

# Fast Method for Computing Pore Size Distributions of Model Materials

Supriyo Bhattacharya and Keith E. Gubbins\*

Center for High Performance Simulation and Department of Chemical and Biomolecular Engineering,  
North Carolina State University, Raleigh, North Carolina 27695-7905

Received September 29, 2005. In Final Form: June 26, 2006

Recently developed atomistic models of highly disordered nanoporous materials offer hope for a much more realistic description of the pore morphology and topology in such materials; however, a factor limiting their application has been the computationally intensive characterization of the models, particularly determination of the pore size distribution. We report a new technique for fast computation of pore size distributions of model materials from knowledge of the molecular coordinates. The pore size distribution (PSD) is defined as the statistical distribution of the radius of the largest sphere that can be fitted inside a pore at a given point. Using constrained nonlinear optimization, we calculate the maximum radii of test particles at random points inside the pore cavity. The final pore size distribution is then obtained by sampling the test particle radii using Monte Carlo integration. The computation time depends on factors such as the number of atoms, the sampling resolution, and the desired accuracy. However, even for large systems, PSDs with very high accuracy (>99.9%) are obtained in less than 24 h on a 3 GHz Pentium IV processor. The technique is validated by applying it to model structures, whose pore size distributions are already known. We then apply this method to investigate the pore structures of several mesoporous silica models such as SBA-15 and mesostructured cellular foams.

## 1. Introduction

The measurement of nitrogen or argon adsorption isotherms at liquid nitrogen temperature<sup>1</sup> and tomographic image analysis<sup>2</sup> are widely used in characterizing nanoporous materials. Using an idealized geometric model of the pores (e.g. slit or cylindrical shapes), BET and density functional theory analysis of the isotherms provide an estimate of the surface area and pore size distribution (PSD) of the material. A central weakness in this approach is the use of an oversimplified pore model that neglects pore tortuosity, connectivity, defects, wall roughness, and heterogeneity. Such pore models may give a reasonable account of adsorption isotherms but often give qualitatively incorrect heats of adsorption and are particularly poor for predicting diffusion rates. As a result, there have been a number of recent attempts to develop more realistic models of the pore morphology and topology, using mimetic atomistic simulation, reconstruction, or other methods. Once developed, the surface area, porosity, and PSD of these model materials can be determined using probe molecules and statistical and geometric analysis.<sup>3,4</sup> Knowledge of the PSD is particularly important since it determines the size of molecules that can be accommodated, selectivity in separations and reactions, etc. A common difficulty in characterizing these realistic model materials is that the analysis to get the PSD is computationally very demanding, and currently this is a limitation in the application of these models. In this paper we report a method for rapid calculation of the PSD that removes this difficulty.

For a model structure, the pore size at a given point can be defined as the largest sphere that encompasses the given point without overlapping the neighboring wall atoms. A schematic in two dimensions is shown in Figure 1. The dotted circles are the various spheres that can be constructed through point *P* without overlapping the wall atoms. The circle in bold represents the largest one among them. Diameter *D* of this sphere is defined as the pore size at point *P*. Thus, the pore size distribution can be estimated by dividing the model space into a finite grid and computing the largest sphere diameter *D* at each grid point. A cumulative histogram *H*(*D*) is then constructed, where *H*(*D*) represents the probability of finding a point in the model space with a pore size greater than or equal to *D*. The pore size distribution *P*(*D*) is the negative of the differential coefficient of *H*(*D*) with respect to *D*; i.e.,

$$P(D) = -\frac{dH(D)}{dD} \quad (1)$$

The only question concerning the above approach is, given a point *P*, how to find the largest sphere diameter *D* encompassing *P*. A suggested approach is to first create a database of spheres with the biggest possible diameter not overlapping the wall atoms, centered at each grid point. To find the pore size at *P*, one has to search through the database in the neighboring grid points of *P* and select the largest sphere that also overlaps *P*. In computing the cumulative pore size distribution *H*(*D*), the histogram bin width is equal to the distance between two adjacent grid points. The histogram bin width is inversely related to the number of grid points per unit length, also called the grid resolution. Every molecular system is associated with an optimum grid resolution which represents the tradeoff between accuracy and computational speed. Resolutions lower than the optimum lead to decreased detail in the distribution curve, whereas higher resolutions lead to enhanced accuracy along with an increase in computation time. The importance of grid resolution will be discussed in section 2. In computer implementations of the above method,<sup>3,4</sup>

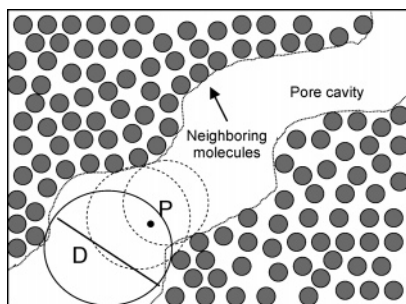
\* To whom correspondence should be addressed. Telephone: 919-515-2262. Fax: 919-515-3465. E-mail: keg@ncsu.edu.

(1) (a) Rouquerol, J.; Rouquerol, F.; Sing, K. S. W. *Adsorption by Powders and Porous Solids: Principles, Methodology and Applications*; Academic Press: New York, 1999 (Feb). (b) Schuth, F.; Sing, K. S. W.; Weitkamp, J. *Handbook of Porous Solids*; Wiley: New York, 2002; Vol. 1.

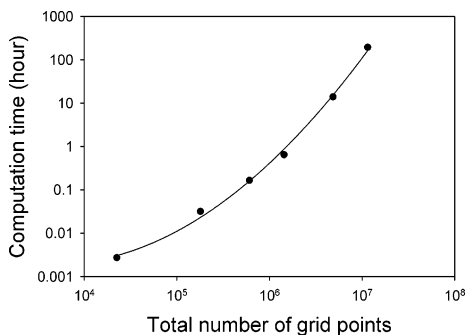
(2) Appoloni, C. R.; Fernandes, C. P.; Innocentini, de Mello, M. D.; Macedo, A. *Mater. Res.* **2004**, *7*, 557.

(3) Gelb, L. D.; Gubbins, K. E. *Langmuir* **1999**, *15*, 305.

(4) Thomson, K. T.; Gubbins, K. E. *Langmuir* **2000**, *16*, 5761.



**Figure 1.** Definition of pore size used in this paper. The middle channel is the pore cavity, and the gray spheres are the wall atoms.  $P$  represents the point at which the pore size is measured.  $D$  is the diameter of the circle in bold.

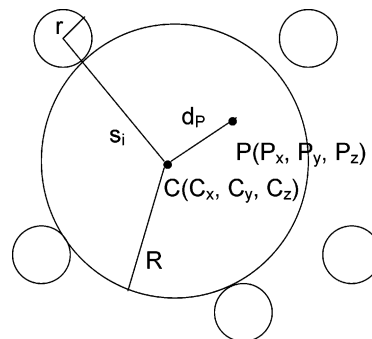


**Figure 2.** Variation of computational time with grid size (on a 3 GHz Pentium IV processor).

both the memory to store the database and the computation time are functions of the grid resolution. As a result, the method works well at low grid resolutions but becomes computationally expensive at higher resolution. Figure 2 shows the variation of computation time with grid resolution for a mesoporous silica model.

In this paper, we propose a technique for fast calculation of the pore size distributions of model materials. To find the largest sphere that can be fitted at a certain point inside the pore, we treat the problem as a constrained nonlinear optimization one and solve it using one of the several nonlinear programming (NLP) algorithms proposed in the literature.<sup>5-7</sup> The sphere diameters are calculated at arbitrary points inside the pore, and the diameters are sampled at regular intervals to generate a distribution. The method has several advantages over the previously discussed method. First of all, there is no need to create a database of sphere diameters over a finite grid, because the method searches through the continuum instead of a grid. Second, fast convergence is achieved through the use of efficient searching schemes instead of probing through each grid point. Combining these two factors, it is possible to achieve a very high degree of accuracy using minimal computing time. Moreover, since the program reports the pore size distribution at regular time intervals, it is possible to observe the evolution of the distribution over time.

The method relies on a constrained optimization approach in solving a geometrical problem. Linear and nonlinear programming models have been successfully used in solving related problems in computational geometry.<sup>8,9</sup> Welzl<sup>8</sup> proposed an algorithm for



**Figure 3.** Largest sphere enclosed by  $n$  atoms:  $r$ , radius of the enclosing atoms;  $R$ , radius of the largest enclosed sphere;  $C$ , center of the enclosed sphere;  $P$ , the point at which the pore size is measured;  $s_i$ , distance from  $C$  to the center of the  $i$ th atom;  $d_p$ , distance between  $C$  and  $P$ .

computing the smallest enclosing disk of a finite set of points in a plane. Gärtner and Schönherr<sup>9</sup> extended Welzl's algorithm to three dimensions for computing ellipsoids and spheres. Our problem is different from the one presented in refs 8 and 9, because instead of the smallest enclosing sphere of a set of points, we are interested in the largest sphere enclosed by a set of points. For the first problem, Welzl employed a linear programming technique combined with an iterative approach. In our case, we have applied nonlinear programming models. The problem presented here can be solved using any suitable NLP scheme, which can handle discontinuous objective functions. We have used the SOLVOPT program written by Kuntsevich and Kappel,<sup>7</sup> because it is well-suited for optimizing nonsmooth objective functions. More information about this program can be found at their website.<sup>7</sup>

## 2. The Problem

Figure 3 depicts the 2-D representation of the problem. Imagine a set of  $n$  spherical atoms in space with radius  $r$ . For simplicity, all the atoms here are assumed to have the same radius, but this treatment can easily be extended to cover atoms with variable radii. Let  $P(P_x, P_y, P_z)$  be the point where the pore size is to be measured. Let  $C(C_x, C_y, C_z)$  be the center of the biggest sphere enclosed by the  $n$  atoms which also overlaps  $P$ . Let  $R$  be the radius of the enclosed sphere,  $s_i$  the distance between  $C$  and the center of the  $i$ th atom, and  $d_p$  the distance between  $C$  and  $P$ . Clearly,  $R$  is a function of  $C$ . Therefore we can write

$$R(C) = \text{MIN}(s_i) - r, \quad i = 1, \dots, n \quad (2)$$

To find the biggest possible sphere radius, we need to find the location of  $C$  for which  $R$  is the maximum. However, there is an additional criterion that  $P$  should be inside the sphere; i.e.,

$$d_p \leq R(C) \quad (3)$$

Therefore, we set up the nonlinear optimization problem as

$$\text{MAXIMIZE } R(C) = \text{MIN}(s_i) - r \text{ subject to constraint: } d_p - R(C) \leq 0 \quad (4)$$

The above problem may be solved using any appropriate NLP routine. In our case, we have used the SOLVOPT<sup>7</sup> program for reasons explained in the previous section.

The basic algorithm is described in Figure 4. We select random points  $P$  and compute the pore sizes at these points. After each sampling interval, we update the cumulative histogram  $H(D)$  and calculate the PSD as the negative of the derivative of  $H(D)$ .

(5) Betts, J. T. *ACM T Math Software* **1980**, *6*, 432.

(6) Press: W. H.; Vetterling, W. T.; Teukolsky, S. A.; Flannery, B. P. *Numerical Recipes in C++*, 2nd ed.; Cambridge University Press: Cambridge, U.K., 2002.

(7) (a) Kappel, F.; Kuntsevich, A. V. *Comput. Optim. Appl.* **2000**, *15*, 193. (b) <http://www.uni-graz.at/imawww/kuntsevich/solvopt/>.

(8) Welzl, E. *Lect. Notes Comput. Sci.* **1991**, *555*, 359.

(9) Gärtner, B.; Schönherr, S. *Inf. Process. Lett.* **1998**, *68*, 33.

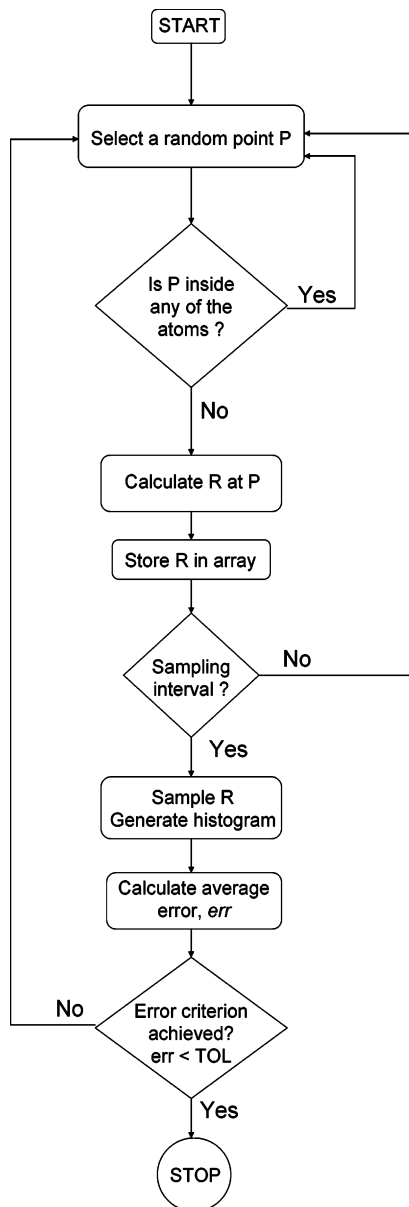


Figure 4. Flowchart for PSD calculation.

The pore size distribution of the accessible volume also depends on the size of the test particle. Reference 4 gives the definition of accessible volume. Pores which are smaller than the diameter of the test particle are not accessible by the test particle. In our calculations, we have used a test particle diameter of 3.409 Å, which corresponds to the Lennard-Jones  $\sigma$  parameter for argon–argon interaction. The average error  $err$  is calculated as the mean of the relative errors between the old and the new probability values. Therefore

$$err = \left\langle \frac{PSD_{new}(D) - PSD_{old}(D)}{PSD_{new}(D)} \right\rangle \quad (5)$$

where  $PSD(D)$  is the probability for the pore size  $D$ . The average error is an estimate of how different a new PSD is compared to the old one. The computation may be terminated when the value of  $err$ , averaged over several Monte Carlo blocks, falls below the specified error limit  $TOL$ . We have used an error limit of 0.01% for all our computations.

The efficiency of this technique depends on the selection of an appropriate grid resolution. The grid resolution should be

large enough not to smooth out the minute details/undulations in the distribution curve. On the other hand, setting the resolution too high will increase the computation time without adding extra information to the distribution. As a rule of thumb, the grid resolution should be smaller than the minimum bandwidth of all the peaks and undulations in the PSD. Initially a very low resolution can be used to get a quick idea about the nature of the distribution curve (e.g., the mean and standard deviation). Then the resolution can be increased in small steps until a small change in grid resolution does not lead to any significant change in the shape of the distribution.

The accuracy of the PSD is highly dependent on the efficiency of the random number generator used in selecting the random points. The values generated by the random number generator must be uniform; i.e., they should have the same probability everywhere within the set. If the generator is biased, i.e., the probability for a certain range of numbers is higher than the rest, the generated PSD will be inaccurate. We have used one of the random number generators from ref 6, which is found to have a uniform probability distribution.

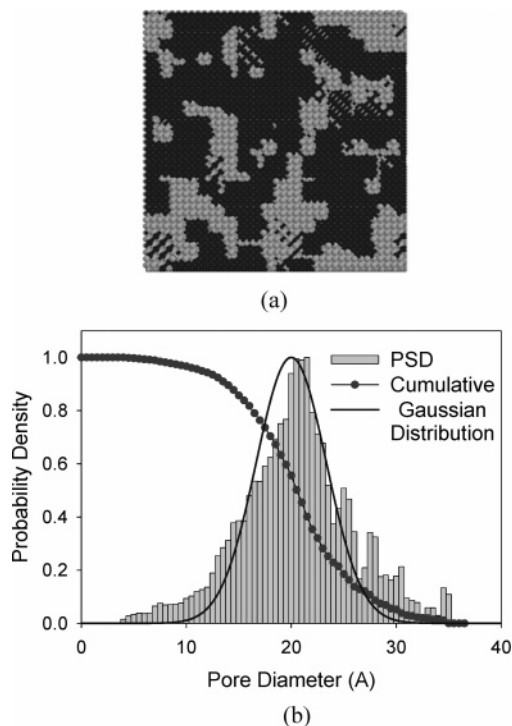
### 3. Results and Discussion

This section is divided into several subsections depending on the topic of interest. First, we show that our method samples the correct pore size distribution by applying in a model, where the PSD is known in advance. Then we discuss the PSDs of several model structures generated by mimetic simulations. Finally, we analyze the performance of our model and compare it with one of the existing methods of calculating pore size distributions. All probability distributions are normalized to unity for clarity. The graphs show both the cumulative histogram  $H$  and the PSD obtained by differentiating  $H$ .

**3.1. Test Model.** Usually, the validity of a numerical method can be tested by applying it to cases where the final result is already known from other sources. To verify our method, we generated a model structure having box dimensions  $120 \times 120 \times 120$  Å, with spherical pores placed at random positions within the box. The diameters of these pores were randomly selected from a Gaussian distribution with a mean of 20 Å and a standard deviation of 3.33 Å. The snapshot of this structure is shown in Figure 5a. We can predict that the pore size distribution of this structure should be very similar to the Gaussian distribution used in generating the pore diameters. Figure 5b shows the cumulative and differential pore size distributions of the model structure as generated by our method. The Gaussian distribution is also shown for comparison. We find that the calculated pore size distribution follows the predicted PSD (Gaussian distribution). However, the calculated distribution shows a bias toward the higher values of pore diameter with a mean at 20.51 Å and a standard deviation of 5.42 Å. The bias arises due to the fact that, while generating the pore cavities, some of the smaller pores were partly or completely engulfed by the larger pores. The reason for the increased standard deviation is that the smaller pores which were partly engulfed by the larger pores appeared as blemishes on the surface of the larger pores contributing to the very small pore diameters. For a system with completely unconnected pores, these differences should be negligible. Additionally, if the system is infinitely large, the calculated PSD should exactly match the Gaussian distribution.

**3.2. Mimetic Models.** In this section, we will discuss the pore size distributions of some of the model structures developed as part of our ongoing project.<sup>10–13</sup> All of these are templated

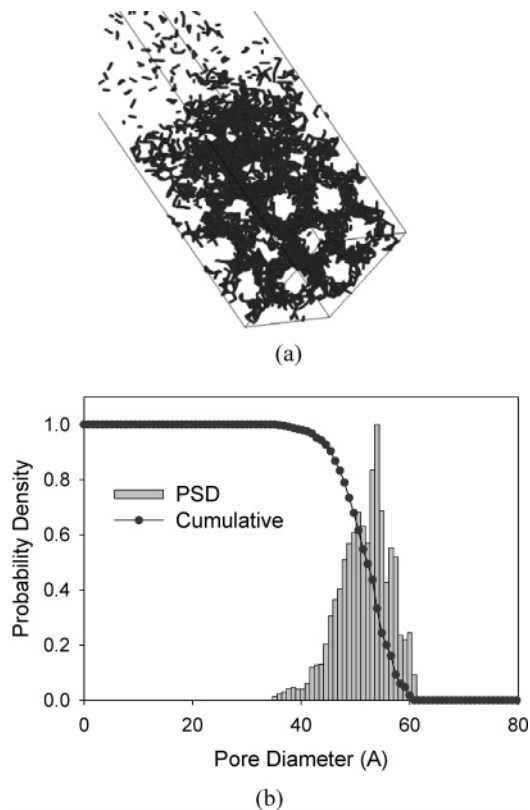
(10) Bhattacharya, S.; Coasne, B.; Hung, F. R.; Gubbins, K. E. *Stud. Surf. Sci. Catal.*; Vol. 160.



**Figure 5.** Test model structure. (a) Cross-sectional view. Dark regions represent the void spaces and the light regions represent the pore walls. (b) Pore size distribution.

mesoporous silica models created by mimicking the synthesis process using lattice Monte Carlo simulations.<sup>14,15</sup> In practice, these materials are synthesized by polymerizing silica around a surfactant–silica mesophase and then extracting the silica structures from the surfactant solution. First, we will describe the PSD of the model SBA-15<sup>16</sup> structure. The pores of the SBA-15 are cylindrical in shape and ordered hexagonally. In practice, the pore size of SBA-15 varies from 50 to 300 Å. Figure 6a shows the simulation box snapshot of the model SBA-15 in the direction of the cylinder axis, and Figure 6b shows the pore size distribution. The pore size of the model SBA-15 varies between 35 and 62 Å with a mean at 51.65 Å. The PSD shows a single peak at 54 Å, which corresponds to the cylindrical mesopore diameter. On either side of this peak, the distribution decays to zero. However, on the left side, toward the smaller pore sizes, the decay is more gradual compared to the right side. The reason is the following. In a lattice model, the individual lattice points are spaced approximately 35 Å apart. To construct the pore wall from such a configuration, we decorated the lattice points with silica spheres having a radius of 17.5 Å. The space where the adjacent spheres touched each other contributed to the smaller pore diameters.

Next consider the pore structure of the mesostructured cellular foams (MCF). Reference 17 gives the details of MCF synthesis. The pore structure of MCF is made up of spherical cells with diameters between 200 and 500 Å, which are connected to each other via windows of smaller diameters (~100 Å). Figure 7a shows a schematic of the MCF structure, and Figure 7b shows



**Figure 6.** SBA-15 structure: (a) simulation box snapshot; (b) pore size distribution.

the simulation box snapshot. Figure 7c shows the pore size distributions for the model MCF structure. Due to the presence of the larger cells and the relatively smaller windows, the pore size distribution of MCF shows two peaks, one at 150 Å (windows) and the other at 450 Å (cell). This is in sharp contrast with the PSD for SBA-15, which consists of only one peak. In the pore size distribution of the model MCF, the windows have a smaller size variation compared to the cells. The MCF cells are produced during synthesis by polymerizing silica around closely packed surfactant micelles, and the points where these spheres touch each other form the windows. While the exact mechanism of window formation is not fully understood, it is clear that the size distribution of these cells is closely related to the cluster size distribution of the surfactant micelles. Following the same logic, the window size distribution is related to the packing behavior of the micelles. A detailed analysis of the cluster size and the packing behavior involves the thermodynamic properties of the simulation system and is beyond the scope of the present discussion.

So far, we have discussed the pore structures of lattice models. We now describe a fully atomistic structure derived from the lattice models. On the basis of previous works by Pellenq and Levitz<sup>18</sup> and Coasne et al.,<sup>19,20</sup> we have developed an atomistic model of SBA-15 from the lattice structure shown in Figure 6. The details of the atomistic modeling can be found in ref 10. Figure 8a shows the schematic of an atomistic SBA-15 pore assembly. The results shown here are averaged over five atomistic SBA-15 pores. The hollow cylinders are the mesopores, and the dark wirings are the micropores connecting the mesopores. Both the micropores, which are less than 10 Å in diameter, and the

(11) Bhattacharya, S.; Gubbins, K. E. *J. Chem. Phys.* **2005**, *123*, 134907.

(12) Siperstein, F. R.; Gubbins, K. E. *Mol. Simul.* **2001**, *27*, 339.

(13) Siperstein, F. R.; Gubbins, K. E. *Langmuir* **2003**, *19*, 2049.

(14) Larson, R. G.; Scriven, L. E.; Davis, H. T. *J. Chem. Phys.* **1985**, *83*, 2411.

(15) Larson, R. G. *J. Phys. II* **1996**, *6*, 1441.

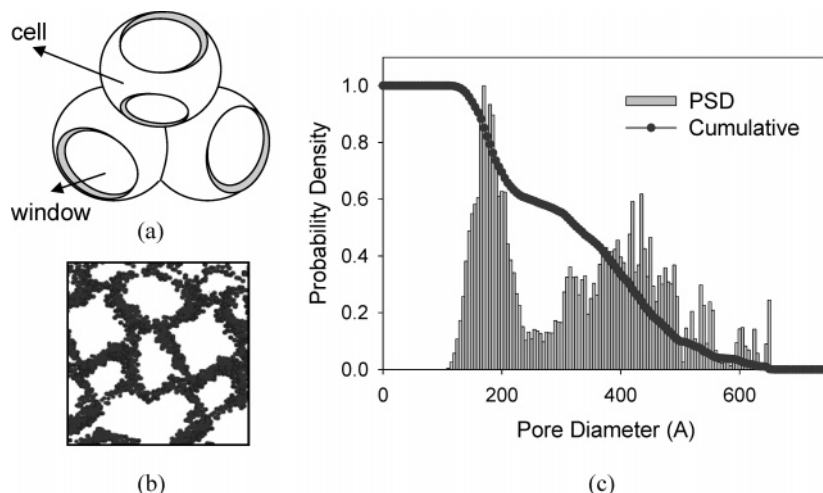
(16) Zhao, D.; Feng, J.; Huo, Q.; Melosh, N.; Fredrickson, G. H.; Chmelka, B. F.; Stucky, G. D. *Science* **1998**, *279*, 548.

(17) Lettow, J. S.; Han, Y. J.; Schmidt-Winkel, P.; Yang, P.; Zhao, D.; Stucky, G. D.; Ying, J. Y. *Langmuir* **2000**, *16*, 8291.

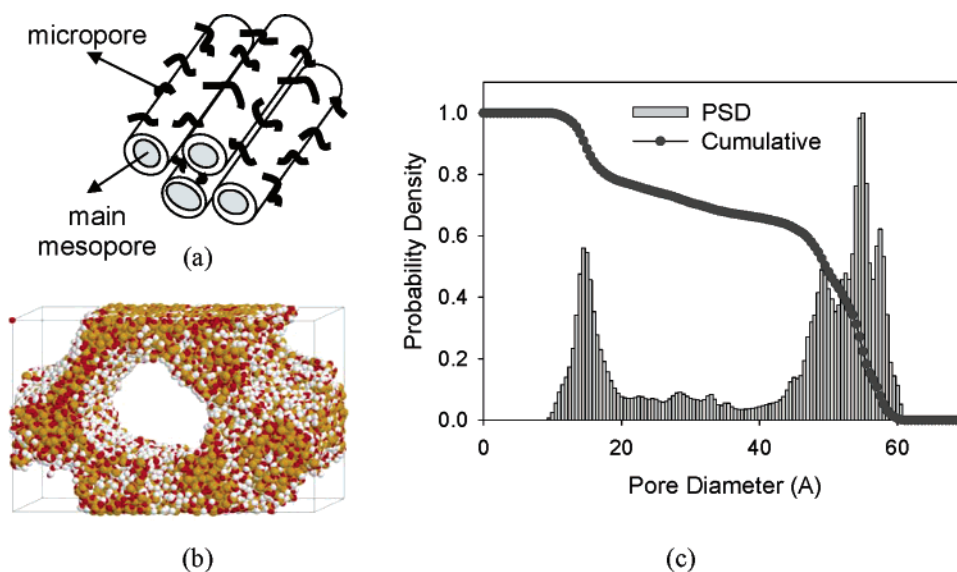
(18) Pellenq, R. J.-M.; Levitz, P. E. *Mol. Phys.* **2002**, *100*, 2059.

(19) Coasne, B.; Pellenq, R. J.-M. *J. Chem. Phys.* **2004**, *121*, 3767.

(20) Coasne, B.; Gubbins, K. E.; Pellenq, R. J.-M. *Part. Part. Syst. Charact.* **2004**, *21*, 149.



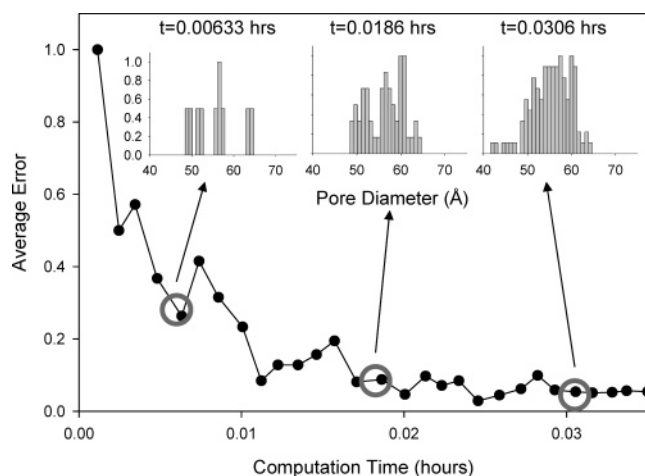
**Figure 7.** MCF structure: (a) schematic of MCF pores; (b) simulation box snapshot; (c) pore size distribution.



**Figure 8.** Atomistic SBA-15 structure: (a) schematic showing the meso- and the micropores; (b) simulation box snapshot (yellow, silicon; red, oxygen; white, hydrogen); (c) pore size distribution.

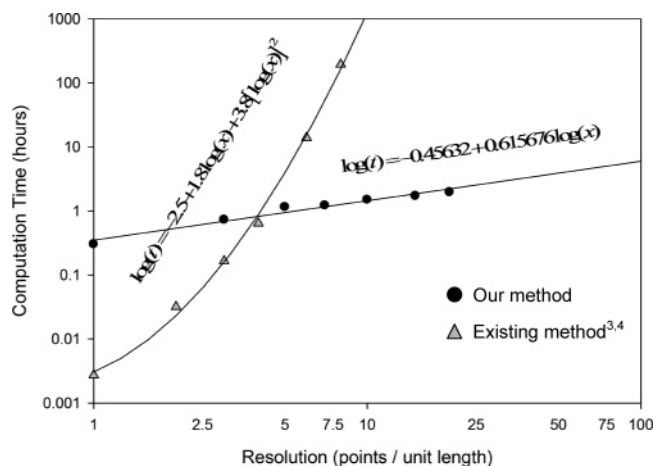
main mesopores constitute the porosity of SBA-15. Figure 8b presents a snapshot of the model SBA-15 structure. The silicon and the oxygen atoms are bonded in an amorphous tetrahedral network with the hydrogen atoms saturating the dangling bonds at the pore surface. In Figure 8b, the locations of the mesopores and the micropores are marked by the white color coded hydrogen atoms. Figure 8c shows the pore size distribution. The PSD consists of one small peak at lower pore diameters (around 15 Å) and a larger peak at 54 Å. The smaller peak at 15 Å represents the micropore diameters, and the larger peak at higher pore sizes represents the mesopores.

**3.3. Performance Analysis.** We will now analyze the performance of the procedure and compare it with one of the existing methods discussed in section 1. All the computations were performed on 3 GHz, Pentium IV processors with 1 GB RAM running RedHat Linux 7.3. The code for calculating the PSDs was compiled using GNU C++. Figure 9 shows the variation of average error with computation time for the lattice model of SBA-15. The definition of average error is given in section 2. The error is initially high but decreases rapidly with time. The rate of decrease of error is also high toward the beginning of the computations but slows down as time progresses. The figure also shows the evolution of the pore size distribution with



**Figure 9.** Average error with computation time. The insets depict the evolution of the PSD over time. Each snapshot corresponds to a different time interval.

computation time. Initially, only a small part of the PSD is sampled, which is shown in the snapshot at  $t = 0.00633$  h. As expected, regions with higher probabilities are sampled first, for example the mesopore peak at  $D = 54$  Å. With time, more points



**Figure 10.** Comparison between our method and the existing method discussed in section 1. Resolution is defined as the number of points per unit length where the pore sizes are calculated.

are sampled and the distribution slowly evolves into its final form.

Figure 10 shows the computation time as a function of grid resolution, both for our method and the existing method discussed in section 1. Resolution is defined as the number of points per unit angstrom, where the pore size distributions are computed. For example, if the resolution is 20, then the PSD is calculated at 0, 0.05, and 0.1 Å, etc. Both axes are plotted in the logarithmic scale. For the existing method, the computation time shows a nonlinear variation with grid resolution. A quadratic regression curve fitted to the data points (shown in Figure 10) indicates that the increase in computation time with grid resolution is initially small at very low resolutions. However, the curve becomes increasingly steeper at higher resolutions, when a small increase in the grid resolution can lead to a very large increase in the computation time. Therefore, it is computationally expensive to calculate PSDs at high resolutions using the existing method. In contrast, for our method, the logarithmic variation of computation time with grid resolution shows a linear trend. The computation time  $t$  (hours) has the functional form  $t = ax^b$ , where  $a$  and  $b$  are constants and  $x$  denotes the resolution (points/unit length). The values of  $a$  and  $b$  can be obtained through regression analysis

and are shown in Figure 10. The value of exponent  $b$  determines the relative increase in computational time with an increase in resolution. For our method, the value of  $b$  is 0.615 676. Comparing the two trends (one for the existing method, the other one for our method), it is observed that, at high resolution, the computation times for our method are orders of magnitude less compared to the existing scheme. However, for very low resolutions ( $<3.96$ ), the existing method is found to be faster compared to the current method. For the current scheme, the computational cost does not increase significantly with grid resolution, as it did for the existing method. Moreover, the method proposed here is ideal for implementing in a parallel architecture, where multiple processors can calculate more number of points in less time and speed up the computation.

#### 4. Conclusion

We have presented a computationally efficient way of computing the pore size distributions of model structures from knowledge of the molecular coordinates. The method uses constrained nonlinear programming techniques for calculating pore sizes at different points within the porous structure. A large number of these points is then sampled to obtain the final pore size distribution. The accuracy of this method has been verified by calculating the pore size distribution of a model structure whose PSD is known in advance. We then demonstrate the practical applicability of the discussed technique by using it to calculate the pore size distributions of several porous models. Finally, performance evaluation shows that the computation time is proportional to  $x^b$ , where  $x$  is the resolution and  $b$  is a constant. Comparing with established techniques, we show that the present method is significantly faster, particularly at high resolutions.

**Acknowledgment.** We would like to thank Jorge Pikunic (grid based PSD algorithm), Benoit Coasne and Francisco Hung (mesoporous silica models) for many helpful discussions. A special note of thanks goes to Franz Kappel and Alexei Kuntsevich for providing the program SOLVOPT. This work was funded by the Department of Energy (DOE) under Grant No. DE-FG02-98ER14847. Supercomputing time was provided by the San Diego Supercomputer Center under Grant NSF/MRAC CHE050047S.

LA052651K