PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Physical Chemistry Chemical Physics Dynamic Article Links ►

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

A molecular dynamics study of oxygen ion diffusion in A-site ordered perovskite $PrBaCo_2O_{5.5}$: data mining the oxygen trajectories

Chi Chen,[a] Dengjie Chen[a] and Francesco Ciucci*[a,b]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

Molecular dynamics (MD) simulations have been widely used to study oxygen ion diffusion in crystals. In the data analysis, one typically calculates the mean squared displacements to obtain the self-diffusion coefficients. Further information extraction for each individual atom poses significant challenges due to the lack of general methods. In this work, oxygen ion diffusion in A-site ordered perovskite $PrBaCo_2O_{5.5}$ is studied using MD simulations and the oxygen migration is analyzed by the k-means clustering machine learning algorithm. The clustering analysis allows the tracking of each individual oxygen jump along with its corresponding location, i.e., oxygen site in BaO, $PrO_{0.5}$ and CoO_2 layers. Therefore it increases the understanding into the factors influencing oxygen diffusion. For example, it is found that the oxygen occupation fraction in the $PrO_{0.5}$ layers increases with temperature, while in the CoO_2 layers it decreases with temperature; the activation enthalpies of oxygen jumps from CoO_2 to CoO_2, CoO_2 to $PrO_{0.5}$ and $PrO_{0.5}$ to CoO_2 are 0.22 eV, 0.54 eV and 0.34 eV respectively, exhibiting anisotropic characteristics. Furthermore, the dwell times of oxygen atoms suggest that the oxygen atoms are highly mobile in $PrO_{0.5}$ layers. Combining the analysis of activation enthalpies and dwell times, it is suggested that the oxygen transport is fast within the CoO_2 layers while the $PrO_{0.5}$ layers work as oxygen vacancy reservoirs.

Introduction

Oxygen diffusion is an important process in many emerging high-temperature applications, such as solid oxide fuel cells (SOFCs)[1] and oxygen transport membranes (OTMs).[2] In SOFCs, the oxygen reduction reaction at the cathode can be facilitated by high oxygen ion conductivity. More importantly, oxygen diffusion through the electrolyte is a key factor affecting the overall cell performance. Therefore, substantial efforts have been devoted to developing highly oxygen conductive materials as electrolytes for SOFCs.[3, 4] Similarly, in OTMs the oxygen ion diffusion across the membrane material is one of the essential processes to oxygen separation. A fundamental understanding of oxygen diffusion in solid state systems is thus of primary importance for the development of the next-generation SOFCs and OTMs.[5]

Experimental techniques including electrical conductivity relaxation[6, 7] and isotope exchange depth profiling[8] are widely used to probe the oxygen diffusion within ceramic materials. However, oxygen diffusion is a complex process, which depends on the crystallographic structure, the composition, and the local strain state of the material.[1] In order to complement the experimental insights, molecular dynamics (MD) simulations are often used because they can provide atomistic understanding into oxygen transport. For example, yttria-stabilized zirconia was calculated to obtain the maximum ionic conductivity at 8 mol% Y_2O_3 by MD simulations,[9] consistently with experimental observations. This maximum conductivity was attributed to the trade-off between the increase of oxygen vacancy and the reduced vacancy mobility brought by Y_2O_3. Fisher et al. used MD simulations to examine the relationship between oxygen ion conductivity and dopant concentration in $Ba_{1-x}Sr_xCo_{1-y}Fe_yO_{2.5}$.[10] This work helped explain why $Ba_{0.5}Sr_{0.5}Co_{0.8}Fe_{0.2}O_{3-\delta}$ was chosen as one of the best cathode materials.[11] Recently, Ciucci et al. calculated oxygen diffusion in $Ba_{1-x}La_xFeO_{3-\delta}$ using MD simulations.[12] The results suggested that the oxygen self-diffusivity decreased with increasing La substitution. Therefore, $Ba_{0.95}La_{0.05}FeO_{3-\delta}$ was chosen as a cathode material for SOFCs. Indeed, $Ba_{0.95}La_{0.05}FeO_{3-\delta}$ exhibited a high performance comparable to the Co-based materials.[12] Besides the studies in simple cubic structures, MD simulations have also been used to investigate anisotropic materials and layered oxides,[13-20] a large family of materials that are interesting as cathodes for SOFCs.[21] For example, Chroneos et al. found that oxygen diffusion in $GdBaCo_2O_{5+\delta}$ is connected to the order/disorder characteristics of Gd/Ba arrangements,[14] which is consistent with experimental observations.[22] Hernandez and Dezanneau et al. studied $NdBaCo_2O_{5+x}$ by MD simulations,[18] reproducing well the oxygen nuclear density obtained by neutron diffraction.

In spite of the power of MD simulations, most data regarding oxygen diffusion is actually discarded. While it is not uncommon to output a few hundred megabytes of data per MD run, only the oxygen diffusivity, a number (a few bytes), is calculated by fitting the mean squared displacements (MSDs).[23, 24] Although the MSD approach is effective for analyzing the impact of compositional modification on the diffusion coefficients,[25, 26] it is an average measure that does not include details on atomic level hopping of mobile species. This approach might not be sufficient

for clarifying the oxygen diffusion mechanisms, especially in materials with complex structure and multiple dopants. Such detailed studies require the utilization of trajectories of all atoms as well as an atomically resolved analysis.[27-29] Unfortunately, this is challenging due to the intrinsically disordered movement of oxygen in space. The other difficulty is that the MD trajectories are usually stored in large datasets. Yet, the data itself may be embedded in a much lower space producing a low-dimensional representation that contains only the relevant features (or collective coordinates).[30] A number of techniques have been developed for reducing the dimensionality of MD simulations including linear methods such as principal component analysis[31] and nonlinear methods, such as local linear embedding,[32] isomaps,[33] sketch maps,[34] and diffusion maps.[35, 36] The dimensionality reduction methods can also be used for intuitive representation and for simulation speed up.[37] However, these approaches have found applications primarily in protein folding and nucleic acid simulations.

In this work the concept of dimensionality reduction is used in order to embed the complex oxygen diffusion in perovskite oxides into a low dimensional clustered space. A recent work focusing on proton conduction in Y-doped $BaZrO_3$ has utilized distances from oxygen atoms with fixed positions to track the proton trajectories.[38] Although this is suitable for proton transport, the standard distance analysis applied therein may not be applied to oxygen diffusion. In fact, due to the oxygen transport such reference locations are no longer fixed and furthermore they may not be placed at the nominal oxygen lattice sites. Additionally, if dopants are introduced or if the diffusion is anisotropic, the reference points may be offset locally from the nominal oxygen sites, as will be shown later in this paper.

Here a general data mining framework is proposed for the analysis of the MD trajectories. Clustering analysis is a general unsupervised classification method that can be used to group data with similar features. This is particular relevant to the study of oxygen diffusion since determining patterns in the trajectories is intrinsically an unsupervised learning problem. The clustering approach is illustrated using the MD simulations of oxygen diffusion in $PrBaCo_2O_{5.5}$ (PBCO). PBCO is chosen because it has great potential as a cathode material for SOFCs.[39, 40] Furthermore, its anisotropic conduction mechanisms due to its layered structure and defect processes have been studied both experimentally[8] and computationally.[16, 41]

Methods

MD simulations

MD simulations are carried out using the LAMMPS package[42] following our previous work, where the model was validated against experimental data.[43] Interactions between ions are described by the Buckingham pairwise potential with coulombic forces,

$$\phi_{\alpha\beta} = \frac{Z_\alpha Z_\beta e^2}{4\pi\epsilon r_{\alpha\beta}} + A_{\alpha\beta} exp\left(-\frac{r_{\alpha\beta}}{\rho_{\alpha\beta}}\right) - \frac{C_{\alpha\beta}}{r_{\alpha\beta}^6} \quad (1)$$

where the subscript denotes the ion pair, Z is the charge, r is the distance between the pair, and $A_{\alpha\beta}$, $\rho_{\alpha\beta}$ and $C_{\alpha\beta}$ are potential parameters. On the right hand side, the first term describes the coulombic force, the second is the repulsive force due to the electron overlap when atoms are close, and the last is the Van der Waals interaction.[23]

The potential parameters are taken from existing publications as listed in **Table 1**, and the cut off range for the short range interaction is set to 11 Å, enough for reproducing the oxygen transport mechanisms.[16]

Table 1. Potential parameters for the Buckingham potential

Ion pairs	$A_{\alpha\beta}$/eV	$\rho_{\alpha\beta}$/Å	$C_{\alpha\beta}$/eVÅ^6	Ref.
O^{2-}---O^{2-}	22764.3	0.1490	43.00	[13]
Ba^{2+}---O^{2-}	1214.4	0.3522	0	[13]
Co^{3+}---O^{2-}	1329.82	0.3087	0	[13]
Pr^{3+}---O^{2-}	1445.2	0.3608	0	[44]

The long-range coulombic interaction is calculated by the Ewald summation with an accuracy of 10^{-4}. Oxygen diffusion in PBCO is simulated in a 8x8x4 supercell[18] where the oxygen vacancies are initially placed at random in the $PrO_{0.5}$ planes.[45] The simulation temperature ranges from 873 to 1573 K and a time step of 1 fs is used for the Verlet integrator. The system is first equilibrated for 100 ps in the NPT ensemble followed by a data collection period of 2 ns in the NVT ensemble with atom trajectories recorded every 1 ps. The Nose-Hoover thermostat was employed in all simulations.

Clustering analysis

The k-means clustering algorithm[46] is used to analyze the trajectories. This method has been used extensively in solving unsupervised machine learning problems. We note that the metal ions are found to only vibrate around their equilibrium sites and are not directly involved in the oxygen migration. Therefore, their movements will not be included in the analysis.

The trajectories of oxygen atoms are stored in $\boldsymbol{T}$, a $n \times 3 \times p$ tensor, where n is the number of oxygen atoms and p is the number of MD output steps. The purpose of the clustering is to map $\boldsymbol{T}$ into smaller matrices as in **Fig. 1** without major loss of information.

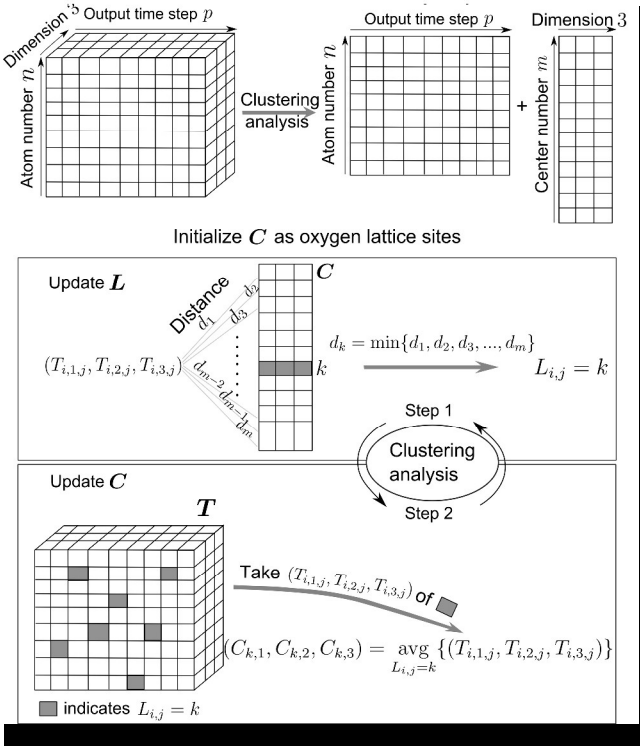

Fig.1 A diagram illustrating the clustering analysis as applied to MD trajectories. The goal of clustering is to reduce the tensor $\boldsymbol{T}$ containing the trajectory data into 2 matrices $\boldsymbol{L}$ and $\boldsymbol{C}$. $\boldsymbol{L}$ stores the labels trajectory and the matrix C contains the coordinates of the cluster centers. The algorithm follows a two-step process by updating the labels and cluster centers iteratively.

Namely, two matrices will be used: $\boldsymbol{C}$, which stores the locations of the cluster centers, and $\boldsymbol{L}$, which tracks the time evolution of each oxygen atom in reference to its corresponding cluster. $\boldsymbol{L}$ is a $n \times p$ matrix that stores the cluster labels of the n atoms at each of the p steps output by the MD simulations. The labels are used to distinguish different clusters and are integers ranging from 1 to m, where m is the number of total clusters or oxygen sites. The coordinates of the cluster centers (one for each label) are stored in $\boldsymbol{C}$, an $m \times 3$ matrix. Each row is mapped to a specific cluster center and the columns indicate x, y, and z coordinates. Typically m is greater than n due to the existence of oxygen vacancies.

The clustering analysis starts by initializing the centers as the nominal crystallographic oxygen lattice sites and then proceeds iteratively as the two-step process illustrated in **Fig. 1**. As a first step, the label trajectories $\boldsymbol{L}$ are updated. One atom snapshot, for example atom i at time step j, is taken out from $\boldsymbol{T}$. The coordinate of this atom is therefore $(T_{i,1,j}, T_{i,2,j}, T_{i,3,j})$. The distances between the captioned atom and all the cluster centers are computed as $\{d_1, d_2, d_3, \dots, d_m\}$ under periodic conditions. If d_k is the minimal distance, then $L_{i,j} = k$. This step corresponds to solving the following optimization problem.

$$L_{i,j} = \underset{k=1\dots m}{\text{argmin}}\ \text{distance}\{(T_{i,1,j}, T_{i,2,j}, T_{i,3,j}), (C_{k,1}, C_{k,2}, C_{k,3})\}$$
$$\text{for } i = 1 \text{ to } n \text{ and } j = 1 \text{ to } p \quad (2)$$

As a second step, $\boldsymbol{C}$ is updated. The k-th cluster center is computed by first selecting the atom snapshots $(T_{i,1,j}, T_{i,2,j}, T_{i,3,j})$ in $\boldsymbol{T}$ whose corresponding labels are $L_{i,j} = k$. Then the coordinates of the selected atom snapshots are averaged to give the new k-th cluster center $(C_{k,1}, C_{k,2}, C_{k,3})$:

$$(C_{k,1}, C_{k,2}, C_{k,3}) = \underset{(i,j) \text{ for } L_{i,j}=k}{\text{avg}} (T_{i,1,j}, T_{i,2,j}, T_{i,3,j})$$
$$\text{for } k = 1 \text{ to } m \quad (3)$$

Steps 1 and 2 are executed iteratively until the distance between the centers computed in two successive loops is lower than a given tolerance.

In order to attach physical meaning to $\boldsymbol{C}$, each of its rows is assigned a site type. This distinguishes oxygen sites in the BaO, CoO_2 and $PrO_{0.5}$ plane. For the perovskites, the oxygen site is located in the octahedral hole surrounded by 4 A-site cations and 2 B-site cations. The site type of the final oxygen sites can be calculated as described in **Fig. 2**. First, the nearest neighbor A-site cations of certain oxygen site are found. Depending on how many Ba are around the site, the oxygen site can be classified into 3 types: type 1 with 4 Ba atoms around the site, type 2 with 2 Ba atoms around, and type 3 with 0 Ba atoms.

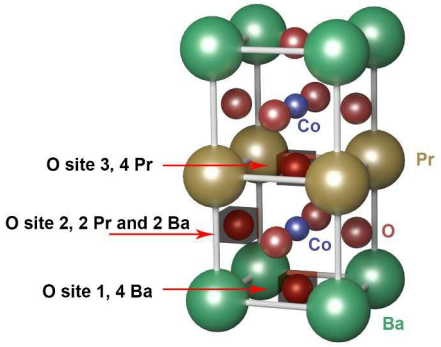

Fig. 2 Oxygen site types defined by the surrounding number of different A-site cations.

The advantage of this method is that it automatically groups the trajectories, classifies the oxygen sites into different categories, and allows easy analysis of the MD trajectory patterns. Besides, this method is general and can be applied to study diffusion patterns of other ionic conductors.

Results and discussion

Oxygen transport trajectories along with the centers for a portion of the whole system at 1573 K are projected onto the ac plane, as shown in **Fig. 3**. The clear anisotropic diffusion characteristics are observable. Oxygen atoms initially in the BaO plane are confined in it and are primarily capable of oscillating around their equilibrium positions. Conversely, the oxygen trajectories across the CoO_2 and the $PrO_{0.5}$ plane are connected. Interestingly, the final cluster centers pertaining to the CoO_2 plane deviate from the initial centers obtained from static crystallography, i.e., the learnt centers are closer to the $PrO_{0.5}$ plane. This shift is caused by the oxygen transport from CoO_2 to $PrO_{0.5}$, which favors oxygen accumulation between these two layers. As shown in **Fig. 3**, the initial centers are biased with respect to the centroid of the oxygen atom snapshot cloud formed by the whole trajectories. This is because the new centers are derived directly from the

computed oxygen trajectories, rather than being inferred from static crystallographic positions. Therefore, the learnt centers from the k-means clustering better describes oxygen locations in comparison to the initial centers.

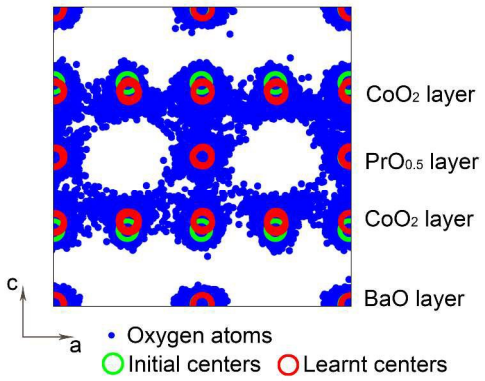

Fig. 3 Learnt cluster centers in comparison to the initial oxygen sites at 1573 K of a 2x8x1 supercell for better reading, projected onto the ac plane.

The trajectories of oxygen projected onto ab planes are shown in **Fig. 4**. In CoO_2 layers, migrations between two different locations are clearly observable, while in the $PrO_{0.5}$ and BaO planes jumps between the same type of sites are extremely unlikely.[13] In the ab plane, however, the learnt centers overlap with the initial centers due to symmetry. The oxygen atoms in the BaO plane are more compactly distributed in comparison to the $PrO_{0.5}$, indicating better transport within the latter.

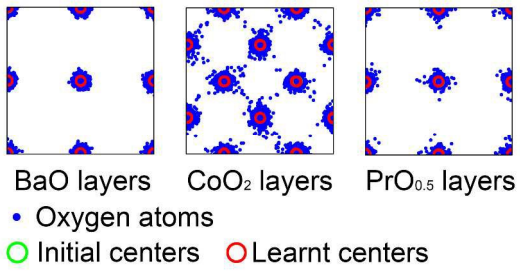

Fig. 4 Oxygen trajectories projected onto the ab plane for BaO layers, CoO_2 layers and $PrO_{0.5}$ layers from left to right. The atom locations are shown at 1573 K for a 2x2 unit cell. The initial centers and the learnt centers overlap.

The movement of oxygen atoms in PBCO has been typically analyzed by following the variation of relevant coordinates.[45] This method is demonstrated in **Fig. 5a** for a single oxygen atom, where 18 jumping events are shown by tracking the evolution of its *x*, *y* and *z* location. One may observe that the change in only one coordinate does not fully describe the oxygen diffusion. This is evidenced by the fact that events 1, 2, 3, 4, 8, 9, 10, 11 and 12 cannot be detected by only the *x* coordinate and events 2, 5, 6, 7, 13, 14, 15 and 17 are not noticeable by monitoring the *y* coordinate alone. Even for the *z* coordinate variations, event 18 is missing. However, when the clustering analysis is applied, the jumping events can be analyzed as a function of only one variable, i.e., the label of the site where the oxygen atom resides.

As shown in **Fig. 5b**, each level indicates a different label and all the jumps can be captured by the label change. Furthermore, by including the site type of each label, the jumps between different types of sites can be analyzed, as shown **Fig. 5c**. It is found therein that solely relying on the *z* coordinate change of **Fig. 5a** misses a jump between sites within the same CoO_2 layer. This is evidenced by the fact that the site type does not change at event 18, as shown in **Fig. 5c**.

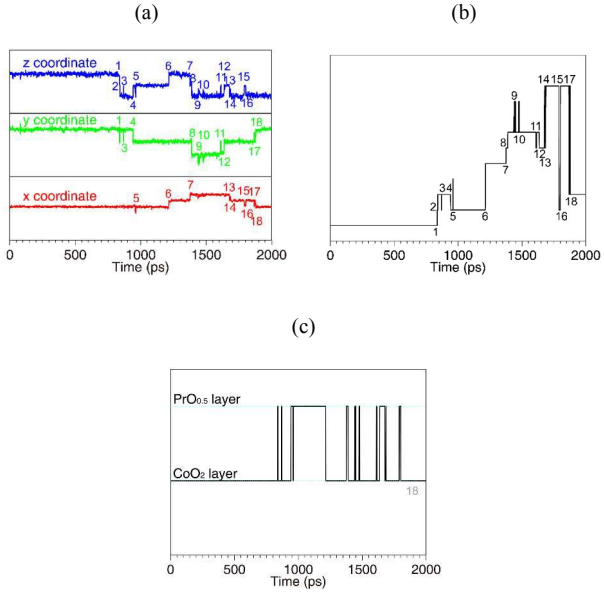

Fig. 5 Coordinate changes for one representative oxygen atom as a function of time with jumping events marked sequentially (a), and corresponding site label (b) and site type(c) evolution.

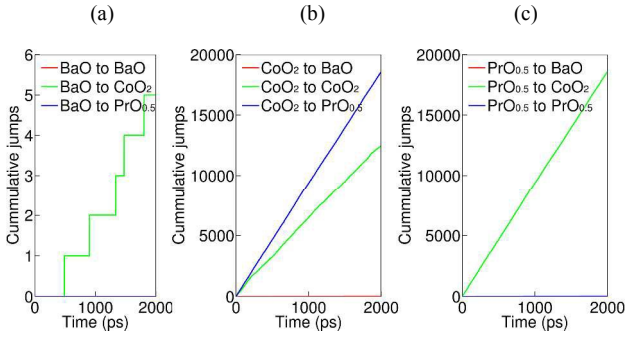

Fig. 6 Cumulative number of jumps over time from oxygen in BaO layers to other layers (a), CoO_2 to other layers (b) and $PrO_{0.5}$ to other layers (c) at 1573 K.

The analysis described in **Fig. 5** can be applied seamlessly to all atoms, leading to the total number of jumps between different layers. **Fig. 6a** shows that only a few hops occur in the BaO planes. In addition, no migration is recorded from the BaO layers to the BaO and $PrO_{0.5}$ layers. For oxygen atoms in the CoO_2 layers, however, frequent leaps are detected from CoO_2 layers to $PrO_{0.5}$ layers as well as to CoO_2 layers themselves as shown in **Fig. 6b**. In addition, hops to the $PrO_{0.5}$ layers are more frequent than to the CoO_2 layers. This may be caused by the higher concentration of oxygen vacancies in the $PrO_{0.5}$ plane as shown in

Fig. 7, thereby providing fast vacancy exchange between CoO_2 and $PrO_{0.5}$ layers. Lastly, the jumps from $PrO_{0.5}$ have only one significant path, as shown in **Fig. 6c**.

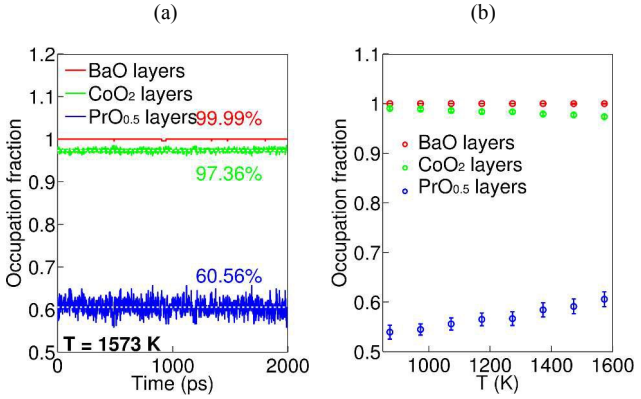

Fig. 7 Oxygen site occupancy fraction in different planes at 1573 K (a) and at different temperatures (b).

The analysis of oxygen migration can be easily used to calculate the occupation fraction of oxygen sites in different planes. **Fig. 7a** shows the evolution of the occupation fraction with simulation time at 1573 K. The system is at equilibrium since only fluctuations around the mean value are observed. The standard deviation in the $PrO_{0.5}$ planes is the highest due to the fewer sites within this plane as well as the higher mobility. This analysis can be applied to other temperatures, as shown in **Fig. 7b**. Interestingly, the oxygen occupation fraction of the $PrO_{0.5}$ planes increases with temperature, while it decreases with temperature for the CoO_2 planes. According to previous static calculations, the configurations with oxygen vacancy located in the $PrO_{0.5}$ planes have lower lattice energies,[45] making them more favorable at low temperature. However, at high temperature, higher energy states can be reached. In turn, this leads to the decrease of vacancy concentrations and to the increase of occupation fraction in the $PrO_{0.5}$ planes. Meanwhile, the occupation in the CoO_2 planes decreases with temperature due to the mass conservation and the lack of transport in the BaO planes.

The temperature effects on the oxygen transport from CoO_2 to CoO_2 layers and CoO_2 to $PrO_{0.5}$ layers are shown in **Fig. 8.** By increasing the temperature, the oxygen transport rates in both planes increase substantially.

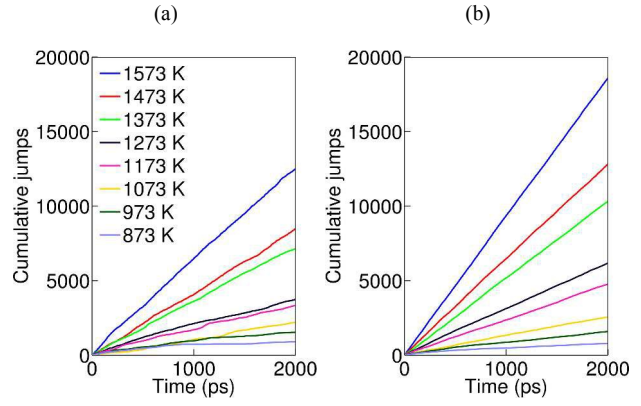

Fig. 8 Temperature dependence of cumulative jump times of oxygen from CoO_2 to CoO_2 layers (a) and CoO_2 to $PrO_{0.5}$ layers (b).

Since the oxygen diffusion is a thermally-activated process, the temperature dependence can be studied by the Arrhenius relation. Previously, it was shown that the oxygen ion conductivity could be calculated according to the equation[47]

$$\sigma = N_0 \frac{q^2}{k_B T} \gamma c(1-c) a_0^2 v_0 exp\left(-\frac{\Delta H_m}{k_B T}\right) \quad (4)$$

where N_0 is the number of equivalent sites per volume, c is the site occupancy fraction, q is the charge of the particle, γ is related to entropy,[47] a_0 is the distance between two equivalent sites, v_0 is a characteristic lattice frequency, ΔH_m is the migration enthalpy, k_B is the Boltzmann constant, and T is the temperature. In analogy to **(4)**, the total number of jumps can be written as follows:

$$\text{Total number of jumps} \propto N_s N_e c_s (1-c_e) exp\left(-\frac{\Delta H_m}{k_B T}\right) \quad (5)$$

where the subscript s and e represent the starting sites and ending sites, and N_s and N_e are the numbers of equivalent sites per formula unit. Site fractions at different temperatures can be obtained from **Fig. 7b**. The total number of hops between select layers are shown as a function of temperature in **Fig. 9a**.

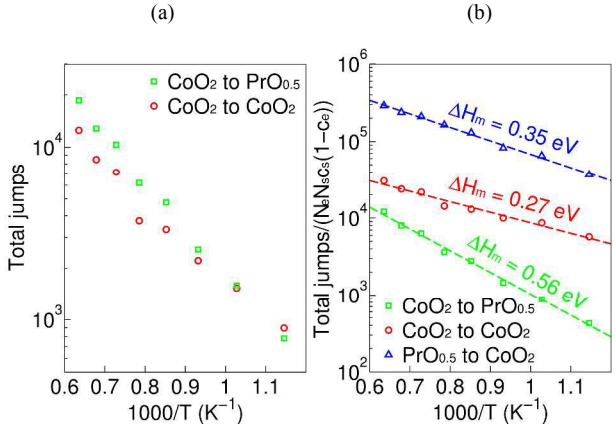

Fig. 9 Total number of oxygen jumps plotted against the reciprocal of temperature (a) and the Arrhenius representation (b).

The total number of jumps throughout versus the reciprocal of temperature roughly follows an exponential relation. If the total number of jumps is divided by the factor $N_s N_e c_s (1-c_e)$, the migration enthalpy between layers can be analyzed, as shown in **Fig. 9b**. The CoO_2 in-plane oxygen migration enthalpy is calculated to be 0.27 eV, which is much smaller than that of the out-of-plane migration 0.56 eV and the $PrO_{0.5}$ to CoO_2 migration, 0.35 eV. Experimental activation energy was determined to be 0.48 eV by fitting the Arrhenius relation of the oxygen diffusion coefficient.[39] This value is in the range of the computational results, suggesting that all 3 types of hopping events contributed to it. The in-plane transport has the lowest activation energy, which again evidences the fact that oxygen diffusion in this material is anisotropic and is constraint to the ab plane. For the oxygen migration from CoO_2 to $PrO_{0.5}$ planes, the activation energy of the forward process is much higher than the backward process. This is because oxygen vacancies tend to stay in the $PrO_{0.5}$ planes due to the lower energies,[45] thus moving oxygen to

this plane is more difficult than moving oxygen out of it.

Another quantity that can be obtained from the clustering analysis is the oxygen dwell time at certain sites. Analyzing the dwell times of atoms at a given site helps explain the interaction between oxygen and its neighbors. This can be particularly useful for studying the effects of dopants and local cation distributions on oxygen diffusion. The probability distribution function (PDF) of the dwell time and the cumulative distribution function (CDF) can be obtained by the Gaussian kernel density estimation with automatic determination of the kernel bandwidth.[48] As shown in **Fig. 10**, most oxygen atoms only stay at certain locations for less than 50 ps. The density decreases exponentially with the dwell time, suggesting that the oxygen atoms are quite mobile in these two planes. In addition, the different slopes for oxygen in these two planes are indicative of different oxygen mobilities; a higher mobility of oxygen is observed in the $PrO_{0.5}$ planes. This is also evidenced by the inset of **Fig. 10**, where a higher density presents at low dwell time for oxygen in $PrO_{0.5}$ planes. The CDFs of the dwell time at different temperatures are shown in **Fig. 11.** With the decrease of temperature, the CDF of dwell time increases at a slower rate. At temperature less than 1273 K, many oxygen atoms stay in the CoO_2 planes longer than the simulation time, as shown in **Fig. 11b**. Generally, the CDF of the $PrO_{0.5}$ increases faster than the other planes at low dwell time, consistently with the probability density distribution in **Fig. 10**. This is also shown in **Fig. 12**, suggesting that the average dwell time of oxygen in the CoO_2 planes is higher than that in the $PrO_{0.5}$ planes. Since oxygen atoms stay in the $PrO_{0.5}$ layers for a shorter time and the oxygen migration enthalpy in the CoO_2 layers is the lowest, it is likely that in PBCO the CoO_2 layers are the main channels for oxygen diffusion and the $PrO_{0.5}$ layers work as oxygen vacancy reservoirs.

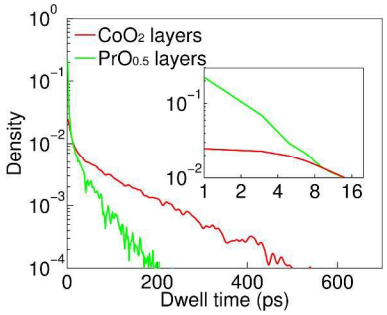

Fig. 10 PDF of the oxygen dwell time in CoO_2 and $PrO_{0.5}$ planes at 1573 K.

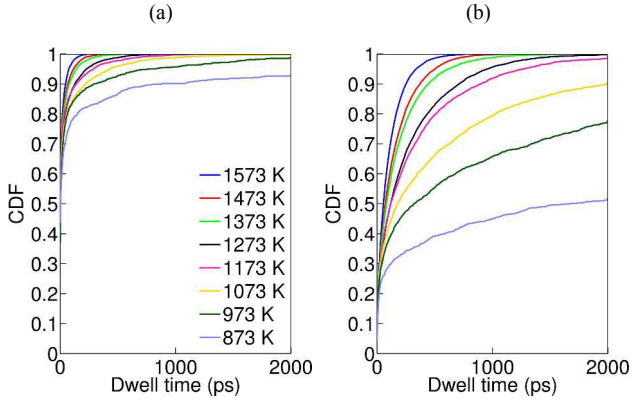

Fig. 11 CDF of oxygen dwell time in $PrO_{0.5}$ (a) and CoO_2 planes (b) at different temperatures.

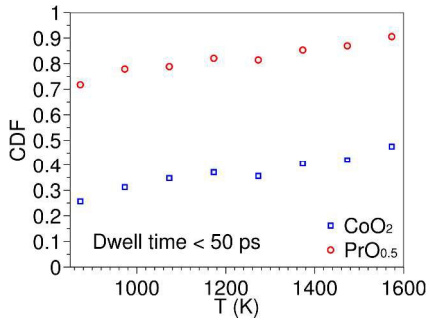

Fig. 12 CDF of dwell time at 50 ps.

Conclusions

An MD study of oxygen ion diffusion in A-site ordered perovskite PBCO is presented. The oxygen trajectories are analyzed by the k-means clustering, a machine learning algorithm, which captures the oxygen jumps between different locations in the material. This algorithm is chosen because the analysis of MD trajectories patterns is intrinsically an unsupervised learning problem. The clustering analysis yields a representation of an oxygen site as a simple label with a site type associated with it, and the oxygen hops between different sites may be interpreted as simple label changes. The clustering results also allow the analysis of individual jump events and types. Based on that, the occupation fraction change at different sites can be determined as a function of temperature. In the $PrO_{0.5}$ planes, the oxygen occupation fraction increases with temperature, while it decreases with temperature in the CoO_2 planes. The clustering analysis also leads to the separation of activation enthalpies of different migrations. It is found that among the diffusion paths the in-plane transport in CoO_2 has the lowest migration barrier (0.27 eV) and the transport from CoO_2 to $PrO_{0.5}$ has the highest barrier (0.56 eV). The dwell time of oxygen at different site types can also be analyzed. In the $PrO_{0.5}$ layers, oxygen atoms stay for shorter time than in the CoO_2 layers. By combining the dwell time analysis with the migration enthalpies, the oxygen migration is shown to be constrained to the CoO_2 layers, while the $PrO_{0.5}$ layers work as vacancy reservoirs. The dwell time analysis is particularly useful for studying the effect of the surrounding atoms on the oxygen transport. Even though the simulations presented here are based on empirical potential models, typical of classical MD, they can be in principle applied to kinetic Monte Carlo simulations, *ab initio* MD, and other types of MD simulations as well. This method is particularly suited for studying the impact of the local atomic arrangements on the ionic transport. Such analysis could be beneficial for the design of perovskites with enhanced conductivity by identifying diffusion bottlenecks and fast diffusion routes. Future theoretical work might use the k-means clustering based hopping analysis and dimensionality reduction methods to design high fidelity reduced models, such as kinetic Monte Carlo models, and therefore extend the simulation time.

Acknowledgments

The authors gratefully acknowledge HKUST for providing start-up funds, and the Research Grants Council of Hong Kong for support through the projects DAG12EG06 and ECS 639713.

Notes and references

[a] *Department of Mechanical Engineering, The Hong Kong University of Science and Technology, Hong Kong, SAR China*
[b] *Department of Chemical and Biomolecular Engineering, The Hong Kong University of Science and Technology, Hong Kong, SAR China*
** Corresponding author, tel.: +852 2358 7187, e-mail address: francesco.ciucci@ust.hk;*

1. A. Chroneos, B. Yildiz, A. Tarancón, D. Parfitt and J. A. Kilner, *Energy Environ. Sci.*, 2011, 4, 2774-2789.
2. J. Sunarso, S. Baumann, J. Serra, W. Meulenberg, S. Liu, Y. Lin and J. Diniz da Costa, *J. Membr. Sci.*, 2008, 320, 13-41.
3. B. C. Steele and A. Heinzel, *Nature*, 2001, 414, 345-352.
4. M. Li, M. J. Pietrowski, R. A. De Souza, H. Zhang, I. M. Reaney, S. N. Cook, J. A. Kilner and D. C. Sinclair, *Nat. Mater.*, 2014, 13, 31-35.
5. R. Merkle and J. Maier, *Angew. Chem. Int. Ed.*, 2008, 47, 3874-3894.
6. C. B. Gopal and S. M. Haile, *J. Mater. Chem. A*, 2014, 2, 2405-2417.
7. F. Ciucci, *Solid State Ionics*, 2013, 239, 28-40.
8. M. Burriel, J. Peña-Martínez, R. J. Chater, S. Fearn, A. V. Berenov, S. J. Skinner and J. A. Kilner, *Chem. Mater.*, 2012, 24, 613-621.
9. Y. Yamamura, S. Kawasaki and H. Sakai, *Solid State Ionics*, 1999, 126, 181-189.
10. C. Fisher, M. Yoshiya, Y. Iwamoto, J. Ishii, M. Asanuma and K. Yabuta, *Solid State Ionics*, 2007, 177, 3425-3431.
11. Z. Shao and S. M. Haile, *Nature*, 2004, 431, 170-173.
12. C. Chen, D. Chen, Y. Gao, Z. Shao and F. Ciucci, *J. Mater. Chem. A*, 2014, 2, 14154-14163.
13. J. Hermet, G. Geneste and G. Dezanneau, *Appl. Phys. Lett.*, 2010, 97, 174102.
14. D. Parfitt, A. Chroneos, A. Tarancón and J. A. Kilner, *J. Mater. Chem.*, 2011, 21, 2183-2186.
15. C. Liu, J. Zhang, L. Wang, Y. Shu and J. Fan, *Solid State Ionics*, 2013, 232, 123-128.
16. I. Seymour, A. Tarancon, A. Chroneos, D. Parfitt, J. Kilner and R. Grimes, *Solid State Ionics*, 2012, 216, 41-43.
17. A. Chroneos, D. Parfitt, J. A. Kilner and R. W. Grimes, *J. Mater. Chem.*, 2009, 20, 266-270.
18. Y. Hu, O. Hernandez, T. Broux, M. Bahout, J. Hermet, A. Ottochian, C. Ritter, G. Geneste and G. Dezanneau, *J. Mater. Chem.*, 2012, 22, 18744-18747.
19. D. Parfitt, A. Chroneos, J. A. Kilner and R. W. Grimes, *Phys. Chem. Chem. Phys.*, 2010, 12, 6834-6836.
20. A. Kushima, D. Parfitt, A. Chroneos, B. Yildiz, J. A. Kilner and R. W. Grimes, *Phys. Chem. Chem. Phys.*, 2011, 13, 2242-2249.
21. A. Tarancón, M. Burriel, J. Santiso, S. J. Skinner and J. A. Kilner, *J. Mater. Chem.*, 2010, 20, 3799-3813.
22. A. Taskin, A. Lavrov and Y. Ando, *Appl. Phys. Lett.*, 2005, 86, 091910-091910-091913.
23. M. S. Islam, *J. Mater. Chem.*, 2000, 10, 1027-1038.
24. K.-D. Kreuer, S. J. Paddison, E. Spohr and M. Schuster, *Chem. Rev.*, 2004, 104, 4637-4678.
25. A. Al-Yasari, A. Jones, A. Orera, D. C. Apperley, D. Driscoll, M. S. Islam and P. R. Slater, *J. Mater. Chem.*, 2009, 19, 5003-5008.
26. M. Khan, M. Islam and D. Bates, *J. Phys. Chem. B*, 1998, 102, 3099-3104.
27. D.-S. Byeon, S.-M. Jeong, K.-J. Hwang, M.-Y. Yoon, H.-J. Hwang, S. Kim and H.-L. Lee, *J. Power Sources*, 2013, 222, 282-287.
28. M.-Y. Yoon, K.-J. Hwang, D.-S. Byeon, H.-J. Hwang and S.-M. Jeong, *J. Power Sources*, 2014, 248, 1085-1089.
29. Y. Wang, M. Klenk, K. Page and W. Lai, *Chem. Mater.*, 2014, 26, 5613-5624.
30. W. M. Brown, S. Martin, S. N. Pollock, E. A. Coutsias and J.-P. Watson, *J. Chem. Phys.*, 2008, 129, 064118.
31. S. Hayward and N. Go, *Annu. Rev. Phys. Chem.*, 1995, 46, 223-250.
32. S. T. Roweis and L. K. Saul, *Science*, 2000, 290, 2323-2326.
33. J. B. Tenenbaum, V. De Silva and J. C. Langford, *Science*, 2000, 290, 2319-2323.
34. M. Ceriotti, G. A. Tribello and M. Parrinello, *Proc. Natl. Acad. Sci.*, 2011, 108, 13023-13028.
35. R. R. Coifman, I. G. Kevrekidis, S. Lafon, M. Maggioni and B. Nadler, *Multiscale Model. Simul.*, 2008, 7, 842-864.
36. A. L. Ferguson, A. Z. Panagiotopoulos, P. G. Debenedetti and I. G. Kevrekidis, *Proc. Natl. Acad. Sci.*, 2010, 107, 13597-13602.
37. M. A. Rohrdanz, W. Zheng and C. Clementi, *Annu. Rev. Phys. Chem.*, 2013, 64, 295-316.
38. A. Cammarata, A. Emanuele and D. Duca, *Chem. Asian J.*, 2012, 7, 1838.
39. G. Kim, S. Wang, A. Jacobson, L. Reimus, P. Brodersen and C. Mims, *J. Mater. Chem.*, 2007, 17, 2500-2505.
40. D. Chen, R. Ran, K. Zhang, J. Wang and Z. Shao, *J. Power Sources*, 2009, 188, 96-105.
41. I. Seymour, A. Chroneos, J. Kilner and R. Grimes, *Phys. Chem. Chem. Phys.*, 2011, 13, 15305-15310.
42. S. Plimpton, *J. Comput. Phys.*, 1995, 117, 1-19.
43. Y. Gao, D. Chen, C. Chen, Z. Shao and F. Ciucci, *J. Power Sources*, 2014, 278, 623-629.
44. M. T. Buscaglia, V. Buscaglia, M. Viviani and P. Nanni, *J. Am. Ceram. Soc.*, 2001, 84, 376-384.
45. J. Hermet, B. Dupe and G. Dezanneau, *Solid State Ionics*, 2012, 216, 50-53.
46. D. Steinley, *Br. J. Math. Stat. Psych.*, 2006, 59, 1-34.
47. J. Kilner, *Solid State Ionics*, 2000, 129, 13-23.
48. Z. Botev, J. Grotowski and D. Kroese, *Ann. Stat.*, 2010, 38, 2916-2957.