journal International*, 197(2):860–874.

[358] Pisani, L. (2011). Simple expression for the tortuosity of porous media. *Transport in Porous Media*, 88(2):193–203.

[359] Plawsky, J. L. (2020). *Transport Phenomena Fundamentals*. CRC press.

[360] Poon, C.-S., Kou, S., and Lam, L. (2006). Compressive strength, chloride diffusivity and pore structure of high performance metakaolin and silica fume concrete. *Construction and Building Materials*, 20(10):858–865.

[361] Prachayawarakorn, S. and Mann, R. (2007). Effects of pore assembly architecture on catalyst particle tortuosity and reaction effectiveness. *Catalysis Today*, 128(1-2):88–99.

[362] Pudney, C. (1998). Distance-ordered homotopic thinning: a skeletonization algorithm for 3d digital images. *Computer Vision and Image Understanding*, 72(3):404–413.

[363] Quiblier, J. A. (1984). A new three-dimensional modeling technique for studying porous media. *Journal of Colloid and Interface Science*, 98(1):84–102.

[364] Quintanilla, J. and Torquato, S. (1997). Local volume fraction fluctuations in random media. *The Journal of Chemical Physics*, 106(7):2741–2751.

[365] Rabbani, A., Ayatollahi, S., Kharrat, R., and Dashti, N. (2016). Estimation of 3-d pore network coordination number of rocks from watershed segmentation of a single 2-d image. *Advances in Water Resources*, 94:264–277.

[366] Rabbani, A., Babaei, M., Shams, R., Da Wang, Y., and Chung, T. (2020). Deepore: a deep learning workflow for rapid and comprehensive characterization of porous materials. *Advances in Water Resources*, page 103787.

[367] Rabbani, A. and Jamshidi, S. (2014). Specific surface and porosity relationship for sandstones for prediction of permeability. *International Journal of Rock Mechanics and Mining Sciences*, 71:25–32.

[368] Ramprasad, R., Batra, R., Pilania, G., Mannodi-Kanakkithodi, A., and Kim, C. (2017). Machine learning in materials informatics: recent applications and prospects. *NPJ Computational Materials*, 3(1):54.

[369] Ramstad, T., Idowu, N., Nardi, C., and Øren, P.-E. (2012). Relative permeability calculations from two-phase flow simulations directly on digital images of porous rocks. *Transport in Porous Media*, 94(2):487–504.

[370] Rao, P. and Schaefer, L. (2019). Lattice boltzmann models for micro-tomographic pore-spaces. *arXiv preprint arXiv:1902.11193*.

[371] Revil, A., Kessouri, P., and Torres-Verdín, C. (2014). Electrical conductivity, induced polarization, and permeability of the fontainebleau sandstone. *Geophysics*, 79(5):D301–D318.

[372] Roberts, A. and Garboczi, E. J. (2002). Computation of the linear elastic properties of random porous materials with a wide variety of microstructure. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 458(2021):1033–1054.

[373] Roberts, A. P. (1997). Statistical reconstruction of three-dimensional porous media from two-dimensional images. *Physical Review E*, 56(3):3203.

[374] Roberts, A. P. and Knackstedt, M. A. (1996). Structure-property correlations in model composite materials. *Physical Review E*, 54(3):2313.

[375] Robnik-Šikonja, M. and Kononenko, I. (2003). Theoretical and empirical analysis of relieff and rrelieff. *Machine Learning*, 53(1-2):23–69.

[376] Röding, M., Ma, Z., and Torquato, S. (2020). Predicting permeability via statistical learning on higher-order microstructural information. *Scientific reports*, 10(1):1–17.

[377] Rubinstein, J. and Torquato, S. (1988). Diffusion-controlled reactions: Mathematical formulation, variational principles, and rigorous bounds. *The Journal of Chemical Physics*, 88(10):6372–6380.

[378] Rubinstein, J. and Torquato, S. (1989). Flow in random porous media: mathematical formulation, variational principles, and rigorous bounds. *Journal of Fluid Mechanics*, 206:25–46.

[379] Sahimi, M. (2011). *Flow and Transport in Porous Media and Fractured Rock: From Classical Methods to Modern Approaches*. John Wiley & Sons.

[380] Saif, T., Lin, Q., Gao, Y., Al-Khulaifi, Y., Marone, F., Hollis, D., Blunt, M. J., and Bijeljic, B. (2019). 4d in situ synchrotron x-ray tomographic microscopy and laser-based heating study of oil shale pyrolysis. *Applied Energy*, 235:1468–1475.

[381] San Wu, Y., van Vliet, L. J., Frijlink, H. W., and van der Voort Maarschalk, K. (2006). The determination of relative path length as a measure for tortuosity in compacts using image analysis. *European Journal of Pharmaceutical Sciences*, 28(5):433–440.

[382] Santos, J. E., Xu, D., Jo, H., Landry, C. J., Prodanović, M., and Pyrcz, M. J. (2020). Poreflow-net: A 3d convolutional neural network to predict fluid flow through porous media. *Advances in Water Resources*, 138:103539.

[383] Saomoto, H. and Katagiri, J. (2015). Direct comparison of hydraulic tortuosity and electric tortuosity based on finite element analysis. *Theoretical and Applied Mechanics Letters*, 5(5):177–180.

[384] Saxena, N., Hofmann, R., Alpak, F. O., Berg, S., Dietderich, J., Agarwal, U., Tandon, K., Hunter, S., Freeman, J., and Wilson, O. B. (2017). References and benchmarks for pore-scale flow simulated using micro-ct images of porous media and digital rocks. *Advances in Water Resources*, 109:211–235.

[385] Saxena, N., Hows, A., Hofmann, R., Alpak, F. O., Freeman, J., Hunter, S., and Appel, M. (2018). Imaging and computational considerations for image computed permeability: Operating envelope of digital rock physics. *Advances in Water Resources*, 116:127–144.

[386] Scanziani, A., Singh, K., and Blunt, M. (2018). Water-wet three-phase flow micro-ct tomograms. http://www.digitalrocksportal.org/projects/167.

[387] Schlüter, S., Berg, S., Rücker, M., Armstrong, R., Vogel, H.-J., Hilfer, R., and Wildenschild, D. (2016). Pore-scale displacement mechanisms as a source of hysteresis for two-phase flow in porous media. *Water Resources Research*, 52(3):2194–2205.

[388] Schlüter, S., Sheppard, A., Brown, K., and Wildenschild, D. (2014). Image processing of multiphase images obtained via x-ray microtomography: a review. *Water Resources Research*, 50(4):3615–3639.

[389] Schlüter, S. and Vogel, H.-J. (2011). On the reconstruction of structural and functional properties in random heterogeneous media. *Advances in Water Resources*, 34(2):314–325.

[390] Schlüter, S., Weller, U., and Vogel, H.-J. (2011). Soil-structure development including seasonal dynamics in a long-term fertilization experiment. *Journal of Plant Nutrition and Soil Science*, 174(3):395–403.

[391] Schmalzing, J., Kerscher, M., and Buchert, T. (1995). Minkowski functionals in cosmology. *arXiv preprint astro-ph/9508154*.

[392] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117.

[393] Schneider, R. (2014). *Convex Bodies: the Brunn–Minkowski Theory*. Number 151. Cambridge university press.

[394] Scholz, C., Wirner, F., Götz, J., Rüde, U., Schröder-Turk, G. E., Mecke, K., and Bechinger, C. (2012). Permeability of porous materials determined from the euler characteristic. *Physical Review Letters*, 109(26):264504.

[395] Schulz, E., Speekenbrink, M., and Krause, A. (2018). A tutorial on gaussian process regression: Modelling, exploring, and exploiting functions. *Journal of Mathematical Psychology*, 85:1–16.

[396] Schwartz, L. M., Garboczi, E. J., and Bentz, D. P. (1995). Interfacial transport in porous media: Application to dc electrical conductivity of mortars. *Journal of Applied Physics*, 78(10):5898–5908.

[397] Scrivener, K. L. (2004). Backscattered electron imaging of cementitious microstructures: understanding and quantification. *Cement and Concrete Composites*, 26(8):935–945.

[398] Sellin, P. and Leupin, O. X. (2013). The use of clay as an engineered barrier in radioactive-waste management–a review. *Clays and Clay Minerals*, 61(6):477–498.

[399] Sethian, J. A. (1996). A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595.

[400] Sha, F., Xiao, L., Mao, Z., and Jia, C. (2019). Petrophysical characterization and fractal analysis of carbonate reservoirs of the eastern margin of the pre-caspian basin. *Energies*, 12(1):78.

[401] Shah, S., Gray, F., Crawshaw, J., and Boek, E. (2016). Micro-computed tomography pore-scale study of flow in porous media: Effect of voxel resolution. *Advances in Water Resources*, 95:276–287.

[402] Shanti, N. O., Chan, V. W., Stock, S. R., De Carlo, F., Thornton, K., and Faber, K. T. (2014). X-ray micro-computed tomography and tortuosity calculations of percolating pore networks. *Acta Materialia*, 71:126–135.

[403] Shearing, P. R., Brandon, N., Gelb, J., Bradley, R., Withers, P., Marquis, A., Cooper, S., and Harris, S. (2012). Multi length scale microstructural investigations of a commercially available li-ion battery electrode. *Journal of The Electrochemical Society*, 159(7):A1023–A1027.

[404] Shen, L. and Chen, Z. (2007). Critical review of the impact of tortuosity on diffusion. *Chemical Engineering Science*, 62(14):3748–3755.

[405] Shields, M., Deodatis, G., and Bocchini, P. (2011). A simple and efficient methodology to approximate a general non-gaussian stationary stochastic process by a translation process. *Probabilistic Engineering Mechanics*, 26(4):511–519.

[406] Shikhov, I. and Arns, C. H. (2015). Evaluation of capillary pressure methods via digital rock simulations. *Transport in Porous Media*, 107(2):623–640.

[407] Shin, H., Lindquist, W. B., Sahagian, D. L., and Song, S.-R. (2005). Analysis of the vesicular structure of basalts. *Computers & geosciences*, 31(4):473–487.

[408] Silberschmidt, V. V. (2008). Account for random microstructure in multiscale models. In *Multiscale Modeling and Simulation of Composite Materials and Structures*, pages 1–35. Springer.

[409] Simpleware (2019). Synopsys. California, USA. Website. http://www.simpleware.com.

[410] Sing, K. S. (1985). Reporting physisorption data for gas/solid systems with special reference to the determination of surface area and porosity (recommendations 1984). *Pure and Applied Chemistry*, 57(4):603–619.

[411] Sing, K. S. and Williams, R. T. (2004). The use of molecular probes for the characterization of nanoporous adsorbents. *Particle & Particle Systems Characterization: Measurement and Description of Particle Properties and Behavior in Powders and Other Disperse Systems*, 21(2):71–79.

[412] Singh, K., Menke, H., Andrew, M., Rau, C., Bijeljic, B., and Blunt, M. J. (2018). Time-resolved synchrotron x-ray micro-tomography datasets of drainage and imbibition in carbonate rocks. *Scientific Data*, 5(1):1–8.

[413] Skordos, P. (1993). Initial and boundary conditions for the lattice boltzmann method. *Physical Review E*, 48(6):4823.

[414] Smith, J., Chen, A., Gostovic, D., Hickey, D., Kundinger, D., Duncan, K., DeHoff, R., Jones, K., and Wachsman, E. (2009). Evaluation of the relationship between cathode microstructure and electrochemical behavior for sofcs. *Solid State Ionics*, 180(1):90–98.

[415] Sobieski, W., Matyka, M., Gołembiewski, J., and Lipiński, S. (2018). The path tracking method as an alternative for tortuosity determination in granular beds. *Granular Matter*, 20(4):72.

[416] Song, W., Yao, J., Li, Y., Yang, Y., and Sun, H. (2017). New pore size distribution calculation model based on chord length and digital image. *Journal of Natural Gas Science and Engineering*, 48:111–118.

[417] Song, Y., Davy, C. A., Kim, T. N., Troadec, D., Hauss, G., Jeannin, L., and Adler, P. M. (2016). Two-scale analysis of a tight gas sandstone. *Physical Review E*, 94(4):043316.

[418] Sors, L. A. S. and Santaló, L. A. (2004). *Integral Geometry and Geometric Probability*. Cambridge university press.

[419] Srisutthiyakorn, N. (2016). Deep-learning methods for predicting permeability from 2d/3d binary-segmented images. In *SEG technical program expanded abstracts 2016*, pages 3042–3046. Society of Exploration Geophysicists.

[420] Srisutthiyakorn, N. and Mavko, G. M. (2017). What is the role of tortuosity in the kozeny-carman equation? *Interpretation*, 5(1):SB57–SB67.

[421] Steinwart, I. and Christmann, A. (2008). *Support Vector Machines*. Springer Science & Business Media.

[422] Stenzel, O., Pecho, O., Holzer, L., Neumann, M., and Schmidt, V. (2016). Predicting effective conductivities based on geometric microstructure characteristics. *AIChE Journal*, 62(5):1834–1843.

[423] Stutzman, P. (2004). Scanning electron microscopy imaging of hydraulic cement microstructure. *Cement and Concrete Composites*, 26(8):957–966.

[424] Sudakov, O., Burnaev, E., and Koroteev, D. (2019). Driving digital rock towards machine learning: Predicting permeability with gradient boosting and deep neural networks. *Computers & Geosciences*, 127:91–98.

[425] Sukop, M. C., Huang, H., Alvarez, P. F., Variano, E. A., and Cunningham, K. J. (2013). Evaluation of permeability and non-darcy flow in vuggy macroporous limestone aquifer samples with lattice boltzmann methods. *Water Resources Research*, 49(1):216–230.

[426] Sun, F., Zielke, L., Marköter, H., Hilger, A., Zhou, D., Moroni, R., Zengerle, R., Thiele, S., Banhart, J., and Manke, I. (2016). Morphological evolution of electrochemically plated/stripped lithium microstructures investigated by synchrotron x-ray phase contrast tomography. *ACS Nano*, 10(8):7990–7997.

[427] Sun, W. and Wong, T.-f. (2018). Prediction of permeability and formation factor of sandstone with hybrid lattice boltzmann/finite element simulation on microtomographic images. *International Journal of Rock Mechanics and Mining Sciences*, 106:269–277.

[428] Sun, W. C., Andrade, J. E., and Rudnicki, J. W. (2011). Multiscale method for characterization of porous microstructures and their impact on macroscopic effective permeability. *International Journal for Numerical Methods in Engineering*, 88(12):1260–1279.

[429] Sundararaghavan, V. and Zabaras, N. (2005). Classification and reconstruction of three-dimensional microstructures using support vector machines. *Computational Materials Science*, 32(2):223–239.

[430] Tahmasebi, P. and Sahimi, M. (2012). Reconstruction of three-dimensional porous media using a single thin section. *Physical Review E*, 85(6):066709.

[431] Taiwo, O. O., Finegan, D. P., Eastwood, D. S., Fife, J. L., Brown, L. D., Darr, J. A., Lee, P. D., Brett, D. J., and Shearing, P. R. (2016). Comparison of three-dimensional analysis and stereological techniques for quantifying lithium-ion battery electrode microstructures. *Journal of Microscopy*, 263(3):280–292.

[432] Tamari, S. (2004). Optimum design of the constant-volume gas pycnometer for determining the volume of solid particles. *Measurement Science and Technology*, 15(3):549.

[433] Tariq, F., Yufit, V., Kishimoto, M., Shearing, P., Menkin, S., Golodnitsky, D., Gelb, J., Peled, E., and Brandon, N. (2014). Three-dimensional high resolution x-ray imaging and quantification of lithium ion battery mesocarbon microbead anodes. *Journal of Power Sources*, 248:1014–1020.

[434] Tembely, M., AlSumaiti, A. M., and Alameri, W. (2020). A deep learning perspective on predicting permeability in porous media from network modeling to direct simulation. *Computational Geosciences*, 24:1541–1556.

[435] Thiedmann, R., Hartnig, C., Manke, I., Schmidt, V., and Lehnert, W. (2009). Local structural characteristics of pore space in gdls of pem fuel cells based on geometric 3d graphs. *Journal of the Electrochemical Society*, 156(11):B1339.

[436] Tian, J., Qi, C., Sun, Y., and Yaseen, Z. M. (2020a). Surrogate permeability modelling of low-permeable rocks using convolutional neural networks. *Computer Methods in Applied Mechanics and Engineering*, 366:103–113.

[437] Tian, J., Qi, C., Sun, Y., Yaseen, Z. M., and Pham, B. T. (2020b). Permeability prediction of porous media using a combination of computational fluid dynamics and hybrid machine learning methods. *Engineering with Computers*, pages 1–17.

[438] Tjaden, B., Brett, D. J., and Shearing, P. R. (2018). Tortuosity in electrochemical devices: a review of calculation approaches. *International Materials Reviews*, 63(2):47–67.

[439] Tjaden, B., Finegan, D. P., Lane, J., Brett, D. J., and Shearing, P. R. (2017). Contradictory concepts in tortuosity determination in porous media in electrochemical devices. *Chemical Engineering Science*, 166:235–245.

[440] Tjaden, B., Lane, J., Withers, P. J., Bradley, R. S., Brett, D. J., and Shearing, P. R. (2016). The application of 3d imaging techniques, simulation and diffusion experiments to explore transport properties in porous oxygen transport membrane support materials. *Solid State Ionics*, 288:315–321.

[441] Torquato, S. (2002). Statistical description of microstructures. *Annual Review of Materials Research*, 32(1):77–111.

[442] Torquato, S. (2006). Necessary conditions on realizable two-point correlation functions of random media. *Industrial & Engineering Chemistry Research*, 45(21):6923–6928.

[443] Torquato, S. (2013). *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*, volume 16. Springer Science & Business Media.

[444] Torquato, S., Beasley, J., and Chiew, Y. (1988). Two-point cluster function for continuum percolation. *The Journal of Chemical Physics*, 88(10):6540–6547.

[445] Torquato, S. and Haslach Jr, H. (2002). Random heterogeneous materials: microstructure and macroscopic properties. *Applied Mechanics Reviews*, 55(4):B62–B63.

[446] Torquato, S. and Lu, B. (1993). Chord-length distribution function for two-phase random media. *Physical Review E*, 47(4):2950.

[447] Tranter, T., Kok, M., Lam, M., and Gostick, J. (2019). pytrax: A simple and efficient random walk implementation for calculating the directional tortuosity of images. *SoftwareX*, 10:100277.

[448] Tye, F. (1983). Tortuosity. *Journal of Power Sources*, 9:89–100.

[449] Van Brakel, J. and Heertjes, P. (1974). Analysis of diffusion in macroporous media in terms of a porosity, a tortuosity and a constrictivity factor. *International Journal of Heat and Mass Transfer*, 17(9):1093–1103.

[450] van der Linden, J. H., Narsilio, G. A., and Tordesillas, A. (2016). Machine learning framework for analysis of transport through complex networks in porous, granular media: a focus on permeability. *Physical Review E*, 94(2):022904.

[451] Vasilyev, L., Raoof, A., and Nordbotten, J. M. (2012). Effect of mean network coordination number on dispersivity characteristics. *Transport in Porous Media*, 95(2):447–463.

[452] Velasco, F. R. (1979). Thresholding using the isodata clustering algorithm. Technical report, MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER.

[453] Vervoort, R. and Cattle, S. (2003). Linking hydraulic conductivity and tortuosity parameters to pore space geometry and pore-size distribution. *Journal of Hydrology*, 272(1-4):36–49.

[454] Viberti, D., Peter, C., Borello, E. S., and Panini, F. (2020). Pore structure characterization through path-finding and lattice boltzmann simulation. *Advances in Water Resources*, page 103609.

[455] Victor, R. and Prodanovic, M. (2016). Low reynolds number velocity simulations in sandstones. http://www.digitalrocksportal.org/projects/29.

[456] Virkar, A. V., Chen, J., Tanner, C. W., and Kim, J.-W. (2000). The role of electrode microstructure on activation and concentration polarizations in solid oxide fuel cells. *Solid State Ionics*, 131(1-2):189–198.

[457] Vogel, H. (1997). Morphological determination of pore connectivity as a function of pore size using serial sections. *European Journal of Soil Science*, 48(3):365–377.

[458] Vogel, H.-J. (2002). Topological characterization of porous media. In *Morphology of Condensed Matter*, pages 75–92. Springer.

[459] Vogel, H.-J. and Roth, K. (2001). Quantitative morphology and network representation of soil pore structure. *Advances in Water Resources*, 24(3-4):233–242.

[460] Vogel, H.-J., Weller, U., and Schlüter, S. (2010). Quantification of soil structure based on minkowski functions. *Computers & Geosciences*, 36(10):1236–1245.

[461] Wang, J., Li, Z., Yan, S., Yu, X., Ma, Y., and Ma, L. (2019). Modifying the microstructure of algae-based active carbon and modelling supercapacitors using artificial neural networks. *RSC Advances*, 9(26):14797–14808.

[462] Wang, Y., Arns, C. H., Rahman, S. S., and Arns, J.-Y. (2018). Porous structure reconstruction using convolutional neural networks. *Mathematical Geosciences*, 50(7):781–799.

[463] Wang, Y., Yao, H., and Zhao, S. (2016). Auto-encoder based dimensionality reduction. *Neurocomputing*, 184:232–242.

[464] Wargo, E., Kotaka, T., Tabuchi, Y., and Kumbur, E. (2013). Comparison of focused ion beam versus nano-scale x-ray computed tomography for resolving 3-d microstructures of porous fuel cell materials. *Journal of Power Sources*, 241:608–618.

[465] Washburn, E. W. (1921). Note on a method of determining the distribution of pore sizes in a porous material. *Proceedings of the National academy of Sciences of the United States of America*, 7(4):115.

[466] Weerts, A., Bouten, W., and Verstraten, J. (1999). Simultaneous measurement of water retention and electrical conductivity in soils: Testing the mualem-friedman tortuosity model. *Water Resources Research*, 35(6):1781–1787.

[467] Weerts, A., Kandhai, D., Bouten, W., and Sloot, P. (2001). Tortuosity of an unsaturated sandy soil estimated using gas diffusion and bulk soil electrical conductivity. *Soil Science Society of America Journal*, 65(6):1577–1584.

[468] Wei, H., Zhao, S., Rong, Q., and Bao, H. (2018a). Predicting the effective thermal conductivities of composite materials and porous media by machine learning methods. *International Journal of Heat and Mass Transfer*, 127:908–916.

[469] Wei, L.-Y. and Levoy, M. (2000). Fast texture synthesis using tree-structured vector quantization. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pages 479–488. ACM Press/Addison-Wesley Publishing Co.

[470] Wei, W., Cai, J., Xiao, J., Meng, Q., Xiao, B., and Han, Q. (2018b). Kozeny-carman constant of porous media: Insights from fractal-capillary imbibition theory. *Fuel*, 234:1373–1379.

[471] Western, A. W., Blöschl, G., and Grayson, R. B. (2001). Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research*, 37(1):83–97.

[472] Wiedenmann, D., Keller, L., Holzer, L., Stojadinović, J., Münch, B., Suarez, L., Fumey, B., Hagendorfer, H., Brönnimann, R., Modregger, P., et al. (2013). Three-dimensional pore structure and ion conductivity of porous ceramic diaphragms. *AIChE Journal*, 59(5):1446–1457.

[473] Wiegmann, A. and Zemitis, A. (2006). Ej-heat: A fast explicit jump harmonic averaging solver for the effective heat conductivity of composite materials.

[474] Wildenschild, D. and Sheppard, A. P. (2013). X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. *Advances in Water Resources*, 51:217–246.

[475] Willingham, T. W., Werth, C. J., and Valocchi, A. J. (2008). Evaluation of the effects of porous media structure on mixing-controlled reactions using pore-scale modeling and micromodel experiments. *Environmental Science & Technology*, 42(9):3185–3193.

[476] Wilson, J. R., Kobsiriphat, W., Mendoza, R., Chen, H.-Y., Hiller, J. M., Miller, D. J., Thornton, K., Voorhees, P. W., Adler, S. B., and Barnett, S. A. (2006). Three-dimensional reconstruction of a solid-oxide fuel-cell anode. *Nature Materials*, 5(7):541.

[477] Winsauer, W. O., Shearin Jr, H., Masson, P., and Williams, M. (1952). Resistivity of brine-saturated sands in relation to pore geometry. *AAPG Bulletin*, 36(2):253–277.

[478] Wu, H., Fang, W.-Z., Kang, Q., Tao, W.-Q., and Qiao, R. (2019). Predicting effective diffusivity of porous media from images by deep learning. *Scientific reports*, 9(1):1–12.

[479] Wu, J., Yin, X., and Xiao, H. (2018a). Seeing permeability from images: fast prediction with convolutional neural networks. *Science bulletin*, 63(18):1215–1222.

[480] Wu, K., Van Dijke, M. I., Couples, G. D., Jiang, Z., Ma, J., Sorbie, K. S., Crawford, J., Young, I., and Zhang, X. (2006a). 3d stochastic modelling of heterogeneous porous media–applications to reservoir rocks. *Transport in Porous Media*, 65(3):443–467.

[481] Wu, Y., Lin, C., Ren, L., Yan, W., An, S., Chen, B., Wang, Y., Zhang, X., You, C., and Zhang, Y. (2018b). Reconstruction of 3d porous media using multiple-point statistics based on a 3d training image. *Journal of Natural Gas Science and Engineering*, 51:129–140.

[482] Wu, Y. S., Frijlink, H., et al. (2006b). The determination of relative path length as a measure for tortuosity in compacts using image analysis. *European Journal of Pharmaceutical Sciences: Official Journal of the European Federation for Pharmaceutical Sciences*, 28(5):433–440.

[483] Wyllie, M., Rose, W. D., et al. (1950). Some theoretical considerations related to the quantitative evaluation of the physical characteristics of reservoir rock from electrical log data. *Journal of Petroleum Technology*, 2(04):105–118.

[484] Wyllie, M. and Spangler, M. (1952). Application of electrical resistivity measurements to problem of fluid flow in porous media. *AAPG Bulletin*, 36(2):359–403.

[485] Xia, Y., Cai, J., Perfect, E., Wei, W., Zhang, Q., and Meng, Q. (2019). Fractal dimension, lacunarity and succolarity analyses on ct images of reservoir rocks for permeability prediction. *Journal of Hydrology*, 579:124198.

[486] Xiong, Q., Baychev, T. G., and Jivkov, A. P. (2016). Review of pore network modelling of porous media: experimental characterisations, network constructions and applications to reactive transport. *Journal of Contaminant Hydrology*, 192:101–117.

[487] Xu, H., Li, Y., Brinson, C., and Chen, W. (2014). A descriptor-based design methodology for developing heterogeneous microstructural materials system. *Journal of Mechanical Design*, 136(5).

[488] Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J., and Madabhushi, A. (2015). Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. *IEEE Transactions on Medical Imaging*, 35(1):119–130.

[489] Xu, P. and Yu, B. (2008). Developing a new form of permeability and kozeny–carman constant for homogeneous porous media by means of fractal geometry. *Advances in Water Resources*, 31(1):74–81.

[490] Yang, K., Li, M., Ling, N. N., May, E. F., Connolly, P. R., Esteban, L., Clennell, M. B., Mahmoud, M., El-Husseiny, A., Adebayo, A. R., et al. (2019). Quantitative tortuosity measurements of carbonate rocks using pulsed field gradient nmr. *Transport in Porous Media*, pages 1–19.

[491] Yang, W., Wang, K., and Zuo, W. (2012). Neighborhood component feature selection for high-dimensional data. *JCP*, 7(1):161–168.

[492] Yang, X.-Y., Chen, L.-H., Li, Y., Rooke, J. C., Sanchez, C., and Su, B.-L. (2017). Hierarchically porous materials: synthesis strategies and structure design. *Chemical Society Reviews*, 46(2):481–558.

[493] Yao, J., Frykman, P., Kalaydjian, F., Thovert, J., and Adler, P. (1993). High-order moments of the phase function for real and reconstructed model porous media: a comparison. *Journal of Colloid and Interface Science*, 156(2):478–490.

[494] Yegnanarayana, B. (2009). *Artificial Neural Networks*. PHI Learning Pvt. Ltd.

[495] Yeong, C. and Torquato, S. (1998). Reconstructing random media. *Physical Review E*, 57(1):495.

[496] Yu, B. and Cheng, P. (2002). A fractal permeability model for bi-dispersed porous media. *International Journal of Heat and Mass Transfer*, 45(14):2983–2993.

[497] Zacharias, N. A., Nevers, D. R., Skelton, C., Knackstedt, K., Stephenson, D. E., and Wheeler, D. R. (2013). Direct measurements of effective ionic transport in porous li-ion electrodes. *Journal of The Electrochemical Society*, 160(2):A306–A311.

[498] Zhang, X. and Knackstedt, M. A. (1995). Direct simulation of electrical and hydraulic tortuosity in porous solids. *Geophysical Research Letters*, 22(17):2333–2336.

[499] Zhao, Y., Liu, S., Elsworth, D., Jiang, Y., and Zhu, J. (2014). Pore structure characterization of coal by synchrotron small-angle x-ray scattering and transmission electron microscopy. *Energy & Fuels*, 28(6):3704–3711.

[500] Zhu, H., Ju, Y., Qi, Y., Huang, C., and Zhang, L. (2018). Impact of tectonism on pore type and pore structure evolution in organic-rich shale: Implications for gas storage and migration pathways in naturally deformed rocks. *Fuel*, 228:272–289.

[501] Zielinski, T. (2012). Inverse identification and microscopic estimation of parameters for models of sound absorption in porous ceramics. In *Proceedings of International Conference on Noise and Vibration Engineering (ISMA2012)/International Conference on Uncertainty in Structural Dynamics (USD2012)*, pages 95–108.