Modeling of bulk kerogen porosity: Methods for control and characterization

Manolis Vasileiadis¹, Loukas D. Peristeras¹,*, Konstantinos D. Papavasileiou¹, Ioannis G. Economou¹,2,*

¹Molecular Thermodynamics and Modeling of Materials Laboratory, Institute of Nanoscience and Nanotechnology, National Center for Scientific Research “Demokritos”, GR-15310 Aghia Paraskevi Attikis, Greece

²Chemical Engineering Program, Texas A&M University at Qatar, P.O. Box 23874, Doha, Qatar

Abstract

Shale gas is an unconventional source of energy, which has attracted a lot of attention during the last years. Kerogen is a prime constituent of shale formations and plays a crucial role in shale gas technology. Significant experimental effort in the study of shales and kerogen has produced a broad diversity of experimentally determined structural and thermodynamic properties even for samples of the same well. Moreover, proposed methods reported in the literature for constructing realistic bulk kerogen configurations have not been thoroughly investigated. One of the most important characteristics of kerogens is their porosity due to its direct connection with their transport properties and its potential as discriminating and classifying metric between samples. In this study, Molecular Dynamics (MD) simulations are used to study the porosity of model...
kerogens. The porosity is controlled effectively with systematic variation of the number and the size of dummy LJ particles that are used during the construction of system's configuration. The porosity of each sample is characterized with a newly proposed algorithm for analyzing the free space of amorphous materials. It is found that with moderately sized configurations, it is possible to construct percolated pores of interest in shale gas industry.

1. Introduction

The continuously increasing energy demands require new alternative sources of energy.\(^1\) Shale gas is an unconventional energy source, showing great potential to play an eminent role in the energy market of the immediate future. Oil shales are mineral formations whose primary constituent is clays that host sedimentary organic matter.\(^2\) Kerogen is usually defined as the insoluble in common organic solvents part of the organic matter found in shale formations capable of generating oil or gas. Based on the hydrogen/carbon (H/C) and oxygen/carbon (O/C) ratios of their elemental analysis, kerogens are classified in three types, namely I, II and III. A complete definition of kerogen, its types and a variety of other relevant information can be found in the review of Vandenbroucke.\(^3\) Briefly, each type of kerogen is represented by specific regions in van Krevelen diagrams, where the H/C is plotted against the O/C ratio. The type of kerogen found in each shale formation is a consequence of many contributing factors such as the presence or not of humic precursors, the level of maturation and the environment, just to name a few. For example, type I kerogen is usually obtained in lacustrine environments, while marine environments usually result in type II kerogen.

Although its existence is known for many years, shale gas attracted a lot of attention over the last decades after the successful efforts to commercially exploit the Barnett shale in Texas, USA\(^4\).
The commercial production of Barnett shale was made possible primarily by the development of two new revolutionary technologies, the extended horizontal drilling\(^5\) and the hydraulic fracturing.\(^5,7\) Despite the optimism about shale gas, there are a number of poorly understood open issues that hinder a widespread production. The well productivity usually decays very fast\(^8,9\) requiring re-stimulation for maximum yield, a process which is not always straightforward and effective.\(^4,10,11\) In addition, a shale play may not always be productive with current technologies available, despite the fact that it is known to host gas.\(^12,13\)

The environmental impact of the shale gas is also still under investigation.\(^14\) Although the combustion of the produced shale gas has lower emissions than other fossil fuels,\(^15\) there are numerous issues related to the techniques that are used for its extraction. Only a proportion of the hydraulic fracturing fluid is recovered from the wells and its final destination is not known.\(^8,9\)

Furthermore, there are reports of high CH\(_4\) concentration in water at regions close to shale fields,\(^16\) while commercial shale gas extraction activities have also been related to increased concentrations of NO\(_x\) and volatile organic compounds (VOC)\(^17\) in nearby regions, as well as with the possibility of pollution by radioactive materials.\(^18\) Finally, investigations have also been conducted with regard to the relation of hydraulic fracturing and increased seismic activity around the shale fields.\(^19\)

It is evident that shale gas production spread is inherently linked with addressing at least a part of the open questions associated with its environmental impact. For this reason, there is significant industrial and academic research effort aiming in settling the various controversies about shale gas technology. Composition, density, porosity, total organic carbon (TOC) and transport properties are only a few of the properties examined in several experimental studies of various shale and kerogen samples of different fields. One of the most distinctive characteristics of shale samples
and kerogen, in particular, is their complex porous network. Various imaging, scattering and adsorption techniques\textsuperscript{20,21} have revealed that pores of different sizes, ranging from micro (<20Å) to macro pores, are found in kerogen and clays of shale samples. Adsorption experiments performed to shale samples, dried or not, have showed that in many cases their maximum CH\textsubscript{4} sorption capacity is linearly correlated to their TOC content.\textsuperscript{22-24} This is a strong indication that organic matter dispersed in shale has the most significant contribution to gas adsorption.

Apart from the experimental research on kerogens, a continuously increasing interest in modeling work is also seen in the literature. To this end, various efforts have been made to study kerogen and other porous carbonaceous materials by means of molecular modeling. A review by Palmer and Gubbins\textsuperscript{25} resulted to the categorization of the methods originally been used for the study of disordered nanocarbons (DNCs) and their porous network. Although these materials have significantly different chemistry compared to kerogen, the methods are not limited to DNCs and can potentially be used for the study of kerogen. This categorization discriminated the methods into i) constitutive, ii) mimetic or iii) reconstruction methods. Constitutive methods include studies on slit pores of graphite,\textsuperscript{25} using coarse grained\textsuperscript{26,27} and other approximate models.\textsuperscript{25,28} Mimetic and reconstruction methods utilize many body potentials like reactive force fields such as REBO\textsuperscript{29} and ReaxFF.\textsuperscript{30} Mimetic techniques include methods such as the quenched molecular dynamics (QMD), while reconstruction methods refer to techniques such as the reverse Monte Carlo method (RMC)\textsuperscript{31} and its variations such as the latest advance hybrid reverse Monte Carlo method (HRMC)\textsuperscript{32} and molecular dynamics hybrid reverse Monte Carlo method (MD-HRMC).\textsuperscript{33} RMC based methods usually employ a reactive force field and aim at constructing a realistic model of the material under study by modifying a configuration of the system via a MC scheme, in a way that experimental information (usually the radial distribution function, \(g(r)\))\textsuperscript{25,33,34} is reproduced.
Bousige et al.\textsuperscript{34} used a MD-HRMC method with the REBO\textsuperscript{29} force field in the study of four different kerogen samples. They constructed kerogen models consisting of C, H and O, which they validated against structural experimental data (e.g. the radial distribution function, pore size distribution) and elastic properties (e.g. Young modulus). This approach produced realistic structures despite the fact that a small proportion of unphysical features such as three membered rings or cumulene chains were also present. Similar HRMC based models have also been used for the study of alkane adsorption,\textsuperscript{35} mixture diffusion in kerogen and the validation of a free volume theory.\textsuperscript{36} A major drawback of the reactive force fields is the high computational cost compared to their classical counterparts. In addition, tedious effort is required for their parameterization in order to describe accurately the specific chemistry of the target system.

A fourth approach refers to the representative molecular models method. It involves the construction of a model molecule of kerogen based on available experimental data (such as elemental analysis or scattering data even of a specific sample) and prior knowledge on molecular fragments that likely exist in kerogen. To the best of our knowledge, the first effort in this direction was made by Behar and Vandenbrouke\textsuperscript{37} who presented representative 2D molecular diagrams of the three types of kerogen at three stages of their maturation. Maturation was reflected in the decreasing size of the molecules of each type, whose initial size (less mature) was in the order of 15,000 carbon atoms. This approach was later extended to 3D structures.\textsuperscript{38} The models of Behar and Vandenbrouke have been also used in the study of the volumetric properties of the bulk Green River immature type I kerogen.\textsuperscript{39} Another 2D model for the Green river kerogen was proposed by Siskin et al.,\textsuperscript{40} with its major difference from the previously mentioned models lying in the kerogen representation, which is performed by six molecules. 3D configurations of the Siskin’s model generated by \textit{ab-initio} calculations have also been published.\textsuperscript{41}
Recently, Ungerer et al.\textsuperscript{42} published a number of kerogen molecules of smaller size, representing different kerogen types at different maturation stage constructed in a similar framework based on experimental data\textsuperscript{43} and prior knowledge about molecular fragments. The geometry of these model molecules was subsequently optimized using semi-empirical quantum chemistry methods, before using the molecules for the calculation of different thermodynamic and volumetric properties via different computational methods. In this approach, the porosity can be imposed using LJ dummy particles,\textsuperscript{44} or by including directly the organic mixture to be studied during the bulk kerogen construction stage.\textsuperscript{45} This type of kerogen models are constructed using equilibrium Molecular Dynamics (MD). Various structural, thermodynamic and transport properties, such as adsorption and/or desorption of small molecule and the diffusivity of pure fluids and fluid mixtures were calculated.\textsuperscript{42,44–49}

Between the four aforementioned broad categories of methods i.e. i) constitutive, ii) memetic iii) reconstructive and iv) representative molecular models, we adopted the latter in this study. In view of the broad diversity of experimentally determined physical properties such as TOC, density, porosity, permeability etc. not only for different shale fields of similar type kerogen but also for different samples of the same well,\textsuperscript{21,24,50–52} it seems appropriate to investigate ways to adapt the methodology of Ungerer et al.\textsuperscript{42} and Collell et al.\textsuperscript{44} so that the resulting porosity has characteristics tailored to the specific sample under study. In this respect, we systematically investigate how several factors of the methodology affect the porosity of the resulted models. We have used the published overmature type II model molecule (type IID) of Ungerer et al.\textsuperscript{42} to construct systems of various sizes employing GAFF\textsuperscript{53} and Dreiding\textsuperscript{54} force fields. We observe the effect of system size on the accessible space characteristics of the bulk. Porosity is controlled by choosing the
number and size of LJ dummy particles and the effect of these choices on the free and accessible volume are assessed.

A number of algorithms for analyzing void space are available in the literature. Greenfield and Theodorou\textsuperscript{55} studied the diffusion pathways in atactic polypropylene using a Delaunay tessellation based scheme. By joining adjacent accessible tetrahedra they traced the void clusters in the system, while repeating the same cluster analysis for different probe radii they were able to identify how these clusters are connected. A more "exotic" analysis by Alinchenko et al.\textsuperscript{56} was used to study the morphology of the void space distributed in a hydrated DMPC bilayer, based on the generalized Voronoi S-tessellation. In this case, the specific features of the underlining S-tessellation complicates and limits the applicability of the method. Finally, based on Voronoi tessellation and by reviewing and implementing a set of algorithms for performing void space geometry-based analysis of crystalline porous materials, Willems et al. developed the zeo++ tool.\textsuperscript{57} In the current study, two methodologies are used for the estimation and characterization of the free and accessible volume: a) a grid based as implemented in Poreblazer\textsuperscript{58} software and b) a Voronoi tessellation based method developed in house and implemented as a python script.

2. Methodology for constructing bulk kerogen models

Bulk kerogen configurations are constructed using MD simulations. All the calculations presented were performed using the II-D molecular model proposed by Ungerer et al.\textsuperscript{42} (Figure 1a), thereby work focused on overmature type II kerogen, which is representative of many productive shale fields such as the Duvernay and Barnett.\textsuperscript{42,43} This is an approximately planar molecule of almost disk-like shape, consisting of seven big fragments. Each of the constituent
fragments is planar, constructed primarily from fused benzene rings, although other five and six-member rings containing oxygen, sulfur and nitrogen atoms are present. The molecular formula of II-D is $\text{C}_{175}\text{H}_{102}\text{N}_{4}\text{O}_{9}\text{S}_{2}$ and its molecular weight is 2468.8 g mol$^{-1}$. Although it is well understood that kerogens are in reality a polydisperse multi-component mixture, the single unit approximation is adopted here. The construction of the initial configuration is based on random placement of kerogen molecules in a cubic box at very low density. Then following Ungerer et al., the system is cooled from a high temperature to the standard ambient temperature of 298.15 K. During the cooling procedure, the pressure was kept constant at 1 atm, although in the original scheme a higher pressure (100 bar) was used during the first stages of the procedure.

Calculations were performed using two force fields, namely, the Dreiding force field$^{54}$ with Gasteiger atomic point charges$^{59}$ and the General AMBER force field (GAFF).$^{53}$ Dreiding was used as a simple, widely available and easy to use force field. GAFF is a force field compatible with the models usually employed for modeling molecules hosted by kerogen, e.g. small organic molecules. Testing on a wide array of molecules showed that GAFF produces minimized structures closer to their crystallographic counterparts, (i.e. lower root mean square displacement) compared to Dreiding.$^{53}$ In order to derive GAFF parameters, we used the published conformation of Ungerer et al.$^{42}$ as the initial structure, which was then optimized at the B3LYP/6-31G* level of theory,$^{60}$ by means of the Gaussian 09$^{63}$ suite of programs. Partial atomic charges were obtained from the optimized geometry at the HF/6-31G* level of theory, according to the Mertz-Kollman population analysis scheme.$^{64,65}$ Then, assuming no significant conformational changes from this geometry during the dynamical simulations, partial atomic charges were derived according to the Restrained Electrostatic Potential (RESP) protocol,$^{66}$ utilizing the ANTECHAMBER$^{67}$ module of the
AMBER12\textsuperscript{68} suite of programs. Atom type definitions for the GAFF force field are shown in Figure 1b.

As an inevitable consequence of the nature of the system, i.e. molecules of considerable size and reduced flexibility, it is anticipated that the ergodic hypothesis\textsuperscript{69} is not necessarily valid. In order to address this issue, and to study the effects of a) the system size, b) the force field used and c) the presence of the LJ particles in the resulted porosity, a number of initial configurations were created whose details are presented in Table 1. The Amorphous Builder\textsuperscript{70} software integrated in Scienomics MAPS platform\textsuperscript{71} was used to place a number of type II-D molecules inside the simulation box with density of 0.01 g cm\textsuperscript{-3}.

In order to impose porosity on the constructed models, LJ dummy particles were used in a manner similar to Collell et al.\textsuperscript{44} The numbers of configurations created for each combination of system size and LJ dummy particle sets for both force fields used are detailed in Table 1. The number of LJ dummy particles varied from one to three, with their diameter varying from 15 to 40 Å. In all cases, the LJ particles had an interaction energy parameter\textsuperscript{49} equal to 0.20 kcal mol\textsuperscript{-1} and mass equal to that of Argon (MW=39.948 g mol\textsuperscript{-1}). In this way, significant parameters of the simulation that affect the characteristics of the resulting porosity (e.g. force field, number and size of dummy particle etc.) were studied, allowing ultimately for the desired porosity to be achieved.

Each of the generated initial configurations was then relaxed using a staged cooling protocol by MD simulations at the isothermal isobaric statistical ensemble (NPT).\textsuperscript{69} The MD simulations were performed with the LAMMPS\textsuperscript{72} MD simulation package. Throughout the NPT simulations the pressure was kept constant at 1 atm. For the 15 molecule system, the simulations started at 2000 K with one NPT equilibration run of 1 ns. The system was then cooled using a stepwise procedure of repeated simulation cycles. Each cycle was composed of two stages, a stage where temperature
was reduced linearly from an initial to a final temperature, followed by a small NPT equilibration stage, as shown in Figure 2. In order to mitigate the computational cost, the 50 and 100 molecule simulations started at 1100 K cooled down to 300 K by following the same two stage simulation cycles approach as in the case of 15 molecule systems. However, the cooling stage of each simulation cycle in these 50 and 100 molecule systems was 0.05 ns long with a cooling rate of 2000 K ns\(^{-1}\), while the NPT equilibration stage lasted for 0.8 ns. The same cooling protocols were used for systems with the corresponding size containing LJ dummy particles. In this case, an extra step of NPT equilibration of 2 ns at 298.15 K and 398.15 K after LJ dummy particle removal was applied. For all the MD simulations, the temperature and pressure were kept constant using Nose-Hoover thermostat and barostat respectively.\(^{73,74}\) The integration time step was set to 1 fs for all calculations. The traditional Ewald summation technique\(^{69,75}\) was used for the calculation of electrostatic interactions in all systems. A simple cut-off at 12 Å was applied for all non-bonded interactions.

Due to the elevated temperature and pressure conditions that exist in a shale gas reservoir,\(^\text{76}\) the effect of pressure on the density was investigated by sequential pressurization of the ambient pressure structures obtained at 298.15 K and 398.15 K to 100 and 250 atm. The obtained kerogen structures were analyzed with tools described in the next section.

3. Calculation and characterization of porosity

The porosity was studied using two approaches; a grid based algorithm as implemented in Poreblazer open source software\(^{58}\) and a Voronoi tessellation based methodology developed here that does not rely on a grid. Each one used for the calculation of different porosity characteristics.
Information obtained by the analysis of the Voronoi network includes: (a) the CH$_4$ accessible volume and surface, (b) the limiting and the maximum pore diameters, and (c) the identification and the analysis of the diffusion paths spanning the configuration, while Poreblazer was used only for the calculation of the pore size distribution (PSD). The limiting pore diameter (LPD) or largest free sphere diameter ($D_l$),$^{77}$ is defined from the critical spherical penetrant diameter, above which, diffusion would not be possible due to the lack of percolated pores, while the maximum pore diameter (MPD) or larger included sphere diameter ($D_i$),$^{77}$ is the maximum penetrant diameter that can fit into the structure. All the aforementioned properties were obtained as averages over time and for different initial configuration of similar characteristics (force field, system size, number and size of dummy LJ particles).

Poreblazer was used with the default settings (12.8 Å cut off distance, diameters of 3.314 Å and 2.58 Å for N$_2$ and He respectively and for all other elements values for the σ and ε parameters of the Universal force field, UFF$^{78}$) except for the grid spacing. The cubelet size was set to 0.7 Å for the 100 molecule simulations and 0.5 Å for the other simulations, in order to reduce the required computational cost. More details for the grid based algorithm and functionality of Poreblazer can be found in relevant literature by Sarkisov and Harrison.$^{58}$

The scheme developed for the detailed characterization of the void space in periodic crystalline or amorphous molecular systems combines the Voronoi tessellation with concepts and algorithms from graph theory. Furthermore, the Dodd and Theodorou$^{79}$ algorithm was used for the analytical calculation of the free/accessible volume. The proposed approach was implemented as a python script with parts of the functionality implemented in FORTRAN for efficiency reasons. The calculation of Voronoi polyhedra was performed by means of voro++ software$^{80}$ written in c++.

In the next section, the constituent steps of the new algorithm are described thoroughly.
3.1. Voronoi tessellation based analysis of the free space

The workflow of the scheme developed for the accessible space analysis is pictorially shown on the flowchart of Figure 3. In the first step of our scheme, the radical Voronoi tessellation is performed using Voro++. The corresponding polyhedra (hereafter referred as Voronoi cells or just cells) are calculated for each atom (Figure 4a). This is a weighted tessellation where the separation plane between two spheres depends on their radius i.e. is the locus of points with equal tangent relative to the two spheres. This property, in the case where the system constituent atoms are of different sizes, ensures that the center of each atom lies inside the corresponding Voronoi polyhedron resulting in a more realistic per atom decomposition of the space required in our analysis. Using the obtained information (i.e. vertices, edges and faces defining the Voronoi cells), we then proceed with the analysis of the void space and the calculation of the quantities mentioned in the previous section. In order to define the accessible volume, a spherical probe is assumed and subsequently the radius of all atoms is increased by the radius of the probe. Here it is worth noting that the radius of the kerogen atoms can be defined according to any force field. In this study the diameters of the UFF force field have been used in all cases, for reasons of compatibility with the calculations performed using Poreblazer. The probe radius assumed is 1.865 Å, which is the radius of CH₄ in TraPPE force field. A description of the way that the aforementioned quantities of interest are calculated follows.

(a) Calculation of the accessible surface and volume

The total accessible volume of the system of N atoms, \( V_{tot}^{acc} \), is the sum of the individual cell contributions:
\[ V_{tot}^{acc} = \sum_{i=1}^{N} V_i^{acc} \]  

where, \( V_i^{acc} \) is the accessible volume of the \( i^{th} \) cell. \( V_i^{acc} \) is calculated as the total volume of the Voronoi cell, decreased by the volume of the fraction of the atom that is within the cell. The latter is obtained by evaluating the remaining sphere (atom) volume, intersected by a number of planes using the Dodd and Theodorou algorithm.\(^7\) In this case, the atom defines the sphere, while the faces of the corresponding Voronoi polyhedron define the cutting planes. The same calculation provides also the cell contribution, \( S_i^{acc} \), to the total accessible surface \( S_{tot}^{acc} \):

\[ S_{tot}^{acc} = \sum_{i=1}^{N} S_i^{acc} \]

(b) Identification of accessible vertices and edges

The remaining analysis requires the identification of the vertices and edges of the constructed polyhedra that do not overlap with systems’ atoms (Figure 4b). These lie in the void space of the material and host the probing molecule. For each accessible vertex, we calculate the maximum probe that it may host; for each accessible edge, we calculate the maximum probe possible to traverse it. These quantities will be used for the identification of diffusion paths along percolated pores, and several characteristic pore diameters (e.g. maximum or limiting pore diameters). Therefore the accuracy by which they are going to be calculated will ultimately determine the accuracy of the subsequent calculations.

Each vertex is checked against its nearest atoms, i.e. the atoms that correspond to the Voronoi cells by which it is shared. Subsequently, for every non-overlapping vertex \( j \), the minimum distance, \( d_j \), from the surface of the four nearest atoms is calculated analytically. An edge is considered as overlapping if at least one of its vertices is either overlapped or intersected by a
nearby sphere. The latter condition is verified by solving analytically the line-sphere intersection problem for the atoms (spheres) of the cells where the edge (line) participates in. In our approach, every non-overlapping edge \( (k) \) is characterized by the passing probe diameter \( (d_{k}^{PPD}) \), which is the maximum diameter of the spherical probe that is able to pass from one constituent vertex to the other. If the edge intersects the triangle defined by its three atom participating cell (Figure 5a), the \( d_{k}^{PPD} \) is calculated analytically as the diameter of the sphere tangent to surfaces of the atoms, with its center lying on the triangle’s plane. Note that the center of this sphere is equidistant from the atoms’ surface. If the edge does not intersect the aforementioned triangle (Figure 5b), a number of equally distributed edge points (including the vertices) are selected and the distance of each point from the surface of the closest atom is calculated. In this case, \( d_{k}^{PPD} \) is specified as the minimum of these distances. In this case, the passing pore diameters calculation is not analytical and the accuracy is affected by the number of points selected along the edge. In this study, a large number of ten points along the edges of this type are chosen (this is almost every 0.1 Å on average).

The non-overlapping vertices and edges form a network spanning the void space of the periodic simulation box (Figure 4c). Due to the periodic boundary conditions (PBC) imposed in our models, a non-overlapping vertex \( v_i \) may have one or more periodic images \( v_i' \). At this point, the periodic indices \( i = (i_x, i_y, i_z) \) of each image \( v_i' \) are identified.

(c) Identification of pores

The analysis of the resulted network leads to the identification of the pores existing in the structure. In the context of graph theory, this network can be seen as a bidirectional, disconnected (in the general case) graph \( G \), while the clusters of interconnected vertices are the constituent connected sub-graphs, \( G_i \).
Topologically, the $G_i$ clusters may be connected due to the periodic boundary conditions imposed in our models (Figure 6). In this sense, a pore is a set of connected clusters. Two clusters, $C_1$ and $C_2$, are considered to be connected if at least one vertex $v_i^{C_1}$ of the second is the periodic image of a vertex $v_i^{C_2}$ belonging to the first or *vice versa*. The pores that exist in the system are reconstructed by merging together their constituent connected clusters. Two clusters $(C_1, C_2)$ are merged by translating the vertices of $C_2$ by the minimum vector separating a shared pair of parent-image vertices $v_i^{C_1}, v_i^{C_2}$ (Figure 7). In order to keep track of the relevant cluster positions, the periodic indices of the image vertices that either themselves or their parents belong to $C_2$, are updated according to the translation performed. Finally, the image vertices of the resulting cluster $(C_{12} = C_1 \cap C_2)$ coinciding with their parents (i.e. their periodic indexes after the update are zero) are fused with their parents i.e. they are removed and the connectivity of their parents is updated properly taking care that no duplicated edges should exist.

The free/accessible volume of the located pores, percolated or not, is calculated based on eq. (1) from the contributions of pores' constituent cells. It is worth noticing that if a Voronoi cell is shared between two or more pores, only the part of the cell belonging to the pore under consideration is taken into account.

(d) *Percolation analysis*

The resulted pores are then examined for percolation. If at least one pair of parent-image vertices $v_i, v_i'$ exist in the pore, then it is considered percolated with the periodic indices $(i_x, i_y, i_z)$ of the image $v_i'$ indicating the direction of percolation. The set of unique periodic indices found for all parent-image pairs reveal the directions of the percolation channels that exist in the pore (Figure 6b). The alternative approach for assessing percolation is to check the spatial characteristics of the
pore. In this case, if a dimension of the pore's bounding box is greater than the corresponding
dimension of the simulation box then the system is considered to be percolated along this direction.

(e) Identification of diffusion paths and their limiting pore diameters

To systematically trace the paths traversing a percolated pore it is useful to define the notion of
an island. The image vertices $v_i'$ of the pore are grouped based on their periodic indices, $i =
(i_x, i_y, i_z)$. For each group the aggregations of interconnected vertices are identified. We refer to
such aggregations as islands and the $m^{th}$ island of group $i$ is denoted as $I_i^m$, with $m = 1, 2, ..., N_i^l$
and $N_i^l$ being the number of islands of group $i$ (Figure 8). The parents of the images belonging to
island $I_i^m$ are also interconnected and define the parent island, hereby denoted as $I_i^m$. Each pair of
parent-image islands defines a percolation channel which is characterized by a maximum size of
the probe able to traverse it. Then, all possible pairs of islands $I_i^m, I_i^{m'}$ with $m, m' = 1, 2, ... N_i^l$
are considered and for every pair a path connecting the two islands is found.

According to graph theory, a path is a sequence of connected edges and the length of the path is
equal to the number of its edges. In a percolated pore, at least a path of length $\ell$ connects a pair of
parent-image vertices, $v_i, v_i'$ (or equivalently a pair of parent-image islands $I_i^m, I_i^{m'}$) of a
percolated pore and is denoted as $P_{v_i, v_i'}^\ell$. The shortest percolated path is the one of the smaller
length:

$$P_{v_i, v_i'}^{\min} = \min_{\ell} P_{v_i, v_i'}^\ell$$ (3)

In the context of amorphous materials, we are interested in identifying the path that has the
potential to allow faster and easier diffusion in the material. This is the shortest path composed of
edges of the larger possible passing pore diameter, $d_k^{PPD}$. In order to have a more complete
exploration of the diffusion possibilities in the material, a path of these characteristics is being
found for every pair $I_i^m, I_i^{m'}$ that has earlier been determined. This task is performed using the
Dijkstra algorithm, as implemented in networkX python library. The unbiased Dijkstra algorithm identifies the path of minimum length connecting two vertices. If weights are assigned to the edges, the algorithm traces the path with the minimum sum of weights. Therefore, in order to find the path that meets our criteria for the pair $I^m_i, I'^m_i$, we choose the parent-image pair with the larger distance from atoms’ surface as the two vertices and we assign appropriate weights, $w_k$, to the edges of the pore:

$$w_k = (M - d^{PPD}_k + m) \alpha$$

(4)

where, $\alpha$ is a scale factor used to increase the separation between the weights that correspond to the smaller, $m$, and the larger, $M$, values of $d^{PPD}_k$ values met in the pore:

$$M = \max_k\{d^{PPD}_k: k \in \text{pore edges}\}$$, and

(5)

$$m = \min_k\{d^{PPD}_k: k \in \text{pore edges}\}.$$  

(6)

It should be noted that according to (5) the larger the $d^{PPD}_k$, the smaller the corresponding weight $w_k$. Therefore, the traced paths meet the criteria set and the maximum probe that can traverse the path corresponds to the edge with the largest $w_k$.

Due to the nature of the systems we are interested in, it is possible that the path traced for the pair $I^m_i, I'^m_i$ does not include the edge corresponding to the real maximum passing probe diameter. This is due to the fact that a path of much smaller length may result to a smaller sum of weights even if the weights of the edges are relatively high. To overcome this problem, the edge with the maximum weight, $w_k$, i.e. the minimum passing probe diameter, located from the Dijkstra algorithm is eliminated from the pore’s graph and the path connecting the islands is recalculated. This procedure is repeated until the graph becomes disjoined (i.e. there is no path connecting these vertices). The free particle diameter corresponds to the examined islands pair, $d^{PPD}_{I^m_i, I'^m_i}$ is the
maximum value of the calculated minimum passing pore diameters of all the traced paths. In turn, the maximum value of the free pore diameters of all the identified channels corresponds to the limiting pore diameter $d_{i, pore}^{PPD}$ of a percolated pore.

(f) **Determination of the maximum and limiting pore diameters**

In the final step, the maximum included sphere, $D_I$, that can be hosted in the structure and the limiting pore size or free sphere diameter, $D_f$, are calculated. $D_I$, is obtained straightforwardly as the maximum value of the distances $d_j$ that were determined earlier for each accessible edge:

$$D_I = \max(\{d_j\})$$

(7)

$D_f$ is the critical penetrant diameter above which diffusion would not be possible due to the lack of accessible percolated pores and therefore, by definition, is the maximum value of the calculated limiting pore diameters:

$$D_f = \max\{d_{i, pore}^{PPD}\}$$

(8)

given that percolated pores exist in the system for the probe diameter in use. Thus, when the accessible volume comprises no percolated pores, the assumed probe radius is reduced gradually until it is sufficiently small in order to be able to freely diffuse in the structure. Then, $D_f$, is calculated accurately.

4. **Results**

All the thermodynamic and structural properties presented here are calculated as averages over time and different initial configuration of the same characteristics, i.e. force field, size and number of LJ particle(s). In this way the effect of thermal motion of the kerogen molecules is reflected on the calculated distributions. Since we are interested for systems relevant to shale gas, the Voronoi
tessellation methodology was applied using a probe of radius of 1.865 Å which is the radius of the united atom representation of CH₄ in the TraPPE force field⁸⁴.

The density as a function of temperature for the different configurations created without the aid of LJ dummy particles is shown in Figure 9 for Dreiding and GAFF force fields, respectively. In both cases, the average value of the density decreases with system size reaching a limiting value already for the 50 molecules system. As expected, the statistical uncertainty becomes narrower with increasing systems size. The GAFF force field produces denser structures for every system size and temperature by a factor of 6.9% on average and a maximum of 16%. The system size effect is evident for both force fields. The obtained densities are expected to remain unaffected upon re-heating of the low temperature structure at least up to 700 K, according to the findings of Ungerer et al.⁴². This is partly verified by heating the 15 II-D structures obtained using GAFF at 298.15 K, to 900 K by a protocol identical to the one in Figure 2. The obtained densities (not presented here) by this procedure are identical to the densities of Figure 9b.

The key characteristics of the void space are summarized in Table 2; namely, the accessible surface (S_{acc}) and accessible volume (V_{acc}), the % porosity, the limiting and maximum pore diameters (LPD and MPD, respectively), the percentage of percolated structures and the volume and surface of the larger pore in the structure, for the different structures constructed. The pore size distributions calculated for the systems constructed without the presence of LJ dummy particle(s) at 298 K are shown in Figure 10 for the two force fields considered. For the system of 15 molecules, the PSD is calculated also at 600 K. In all cases, the smallest pore diameter is approximately 2.5 Å. The maximum pore diameter observed using Dreiding force field is roughly 12 Å for all sizes, while for structures constructed with GAFF and 15 or 50 II-D molecules approach the value of 10 Å and jumps to 14 Å for the 100 II-D molecules systems. For both force
fields, temperature is found to have negligible effect on the pore sizes of the constructed kerogen models, a finding suggesting limited mobility of the bulky II-D molecules in a congested environment, even at temperatures as high as 600 K. As the system size increases, larger pores appear in the configurations, while only for the GAFF 100 II-D molecule system this is combined with an increased maximum pore diameter from 12 to 14 Å. This indicates that the PSD is system size dependent and suggest that it is necessary to move to relatively large system sizes consisting of at least several tens of molecules when simulating kerogen with this type of models.

Even though the pore size distribution describes the type of pores of the material, it is not sufficient to provide a complete picture of the pathways available for the diffusion of gases in kerogens. In this respect, the limiting pore diameter or free sphere radius is necessary as a rough estimate for the size of the molecules possible to traverse the percolated paths spanning the structure. Normalized LPD distributions (LPDD) defined equivalently to PSD and calculated using standard binning procedure are shown in Figure 10c and 10d for the various system sizes for both force fields at 298 K and for the 15 II-D system also at 600 K. In all cases the LPD value does not exceed the diameter of TraPPE CH₄ i.e. 3.73 Å which indicates that even the diffusion of small molecules such as CH₄ is difficult to occur especially in a static structure. The effect of system size in the LPDD is in general considered minor.

Considering the diversity of experimental results, a number of kerogen models were constructed with the aid of LJ dummy particles aiming to impose in a controlled manner the major characteristics of the porosity. This effort focused to 50 II-D structures using GAFF since the bulk density reached a limiting value for this system size and GAFF is more accurate in the structural reproduction of small organic molecules. Figure 11 shows the effect of the LJ dummy particles and pressure on the calculated density for these configurations at 298 K. As expected, the structures
generated with the aid of LJ dummy particles are less dense with decreasing density as the size of the particle(s) increase. This behavior is not observed for the structures constructed with one 30 Å and one 40 Å LJ particle where the structure constructed with 40 Å LJ paricle is denser. The majority of the models are denser than 1.1 g cm$^{-3}$ with the smallest observed density of 1.04 g cm$^{-3}$ obtained when two 30 Å LJ particles are used. With the exception of the case of the one LJ particle with diameter 20 Å, the increase of the pressure from 1 to 250 atm results in an increase of the density around 0.1% indicating the small effect of the pressure on the density.

The PSD and LPDD of the 50 molecules GAFF structures constructed with the aid of different number of 15 Å LJ dummy particles are shown in Figure 12, together with the distributions (defined and calculated equivalently to LPDD) of the maximum pore diameter (MPDD) and the % porosity (PD). The % porosity is calculated as the fraction of CH$_4$ accessible volume with respect to the total volume of the system. The use of LJ dummy particle(s) increases the accessible volume in the structure as it is evident from PD (Figure 12a). This extra space is distributed in larger pores, as it is indicated by the broadened PSD (Figure 12b) with maximum pore diameters approaching 20 Å. The presence of bigger pores is also reflected in the MPD distribution (Figure 12d). In general, inclusion of more LJ particles allows the construction of bigger pores in addition to the small pores (~2-5 Å) that remain in the structure. On the other hand, LPD distribution (Figure 12c) is not significantly affected by the increase in the number of LJ particles.

The effect of the variation of the LJ dummy particle diameter on the characteristics of the accessible volume is shown in Figure 13. Porosity can also be increased by increasing the LJ dummy particle size as it is intuitively expected. While this is true in principle, there is an upper limit in the porosity that can be achieved. When increasing the size from 30 Å to 40 Å, no significant increase in the induced porosity is observed (Figure 13a). In every case, no more than 10% porosity
is being obtained. Comparing to increasing the number of LJ particles, variation of their size is more effective in inducing porosity in the sense that equal or bigger porosity can be achieved with a small increase of the size of one particle. Furthermore, increasing the LJ dummy particle diameter from 15 to 30 Å results in larger pores (with bigger change from 15 to 20 Å), as it is indicated by the PSDs of Figure 13b. The broadening trend of the PSD is not observed when the diameter of the LJ particle is increased from 30 to 40 Å. This suggests an upper limit in the pore sizes that can be constructed using one LJ particles with approximate value of 23 Å. The effect of increasing the LJ particle size on pore diameters shown by PSDs is also reflected on MPDD, which is translated towards bigger pores. In agreement with previous findings this trend ends when LJ particle size increases from 30 to 40 Å (Figure 13d). The cease of the LJ particle size effect on pore diameter is a consequence of the size of the II-D molecule, whose diameter is approximately 28 Å in its gas phase conformation and can be even smaller in the bulk. Overall, the LJ particle diameter variation has a more pronounced effect on LPD distribution (Figure 13c) compared to changing the number of smaller size LJ particles. Limiting pore distributions become broader and less sharp, reaching values greater than the diameter of CH₄ (something which couldn’t be achieved with variation of then number of LJ dummy particles) up to 5 Å. It is evident that the increase of LJ dummy particle size is a much more effective in controlling the major characteristics of porosity, being the only way that allows the creation of diffusion channels in the bulk (i.e. limiting pore diameters greater than the diameter of CH₄).

The influence of particles of 20 and 30 Å size on the limiting pore is shown on Figure 14. It is clear that an increased number of larger particles can result in well-shaped, wide percolated pores that allow large size molecule diffusion. A considerable number of pores with limiting diameters higher than the diameter of CH₄ and up to 10 Å are created.
Overall no single model shares all types of micropores that are known to exist in kerogen. Despite this, the various combinations of LJ dummy particle numbers and sizes have made it possible to create structures of different porosity characteristics that are all present in kerogen. These characteristics are closely linked to phenomena that are of interest in modeling systems relevant to shale gas.

Although direct quantitative comparisons between the calculated quantities presented here and experimental results is not possible, all the models that have been generated are in general within or close to experimental range. The majority of the models give densities between 1.1 and 1.3 g cm\(^{-3}\), values very close the commonly observed densities for type II kerogen\(^{52}\). PSDs are in accordance with the experimental pores sizes of micropores, having diameters of a few Å.\(^{20,34}\)

Ultimately, to the best of our knowledge, the only information on limiting pores is related to Barnett shale, where experiments have proven the existence of pore throats of less than 5 Å,\(^4,7\) a finding that corroborates our calculations.

Percolation analysis reveals that CH\(_4\) accessible pores can be percolated only when a LJ dummy particle is used. Other simulation parameters such as the system size or the force field seem to have minor effect on the more important characteristics of the porosity. The percentage of configurations comprising percolated pores is significant only in the cases of 3x20 Å, 2x30 Å and 1x40 Å LJ dummy particle sets (Table 2). When probing the accessible volume with a 2 Å particle diameter, all approaches with both force fields result in a large number of percolated configurations, exhibiting only a single pore in the vast majority of the cases (more than 99%). For this reason, analysis of the diffusion paths is performed with probe diameters equal to 2 Å, so that pore characteristics of all procedures are acquired. Figure 15a shows the % of each GAFF atom type (see Figure 1b) located at the larger percolated pore surface, as a fraction of the total number
of atoms of this type found in the configuration. There is clear tendency of some specific atom
types to lie predominantly at the pore surface instead of the bulk. For example, more than half of
the pyrrolic hydrogens (atom type hn) in each configuration are found at the surface of pores. This
finding may have important implications when studying the diffusion of polar molecules. The total
height of the bars on Figure 15a, is related to the total number of atoms by which the pore surface
is constructed. Therefore, a correlation between the total height of bars for each modeling approach
and the corresponding larger pore surface area (see penultimate column of Table 2) is anticipated.
Such a correlation between pore surface and number of pore surface atoms is observed for all cases
except for the case of one 40 Å dummy particle. The bars for the structures constructed with a 40
Å LJ particle are shorter than the bars of the a 30 Å LJ particle structures despite the fact that pore
surface is higher for the 40 Å LJ particle. This is a consequence of the parallel alignment of the
molecules in the void surface (unlike the perpendicular alignment of the previous cases), as it is
visually verified.

The fraction of the atoms of the pore that are found at limiting diameter regions is shown in
Figure 15b. We define as limiting regions the pore segments which are accessible by probes up to
20 % larger than CH₄ i.e. of 4.476 Å diameter. The reduction of bars in Figure 15a to the bars in
Figure 15b does not follow a specific trend. Finally, Figure 15c shows the fraction of each atom
type in the limiting region, with respect to the total number of atoms in the limiting region. There
is a measurable fraction of pyrrolic groups, indicated by the presence of hn atoms, while the
limiting region environment is dominated by aromatic carbon and hydrogen ca and ha respectively.

5. Conclusions
Dreiding and GAFF force fields were used for the construction of bulk kerogen models using a model kerogen molecule, namely II-D, available in the literature. A staged cooling technique was used for the construction of systems of varying size up to 100 II-D molecules. Different sets of LJ particles were used during the cooling stage for imposing porosity in a systematic and controlled manner. Density was calculated and porosity was analyzed using two algorithms. The force field was found to have a minor effect on the calculated properties, which is consistent with calculated densities reported in the literature for other force fields. Systems size is found to have some effect on densities and the created pores. While decreasing the system size increases the computational efficiency, it seems that it is necessary to have a considerable system size of at least 50 II-D molecules in order to have a representative PSD and LPD distributions. The method of dummy LJ particles is found to be appropriate in the effort to have a representative porous space in the bulk. While in every case the ultralow size pores of below 10 Å are present, including LJ dummy particles allows the modeling of the remaining spectrum of micro-pores (up to 20 Å). Increasing the number of LJ particles has a broadening effect on the PSD, but does not seem very effective in altering the LPD distribution. On the other hand, the increasing size of a single LJ particle seems the appropriate way to affect in a controlled way both PSD and LPD distribution. Certain combinations of size and number of particles allow the construction of pores large enough to study the diffusion of several molecules of considerable size.

Elemental analysis of the pore surface indicates a variation in the preference of certain atom types to occupy the surface of the pore instead of the bulk. Overall, important parameters in the construction method for bulk kerogen models were investigated on mature type II kerogen and effective ways that allow one to construct models of tailored porosity in accordance with the
diverse experimental results were proposed. In the future, the diffusion of small molecules in the kerogen structure will be examined.

AUTHOR INFORMATION

Corresponding Authors

*Emails: i.economou@inn.demokritos.gr, l.peristeras@inn.demokritos.gr.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

All the results reported here were performed under ShaleXenvironmenT, H2020 project, funded by the European Commission Horizon 2020 Framework Programme for Research and Innovation (Grant Agreement no 640979).

ACKNOWLEDGMENT

We are grateful to European Commission for funding our research under Horizon 2020 Programme for Research and Innovation and to Professor Andrey G. Kalinichev and Professor Martin Lisal for fruitful discussions on modeling kerogens. This work was supported by computational time granted from the Greek Research & Technology Network (GRNET) in the National High Performance Computing facility - ARIS - under project ID002043. We are also thankful to the High Performance Computing Center of Texas A&M University at Qatar for generous resource allocation. Scienomics SARL is acknowledged for providing MAPS software
used to generate initial structures of the systems simulated. The code for the Voronoi based
characterization of the accessible space of an amorphous material is available upon request.
Table 1: Details of the number of configurations constructed, categorized based on (a) the size of the system, (b) the size and the number of the LJ particles and (c) the applied force field. The bar (-) represents not attempted calculations.

<table>
<thead>
<tr>
<th># II-D molecules</th>
<th>LJ particle</th>
<th># configurations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>σ (Å)</td>
<td># particles</td>
</tr>
<tr>
<td>15</td>
<td>None</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>50</td>
<td>None</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>1</td>
</tr>
<tr>
<td>100</td>
<td>None</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 2. Summary of the average values of the major characteristics of the constructed bulk kerogen models. Numbers in parentheses are the corresponding standard deviations. The Dreiding force filed is abbreviated with D and the GAFF force field with G.

<table>
<thead>
<tr>
<th># II-D.</th>
<th>LJ particles</th>
<th>$S_{ac}$ (m² g⁻¹)</th>
<th>$V_{ac}$ (10⁻¹⁰ m³ g⁻¹)</th>
<th>% porosity</th>
<th>LPD (Å)</th>
<th>MPD (Å)</th>
<th>% percolated configurations</th>
<th>Larger pore† (GAFF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>None</td>
<td>188 (59)</td>
<td>145 (53)</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td></td>
<td>88 (53)</td>
<td>74 (50)</td>
<td>1.0 (0.6)</td>
<td>0.9 (0.6)</td>
<td>2.3 (0.3)</td>
<td>2.1 (0.3)</td>
<td>7.4 (1.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>257 (10)</td>
<td>212 (20)</td>
<td>2.1 (0.1)</td>
<td>1.5 (0.2)</td>
<td>2.6 (0.3)</td>
<td>2.2 (0.2)</td>
<td>11.7 (0.3)</td>
</tr>
<tr>
<td>15.0</td>
<td>1</td>
<td>339 (59)</td>
<td>312 (86)</td>
<td>3.5 (0.9)</td>
<td>2.8 (0.6)</td>
<td>11.6 (1.2)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>270 (33)</td>
<td>222 (48)</td>
<td>2.7 (0.6)</td>
<td>2.4 (0.5)</td>
<td>11.4 (1.7)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>253 (35)</td>
<td>228 (50)</td>
<td>2.7 (0.6)</td>
<td>2.5 (0.3)</td>
<td>12.7 (0.8)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>20.0</td>
<td>1</td>
<td>289 (46)</td>
<td>328 (96)</td>
<td>3.9 (1.0)</td>
<td>2.5 (0.4)</td>
<td>16.5 (2.4)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>381 (55)</td>
<td>616 (123)</td>
<td>6.9 (1.3)</td>
<td>3.0 (1.0)</td>
<td>18.1 (1.2)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>436 (39)</td>
<td>813 (77)</td>
<td>8.9 (0.7)</td>
<td>4.7 (2.3)</td>
<td>19.6 (2.0)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>30.0</td>
<td>1</td>
<td>379 (48)</td>
<td>564 (740)</td>
<td>6.4 (0.8)</td>
<td>2.8 (0.4)</td>
<td>19.9 (1.2)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>498 (35)</td>
<td>1196 (231)</td>
<td>12.4 (2.1)</td>
<td>4.9 (1.5)</td>
<td>(23.7) (2.5)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>40.0</td>
<td>1</td>
<td>350 (42)</td>
<td>561 (103)</td>
<td>6.4 (1.0)</td>
<td>3.0 (0.8)</td>
<td>19.8 (2.6)</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>100</td>
<td>None</td>
<td>271 (6)</td>
<td>232 (270)</td>
<td>2.0 (0.1)</td>
<td>2.0 (0.5)</td>
<td>2.5 (0.1)</td>
<td>2.2 (0.2)</td>
<td>10.4 (0.4)</td>
</tr>
</tbody>
</table>

†CH₄ based
Figure 1: (a) Top and side (inset upper right) views of the molecular model for the II-D type kerogen. The molecular formula of II-D model is $C_{175}H_{102}N_4O_9S_2$ and its molecular weight is 2468.8 g mol$^{-1}$. It consists of five and six member fused rings assembled in a disk-like shape. Color code: green (C), white (H), blue (N), red (O) and yellow (S). (b) Model view where the GAFF force field atom types are displayed: black (c3), grey (ca), purple (h1), light blue (h2), ochre (h4), white (ha), light orange (hc), lime (hn), blue (nb), red (os), yellow (ss).
Figure 2. Staged cooling protocol for the construction of 15 II-D bulk kerogen models demonstrated for a run using GAFF. The system is cooled by repeated simulation cycles each one composed by two stages: a linear cooling stage followed by a NPT equilibration. As temperature is reduced to 1000 K, each simulation cycle comprises a 0.3 ns cooling stage, with a rate equal to 833 K ns$^{-1}$. As we move from 1000 to 298.15 K, each cooling stage is 0.1 ns long and has a cooling rate of 1000 K ns$^{-1}$. Throughout this procedure, the NPT equilibration stages are 1 ns long. The density evolution along the run is also shown for comparison.
Figure 3: Flowchart of the methodology developed for the estimation and characterization of the accessible space in an amorphous material.
Figure 4: (a) An indicative polyhedron resulted from the Voronoi tessellation of the system. (b) The vertices, $v_i$, and the edges, $e_k$, of each cell that do not overlap with the atom that corresponds to the Voronoi cell are identified. The radius of the atom has been previously increased by the radius of the spherical probe considered. (c) The set of non-overlapping vertices and edges forms a graph/network, which topologically span the simulation box. Their corresponding Voronoi polyhedra are also drawn in blue (scaled down by an affine transformation with respect to their centers). Due to the periodic boundary conditions applied, pairs of parent-image vertices exist in the system (image vertices colored green here).
Figure 5: Each non-overlapping edge, $e_k$, (red line), is assigned with the passing probe diameter i.e. the maximum diameter of the spherical probe which is possible to pass from the one constituent vertex to the other depending on the relative position of edge with the atoms by which it is shared: (a) $e_k$ intersects with the triangle formed by the atoms by which $e_k$ is shared and (b) $e_k$ does not intersect with the triangle created by the atoms whose cells share $e_k$. 
Figure 6: (a) A pore is a set of clusters (sub-graphs) connected through the periodic boundary condition applied in our system. (b) The constituent clusters are merged to reconstruct the corresponding pore. The resulted pore is percolated if image vertices (marked with green color) still exist while their periodic image indicates the direction of the channel.
Figure 7: (a) Two clusters, Cluster1 ($C_1$) and Cluster2 ($C_2$) of accessible vertices in a fully periodic system. The image vertices of parent-image pairs are shown in green. Since the system is fully periodic, cluster $C_1$ is connected (forming a pore) with the periodic image of $C_2$. This can be realized as an action of merging the two clusters (into a pore) after translation of $C_2$ by the vector separating the image vertices of $C_1$ by their parents in $C_2$ (b). After merging the clusters, the periodic indices of the surviving image vertices are properly updated indicate the direction of the percolated channel.
Figure 8: Image vertices (highlighted in green) of the pore are grouped based on the periodic indices identifying existing channels. Then, interconnected vertices of the same channel are aggregated and form islands.
Figure 9. Density of kerogen structure as a function of temperature along the cooling procedure for (a) Dreiding and (b) GAFF force fields.
Figure 10 Average pore size (a and b) and limiting pore diameter distributions (c and d) for structures of different size (15 to 100 II-D molecules) at 298 K constructed using Dreiding (a, c) and GAFF (b, d) force fields. The PSD of the 15 II-D system at 600 K is also presented.
Figure 11. Density as function of number and size of LJ dummy particles (LJP) for runs with 50 II-D molecules using GAFF force field.
Figure 12. Accessible volume characteristics for 50 II-D bulk kerogen systems modeled using GAFF and constructed with the aid of different number of 15 Å LJ dummy particles. (a) PD, (b) PSD, (c) LPDD and (d) MPDD.
Figure 13. Accessible volume characteristics for 50 II-D bulk kerogen systems modeled using GAFF and constructed with the aid of varying size (15 to 40 Å) of a single LJ dummy particle. (a) PD, (b) PSD, (c) LPDD and (d) MPDD.
Figure 14. LPDD for different number of LJ particles of size (a) 20 and (b) 30 Å.
Figure 15. GAFF atom type analysis of the surface of the pores. (a) Percentage of atoms of each type that form the surface with respect to the total number of atoms of this type in the whole structure, (b) Percentage of the surface atoms that are found at the limiting regions of the pore (limiting regions are defined as a radius of $3.73 \cdot 1.2 = 4.476 \text{ Å}$ around the pore). (c) Percentage of the atoms of the limiting regions that are of specific type.
References


