This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

Massively Parallel GPU-Accelerated String Method for Fast and Accurate Prediction of Molecular Diffusivity in Nanoporous Materials

Journal:	*ACS Applied Nano Materials*
Manuscript ID	an-2021-00727e.R2
Manuscript Type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Zhou, Musen; University of California Riverside Bourns College of Engineering, Chemical and Environmental Engineering Wu, Jianzhong; University of California Riverside, Department of Chemical and Environmental Engineering

SCHOLARONE™
Manuscripts

Massively Parallel GPU-Accelerated String Method for Fast and Accurate Prediction of Molecular Diffusivity in Nanoporous Materials

Musen Zhou and Jianzhong Wu*

Department of Chemical and Environmental Engineering, University of California, Riverside, Riverside, CA, 92521, United States

Abstract

The diffusivity of guest molecules in nanoporous materials is instrumental for practical applications ranging from gas separation to catalysis and energy storage. Conventional methods to predict diffusion coefficients are computationally demanding in particular for polyatomic molecules with small diffusivity in nanoporous materials. In this work, we have implemented a massively parallel GPU-accelerated string method to calculate the minimum energy path (MEP) for the diffusion of polyatomic molecules in nanoporous materials. The GPU parallelization enables fast prediction of molecular diffusivity in nanoporous material, which speeds up the computation by a factor of over 500 in comparison with serial CPU calculations. The massively parallel GPU-accelerated string method yields diffusion coefficients in excellent agreement with results from molecular dynamics while reduces the computational cost by several orders of magnitude. It will thus open up opportunities for high-throughput screening and inverse design nanoporous materials.

Keywords: Diffusion, transition-state theory, nanoporous materials, high-throughput screening, gas separation

* To whom correspondence should be addressed. Email: jwu@engr.ucr.edu

1. Introduction

Recent years have seen the rapid development of nanoporous materials with a vast variety of building blocks.[1-6] Nanoporous materials (e.g., metal-organic frameworks) can now be designed and synthesized by assembling organic ligands and metal cluster with appropriate topology. As a result, large materials databases become commonplace promising data-driven applications via high-throughput screening and computational design.[7-9] Transport properties such as diffusion coefficient are closely related to many important applications. For example, diffusivity dictates the performance of nanoporous materials, including zeolites, metal-organic and covalent organic frameworks, for gas separation and ion sieving.[10-13] Efficient computational methods for fast yet accurate prediction of transport properties are always in great demand for searching the best nanoporous materials in a structural database and/or for the inverse design.[10, 14]

For guest molecules in a nanoporous material, the diffusion coefficients can be measured using experimental techniques such as quasi-elastic neutron scattering (QENS) and pulsed-field gradients-nuclear magnetic resonance (PFG-NMR).[15, 16] Such experiments are laborious and not suitable for high-throughput operations due to the time-consuming nature of sample preparation and measurement.[17, 18] As a result, experimental data are rarely available for molecular diffusion coefficients of chemical species in large libraries of nanoporous materials. Alternatively, diffusion coefficients can be predicted from a number of theoretical methods.[14, 19-21] Among them, molecular dynamics (MD) simulation has been most widely used to investigate the diffusion of gas molecules in nanoporous materials. Despite its popularity, construction of a diffusion-coefficient database by 'brute force' MD simulation is computationally prohibitive. The task is challenging in particular when one is concerned with the separation of organic molecules (e.g., paraxylene/orthoxylene and benzene/cyclohexane) using nanoporous materials due to the slow diffusivity (less than 10^{-12} m^2/s).

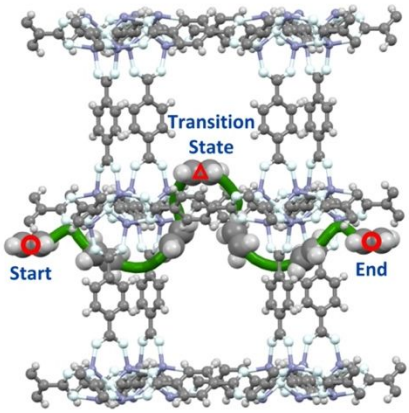


Figure 1. Minimum energy path for the center of mass (green line) of ethene in MOF-5. Molecular configuration of ethene at different position along the minimum energy path is also shown. The size of atoms is rescaled and for the purpose of illustration only. Grey, white, ice blue and purple represent carbon, hydrogen, oxygen and zinc, respectively.

Many theoretical attempts have been made to circumvent the computational limit of MD simulation in predicting diffusion coefficients.[10, 11] A well-established alternative is by using the transition-state theory (TST).[21] While diffusivity is typically calculated from the Einstein equation via mean-square displacement (MSD) over long equilibrium steps in MD simulation, TST predicts diffusion coefficients based on a minimum energy path (MEP) that is solely determined by the energy landscape of guest-host interactions (shown in Figure 1). Mathematical tools such as nudged elastic band (NEB) and string methods have been commonly used to calculate the MEP.[22-26] While NEB is mostly used in quantum-mechanical calculations of transport properties such as ion diffusivity, the string method is more suitable to obtain the highly curved MEP dictating gas diffusion in nanoporous materials.[27-31] More specifically, the string method is able to identify the diffusion pathways based on the energy gradients such that each path follows an exactly minimum energy route. Besides NEB and string methods, other mathematical tools, such as tunnel and transition-state search, cluster analysis and grid searching, are also promising.[11, 32]

Computationally, TST is able to predict diffusion coefficients much more efficient than molecular simulation because it entails no thermal fluctuations or atomic motions. Regrettably, existing applications of TST methods are mostly limited to the diffusion of simple gas molecules as represented by the single-site Lennard-Jones (LJ) potential. Not only is the extension of the MEP calculation to polyatomic molecules mathematically challenging, but the computational efficiency is severely compromised due to the rapid increase of dimensionality in representing MEP for polyatomic molecules.

In a previous work,[33] we demonstrated that, given a fine-enough three-dimensional potential grid, the string method can be used to accurately assess the minimum energy path (MEP) for the diffusion of simple gas molecules in nanoporous materials. However, the same procedure is not directly applicable to polyatomic molecules because the memory of a typical desktop computer is infeasible to handle the external potential using a multi-dimensional grid with a sufficiently fine resolution essential in MEP calculations. If the external potential is calculated on-the-fly as the string evolves, it would be an enormous computational burden for serial implementation with conventional central processing unit (CPU). Different from CPU, a graphic processing unit (GPU) has many more arithmetic logic units (ALUs, a.k.a. threads) thereby it is capable of high-throughput data processing. Inspired by recent progress of massively parallel GPU-acceleration of simulation methods with excellent performance,[34-36] we have implemented in this work a massively parallel GPU-accelerated string method for predicting the diffusivity of polyatomic molecule in a large library of nanoporous materials. The algorithm speeds up the theoretical predictions of diffusion coefficients with the string method by a factor of ~500 in comparison with serial CPU implementation. Importantly, the theoretical results are in excellent agreement with MD simulation data for a number of materials. We also benchmark the computational efficiency

for high-throughput screening of metal-organic frameworks (MOFs) for ethane/ethylene separation. More than 90% of calculations of the diffusion coefficient in the nanoporous materials can be completed within 30 seconds. By analyzing over 3080 structures from the CoRE MOF 2019 library, we are able to identify promising materials and desirable structural features leading to the highest membrane selectivity. We expect that the expanded computational capability will likely open up avenues for the construction of a large computational database for molecular diffusivity thus empowering data-driven approaches to the inverse design of nanoporous materials.

2. Methods and Models

2.1 Transition-State Theory

According to the transition-state theory[21], the self-diffusion coefficient for a guest molecule inside a nanoporous material can be calculated from

$$D_0 = \frac{1}{2} k a^2 \tag{1}$$

where D_0 is the self-diffusion coefficient, k is the hopping rate (i.e., transmission rate), and a is the hopping distance between two neighboring cages. At infinite dilution, the hopping rate can be obtained from the minimum energy path (MEP) for the molecular diffusion following the Bennett-Chandler approach

$$k = \sqrt{\frac{k_B T}{2\pi m}} \frac{\exp\left[-\beta V^{ext}\left(s^*\right)\right]}{\int_0^1 \exp\left[-\beta V^{ext}\left(s\right)\right] ds} \tag{2}$$

where k_B is the Boltzmann constant, T stands for the absolute temperature, $\beta = 1/\left(k_B T\right)$, m represents the molecular mass, V^{ext} is the potential energy due to the interaction of the guest molecule with the porous material. In the transition state theory, the minimum energy path is described in terms of a dimensionless variable s, which represents the normalized reaction

coordinate for the molecular transition between neighboring cages. In general, s depends on the molecular configuration and the center of mass of the guest molecule.

In this work, we assume that both the nanoporous material and the guest molecule have fixed structures. As a result, the reaction coordinate can be uniquely defined by six collective variables $s(\theta_1,...,\theta_6)$. The first three variables are related to the molecular position and the other three variables represent the Euler angles of the guest molecule, i.e., $s(\theta_1,...,\theta_6)=s(\mathbf{r},\omega)$, where $\mathbf{r}=(x,y,z)$ represents the position for the molecular center of mass (COM), and $\omega=(\alpha,\beta,\gamma)$ describes how a polyatomic molecule is oriented relative to its original input structure (as shown in Figure S1).

To implement the string method numerically, we describe the minimum energy path by using a series of discrete points referred to as images. At each point/image, the dimensionless variables s can be expressed in terms of $\mathbf{r}$ and ω

$$s_i=\sqrt{\left(\frac{l_{\mathbf{r},i}}{L_{\mathbf{r}}}\right)^2+\left(\frac{l_{\omega,i}}{L_{\omega}}\right)^2} \tag{3}$$

where $l_{\mathbf{r},i}=\sum_{j=2}^{i}\left|\mathbf{r}_j-\mathbf{r}_{j-1}\right|$ and $l_{\omega,i}=\sum_{j=2}^{i}\left|\omega_j^*-\omega_{j-1}^*\right|$ are the string arc lengths for the spatial and rotational variables at image i, while $L_{\mathbf{r}}$ and L_{ω} are the arc lengths of the entire string for $\mathbf{r}$ and ω, respectively.

In this work, we use atomistic models for both nanoporous materials and guest molecules. As in a standard molecular force field, the non-bonded interactions are described by the Lennard-Jones (LJ) potential. In addition to the short-range repulsion and van der Waals (vdW) attraction, the electrostatic interactions due to atomic partial charges are accounted for with the Coulomb potential. For atoms in the framework materials, the universal force field (UFF) is adopted for the

LJ parameters, while the charge equilibration method from the RASPA software package is used to assign the point charges of individual atoms.[37-39] The unit cell of each framework material is duplicated along the axis so that the edge length is more than two times the cutoff distance. Whereas the diffusion path may vary with the loadings due to gas-gas interactions, we expect that the effect is relatively insignificant because the free-energy landscape is dominated by the external energy. The force-field parameters for guest molecules considered in this work are available in Supporting Information. Structural properties of nanoporous materials, such as pore diameters and void fraction, are calculated with Zeo++.[40] Images and videos of atomistic molecular structures presented in this work are rendered from visual molecular dynamics (VMD) and Mercury.[41, 42]

Given the position and configuration of a guest molecule, the external potential accounts for its interaction with the nanoporous material and is given by

$$V^{ext} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_f} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}} \quad (4)$$

where ε and σ stand for the LJ energy and size parameters, respectively, ε_0 stands for the vacuum permeability, N_g and N_f are the number of atoms in each guest molecule and that from the nanoporous material. In calculation of V^{ext}, we use the Ewald summation method for electrostatic interactions, and the Lorentz-Berthelot mixing rule is used for the energy and size parameters between different atoms. The periodic boundary conditions are applied to all directions with the vdW interactions truncated and shifted to zero at 12.9 Å.

2.2 String Method

Within the framework of the transition-state theory (TST), both the computational cost and accuracy are critically dependent on the construction of the minimum energy path (MEP). In our previous work,[33] we demonstrated that the string method provides an efficient way to identify

MEPs, leading to an accurate prediction of self-diffusion coefficients for simple gas molecules in nanoporous materials. In principle, the string method is equally applicable to more complicated polyatomic molecules with the minimum energy path obtained by evolving discrete points (a.k.a. images) along "a string" towards the direction of decreasing the external potential. As the number of atoms in the guest molecule increases, the minimum energy path becomes much more difficult to calculate due to the drastic increase of pairwise interactions and the images in the reaction coordinate. In this work, we employ a simplified yet more accurate version of the string method to obtain the minimum energy path.[23] Compared with the original string method, the simplified string method is numerically more stable and accurate, yet it is computationally more efficient.[23]

According to the simplified string method,[23] the evolution of the normalized reaction coordinate is driven by the full gradient of the external potential

$$\frac{ds_i}{dt} = -\nabla V^{ext}(s_i) = -\left[\frac{\partial V^{ext}(s_i)}{\partial x} + \frac{\partial V^{ext}(s_i)}{\partial y} + \frac{\partial V^{ext}(s_i)}{\partial z} + \frac{\partial V^{ext}(s_i)}{\partial \alpha} + \frac{\partial V^{ext}(s_i)}{\partial \beta} + \frac{\partial V^{ext}(s_i)}{\partial \gamma}\right] \quad (5)$$

where t is a fictitious time used in the iteration to search for the minimum energy path, and s_i represents image i on the string. During each iteration, the guest molecule is first updated according to

$$\theta_{i,k}^{\square}(t) = \theta_{i,k}(t) - \Delta t \frac{\partial V^{ext}(s_i)}{\partial \theta_k}|_t, \qquad k = 1,...,6 \quad (6)$$

where θ_i represents a molecular coordinate (position or angle) corresponding to image i, the partial derivative is evaluated at fictious time t, and the superscript $^{\square}$ represents the updated string. Throughout this work, the forward Euler method is used to calculate the derivative of the external potential with respect to θ_i, and Δt is set as 1×10^{-4}.

To identify the diffusion path, we first calculate the energy landscape for the guest molecule at the entrance plane via a discrete grid. The position and orientation of the guest molecule that minimize the external potential are used as the starting image of the string. Due to the periodic boundary conditions, the guest molecule has the configuration at the starting and ending images. Their difference lies only in the reaction coordinate, i.e., parameter *s* along the direction of the minimum energy path. The initial string is generated by positioning the images evenly between the starting and ending points. After that, each iteration updates the position and orientation of the guest molecule according to the normalized reaction coordinate *s* (eq 6). After string evolution in each step, interpolation/reparameterization (eq 7) is needed to retain the continuous shape of string through the nanoporous materials.

When a guest molecule diffuses through a nanoporous material, the preferred molecular orientation depends on the position at the molecular center of mass (COM). As a result, different images have orientations independent from each other. We only need to interpolate the molecular COM position of the evolved images according to the arc length

$$\mathbf{r}_i\left(t+\Delta t\right)=\mathbf{r}_{k-1}^{\square}\left(t\right)+\left[\frac{i-1}{N-1}L_{\mathbf{r}}^{\square}\left(t\right)-l_{\mathbf{r},k-1}^{\square}\left(t\right)\right]\frac{\mathbf{r}_k^{\square}\left(t\right)-\mathbf{r}_{k-1}^{\square}\left(t\right)}{\left|\mathbf{r}_k^{\square}\left(t\right)-\mathbf{r}_{k-1}^{\square}\left(t\right)\right|} \tag{7}$$

where $l_{\mathbf{r},i}^{\square}=\sum_{j=2}^{i}\left|\mathbf{r}_j^{\parallel}-\mathbf{r}_{j-1}^{\parallel}\right|$ is the string arc length for spatial coordinates (*x*, *y*, *z*) at image *i* after evolution, $L_{\mathbf{r}}^{\square}$ is the entire string arc length for molecular COM position after evolution, and *N* is the number of images used in the string. To prevent the abrupt change of the molecular orientation, a smooth function is used for interpolating between neighboring images[22]

$$\theta_i=\left(1-\delta\right)\theta_i+\frac{\delta}{2}\left(\theta_{i-1}+\theta_{i+1}\right) \tag{8}$$

where δ is the parameter to control the degree of smoothness. A small number, $\delta = 1 \times 10^{-4}$, is used in this work to ensure the accuracy of molecular orientation that minimizes the external potential. In Supporting Information, we provide animations for the entire string evolution process for ethene in MOF-5.

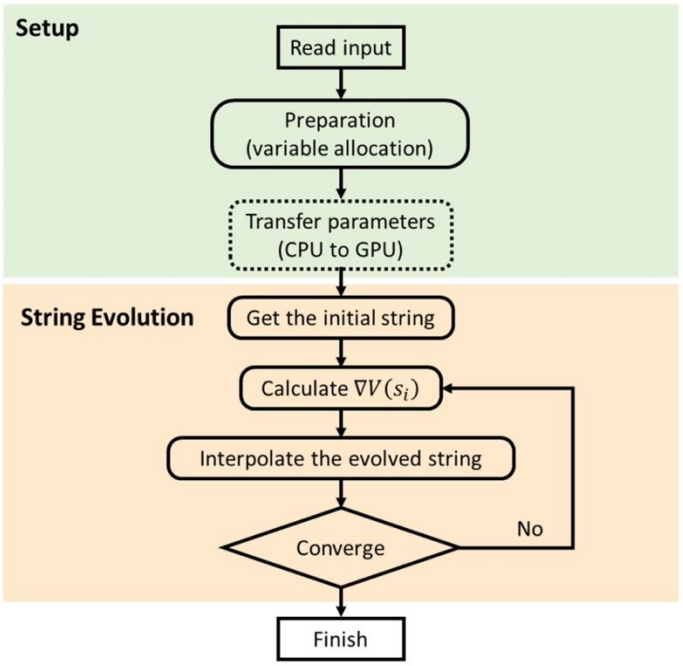
2.3 GPU Implementation

In implementing the string method to predict the diffusivity of polyatomic molecules in nanoporous materials, the computational cost is mostly affiliated with the calculation of the external potential and its derivates. In our previous work where the string method was applied to a single LJ particle, the external potential was pre-calculated by placing the guest molecule in a three-dimensional grid. The external potential at any point can be interpolated with a linear scheme. The procedure is not directly applicable to polyatomic molecules because the grid size grows exponentially with the dimensionality. The calculation of the external potential as a function of position and orientation in a fine grid is not only computationally prohibitive, but it is also logistically challenging to store the potential energy data over such a large grid. Without the support of a supercomputer with an enormous memory, the external potential has to be calculated on-the-fly along with the string evolution.

In this work, we implement the massively parallel simplified string method with the graphic processing unit (GPU). Figure 2(A) shows a schematic procedure. Compared with central processing unit (CPU), GPU is designed for high-throughput data processing and paralleled tasks. GPU enables the parallel calculation of the external potential on-the-fly and is much less memory demanding because it involves only the external potential related to the images along the string instead of entire energy landscape for the guest molecule inside framework material. As the external potential and its derivates are calculated for each image on the string independent of each

other, a paralleled implementation can significantly improve the speed and reduce the computational cost.

(A)



(B)

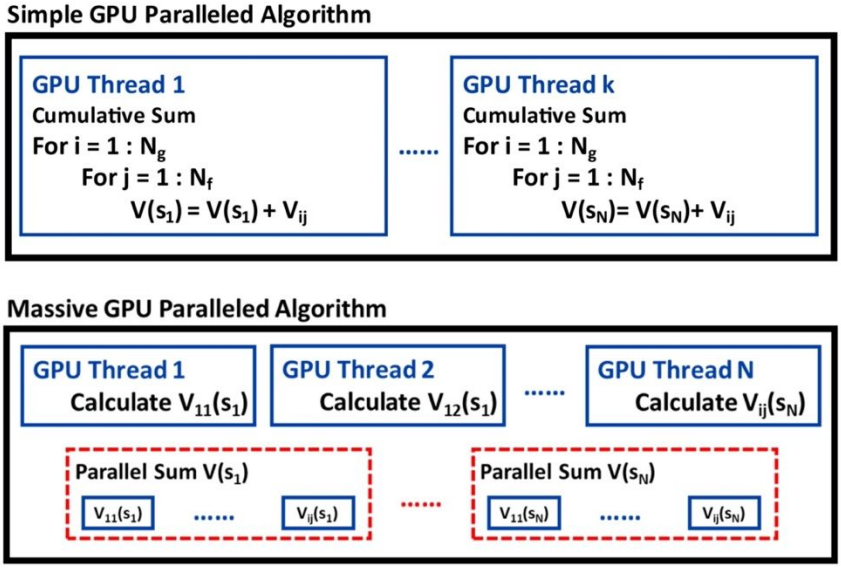


Figure 2. (A) Computational flowchart for the GPU implementation of the simplified string method. (B) Schematic illustration of simple and massive GPU paralleled algorithms for calculating the external potential for images on the string. Blue box represents the GPU thread.

We minimize the memory transfer between the host (CPU) and the device (GPU) in implementation of the string method. As shown in Figure 2, the memory transfer takes place only for reading the input information (such as the force-field parameters and the atomic structures of the nanoporous material and guest molecule), and for checking the convergence and outputting the final string configuration. The calculation is carried out only at the GPU device throughout the string evolution. The most expensive step in string evolution lies in the calculation of the derivatives of the external potential. Again, parallel implementation can significantly reduce the computational cost. As shown in Figure 2(B), two different GPU paralleled algorithms have been implemented and tested in this work. To calculate the derivatives of the external potential, we may assign a GPU thread for a given set parameters (x, y, z, α, β, γ). For a given set model parameters, the external potential is calculated on a single GPU thread by the cumulative summation of all pair potentials (V_{ij}) between atom i from the polyatomic molecule (guest) and atom j from the nanoporous material (host). While this parallel scheme is intuitive and simple to implement, it does not utilize all available GPU threads especially when the number of images on the string is relatively small. Alternatively, we can assign one GPU thread for each pair of the interatomic potential (V_{ij}) and calculate the external potential by the summation of V_{ij} for all atomic pairs evaluated via multiple GPU threads. In this work, we use the CUDA UnBound (CUB) library, a configurable C++ template library developed by Nvidia for Compute Unified Device Architecture (CUDA), to carry out the summation of V_{ij} on GPU.[43] With this massively parallel implementation, all GPU threads can be fully utilized even when the number of string images is relatively small.

One caveat of the massively paralleled implementation is that it leads to a higher demand of the memory usage. Because it stores all the pairwise interaction before summation, the massive GPU paralleled algorithm limits its capability handling a large set of string images. It is also worth mentioning that similar GPU-accelerated algorithms can also be developed to calculate the potential energy surface of polyatomic molecule in nanoporous materials and we have demonstrated such implementation for molecules modeled by single LJ site in our previous work.[34] The thermodynamic quantities would enable a rapid evaluation of properties, such as zero-coverage adsorption amount, in nanoporous materials for gas storage and separation.

3. Results and Discussion

3.1 GPU Speedup

To benchmark different GPU-accelerated parallel methods for implementing the string method, we take the diffusion of an ethene molecule in MOF-5 as a model system. Figure 3 compares the computational costs for two implementations of GPU-accelerated parallel algorithms as a function of the system size as measured in terms of the number of V^{ext} calculations (# of V^{ext}) and the number of V_{ij} calculations (# of V_{ij}). Here, the speedup factor is obtained by the comparison of the performance for Nvidia Tesla P100, which is used for all our GPU calculations, with that for the serial CPU implementation on Intel Xeon E5-2640. Both simple and massive paralleled GPU implementations outperform the serial CPU implementation regardless of the system size. The number of V^{ext} calculations is solely determined by the number of images on the string, while the number of V_{ij} calculations for each image depends on the number of atoms at the guest molecule and the number of atoms from the nanoporous material. When the number of images on the string (equivalently, the number of V^{ext} calculations) increases, the speedup factor rises exponentially for the simple GPU paralleled implementation until it reaches a plateau after all GPU threads are

utilized. For the GPU device tested in this work, the speedup factor approaches an asymptotic limit when it processes more than 100,000 images along the diffusion pathway (viz., the string). The maximum speedup by the simple GPU paralleled algorithm is about 2500 folds of the CPU serial implementation.

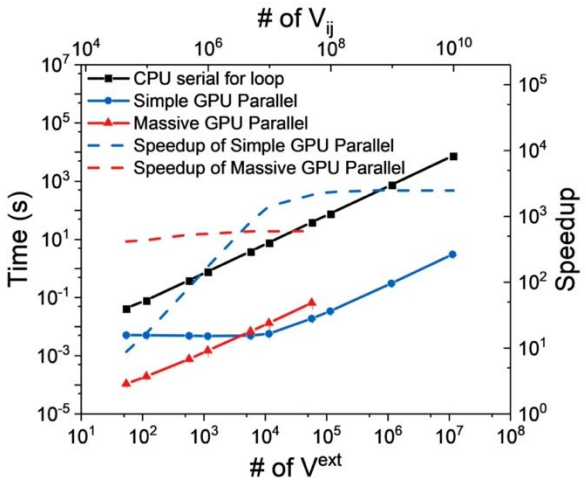


Figure 3. Comparison of the computational time versus the number of the total potential (V^{ext}) and pair potential (V_{ij}) evaluations for predicting ethene diffusion in MOF-5. The speedup factor is benchmarked with CPU calculations conducted on Intel Xeon E5-2640. All GPU calculations are carried out on Nvidia Tesla P100.

For the massive GPU paralleled implementation, the computational cost increases exponentially with the number of string images irrespective of the system size. Because the fully paralleled algorithm maximizes the usage of all active GPU threads, the speedup factor, which is around 500, is almost independent of the number of string images. As mentioned above, the massive GPU paralleled implementation consumes more memory space than the simple GPU paralleled implementation (as shown in Figure S2). Thus, the upper limit for the number images that can be processed by the massive GPU paralleled implementation is much lower than that for simple GPU paralleled implementation. For most nanoporous materials considered in this work,

the unit cell size varies from 10 Å to 30 Å, such that a string with hundreds of images would be sufficient to preserve all atomistic details along the minimum energy path (MEP). As a result, the massive GPU paralleled implementation has a better performance and thus used in all the following calculations.

3.2 Calibration with MD simulation

We see above that excellent computational performance can be achieved by massively parallel GPU-accelerated implementation of the string method. But how accurate is the theoretical procedure for predicting diffusivity coefficients in comparison with conventional methods? In this section, we compare our theoretical predictions for the diffusion coefficients of 8 polyatomic molecules in MOF-5, a well-studied metal organic framework (MOF), with those from molecular dynamics (MD) simulations. As mentioned before, our theoretical predictions are based on the transition-state theory (TST) with the minimum energy path (MEP) calculated from the massive GPU paralleled implementation of the string method.

As shown in Figure 4(A), the diffusion coefficients calculated in this work agree well with those from MD simulation over a broad range of values. The good agreement affirms the accuracy of the minimum energy path obtained from the GPU-accelerated calculations. Overall, TST predicts the self-diffusivity of various polyatomic molecules in MOF-5 slightly higher than that from MD simulation. The systematic error is introduced probably because TST neglects the barrier recrossing in molecular hopping.[44] The barrier recrossing of gas molecule becomes more significant especially at finite loadings. In this case, the dynamically corrected TST (dcTST) can better estimate the hopping rate ($k_{dcTST}=\kappa k_{TST}$) by correcting the recrossing event with the transmission coefficient (κ).[45, 46] Another possible reason is that most MD simulations are not carried in the single molecule limit thus the simulation results are affected by intermolecular

interactions between the guest molecules. In principle, TST can be used to predict diffusivity coefficients at finite gas pressure if MEP is replaced by the free-energy landscape. Alternatively, the self-diffusivity at finite loading can be calculated from the diffusivity coefficient at infinite dilution in combination with the excess-entropy scaling method.[33, 47] It is worth mentioning that, in comparison with MD, one of the most significant advantages of TST is computational efficiency. While it takes up to several thousands of CPU hours to simulate the diffusion coefficient of CO_2 (at the scale of 10^{-9} m^2/s), a relatively small polyatomic molecule, the same calculation can be finished in this work within 30 seconds for each material by using a single GPU card. Although both gas molecules (guest) and framework materials (host) are assumed to be rigid in this work, the string method can also be used to obtain the minimum energy path when flexibility of gas molecules and framework is significant such as large molecule squeezing through tight aperture. In the latter case, the computation will be more demanding because we need to consider both the guest-host interactions and the intramolecular potential.

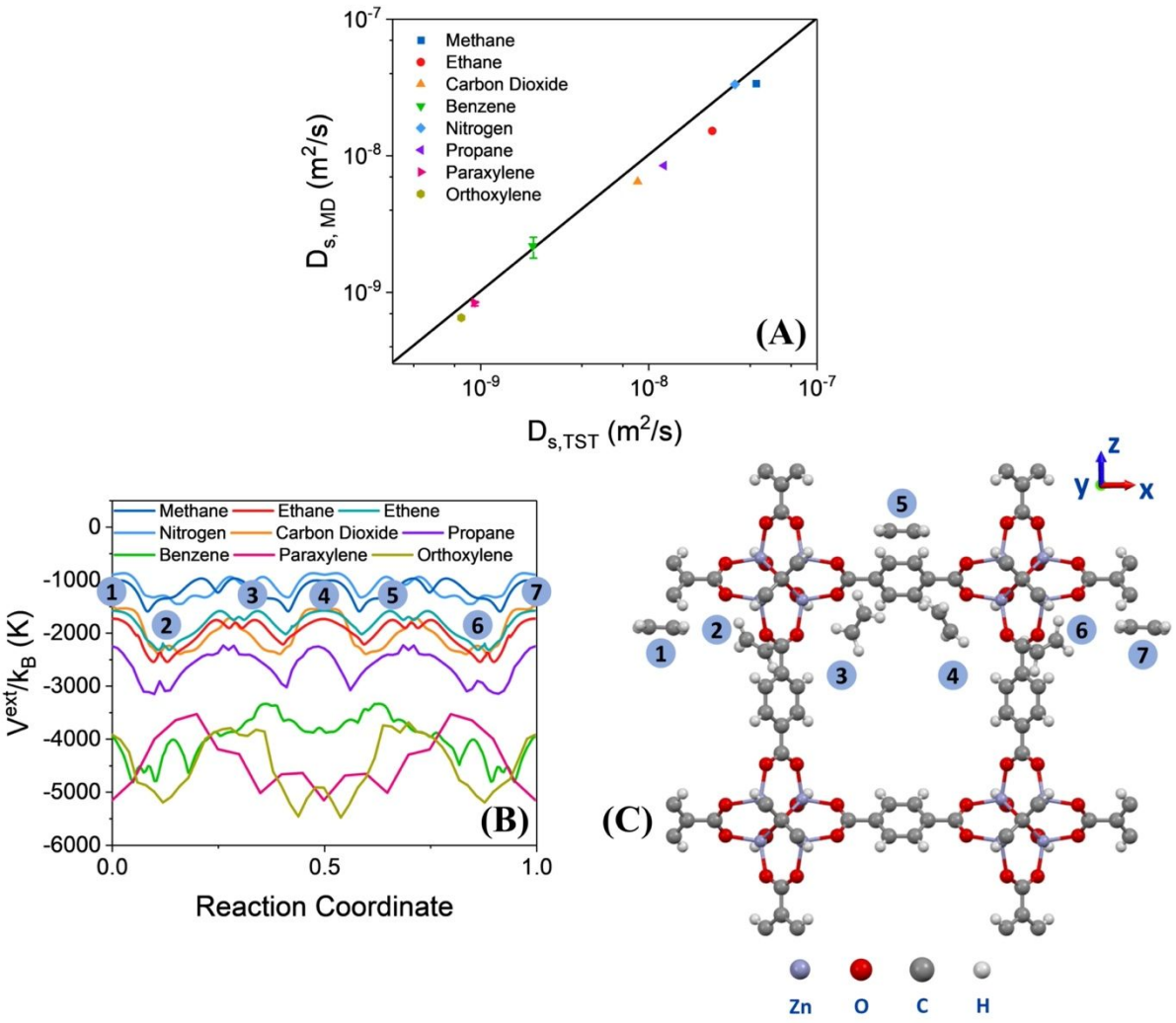


Figure 4. (A) Comparison of diffusion coefficients for 8 polyatomic molecules in MOF-5 predicted by the transition-state theory (TST) and by molecular dynamics (MD) simulations. The MD results are from the literature.[16, 48-51] (B) The minimum energy paths calculated from the GPU paralleled implementation of the simplified string method versus the reaction coordinate. Here the numbers 1-7 stand for images along the minimum energy path in MOF-5 for the ethene molecule. (C) The positions and orientations of the ethene molecule corresponding to the 7 images labeled in (B). (Ethene is modeled as a diatomic molecule according to TraPPE-UA force field. The molecular structures of ethene in (C) are only for illustration purpose).

Figure 4 (B) presents the minimum energy paths (MEPs) for 8 polyatomic molecules tested in this work. For small molecules such as ethane, carbon dioxide and propane, their energy landscapes along the diffusion path have a similar shape and exhibit only slightly different potential energy barriers, suggesting a similar hopping pattern for the diffusion of small molecules in MOF-5. According to the molecular orientation along the MEP shown in Figure 4(C), the center of mass (COM) for the ethene molecule stays close to the metal cluster and organic linker in MOF-5 instead of going through MOF-5 in the center of pore in order to maintain the minimized hots-guest interactions. When the ethene molecule enters across the pore, its orientation also changes so that the smaller edge of the molecular plane would be directed toward the angle minimizing the external potential.

Figure 4(B) shows the minimum energy paths (MEPs) for benzene, paraxylene and orthoxylene. These paths are significantly different from those for smaller molecules such as ethene due to the molecular size. Even when comparing the potential energies and molecular orientation along MEPs of large molecules, they are quite different from each other because, for large molecules, even slight modifications (add/relocate) on the functional group would lead to significant difference in their preferred orientation at the energy barrier and rotational and translational activation energy along the reaction coordinate. In Figure 5, we compare the position and molecular orientation of p-type xylene with those corresponding to o-type xylene along the MEP in MOF-5. As discovered in NMR studies and by MD simulation,[52] the p-type xylene molecule, especially its methyl group, is located around the pore center when the hopping takes place between neighboring cages (shown in Figure S3). For the o-type xylene, the COM position is close to the metal cluster (viz., at the corner of pore) before hoping to the neighboring cages. In addition, the molecular orientation of p-type xylene changes much less than o-type xylene along

the MEP because p-type xylene has higher activation energy due to the rotational move as shown in Figure 5.

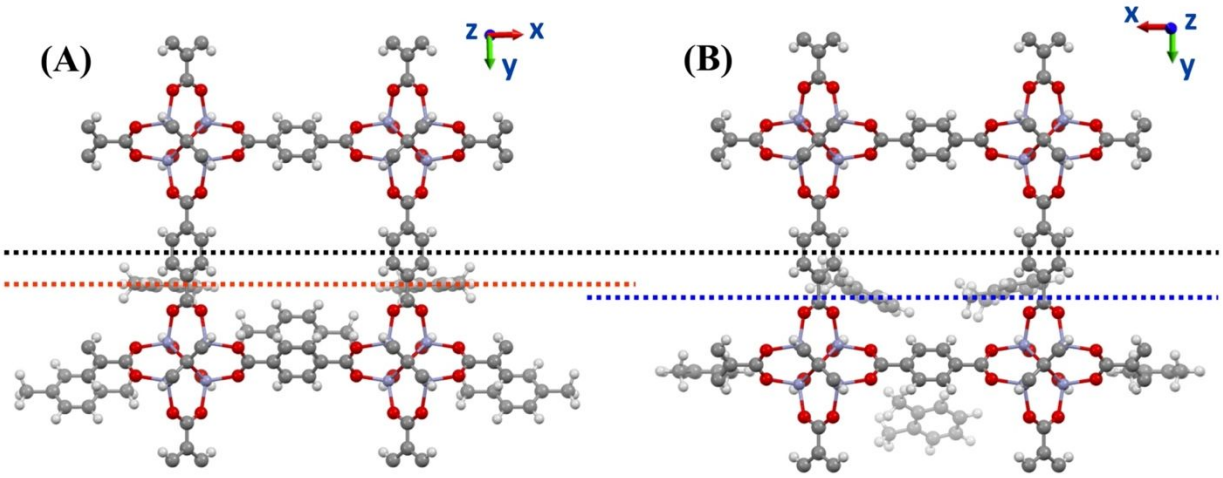


Figure 5. Molecular positions and orientations of p-type xylene (A) and o-type xylene (B) in MOF-5 along the minimum energy path. (P-type and o-type xylene are modeled as eight united group sites according to TraPPE-UA force field. The detailed guest molecular structures are for illustration purpose only). Black, red and blue dashed line represents the center of pore in MOF-5, center of mass for p-type xylene when crossing the pore and center of mass for o-type xylene when crossing the pore, respectively.

3.3 High-throughput Screening

Efficient prediction of diffusivity will likely open up opportunities for high-throughput screening and, eventually, for the inverse design of nanoporous materials for practical applications such as gas separation. Here, we demonstrate the capability of massively parallel GPU-accelerated string method for high-throughput screening of MOFs useful for ethane/ethene separation, a challenging yet important task in the chemical industry. Since ethane and ethene molecules have similar physical characteristics such as the size and shape, previous studies indicate that MOFs with pore limit diameter (PLD) between 3 Å and 4 Å are most efficient in terms of selectivity.[53]

We use PLD as an initial criterion to select 3080 candidates from the computational-ready, experimental (CoRE) MOF 2019 database which covers over 14 000 porous structures.[54]

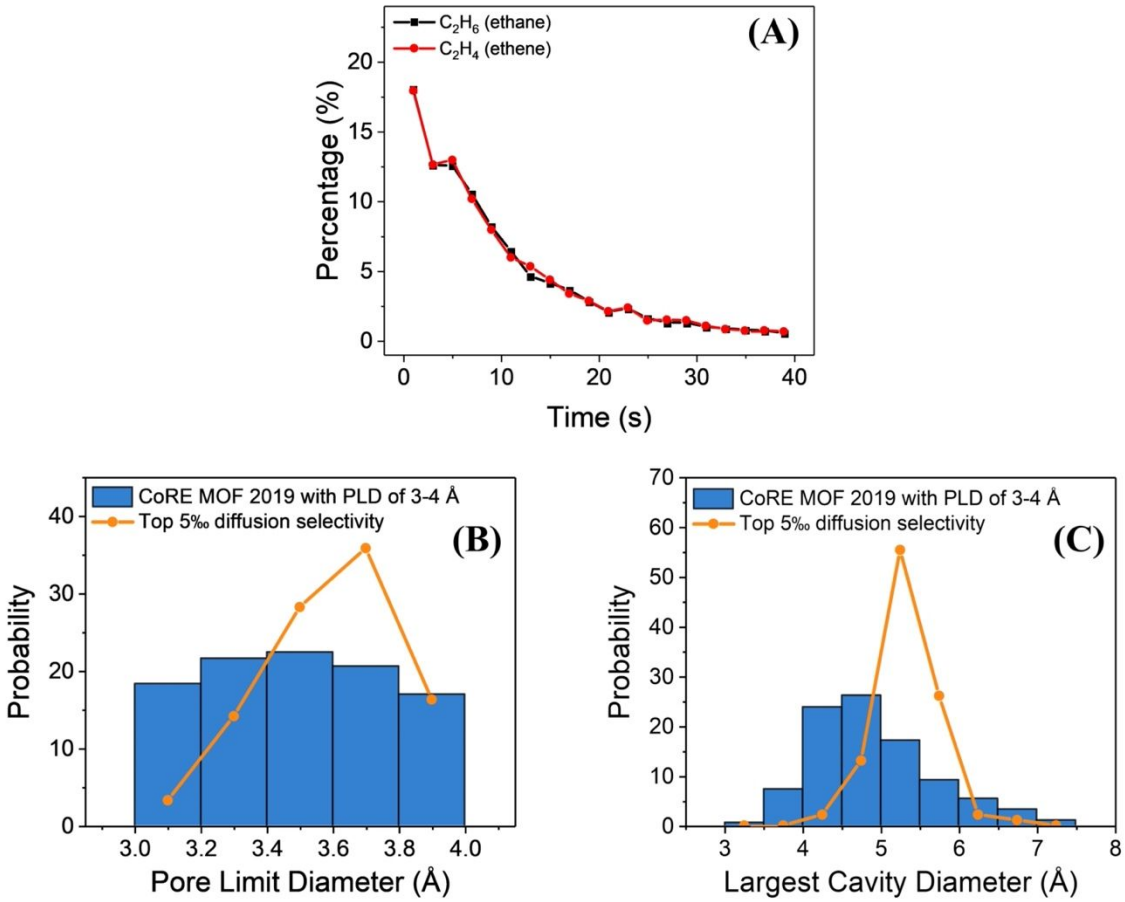


Figure 6. (A) Distribution of the computational time in high-throughput screening of 3080 MOF candidates for the separation of ethane/ethene gases at room temperature. Distributions of (B) the pore limit diameter (PLD) and (C) largest cavity diameter (LCD) of those MOFs with the highest diffusion selectivity.

As shown in Figure 6(A), the diffusivity calculation for most MOFs (more than 90%) can be accomplished within less than 30 seconds, which is significantly faster than conventional methods such as molecular dynamics simulation (up to hundreds of CPU hours per material).[33] The massively parallel GPU-accelerated string method is also much faster than emerging methods that

search only the tunnel space or the transition state with the polyatomic molecules represented by a single-site LJ model, which typically cost ~0.5 CPU hour per material.[10] The computational cost of the string method depends not only on how many iterations it would take to identify the minimum energy path but also on the number of atoms in the system including those from both the guest molecule and the framework material (host). Ethene (3.23×4.18×4.84 Å) and ethane (3.81×4.08×4.82 Å) have similar molecular shape and size. Because both are modeled as diatomic molecules in the TraPPE force field, the computational costs of finding the minimum energy paths for ethane and ethene are almost identical.[55]

In Figure 6(B) and (C), we present the distributions of the pore limit diameter (PLD) and the largest cavity diameter (LCD) for MOFs in the database with top 0.5% diffusion selectivity for the separation of ethane/ethene gases at room temperature. Here, the diffusion selectivity is calculated from the ratio of diffusivity, $S_{diff,1/2} = D_{0,1} / D_{0.2}$ and 1 and 2 refers to ethene and ethane, respectively. For materials in the CoRE MOF 2019 database, the PLDs are evenly distributed between 3 Å and 4 Å, whereas the distribution of MOFs with the highest diffusion selectivity has a notable peak between 3.4 Å and 3.8 Å. Although the range of the PLD from 3.4 Å to 3.8 Å is significantly smaller than the kinetic diameter of ethane (4.443 Å) and ethene (4.163 Å) molecules derived from the second virial coefficients, it falls into the perfect range for separating ethene from ethane according to the molecular size and shape. The peak value (3.4-3.8 Å) is larger than the smallest edge of an ethene molecule (3.28 Å) but smaller than that of ethane (3.81 Å).[53] Compared with the LCD distribution of MOFs in the background, which follows approximately a normal distribution with the mean between 4.5 Å and 5 Å, MOFs with top 0.5% diffusion selectivity has a slightly higher mean, between 5 Å and 5.5 Å, in the LCD distribution. According to our previous work,[29, 56, 57] a nanoporous material with the LCD larger than the molecular size would impose

more attraction along the minimum energy path, which is beneficial to achieve the diffusivity coefficient at the scale of practical interest.

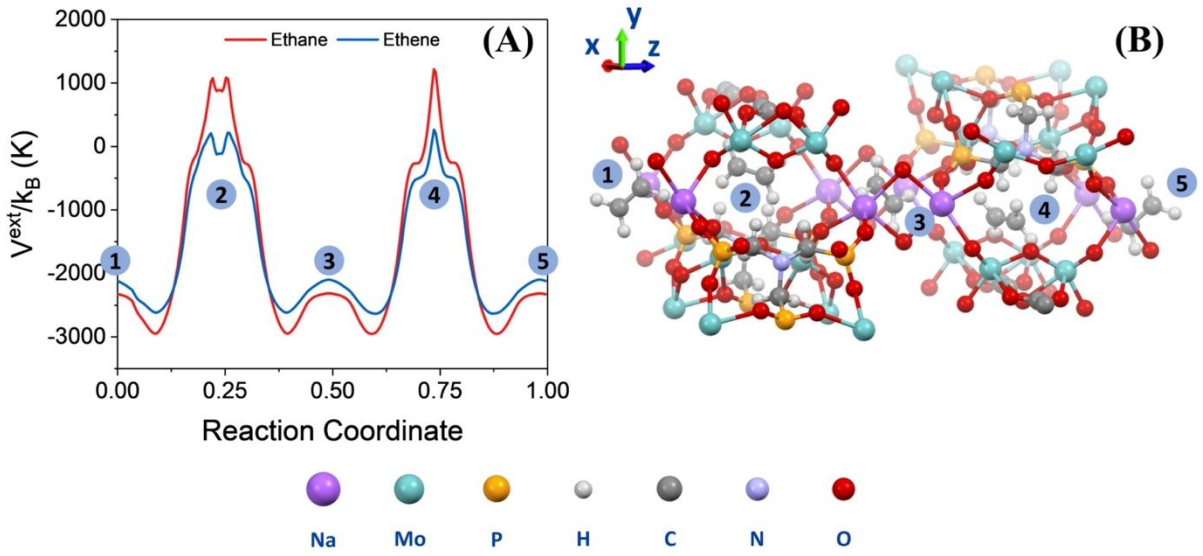


Figure 7. (A) Energy landscape along the minimum energy path for ethane and ethene in YIGFIF, a nanoporous material with the highest diffusion selectivity for the separation of ethane and ethene gases at 300 K. (B) The molecular position and orientation of an ethene molecule along the minimum energy path in YIGFIF.

In the Table S2, we present the diffusion coefficients and structural properties of top 10 MOFs with the highest diffusivity. Among all MOFs investigated in this work, YIGFIF has a diffusion selectivity of 57.69, which is the highest for the separation of ethane/ethene at 300 K. Figure 7 shows the energy landscape of ethane and ethene molecules along the minimum energy path in YIGFIF. For ethane diffusing along the MEP in YIGFIF, the COM position is almost identical to that for ethene. As shown in Figure 7(B), the rotation of an ethene molecule inside YIGFIF is restricted due to strong confinement along the MEP in YIGFIF (PLD: 3.38 Å and LCD 5.02 Å). However, the nanoporous material exerts a repulsive energy on ethane at the transition state much stronger than that on ethene. The larger diffusion barrier may be attributed to the minimum cross-

section area of YIGFIF (~3.38 Å ×4.94 Å), which can be utilized to sieve ethane and ethene with an excellent diffusion selectivity. Although YIGFIF is the most promising MOF candidate for the separation of ethane/ethene as the membrane materials according to our screening, it has not yet been experimentally tested for any practical applications.[58] Compared with conventional ethene-selective adsorbent materials (selectivity up to 48.7), membrane separation with YIGFIF would be much less energy-intensive for industrial application with around 20% higher separation selectivity.[53, 59] For the state-of-art ethane-selective adsorbent materials (selectivity up to 4.4), YIGFIF can achieve a much higher separation selectivity of ethane/ethene.[60]

4. Conclusion

In this work, we have implemented a GPU-accelerated string method to calculate the minimum energy path (MEP) for polyatomic molecules in nanoporous materials. The MEP calculation is essential for predicting diffusivity using the transition-state theory. Both simple GPU parallel algorithm and massive GPU parallel algorithm are tested and benchmarked with serial CPU calculations. Compared with the serial CPU implementation on Intel Xeon E5-2640, GPU implementations on Nvidia Tesla P100 may speedup the diffusivity calculation up to ~2500 folds via the simple GPU parallel algorithm. The outstanding performance is attributed to massive threads available on GPU and the minimized memory transfer between CPU (host) and GPU (device). Although the simple GPU paralleled implementation can achieve up to three orders of magnitude speedup compared to the serial CPU implementation, the speedup factor depends on the number of images on the diffusion pathway (viz. the reaction coordinate represented by a string) and is much lower than that could be achieved by massive GPU paralleled implementation, especially when the number of images on the string is relatively small due to the insufficient usage of GPU threads. For massive GPU parallel implementation, a constant speedup factor around 500

is achieved regardless of the number of images on the string, an indication of excellent parallelization for MEP calculations. Because most nanoporous materials have the largest edge in the unit cell less than 30 Å, the massive GPU paralleled implementation of the string method is more advantageous for calculation of MEP for polyatomic molecules in nanoporous materials.

The diffusion coefficients of guest molecules in nanoporous materials can be calculated from MEP via the transition-state theory (TST). The results are compared with the diffusion coefficients from molecular dynamics (MD) simulation for 8 polyatomic molecules in MOF-5. Excellent agreement between theory and simulation is achieved, further indicating the accuracy of MEP obtained by the GPU-accelerated string method. Because TST underestimates the diffusion barrier and recrossing of molecular hopping, and because most MD simulations are not run in the single-molecule limit, the diffusivity from TST is slightly larger than that from simulation. While the diffusion of small molecules (e.g., ethene, nitrogen and carbon dioxide) in MOF-5 shares a similar hopping pattern, large molecules such as p-type xylene and o-type xylene have significantly different trajectories for their positions and orientations along the minimum energy path. Different from MD simulation whereby diffusivity is calculated from the statistical average of molecular movements in random, the minimum energy path calculated from GPU-accelerated massively parallel string method offers the microscopic details of molecular hopping that can be utilized to guide the rational design of nanoporous materials for the separation of polyatomic molecules.

Finally, we have demonstrated the capability of massively parallel GPU-accelerated string method for high-throughput screening MOFs for the separation of ethane and ethene, two polyatomic molecules of similar size and shape that are of tremendous importance for the chemical industry. In the high-throughput screening calculations, 3080 MOFs are selected from computational-ready, experimental MOF database CoRE MOF 2019 according to their pore limit

diameters (PLDs). The massively parallel GPU-accelerated string method is used to calculate the diffusion selectivity for ethane/ethene separation. The diffusivity calculation can be completed within 30 seconds per material for more than 90% of MOFs, which is significantly faster than MD simulation (cost ~ hundreds of CPU hours per material). The GPU calculation outperforms even emerging methods such as TuTraSt (cost ~ 0.5 CPU hour for a single LJ site).[10, 33] While the calculation of slow diffusion behavior (less than 1×10^{-12} m^2/s) is computationally prohibitive for molecular dynamics, the computational cost for the string method does not change with the scale of the diffusion coefficient. According to our high-throughput screening calculations, MOFs with the pore limit diameter (PLD) from 3.4 to 3.8 Å and the largest cavity diameter (LCD) between 5 and 5.5 Å can efficiently separate ethene (with the molecular dimensions of 3.23×4.18×4.84 Å) from ethane (with the molecular dimensions of 3.81×4.08×4.82 Å) by their molecular sizes and shapes. YIGFIF (PLD= 3.38 Å and LCD=5.02 Å) from the CoRE MOF 2019 database has been identified with the highest diffusion selectivity for ethene/ethane separation, with a theoretical selectivity that can reach up to 57.69 at 300 K.

The GPU-accelerated massively parallel implementation of string method enables efficient and accurate calculation of diffusion coefficients for polyatomic molecules in nanoporous materials. We expect that the computational platform will be generally useful for high-throughput screening of nanoporous materials, for example, as the membrane for the separation of polyatomic molecules. It can also be used to construct high-fidelity properties database for the inverse design of nanoporous materials.

Supporting Information

It provides the force field parameters for all polyatomic molecules considered in this work, the GPU memory usages of different GPU paralleled implementations and the structural properties of

MOFs with the highest diffusion selectivity for the separation of ethane and ethene at 300 K. Movies for the string evolution of ethene in MOF-5 are also provided. For download and use of the computer codes for the string method implemented in this work, visit the GitHub repository (https://github.com/MusenZhou/GaSSM).

Acknowledgement

This work is financially supported by the National Science Foundation's Harnessing the Data Revolution (HDR) Big Ideas Program under Grant No. NSF 1940118.

Reference

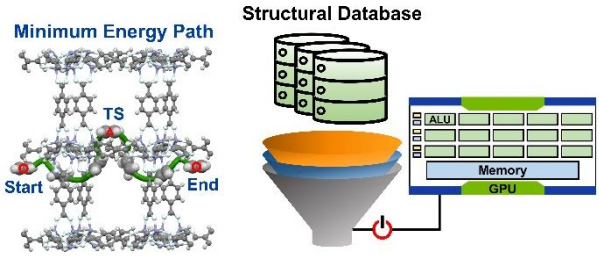
1. Rangnekar, N.; Mittal, N.; Elyassi, B.; Caro, J.; Tsapatsis, M., Zeolite membranes - a review and comparison with MOFs. *Chem Soc Rev* **2015,** *44* (20), 7128-54.
2. Al-Rowaili, F. N.; Jamal, A.; Shammakh, M. S. B.; Rana, A., A Review on Recent Advances for Electrochemical Reduction of Carbon Dioxide to Methanol Using Metal-Organic Framework (MOF) and Non-MOF Catalysts: Challenges and Future Prospects. *Acs Sustainable Chemistry & Engineering* **2018,** *6* (12), 15895-15914.
3. Ali, M.; Pervaiz, E.; Noor, T.; Rabi, O.; Zahra, R.; Yang, M., Recent advancements in MOF-based catalysts for applications in electrochemical and photoelectrochemical water splitting: A review. *International Journal of Energy Research* **2021,** *45* (2), 1190-1226.
4. Li, Y.; Wu, Q.; Guo, X.; Zhang, M.; Chen, B.; Wei, G.; Li, X.; Li, X.; Li, S.; Ma, L., Laminated self-standing covalent organic framework membrane with uniformly distributed subnanopores for ionic and molecular sieving. *Nat Commun* **2020,** *11* (1), 599.
5. Fan, H.; Peng, M.; Strauss, I.; Mundstock, A.; Meng, H.; Caro, J., MOF-in-COF molecular sieving membrane for selective hydrogen separation. *Nat Commun* **2021,** *12* (1), 38.
6. Li, J.; Zhou, X.; Wang, J.; Li, X. F., Two-Dimensional Covalent Organic Frameworks (COFs) for Membrane Separation: a Mini Review. *Industrial & Engineering Chemistry Research* **2019,** *58* (34), 15394-15406.
7. Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; S., L.; D., V. K.; S., L.; Milisavljevic, M.; Zhang, H.; Camp, J.; Slater, B.; Siepmann, J. I.; Sholl, D. S.; Snurr, R. Q., Computation-Ready Experimental Metal-Organic Framework (CoRE MOF) 2019 Dataset. Zenodo, 2019.
8. Moghadam, P. Z.; Li, A.; Wiggin, S. B.; Tao, A.; Maloney, A. G. P.; Wood, P. A.; Ward, S. C.; Fairen-Jimenez, D., Development of a Cambridge Structural Database Subset: A Collection of Metal-Organic Frameworks for Past, Present, and Future. *Chemistry of Materials* **2017,** *29* (7), 2618-2625.
9. Tranchemontagne, D. J.; Mendoza-Cortes, J. L.; O'Keeffe, M.; Yaghi, O. M., Secondary building units, nets and bonding in the chemistry of metal-organic frameworks. *Chem Soc Rev* **2009,** *38* (5), 1257-83.
10. Mace, A.; Barthel, S.; Smit, B., Automated Multiscale Approach To Predict Self-Diffusion from a Potential Energy Field. *J Chem Theory Comput* **2019,** *15* (4), 2127-2141.

11. Haldoupis, E.; Nair, S.; Sholl, D. S., Efficient calculation of diffusion limitations in metal organic framework materials: a tool for identifying materials for kinetic separations. *J Am Chem Soc* **2010,** *132* (21), 7528-39.
12. Abraham, J.; Vasu, K. S.; Williams, C. D.; Gopinadhan, K.; Su, Y.; Cherian, C. T.; Dix, J.; Prestat, E.; Haigh, S. J.; Grigorieva, I. V.; Carbone, P.; Geim, A. K.; Nair, R. R., Tunable sieving of ions using graphene oxide membranes. *Nature nanotechnology* **2017,** *12* (6), 546-550.
13. Wijmans, J. G.; Baker, R. W., The Solution-Diffusion Model - a Review. *Journal of Membrane Science* **1995,** *107* (1-2), 1-21.
14. Dubbeldam, D.; Snurr, R. Q., Recent developments in the molecular modeling of diffusion in nanoporous materials. *Molecular Simulation* **2007,** *33* (4-5), 305-325.
15. Jobic, H.; Laloue, N.; Laroche, C.; van Baten, J. M.; Krishna, R., Influence of isotherm inflection on the loading dependence of the diffusivities of n-hexane and n-heptane in MFI zeolite. Quasi-elastic neutron scattering experiments supplemented by molecular simulations. *J Phys Chem B* **2006,** *110* (5), 2195-201.
16. Ford, D. C.; Dubbeldam, D.; Snurr, R. Q.; Kunzel, V.; Wehring, M.; Stallmach, F.; Karger, J.; Muller, U., Self-Diffusion of Chain Molecules in the Metal-Organic Framework IRMOF-1: Simulation and Experiment. *J Phys Chem Lett* **2012,** *3* (7), 930-3.
17. Urbanczyk, M.; Kozminski, W.; Kazimierczuk, K., Accelerating diffusion-ordered NMR spectroscopy by joint sparse sampling of diffusion and time dimensions. *Angew Chem Int Ed Engl* **2014,** *53* (25), 6464-7.
18. Price, K. E.; Lucas, L. H.; Larive, C. K., Analytical applications of NMR diffusion measurements. *Anal Bioanal Chem* **2004,** *378* (6), 1405-7.
19. Altintas, C.; Keskin, S., Molecular simulations of MOF membranes for separation of ethane/ethene and ethane/methane mixtures. *RSC Adv* **2017,** *7* (82), 52283-52295.
20. Avci, G.; Velioglu, S.; Keskin, S., High-Throughput Screening of MOF Adsorbents and Membranes for H2 Purification and CO2 Capture. *ACS Appl Mater Interfaces* **2018,** *10* (39), 33693-33706.
21. Kärger, J.; Ruthven, D. M.; Theodorou, D. N., *Diffusion in Nanoporous Materials, 2 Volume Set*. John Wiley & Sons: 2012.
22. Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G., String method in collective variables: minimum free energy paths and isocommittor surfaces. *J Chem Phys* **2006,** *125* (2), 24106.
23. E, W.; Ren, W.; Vanden-Eijnden, E., Simplified and improved string method for computing the minimum energy paths in barrier-crossing events. *J Chem Phys* **2007,** *126* (16), 164103.
24. Weinan, E.; Ren, W. Q.; Vanden-Eijnden, E., String method for the study of rare events. *Physical Review B* **2002,** *66* (5), 052301.
25. Jónsson, H.; Mills, G.; Jacobsen, K. W., Nudged elastic band method for finding minimum energy paths of transitions. **1998**.
26. Sheppard, D.; Xiao, P.; Chemelewski, W.; Johnson, D. D.; Henkelman, G., A generalized solid-state nudged elastic band method. *J Chem Phys* **2012,** *136* (7), 074103.
27. Cvitas, M. T.; Althorpe, S. C., Locating Instantons in Calculations of Tunneling Splittings: The Test Case of Malonaldehyde. *J Chem Theory Comput* **2016,** *12* (2), 787-803.
28. Zhou, M.; Vassallo, A.; Wu, J., Toward the inverse design of MOF membranes for efficient D2/H2 separation by combination of physics-based and data-driven modeling. *Journal of Membrane Science* **2020,** *598*, 117675.

29. Zhou, M. S.; Tian, Y.; Fei, W. Y.; Wu, J. Z., Fractionation of Isotopic Methanes with Metal-Organic Frameworks. *J Phys Chem C* **2019,** *123* (12), 7397-7407.
30. Mahdizadeh, S. J.; Goharshadi, E. K., Multicomponent gas separation and purification using advanced 2D carbonaceous nanomaterials. *Rsc Advances* **2020,** *10* (41), 24255-24264.
31. Kocer, C. P.; Griffith, K. J.; Grey, C. P.; Morris, A. J., Lithium Diffusion in Niobium Tungsten Oxide Shear Structures. *Chem Mater* **2020,** *32* (9), 3980-3989.
32. Coifman, R. R.; Kevrekidis, I. G.; Lafon, S.; Maggioni, M.; Nadler, B., Diffusion Maps, Reduction Coordinates, and Low Dimensional Representation of Stochastic Systems. *Multiscale Modeling & Simulation* **2008,** *7* (2), 842-864.
33. Tian, Y.; Xu, X.; Wu, J., Thermodynamic Route to Efficient Prediction of Gas Diffusivity in Nanoporous Materials. *Langmuir* **2017,** *33* (42), 11797-11803.
34. Zhou, M.; Wu, J., A GPU implementation of classical density functional theory for rapid prediction of gas adsorption in nanoporous materials. *J Chem Phys* **2020,** *153* (7), 074101.
35. Acun, B.; Hardy, D. J.; Kale, L. V.; Li, K.; Phillips, J. C.; Stone, J. E., Scalable Molecular Dynamics with NAMD on the Summit System. *IBM J Res Dev* **2018,** *62* (6), 1-9.
36. Nitsche, M. A.; Ferreria, M.; Mocskos, E. E.; Gonzalez Lebrero, M. C., GPU Accelerated Implementation of Density Functional Theory for Hybrid QM/MM Simulations. *J Chem Theory Comput* **2014,** *10* (3), 959-67.
37. Rappe, A. K.; Goddard, W. A., Charge Equilibration for Molecular-Dynamics Simulations. *J Phys Chem-Us* **1991,** *95* (8), 3358-3363.
38. Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M., Uff, a Full Periodic-Table Force-Field for Molecular Mechanics and Molecular-Dynamics Simulations. *Journal of the American Chemical Society* **1992,** *114* (25), 10024-10035.
39. Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q., RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Molecular Simulation* **2016,** *42* (2), 81-101.
40. Willems, T. F.; Rycroft, C.; Kazi, M.; Meza, J. C.; Haranczyk, M., Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous and Mesoporous Materials* **2012,** *149* (1), 134-141.
41. Humphrey, W.; Dalke, A.; Schulten, K., VMD: visual molecular dynamics. *J Mol Graph* **1996,** *14* (1), 33-8, 27-8.
42. Macrae, C. F.; Sovago, I.; Cottrell, S. J.; Galek, P. T.; McCabe, P.; Pidcock, E.; Platings, M.; Shields, G. P.; Stevens, J. S.; Towler, M., Mercury 4.0: From visualization to analysis, design and prediction. *Journal of applied crystallography* **2020,** *53* (1), 226-235.
43. Merrill, D., Cuda unbound (cub) library. *NVIDIA-Labs* **2015**.
44. Vanden-Eijnden, E.; Tal, F. A., Transition state theory: variational formulation, dynamical corrections, and error estimates. *J Chem Phys* **2005,** *123* (18), 184103.
45. Beerdsen, E.; Dubbeldam, D.; Smit, B., Understanding diffusion in nanoporous materials. *Physical review letters* **2006,** *96* (4), 044501.
46. Dubbeldam, D.; Beerdsen, E.; Vlugt, T. J. H.; Smit, B., Molecular simulation of loading-dependent diffusion in nanoporous materials using extended dynamically corrected transition state theory. *Journal of Chemical Physics* **2005,** *122* (22), 224712.
47. Dyre, J. C., Perspective: Excess-entropy scaling. *J Chem Phys* **2018,** *149* (21), 210901.
48. Skoulidas, A. I.; Sholl, D. S., Self-diffusion and transport diffusion of light gases in metal-organic framework materials assessed using molecular dynamics simulations. *Journal of Physical Chemistry B* **2005,** *109* (33), 15760-15768.

49. Cabrales-Navarro, F. A.; Gomez-Ballesteros, J. L.; Balbuena, P. B., Molecular dynamics simulations of metal-organic frameworks as membranes for gas mixtures separation. *Journal of Membrane Science* **2013,** *428*, 241-250.
50. Amirjalayer, S.; Schmid, R., Mechanism of benzene diffusion in MOF-5: A molecular dynamics investigation. *Microporous and Mesoporous Materials* **2009,** *125* (1-2), 90-96.
51. Witherspoon, V. J.; Yu, L. M.; Jawahery, S.; Braun, E.; Moosavi, S. M.; Schnell, S. K.; Smit, B.; Reimer, J. A., Translational and rotational motion of C8 aromatics adsorbed in isotropic porous media (MOF-5): NMR studies and MD simulations. *The Journal of Physical Chemistry C* **2017,** *121* (28), 15456-15462.
52. Witherspoon, V. J.; Yu, L. M.; Jawahery, S.; Braun, E.; Moosavi, S. M.; Schnell, S. K.; Smit, B.; Reimer, J. A., Translational and Rotational Motion of C8 Aromatics Adsorbed in Isotropic Porous Media (MOF-5): NMR Studies and MD Simulations. *J Phys Chem C* **2017,** *121* (28), 15456-15462.
53. Bao, Z.; Wang, J.; Zhang, Z.; Xing, H.; Yang, Q.; Yang, Y.; Wu, H.; Krishna, R.; Zhou, W.; Chen, B., Molecular Sieving of Ethane from Ethylene through the Molecular Cross-Section Size Differentiation in Gallate-based Metal–Organic Frameworks. *Angewandte Chemie International Edition* **2018,** *57* (49), 16020-16025.
54. Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; Lee, S.; Zhang, H. D.; Vogiatzis, K. D.; Milisavljevic, M.; Ling, S. L.; Camp, J. S.; Slater, B.; Siepmann, J. I.; Sholl, D. S.; Snurr, R. Q., Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal-Organic Framework Database: CoRE MOF 2019. *Journal of Chemical and Engineering Data* **2019,** *64* (12), 5985-5998.
55. Wu, Y.; Chen, H. Y.; Liu, D. F.; Qian, Y.; Xi, H. X., Adsorption and separation of ethane/ethylene on ZIFs with various topologies: Combining GCMC simulation with the ideal adsorbed solution theory (IAST). *Chemical Engineering Science* **2015,** *124*, 144-153.
56. Zhou, M.; Vassallo, A.; Wu, J., Toward the inverse design of MOF membranes for efficient D2/H2 separation by combination of physics-based and data-driven modeling. *Journal of Membrane Science* **2020,** *598*.
57. Wang, J. Q.; Zhou, M. S.; Lu, D. N.; Fei, W. Y.; Wu, J. Z., Computational screening and design of nanoporous membranes for efficient carbon isotope separation. *Green Energy & Environment* **2020,** *5* (3), 364-373.
58. Yang, L.; Zhou, Z.; Ma, P. T.; Wang, J. P.; Niu, J. Y., Self-assembly of two ring-shaped hexanuclear Mo(VI) clusters. *Crystengcomm* **2013,** *15* (27), 5452-5457.
59. Yang, S. H.; Ramirez-Cuesta, A. J.; Newby, R.; Garcia-Sakai, V.; Manuel, P.; Callear, S. K.; Campbell, S. I.; Tang, C. C.; Schroder, M., Supramolecular binding and separation of hydrocarbons within a functionalized porous metal-organic framework. *Nature Chemistry* **2015,** *7* (2), 121-129.
60. Li, L.; Lin, R. B.; Krishna, R.; Li, H.; Xiang, S.; Wu, H.; Li, J.; Zhou, W.; Chen, B., Ethane/ethylene separation in a metal-organic framework with iron-peroxo sites. *Science* **2018,** *362* (6413), 443-446.

TOC Graphic


Minimum Energy Path
TS
Start
End
Structural Database
ALU
Memory
GPU