PHYSICAL REVIEW E **106**, 055304 (2022)

Pore-network extraction using discrete Morse theory: Preserving the topology of the pore space

Andrey S. Zubov,[1] Dmitry A. Murygin,[2] and Kirill M. Gerke[2,*]
[1]*Joint Institute for Nuclear Research, 141980 Dubna, Russia*
[2]*Schmidt Institute of Physics of the Earth of Russian Academy of Sciences, 107031 Moscow, Russia*

(Received 30 July 2022; accepted 17 October 2022; published 9 November 2022)

Pore-scale modeling based on the 3D structural information of porous materials has enormous potential in assessing physical properties beyond the capabilities of laboratory methods. Such capabilities are pricey in terms of computational expenses, and this limits the applicability of the direct simulations to a small volume and requires high-performance computational resources, especially for multiphase flow simulations. The only pore-scale technique capable of dealing with large representative volumes of porous samples is pore-network (PNM) based modeling. The problem of the PNM approach is that 3D pore geometry first needs to be simplified into a graph of pores and throats that conserve topological and geometrical properties of the original 3D image. While significant progress has been achieved in terms of geometry representation, no methodology provides full conservation of the topological features of the pore structure. In this paper we present a pore-network extraction algorithm for binary 3D images based on discrete Morse theory and persistent homology that by design targets topology preservation. In addition to methodological developments, we also clarify the relationship between topological characteristics of constructed Morse chain complex and pore-network elements. We show that the Euler numbers calculated for PNMs based on our methodology coincide with those obtained using the direct topological analysis. The characteristics of the extracted pore network are calculated for several 3D porous binary images and compared with the results of maximum inscribed balls-based and watershed-based approaches as well as a hybrid approach to support our methodology.

DOI: 10.1103/PhysRevE.106.055304

I. INTRODUCTION

The physical properties of porous and other composite materials are governed by their structure and the spatial distribution of constituting phases, and can be assessed based on structural information using statistical relationships or pore-scale simulations [1–3]. During last two decades such assessment approaches received significant attention due to the progress in the imaging of the material samples. X-ray computed tomography (XCT) is among the most convenient 3D image acquisition techniques [4] but may require scanning electron microscopy (SEM) combined with focused ion beam milling (FIB-SEM) [5,6] to improve imaging resolution vs sample size imaging trade-off [7]. Direct imaging methods can be combined with stochastic reconstructions [8–12] to produce 3D information from limited amount of data or periodize pore structure to compute tensorial properties [13,14], that can be further incorporated into a multiscale structure acquisition framework [7,15–17].

The data from even single-scale imaging are usually vast, and the size of resulting 3D digital image can easily exceed 2000^3 voxels, not to mention the results of the multiscale image fusion. Direct flow simulations, both single and multiphase cases, require enormous amount of computational power, up to days of execution on 10^2–10^3 CPUs in parallel. Such expenses make routine simulations with direct methods impractical and hamper the development of digital core technologies. The so-called pore-network approach (PNM) is much more computationally cheap due to simpler description of flow phenomena with the help of semiempirical models for single-phase flow and multiphase fluid-solid interactions. For example, the Poiseuille flow approximation combined with harmonic averaging of hydraulic conductances allows for efficient flow velocity assessment: the single-phase flow problem can be attacked by solving a linear system of equations [18], as opposed by a need to perform thousands of iterations in direct (explicit scheme) Stokes solvers based on a light-weight finite-difference method [19] or lattice Boltzmann method [20]. For two-phase flow simulations, PNMs utilize such models as piston displacement [18,21,22], pore-body filling [18,21–23] and snap-off [18,21,22,24], and wettability alteration and aging influence [18,25,26]; these semiempirical models are usually the results of experimental observations [27]. Miao *et al.* [28] proposed a novel framework that uses machine learning based on direct pore-scale simulations to parametrize the flow processes; this framework was applied in different PNM models [29–31]. The way multiphase flow processes are described depends on the type of the PNM that can be divided into two classes: quasistatic [18,21,22,26] and dynamic [32]. The former assume that capillary forces are dominant and are extremely fast in their computations, but lack time resolution. The latter type tracks menisci positions and evolution with time, but solves a much more evolved system of nonlinear equations that slows down all computations [33].

---

[*] kg@ifz.ru

To utilize semiempirical models or machine learning to describe different displacement mechanisms or properties, PNM first requires partition of the pore space into separate elements. In other words, the PNM approach needs to simplify the geometry of the real 3D pore network by representing it with the help of pore and throat elements while conserving the geometry and topology. This process that requires 3D pore geometry image as an input data is called pore-network extraction. At this moment, the dominant state-of-the-art extraction approaches include watershed-based [34,35], median axis-based [36], maximal inscribed balls [37,38] algorithms, and their hybrids [39,40]. All these methods differ in the way they define pores and throats and segment them out of the 3D pore space. The classical approach to represent pore geometries is the circle-triangle-square or CTS approach [18,41] was critiqued for its inability to represent nonconvex shapes [42]. The machine learning based framework showed that one can remove the simplification of the shapes completely with the help of direct simulations [28].

Combined with single- and multiphase flow simulation physics, PNM approach can be effectively applied to solve numerous other important problems in porous media. Some notable examples include, but are not limited to, the analysis of chemical reactions, transport of mass and energy, phase change, combustion, drying, and dispersion modeling [43–51].

From the geometrical point of view all extraction methods are relatively the same. But the major difference in the results they produce is the topological features of the extracted PNMs. It was recently argued that current extraction methods do not preserve topology [52] and it was shown that major popular approaches indeed fail in preserving the pore network's Euler numbers [40].

In addition to technological aspects of the pore-network extraction procedure, it also adds computational burden to the overall pore-scale modeling pipeline. In other words, while flow modeling with PNMs is orders of magnitude faster compared to direct methods, extraction is still costly, especially for large and complex 3D pore geometries. The latter problem can be effectively solved with the domain decomposition approach [53]. Pore networks can be stochastically generated [54] or reconstructed [55]; these methods still need some initial PNM statistics to rely on, which usually is the result of the extraction procedure. This all means that pore-network extraction procedure should be fast and computationally efficient.

To preserve the topological structure of the initial 3D image we base our pore-network extraction algorithm on discrete Morse theory. This theory was introduced by Forman [56] and then further developed and applied to the image analysis by other researchers. For example, Lewiner *et al.* [57] proposed the algorithms for building of an optimal discrete Morse function and Morse gradient vector field without initial scalar function, while King *et al.* [58] developed the method for building a scalar function defined at the vertices of simplicial complex. These developments facilitated methods based on weighted graph representation of Morse function [59], divide-and-conquer algorithm [60], and the heuristic algorithm [61] that were proposed to construct gradient vector field for arbitrary Morse function defined on the initial complex. Robins *et al.* [62] created the ProcessLowerStar algorithm, the first for which the correctness of constructed gradient vector field in 3D case was strictly proved. The effective methods for Morse chain complex extraction are considered in works [60,63,64]. One can also mention a series of papers [65–67] in which Morse-Smale complex is constructed for piecewise linear functions; the algorithms used are based on Banchoff's extension of continuous Morse theory [68]. Building such complexes is the starting point of the pore-network extraction based on discrete Morse theory.

The persistent homology is a very effective tool in topological data analysis closely related to the image analysis with discrete Morse approach. In this paper we apply the technique of simplification [69] developed within persistent homology theory. Another widely used related method is the analysis of persistent diagrams for revealing their correlations with different geometrical, topological and physical properties of porous media [70–74]. Edelsbrunner *et al.* [65] were the first to explore the connections between Morse theory and persistent homology, and the persistence homology algorithms found their way into practically all the latest works on image analysis in the framework of discrete Morse theory. Bauer *et al.* [75] proposed to take into account persistent pairs during gradient field construction. Persistence diagrams describing birth and death of persistent pairs obtained from Morse chain complex were analyzed to extract morphological and hydraulic characteristics of a variety of porous media samples [73,74]. Discrete Morse theory was also recently successfully applied for the image analysis in astronomy [76,77], topography [78], and medical biology [79]. The results of image analysis using discrete Morse theory and persistent homology are becoming a trend in the machine learning or deep learning algorithms where they are converted into training features [72,73,78,79].

Our study was inspired by the stimulating work of Delgado-Friedrichs *et al.* [64], who have shown how to make the partition of pore space into the basins corresponding to the different pores and also to extract 1- and 2-skeletons, and the recent results of Gerke *et al.* [40] that indicated topological problems within all major extraction algorithms. The objective of this paper is to establish a robust and efficient pore-network extraction procedure for binary 3D images based on Morse theory and persistent homology that achieve the goal of conserving the topological properties of the original pore space.

II. THEORETICAL BACKGROUND

As it is not possible to provide a detailed description of the topological theory here, in this section we provide only a brief overview and illustrate the introduced concepts with simple examples, where possible. For a very good introduction to discrete Morse theory and persistent homology see Scoville's book [80]; also Sousbie *et al.* [76] provide a very intuitive matching of discrete Morse theory against continuous Morse theory that may help in getting over the steepest part of the learning curve.

A. Cubical complexes

In topology, a complex can be treated as a decomposition of a space into cells of different dimensions. Here we

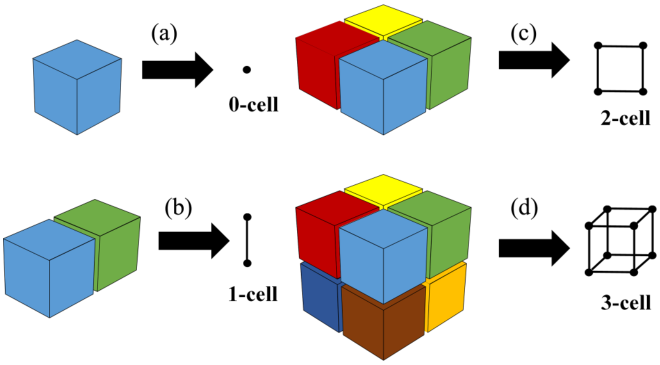

FIG. 1. Mapping voxels to the cells of a cubical complex: (a) 0-cell; (b) 1-cell; (c) 2-cell; (d) 3-cell.

consider cubical complexes, which describe 3D grid structures [81]. One can define 0-, 1-, 2-, and 3-cells as vertices, edges, squares, and cubes of the grid. We denote a cell belonging to cubical complex K as $\alpha^{(p)}$, where p is cell dimension ($p = 0, 1, 2, 3$). A cell $\alpha^{(p)}$ is a face of another cell $\beta^{(q)}$, if $p < q$ and the vertices of $\alpha^{(p)}$ are a subset of the vertices of $\beta^{(q)}$. We can also say that $\beta^{(q)}$ is a coface of $\alpha^{(p)}$. If $q = p + 1$, we call $\alpha^{(p)}$ a facet of $\beta^{(p+1)}$ and $\beta^{(p+1)}$ a cofacet of $\alpha^{(p)}$.

For a given cubical complex K, a filtration of K is defined as the sequence of complexes: $\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K$. If Ω is a set of all vertices (0-cells) of K, and function $f : \Omega \to \mathbb{R}$, one can extend this function to p-cells with $p > 0$ by the following way: each cell is assigned the maximum function value of the vertices it contains. Then the filtration of K with respect to f is defined by the sublevel complexes $K(t) = f^{-1}(-\infty, t]$. We start with an empty complex and at each step of the filtration one cell is added. The cell to be added is chosen according to the following rule: the value of function f at this cell is not less than function values of cells which have already been added at previous filtration steps. If two cells $\alpha^{(p)}$ and $\beta^{(q)}$ of different dimensions p and q have equal function values, then the cell with a smaller dimension is added first. The filtration concept plays an important role in the persistent homology of cubical complex (see the next subsection).

In 3D image analysis we can map voxel structure to a cubical complex as the following: each voxel is mapped to a vertex (0-cell) of a cubical complex. Then any 0-cell will represent 1 voxel, 1-cell – 2 voxels, 2-cell – 4 voxels, and 3-cell – 8 voxels, as is shown in Fig. 1.

Let us demonstrate the concepts presented in this section using a toy example of the cubical complex presented in Fig. 2. For simplicity, we consider only the 2D case; therefore 3-cells are absent in this complex.

We enumerate vertices of the considered complex by integers $i = 0,\ldots,11$. Then we denote 0-cells as $c(i)$, 1-cells as $c(i,\ j)$, 2-cells as $c(i,\ j,\ k,\ l)$. It is easy to show that our cubical complex contains 12 0-cells, 17 1-cells, and six 2-cells. Taking, for example, 1-cell $c(7,\ 11)$ we can find its facets, 0-cells $c(7)$ and $c(11)$, as well as its cofacets, 2-cells $c(7,\ 8,\ 10,\ 11)$ and $c(6,\ 7,\ 11,\ 9)$. To illustrate the filtration of cubical complex, presented in Fig. 2, let us define function f at its cells. Note that this function can be chosen arbitrarily, and for simplicity we consider the following example:

$$
\begin{aligned}
f(c(i)) &= i,\\
f(c(i,\ j)) &= \max(i,\ j),\\
f(c(i,\ j,\ k,\ l)) &= \max(i,\ j,\ k,\ l),
\end{aligned}
\tag{1}
$$

where $i, j, k, l = 0,\ldots,11$. Then, for example, $f(c(5)) = 5$, $f(c(6, 9)) = 9$ and $f(c(4, 1, 8, 7)) = 8$. Now we get the following sequence of sublevel complexes K_i: $K_0 = \{c(0)\}$, $K_1 = \{c(0)), c(1)\}$, $K_2 = \{c(0), c(1), c(2)\}$, $K_3 = \{c(0), c(1), c(2), c(0, 2)\}$ and so on, until we add all the cells of complex K.

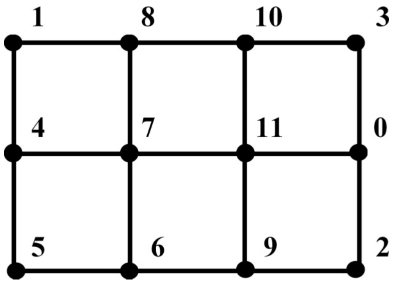

FIG. 2. Example of a cubical complex. Vertices are enumerated by integers from 0 to 11.

B. Homology of cubical complex

A set of facets of any p-dimensional cell $\alpha^{(p)}$ is called its boundary and denoted by $\partial_p \alpha$. Let a p-chain be a formal sum of p-cells with $\mathbb{Z}_2$ (modulo 2 addition) coefficients, and we can add two p-chains by adding their values. This forms a group (called chain group) $C_p(K, \mathbb{Z}_2)$. Then a given cubical complex K can be represented as chain complex, defined by the following sequence:

$$
C(K) : C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0,
\tag{2}
$$

where boundary operator ∂_p is $n_p \times n_{p-1}$ binary matrix, whose columns are the boundaries of p-dimensional cells. Here n_p is the number of such cells.

One can also define cycle group $Z_p(K) = \text{Ker}\partial_p$, boundary group $B_p(K) = \text{Im}\partial_{p+1}$ and homology group $H_p(K) = Z_p(K)/B_p(K)$. The rank of the pth homology group H_p is called the pth Betti number and can be calculated using dimensions of vector spaces $\text{Ker}\partial_p$ and $\text{Im}\partial_{p+1}$ as

$$
b_p = \dim(\text{Ker}\partial_p) - \dim(\text{Im}\partial_{p+1}).
\tag{3}
$$

Betti numbers provide direct topological description of a given 3D complex K: $b_0(K)$ is the number of connected components in the complex, $b_1(K)$ and $b_2(K)$ are the numbers of 1D and 2D holes, respectively. Another widely used topological characteristic, the Euler number χ, can be expressed through Betti numbers as

$$
\chi(K) = b_0(K) - b_1(K) + b_2(K).
\tag{4}
$$

Now let us explain how can we can calculate Betti number for the cubical complex, presented in Fig. 2. Here, we use the pipeline as described in [80]. In the 2D case we start from C_2 in chain complex [Eq. (2)]. We have six 2-cells, 17 1-cells, and 12 0-cells. Then ∂_2 is the matrix of size 6 × 17; to construct it

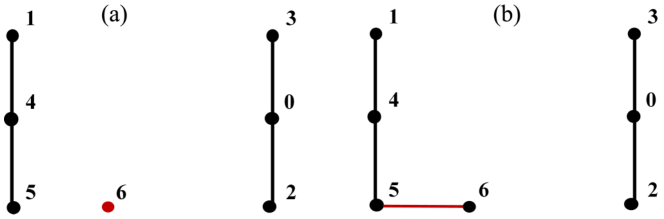

FIG. 3. Two steps of filtration with respect to the function f given by Eq. (1) for the cubical complex, presented in Fig. 2: (a) creator 0-cell $c(6)$ is added; (b) destroyer 1-cell $c(5, 6)$ is added.

we should enumerate 2-cells from $i = 1$ to 6, and 1-cells from $j = 1$ to 17. Then the matrix element $\partial_2(i, j) = 1$, if 1-cell j is a facet of 2-cell j, and $\partial_2(i, j) = 0$ in the opposite case. The similar procedure is applied to construct ∂_1 matrix of size 17×12. Since 0-cell has no facets, ∂_0 is a matrix of size 12×1, which contains only zeros. Now we can calculate image (Im) and kernel (Ker) dimensions for constructed matrices and obtain Betti numbers using Eq. (3).

The persistent homology studies evolution of topological features in time, during the filtration process. The main idea of this theory is grouping all cells added during the filtration into the closest in time pairs which do not change Betti numbers. Note that adding a new cell $\alpha^{(p)}$ to the complex will increase by one the value of b_p, so it will contribute to the first term in Eq. (3) or decrease by one the value of b_{p-1}, thus contributing to the second term. In the first case, cell $\alpha^{(p)}$ is called a creator, in the second a destroyer.

The examples of creator and destroyer cells are presented in Fig. 3. Adding 0-cell $c(6)$ leads to the change in the complex's topology. We have two connected components before adding this cell, and three components after its addition. So 0-cell $c(6)$ creates a new topological feature. At the next filtration step, we add 1-cell $c(5, 6)$, and the number of connected components becomes 2 again. Thus, 1-cell $c(5, 6)$ destroys the topological feature created in the previous filtration step.

Closest in time pairs "creator-destroyer" $\alpha^{(p)}(t_1)$ and $\beta^{(p+1)}(t_2)$ which do not change b_p are called persistent pairs, and the difference $f(\beta) - f(\alpha)$ is called persistence. It characterizes the lifetime of the topological feature described by the corresponding Betti number. If we set some persistent limit and remove from the complex persistent pairs with the persistence less than this limit, it does not affect the homology of pairs with the larger persistence. This technique can be used for noise filtration and size reduction of the complex considered. The list of persistent pairs for the complex can be obtained using reduced matrix algorithm [69,82].

C. Discrete Morse theory

Discrete Morse theory [56] can be treated as the expansion of Morse theory [83] to discrete data. While Morse theory considers smooth scalar functions defined over continuous manifolds, discrete Morse theory is applicable to scalar data on cells of different dimensions. Therefore, one should modify the smoothness condition to be used in cell complexes. Discrete Morse function defined as the following meets these requirements.

A discrete function f defined over a cubical complex K associates a real value $f(\alpha^{(p)})$ to each cell $\alpha^{(p)} \in K$. The discrete function f is a discrete Morse function if and only if, for every $\alpha^{(p)} \in K$, f takes a value less than or equal to $f(\alpha^{(p)})$ on at most one cofacet of $\alpha^{(p)}$ and takes a value greater than or equal to $f(\alpha^{(p)})$ on at most one facet of $\alpha^{(p)}$.

A cell $\alpha^{(p)}$ is critical, if all its cofacets take strictly greater values and all its facets are strictly lower. In other words, locally, a cell has a higher value than its facets and a lower value than its cofacets, and there can be only one exception at most in each case.

Forman showed [56] that for any discrete Morse function f, cell $\alpha^{(p)}$ can be noncritical if one of the following exclusive conditions is satisfied: exactly one of cofacets of this cell and all its facets have a lower or equal value or exactly one of its faces and all its cofacets have a higher or equal value. As a result, each noncritical cell $\alpha^{(p)}$ may be paired either with a noncritical cell that is a cofacet of $\alpha^{(p)}$, or with a noncritical cell that is a facet of $\alpha^{(p)}$. A collection of such pairs and a set of critical cells form discrete gradient vector field V, the discrete equivalent of gradient vector field in continuous Morse theory.

The discrete equivalent of of an integral line is a V-path, the sequence of cells $\alpha_0^{(p)}, \beta_0^{(p+1)}, \alpha_1^{(p)}, \beta_1^{(p+1)}, \alpha_2^{(p)}, \ldots, \beta_r^{(p+1)}, \alpha_{r+1}^{(p)}$ such that for each $i = 0, \ldots, r$, cells $\alpha_i^{(p)}$ and $\beta_i^{(p+1)}$ form the pair belonging to V, $\alpha_i^{(p)}$ is facet of $\beta_i^{(p+1)}$, $f(\alpha_i^{(p)}) \geqslant f(\beta_i^{(p+1)})$ and $\alpha_i^{(p)} \neq \alpha_{i+1}^{(p)}$. One can denote such a pair as $V(\alpha_i^{(p)}) = \beta_i^{(p+1)}$. All V-paths are acyclic if and only if f is a discrete Morse function [56].

As in the case of continuous Morse theory, in discrete Morse theory the dimension of critical cell equals to the index of criticality. Therefore, for cubical complexes critical 0-cells (vertices) are minima, critical 1-cells (edges) are 1-saddle points, critical 2-cells (squares) are 2-saddle points, and critical 3-cells (cubes) are maxima. V-paths starting at facets of a critical cell are discrete equivalents of descending manifolds. V-paths terminating at a critical cell are discrete equivalents of ascending manifolds.

To construct the discrete gradient vector field V, one can use the algorithm ProcessLowerStar. The detailed description of this algorithm, including corresponding pseudocode and the illustration of major steps with a toy example can be found in Robins *et al.* [62]. Here we present only the main idea of ProcessLowerStar. In this algorithm each 0-cell (vertex) α of cubical complex K generate a set of cells $L(\alpha)$ called Lower Star such that sum of this sets over all vertices is equal to K, and for any 0-cells α and β, $L(\alpha) \cap L(\beta) = \emptyset$. Then discrete gradient vector field is constructed for each set $L(\alpha)$ by searching for pairs $(\alpha_i^{(p)}, \beta_i^{(p+1)})$; cells of smaller dimensions are paired before the cells of larger dimensions. Cells remaining unpaired are treated as critical. Note that Lower Star partition of K is possible only if the values of Morse function f are unique at all vertices of K [62].

We construct gradient vector for the cubical complex, presented in Fig. 2, using function f given by Eq. (1), and the ProcessLowerStar algorithm; the results are presented in Fig. 4. Arrows show pairs of cells for which $V(\alpha) = \beta$, for example, $V(c(4)) = c(4, 1)$, $V(c(7, 11)) = c(6, 7, 11, 9)$. One can observe two critical 0-cells $c(0)$ and $c(1)$, two critical 1-cells $c(8, 12)$ and $c(6, 9)$, and one critical 2-cell

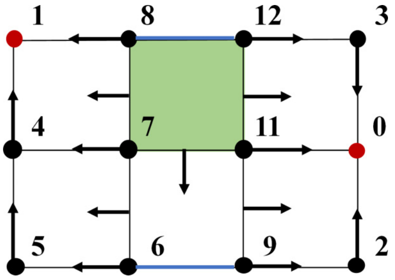

FIG. 4. Gradient vector field V for the cubical complex, presented in Fig. 2. Arrows show pairs of cells, for which $V(\alpha) = \beta$. Critical 0-, 1-, and 2-cells are denoted by red, blue, and green, respectively.

$c(7, 8, 12, 11)$. Note that there are no cycles in the constructed vector field, or in other words, no cell that has both arriving and departing arrows. In this case function f, as given by Eq. (1), is indeed the Morse function.

D. Homology of Morse chain complex

Similarly to chain complex C [Eq. (2)], one can define Morse chain complex M as the sequence

$$M(K) : M_3 \xrightarrow{\partial_3} M_2 \xrightarrow{\partial_2} M_1 \xrightarrow{\partial_1} M_0 \xrightarrow{\partial_0} 0. \tag{5}$$

Here $M_p \subseteq C_p(K, \mathbb{Z}_2)$ is the chain group generated by critical p-cells. Boundary operator ∂_p is $n_p \times n_{p-1}$ binary matrix, in which $(\partial_p)_{\alpha\beta} = 1$, if the number of V-paths between faces of critical p-cell $\alpha^{(p)}$ and the critical $(p-1)$-cell $\beta^{(p-1)}$ is odd. In the opposite case, $(\partial_p)_{\alpha\beta} = 0$. Here n_p is the number of critical p-cells. Morse chain complexes can be treated as the discrete analogues of Morse-Smale complexes considered in continuous Morse theory.

Forman [56] proves that the homology of C defined by Eq. (2) is always isomorphic to the homology of M. Therefore such topological characteristics as Betti and Euler numbers calculated for C are equal to ones calculated for M. This means that for Betti number calculations we can use Eq. (3) and take boundary operators ∂_p depending only on critical cells, as was described above. Also, the persistent homology of C coincides with the persistent homology of M [62]. Since M is usually much smaller than C, using discrete Morse theory leads to more memory- and time-efficient algorithms for persistent homology.

Both chain complexes and Morse chain complexes belong to a more general class, CW-complex, in which the cells are homeomorphic to euclidean balls [84].

One can build Morse chain complex using gradient vector field V simply by following V-paths from the faces of each critical cell $\alpha^{(p)} \in M$, and determining which critical $(p-1)$-cells the V-paths terminate at. Modified breadth-first search algorithm [64] can be used to solve this problem.

Let us consider Morse chain complex for the cubical complex, presented in Fig. 2. We can find V-paths for critical cells from the analysis of gradient vector field, presented in Fig. 4. We have one critical 2-cell, two critical 1-cells, and two critical 0-cells. There is a single V-path between critical 2-cell $c(7, 8, 12, 11)$ and critical 1-cell $c(8, 12)$: {c(8, 12)}. There is a single V-path between critical 2-cell $c(7, 8, 12, 11)$ and critical 1-cell $c(6, 9)$: {c(7, 11), $c(6, 7, 11, 9)$, $c(6, 9)$}. Thus,

$$\partial_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{6}$$

There is a single V-path between critical 1-cell $c(8, 12)$ and critical 0-cell $c(0)$: $\{c(12), c(12, 3), c(3), c(3, 0), c(0)\}$. Also, between critical 1-cell $c(8, 12)$ and critical 0-cell $c(1)$: $\{c(8), c(8, 1), c(1)\}$. Also, between critical 1-cell $c(6, 9)$ and critical 0-cell $c(0)$: $\{c(9), c(9, 2), c(2), c(2, 0), c(0)\}$. There is a single V-path between critical 1-cell $c(6, 9)$ and critical 0-cell $c(1)$: $\{c(6), c(6, 5), c(5), c(5, 4), c(4), c(4, 1), c(1)\}$. In this case,

$$\partial_1 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \tag{7}$$

Since we have two critical 0-cells, ∂_0 is a zero 2×1 matrix:

$$\partial_0 = [0, 0]. \tag{8}$$

To calculate image and kernel dimensions of matrix (∂_p), we can use well-known linear algebra formulas

$$\dim(\mathrm{Im}\partial_p) = \mathrm{Rank}\partial_p \tag{9}$$

and

$$\dim(\mathrm{Ker}\partial_p) = n_p - \mathrm{Rank}\partial_p. \tag{10}$$

Since $\mathrm{Rank}\partial_2 = 1$, $\mathrm{Rank}\partial_1 = 1$ and $\mathrm{Rank}\partial_0 = 0$, using Eq. (3) we can compute Betti numbers for the considered complex: $b_0 = 1$, $b_1 = 0$. Afterwards, using Eq. (4) we evaluate the Euler number value: $\chi = 1$.

Some of critical cells found in discrete Morse approach can be canceled according to a given persistent limit. This process is usually called simplification. Forman [56] shows that the cancellation is possible if there is exactly one V-path from the boundary of $\alpha^{(p)}$ to $\beta^{(p-1)}$. In this case, after the cancellation one should reverse V-path between $\alpha^{(p)}$ and $\beta^{(p-1)}$ and the modified vector field satisfies the condition of acyclic V-paths. After V-paths reversing loops arise if there is more than one V-path. The condition of exactly one V-path existence provides us with a possibility to abandon the reduced matrix algorithm for the derivation of persistent pairs from Morse chain complex, and use more time-efficient algorithm based on a cancellable close pair concept [64].

A cancellable close pair $(\alpha^{(p)}, \beta^{(p+1)})$ is a pair of critical cells such that there is a single V-path from the boundary of $\beta^{(p+1)}$ to $\alpha^{(p)}$, with the condition that any other critical cell $\gamma^{(p)}$ that is reachable via a V-path from the boundary of $\beta^{(p+1)}$ has $f(\gamma^{(p)}) \leqslant f(\alpha^{(p)})$, and for any critical cell $\delta^{(p+1)}$ for which there is a V-path terminating at $\alpha^{(p)}$ has $f(\delta^{(p+1)}) \geqslant f(\beta^{(p+1)})$. As it is proven in [64], a cancellable close pair is the persistent pair and its cancellation does not affect the homology of pairs with the persistence larger than a given persistence limit.

When the list of cancellable close pairs for the initial gradient vector field is found for the fixed value of persistent limit, they are removed and the gradient vector field is modified by the reversing of corresponding V-paths. New cancellable close pairs can be found again with the same value of persistent limit parameter. Thus, several simplification steps are required, until the list of cancellable close pairs becomes empty.

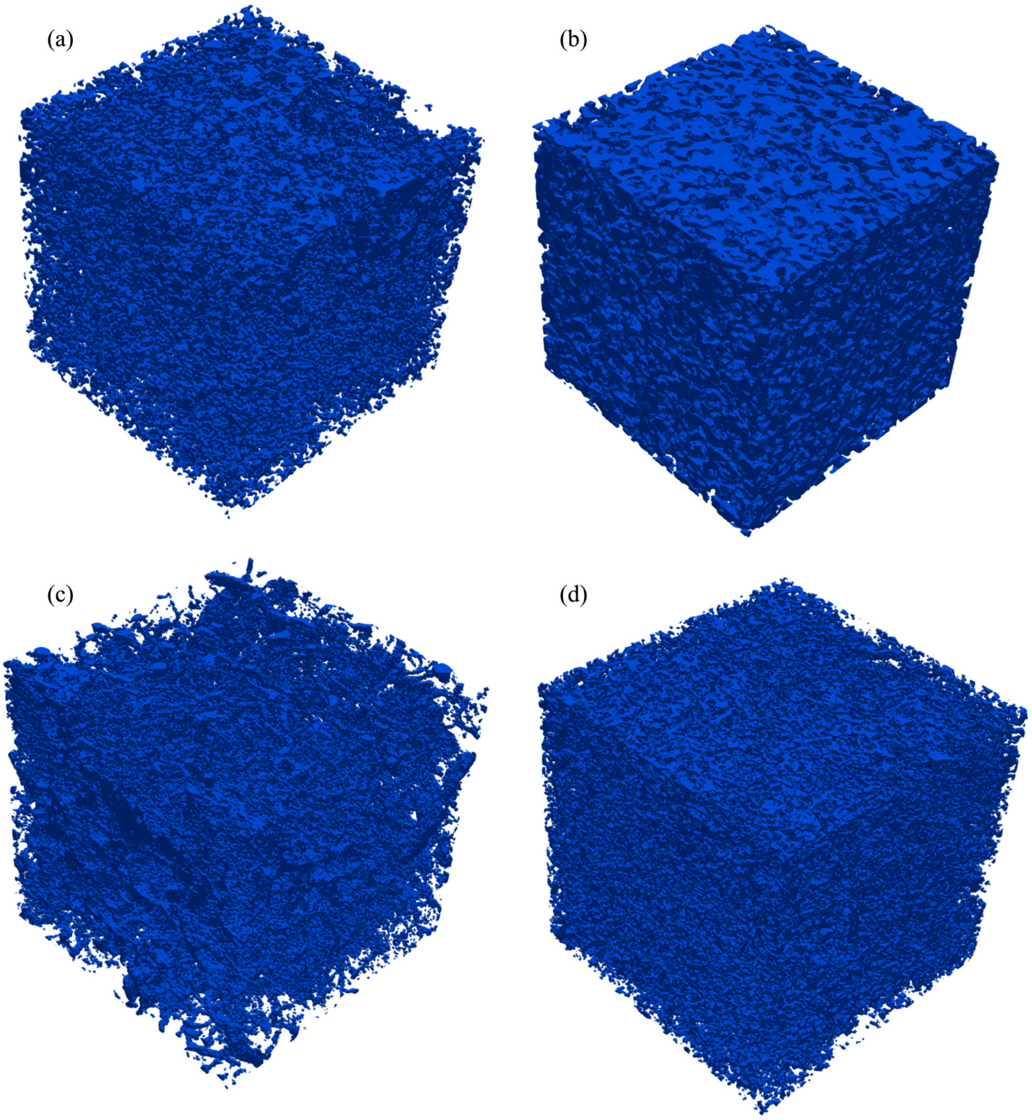

FIG. 5. Binary structure of samples presented in Table I: (a) carbonate; (b) ceramic; (c) soil; (d) sandstone. Pore phase is shown by blue.

E. Euler number in discrete Morse theory

Let us consider a list of persistent pairs "creator-destroyer" and a set of unpaired creators (if any) derived from Morse chain complex, N_p, N_p^C, N_p^D are numbers of critical cells, creators and destroyers, respectively, for $p = 0, 1, 2, 3$. Then the expression for pth Betti number [Eq. (3)] can be rewritten as

$$b_p = N_p^C - N_{p+1}^D. \tag{11}$$

The number of critical cells can be expressed through number of creators and destroyers as the following:

$$N_0 = N_0^C;\ N_1 = N_1^C + N_1^D;\ N_2 = N_2^C + N_2^D,\ N_3 = N_3^D. \tag{12}$$

Here we take into account that all 0-cells can be only creators, and all 3-cells can be only destroyers. Now substituting the expression [Eq. (11)] into the formula for Euler number [Eq. (4)] and taking into account [Eq. (12)] one can obtain

$$\chi = N_0 - N_1 + N_2 - N_3. \tag{13}$$

This formula depends only on numbers of critical cells, which can be obtained just after gradient vector field construction. So discrete Morse theory has a simple in-built algorithm for the calculation of such important topological characteristic as Euler number. We should also stress here that the value of χ is not changed after the simplification. This important property is provided by the usage of the persistent homology theory.

The formula for Euler number [Eq. (13)] can be also be derived in a different manner, starting from the expression for Euler number for the set of voxels [85]:

$$\chi = \#\text{vertices} - \#\text{edges} + \#\text{faces} - \#\text{volumes}, \tag{14}$$

where # means "number of." Taking into account mapping voxels to the cubical complex, one can obtain the expression

TABLE I. Porosity 3D image characteristics used to test the pore-network extraction algorithm based on the discrete Morse theory.

Sample	Porosity	Connected porosity	Imaging resolution, μm	3D image volume, voxels
Carbonate	0.073754	0.05045	5.38592	700^3
Ceramic	0.37666	0.37655	2.24	500^3
Soil	0.098103	0.084112	20.96	700^3
Sandstone	0.16952	0.16687	5.16734	700^3

depending on the numbers of p-cells N_{pc} ($p = 0, 1, 2, 3$):

$$\chi = N_{0c} - N_{1c} + N_{2c} - N_{3c}. \tag{15}$$

Since at the transition from chain complex C to topologically equivalent Morse chain complex M, p-cells are replaced by critical p-cells, one can finally get Eq. (13).

It is easy to show that application of Eq. (13) and Eq. (15) to the cubical complex, presented in Fig. 4, leads to the same result for the Euler number, $\chi = 1$, as we obtain in the previous subsection using Eq. (4).

III. APPLICATION OF DISCRETE MORSE THEORY AND PERSISTENT HOMOLOGY TO PORE-NETWORK EXTRACTION

In this section, we demonstrate how to apply discrete Morse theory and persistent homology to the analysis of 3D binary images. First, we describe all methodological details of the extraction procedure, which can be separated into several major steps.

The first step is to construct the discrete Morse function and gradient vector field from the binary image. The immediate extraction of pore-network elements from the constructed field would lead to unrealistically large number of them, therefore the preliminary simplification is required. Such simplification is the second step of the pore-network extraction algorithm. The last step is the derivation of pore-network elements from final vector field modified after simplification.

Our early attempts were based on C++ code Diamorse [86] that can perform the first and the second steps of the pipeline described above, with a final step performed by a code algorithmically similar to that used by Gerke *et al.* [40]. A fully parallelized in-house C++ code was developed afterwards to implement the whole PNM extraction pipeline, but some efficient solutions such as hashes for facets were adopted from Diamorse in a more of less intact form.

Four different 3D binary images were used to test the pore-network extraction algorithm based on the discrete Morse theory. The major characteristics of these images are presented in Table I, and their binary structure is shown in Fig. 5. The variety of samples is wide enough to cover different geological genesis, e.g., sandstone and carbonate rocks studied in petrophysical applications [40]; soil as a regular object in soil sciences and hydrology [87]; and artificial porous media (ceramic sample) relevant for industrial applications [88]. Moreover, the porosity of the chosen sample collection varies between 5% and 37%, thus, providing a possibility to test our pore-network extraction model on both tight and highly porous examples.

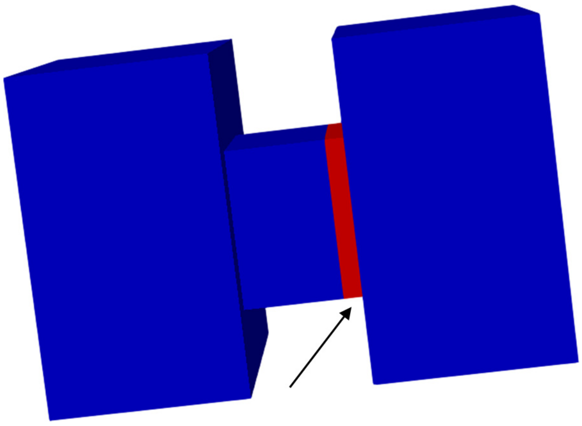

FIG. 6. Two pore boxes surrounded by solid space. Calculations are performed with cell ID used as the additional comparison key for equal EDT values. Pore and throat voxels are presented by blue and red, respectively.

A. Construction of a gradient vector field from a binary image

The input of the pipeline consists of a 3D binary image. It is represented by a 3D array $\Omega = L \times M \times N$. Solid and pore phases are coded by unity and zero. Voxel structure of 3D image is mapped to a cubical complex K as has been described in Sec. II A (see also Fig. 1). Each cell of the cubical complex is denoted by unique ID, linear in l, m, n [86].

Mapping voxel structure of 3D image to a cubical complex results in 6-connectivity between voxels. If one wishes to utilize 26-connectivity condition, mapping to a simplicial complex is required. We prefer 6-connectivity due to ambiguity of pore connectivity in the 26-connectivity model; moreover, high-resolution images (with discretizations higher than a couple of voxels for a pore element) should provide enough information for 6-connectivity to work properly. Note that the discrete Morse theory can be successfully applied to both cubical and simplicial complexes.

To construct gradient vector field one needs to define Morse function $f : \Omega \rightarrow \mathbb{R}$. We take signed Euclidean distance transform (EDT) function as f, with negative values in pore phase. It guarantees us that minima of f coincide with the pore centers. The algorithm from the work of Meijster *et al.* [89] was implemented to calculate EDT values.

The uniqueness of Morse function values at the vertices of the cubical complex can be ensured by several ways. Robins *et al.* [62] proposed to use small perturbation, linear in l, m, n, to break the equality of Morse function values. In the original Diamorse code [86] cell ID is used as the additional comparison key: in the case when the values of f at two 0-cells α and β are equal, it is defined $f(\alpha) > f(\beta)$ if cell ID$(\alpha) >$ cell ID(β). Both approaches lead to systematic shift in the properties of extracted pore-network elements. See, for example, Fig. 6, where EDT values are equal at the central axis voxels of the neck between the two box-shaped pores. Using cell ID as the additional comparison key leads

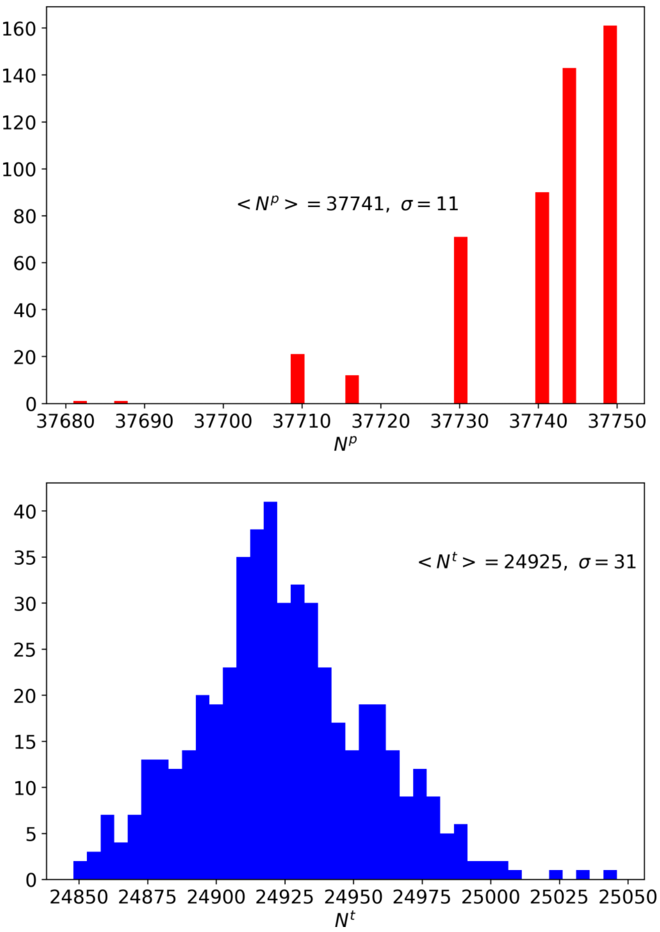

FIG. 7. Distributions of pores (red) N^p and throats (blue) N^t numbers caused by small random perturbation of Morse function. The results as shown are the results of 500 pore-network extractions from the carbonate sample. Mean values and standard deviations are indicated.

to the shift of throat to the border of the right box in this example.

We propose to add small random perturbation to all Morse function values. For any 0-cell $\alpha \in K$ we use

$$f(\alpha) = f(\alpha) + \epsilon r, \tag{16}$$

where $r \in \mathbb{R}$ is a random number in (0,1) interval, and ϵ should be taken small enough to not exceed the minimal nonzero difference between Morse function values at any two vertices of K. The proposed approach seems to be more realistic and ensures the absence of any systematic shifts in the extracted pore-network properties. Moreover, it accounts for the uncertainty of local EDT due to XCT resolution limitation for real porous media samples. In the present study we fix the parameter as $\epsilon = 0.001$ voxel size units for all extractions.

The small disadvantage of the proposed approach is that the extracted pore-network characteristics demonstrate the statistical behavior (see Fig. 7 as an example) even at very small values of ϵ. In this sense, discrete Morse theory is not stable to small perturbations. However, whether the other pore-network extraction models which use EDT values are stable is still open. We are not aware of any research results in this area. As one can see in Fig. 7, the standard deviations both for pores and throats number are very small, so we adopt the expression in Eq. (16) for the Morse function in all our calculations.

TABLE II. Execution times (in seconds) for several binary 3D images with two different approaches for gradient field construction (I - with total complex, II - with pore complex), case a - using one thread for computation, case b - 32 threads.

Sample	Ia	Ib	IIa	IIb
Carbonate	1825 s	278 s	356 s	37 s
Ceramic	486 s	98 s	250 s	40 s
Soil	1403 s	192 s	346 s	46 s
Sandstone	1829 s	312 s	405 s	56 s

There is another way to decrease the deviations of pore-network characteristics distribution if the additional comparison key is used when EDT values in two vertices are equal. The possible solution is to compare the sums of EDT values at the neighboring vertices. Small random perturbation can be added only if these sums are still equal. This approach has more physical sense than the simple comparison of cell IDs, but, as we found, decrease the speed of run.

Since during the extraction of pore-network properties we are interested only in the pore phase, it is necessary to consider only those cells which have negative Morse function values. This can be implemented using two different strategies. The first one operates during the gradient vector field construction followed by the simplification for the total cubical complex; we choose only those cells that belong to the pore phase at the stage of pore-network properties extraction. In the second approach, this choice is made at the first step; the gradient vector field is constructed only for pore complex, in other words, for cells which Morse function value is negative. The exclusion of the solid phase can be performed during the Lower Star partition of cubical complex. We include into the Lower Star sets only those cells which belong to the pore phase. Thus obtained gradient vector field are the starting point for the next steps of the pore-network extraction algorithm.

We checked that these two approaches lead to the same results. Nevertheless, the execution time in the second case is several times less (see Table II) due to the large decrease of cells included in the gradient vector field. So we can conclude that using the pore complex at the stage of the gradient vector field construction is preferable in practical calculations than using the total complex. The results presented in Table III are obtained using computer station with a double Intel Xeon Gold 6248R CPU at 3.00 GHz. The simple OpenMP parallelization approach showed significant speed-ups, but its efficiency is far from linear for 32 threads. We estimate the upper bound for memory requirements as $20N$ bytes, where N is the total number of voxels within the sample. The exact amount of the used memory depends on the number of extracted PNM components.

During the simplification procedure numerous critical cells obtained from the initial gradient vector field are canceled

TABLE III. Pores, throats, junction numbers, values $\Delta\chi_{\text{intr}}$, and Euler numbers calculated for several binary images.

Sample	N^p	N^t	N^j	$\Delta\chi_{\text{intr}}$	χ
Carbonate	37 741	24 925	935	−545	13 206
Ceramic	9238	33 299	7221	2145	−14 695
Soil	49 464	27 678	2660	−3505	20 941
Sandstone	84 401	67 253	2038	−4838	14 348

according to a given persistent limit. As one can see in Fig. 8(a), this parameter has a significant influence on the results of pore-network extraction. So, its proper choice is important. We calculate the dependence of neighboring and nonneighboring cancellable close pairs numbers on the persistent limit. We call the cancellable close pair neighboring, if at last two vertices belonging to critical $(p+1)$- and p-cells, which form this pair, coincide or are located at 6-connectivity neighborhood of each other. In the opposite case, we call the cancellable close pair nonneighboring. Figures 8(b) and 8(c) show that the canceling of neighboring pairs is the dominant process at simplification. This is an intuitively logical result, as the neighboring critical cells usually map neighboring topological properties (for example, pore and throat centers) and such neighborhood is unreal in most cases and should be removed by simplification. While the number of nonneighboring cancellable close pairs smoothly grows with increase of persistent limit, the number of neighboring close pairs comes to the plateau at the value of persistent limit equal to 1. Such behavior of this value is observed for all the samples considered in this work (see Fig. 9). So we fix the persistent limit parameter at the value equal to 1 in all calculations. The same value of this parameter was chosen by Deglado-Friedrichs *et al.* [64], yet no explanation was provided. Taking the larger values of persistent limit can lead to the loss of actual pore-network characteristics described by the critical cells (note that Euler characteristics are unchanged by simplification).

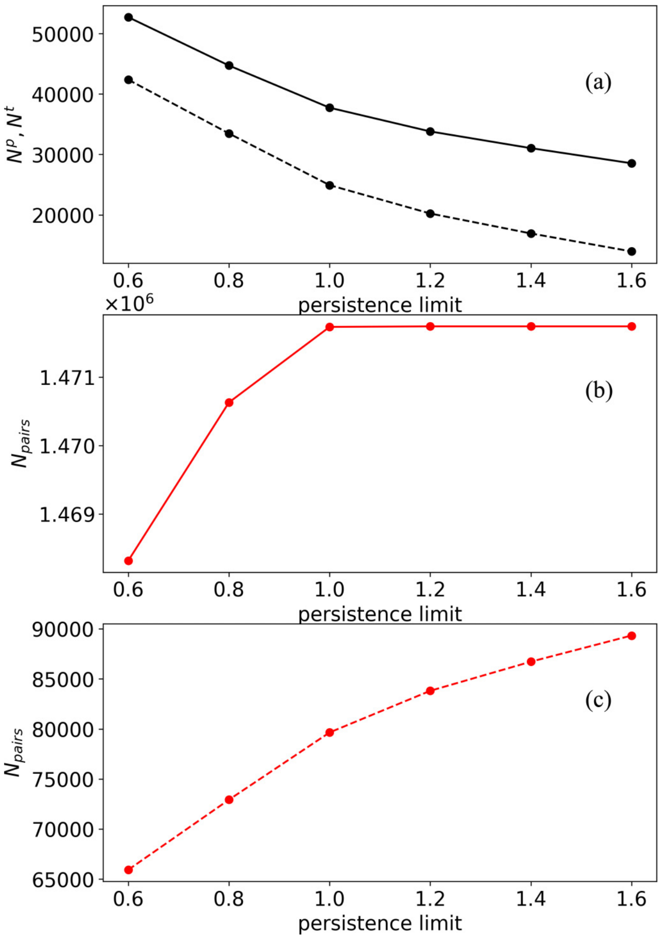

FIG. 8. (a) The dependence of pores N^p (black solid line) and throats N^t(dashed solid line) numbers on the persistent limit parameter. (b, c) The dependence of neighboring (red solid line) and nonneighboring (red dashed line) cancellable close pairs numbers N_{pairs} on the persistent limit parameter. All calculations shown are carried out for the carbonate sample.

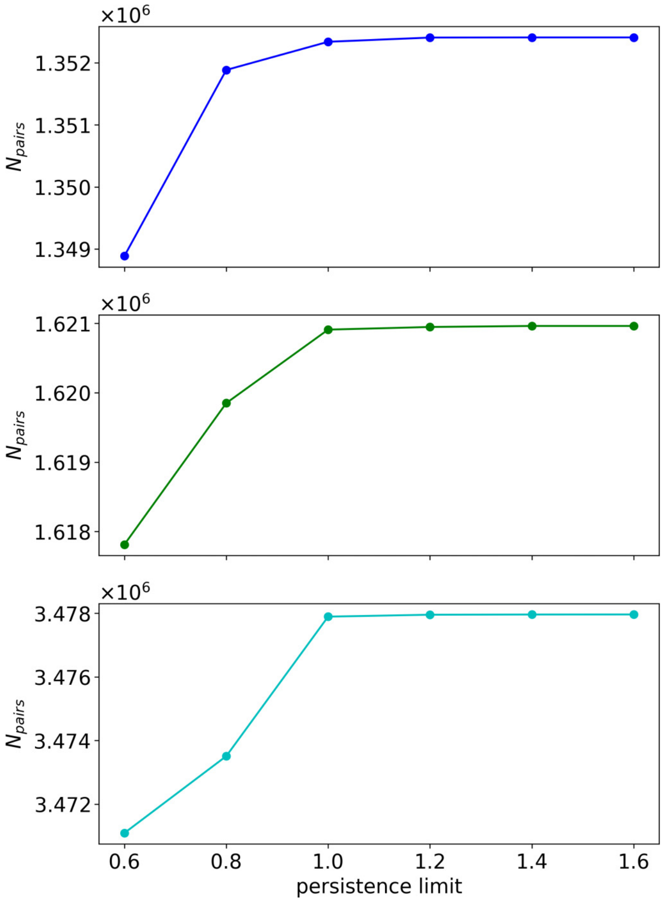

FIG. 9. The dependence of neighboring cancellable close pairs numbers on the persistent limit parameter for ceramic (blue), soil (green), and sandstone (cyan) samples.

B. Extraction of pore bodies and pore throats

In this subsection we describe in details the algorithm for the extraction of pore-network characteristics from the gradient vector field constructed for the initial pore complex using perturbed EDT as the discrete Morse function and the ProcessLowerStar algorithm, after the constructed vector field is transformed during several simplification steps according to the chosen persistent limit. Such final modified gradient vector field is the starting point for the extraction of separate pore-network elements.

As it is shown in the work [64], any 0-cell of cubical complex belongs to the stable set of exactly one critical 0-cell which corresponds to the local minimum. The stable set of critical cell α consists of all 0-cells from V-paths taken in reversed directions and starting at the cofacets of a α. This provides us with the natural partitioning to separate pores of all pore voxels mapped by 0-cells. First, we mark all critical 0-cells of the gradient vector field by unique IDs, then construct the stable set of each critical 0-cell α and mark all 0-cells from this set by the same ID as ID of α. A modified breadth-first search algorithm [64] is used for the construction of stable sets.

By the end of this procedure, each voxel from the pore phase of the input binary image is related to some pore with the unique ID, the example of such partitioning by different pore IDs is presented in Fig. 10.

The coordinates of pore center for a given pore ID are easily found as the coordinates of the center of the voxel mapped by the corresponding critical 0-cell, and the radius of pore r_p is simply the absolute EDT value calculated at this point. Pore volume V_p is calculated as the product of the number of 0-cells with a given ID and the elementary voxel volume.

To parametrize each pore-network element we find the pore boundary for the calculations of pore surface area S_p and pseudodiameter D_p. The pore boundary can be easily found as the set of all pore voxels for which at last one of 6-connectivity neighboring voxels does not belong to this pore. For 0-cell α, its 6-connectivity neighborhood is the set of all the facets of all 1-cell cofacets of α except for α itself. If at least one of found facets does not belong to the given pore, α is included to the pore boundary.

The value of pore surface area S_p is calculated as the product of the elementary voxel area and number of such 1-cell cofacets of the pore boundary 0-cells, as one of their facets does not belong to the given pore. The pore pseudodiameter D_p is defined as the maximal length of edges passing through the pore center, and starting and terminating at the surface of the pore boundary voxels. Using the values V_p, S_p and D_p, one can calculate the pore shape factor $G_p = V_pD_p/S_p^2$. This is a classical pore morphological approach, e.g., [38,40], but with known caveats (for example, this way surface area depends on the resolution the 3D input image). Here we focus on topological features of the pore network and leave the perfection of the geometrical representation for the future studies.

Using the obtained pore partitioning of the cubical pore complex, it is easy to extract throats between the pores with different IDs. We define an element of a throat's set as two voxels mapped by 1-cell, with its vertices being marked by different pore IDs. Then passing through all 1-cells of given complex, we can find the sets of 1-cells, which are marked by the same pairs of pore IDs (e.g., ID1, ID2). All the voxels corresponding to the vertices of 1-cells marked by (ID1, ID2) form one or several throats which connect pores marked by ID1 and ID2.

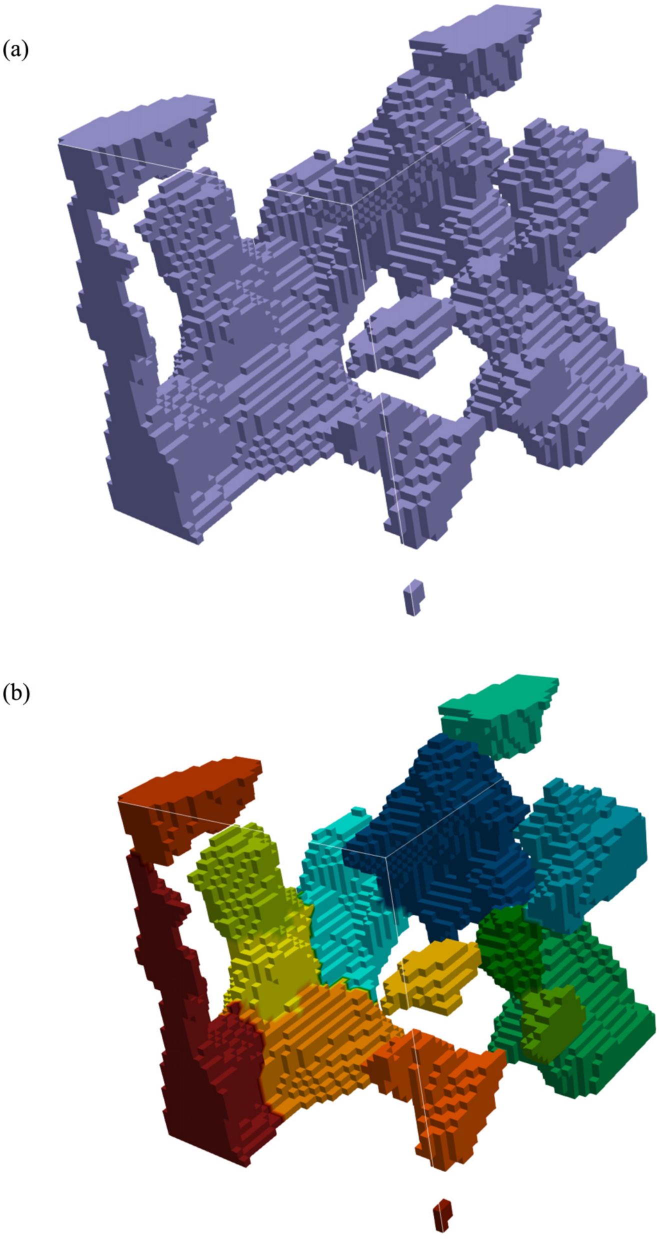

FIG. 10. (a) Pore phase of binary image, 100^3 voxels subcube of soil sample. (b) Partitioning of pore phase by separate pores using discrete Morse theory. Different colors represent different pore IDs.

The number of unique throats between two pores is equal to the number of unconnected components in the set T of 1-cells marked by unique (ID1, ID2). We use the following algorithm to mark the throats:

(a) 2-cell cofacets of 1-cell $\beta \in T$ are 1x1 squares with β as one of their edges. Now we can define the neighbors of β as 1-cell facets of these 2-cells, except for β itself.

(b) Cell β and all the neighbors of β which also belong to T are added to the same component.

(c) Taking any 1-cell from T as starting point and making breadth-first search we find out the subset t of T which corresponds to the separate throat.

(d) Set difference $T - t$ is the input for the next unconnected component extraction step.

The coordinates of throat center are found as the coordinates of the center of the voxel mapped by the vertex of 1-cell from t with the minimal (recall that pore complex has negative EDT values) EDT value. If critical 1-cell $\beta_c \in t$, one of its vertices contributes to the throat center.

Since in discrete Morse theory throats are considered as pairs of voxels mapped by 1-cells, the surface between the voxels in the pair contributes to the throat area. Then the throat area S_t is calculated as the product of elementary voxel area and number of 1-cells in t. We define border 1-cells as such 1-cells neighboring to 1-cells from t that do not belong to t itself. The perimeter of the throat P_t is then calculated as the product of the number of border cells and elementary voxel length. Using the values S_t and P_t, one can calculate the classical throat shape factor as $G_t = S_t/P_t^2$ [41].

Correct throat identification is a challenge in most algorithms used to extract the pore network from a 3D binary image. Skeleton-based algorithms [90–93] define a throat as the smallest cross-section voxel area belonging to the pore phase with the center at medial axis path (skeleton) between two pore centers. First, the segment of interest in the skeleton is chosen, then pore areas crossing all voxels of the segment are minimized. Only the planes perpendicular to the segment direction can be considered [92,93] to speed up the calculations. The found throats are always planar and the correction for nonplanarity can be made at the last step of the algorithm [93]. One more problem which should be solved in the skeleton-based approach is possible throat intersection, which is forbidden by hydraulic reasons. So postprocessing of intersecting events is necessary [92]. Classical watershed-based segmentation to obtain the throats does not provide correct result, as has been extensively discussed in skeleton-based extraction literature mentioned earlier.

Here we can indicate the advantages of throat extraction using the discrete Morse theory over skeleton-based algorithms. First, preliminary skeleton extraction is not required in our algorithm. In addition, we do not encounter the computational difficulties as related to searching for a huge number of 2D planes and our algorithm does not require any approximations to avoid such difficulties. As one can see in Fig. 11, discrete Morse theory provides the correct description of nonplanar throats. The throat crossing problem does not exist in our approach, as we consider only unconnected components with unique pair of pore IDs as separate throats. No extraction algorithm except for Morse theory based, and the one described by Jun [93] can produce a throat as shown in Fig. 11.

In the extracted pore network more than two pores can have common points or borders. Such cases called junctions are also taken into account in our pore-network extraction algorithm. In 6-connectivity, provided by cubical complex structure, not more than four voxels can have common points, so only 3- and 4-junctions are possible in our approach. We extract junction after the pore and throat segmentation steps.

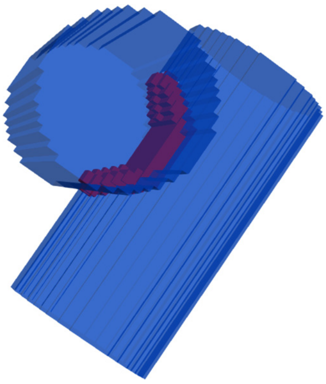

FIG. 11. The throat (red) found in pore space (blue) formed by two crossing cylinders using discrete Morse theory.

First, all the loops including three and four pores are found. They are further treated as junction candidates. Consider, for example, a loop of three pores marked by ID1, ID2, and ID3 and take the throats set T marked by the pair (ID1, ID2). We find the vertices of the cofacets of all 1-cells which belong to this set. If the vertex belongs to the pore marked by ID3, we get 3-junction point. The similar approach is used to detect 4-junctions. Note that for all the samples considered in this paper only 3-junctions were observed.

Note that our toy example of cubical complex, as shown in Fig. 2, does not represent any real porous media image. The gradient vector field for this cubical complex (Fig. 4) is constructed with a Morse function that differs from perturbated EDT, which we used for 3D binary image analysis. The choice of such a toy example was dictated by our inability to draw clear 3D real examples, but proved helpful in explaining the major steps of the discrete Morse theory and persistent homology applications. Nevertheless, it is formally possible to extract pores and throats from gradient vector field, presented in Fig. 4. To get the pore separation of all 0-cells, one needs to find all V-paths taken in the reversed direction which start at the cofacets of critical 0-cells (pore centers), and for each critical 0-cell select 0-cells which are included in found V-paths. In our example we have two critical 0-cells: $c(0)$ and $c(1)$. It is easy to find the required reversed V-paths from the analysis of Fig. 4. Next, we search for the direct V-pathes, in doing so we should move according the directions of arrows in Fig. 4, and in the case the reversed V-pathes, we should move in the opposite direction. Then for critical 0-cell $c(0)$ we obtain the following V-paths: $\{c(0,3), c(3), c(12,3), c(12)\}$, $\{c(0,11), c(11)\}$, $\{c(0,2), c(2), c(9,2), c(9)\}$. Taking 0-cells from the found V-paths, we get the following set of voxels, forming the first pore: $P_1 = \{c(0), c(3), c(12), c(11), c(2), c(9)\}$. For critical 0-cell $c(1)$ we get the following V-paths: $\{c(1,8), c(8)\}$, $\{c(1,4), c(4), c(4,7), c(7)\}$, $\{c(1,4), c(4), c(4,5), c(5), c(5,6), c(6), c(6,7), c(7)\}$. Taking 0-cells from the found V-paths, we get the following set of voxels, forming the second pore: $P_2 = \{c(1), c(8), c(4), c(7), c(5), c(6)\}$. The found 0-cells belonging to pores P_1 and P_2 are denoted in Fig. 12 by yellow and purple, respectively. Searching for 1-cells,

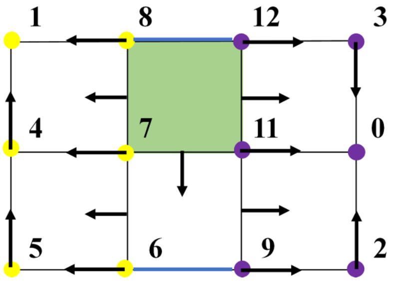

FIG. 12. Gradient vector field V and pore separation for the cubical complex, presented in Fig. 2. Arrows show pairs of cells, for which $V(\alpha) = \beta$. Critical 1- and 2-cells are denoted by blue and green, respectively. 0-cells belonging to different pores are denoted by yellow and purple.

which vertices belong to P_1 and P_2, give us the throat set $t(1, 2) = \{c(8, 12), c(7, 11), c(6, 9)\}$.

C. Topology of the extracted pore network and critical cells

As the pore-network is the graph structure, for the analysis of its topology it is natural to start from Euler number formula for a graph:

$$\chi = V - E + F, \tag{17}$$

with V vertices, E edges, and F faces, formed by loops between vertices.

In the case of pore-network graph, vertices are pores, edges are throats, and only those loops between pores should be counted which do not include the solid phase. Such loops correspond to junctions. Then Euler number of the pore network can be calculated as

$$\chi = N^p - N^t + N^j, \tag{18}$$

where N^p, N^t, N^j are numbers of pores, throats, and junctions.

As the Eq. (18) suggests, it is proposed that each pore contributes to total Euler number with the summand equal to 1. In reality, this condition is not fulfilled, as pores can have complicated intrinsic topological structure leading to the Euler number other than 1 for some pore summands. The examples of ring and cavity single pore structures are presented in Figs. 13(a) and 13(b). It can be argued that some of the cases in Fig. 13 are related to a case where solid phase "hangs" within the pore space and is, thus, unphysical. While this is indeed true in some situations, creating a universal solution that can take into account any possible situation (e.g., multiphase flow in nonbinary porous media) is our target here. Based on these considerations, the expression in Eq. (18) should be modified as

$$\chi = N^p - N^{th} + N^j + \Delta\chi_{\text{intr}}, \tag{19}$$

where the member $\Delta\chi_{\text{intr}}$ characterizes the contribution of single pores intrinsic topological structures.

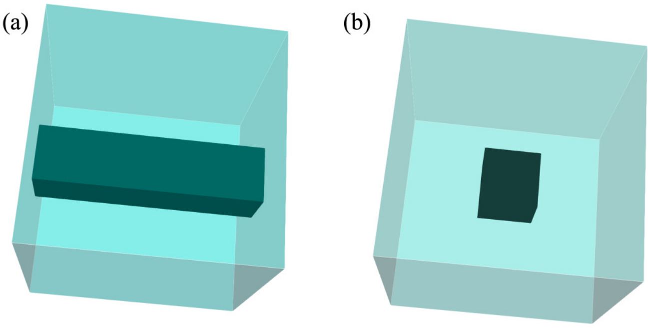

FIG. 13. Ring (a) and cavity (b) structures in a single pore. Pore phase, formed by the light blue box, is surrounded by solid phase. Solid phases inside the pores are presented by black boxes. Euler numbers are equal to 0 (a) and 2 (b).

Numbers of different pore-network elements and the values $\Delta\chi_{\text{intr}}$, calculated for all samples are presented in Table III. The number of pores, throats, and junctions can be calculated during the pore-network extraction, using the algorithms described in the previous subsection. The direct calculation of the values $\Delta\chi_{\text{intr}}$ is the problem. First, some pores can have very complicated intrinsic topological structures other than simple rings and cavities presented in [Figs. 13(a) and 13(b)]. Second, several pores can share one ringlike structure. Imagine, for example, the rings of different outer diameters, corresponding to pores, put on the rod, corresponding to the solid phase. Fortunately, we can extract these values using the equality of Euler numbers calculated with formulas in Eq. (19) and Eq. (13), since the cubical complex and pore network extracted from it using discrete Morse theory should have the same topological properties.

Based on the analysis of the data in Table III we can conclude that the contribution of junctions and intrinsic pore topology to final results depends on the sample porosity and pore sizes. While both N^j and $\Delta\chi_{\text{intr}}$ values are less than 10 % from total pore number for low-porosity carbonate, soil, and sandstone samples, the relative values of N^j and $\Delta\chi_{\text{intr}}$ come up to, respectively, 78% and 23% for the highly porous ceramic sample. Whereas the sandstone sample has higher porosity than the soil sample, its relative values N^j and $\Delta\chi_{\text{intr}}$ are smaller, as by contrast to soil, it has many small size pores. Most of these pores both have no complicated intrinsic topology and barely form junctions. Note, also that $\Delta\chi_{\text{intr}}$ is positive for ceramic sample and negative for other samples. It means, that pores with complicated intrinsic topology mainly have cavity-like structure in ceramic and ringlike structure in other considered media.

The topology of cubical complex, as was discussed in Sec. II, is characterized by critical cells, and the expression for Euler number is given by Eq. (13). The choice of signed EDT as the discrete Morse functions guarantees us the absence of maxima and corresponding critical 3-cells in the pore phase, and $N_3 = 0$. The numbers of critical cells obtained using discrete Morse theory and Euler number values for several binary 3D images are presented in Table IV.

According to our findings here, one can conclude that each extracted pore is characterized by critical 0-cell, which determines the position of the pore center. This leads to the equality $N^p = N_0$.

In the case of throats, the picture is more complicated. The results of throats' extraction for several binary images are presented in Table V. Here we observe that the throats containing one critical 1-cell provide the dominant contribution to the throats' total number. Nevertheless, the contribution of throats

TABLE IV. Numbers of critical cells N_i $(i = 0, 1, 2)$ obtained using discrete Morse theory and Euler number values for several binary 3D images.

Sample	N_0	N_1	N_2	χ
Carbonate	37 741	25 252	717	13 206
Ceramic	9238	30 160	6227	−14 695
Soil	49 464	36 994	8471	20 941
Sandstone	84 401	73 192	3139	14 348

containing more than one critical 1-cell and without critical 1-cells ranges from 5% to 25% for different samples and, thus, cannot be neglected.

The throat containing n critical 1-cells with $n > 1$ are characterized by these cells and a set of $n - 1$ critical 2-cells to provide the correct Euler number, as it is shown in Fig. 14. As observed from Fig. 14(d), several critical 1-cells correspond to the parts of throats having curvatures of different sign, and critical 2-cells indicate that sign of the curvature has changed. The throat without critical 1-cells can be characterized by critical 1-cell with both vertices belonging to one of neighboring pores or can indicate junction [see Figs. 15(b) and 15(e)]. Large contribution of N_0^t in total throats number for ceramic and soil is explained by the large number of junctions within these samples.

Critical 1- and 2-cells also contribute to the characterization of single pores with complicated topology and junctions. There are two possible scenarios of junction characterization by critical cells (see Fig. 15). Three critical 0-cells as pore centers, three critical 1-cells as throat centers, and one critical 2-cell provide the correct Euler number equal to one [Figs. 15(c) and 15(e)]. Another possibility is the presence of three critical 0-cells and only two critical 1-cells, leaving one throat free from 1-critical cells [Figs. 15(b) and 15(d)]. As seen from Fig. 15(d), this scheme of critical cells is actualized when the size of one throat (**ab** in the considered example) is much less than the sizes of other two throats. At this point it is not clear how to incorporate junctions into the pore network. The easiest solution would be to represent them with a pore body with a size equal or smaller than the surrounding throats; this is correct from at least topological considerations. On the other hand, we argue that direct single and, especially, multiphase flow simulations are needed to understand the influence of the junctions on the flow properties that could reveal a better idea how to include them into the PNM model.

TABLE V. Numbers of different type throats extracted using discrete Morse theory for several binary 3D images. N_i^t is the number of throat cell subsets containing exactly i critical cells $(i = 0, 1)$. $N_{>1}^t$ is the number of throat cell subsets containing more than one critical 1-cell. N_{total}^t is the total number of throats.

Sample	N_0^t	N_1^t	$N_{>1}^t$	N_{total}^t
Carbonate	943	23 616	366	24 925
Ceramic	6576	24 588	2135	33 299
Soil	2614	23 074	1990	27 678
Sandstone	2033	64 246	974	67 253

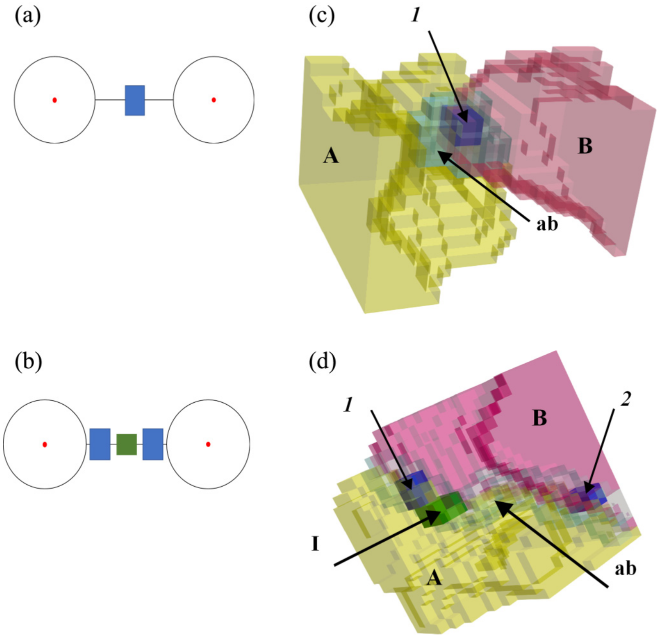

FIG. 14. (a, b) Scheme of topology features revealed in the cubical complex and the pore network in the case of two pores connected by the throat: (a) one critical cell at the throat; (b) two critical 1-cells and one critical 2-cell at the throat, which is topologically equivalent to case (a). Pores and channels are presented by circles and lines, respectively. Critical 0-, 1-, and 2-cells are presented by red points, blue rectangles and green squares, respectively. (c, d) Actualization of schemes (a) and (b), respectively, in soil sample. **A** (yellow), **B** (red) are pores, **ab** (light blue) is a throat between them, ***1***, ***2*** (dark blue) are critical 1-cells, **I** (green) is a critical 2-cell. Euler number is equal to 1.

For a single pore with ring structure [Fig. 13(a)] one critical 0-cell and one critical 1-cell provide the correct Euler number. A single pore with cavity structure [Fig. 13(b)] is characterized by 1 critical 0-cell, k critical 1-cells, and $k + 1$ critical 2-cells, where the value $k \geqslant 0$ depends on the position, form, and relative volume of a cavity.

D. Calculated pore-network characteristics: Comparison to other extraction methods

The results of pore-network extraction from real 3D porous media images shown in Fig. 5 using our approach are depicted as conventional ball-and-stick diagrams in Fig. 16. Pore-network characteristics calculated for these samples are presented in Table VI. The results as shown in the table, represent our method presented in this paper (Morse), as compared with PNMs obtained using maximum inscribed ball (MIB) [38], watershed-based (WS) [35], and hybrid [40] approaches. We have also compared pore and throat radius distributions, and coordination number distributions obtained using different pore-network extraction models for all these samples (Figs. 17–20). We normalized the presented distributions separately for each sample and extraction method. According to the MIB-based algorithm of pore-network extraction [37,38], first, a maximal inscribed ball is found by inflation and deflation at each pore voxel of the considered

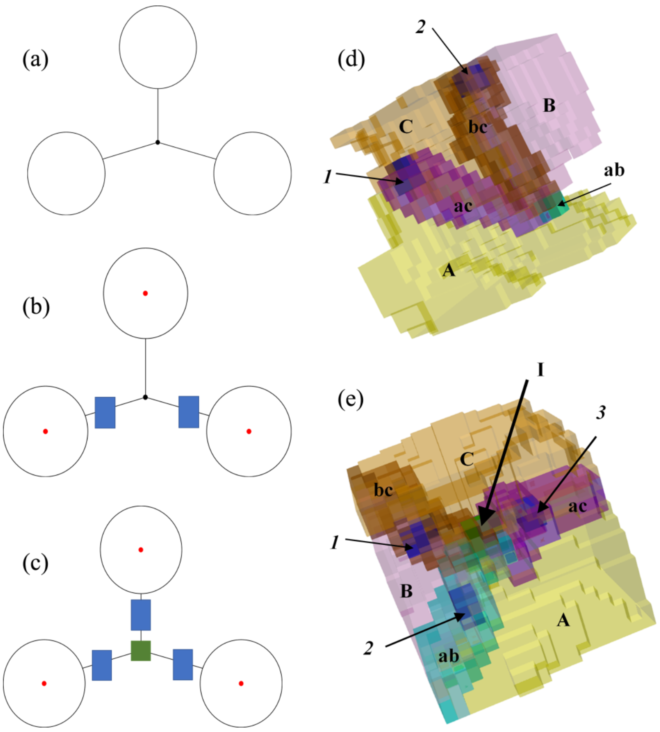

FIG. 15. Scheme of junction point revealed in the pore network (a) and the cubical complex (b, c). In both systems Euler number is equal to 1. Pores, channels, and junction are presented by circles, lines, and black point, respectively. Critical 0-, 1-, and 2-cells are presented by red points, blue rectangles, and green squares, respectively. (d, e) Actualization of schemes (b) in soil sample and (c) in ceramic sample, respectively. **A** (yellow), **B** (red), and **C** (red) are pores, **ab** (light blue), **bc** (brown), and **ac** (purple) are throats between them, ***1***, ***2***, ***3*** (dark blue) are critical 1-cells, **I** (green) is a critical 2-cell. Euler number is equal to 1.

TABLE VI. Pore N^p and throat N^t numbers, mean pore $< R_p >$ and throat $< R_t >$ radii, and mean coordination numbers Z calculated for several samples using different pore-network extraction methods.

Characteristic	Morse	MIB	WS	Hybrid
Carbonate				
N^p	37 741	29 355	12 685	38 496
N^t	24 925	25 111	5071	36 462
$< R_p >$, μm	14.3	10.4	12.0	13.4
$< R_t >$, μm	8.8	6.1	9.3	9.3
$< Z >$	1.3	1.7	0.8	1.9
Ceramic				
N^p	9238	9636	5174	13 778
N^t	33 299	32 918	14 469	38 095
$< R_p >$, μm	11.4	10.3	9.9	10.9
$< R_t >$, μm	5.1	6.1	12.6	6.6
$< Z >$	7.2	6.8	5.6	5.5
Soil				
N^p	49 464	35 690	3908	46 705
N^t	27 678	37 026	10 540	50 157
$< R_p >$, μm	28.5	37.6	43.6	50.1
$< R_t >$, μm	27.8	23.1	42.6	35.1
$< Z >$	1.1	2.0	0.7	2.1
Sandstone				
N^p	84 401	54 518	18 497	80 017
N^t	67 253	85 570	73 003	120 557
$< R_p >$, μm	8.0	11.2	31.6	13.6
$< R_t >$, μm	4.7	5.9	2.6	8.7
$< Z >$	1.6	3.1	7.9	3.0

binary 3D image (the ball inflation can be substituted with EDS, something that became a *de facto* inscribed ball in later works). After all MIBs are placed, they are sorted and clustered in a hierarchy. The MIBs of largest radii become the parent of family trees, and a new MIB with smaller or equal radius is added to this tree as a child if it has common voxels with the MIBs from previous generation. This clustering provides segmentation of pore space into pores with the centers at the centers of parent MIBs. Set of voxels belonging to different families is treated as the throat between corresponding pores, and the throat radius is defined as the minimal MIB radius value within all the throat voxels.

WS-based algorithms [34,35] start with distant transform for each pore voxel of the considered binary 3D image. Then a Gaussian filter and a maximum filter are applied to the obtained distance map and peaks of the resulting smooth scalar function defined over pore voxels are analyzed and processed. Filtering also helps in saddle points removal and merging of the nearby peaks. The rest of the peaks are taken as the seeds of watershed segmentation provided by a marker-based algorithm [94]. Obtained basins and watersheds are treated as the pores and throats, respectively.

The hybrid approach [40] to pore-network extraction combines methods of MIB-based and WS-based algorithms. The centers of pores are found using MIB clustering, and segmentation into pores is carried out using the watershed algorithm [95].

Compared to MIB, hybrid approach leads to the larger number of pores due to the slightly different methods for MIB finding. In [40] a looser criterion to maximum ball placement is applied to get the speed of execution compared to the algorithm presented in the classical paper [38]. The difference in pore numbers strongly depends on the simplification parameters used in different pore-network extraction models: persistent limit in Morse-based model and parameters of filters in WS-based model. The number of pores is minimal in WS-based model, so we can conclude it provides the most intensive merging of nearby pores. The pore number calculated using Morse-based model can be larger or smaller than the values of this characteristic calculated using MIB-based and hybrid models depending on the sample considered. It is influenced by two differently directed trends. First, the expansion of MIBs clusters is carried out with 26-connectivity condition in MIB-based and hybrid models, while in our algorithm 6-connectivity is used. This leads to the larger number of pores in the Morse-based model. In WS-based model the preprocessing of peaks before the watershed segmentation is performed with the 26-connectivity condition as well. Second, no pore merging is applied in MIB-based and hybrid models, while simplification in Morse-based model reduces the number of pores. Nevertheless, one can point out the closeness of N_p values found using Morse-based and hybrid approaches

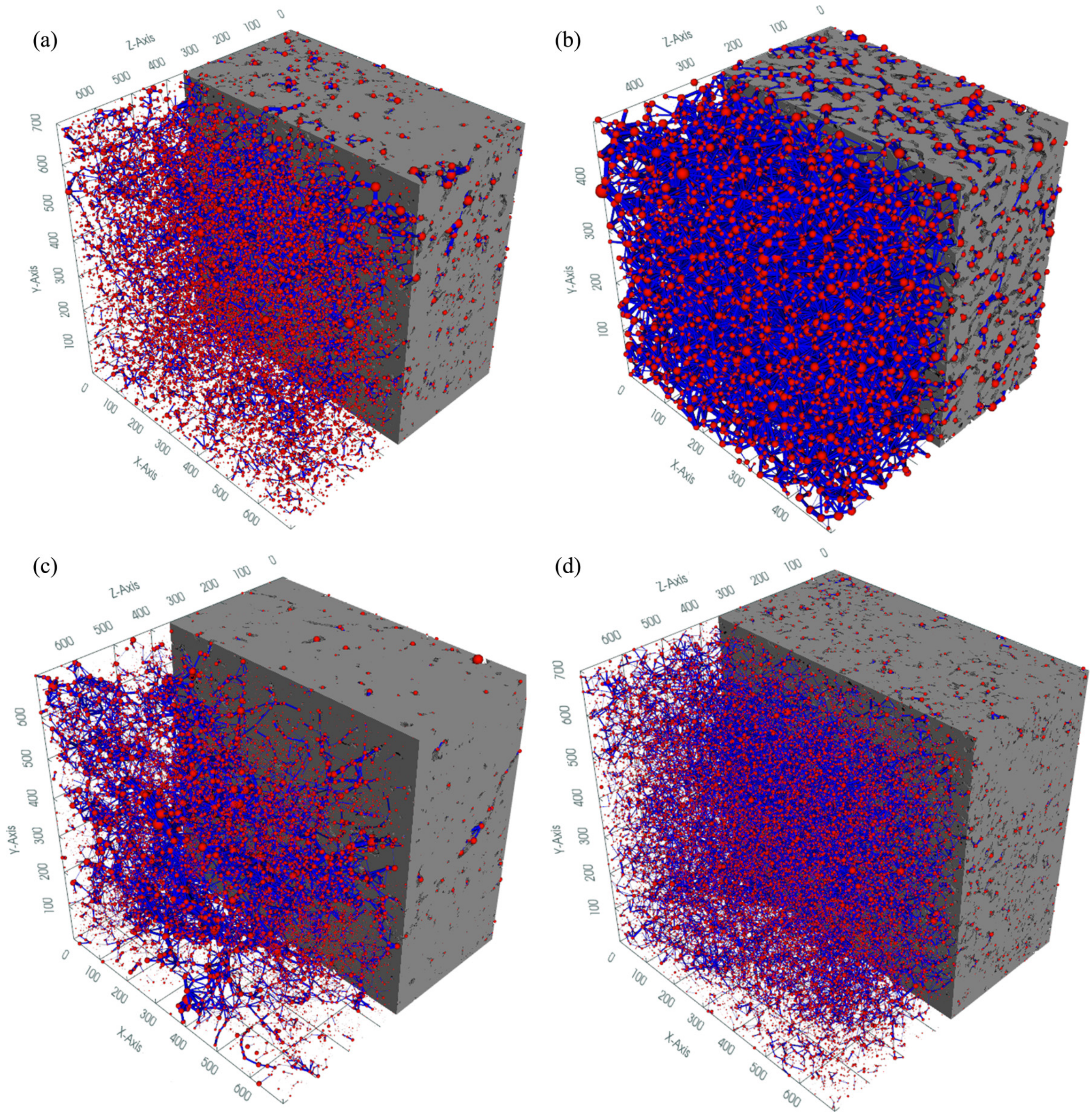

FIG. 16. Ball-and-stick diagrams of pore networks extracted from XCT 3D images of studied samples: (a) carbonate; (b) ceramic; (c) soil; (d) sandstone. Pores and throats are presented by red balls and blue cylinders, respectively. Ball and cylinder radii are proportional to the radii of corresponding pores and throats.

for all the considered samples except for highly porous ceramic.

The increase in the pore number leads to increase in throat count, yet the increment is not linear and depends on the local pore geometry for each sample. Trends of increasing throat number due to 6-connectivity in our model and its decreasing caused by simplification are also actual. Note, however, that simplification has the stronger influence on N^t value than on N^p value, since when two pores are merged there can be more than one throat between them. As Euler number is conserved in Morse-based algorithms, single-pore ring structures appear in the pore network after simplification in this case. This case is schematically presented in Fig. 21. Before the simplification, we have $N^p = 2$, $N^t = 2$, $\Delta\chi_{\text{intr}} = 0$, and after the simplification $N^p = 1$, $N^t = 0$, $\Delta\chi_{\text{intr}} = -1$. So the pore number is decreased by 1, and the throat number is decreased by 2. Thus, the values of N^t calculated in our model are smaller than those obtained using hybrid approach. The smaller number of throats lead to the smaller coordination number.

The maximum in coordination number distribution obtained using Morse-based model is observed at zero for all samples except for ceramic. To clarify its origin we calculated numbers of nonconnected pores $N^p_{\text{con}=0}$ in different models and compared them against the total number of closed pore space components N^p_{closed}, obtained with 6- and 26-connectivity conditions. The results are presented in Table VII. One can see from this table that almost all closed pore components consist of only one pore due to simplification in Morse-based mode, since in other models they are likely to be segmented into several pores. In the ceramic the sample number of closed

TABLE VII. Numbers of pores $N^p_{\text{con}=0}$, calculated in different pore-network extraction models and the total numbers of closed pore space components N^p_{closed} obtained with 6- (6-con.) and 26-connectivity (26-con.) conditions in the samples considered.

Sample	$N^p_{\text{con}=0}$, Morse	$N^p_{\text{con}=0}$, MIB	$N^p_{\text{con}=0}$, WS	$N^p_{\text{con}=0}$, hybrid	N^p_{closed}, 6-con.	N^p_{closed}, 26-con.
Carbonate	16 970	7530	7845	7404	19 248	18 666
Ceramic	324	68	111	74	324	264
Soil	36 879	8061	7358	8265	37 889	29 821
Sandstone	46 839	6236	129	6378	47 806	39 540

pore space components is very small, and its influence is negligible.

From the analysis of Table VII, one can also note rather large difference between the values N^p_{closed} calculated with 6- and 26-connectivity conditions for soil and sandstone samples. Small volume closed pore components entirely belong to one pore in all the compared pore-network extraction models, therefore their excess in 6-connectivity analysis leads to the shift of pore radius distribution to small values in Morse-based approach for these samples. Average throat radii calculated in our model are rather close to the results obtained using the other models.

Euler numbers calculated using discrete Morse theory (Morse on figures) based on Eq. (13) for several samples are presented in Table VIII. The values obtained with approximate expression as the difference between pores and throats numbers:

$$\chi \approx N^p - N^t \tag{20}$$

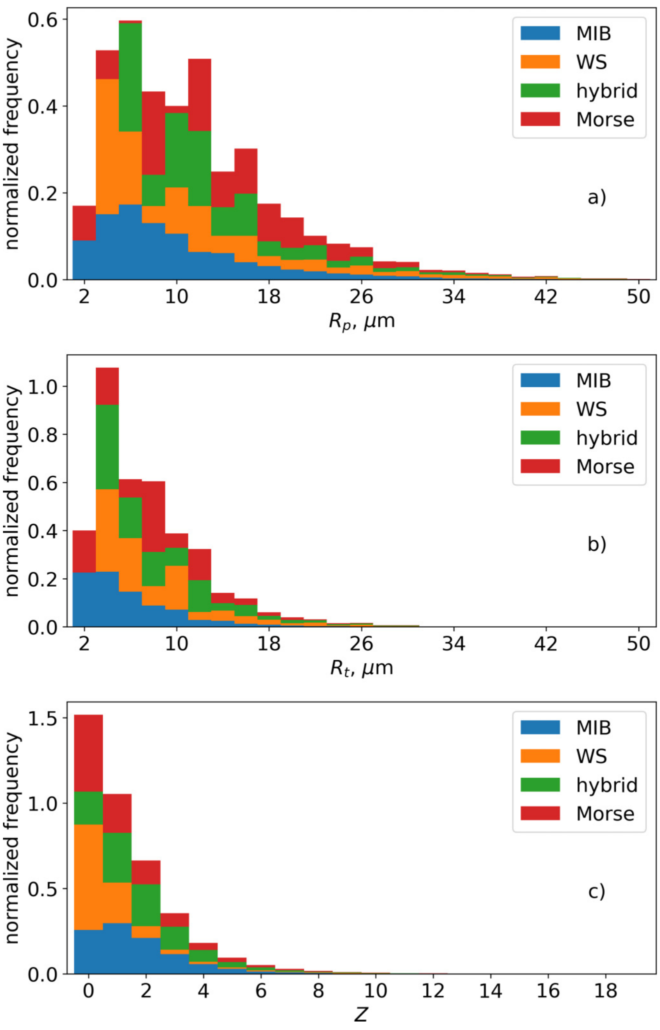

FIG. 17. Pore radius (a), throat radius (b), and coordination number (c) distributions extracted using different models for carbonate sample.

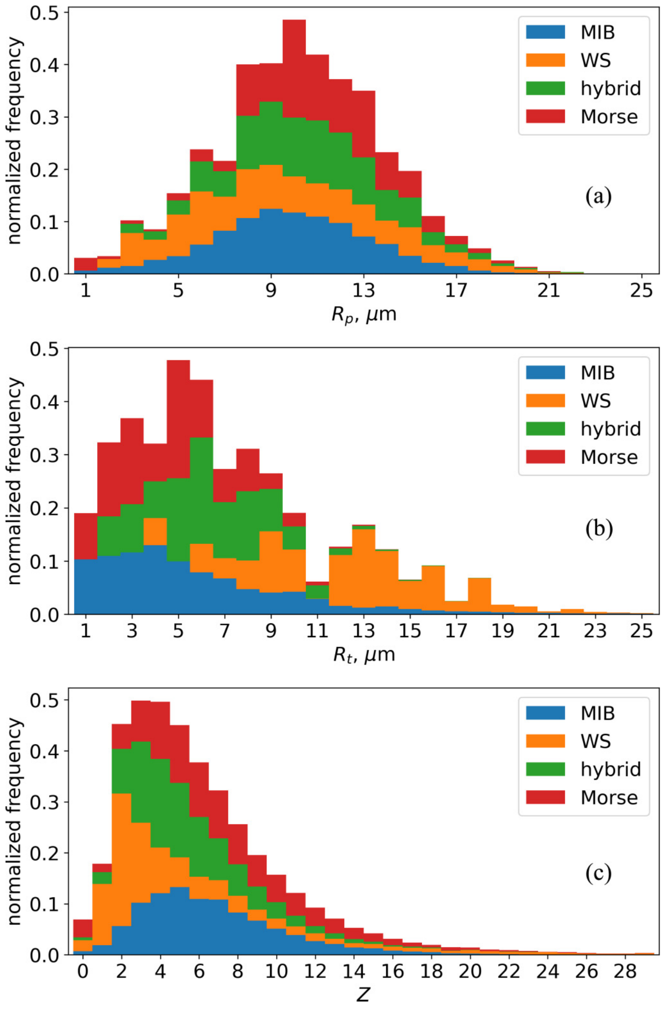

FIG. 18. Pore radius (a), throat radius (b), and coordination number (c) distributions extracted using different models for ceramic sample.

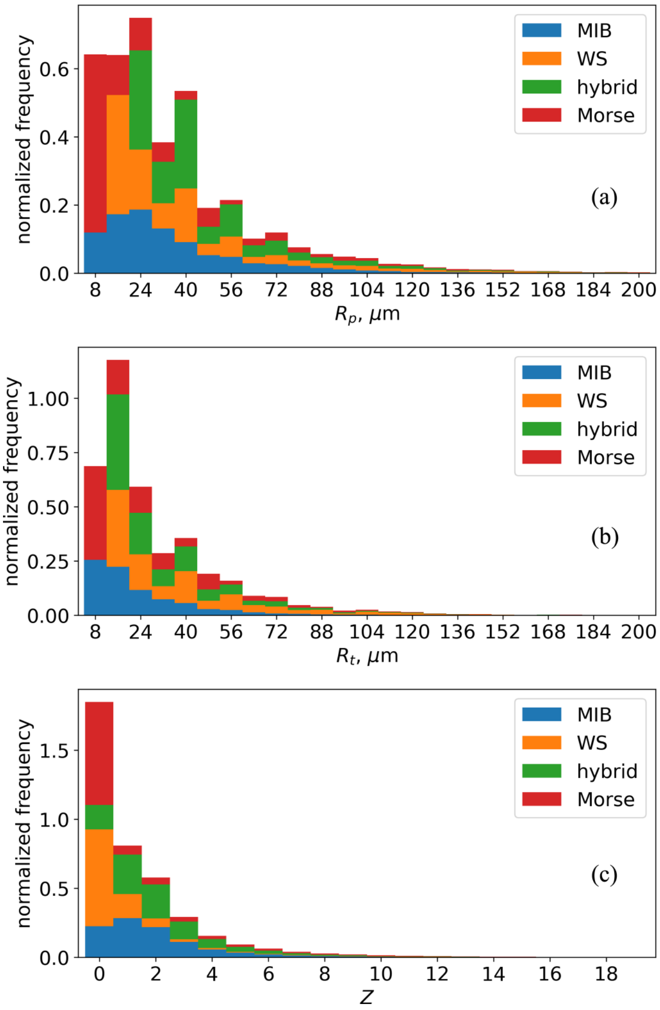

FIG. 19. Pore radius (a), throat radius (b), and coordination number (c) distributions extracted using different models for soil sample.

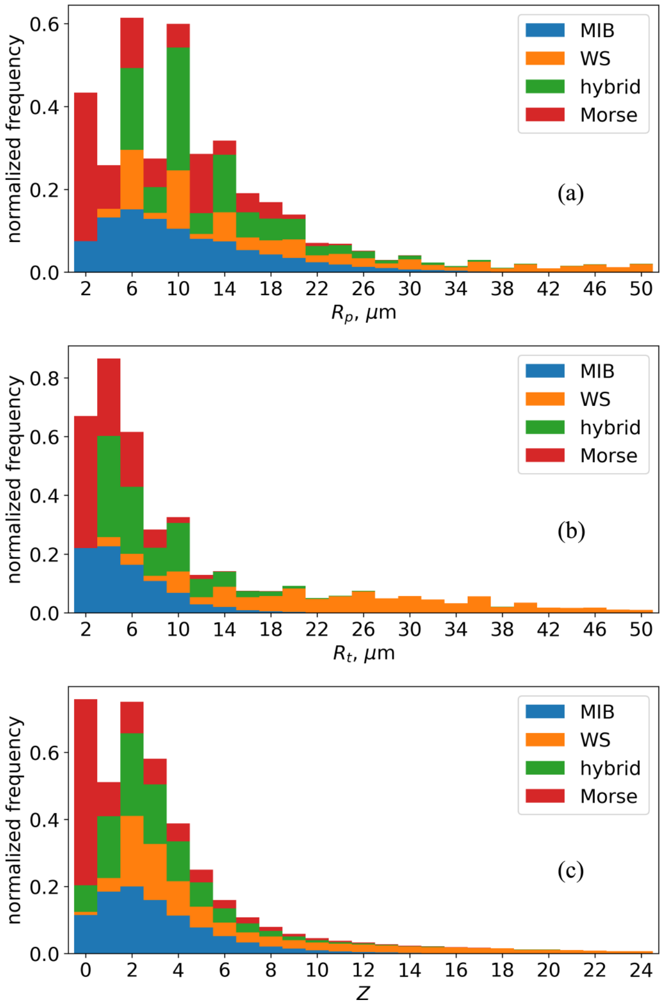

FIG. 20. Pore radius (a), throat radius (b), and coordination number (c) distributions extracted using different models for sandstone sample.

are also presented. Our results are now compared with those assessed using other pore-network extraction models (MIB, WS, hybrid; see figures).

In Table VIII we also present Euler numbers calculated with the help of open source library EulerCharacteristic.jl [96] and software packages for image analysis and Minkowski functionals evaluation, MorphoLibJ [97] for ImageJ and Quantim code [85].

TABLE VIII. Euler numbers calculated for studied samples using different PNM extraction models and different software packages for image analysis.

	Carbonate	Ceramic	Soil	Sandstone
PNM				
Morse, (13)	13 206	−14 695	20 941	14 348
Morse, (20)	12 816	−24 061	21 718	17 148
MIB, (20)	4989	−22 491	−318	−29 501
WS, (20)	6714	−9295	−6632	−54 506
Hybrid, (20)	3037	−23 212	−2256	−38 281
3D Image				
EulerC.jl	13 206	−14 695	20 941	14 348
MorphoLibJ	13 206	−14 695	20 941	14 348
Quantim	11 622	−14 949	20 903	11 319

From the analysis of Table VIII one can note the coincidence of values calculated with formula (13) and obtained using EulerCharacteristic.jl and MorphoLibJ package. The rather small disagreement of these data with results obtained using Quantim code can be explained by boundary effects. As our calculations of Euler numbers using formulas in Eq. (13), EulerCharacteristic.jl and MorphoLibJ calculations are carried out with open image boundaries, stereological rule [98] is applied to boundaries in Quantim. In the case of open boundaries, all objects that contact boundaries are included. Due to stereological rule, all objects that contact the top, left, and front boundary are included but all objects that contact the bottom, right, and back boundary are excluded.

The comparison of Euler numbers values calculated in our approach using Eq. (13) and Eq. (20) let us conclude that using the approximate expression (20) leads to the small errors in the case of low porosity samples, but with the increase of

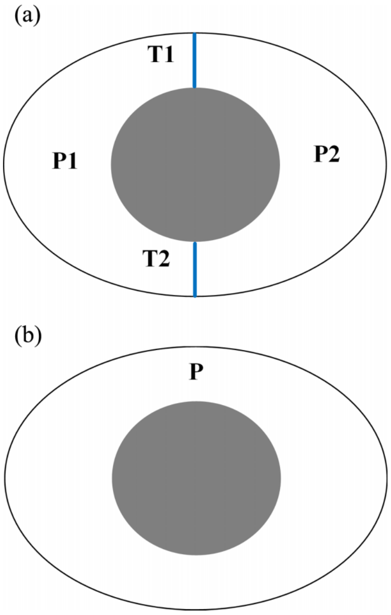

FIG. 21. Schematic picture of pores and throats before (a), and after (b) simplification. Gray area indicates the solid phase. First, pores **P1** and **P2** are connected by two throats **T1** and **T2**. Then simplification can merge these pores into one pore **P** with intrinsic ring structure, as in Fig. 13(a). Euler number is equal to 0 in both cases.

porosity the deviations from the correct Euler number grow up considerably.

Morse-based, MIB-based, and hybrid models lead to close Euler number values calculated using the formula in Eq. (20). In other cases, the large deviation from the correct values is observed for Euler numbers calculated using other pore-network extraction algorithms other than Morse-based. For soil and sandstone samples other methods even produce the incorrect sign of χ. This fact confirms the importance of the accurate topology preservation in pore-network extraction—this is provided by discrete Morse theory solution presented in this paper.

IV. SUMMARY AND OUTLOOK

In this paper we present pore-network extraction algorithm for binary 3D images based on discrete Morse theory and persistent homology. The main advantage of our method is conserving the topological properties of the original 3D image in the extracted pore network. This is confirmed by the coincidence of Euler numbers calculated in PNM and on 3D image using direct topological analysis.

The first step of our algorithm of pore-network extraction is the choice of discrete Morse function and the construction of gradient vector field. To avoid any systematic shifts in final results, we propose to use the randomly perturbed signed EDT function as Morse function. ProcessLowerStar algorithm [62] was applied to construct gradient vector field.

The further simplification of gradient vector field leads to the removal of persistent pairs which consist of critical p- and $p-1$-cells. The part of pairs to be removed is controlled by persistent limit parameter. We showed that the characteristics of the extracted pore network are very sensitive to the value of this parameter and proposed the method how to fix it. The choice of persistent limit equal to 1 seems to be reasonable for all the samples considered in this paper. However, the influence of this parameter on final results can be studied further for larger set of 3D images.

The gradient vector field transformed after simplification is the starting point for segmentation of pore space to separate pores. The stable set of 0-cells is constructed by V-path traversals for each critical 0-cell. All 0-cells which belong to this set correspond to the voxels forming the separate pore. 1-cells with vertices belonging to different pores are included into the set corresponding to the throat between these pores. Junction points are found from the analysis of loops including three and four pores.

The topological properties of the initial cubical complex are the same as for Morse chain complex extracted from gradient vector field. They are characterized by critical cells. For the extracted pore network, topology is defined by configuration of pores, throats, and junctions. We clarified the relationship between these pore-network characteristics and critical cells. Critical 0-cells coincide with pore centers, critical 1-cells correspond to throats, critical 2-cells can characterize throats containing more than one critical 1-cell and can indicate junctions. Also, critical 1- and 2-cells contribute into the characterization of single pores with complicated intrinsic topology.

We compared pore-network characteristics calculated for several samples using discrete Morse theory with ones calculated using MIB-based, WS-based, and hybrid models. While for some specific samples and characteristics the other models lead to the results close to ones obtained in our approach, in general, the deviations are quite large. The difference between Euler numbers calculated using discrete Morse theory and other models is especially remarkable.

The results of our analysis and the developed approach open the way for deeper understanding of pore-network topological features on the flow simulations. It was shown already that the permeability can be predicted with high accuracy with PNMs. On the other hand, multiphase flow does not seem to be simulated adequately in modern pore-network simulators [40,52]. Further development of our PNM extraction coupled with a novel PNM flow simulator to study such influences is the focus of ongoing research.

ACKNOWLEDGMENTS

This research is a part of Rosneft Digital core technology development. We thank Dr. A. Pashali and Dr. K. Toropov for the opportunity to publish this research. The collaborative effort of the authors is within the FaT iMP (Flow and Transport in Media with Pores) research group. We are grateful to our colleagues A. Ananev and M. Karsanina for useful discussions and help with figure creation.

[1] M. Sahimi, *Heterogeneous Materials I: Linear Transport Properties and Optical Properties* (Springer-Verlag, New York, 2003)

[2] S. Torquato, *Random Heterogeneous Materials: Microstructure and Macroscopic Properties* (Springer-Verlag, New York, 2002).

[3] M. J. Blunt, *Multiphase Flow in Permeable Media: A Pore-Scale Perspective* (Cambridge University Press, Cambridge, 2017).

[4] V. Cnudde and M. N. Boone, High-resolution x-ray computed tomography in geosciences: A review of the current technology and applications, Earth-Sci. Rev. **123**, 1 (2013).

[5] S. Hemes, G. Desbois, J. L. Urai, B. Schröppel, and J.-O. Schwarz, 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 Mesoporous Mater. **208**, 1 (2015).

[6] K. M. Gerke, E. V. Korostilev, K. A. Romanenko, and M. V. Karsanina, Going submicron in the precise analysis of soil structure: A FIB-SEM imaging study at nanoscale, Geoderma **383**, 114739 (2021).

[7] K. M. Gerke, M. V. Karsanina, and D. Mallants, Universal stochastic multiscale image fusion: An example application for shale rock, Sci. Rep. **5**, 15880 (2015).

[8] P. Adler, C. Jacquin, and J. Quiblier, Flow in simulated porous media, Int. J. Multiphase Flow **16**, 691 (1990).

[9] C. L. Y. Yeong and S. Torquato, Reconstructing random media, Phys. Rev. E **57**, 495 (1998).

[10] P. Tahmasebi and M. Sahimi, Cross-Correlation Function for Accurate Reconstruction of Heterogeneous Media, Phys. Rev. Lett. **110**, 078002 (2013).

[11] A. Derossi, K. M. Gerke, M. V. Karsanina, B. Nicolai, P. Verboven, and C. Severini, Mimicking 3D food microstructure using limited statistical information from 2D cross-sectional image, J. Food Eng. **241**, 116 (2019).

[12] M. V. Karsanina and K. M. Gerke, Hierarchical Optimization: Fast and Robust Multiscale Stochastic Reconstructions with Rescaled Correlation Functions, Phys. Rev. Lett. **121**, 265501 (2018).

[13] R. Guibert, M. Nazarova, P. Horgue, G. Hamon, P. Creux, and G. Debenest, Computational permeability determination from pore-scale imaging: Sample size, mesh and method sensitivities, Transp. Porous Media **107**, 641 (2015).

[14] K. M. Gerke, M. V. Karsanina, and R. Katsman, Calculation of tensorial flow properties on pore level: Exploring the influence of boundary conditions on the permeability of three-dimensional stochastic reconstructions, Phys. Rev. E **100**, 053312 (2019).

[15] M. V. Karsanina, K. M. Gerke, E. B. Skvortsova, A. L. Ivanov, and D. Mallants, Enhancing image resolution of soils by stochastic multiscale image fusion, Geoderma **314**, 138 (2018).

[16] X. Li, Q. Teng, Y. Zhang, S. Xiong, and J. Feng, Three-dimensional multiscale fusion for porous media on microtomography images of different resolutions, Phys. Rev. E **101**, 053308 (2020).

[17] Y. Wu, P. Tahmasebi, C. Lin, L. Ren, and C. Dong, Multiscale modeling of shale samples based on low- and high-resolution images, Marine Petroleum Geol. **109**, 9 (2019).

[18] P. H. Valvatne and M. J. Blunt, Predictive pore-scale modeling of two-phase flow in mixed wet media, Water Resour. Res. **40**, W07406 (2004).

[19] K. M. Gerke, R. V. Vasilyev, S. Khirevich, D. Collins, M. V. Karsanina, T. O. Sizonenko, D. V. Korost, S. Lamontagne, and D. Mallants, Finite-difference method Stokes solver (FDMSS) for 3D pore geometries: Software development, validation and case studies, Comput. Geosci. **114**, 41 (2018).

[20] S. Khirevich and T. W. Patzek, Behavior of numerical error in pore-scale lattice Boltzmann simulations with simple bounce-back rule: Analysis and highly accurate extrapolation, Phys. Fluids **30**, 093604 (2018).

[21] T. W. Patzek, Verification of a complete pore network simulator of drainage and imbibition, SPE J. **6**, 144 (2001).

[22] S. Bakke and P.-E. Øren, 3-D pore-scale modelling of sandstones and flow simulations in the pore networks, SPE J. **2**, 136 (1997).

[23] L. Ruspini, R. Farokhpoor, and P. Øren, Pore-scale modeling of capillary trapping in water-wet porous media: A new cooperative pore-body filling model, Adv. Water Resour. **108**, 1 (2017).

[24] J. Roof, Snap-off of oil droplets in water-wet pores, Soc. Petroleum Eng. J. **10**, 85 (1970).

[25] A. R. Kovscek, H. Wong, and C. J. Radke, A pore-level scenario for the development of mixed wettability in oil reservoirs, AIChE J. **39**, 1072 (1993).

[26] M. van Dijke and K. Sorbie, Existence of fluid layers in the corners of a capillary with non-uniform wettability, J. Colloid Interface Sci. **293**, 455 (2006).

[27] R. Lenormand, C. Zarcone, and A. Sarr, Mechanisms of the displacement of one fluid by another in a network of capillary ducts, J. Fluid Mech. **135**, 337 (1983).

[28] X. Miao, K. M. Gerke, and T. O. Sizonenko, A new way to parameterize hydraulic conductances of pore elements: A step towards creating pore-networks without pore shape simplifications, Adv. Water Resour. **105**, 162 (2017).

[29] A. Rabbani and M. Babaei, Hybrid pore-network and lattice-Boltzmann permeability modelling accelerated by machine learning, Adv. Water Resour. **126**, 116 (2019).

[30] M. Tembely, A. M. AlSumaiti, and W. S. Alameri, Machine and deep learning for estimating the permeability of complex carbonate rock from x-ray micro-computed tomography, Energy Rep. **7**, 1460 (2021).

[31] N. Misaghian, M. Agnaou, M. A. Sadeghi, H. Fathiannasab, I. Hadji, E. Roberts, and J. Gostick, Prediction of diffusional conductance in extracted pore network models using convolutional neural networks, Comput. Geosci. **162**, 105086 (2022).

[32] A. Aghaei and M. Piri, Direct pore-to-core up-scaling of displacement processes: Dynamic pore network modeling and experimentation, J. Hydrol. **522**, 488 (2015).

[33] V. Joekar-Niasar and S. Hassanizadeh, Analysis of fundamentals of two-phase flow in porous media using dynamic pore-network models: A review, Crit. Rev. Environ. Sci, Tech. **42**, 1895 (2012).

[34] A. Rabbani, S. Jamshidi, and S. Salehi, An automated simple algorithm for realistic pore network extraction from micro-tomography images, J. Petroleum Sci. Eng. **123**, 164 (2014).

[35] J. T. Gostick, Versatile and efficient pore network extraction method using marker-based watershed segmentation, Phys. Rev. E **96**, 023307 (2017).

[36] M. Prodanović, W. Lindquist, and R. Seright, Porous structure and fluid partitioning in polyethylene cores from 3D x-ray microtomographic imaging, J. Colloid Interface Sci. **298**, 282 (2006).

[37] D. Silin and T. Patzek, Pore space morphology analysis using maximal inscribed spheres, Physica A **371**, 336 (2006).

[38] H. Dong and M. J. Blunt, Pore-network extraction from micro-computerized-tomography images, Phys. Rev. E **80**, 036307 (2009).

[39] A. Q. Raeini, B. Bijeljic, and M. J. Blunt, Generalized network modeling: Network extraction as a coarse-scale discretization of the void space of porous media, Phys. Rev. E **96**, 013312 (2017).

[40] K. M. Gerke, T. O. Sizonenko, M. V. Karsanina, E. V. Lavrukhin, V. V. Abashkin, and D. V. Korost, Improving watershed-based pore-network extraction method using maximum inscribed ball pore-body positioning, Adv. Water Resour. **140**, 103576 (2020).

[41] T. Patzek and D. Silin, Shape factor and hydraulic conductance in noncircular capillaries: I. One-phase creeping flow, J. Colloid Interface Sci. **236**, 295 (2001).

[42] A. V. Ryazanov, M. I. J. van Dijke, and K. S. Sorbie, Two-phase pore-network modelling: Existence of oil layers during water invasion, Transp. Porous Media **80**, 79 (2009).

[43] X. Li and Y. Yortsos, Theory of multiple bubble growth in porous media by solute diffusion, Chem. Eng. Sci. **50**, 1247 (1995).

[44] C. Lu and Y. C. Yortsos, Pattern formation in reverse filtration combustion, Phys. Rev. E **72**, 036201 (2005).

[45] A. G. Yiotis, A. G. Boudouvis, A. K. Stubos, I. N. Tsimpanogiannis, and Y. C. Yortsos, Effect of liquid films on the isothermal drying of porous media, Phys. Rev. E **68**, 037303 (2003).

[46] M. Chen, Y. C. Yortsos, and W. R. Rossen, Pore-network study of the mechanisms of foam generation in porous media, Phys. Rev. E **73**, 036304 (2006).

[47] V. Surasani, T. Metzger, and E. Tsotsas, Consideration of heat transfer in pore network modelling of convective drying, Int. J. Heat Mass Transf. **51**, 2506 (2008).

[48] O. Maalal, M. Prat, and D. Lasseux, Pore network model of drying with Kelvin effect, Phys. Fluids **33**, 027103 (2021).

[49] R. Acharya, S. Van der Zee, and A. Leijnse, Approaches for modeling longitudinal dispersion in pore-networks, Adv. Water Resour. **30**, 261 (2007).

[50] Y. Gong and M. Piri, Pore-to-core upscaling of solute transport under steady-state two-phase flow conditions using dynamic pore network modeling approach, Transp. Porous Media **135**, 181 (2020).

[51] F. Larachi, R. Hannaoui, P. Horgue, F. Augier, Y. Haroun, S. Youssef, E. Rosenberg, M. Prat, and M. Quintard, X-ray microtomography and pore network modeling of single-phase fixed-bed reactors, Chem. Eng. J. **240**, 290 (2014).

[52] T. G. Baychev, A. P. Jivkov, A. Rabbani, A. Q. Raeini, Q. Xiong, T. Lowe, and P. J. Withers, Reliability of algorithms interpreting topological and geometric properties of porous media for pore network modelling, Transp. Porous Media **128**, 271 (2019).

[53] Z. A. Khan, A. Elkamel, and J. T. Gostick, Efficient extraction of pore networks from massive tomograms via geometric domain decomposition, Adv. Water Resour. **145**, 103734 (2020).

[54] Z. Jiang, M. Van Dijke, K. Wu, G. D. Couples, K. S. Sorbie, and J. Ma, Stochastic pore network generation from 3D rock images, Transp. Porous Media **94**, 571 (2012).

[55] S. Park, S. Lim, P. Siriviriyakul, and J. S. Jeon, Three-dimensional pore network characterization of reconstructed extracellular matrix, Phys. Rev. E **101**, 052414 (2020).

[56] R. Forman, Morse theory for cell complexes, Adv. Math. **134**, 90 (1998).

[57] T. Lewiner, H. Lopes, and G. Tavares, Optimal discrete Morse functions for 2-manifolds, Comput. Geom, **26**, 221 (2003).

[58] H. King, K. Knudson, and N. Mramor, Generating discrete Morse functions from point data, Exp. Math. **14**, 435 (2005).

[59] J. Reininghaus, D. Günther, I. Hotz, S. Prohaska, and H.-C. Hege, Tadd: A computational framework for data analysis using discrete Morse theory, in *International Congress on Mathematical Software*, edited by K. Fukuda, J. van der Hoeven, M. Joswig, and N. Takayama (Springer, 2010), pp. 198–208.

[60] A. Gyulassy, P.-T. Bremer, B. Hamann, and V. Pascucci, A practical approach to Morse-Smale complex computation: Scalability and generality, IEEE Trans. Visual. Comput. Graphics **14**, 1619 (2008).

[61] A. G. Gyulassy, Combinatorial construction of Morse-Smale complexes for data analysis and visualization, Ph.D. thesis, University of California, Davis, 2008.

[62] V. Robins, P. J. Wood, and A. P. Sheppard, Theory and algorithms for constructing discrete morse complexes from grayscale digital images, IEEE Trans. Pattern Anal. Mach. Intell. **33**, 1646 (2011).

[63] D. Günther, J. Reininghaus, H. Wagner, and I. Hotz, Efficient computation of 3D Morse–Smale complexes and persistent homology using discrete Morse theory, Visual Comput. **28**, 959 (2012).

[64] O. Delgado-Friedrichs, V. Robins, and A. Sheppard, Skeletonization and partitioning of digital images using discrete Morse theory, IEEE Trans. Pattern Anal. Mach. Intell. **37**, 654 (2014).

[65] H. Edelsbrunner, J. Harer, and A. Zomorodian, Hierarchical morse complexes for piecewise linear 2-manifolds, in *Proceedings of the Seventeenth Annual Symposium on Computational Geometry* (Association for Computing Machinery, New York, 2001), pp. 70–79.

[66] H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, Morse-smale complexes for piecewise linear 3-manifolds, in *Proceedings of the Nineteenth Annual Symposium on Computational Geometry* (Association for Computing Machinery, New York, 2003), pp. 361–370.

[67] A. Gyulassy, V. Natarajan, V. Pascucci, and B. Hamann, Efficient computation of Morse-Smale complexes for three-dimensional scalar functions, IEEE Trans. Visual. Comput. Graphics **13**, 1440 (2007).

[68] T. F. Banchoff, Critical points and curvature for embedded polyhedral surfaces, Am. Math. Mon. **77**, 475 (1970).

[69] H. Edelsbrunner, D. Letscher, and A. Zomorodian, Topological persistence and simplification, in *Proceedings 41st Annual Symposium on Foundations of Computer Science* (IEEE, 2000), pp. 454–463.

[70] M. Saadatfar, H. Takeuchi, V. Robins, N. Francois, and Y. Hiraoka, Pore configuration landscape of granular crystallization, Nat. Commun. **8**, 15082 (2017).

[71] F. Jiang, T. Tsuji, and T. Shirai, Pore geometry characterization by persistent homology theory, Water Resour. Res. **54**, 4150 (2018).

[72] I. Obayashi, Y. Hiraoka, and M. Kimura, Persistence diagrams with linear machine learning models, J. Appl. Comput. Topol. **1**, 421 (2018).

[73] C. Moon, S. A. Mitchell, J. E. Heath, and M. Andrew, Statistical inference over persistent homology predicts fluid flow in porous media, Water Resour. Res. **55**, 9592 (2019).

[74] M. M. Thakur, F. Kim, D. Penumadu, and A. Herring, Pore space and fluid phase characterization in round and angular partially saturated sands using radiation-based tomography and persistent homology, Transp. Porous Media **137**, 131 (2021).

[75] U. Bauer, C. Lange, and M. Wardetzky, Optimal topological simplification of discrete functions on surfaces, Discrete Comput. Geom. **47**, 347 (2012).

[76] T. Sousbie, C. Pichon, and H. Kawahara, The persistent cosmic web and its filamentary structure—II. Illustrations, Mon. Not. R. Astron. Soc. **414**, 384 (2011).

[77] N. Shivashankar, P. Pranav, V. Natarajan, R. van de Weygaert, E. P. Bos, and S. Rieder, Felix: A topology based framework for visual exploration of cosmic filaments, IEEE Trans. Visual. Comput. Graphics **22**, 1745 (2015).

[78] T. K. Dey, J. Wang, and Y. Wang, Road network reconstruction from satellite images with machine learning supported by topological methods, in *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (Association for Computing Machinery, New York, 2019), pp. 520–523.

[79] S. Banerjee, L. Magee, D. Wang, X. Li, B.-X. Huo, J. Jayakumar, K. Matho, M.-K. Lin, K. Ram, M. Sivaprakasam *et al.*, Semantic segmentation of microscopic neuroanatomical data by combining topological priors with encoder–decoder deep networks, Nat. Mach. Intell. **2**, 585 (2020).

[80] N. A. Scoville, *Discrete Morse Theory*, Student Mathematical Library, Vol. 90 (American Mathematical Society, Providence, RI, 2019).

[81] V. A. Kovalevsky, Finite topology as applied to image analysis, Comput Vision Graphics Image Proc. **46**, 141 (1989).

[82] C. Chen and M. Kerber, Persistent homology computation with a twist, in *Proceedings 27th European Workshop on Computational Geometry*, Vol. 11 (2011), pp. 197–200.

[83] J. Milnor, Morse theory.(am-51), Vol. 51, in *Morse Theory.(AM-51)*, Annals of Mathematic Studies, Vol. 51 (Princeton University Press, Princeton, 1969).

[84] A. Hatcher, *Algebraic Topology* (Cambridge University Press, Cambridge, 2002).

[85] H.-J. Vogel, U. Weller, and S. Schlüter, Quantification of soil structure based on Minkowski functions, Comput. Geosci. **36**, 1236 (2010).

[86] O. Delgado-Friedrichs, Diamorse-digital image analysis using discrete Morse theory and persistent homology, https://github.com/AppliedMathematicsANU/diamorse (2016).

[87] K. M. Gerke and M. V. Karsanina, How pore structure non-stationarity compromises flow properties representativity (REV) for soil samples: Pore-scale modelling and stationarity analysis, Eur. J. Soil Sci. **72**, 527 (2021).

[88] K. Gerke, D. Korost, R. Vasilyev, M. Karsanina, and V. Tarasovskii, Studying structure and determining permeability of materials based on x-ray microtomography data (using porous ceramics as an example), Inorg. Mater. **51**, 951 (2015).

[89] A. Meijster, J. B. Roerdink, and W. H. Hesselink, A general algorithm for computing distance transforms in linear time, in *Mathematical Morphology and Its Applications to Image and Signal Processing*, edited by J. Goutsias, L. Vincent, and D. S. Bloomberg (Springer, 2002), pp. 331–340.

[90] W. B. Lindquist, A. Venkatarangan, J. Dunsmuir, and T.-f. Wong, Pore and throat size distributions measured from synchrotron x-ray tomographic images of Fontainebleau sandstones, J. Geophys. Res.: Solid Earth **105**, 21509 (2000).

[91] H. Shin, W. B. Lindquist, D. L. Sahagian, and S.-R. Song, Analysis of the vesicular structure of basalts, Comput. Geosci. **31**, 473 (2005).

[92] J.-W. Kim, D. Kim, and W. Lindquist, A re-examination of throats, Water Resour. Res. **49**, 7615 (2013).

[93] K.-T. Jun, Throat finding algorithms based on throat types, arXiv:1412.1401.

[94] S. Beucher and F. Meyer, The morphological approach to segmentation: the watershed transformation, in *Mathematical Morphology in Image Processing*, edited by E. Dougherty and B. J. Thompson (CRC Press, Boca Raton, FL, 2018), pp. 433–481.

[95] R. Barnes, C. Lehman, and D. Mulla, Priority-flood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models, Comput. Geosci. **62**, 117 (2014).

[96] V. Postnicov, Library for direct computation and annealing update of euler characteristic from 3D binary images, https://github.com/fatimp/EulerCharacteristic.jl (2022).

[97] D. Legland, I. Arganda-Carreras, and P. Andrey, Morpholibj: Integrated library and plugins for mathematical morphology with ImageJ, Bioinformatics **32**, 3532 (2016).

[98] W. Nagel, J. Ohser, and K. Pischang, An integral-geometric approach for the Euler-Poincaré characteristic of spatial images, J. Microsc. **198**, 54 (2000).