2021

Assessing the Accuracy of Quantum Monte Carlo Pseudopotentials for CO2 Capture in Metal Organic Frameworks

Chloe Renfro
Bowdoin College

Assessing the Accuracy of Quantum Monte Carlo Pseudopotentials for CO_2 Capture in Metal Organic Frameworks

An Honors Paper for the Department of Chemistry

By Chloe Renfro

Bowdoin College, 2021

Table of Contents

List of Figures

List of Tables

List of Files

Acknowledgements

I am extremely grateful to my research advisor, Professor Allison Dzubak, for her guidance and constant assistance on my many research and computing grid woes. Despite the cancellation of summer research and the difficult research conditions thanks to Covid, we persevered. I would also like to thank DJ Merrill without whose assistance on the computing grid, many errors would have likely gone unresolved. I am also appreciative of Professor Jeffrey Nagle and Professor Richard Broene for their comments and useful suggestions on my paper.

Abstract

As global emissions of CO_2 and other greenhouse gases rises, global warming persists as an imminent threat to the environment and every day lives. To reduce greenhouse gas emissions in the atmosphere, there is a need to design materials to separate and capture the different gasses. Current gas capturing technologies lack efficiency and have extensive energy costs. A class of materials for CO_2 capture is Molecular Organic Frameworks (MOFs). In order for a MOF to be efficient for this type of separation, the MOF needs to be able to selectively bind to the gas, while also not suffering a high energy cost to remove the gas and reuse the material. Computationally calculated binding energies are used to determine the usefulness of a MOF at capture and separation of a certain gas. Each computational method has its advantages and limitations. In this work, diffusion quantum Monte Carlo is being explored. This paper focuses on the accuracy of recently developed pseudopotentials for DMC use. These pseudopotentials have been tested on smaller molecules but have not been systematically tested for systems such as MOFs. Results from a DMC calculation of Zn-MOF-74 show a binding energy of -18.02 kJ/mol with an error bound of 16.74 kJ/mol. In order to assess the accuracy of the DMC results for binding energies of this magnitude the uncertainty need to be reduced, a subject of ongoing work.

1. Introduction and Background

Climate change is a prevailing issue that has huge impacts on the planet and future lives. The reduction of mountain snow in the western US, drier summers leading to more wildfires, [1] and an increase in respiratory diseases [2] have all been discovered as negative effects of an increase of greenhouse gasses. A leading cause of climate change is the increase in greenhouse gasses, such as carbon dioxide in the atmosphere. Figure 1 shows the increase of greenhouse gas emissions from 1970-2020. [3] There is a clear and steady increase of carbon dioxide in the air, which can be attributed to human activity. A 2019 report has shown that although annual there has been a rising level of global consciousness for CO_2 emissions, the overall levels of CO_2 are still rising in the atmosphere. [4] The gasses trap heat from the sun in the atmosphere leading to an overall temperature increase of the earth's surface, known as global warming. Global warming is a clear threat to human lives and has led to more intense weather phenomena, as well as more extreme temperatures. This has disastrous consequences on ecosystems and daily life.

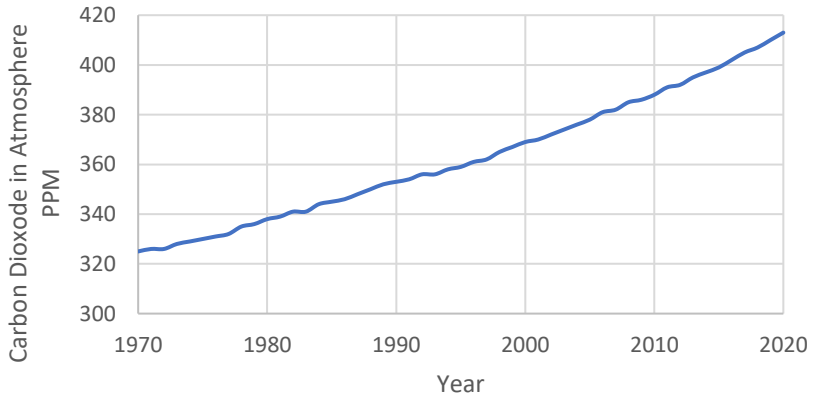


Figure 1: A graph of carbon dioxide emissions in the atmosphere from 1970 to 2020. Data from the National and Atmospheric Administration. [3]

To combat global warming, many technologies have been invented and implemented to capture carbon dioxide in the air. One of these methods is underground sequestration of carbon dioxide. Carbon dioxide is captured through geological means, in sediment beds and reservoirs, and stored underground. It has been suggested that nearly 40% of atmospheric carbon dioxide could be captured in this manner. The method is still being studied and developed. [5] However, a major contributor to carbon dioxide emissions comes from factories and smoke stacks. Underground carbon dioxide sequestration technologies are not equipped to address this release, but underground storage is still a viable option once the carbon dioxide has been captured. [5]

The prevalent method for gaseous carbon dioxide capture from factories, powerplants, and smoke stacks in general, is chemical absorption. The effluent gas is introduced to a liquid solution that absorbs the carbon dioxide from the gas mixture. The components of the gas mixture determine how long the treatment takes. Additionally, the absorbent must be customized to be able to separate the carbon dioxide from the other gasses. After the carbon dioxide is removed from the gas mixture, it is separated from the absorbent. The carbon dioxide is then stored to be transported and the absorbent is regenerated and then is able to be reused in the gas separation process. Figure 2 shows a flow chart of the chemical absorption process. [6] Although able to successfully capture and release CO_2, this method is not very common. It is costly and inefficient to capture carbon dioxide this way, due to the energy needed, and thus is not widely deployed. A major contributor to the energy cost can be attributed to heating the aqueous solution to regenerate the absorbent. One of the problems being explored reduces the energy required to regenerate the absorbent. Porous materials that can be used as adsorbents are a potential solution.

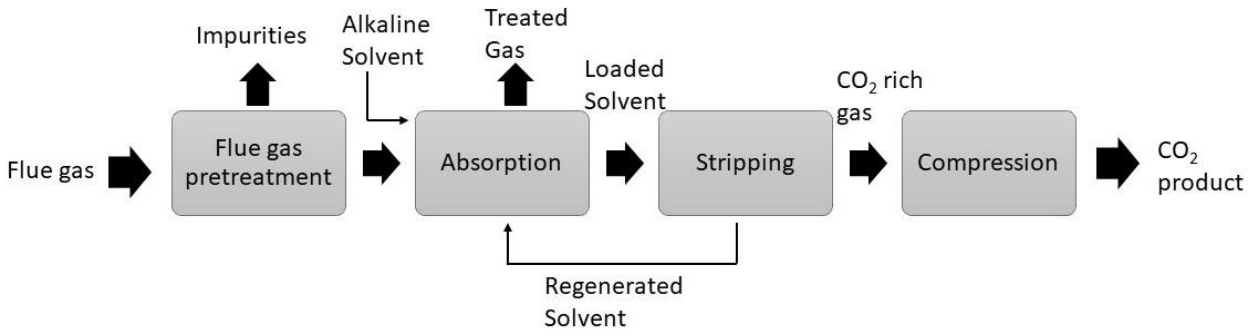
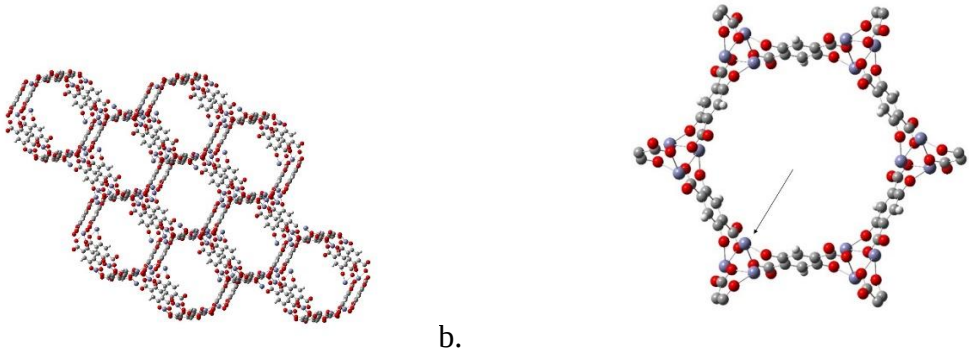


Figure 2: The chemical absorption process for removing carbon dioxide from the air with a liquid absorbent. Adapted from Leung et al. [6]

New materials are being explored for their ability to capture carbon dioxide before it is released into the atmosphere. A subset of these materials are called metal organic frameworks, or MOFs. MOFs consist of repeating units of a metal center, bridged with organic molecules. MOFs are being explored because they have the ability to be optimized for gas capture and separation, by customizing different topologies, metals, and organic linkers. MOFs are particularly promising because of their potential optimization, due to the vast number of MOFs that can be made, as well as their practicality, efficiency, and reusability.

A promising MOF for CO_2 capture is MOF-74, or M_2(dobdc) where M is a metal and dobdc = 2,5-dioxido-1,4-benzenedicarboxylate; it is shown in Figure 3. MOF-74 has been thoroughly characterized through diffraction experiments and its adsorption of carbon dioxide has been extensively studied. It was found that carbon dioxide primarily adsorbs to the transition metal sites in a nearly linear structure, with almost 180 degree bond angles in the carbon dioxide molecule during the adsorption process. [7] Knowing the adsorption sites allows for customization and gives greater insight into adsorption mechanisms. MOF-74 was also shown to be highly selective in adsorbing carbon dioxide. Figure 4 shows the effectiveness of Mg-MOF-

74 in selectively capturing carbon dioxide over methane [8] Based on the figure, Mg-MOF-74 has been able to be optimized to selectively capture carbon dioxide over methane.



a. b.

Figure 3: Structure of MOF-74. The transition metal site is represented in blue, carbon in gray, oxygen in red, and hydrogen in white. Figure **a** shows the porous MOF material. Figure **b** illustrates a 2 dimensional slice of the MOF, the exposed transition metal sites (arrow) provide the primary adsorption sites for carbon dioxide.

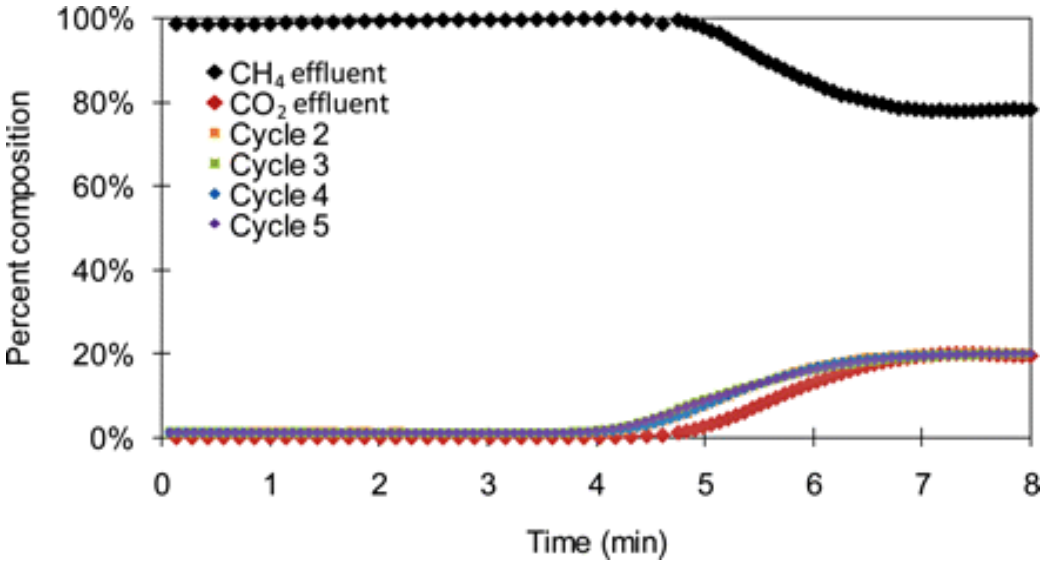


Figure 4: The adsorption of CO_2 vs CH_4 by the Mg-MOF-74 as a function of time. Figure from Britt et al. [8]

Zn_2(dobpdc), an extended linker of Zn-MOF-74 where dobpdc is 4,4'-dioxidobiphenyl-3,3'-dicarboxylate, has been examined for its ability to capture carbon dioxide. [9] This MOF was shown to successfully separate carbon dioxide from a mixture of gasses. Zn_2(dobpdc) illustrates the potential effectiveness of MOFs for capturing greenhouse gasses. One of the promising features of MOFs is that the materials can be optimized to capture different gasses. This key property lies in the adsorption sites within the molecules. [10] It has been shown that the transition metal sites, which are the primary adsorption sites, dictate how well gases are absorbed and separated. The attractive forces between the transition metal and the carbon dioxide molecule are the key to the adsorption and separation processes. [10] MOFs have been shown to be strong candidates for carbon dioxide capture in the air, due to their ability to be customized. Tuning adsorption of the MOF can be accomplished through a judicious choice of the transition metal; to design and tune these materials, computational modeling can be used.

Computational modelling, which uses computers and mathematical algorithms to model molecules and reactions, can be used to model MOFs. Due to the magnitude of possible MOFs available, computational modelling can save on both time and cost when determining if the desired molecule will selectively adsorb carbon dioxide. Since the metal sites are the primary adsorption sites, these centers are where computational calculations are focused. [10] The computational modelling of MOFs has been explored for decades, using different modelling methods and theories. Density functional theory has been the main focus of Quantum chemical MOF modelling thus far. [11]

Density functional theory (DFT) is the most common quantum chemical method for modelling of MOFs. [11] DFT is a method that treats the individual electrons as an electron density cloud to solve the Schrödinger equation. Although formally DFT is an exact method, the

exact exchange-correlation functional is unknown, so the development of various approximate functionals is an active area of research. It is generally unknown how a given functional will perform until it has been rigorously benchmarked. The systems studied have many different electrons that interact with each other, but DFT treats these as an electron density, rather thanconsidering explicit electron-electron correlations. [12] This treatment of electrons causes errosr due to missing electron-electron interactions.

It has been shown that some DFT methods may not be well-suited for MOFs with coordinatively unsaturated transition metal sites, such as in MOF-74. [13] Table 1 shows select DFT calculated CO_2, N_2, and H_2O adsorption enthalpy energies in $Cu(HCOO)_2$, copper formate tetrahydrate, of different functionals. [14] The functional can have a great impact on the calculated energy, and so the choice of functional and other parameters is extremely important in determining energies. However, the corrections use empirical parameters, which rely on experimental data to make the necessary corrections and choose the correct functional. These experimental-based corrections cause DFT to lose its predictive power. Other studies have shown that DFT calculated binding energies for O_2/N_2 in $M_3(btc)_2$, where btc is benzene-1,3, 5-tricarboxyalte, and M_2(dobdc) can differ from experimental values by nearly 300 kJ/mol. [15] With modern, high performance, massively parallel, super computers, systematically improvable methods, such as quantum Monte Carlo (QMC), can be used to model systems such as MOFs. These methods are more computationally demanding, so have not been plausible for larger molecules and materials until recent high powered computing machines have become more widely available. Diffusion quantum Monte Carlo (DMC) has been shown to overcome some of the limitations faced by DFT for modelling MOFs. [16] Diffusion quantum Monte Carlo is a method very recently being explored for gas adsorption in MOFs, but recently a few studies have

been completed, showing its possible accuracy. A 2017 study has shown that Quantum Monte Carlo can be used to successfully model MOFs, through studying CO_2 adsorption in M_2(dobdc) MOFs, where M represents Mg, Mn, Fe, Co, Ni, Cu, and Zn. [17]

Table 1: Differences between DFT calculated and experimentally figured CO_2, CO, and N_2 adsorption enthalpies for Mg-MOF-74. Table adapted from Grajciar et al. [14]

Functional	SVWN5	PBE	RPBE	PBE0	M05-2X	B2PLYP	BLYP-D3	optPBE
CO_2	-9.8	-3.6	0.9	-5.2	-10.8	-6.3	-6.9	-10.9
H_2O	-33.0	-14.9	-4.4	-21.8	-36.3	-22.8	-20.4	-22.2
N_2	-8.9	-3.0	1.2	-4.0	-8.0	-4.7	-7.7	-9.2

Diffusion quantum Monte Carlo is an exact method, but we also rely on approximations in its implementation. As an explicitly correlated method, DMC does not require empirical corrections to account for dispersion interactions. [18] The three major approximations are the fixed node approximation, the usage of pseudopotentials, and the locality approximation. [19] Pseudopotentials mimic the potential in the bonding electrons, but replace the potential near the nucleus, which simplifies the calculations. Figure 6 shows a representation of the pseudopotential energy versus the actual energy for a non-descript molecule. [20] A solution that was explored to address the DFT failings was to use more accurate pseudopotential sets, because it can correct an error introduced by the 3d shells of transition metals. [21] Similar approaches are being explored for DMC. In DMC pseudopotentials become a primary area where computational errors can be introduced. Introducing non-local pseudopotentials causes a localization error for 3d transition metals. [19] It has been shown that pseudopotentials can be

designed that reduce this error to nearly 0 for Cerium, and thus expected to be able to expand to transition metals. [22]

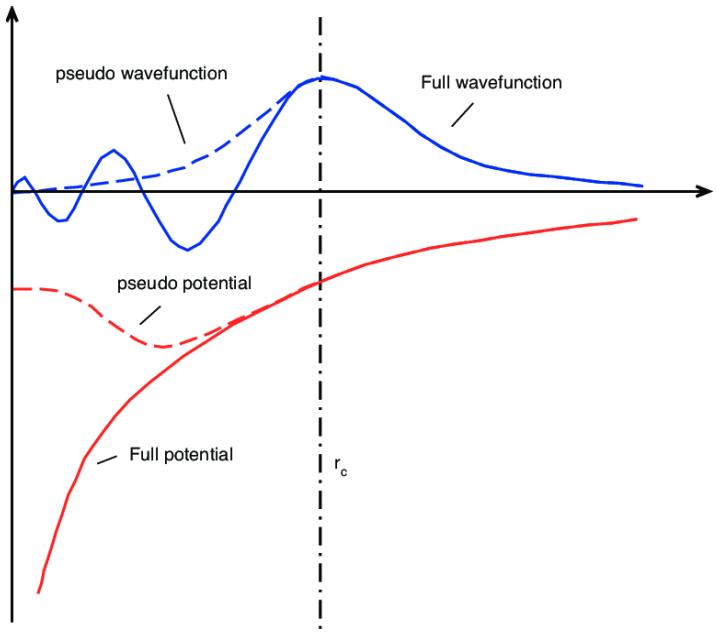


Figure 6: Illustration of a pseudopotential in which closer to the nucleus than some cutoff value, r_0, the full potential is replaced by a pseudopotential. The x-axis is the distance from the nucleus and the y-axis represents energy. A pseudopotential simplifies the electron energy near the nucleus while retaining the potential at the bonding electrons. This simplifies calculations by introducing an approximation. Figure from Rasander [19]

QMC pseudopotential development is an active area of research. Pseudopotentials are being developed specifically for use in DMC calculations. Newly developed pseudopotentials have been benchmarked for the accuracy on small molecules, but their accuracies are unknown for materials like MOFs. [23]

This paper focuses on assessing the accuracy of pseudopotentials for DMC calculations of MOF-CO_2 binding energies. Using MOF-74, calculated CO_2 binding energies will be

compared to experimental values to determine the accuracy of the DMC calculations. The 3d transition metal series will be explored, starting with Zn-MOF-74. Having an accurate method with pseudopotentials that do not introduce much of an error, will allow for more MOFs to be optimized for gas separation and carbon dioxide capture. The adsorption energy will be calculated because the MOF needs to have a balance of selectivity and energy efficiency. If the energy is too high, then recovering the carbon dioxide from the MOF will cost too much energy. If the energy is too low, the MOF will not efficiently capture the desired gas. Thus, calculating adsorption energies allows one to see if the particular MOF is suitable.

2. Methods

2.1 CO_2 Binding in MOF-74

The binding energy of CO_2 to the coordinatively unsaturated transition metal of the MOF is calculated as:

$$E_{binding} = E_{MOF+CO2} - E_{MOF} - E_{CO2} \quad (1)$$

The structure of the MOF with CO_2 was ascertained from crystal structure data, for both Zn-MOF-74 [7] and Cu-MOF-74. [7] A model cluster was taken from this structure, by focusing on only one CO_2 binding location, and capping the peripheral truncated structures with hydrogen atoms (maintaining original bond angles). Figure 7 shows the final model cluster used for Zn-MOF-74. For Cu-MOF-74, the copper atoms that were not interacting with the oxygen were replaced with zinc. Zinc was chosen because of its similar size to copper and its full d-electron shell.

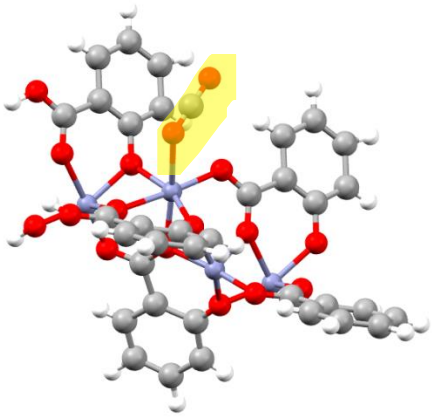


Figure 7: Model cluster of Zn-MOF-74 with CO_2 binding to the coordinatively unsaturated metal site. Zinc is represented in blue, carbon in gray, oxygen in red, and hydrogen in white.

2.2 Generation of Trial Wavefunctions

Initial trial wavefunctions were generated at the DFT level using the PBE0 functional [24] with GAMESS software. Since the DFT software was used to generate trial wavefunctions, the accuracy of the energy was not highest priority, so the default functional was sufficient. [25] Recently developed pseudopotentials were employed (carbon, hydrogen, oxygen) [26] (zinc) [27]. A basis set scan was preformed from the cc.VDZ to the cc.aug.V6Z level (carbon, hydrogen, oxygen) [26] (zinc, copper) [27]. Calculations were performed by using 8 cores on the Bowdoin High Performance Computing grid.

2.3 Wavefunction Optimization

In order to be used for DMC calculations, the output file of the DFT calculations, which gave the DFT calculated wavefunction, had to be converted to a format readable by the QMCPACK software. [28] Convert for QMC, a built in QMCPACK function, was used for this purpose. From the conversion, an xml structure file was created, along with the DFT generated wavefunction and input file for DMC calculations. The input file used for these calculations includes the Jastrow factor. [29] The Jastrow factor is used because it accounts for the interactions of moving electrons. Using only the optimization blocks of the input file provided by the convert for qmc, the wavefunction is optimized. In the example initial VMC block, the qmc method is VMC, variational Monte Carlo. The move is pbyp, particle by particle. Under estimator name, hdf5 is set to “no”, which means that the wavefunction is in xml form, not in an hdf5 file. The warmupSteps, used for equilibration, are set to 100. The blocks, which represent the number of VMC initializations run, is set to 20. The steps, which are measurements per block, are set to 50, with 8 sub-steps. Sub-steps move the electrons once, but does not evaluate the energy at the end of the sub-step. The timestep is set to 0.5, referencing the statistical

efficiency, and the usedrift is set to “no”, which means that a gaussian distribution is used to move the electrons. Under the example initial VMC optimization section, the loop max is set to 4, which means that the optimization will be run at most 4 times. The qmc method is linear, which is the optimization method, with the particle by particle move type. The hdf5 is set to “no”, the warmupSteps is 100, the blocks are 20, the timestep is 0.5, the substeps is 4, and usedrift is “no”. The walkers, which is the walkers used per MPI, is 1. The samples are set to 16000. The MinMethod is OneShiftOnly with the minwalkers set to 0.0001. The OneShiftOnly speeds optimization by taking its sample if the effective weight is greater than the minwalkers. The follow-up VMC optimization has a loop max of 10. The qmc method is linear with a particle by particle move. The filetype is xml, with 100 warmupSteps, 20 blocks, 4 substeps a 0.5 time step and no usedrift. The samples are set to 64000, with 1 walker per GPU, the OneShiftOnly MinMethod and a minwalker of 0.3.

2.4 DMC calculations

To perform the DMC calculations, the optimized wavefunction that has the lowest energy is chosen. The structure and input file from the conversion procedure are used, along with the same cc.ECP files. From the input file, the optimization blocks are deleted, and only the active DMC blocks are kept. The example initial VMC measuring block is the same as it was for the optimization section. For the Production VMC and DMC section, first the VMC production block runs, which produces the VMC energy results. The parameters are the same as they were for the example initial VMC measuring block, with an added sample of 16000. The DMC calculations, the qmc method was set to dmc, with a particle by particle move and a checkpoint of 20. The hdf5 is set to “no”, so the wavefunction is given in xml format. The targetwalkers, which is the total number of walkers used in the DMC production, is set to 1600. The

reconfiguration is set to “no”, so the reconfiguration technique was not used. The warmupSteps is set to 100, with a timestep of 0.005, run with 100 steps and 100 blocks. The *nonlocalmoves* is set to yes, so the pseudopotentials are evaluated with a nonlocal move algorithm. [30] For DMC calculations, the number of cores effects the speed and completeness of the calculations. For the CO_2 molecule, 8 cores were used and the calculations went to completion. For the MOF, 120 cores were used and only 62 of 100 blocks of the DMC calculations were completed. For the MOF and CO_2, 130 cores were used and only 58 blocks were completed. Due to computing grid restrictions, tests on higher number of cores were not able to be completed at this time.

3. Results and Discussion

To calculate the binding energies, energies must be calculated for the MOF and CO_2, the MOF alone, and CO_2 alone (Eq. 1). Zn-MOF-74 and Cu-MOF-74 are the two MOFs used for this paper. In the process of generating trial wavefunctions, binding energies at the DFT level can be compared with energies calculated at the DMC level, and experimental values.

3.1 CO_2 Molecule

3.1.1 CO_2 Molecule DFT Calculations

The DFT calculated energies for CO_2 are shown in Table 2. Figure 8 shows the basis set vs time and Figure 9 shows the basis set vs calculated energy. Based on these figures, cc.VQZ is likely the best basis set to use for CO_2. Similar comparisons will be made for the MOF calculations. From Figure 9 it is easily seen that the energy is nearing the convergence point for the basis sets, so not much accuracy is lost since the energy difference between cc.VQZ and aug.cc.V6Z is minimal. Table 2 shows that the difference in energy is only approximately 2.09 kJ/mol. This is very minimal as opposed to the 71.2 kJ/mol difference between the cc.VDZ and aug.cc.V6Z basis sets. From Table 2, it can be seen that there is however a huge computational gain from choosing cc.VQZ since the computational time is 5.47% of the aug.cc.V6Z basis set. Thus the cc.VQZ basis set is a reasonable balance between computational cost and accuracy. This is an important consideration for the larger MOF systems discussed subsequently for which calculations involving the larger basis sets are not feasible.

Table 2: Energy (E_h), Energy difference of the aug.cc.V6Z basis set (kJ/mol), and computational time (s) of 9 different basis sets, used for a DFT calculation of CO_2.

Number of Basis Functions	Basis Set	Energy (E_h)	Energy Difference (kJ/mol)	Time (s)
42	cc.VDZ	-37.76102556	71.196082	2.3
72	aug.cc.VDZ	-37.77199439	42.397417	4.2
102	cc.VTZ	-37.78322468	12.912297	9.6
162	aug.cc.VTZ	-37.7858050	6.1375602	27.0
207	cc.VQZ	-37.78734587	2.0921228	48.5
312	cc.V5Z	-37.78769042	1.187505	160.9
312	aug.cc.VQZ	-37.78776441	0.9932416	182.9
417	aug.cc.V5Z	-37.787785	0.9391757	463.6
417	cc.V6Z	-37.788112	0.080042	424.5
522	aug.cc.V6Z	-37.78814271	0	887.1

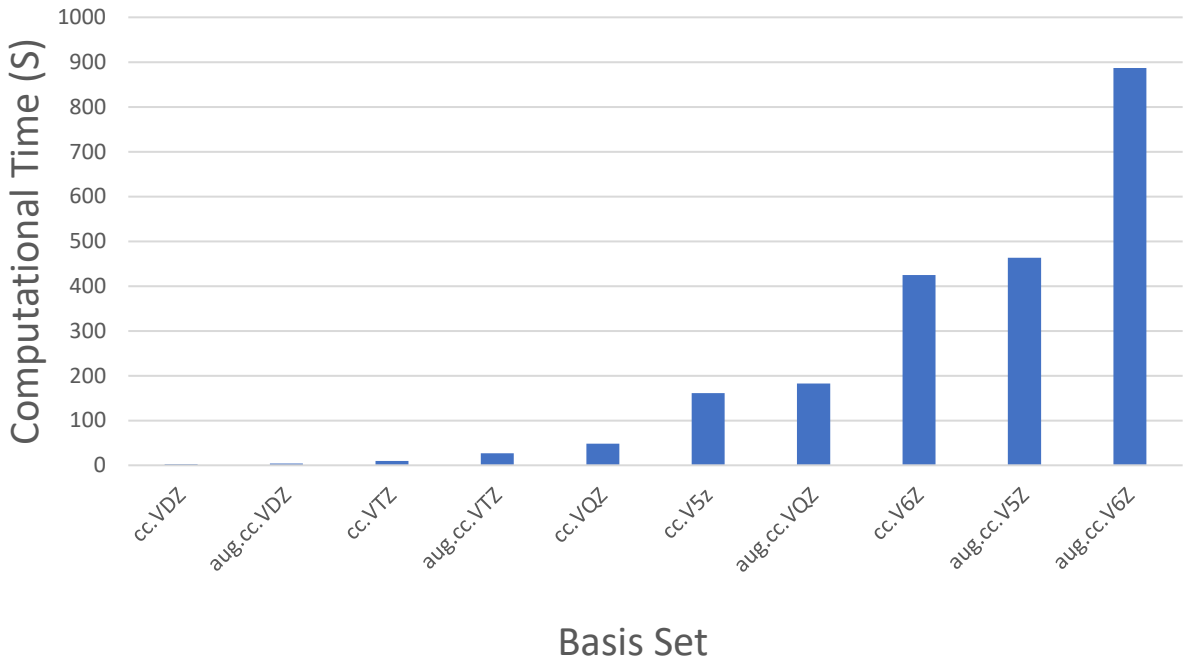


Figure 8: The computational time (s) as a function of the size of basis set employed, for a DFT calculation of CO_2.

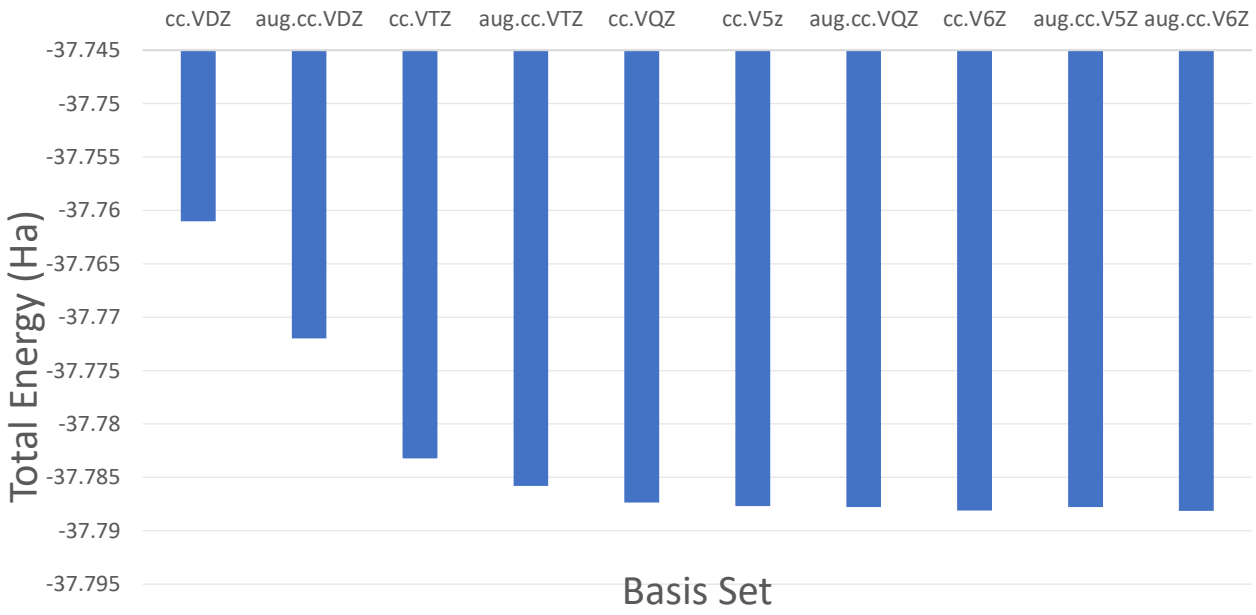


Figure 9: Total Energy (E_h) as a function of the size of the basis set employed for a DFT calculation of CO_2.

3.1.2 CO_2 Molecule DMC Calculations

For CO_2, the DFT calculated wavefunctions for the cc.VDZ and cc.VTZ were optimized and then used for DMC calculations. Table 3 shows the DMC calculated CO_2 energies with error bounds.

Table 3: Energy (E_h) of cc.VDZ and cc.VTZ basis sets, used for a DMC calculation of CO_2.

Basis Set	Energy (E_h)	Error (E_h)
cc.VDZ	-37.761833	0.000404
cc.VTZ	-37.767865	0.000393

3.2 Zn-MOF-74

3.2.1 Zn-MOF-74 DFT Calculations

The DFT energies of the Zn-MOF-74 basis sets were calculated in the process of generating the wavefunctions. The aug.cc.V6Z and cc.V6Z basis sets were not available for the zinc atom. Due to system restraints, the cc.V5Z and aug.cc.V5Z basis sets exceed the computing grid memory bounds. Because the computational running time is directly related to the number of basis functions, which increases with the amount of atoms in the molecule, larger molecules have longer calculation running times. So far, calculations have been completed only for the cc.VDZ and cc.VTZ basis sets. To calculate the binding energies, the wavefunctions and corresponding energies for both the MOF with and without CO_2 were calculated. Table 4 (with CO_2) and Table 5 (without CO_2) show the results, and the difference between the basis set and lowest energy basis set (cc.VTZ). The difference is not-minimal, being 983.5 kJ/mol and 933 kJ/mol, respectively. In our results, the basis set size has a huge impact on the accuracy of the

calculations, and it has been shown through other studies that choosing a basis set is important, due to the dependence of the calculated on energy on the basis set size. [31]

Table 4: Energy (E_h) and computational time (s) of basis sets used for DFT Zn-MOF-74 with CO_2 calculations

Number of Basis Functions	Basis Set	Energy (Harttree)	Energy Difference (kJ/mol)	Time (s)
1158	cc.VDZ	-1396.531814	983.46	78053
2478	cc.VTZ	-1396.906393	0	619652

Table 5: Energy (E_h) and computational time (s) of basis sets used for DFT Zn-MOF-74 calculations

Basis Functions	Basis Set	Energy (E_h)	Energy Difference (kJ/mol)	Time (s)
1116	cc.VDZ	-1358.767205	932.00	61830.8
2376	cc.VTZ	-1359.122562	0	345880.3

3.2.2 Zn-MOF-74 DFT Binding Energy Results and Discussion

Using Equation 1, the binding energy for Zn-MOF-74 can be calculated at the DFT level. Table 6 shows the resulting energies, in both E_h and kJ/mol, for the cc.VDZ and cc.VTZ basis sets. The accuracy of the DFT values are evaluated using a comparison with experimental isosteric heat of adsorption, *Q*, values. [7]

Table 6: Binding Energy (E_h and kJ/mol) of CO_2 with Zn-MOF-74 as a function of basis set at the DFT level of theory with experimental Q value comparison

Basis Set	Binding Energy (E_h)	Binding Energy (kJ/mol)	Experimental *Q* Value (kJ/mol)	Experimental and Theory Difference (kJ/mol)
cc.VDZ	-0.00358336	-9.41	-26.8	17.4
cc.VTZ	-0.01183607	-1.59	-26.8	25.2

As can be seen in Table 5, the DFT calculated binding energies are dependent upon basis set size. The energy changes by a factor of nearly 6 going from the cc.VDZ to the cc.VTZ basis set. The difference between the cc.VDZ binding energy and the experimental Q value is 17. kJ/mol. Whereas, the difference between the cc.VTZ and experimental values is 25.2 kJ/mol. We'd expect the binding energy to get more accurate (closer to experimental value) as the basis set size increases, because it more accurately describes the atoms behaviors. However, this is not the trend we observe. While the individual molecules energy gets lower as predicted, the binding energy was actually smaller for the cc.VTZ than the cc.VDZ. This can likely be attributed to a lack of basis set superposition error (BSSE) correction. [32] This correction was not applied because the DFT binding energies are not the focus of this research. When comparing each molecule energy for the two basis sets, it is seen that the CO_2 binding energy is more affected by increasing the basis set. The energy decreases by 0.059% from the cc.VDZ to the cc.VTZ basis set. However, for the MOF with CO_2 the decrease was only 0.027% and 0.026% for the MOF alone. So, the energy decreases by a factor of twice as much for CO_2 than the MOFs. For DFT calculations, the basis set has a much bigger impact on CO_2 than the MOF, presumably because

of the size of the molecules. The DFT binding energies were not extremely accurate, presumably due to the lack of correction factors.

3.2.3. Zn-MOF-74 DMC Calculations

For the cc.VDZ basis set, DMC calculations were performed for the MOF with and without CO_2. Table 7 summarizes the energy and error bounds.

Table 7: Energy (E_h) of the cc.VDZ basis sets used for DMC Zn-MOF-74 calculations

	Energy (E_h)	Error (E_h)
MOF with CO_2	-1396.303244	0.003679
MOF	-1358.534547	0.005191

3.3 Zn-MOF-74 DMC Binding Energy Results and Discussion

The binding energy for the cc.VDZ basis set of CO_2 and Zn-MOF-74 was calculated and Table 8 summarizes the results. The results are compared against experimental values.

Table 8: Binding Energy (E_h and kJ/mol) of CO_2 with Zn-MOF-74 at the DMC level of theory

	Binding Energy (E_h)	Error (E_h)	Binding Energy (kJ/mol)	Error (kJ/mol)	Experimental *Q* Value
cc.VDZ	-0.006864	0.006375	-18.02	16.74	-26.8

The DMC calculated binding energy was -18.02 kJ/mol with an error of 16.74 kJ/mol. The experimental Q value is within the error bounds of the DMC calculated binding energy. The DFT

calculated binding energy of the cc.VDZ basis set was -9.41 kJ/mol, so the DMC calculations lowered the value by nearly twice. For the cc.VDZ basis set, the DMC calculations improved the DFT results, and the binding energies error encapsulated the experimental Q. However, the error was almost as large as the actual binding energy, so the results are not as precise as desired. For DMC calculations, the number of cores directly correlates to the speed and completeness of the calculations. For our calculations, the jobs would run for 4 days before stopping. For CO_2, using 8 cores ensured that the calculations completed all 100 blocks specified by the input file. However, for the MOF calculations, 8 cores were not sufficient. For the MOF alone, 120 cores were used, and the DMC calculations were able to finish 62 blocks before the calculations stopped due to time restrictions. For the MOF+CO_2, 130 cores were used and 58 blocks were completed. To minimize the binding energy and error bounds, the calculations need to be run with more cores. Due to current computing restrictions, this was not possible to test for this paper. However, using more cores will lower the energy and error bound for the individual calculations (since more blocks can be completed), and this will hopefully also lower the binding energy.

3.4 Cu-MOF-74

3.4.1 Cu-MOF-74 DFT Calculations

The Cu-MOF-74 energies were also calculated in the process of generating the trial wavefunctions at the DFT level. The calculations were performed for a single adsorption site, so the non-CO_2 binding metal sites were replaced with zinc, due to its full 3d electron subshell. Table 9 (with CO_2) and Table 10 (without CO_2) show the results for the cc.VDZ basis set. Due to time and computing restrictions, cc.VDZ and cc.VTZ are the only finished basis sets.

Table 9: Energy (E_h and kJ/mol) and computational time (s) of basis sets used for DFT Cu-MOF-74 with CO_2 calculations

Number of Basis Functions	Basis Set	Energy (E_h)	Energy Difference (kJ/mol)	Time (s)
1158	cc.VDZ	-1366.967665	890.78	92142.7
2478	cc.VTZ	-1367.306947	0	1400117

Table 10: Energy (E_h and kJ/mol) and computational time (s) of basis sets used for DFT Cu-MOF-74 calculations

Number of Basis Functions	Basis Set	Energy (E_h)	Energy Difference (kJ/mol)	Time (s)
1116	cc.VDZ	-1329.200163	839.54	84882.6
2376	cc.VTZ	-1329.519927	0	636478.1

As shown in tables 9 and 10, the energy difference is highly correlated to the basis set size, as it was for Zn-MOF-74.

3.4.2 Cu-MOF-74 DFT Binding Energy Results

Using Equation 1, the binding energy for a single CO_2 at a copper adsorption site can be calculated. Table 11 shows the energies along with a comparison of experimental *Q* values. [7]

Table 11: Binding Energy (E_h and kJ/mol) of basis sets used for Cu-MOF-74 calculations

Basis Set	Binding Energy (E_h)	Binding Energy (kJ/mol)	Experimental *Q* Value (kJ/mol)	Experimental and Theory Difference (kJ/mol)
cc.VDZ	-0.00647644	-17.00	-22.1	5.1
cc.VTZ	-0.015024843	-9.96	-22.1	12.1

The Cu-MOF-74 binding energies show a similar relationship to the Zn-MOF-74 DFT binding energies. The basis set size is highly correlated to the binding energy, and the experimental *Q* values differ from the computational binding energies by 5-12 kJ/mol.

4. Conclusions

DMC calculations using a trial wavefunction obtained from PBE0/cc.VDZ with an optimized 3-body Jastrow term result in a binding energy of -18.02 kJ/mol with an error of 16. 74 kJ/mol for CO_2 in Zn-MOF-74. This compares well to the experimental *Q* value of -26.8 kJ/mol, and is an improvement upon a DFT calculated binding energy of -9.41 kJ/mol. However, the error bars of the calculations are similar in magnitude to the actual binding energy, and so need to be reduced. A possible solution is to run the DMC calculations on more computing cores, so it is able to calculate more samples and reduce the uncertainty. A more complete DMC calculation could potentially also lead to a more accurate binding energy and a smaller error bound.

5. Future Work

To accurately access the pseudopotentials for use in large molecules DMC, such as MOF, calculations, the DMC binding energies need to be calculated for the rest of the 3d transition metal series. The binding energies can be compared against experimental values to determine if the DMC calculations are accurate. Additional basis sets also need to be tested to determine the impact the basis set size has on DMC calculations. Tests need to be run for the DMC calculations, using more cores on the computing grids, to see if more DMC iterations will decrease the error bound, and potentially lead to a more accurate binding energy.

Appendix

DFT Errors and Resolutions

CO_2 Symmetry Games error

During DFT calculations of CO_2 an error was encountered. The symmetry was originally designates as D_{2h}, to show it was a linear molecule that reflected over the point(0, 0, 0). Oxygen was only defined once due to symmetry. An error message was returned however. File A1 shows the input file (ECPs and psuedopotentials are omitted) and File A2 shows the returned error message.

```
$contrl coord=unique dfttyp=pbe0 ecp=read exetyp=run icharg=0
        ispher=1 maxit=200 mult=1 runtyp=energy scftyp=uhf $end
 $system memory=150000000 $end
 $scf    dirscf =.true. $end
 $guess  guess=hcore prtmo=.true.  $end
 $ECP
C-ccECP GEN 2 1
O-ccECP GEN 2 1
Dnh 2

Carbon 12    0.0000000000    0.00000000000     0.000000000000

Oxygen 8     1.160000000    0.00000000000    0.0000000000000

$END
```

File 1: Input scf file for a CO_2 molecule, designating a D_{2h} symmetry about the x-axis

```
FIRST CHARACTER OF ANY KEYWORD >       < MUST ALWAYS BE A LETTER
 **** ERROR READING INPUT GROUP $CONTRL   *****
 THE PROBLEM IS WITH THIS INPUT LINE, NEAR THE X MARKER
       ISPHER=1 MAXIT=200 MULT=1 RUNTYP=ENERGY SCFTYP=UHF $END
```

X

THE ONLY KEYWORDS ACCEPTED IN THIS GROUP ARE:

SCFTYP RUNTYP EXETYP ICHARG MULT UNITS

INTTYP LOCAL MAXIT NPRINT IREST NORMF

NORMP ITOL ICUT NZVAR NOSYM GEOM

MPLEVL AIMPAC PP ECP PLTORB MOLPLT

COORD FRIEND NOSO CITYP CCTYP ISPHER

QMTTOL RELWFN NUMGRD GRDTST GRDTYP VBTYP

DFTTYP TDDFT ISKPRP NFFLVL ETOLLZ IAHARD

CIMTYP CASINO PMTD1 DGRID

TYPING ERROR IN $CONTRL INPUT - CHECK NEAR $ MARKER

EXECUTION OF GAMESS TERMINATED -ABNORMALLY- AT Tue Sep 29 11:19:19 2020

File 2: Output Error Message for a DFT calculation of CO_2 with a D_{2h} symmetry designation

The error message suggest that a keyword in the $CONTRL group was not a recognizable term. However, all terms were in the list of acceptable keywords. So, as a solution, the file was rewritten to include no symmetry, and instead describe each atom individually with a symmetry group of C1 (no symmetry). This was successful and corrected the error introduced by assuming symmetry.

MOF ECP reading error

The MOF files were compiled by taking the structure coordinates and inputting them into a file the same form as the CO_2 molecule. After submitting the job, it finished and returned with the error message in file 3.

**On entry to DGEMM parameter number 3 had an illegal value

**On entry to DGEMM parameter number 3 had an illegal value

**On entry to DGEMM parameter number 3 had an illegal value

Mpiexec has exited due to process rank 3 with PID 16262 on

node moose 15 exiting improperly. There are three reasons this could occur:

1. this process did not call “init” before exiting, but other in the job did. This can cause a job to hang indefinitely while it waits for all processes to call “init”. By rule, if one process calls “init”, then ALL processes must call “init” prior to termination.

2. this process called “init”, but exited without calling “finalize”. By rule, all processes that call “init” MUST call “finalize” prior to exiting or it will be considered an “abnormal termination”

3. this process called “MPI_Abort” or “orte_abort” and the mca parameter orte_create_session_dirs is set to false. In this case, the run-time cannot detect that the abort call was an abnormal termination. Hence, the only error message you will receive is this one.

This may have cause other processes in the application to be terminated by signals sent by mpiexec (as reported here).

You can avoid this message by specifying -quiet on the mpiexec cpmmand line.

File 3: DFT MOF error DGEMM parameter 3

This error was unknown to us and a solution was unable to be found from prior research. To test if the issue was with the model cluster or the specification of the ECP’s, a sample file was run, in which the basis sets were not specified. The cc.VTZ basis set was used because the GAMESS software did not have the cc.VDZ basis set File 4 shows the sample $contrl section from the test file.

$contrl coord = unique dfttyp= pbe0 exetyp=run icharg=0

Ispher=1 maxit=200 mult=1 runtyp=energy scftyp=uhf $end

$system memory = 150000000 $end

$BASIS GBASIS=VTZ $END

$scf guess=huckel prtmo=.true $end

File 4: $Control section for built in ecp test tun of Zn-MOF-74

This also returned an error message, it is shown in File 5.

THE POINT GROUP OF THE MOLECULES IS C1

THE ORDER OF THE PRINCIPAL AXIS IS 0

DFT GRIDS ARE AVAILABLE ONLY FOR INTEGER ATOMIC

NUMBERS, SO -SPARKLES- ARE NOT ALLOWED.

File 5: error message for MOF built-in ecp basis test

To test if the issue was with the basis set, we changed the $BASIS to GBASIS=STO NGAUSS =3. This file successfully ran, showing that the issue was not with the model cluster, but rather with the EECPs and basis functions.

The file was rerun, with the ECP's desired inputted and $CONTRL changed to ecp=read. The $BASIS had to be changed to SBKJC (a ECP based pseudopotential set) because STO is a non-ecp basis set. This returned the same error as in file 3. This showed that the issue was not with the potentials, because the file ran calculated everything up to the huckel guess. Under the $guess section, we changed guess=huckel to guess=hcore, and added the desired psuedopotentials. This test run was successful for the cc.VDZ basis set.

The rest of the basis sets were tested. The results showed that there was still an issue with the input file. For all basis sets, the energy of the MOF was 129.7433802279 E_h and the energy of the MOF + CO_2 was 53.4469648275 E_h. The energies should be different and getting progressively lower with an increasing number of functions in the basis set. Within the output

file for the MOF+CO_2, the information given said that the ECP removed 522 core electrons and protons, leaving the system with -26 electrons. The MOF alone had a similar message. This suggested that something was still wrong with the ECPs, because it should not be removing more electrons than exist. Looking further into the output file, the software was recognizing non-atoms as atoms, and thus subtracting more electrons than it should, as shown in File 6. ZCORE represents the number electrons subtracted from the total for the specified atom.

PARAMETERS FOR “O-CCECP “ ON ATOM 1 WITH ZCORE 2 AND LMAX 1 ARE

PARAMETERS FOR “ZN-CCECP’ ON ATOM 2 WITH ZCORE 10 AND LMAX 2 ARE

PARAMETERS FOR “3” ON ATOM 4 WITH ZCORE 0 AND LMAX 0 ARE

PARAMETERS FOR “6.000000” ON ATOM 5 WITH ZCORE 12 AND LMAZ 0 ARE

PARAMETERS FOR “73.85984” ON ATOM 6 WITH ZCORE 14 AND LMAX 0 ARE

File 6: Paramater specification errors for DFT calculations of Zn-MOF-74

File 6 shows that the GAMESS software was reading in the actual ECPs as atoms. To rectify this, each ECP needs to only be defined under the first time that element appears. The pseudopotentials, still need to be defined under each atom at its coordinate. File 7 shows this for the cc.VDZ basis set of Zn-MOF-74 with CO_2.

```
$ECP
O-ccECP GEN 2 1
3
  6.000000 1  12.30997
  73.85984 3  14.76962
 -47.87600 2  13.71419
1
```

85.86406 2 13.65512
Zn-ccECP GEN 10 2
4
20.00000000 1 35.80797616
716.15952323 3 34.53646083
-204.68393323 2 28.62830178
0.76026614 2 7.96239682
2
431.70804302 2 35.02141356
95.87640437 2 14.63498691
2
313.57770563 2 42.22979234
74.01270048 2 14.57429304
O-ccECP GEN 2 1
C-ccECP GEN 2 1
3
4.00000 1 14.43502
57.74008 3 8.39889
-25.81955 2 7.38188
1
52.13345 2 7.76079
C-ccECP GEN 2 1
C-ccECP GEN 2 1
H-ccECP GEN 0 1
3
1.00000000000000 1 21.24359508259891
21.24359508259891 3 21.24359508259891
-10.85192405303825 2 21.77696655044365
1
0.00000000000000 2 1.000000000000000

O-ccECP GEN 2 1

O-ccECP GEN 2 1

O-ccECP GEN 2 1

C-ccECP GEN 2 1

C-ccECP GEN 2 1

C-ccECP GEN 2 1

C-ccECP GEN 2 1

H-ccECP GEN 0 1

Zn-ccECP GEN 10 2

File 7: Example ECP section for the cc.VDZ basis set of Zn-MOF-74 with CO_2

This fixed the error, and allowed the calculations to run smoothly.

Sample DFT files

File 8 shows the example input file for the cc.VDZ basis set of the CO_2 molecule. The MOF files look similar except they use hcore instead of huckel, and only the first atom of each element has its ECP defined. For CO_2, defining the ECP for both oxygens is inconsequential because the 2nd oxygen is at the end of the section, so the software stops reading in the input file when it reaches the oxygen.

$contrl coord=unique dfttyp=pbe0 ecp=read exetyp=run icharg=0
ispher=1 maxit=200 mult=1 runtyp=energy scftyp=uhf $end
$system memory=150000000 $end
$scf dirscf =.true. $end
$guess guess=huckel prtmo=.true. $end
$ECP
C-ccECP GEN 2 1
3
4.00000 1 14.43502

57.74008 3 8.39889

-25.81955 2 7.38188

1

52.13345 2 7.76079

O-ccECP GEN 2 1

3

 6.000000 1 12.30997

 73.85984 3 14.76962

-47.87600 2 13.71419

1

 85.86406 2 13.65512

O-ccECP GEN 2 1

3

 6.000000 1 12.30997

 73.85984 3 14.76962

-47.87600 2 13.71419

1

 85.86406 2 13.65512

 $END

 $DATA

Carbon Dioxide

C1

C 6   0.0000000000   0.00000000000   0.000000000000

s 9 1.00

1 13.073594   0.0051583

2 6.541187   0.0603424

3 4.573411   -0.1978471

4 1.637494   -0.0810340

5 0.819297   0.2321726

6 0.409924   0.2914643

7 0.231300 0.4336405

8 0.102619 0.2131940

9 0.051344 0.0049848

s 1 1.00

1 0.127852 1.000000

p 9 1.00

1 9.934169 0.0209076

2 3.886955 0.0572698

3 1.871016 0.1122682

4 0.935757 0.2130082

5 0.468003 0.2835815

6 0.239473 0.3011207

7 0.117063 0.2016934

8 0.058547 0.0453575

9 0.029281 0.0029775

p 1 1.00

1 0.149161 1.000000

d 1 1.00

1 0.561160 1.000000

O 8 1.160000000 0.00000000000 0.0000000000000

S 9

1 54.775216 -0.0012444

2 25.616801 0.0107330

3 11.980245 0.0018889

4 6.992317 -0.1742537

5 2.620277 0.0017622

6 1.225429 0.3161846

7 0.577797 0.4512023

8 0.268022 0.3121534

9 0.125346 0.0511167

S 1

1 0.258551 1.0000000

P 9

1 22.217266 0.0104866

2 10.74755 0.0366435

3 5.315785 0.0803674

4 2.660761 0.1627010

5 1.331816 0.2377791

6 0.678626 0.2811422

7 0.333673 0.2643189

8 0.167017 0.1466014

9 0.083598 0.0458145

P 1

1 0.267865 1.0000000

D 1

1 1.232753 1.0000000

O 8 - 1.160000000 0.00000000000 0.0000000000000

S 9

1 54.775216 -0.0012444

2 25.616801 0.0107330

3 11.980245 0.0018889

4 6.992317 -0.1742537

5 2.620277 0.0017622

6 1.225429 0.3161846

7 0.577797 0.4512023

8 0.268022 0.3121534

9 0.125346 0.0511167

S 1

```
1   0.258551     1.0000000
P 9
1  22.217266     0.0104866
2  10.74755      0.0366435
3   5.315785     0.0803674
4   2.660761     0.1627010
5   1.331816     0.2377791
6   0.678626     0.2811422
7   0.333673     0.2643189
8   0.167017     0.1466014
9   0.083598     0.0458145
P 1
1   0.267865     1.0000000
D 1
1   1.232753     1.0000000

$END
```

File 8: Example DFT input file of CO_2 using the cc.VDZ basis set

Sample DMC Optimization files

When converting the DFT files for use in the DMC calculations, the software produces an input file that can be modified for wavefunction optimization or DMC production. File 9 shows the input file used for the optimization of the cc.VDZ basis set of CO_2.

```
<?xml version="1.0"?>
<simulation>
  <!--
Example QMCPACK input file produced by convert4qmc
It is recommend to start with only the initial VMC block and adjust
parameters based on the measured energies, variance, and statistics.
```

```
-->
  <!--Name and Series number of the project.-->
  <project id="opt" series="0"/>
  <!--Link to the location of the Atomic Coordinates and the location of the Wavefunction.-->
  <include href="scf.structure.xml"/>
  <include href="scf.wfj.xml"/>
  <!--Hamiltonian of the system. Default ECP filenames are assumed.-->
  <hamiltonian name="h0" type="generic" target="e">
    <pairpot name="ElecElec" type="coulomb" source="e" target="e" physical="true"/>
    <pairpot name="IonIon" type="coulomb" source="ion0" target="ion0"/>
    <pairpot name="PseudoPot" type="pseudo" source="ion0" wavefunction="psi0"
format="xml">
      <pseudo elementType="C" href="C.ccECP.xml"/>
      <pseudo elementType="O" href="O.ccECP.xml"/>
    </pairpot>
  </hamiltonian>
  <!--
Example initial VMC to measure initial energy and variance
-->
  <qmc method="vmc" move="pbyp" checkpoint="-1">
    <estimator name="LocalEnergy" hdf5="no"/>
    <parameter name="warmupSteps">100</parameter>
    <parameter name="blocks">20</parameter>
    <parameter name="steps">50</parameter>
    <parameter name="substeps">8</parameter>
    <parameter name="timestep">0.5</parameter>
    <parameter name="usedrift">no</parameter>
  </qmc>
  <!--
Example initial VMC optimization
```

```
Number of steps required will be computed from total requested sample
count and total number of walkers
-->
  <loop max="4">
    <qmc method="linear" move="pbyp" checkpoint="-1">
      <estimator name="LocalEnergy" hdf5="no"/>
      <parameter name="warmupSteps">100</parameter>
      <parameter name="blocks">20</parameter>
      <parameter name="timestep">0.5</parameter>
      <parameter name="walkers">1</parameter>
      <parameter name="samples">16000</parameter>
      <parameter name="substeps">4</parameter>
      <parameter name="usedrift">no</parameter>
      <parameter name="MinMethod">OneShiftOnly</parameter>
      <parameter name="minwalkers">0.0001</parameter>
    </qmc>
  </loop>
  <!--
Example follow-up VMC optimization using more samples for greater accuracy
-->
  <loop max="10">
    <qmc method="linear" move="pbyp" checkpoint="-1">
      <estimator name="LocalEnergy" hdf5="no"/>
      <parameter name="warmupSteps">100</parameter>
      <parameter name="blocks">20</parameter>
      <parameter name="timestep">0.5</parameter>
      <parameter name="walkers">1</parameter>
      <parameter name="samples">64000</parameter>
      <parameter name="substeps">4</parameter>
      <parameter name="usedrift">no</parameter>
```

```
    <parameter name="MinMethod">OneShiftOnly</parameter>
    <parameter name="minwalkers">0.3</parameter>
   </qmc>
  </loop>
</simulation>
```

File 9: Wavefunction optimization input file

File 10 shows the script used to submit the job. Opt.in.xml refers to File 9.

```
#!/bin/bash
#$ -cwd
#$ -j y
#$ -S /bin/bash
#$ -M crenfro@bowdoin.edu -m b -m e

export OMP_NUM_THREADS=1
module load qmcpack-mpi
mpiexec -n $NSLOTS qmcpack opt.in.xml
```

File 10: Wavefunction optimization optimization script

Sample DMC Calculation files

File 11 shows the input file used for the DMC calculations. It is the same file produced by the convert to qmc used in the optimization step, except the optimization blocks are deleted and the production blocks are kept.

```
<?xml version="1.0"?>
<simulation>
  <!--Name and Series number of the project.-->
  <project id="qmc" series="0"/>
  <!--Link to the location of the Atomic Coordinates and the location of the Wavefunction.-->
```

```
  <include href="scf.structure.xml"/>
  <include href="opt.s011.opt.xml"/>
  <!--Hamiltonian of the system. Default ECP filenames are assumed.-->
  <hamiltonian name="h0" type="generic" target="e">
    <pairpot name="ElecElec" type="coulomb" source="e" target="e" physical="true"/>
    <pairpot name="IonIon" type="coulomb" source="ion0" target="ion0"/>
    <pairpot name="PseudoPot" type="pseudo" source="ion0" wavefunction="psi0"
format="xml">
      <pseudo elementType="C" href="C.ccECP.xml"/>
      <pseudo elementType="O" href="O.ccECP.xml"/>
    </pairpot>
  </hamiltonian>
  <!--
Example initial VMC to measure initial energy and variance
-->
  <qmc method="vmc" move="pbyp" checkpoint="-1">
    <estimator name="LocalEnergy" hdf5="no"/>
    <parameter name="warmupSteps">100</parameter>
    <parameter name="blocks">20</parameter>
    <parameter name="steps">50</parameter>
    <parameter name="substeps">8</parameter>
    <parameter name="timestep">0.5</parameter>
    <parameter name="usedrift">no</parameter>
  </qmc>
  <!--
Production VMC and DMC
Examine the results of the optimization before running these blocks.
e.g. Choose the best optimized jastrow from all obtained, put in
wavefunction file, do not reoptimize.
-->
```

```
  <qmc method="vmc" move="pbyp" checkpoint="-1">
    <estimator name="LocalEnergy" hdf5="no"/>
    <parameter name="warmupSteps">100</parameter>
    <parameter name="blocks">200</parameter>
    <parameter name="steps">50</parameter>
    <parameter name="substeps">8</parameter>
    <parameter name="timestep">0.5</parameter>
    <parameter name="usedrift">no</parameter>
    <!--Sample count should match targetwalker count for DMC. Will be obtained from all nodes.-->
    <parameter name="samples">16000</parameter>
  </qmc>
  <qmc method="dmc" move="pbyp" checkpoint="20">
    <estimator name="LocalEnergy" hdf5="no"/>
    <parameter name="targetwalkers">16000</parameter>
    <parameter name="reconfiguration">no</parameter>
    <parameter name="warmupSteps">100</parameter>
    <parameter name="timestep">0.005</parameter>
    <parameter name="steps">100</parameter>
    <parameter name="blocks">100</parameter>
    <parameter name="nonlocalmoves">yes</parameter>
  </qmc>
</simulation>
```

File 11: Sample input file for DMC production for CO_2 molecule

The opt.s011.opt.xml file is the lowest energy wavefunction produced by the optimization process. The job is submitted with a script, which is shown in File 12, where qmc.in.xml is file 11.

```
#!/bin/bash
```

```
#$ -cwd
#$ -j y
#$ -S /bin/bash
#$ -M crenfro@bowdoin.edu -m b -m e

export OMP_NUM_THREADS=1
module load qmcpack-mpi
mpiexec -n $NSLOTS qmcpack qmc.in.xml
```

File 12: Sample submission script for DMC production

References

1. Gergel, D. R.; Nijssen, B.; Abatzoglou, J. T.; Lettenmaier, D. P.; Stumbaugh, M. R.. Effects of climate change on snowpack and fire potential in the western USA. *Climatic Change* 2017, 141 (2), 287-299.
2. George, M.; Bruzzese, J.-M.; Matura, L. A., Climate Change Effects on Respiratory Health: Implications for Nursing. Journal of Nursing Scholarship 2017, 49 (6), 644-652.
3. Lindsey, R. Climate Change: Atmospheric Carbon Dioxide. *NOAA* Climate.gov. https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide
4. Olivier, J.; Peters, J.. Trends in Global CO2 and Total Greenhouse Gas Emissions; 2019 report; report 4068; *PBL Netherlands Environmental Assessment Agency*, 2020; 4-6.
5. Suekane, T.; Nobuso, T.; Hirai, S.; Kiyota, M., Geological storage of carbon dioxide by residual gas and solubility trapping. *International Journal of Greenhouse Gas Control* 2008, 2 (1), 58-64.
6. Leung, D. Y. C.; Caramanna, G.; Maroto-Valer, M. M., An overview of current status of carbon dioxide capture and storage technologies. *Renewable and Sustainable Energy Reviews* 2014, 39, 426-443.
7. Queen, W. L.; Hudson, M. R.; Bloch, E. D.; Mason, J. A.; Gonzalez, M. I.; Lee, J. S.; Gygi, D.; Howe, J. D.; Lee, K.; Darwish, T. A.; James, M.; Peterson, V. K.; Teat, S. J.; Smit, B.; Neaton, J. B.; Long, J. R.; Brown, C. M., Comprehensive study of carbon dioxide adsorption in the metal–organic frameworks M2(dobdc) (M = Mg, Mn, Fe, Co, Ni, Cu, Zn). *Chemical Science* 2014, 5 (12), 4569-4581.

8. Britt, D.; Furukawa, H.; Wang, B.; Glover, T. G.; Yaghi, O. M., Highly efficient separation of carbon dioxide by a metal-organic framework replete with open metal sites. *Proceedings of the National Academy of Sciences* 2009, 106 (49), 20637-20640.
9. Forse, A. C.; Gonzalez, M. I.; Siegelman, R. L.; Witherspoon, V. J.; Jawahery, S.; Mercado, R.; Milner, P. J.; Martell, J. D.; Smit, B.; Blümich, B.; Long, J. R.; Reimer, J. A., Unexpected Diffusion Anisotropy of Carbon Dioxide in the Metal–Organic Framework Zn2(dobpdc). *Journal of the American Chemical Society* 2018, 140 (5), 1663-1673.
10. Uzun, A.; Keskin, S., Site characteristics in metal organic frameworks for gas adsorption. *Progress in Surface Science* 2014, 89 (1), 56-79.
11. Daglar, H.; Keskin, S., Recent advances, opportunities, and challenges in high-throughput computational screening of MOFs for gas separations. *Coordination Chemistry Reviews* **2020,** *422*, 213470.
12. Fiolhais, C.; Nogueira, F.; Marques, M. A. L., A Primer in Density Functional Theory. *Springer Berlin Heidelberg*: 2003.
13. Cohen, A. J.; Mori-Sánchez, P.; Yang, W., Insights into Current Limitations of Density Functional Theory. *Science* 2008, 321 (5890), 792-794.
14. Grajciar, L.; Nachtigall, P.; Bludský, O.; Rubeš, M., Accurate Ab Initio Description of Adsorption on Coordinatively Unsaturated Cu2+ and Fe3+ Sites in MOFs. *Journal of Chemical Theory and Computation* 2015, 11 (1), 230-238.
15. Parkes, M. V.; Sava Gallis, D. F.; Greathouse, J. A.; Nenoff, T. M., Effect of Metal in M3(btc)2 and M2(dobdc) MOFs for O2/N2 Separations: A Combined Density Functional

Theory and Experimental Study. *The Journal of Physical Chemistry C* 2015, 119 (12), 6556-6567.

16. Benali, A.; Luo, Y.; Shin, H.; Pahls, D.; Heinonen, O., Quantum Monte Carlo Calculations of Catalytic Energy Barriers in a Metallorganic Framework with Transition-Metal-Functionalized Nodes. *The Journal of Physical Chemistry C* 2018, 122 (29), 16683-16691.
17. Vlaisavljevich, B.; Huck, J.; Hulvey, Z.; Lee, K.; Mason, J. A.; Neaton, J. B.; Long, J. R.; Brown, C. M.; Alfè, D.; Michaelides, A.; Smit, B., Performance of van der Waals Corrected Functionals for Guest Adsorption in the M2(dobdc) Metal–Organic Frameworks. *The Journal of Physical Chemistry A* 2017, 121 (21), 4139-4151.
18. Grüneis, A.; Hirata, S.; Ohnishi, Y.-y.; Ten-no, S., Perspective: Explicitly correlated electronic structure theory for complex systems. *The Journal of Chemical Physics* 2017, 146 (8), 080901.
19. Dzubak, A. L.; Krogel, J. T.; Reboredo, F. A., Quantitative estimation of localization errors of 3d transition metal pseudopotentials in diffusion Monte Carlo. *The Journal of Chemical Physics* 2017, 147 (2), 024102.
20. Rasander, Mikael. (2010). A Theoretical Perspective on the Chemical Bonding and Structure of Transition Metal Carbides and Multilayers. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology*.
21. Tan, H.; Li, Y.; Zhang, S.; Duan, W., Effect of Hartree–Fock pseudopotentials on local density functional theory calculations. *Physical Chemistry Chemical Physics* 2018, 20 (27), 18844-18849.

22. Krogel, J. T.; Kent, P. R. C., Magnitude of pseudopotential localization errors in fixed node diffusion quantum Monte Carlo. *J. Chem. Phys*. 2017, 146 (24), 244101/1.
23. Krogel, J. T.; Santana, J. A.; Reboredo, F. A., Pseudopotentials for quantum Monte Carlo studies of transition metal oxides. Physical Review B 2016, 93 (7), 075143.
24. J.P.Perdew, K.Burke, M.Ernzerhof, *Phys.Rev.Lett*. 77, 3865-8(1996); Err. 78,1396(1997)
25. Barca, G. M. J.; Bertoni, C.; Carrington, L.; Datta, D.; Silva, N. D.; Deustua, J. E.; Fedorov, D. G.; Gour, J. R.; Gunina, A. O.; Guidez, E.; Harville, T.; Irle, S.; Ivanic, J.; Kowalski, K.; Leang, S. S.; Li, H.; Li, W.; Lutz, J. J.; Magoulas, I.; Mato, J.; Mironov, V.; Nakata, H.; Pham, B. Q.; Piecuch, P.; Poole, D.; Pruitt, S. R.; Rendell, A. P.; Roskop, L. B.; Ruedenberg, K.; Sattasathuchana, T.; Schmidt, M. W.; Shen, J.; Slipchenko, L.; Sosonkina, M.; Sundriyal, V.; Tiwari, A.; Vallejo, J. L. G.; Westheimer, B.; Włoch, M.; Xu, P.; Zahariev, F.; Gordon, M. S., Recent developments in the general atomic and molecular electronic structure system. The Journal of Chemical Physics 2020, 152 (15), 154102.
26. Bennett, M. C.; Melton, C. A.; Annaberdiyev, A.; Wang, G.; Shulenburger, L.; Mitas, L., A new generation of effective core potentials for correlated calculations. The Journal of Chemical Physics 2017, 147 (22), 224106.
27. Annaberdiyev, A.; Wang, G.; Melton, C. A.; Bennett, M. C.; Shulenburger, L.; Mitas, L., A new generation of effective core potentials from correlated calculations: 3d transition metal series. The Journal of Chemical Physics 2018, 149 (13), 134108.
28. Kim, J.; Baczewski, A. D.; Beaudet, T. D.; Benali, A.; Bennett, M. C.; Berrill, M. A.; Blunt, N. S.; Borda, E. J. L.; Casula, M.; Ceperley, D. M.; Chiesa, S.; Clark, B. K.; Clay, R. C.; Delaney, K. T.; Dewing, M.; Esler, K. P.; Hao, H.; Heinonen, O.; Kent,

P. R. C.; Krogel, J. T.; Kylänpää, I.; Li, Y. W.; Lopez, M. G.; Luo, Y.; Malone, F. D.; Martin, R. M.; Mathuriya, A.; McMinis, J.; Melton, C. A.; Mitas, L.; Morales, M. A.; Neuscamman, E.; Parker, W. D.; Pineda Flores, S. D.; Romero, N. A.; Rubenstein, B. M.; Shea, J. A. R.; Shin, H.; Shulenburger, L.; Tillack, A. F.; Townsend, J. P.; Tubman, N. M.; Van Der Goetz, B.; Vincent, J. E.; Yang, D. C.; Yang, Y.; Zhang, S.; Zhao, L., QMCPACK: an open sourceab initioquantum Monte Carlo package for the electronic structure of atoms, molecules and solids. Journal of Physics: Condensed Matter 2018, 30 (19), 195901.

29. Drummond, N. D.; Towler, M. D.; Needs, R. J., Jastrow correlation factor for atoms, molecules, and solids. *Physical Review B* **2004,** *70* (23), 235119.
30. Casula, M., Beyond the locality approximation in the standard diffusion Monte Carlo method. *Physical Review B* 2006, 74 (16), 161102.
31. Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J., Hartree−Fock and Density Functional Theory ab Initio Calculation of Optical Rotation Using GIAOs: Basis Set Dependence. The Journal of Physical Chemistry A 2000, 104 (5), 1039-1046.
32. Kruse, H.; Grimme, S., A geometrical correction for the inter- and intra-molecular basis set superposition error in Hartree-Fock and density functional theory calculations for large systems. *The Journal of Chemical Physics* **2012,** *136* (15), 154101.