Multiple Cubes Growth Algorithms for Simple Representative Elementary Volume Determination on 3D Binary Images

R. I. Kadyrov¹

Kazan Federal University (Institute of Geology and Petroleum Technologies)

¹ ORCID: 0000-0002-7566-6312, rail7777@gmail.com

Abstract
Porosity analysis is fundamental for understanding various material properties and transport phenomena in scientific and engineering disciplines. This study delves into the challenging task of determining the representative elementary volume (REV) in porous media, crucial for accurate analyses. Two novel algorithms, Center-Corner Cubes Growing (3CG) and Random Cubes Growing (RCG), were proposed and tested on synthetic body-centered cubic (BCC) sphere packing and natural porous structures of Berea sandstone and Indiana limestone, obtained using µCT. First algorithm (3CG) operates by analyzing porosity within cubes growing from each of the eight corners and a central region of a 3D binarized stack. In contrast, the Random Cube Growing (RCG) algorithm randomly selects seed points within the 3D stack and grows cubic regions around them. Both algorithms systematically compute porosity for various cube sizes, determining the average porosity and standard deviation for each extent. These visual analytics tools contribute to identifying the specific size ranges where porosity curves converge and stabilize, indicating potential REV within the material. While 3CG simplifies the approach by focusing on a limited number of curves, RCG provides a broader view, capturing diverse porosity patterns. The absence of consistent local minima in certain cases indicates high porosity heterogeneity and the impossibility of achieving REV in certain sample sizes.

Keywords: Representative Elementary Volume (REV), Porosity, Center-Corner Cubes Growing (3CG), Random Cubes Growing (RCG), Plot, Visualization.

1. Introduction
Porosity analysis has a great importance among scientific and engineering fields due to its capacity to explain fundamental material properties and transport phenomena [1]. The quantification of porosity, reflecting void spaces within a given volume, underlies the characterization of permeability, mechanical strength, and thermal conductivity. In geology, porosity measurements are used for hydrocarbon reservoir assessments and groundwater flow predictions [2]. Similarly, in civil engineering, porosity insights influence the strength and stability of concrete structures [3,4]. In metallurgy, porosity evaluation is crucial for evaluating the structural integrity of cast metals [5]. In the realm of biomaterials, porosity analysis informs the design and performance estimation of implants and scaffolds, directly influencing their biocompatibility and mechanical behavior [6]. Furthermore, environmental sciences benefit from porosity quantification in understanding soil water retention and contaminant transport within subsurface environments [7,8].

Three-dimensional (3D) imaging techniques have become key tools for capturing complex porous structures, significantly enhancing the depth of insight into porous media characterization. Natural and engineered porous materials, often characterized by irregular geometries and interconnected voids, need imaging techniques capable of accurately representing their complex structures. 3D imaging techniques address this challenge by providing spatially re-
solved data, enabling the representation of porosity distribution, pore connectivity, pore shapes and pore-throat sizes. Methods such as X-ray computed tomography (CT) [9,10], focused ion beam scanning electron microscopy (FIB-SEM) [11], optical coherence tomography (OCT) [12] and some other techniques, have revolutionized the science by enabling the direct visualization of intricate porous architectures.

One of the cornerstone concepts in the analysis of porous media and its characteristics is the notion of the representative elementary volume (REV), a fundamental unit that underpins the investigation of heterogeneous structures. The concept of the REV can be succinctly described as the minimal sample volume within a heterogeneous material wherein the statistical properties stabilize to catch representative and averaged characteristics of the whole material [13]. Consequently, analysis conducted on the REV yields insights that approximate the material's behavior on a larger scale. The REV concept bridges the gap between microscopic pore-scale interactions and macroscopic behavior. The determination of the REV holds profound significance, as it directly influences the choice of spatial resolution for imaging and numerical simulations.

Various studies have explored the concept and dimensions of the Representative Elementary Volume (REV) at the microscopic level in different types of porous materials [14–16]. Zhang et al. [17] introduced a statistical approach to study REV and obtained it for crushed glass beads and Brent sandstone. Al-Raoush and Papadopoulos [18] highlighted the inadequacy of REV based on porosity for natural sand systems in terms of particle size distribution, local void ratio, and the coordination number. Peyman et al. [19] demonstrated that permeability’s REV is larger than that for static properties like porosity and specific surface area because it accounts for the tortuosity and connectedness of the flow paths. They also noted that for carbonate samples, the REV appeared to be larger than the image size.

The determination of the appropriate size for REV poses a fundamental challenge in the analysis of porous media. This challenge stems from the intrinsic heterogeneity of these materials, where properties such as porosity and permeability exhibit spatial variations across multiple scales [17,20,21]. The REV size directly impacts the level of detail captured in analyses, and selecting an inappropriate value may result in misleading or inaccurate representations of the porous medium’s behavior. A REV that is too large may oversimplify the material, neglecting crucial heterogeneity, while a REV that is too small may overly complicate the analysis, introducing noise from local fluctuations [22]. Consequently, finding the balance between these extremes is paramount to achieving meaningful results. Additionally, the REV size is influenced by factors such as material type, imaging resolution, and the specific property under investigation. This multifaceted nature underscores the challenge of determining a universally applicable REV size, emphasizing the need for a case-by-case assessment. As noted by Gerke and Karasina [23], the plateau in any physical property is a necessary, but insufficient, condition for the REV, which requires pore structure stationarity as an additional criterion. The impact of selecting an incorrect REV size can be profound, potentially leading to erroneous conclusions in various fields.

Nevertheless, given the inherent complexity of porous media, the REV’s search is a quintessential step toward achieving meaningful and accurate analyses that transcend the intricate pore-level dynamics, offering instead a representative and tractable framework for the study of porous materials. The most common approach for measuring the REV of porosity involves the growth of a central cube [10,23–25]. This method entails selecting a cubic region of interest (ROI) within the material or rock sample, constructing a cube in the central area that expands with a specific increment toward the ROI’s boundaries, and measuring the porosity coefficient within the cube at each increment. Subsequently, a plot is generated illustrating the change in the porosity coefficient as a function of the linear size of the cube (Fig. 1). The REV is determined as the point on the curve where fluctuations diminish, and the curve levels off, reaching a plateau. However, this approach has a significant drawback: it is often challenging to ascertain whether the curve has truly leveled off for a specific minimum cube size or if the leveling is a local phenomenon influenced by sampling randomness. In our
study, we propose two alternative algorithms to address this issue: Corner-Center Cubes Growing (3CG) and Random Cubes Growing (RCG). Both algorithms are visual analytic tools designed to assist in the REV determination on the plot.

Fig. 1. Illustration of the concept of the Representative Elementary Volume (REV) for porosity: a cubic ROI starts to grow from the center of a 3D image with segmented pore space; after each growing step, the porosity value is measured inside the ROI volume (A, B and C cubes) and marked on the plot (A, B and C points); the region where the line ceases to fluctuate significantly and levels off is conventionally attributed to the REV

2. Materials and methods

2.1 Binary image stack preparation

The algorithms proposed in this study can operate with any voxel-based models derived from various methods of 3D analysis of internal structures (such as µCT, FIB-SEM, etc.). However, these volumes should be binarized and represented as a stack of images in formats .tif or .png. The segmented structure of the pore space should have a value of 1 (white), while the matrix should have a value of 0 (dark). Furthermore, to ensure the proper functioning of the algorithms, it is essential to have information about voxel size (mm).

For the purpose of algorithm testing, we employed three distinct 3D images. The first binary stack comprises a synthetic volume exhibiting a body-centered cubic (BCC) arrangement of spherical structures (see Fig. 2A). The volume dimensions are 100×100×100 voxels, with each voxel being simplistically set at 1 mm in size. The diameter of each sphere is 28 mm. The second stack represents a segmented pore space of Berea sandstone with dimensions of 400×400×400 voxels. It was managed using µCT with a resolution of 3.51 µm (see Fig. 2B) [26,27]. The third stack shares similar dimensions of 400×400×400 voxels and depicts the segmented pore space of Indiana limestone, also obtained through µCT with a resolution of 7.8 µm (see Fig. 2C). Microtomography for both core samples was conducted using the General Electric v|tome|x S240 micro- and nanofocus X-ray research system for computed tomography. The following analytical parameters were set for the scanning: a current of 110 µA, voltage of 110 kV, a number of projections of 1200, an average of 3, and a timing of 200 ms for each projection. The reconstruction of the 3D model was achieved using phoenix datos|x software.
2.2 Center-Corner Cubes Growing (3CG) algorithm

The first algorithm designed for the analysis of porous structures within three-dimensional (3D) image stacks is Center-Corner Cubes Growing (3CG). It begins by importing essential libraries, including NumPy for numerical operations, Matplotlib for data visualization, and scikit-image (skimage) for image processing. The code is structured to handle both TIFF (*.tif) and PNG (*.png) formatted images, enhancing its versatility. The script compiles a list of file paths for all images located in the specified directory, facilitating batch processing of multiple datasets. The input parameters also include the voxel size, representing the spatial resolution of the 3D image and `zoom_from` and `zoom_to` values, which control the limits of cube size ranges for analysis (e.g., from 10 to 50 voxels).

Within the 3CG algorithm, a function for porosity calculation takes center stage. It operates with two critical parameters: the 3D image stack and the voxel size. The essence of the function lies in an iterative process where the porosity within cubes growing from each of the eight corners of the 3D stack, as well as from a central region, is systematically analyzed (Fig. 3). For each cube size, ranging from 2 voxels to the size of the image stack, the function computes the cube's volume, counts the voxels containing pores, and calculates porosity as the ratio of pore volume to cube volume.
Fig. 3. The principle of the 3CG algorithm: cubes begin to grow from the corners (blue) and center (green) of the 3D image stack until it is completely filled, and at each iteration step, the porosity coefficient is measured inside them.

The algorithm is adapted to work with both stacks having an even number of elementary cells (e.g., 100×100×100) and stacks with odd dimensions (e.g., 101×101×101). The script generates an informative graph that provides a visual representation of porosity trends. This graph includes curves for cubes growing from the corners and the center, as well as a curve representing the average porosity for all curves. Additionally, the algorithm incorporates an option to calculate the minimum of standard deviations for the porosity curves, which can be useful for the estimation of the Representative Elementary Volume (REV). The region where the porosity curves converge can be interpreted as the REV size, and the point of intersection of the minimum standard deviation for this region with the average porosity curve can be proposed as the REV porosity value. The python code of 3CG algorithm is available on the web-source [28].

2.3 Random Cubes Growing (RCG) algorithm

The second algorithm is designed to analyze a porous structure and calculate the porosity of randomly growing cubic regions within the image stack. The code begins by importing the necessary libraries, including NumPy for numerical operations, Matplotlib for plotting, and other modules for image processing and file handling. It also specifies input parameters such as the path to the image stack, the voxel size (in mm), the number of random seed points inside the image stack for analysis and the limits of cube size ranges for analysis (controlled by `zoom_from` and `zoom_to` inputs).

The code then reads the image files, stacks them into a 3D array, and determines the dimensions of the image stack (x, y and z). It generates random seed points within the image volume, which serve as starting points for analyzing cubic regions. For each seed point, it iterates through iteratively growing cube sizes within a specified zoom range. The RCG algorithm sequentially extracts the cubic region of interest (ROI) from the image stack, calculates its porosity using the porosity calculation function, and stores the results in an array. During execution, the algorithm checks that the seed points do not extend beyond the boundaries of the image stack. If the seed point is outside the boundary in any dimension (x, y, z), it is adjusted to remain inside the boundary. Similarly, the algorithm checks that the endpoint coordinates of the growing cubes are within the boundary of the image stack (Fig. 4). If the endpoints are not inside the image volume, the cube growth stops.

The algorithm computes the mean porosity and standard deviation for each cube size across all seed points. The REV is identified as the cube size with the lowest standard deviation in porosity, indicating a statistically representative volume. The REV porosity is then calculated at this cube size. Finally, the code plots the porosity values for each seed point and the mean porosity as a function of cube size. As option, it also can mark the minimal standard
deviation size for porosity curves and the corresponding proposed REV porosity on the plot, providing a visual representation of the analysis results. The python code of RCG algorithm is available on the web-source [29].

![Fig. 4](image.png)

**Fig. 4.** The principle of the RCG algorithm: cubes begin to grow from the random seed points (black) inside the 3D image and stop growing when they become outside of it. Here, A is still growing inside the 3D image in all 6 directions, while B and C faced the borders of the 3D image and stopped their increasing.

### 3. Results and discussion

The 3CG and RCG algorithms were tested on the binarized stacks of BCC sphere packing, porous structures of Berea sandstone and Indiana limestone. The application of the 3CG algorithm to the synthetic volume of BCC sphere packing results in the depiction of two intersecting curves on the graph at five distinct points (Fig. 5A). One curve corresponds to the cube growing from the center of the image stack. The other curve comprises all eight porosity curves for cubes growing from the corners of the 3D image, which, due to the symmetry of the synthetic BCC structure, merge into a single curve. Consequently, the curve of average porosity gravitates towards this line. Notably, the intersection points differ: there are three points where the curves initially converge and then intersect, and two points where the lines cross each other in a cross shape. These three points, approximately at 33 mm, 66 mm, and 99 mm, represent the representative volumes. By limiting the upper cube size to approach the region of the first convergence and intersection of the curves, the embedded function for determining the minimum standard deviation indicates a size for this representative volume of 36.00 mm. At the intersection point with the mean porosity line, the porosity value for this volume is 66.42% (Fig. 5B). The computational time of the 3CG algorithm to BCC sphere packing stack was 0.64 s, memory consumption reached 191.63 MB.
Fig. 5. Porosity variations versus cube sizes in 3CG and RCG algorithms for BCC Sphere Packing: A – 3CG algorithm for full stack; B – zoomed plot for 3CG algorithm by limiting the cube sizes up to 50 voxels; C – RCG algorithm with 25 seeds for full stack; D – zoomed plot for RCG algorithm with 25 seeds by limiting the cube sizes up to 50 voxels

The results obtained from the RCG algorithm with 25 random seed points exhibit a similar pattern: we observe three points where the curves intersect and two regions where the bundles of curves narrow down but do not intersect (Fig. 5C). These three regions correspond to representative volumes approximately at 33 mm, 66 mm, and 99 mm. Approaching the first intersection region and determining the minimum standard deviation provides us with a point for the representative volume with dimensions of 33 mm and a porosity value of 68.12% (Fig. 5D). The computational time of the RCG algorithm with 25 seed points achieved 1.36 s, memory consumption was 116.66 MB.

Thus, both algorithms demonstrated relatively similar results: the minimum representative volume has dimensions of 36 mm and 33 mm, with porosity values of 66.42% and 68.12%, respectively. However, due to the fact that 3CG relied on only two curves in its calculations, its results proved to be less precise. The determined representative volume should encompass a sphere and an additional region around it with eight 1/8 sphere parts and have a cell size of 33 mm (see Fig. 6A). Additionally, it is well established that the BCC structure should have a packing fraction of 68.02% [30], which aligns closely with the obtained porosity values, especially from the RCG algorithm. Nevertheless, it is worth noting that the obtained representative volume does not constitute the theoretically minimal possible representative volume (i.e., REV) for porosity. There exists a smaller elementary volume with a size of about 17 mm (see Fig. 6B) that would have the same porosity and repeat in the BCC structure. However, such a volume cannot be directly detected by the presented algorithms based on cube growth.
The representative volume (33 mm) for the BCC structure determined by the cube growth algorithms (A) and theoretical REV (about 17 mm) for BCC structure (B).

The application of the 3CG algorithm to Berea sandstone reveals a sharp convergence of 9 curves (for 8 corner cubes and 1 central cube) within the range of 0.4–0.5 mm (Fig. 7A). Beyond this range, the curves continue to converge, albeit significantly slower. Approaching this specific region allows the identification of a local minimum in standard deviation at the value of 0.5 mm, where the average porosity is 21.41% (Fig. 7B). The computational time of the 3CG algorithm to Berea sandstone stack was 43.65 s, memory consumption was 194.09 MB. The RCG algorithm with 25 seed points shows similar outcomes: all curves converge notably in the 0.4–0.5 mm range (Fig. 7C), and approaching this region also reveals a local minimum in standard deviation at 0.47 mm with an average porosity of 20.95% (Fig. 7D). Consequently, these identified points can be interpreted as the size of the REV. Additionally, it is noteworthy that the average porosity line after the REV points for both algorithms generally exhibits a horizontal direction with slight fluctuations in values. The computational time of the RCG algorithm with 25 seed points achieved 230.04 s, memory consumption reached 134.91 MB.

Utilizing the 3CG algorithm for the Indiana limestone stack demonstrates a prolonged convergence of 9 curves towards the 2.0–2.5 mm range (Fig. 8A). However, approaching this specific region does not enable the identification of a local minimum in standard deviation until its end (Fig. 8B). The computational time of the 3CG algorithm to Indiana limestone stack was 43.13 s, memory consumption reached 256.67 MB. The RCG algorithm with 25 seed points shows similar outcomes: all curves notably converge towards the 2.5 mm range.
(Fig. 8C), yet approaching this region also fails to pinpoint a local minimum in standard deviation within this range until the end of the curve (Fig. 8D). Additionally, the average porosity line continues to decrease slowly until the very end. Thus, it can be inferred that there is a high heterogeneity in the porosity structure and that there is no REV within the size of the stack. The computational time of the RCG algorithm with 25 seed points achieved 257.93 s, memory consumption was 442.67 MB.

The 3CG and RCG algorithms demonstrated the ability to extract much more information about porous structures than the standard approach with only a center-growing cube. However, they come with their own distinct advantages and limitations. The 3CG algorithm, by relying on a limited number of curves, provides a straightforward approach to understanding porosity patterns. Its simplicity allows for quick initial assessments and can offer valuable insights into relatively homogenous materials. However, this simplicity becomes a limitation when dealing with highly heterogeneous porous structures, as it might not capture the full complexity of the material. Moreover, in artificial materials that have spatial symmetry, the porosity curve for corner cubes can merge into one curve, because of which the accuracy of the analysis decreases.

On the other hand, the RCG algorithm, utilizing a larger set of seed points, offers a more comprehensive view of porosity variations. This broader scope allows it to capture a wider range of porosity patterns, making it better suited for materials with significant heterogeneity. Its ability to converge multiple curves to specific points indicates a level of stability in porosity characteristics within certain size ranges. This makes it valuable for identifying REV in materials with complex porous structures. However, this broader approach comes with computational challenges. The increased complexity and the larger number of calculations required can make the RCG algorithm computationally intensive and time-consuming, especially when dealing with extensive datasets or high-resolution images.

The algorithmic outcomes, particularly the convergence patterns and the identification of local minima in standard deviation, provide valuable insights into the porosity analysis of natural materials. When the curves representing porosity converge and exhibit local minima, it indicates specific size ranges within the material where the porosity structure stabilizes or exhibits consistent behavior. These points, often referred to as the REV, signify volumes where the material's porosity characteristics are reliably captured for analysis.

However, the absence of clear and consistent local minima in the algorithms' outcomes suggests a high degree of porosity heterogeneity within the studied materials. In practical terms, this heterogeneity means that there is no specific volume within the material where
the porosity behaves uniformly or predictably. As a result, the porosity analysis might be challenging and may require more complex modeling or advanced algorithms to accurately capture and interpret the porous structure of these materials. These outcomes underscore the importance of understanding the heterogeneity of porous media and adapting analysis techniques accordingly to its limitations.

4. Conclusion

In conclusion, the study presents algorithms, 3CG and RCG, designed to tackle the intricate task of determining the Representative Elementary Volume (REV) in porous media. These algorithms, although slightly diverse in their approaches, offer valuable insights into the stabilization of porosity patterns within specific size ranges. While the algorithms perform computations automatically, the final decision regarding the REV may involve human interpretation of the visualized data on the plot. The outcomes highlight the inherent complexity of porous materials, emphasizing the need for tailored methodologies to capture their diverse characteristics accurately. While 3CG provides a fast straightforward analysis, RCG offers a comprehensive view, particularly beneficial for materials with significant heterogeneity. The absence of consistent local minima in certain cases underscores the challenges posed by REV search in highly heterogeneous structures. These findings contribute to the evolving field of porous media characterization, offering crucial guidance for meaningful scientific interpretations and engineering applications.

Acknowledgements

This work was funded by the subsidy allocated to Kazan Federal University for the state assignment in the sphere of scientific activities, project № FZSM-2023-0014.

References