Active Space Methods in Electronic Structure Theory and Applications to Gas Separations in Metal-Organic Frameworks

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Samuel Joseph Stoneburner

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Laura Gagliardi, advisor

June 2019

Acknowledgements

My gratitude goes out to my advisor, Laura Gagliardi, for conversations, explanations, pieces of advice and insight, amazing collaborative opportunities, and countless other forms of mentorship and support. I would also like to thank the others who have mentored me within the group, Allison Dzubak, Giovanni Li Manni, and especially Konstantinos (Kostas) Vogiatzis.

There have been many collaborators and co-authors, and they all have my appreciation, but in particular I wish to thank WooSeok Jeong, whose tireless quest for understanding is inspiring, and Don Truhlar, who has taught me so much about precision in technical writing.

On a more personal note, I want to thank my wife, Jamie, for keeping me grounded and providing necessary perspective on so many parts of life, and my children, Halili, Jocelyn, John, and Zuri, for the sheer joy of watching them grow and learn over these many months.

Finally, I give highest thanks to God for everything, and especially for making the universe run on math, without which none of the work in this dissertation would have been possible.

Dedication

This dissertation is dedicated to my wife, Jamie, who knew I was a science nerd and married me anyway.

Table of Contents

List of Tables

.

List of Figures

.

List of Abbreviations

ADF..... Amsterdam Density Functional

AFM.....antiferromagnetic

AO..... atomic orbital

BSSE.....basis set superposition error

CAS..... complete active space

CASCI.....complete active space configuration interaction

CASPT2 .....complete active space second-order perturbation theory

CASPT2/CC..... CASPT2 with coupled-cluster semi-core correlation

CASSCF.....complete active space self-consistent field theory

CC .....coupled-cluster theory

CCSD .....coupled-cluster singles and doubles

CCSD(T)..... coupled-cluster singles and doubles with perturbative triples

CI.....configuration interaction

CI.....configuration interaction

CoRE database..... computation-ready experimental database of MOF structures

CPO..... correlated participating orbitals

CSD..... Cambridge Structural Database

CSF .....configuration state functions

DEA-EOMCC..... doubly electron-attached equation-of-motion coupled-cluster theory

DFT .....density functional theory

DFT/CC ..... density functional theory with coupled-cluster corrections

DMRG ..... density matrix renormalization group

EA ..... electron affinity

EBS ..... extended broken-symmetry

eCPO ..... extended correlated participating orbitals

EOM-CC ..... equation-of-motion coupled-cluster theory

EQeq ..... extended charge equilibration

FCI ..... full configuration interaction

FCIQMC ..... full configuration interaction quantum Monte Carlo

FeP ..... iron porphyrin

FM ..... ferromagnetic

GAS ..... generalized active space

GASPT2 ..... generalized active space second-order perturbation theory

GASSCF ..... generalized active space self-consistent field theory

GCMC ..... grand canonical Monte Carlo

GGA ..... generalized gradient approximation

GUGA ..... graphical unitary group approach

HF ..... Hartree-Fock

HLE ..... high local exchange

HOMO ..... highest occupied molecular orbital

HS ..... high-spin

icMRCI ..... internally-contracted multireference configuration interaction

IP ..... ionization potentials

IPEA shift..... ionization potential electron affinity shift

irrep ..... irreducible representation

IS ..... intermediate-spin

KS-DFT..... Kohn-Sham density functional theory

LCD..... largest cavity diameters

LDA ..... local density approximation

LMO..... localized molecular orbitals

LS ..... low-spin

LUMO..... lowest occupied molecular orbital

MBPT..... many-body perturbation theory

MC-PDFT ..... multiconfiguration pair-density functional theory

mCPO..... moderate correlated participating orbitals

MCSCF ..... multiconfiguration self-consistent field

MkCCSD..... Mukherjee coupled-cluster singles and doubles

MO ..... molecular orbital

MOF ..... metal-organic framework

MP2..... Møller-Plesset perturbation theory, second order

MRCI ..... multireference configuration interaction

MRCISD ..... multireference configuration interaction singles and doubles

MRPT2..... multireference second order perturbation theory

MUD ..... mean unsigned deviations

MUE.....mean unsigned error

nCPO..... nominal correlated participating orbitals

OMS.....open metal sites

ORMAS .....occupation restricted multiple active spaces

PAW..... projector augmented wave

PLD ..... pore limiting diameters

PT2.....second-order perturbation theory

PT2.....second-order perturbation theory

QTAIM ..... Quantum Theory of Atoms in Molecules

RAS..... restricted active space

RASCI.....restricted active space configuration interaction

RASPT2 .....restricted active space second-order perturbation theory

RASSCF.....restricted active space self-consistent field theory

ROHF .....restricted open-shell Hartree-Fock

ROS-MP2.. Restricted open shell Møller-Plesset perturbation theory, second order

SA ..... state-average

SBU.....secondary building unit

SCF ..... self-consistent field

SOMO ..... singly occupied molecular orbitals

SOMO ..... singly occupied molecular orbitals

UBD(T) ..... spin-unrestricted Brueckner doubles with perturbative triples

UHF..... unrestricted Hartree-Fock

VASP .....Vienna Ab Initio Simulation Package

VEE.....vertical excitation energy

WABS.....weighted-average broken-symmetry

WFT.....wave function theory

ZORA.....zeroth-order regular approximation

πCPO..... π correlated participating orbitals

1 Introduction

The theoretical background in Sections 1.2.5, 1.2.6, and 1.2.8 includes some content adapted from a collaborative Perspective article written by Carlo Alberto Gaggioli, Samuel J. Stoneburner, Christopher J. Cramer, and Laura Gagliardi. The article has been submitted for publication.[1] In that work Samuel J. Stoneburner wrote much of the content on multireference calculations and made other contributions throughout the manuscript.

1.1 Organization of this Dissertation

Section 1 of this dissertation is introductory material. Following this summary in Section 1.1, theoretical background is presented in Section 1.2 that describes many of the electronic structure methods used throughout this dissertation. Section 1.3 describes some of the challenges in choosing orbitals for active spaces, which is necessary for much of the research presented in this dissertation, and Section 1.4 comments on the practical applications that will be addressed.

Section 2 features the systematic exploration of active spaces from the standpoint of theoretical development, beginning with Section 2.1 and the benchmarking of generalized active space self-consistent field theory (GASSCF)[2] and SplitGAS[3] on a variety of systems. Section 2.2 considers the application of the "correlated participating orbital" (CPO)[4] active space selection scheme to complete active space self-consistent field theory (CASSCF)[5,6] and restricted active space self-consistent field theory (RASSCF)[7] followed by second-order perturbation theory (CASPT2[8,9] and RASPT2,[10] respectively) for singlet-triplet splittings of diradical organic molecules. An additional type of CPO, "πCPO", is introduced as an effective and economical option for π-system excitation energies.

Section 3 consists of assessments of multiconfiguration pair-density functional theory MC-PDFT[11,12] and includes additional work with active space selection. Section 3.1 is a direct continuation of the work presented in Section 2.2 and demonstrates that with the same systems and CPO active space selection scheme MC-PDFT can provide good agreement with CASPT2 at a much lower computational cost. Section 3.2 further demonstrates the computational affordability of MC-PDFT through the calculation of the full spin ladder of Fe_2S_2 compounds for which second-order perturbation theory could only be performed for high-spin states. Section 3.3 considers the relative spin-state

energies of several other iron complexes and examines the effects of including high local exchange (HLE)[13] modifications to the MC-PDFT exchange and correlation energies.

Section 4 contains the practically motivated work of this dissertation, namely the application of electronic structure theory to gas separations in metal-organic frameworks (MOFs). Section 4.1 addresses CO_2 capture in a copper paddle-wheel MOF and includes comparative tests of multiple active space choices to ensure that the minimal necessary active space is used for the calibration of force field parameters. The remaining three subsections all focus on metal-catecholates, with Section 4.2 using KS-DFT and CASPT2 in comparisons of different first-row transition metals for the capture of toxic NO and Section 4.3 applying the same approach for O_2/N_2 separation. Section 4.3 in particular includes detailed analysis of various active space choices for the CASPT2 results. Inspired by the promising results presented in Section 4.3, Section 4.4 describes a screening study to identify specific MOF structures for metal-catecholate modification as synthetic targets for the purpose of O_2/N_2 separation. The combination of Section 4.3 and Section 4.4 serves as an example of how electronic structure calculations on cluster models can be the initial step towards finding solutions to practical problems.

1.2 Theoretical Background and Methods

1.2.1 Wave Function Theory

At the heart of quantum chemistry is the non-relativistic time-independent Schrödinger equation,[14–16]

$$H|\Psi\rangle = E|\Psi\rangle \tag{1}$$

where $|\Psi\rangle$ is the wave function, H is the Hamiltonian operator, and E is the energy, an eigenvalue of the Hamiltonian operator applied to the wave function. Solving this equation analytically is known to be intractable for systems containing more than one electron,[17] but a great deal of progress has been made in finding simplifying assumptions and alternative approaches that allow for practically feasible solutions while still approaching or achieving chemical accuracy.

The Hamiltonian operator is

$$H = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2}$$

Proceeding from left to right, the terms correspond to the kinetic energy of the electrons, the kinetic energy of the nuclei, the attractive potential between the electrons and nuclei, the repulsive potential between electrons, and the repulsive potential between nuclei. (Unless otherwise noted, equations are in atomic units for simplicity of presentation.) Here N is the number of electrons, M the number of nuclei, M_A the mass of nucleus A divided by the mass of an electron, Z_A the nuclear charge of the nucleus A, r_{iA} the distance between the ith electron and nucleus A, r_{ij} the distance between the ith and jth electrons, and R_{AB} the distance between nuclei A and B. The Hamiltonian is often simplified by the Born-Oppenheimer approximation[18] (i.e., treating the nuclei as fixed

point charges, based on the assumption that the electrons are moving much faster than the nuclei).[15] This enables neglecting the kinetic energy of the nuclei and treating the nuclear-nuclear repulsion as a constant:

$$E_{nucl} = \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{3}$$

All that remains, then, is the electronic energy, the Hamiltonian for which is[15]

$$H_{elec} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{4}$$

The electron has three spatial coordinates, **r**, as well as one spin coordinate, ω, with the four coordinates combined being

$$\mathbf{x} = \{\mathbf{r}, \omega\} \tag{5}$$

The wave function itself, Ψ, can then be written as a function of spin orbitals, $\chi(\mathbf{x})$, (where the spin orbital, χ, is a product of a spatial function and a spin function), with a wave function corresponding to N electrons being $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$. A first approximation is to treat the electrons as non-interacting, with the corresponding Hamiltonian simply being

$$H_{non-interacting} = \sum_{i=1}^{N} h(i) \tag{6}$$

where $h(i)$ is an operator corresponding to the kinetic and potential energy of the non-interacting ith electron. The spin orbitals are then eigenfunctions of $h(i)$, so that

$$h(i)\chi_j(\mathbf{x_i}) = \varepsilon_j \chi_j(\mathbf{x_i}) \tag{7}$$

where ε_j is the eigenvalue, the energy of electron i in spin-orbital j. The wave function can then be written as a Hartree product,[15] i.e.,

$$\Psi^{HP}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) \ldots \chi_k(\mathbf{x}_N) \tag{8}$$

In order to satisfy the Pauli exclusion principle, the wave function must be antisymmetric, i.e., exchanging any two electrons should result in a change of sign[15–17]:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_N) = -\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N) \tag{9}$$

This is accomplished by taking a linear combination of Hartree products, which is then often written as a Slater determinant,[15,16,19]

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & & \chi_k(\mathbf{x}_2) \\ \vdots & & \ddots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix} \tag{10}$$

where $(N!)^{-1/2}$ is a normalization factor, χ is a spin orbital, and $\mathbf{x}$ is the coordinates of electron. The antisymmetry requirement is inherently satisfied, as all determinants change sign upon the exchange of any two rows (i.e., electrons).[17] A common shorthand notation treats the normalization factor as implicit and gives the diagonal of the Slater determinant in ket form[15]:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = |\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2)\cdots\chi_k(\mathbf{x}_N)\rangle \tag{11}$$

Further shorthand notation implicitly assumes the electrons are always labeled as $\mathbf{x}_1$, $\mathbf{x}_2$, $\cdots \mathbf{x}_N$, yielding[15]

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = |\chi_i\chi_j\cdots\chi_k\rangle \tag{12}$$

1.2.2 Hartree-Fock

One of the earliest and most influential attempts to address the Schrödinger equation is Hartree-Fock theory.[20,21] While not sufficiently accurate in and of itself for anything more than qualitative information,[16,17] it is the basis of many of the more refined methods currently methods in wide use.[2,6,7,15,17,22–25] Hartree-Fock relies on a single Slater determinant to describe an approximate wave function,

$$\Psi_0 = |\chi_i \chi_j \cdots \chi_k\rangle \tag{13}$$

and finds spin orbitals that minimize the energy

$$E_0 = \langle \Psi_0 | H | \Psi_0 \rangle \tag{14}$$

The energy is minimized to employ the variational principle, i.e., that an approximate wave function will always yield an energy that is higher than the actual energy, and therefore the lowest energy found will be the most accurate.[15,26] The energy of individual spin orbitals is found using the Fock operator

$$f(i) = h(i) + v^{HF}(i) \tag{15}$$

where $h(i)$, the core-Hamiltonian operator, is the sum of the kinetic energy and electron-nuclear repulsion terms,

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} \tag{16}$$

and $v^{HF}(i)$ is an average potential (or field), replacing the two-electron potential,

$$\sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{17}$$

This use of an average potential instead of explicit two-electron potentials is the most significant simplification of Hartree-Fock.[15]

The average potential, $v^{HF}(i)$, has two components, Coulomb and exchange, which for any given electron require summing over all other electrons.[15] For the first electron (assumed to be in spin orbital χ_a), the average potential is

$$v^{HF}(1) = \sum_b (J_b(1) - K_b(1)) \tag{18}$$

where $J_b(1)$ is the Coulomb operator, which simplifies electron-electron repulsion considerations by replacing the instantaneous position of the electron in spin orbital χ_b

(here labeled electron 2, with coordinates $\mathbf{x}_2$) with an average interaction over all space and spin coordinates,

$$J_b(1)\chi_a(1) = \left[\int \chi_b^*(2)\frac{1}{r_{12}}\chi_b(2)d\mathbf{x}_2\right]\chi_a(1) \tag{19}$$

thus giving a one-electron potential instead of a two-electron potential. The other term, $K_b(1)$, is the exchange operator, and represents a purely quantum phenomenon, specifically, the effect of "exchanging" electrons between spin orbitals:

$$K_b(1)\chi_a(1) = \left[\int \chi_b^*(2)\frac{1}{r_{12}}\chi_a(2)d\mathbf{x}_2\right]\chi_b(1) \tag{20}$$

(Note that while the Coulomb integral in $J_b(1)$ is local, i.e., it depends only on χ_b across all space and can therefore be calculated at the location of electron 1 ($\mathbf{x}_1$), the exchange integral depends on χ_a as well as χ_b across all space, and is therefore nonlocal.)[15] The Fock operator can thus be written

$$f(1) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_b (J_b(1) - K_b(1)) \tag{21}$$

effectively replacing the Hamiltonian operator for a given electron, yielding the Hartree-Fock equation for a single electron:

$$f|\chi_a\rangle = \varepsilon_a|\chi_a\rangle \tag{22}$$

where the spin orbital $|\chi_a\rangle$ is an eigenvalue of the Fock operator f and ε_a is the corresponding eigenvalue, i.e., the energy of the electron in spin orbital $|\chi_a\rangle$.[15] Returning to Equation (14), then, the total energy is

$$E_0 = \sum_a \varepsilon_a - \frac{1}{2}\sum_b \big(J_b(1) - K_b(1)\big) + E_{nucl} \tag{23}$$

Even with these simplifications, the dependence of v^{HF} on the spin orbitals renders the Hartree-Fock equations non-linear, necessitating an iterative procedure in finding the solution. This is accomplished by making an initial guess for the spin orbitals, calculating the average potential, and solving the Hartree-Fock Equation (22). The

solution provides a new set of spin orbitals, and the whole process begins again, continuing until there is no longer a significant change the potential field, v^{HF}, or in the energy from one iteration to the next. This is known as self-consistency, and Hartree-Fock-based methods are often referred to as self-consistent field (SCF) methods.[15]

Unfortunately, the average potential's description of Coulomb repulsion between two electrons glosses over a considerable amount of detail,[17] giving rise to what is commonly known as electron correlation. Correlation is often divided into two categories, usually named dynamical and non-dynamical (or static), but it is important to understand that these distinctions arise from which methods are used to recover some portion of the Coulomb correlation, and that there is inherent overlap between them.[17] Non-dynamical correlation is typically associated with near-degenerate electronic states that cannot be adequately described by one determinant. This is most famously illustrated by the failure of Hartree-Fock and other single-determinant methods to correctly describe the dissociation of the hydrogen molecule.[17] Solutions to this problem involve using multiple determinants, i.e., multireference methods, and they will be addressed in a later section. Dynamical correlation is assigned to the fact that the average field allows electrons to get closer together than true Coulombic repulsion would allow.[27] It is commonly addressed with methods that make use of the virtual space (i.e., the unoccupied orbitals), and these can use either a single or multireference determinant. Coupled-cluster is one of the most successful methods for treating dynamical correlation, as it treats all other relevant determinants as "excitations" of the reference determinant, but it is often cost-prohibitive for systems of interest and does not perform well if multiple reference determinants are needed.[17,28]

Perturbation theory is also often used as a post-Hartree-Fock method, with the most commonly known being Møller-Plesset. The Hamiltonian is partitioned as

$$H = H_0 + \lambda V \tag{24}$$

where H_0 is the Fock operator (plus E_{nucl}) and λV is the correlation potential, with λ being an ordering parameter for Taylor expansions of the eigenfunctions and eigenvalues[15]:

$$E_i = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \cdots \tag{25}$$

$$|\Phi_i\rangle = |i\rangle + \lambda \left|\Psi_i^{(1)}\right\rangle + \lambda^2 \left|\Psi_i^{(2)}\right\rangle + \cdots \tag{26}$$

The Taylor expansion is often truncated, (e.g., at the second order term, giving rise to names such as Møller-Plesset perturbation theory, second order, or MP2).[22] The major disadvantage to perturbation-based methods is that they are no longer variational, and therefore it can no longer be assumed that the lowest energy calculated is the most accurate.[29]

1.2.3 Density Functional Theory

One especially popular way of capturing correlation sidesteps many of the difficulties with wave function theory (WFT) all together. Density functional theory (DFT) determines the ground state energy simply from the probability density of the electronic wave function, commonly known as electron density.[27] Hohenberg and Kohn[30] proved that electronic density uniquely determines the Hamiltonian operator (or, as it is most commonly stated, that the energy is a functional of the density), and that the energy can be variational. They then proved the existence of a universal density functional from which the energy could then be calculated:

$$E = \underset{\rho \to N}{min} \left(F[\rho] + \int \rho(\mathbf{r})\, V_{Ne} d\mathbf{r} \right) \tag{27}$$

Here E is the total energy, $\underset{\rho \to N}{min}$ indicates the use of the variational principle to find the density, ρ, for N electrons that minimizes the total energy, $F[\rho]$ is the universal

functional, and $\int \rho(\mathbf{r})\, V_{Ne} d\mathbf{r}$ accounts for the nuclear-electron attraction potential. This potential is system-dependent,[27] but is easily calculated based on the positions of the nuclei:

$$V_{Ne} = -\sum_{i}^{N}\sum_{A}^{M}\frac{Z_A}{r_{iA}} \tag{28}$$

The universal density functional, $F[\rho]$, is not system-dependent (hence, "universal"). It has three major components:

$$F[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V_H[\rho(\mathbf{r})] + E_{ncl}[\rho(\mathbf{r})] \tag{29}$$

where $T[\rho(\mathbf{r})]$ is the kinetic energy, $V_H[\rho(\mathbf{r})]$ is the classical Coulomb repulsion, and $E_{ncl}[\rho(\mathbf{r})]$ is the collection of non-classical contributions, including self-interaction correction (which conveniently cancels in Hartree-Fock, but not necessarily in DFT), exchange, and electron correlation.[27] Unfortunately, the kinetic and non-classical functionals are not known,[27] and the universal density functional is unlikely to ever be known.[31,28]

Kohn and Sham[23] proposed methods for approximating the density functional using self-consistent methods similar to Hartree-Fock. This approach is known as Kohn-Sham density functional theory (KS-DFT). (Kohn-Sham DFT is the basis of all formulations that will be discussed in this section, and thus any subsequent references to "DFT" within this section are intended to mean "Kohn-Sham DFT", but this is not intended to deny the existence of other formulations or approaches.[11,32–36]) In Kohn-Sham DFT, a single Slater determinant is used to describe a system of N *non-interacting* electrons with the same electron density as the real (interacting) system. This system has the Hamiltonian

$$H_s = -\sum_{i}^{N}\frac{1}{2}\nabla_i^2 + \sum_{i}^{N} V_s(\mathbf{r}_i) \tag{30}$$

where $\frac{1}{2}\nabla_i^2$ is the kinetic energy of the ith electron and $V_s(\mathbf{r}_i)$ is the potential energy as a function of the electron's position.[27] V_s is an effective local potential, the nature of which is discussed in succeeding paragraphs.

As in Hartree-Fock, a determinant is constructed from a set of one-electron spin orbitals, ϕ_i, which guarantees the resulting wave function is antisymmetric. (The change in notation from χ to ϕ is to emphasize that the conceptual similarity with Hartree-Fock is not mathematically or physically identical.) The density is known from the orbitals:

$$\rho(\mathbf{r}) = \sum_i^N |\phi_i|^2 \tag{31}$$

where $\rho(\mathbf{r})$ is the density as a function of spatial coordinates. From the Hamiltonian in Equation (30) we get the Kohn-Sham operator

$$f^{\mathrm{KS}} = -\frac{1}{2}\nabla_i^2 + V_s(\mathbf{r}_i) \tag{32}$$

and so the energy of each Kohn-Sham orbital is

$$f_i^{\mathrm{KS}}\phi_i = \varepsilon_i\phi_i \tag{33}$$

As the electrons are "non-interacting", the total energy is a simple sum of the individual energies. The challenge, then, lies in finding densities that correspond to the density of the real, interacting system.[27]

In order to incorporate the non-interacting framework, Kohn and Sham[23] rearranged the density functional to separate terms arising from electron-electron interaction:

$$F[\rho] = T_s[\rho] + V_H[\rho] + E_{XC}[\rho] \tag{34}$$

where $T_s[\rho]$ is the *non-interacting* kinetic energy,

$$T_s[\rho] = -\frac{1}{2}\sum_i^N \langle\phi_i|\nabla^2|\phi_i\rangle \tag{35}$$

and $V_H[\rho]$ is simply the classical Coulomb repulsion,

$$V_H[\rho] = \frac{1}{2}\sum_i^N \sum_j^N \int\int |\phi_i(\mathbf{r}_1)|^2 \frac{1}{r_{12}} |\phi_i(\mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2 \tag{36}$$

The error resulting from the non-interacting kinetic energy approximation is combined with the non-classical effects in a new term, the exchange-correlation functional:

$$E_{XC}[\rho] = (T[\rho] - T_s[\rho]) + (V_{ee}[\rho] - V_H[\rho]) \tag{37}$$

where $E_{XC}[\rho]$ is the exchange-correlation functional, $T[\rho]$ is the true kinetic energy, $T_s[\rho]$ is the non-interacting kinetic energy, $V_{ee}[\rho]$ is the true electron-electron interaction potential, and $V_H[\rho]$ is the classic Coulomb electronic repulsion potential.[27] For the Kohn-Sham operator in Equation (32) to capture all of the energy, then, it is necessary for $V_s(\mathbf{r}_i)$ to account for the electron-nuclear potential (Equation (28)), the Coulomb repulsion, $V_H[\rho]$ (Equation (36)), and the exchange-correlation, $E_{XC}[\rho]$ (Equation (37)), so for a single electron

$$V_s(\mathbf{r}) = = -\sum_A^M \frac{Z_A}{r_{1A}} + \int \frac{\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + V_{XC}(\mathbf{r}_1) \tag{38}$$

V_{XC} is the potential arising from the exchange correlation, E_{XC}, and is defined simply as its derivative with respect to the density,

$$V_{XC} \equiv \frac{\delta E_{XC}}{\delta \rho} \tag{39}$$

The exact forms of V_{XC} and E_{XC} are unknown.[27]

As all unknown factors in Kohn-Sham DFT are placed in the exchange-correlation functional, and the exact form of the functional is unknown, finding effective approximations of the exchange-correlation functional forms the backbone of DFT development. One of the earliest approximations, which Kohn and Sham included in their original paper,[23] was to use the model system of the homogenous electron gas, which is known as local density approximation (LDA).[27] Including the gradient of the charge

density, $\nabla\rho(\mathbf{r})$, gives rise to the generalized gradient approximation (GGA) functionals. Meta-GGA functionals also include the Laplacian, $\nabla^2\rho(\mathbf{r})$, or the kinetic energy densities. Hybrid functionals, which are usually based on GGA or meta-GGA functionals, include some Hartree-Fock exchange.[37] Double hybrid functionals include Hartree-Fock exchange and MP2 correlation.[38,39] If the exact functional were known, DFT would give exact energies,[40] and there are many situations in which the current approximations perform extremely well.[28,37,41] Unfortunately, existing functionals do not have consistently good accuracy when non-dynamical correlation plays a large role, so in practice multiconfigurational methods are required in those cases.[11,27,28,42–44]

1.2.4 CSFs and CI

Before discussing multiconfigurational methods, a brief explanation of configuration state functions (CSFs) will be given. A significant disadvantage to Slater determinants, upon which so many of the above methods rest, is that they are not spin eigenfunctions, that is, they are eigenfunctions of the spin projection operator, S_z, but not necessarily of the total spin operator, S^2. CSFs are linear combinations of Slater determinants that satisfy both, i.e.,

$$S_z|k;S,m\rangle = m|k;S,m\rangle, m = -S, -S+1, \dots, +S \tag{40}$$

$$S^2|k;S,m\rangle = S(S+1)|k;S,m\rangle \tag{41}$$

where k is the CSF index, S is the spin quantum number, and m is the spin projection eigenvalue,

$$m = \frac{1}{2}\left(N_\alpha - N_\beta\right) \tag{42}$$

with N_α and N_β being the number of α- and β-spin electrons, respectively.[45] In multiconfigurational methods, the wave function is often expressed as a linear combination of CSFs:

$$|0\rangle = \sum_{\mu=1}^{\mu_{tot}} C_\mu^{(0)} |\mu\rangle \tag{43}$$

where $|0\rangle$ is the wave function, $|\mu\rangle$ is a given CSF, $C_\mu^{(0)}$ is a corresponding weighting coefficient, and μ_{tot} is the total number of CSFs. The coefficients are typically found using configuration interaction (CI), in which every configuration under consideration is treated as an excitation from a reference wave function. (CI can use a reference wave function that is single reference or multireference, and in a Slater determinant or CSF basis, but for the sake of discussion a multireference wave function in a CSF basis will be presented.) The coefficients can be treated as eigenvalues of a Hamiltonian matrix,

$$\mathbf{HC} = E\mathbf{C} \tag{44}$$

Where $\mathbf{C}$ is a vector containing the CSF coefficients and $\mathbf{H}$ is the Hamiltonian matrix in the CSF basis.[5,29]

In keeping with convention for CSFs, the electronic Hamiltonian of Equation (4) is written in second quantization:

$$H_{elec} = \sum_{p,q} h_{pq} E_{pq} + \frac{1}{2} \sum_{p,q,r,s} g_{pqrs} e_{pqrs} \tag{45}$$

where h_{pq} is the one-electron integral, corresponding to the kinetic energy and electron-nuclear attraction,

$$h_{pq} = \int \phi_p^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 - \sum_A \frac{Z_A}{r_1} \right) \phi_q(\mathbf{r}) d\mathbf{r} \tag{46}$$

and g_{pqrs} is the two-electron integral, corresponding to electron-electron repulsion,

$$g_{pqrs} = \int\int \frac{\phi_p^*(\mathbf{r}_1)\phi_r^*(\mathbf{r}_2)\phi_q(\mathbf{r}_1)\phi_s(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \tag{47}$$

The integrals have corresponding operators, with E_{pq} being the one-electron excitation operator,

$$E_{pq} = \sum_{\sigma} a_{p\sigma}^{\dagger} a_{q\sigma} \tag{48}$$

and e_{pqrs} being the two-electron excitation operator,

$$e_{pqrs} = E_{pq}E_{rs} - \delta_{qr}E_{ps} = \sum_{\sigma\tau} a_{p\sigma}^{\dagger} a_{r\tau}^{\dagger} a_{s\tau} a_{q\sigma} \tag{49}$$

where $a^{\dagger}$ is the creation operator, a is the annihilation operator, p, q, r, and s refer to molecular orbitals φ_p, φ_q, φ_r, and φ_s, and σ and τ refer to spin. The wave function is found by minimizing the energy, which in general requires optimized orbitals and CI coefficients, but for CI by itself the orbitals are not optimized.[5,29,45]

Assuming a normalized wave function, the energy is

$$\begin{aligned} E &= \langle 0|H_{elec}|0\rangle \\ &= \sum_{p,q} h_{pq}\langle 0|E_{pq}|0\rangle + \frac{1}{2}\sum_{p,q,r,s} g_{pqrs}\ \langle 0|e_{pqrs}|0\rangle \\ &= \sum_{p,q} h_{pq} D_{pq}^{(00)} + \frac{1}{2}\sum_{p,q,r,s} g_{pqrs}\ P_{pqrs}^{(00)} \end{aligned} \tag{50}$$

where

$$D_{pq}^{(00)} = \langle 0|E_{pq}|0\rangle = \sum_{\mu,\nu} C_{\mu}^{(0)} C_{\nu}^{(0)} \langle \mu|E_{pq}|\nu\rangle \tag{51}$$

is the first-order reduced density matrix and

$$\begin{aligned} P_{pq}^{(00)} &= \langle 0|E_{pq}E_{rs} - \delta_{qr}E_{ps}|0\rangle \\ &= \sum_{\mu,\nu} C_{\mu}^{(0)} C_{\nu}^{(0)} \langle \mu|E_{pq}E_{rs} - \delta_{qr}E_{ps}|\nu\rangle \end{aligned} \tag{52}$$

is the second-order reduced density matrix, with μ and ν referring to CSFs and $C^{(0)}$ being the coefficients.[5,29,45]

The Hamiltonian matrix elements, therefore, are

$$H_{\mu\nu} = \langle \mu | H_{elec} | \nu \rangle \tag{53}$$

with μ and ν being CSFs. Solving the eigenvalue problem of Equation (44) has been the subject of much attention, and there are many different methods in existence. Many implementations use the graphical unitary group approach (GUGA)[46], an efficient means of specifying the expansion space[45] and dealing with the coupling between CSFs.[7,47] Even more common are Davidson's algorithms for the efficient diagonalization of large matrices,[48–50] an especially important feature of which is the expansion vector, corresponding in this case to the CI coefficients (C), and the matrix vector product, σ,

$$\sigma_\mu = \sum_\nu H_{\mu\nu} C_\nu \tag{54}$$

which is used to update the CI coefficients in "direct CI" methods.[5,51] Even with these advances, however, CI can be quite computationally expensive, especially if every configuration in the given space is considered (i.e., a "Full CI"). However, there are many methods of restricting the CI space, either by limiting excitations or by orbital partitioning. Some methods that take the latter approach while also optimizing the orbitals are addressed in the following section.

1.2.5 MCSCF

In multiconfigurational methods, a zeroth-order wave function that inherently accounts for non-dynamical correlation is found using multiconfiguration self-consistent field (MCSCF), with dynamical correlation being again accounted for by a post-SCF (or, in this case, post-MCSCF) method.[9,45] The most popular MCSCF method is complete active space self-consistent field, or CASSCF.[5,6] CASSCF limits the full CI calculations by dividing the one-electron orbital space into inactive, active, and secondary subspaces. (Throughout this section the convention will be observed where active spaces are

identified by the number of active electrons N in the number of active orbitals n, or (N, n).) In practice, orbitals are placed in their respective subspaces by the user based on chemical intuition and experience. The active orbitals undergo full CI, i.e., all possible configurations within spin and spatial symmetry constraints are considered. The active space typically include any orbitals expected to be involved in a reaction or other interaction, but for small enough systems they can include the entire set of valence orbitals (or even more).[47,52] The inactive orbitals are kept doubly occupied in all configurations; core orbitals are almost always placed in the inactive space.[5] Secondary orbitals, also known as virtual orbitals, are unoccupied.[6] However, orbital rotations between subspaces (i.e., inactive-active, active-virtual, and inactive-virtual) are allowed when the orbitals are optimized as part of the variational procedure (with the details varying widely by implementation).[45] In CASSCF the energy is invariant to rotations *within* a given subspace, so such rotations do not need to be included.[5]

Unfortunately, the scaling of the CI expansion inherently limits the size of the active space that can be feasibly pursued,[47] and active spaces larger than eighteen electrons in eighteen orbitals, or (18,18), are still impossible for most computers,[53] although (20,20) has been achieved using massive parallelization.[54] However, it is known that over 99% of the configurations in the full CI space are what Ruedenberg called "deadwood", that is, they make no significant contribution to the wave function.[55–57] It is therefore possible to use larger active spaces if significant portions of the deadwood configurations can be eliminated in advance. Doing so has been the focus of decades of research, with the result of many techniques for limiting the CI expansion.[2] One common method is restricted active space self-consistent field theory (RASSCF),[7] which establishes three subspaces within the active space. The first (RAS1) consists of doubly occupied orbitals with a limited number of holes permitted. The third space (RAS3) contains unoccupied orbitals with a limited number of electrons permitted to be excited

into them. The second space (RAS2) still undergoes a full CI. Electrons can be excited from RAS1 into RAS2 or RAS3, or from RAS2 to RAS3 (given the constraints on RAS1 and RAS3), and all excitations are considered within RAS2. An even more flexible method generalized the idea behind RASSCF and was therefore named generalized active space self-consistent field theory (GASSCF).[2] Here an arbitrary number of subspaces is permitted (instead of the maximum of three), and each one is assigned a minimum and maximum occupancy (Figure 1). These occupancies are additive across subsequent subspaces. In both RASSCF and GASSCF, orbital rotations between the subspaces need to be considered, but the energy remains invariant within subspaces.[45] Note that CASSCF, RASSCF, and GASSCF can be performed as CI-only calculations as well by omitting the orbital optimization, leading to CASCI, RASCI, and GASCI, respectively.

The methods discussed in this section play a significant role in most of the work discussed in this dissertation. Additionally, it is worth noting that there has also been considerable progress in the development of density matrix renormalization group (DMRG)[58–60] as a way to perform complete active space calculations with much larger active spaces than were previously possible. Currently up to 84 electrons in 84 orbitals (84,84) has been performed,[61] and over 100 active orbitals are feasible.[62] While DMRG was not employed in any of the work featured in this dissertation, the large active spaces possible with DMRG are relevant to the challenging problem of selecting active orbitals, which will be addressed in Section 1.3. DMRG also has been successfully combined with MC-PDFT, which provides a way to obtain quantitative results with large active spaces (see Section 1.2.8).

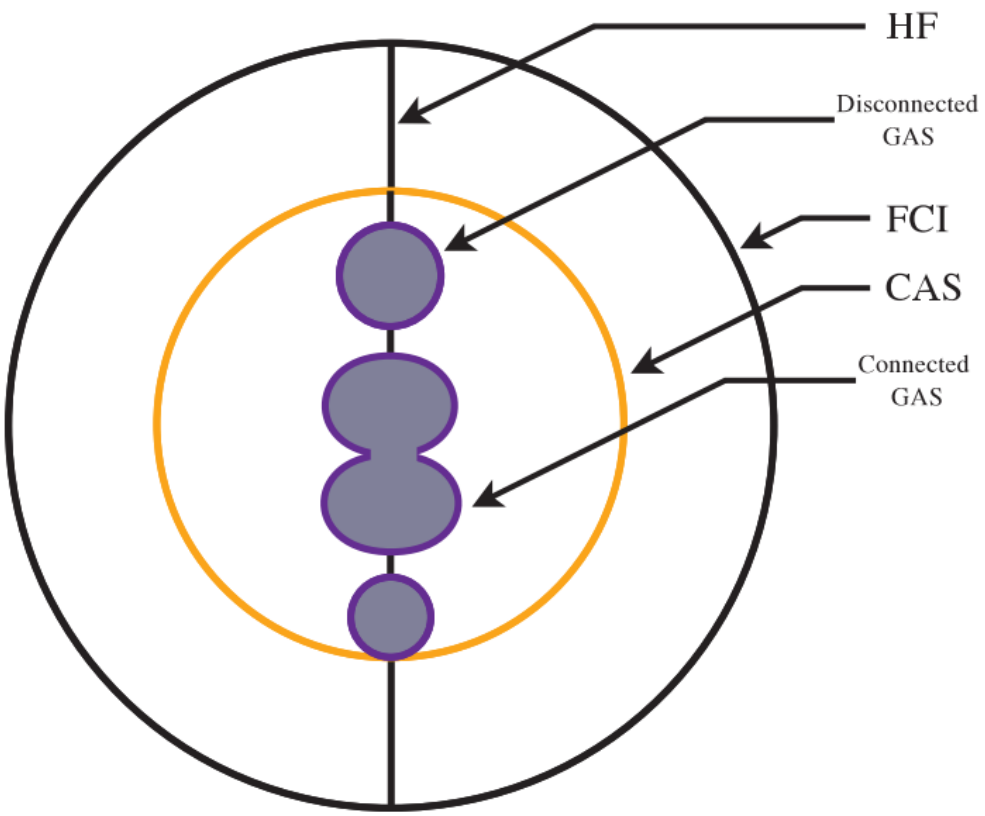


Figure 1: Schematic representation of the full CI space (large circle) subdivided into CAS (yellow) and GAS spaces (purple). Reprinted from Ma, D.; Li Manni, G.; Gagliardi, L. The Generalized Active Space Concept in Multiconfigurational Self-Consistent Field Methods. *J. Chem. Phys.* **2011**, *135*, 044128,[2] with the permission of AIP Publishing.

1.2.6 CASPT2

In all of the active space MCSCF methods discussed in Section 1.2.5, some dynamical correlation is neglected and obtaining quantitative results requires the application of a post-MCSCF method. The most popular[63] is second-order perturbation theory, or CASPT2[8,9] (when applied to a CASSCF reference wave function). CASPT2 operates similarly to MP2 (which was introduced at the end of Section 1.2.2), and by design reduces to MP2 when there are no orbitals in the active space.[8,9]

CASPT2 determines the second-order energy E_2 from the first-order wave function $|\Psi_1\rangle$ which it obtains by projecting single and double excitations from the reference (CASSCF) wavefunction $|0\rangle$ onto a modified zeroth-order Hamiltonian H_0:

$$H_0 = P_0 F P_0 + P_K F P_K + P_{SD} F P_{SD} + P_{TQ\ldots} F P_{TQ\ldots} \tag{55}$$

where

$$F = \sum_{pq} f_{pq} E_{pq} \tag{56}$$

is a one-electron operator,

$$P_0 = |0\rangle\langle 0| \tag{57}$$

is the projection operator onto V_0, the configurations subspace spanned by the CASSCF wave function, P_K is the projection operator onto V_K, the configurations subspace spanned by the orthogonal complement to the CASSCF wave function, P_{SD} is the projection operator onto V_{SD}, the configurations subspace spanned by single and double excitations from the CASSCF wave function excluding those already accounted for by V_0 and V_K, and, $P_{TQ\ldots}$ is the projection operator onto $V_{TQ\ldots}$, the configuration subspace containing all higher-order excitations not already accounted for by the other three subspaces. The $V_{TQ\ldots}$ configurations do not contribute to the second-order energy and the

configurations in V_K do not interact with the reference wave function $|0\rangle$, leaving only the V_{SD} configurations contributing to the first-order wave function.[9,64]

Defining a perturbation operator V as

$$V = H - H_0 \tag{58}$$

where H is the Hamiltonian operator and H_0 is the zeroth-order Hamiltonian, the first-order wave function can be found:

$$|\Psi_1\rangle = \sum_{i=1}^{i_{tot}} C_i|i\rangle = -\sum_{i=1}^{i_{tot}} \frac{\langle i|V|0\rangle}{E_i^{(0)} - E^{(0)}}|i\rangle \tag{59}$$

where i_{tot} is the total number of configurations in V_{SD}, each $|i\rangle$ is a specific configuration, C_i is a coefficient similar to the CI coefficients of $|0\rangle$ in Equation (43), $E_i^{(0)}$ is the eigenvalue of the zeroth-order Hamiltonian H_0 applied to $|i\rangle$, and $E^{(0)}$ is the eigenvalue of $|0\rangle$. The second-order energy, E_2, is therefore[9,64]

$$E_2 = \langle 0|V|\Psi_1\rangle = \sum_{i=1}^{i_{tot}} C_i\langle 0|V|i\rangle = -\sum_{i=1}^{i_{tot}} \frac{|\langle 0|V|i\rangle|^2}{E_i^{(0)} - E^{(0)}} \tag{60}$$

The early implementations of CASPT2 were found to underestimate open-shell energies, resulting in corresponding errors in bond energies and excitation energies when the number of paired electrons changes. These errors were assigned to problems with the Fock matrix, i.e., the matrix with elements f_{pq} from Equation (56), for partially occupied (i.e., active) orbitals. The diagonal elements of the Fock matrix correspond to the negative of the ionization potential (IP) for doubly occupied (inactive) orbitals and the negative of the electron affinity (EA) for unoccupied (virtual) orbitals, but for active orbitals the corresponding diagonal elements are weighted averages of –IP and –EA. For excitations into or out of partially occupied orbitals, the denominator in Equation (60) for the calculation of the second-order energy becomes too small. Ghigo et al. proposed a correction in the form of a shift to the affected elements that would ideally replace the

energy with –EA when exciting into an active orbital and with –IP when exciting out of an active orbital. However, lacking a general way to accomplish that goal, they employed an average shift parameter ϵ with an empirically determined value.[65] This protocol is generally known as the IPEA shift and it is now standard practice to employ it with a default value of 0.25 a.u. for ϵ.[64] However, it is a matter of ongoing debate as to what value is appropriate for different classes of systems, or even whether IPEA shift should be used at all.[64,66] The IPEA shift can be used as an empirical parameter that is optimized for individual problems,[67] but this approach has been criticized by those seeking to preserve *ab initio* character in CASPT2.[64] CASPT2 also suffers from an "intruder state" problem, in which at certain geometries there can be erroneously large contribution from certain configurations due to near-zero values in the energy denominator. This can often be resolved by the application of an imaginary level shift to the denominator.[68]

Despite these challenges, second-order perturbation theory has been combined not only with CASSCF (i.e., CASPT2), but also with RASSCF, GASSCF, and DMRG (RASPT2,[10] GASPT2,[69] and DMRG-CASPT2,[70] respectively). However, the computational cost of CASPT2 for large active space sizes is even greater than for the preceding static correlation step due to the need to calculate higher-order (up to fourth) reduced density matrices that scale as poorly as N^8, where N is the number of active orbitals.[70–72] The CASPT2 step in a CASSCF/CASPT2 calculation begins to dominate timing and memory requirements at about fourteen electrons in fourteen orbitals (14,14),[71,72] making it impractical for active spaces with more than 14 orbitals,[70] despite CASSCF calculations of (18,18) active spaces being possible.[53] Some larger active spaces using a restricted active space reference function are included in Sections 2.2 and 3.2, but, as will be seen in those sections, RASPT2 scales very poorly. Even DMRG-CASPT2 is feasible only for up to 30 orbitals,[70] in contrast to 100 or more for the DMRG step.[62]

CASPT2 and RASPT2 play a significant role in the work presented in this dissertation. As mentioned above, Sections 2.2 and 3.2 include cases where CASPT2 or RASPT2 calculations would have been impractical from the standpoint of computational expense. Similar concerns play an indirect role in active space selection in Sections 4.2 and 4.3 (see Section 1.3 for further details). All work in Section 3 involves comparisons between MC-PDFT (see Section 1.2.8) and CASPT2 and/or RASPT2, and all work in Section 4 (on gas separations in MOFs) relies on CASPT2 results for validation of other methods employed.

1.2.7 SplitGAS

In contrast to the perturbative inclusion of single and double excitations of CASPT2, SplitGAS[3] offers the potential to recover dynamic correlation through the perturbative enlargement of the active space, which allows higher excitations while improving affordability with constraints on how configurations are permitted to interact. SplitGAS extends the active space by separating it into a primary (P) space and an extended (Q) space. The P space corresponds to the active spaces of CASSCF, RASSCF, or GASSCF, and can be further subdivided as those respective ansatzes permit. The Hamiltonian matrix for the full P+Q space would be intractable, so it is approximated by a dressed matrix U with elements

$$U_{mn} = H_{mn} - \sum_{\gamma} \frac{H_{m\gamma} H_{\gamma n}}{H_{\gamma\gamma} - E} \tag{61}$$

where m and n refer to CSFs in the P space and γ refers to CSFs in the Q space. Every element U_{mn} is perturbed by interaction with every Q space CSF γ, thus accounting for P-Q interaction, while Q-Q interaction is taken only on the diagonal (Figure 2).

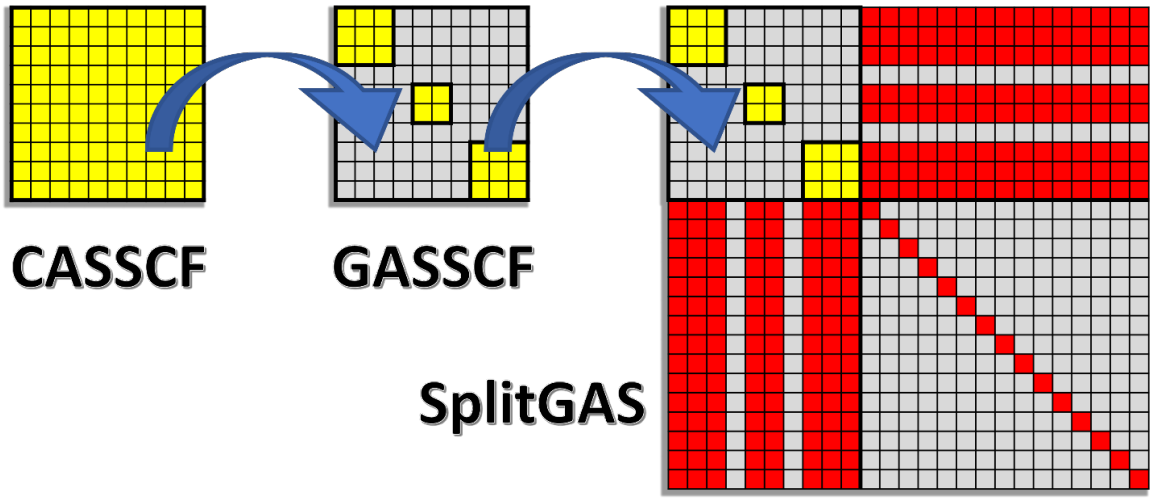


Figure 2: Schematic of the CASSCF, GASSCF, and SplitGAS active spaces, with the P space in yellow and the Q space in red. Reproduced with permission from Vogiatzis, K. D.; Li Manni, G.; **Stoneburner, S. J.**; Ma, D.; Gagliardi, L. Systematic Expansion of Active Spaces beyond the CASSCF Limit: A GASSCF/SplitGAS Benchmark Study. *J. Chem. Theory Comput.* **2015**, *11*, 3010–3021.[73] Copyright 2015 American Chemical Society.

Therefore, while the U matrix is dressed with elements corresponding to the P and Q space elements of H, the dressed U matrix has the dimensions of the P space only. Thus, in matrix form,

$$U^{PP} = H^{PP} - H^{PQ}(H_{diag}^{QQ} - E)^{-1}H^{QP} \tag{62}$$

This is the diagonal approximation of Löwdin's partitioning technique[74]. The CI coefficients are obtained through a direct CI method via a σ-vector of size (P):

$$\sigma_m = \sum_n^P U_{mn}C_n = \sum_n^P \left(H_{mn} - \sum_\gamma^Q \frac{H_{m\gamma}H_{\gamma n}}{H_{\gamma\gamma} - E} \right) C_n \tag{63}$$

with matrix form

$$\sigma^P = [H^{PP} - H^{PQ}\left(H_{diag}^{QQ} - E\right)^{-1}H^{QP}]C^P \tag{64}$$

where C_n and C^P are the CI coefficients and the corresponding matrix, respectively. The perturbative correction is included only periodically, in what are termed "macroiterations", with the intermediate iterations being normal CI iterations in the P

space. In order to prevent memory overload (by requiring the storage of a complete P or Q space-sized vector), the σ-vector is broken up into batches organized by CSF occupation class, which are ways of organizing the CSFs allowed under the minimum and maximum occupancies for each space by the actual occupancies of each space in the given CSF.[3,51]

SplitGAS is the primary focus of Section 2.1, which benchmarks the performance of SplitGAS for several different systems.

1.2.8 MC-PDFT

Both CASPT2 and SplitGAS work within the framework of wave function theory and interacting electronic configurations, but another approach is to use MCSCF together with density functional theory. Many attempts have been made to combine multiconfigurational WFT with DFT in hopes that static correlation could be dealt with by WFT and the dynamical correlation could be included with DFT in a much less computationally expensive manner than methods such as CASPT2. However, some degree of dynamical correlation energy is included in the WFT energy that is also included in the DFT energy, and it is extremely difficult to eliminate this doubly-counted energy in combined WFT and DFT methods. The proper way to interface the WFT and DFT methods is also not straightforward, as the spin densities of multiconfigurational wave functions are incompatible with extant KS-DFT functionals.[11] Multiconfiguration pair-density functional theory (MC-PDFT)[11,12] resolves both of the above difficulties by taking only the kinetic energy and classical Coulomb contributions to the wave function energy and calculating the rest of the energy by applying an on-top density functional to the density and on-top pair density obtained from the wave function calculation.

The MCSCF energy expression is the same as the one given in Equation (50):

$$E_{MCSCF} = \langle 0|H_{elec}|0\rangle \tag{50}$$

$$= \sum_{p,q} h_{pq}\langle 0|E_{pq}|0\rangle + \frac{1}{2}\sum_{p,q,r,s} g_{pqrs}\ \langle 0|e_{pqrs}|0\rangle$$

$$= \sum_{p,q} h_{pq} D_{pq}^{(00)} + \frac{1}{2}\sum_{p,q,r,s} g_{pqrs}\ P_{pqrs}^{(00)}$$

where

$$H_{elec} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \tag{4}$$

The first term in the energy expression corresponds to the one-electron terms in H_{elec} (i.e., the electronic kinetic energy and the electron-nuclear attraction) and depends on the one-electron density matrix. The second term in the energy expression corresponds to the electron-electron interaction and depends on the two-electron density matrix. In MC-PDFT, this second term is replaced by two terms:

$$E_{MC-PDFT} = \sum_{p,q} h_{pq} D_{pq}^{(00)} + \frac{1}{2}\sum_{p,q,r,s} g_{pqrs} D_{pq}^{(00)} D_{rs}^{(00)} + E_{ot}[\rho, \Pi] \tag{65}$$

The classical Coulomb repulsion between electrons is now obtained using a product of one-electron density matrices instead of using the two-electron density matrix.[11] The difference between the exact kinetic energy and the MCSCF-calculated kinetic energy and the energy arising from exchange and correlation is all included in the last term, the energy calculated with the on-top density functional. Because no correlation energy contribution from MCSCF is included, there is no double-counting of correlation energy.[75]

The on-top density functional is written in terms of ρ and Π, the electron density and on-top pair density, respectively. Π can be understood as the probability of finding two electrons in one place[76] and is defined[77] as

$$\Pi(\mathbf{r}) = \binom{N}{2} \int |\Psi(x_1, x_2, \ldots, x_N)|^2 \mathrm{d}\sigma_1, \mathrm{d}\sigma_2, \ldots, \mathrm{d}\sigma_N \, \mathrm{d}\mathbf{r}_3, \mathrm{d}\mathbf{r}_4, \ldots, \mathrm{d}\mathbf{r}_N |_{\mathbf{r}_1=\mathbf{r}_2=\mathbf{r}} \tag{66}$$

where N is the number of electrons, $\mathbf{r}_i$ is the three spatial coordinates of electron i, σ_i is the spin coordinate of electron i, and x_i is a space-spin variable $(\mathbf{r}_i, \sigma_i)$. The first generation of pair-density functionals are translations of GGA KS-DFT functionals. The original GGA functionals E_{xc} depend on the total density ρ,

$$\rho(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r}) \tag{67}$$

the spin magnetization density m,

$$m(\mathbf{r}) = \rho_\alpha(\mathbf{r}) - \rho_\beta(\mathbf{r}) \tag{68}$$

and their respective gradients ρ' and m'. (In Equations (67) and (68), $\rho_\alpha(\mathbf{r})$ and $\rho_\beta(\mathbf{r})$ are the densities of spin-up and spin-down electrons, respectively.) The translation leaves the dependence on ρ and ρ' unchanged while replacing m and m':[11]

$$E_{ot}[\rho(\mathbf{r}), \Pi(\mathbf{r})] = E_{xc}\left(\rho(\mathbf{r}), \begin{cases} \rho(\mathbf{r})(1-R)^{1/2} & if\ R \le 1 \\ 0 & if\ R > 1 \end{cases}, \rho'(\mathbf{r}), \begin{cases} \rho'(\mathbf{r})(1-R)^{1/2} & if\ R \le 1 \\ 0 & if\ R > 1 \end{cases}\right) \tag{69}$$

where R is a functional of ρ and Π:[11]

$$R(\mathbf{r}) = \frac{4\Pi(\mathbf{r})}{\rho(\mathbf{r})^2} \tag{70}$$

Because the spin-dependent m has been replaced with spin-free ρ and Π, there is no concern about the spin symmetry of the MCSCF wave function leading to spin densities unusable by DFT.[11,78]

Note that for a single Slater determinant $m = \rho(\mathbf{r})(1-R)^{1/2}$ and $R \le 1$,[79] but R can be greater than 1 for multiconfiguration wave functions. Therefore, the new terms in Equation (69) are set to 0 when R > 1 in order to ensure that the square roots are real-valued.[11] This form results in a discontinuity which precludes the inclusion of the gradient of the on-top pair density in the translation scheme of Equation (69). In order to

resolve the discontinuity an alternative "fully translated" scheme was proposed that ensures smoothness with a more complicated stepwise function with several additional parameters (see Ref. [78] for more details). So far the local spin-density approximation LSDA,[23] the Becke-Lee-Yang-Parr (BLYP)[80,81] functional, the Perdew-Berke-Ernzerhof (PBE)[82] functional, and the revised PBE (revPBE)[83] functional have been translated (tLSDA, tPBE, trevPBE, and tBLYP, respectively)[11] and fully translated (ftLSDA, ftPBE, ftrevPBE, and ftBLYP, respectively).[78]

Section 3 is exclusively devoted to papers with MC-PDFT, all of which demonstrate a significant cost advantage of MC-PDFT over CASPT2 and RASPT2. Section 3.1 and 3.3 include timing comparisons between CASPT2 and tPBE for organic diradicals and iron complexes, respectively, and Section 3.2 shows that MC-PDFT can be applied to spin states of diiron compounds that are completely intractable with RASPT2. Sections 3.1 and 3.2 also both serve to demonstrate that the limiting factor in the affordability of an MC-PDFT calculation is the preceding MCSCF step. With regard to that issue, the combination of MC-PDFT with DMRG (DMRG-PDFT) was recently developed and demonstrated on the singlet-triplet gaps of polyacenes and polyacetylenes with active spaces as large as (30,30),[84] and another study on iron porphyrin included a (34,35) active space for DMRG-PDFT.[85] DMRG-PDFT is comparable to DMRG alone in terms of computational expense,[84] unlike the CASPT2 portion of DMRG-CASPT2,[70] and thus DMRG-PDFT could theoretically be applied to any active spaces that can be affordably treated with DMRG alone (i.e., over 100 active orbitals[62]). While DMRG-PDFT was too recent a development to be included in the work presented in this dissertation, it would be a logical next step, especially for the iron-sulfide systems presented in Section 3.2.

1.3 Active Space Selection

CASPT2, SplitGAS, and MC-PDFT all require the user to choose active spaces for the preceding MCSCF calculation. (SplitGAS also requires similar decisions during the execution of the method itself, especially regarding the secondary (Q) space.) The proper selection of active orbitals has been the topic of much discussion, and a complete review is not possible here. In brief, the standard approach is trial-and-error,[86] with guess active spaces selected based on the user's knowledge and experience (often referred to as "chemical intuition") and successively improved based on the performance of the guesses. The *ad hoc* nature of this procedure means that a high level of expertise is required[28,52,87] and hinders efforts to make systematic improvements in active space methods.[88,89]

A variety of ways to bring more order and user-friendliness to active space selection have been suggested, ranging from simple guidelines[47] to more involved decision-making schemes or automatization protocols based on atomic orbital contributions,[25,86] molecular orbital occupation numbers,[90] molecular orbital energies,[91] or molecular orbital entanglement entropies.[88,89,92] However, the recommendations generated with these procedures are not always affordable and few (if any) of them have yet demonstrated whether they can be generally applied. In practical application, therefore, it is still common for active spaces to be selected through trial-and-error.[86]

The research presented in this dissertation provides examples of some of these approaches and illustrates the challenges involved in active space selection. The practically motivated work in Section 4 involved significant trial-and-error, although occupation numbers play a role in the final selection. In particular, Section 4.3 provides detailed reasoning behind the active space selections for an array transition metal-catecholates and includes some discussion of alternative choices that were deemed less

satisfactory. Section 3.2 demonstrates one of the more challenging aspects of active space selection, namely, that the user has to make affordability of the calculation a primary concern. Related to that, Sections 2.1 and 4.1 include some comparisons of different active space sizes for the systems in question. Section 2.1 also features systematic expansion of some selected active spaces using CASSCF, GASSCF, and SplitGAS. The most systematic examples, however, are in Section 2.2 and 3.1, which employ the CPO active space selection scheme with CASPT2 and MC-PDFT, respectively. Even in those cases, computational cost is again a concern and necessitates the use of RASSCF and RASPT2 for the larger active spaces.

1.4 Applications: Gas Separations in Metal-Organic Frameworks

In Section 4 of this dissertation, KS-DFT and multireference WFT are used in practical applications for gas separations in metal-organic frameworks (MOFs).[93] MOFs are a class of material under intense scrutiny due to high expectations for a wide variety of potential applications. MOFs are the product of considerable effort toward designed crystalline solids, namely, a process named "reticular synthesis" by Yaghi *et al.*[94] as a subclass of crystal engineering. MOFs are constructed from two primary classes of components: metals or metal-containing nodes, known as secondary building units (SBUs), and organic linkers. The SBUs are composed of metals with carbons and oxygens added in such a way as to provide directionality for binding to the linkers.[94,95] This, combined with the rigidity of the linkers, enables stable and consistent forming of repeated geometrical structure.[94,96] There are many potential applications for MOFs, including catalysis,[97–103] gas separation,[103–112] gas storage,[113,114] drug delivery,[115,116] and chemical sensing.[117–121] The body of literature devoted to MOFs is far too large to

address in detail, but many reviews have been published that may be of use to interested readers.[28,101–103,110–112,120–125]

MOFs have several advantages that have made them of interest to researchers. Their rigid structure provides high porosity[94,96] and prevents the aggregation that occurs in some other nanoscale compounds, leading to high surface area.[96] The variety of possible building blocks means that there is an infinite number of possible MOFs, suggesting many possibilities for fine-tuning application and functionality.[117] The rigid structure also lends itself well toward containing strongly bound metal centers with open coordination sites, which provides many unique opportunities for chemical binding and catalysis.[28] While these sites are most often part of the SBUs,[105] there is also a growing interest in generating such sites on the linkers using post-synthetic modification.[126,127] More generally, many new MOFs can be developed in a logical, orderly fashion,[94] and MOF design therefore readily lends itself to synergy with theoretical methods as supplements to existing experimental work and as the directors of the most likely useful directions.[28,118] Section 4.1 is in the former category, while Section 4.2, 4.3, and 4.4 are in the latter category. Some brief comments pertaining to the work presented in Section 4 are given here, but more detailed background information can be found in the individual introductory sections of each paper (Sections 4.1.2, 4.2.2, 4.3.2, and 4.4.2).

Section 4.1 is about working to theoretically account for the good experimental performance of a copper paddle-wheel MOF for capturing CO_2. CO_2 capture has been an especially popular topic in metal-organic frameworks due to the desire to reduce or eliminate CO_2 atmospheric output in industrial processes.[103] In previous work it had been shown that KS-DFT failed to reproduce experimental results. CASPT2 is thus used to confirm the hypothesis that correlation between the two Cu atoms in the paddle-wheel leads to multireference character, and thus a higher level of theory is necessary even for

calculations on the whole MOF structure. The CASPT2 calculations also validate the choice of theory used in further periodic calculations. The work shows how high-level WFT can play a pivotal role for challenging systems even when working on large-scale systems such as MOFs.

Sections 4.2, 4.3, and 4.4 are all predictive in nature and feature metal-catecholates, which can be added to MOFs post-synthetically to improve targeted functionality.[128] Section 4.2 uses a combination of KS-DFT and CASPT2 to screen first-row transition metals in a model catecholate system for toxic gas capture, which has been a target application of MOFs for some time[110,111] and for which metal-catecholates have had demonstrable success.[129] A similar protocol is used in Section 4.3 in order to screen for O_2-selective air separation potential, which is a high-value target due to the many applications that require pure O_2 gas.[130] Based on the work in Section 4.3 and additional KS-DFT calculations, Section 4.4 screens a large database of experimental structures[131] and identifies specific MOFs and metal-catecholates as synthetic targets. Section 4.3 and Section 4.4 are collectively motivated by an interest in carrying predictions forward on a path towards an industrial product, which is an ongoing concern in MOF development.[124]

In summary, Sections 2 and 3 look at ways to improve high-level multireference WFT methods involving active space selection, while Section 4 demonstrates multiple ways in which methods such as those considered in Sections 2 and 3 can make contributions toward solutions to real-world problems.

List of Papers Featured in this Dissertation

Section 2.1: Vogiatzis, K. D.; Li Manni, G.; **Stoneburner, S. J.**; Ma, D.; Gagliardi, L. Systematic Expansion of Active Spaces beyond the CASSCF Limit: A GASSCF/SplitGAS Benchmark Study. *J. Chem. Theory Comput.* **2015**, *11*, 3010–3021.

Section 2.2: **Stoneburner, S. J.**; Shen, J.; Ajala, A. O.; Piecuch, P.; Truhlar, D. G.; Gagliardi, L. Systematic Design of Active Spaces for Multi-Reference Calculations of Singlet–Triplet Gaps of Organic Diradicals, with Benchmarks Against Doubly Electron-Attached Coupled-Cluster Data. *J. Chem. Phys.* **2017**, *147*, 164120.

Section 3.1: **Stoneburner, S. J.**; Truhlar, D. G.; Gagliardi, L. MC-PDFT Can Calculate Singlet-Triplet Splittings of Organic Diradicals. *J. Chem. Phys.* **2018**, *148*, 064108.

Section 3.2: Presti, D.; **Stoneburner, S. J.;** Truhlar, D. G.; Gagliardi, L. Full Correlation in a Multiconfigurational Study of Bimetallic Clusters : Restricted Active Space Pair-Density Functional Theory Study of [2Fe-2S] Systems. *J. Phys. Chem. C* **2019**, *123*, 11899–11907.

Section 3.3: **Stoneburner, S. J.**; Truhlar, D. G.; Gagliardi, L. Affordable and Accurate Transition Metal Spin-State Energetics via MC-PDFT Using tPBE with High Local Exchange. **2019**, to be submitted for publication.

Section 4.1: Ongari, D.; Tiana, D.; **Stoneburner, S. J.**; Gagliardi, L.; Smit, B. Origin of the Strong Interaction between Polar Molecules and Copper(II) Paddle-Wheels in Metal Organic Frameworks. *J. Phys. Chem. C* **2017**, *121*, 15135–15144.

Section 4.2: **Stoneburner, S. J.**; Livermore, V.; McGreal, M. E.; Yu, D.; Vogiatzis, K. D.; Snurr, R. Q.; Gagliardi, L. Catechol-Ligated Transition Metals: A Quantum

Chemical Study on a Promising System for Gas Separation. *J. Phys. Chem. C* **2017**, *121*, 10463–10469.

Section 4.3: **Stoneburner, S. J.**; Gagliardi, L. Air Separation by Catechol-Ligated Transition Metals: A Quantum Chemical Screening. *J. Phys. Chem. C* **2018**, *122*, 22345–22351.

Section 4.4: Demir, H.; **Stoneburner, S. J.**; Jeong, W.; Ray, D.; Zhang, X.; Farha, O. K.; Cramer, C. J.; Siepmann, J. I.; Gagliardi, L. Metal-Organic Frameworks with Metal Catecholates for O_2/N_2 Separation. *J. Phys. Chem. C* **2019**, *123*, 12935-12946.

2 Systematic Active Space Selection

2.1 Systematic Expansion of Active Spaces beyond the CASSCF Limit: A GASSCF/SplitGAS Benchmark Study

This section describes the outcome of a collaborative research project carried out by Konstantinos D. Vogiatzis, Giovanni Li Manni, Samuel J. Stoneburner, and Dongxia Ma (and advised by Laura Gagliardi). A report on this research project has been published.[73]

Samuel J. Stoneburner performed CASSCF, GASSCF, and SplitGAS calculations on furan and pyrrole.

2.1.1 Overview

The applicability and accuracy of the generalized active space self-consistent field, GASSCF, and SplitGAS methods are presented. The GASSCF method enables the exploration of larger active spaces than the conventional complete active space SCF, CASSCF, by fragmentation of a large space into subspaces and by controlling the interspace excitations. In the SplitGAS method, the GAS configuration interaction, CI, expansion is further partitioned in two parts: the principal, which includes the most important configuration state functions, and an extended, containing less relevant but not negligible ones. An effective Hamiltonian is then generated, with the extended part acting as a perturbation to the principal space. Excitation energies of ozone, furan, pyrrole, nickel dioxide, and copper tetrachloride dianion are reported. Various partitioning schemes of the GASSCF and SplitGAS CI expansions are considered and compared with the complete active space followed by second-order perturbation theory, CASPT2, and multi reference CI method, MRCI, or available experimental data. General guidelines for the optimum applicability of these methods are discussed together with their current limitations.

2.1.2 Introduction

The treatment of the correlated motions of electrons constitutes one of the still unsolved challenges of modern electronic structure theory.[17] Correlation energy arises from the erroneous description of the Coulomb repulsion between electrons[132] in the Hartree-Fock theory and can be classified in two types, dynamical correlation and non-dynamical or static correlation. Dynamical correlation refers to the instantaneous electron repulsion and in wave-function methods it is usually recovered by electron excitations from the zero-order wave function to the virtual space. Effective approximations from

many-body perturbation theory (MBPT)[133] and coupled-cluster theory (CC)[134] provide a quantitative description of dynamical correlation. In density functional theory (DFT), dynamical correlation is instead accounted for by the functional choices. Non-dynamical correlation is important for atoms or molecules with nearly degenerate orbitals and with ground or excited states that cannot be described with a single Slater determinant. Typically, such electronic systems are described by a zero-order multiconfigurational wave function, which introduces the non-dynamical correlation.[5,45]

One of the most commonly used multiconfigurational methods are the complete active space self-consistent field (CASSCF)[6] method and its extension through second-order perturbation theory (CASPT2).[9] In CASSCF, the configuration space is determined by the electrons and orbitals included in the active space. A full configuration interaction (FCI) wave function is generated within this space and the orbitals are variationally optimized simultaneously with the configuration coefficients. Redundancies in the orbital rotations are eliminated by accounting only for inactive-active, active-virtual and inactive-virtual rotations. CASPT2 includes dynamical correlation and it has been successful in the description of excited states,[135] magnetic properties,[136] and dissociation energies.[67,137,138] The factorial scaling of the size of the Hilbert space with respect to the size of the active space limits the applicability of CASSCF/CASPT2, which cannot be applied to molecular systems that need active spaces larger than 16 electrons in 16 orbitals.

The exploration of larger active spaces beyond the standard multiconfigurational wave function-based methods is an active field of research. Partitioning of the complete active space (CAS) to smaller subspaces subspaces can significantly reduce the number of configuration state functions (CSFs) without affecting the accuracy of the results. Typically, most of the configurations in a configuration interaction (CI) expansion have a

minor contribution to the total wave function, i.e. they have a nearly zero coefficient in the configuration expansion. A carefully chosen partition can exclude such configurations, and therefore allow the use of larger active spaces. The selection of dominant configurations for a specific state simplifies the wave function without affecting the accuracy of the method. This concept has been successfully used in multireference CI (MRCI) studies.[139–144]

The generalized active space (GAS)[2,145] scheme is an attractive technique for eliminating negligible configurations from the configuration space. In GAS, multiple active spaces are chosen, and inter-space excitations are controlled by the user's choice. A suitable choice of the GAS spaces removes the irrelevant configurations from the CI expansion and keeps only configurations with relevant weights. With GASSCF larger active spaces are accessible than in conventional CASSCF, like for example the GAS5(20,32) for the Gd dimer:[2] 20 electrons in 32 orbitals with five GAS spaces. The occupation restricted multiple active spaces (ORMAS)[146] method relies upon a similar concept. The restricted active space (RAS) scheme used in RASSCF[7,143] and RASPT2[10] is a special case of the GAS scheme, where only three subspaces are chosen, namely RAS1, RAS2 and RAS3. In the reference space, the RAS1 orbitals are fully occupied, the RAS3 orbitals are empty, and the RAS2 orbitals have occupation ranging from 0 to 2. Restrictions are imposed on the RAS1 and RAS3 excitation level: a predefined maximum number of *holes* in RAS1 and a maximum number of *particles* in RAS3 determine the size of the wave function together with the excitations in the RAS2 space.

A further simplification of the CI expansion is achieved by the SplitCAS[147] method, which is based on partition techniques and targets the recovery of the missing dynamical correlation. In SplitCAS, a CAS CI expansion is partitioned in two subspaces, a principal space (P) and a much larger extended space (Q). Löwdin’s partitioning

techniques[74] approximated to the second order are used to reduce the initial eigenvalue problem to a problem of size equal to the size of the P-space. Davidson's "Reduced Model Space" approach[148,149] and Shavitt's "B_k" method[150] constitute modifications of this partition technique and share similarities with SplitCAS. Recently, the Löwdin's partitioning technique was applied successfully in the spin-flip framework through a quasi-degenerate approximation.[151,152] The more flexible SplitGAS[3] method allows orbital based partitioning of the P- and Q-spaces, as the GAS scheme is used to select configurations to add in P or in the Q part of the CI expansion. The SplitCAS and SplitGAS approaches capture both static and dynamical correlation, as shown in the case of the Cr_2 dimer.[3]

In this article, a benchmark study of the GASSCF and SplitGAS methods is presented. The effective truncation of an active space to smaller spaces is examined at the GASSCF level. SplitGAS uses this partition concept while, at the same time, it captures the dynamical correlation effects that are missing from GASSCF. These two methods can thus be considered as complementary to each other, and are examined together. In the articles in which the methods were first introduced,[2,3] successful calculations of bond dissociation energy curves were reported. In this work, we focus mainly on excited states. A detailed analysis of small, well-documented systems, such as ozone, furan, and pyrrole allows the exploration of different active space partitions strategies. Analysis of molecules containing a transition metal (copper tetrachloride and linear nickel dioxide) will also be presented.

Our GASSCF results demonstrate the systematic reduction of the CI expansion without loss of accuracy with respect to the equivalent CAS calculations. The exponential scaling in GAS is reduced, but not eliminated. Therefore, although GASSCF can be used in combination with larger active spaces compared to CASSCF, most of dynamic

correlation is still not recovered and a quantitative description of excitation energies is out of reach. Conversely, SplitGAS, allowing much larger active spaces, achieves good agreement with available experimental data. For specific cases, when experimental data are not available, our results are compared to those obtained at MRCI level of theory. The paper is organized as follows: In Section 2.1.3 computational methods are discussed, in Section 2.1.4 results for each system are discussed, and lastly Section 2.1.5 includes a discussion about the applicability of the methods and suggestions for their further development.

2.1.3 Computational Methods

All CASSCF, CASPT2 and GASSCF calculations were performed with the MOLCAS-7.8 software package.[153] The SplitGAS method is implemented in the LUCIA code[154] and uses one- and two-electron integrals in the molecular orbital (MO) basis obtained from the MOTRA module of MOLCAS-7.8. Internally-contracted MRCI (icMRCI)[155,156] calculations were performed with the MOLPRO program package,[157] version 2012.1. In all calculations scalar relativistic effects were included using a second-order Douglas-Kroll-Hess Hamiltonian. A shifted zeroth-order Hamiltonian was applied to all CASPT2 calculations (IPEA shift)[65] by using the default value of 0.25 au.

The triple-ζ-quality relativistic all-electron ANO-RCC-VTZP basis sets[158,159] were used for ozone, furan, pyrrole and linear nickel dioxide, contracted to [4*s*3*p*2*d*1*f*] for oxygen, carbon, and nitrogen, [6*s*5*p*3*d*2*f*1*g*] for nickel and [3*s*2*p*1*d*] for hydrogen. For the copper tetrachloride dianion, the ANO-RCC-VTZP base set was used, contracted to [7*s*6*p*5*d*3*f*2*g*1*h*] for copper and [5*s*4*p*2*d*1*f*] for chlorine. In the case of nickel dioxide, some basis set dependence was explored by repeating the calculations with the

quadruple-ζ-quality ANO-RCC-VQZP basis set (contracted to 7*s*6*p*4*d*3*f*2*g* for nickel and 5*s*4*p*3*d*2*f* for oxygen).

The notation "CAS(*n*,*m*)" is followed for the description of the active spaces, where *n* is the number of electrons and *m* the number of orbitals. For the GASSCF calculations, a "GAS-*k*(*n*,*m*)-*i*e,*j*e,…,*m*e" notation is used, where *k* corresponds to the number of subspaces, *i* to the maximum number of electrons that can be excited from the first subspace GAS1 to the other subspaces, *j* to the maximum number of electrons that can be excited from GAS2 to the remaining subspaces etc., with $m = k - 1$ (because no electrons can be excited to the last GAS subspace) For example, the notation GAS-3(18,12)-1e,2e means that an active space composed by 18 electrons in 12 orbitals is divided in 3 subspaces, where single electron excitations are allowed from GAS1 to GAS2 and GAS3, and single and double excitations from GAS2 to GAS3. For the SplitGAS calculations, the "SplitGAS-*k*(n_P,m_P)//(n_T,m_T)" notation is used, where *k* corresponds to the number of subspaces; n_P and m_P are the number of electrons and the number of orbitals in the P space, respectively; T = (P+Q) represents the total molecular space and n_T and m_T are the number of electrons and the number of orbitals in the T space, respectively.

Two different types of MOs are included in the Q-space. The first are doubly occupied valence MOs, which should be correlated for a quantitative calculation of excitation energies. Their choice is system-specific and is based on the chemical nature of the molecule under study, as explained in the next paragraphs. The second (and larger) group of MOs included in the Q-space is composed of virtual orbitals. The selection of the correlated virtual space is based on an orbital energy threshold. The choice of the energy threshold and the truncation of the complete virtual space is discussed in the next paragraphs.

2.1.4 Results

2.1.4.1 Ozone

This system has been the subject of numerous studies. The availability of both experimental[160–167] and theoretical[4,168–183] reference values makes it an ideal molecular system for benchmarking the GASSCF and SplitGAS methods. A set of vertical excitation energies (VEEs), ionization potentials (IPs), and an electron affinity (EA) were computed (Table 1) and compared with available experimental data.

Ozone has C_{2v} symmetry at equilibrium. In the calculations the three O atoms are placed on the *xy*-plane. MOs composed of combinations of the 1*s* (core), 2*s* and the in-plane $2p_x$ and $2p_y$ atomic orbitals (AOs) belong to either the a_1 or b_2 irreducible representations (irreps) of the C_{2v} point group (Figure S1). The valence 2*s*, $2p_x$, and $2p_y$ AOs form the σ backbone of the molecule; three bonding σ ($3a_1$, $5a_1$, $4b_2$), three antibonding σ^* ($4a_1$, $5b_2$, $7a_1$) and three non-bonding n_σ ($2b_2$, $6a_1$, $3b_2$) MOs. Combinations of the three out-of-plane $2p_z$ AOs belong to either b_1 or a_2 irreps and form the π system of the molecule. These are the bonding π ($1b_1$), antibonding π^* ($2b_1$) and non-bonding n_π ($1a_2$) orbitals. All valence electrons are correlated in the largest space considered in this study, which is of size (18,54). As shown in Table 1, all VEEs involve one or two electron excitations from the doubly occupied MOs of the ground state to the antibonding π^*. The three IPs involve an electron loss from the non-bonding $3n_\sigma$ ($6a_1$), $4n_\sigma$ ($4b_2$), and n_π ($1a_2$) MOs, respectively. The lowest anionic state, $^2B_2$, used in the calculation of the EA, has an extra electron in the antibonding π^* ($2b_1$) MO.

Table 1. VEEs, IPs and EA of ozone obtained at different levels of theory. The mean unsigned error (MUE) compared to the experimental values is reported for all methods. All energies in eV.

State	Orbitals involved	CASSCF (12,9)	GAS-2 (12,9)-1e	CASPT2 (12,9)	MRCI	SplitGAS-6 (12,9)//(18,40)	Reference
2^1A_1	$(6a_1)^2 \rightarrow (2b_1)^2$, $(4b_2)^2 \rightarrow (2b_1)^2$	4.57	4.54[a]	4.50	4.35[b]	4.59	4.5[e]
1^1A_2	$4b_2 \rightarrow 2b_1$	2.36	2.36	2.02	2.12[b]	2.13	1.6[f] and 1.92[g]
1^1B_1	$6a_1 \rightarrow 2b_1$	2.60	2.59	2.11	2.15[b]	2.24	2.1[g]
1^1B_2	$1a_2 \rightarrow 2b_1$	5.78	5.79	4.91	5.19[b]	5.27	4.86[g,h]
1^3A_2	$4b_2 \rightarrow 2b_1$	2.20	2.19	1.84	1.90[b]	1.75	1.80[g]
1^3B_2	$1a_2 \rightarrow 2b_1$	1.76	1.75	1.75	1.74[b]	1.56	1.43 and 1.29[g]
1^3B_1	$6a_1 \rightarrow 2b_1$	2.24	2.24	1.71	1.73[b]	1.65	1.67[g]
$^2A_1$	e^- lost from $6a_1$	11.49	11.49	12.66	12.44[c]	12.29	12.73[i]
$^2B_2$	e^- lost from $4b_2$	11.42	11.42	12.78	12.49 [c]	12.38	13.00 [i]
$^2A_2$	e^- lost from $1a_2$	11.97	12.01	13.64	13.17 [c]	13.09	13.54[i]
$^2B_2$	e^- added on $2b_1$	3.51	3.53	1.29	1.71 [d]	1.52	2.10[k]
MUE		0.82	0.81	0.14	0.25	0.29	-

[a] From GAS2(12,9)-2e [b] Ref. [177]. [c] Ref. [174]. [d] Ref. [175]. [e] Ref. [171]. [f] Ref. [164]. [g] Ref. [176]. [h] Ref. [162]. [i] Ref. [183]. [k] Ref. [184].

Two CAS choices were considered: CAS(12,9) and CAS(18,12). The CAS(12,9) is composed of the nine MOs containing the $2p_x$, $2p_y$ and $2p_z$ AOs. The 1^1A_1 ground state has a multireference character, with the valence electronic structure of the dominant configuration (83%) being $(2\sigma)^2(2n_\sigma)^2(3n_\sigma)^2(4n_\sigma)^2(\pi)^2(n_\pi)^2$ or $(3b_2)^2(5a_1)^2(6a_1)^2(4b_2)^2(1b_1)^2(1a_2)^2$. In the CAS(18,12), the 1σ ($3a_1$), $1\sigma^*$ ($4a_1$) and $1n_\sigma$ ($2b_2$) MOs (mostly linear combination of 2s AOs) have also been correlated.

In the GASSCF framework, the (12,9) active space was divided into two subspaces; the first includes the six σ orbitals and the second the three π orbitals. Only single electron excitations were allowed between the two subspaces. This constrained active space is labelled GAS-2(12,9)-1e, by following the notation introduced in the previous section. This partition leads to a reduction of the number of CSFs by about 33% compared to the corresponding CAS(12,9) calculation. For example, for the 1^1A_1 ground state, 666 CSFs are generated from CAS(12,9), while GAS-2(12,9)-1e generates 444. It is worth mentioning that with the GAS-2(12,9)-1e choice, the $1^1A_1 \rightarrow 2^1A_1$ excitation cannot be investigated, as it would require a 2-electron excitation from GAS1 to GAS2. For this excitation a GAS2(12,9)-2e scheme has been chosen (see below). Excellent agreement for the VEEs, IPs, and the EA is observed between the parent CAS(12,9) and GAS-2(12,9)-1e calculations, as shown in Table 2.

Table 2. VEEs, IPs and EA of ozone calculated from CASSCF(12,9), CASSCF(18,12) and the corresponding GASSCF at different levels of approximate active spaces. All energies in eV. The number of CSFs of the 1^1A_1 ground state is given in the last row.

State	**CAS(12,9)**	**GAS-2(12,9)-1e**	**CAS(18,12)**	**GAS-3(18,12)-1e,1e**	**GAS-3(18,12)-1e,2e**	**GAS-3(18,12)-2e,2e**
2^1A_1	4.57	4.54[a]	4.42	*7.98*	4.47	4.45
1^1A_2	2.36	2.36	2.36	2.37	2.42	2.38
1^1B_1	2.60	2.59	2.45	2.48	2.53	2.50
1^1B_2	5.78	5.78	5.68	5.75	5.70	5.68
1^3A_2	2.20	2.19	2.17	2.18	2.23	2.19
1^3B_2	1.76	1.75	1.82	1.74	1.80	1.81
1^3B_1	2.24	2.24	2.06	2.08	2.13	2.07
$^2A_1$	11.49	11.49	11.13	11.26	11.26	11.13
$^2B_2$	11.42	11.42	11.27	11.39	11.38	11.28
$^2A_2$	11.97	12.01	11.85	12.01	11.87	11.84
$^2B_2$	3.51	3.53	3.54	3.55	3.59	3.57
CSFs	666	444	4067	1518	1935	3230

[a] From GAS2(12,9)-2e

For the (18,12) active space, the three low-lying 1σ, $1\sigma^*$ and $1n_\sigma$ MOs form the GAS1 active space. As in the GAS-2(12,9)-1e level, the σ-type ($2p_x$, $2p_y$) and π-type MOs belong to different subspaces (GAS2 and GAS3, respectively). Three different interspace excitation schemes have been explored: In GAS3(18,12)-1e,1e single excitations from GAS1 and from GAS2 are allowed. In GAS-3(18,12)-1e,2e single excitations out of GAS1 and up to double excitations out of GAS2 are allowed. Finally, in GAS-3(18,12)-2e,2e up to double excitations out of GAS1 and GAS2 are allowed. The GAS3(18,12)-1e,1e scheme achieves good accuracy in comparison to the parent CAS(18,12) calculation; VEEs, IPs and EA deviate less than 0.2 eV from CASSCF(18,12), except for the 2^1A_1 state, which involves a two-electron excitation. The GAS-3(18,12)-1e,2e and GAS-3(18,12)-2e,2e further reduce the errors from the parent CAS(18,12). One should notice that these schemes do not fall into the RAS formalism, despite dividing the active space in three subspaces, because the third subspace, GAS3, space contains occupied orbitals.

All VEEs, IPs, and EA values reported in Table 2 show large deviations from the experimental values. The main reason for this disagreement is the lack of correlation energy, which is only partially recovered with the current choice of active space, both at CASSCF and GASSCF levels. Previous theoretical studies[169,172,174,175,177] have shown that non-dynamical correlation alone is not sufficient for a quantitative description of the excitation energies of ozone. In these studies, dynamical correlation is included in the multiconfigurational treatment of the wave function by means of perturbation theory (CASPT2 level)[169] or configuration-interaction (MRCI level),[174,175,177] or by the EOM-CCSDT formalism.[172] In the present work, dynamical correlation has been added using the SplitGAS method.

In the SplitGAS formalism a combined (P+Q) space of size (18,40) was used. The P space is composed of the CAS(12,9) as described above. The extended space in SplitGAS is formed by single excitations out of the 1σ, $1\sigma^*$, and $1n_\sigma$ MOs and single and double excitations into 28 additional canonical virtual orbitals. The choice of only single excitations out of the 1σ, $1\sigma^*$, and $1n_\sigma$ orbitals was motivated by the fact that no significant differences between the GAS-3(18,12)-1e,2e and the GAS-3(18,12)-2e,2e levels were observed at the GASSCF level of theory. In a preliminary investigation (see Supporting Information) various levels of truncation of the virtual orbital space, based on orbital energies, were applied to form the extended (Q) of SplitGAS. Threshold values of 1.0, 2.0, 4.0 and 5.0 E_h (Hartree) were considered (Table S1). No significant changes in the VEEs, IPs and EA were observed for a Q-space that included virtual MOs with energies above 2.0 E_h. Conversely, the threshold value of 1.0 E_h was too poor for a quantitative calculation of dynamical correlation. This result lead us to choose 28 virtual orbitals to form the extend space. Table S2 summarizes the number of electrons and orbitals in each subspace. In order to reduce memory requirements, virtual orbitals belonging to different irreps were divided in separate subspaces. In total, six subspaces were considered, SplitGAS-6(12,9)//(18,40). This level of theory includes more than 6 million CSFs in the (P+Q)-space for the 1^1A_1 ground state.

The VEEs, IPs, and EA obtained at the SplitGAS-6(12,9)//(18,40) for the 11 states considered are reported in Table 1. For comparison, experimental and computational results from the CASSCF(12,9), GASSCF-2(12,9)-1e, CASPT2(12,9) and MRCI level of theory are also included. The mean unsigned error (MUE) for all theoretical methods in comparison to the experimental values is reported. As previously discussed, CASSCF(12,9), and GASSCF-2(12,9)-1e show similar behavior (the MUE is 0.82 and 0.81 eV, respectively). CASPT2(12,9) captures most of the missing correlation energy and achieves good agreement with the experimental results (MUE = 0.14eV).

MRCI and SplitGAS-6(12,9)//(18,40) have a similar accuracy, with a MUE of 0.25 and 0.29 eV, respectively.

All methods including dynamical correlation considered in this study, i.e. CASPT2, MRCI, and SplitGAS, achieve similar accuracy for the VEEs (on average deviations less than 0.2 eV from the reference experimental data). However, the excitation from the ground state to the 1^1B_2 state is a notable exception. The VEE to the 1^1B_2 state calculated with MRCI and SplitGAS deviate from both the experimental reference value of 4.86 eV[167] and the CASPT2(12,9) value of 4.91 eV. Previous theoretical results for this specific state are discussed (Table 3). The 1^1B_2 excited state belongs to the Hartley band which absorbs in the region between 4.2 and 5.6 eV.[165] The Hartley band has the strongest absorption in the spectrum of O_3 below 5.8 eV with an oscillator strength *f* of about 0.1.[160] A large oscillation strength was calculated at the MRCI level by Elliot et al.[177] for the 1^1B_2 state. The reference value of 4.86 eV used by Musiał *et al*.[167,172] falls almost in the middle of the absorption region of the Hartley band. CASPT2(12,9) calculations of Borowski *et al*.[169] underestimated the VEE of 1^1B_2 state by about 0.2eV (4.69 eV). A better agreement can be achieved by using a shifted zeroth-order Hamiltonian (IPEA shift)[65] in CASPT2 (4.91 eV, this work). Results obtained using equation-of-motion coupled-cluster theory with single, double, and triple excitations (EOM-CCSDT)[172], or MRCI by selecting specific configurations[171] or by using the full configuration space[177] lie between 5.13 eV and 5.19 eV. SplitGAS-6(12,9)//(18,40), which spans the Q-space based on an energy threshold of 2 E_h, predicts the VEE of 1^1B_2 state to be 5.27 eV. This value is slightly higher than the corresponding EOM-CCSDT or MRCI energies. Better agreement can be achieved if a 4 E_h threshold is applied on the Q-space; 5.18 eV from the SplitGAS-6(12,9)//(18,54) level. By localizing the virtual MOs (LMOs) with the Cholesky localization scheme[185] as implemented in MOLCAS, we were able to select MOs with a predominant AO contribution. In the SplitGAS-6(12,9)//(18,42)

(LMO) calculation, MOs with 3*s*3*p*4*s*3*d* contribution were included in the extended Q-space. This methodology reduces the VEE of 1^1B_2 state by about 0.1 eV (5.17 eV). The $^3B_2$ state is another state where disagreement between the theoretical methods was found. CASPT2 (1.75 eV) and MRCI (1.74 eV) deviate significantly from the two experimental reference values (1.43 and 1.29 eV), while SplitGAS (1.56 eV) and EOM-CCSDT (1.54 eV) are in better agreement.

Table 3. Comparison between different levels of theory for the VEE from the 1^1A_1 ground state to the 1^1B_2 excited state of ozone. The number of CSF for the MR methods is also given.

Method	1^1B_2 VEE (eV)	Number of CSFs
CASPT2[a]	4.69	626
CASPT2 (This work)	4.91	626
EOM-CCSDT[b]	5.13	-
MRCI[c]	5.16	10 300
MRCI[d]	5.19	~ 39 10^6
SplitGAS-6(12,9)//(18,40)	5.27	~ 6 10^6
SplitGAS-6(12,9)//(18,42) (LMOs)	5.17	~ 8 10^6
SplitGAS-6(12,9)//(18,54)	5.18	~ 25 10^6
Reference[e,f]	4.86	

[a] Ref. [169]. [b] Ref. [172]. [c] Ref. [171]. [d] Ref. [177]. [e] Ref. [176]. [f] Ref. [167].

Finally, IPs calculated at the SplitGAS-6(12,9)//(18,40) level deviate from the experimental reference values[183] by about 0.3 to 0.6 eV and by about 0.1 eV from IPs obtained at the MRCI level by Schmelz *et al*.[174] CASPT2 results are in a better agreement with the reference IPs. A similar behavior is observed for the EA, with the only exception that CASPT2(12,9) significantly underestimates the experimental value by about 0.8 eV.

2.1.4.2 Furan/Pyrrole

The next systems to be discussed are two aromatic five-member molecules, furan and pyrrole. The calibration of our methodology for these two molecules yields information that may be useful for more complicated organic molecules. Furan and pyrrole are isoelectronic and exhibit similar absorption patterns in their UV spectra, i.e., their absorption bands are attributed to the same electronic excitations. Their fully conjugated planar system leads to a weak aromatic 6π-electron character. In C_{2v} symmetry, and for the 1^1A_1 ground state, furan and pyrrole have MOs filled up to $9a_1$, $2b_1$, $1a_2$, and $6b_2$. The occupied π-orbitals are the $1b_1$, $2b_1$, and $1a_2$, with the lowest π^* orbitals being the $3b_1$ and $2a_2$. The lowest singlet and triplet valence excited states considered in this benchmark study involve π-to-π* excitations. The VEEs from the ground state to the singlet 2^1A_1, 3^1A_1 and 1^1B_2 and the triplet 1^3A_1 and 1^3B_2 states are examined. Rydberg orbitals and states are excluded from this study. As discussed by others[186], the difference in CASPT2 energies from state-specific and state-averaged orbitals is less than 0.1 eV. For simplicity, all results reported in this study are obtained with state-specific CASSCF orbitals. Our results are compared with the experimental data of Flicker et al.[187] Finally, we have recalculated the VEEs of interest with MRCISD(6,5) by using the same CAS(6,5) wave function as in our SplitGAS calculations.

Table 4 includes VEEs for the first five π-to-π* excited states of furan calculated with CASSCF and GASSCF. The size of the active spaces has been systematically expanded from (6,5) up to (6,39). The CAS(6,10) includes three b_1 and two a_2 extra virtual MOs. The CAS(6,15) includes five additional virtual MOs (three b_1, two a_2). An increase of the active space results in a decrease of the VEEs of the 1^1B_2 and 3^1A_1 states by about 0.5 eV. Smaller differences are observed for the 2^1A_1, 1^3A_1 and 1^3B_2 states (0.1 eV or less). Energy thresholds of 2.0 and 3.0 E_h , respectively, for the canonical virtual

orbitals belonging to b_1 and a_2 irreps lead to the CAS(6,27) and CAS(6,39), respectively. GAS2-type CI expansions have been constructed as follows: the CAS(6,*X*) space (*X* = 10, 15, 27, or 39) is divided in two subspaces. The first subspace contains the initial five π-to-π* orbitals of the minimum CAS(6,5) space while the second subspace contains the additional virtual orbitals. Single or single and double excitations are allowed from GAS1 to GAS2. These are reported in Table 4 as GAS-2(6,*X*)-1e and GAS-2(6,*X*)-2e, respectively. No significant differences are observed for the VEEs when comparing the excitation energies calculated from CASSCF(6,10) and CASSCF(6,15) with the corresponding results from GAS2, although significant reduction of the CI expansion is achieved by the GAS choices. For example, the GAS-2(6,15)-1e and GAS-2(6,15)-2e contain only 1% and 9% of the total number of CSFs, respectively. Increasing the number of unoccupied orbitals (*X* = 27 or 39) does not affect the VEEs. Among the different approaches presented in Table 4, GAS-2(6,15)-1e is the best choice compared to the parent CAS(6,15) energies, in terms of efficiency and accuracy. The results with a GAS-2(6,15)-1e space with additional subspaces, containing doubly occupied and virtual MOs of σ type, are reported in the Supporting Information (Table S3).

Table 4. Expansion of active spaces for furan at the CASSCF and GASSCF level where only excitations from the π-to-π* orbitals are considered. All energies in eV. The number of CSFs of the $^1A_1$ ground state is given.

	CASSCF (6,5)	CASSCF (6,10)	CASSCF (6,15)	CASSCF (6,27)	CASSCF (6,39)
CSFs	28	2 510	31 990	1 198 535	11 290 082
1^1B_2	8.09	7.73	7.56	7.51	7.49
2^1A_1	6.85	6.85	6.80	6.79	6.78
3^1A_1	10.25	9.86	9.75	9.70	9.69
1^3A_1	5.47	5.49	5.45	5.44	5.44
1^3B_2	4.02	4.11	4.11	4.12	4.12
		GAS-2(6,10)-1e	GAS-2(6,15)-1e	GAS-2(6,27)-1e	GAS-2(6,39)-1e
CSFs	-	217	406	856	1 321
1^1B_2	-	7.78	7.65	7.63	7.64
2^1A_1	-	6.81	6.77	6.77	6.77
3^1A_1	-	9.85	9.77	9.75	9.76
1^3A_1	-	5.45	5.41	5.40	5.40
1^3B_2	-	4.10	4.08	4.07	4.07
		GAS-2(6,10)-2e	GAS-2(6,15)-2e	GAS-2(6,27)-2e	GAS-2(6,39)-2e
CSFs	-	829	2 819	12 431	28 913
1^1B_2	-	7.75	7.58	7.53	7.51
2^1A_1	-	6.87	6.80	6.79	6.78
3^1A_1	-	9.88	9.77	9.73	9.71
1^3A_1	-	5.52	5.45	5.44	5.44
1^3B_2	-	4.12	4.11	4.12	4.11

Table 5 shows SplitGAS results for the low-lying states of furan and pyrrole. In all cases the correlated space is divided in six subspaces. The first subspace includes the occupied MOs of the σ system (six a_1 MOs and four b_2 MOs), composed of the 2*s*, $2p_x$, and $2p_y$ AOs of carbon, and oxygen or nitrogen for furan and pyrrole, respectively. Single and double excitations are allowed from this subspace. The second subspace includes the π system of the aromatic molecules and forms the P-Space of the SplitGAS calculations. The remaining four spaces are composed of low-lying canonical virtual orbitals and they form the Q-Space, chosen according to orbital energy thresholds. The values of 1.0 and 2.0 E_h have been examined, resulting in active spaces (P+Q) with 26 electrons in 55 and

84 orbitals in total, respectively. As can be seen in Table 5, our results considerably deviate from the experimental VEEs for both molecules, with the errors exceeding the 1 eV in some cases. The reason for this divergence is the incorporation of Rydberg-type orbitals inside the Q-space, leading to unphysical mixing of Rydberg and valence excited states. Rydberg orbitals could be deleted from the orbital space after separately optimizing the Rydberg excited states.[188,189] In the present work, however, this procedure has been avoided, as our goal is not necessarily the theoretical reproduction of the experimental spectra, but rather a comparison of SplitGAS and MRCISD (singles and doubles) results, by using the same basis set (ANO-RCC-VTZP) and same primary (6,5) space. Reasonable agreement between the two methods is observed for most of the states. For the majority of the excited states, the larger Q-space shifts the VEEs closer to the MRCISD values. The biggest deviation was found for the VEE of the 2^1A_1 state. For both aromatic molecules, the VEE to this particular state is not affected by the expansion of the Q-space based on an energy threshold.

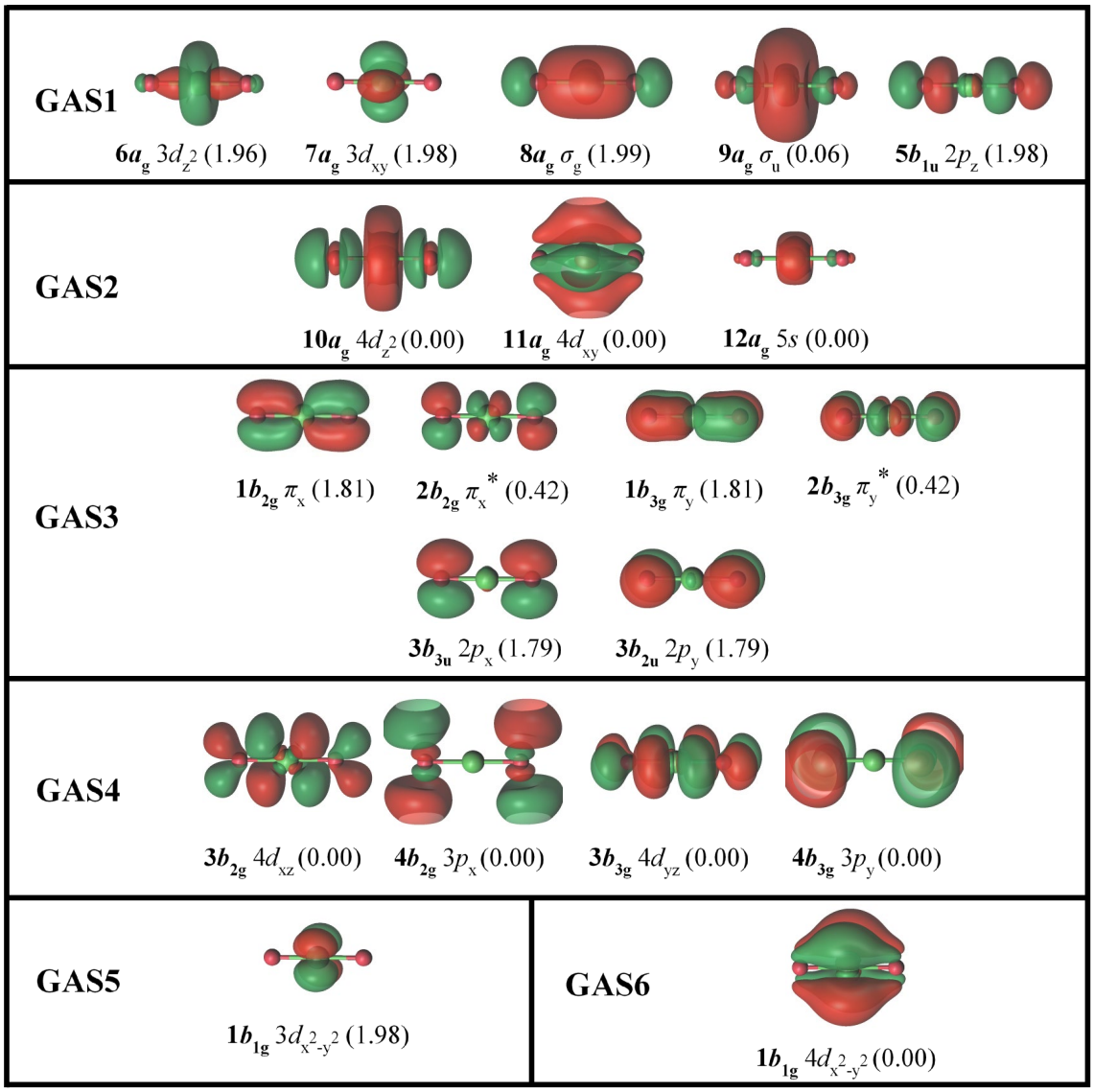
Table 5. Expansion of active spaces for furan and pyrrole for SplitGAS. Experimental and MRCI VEEs are given for comparison. All energies in eV.

	State	Exp.	MRCI(6,5)	SplitGAS-6 (6,5)//(26,55) 1 Hartree	SplitGAS-6 (6,5)//(26,84) 2 Hartree
Furan	$1\ ^1B_2$	6.06	6.93	6.94	6.85
	$2\ ^1A_1$		6.79	7.13	7.13
	$3\ ^1A_1$	7.82	9.11	9.21	9.10
	$1\ ^3A_1$	5.22	5.58	5.84	5.79
	$1\ ^3B_2$	4.05	4.26	4.38	4.33
Pyrrole	$1\ ^1B_2$		7.02	7.15	7.03
	$2\ ^1A_1$	5.98	6.55	6.92	6.91
	$3\ ^1A_1$	7.54	8.66	8.78	8.63
	$1\ ^3A_1$	5.10	5.62	5.85	5.81
	$1\ ^3B_2$	4.21	4.55	4.71	4.67

2.1.4.3 Nickel Dioxide

The nickel atom forms three different conformers with molecular oxygen, a superoxo, a peroxo and a linear isomer. A recent, detailed MRCI study by Hübner and Himmel[190] showed that the linear ONiO isomer is the lowest energy form of $Ni(O_2)$. This conformation is about 1.53 eV more stable than the cyclic $Ni(O_2)$. Experimental evidence for the existence of the linear ONiO has been found by reaction of laser-ablated Ni atoms with O_2 in argon[191] and neon[192] matrices. The electronic structure of various conformers of the anionic species, $ONiO^-$, has been studied by photoelectron spectroscopy.[193,194] Early theoretical works were focused on the cyclic $Ni(O_2)$ isomer.[195–197] DFT calculations predicted a linear $^1\Sigma_g^+$ ground state[191,198] and placed the peroxo $^3B_1$ state 0.73 eV above the $^1\Sigma_g^+$ state.[191]

(a)


GAS1
6a_g 3d_z2 (1.96)
7a_g 3d_xy (1.98)
8a_g σ_g (1.99)
9a_g σ_u (0.06)
5b_1u 2p_z (1.98)
GAS2
10a_g 4d_z2 (0.00)
11a_g 4d_xy (0.00)
12a_g 5s (0.00)
GAS3
1b_2g π_x (1.81)
2b_2g π_x* (0.42)
1b_3g π_y (1.81)
2b_3g π_y* (0.42)
3b_3u 2p_x (1.79)
3b_2u 2p_y (1.79)
GAS4
3b_2g 4d_xz (0.00)
4b_2g 3p_x (0.00)
3b_3g 4d_yz (0.00)
4b_3g 3p_y (0.00)
GAS5
1b_1g 3d_x2-y2 (1.98)
GAS6
1b_1g 4d_x2-y2 (0.00)

(b)

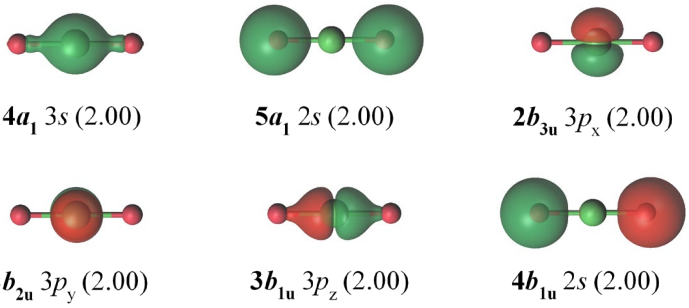

4a_1 3s (2.00)
5a_1 2s (2.00)
2b_3u 3p_x (2.00)
2b_2u 3p_y (2.00)
3b_1u 3p_z (2.00)
4b_1u 2s (2.00)

Figure 3. (a) The natural orbitals of linear nickel dioxide at equilibrium bond distance of the $^1\Sigma_g^+$ ground state divided in six subspaces for the GASSCF-6(18,20)-2e1e2e0e2e calculations. GAS1, GAS3 and GAS5 are the same subspaces used for the GAS-3(18,12) scheme. At that level of theory, they are named as GAS1, GAS2 and GAS3, respectively. (b) The six orbitals included in the SplitGAS-8(18,12)//(30,35) calculation. Orbital labels and occupation numbers from the parent CAS(18,12) calculation for the $^1\Sigma_g^+$ ground state are listed below each orbital.

The adiabatical excitation energies of the linear ONiO isomer are examined. In particular, the four lowest states, $^1\Sigma_g^+$, $^3\Pi_g$, $^1\Pi_g$, and $^5\Pi_u$, are calculated by symmetrically varying the $R_{Ni\text{-}O}$ bond distances. The excitation from the $^1\Sigma_g^+$ ground state to the $^1\Pi_g$ state involves an electron excitation from the $3d_{z2}$ orbital ($6a_g$) to the π^* orbital ($2b_{3g}$ or $2b_{3g}$). The first triplet state is described by a spin-flip of the electron occupying the π^* orbital of the $^1\Pi_g$ state. The quintet $^5\Pi_u$ state involves two-electron excitations to the π^* orbitals; one-electron excitation from the $3d_{z2}$ orbital and one-electron from the $2p_x/2p_y$ orbitals.[190] An active space of (18,12) was used in the CASSCF, and GASSCF calculations (Figure 3a), which includes the six 4*s*3*d* orbitals of nickel and the six 2*p* orbitals of the two oxygen atoms.

At the GASSCF level, the CAS(18,12) is partitioned in three subspaces based on symmetry considerations. The first space (GAS1) is composed of the σ-type MOs, which are the bonding $8a_g$, the antibonding $9a_g$, and the non-bonding $5b_{1u}$ orbitals (Figure 3a). GAS1 also includes the $3d_{z2}$ and $3d_{xy}$ orbitals belonging to the same irreducible representation as the σ-type (σ and σ^*) MOs. GAS2 includes the π-type MOs, i.e. the two components of the π bonding $1b_{2g}$ and $1b_{3g}$, the π^* antibonding $2b_{2g}$ and $2b_{3g}$, and the non-bonding $3b_{2u}$ and $3b_{3u}$. GAS3 includes only the remaining $3d_{x2\text{-}y2}$ ($1b_{1g}$) orbital. No excitations from or to GAS3 were allowed. Thus, this partition can be considered equivalent to a CAS(16,11); nonetheless, the (18,12) nomenclature is kept for consistency with subsequent partitioning schemes (vide intra). Two different excitation schemes were applied within the three subspaces. The first includes CSFs generated by single electron

excitations from GAS1 to GAS2 (GASSCF3(18,12)-1e0e); the second adds two-electron excitations from GAS1 to GAS2 (GASSCF-3(18,12)-2e0e).

The impact of the double-shell[199] of Ni was examined by extending the (18,12) active space to (18,20). The additional 8 virtual MOs are linear combinations of the six 5*s*4*d* of Ni and the $3p_x/3p_y$ orbitals of the oxygen atoms. The full (18,20) space is partitioned in 8 subspaces, as shown in Figure 3a. Orbitals $6a_g$, $7a_g$, $8a_g$, $9a_g$, and $5b_{1u}$ form GAS1 (as in the GAS-3(18,12) scheme). The $10a_g$, $11a_g$, and $12a_g$ MOs, which are dominated by contributions from the 5*s*, $4d_{z2}$ and $4d_{xy}$ atomic orbitals of Ni, form the GAS2. GAS3 is formed by the $1b_{2g}$, $2b_{2g}$, $1b_{3g}$, $2b_{3g}$, $3b_{2u}$, and $3b_{3u}$ orbitals (GAS2 in the GAS-3(18,12) scheme). GAS4 is composed by $4d_{xz}$ and $4d_{yz}$ orbitals ($3b_{2g}$ and $3b_{3g}$, respectively) and the double-shell of the $2p_x/2p_y$ orbitals of the oxygen atoms ($4b_{2g}$ and $4b_{3g}$, respectively). GAS5 contains orbital $1b_{1g}$ while the $2b_{1g}$ ($4d_{x2-y2}$) forms the GAS6 space. The minimum and maximum electron occupations of each subspace are reported in Table S4.

Single and double excitations are allowed from each subspace of the initial CAS(18,12) to the corresponding additional subspaces of the second shell, i.e. from GAS1 to GAS2, from GAS3 to GAS4, and from GAS5 to GAS6. An excitation level between subspaces of different symmetry similar to that in the GAS-3(18,12)-1e0e is followed in this scheme. Only single excitations are allowed between the first two (GAS1 and GAS2) and the next two subspaces (GAS3 and GAS4). No excitations are allowed from the first four to the last two subspaces (GAS5 and GAS6). This level of theory is abbreviated as GASSCF-6(18,20)-2e1e2e0e2e and includes about 18 million CSFs. The corresponding CASSCF(18,20) calculation would be prohibitive (more than 1 billion CSFs).

Table 6. Adiabatic excitation energies (in eV) and Ni-O bond distances (in Å) for the $^1\Sigma_g^+$, $^3\Pi_g$, $^1\Pi_g$ and $^5\Pi_u$ states of linear nickel dioxide at the different levels of theory considered in this study.

Method	CSFs of $^1\Sigma_g^+$	$^1\Sigma_g^+$		$^3\Pi_g$		$^1\Pi_g$		$^5\Pi_u$	
		E	R_{Ni-O}	*E*	R_{Ni-O}	*E*	Ni-O	*E*	R_{Ni-O}
GASSCF-3(18,12)-1e0e	709	0	1.63	0.52	1.70	0.84	1.73	-	-
GASSCF-3(18,12)-2e0e	1 167	0	1.63	0.58	1.70	0.89	1.73	0.84	1.75
CASSCF(18,12)	2 108	0	1.63	0.59	1.70	0.91	1.73	0.89	1.75
GASSCF-6(18,20)-2e1e2e0e2e	18 387 635	0	1.63	0.55	1.68	0.97	1.68	0.81	1.73
CASPT2(18,12)	2 108	0	1.61	0.31	1.61	0.69	1.60	1.27	1.68
SplitGAS-8(18,12)//(30,35)	39 763 303	0	1.59	0.33	1.61	0.82	1.61	0.64	1.68
SplitGAS-8(18,12)//(30,35)[a]	39 763 303	0	1.62	0.42	1.65	1.01	1.68	0.63	1.70
SplitGAS-6(18,12)//(18,29)[a,b]	3 407 698	0	1.62	0.49	1.65	1.05	1.68	0.70	1.70
MRCI(18,12)[c]		0	1.61	0.54	1.64	0.94	1.63	1.17	1.71
Other (DFT studies)			1.586 - 1.613[d,e,f,g]	0.35[e] 0.41[f]				1.01[f]	

[a] Same basis as MRCI(18,12) (contracted as 7*s*6*p*4*d*3*f*2*g* for Ni, 5*s*4*p*3*d*2*f* for O).
[b] No *semi-core* correlation. See text for details.
[c] Ref. [190]. [d] Ref. [198]. [e] Ref. [191]. [f] Ref. [192]. [g] Ref. [200].

Table 6 reports the adiabatic excitation energies and the $R_{Ni\text{-}O}$ bond distances of the minimum of each state ($^1\Sigma_g^+$, $^3\Pi_g$, $^1\Pi_g$, and $^5\Pi_u$). Previous MRCI[190] and DFT[191,192,198,200] results are also included for comparison. The MRCI data were obtained by using CASSCF(18,12) orbitals. It should be noted that in the MRCI reference data, the *semi*-core correlation is not included, i.e. the 3*s*3*p* orbitals are not correlated, and effects from higher excitations (e.g. second *d*-shell) are neglected. The authors[190] give a rough estimate of the error introduced from their computational approach of about 0.13 eV, which is essentially from the 3*s*3*p* *semi*-core-valence correlation. The ANO-RCC basis set, contracted as 7*s*6*p*4*d*3*f*2*g* for nickel and 5*s*4*p*3*d*2*f* for oxygen, was used in the MRCI study. In this study, as already discussed in the Computational Methods section, two different contraction schemes were employed, namely the same as in the MRCI study and a smaller one (6*s*5*p*3*d*2*f*1*g* for nickel, 4*s*3*p*2*d*1*f* for oxygen). The basis set effects are discussed at the end of this section.

The wave function constructed by the CASSCF(18,12) level contains 2108 CSFs. Excitation energies obtained at this level are in agreement with MRCI data, while the bond distance predicted by CASSCF(18,12) is 0.02 Å longer that the MRCI value. The partition schemes introduced by the GASSCF-3(18,12)-1e0e and GASSCF-3(18,12)-2e0e result in adiabatic excitation energies comparable with the parent CASSCF(18,12). The only disagreement is for the $^5\Pi_u$ state for the GAS scheme that allows a single electron excitation between different subspaces (GASSCF-3(18,12)-1e0e). This state involves two-electron excitations to the π^* orbitals, while the GASSCF-3(18,12)-1e0e level allows only single-excitations. The position of the minimum on the potential energy curves of the four states is not affected by the partition of the CAS(18,12). Using a larger active space (18,20) within the GASSCF framework does not significantly affect the excitation energies. On the other hand, the equilibrium $R_{Ni\text{-}O}$ bond distances for the three excited states are in a better agreement with the MRCI results.

Dynamical correlation obtained by excitations to the virtual space is introduced by CASPT2 and SplitGAS (Table 6). The CSFs included in the SplitGAS-8(18,12)//(30,35) calculation are constructed by taking into account an active space composed of 30 electrons in 35 orbitals (in P+Q space). The P-space is the CAS(18,12) described in the previous paragraph and shown in Figure 3a. *Semi*-core MOs of Ni (3*s*3*p*) and the two 2*s* of the oxygen atoms (Figure 3b) and their electrons were used to build the Q-Space. Seventeen additional low-lying canonical virtual orbitals, which include the double-shell MOs described for the GAS-6(18,20) scheme, are also used to generate configurations in the Q-Space. Up to two holes were allowed in the MOs formed by the 2*s* atomic orbitals of the oxygen atoms ($5a_g$, $4b_{1u}$), while only single holes were allowed in the 3*s*3*p* of Ni. These constraints were based on the fact that the 3*s*3*p* MOs of Ni are much lower in energy than the 2*s* of oxygens. The 17 additional virtual orbitals used for the SplitGAS had mostly 4*p*5*s*4*d* character of Ni and the 3*s*3*p* of O. In total, 30 electrons in 35 orbitals were correlated based on the restrictions described above. This partition of the CAS(30,35) generated about 40 million CSFs for the $^1\Sigma_g^+$ ground state and is referred to as SplitGAS-8(18,12)//(30,35).

The SplitGAS-8(18,12)//(30,35) adiabatic excitation energy to the $^3\Pi_g$ state of 0.33 eV agrees with the 0.31 eV energy from CASPT2(18,12) and with previous DFT studies (0.35-0.41 eV), but it is about 0.2 eV lower compared to the MRCI value. The SplitGAS excitation energy to the first singlet $^1\Pi_g$ excited state (0.82 eV) is in-between the MRCI (0.94 eV) and the CASPT2(18,12) (0.69 eV) values. Finally, the SplitGAS-8(18,12)//(30,35) calculations predict that the quintet $^5\Pi_u$ state is lower than the $^1\Pi_g$ state by about 0.2 eV. On the other hand both CASPT2(18,12) and MRCI predict the quintet state to lie higher than the $^1\Pi_g$ state, and about 1.2 eV higher than the ground state. SplitGAS-8(18,12)//(30,35) and CASPT2(18,12) predict similar equilibrium distances for

the $^1\Sigma_g^+$, $^3\Pi_g$, $^1\Pi_g$, and $^5\Pi_u$ states. The $R_{Ni\text{-}O}$ distances for these states are shorter than those predicted by the CAS or GAS methods, which lack dynamical correlation.

In Ref. [190], the ANO-RCC-VQZP (contracted as 7*s*6*p*4*d*3*f*2*g* for nickel and 5*s*4*p*3*d*2*f* for oxygen) was employed. In order to make the comparison more consistent we have repeated the calculations described above with the same basis set as in Ref. [190]. These calculations were performed with and without correlating the *semi-core* orbitals (SplitGAS-8(18,12)//(30,35) and SplitGAS-6(18,12)//(18,29), respectively) As can be seen in Table 6, SplitGAS and MRCI predict similar excitation energies for the first two excited states. However, the difference between the two methods for the quintet $^5\Pi_u$ state is about 0.5 eV. Finally, the effect on SplitGAS of the *semi-core* correlation is estimated to be about 0.07 eV.

2.1.4.4 Copper Tetrachloride Dianion

We computed excitation energies to the ligand field states and to a charge transfer state of the copper tetrachloride dianion $[CuCl4]^{2-}$. We used the same basis set (ANO-RCC basis set contracted to [7*s*6*p*5*d*3*f*2*g*1*h*] for Cu and [5*s*4*p*2*d*1*f*] for Cl) and geometry (planar D_{4h} with $R_{Cu\text{-}Cl}$ = 2.291 Å) as used in previous studies,[201–203] in which the CASPT2(11,11) level of theory was employed.

In the present study, the calculations were performed by imposing D_{2h} symmetry, the highest allowed symmetry in MOLCAS. The orbitals and states are labeled by using the irreps of the D_{2h} point group. The (11,11) active space includes the 3*d* and 4*d* orbitals of Cu, and a symmetry adapted orbital composed by four 3*p* orbitals of the chlorine atoms pointing toward the metal center. The singly occupied MO has σ^* antibonding character with predominant contribution from the Cu $3d_{x2\text{-}y2}$ atomic orbital. Excitations from the remaining four 3*d* orbitals to the σ^* antibonding MO correspond to the 1^2B_{1g},

$1^2B_{2g}/1^2B_{3g}$ and 2^2A_g ligand field states. A single excitation from the σ Cu-Cl bonding orbital to the σ^* antibonding orbital gives rise to the 3^2A_g charge transfer state.

Table 7. Low energy ligand field and charge transfer VEEs of $[CuCl_4]^{2-}$ obtained at different levels of theory. All energies in eV. The number of CSFs of the 1^2A_g ground state is given in the last row.

State	CASSCF (11,11)	GASSCF-4 (11,11)-0e0eN_{ex}e[b]	CASPT2 (11,11)	SplitGAS7 (11,11)//(33,37)	Exp.[a]
1^2B_{1g}	2.24	2.25	1.52	1.45	1.55
$1^2B_{2g}/1^2B_{3g}$	2.54	2.55	1.77	1.74	1.76
2^2A_g	1.77	1.77	2.00	1.77	
3^2A_g	6.58	6.59	4.60	6.41	
CSFs	26 256	6 912		79 307 583	

[a] Ref. [204]. [b] $N = 1$ for 1^2B_{1g}, 1^2B_{2g}, and 1^2B_{3g} states, $N_{ex} = 0$ for 2^2A_g, and 3^2A_g states.

Table 7 reports the ligand-field and charge-transfer excitations of $[CuCl_4]^{2-}$ obtained with the different methods discussed in this study. CASPT2(11,11) energies are in good agreement with the experimental values[204] for the first two ligand-field states (1^2B_{1g}, $1^2B_{2g}/1^2B_{3g}$). The GASSCF-4(11,11)-0e0eN_{ex}e partition ($N_{ex} = 0$ or 1, *vide infra*) is based on symmetry considerations: MOs belonging to the same irrep form a separate subspace (Figure 4). GAS1 includes two b_{1g} orbitals, the $3d_{xy}$ and the isosymmetric $4d_{xy}$ (double-shell). Similarly, GAS2 includes two b_{2g} orbitals ($3d_{xz}$ and $4d_{xz}$), and GAS4 two b_{3g} orbitals ($3d_{yz}$ and $4d_{yz}$). GAS4 includes the five a_g orbitals occupied by 5 electrons in the 1^2A_g ground state. These are the σ bonding (with predominant $3p$ contribution from the Cl atoms) and σ^* antibonding (predominant contribution from the Cu $3d_{x2-y2}$ atomic orbital) character MOs, the $3d_{z2}$ orbital, and the two corresponding $4d$ orbitals of Cu (double-shell). Different considerations are accounted for in the calculation of the VEEs. No electron excitations are allowed in the calculations of the 2^2A_g and 3^2A_g states ($N_{ex} = 0$). The 1^2B_{1g}, 1^2B_{2g}, and 1^2B_{3g} states involve an electron excitation from the $3d$ orbitals

of GAS1, GAS2 or GAS3, respectively, to the $3d_{x2-y2}$ orbital ($15a_g$) of GAS4. Thus, single electron excitations ($N_{ex} = 1$) are allowed from *one* subspace to GAS4 *per state*. For simplicity, this level of theory is referred to as GASSCF-4(11,11)-0e0eN_{ex}e ($N_{ex} = 0$ or 1). No significant deviations (≥ 0.01 eV) from those obtained with the parent CASSCF(11,11) level (Table 7) occur, while the size of the CI expansion is reduced by almost 75%. It should be noted that both CASSCF and GASSCF densities predict a 0.82 ground-state spin population for Cu, while the experimental value is 0.62.[205]

Two different P-spaces were employed in the SplitGAS calculations. In the first case, the full CAS(11,11) forms the P-space. The Q-space is composed of single holes created in the 3*p* orbitals of the four chlorine atoms, and double particles in the 15 lowest virtual orbitals, which are split in 5 subspaces. This level of theory is abbreviated as SplitGAS-7(11,11)//(33,37) and yields ligand field excitation energies in agreement with the experiment available values (Table 7). Conversely, no significant contribution to the dynamical correlation for the charge-transfer state (3^2A_g) is observed. The SplitGAS-7(11,11)//(33,37) excitation energy for this state is similar to the one obtained at the CASSCF(11,11) level of theory while the CASPT2(11,11) value is about 2eV lower (See the discussion in the next section). No experimental value is available.

The number of CSFs for the 1^2A_g ground state generated from the SplitGAS-7(11,11)//(33,37) is about 80 million. Thus, this choice of fragmentation of the (33,37) orbital space yields a very large number of CSFs which does not allow us the flexibility needed for a detailed and systematic exploration of the limits of SplitGAS. For example, the addition of only 9 extra virtual MOs increases the size of the CI expansion to almost 200 million CSFs. Therefore, the CAS(11,11) as P-space is not further considered.

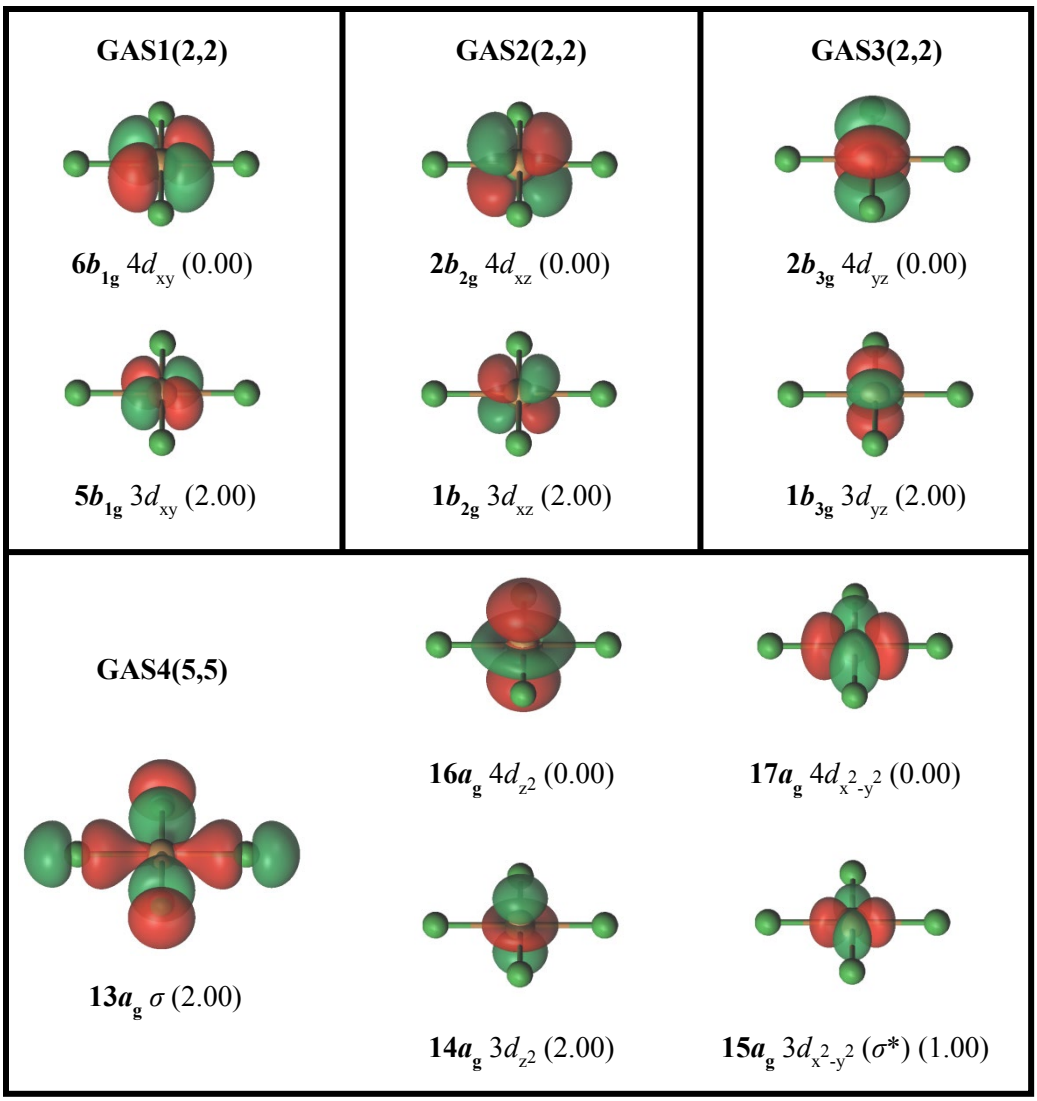


Figure 4. The four subspaces (GAS1(2,2), GAS2(2,2), GAS3(2,2) and GAS4(5,5)) used in the GASSCF-4(11,11) calculations for $[CuCl_4]^{2-}$. Single electron excitations from GAS1, GAS2 or GAS3 to GAS4 are allowed for the 1^2B_{1g}, 1^2B_{2g}, and 1^2B_{3g} excited states, respectively. Orbital labels and occupation numbers for the ^2A_g ground state are listed below each orbital.

A different partition of the active space allowed a significant simplification of the CI problem and the exploration of larger spaces. The CAS(11,11) is divided in two subspaces: the σ/σ* MOs and the four 3*d* orbitals of Cu form the P-space of size (11,6), while the second *d*-shell of Cu is shifted into Q-space. The shift of the second *d*-shell

from P-space to Q-space enables the systematic expansion of the Q-space size. However, this choice yielded two problems. Firstly, the CI optimization of SplitGAS reached no convergence for the 2^2A_g state. Secondly, the initial six MOs included in the P-space differ for the 1^2B_{1g}, 1^2B_{2g} and 1^2B_{3g} states from the MOs of the 1^2A_g ground state; the optimized MOs of CASSCF(11,6), the preceding step of the SplitGAS method, were not the same for the ground and the three excited states. This unbalanced treatment of the P-space lead to erroneous VEEs. This issue was overcome by using the MOs of the 1^2A_g ground state as P-space for the SplitGAS calculations of the 1^2B_{1g}, 1^2B_{2g} and 1^2B_{3g} states.

Table 8 reports the results obtained by systematically increasing the CI expansion of the ground and excited states of $[CuCl_4]^{2-}$. This is achieved by including more orbitals in the Q-space. These orbitals can be either doubly occupied valence orbitals or empty virtual orbitals. In the first case (Table 8b), the number of allowed excitations (*holes*) from these MOs to the (11,6) P-space and to the virtual Q-space is examined. A calculation with a minimal (11,11) P+Q space, SplitGAS-4(11,6)//(11,11), is the starting point of the systematic expansion of the Q-space and it is reported for completeness (Table 8a). This level of theory fails to provide quantitative results for the excitation energies under study. Table 8b shows the systematic increase of the valence orbitals in which one- or two-holes are created. For the first two ligand-field states (1^2B_{1g} and 1^2B_{2g}/1^2B_{3g}), the excitation energies vary between 1.25 and 1.33 eV, and 1.56 and 1.64 eV, respectively. These energies are independent of the number of correlating orbitals and the number of excitations (one or two holes). The excitation energy obtained for the 3^2A_g charge transfer state differs by about 2 eV from the CASPT2(11,11) value. In Table 8c, the VEEs upon expansion of the virtual space included in the Q-space are reported. The choice of the MOs included in the Q-space was based on energy considerations. As in the case where the number of holes was systematically increased, the two ligand field excitations are in agreement with the reference values. For the 1^2B_{1g} state, the difference

from the reference value is 0.02 eV, and for the $1^2B_{2g}/1^2B_{3g}$ state the difference is 0.18 eVThe latter deviation is likely due to the initial choice of P-space. The SplitGAS-7(11,11)//(33,37) level is significantly more accurate, where the five 4*d* orbitals of Cu are included in the P-space. Conversely, the VEE of the charge transfer state is systematically decreased from 6.55 eV (SplitGAS-7(11,6)//(49,60)) to 5.10 eV (SplitGAS-7(11,6)//(49,206)), approaching the value of CASPT2 (4.60 eV). Despite this fortuitous decrease of the 3^2A_g VEE, we conclude that the current implementation of the method cannot accurately calculate charge transfer states when a small Q-space is chosen. Suggestions as to how this problem can be overcome are provided in the next section.

Table 8. Systematic expansion of the Q-space of SplitGAS for the low energy ligand field and charge transfer VEEs of $[CuCl_4]^{2-}$. All energies in eV. The number of CSFs of the 1^2A_g ground state is given in the last column.

Level		1^2B_{1g}	1^2B_{2g}/ 1^2B_{3g}	3^2A_g	CSFs
(a) Minimal P- and Q-spaces					
SplitGAS-4(11,6)//(11,11)		1.06	1.38	8.49	546
(b) Expansion of *holes*					
SplitGAS-6(11,6)//(11,41)	No *holes*	1.39	1.64	6.88	12 408
SplitGAS-7(11,6)//(33,52)	1*h* from 3*p* Cl	1.30	1.57	6.63	86 079
	2*h* from 3*p* Cl	1.25	1.56	6.89	222 630
SplitGAS-7(11,6)//(41,56)	1*h* from 3*s*3*p* Cl	1.31	1.57	6.60	113 171
	2*h* from 3*s*3*p* Cl	1.28	1.59	6.74	369 392
SplitGAS-7(11,6)//(49,60)	1*h* from 3*s*3*p* Cl	1.33	1.59	6.55	140 199
	1*h* from 3*s*3*p* Cu				
	2*h* from 3*s*3*p* Cl	1.29	1.60	6.70	552 936
	2*h* from 3*s*3*p* Cu				
SplitGAS-8(11,6)//(73,72)	2*h* from 3*s*3*p* Cl	1.29	1.60	6.71	1 158 826
	2*h* from 3*s*3*p* Cu				
	1*h* from 2*p* Cl				
(c) Expansion of *virtual space*					
SplitGAS-7(11,6)//(49,60)	Threshold: 1.0 E_h	1.33	1.59	6.55	140 199
SplitGAS-7(11,6)//(49,74)	1.5 E_h	1.33	1.67	6.40	1 086 697
SplitGAS-7(11,6)//(49,94)	2.0 E_h	1.34	1.65	6.41	1 848 705
SplitGAS-7(11,6)//(49,122)	2.5 E_h	1.27	1.58	6.50	4 237 455
SplitGAS-7(11,6)//(49,148)	4.0 E_h	1.36	1.97	6.00	6 816 011
SplitGAS-7(11,6)//(49,168)	10.0 E_h	1.50	1.85	5.50	9 215 204
SplitGAS-7(11,6)//(49,206)	All Virtual MOs	1.57	1.94	5.10	14 757 732
Exp.		1.55	1.76		

2.1.5 Discussion and Conclusions

The GASSCF method has several interesting features. It can be used to explore large active spaces that would not affordable in CASSCF to detect the contribution of specific excitations (eg. π to σ^* or π^* orbitals). Moreover, it allows the use of more compact CI expansions than in the corresponding CASSCF calculations. A systematic extension of the size of the GAS spaces or the number of excitations between GAS

spaces results in the capture of a larger portion of static correlation energy. In order to include dynamical correlation, one should include *n*-electron excitations to the virtual MO space. Two bottlenecks hinder the applicability of the method: the exponential scaling of the CI space, since this is a conventional multireference CI expansion, and the lack of a complete incorporation of dynamical correlation. SplitGAS addresses the latter issue by generating larger and more accurate multi configurational wave functions. The applicability of the method for the calculation of excited states was demonstrated in this study.

Ozone was the first test case examined. SplitGAS-6(12,9)//(18,40) showed a very good performance for vertical excitation energies with deviations from experimental values not larger than 0.2 eV. The only exception was for the 1^1B_2 state, which was discussed more thoroughly by comparing data from methods that account for both static and dynamical correlation. For ionization energies and electron affinities, SplitGAS performs similarly to MRCI singles-and-doubles (SD). The π-to-π^* excitations of furan and pyrrole and the four lowest states of nickel dioxide were investigated. For all these systems, the accuracy of SplitGAS is comparable with MRCISD. In both computational methods, the same primary space was used, while the formation of the CI-expansion of the extended space differs. Finally, the ligand field excitations and a charge transfer state of $[CuCl_4]^{2-}$ were examined by means of the SplitGAS method.

For the case of $[CuCl_4]^{2-}$, when a limited primary space (11,6) was used, the results are affected by unbalanced active spaces for specific excited states. In other words, the MOs included in the P-space were not the same for all states. This lead to erroneous results and treatment of ground and excited states on a different footing. We overcome this problem by always starting the SplitGAS calcultions from the MOs of the ground state. An alternative approach is the localization of the virtual space, which

allows the choice of MOs with favorable predominant character (e.g. the second 4*d* shell for a 3*d*-transition metal). This study shows that this approach does not significantly increase the accuracy of the method. Thus, for a more robust treatment of the correlated P- and Q-spaces, the implementation of an orbital optimization algorithm is necessary, and we are currently working on it. In a SplitGAS-MCSCF-type methodology, the electronic energy is minimized with respect to both the CI-expansion and the atomic orbital coefficients.

SplitGAS, like all methods based on perturbation theory, is non-variational. Therefore, increasing the size of the Q-space may affect different states in different ways. As a consequence, energy differences are not necessarily reproduced more accurately by expanding the Q-space. This is particularly true when the choice of the size of the Q-space is based on an arbitrary energy threshold and not on a choice of physically relevant virtual MOs (eg. ligand virtual MOs, Rydberg orbitals, etc.). In the $[CuCl_4]^{2-}$ case, for example, a systematic extension of the Q-space does not guarantee a systematic increase (or decrease) of the excitation energy (Table 8). On the contrary, fluctuations were observed. We believe that a more systematic choice of orbitals included in the Q-space can provide a robust solution that alleviates this discrepancy. This choice should be based on the type of molecular orbitals (determined by their atomic orbital contribution) desired in the extended space. For a diatomic molecule, the atomic orbitals included in Q-space can be carefully chosen based on symmetry considerations. This has been effectively demonstrated by the calculation of the dissociation energy curve of the chromium dimer at the SplitGAS-6(12,12)//(24,48) level.[3] For polyatomic systems, like the $[CuCl_4]^{2-}$ molecule, an efficient orbital optimization procedure should be included, similarly to the CASSCF method. We are making progress in this direction. In such a way, a more physically motivated choice of the orbitals will make the SplitGAS method considerably more accurate.

2.1.6 Additional Information

Supporting Information: Natural orbitals of the O_3 and furan molecules; SplitGAS and GASSCF results for various systems with different active spaces; input example for GASSCF calculations; and equilibrium geometries and point groups. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00191.

This work was supported in part by the (U.S.) Department of Energy (DOE), Office of Basic Energy Sciences, under SciDAC Grant No. DE-SC0008666.

2.2 Systematic Design of Active Spaces for Multi-Reference Calculations of Singlet–Triplet Gaps of Organic Diradicals, with Benchmarks Against Doubly Electron-Attached Coupled-Cluster Data

This section describes the outcome of a collaborative research project carried out by Samuel J. Stoneburner, Jun Shen, and Adeayo O. Ajala (and advised by Piotr Piecuch, Donald G. Truhlar, and Laura Gagliardi). A report on this research project has been published.[206]

Samuel J. Stoneburner performed all CASSCF/CASPT2 and RASSCF/RASPT2 calculations and wrote the corresponding sections of the manuscript.

2.2.1 Overview

Singlet–triplet gaps in diradical organic π-systems are of interest in many applications. In this study, we calculate them in a series of molecules, including cyclobutadiene and its derivatives and cyclopentadienyl cation, by using correlated participating orbitals within the complete active space (CAS) and restricted active space (RAS) self-consistent field frameworks, followed by second-order perturbation theory (CASPT2 and RASPT2). These calculations are evaluated by comparison with the results of doubly electron-attached (DEA) equation-of-motion (EOM) coupled-cluster (CC) calculations with up to 4-particle–2-hole (4*p*-2*h*) excitations. We find active spaces that can accurately reproduce the DEA-EOMCC(4*p*-2*h*) data while being small enough to be applicable to larger organic diradicals.

2.2.2 Introduction

Organic diradicals are of interest as reaction intermediates[207] and in a variety of applications, including photochemical pathways,[208] molecular magnets,[209] magnetic resonance imaging,[210] spintronics,[211,212] nonlinear optics,[213] and photovoltaics.[214–218] One of the most important characteristics of diradical molecules is the energy gap between their lowest singlet and triplet states, ΔE_{ST}. The persistence of magnetic properties at room temperature typically requires a triplet ground state with ΔE_{ST} of at least a couple of kcal/mol[219] and the magnitude of ΔE_{ST} plays a direct role in singlet fission.[214] However, determining accurate values of ΔE_{ST} for diradicals remains a challenge, even when high-level *ab initio* methods are employed.[220–225] This is because diradicals feature low-lying open-shell singlet states with nearly degenerate singly occupied molecular orbitals (SOMOs)[226–228] and challenging closed-shell singlets with multiple significantly contributing configuration state functions,[225,229] and the treatment of these states has to be

balanced with the treatment of triplet states that have a single-reference nature. One can try to use conventional single-reference methods, such as the coupled-cluster (CC)[230–235] approaches of the CCSD[236] or CCSD(T)[237] type, or use the Kohn-Sham (KS) density functional theory (DFT) with symmetry-broken solutions,[221] but these approaches can lead to spin-contaminated results and an erratic description of the multi-determinantal singlet states.[225,238,239] In this work, we turn to multireference methods and new generations of particle-non-conserving single-reference CC schemes that can address deficiencies of other quantum chemistry approaches in applications involving diradicals in a computationally manageable fashion.

The most widely used multireference methods are based on complete active space self-consistent field (CASSCF)[5,6,45] reference states. In the CASSCF method, the wave function is defined by partitioning MOs into three disjoint sets, namely, the inactive, active, and external orbitals. The inactive orbitals are kept doubly occupied and the external orbitals are kept empty during the calculations. The electrons in the active orbitals are allowed to distribute in all possible ways, generating a full configuration interaction (CI) state within the active space.[5,240] In order to obtain reliable results, the active orbitals should be chosen such that the configuration state functions (CSFs) included in the CASSCF calculation dominate the electronic states of interest, capturing the correlation effects due to electronic near-degeneracies. CASSCF should provide a good treatment of static correlation, but it neglects most of the dynamical correlation effects that originate from short-range electron-electron repulsion and long-range dispersion interactions.[241] In the present work, the missing dynamic correlations are added with the help of multireference perturbation theory,[45] following the complete active space second order perturbation theory (CASPT2) model.[8,9]

CASPT2 allows one to handle electronic near-degeneracies and dynamic correlations in a reasonably accurate and balanced manner if adequate active spaces can be found and used. However, in analogy to full CI, the number of CSFs in the active space scales factorially with the numbers of active orbitals and electrons.[53] As a result, the CASSCF and CASPT2 approaches with active spaces larger than 16 electrons in 16 orbitals are unaffordable with current standard programs.[53] Thus, it is desirable to consider less expensive alternatives to CASSCF for generating reference wave functions for the subsequent multireference perturbation theory, CI, and CC calculations. One such alternative is offered by the restricted active space SCF (RASSCF) approach,[7] which decomposes the active orbital space into three subspaces, abbreviated as RAS1, RAS2, and RAS3, so that the numbers of CSFs used in the CI diagonalization steps are much smaller than those characterizing CASSCF calculations. In RAS1, all orbitals are doubly occupied except for electronic excitations up to a certain excitation rank (typically, two) into RAS2 and RAS3. The active orbitals in RAS3 are unoccupied except for electronic excitations up to a certain excitation rank (once again, typically, two) from RAS1 and RAS2. The remaining active electrons are distributed among the available RAS2 orbitals in all possible ways. RASSCF allows much larger active spaces than those that can presently be used in CASSCF computations, but the calculations can still become unaffordable as the system size increases, so finding ways to minimize the numbers of active electrons and active orbitals in multireference work remains an important objective.[47,52,88,90] In a typical application, the choice of active space is made by chemical intuition and trial and error. This makes the results of multireference calculations user-dependent and the choice of adequate active space can be labor intensive. Here, we instead consider a more systematic procedure, namely the "correlated participating orbitals" (CPOs) scheme proposed in Ref. [25]. Originally developed for reactions and

barrier heights, the CPO scheme has recently been systematically and successfully applied to singlet-triplet splittings in divalent radicals.[242,243]

The main objective of the present study is to explore the usefulness of CPOs in CASPT2 and RASPT2 calculations of the singlet–triplet gaps in a series of organic diradical π-systems that were previously explored by Saito *et al*.[239] using the restricted and unrestricted CCSD and CCSD(T) methods, the state-specific multireference CCSD approach of Mukherjee and co-workers,[244] abbreviated as MkCCSD, and unrestricted KS-DFT approaches employing selected exchange–correlation functionals. We systematically examine three CPO-type active spaces and their subdivisions with the goal of finding active spaces that can provide a reliable description of the systems examined in this work and that can serve as the basis for a more general recipe, which might be used in CASPT2 and RASPT2 calculations for other diradical organic π-systems in the future.

The significant disagreements among the different methods employed by Saito *et al*.[239], as well as in various other papers (e.g., Refs. [245–247]), show that the systems examined by these authors and in the present study are computationally very challenging. In particular, the various single- and multireference CC results for the singlet–triplet gaps reported in Ref. [239] are not consistent enough to serve as reliable reference values to benchmark our CPO-based CASPT2 and RASPT2 methods. To address this concern, we performed in this work new benchmark calculations using methods based on the doubly electron-attached (DEA) equation-of-motion (EOM) CC formalism, [248–253] which belongs to a broader category of particle non-conserving EOMCC theories (see Refs. [254] and [134] for selected reviews). The DEA-EOMCC framework allows one to determine ground and excited states of systems, such as diradicals, that are formally obtained by attaching two electrons to closed shells. In addition to the usual features of the CC/EOMCC methodology, such as fast convergence toward the exact, full CI limit, size extensivity in

describing the underlying ground states, and size intensivity of the excitation (in this case, electron attachment) energies, the DEA-EOMCC calculations produce wave functions that are automatically adapted to the spin symmetry, i.e., one avoids the spin-contamination issues that arise when the conventional single-reference CC and EOMCC approaches using the unrestricted Hartree-Fock (UHF) or restricted open-shell Hartree-Fock (ROHF) references are exploited. Because of our interest in providing reliable data for benchmarking the CPO-based CASPT2 and RASPT2 approaches, we focus on the DEA-EOMCC calculations with up to four-particle–two-hole (4*p*-2*h*) components in the corresponding electron-attachment operator, which, as shown in Refs. [251–253], provide a nearly exact description of the electronic spectra of diradicals. Since the full DEA-EOMCC(4*p*-2*h*) calculations for systems with larger numbers of electrons are prohibitively expensive, we use the more practical active-space DEA-EOMCC models, in which one selects the leading 4*p*-2*h* or 4*p*-2*h* and 3*p*-1*h* contributions with the help of small subsets of active orbitals.[251–253] As shown in Refs. [251–253], the DEA-EOMCC approaches with an active-space treatment of 4*p*-2*h* or 4*p*-2*h* and 3*p*-1*h* components accurately reproduce the results of the full DEA-EOMCC(4*p*-2*h*) calculations at the small fraction of the computational cost, so they are well suited for generating reliable data for benchmarking the CPO-based CASPT2 and RASPT2 methods in this work. Although, as shown in this work, the effect of the basis set on the calculated singlet–triplet values is small, we consider it as well by combining the highest-level DEA-EOMCC(4*p*-2*h*)-type data obtained with a smaller basis set with the results of the larger-basis set DEA-EOMCC(3*p*-1*h*) calculations, in which 4*p*-2*h* terms are neglected.

In summary, the main objective of this work is to test various choices of the CPO active spaces and their RAS subdivisions, so that we can find optimum spaces that predict singlet–triplet gaps that are in good agreement with the DEA-EOMCC benchmark data, while being small enough to be applicable to larger organic diradicals.

By having access to the DEA-EOMCC information at the 3*p*-1*h* and 4*p*-2*h* levels, we can also comment on the importance of 4*p*-2*h* contributions in studies of diradicals.

2.2.3 Computational Details

2.2.3.1 Molecular Systems Examined in this Study

The following diradical systems, shown in Figure 5, are considered in this work: D_{4h}-symmetric form of cyclobutadiene (**1**), D_{5h}-symmetric cyclopentadienyl cation (**2**), and five cyclobutadiene derivatives with polar substituents, including C_1-symmetric aminocyclobutadiene (**3**), C_1-symmetric formylcyclobutadiene (**4**), C_1-symmetric 1-amino-2-formyl-cyclobutadiene (**5**), C_{2v}-symmetric 1,2-bis(methylene)cyclobutadiene (**6**), and D_{2h}-symmetric 1,3-bis(methylene)cyclobutadiene (**7**). All geometries were taken from Saito et al.[239] Notice that, even for the more symmetric systems, all calculations were performed in the C_1 point group. Each system features two degenerate (systems **1** and **2**) or nearly degenerate (the remaining systems) singly occupied π orbitals centered primarily on the carbon rings. The singly occupied orbitals of system **1** are shown in Figure 6. In each case, the lowest-energy singlet and triplet states differ by a spin-flip π→π* transition and the corresponding energy gap is defined by

$$\Delta E_{ST} = E_{singlet} - E_{triplet} \tag{71}$$

where a negative number indicates that the singlet is lower in energy. In the case of system **1**, the singlet is $^1B_{1g}$ and the triplet is $^3A_{2g}$. (However, we run the calculations without imposing symmetry constraints).[255]

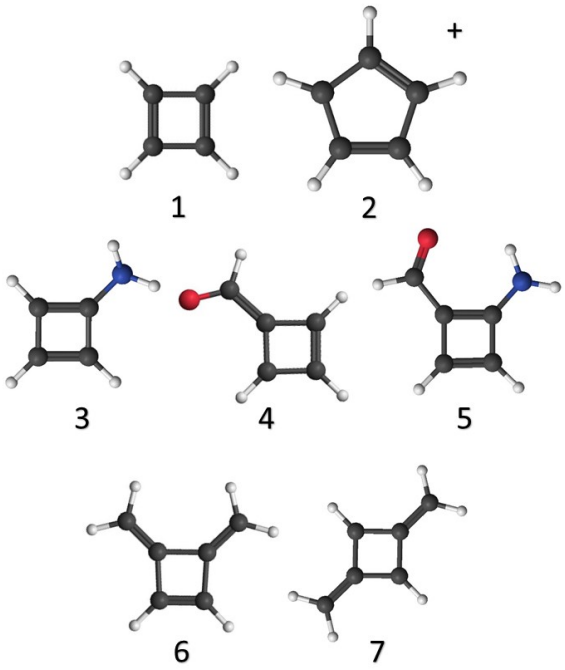


Figure 5*:* Diradical systems under investigation. **1**: C_4H_4**, 2**: $C_5H_5^+$**, 3**: $C_4H_3NH_2$**, 4**: C_4H_3CHO**, 5**: $C_4H_2NH_2CHO$**, 6**: C_4H_2-1,2-$(CH_2)_2$**, 7**: C_4H_2-1,3-$(CH_2)_2$**.**

The orbitals involved in the singlet–triplet transition, whose occupancies change in the dominant CSFs, are always the frontier orbitals, which are the SOMOs except for system **5**. These SOMOs are singly occupied π orbitals located primarily on the carbon ring, although the substituents are also involved for systems **6** and **7**. The lowest singlet of system **5** has frontier orbitals with occupation numbers close to two or zero, so for the singlet state of system **5** the frontier orbitals are the highest occupied and lowest unoccupied MOs**.**

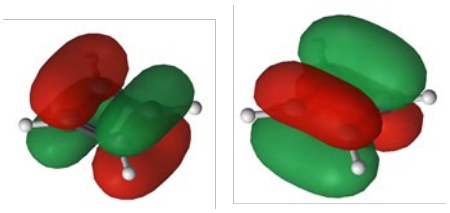


Figure 6: Singly occupied π orbitals for system **1**.

2.2.3.2 Benchmark DEA-EOMCC Calculations

The DEA-EOMCC methods aim at the determination of ground and excited states of systems, such as diradicals, which can be obtained by attaching two electrons to the corresponding closed-shell cores. This is accomplished using the wave function ansatz $\left|\Psi_\mu^{(N)}\right\rangle = R_\mu^{(+2)}\left|\Psi_0^{(N-2)}\right\rangle$, where $\left|\Psi_\mu^{(N)}\right\rangle$ is the ground ($\mu = 0$) or excited ($\mu > 0$) state of the N-electron diradical of interest, $\left|\Psi_0^{(N-2)}\right\rangle = e^T\left|\Phi^{(N-2)}\right\rangle$ is the CC ground state of the (N – 2)-electron closed-shell core (with T and $\left|\Phi^{(N-2)}\right\rangle$ representing the corresponding cluster operator and reference determinant), and $R_\mu^{(+2)} = R_{\mu,2p} + R_{\mu,3p\text{-}1h} + R_{\mu,4p\text{-}2h} + \cdots$ is the operator attaching two electrons to $\left|\Psi_0^{(N-2)}\right\rangle$ using the 2p component $R_{\mu,2p}$, while allowing the relaxation of the remaining electrons via its 3p-1h ($R_{\mu,3p\text{-}1h}$), 4p-2h ($R_{\mu,4p\text{-}2h}$), and other many-body components.

As shown in Refs. [251–253], the level of the DEA-EOMCC theory that provides a very accurate description of diradical electronic spectra, including energy gaps between the low-lying singlet and triplet states, is DEA-EOMCC(4p-2h), where the electron-attaching operator $R_\mu^{(+2)}$ is truncated at the 4p-2h component $R_{\mu,4p\text{-}2h}$. Unfortunately, the most expensive steps of full DEA-EOMCC(4p-2h) scale as $n_o^2 n_u^6$, where n_o (n_u) is the number of orbitals occupied (unoccupied) in the underlying reference determinant $\left|\Phi^{(N-2)}\right\rangle$, limiting the DEA-EOMCC(4p-2h) calculations to small systems. However, as demonstrated in Refs. [251–253], it is sufficient to use small subsets of orbitals unoccupied in $\left|\Phi^{(N-2)}\right\rangle$ to select the dominant 4p-2h terms, with virtually no loss in accuracy and at the

small fraction of the cost of parent DEA-EOMCC(4*p*-2*h*) computations. The resulting DEA-EOMCC(4*p*-2*h*)$\{N_u\}$ approach, where $N_u << n_u$ designates the number of active unoccupied orbitals used to select the leading 4*p*-2*h* contributions, which belongs to a larger family of the active-space CC and EOMCC theories,[256] reduces the $n_o^2 n_u^6$ steps of its full DEA-EOMCC(4*p*-2*h*) parent to a more manageable $N_u^2 n_o^2 n_u^4$ level. One can use similar ideas to select the dominant 3*p*-1*h* contributions, either within the DEA-EOMCC(4*p*-2*h*)$\{N_u\}$ scheme, or within its lower-level DEA-EOMCC(3*p*-1*h*) counterpart where 4*p*-2*h* terms are neglected, replacing the $n_o n_u^5$ steps associated with 3*p*-1*h* contributions by the less expensive $N_u n_o n_u^4$ operations.[253] As shown in Ref. [253], the resulting DEA-EOMCC(3*p*-1*h*,4*p*-2*h*)$\{N_u\}$ and DEA-EOMCC(3*p*-1*h*)$\{N_u\}$ approaches accurately reproduce the corresponding DEA-EOMCC(4*p*-2*h*) or DEA-EOMCC(4*p*-2*h*)$\{N_u\}$ and DEA-EOMCC(3*p*-1*h*) data at the small fraction of the computational costs.

The highest level of the DEA-EOMCC theory used in this work is DEA-EOMCC(4*p*-2*h*)$\{N_u\}$. In carrying out the DEA-EOMCC(4*p*-2*h*)$\{N_u\}$ calculations, we followed Saito et al.[239] and used the cc-pVDZ basis set.[257] In order to examine the dependence of our results on the basis set, we also used the larger maug-cc-pVTZ basis.[258] For the larger systems considered in this study, namely, cyclobutadiene derivatives, the DEA-EOMCC(4*p*-2*h*)$\{N_u\}$/maug-cc-pVTZ calculations using our present codes turned out to be quite expensive, so to estimate the DEA-EOMCC(4*p*-2*h*)$\{N_u\}$/maug-cc-pVTZ results we adopted a simple extrapolation scheme, abbreviated as DEA-EOMCC[4*p*-2*h*], where we calculate the final energies as follows:

$$E[4p\text{-}2h] = E(4p\text{-}2h)\{N_u\}/\mathrm{DZ} + E(3p\text{-}1h)\{N_u\}/\mathrm{mTZ} - E(3p\text{-}1h)\{N_u\}/\mathrm{DZ} \qquad (72)$$

The first term on the right-hand side of Equation (72) is the DEA-EOMCC(4*p*-2*h*)$\{N_u\}$/cc-pVDZ energy. The effect of going from the cc-pVDZ basis set (abbreviated as DZ) to maug-cc-pVTZ (abbreviated as mTZ) is estimated by forming the difference of

energies obtained in the DEA-EOMCC(3p-1h)\{N_u\}/maug-cc-pVTZ and DEA-EOMCC(3p-1h)\{N_u\}/cc-pVDZ calculations.

In addition to the calculations entering Equation (72), we performed the full DEA-EOMCC(3p-1h) and active-space DEA-EOMCC(3p-1h,4p-2h)\{N_u\} computations using the cc-pVDZ basis set (all seven systems) and the DEA-EOMCC(3p-1h,4p-2h)\{N_u\}/maug-cc-pVTZ calculations for the smallest system **1**. We carried out these extra computations to validate Equation (72), especially the usefulness of the DEA-EOMCC(3p-1h)\{N_u\} approach in estimating the effect of going from the cc-pVDZ basis set to maug-cc-pVTZ (see Section 2.2.4.1 for a discussion). Following Refs. [251–253], in all of the DEA-EOMCC calculations performed in this work, the ground states of the underlying (N – 2)-electron closed-shell cores were obtained using CCSD.

All of the DEA-EOMCC calculations reported in this work and the underlying CCSD computations were performed using the restricted Hartree-Fock (RHF) MOs corresponding to the (N – 2)-electron closed-shell cores. In this way, we could maintain all of the relevant symmetries throughout the calculations. We tested the usage of other orbitals, such as the N-electron ROHF MOs obtained for the triplet states of diradicals examined in this work, but, in agreement with Refs. [251–253], the resulting singlet–triplet gaps, especially those obtained with the highest DEA-EOMCC(4p-2h)-type levels, turned out to be virtually independent of the type of MOs used in the calculations. As in Ref. [239], in all of the post-HF calculations, the core orbitals correlating with the 1s shells of the C, N, and O atoms were kept frozen and the spherical components of d and f basis functions were employed throughout.

In carrying out the various DEA-EOMCC computations, we followed the strategy employed in Ref. [239]. Thus, we used the D_{4h} point group for system **1**, the D_{5h} group for system **2**, and C_1 for the remaining systems **3**–**7**. In each case, the closed-shell (N – 2)-

electron reference system used to set up the DEA-EOMCC calculations was obtained by vacating the two valence partly occupied orbitals that define the singlet and triplet states of interest, which are exactly degenerate in systems **1** and **2** and nearly degenerate in systems **3**–**7**. For example, the (N – 2)-electron reference dication used in the DEA-EOMCC calculations for system **1** was obtained by vacating the two valence SOMOs of e_g symmetry. For system **2,** we vacated the degenerate valence e_1'' shell, etc. Consistent with the structure of the valence π shells in systems **1**–**7**, which consist of one doubly occupied, two partly occupied, and one unoccupied MOs in systems **1**, **3**, and **4** and one doubly occupied, two partly occupied, and two unoccupied MOs in systems **2** and **5**–**7**, the active spaces needed to perform the DEA-EOMCC(3*p*-1*h*)$\{N_u\}$, DEA-EOMCC(3*p*-1*h*,4*p*-2*h*)$\{N_u\}$, and DEA-EOMCC(4*p*-2*h*)$\{N_u\}$ calculations were defined in the following manner. For systems **1**, **3**, and **4**, we used the N_u = 3 MOs, which are the three lowest-energy unoccupied orbitals in the respective $\mathbf{1}^{2+}$, $\mathbf{3}^{2+}$, and $\mathbf{4}^{2+}$ reference dications. For systems **2**, **5**, **6**, and **7**, we used the N_u = 4 orbitals, which are the four lowest-energy unoccupied MOs in the respective (N – 2)-electron $\mathbf{2}^{2+}$, $\mathbf{5}^{2+}$, $\mathbf{6}^{2+}$, and $\mathbf{7}^{2+}$ species. We verified the appropriateness of the above active orbital choices by comparing the full DEA-EOMCC(3*p*-1*h*)/cc-pVDZ and active-space DEA-EOMCC(3*p*-1*h*)$\{N_u\}$/cc-pVDZ data (see Section 2.2.4.1 for further discussion).

All of the DEA-EOMCC calculations reported in this work were performed using the codes developed in Refs. [251–253], interfaced with GAMESS[259] and taking advantage of the spin-free CCSD GAMESS routines[260] and the routines used in some of our earlier EOMCC studies.[181,261,262]

2.2.3.3 CASPT2 and RASPT2 Calculations

CASPT2[8] and RASPT2[10] calculations, including the underlying reference state calculations by CASSCF[6] and RASSCF,[7] were performed using the maug-cc-pVTZ[263] and ANO-RCC-VTZP[264] basis sets with Cholesky decomposition[185] using a developer version of Molcas 8.1.[53,153,265] Orbitals were visualized using Luscus 0.8.3.[266] All calculations were performed without symmetry restrictions, i.e., in C_1 symmetry.

For CASSCF calculations, the active space notation is (*n,N*), where *n* is the number of active electrons, and *N* is the number of orbitals in the active space. For RASSCF calculations, the active space notation is ($n,h,p;N_1,N_2,N_3$), where *n* is the total number of active electrons, has is the maximum number of holes in RAS1, *p* is the maximum number of particles in RAS3, and N_i is the number of orbitals in R*i*.

CASPT2 and RASPT2 calculations were performed with an imaginary shift of 0.1 hartrees to alleviate intruder state problems. The default IPEA shift of 0.25 hartrees[65] was used to compensate for the systematic overestimation of correlation energy in CASPT2. (In the supplementary material we give some comparison results obtained without an IPEA shift.)

2.2.3.3.1 CPO definitions

The CPO scheme is based on the idea that the active space should consist of "participating" orbitals, i.e., the orbitals most strongly involved in the process of interest, plus one correlating orbital for each participating orbital.[25] Participating orbitals are identified based on the orbital occupations from the dominant configurations, not the occupation numbers from the zeroth-order wave function. For all systems studied other than system **5**, considering the occupation numbers from the zeroth-order wave function would erroneously suggest that there is no difference between the singlet and the triplet,

as the frontier orbitals are singly occupied in both cases. However, although the triplet has a single dominant configuration state function, the singlet has two dominant configurations: one in which one of the frontier orbitals is doubly occupied, and one in which the other is doubly occupied. More specific information regarding the wave functions and dominant configurations is included in Section 2.2.4.2 and in the supplementary material.

The original CPO scheme had three choices: nominal, moderate, and extended, abbreviated as nom-, mod-, and ext-CPO.[25] In the present article we introduce a fourth option for π-systems that lies between nominal and moderate, referred to as "π-CPO". These four choices will be referred to for the rest of the paper as nCPO, πCPO, mCPO, and eCPO. In nCPO, active orbital are the frontier orbitals and their correlating orbitals. See Figure 6 and Figure 7 for examples of frontier orbitals and their correlating orbitals, respectively, for system **1**. In the other CPO options, as discussed next, we add additional orbitals on the atoms on which the frontier orbitals reside; these atoms are called "participating atoms". The substituent carbons of systems **6** and **7** are participating atoms, but the substituents of systems **3**, **4**, and **5** are not.

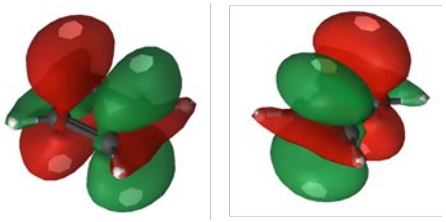


Figure 7*:* Correlating π' orbitals of the nCPO scheme for system **1**.

In the πCPO scheme, all valence π orbitals of participating atoms are active., and correlating orbitals are added as needed to ensure that each singly or doubly occupied orbital is paired with an unoccupied orbital. See Figure 8 for examples of participating π orbitals in system **1**.

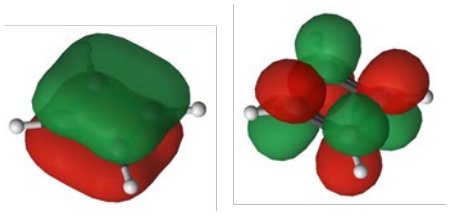


Figure 8*:* Included π orbitals in the πCPO scheme for system **1**.

In the mCPO scheme, all valence p orbitals of participating atoms are active, and correlating orbitals are added as needed to ensure that each singly or doubly occupied orbital is paired with an unoccupied orbital. Where there is significant *s*-*p* mixing, the *s* orbitals are taken to be those of lowest energy (one for each participating atom), and the rest are treated as *p* orbitals. For example, in system **1**, there are four participating atoms, and the four lowest valence orbitals are considered to be the *s* orbitals. See Figure 9 for examples of included *p* orbitals for system **1**.

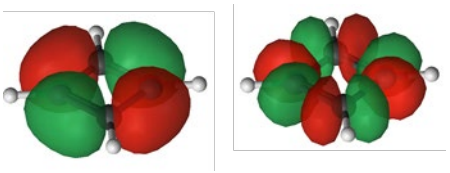


Figure 9*:* Included *p* orbitals of the mCPO scheme for system **1**.

In the eCPO scheme, all valence *p* and *s* orbitals of participating atoms are active, and correlating orbitals are added as needed to ensure that each singly or doubly occupied orbital is paired with an unoccupied orbital. See Figure 10 for examples of included *s* orbitals for system **1**.

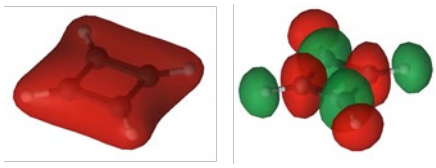


Figure 10*:* Included *s* orbitals of the eCPO scheme for system **1**.

2.2.3.3.2 RAS subdivisions

As discussed in the introduction, active spaces larger than sixteen electrons in sixteen orbitals are unaffordable in current standard implementations of CASSCF and CASPT2. However, the active spaces chosen using eCPO would be larger than the (16,16) limit, even for systems as small as those studied here. For system **1**, the smallest system under consideration, the active space chosen with eCPO would be (20,22), with 6 x 10^{10} CSFs for the singlet state. In order to employ such large active spaces, RASPT2 was employed. We also employed RASPT2 for some systems where CASPT2 is affordable because goal is to test the accuracy of RASPT2 against the benchmarks that are available on small systems so we know whether it is expected to be accurate for large systems where CASPT2 is not affordable.

In RASSCF calculations in this work, two excitations were permitted from RAS1 and two excitations were permitted into RAS3. Just as CPO provides a scheme for choosing which orbitals are included in the active space, it is also helpful to have a systematic way of choosing how to divide the active space into the three RAS subspaces.

For the systems studied here we define two schemes, "valence-π" and "limited-π". In the "valence-π" scheme, RAS2 includes any valence π orbitals that are included in the given CPO level; any π orbitals that are higher in energy than the valence orbitals are placed in RAS3, along with any other unoccupied orbitals. Any doubly occupied orbitals that are not part of the π-system are placed in RAS1.

For eCPO with systems **6** and **7**, the valence-π RAS division scheme results in active spaces of (30,2,2;12,6,14). The RASPT2 portion would have taken more than 22 days of CPU time, so valence-π was not performed for systems **6** and **7**. Instead, only the "limited-π" scheme was employed. Limited-π is similar to valence-π, but instead of all valence π orbitals being in RAS 2, occupied π orbitals below the highest two occupied π orbitals are in RAS1 and their correlating orbitals are in RAS3. Systems **1** through **5** have only two occupied π orbitals, so the limited-π active spaces are identical to the valence-π active spaces for those systems.

For systems **6** and **7**, convergence could not be achieved with limited-π RASSCF using an active space defined by πCPO, so instead RCI calculations were performed using orbitals from valence-π RASSCF with an active space defined by πCPO. Additionally, the CASSCF active space defined by mCPO for system **2** was prohibitively expensive to perform, and instead CCI calculations were performed using orbitals from valence-π RASSCF with an active space defined by mCPO.

2.2.4 Results and Discussion

2.2.4.1 DEA-EOMCC Benchmark Calculations

The results of our various DEA-EOMCC calculations for the singlet–triplet gaps in systems **1**–**7** are summarized in Table 9. Our highest-level calculated DEA-EOMCC(4*p*-2*h*){N_u}/cc-pVDZ data and their extrapolation to the larger maug-cc-pVTZ

basis set using Equation (72), abbreviated as DEA-EOMCC[4*p*-2*h*], which we treat in this work as best estimates of the $\Delta E_{S\text{-}T}$ values of interest, indicate that systems **1** and **3**–**6** have singlet ground states, whereas the ground states of systems **2** and **7** are triplets. As shown in Table 10, where we compare our extrapolated DEA-EOMCC[4*p*-2*h*] values with the singlet–triplet gaps resulting from the symmetry-broken, UHF-based, calculations using the single-reference CCSD(T) (UCCSD(T)) approach and its Brueckner-orbital UBD(T) analog,[267] and the multireference MkCCSD computations using the ROHF and CASSCF orbitals, our DEA-EOMCC(4*p*-2*h*)-level results agree in this regard with the findings of Saito et al.[239]

Table 9: The various DEA-EOMCC results for the singlet–triplet gaps $\Delta E_{S\text{-}T}$ (in kcal/mol) in systems **1**–**7**.

Method	**1**	**2**	**3**	**4**	**5**	**6**	**7**
cc-pVDZ							
(3*p*-1*h*){N_u}	-1.37	16.38	0.44	-1.24	-6.90	-81.63	18.26
(3*p*-1*h*)	-1.42	16.06	0.34	-1.32	-7.46	-81.84	17.95
(3*p*-1*h*,4*p*-2*h*){N_u}	-4.98	14.25	-3.22	-4.32	-4.82	-78.42	20.03
(4*p*-2*h*){N_u}	-5.04	13.91	-3.30	-4.40	-5.49	-78.75	19.76
maug-cc-pVTZ							
(3*p*-1*h*){N_u}	-0.53	16.35	1.09	-0.48	-7.09	-80.56	16.98
[4*p*-2*h*][a]	-4.20[b]	13.88	-2.65	-3.65	-5.68	-77.68	18.49
N_u	3	4	3	3	4	4	4

[a] Best estimate defined by the extrapolation formula given by Equation (72).
[b] The DEA-EOMCC(3*p*-*1h*,4*p*-2*h*){3}/maug-cc-pVTZ calculation gives -4.08 kcal/mol.

Before making further comparisons between the results of our DEA-EOMCC calculations and the $\Delta E_{S\text{-}T}$ values reported in Ref. [239], we comment on the extrapolation procedure defined by Equation (72), which is used in this work to provide reference data

for benchmarking the CPO-based CASPT2 and RASPT2 schemes. We begin with the choice of active orbitals used to select the dominant 3*p*-1*h* and 4*p*-2*h* contributions in the DEA-EOMCC(3*p*-1*h*)$\{N_u\}$ and DEA-EOMCC(4*p*-2*h*)$\{N_u\}$ computations. A comparison of the results of the full DEA-EOMCC(3*p*-1*h*) and active-space DEA-EOMCC(3*p*-1*h*)$\{N_u\}$ calculations using the cc-pVDZ basis set demonstrates that our choice of active orbitals allowing us to select the dominant higher−than−2*p* contributions in the DEA-EOMCC wave function ansatz is appropriate. Indeed, as shown in Table 9, the differences between the singlet−triplet gaps resulting from the DEA-EOMCC(3*p*-1*h*)/cc-pVDZ and DEA-EOMCC(3*p*-1*h*)$\{N_u\}$/cc-pVDZ calculations are very small, ranging from 0.05 kcal/mol for system **1** to 0.56 kcal/mol for system **5**, where the DEA-EOMCC(3*p*-1*h*)/cc-pVDZ gap value is -7.46 kcal/mol. In fact, one observes similarly small differences when comparing the results of the higher-level DEA-EOMCC(4*p*-2*h*)$\{N_u\}$ calculations, in which 4*p*-2*h* terms are treated using active orbitals, but 3*p*-1*h* terms are treated fully, with the results obtained with the DEA-EOMCC(3*p*-1*h*,4*p*-2*h*)$\{N_u\}$ approach, in which both types of terms are treated using active orbitals. These observations are consistent with the well-known characteristic of the active-space CC and EOMCC methods, including the active-space DEA-EOMCC approaches employed in this study, which is their ability to reproduce the results of the parent CC/EOMCC calculations with small numbers of active-orbitals used in selecting higher-order excitations.[251–253,256] We can certainly conclude that the use of the active-space DEA-EOMCC(3*p*-1*h*)$\{N_u\}$ approach in Equation (72), as a substitute for the considerably more expensive full DEA-EOMCC(3*p*-1*h*) parent in estimating the effect of going from the cc-pVDZ basis set to the maug-cc-pVTZ basis, is an appropriate procedure.

Equation (72) is also justified by the fact that the effect of going from the smaller cc-pVDZ basis to the larger maug-cc-pVTZ basis set on the calculated $\Delta E_{S\text{-}T}$ values is generally rather small, implying that it is safe to estimate it using the lower-level DEA-

EOMCC(3*p*-1*h*){N_u} method, as opposed to the significantly more expensive DEA-EOMCC(4*p*-2*h*){N_u} approach. Indeed, as shown in Table 9, the differences between the singlet–triplet gaps resulting from the DEA-EOMCC(3*p*-1*h*){N_u}/cc-pVDZ and DEA-EOMCC(3*p*-1*h*){N_u}/maug-cc-pVTZ calculations range from 0.03 kcal/mol for system **2** to 1.28 kcal/mol for system **7**, where the DEA-EOMCC(3*p*-1*h*)/cc-pVDZ gap value is 18.26 kcal/mol, for an average of 0.69 kcal/mol. Furthermore, although we were unable to perform the DEA-EOMCC/maug-cc-pVTZ calculations with 3*p*-1*h* and 4*p*-2*h* terms in the electron-attaching $R_\mu^{(+2)}$ operator using our existing codes for all of the systems examined in this work, we managed to obtain the DEA-EOMCC(3*p*-*1h,*4*p*-2*h*){3}/maug-cc-pVTZ value for the singlet–triplet gap in system **1**, obtaining -4.08 kcal/mol (see Table 9). Our extrapolation of the DEA-EOMCC(4*p*-2*h*)/maug-cc-pVTZ-level result based on Equation (72) gives -4.20 kcal/mol, in virtually perfect agreement with the DEA-EOMCC(3*p*-*1h,*4*p*-2*h*){3}/maug-cc-pVTZ calculation. This means that Equation (72) works well, allowing us to capture the effect of high-order 4*p*-2*h* correlations and the effect of going from the cc-pVDZ basis set to maug-cc-pVTZ in an accurate and computationally manageable manner.

Table 10: A comparison of the ΔE_{ST} values (in kcal/mol) characterizing systems **1**–**7** obtained with the DEA-EOMCC[4*p*-2*h*] extrapolation defined by Equation (72) and in the DEA-EOMCC(4*p*-2*h*)\{N_u\}/cc-pVDZ calculations with the UCCSD(T), UBD(T), ROHF-MkCCSD, and CASSCF-MkCCSD results reported by Saito *et al*.[239]

Molecule	DEA-EOMCC [4*p*-2*h*]/ (4p-2h)\{N_u\}	Saito *et al*.[239]			
		UCCSD(T)	UBD(T)	ROHF-MkCCSD	CASSCF-MkCCSD
1	-4.2/-5.0	-4.8	-5.1	-8.6	-8.1
2	13.9/13.9	14.8	14.0	13.5	9.4
3	-2.7/-3.3	-3.2	-3.6	-6.5	-7.3
4	-3.6/-4.4	-4.5	-4.5	-7.1	-6.9
5	-5.7/-5.5	-0.6	-0.9	-2.7	-4.5
6	-77.7/-78.8	-82.7	-79.8	-82.7	-84.2
7	18.5/19.8	15.0	17.1	20.0	19.5
MUE[a]	0.0/0.7	2.4	1.6	3.1	3.6

[a]Mean unsigned errors relative to the extrapolated DEA-EOMCC[4*p*-2*h*] results using Equation (72).

Having established the validity of Equation (72), which, given the above analysis and previous extensive studies of the DEA-EOMCC approaches with up to 4*p*-2*h* excitations,[251–253] is expected to produce singlet–triplet gap values in systems **1**–**7** to within 1 kcal/mol or better, we comment on our best DEA-EOMCC[4*p*-2*h*] (and the corresponding DEA-EOMCC(4*p*-2*h*)\{N_u\}/cc-pVDZ) ΔE_{ST} values. First, it is important to note that although bulk of the correlation effects is captured at the DEA-EOMCC(3*p*-1*h*) level, the high-order 4*p*-2*h* effects can be quite substantial. When we compare the extrapolated DEA-EOMCC[4*p*-2*h*] and calculated DEA-EOMCC(3*p*-1*h*)\{N_u\}/maug-cc-pVTZ gap values, or, equivalently, the DEA-EOMCC(4*p*-2*h*)\{N_u\}/cc-pVDZ and DEA-EOMCC(3*p*-1*h*)\{N_u\}/cc-pVDZ data, the 4*p*-2*h* effects range, in absolute value, from 1.4 kcal/mol in system **5** to 3.7 kcal/mol in systems **1** and **3**. Although they typically reduce the total electronic energies of the individual states, their net effect on the calculated

singlet–triplet gaps can go either way. Indeed, we may encounter lowering of the signed ΔE_{ST} values due to 4*p*-2*h* correlations, as in systems **1**–**4**, or we can find cases where the signed singlet–triplet gaps defined by Equation (71) increase, as in systems **5**–**7**. In some cases, the 4*p*-2*h* effects can change a singlet–triplet gap near zero to a considerably larger absolute value, as in systems **1** and **4**, but there also are situations, such as system **3**, where 4*p*-2*h* correlations change state ordering and the sign of ΔE_{ST}. It is quite clear from the results shown in Table 9 that one has to account for the high-order 4*p*-2*h* effects within the DEA-EOMCC framework to obtain reasonably converged values of the singlet–triplet gaps in diradicals. This is consistent with our earlier DEA-EOMCC studies reported in Refs. [251–253].

High accuracy of our extrapolated DEA-EOMCC[4p-2h] data and the underlying DEA-EOMCC(4*p*-2*h*)$\{N_u\}$/cc-pVDZ calculations, which include sophisticated 4*p*-2*h* terms, in addition to their lower-rank 2*p* and 3*p*-1*h* counterparts, on top of CCSD, implies that we should be able to judge other methods. Before discussing our assessment of the various CPO-based CASPT2 and RASPT2 calculations in Section 2.2.4.2, we comment on the UCCSD(T), UBD(T), ROHF-MkCCSD, and CASSCF-MkCCSD computations reported by Saito et al.[239] As already pointed out, all of these methods agree in predicting correct state ordering. Unfortunately, as shown in Table 10, they disagree, sometimes rather significantly, in quantitative predictions. In the case of systems **1**–**4**, there is a great deal of consistency among the singlet–triplet gap values provided by UCCSD(T) and UBD(T) and those obtained in our DEA-EOMCC(4*p*-2*h*)$\{N_u\}$/cc-pVDZ and DEA-EOMCC[4*p*-2*h*] calculations, which agree to within ~1 kcal/mol, but one cannot say the same about the MkCCSD data, which seem to have rather large errors, on the order of 3–4 kcal/mol, displaying a significant dependence of the resulting ΔE_{ST} values on the type of orbitals used in the calculations in the case of system **2**. The poor performance of MkCCSD for system **1** is reinforced by the results of the multireference averaged

quadratic CC calculations,[268] reported in Ref. [239] as well, which give ΔE_{ST} of -5.5 kcal/mol, in good agreement with our highest-level DEA-EOMCC[4p-2h] and DEA-EOMCC(4*p*-2*h*){N_u}/cc-pVDZ calculations and the UCCSD(T) and UBD(T) data, but in sharp disagreement with the ROHF- and CASSCF-based MkCCSD values. Based on the results for systems **1**–**4** and the mean-unsigned error (MUE) values relative to DEA-EOMCC[4*p*-2*h*] reported in Table 10, one might recommend the use of the symmetry-broken UCCSD(T) and UBD(T) methods in the calculations of singlet–triplet gaps in diradicals, but the results for system **5**, where errors relative to DEA-EOMCC(4*p*-2*h*){N_u}/cc-pVDZ and DEA-EOMCC[4*p*-2*h*] in the UCCSD(T) and UBD(T) ΔE_{ST} values are on the order of 5 kcal/mol, show that this would be misleading. The agreement among the UCCSD(T), UBD(T), ROHF-MkCCSD, and CASSCF-MkCCSD ΔE_{ST} values improves, when systems **6** and **7** are examined, but one still observes substantial differences among the results obtained with these four methods, on the order of 4-5 kcal/mol, which do not allow us to use them to benchmark our CPO-based CASPT2 and RASPT2 approaches. Our extrapolated DEA-EOMCC[4p-2h] data and the underlying DEA-EOMCC(4*p*-2*h*){N_u}/cc-pVDZ calculations are considerably more reliable in this regard.

2.2.4.2 CASPT2 and RASPT2 calculations

CASPT2 and RASPT2 results using the maug-cc-pVTZ basis set are presented in Table 11. The various active spaces and their sizes are presented in Table 12. Results with the ANO-RCC-VTZP basis set are similar and are presented in the supplementary material.

The MUEs relative to the benchmark DEA-EOMCC[4*p*-2*h*] data shown in Table 11 are significantly larger for nCPO than for the other CPO choices. The MUEs of πCPO,

mCPO, and eCPO are all under 1.0 kcal/mol, while the MUEs of nCPO are between 6 and 8 kcal/mol. Moreover, the maximum error of nCPO is above 16 kcal/mol. These poor results indicate that the nCPO calculations do not properly reflect the multireference character of the singlets. As explained in section 2.2.3.3.1, the triplet states are dominated by a single-configuration, while the singlet states have two dominant configurations; the weights of these two configurations vary depending on system and active space, but they are roughly equal except for all systems other than system **5**. With nCPO, however, one configuration frequently outweighs the other, resulting in an inaccurate description of the wave function.

Although nCPO performs poorly, it most closely corresponds to the (2,2) active spaces used in the multireference CC calculations of Saito *et al*.[239] For both the singlet and triplet states, the π orbital directly below the nominal participating orbitals has an occupation number between 1.90 and 1.94, and the π orbital directly above the nominal participating orbitals has an occupation number between 0.06 and 0.10. See Figure 8 for examples of these orbitals from System **1**. Active spaces chosen with nCPO and the (2,2) active space used in Saito *et al*.[239] force these orbitals to have occupation numbers of 2.00 and 0.00, respectively. In contrast, the DEA-EOMCC calculations used for our reference values permit holes in the occupied π orbital and excitations into the unoccupied π orbital (among other possible holes and excitations that are more case-dependent), as do all CASPT2 and RASPT2 calculations presented here other than those using nCPO active spaces.

For πCPO, mCPO, and eCPO, there is little difference among the MUEs, as all MUEs are under 1.0 kcal/mol. The maximum errors vary depending on how the active space is divided, but in all cases they are 2.6 kcal/mol or below. We conclude that little is to be gained by using the eCPO active space, which roughly corresponds to a full-valence

active space for participating atoms and is unaffordable for CASPT2, even with the small systems studied here. The mCPO active space is also unaffordable for systems **6** and **7**, which are still very small systems, as the CASSCF active space would be (18,20) and have 5 x 10^9 CSFs for the singlet. Overall, πCPO offers the best balance between affordability and accuracy, as it provides a comparable level of accuracy with mCPO and eCPO, especially since Table 12 shows that it requires active spaces many orders of magnitude smaller than the order of 10^7 to 10^{11} CSFs required by mCPO and eCPO. Good accuracy is achieved with πCPO because only π orbitals have occupation numbers less than 1.97 or more than 0.03, regardless of whether additional orbitals are included in the active space. Therefore, πCPO allows for a sufficient description of the multireference character of these systems, and mCPO and eCPO add considerable expense for no significant benefit.

Valence-π RASPT2 has very similar MUEs to the corresponding CASPT2 calculations. All of the orbitals in RAS1 or RAS3 in valence-π RASSCF have occupation numbers between 2.00 and 1.97 or between 0.03 and 0.00, just as for CASSCF. Encouragingly, limited-π RASPT2 also enjoys similar accuracy to CASPT2, even though some orbitals placed in RAS1 or RAS3 have more intermediate occupation numbers associated with multireference character. This suggests that for larger systems featuring many more π orbitals, it may be possible to use limited-π RASPT2 with πCPO to keep computational costs low by placing most of the active orbitals in RAS1 or RAS3.

Table 11: CASPT2/RASPT2 results. All values are ΔE_{ST} (kcal/mol), where a negative number indicates that the singlet is lower.

System	Active space	CASPT2	RASPT2 Valence-π	RASPT2 Limited-π	DEA-EOMCC (4p2h)
1: C_4H_4	eCPO	*a*	-3.8	-3.8	-4.2
	mCPO	-4.3	-4.0	-4.0	
	πCPO	-4.4	-4.4	-4.4	
	nCPO	-12.0	-12.0	-12.0	
2: $C_5H_5^+$	eCPO	*a*	14.5	14.5	13.9
	mCPO	13.5[b]	13.7	13.7	
	πCPO	14.9	15.0	15.0	
	nCPO	21.9	20.5	20.5	
3: $C_4H_3NH_2$	eCPO	*a*	-2.2	-2.2	-2.7
	mCPO	-2.7	-2.8	-2.8	
	πCPO	-2.5	-2.5	-2.5	
	nCPO	8.0	13.5	13.5	
4: C_4H_3CHO	eCPO	*a*	-3.5	-3.5	-3.6
	mCPO	-4.0	-4.0	-4.0	
	πCPO	-3.9	-3.6	-3.6	
	nCPO	8.6	7.7	7.7	
5: C_4H_2 NH_2CHO	eCPO	*a*	-7.2	-7.2	-5.7
	mCPO	-6.3	-6.3	-6.3	
	πCPO	-6.7	-7.4	-7.4	
	nCPO	-3.1	3.6	3.6	
6: C_4H_2-1,2-$(CH_2)_2$	eCPO	*a*	*a*	-75.9	-77.7
	mCPO	*a*	-75.1	-77.0	
	πCPO	-75.5	-75.3	-75.4[b]	
	nCPO	-81.9	-81.9	-81.9	
7: C_4H_2-1,3-$(CH_2)_2$	eCPO	*a*	*a*	18.7	18.5
	mCPO	*a*	18.8	18.4	
	πCPO	18.6	18.4	19.0[b]	
	nCPO	18.3	18.3	18.3	
MUE[c]	eCPO	-	0.6	0.7	Reference
	mCPO	0.3	0.6	0.3	
	πCPO	0.7	0.8	0.9	
	nCPO	6.5	7.9	7.9	

[a]Not available. [b]CI only rather than CASSCF or RASSCF. [c]MUEs exclude absent data.

Table 12: CASSCF/RASSCF active space sizes and numbers of configuration state functions (CSFs), including cases that were too large to attempt. Exact numbers, including RASSCF limited-π, are presented in the supplementary material.

System	Active space	CASSCF Size	CSFs Sing.	CSFs Trip.	RASSCF Valence-π Size	CSFs Sing.	CSFs Trip.
1: C_4H_4	eCPO	(20,22)[a]	$6x10^{10}$[a]	$1x10^{11}$[a]	(20,2,2;8,4,10)	$2x10^5$	$3x10^5$
	mCPO	(12,14)	$2x10^7$	$4x10^7$	(12,2,2;4,4,6)	$2x10^4$	$3x10^4$
	πCPO	(4,6)	105	105	(4,0,2;0,4,2)	96	97
	nCPO	(2,4)	10	6	(2,0,2;0,2,2)	10	6
2: $C_5H_5^+$	eCPO	(24,26)[a]	$1x10^{13}$ [a]	$3x10^{13}$ [a]	(24,2,2;10,5,11)	$9x10^5$	$2x\ 10^6$
	mCPO	(14,16)[b]	$3x10^8$ [b]	$5x10^8$ [b]	(14,2,2;5,5,6)	$8x10^4$	$1x10^5$
	πCPO	(4,6)	105	105	(4,0,2;0,5,1)	105	105
	nCPO	(2,4)	10	6	(2,0,2;0,2,2)	10	6
3: $C_4H_3NH_2$	eCPO	(20,22)[a]	$6x10^{10}$ [a]	$1x10^{11}$ [a]	(20,2,2;8,4,10)	$2x10^5$	$3x10^5$
	mCPO	(12,14)	$2x10^7$	$4x10^7$	(12,2,2;4,4,6)	$2x10^4$	$3x10^4$
	πCPO	(4,6)	105	105	(4,0,2;0,4,2)	96	97
	nCPO	(2,4)	10	6	(2,0,2;0,2,2)	10	6
4: C_4H_3CHO	eCPO	(20,22)[a]	$6x10^{10}$ [a]	$1x10^{11}$ [a]	(20,2,2;8,4,10)	$2x10^5$	$3x10^5$
	mCPO	(12,14)	$2x10^7$	$4x10^7$	(12,2,2;4,4,6)	$2x10^4$	$3x10^4$
	πCPO	(4,6)	105	105	(4,0,2;0,4,2)	96	97
	nCPO	(2,4)	10	6	(2,0,2;0,2,2)	10	6

[a]Not possible due to size. Number of CSFs calculated with Weyl's formula.[47]
[b]CI only rather than CASSCF or RASSCF.
[c]Not attempted due to excessive time required.

Table 12 (cont.): CASSCF/RASSCF active space sizes and numbers of configuration state functions (CSFs), including cases that were too large to attempt. Exact numbers, including RASSCF limited-π, are presented in the supplementary material.

System	Active space	CASSCF			RASSCF Valence-π		
		Size	CSFs Sing.	CSFs Trip.	Size	CSFs Sing.	CSFs Trip.
5: C_4H_2 NH_2CHO	eCPO	(20,22)[a]	$6x10^{10}$ [a]	$1x10^{11}$ [a]	(20,2,2;8,4,10)	$2x10^5$	$3x10^5$
	mCPO	(12,14)	$2x10^7$	$4x10^7$	(12,2,2;4,4,6)	$2x10^4$	$3x10^4$
	πCPO	(4,6)	105	105	(4,0,2;0,4,2)	96	97
	nCPO	(2,4)	10	6	(2,0,2;0,2,2)	10	6
6: C_4H_2-1,2-$(CH_2)_2$	eCPO	(30,32)[a]	$4x10^{16}$ [a]	$9x10^{16}$ [a]	(30,2,2;12,6,14) [c]	$7x10^6$ [c]	$7x10^6$ [c]
	mCPO	(18,20)[a]	$5x10^9$ [a]	$1x10^{10}$ [a]	(18,2,2;6,6,8)	$2x10^5$	$1x10^6$
	πCPO	(6,8)	1176	1512	(6,0,2;0,6,2)	1015	1317
	nCPO	(2,4)	10	6	(2,0,2;0,2,2)	10	6
7: C_4H_2-1,3-$(CH_2)_2$	eCPO	(30,32)[a]	$4x10^{16}$ [a]	$9x10^{16}$ [a]	(30,2,2;12,6,14)[c]	$7x10^6$ [c]	$7x10^6$ [c]
	mCPO	(18,20)[a]	$5x10^9$ [a]	$1x10^{10}$ [a]	(18,2,2;6,6,8)	$2x10^5$	$1x10^6$
	πCPO	(6,8)	1176	1512	(6,0,2;0,6,2)	1015	1317
	nCPO	(2,4)	10	6	(2,0,2;0,2,2)	10	6

[a]Not possible due to size. Number of CSFs calculated with Weyl's formula.[47]
[b]CI only rather than CASSCF or RASSCF.
[c]Not attempted due to excessive time required.

2.2.5 Conclusions

Singlet–triplet gaps in several diradical organic π-systems, including cyclobutadiene and its derivatives and cyclopentadienyl cation, were calculated using the CPO-based CASPT2 and RASPT2 approaches benchmarked against high-level DEA-EOMCC data including up to 4*p*-2*h* excitations. The goal was to develop a systematic way to choose and sub-divide active spaces within the CPO framework and find active spaces that can accurately reproduce the DEA-EOMCC(4*p*-2*h*)-level data, while being small enough to be applicable to larger organic diradicals.

To generate benchmark data for assessing the accuracy of the CPO-based CASPT2 and RASPT2 approaches, we performed a large number of DEA-EOMCC calculations, including full DEA-EOMCC(3*p*-1*h*), active-space DEA-EOMCC(3*p*-1*h*){N_u} computations, in which the high-order 4*p*-2*h* terms are neglected calculations, active-space DEA-EOMCC(3*p*-1*h*,4*p*-2*h*){N_u} and DEA-EOMCC(4*p*-2*h*){N_u} calculations, in which 4*p*-2*h* effects are accounted for, and we developed a useful extrapolation scheme that allowed us to capture 3*p*-1*h* and 4*p*-2*h* correlations and the effect of going from the cc-pVDZ basis set to its larger maug-cc-pVTZ counterpart in an accurate and computationally manageable manner. While generating the benchmark DEA-EOMCC information, we investigated the role of high-order 4*p*-2*h* effects, showing that they can be quite important in obtaining accurate singlet–triplet gaps in diradicals, confirming the earlier findings in this regard.[251–253]

We find that the CPO scheme is quite successful for these systems; eCPO and mCPO are highly accurate, with MUEs of 0.3–0.7 kcal/mol, but would usually be cost-prohibitive for systems of practical interest. At the other end of quality spectrum, nCPO is insufficient, with MUEs of 6.5 and 7.9 kcal/mol. However πCPO has MUEs of 0.7-0.9

kcal/mol, almost as good as mCPO and eCPO, and it is much more affordable and thus shows promise for calculation of π-system excitations in larger systems.

Examination of occupation numbers demonstrated that π orbitals in general, not merely the two nominal participating orbitals, are important contributors to the multireference character, but orbitals outside of the π system are effectively either doubly occupied or unoccupied. These observations explain why eCPO and mCPO do not show improvements in accuracy over πCPO, but nCPO is inaccurate. This observation also explains why methods using only a (2,2) active space have not been able to achieve consistent and accurate results. Based on this data, πCPO is recommended for these sorts of systems, providing significant cost savings over full-valence approaches to selecting active spaces. Even greater savings can be obtained by using RASSCF to further reduce the cost of the CI expansion, especially in light of the fact that the RASPT2 MUEs are only 0.1–0.2 kcal/mol higher than CASPT2 MUEs for πCPO.

2.2.6 Additional Information

See supplementary material for active spaces (including visualized orbitals and occupation numbers) and absolute energies for all CASPT2 and RASPT2 calculations presented, exact numbers of CSFs, including for limited-π RASSCF, and ANO-RCC-VTZP results, including a select comparison with a zero IPEA shift. The supplementary material is available on the AIP Publications website at DOI: 10.1063/1.4998256.

The authors thank Molly Anderson for helpful assistance. This work was supported in part (S.J.S., D.G.T., and L.G.) by the NSF Grant CHE-1464536 for the multireference calculations (CASSCF, CASPT2 and MC-PDFT), and in part (P.P., J. S., and A.O.A) by the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy under award DE-FG02-01ER15228 for the coupled cluster calculations.

3 Spin Energetics in MC-PDFT

3.1 MC-PDFT Can Calculate Singlet-Triplet Splittings of Organic Diradicals

This section describes the outcome of a collaborative research project carried out by Samuel J. Stoneburner (and advised by Donald G. Truhlar and Laura Gagliardi). A report on this research project has been published.[269]

Samuel J. Stoneburner performed all calculations and wrote the manuscript.

Reproduced from **Stoneburner, S. J.**; Truhlar, D. G.; Gagliardi, L. MC-PDFT Can Calculate Singlet-Triplet Splittings of Organic Diradicals. *J. Chem. Phys.* **2018**, *148* , 064108,[269] with the permission of AIP Publishing.

3.1.1 Overview

The singlet–triplet splittings of a set of diradical organic molecules are calculated using multiconfiguration pair-density functional theory (MC-PDFT) and the results are compared with those obtained by Kohn-Sham density functional theory (KS-DFT) and complete active space second-order perturbation theory (CASPT2) calculations. We found that MC-PDFT, even with small and systematically defined active spaces, is competitive in accuracy with CASPT2, and it yields results with greater accuracy and precision than Kohn-Sham DFT with the parent functional. MC-PDFT also avoids the challenges associated with spin contamination in KS-DFT. It is also shown that MC-PDFT is much less computationally expensive than CASPT2 when applied to larger active spaces, and this illustrates the promise of this method for larger diradical organic systems.

3.1.2 Introduction

Our objective here is to validate multiconfiguration pair-density functional theory (MC-PDFT) for predicting the energy splitting between singlet and triplet states of organic diradical molecules. Diradicals have two unpaired electrons in degenerate or nearly degenerate molecular orbitals. For the present study we define the singlet–triplet splitting (ΔE_{ST}) as the vertical transition energy between the lowest-energy triplet state and the lowest-energy singlet state (positive if the triplet is lower in energy, negative if the singlet state is lower in energy). Calculating ΔE_{ST} can be challenging, since the lowest singlet state is inherently multiconfigurational due to the near-degeneracy of the singly occupied molecular orbitals (SOMOs).[221] Inherently multiconfigurational systems usually require careful balancing of dynamic and static correlation energy, and in wave function theory this typically requires either expensive post-self-consistent-field (SCF)

dynamic correlation calculations to be added to a multiconfiguration self-consistent-field (MCSCF) reference function or even more expensive high excitations (for example, quadruples) to be added to a single-configuration SCF reference function.

If one wishes to use a theory with a single-configuration SCF wave function, such as Kohn-Sham (KS) density functional theory (DFT), it is important to understand that open-shell determinants are generally not spin eigenfunctions. As a result, the expectation value of the S^2 operator (where S is total electron spin) is not equal to the correct value (i.e., $S(S+1)$) for a given spin multiplicity;[270,271] this is commonly called "spin contamination." In spin-unrestricted KS-DFT, the expectation value of S^2 is calculated from the determinant,[271] and nonzero amounts of spin contamination are always present.[270] (A review of spin in open-shell KS-DFT calculations is beyond the scope of this work; interested readers may consult the detailed discussion in Ref. [271].) In order to obtain accurate singlet and triplet energies in KS-DFT, one uses broken-spin-symmetry functions[239] in which the SCF determinants are mixtures of singlet and triplet states. For diradicals, spin contamination is much greater than for simple radicals (e.g., the spin contamination is much greater for very stretched H_2, which is a diradical, than for Li atom, which is a simple radical). There have been numerous efforts to "decontaminate" broken-spin-symmetry results, and they have met with some success, but the methods are not completely satisfactory.[239,272–280]

Broken spin-symmetry is not required when the SCF function is a properly symmetrized MCSCF wave function, such as a complete active space self-consistent field (CASSCF)[6] or a restricted active space self-consistent field (RASSCF)[7] wave function. A CASSCF wave function partitions the molecular orbitals into three categories: inactive, active, and secondary. Active orbitals are permitted to take any occupation, while inactive and secondary orbitals are kept doubly occupied and unoccupied, respectively.

Static correlation, i.e., correlation arising from inherently multiconfigurational character, is accounted for by doing full configuration interaction (CI) within the active space. However, dynamic correlation (due to fully representing the Coulomb hole at short electron-electron distances or correlating subsystem multipole moments for separated subsystems) is not fully accounted for in CASSCF, and this necessitates a post-SCF calculation such as second-order perturbation theory (CASPT2).[8,9] Similar remarks apply to RASSCF and restricted active space second-order perturbation theory RASPT2,[10] with the difference being that some less important excitations are left out in the MCSCF step. While CASPT2 and RASPT2 methods provide good accuracy for many problems, they rapidly become unaffordably expensive as active-space size increases. Multiconfiguration pair-density functional theory (MC-PDFT)[11] was created to combine the benefits of multiconfigurational self-consistent field calculations, such as CASSCF or RASSCF, for representing the character and spin symmetry of inherently multiconfigurational states with the speed and cost advantages of density functional theory (DFT) for calculating the dynamic correlation energy. MC-PDFT employs an on-top density functional (which is a functional of the total electron density and the on-top pair density, the probability that two electrons are found at a given point); the kinetic energy and classical electrostatic energy are taken from the MCSCF results, while the electron exchange and additional correlation energy (where the latter includes the difference between the accurate kinetic energy and the MCSCF kinetic energy) are provided by the on-top functional. Because the electron density and on-top densities are taken from the MCSCF wave function, which is a spin eigenfunction, there is no spin contamination.

MC-PDFT has already been shown to be competitive in accuracy with the more expensive CASPT2 method for a variety of cases, including electronic excitation energies of organic molecules and transition-metal bond dissociation energies.[40,78,238,281] Additionally, systematic choices for the active space in the preceding MCSCF calculation

have been explored with the correlated participating orbital (CPO) scheme.[25] CPO has been successfully used in conjunction with MC-PDFT for both transition metal bond dissociation energies[243] and simple divalent diradicals.[242] In a recent paper[206] we described the application of the CPO scheme to anti-aromatic diradical systems to design active spaces for CASSCF/CASPT2 and RASSCF/RASPT2 calculations. Due to the wide variances in other published ΔE_{ST} values, new reference calculations were performed using doubly electron-attached (DEA) equation-of-motion (EOM) coupled-cluster (CC) calculations. The CASPT2 and RASPT2 results compared well with the DEA-EOM-CC reference results for most of the active spaces considered. CASPT2 and RASPT2 had mean unsigned errors (MUEs) of less than 1 kcal/mol, while the multireference coupled-cluster results of Saito et al. had MUEs of over 3 kcal/mol. We now show that similarly accurate results can be obtained using MC-PDFT, but at a much lower cost.

3.1.3 Computational Details

The seven systems considered in this work are shown in Figure 11, namely square cyclobutadiene (**1**), cyclopentadienyl cation (**2**), aminocyclobutadiene (**3**), formylcyclobutadiene (**4**), 1-amino-2-formylcyclobutadiene (**5**), 1,2 basis(methylene)cyclobutadiene (**6**), and 1,3-basis(methylene)cyclobutadiene (**7**). These systems were chosen in our previous paper[206] for the sake of comparison to the multireference coupled-cluster work of Saito et al.[239] To be consistent with the benchmark results to which we compare, geometries were taken from Saito et al.[239] (Their geometries correspond to the triplet for systems **1** to **5**, but for the other two systems it is not stated in their work whether the geometries are for the singlet or the triplet.) The singlet-triplet energy gap, ΔE_{ST}, was calculated as the difference between the lowest singlet state and the lowest triplet state:

$$\Delta E_{ST} = E_{singlet} - E_{triplet} \quad (73)$$

where a positive/negative sign indicates the triplet/singlet is lower.

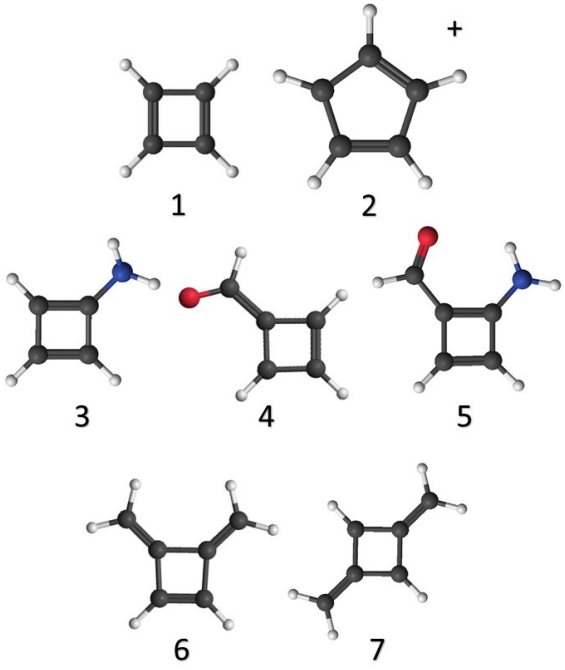


Figure 11: Diradical systems under investigation. **1**: C_4H_4**, 2**: $C_5H_5^+$**, 3**: $C_4H_3NH_2$**, 4**: C_4H_3CHO**, 5**: $C_4H_2NH_2CHO$**, 6**:C_4H_2-1,2-$(CH_2)_2$**, 7**:C_4H_2-1,3-$(CH_2)_2$**.** Reproduced from S. J. Stoneburner, J. Shen, A. O. Ajala, P. Piecuch, D. G. Truhlar, and L. Gagliardi, J. Chem. Phys. 147, 164120 (2017), with the permission of AIP Publishing.[206]

MC-PDFT calculations were performed using the maug-cc-pVTZ[263] and ANO-RCC-VTZP[264] basis sets with Cholesky decomposition[185] using Molcas 8.2.[53] For MC-PDFT, several on-top functionals were used: translated[11] and fully translated[78] versions of the local spin-density approximation LSDA,[23] the Becke-Lee-Yang-Parr (BLYP)[80,81] functional, the Perdew-Berke-Ernzerhof (PBE)[82] functional, and the revised PBE (revPBE)[83] labeled as tLSDA, tBLYP, tPBE, and trevPBE for the translated and ftLSDA, ftBLYP, ftPBE, and ftBLYP for the fully translated functionals.

We used the CPO active spaces developed in our previous work.[206] The full descriptions of these spaces can be found there, but we briefly summarize them here. In a CPO scheme, one first selects a set of primary orbitals, and then one adds a correlating orbital to each primary orbital that is not already correlated. The four levels of the CPO scheme employed here and their corresponding choices of primary orbitals are as follows:

- nominal CPO (nCPO): The primary orbitals are the frontier orbitals, which in all cases except **5** are the two singly occupied molecular orbitals (SOMOs). For system **5**, one of the frontier orbitals is doubly occupied and the other is unoccupied, so the frontier orbitals are the highest occupied molecular orbital (HOMO) and the lowest occupied molecular orbital (LUMO) instead of SOMOs.
- πCPO: The primary orbitals include all π orbitals of all atoms involved in the frontier orbitals.
- moderate CPO (mCPO): The primary orbitals include all p orbitals of all atoms involved in the frontier orbitals.
- extended CPO (eCPO): The primary orbitals include all p and s orbitals of all atoms involved in the frontier orbitals.

We used both CASSCF and RASSCF with the above active spaces. While in CASSCF, all active orbitals are in a single active space, in RASSCF, there are three subspaces: RAS1 contains doubly occupied orbitals, with limited excitations into RAS2 and RAS3 permitted. RAS3 contains unoccupied orbitals, with limited excitations from RAS1 and RAS2 permitted. RAS2 orbitals can have any occupancy, i.e., they are treated with full CI. We employed the same space subdivisions as in our previous work:[206]

- Valence-π RASSCF: All valence π orbitals are in RAS2. All remaining doubly occupied active orbitals are in RAS1, and all remaining unoccupied active orbitals are in RAS3.
- Limited-π RASSCF: Similar to valence-π, except that only the highest two occupied (singly or doubly) π orbitals are in RAS2. Any additional occupied π orbitals are in RAS1, and their correlating orbitals are in RAS3.

In both kinds of RASSCF, up to 2 electrons are permitted to be excited from RAS1, and up to 2 electrons are permitted to be excited into RAS3.

As discussed in our previous work,[206] CI-only calculations (i.e., CI without SCF) were used in three cases where CASSCF or RASSCF was not possible. For systems **6** and **7** with limited-π RASSCF-πCPO, the RASSCF calculations did not converge, so CI-only calculations were performed using orbitals from the valence-π RASSCF-πCPO active space. The CASSCF-mCPO calculation for system **2** was prohibitively expensive, so CI-only calculations were performed using orbitals from the valence-π RASSCF-mCPO active space.

Gaussian 09[282] with the maug-cc-pVTZ basis set was used for KS-DFT calculations using the PBE and PBE0[283] exchange-correlation functionals. All KS-DFT calculations were spin-unrestricted. Two methods were used to calculate ΔE_{ST} from KS-DFT energies: The first method, called the variational method, uses Equation (73) with the variationally lowest-energy solution, and the second method, called the weighted-average broken-symmetry (WABS) method, uses the Yamaguchi formula[239,284–286] for the singlet energy, which results in

$$\Delta E_{ST} = \frac{-2S_{\text{triplet}}(E_{\text{triplet}} - E_{\text{singlet}})}{\langle S^2\rangle_{\text{triplet}} - \langle S^2\rangle_{\text{singlet}}} \tag{74}$$

where $\langle S^2\rangle$ is the expectation value of the square of the total electron spin (calculated from the Slater determinant) and S_{triplet} is calculated by

$$S_{\text{triplet}}\left(S_{\text{triplet}} + 1\right) = \langle S^2\rangle_{\text{triplet}} \tag{75}$$

Note that a similar equation was used in a previous paper,[242] but here we insert a negative sign to conform to the convention used here that ΔE_{ST} is negative when the singlet is lower in energy than the triplet.

3.1.4 Results and Discussion

In all cases, the MUEs of nCPO calculations are much larger than all the others. As shown in our previous work,[206] nCPO-based active spaces are insufficient. While most of the multiconfigurational character is due to variable occupancy of the two frontier orbitals, the second-highest occupied π orbital and second-lowest unoccupied π orbital also lead to significant multiconfigurational character, which necessitates at least a πCPO active space. We also note that functionals based on the simple LSDA approximation perform worse than the functionals based on generalized gradient approximation (GGA) functionals,[37] that is, based on BLYP, PBE, and revPBE. Therefore nCPO results, tLSDA results, and ftLSDA results are relegated to the supplementary material and are not discussed further here. Mean unsigned errors (averaged over the seven molecules) for the other three active spaces with the other six functionals are in Table 13 for the maug-cc-pVTZ basis set. Results with the ANO-RCC-VTZP basis set are similar and are presented in the supplementary material.

Table 13 shows that the fully translated functionals perform noticeably better than their translated counterparts. The functional yielding the smallest MUEs is ftPBE. None

of the MUEs for ftPBE exceed 2.2 kcal/mol, and they vary by only 0.6 kcal/mol, thus demonstrating both accuracy and consistency. The other fully translated functionals that we employed, ftrevPBE and ftBLYP, are also quite good, but the MUEs of ftrevPBE span a larger range than those of ftPBE, and MUEs of ftBLYP are slightly higher than those for ftPBE. For brevity, molecule-by-molecule results will be discussed only for ftPBE, but molecule-by-molecule results for the other functionals are presented in the supplementary material.

Table 13: MC-PDFT mean unsigned errors (MUEs,[a] kcal/mol) for various translated functionals, active spaces, and active space subdivisions with the maug-cc-pVTZ basis set.

Functional	Active space	CASSCF	RASSCF Valence-π	RASSCF Limited-π
tPBE	eCPO	*b*	3.3	2.9
	mCPO	2.8	1.7	1.6
	πCPO	3.4	3.4	3.4
ftPBE	eCPO	*b*	1.9	1.7
	mCPO	2.2	1.8	1.7
	πCPO	2.0	2.0	2.0
trevPBE	eCPO	*b*	3.3	3.1
	mCPO	2.8	1.9	1.8
	πCPO	3.5	3.5	3.5
ftrevPBE	eCPO	*b*	1.5	1.4
	mCPO	2.6	2.2	2.0
	πCPO	1.8	1.8	1.8
tBLYP	eCPO	*b*	3.7	3.2
	mCPO	3.1	1.9	1.8
	πCPO	3.4	3.4	3.4
ftBLYP	eCPO	*b*	2.3	2.1
	mCPO	2.6	2.2	2.0
	πCPO	2.0	2.0	2.0

[a] The MUEs exclude unavailable data, which are cases that we were unable to complete due to computer time or memory requirements. MUEs are in reference to the DEA-EOM-CC reference values published previously.[206]

[b] Unavailable.

Table 14: Singlet-triplet gaps (ΔE_{ST}, kcal/mol) by ftPBE, CASPT2/RASPT2, KS-DFT, and benchmark calculations[a]

System		CAS-		Val-π RAS-		Lim-π RAS-					
		ftPBE	PT2[206]	ftPBE	PT2[206]	ftPBE	PT2[206]		PBE	PBE0	Benchmark[206]
1	e	*b*	*b*	-3.6	-3.8	-3.6	-3.8				-4.2
C_4H_4	m	-5.5[c]	-4.3[c]	-4.4	-4.0	-4.4	-4.0	var.	-1.8	-4.8	
	π	-2.5	-4.4	-2.5	-4.4	-2.5	-4.4	WABS	-3.6	-10.3	
2	e	*b*	*b*	11.3	14.5	11.3	14.5				13.9
$C_5H_5^+$	m	10.6	13.5	10.6	13.7	10.6	13.7	var.	5.0	4.5	
	π	14.8	14.9	14.8	15.0	14.8	15.0	WABS	10.2	9.3	
3	e	*b*	*b*	-2.1	-2.2	-2.1	-2.2				-2.7
$C_4H_3NH_2$	m	-3.3	-2.7	-3.7	-2.8	-3.7	-2.8	var.	-0.8	-3.4	
	π	-0.7	-2.5	-0.7	-2.5	-0.7	-2.5	WABS	-1.6	-7.1	
4	e	*b*	*b*	-2.5	-3.5	-2.5	-3.5				-3.6
C_4H_3CHO	m	-3.8	-4.0	-6.2	-4.0	-6.2	-4.0	var.	-1.4	-4.1	
	π	-2.1	-3.9	-2.1	-3.6	-2.1	-3.6	WABS	-2.9	-8.8	

[a] Val-π denotes Valence-π; Lim-π denoted Limited-π; e, m, and π denote eCPO, mCPO, and πCPO.

[b] Not available.

[c] CI optimization only.

[d] Mean unsigned errors excluding absent data.

Table 14 (cont.): Singlet-triplet gaps (ΔE_{ST}, kcal/mol) by ftPBE, CASPT2/RASPT2, KS-DFT, and benchmark calculations[a]

System		CAS-		Val-π RAS-		Lim-π RAS-			PBE	PBE0	Benchmark[206]
		ftPBE	PT2[206]	ftPBE	PT2[206]	ftPBE	PT2[206]				
5	e	*b*	*b*	-10.0	-7.2	-10.0	-7.2				-5.7
$C_4H_2NH_2CHO$	m	-11.4	-6.3	-9.8	-6.3	-9.8	-6.3	var.	-8.7	-4.9	
	π	-10.5	-6.7	-10.5	-7.4	-10.5	-7.4	WABS	-9.1	-5.1	
6	e	*b*	*b*	*b*	*b*	-78.9	-75.9				-77.7
C_4H_2-1,2-$(CH_2)_2$	m	*b*	*b*	-77.0	-75.1	-77.1	-77.0	var.	-74.3	-73.5	
$(CH_2)_2$	π	-77.6	-75.5	-77.6	-75.3	-77.6[c]	-75.4[c]	WABS	-74.1	-73.1	
7	e	*b*	*b*	*b*	*b*	17.3	18.7				18.5
C_4H_2-1,3-$(CH_2)_2$	m	*b*	*b*	17.5	18.8	18.4	18.4	var.	7.4	12.2	
$(CH_2)_2$	π	15.4	18.6	15.4	18.4	15.4[c]	19.0[c]	WABS	13.9	22.2	
MUE[d]	e	*b*	*b*	1.9	0.6	1.7	0.7				
	m	2.2	0.3	1.8	0.6	1.7	0.3	var.	4.7	3.2	
	π	2.0	0.7	2.0	0.8	2.0	0.9	WABS	2.5	4.2	

[a] Val-π denotes Valence-π; Lim-π denoted Limited-π; e, m, and π denote eCPO, mCPO, and πCPO.

[b] Not available.

[c] CI optimization only.

[d] Mean unsigned errors excluding absent data.

Table 15: Expectation values of the square of the total electron spin $\langle S^2 \rangle$ from ftPBE[a] and PBE.[b]

System	2*S*+1	$\langle S^2 \rangle$ Pure	ftPBE[a] (MC-PDFT)	PBE (KS-DFT)
1: C_4H_4	1	0.00	0.00	1.03
	3	2.00	2.00	2.00
2: $C_5H_5^+$	1	0.00	0.00	1.02
	3	2.00	2.00	2.00
3: $C_4H_3NH_2$	1	0.00	0.00	1.03
	3	2.00	2.00	2.00
4: C_4H_3CHO	1	0.00	0.00	1.03
	3	2.00	2.00	2.01
5: $C_4H_2NH_2CHO$	1	0.00	0.00	0.08
	3	2.00	2.00	2.00
6: C_4H_2-1,2-$(CH_2)_2$	1	0.00	0.00	0.00
	3	2.00	2.00	2.01
7: C_4H_2-1,3-$(CH_2)_2$	1	0.00	0.00	0.95
	3	2.00	2.00	2.03

[a]MC-PDFT $\langle S^2 \rangle$ values are pure by construction. [b]PBE0 results are similar to PBE ones and are presented in the supplementary material.

Calculated singlet-triplet gaps (ΔE_{ST}) for individual molecules are presented in Table 14. CASPT2 and RASPT2 singlet-triplet splittings from our previous work[206] are included for comparison, along with KS-DFT ΔE_{ST} results based on the parental functionals. Expectation values of S^2 for PBE and ftPBE are presented in Table 15. As discussed in the introduction, MC-PDFT spin states are pure by construction because the electron density and on-top density are taken from an MCSCF wave function that is a spin eigenfunction. In the KS-DFT calculations, the calculated S^2 expectation values for the singlet states of **1**, **2**, **3**, **4**, and **7** are approximately halfway between the values corresponding to the pure singlet and the pure triplet; this reflects the spin contamination that is known[239] to be a common problem in KS-DFT for intermediate spin states.

The MUEs for KS-DFT using the variational method are noticeably larger than those of PT2 or ftPBE. While WABS is generally expected to perform better than the variational method,[242] here it is inconsistent: The MUEs for KS-DFT using the WABS method are better than for the variational method with PBE, but are worse than those with PBE0. The differences are much less dramatic for systems **5** and **6**, which have less spin contamination in the singlet states. In all cases, the MUEs of KS-DFT are greater than the MUEs of ftPBE, which does not require or use any broken-spin-symmetry treatment. The maximum errors using the variational method for KS-DFT are 11.0 kcal/mol for PBE (system **7**) and 9.4 kcal/mol for PBE0 (system **2**). The maximum errors for KS-DFT are smaller with the WABS method, with 4.6 and 6.1 kcal/mol for PBE and PBE0, respectively, but they are still quite large. The highest error for ftPBE is 5.8 kcal/mol (Table 14: CASSCF, mCPO, system **5**), but for other than for system **5**, ftPBE's error never exceeds 3.2 kcal/mol, which is not far above PT2's maximum error of 2.6 kcal/mol. Thus, ftPBE is shown to have significantly greater reliability than the Kohn-Sham methods for the cases studied, while also entirely avoiding all complications arising from spin contamination or broken spin symmetry. Additionally, ftPBE performs competitively with CASPT2 and RASPT2. With one exception (Table 14: CASSCF, mCPO), ftPBE MUEs are within 1.4 kcal/mol of the MUEs of CASPT2 and RASPT2.

In comparing MC-PDFT to CASPT2 and RASPT2, computational costs should also be considered. Timings for all active spaces are presented in the supplementary material. For the largest active space, which is eCPO with limited-π RASPT2 for triplet **6**, the post-SCF perturbation theory calculations required several days to complete on a single processor, while the post-SCF ftPBE step required less than a half hour on a single processor, with a savings of a factor of 410. Note that the post-SCF steps of CASPT2, RASPT2, and MC-PDFT are all preceded by MCSCF calculations, which take considerable time for large active spaces. The total time (SCF plus post-SCF) for the

longest calculation was 217 hours for RASPT2, but only 13 hours for MC-PDFT, a savings of a factor of 17. MC-PDFT also has the advantage of requiring far less memory, as discussed in previous work.[242]

3.1.5 Conclusions

Our calculations of singlet-triplet splittings for several organic diradical systems demonstrate that MC-PDFT is effective for these systems. When compared to benchmark calculations from our previous paper,[206] the ftPBE on-top functional shows better consistency and accuracy than PBE, the KS-DFT parent functional, and it eliminates the complexities of spin contamination and broken-spin-symmetry in KS-DFT. The accuracy of ftPBE is competitive with CASPT2 and RASPT2, but at a much lower computational cost. As active spaces become very large, CASPT2 and RASPT2 take much more time than the preceding CASSCF and RASSCF calculations. In contrast, MC-PDFT time requirements are small in proportion to the CASSCF and RASSCF requirements, and this shows that MC-PDFT can be used for much larger active spaces than CASPT2 or RASPT2. The quality of results using the small πCPO active space scheme is particularly encouraging, as this suggests MC-PDFT could be used on similar, but larger, systems requiring much larger active spaces which would be unaffordable with CASPT2 or RASPT2.

3.1.6 Additional Information

See supplementary material for hardware information, sample input files, ANO-RCC-VTZP results, nCPO results, tLSDA and ftLSDA results, PBE0 spin expectation values, and all absolute energies, active space sizes, numbers of CSFs, and timings for all functionals and active spaces. The supplementary material is available on the AIP Publications website at DOI: 10.1063/1.5017132.

This work was supported in part by AFOSR grant FA9550-16-1-0134.

3.2 Full Correlation in a Multiconfigurational Study of Bimetallic Clusters: Restricted Active Space Pair-Density Functional Theory Study of [2Fe-2S] Systems

This section describes the outcome of a collaborative research project carried out by Davide Presti and Samuel J. Stoneburner (and advised by Donald G. Truhlar and Laura Gagliardi). A report on this research project has been published.[287]

Samuel J. Stoneburner performed many of the calculations and contributed to the manuscript, working closely with Davide Presti.

3.2.1 Overview

Iron-sulfur clusters play a variety of important roles in protein chemistry, and understanding the energetics of their spin ladders is an important part of understanding these roles. Computational modeling can offer considerable insight into such problems; however, calculations performed thus far on systems with multiple transition metals have typically either been restricted to a single-configuration representation of the density, as in Kohn-Sham theory, or been limited to correlating excitations only within an active space, as in active-space self-consistent field methods. For greater reliability, a calculation should include full correlation, i.e., not only correlation internal to the active space but also external correlation, and it is desirable to combine this full electron correlation with a multiconfigurational representation of the wave function; but this has been impractical thus far. Here we present an affordable way to do that by using restricted-active-space pair-density functional theory. We show that with this method it is possible to compute the entire spin ladder for systems containing two Fe centers bridged by two S atoms. On the other hand, with second-order perturbation theory only the high-spin states can be computed. A key result is that, in agreement with some experiments, we find a high-spin ground state for a relaxed reduced $[Fe_2S_2(SCH_3)_4]^{3-}$ cluster, which is a novel result in computational studies.

3.2.2 Introduction

Iron-sulfur (Fe-S) compounds are widely studied because of their key role in biochemistry;[288] Fe-S clusters are the prosthetic groups in many metalloproteins, such as ferredoxins, hydrogenase, NADH-dehydrogenase, and nitrogenase,[288–292] and they have significant roles involving electron transport in the metabolic pathways for both prokaryotic and eukaryotic cells.[293,294] Furthermore, Fe-S clusters are protagonists of the so-called "iron-sulfur world";[295–298] prebiotic Fe-S clusters containing organic linkers (such as glutathione[299] and other analogs, mostly thiolates[299,300]) have been demonstrated to be able to form in extreme conditions, and their formation has been shown to benefit from the presence of the UV-light radiation (i.e. photo oxidizing/photolytic environment) provided by the "young Sun".[299] This may have led to the formation of more complex Fe-S-based prebiotic molecules capable of organizing themselves into iron-sulfur cluster-peptide complexes, enzymes, and proteins.

In the present work, we consider a particular class of Fe-S clusters: those containing two Fe centers bridged by two S atoms (referred to as [2Fe-2S] for convenience); these clusters present an intermediate situation between the single-center Fe clusters, from which they are formed, and the three- and four-center Fe clusters which, together with [2Fe-2S], are involved in the formation of iron-sulfur cluster-peptide complexes. Their UV-light-driven synthesis, mentioned above, was elegantly described in a recent experimental paper.[299] Experimentally [2Fe-2S] clusters have been characterized extensively with UV-Visible optical absorption spectra, Mössbauer spectra, and EPR spectroscopy.[291,294,299–302] However, a direct assignment of low-lying electronic states is still unattainable in larger systems.[303] Theoretical studies have contributed significantly to the understanding of their electronic structure,[16–26] but as pointed out in the work of Sharma et al.,[303] the theoretical models adopted until recently are not

straightforward and have given conflicting answers to key questions. Sharma et al.[303] pointed out the possibility that one could obtain more definitive results by using *ab initio* calculations free of the assumptions made in a phenomenological model, in particular avoiding use of the Heisenberg double-exchange model.

Complexes containing transition metals, especially multiple transition metal atoms, frequently have intrinsically multiconfigurational character, i.e., they cannot be well described by a single configuration state function (CSF). This characteristic is often called "static correlation" or "strong correlation", and the work of Sharma et al.[303] featured a thorough treatment of the static correlation of iron-sulfur systems. However, in order to make comparisons with experiment, it is also important to treat dynamic correlation at a high level. Dynamic correlation represents a relatively small part of the total electronic energy, but it is often of utmost importance in quantitatively determining relative energies such as the energy ordering of different electronic states.

Kohn-Sham density functional theory (KS-DFT) in principle includes the full correlation energy (static and dynamic) without separating it; however, the available density functionals treat the correlation energy better in systems well described by a single configuration state function, probably because KS-DFT is based on a single-configuration representation of the electron density. Nevertheless it is often the best theory available and is therefore widely used for these systems.[304,306–311] Recent applications using extended broken-symmetry (EBS) KS-DFT[314] and variational Monte-Carlo[315] were used to obtain a reasonable ground-state relaxed geometry for these compounds.

Multiconfiguration wave function methods have also been employed,[309,310] and they represent an attempt to overcome the limitation of single-CSF methods. Examples of multiconfiguration wave function methods include multiconfiguration self-consistent-

field (MCSCF) methods such as the complete-active-space self-consistent-field method (CASSCF),[6] multireference configuration-interaction (MRCI) based on an MCSCF or CASSCF reference function,[155] and multireference second order perturbation theory (MRPT2) such as complete active space perturbation theory[8–10] (CASPT2). CASSCF involves full configuration interaction (full CI) in an active space, which can provide only a qualitatively correct description of the electronic wave function. The included correlation, that which is representable internal to the active space, consists of the static correlation and a small portion of the dynamic correlation; the remaining dynamic correlation, involving excitations into the virtual space, is the external correlation; CASSCF lacks most of the dynamic correlation because excitations to generate the electronic configurations are limited to the active space, which has a limited number of orbitals. The external correlation is usually treated by a post-SCF method, using the SCF orbitals to build a reference wave function. When the reference wave function is multiconfigurational, the method is called a multireference method. MRCI treats the full electron correlation when including sufficiently high excitations into the virtual space, but that is prohibitively expensive in most cases, so excitations are usually limited to single and double excitations (MRCISD). CASPT2 is also limited to double excitations in the virtual space, and although it is less expensive than MRCISD, it is also usually too expensive for large or complex systems.

In a recent work,[303,313] state-of-the-art methods for the treatment of large active spaces, such as density-matrix renormalization group (DMRG) as a solver for complete active-space configuration interaction (CASCI), were employed by Sharma et al.[303] to explore the spin-ladder of the low-energy spectrum of [2Fe-2S] and [4Fe-4S] clusters; this allowed the use of larger active spaces than are practical for conventional CASSCF. But even with the largest practical active spaces, a large fraction of the dynamic correlation energy is missing in CASSCF.

While there have been many theoretical studies, to the best of our knowledge no one has yet treated these iron-sulfur clusters with conventional post-MCSCF dynamic correlation treatments such as MRCI or CASPT2. The lack of such applications to [2Fe-2S] clusters is due to their high or impractical cost. While the time required for an MRPT2 calculation can be reduced with the use of a parallelized code, the amount of memory becomes prohibitive for calculations involving the number of CSFs that would typically be required to calculate low-spin states involving transition metal compounds.

The objective of the present work is to combine the advantages of a multiconfigurational representation of the density with an affordable treatment of the full correlation energy. We do this by using multiconfiguration pair-density functional theory (MC-PDFT),[11] which evaluates an on-top density functional using the density and on-top pair density of a multiconfigurational wave function. Unlike the energy calculation in MRPT2, where correlation energy is separated into an internal part treated by MCSCF and an external part involving a perturbative treatment of excitations into the virtual space, with MC-PDFT one computes the unpartitioned (i.e., full) electron correlation energy in a single step by using the multiconfigurational kinetic energy, density, and on-top pair density and an on-top density functional. Only kinetic energy and classical Coulomb contributions are taken from the MCSCF wave function; the MCSCF energy is not used. Therefore there is no double counting of the correlation energy contribution to the electron-electron repulsion.[11,75,78,316,317]

In the present work we calculated the low-energy spin ladder for both oxidized and reduced [2Fe-2S] clusters and compared the results to literature data. The computational capabilities of MC-PDFT are illustrated for large active space with 22 electrons in 26 orbitals, giving rise, for example, to 6.3×10^6 CSFs for the triplet spin configuration. We achieve this large number of CSFs by using a recently developed

protocol, RASCI-PDFT,[318] based on the restricted active space self-consistent-field (RASSCF)[7]/configuration interaction (RASCI)[143] method. For comparison, RASCI-PT2[202] energies were computed whenever possible, but they were unaffordable except for a few cases; this shows that RASCI-PDFT can handle problems larger than those feasible with previously available post-MCSCF methods.

3.2.3 Computational Methods

All calculations were performed for gas-phase species using a parallel build of Molcas 8.2.[53] Symmetry constraints were not imposed. The ANO-RCC-VTZP large basis-set[264] was used for iron and sulfur atoms, and the ANO-RCC-MB minimal basis-set[264] was used for C and H atoms of the ligands. The Douglas-Kroll-Hess Hamiltonian[319–321] was employed throughout to account for scalar relativistic effects.

In the wave function calculations, external correlation was included perturbatively with the zero-order wave function obtained by RASSCF[7] and RASCI[143] (see Sections 3.2.3.2.1 and 3.2.3.2.2 for further details) by using second-order perturbation theory (to be indicated as RASCI-PT2). RASSCF calculations involve simultaneous optimization of both orbitals and configuration interaction coefficients. In the present RASCI calculations, the orbitals optimized by RASSCF for the high-spin configuration are used for restricted-active-space configuration interaction calculations for the whole spin-ladder. The active spaces employed are described in detail in the next section. RASSCF and RASCI excitation energies were computed with the state-average[322] (SA) approach. RASCI-PT2 calculations were carried out as single-state calculations (based on state-averaged orbitals) with an imaginary shift[68] of 0.1 hartrees and with the default value (0.25 hartrees) for the IPEA shift.[65] For spin multiplicity M (which equals $2S + 1$, where S is the total electron spin) lower than 8, memory allocation for MRPT2 calculations

exceeded the maximum available for our hardware, thus rendering the calculations unfeasible.

The density functional calculations are carried out with pair-density functional theory[11,241] based on RASCI wave functions and therefore are called RASCI-PDFT. The RASCI-PDFT calculations used the translated PBE (tPBE) on-top functional.[11]

Orbitals are depicted with a surface isodensity value of 0.4 a.u. Plots and figures were produced with Gnuplot 5.0, Luscus,[266] and Vesta.[323]

3.2.3.1 Cluster Models

We write superscripts $\pm q$ for chemical oxidation states, and superscripts $q\pm$ for physical charges. The central fragment of the compounds investigated here are the oxidized cluster $[2Fe\text{-}2S]^{2+}$, where both metal centers are formally in the oxidation state Fe^{+3}, and the reduced cluster is $[2Fe\text{-}2S]^{1+}$, where one center is Fe^{+3}, and the other one is Fe^{+2}. The oxidized cluster has ten unpaired electrons in the high-spin (HS) configuration, and the reduced cluster has nine. Since [2Fe-2S] clusters have four thiolate linkers, the full systems under study (shown in Figure 12) are denoted as $[Fe_2S_2(SCH_3)_4]^{2-/3-}$. The unrelaxed structures from Sharma et al.[303] were used for both the oxidized and reduced forms. For the reduced cluster, $Fe_2S_2(SCH_3)_4]^{3-}$, the slightly asymmetric relaxed structure, obtained by broken-symmetry KS-DFT and reported in Ref. [303], was also considered.

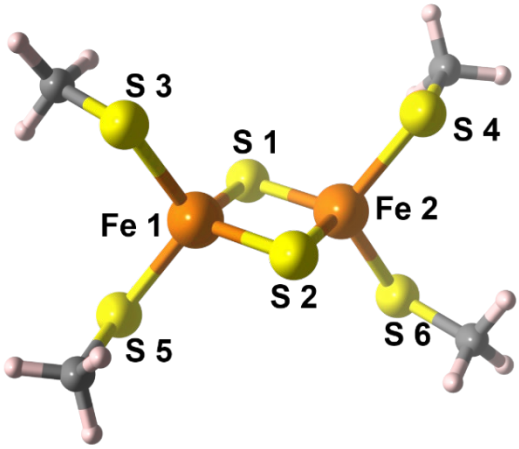


Figure 12: Unrelaxed structure and atom labeling of the $[Fe_2S_2(SCH_3)_4]^{2-/3-}$ cluster as taken from Ref. [303] (derived, in turn, from the experimental structure of Ref. [324])

3.2.3.2 Active space determination

3.2.3.2.1 RASSCF

The active spaces used in the present study are defined based on the ones reported in Ref. [303]. The high-spin (HS) spin state of the $[Fe_2S_2(SCH_3)_4]^{2-}$ oxidized cluster (M = 11) was used to generate the active spaces used for both oxidized $[Fe_2S_2(SCH_3)_4]^{2-}$ and reduced $[Fe_2S_2(SCH_3)_4]^{3-}$ clusters. See the SI for details of the protocol. While it is possible that the use of orbitals optimized only for high-spin when calculating the full spin ladder could introduce some bias favoring higher spin states, we note that the DMRG calculations in Ref. [303] also used orbitals optimized for a high-spin state. After a comparison with results from orbitals optimized for each spin state, the authors concluded that the effect on relative spin energetics was minimal and performed CI-only calculations with the high-spin orbitals thereafter.

The active spaces in RASSCF and RASCI calculations are labeled as (n,h,e;o_1,o_2,o_3), with n being the total number of electrons considered in the RAS1, RAS2, and RAS3 subspaces, h being the maximum number of holes allowed in RAS1, e

being the maximum number of electrons allowed to be excited into RAS3, and o_1, o_2, and o_3 being the number of orbitals in RAS1, RAS2, and RAS3, respectively. The active spaces we use are explained in Table 16. The orbitals in the active spaces for the oxidized case are depicted in Figure S2, and those for the reduced case are in Figures S3 and S4. (Figures and tables with a prefix S are in Supporting Information.)

Table 16. Active spaces employed in this work[a]

Cluster	Name	*n,e,h;* o_1,o_2,o_3	Description of orbitals
either	all		RAS1: bridging sulfur 3*p* orbitals
			RAS2: all 3*d* orbitals
oxidized	(22,26)	22,1,1; 6,10,10	RAS3: all 3*d*′ [b] orbitals (correlating orbitals)
	(22,21)	22,1,1; 6,10,5	RAS3: five 3*d*′ [b] orbitals with highest occupation numbers
reduced	(23,26)	23,1,1; 6,10,10	same as (22,26)
	(23,22)	23,1,1; 6,10,6	RAS3: six 3*d*′ [b] orbitals with highest occupation numbers (unrelaxed geometry)
	(23,21)	23,1,1; 6,10,5	RAS3: five 3*d*′ [b] orbitals with highest occupation numbers (relaxed geometry)

[a]In all cases, these orbitals were obtained by a state-averaged RASSCF calculation on the lowest ten roots of the high-spin cluster (multiplicity $M = 11$ for the oxidized; multiplicity $M = 10$ for the reduced). In the reduced case, the orbitals of the oxidized case are used as starting orbitals for the SCF iterations.

[b]The orbitals labeled 3*d*′ are a second subshell of d orbitals that contribute to the correlation of the first subshell of d orbitals; they are sometimes referred to as "4*d*" or "4*d*-like".

3.2.3.2.2 RASCI

Having obtained the orbitals by RASSCF calculations on the high-spin states, we performed RASCI calculations for the full spin ladder of each cluster; these calculations consisted of reoptimizing only the CI coefficients for every possible spin multiplicity using the orbitals optimized for the high-spin states of the cluster. The RASCI wave

functions were also used as references for further calculations by MC-PDFT and MRPT2, and these calculations are labeled as RASCI-PDFT and RASCI-PT2, respectively. The RASCI energies without either MC-PDFT or MRPT2 are labeled simply "RASCI".

Since the number of CSFs grows rapidly with active space size (see Table 17), and since the RAS2 subspace already has 10 electrons in 10 orbitals, double excitations between RAS subspaces would have resulted in unaffordably large calculations for all spin multiplicities other than the highest-spin configurations ($M = 11$ for the oxidized cluster and $M = 10$ for the reduced cluster). We gauged the error in RASCI excitation energies due to the use of single excitations between subspaces by also considering two holes in RAS1 for the high-spin configuration, while keeping only single excitations to RAS3 to limit the number of CSFs. We report in Table S1 the difference between the excitation energies obtained for the unrelaxed oxidized and reduced clusters with two holes or one hole allowed in RAS1. The mean differences are in the range -0.03 to 0.19 eV with the chosen active spaces.

Table 17. Number of configuration state functions (CSFs) for each spin configuration with 26 orbitals.

oxidized		reduced	
M	CSFs	*M*	CSFs
11	6,221	10	28,880
9	134,739	8	378,280
7	1,067,605	6	2,012,780
5	3,790,875	4	4,950,990
3	6,260,760†	2	5,241,060
1	3,409,164		

† This active space for the triplet state has 2.25 times more CSFs than the generalized active space (GAS)-PDFT calculation for triplet dodecacene that included 50 electrons in 50 orbitals.[238]

Many of the *3p* orbitals of bridging sulfurs are mixed with S–CH_3 σ-type bonding orbitals, especially in the case of the reduced cluster, and this makes the active spaces including 26 orbitals comparable to the large one with 32 orbitals employed in Ref. [303]. The main remaining difference is that in this work the 4*s* orbitals of Fe were excluded because of their being highly mixed with 4*d* and 5*d* orbitals, which would have required several additional correlating orbitals in the active space in order to maintain balance. Furthermore, here the large ANO-RCC-VTZP basis-set is used (with ANO-RCC-MB for C and H), whereas the orbitals of Ref. [303]were obtained by a small basis set in BP86/def2-SV(P) calculations. Note that both our orbitals and those in Ref. [303] were optimized for a high-spin state, as stated in Section 3.2.3.2.1.

3.2.4 Results and Discussion

3.2.4.1 $[Fe_2S_2(SCH_3)_4]^{2-}$ oxidized cluster

The oxidized cluster has two centers that are formally Fe^{+3} centers; thus each of these centers has five unpaired electrons. The centers are antiferromagnetically (AFM) coupled[301] in the ground state of the oxidized cluster. The AFM state is an overall singlet, whereas the ferromagnetically (FM) coupled undectet is the highest spin state. We remind the reader that the orbitals are optimized with the FM configuration due to the significantly smaller number of configurations (see Table 17) in that spin state, and that the authors of Ref. [303] found little benefit in optimizing the orbitals for each spin state separately. Figure 13a shows the RASCI spin ladders for vertical excitation energies using both the (22,26) and the (22,21) active spaces. Energies are reported in the text in eV units for easier readability, whereas they are plotted in figures in cm^{-1} for easier comparison with Ref. [303]. Note that 1 eV = 8065.5 cm^{-1}.

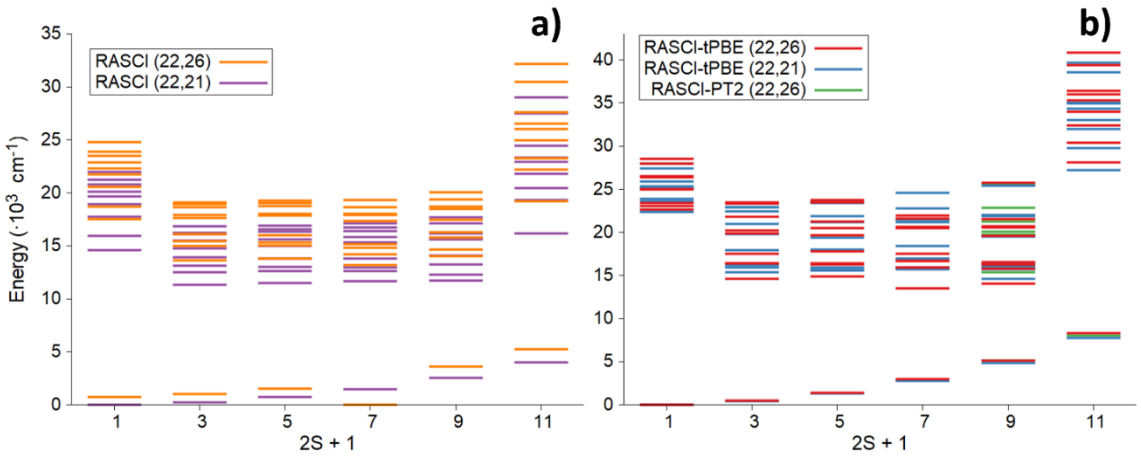


Figure 13: Vertical-excitation spin ladders. (a) RASCI (ten roots for each spin configuration) computed with two active spaces: (22,26) and (22,21). (b) RASCI-PDFT and RASCI-PT2 (nonet and undectet ground-state only). The singlet ground state is taken as the zero of energy in this figure, and for a visual comparison of excitations energies, the energy of the RASCI-PT2 nonet ground state is shifted to the energy of the RASCI-PDFT nonet ground state. In the figure, RASCI-PDFT is labeled RASCI-tPBE.

There is no tabulated data in Ref. [303] for the oxidized clusters, but Figure 2 therein depicts the ground-state energy levels for each spin state of the oxidized cluster. The DMRG-CASCI values in Ref. [303] span a larger range (approximately 8,000 cm^{-1}) than present RASCI calculations (about 5,000 cm^{-1} in Figure 13a), but it is difficult to say which results are more accurate. While the DMRG-CASCI calculation of Ref. [303] includes more orbitals in the active space (specifically, the 4*s* subshells of the Fe atoms) and a more complete treatment of the active space, a much larger basis set is adopted for the present work. Note that both the DMRG-CASCI and our RASCI calculations used orbitals optimized for the high-spin configuration, and therefore the differences in results cannot be explained by a possible high-spin bias in the orbitals. Figure 13a of the present work shows a sizable gap between the ground and the first excited state for all spin states, while the higher excited states form a denser manifold. We will see below that this trend is also found in post-RASCI energies computed by RASCI-PDFT and RASCI-PT2. Figs. 2a and 2b have slightly different wavenumber scales, but comparison of numerical

relative energies shows that RASCI-PDFT has a mildly wider spread of excitation energies than does RASCI.

RASCI (22,21) energies give a trend that is comparable with RASCI (22,26), but while RASCI (22,26) incorrectly predicts the lowest energy spin state to be a septet, RASCI (22,21) correctly orders the different spin states increasing from singlet to undectet, as has been inferred from experiment. Although the RASCI calculation without external correlation does not obtain correct spin state order with the largest active space (i.e. including some but not all 3*d'* orbitals), Figure 13b shows encouragingly that RASCI-PDFT does yield the correct spin state ordering for the largest active space. Furthermore, for each spin state, the lowest states for each of the spins are almost superimposed in the RASCI-PDFT calculations, showing that RASCI-PDFT is less sensitive to active space choice than is RASCI.

For comparison, single-state RASCI-PT2 are reported only for the states for which there are converged results (the undectet ground state and the nonet ground and excited states), i.e., $M = 9$ and $M = 11$.

3.2.4.2 $[Fe_2S_2(SCH_3)_4]^{3-}$ reduced cluster

Calculations for the reduced cluster are reported at two geometries: the unrelaxed one (discussed in Section 3.2.4.2.1), which is the same slightly asymmetric structure as was used for the oxidized cluster, and the relaxed one (discussed in Section 3.2.4.2.2), which is most asymmetric. Both structures are from Ref. [303]. Table 18 summarizes the vertical excitation energies of the lowest five electronic states (which are supposed to be more relevant experimentally at room temperature) for each spin configuration by RASCI, and compared to the DMRG-CASCI results of Sharma et al.[303]

The work of Sharma et al. involved using the DMRG solver for CASCI with very large active spaces. Since CASSCF, even with a large active space, does not include most of the dynamic correlation, it is comparable to our RASCI in that neither calculation includes external correlation, whereas our RASCI-PDFT calculation includes dynamic correlation with no active space limitation. However, the DMRG-CASCI provides an approximation to full CI within the active space, whereas our calculations involve restricted CI as explained in Section 3.2.3.2.1. Conversely, the basis set used for DMRG-CASCI was unusually small [def2-SV(P)] for a wave function calculation, whereas we use an extended basis set. Taking all of these factors into account, we cannot say definitively which set of results is more accurate, because we lack an external reference by which to judge between them, and the different approaches each have advantages and disadvantages in terms of the severity of their approximations.

3.2.4.2.1 Reduced cluster with unrelaxed geometry

The present RASCI calculations with the (23,22) active space on the unrelaxed geometry are not expected to capture most of the dynamic correlation due to the limitations of the active space, i.e., the number of orbitals as well as the restriction to having only single excitations from RAS1 and to RAS3.

Table 18 RASCI vertical excitation energies (eV) for the lowest five electronic states with respect to the DMRG-CASCI results of Sharma et al.[303] for the unrelaxed geometry of the reduced cluster. The lowest energies in each case are used as the zeroes of energy.

	RASCI (23,1,1;6,10,10)					RASCI (23,1,1;6,10,6)				
2S+1	State 1	State 2	State 3	State 4	State 5	State 1	State 2	State 3	State 4	State 5
2	0.07	0.24	0.40	0.47	0.53	1.27	1.46	1.74	1.96	2.16
4	0.04	0.31	0.35	0.46	0.62	1.24	1.62	1.67	2.12	2.13
6	0.07	0.21	0.34	0.64	0.72	1.24	1.62	1.83	2.15	2.29
8	0.09	0.17	0.37	0.87	0.95	1.30	1.60	2.08	2.23	2.34
10	0.00	0.34	0.50	1.14	1.24	0.00	0.31	0.33	0.98	1.13
	DMRG-CASCI, Sharma et al.[303]									
2S+1	State 1	State 2	State 3	State 4	State 5					
2	0.00	0.04	0.14	0.33	0.53					
4	0.02	0.07	0.21	0.53	0.55					
6	0.04	0.08	0.36	0.58	0.66					
8	0.08	0.16	0.58	0.67	0.79					
10	0.13	0.36	0.91	0.92	1.05					

Table 18 and Figure 14a and c show that the larger active space (23,26) yields results more similar to the DMRG-CASCI results than the smaller (23,22) active space in terms of vertical excitation energies, although the differences are still sizable. These differences cannot be simply explained with the choice of employing the orbitals optimized only for the dectet spin state, as the DMRG-CASCI used high-spin orbitals as well.[303] Although the RASCI calculations present a different energetic ordering of the ground state of each spin state with respect to DMRG-CASCI, the ordering at the RASCI-tPBE level with the (23,26) active space is similar to the DMRG-CASCI results (see Table S3 and Figure 14d).

The (23,26) active space captures additional dynamic correlation even at the RASCI level due to the inclusion of all the correlating 3*d′* orbitals of iron, which have

previously been found to be important in first-transition-row metals with more than half-filled 3*d* subshells.[10,325,326] Also, the RASCI excitation energies (Figure 14a and c) and their spacings, i.e., both the global energy span and the spacings within each multiplet manifold, are more similar to those computed by DMRG-CASCI than was the case for the (23,22) active space.

When approximating the full correlation energy by RASCI-PDFT (Figure 14d), the situation does not change considerably. The finding that there are only small differences between the (23,26) active space and the active space of Sharma et al. indicates that the inclusion of the 4*s* orbitals (part of the so-called 'double-shell' effect, together with 3*d'* orbitals) in the active space is unneeded, as already reported in similar cases.[325,326]

The cluster is predicted by RASCI-PDFT to have a quartet ground state (Figure 14d) that is only 0.01 eV more stable than the doublet (see Table S3). As seen in the next subsection, this occurs for the relaxed structure as well. Note that we choose the doublet as the reference energy (zero of energy) in Table 18 and in Figure 14 (excepting Figure 14b), but the 0.01 eV is a small difference and below the expected precision of the method, thus we cannot make a reliable claim regarding whether the doublet or the quartet is the true ground state.

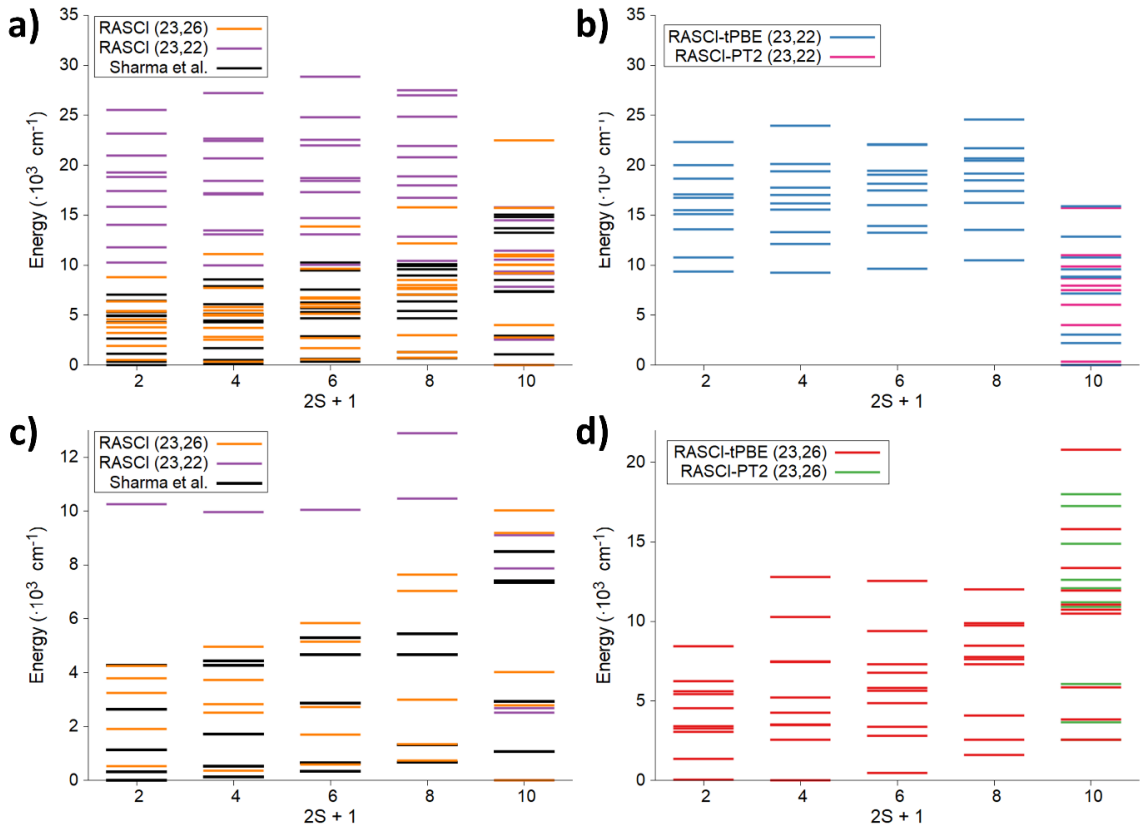


Figure 14: (a) RASCI spin ladder for the unrelaxed $[Fe_2S_2(SCH_3)_4]^{3-}$ cluster, computed with the two active spaces. The DMRG-CASCI results of Sharma et al.[303] are also shown. (b) RASCI-PDFT spin ladder and RASCI-PT2 dectet results with the (23,22) active space. Panel (c) shows a closer view than (a) on the lowest five states for the RASCI (23,26) and the DMRG-CASCI data. (d) RASCI-PDFT spin ladder and RASCI-PT2 dectet results with the (23,26) active space. The doublet ground state is taken as zero. As in Figure 2, the RASPT2 energies of panel (d) are shifted to the RASCI-tPBE dectet ground state. In the figure, RASCI-PDFT is labeled RASCI-tPBE.

3.2.4.2.2 Reduced cluster with relaxed geometry

Table 19 reports the lowest vertical excitation energies of the relaxed reduced cluster (full data is available in Table S4). In contrast to the results for the unrelaxed cluster, the (23,21) active space RASCI results are much closer to those of the (23,26) active space and to those obtained by DMRG-CASCI for low-energy states, although there are very large differences for higher-energy states (see Table S4 and Figure S5).

There is also one fewer 3*d′* orbital in the (23,21) active space than in the corresponding unrelaxed case (23,22) due to its very small occupation (0.01). However, the inclusion of all correlating *3d′* orbitals in the (23,26) active space makes a remarkable improvement in the description of the internal correlation contribution to higher excitation energies. Energy levels corresponding to Table 19 are shown in Figure S5b.

Table 19. RASCI vertical excitation energies (eV) for the lowest five electronic states with respect to the DMRG-CASCI results of Sharma et al.[303] for the relaxed geometry of the reduced cluster. The lowest energies in each case are set to zero.

	RASCI (23,1,1;6,10,10)					RASCI (23,1,1;6,10,5)				
2S+1	State 1	State 2	State 3	State 4	State 5	State 1	State 2	State 3	State 4	State 5
2	0.04	0.15	0.38	0.42	0.50	0.00	0.12	0.35	0.38	0.46
4	0.05	0.15	0.37	0.42	0.51	0.01	0.13	0.36	0.39	0.47
6	0.04	0.15	0.37	0.41	0.55	0.03	0.15	0.37	0.42	0.50
8	0.02	0.18	0.38	0.44	0.61	0.04	0.19	0.40	0.46	0.56
10	0.00	0.23	0.40	0.50	0.69	0.04	0.24	0.46	0.51	0.64
	DMRG-CASCI, Sharma et al.[303]									
2S+1	State 1	State 2	State 3	State 4	State 5					
2	0.00	0.15	0.26	0.47	0.50					
4	0.03	0.15	0.29	0.52	0.56					
6	0.08	0.16	0.35	0.57	0.61					
8	0.10	0.17	0.42	0.64	0.67					
10	0.12	0.23	0.53	0.74	0.80					

For the relaxed reduced cluster, for which the relative energies are depicted in Figure 15a and b, the most notable difference from the unrelaxed cluster is that RASCI-PDFT predicts the ground state of the spin ladder to be the dectet (ferromagnetic) spin configuration in both active spaces. This is also in contrast with previous DMRG-CASCI calculations[303] for both unrelaxed and relaxed structures. This dectet ground state is

compatible with some experimental observations of a high-spin ground state in valence-delocalized [2Fe-2S] reduced clusters,[327–329] where the two iron centers are $Fe^{+2.5}$. The reversed stability is not evident at the RASCI level (see Figure S5 as compared to Figure 15), which suggests that a beyond-MCSCF treatment (i.e., inclusion of external correlation energy) is necessary. Such a treatment was absent in the DMRG-CASCI calculations of Ref. [303]. The experimental cases very often feature a low-spin ground state, but more generally there may be a strong dependence on geometry (see also Section 3.2.4.3). Thus, no general statements can be made about the ground spin of all [2Fe-2S] clusters. Regardless, our results give evidence to the potentially dramatic effects of including (or, conversely, ignoring) dynamic correlation.

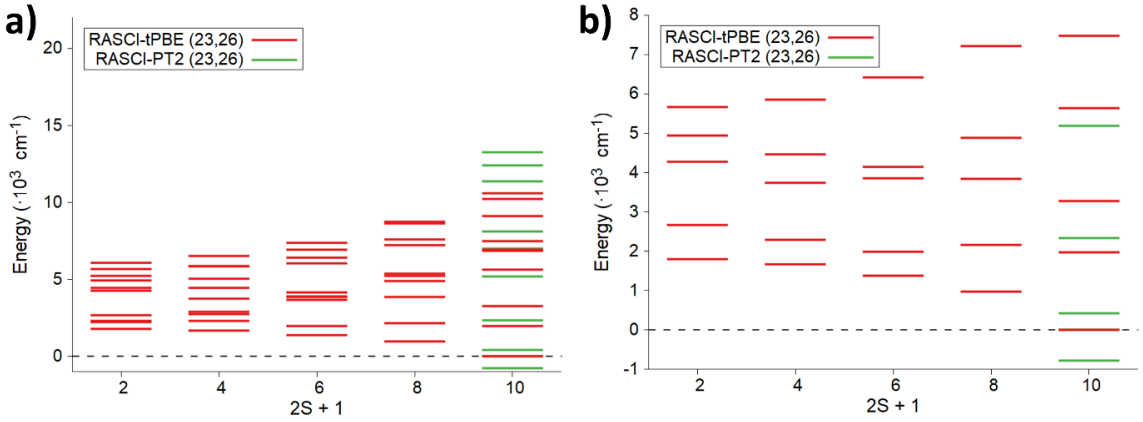


Figure 15: (a) RASCI-PDFT spin ladder for the relaxed $[Fe_2S_2(SCH_3)_4]^{3-}$ cluster, computed with the large active space. The RASCI-PT2 dectet results are also shown. The RASCI-PT2 energies are shifted to the RASCI-PDFT dectet ground state, which is taken as the zero of energy. In the figure, RASCI-PDFT is labeled RASCI-tPBE. (b) A closer look at the lower-energy states of panel (a) (note that for the highest-spin case, the RASCI-tPBE and RASCI-PT2 states at 0.00 eV are superimposed).

RASCI-PT2 calculations either did not converge or were not feasible except for the dectet. The corresponding energy values (reported in Tables S2-S4) are shown in Figure 15. Note that Figure 15 shows what appears to be a negative excitation energy. This is because the RASCI-PT2 states are not ordered the same as RASCI and RASCI-PDFT. Using the labeling scheme in Table S4, the "State 1" dectet is the ground state at the RASCI and RASCI-PDFT levels. To make comparison of relative energies easier, the RASCI-PT2 energies are shifted so that the energy of State 1 is zero. However, the State 2 RASCI-PT2 energy is lower than the State 1 RASCI-PT2 energy and so appears to be negative after the shift (see also Table S4).

3.2.4.3 Configurations and populations

Looking at the main electronic configurations of the lowest energy high-spin states (Table S5), we note that the dominant configurations in all systems have weights (squared CI coefficients) far from unity. This means that the corresponding wave functions have significant multiconfigurational character. The lowest-energy undectet (high) spin state of the oxidized cluster has three contributions with weight greater than 0.10, with one clearly dominant but having a relatively small weight (0.53). This is also seen for the lowest-energy dectet state of the reduced cluster at the same (unrelaxed) geometry (weight: 0.62). In contrast, the weight of the dominant configuration for the lowest-energy dectet state at the relaxed geometry is only 0.21 and is accompanied by three other high-spin configurations that have similar weights (0.12, 0.16, and 0.11). The state-specific configurations (Table S5) also show a slightly diminished occupancy of 3*p* orbitals of the bridging sulfurs. The spin densities reported in Tables S6-S8 of the SI show that in the high-spin cases one of the unpaired electrons is delocalized over the bridging sulfurs for the oxidized cluster, whereas for the two reduced clusters it is

delocalized over the entire [2Fe-2S] fragment. The reduced cluster shows a nearly symmetric electronic structure at the unrelaxed geometry and a slightly asymmetric structure at the relaxed geometry.

Formally the oxidation states for the oxidized and reduced clusters are Fe^{+3}/Fe^{+3} and Fe^{+3}/Fe^{+2}, respectively, but the orbital occupation numbers in the ground states (see Figures S6-S8) suggest the reduced cluster is better described as $Fe^{+2.5}/Fe^{+2.5}$, which is in agreement with some experimental conclusions.[327–329] The additional electron in the reduced cluster as compared to the oxidized cluster is found in a bonding/antibonding pair of 3*d* orbitals and is evenly distributed over the two Fe centers. At the unrelaxed geometry it appears almost entirely in the δ/δ* combination of the Fe $3d_{yz}$ orbitals, with the additional occupancy being added evenly between the δ and δ* orbitals. In contrast, at the relaxed geometry it is in the σ bonding combination of the Fe $3d_{z^2}$ orbitals, and with the σ* antibonding orbital being less occupied than in the oxidized case or in the reduced unrelaxed cluster. These observations may be explained by the relaxed geometry having an Fe-Fe distance of 2.91 Å, which is longer than 2.69 Å of the unrelaxed geometry; the greater distance correlates with the more favorable σ-bonding interaction.

3.2.5 Conclusions

We have shown that RASCI-PDFT is an affordable way to approximate the full electron correlation in systems like [2Fe-2S] clusters that require multireference wave functions containing a large number of orbitals in the active space (26) and a large number of a configuration state functions (over six million). Our results with RASCI-PDFT show that correlation effects beyond those included in CASSCF, RASSCF, CASCI, and RASCI, even with the large active spaces allowed by DMRG solvers, can be considerable in these systems. These correlation effects cannot currently be evaluated by

standard post-MCSCF methods such as RASCI-PT2, which remains unaffordable for most of the low-spin multiplets.

Both oxidized and reduced clusters at their unrelaxed geometry have a low-spin ground state. However, in contrast with previous calculations, a reversed spin-ladder energy ordering is obtained with RASCI-PDFT for the reduced [2Fe-2S] cluster at the relaxed geometry. This is in agreement with some previous experimental observations of a high-spin ground state in valence-delocalized [2Fe-2S] clusters.

RASCI-PDFT makes the treatment of full correlation with bimetallic systems affordable, as demonstrated by the calculation of full spin ladders where only the highest-spin cases were feasible with RASPT2. A suitable improvement over RASCI-PDFT, and a good compromise between accuracy and feasibility towards larger active spaces and compounds, would be the use of the recently developed DMRG-PDFT[84] method. The combination of DMRG's powerful ability to reduce the computational cost scaling in a large active space with MC-PDFT's evaluation of the full electron correlation will allow a quantitative description of larger iron-sulfur clusters, e.g. [4Fe-4S], as well as an unprecedented possibility to investigate very large systems with a multiconfigurational reference function combined with both internal and external dynamic correlation.

3.2.6 Additional Information

Supporting Information: Orbitals and active spaces, detailed active space search protocol, vertical excitation energies and spin ladders, RASCI dominant configurations, partial atomic charges and spin populations, structures, SA-RASCI absolute energies. (PDF) RASCI optimized state-average natural orbitals in the Molcas format (RasOrb, plaint text). (ZIP archive). The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.9b00222.

We thank C. A. Gaggioli for helpful discussions. This work was supported in part by the National Science Foundation by grant no. CHE-1464536.

3.3 Affordable and Accurate Transition Metal Spin-State Energetics via MC-PDFT Using tPBE with High Local Exchange

This section describes the outcome of a collaborative research project carried out by Samuel J. Stoneburner (and advised by Donald G. Truhlar and Laura Gagliardi). A report on this research project will be submitted for publication.[330]

Samuel J. Stoneburner performed all calculations and wrote the manuscript.

Reproduced from **Stoneburner, S. J.**; Truhlar, D. G.; Gagliardi, L. Affordable and Accurate Transition Metal Spin-State Energetics via MC-PDFT Using tPBE with High Local Exchange. **2019**, to be submitted for publication.[330]

3.3.1 Overview

The spin energetics of transition metal complexes are explored with a variety of electronic structure methods, but most of them require compromising between accuracy and affordability. In this work, spin splittings of several iron complexes are studied with multiconfigurational pair-density functional theory (MC-PDFT), a method that inexpensively includes full correlation without distinguishing between core and valence, and are compared to previously published complete active space second-order perturbation theory (CASPT2) and CASPT2 with coupled-cluster semi-core correlation (CASPT2/CC) results. In contrast to CASPT2's systematic overestimation of high-spin stability with respect to the CASPT2/CC reference, MC-PDFT with the tPBE on-top functional in its original form overstabilizes low-spin states which can be largely corrected by modifying the exchange and correlation contributions to the on-top density functional using high local exchange (tPBE-HLE). Moreover, tPBE-HLE correctly predicts most experimental ground states, while CASPT2 incorrectly predicts high-spin ground states in all cases. MC-PDFT also demonstrates a performance advantage over CASPT2, with tPBE being faster than CASPT2 by a factor of 50 in a moderately-sized example.

3.3.2. Introduction

Understanding the spin-state energetics of transition metal complexes is important for applications such as catalysis[331–333] and magnetic properties.[334,335] In particular, iron complexes such as porphyrins have been extensively studied with theory[335–353] and are also often chosen as model systems in electronic structure method benchmarking,[66,85,343,345,348,354–360] in part due to an abundance of experimental references.[334,335,361–369]

In general, the most popular electronic structure method for molecular systems is Kohn-Sham density functional theory (KS-DFT),[23] as it is often the only choice capable of delivering sufficient accuracy at an affordable cost.[370] However, the open-shell nature of many transition metal complexes present difficulties for single-reference methods, including KS-DFT. For open-shell systems (other than high-spin) it is typically necessary to use a spin-unrestricted or a broken-symmetry formalism in order to obtain reasonable energetics.[271] The resulting ambiguities regarding the spin state (i.e., "spin contamination") are undesirable for comparing energies of different spin states, which is an especially challenging problem with Fe complexes.[135] Moreover, local density functionals are known to overstabilize low-spin states[371–376] due to the exchange leading to overly delocalized binding orbitals,[372,377] although some functionals have overcome this difficulty.[13,378] In the context of iron porphyrins, strong functional dependence for spin-state energetics with both local and hybrid functionals has been observed over the past several years.[337–340,347–350] For example, Ghosh observed local density functionals failed to properly describe the spin-state energetics and that spin densities varied widely

between the different functionals.[338] Pierloot and coworkers have noted significant inconsistencies in results obtained from hybrid functionals as well.[339,349,350]

Wave function theory (WFT) provides a menu of options for cases where KS-DFT results are ambiguous, but they are not without their own challenges. For benchmarking purposes, coupled-cluster (CC) methods such as coupled-cluster with singles, doubles, and perturbative triples (CCSD(T))[237] are often preferred, although the common practice of using restricted-spin open-shell KS-DFT to generate input orbitals for restricted-spin CCSD(T) has recently been called into question.[379] CCSD(T) is also quite expensive: Radoń provided benchmark values for a variety of iron porhyrin model systems, but primarily on smaller "mimics" due to computational cost. Several DFT functionals were also employed for both the model systems and their mimics. After taking linear fits of the DFT data for each spin excitation energy for a given system, Radoń used the differences in the linear fits to estimate the changes in excitation energies due to the use of the mimics. The CCSD(T) values for full porphyrin systems (excepting the undecorated porphyrin, FeP, and FeP(Cl), which also received CCSD(T) treatment) were then estimated by applying the adjustments calculated from the DFT trendlines to the CCSD(T) results from the mimics.[340] In addition to expense, the reliability of CC methods in general suffers in cases with significant multireference character. Several diagnostic measures have been applied to these systems to determine whether CC results can be considered reliable.[66,340,356]

Multireference effects can be explicitly included by using the multiconfigurational self-consistent field (MCSCF) methods such as complete active space self-consistent field theory (CASSCF).[5] In CASSCF, all spin- and (optionally)

symmetry-permitted electronic configurations are included for a user-specified number of electrons in a select set of orbitals. Currently the practical limit on active space size is eighteen electrons in eighteen orbitals, or (18,18).[53] (Here we use the convention (*n*,*N*) to represent an active space with *n* electrons in *N* orbitals.) Restricted active space self-consistent field (RASSCF)[7] enables larger active spaces by excluding a partial set of configurations, and active spaces with a full CAS approach as large as (40,38)[357] are now possible with full configuration interaction quantum Monte Carlo CASSCF (FCIQMC-CASSCF, also referred to as Stochastic-CASSCF),[360] or as large as (84,84)[61] with density matrix renormalization group (DMRG).[58–60]

With all of the above multireference methods some degree of correlation remains unaddressed because of excluded configurations, and thus quantitative accuracy requires an additional level of theory. Most common[63] is second-order perturbation theory (CASPT2,[8,9] RASPT2,[10] and DMRG-CASPT2[70]), which includes much of the unaddressed correlation by pertubatively adding single and double excitations. While CASPT2 has long been the method of choice for systems with strong multireference character, it is computationally expensive,[71,72] and it has been observed that it overstabilizes high-spin states by as much as 10 kcal/mol for first-row transition metal complexes.[66,376] Pierloot et al. studied this systematic error in first-row transition metal complexes and concluded that it is due to inconsistent treatment of semicore correlation (i.e., correlation associated with excitations from 3*s* and 3*p* orbitals).[66] Phung et al. then proposed a combined CASPT2/CC method featuring CASPT2 for valence correlation and CCSD(T) or CCSD[236] for the semicore correlation and concluded that the high-spin overstabilization, while not completely eliminated, was reduced to about 2 kcal/mol. The

authors also noted that application of CASPT2/CC is expected to be restricted to cases where CCSD is applicable, namely when there is a single dominant configuration.[356] A recent benchmarking study by Radoń on a different set of iron systems found similarly small errors in CASPT2/CC, albeit less systematically in favor of the high-spin states.[380]

An economical alternative to CASPT2 is multiconfiguration pair-density functional theory (MC-PDFT),[11] which takes the electron density and pair density from a preceding MCSCF calculation and applying an on-top density functional analogous to the exchange-correlation functional of KS-DFT. Because the MCSCF wave function is a spin eigenfunction, spin contamination and related concerns are completely avoided.[269] MC-PDFT has been used with CASSCF, RASSCF, and DMRG wavefunctions in applications to a variety of transition metal complexes,[40,71,72,78,243,377,381–383] including unmodified iron porphyrin (FeP).[85]

The first generation of MC-PDFT functionals are translations of KS-DFT functionals, e.g., tPBE[11] from PBE.[82] Recently in KS-DFT more accurate band gaps and molecular Rydberg excitations were obtained by multiplying the exchange contribution by a factor of 1.25 and the correlation by a factor of 0.5 (labeled "high local exchange" or HLE),[13,378] and Sharma et al. found that the same adjustments to tPBE (i.e., tPBE-HLE) made improvements in the calculated excitation energies of benzene.[384]

In this work, we examine the performance of tPBE and tPBE-HLE for spin-changing excitation energies of iron complexes. In light of the demonstrated high-spin overstabilization of CASPT2 for transition metal excitation energies, we forego using CASPT2 as a reference and instead use the CASPT2/CC values of Phung et al.[356] Accordingly, we have chosen to focus on the same iron complexes and associated spin

states featured in their work (Figure 16). The complexes consist of iron porphyrins (FeP), some modified with additional ligands, and their respective "mimics" with two amidine ligands (FeL_2, used also in the coupled-cluster work of Radoń[340]). A model of a synthetic nonheme oxidant, $[Fe(NH_3)_5O]^{2+}$, was also included.[385,386]

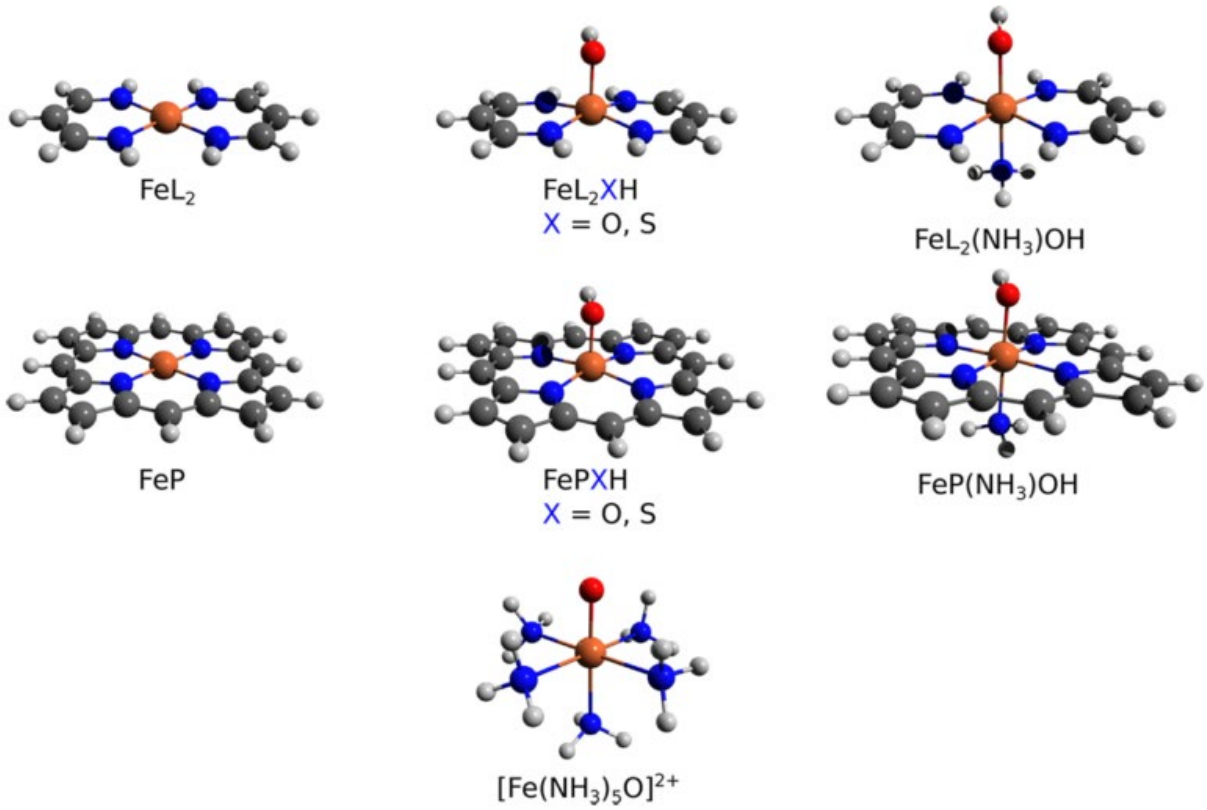


Figure 16: Iron complexes considered in this work. Reprinted with permission from Ref. [356]. Copyright 2018, American Chemical Society.

3.3.3 Computational Methods

The spins states and systems studied in Ref. [356] were calculated with MC-PDFT[11] in Molcas 8.2[53] using tPBE,[11] the translated on-top version of PBE.[82] Translated on-top versions of revPBE[83] and BLYP[80,81] (trevPBE and tBLYP,[11] respectively) were also used, with results reported in the Supporting Information (Table S1). High local exchange (HLE)[13,378,384] values were obtained by multiplying the individual exchange and correlation contributions by 1.25 and 0.5, respectively. Following the convention of Ref. [356], spin-splittings (ΔE) are calculated as follows:

$$\Delta E = E_{\mathrm{IS,LS}} - E_{\mathrm{HS}} \tag{76}$$

where "IS,LS" indicates either intermediate or low spin, "HS" indicates high spin, and E is the energy. The CASPT2 and CASPT2/CC results from Ref. [356] are with both valence and semi-core correlation energy (ΔE(+sp) as defined in Refs. [66,356]). Note that Equation (76) results in the sign convention that a negative ΔE value indicates the lower spin is favored.

Input and orbital files were obtained from the authors of Ref. [356] to ensure the same orbitals would be used for CASSCF[5] and CASPT2[8,9] calculations. The active spaces for all systems other than $[Fe(NH_3)_5O]^{2+}$ were selected in Ref. [66] and included all 3*d* and 4*d* orbitals of the Fe along with any ligand orbitals judged to have important covalent interactions with the 3*d* orbitals of the Fe. It should be noted that for some low-spin states the 4*d* orbitals corresponding to unoccupied 3*d* orbitals were omitted from the active space to prevent mixing or rotation into other virtual orbitals. The active spaces $[Fe(NH_3)_5O]^{2+}$ were selected in Ref. [356] according to a similar philosophy, with 4*d*

orbitals included in the active space only for occupied 3*d* orbitals. More detailed explanations can be found in the original references.[66,356]

In Ref. [356], a variety of correlation consistent (cc) basis sets were used in order to extrapolate to the complete basis set (CBS) limit. In order to minimize questions of basis set dependency we used "awC5Z/aTZ", the basis set found to be closest to the CBS limit. "awC5Z/aTZ" is a shorthand notation used in Ref. [356] for the following:

Fe: aug-cc-pwCV5Z-DK[387]

H: cc-pVTZ-DK[257,388]

All other atoms: aug-cc-pVTZ-DK[257,388,389]

In the above names, "cc" refers to "correlation consistent", meaning that as basis set size is increased basis functions with similar contributions to the correlation are added, "p" refers to the inclusion of polarization basis functions to enable reduced symmetry about the nucleus, "wCV" is "weighted core-valence" and means that both valence and core-valence correlation are included without core-core correlation (in contrast to a simple "V", meaning valence-only). "aug" means the basis set is augmented with additional diffuse functions to describe weak interactions, "TZ" and "5Z" refer to the number of basis functions included and mean "triple-zeta" and "quintuple-zeta", respectively, and "DK" means that the basis sets used were versions for Douglas-Kroll-Hess relativistic treatment. Note that Phung et al. chose smaller basis sets for the ligands than for the metals because of the expected importance of the (localized) 3*d* orbitals.[356]

Cholesky decomposition[185] was used with a threshold of 10^{-6} a.u. to reduce the memory and storage requirements of the two-electron integrals. A second-order Douglas-Kroll-Hess Hamiltonian[320,390,391] was employed to account for scalar relativity. All

CASPT2 calculations had the default ionization potential electron affinity (IPEA) shift[65] of 0.25 a.u. to correct systematic errors and an imaginary shift[68] of 0.1 a.u. to reduce intruder states. All of these options are the same as in Ref. [356]. However, while Ref. [356] used Molcas 8.1,[53] we used Molcas 8.2[53] because of its more complete implementation of MC-PDFT.

3.3.4 Results and Discussion

Table 20: ΔE of all complexes studied (kcal/mol). Negative values indicate the lower spin state is favored.

System	States	tPBE	tPBE-HLE	CASPT2[35,6]	CASPT2/CC[35,6]
FeL_2	^1A_g-^5A_g	14.25	24.83	37.84	31.92
	$^3B_{1g}$-^5A_g	-13.46	0.51	-2.28	-4.80
	$^3B_{3g}$-^5A_g	-16.67	-4.04	-4.63	-6.83
FeL_2OH	$^2A''$-$^6A'$	-17.59	0.98	10.25	6.55
	$^4A''$-$^6A'$	-13.03	-0.23	6.90	4.68
FeL_2SH	$^2A''$-$^6A'$	-29.86	-15.95	-9.21	-10.22
	$^4A''$-$^6A'$	-17.26	-5.64	0.31	-3.17
$FeL_2(NH_3)OH$	$^2A''$-$^6A'$	-31.09	-10.94	-4.63	-8.58
	$^4A''$-$^6A'$	-6.95	5.42	12.12	9.69
FeP	$^1A_{1g}$-$^5A_{1g}$	15.31	23.84	38.03	32.41
	$^3A_{2g}$-$^5A_{1g}$	-8.43	1.09	3.74	0.34
	^3E_g-$^5A_{1g}$	-7.28	0.25	5.65	2.29
FePOH	$^2A''$-$^6A'$	-9.25	5.67	18.81	14.27
	$^4A''$-$^6A'$	-8.20	3.34	12.40	9.85
FePSH	$^2A''$-$^6A'$	-16.41	-7.85	2.13	0.31
	$^4A''$-$^6A'$	-11.13	-1.81	6.53	2.71
$FeP(NH_3)OH$	$^2A''$-$^6A'$	-28.73	-12.32	1.31	-3.59
	$^4A''$-$^6A'$	-6.50	4.69	14.88	11.88
$[Fe(NH_3)_5O]^{2+}$	$^3A''$-^5A	-9.97	-1.23	2.50	-0.42
MUD with respect to PT2/CC		16.40	5.07	3.33	-
MUDs, LS-HS only[a]		20.81	6.85	3.93	-
MUDs, IS-HS only[a]		13.19	3.78	2.90	-

a: LS (low-spin): singlets and doublets
IS (intermediate-spin): triplets and quartets
HS (high-spin): quintets and sextets

Our primary results are presented in Table 20. The CASPT2 and CASPT2/CC ΔE results are from Ref. [356] and we use CASPT2/CC as our reference. Note that the CASPT2 results are consistently more positive than CASPT2/CC, as discussed in Refs. [66] and [356]. In contrast, tPBE results are consistently more negative than CASPT2/CC, similar to the low-spin overstabilization that would be expected in local KS-DFT functionals such as PBE. However, the mean unsigned deviations (MUD) for tPBE is much larger than for CASPT2 (16.40 vs. 3.33 kcal/mol, respectively). Applying the HLE rescaling of exchange and correlation alleviates the problem, with an MUD of only 5.07 kcal/mol. We also note that trevPBE-HLE and tBLYP-HLE (Table S1) yield similarly good MUDs (3.59 and 5.08 kcal/mol, respectively). The individual unsigned deviations for tPBE-HLE are all smaller than for tPBE, usually by at least a factor of two. With only a few exceptions, the tPBE-HLE deviations are negative, indicating that for the most part the adjustments to exchange and correlation do not overcompensate for the low-spin overstabilization in tPBE. The negative deviations are also encouraging in light of Phung et al.'s comment that CASPT2/CC likely overestimates the stability of high-spin states by about 2 kcal/mol, implying that the true answers may be slightly more negative than the CASPT2/CC reference.

The systematic nature of the deviations is reinforced when they are separated by LS-HS vs. IS-HS (where LS is singlet or doublet, IS is triplet or quartet, and HS is quintet or sextet): the deviations for IS-HS are almost always of smaller magnitude than LS-HS deviations for the same system. The trend is not dependent on whether the LS-HS case or the IS-HS case(s) have larger magnitudes in the CASPT2/CC reference values, further reinforcing that the trend arises from systematic low- or high-spin overstabilization rather

than being a simple correlation between the magnitudes of the reference values and the deviations.

An analysis of the differences in absolute energies between tPBE and tPBE-HLE shows that HLE accomplishes its correction of tPBE's low-spin overstabilization primarily by stabilizing the high-spin state, the energy of which is always decreased by 7-20 kcal/mol more than any other given state. (Specific values are presented in Section S4 of the Supporting Information.)

3.3.4.1 Ground Spin States

Table 21: Ground states of molecules studied compared across different methods and experiment. For convenience, correct predictions by each method are highlighted in green.

System	Exp.[a]	tPBE	tPBE-HLE	CASPT2[356, b]	CASPT2/CC[356, b]
FeP	Triplet[c]	Triplet	Quintet	Quintet	Quintet
FePSH	Doublet[d]	Doublet	Doublet	Sextet	Sextet
$FeP(NH_3)OH$	Doublet[e]	Doublet	Doublet	Sextet	Doublet
$[Fe(NH_3)_5O]^{2+}$	Triplet[f]	Triplet	Triplet	Quintet	Triplet

a: Experimental work was on more complicated systems of which the systems studied here can be understood as models. For more detail see Ref. [356].
b: CASPT2/CC correctly predicts all of the experimental ground spin states in the CBS-extrapolated results presented in Table 3 of Ref. [356], while CASPT2 predictions are unchanged from those shown here.
c: Refs. [366–368]
d: Ref. [369]
e: Ref. [361,392]
f: Ref. [393]

In Ref. [356] Phung et al. listed several points of comparison with experiment in cases for which their CBS-extrapolated CASPT2 and CASPT2/CC results disagreed

regarding the ground state. We summarize these comparisons in Table 21 along with the corresponding results for tPBE and its variants. As our calculations are implicitly at 0K we focus exclusively on the experimental ground states rather than the overall state ordering, which is generally not provided from the experimental work cited in Table 21 except at higher temperatures or when varying experimental conditions such as pH. That being said, agreement with experiment at ground state is quite encouraging: while CASPT2 erroneously predicts a high-spin ground state in all cases, tPBE and tPBE-HLE correctly identify the low- or intermediate- (as appropriate) spin states as ground with the exception of tPBE-HLE with FeP.

3.3.4.2 Active Space Dependency for FeP

In this work we have constrained ourselves to the active spaces used by Phung et al.[356] to ensure direct comparison with the CASPT2/CC reference values. However, the effect of active space selection for FeP has been noted on multiple occasions (e.g., Refs. [66,85,357]) and deserves some comment here.

CASPT2 incorrectly predicts the quintet $^5A_{1g}$ state to be below the triplet ^3E_g and $^3A_{2g}$ states in energy with both (8,11) and (16,15) active spaces.[66] However, Li Manni et al. recently demonstrated with Stochastic-CASSCF that the ordering of the ^3E_g and $^5A_{1g}$ states can be corrected with an active space of (32,34) or (40,38), even with no post-MCSCF procedure included.[357] To the best of our knowledge, CASPT2 has not yet been attempted with active spaces that large. (While RASPT2 with a (34,35) active space was reported in Ref. [355], as some of the same authors noted in Ref. [66] there was a basis set error that significantly affected the results.) However, Zhou et al. tested tPBE with a

variety of active spaces and found that tPBE triplet-quintet gaps (with both triplet states) retained the same signs and varied by less than 2.2 kcal/mol between CASSCF-(8,11) and DMRG-CASSCF-(34,35),[85] demonstrating that MC-PDFT is quite stable with regard to active space selection for FeP. Additionally, their (8,11) results are within 0.15 kcal/mol of the tPBE results presented here, despite the different basis sets used. Given that HLE is a simple rescaling of exchange and correlation, it is expected that the stability with regard to active space would apply to tPBE-HLE as well. We tested this hypothesis by applying the HLE adjustments to the tPBE calculations from Zhou et al.[85] and confirmed that the stability with regard to active space observed for tPBE carries forward to tPBE-HLE (see Section S2 of the Supporting Information for more detail).

3.3.4.3 Timing Comparisons

One of the primary advantages of MC-PDFT over second-order perturbation theory is that it can be performed at a much lower computational cost, as shown in Table 22. (The MC-PDFT column is labeled simply "tPBE" because the HLE adjustments are made entirely post-hoc and do not add any computational time.) Both CASPT2 and MC-PDFT were performed in the same run. The remaining time contributing to the "Total" column primarily consists of calculating the one- and two-electron integrals (which is lengthy due to the large basis set used) and the CASSCF calculations. Note that as the total calculation time increases it is increasingly dominated by the CASPT2 step. In the most extreme case shown, CASPT2 takes over 70% of the total time and requires over eight hours, while MC-PDFT takes only nine minutes (i.e., MC-PDFT is faster by a factor of 50). Furthermore, there were several systems we did not include in Table 22

because they could be treated in only parallel, typically due to the memory requirements of the integral generation, and thus timing comparisons were not possible. Our previous work on organic diradicals,[269] however, leads us to expect the relative cost benefit of MC-PDFT over CASPT2 would only increase for these other systems.

Table 22: Comparison of serial computational time for CASPT2 and MC-PDFT (h:mm), sorted by total computation time.

System	State	Basis functions	CSFs	Total comp. time	CASPT2	tPBE
FeL_2	^5A_g	860	5,476	0:59	0:18	0:02
FeL_2	^1A_g	860	3,540	1:00	0:17	0:02
FeL_2	$^3B_{1g}$	860	12,720	1:03	0:19	0:02
FeL_2	$^3B_{3g}$	860	12,740	1:09	0:17	0:02
FeP	$^5A_{1g}$	1532	5,476	7:00	3:18	0:10
FeL_2SH	$^2A''$	924	52,272	7:09	2:31	0:16
FeP	$^1A_{1g}$	1532	8,290	7:21	3:30	0:11
FeL_2SH	$^6A'$	924	156,156	7:32	2:53	0:16
FeP	^3E_g	1532	12,740	7:37	3:46	0:11
FeP	$^3A_{2g}$	1532	12,720	7:42	3:46	0:11
FeL_2SH	$^4A''$	924	429,534	7:52	3:09	0:16
FeL_2OH	$^2A''$	920	490,776	8:03	3:17	0:16
$[Fe(NH_3)_5O]^{2+}$	$^3A''$	746	5,271,210	8:17	4:23	0:10
FeL_2OH	$^4A''$	920	2,928,170	10:09	5:10	0:17
$[Fe(NH_3)_5O]^{2+}$	$^5A'$	746	8,509,200	11:11	8:03	0:09

3.3.5 Conclusions

While CASPT2 demonstrates a consistent overstabilization of high-spin states and the tPBE on-top functional demonstrates an even stronger systematic overstabilization of low-spin states in comparison to CASPT2/CC reference values, tPBE with high local exchange (tPBE-HLE) yields deviations from the reference only slightly larger than CASPT2 (5.07 and 3.33 kcal/mol, respectively). Furthermore, tPBE-HLE correctly identifies most of the experimental ground states while CASPT2 instead predicts high-spin ground states in all cases. We find that high local exchange significantly improves MC-PDFT's description of spin-state energetics in iron complexes.

As was shown in the timing comparison, MC-PDFT can provide significant cost advantages over CASPT2 as well. The systems in this study were of low-to-moderate multireference character, but for systems with greater multireference character or simply at a larger scale it will be important to employ methods that can give accurate results at an affordable cost. MC-PDFT does not suffer from a lack of MR treatment (as do coupled-cluster-based methods), nor does it become prohibitively expensive as quickly as CASPT2. Moreover, the recent development of DMRG-PDFT[84] enables the use of tPBE-HLE even for large active spaces such as may be necessary for treating systems with multiple transition metal centers.

3.3.6 Additional Information

Supporting Information: ΔE results for trevPBE and tBLYP, with and without HLE. tPBE-HLE results for FeP with the active spaces of Ref. [85]. MC-PDFT absolute energies. Relative changes in energy due to HLE. M-diagnostic results. This material is available upon request and will be available online free of charge upon publication.

The authors thank Kristine Pierloot for providing files associated with Ref. [356] and Chen Zhou for providing files associated with Ref. [85]. The authors also thank Larry Que for helpful conversation.

This work was supported in part by the National Science Foundation by grant no. CHE-1464536.

4 Gas Separations in Metal-Organic Frameworks

4.1. Origin of the Strong Interaction between Polar Molecules and Copper(II) Paddle-Wheels in Metal Organic Frameworks

This section describes the outcome of a collaborative research project carried out by Daniele Ongari, Davide Tiana, and Samuel J. Stoneburner (and advised by Laura Gagliardi and Berend Smit). A report on this research project has been published.[394]

Samuel J. Stoneburner performed the multireference calculations.

4.1.1 Overview

The copper paddle-wheel is the building unit of many metal organic frameworks. Because of the ability of the copper cations to attract polar molecules, copper paddle-wheels are promising for carbon dioxide adsorption and separation. They have therefore been studied extensively, both experimentally and computationally. In this work we investigate the copper-CO_2 interaction in HKUST-1 and in two different cluster models of HKUST-1: mono-copper $Cu(formate)_2$ and di-copper $Cu_2(formate)_4$. We show that density functional theory methods severely underestimate the interaction energy between copper paddle-wheels and CO_2, even including corrections for the dispersion forces. In contrast, a multireference wave function followed by perturbation theory to second order, using the CASPT2 method, correctly describes this interaction. Restricted open shell Møller-Plesset 2 method (ROS-MP2, equivalent to (2,2) CASPT2) was also found to be adequate in describing the system and was used to develop a novel force field. Our parametrization is able to predict the experimental CO_2 adsorption isotherms in HKUST-1, and it is shown to be transferable to other copper paddle-wheel systems.

4.1.2 Introduction

Metal organic frameworks (MOFs) are a class of three-dimensional nanoporous materials composed of metal nodes connected by organic ligands. The oriented coordination bond between these two components is responsible for the structure of the crystal. The possibility of combining different metals with different ligands provides a large variety of MOF structures. More than ten thousand structures have already been synthesized,[96] but this is only a small fraction of the hundreds of thousands of structures that have been predicted computationally.[395]

MOFs have attracted considerable attention in the past decade for various applications, including gas adsorption and storage,[396] gas separation,[397] fuel production,[398] chemical sensing,[118] and catalysis.[399]

Computational modeling is extensively used to investigate the properties of synthesized materials for a given application, and to predict the performance of hypothetical structures. In the case of gas adsorption, the quality of the model directly derives from the accuracy with which one can describe the microscopic interactions between the guest molecules and the framework. Density functional theory (DFT) calculations are routinely used for this purpose.[28,400,401] However, weak interactions, due to dispersion forces arising from electron correlation, are poorly described by standard DFT methods. Corrections need to be introduced for this purpose (see the recent review of Grimme *et al.*[402] and references therein). Alternatively, post Hartree-Fock methods can be employed to evaluate interaction energies with high accuracy. However, because of the unfavorable scaling with the size of the system, they can hardly be used directly to compute interaction energies in MOFs, whose unit cells typically contains hundreds of atoms.[28]

This work focuses on the interaction between the carbon dioxide molecule and the copper (II) paddle-wheel, which is a metal organic structure composed of two copper cations connected to four carboxylates anions in a square planar coordination geometry. The smallest example of this structure is the $Cu_2(formate)_4$ molecule (Figure 17, left).

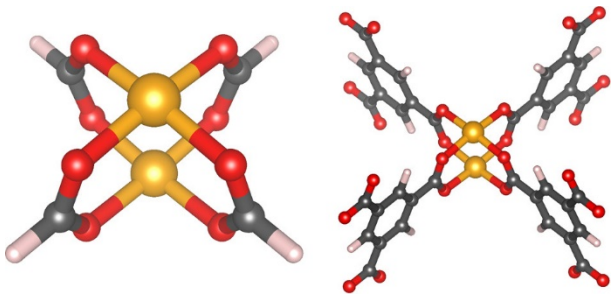


Figure 17: Copper paddle-wheel structure is composed of two coppers atoms bridged though four di-carboxilate anion. $Cu_2(formate)_4$ (left) represents the simplest paddle-wheel geometry possible. Di-copper benzil-1,2,3-trimetylcarboxylate, $Cu_2(BTC)_4$ (right), is the building unit of HKUST-1 framework: each BTC has three caboxylate groups that allow the creation of a three-dimensional network.

The copper paddle-wheel is the building unit of many MOFs, including HKUST-1 ($Cu_3(BTC)_2$). The structure of HKUST-1 presents three pores (Figure 18, left) and several characteristic adsorption sites for CO_2 (Figure 18, right). The biggest pore is characterized by the presence of twelve open metal sites (OMSs) i.e. unsaturated copper cations which are obtained after solvent removal and which are able to attract polar molecules through electrostatic interaction.

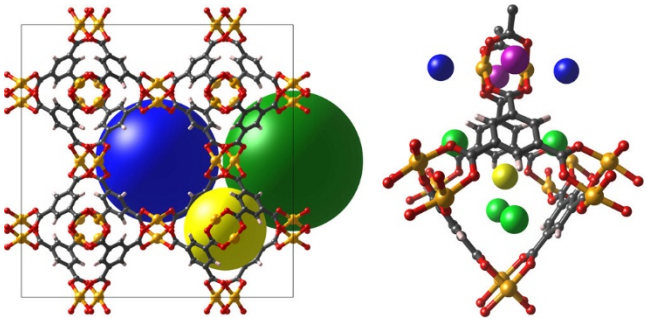


Figure 18: Three different pores in HKUST-1 (left): big pore (blue), medium pore with open metal sites (green), small pore (yellow). Characteristic sites of adsorption for CO_2 (right): open metal site (blue), small pore window (green), small pore center (yellow) and large pore corner (purple).

HKUST-1 is one of the earlier reported MOFs.[403] It is among the best performers for natural gas storage,[404] and it has also attracted interest for gas separation[405–407] and heterogeneous catalysis.[408–410] Because of its popularity, many experimental data are available for this framework. Wu *et al.*[411] conducted *in situ* neutron diffraction studies for CO_2 adsorption in HKUST-1 that show that at low loading of CO_2 and low temperature (20 K) the open metal site is the strongest adsorption site because it is the only one to be occupied at a 1:1 CO_2:Cu ratio of loading. They were also able to rank the strength of the secondary sites by increasing the amount of CO_2 and observing the filling in each site: small pore windows sites and center sites are the second and the third, respectively, and large pore corner sites are the fourth in terms of order of filling and therefore interaction energy strength.

In a recent work, Grajciar *et al.*[412] showed that DFT dispersion corrected methods, e.g. Grimme's pairwise correction for dispersions (D2[413] and D3[414]) and van der Waals density functionals (vdW-DF[415] and vdW-DF2[416]), underestimate the strength of the open metal site and are not able to reproduce the experimental adsorption data obtained by Wu *et al.*

The van der Waals density functional methods, in particular, were used previously by our group to compute the CO_2 binding energy in MOF-74 for different metals[400,417–419] and to parametrize the associated force field.[420,421] A good agreement with experiment was always observed, giving rise to the question of why the same ab-initio methods are not able to model correctly the CO_2 interaction with the open metal site in a copper paddle-wheel framework. This underestimation of the interactions in HKUST-1 motivated Grajciar *et al.* to employ a DFT - Coupled Clusters corrected (DFT/CC) method[422] to study this system and obtain a tailor-made correction for the CO_2 interaction with HKUST-1. In DFT/CC, the error associated with the PBE density functional is

corrected by a term dependent on the pairwise distance between the CO_2 atoms and the atoms of the framework. This term was estimated from the difference between the DFT and the CCSD(T) computed one-dimensional potential energy curves of CO_2 interacting with some other reference molecules, i.e. H2, benzene, CO_2, and Cu(formate)$_2$.

It is known that a copper-copper magnetic interaction is present in HKUST-1,[423] and consequently the correlation between the electrons of the two coppers can affect the interaction with the CO_2. Because of this, we investigated the legitimacy of transferring the DFT error for CO_2 interaction from the mono-copper system Cu(formate)$_2$ to the di-copper paddle-wheel structure (and to the HKUST-1 framework) by using multireference wave function methods. These methods are critical for accurately modeling systems with a relevant magnetic coupling such as the Cu paddle-wheel.[424–426] Accordingly, we explored in this work the adequacy of different quantum methods for describing the electronic structure of the system and the interaction between the metal cation and carbon dioxide.

Furthermore, we used our insights to develop a classical force field that is able to accurately describe the Cu paddle-wheel interaction with CO_2 and model the adsorption in MOFs containing this building unit. It was estimated[427] that among 4764 three-dimensional MOF structures from the Cambridge Structural Database[428] (as refined in the CoRE MOF database),[131] 4.2% of them contain the Cu paddle-wheel, and another 3.5% contai the paddle-wheel motif formed by other cations. Cu paddle-wheels are a recurrent building unit among the different MOFs, and with a reliable and transferable force field it would be possible to also screen these frameworks and identify their performances for CO_2 adsorption.

4.1.3 Computational Methods

The periodic calculations were performed using the Perdew-Burke-Ernzerhof GGA method PBEsol[429] to optimize the framework, and the second version of van der Waals dispersion corrected density functional vdW-DF2[416] to compute the interactions. The plane wave Quantum Espresso 5.4 package[430] was employed. We adopted the projector augmented wave (PAW) method[431,432] with a cutoff energy of 60 Ry for the wave function and 300 Ry for the electron density. Due to the dimension of the unit cell of HKUST-1 a Γ-point sampling of the Brillouin zone integration was used, with a smearing occupation of 0.02 Ry.

For the cluster calculations, geometry optimizations were performed using the unrestricted M06-L/cc-pVDZ[433] level of theory and subsequent single point energy difference calculations were performed using restricted open shell MP2,[434] (ROS-MP2) and unrestricted M06-L and M06.[433] The Gaussian-09 package[282] was employed. We tested the convergence of the basis set using cc-pVDZ, AUG-cc-pVDZ, cc-pVTZ and AUG-cc-pVTZ[257,387,388,435] A spin multiplicity of three was used to model the magnetic state of the copper paddle-wheel clusters. To account for the error in computing the interaction due to the basis set superposition, the counterpoise method by Boys and Bernardi was employed.[436] For the ROS-MP2 calculations the frozen orbitals are the 1s for C and O, and 1s, 2s, 2p, 3s and 3p for Cu.

Multireference calculations were performed on the cluster models using the complete active space self-consistent field method (CASSCF)[6] followed by second-order perturbation theory (CASPT2)[8] using Molcas 8.2.[53] All CASSCF/CASPT2 calculations were performed without symmetry. Relativistic basis sets of atomic natural orbital type (ANO-RCC)[264] were employed for all the atoms. To explore basis set convergence, three different basis sets of increasing size were tested. The first one, BS1, is of double-ζ

quality plus polarization; the second one, BS2, is of triple-ζ quality plus polarization on Cu, O, C atoms and double-ζ quality plus polarization on H atoms; the third one, BS3, is of quadruple-ζ quality plus polarization for Cu and CO_2, triple-ζ quality plus polarization on the remaining C and O atoms, and double-ζ quality plus polarization on H atoms. Scalar relativistic effects were included using the Douglas-Kroll-Hess Hamiltonian.[391] The computational cost arising from the two-electron integrals was drastically reduced by employing the Cholesky decomposition technique.[437] The decomposition threshold was chosen to be 10^{-4}, as this should correspond to an accuracy in total energies of the order of mHartree or higher. In the CASPT2 calculations, in order to prevent possible intruder states, an imaginary shift of 0.1 au was added to the zero-order Hamiltonian. The default IPEA shift of 0.25 au was used. The default choices of the program were employed for freezing orbitals, resulting in the 1s orbitals of C and O being frozen, along with the 1s, 2s, 2p, and 3s orbitals of Cu.

For the cluster models the interaction energy between the framework and the CO_2 molecule was computed as the difference between the energy of the super-system, the framework plus CO_2, and the energies of the two isolated components, namely CO_2 and the framework.

The Raspa 2.0 package[438] was employed for the force field calculations. In all the simulations TraPPE[439] Lennard-Jones parameters and charges were used to model CO_2-CO_2 interactions, while different sets of parameters were used to model the framework-CO_2 interaction, as discussed within the results. The details of the simulations are provided in the Supporting Information.

4.1.4 Results and Discussion

4.1.4.1 Comparison of Simulated and Experimental Isotherms in HKUST-1

The CO_2 isotherms computed with the standard force field i.e. UFF,[440] DREIDING,[441] and TraPPE,[439] are found to be in strong disagreement with the experimental data in the range of pressure from zero to one bar. Figure 19 shows the simulated isotherms, computed using the Grand Canonical Monte Carlo (GCMC) technique with different sets of parameters for the dispersion forces and the corresponding experimental isotherms.

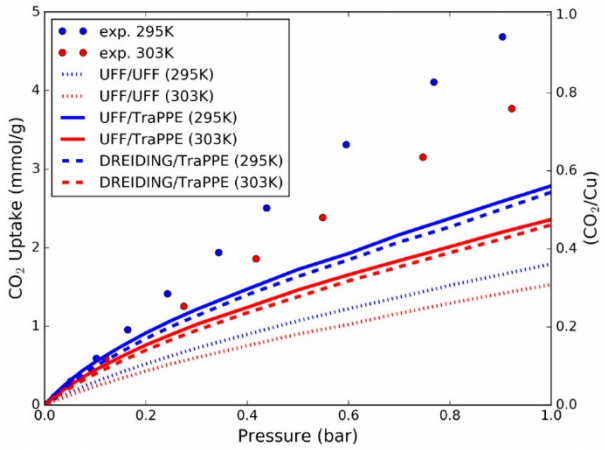


Figure 19: Comparison of experimental (295[405] and 303 K[422]) and simulated adsorption isotherms. TraPPE[439] Lennard-Jones parameters and charges are used for CO_2 interactions. To compute the dispersion forces acting between CO_2 guest molecules and the crystal, three commonly used approaches are compared. First we used Lennard-Jones parameters from UFF[440] (Lorentz-Berthelot mixing rules). Then we used UFF/TraPPE and DREIDING/TraPPE parameters[441] (notation: $FF_{framework}/FF_{adsorbate}$). The point charges for the framework atoms are extracted from a PBEsol DFT calculation using the REPEAT scheme;[442] in the Supporting Information we reported the charges' values, and we compared them with the values obtained by using Bader's method.[443] The framework is assumed to be rigid in all the simulations.

All the simulations underestimate the uptake of CO_2, which means that the force field underestimates the adsorbate-host interactions. The force field interaction energies for specific sites are compared to those obtained by DFT calculations in Table 23. The binding energy for each site, corresponding to the optimized position of a CO_2 molecule in the open metal site, in the small pore window site, and in the small pore center site, are reported.

Table 23: Interaction energy (kJ/mol) between CO2 and HKUST-1, for different adsorption sites[a]

Method	Open metal	Window	Center
FF (UFF/UFF)	-19.3	-25.7	-26.3
FF (UFF/TraPPE)	-19.0	-27.5	-29.0
FF (DREIDING/TraPPE)	-19.4	-27.2	-28.5
DFT (vdW-DF2)	-22.1	-30.2	-26.3
DFT (PBEsol)	-12.1	-6.7	-0.8
DFT/CC (Grajciar *et al.*[422])	-28.2	-23.1	-23.2

[a]The open metal site in the apical position of copper paddle-wheel, the window and the center of small octahedral pores. Force field and periodic DFT calculations are compared. Results obtained with PBEsol are also reported to appreciate the contribution of the dispersion corrections introduced by the vdW-DF2 non local functional.

From neutron diffraction *in situ* experiments by Wu *et al.* we know that OMSs are the first filled sites, then windows and cage sites get populated by CO_2. This observation proves that OMSs have the strongest binding energy. Despite the fact that UFF/UFF, UFF/TraPPE and DREIDING/TraPPE force field are giving similar results to the vdW-DF2 method, and this could in principle validate the force fields, we clearly see from Table 23 that in all four of these cases the OMS is predicted to be the weakest site. As a consequence, these standard methods erroneously predict that the OMS is poorly occupied, as its interaction energies with CO_2 are $\sim 4k_bT$ and $\sim 60k_bT$ weaker than

other sites at 303K and 20K, respectively. Standard force fields are known to incorrectly model the strong interaction of adsorbate molecules with OMSs in MOFs,[444] but vdW-DF2 is also showing the same problem in the case of copper paddle-wheel, while it was found to model accurately the open metal site interaction with CO_2 for other MOFs.[417,420,421]

There are different assumptions in these calculations that may not hold for this system, therefore the interaction energy between carbon dioxide and the copper atom in HKUST-1 was also computed using other approaches. We considered the introduction of the Hubbard correction[445] to model the *d* orbitals of copper, because it was shown to influence the CO_2 interaction with the OMS in MOF-74.[417,446] The value of U=3.8 eV, which can reproduce the experimental oxidation energy of copper,[447] was used. Also different versions of the van der Waals density functional were compared to vdw-DF2 method, i.e. vdW-DF[448] and revised-vdW-DF2.[449] In all the cases the geometry of CO_2 was optimized keeping the framework rigid, as obtained from the PBEsol calculation. The results are reported in Table 24. No significant deviations in the interaction energy were found, the only slightly increased value being obtained with vdW-DF, which is known to systematically overestimate dispersion interactions.[450]

Table 24: CO_2 open metal site interaction energies in HKUST-1 computed with different dispersion corrected DFT methods.

Method	Open metal site interaction
vdW-DF	-24.9 kJ/mol
vdW-DF2	-22.1 kJ/mol
vdW-DF2+U	-21.4 kJ/mol
vdW-DF2-rev	-20.2 kJ/mol

Finally, the rigid framework assumption was neglected, performing a full optimization of the framework's atoms with the adsorbed molecule in the OMS, using vdW-DF2. No significant deviation in the binding energy was found: -1.6 kJ/mol of difference from the rigid calculation. Moreover, we noticed an exaggerated distortion of the copper paddle-wheel structure which has not been reported experimentally, suggesting the inadequacy of the vdW-DF2 method to optimize the crystal geometry. The rigidity of the adsorbent was therefore assumed as reasonable.

4.1.4.2 Interactions Computed in the Cluster Models

To understand why vdW-DF2 method underestimates the CO_2-Cu interaction in HKUST-1, we analysed two smaller representative clusters, Cu(formate)$_2$ and Cu_2(formate)$_4$. The interaction energy with carbon dioxide was scanned at different distances by keeping the CO_2 molecule perpendicular to the CuO_4 plane, as shown in Figure 20.

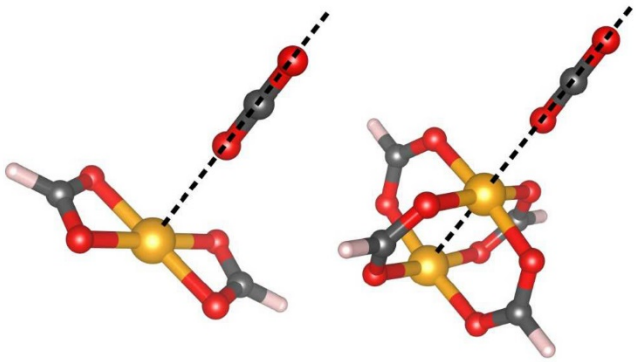


Figure 20: Path representation of linear scans of CO_2 interacting with Cu(formate)$_2$ (left) and Cu_2(formate)$_4$ (right). The dotted line, along which the CO_2 molecule is displaced, is perpendicular to the CuO_4 plane.

This configuration, referred here as "linear", was chosen to decrease the number of degrees of freedom for the CO_2 position to just one i.e. the copper-oxygen distance in the axial direction. This configuration also minimizes all the pairwise contributions of the interaction but the copper-oxygen one, which is the one vdW-DF2 is failing to model properly. Within HKUST-1, the optimal linear configuration corresponds to a distance of 2.65 and a binding energy of -13.4 kJ/mol, computed using vdW-DF2. Figure 21 and Figure 22 show the interaction energy of CO_2 as a function of the Cu-O distance computed with different methods in $Cu(formate)_2$ and $Cu_2(formate)_4$ respectively.

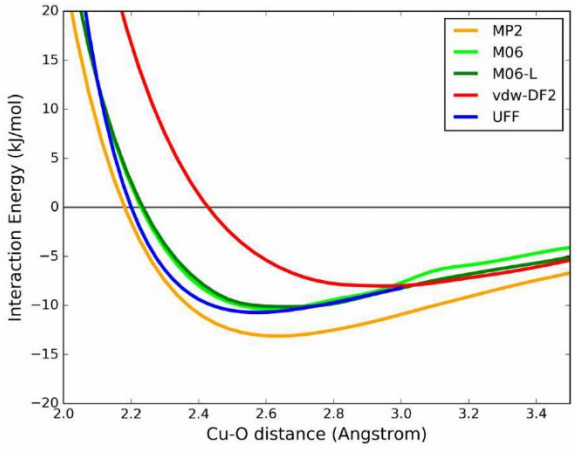


Figure 21: Interaction energy profile for the CO_2-$Cu(formate)_2$ linear scan: the interaction energy is plotted as a function of the distance between the copper atom and the CO_2 molecule's oxygen.

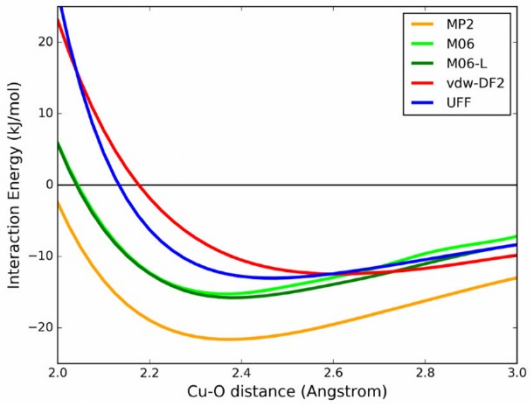


Figure 22: Interaction energy profile for the CO_2-$Cu_2(formate)_4$ linear scan: the interaction energy is plotted as a function of the distance between the CO_2 molecule's oxygen and the closest copper.

The inspection of the energy profiles reported in Figure 21 for $Cu(formate)_2$, shows that the CO_2-copper binding energies differ within 4 kJ/mol among the various methods, ranging between -8.0 kJ/mol (vdW-DF2) and -13.1 kJ/mol (MP2). The minimum energy distance for vdW-DF2 is longer than with the other methods, 2.9 instead of 2.5-2.6 . The M06 and M06-L functionals produce similar energy profiles. Hence, the inclusion of the semi-local contribution with Hartree-Fock exchange present in M06 has a minor effect. It is also interesting to note the overall good agreement with the UFF force field. The attraction computed by the force field is mainly due to the coulombic (REPEAT-TraPPE) interaction, with only a small influence of dispersion forces: the electrostatic contributions represent 96% of the interaction at the optimal distance of 2.50 .

In the $Cu_2(formate)_4$ case, vdW-DF2, M06 and M06L underestimate the interaction energy compared with ROS-MP2 by 9.1 kJ/mol, 6.4 kJ/mol and 5.8 kJ/mol, respectively. Moreover, if compared to the mono-copper system, the ROS-MP2 calculation leads to a binding energy which is 8.5 kJ/mol more stable in this di-copper system.

In a second series of calculations we optimized the position of the CO_2 molecule, keeping the $Cu_2(formate)_4$ cluster rigid. The CO_2 molecule creates an angle with the copper-copper line of 109 (vdW-DF2) to 116 (M06-L) due to both the interaction of the lone pair of CO_2 oxygen with the copper and the partially positive CO_2 carbon with partially negative oxygen from the paddle-wheel. This optimized configuration is referred here as "tilted" position, because of the CO_2 inclination with respect to the the CuO_4 plane. The interaction energies between $Cu_2(formate)_4$ and the linear and tilted configurations of CO_2 computed with different methods are reported in Table 25.

Table 25: Energy of interaction (kJ/mol) between $Cu_2(formate)_4$ and CO_2 in linear and tilted conformation.[a]

	linear CO_2		tilted CO_2		
Method	Interaction Energy (kJ/mol)	Cu-O dist. (Å)	Interaction Energy (kJ/mol)	Cu-O dist. (Å)	Cu-O-O angle (deg)
FF(UFF/UFF)	-13.0	2.5	-14.3	2.5	127.4
ROS-MP2/cc-pVTZ	-18.2 (-24.5)	2.4	-22.9 (-31.3)	(M06L opt)	(M06L opt)
ROS-MP2/ANO-RCC(BS2)	-20.4 (-38.0)	2.4	-24.8 (-43.3)	(M06L opt)	(M06L opt)
ROS-MP2/aug-cc-pVTZ	-21.6 (-27.1)	2.4	-27.2 (-33.1)	(M06L opt)	(M06L opt)
M06/aug-cc-pVTZ	-15.2 (-17.7)	2.4	-21.9 (-25.3)	2.4	114.5
M06L/aug-cc-pVTZ	-15.8 (-18.6)	2.4	-23.3 (-26.0)	2.4	115.9
vdW-DF2/cutoff=60Ry	-12.5	2.6	-18.4	2.6	109.9

[a]For all the calculations that employ gaussian basis functions, the energies obtained without counterpoise correction are reported in parentheses. ROS-MP2 calculations without augmented basis function are included to show the variability due to their exclusion in computing interactions.[451] ROS-MP2/ANO-RCC calculations are also compared with CASPT2 results in Section 3.3: for consistency we used the same basis set as BS2, with triple-ζ quality plus polarization on Cu, O, C atoms and double-ζ quality plus polarization on H atoms.

On the basis of quantum calculations, the tilted conformation binding energy is ca. 5.5-7.5 kJ/mol larger than the linear conformation binding energy. The force field model, based on pairwise interactions, underestimates this difference at only 1.3 kJ/mol.

Finally, we tested the possible additive effect on the CO_2 binding energy by adding a second CO_2 molecule bonded symmetrically on the other copper of $Cu_2(formate)_4$. The binding energies computed for this system don't show any significant deviation (-21.0 kJ/mol and -26.9 for the linear and tilted conformations respectively) and therefore any additive effect can reasonably be neglected.

To summarize, vdW-DF2 underestimates the CO_2-$Cu_2(formate)_4$ binding energy by 9.1 and 8.8 kJ/mol, respectively for the linear and tilted configurations, if compared to the ROS-MP2/aug-cc-pVTZ calculations. Considering the ROS-MP2 results, we are now able to improve our model for HKUST-1 and similar copper paddle-wheel MOFs.

4.1.4.3 Multireference Calculations

To have more insight into the interaction between CO_2 and $Cu_2(formate)_4$, we performed wave-function based multireference complete active space calculations, followed by second order perturbation theory.

Several different active spaces were attempted, but only the (2,2) orbitals (Figure 23) were clearly necessary. In terms of energy, the next four highest occupied orbitals and their corresponding unoccupied orbitals were all π/π^* orbitals of the O and C atoms of the paddlewheels, resulting in the (10,10) active space shown in Figure 24. Including only some of them resulted in an arbitrarily imbalanced active space that did not realistically reflect the symmetry of the molecule or the near-degeneracy of the orbitals. Including additional orbitals beyond or instead of those of the (10,10) yielded no interesting results – the occupation numbers of the orbitals were very close to 2 or 0, and

the binding energies were almost always quite close to those calculated with the (2,2) active space. In order to maintain balance in calculating the interaction energies, the isolated CO_2 was calculated with a (0,0) active space, while the $Cu_2(formate)_4$ and $Cu_2(formate)_4$-CO_2 were calculated with either (2,2) or (10,10). CO_2 is a closed-shell singlet, so there is no expectation of any multireference character in the isolated CO_2 subsystem.

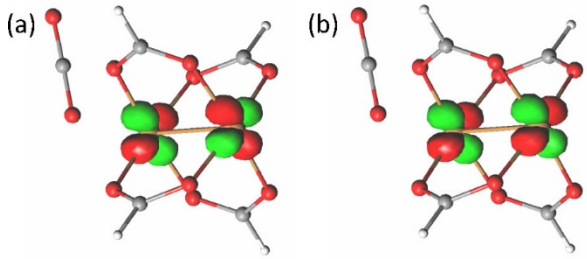


Figure 23: The two molecular orbitals MO $_1$ (a) and MO $_2$ (b), in the tilted di-copper system at equilibrium, with their occupation number in parentheses. In the linear system they look similar. Their occupation number is 1. They correspond to an overall configuration of 0.51 MO $\frac{2}{1}$ + 0.49 MO $\frac{2}{2}$.

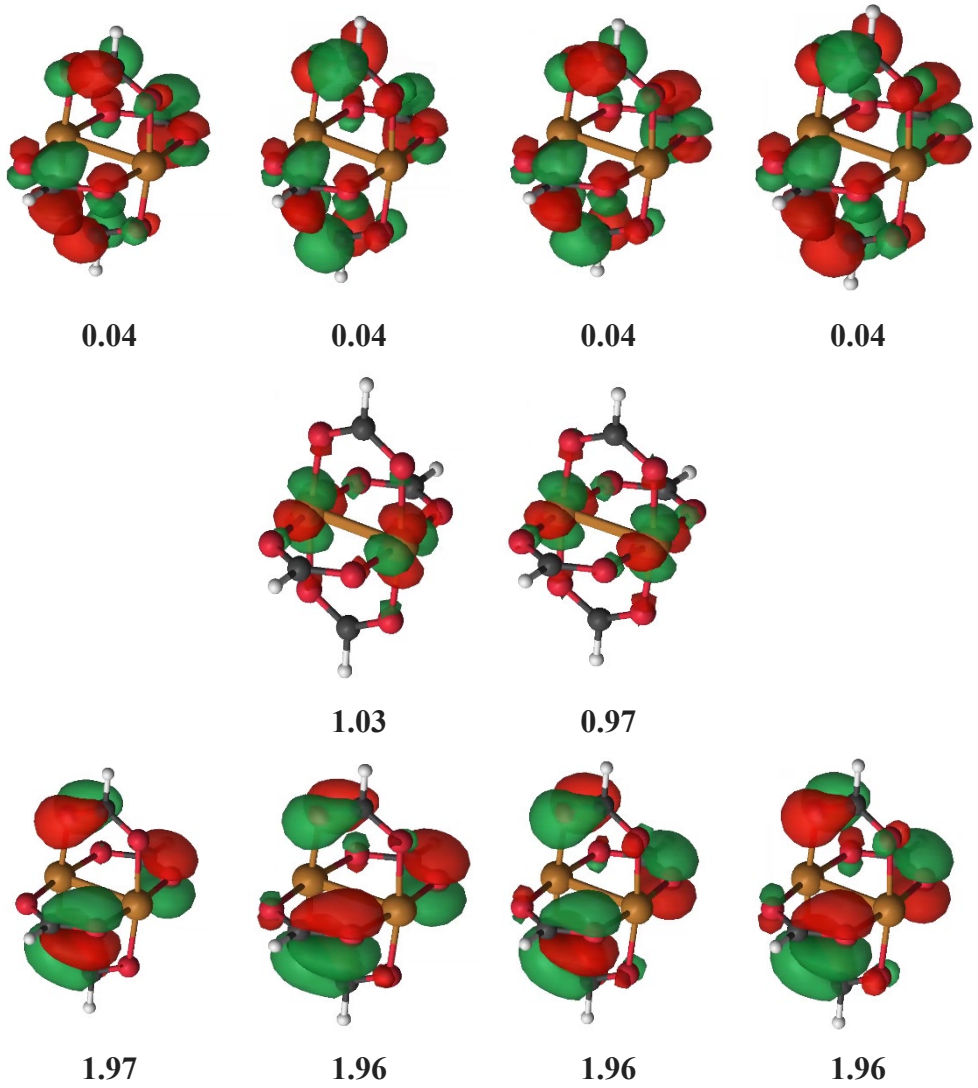


Figure 24: Orbitals of the (10,10) active space, BS2 basis set, with the Cu_2(formate)$_4$. The (10,10) active space looked quite similar for other basis sets and for the Cu_2(formate)$_4$-CO_2 supersystems. Occupation numbers are below each orbital.

The (2,2) CASSCF calculation is equivalent to a restricted open-shell (ROS)-HF calculation, while the (2,2) CASPT2 calculation is equivalent to the ROS-MP2 calculation. Notice that a singlet CASSCF (2,2) active space indeed corresponds to a multireference calculation in the sense that it generates a wave function that is the combination of two configuration state functions (or Slater determinants). Both the

singlet and triplet lowest spin states were explored. In all cases the singlet state is the ground state and it lies 3 kJ/mol lower than the triplet state. This result is in good agreement with the experimental values obtained for MOF-11: 3.4 and 5.3 kJ/mol respectively for the water bound and the anhydrous structure.[452] It is also in good agreement with the 3.2 kJ/mol Maurice et al. calculated with DDCI3 on a similar system, copper acetate monohydrate.[453]

In the following we will discuss the energetics and electronic structure configurations of the singlet. However, as discussed above and also in the literature,[424–426] it is reasonable to expect that the open-shell singlet and the triplet potential energy surfaces have a parallel shape. The singlet state is a linear combination of two electronic configurations with 50% weight each (Table 26). The first configuration corresponds to orbital MO_1 doubly occupied (MO_1^2, Figure 23a) and the second to orbital MO_2 doubly occupied (MO_2^2, Figure 23b). In the (2,2) calculations these orbitals are the only ones included in the active space. They have an average occupation number of about 1 each (because each of them has only a 50% probability of being doubly occupied). In the (10,10) calculation, the other orbitals included in the active space are π and π^* orbitals on the O and C atoms of the paddles. They have occupations of 2 and 0, respectively, within each pair.

Table 26: Wave function description in the case of the (2,2) active space for both the titled and linear systems at equilibrium. Dominant electronic configurations with their weight in the total wave function.

Electronic configuration	% Weight
$MO1^2\ MO2^0$	0.51
$MO1^0\ MO2^2$	0.49

The binding energies are reported in Table 27. In the di-copper system, the binding energy of CO_2 to Cu is significantly larger than in the mono-copper case, as already discussed in Section 3.2. This behaviour can be explained by inspection of the electronic configuration of the Cu_2 system. The two Cu atoms are close enough to have electronic communication and the overall wave function is a superposition of two electronic configurations. A multiconfigurational method is therefore needed to correctly describe this system in the singlet ground state. The mono-copper system, on the other hand, has a single configuration, which is reasonably well described by MP2. The triplet state of the Cu_2 system is also single-configurational.

Table 27: CASPT2 interaction energies (kJ/mol) between $Cu_2(formate)_4$ and CO_2 in linear and tilted conformations for different active spaces and different basis sets for the singlet ground state.[a]

configuration	Active Space	BS1	BS2	BS3
linear	(2,2)	-15.0 (-43.2)	-18.7 (-33.5)	-20.2 (-31.7)
linear	(10,10)	-14.8 (-46.6)	-18.6 (-36.8)	-20.1 (-35.0)
tilted	(2,2)	-17.7 (-49.3)	-23.5 (-40.0)	-25.8 (-39.6)
tilted	(10,10)	-15.8 (-51.1)	-21.8 (-41.7)	-23.9 (-41.1)

[a]The distance between CO_2 and copper is 2.4 for both the linear and the tilted conformations. Values include counterpoise correction. Values without counterpoise correction are in parentheses.

The interaction energies for the singlet and triplet states are very similar (within 1 kJ/mol) and only the singlet energies are reported in Table 27. Our results show that an active space of (2,2) followed by PT2, equivalent to ROS-MP2, is sufficient to describe the binding of this system, as the binding energy does not change by more than 2 kJ/mol when increasing the active space to (10,10).

Basis set effects were explored for the CASPT2 calculations. Table 27 shows that going from BS1 to BS2 the uncorrected binding energy decreases by about 10 kJ/mol, while it remains almost unchanged going from BS2 to BS3. The counterpoise-corrected binding energies change by 3-6 kJ/mol going from BS1 to BS2, while again undergoing little change when going from BS2 to BS3. The CASPT2 results with the (2,2) active space reported in Table 27 should be compared to the ROS-MP2/ANO-RCC (BS2) results reported in Table 25. The only difference between these two sets of results is that those in Table 25 are obtained for the triplet, while those in Table 27 are obtained for the open-shell singlet and with unfrozen 3p orbitals for Cu. The two sets of values including counterpoise corrections differ by less than 2 kJ/mol, and more generally the most accurate CASPT2/BS3 energies agrees well with the ROS-MP2/aug-cc-pVTZ values, especially in the linear conformation (difference of 1.5 kJ/mol).

4.1.4.4 Correction of the Force Field

In order to model properly the interaction of the carbon dioxide with the open metal site in a classical force field, we needed to correct the potential energy curve based on our first principle calculations. The most representative path for different CO_2-Cu distances is the one where the energy is mainly influenced by the interaction with the cation rather than the interaction with other atoms of the cluster (or framework). Hence, we fitted the linear CO_2-Cu_2(formate)$_4$ curve obtained with the ROS-MP2/aug-cc-pVTZ method to obtain the new parameters for the force field. Only the Cu-O van der Waals potential was tuned, while keeping the standard UFF parameters for all other atoms pairs and REPEAT (PBEsol derived) point charges to model electrostatic interactions. For the Cu-O interaction, a Buckingham potential was adopted to correctly represent the repulsion at short distance and an r^{-8} attractive term was added to account for the

stabilization observed in the ROS-MP2 calculations. The details about the fitting and the coefficient for the Cu-O potential are reported in Supporting Information. The optimal CO_2 interaction with $Cu_2(formate)_4$, which corresponds to the tilted conformation, computed with the fitted force field parameters has a value of -23.2 kJ/mol. This result is consistent with the UFF difference between the linear and tilted configurations of -1.3 kJ/mol. We notice that by applying this relatively simple but effective correction, obtained without modifying the pairwise interaction with other atoms and without introducing a specific contribution based on the Cu-CO_2 angle, the minimum interaction energy obtained for $Cu_2(formate)_4$ is in fair agreement with the ROS-MP2 result of -27.2 kJ/mol.

Finally, we replicated the GCMC simulations in HKUST-1, using our UFF parameters with the corrected Cu-O potential. The comparison with experimental data is reported in Figure 25. The simulations are still slightly underestimating the measured uptake, and this is reflecting the previously mentioned underestimation of ca. 4 kJ/mol for the interaction energy in the optimal tilted configuration. However, the assumptions made for the force field are sufficient to obtain a good representation of the uptake around ambient temperature, and a significant improvement with respect to the standard force fields.

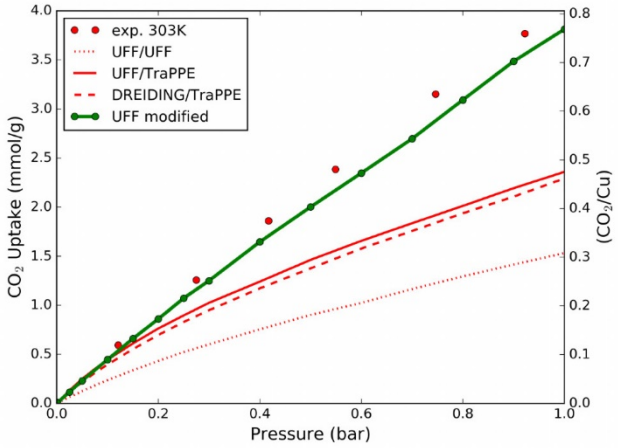


Figure 25: Comparison between the experimental[422] and simulated isotherms for CO_2 inside HKUST-1 at 303K. The modified UFF forcefield is obtained by fitting the Cu-O potential on ROS-MP2 calculations.

The minimum energy of interaction computed with our new force field in the three main adsorption sites of HKUST-1, i.e. open metal, small pore window, and small pore center sites, are now ranked correctly: -27.3 kJ/mol, -26.8 kJ/mol, and -26.8 kJ/mol respectively, and the OMS stability is not underestimated any longer compared to the *in situ* experimental results.

As a starting point for our correction, we used UFF/UFF mixed parameters instead of UFF/TraPPE or DREIDING/TraPPE, because from the simulated isotherm (Figure 25) we can observe that these last force fields are already predicting the experimental uptake at very low pressure (below 0.1 bar), even if the open metal site interaction is strongly underestimated. This is an artifact due to a fortuitous error cancellation with the overestimation of the interaction in other sites, i.e. the small pores centers (see Table 23), which are already saturated at 0.82 mmol/g, as clearly shown by the deviation of the simulated isotherm from the experimental one. Therefore, employing the conventionally used UFF/UFF, UFF/TraPPE or DREIDING/TraPPE mixed parameters to describe the guest-host interaction in an analysis of the site occupancy

would lead to wrong conclusions, i.e. that in HKUST-1 the open metal sites are very poorly occupied at low uptake.[454] Ulterior comparisons with experimental data is provided in the Supporting Information: CO_2 uptake at higher pressure and different temperatures,[455] and the heat of desorption in function of the uptake.[422,456,457]

4.1.4.5 Investigation of the "Double" Open Metal Site Interaction in Cu-TDPAT

To further test the reliability of our force field, we investigated another interesting copper paddle-wheel metal organic framework, Cu-TDPAT, firstly synthesized by Li *et al.*[458] The crystalline structure is characterized by the presence of strong adsorption sites for CO_2, where both oxygens of the guest molecule are attracted to two different copper cations (Figure 26), leading to an interaction energy which is roughly double with respect to the conventional single open metal site of copper paddle-wheel.

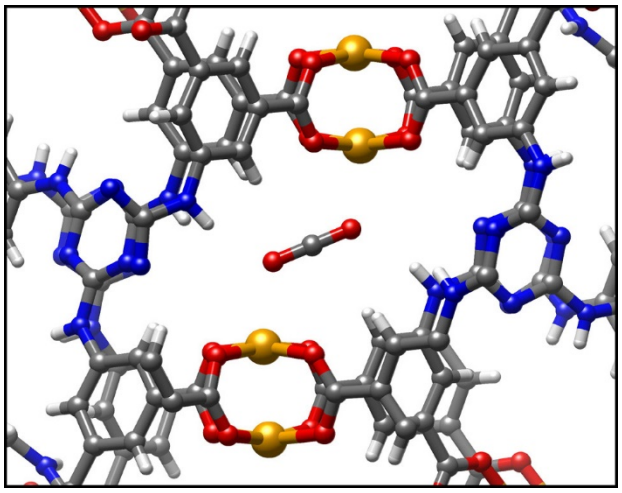


Figure 26: CO_2 molecule adsorbed in the double open metal site of Cu-TDPAT.

Because of this reason, Cu-TDPAT is one of the top performing MOFs for both gravimetric and volumetric CO_2 uptake at ambient pressure.[459] The conventional unit cell of Cu-TDPAT contains 48 copper cations: 24 of them compose 12 double open metal sites while the remaining 24 atoms compose 24 single open metal sites, with a conformation very similar to the OMS of HKUST-1. Due to the large dimension of the unit cell (960 atoms), the crystal is too big to perform a DFT calculation with an accuracy comparable with our previous calculation on HKUST-1. Consequently, we employed the Extended Charge Equilibration (EQeq) method[460] to compute the partial charges of the framework. This method is able to self-consistently compute point charges for MOFs, with results very similar to the charges obtained by fitting the electrostatic potential from a quantum calculation, e.g. REPEAT. HKUST-1 itself was successfully tested in the original paper presenting the EQeq method.[460] Compared to the quantum electrostatic potential fitting, this method is drastically faster (a few minutes instead of hours for HKUST-1) and is applicable to a unit cell containing a large number of atoms, which is practically forbidden to DFT calculations. The result obtained for the copper paddle-wheel is consistent with our PBEsol calculation in HKUST-1. Using the EQeq method the average point charges for Cu-TDPAT are 0.905 and -0.398 for the copper and the carboxylic oxygen respectively, versus 0.914 and -0.57 for HKUST-1 computed using REPEAT. With the new set of parameters, we compared the results of the GCMC simulations to experimental data (Figure 27).

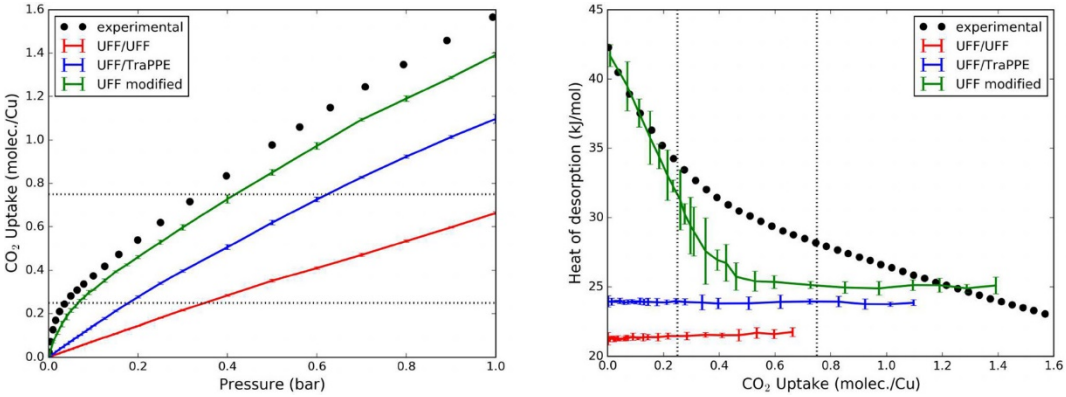


Figure 27: Comparison of experimental[458] and simulated adsorption of CO_2 in Cu-TDPAT at 298K using different set of parameters. The force field developed in this method is reported as "UFF modified", while UFF/UFF and UFF/TraPPE are conventionally used standard sets of parameters. In both plots the uptake is converted to CO_2 molecules per copper ratio, and the equivalence to the number of double open metal sites (0.25 CO_2/Cu) and the number of total open metal sites (0.75 CO_2/Cu) is highlighted with a dotted line. The experimental heat of desorption (black dots, right picture) has been computed through Langmuir-virial method while the simulated values (colored lines) are computed from the guest molecules number fluctuation in the GCMC simulation.

This comparison shows a good agreement, as for HKUST-1, which gives us some confidence in the transferability of our force field to model CO_2 adsorption. Moreover, it becomes more evident how UFF/UFF and UFF/TraPPE parametrizations do not capture the strong interaction between CO_2's oxygen and copper.

4.1.5 Conclusions

In this work we have shown that the Cu-Cu interaction in copper paddle-wheel systems is the reason why DFT methods, even when they include dispersion corrections, systematically underestimate the interaction between CO_2 and copper paddle-wheel motif. Our calculations confirm the presence of the copper-copper coupling, influencing the attraction of the CO_2, and suggest that the mono copper cluster $Cu(formate)_2$ is not a realistic model to describe this interaction.

One thus needs an electronic structure theory that properly describes the Cu-Cu interaction, such as the ROS-MP2 wave function. We show that if this interaction is included in our calculations, the prediction of the binding energies is in better agreement with the experimental data. To justify the choice of ROS-MP2 method, which is equivalent to a (2,2) CASPT2 calculation, we performed a number of multireference calculations over a variety of active spaces, basis sets, and spin states. We concluded that the ROS-MP2 level of theory is good enough to model the $Cu_2(formate)_4$-CO_2 interaction.

Using the ROS-MP2 results, we reparametrized the UFF pairwise potential to correctly model the interaction of CO_2 with the open metal site in HKUST-1, which was severely underestimated by conventional force fields. The results obtained from our new force field agree with experimental isotherms as well as with *in situ* PXRD studies, which found the open metal site to be the strongest adsorption site for CO_2 rather than the small pore center site. The correction proposed in this work acts in proximity of the open metal site, and this local character of the correction terms allows us to transfer the same parameters to other MOFs containing the copper paddle-wheel motif.

To test this transferability, we employed our force field to model the CO_2 interaction with the “double” open metal sites of Cu-TDPAT framework, and we showed

a significant improvement with respect to conventional UFF parameters in this case as well. The modified set of parameters proposed makes it possible now to accurately describe the adsorption behavior for this class of MOFs. In this study we have focused on CO_2, but similar effects can be expected for other polar molecules.

4.1.6 Additional Information

Supporting Information: Supplementary results for the multireference calculations. Details of the classical simulations, coefficients of the force field and comparison between the ROS-MP2 and the fitted Cu-O potential. Atomic coordinates for the cluster models. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.7b02302.

The research of D.O., D.T. and B.S. was partially supported by the National Center of Competence in Research (NCCR) “Materials’ Revolution: Computational Design and Discovery of Novel Materials (MARVEL)” of the Swiss National Science Foundation (SNSF) and by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 666983, MaGic). Part of the computational work was made possible through a grant from the Swiss National Supercomputing Center (CSCS) under Project no. s611. S.J.S. and L.G. were supported as part of the Inorganometallic Catalyst Design Center, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, under Award DE-SC0012702.

4.2. Catechol-Ligated Transition Metals: A Quantum Chemical Study on a Promising System for Gas Separation

This section describes the outcome of a collaborative research project carried out by Samuel J. Stoneburner, Vanessa Livermore, Meghan E. McGreal, Decai Yu, and Konstantinos D. Vogiatzis, (and advised by Randall Q. Snurr and Laura Gagliardi). A report on this research project has been published.[461,462]

Samuel J. Stoneburner performed the DFT calculations, most of the CASSCF/CASPT2 calculations, and wrote the manuscript.

4.2.1 Overview

Metal-organic frameworks (MOFs) have received a great deal of attention for their potential in atmospheric filtering, and recent work has shown that catecholate linkers can bind metals, creating MOFs with monocatecholate metal centers and abundant open coordination sites. In this study, M-catecholate systems (with M = Mg^{2+}, Sc^{2+}, Ti^{2+}, V^{2+}, Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}) were used as computational models of metalated catecholate linkers in MOFs. Nitric oxide (NO) is a radical molecule that is considered an environmental pollutant and is toxic if inhaled in large quantities. Binding NO is of interest in creating atmospheric filters, both at the industrial and personal scale. The binding energies of NO to the metal-catecholate systems were calculated using density functional theory (DFT) and complete active space self-consistent field (CASSCF) followed by second order perturbation (CASPT2). Selectivity was studied by calculating the binding energies of additional guests (CO, NH_3, H_2O, N_2, and CO_2). The toxic guests have stronger binding than the benign guests for all metals studied, and NO has significantly stronger binding than other guests for most of the metals studied, suggesting that metal-catecholates are worthy of further study for NO filtration. Certain metal-catecholates also showed potential for separation of N_2 and CO_2 via N_2 activation, which could be relevant for carbon capture or ammonia synthesis.

4.2.2 Introduction

Nitric oxide (NO) is a radical molecule known to have an important signaling role in biology.[463] However, it is toxic in large quantities, with effects including DNA degradation and/or possible mutagenesis and lung damage.[464] NO can also do indirect damage by reacting with atmospheric H_2O or NH_3 to form HNO_3,[110] with O_2 to form NO_2, and with superoxide anions to form peroxynitrite,[464] all of which are toxic when inhaled. Personal protection typically takes the form of a gas mask, with underlying technology that is still fundamentally the same a century after its inception.[465] Most of the recent research into NO capture has been in contexts that do not readily relate to personal protection, namely, flue gas NO capture and drug delivery. In flue gas cleansing, NO is typically converted to the more soluble NO_2, which is useful in an industrial setting.[466,467] However, the greater toxicity of NO_2 makes this approach unsuitable for a personal filter. Those seeking to use NO medicinally typically desire weak binding that allows NO to be replaced with water,[468] but in a gas mask the toxic guests should bind *preferentially* over water.

Gas mask filters typically contain activated carbon impregnated with additional materials for specific functionality, such as metal salts, acids, and amines.[110,465] The binding of these materials is fairly loose, and over time impregnates with complementary functionalities, such as acids and bases, interact and render the filter less effective.[110] Therefore there is a need for a new class of support that can provide a variety of active sites in a rigid structure. Ideally such a structure would include uncoordinated metal sites which would then be able to bind NO irreversibly.

Metal-organic frameworks (MOFs) are a family of hybrid materials constructed from inorganic building units and organic linkers. They typically possess a rigid structure that provides high porosity, which is advantageous for filtering gases.[94,96] There are a

wide variety of possible building blocks, which enable fine-tuning of application and functionality.[117] Often MOFs have metal centers with open coordination sites, which provides many unique opportunities for chemical binding and catalysis.[28] MOFs have already been considered for filtering a variety of airborne toxins, including carbon monoxide, hydrogen sulfide, and ammonia.[110] Previous work with NO binding to MOFs has primarily been in other contexts. In CO_2 separation, NO has been studied as a potential contaminant in a variety of functionalized IRMOF ligands,[469] in MIL-101 (MIL = Materials Institute Lavoisier),[470] MOF-5,[471] M-MOF-74,[470,472,473] and MOF-177.[471] Biomimetic redox reactions with NO using Fe^{2+}-MOF-5 have been studied,[474] but these would yield the even more toxic NO_2. NO's importance as a biosignalling molecule has also led to an interest in using MOFs to deliver NO within the human body. Approaches have included binding to HKUST-1,[475,476] MIL-88(Fe),[468] M-MIL-100,[477] and M-MOF-74,[115,477,478] as well as postsynthetic incorporation of diazeniumdiolate to a variety of MOFs in attempts to chemically trigger NO release.[479,480] Most MOFs considered so far have been found to bind water favorably over NO,[110] and they would therefore not be suitable for gas masks, as any toxin that can be displaced by water would not be effectively filtered.

MOFs have also been considered for CO_2/N_2 separation. Most research has been driven by interest in CO_2 capture or reactivity, in which case it is desirable to bind CO_2 and allow N_2 to pass through. Examples include UiO-66 (UiO = University of Oslo),[481] [Cu(bcppm)H_2O] (bcppm = Bis(4-(4-carboxyphenyl)-1H-pyrazolyl)methane),[482] certain members of the bio-MOF series,[483] NH_2-MIL-53(Al),[484] and others.[485–491] However, if N_2 activation is of interest, for example as part of the catalytic production of ammonia,[492,493] it can be desirable for N_2 to bind more strongly than CO_2. There are not many cases in the literature of MOFs that activate N_2, although it has been predicted that V-MOF-74 would be capable of doing so.[106]

One of the primary concerns in using MOFs for gas filtration is accessibility at the metal site.[110] A possible solution is to metalate the organic linkers, resulting in an open metal site that reaches into the pore of the MOF. Nguyen and coworkers successfully modified CatBrO-MOFs with V^{IV} ions and demonstrated that the vanadyl monocatecholate group can perform oxidation chemistry.[494] Post-synthesis modification of the highly robust UiO-66 MOF to create a material exhibiting catecholate groups was performed by Cohen and coworkers.[495] Metalation of the new MOF (referred to as UiO-66-Cat) with Cr^{3+} and Fe^{3+} cations resulted in active metal centers that can perform oxidation of alcohols. UiO-66-Cat with Ti^{4+} was synthesized and studied for cyclohexene oxidation.[496] Metalation of catechol-functionalized porous organic polymers with Mg^{2+}, Mn^{2+}, Cu^{2+}, Zn^{2+}, and Ta^{5+} produced materials with enhanced heats of hydrogen adsorption or high uptake of ammonia.[129,497,498] There have also been several computational studies of hydrogen binding to metal catecholates and related species.[499,500] A recent study on metalated MOF linkers for CH_4 storage included Mg- and Ca-catecholate linkers. The authors concluded that a bent Mg-catecholate geometry enables the greatest number of CH_4 molecules to be adsorbed to a single site.[501]

As monocatecholate low-valence metal centers are scarce in the literature, metalation of catecholate-based MOF linkers presents an unprecedented opportunity for exploration of low-coordinate environments. A notable exception to this scarcity is the study of Abramov et al., which used EPR spectroscopy to monitor NO binding to Zn- and Pb-catecholato complexes.[502] A fast screening of combinations between different metal-catecholates and gas molecules can be performed by electronic structure calculations. Kim *et al.* conducted computational studies on the relative binding of ammonia and water using a wide variety of metal catecholates, finding that several mid to late transition metals strongly preferred ammonia over water.[104] In this work, we have computationally studied the binding of NO to a series of first-row transition metal ions, as well as Mg^{2+},

bound to a catecholate using density functional theory (DFT), coupled-cluster and multireference calculations. NO is a doublet radical and is therefore open-shell, as are some of the metal centers under consideration. These systems have many close-lying electronic states and are likely to pose challenges to single-reference methods such as DFT, and so multireference treatment becomes necessary. We have also considered selectivity by calculating the binding energies of several additional guests. CO and NH_3 are other toxic guests of potential interest for this application. H_2O, as mentioned, needs to have weaker binding than the toxic guests of interest for a material using catecholates to be useful in a gas mask. Similarly, N_2, and CO_2 will be present in large quantities in any realistic conditions, and so they also must have relatively weak binding.

4.2.3 Computational Methods

Catecholate (sometimes abbreviated as "cat" throughout the text) was chosen as a model system for the catecholate linkers in similarity to previous work.[104] While Maihom *et al*.[503] used a larger model, their catecholate linker was shorter than those which have served as the inspiration for this work, and was more susceptible to longer range interactions The cat-metal-guest supersystem (Figure 28) and the separate subsystems (the bare cat-metal and the guest alone) were optimized using M06-L,[504] M06,[433] PBE,[82] and PBE0[283] for M = Mg^{2+}, Sc^{2+}, Ti^{2+}, V^{2+}, Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}. PBE and PBE0 calculations were performed with Grimme's D3(BJ) dispersion correction.[414] Vibrational frequency analysis was performed to confirm that the converged solutions were local minima. Gaussian09[282] with the def2-TZVPP basis set[505] was used for all DFT calculations. For reference, coupled-cluster calculations with interference effects[506,507] were performed on the PBE0-D3(BJ) geometries of cat-Mg, CO, NH_3, H_2O, N_2, CO_2, and their respective supersystems. This choice was based on a

recent study[508] where PBE0-D3(BJ) was found to provide accurate interaction energies between closed-shell transition metals and small gas molecules. The benchmark study revealed that M06 had performed better for the systems in question, and so M06 geometries were used for subsequent single-point complete active space self-consistent field (CASSCF)[6] followed by second-order perturbation theory (CASPT2)[9] calculations. Molcas 7.8[509] and the ANO-RCC-VTZP[264] basis set with Cholesky decomposition[185] were used for all CASSCF/CASPT2 calculations. The default value of the IPEA shift[65] (0.25 eV) and an imaginary shift of 0.1 a.u were used in all CASPT2 calculations. Details of the benchmarking calculations and geometry choices are presented in the SI.

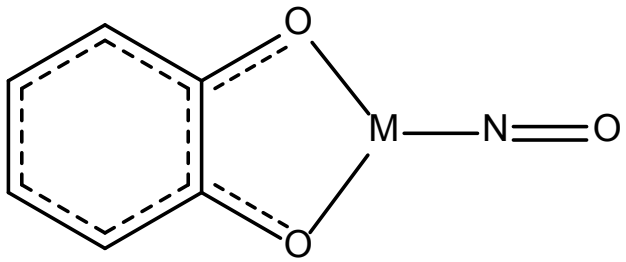


Figure 28. NO bound to the cat-M complex.

The free energy of adsorption of the guest to the cat-M was calculated as the energy of the supersystem minus the energies of the bare cat and the isolated guest:

$$\Delta G_{ads} = G_{cat\text{-}M\text{-}guest} - G_{cat\text{-}M} - G_{guest} \quad (77)$$

Free energies at 298.150 K and 1 atm are presented, but electronic energies, zero-point energy corrections, and enthalpies are available in the SI. To ensure that basis set superposition error (BSSE)[436] was not a significant factor, the CASPT2 binding energy for NO to cat-Mn was also calculated using ANO-RCC-QZVP. The results differed by only 0.5 kcal/mol from those with the ANO-RCC-VTZP basis, and so it was concluded that BSSE was minimal and counterpoise corrections were not added.

For all cases other than Co, the active spaces used in the CASSCF/CASPT2 calculations included the catecholate carbon-centered π-system orbitals and any orbitals of intermediate occupancy (e.g., singly occupied 3*d* orbitals). For Co, all 3*d* orbitals were included due to changes in *d*-orbital occupancy and mixing. The details of active space choices are available in the SI.

4.2.4 Results and Discussion

M06 results were found to be closest to the coupled-cluster reference for Mg, but M06-L and PBE-D3 had better agreement with CASPT2 overall (Tables S11-S13, SI). M06-L suffered from convergence difficulties with certain systems (see SI), so PBE-D3 is presented throughout the paper.

Free energies of adsorption of the guests to the M-cats are presented in Figure 29. It should be noted that while a previous MP2 study predicted much higher energies for cat-Cu binding to both H_2O and NH_3 than we present here,[104] the relative difference between these two guests is about the same. Therefore, the point made in that study regarding the cat-Cu system's strong selectivity for NH_3 over H_2O still holds. The discrepancies in individual energies are likely due to the highly multiconfigurational nature of the cat-Cu system, which cannot be described at the MP2 level. It should be noted that Kohn-Sham DFT, which we used in this work, does not necessarily suffer from the same difficulty. Despite KS-DFT being formally a single-reference method, the exact density functional would correctly account for *all* electron correlation.[28] In practice, then, a given density functional may perform well or poorly on a given multireference system depending on how well it approximates the exact functional in that particular case.

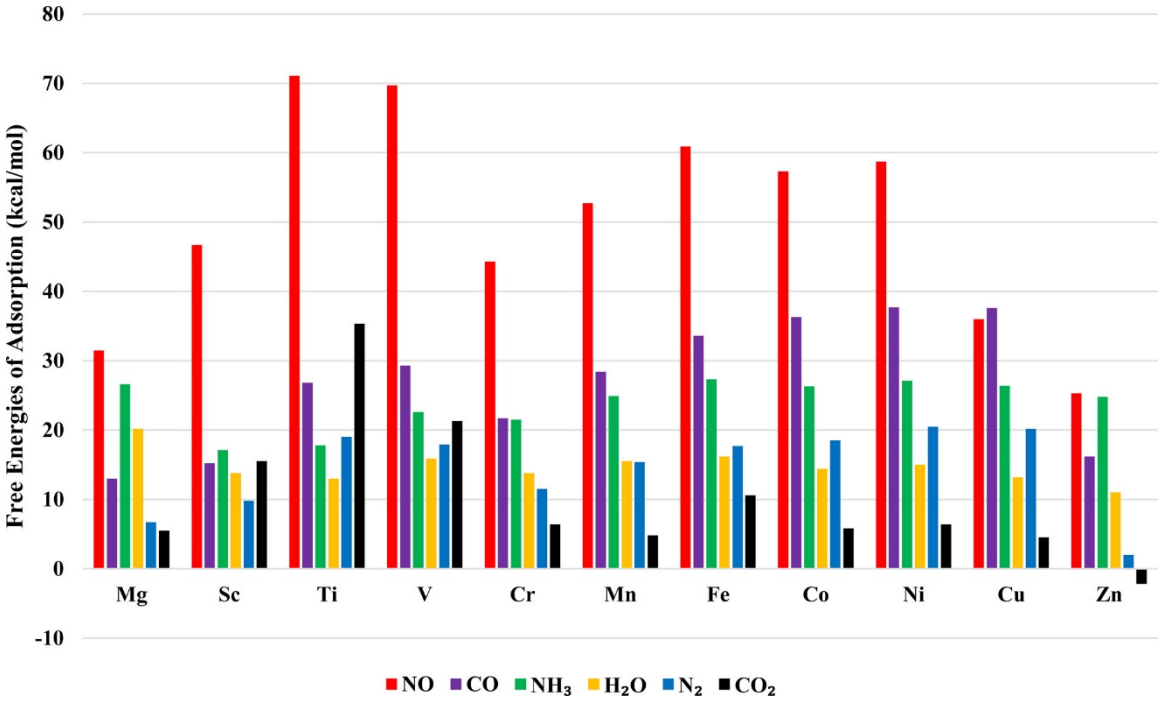


Figure 29. PBE-D3(BJ) free energies of adsorption (kcal/mol) of guests to cat–M systems. Signs are reversed for ease of viewing.

4.2.4.1 NO Binding and Selectivity

Table 28. PBE-D3(BJ) and CASPT2[a] free energies of adsorption (kcal/mol) of NO to cat-M systems.

	PBE-D3	CASPT2
Mg	-31.5	-30.6
Sc	-46.7	-53.0
Ti	-71.1	-65.7
V	-69.7	-72.6
Cr	-44.3	-37.3
Mn	-52.7	-49.7
Fe	-60.9	-57.0
Co	-57.3	-77.1
Ni	-58.7	-45.7
Cu	-36.0	-39.1
Zn	-25.3	-30.6

[a]CASPT2 free energies were obtained by adjusting the electronic energies by the PBE-D3(BJ) thermal free energy corrections.

NO free energies of adsorption (Table 28) are calculated to be greater than all others by at least 10 kcal/mol for all metals studied other than Mg, Cu, and Zn. The strong binding of NO is due to π-backbonding from 3*d* orbitals on the metal into the π* orbitals of the NO (Figure 30). In some cases, the catecholate also reduces the metal (Figure 31) and provides an additional electron available for the π-interaction with the NO. The binding motif shown is typical for most metals. Mg and Zn do not have the same availability of 3*d* orbitals, and their interaction is more of σ character. The Cu^{2+} cation in cat-Cu, being reduced by the catecholate, has a full 3*d* subshell similar to Zn, and so it also has weaker NO binding. Additionally, cat-Ti and cat-V have metals bent out of the plane of the catecholate, and the NO was bound in a side-on fashion (Figures S18 and S19), rather than the end-on seen for the other metals. In those cases, binding was a combination of π and δ-interaction (Figure 32).

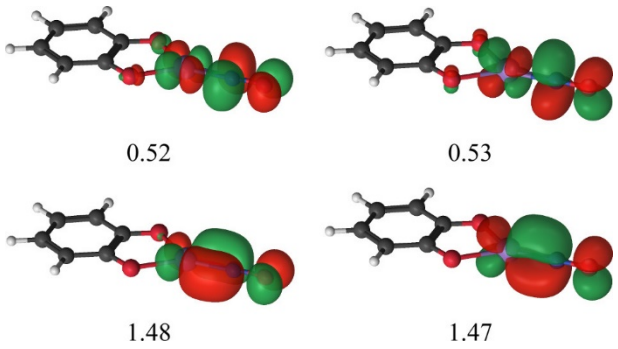


Figure 30. π (bottom) and π* (top) of 3*d*-Mn and NO π*CASSCF orbitals in cat-Mn-NO. Occupation numbers are below each orbital. The intermediate occupancy of the orbitals demonstrates the multiconfigurational nature of these systems.

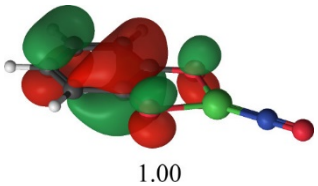


Figure 31. Singly occupied cat CASSCF orbital responsible for reducing the metal.

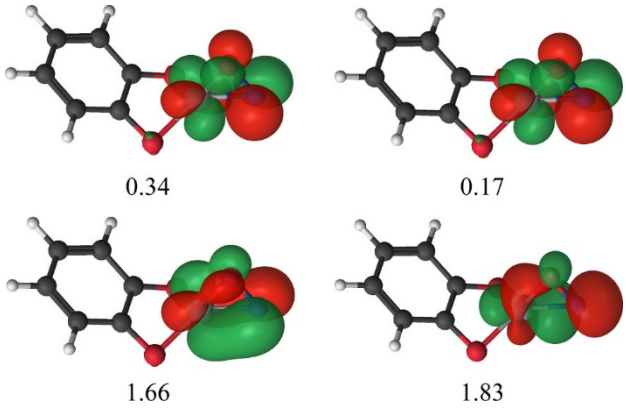


Figure 32. δ (bottom left), δ* (top left), π (bottom right), and π* (top right) of 3*d*-Mn and NO π* CASSCF orbitals in cat-V-NO. Occupation numbers are below each orbital.

4.2.4.2 N_2 and CO_2 Binding

Table 29. PBE-D3(BJ) and CASPT2[a] free energies of adsorption (kcal/mol) of N_2 and CO_2 to cat-M systems.

	N_2		CO_2	
	PBE-D3	CASPT2	PBE-D3	CASPT2
Mg	-6.7	-2.4	-5.5	-13.2
Sc	-9.8	-9.5	-15.5	-21.5
Ti	-19.0	-15.5	-35.3	-38.7
V	-17.9	-16.4	-21.3	-22.0
Cr	-11.5	-9.8	-6.4	-12.8
Mn	-15.4	-9.0	-4.8	-10.6
Fe	-17.7	-13.3	-10.6	-15.9
Co	-18.5	-19.1	-5.8	-3.9
Ni	-20.5	-27.1	-6.4	-12.4
Cu	-20.2	-24.4	-4.5	-8.6
Zn	-2.0	-14.0	2.2	-18.4

[a]CASPT2 free energies were obtained by adjusting the electronic energies by the PBE-D3(BJ) thermal free energy corrections.

N_2 was found to have significantly (> 5 kcal/mol) larger free energies of adsorption than CO_2 for Co, Ni, and Cu (Table 29). For the majority of the metals studied, N_2 was found to have a binding energy between -15 and -21 kcal/mol. The strength of the metal-N_2 bond is assigned to N_2 activation and backbonding.

The N_2 stretching frequency was reduced by 148, 90, and 65 wavenumbers for Co, Ni, and Cu, respectively, compared to the free N_2. While it may seem counter-intuitive that the N_2 stretching frequency is highest for Co, which had the lowest binding energy, the difference in binding energies across these three metals is only two kcal/mol, and it would be hard to justify any particular trends in stretching frequency changes across a difference that small. The changes in stereching frequency are comparable to those found by Lee et al. in V-MOF-74, which was also calculated to backbind N_2.[106] It was also found in that paper that while V-MOF-74 is expected to activate N_2, Fe-MOF-74 is not, while in this study we found that cat-V and cat-Fe are both expected to activate N_2. The difference is due to the coordination environments: In M-MOF-74, the metal is hexacoordinated, with the N_2 being the sixth ligand. In the case of Fe-MOF-74, there is a σ* antibonding interaction between the N_2 and the $3d_{z^2}$ orbital of the Fe such that N_2 cannot get close enough to backbond. In the case of the cat-Fe we present here, Fe is only tricoordinated, the bonding is of similar fashion to the NO discussed above, and the $3d_{z^2}$ orbital is non-interacting.

CO_2 typically has more limited orbital interaction with the metal. In most cases the optimized structure is a linear CO_2 in the plane of the catecholate. The primary exceptions are Sc and Ti, for which CO_2 takes on a bent configuration perpendicular to the plane (Figures S39 and S40).

4.2.5 Conclusions and Outlook

In the interest of finding better potential materials for gas mask filters, the binding of NO to first-row transition metal catecholates was studied using electronic structure theory calculations. Metal catecholates can be formed in the frameworks of porous materials by post-synthetic modification. They have a low coordination number, which allows them to potentially bind multiple gas molecules and enhance the sorption performance of the material. Computational work showed strong NO binding for most of the first-row transition metal ions, which was assigned to two π-bonding interactions between 3*d* orbitals of the metal and the π* orbitals of the NO. Mg, Cu, and Zn were found to have weaker interaction due to a lack of partially occupied 3*d* orbitals available for the π interaction.

Selectivity was also addressed by calculating binding energies for other toxic guests (CO and NH_3), and for benign guests (H_2O, N_2, and CO_2). Toxic gases had stronger binding than benign gases in many cases, but only NO had stronger binding than all benign gases for all metals studied. In particular, NO had stronger binding than H_2O, which is the opposite of what has been found in most of the MOFs previously studied.

Several systems were found to have potential for N_2/CO_2 separation, namely, Co-, Ni-, and Cu-catecholates, which are predicted to show preferential adsorption of N_2, i.e. the inverse of the selectivity displayed by most adsorbents.

The relationship identified between the electronic configuration of the transition metals and the binding energy of the catecholates to NO and other guests offers broad opportunities for discovery of new materials for gas separation guided by theory.

4.2.6 Additional Information

Supporting Information: Additional information about the optimized geometries, spin states, and active spaces. All free energies and enthalpies of adsorption. All electronic binding energies with and without zero-point energy corrections. All absolute energies used in calculation of the binding energies. QZVP basis set comparison details. Coupled-cluster reference calculation details. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.7b02685.

The authors thank Dale Pahls for useful discussion and Gary Bondarevsky for insight in preliminary exploration. The authors also acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. URL: http://www.msi.umn.edu. This research is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-12ER16362.

4.3. Air Separation by Catechol-Ligated Transition Metals: A Quantum Chemical Screening

Reproduced with permission from **Stoneburner, S. J.**; Gagliardi, L. Air Separation by Catechol-Ligated Transition Metals: A Quantum Chemical Screening. *J. Phys. Chem. C* **2018**, *122* (39), 22345–22351.[326]

4.3.1 Overview

The separation of O_2 and N_2 from air is of great importance in a variety of industrial contexts, but the primary means of accomplishing the separation is cryogenic distillation, an energy-intensive process. A material that could enable air separation to occur at conventional temperatures would be of great economic and environmental benefit. Metalated catecholates within metal-organic frameworks have been considered for other gas separations and are shown here to have significant potential for air separation. Calculations of interaction energies between catecholates with first-row transition metals and guests O_2 and N_2 were performed using density functional theory and multireference complete active space self- consistent field followed by second order perturbation theory. A general recipe is offered for active space selection for metalated-catecholate systems. The multireference results are used to rationalize O_2 binding in terms of redox activity with the metalated catecholate. O_2 is predicted to bind more strongly than N_2 for all cases except Cu^{2+}, with general agreement in the binding trends among all methods.

4.3.2 Introduction

The separation of N_2 and O_2 into pure gases is critical for a wide variety of applications. Nitrogen is used as a feedstock in ammonia synthesis. It is also employed in many contexts as an inert atmosphere, and as a refrigerant in its liquid form.[510] Oxygen is used by the steel industry in massive quantities, but it is also an important reactant in the manufacturing of a wide variety of chemicals.[511,512] Moreover, oxygen is a key component of oxy-fuel combustion, which improves efficiency and greatly reduces the NO_x pollutants in flue gas.[513,514] Oxy-fuel combustion is also an approach to carbon capture.[515,516] The use of oxy-fuel eliminates the need to separate N_2 from CO_2 post-

combustion, which is the greatest difficulty in standard post-combustion carbon capture.[517] The current primary method for obtaining pure N_2 and O_2 gases is cryogenic distillation, an energy-intensive process.[130] An oxy-fuel CO_2 capture system using cryogenic air separation is expected to increase fuel consumption by up to 35%,[518] rendering O_2 prohibitively expensive for oxy-fuel-based carbon capture.[517] An approach that would enable air separation at more convenient temperatures would make oxy-fuel more economically feasible and reduce the cost of many other processes and products. Alternative methods to cryogenic distillation largely consist of adsorption methods featuring membranes or zeolites, but current materials do not yield high enough purity for many industrial applications.[519,520]

Metal-organic frameworks (MOFs) have been the focus of considerable research in the quest for improved materials for applications such as catalysis, gas storage, and gas separations.[28,108] MOFs are nanoporous materials constructed from metal-containing nodes and organic linkers, with a wide variety of possible components and potential structures.[94] While much research has been devoted to MOFs useful for CO_2-related separations,[28] research on MOFs for air separations has been limited so far.[517,521] Certain MOFs have shown some promise in O_2/N_2 selectivity, especially $M_3(btc)_2$ (btc= 1,3,5-benzenetricarboxylate)[513,521] and $M_2(dobdc)$ ($dobdc^{4-}$ = 2,5-dioxido-1,4-benzenedicarboxylate).[513,522,523] It has been found that early first-row transition metals have better O_2/N_2 selectivity in these MOFs,[513] although $Fe_2(dodbc)$ has also received special attention.[522] However, these MOFs have been criticized as being unstable in air due to oxidation of their open metal sites.[524,525] MIL-101 (1,4-benzenedicarboxylate linkers),[524] Sc-MIL-100 (benzenetricarboxylate linkers),[520] Co-BTTri (H3BTTri = 1,3,5-tri(1,2,3-triazol-5-yl)benzene), and Co-BDTriP (H3BDTriP = 5,5′-(5-(1H-pyrazol-4-yl)-1,3-phenylene)bis(1H-1,2,3-triazole))[525] are more stable and also show good selectivity, although MIL-101 had notably reduced capacity in comparison to $Fe_2(dodbc)$.[524] Cr_^{t}Bu-

bdc (bdc^{2-} = benzene- 1,3-dicarboxylate) binds O_2 much more strongly than N_2, but the O_2 binding is irreversible.[526]

Metal-catecholates are formed in MOFs by deprotonating a catechol or similar group and adding a metal ion.[494] We recently reported a screening of metal-catecholates (or cat-M, where M is a metal ion) for NO filtration.[461] Metal-catecholates and other related systems have been studied for applications in gas separation,[104,129,527,528] gas storage,[497,499–501] and catalysis,[494–496,498,503,529–531] but to the best of our knowledge this is the first time they have been considered for air separation.

We employ a computational screening process to determine which metals are most favorable for the desired separation. In so doing we not only save considerable time and effort in comparison with experimentally synthesizing and testing each metal, but we also obtain insight into the electronic structures of these systems and how they drive differences in binding. This insight can be used to predict new favorable separations. In order to facilitate comparisons across the period, we considered only the 2+ oxidation state for each metal. After a general explanation of the computational methods, we present a recommendation for selecting active space orbitals for the systems studied in this work. We continue with an analysis of the differences in O_2 and N_2 binding and what we believe to be the causes. We finish with some brief comments about how these results apply to NO filtration and the work we previously published in that context.[461]

4.3.3 Computational Methods

Binding of N_2 to metal-catecholates was considered in our previous paper.[461] Binding of O_2 was treated here using the same protocol as in that work. Catecholate (abbr. "cat") was used as a model system for catecholate linkers. The metals considered were Mg2+, Sc2+, Ti2+, V2+, Cr2+, Mn2+, Fe2+, Co2+, Ni2+, Cu2+, and Zn2+, written generally as "M". O2, cat-M, and cat-M-O2 were optimized using M06-L,[504] M06,[433] PBE,[82] and PBE0[283] using Gaussian09[282] with the def2-TZVPP basis set.[505] Grimme's D3(BJ) dispersion correction[414] was used for PBE and PBE0. Structures were confirmed to be local minima by vibrational frequency analysis. All DFT calculations were spin-unrestricted and performed without symmetry.

Our previous work[461] found that M06 energies were in best agreement with a coupled-cluster benchmark on cat-Mg, so M06 geometries were used in single-point complete active space self-consistent field (CASSCF)[6] calculations followed by second-order perturbation theory (CASPT2),[9] as well as the analogous restricted active space methods (RASSCF[7] and RASPT2,[10] respectively). CASSCF/CASPT2 and RASSCF/RASPT2 calculations were performed in Molcas 8.2[53,153,265,532] without symmetry using the ANO-RCC-VTZP[264] basis set. Cholesky decomposition[533] with the default decomposition threshold of 1.0d-4 a.u. (corresponding to an error of 0.01 kcal/mol·electron)[65] was applied to the two-electron integrals to improve computational efficiency. The CASPT2 and RASPT2 (collectively referred to as "PT2") calculations had the default IPEA shift[65] of 0.25 hartrees to compensate for systematic overestimation of correlation energy and an imaginary shift of 0.1 hartrees to prevent intruder states. Orbitals were visualized with Luscus 0.8.3.[266]

We calculated DFT free energies at 298.150 K and 1 atm. However, the geometries were not optimized at the PT2 level, so thermodynamic adjustments are not

available for PT2 energies. In order to facilitate comparisons between methods, we report electronic energies:

$$\Delta E_b = E_{\text{cat-M-guest}} - E_{\text{cat-M}} - E_{\text{guest}} \quad (78)$$

where "cat" refers to the catecholate, "M" refers to the metal ion, and "guest" is the gas molecule of interest. The free energies of adsorption from the DFT calculations are available in the Supporting Information, but in general they are about 10 kcal/mol smaller than the electronic binding energies.

4.3.3.1 Active Space Selection

In a complete-active-space SCF (CASSCF) calculation, the users define three sets of molecular orbitals: inactive, active, and external. The inactive orbitals remain doubly occupied during the entire calculation, and the external orbitals remain unoccupied. The active orbitals contain the remaining electrons that are not contained in the inactive orbitals. A full configuration interaction wave function is generated by distributing the active electrons in the active orbitals in all possible ways within given user-defined spin and spatial symmetry constraints. With an appropriate active space, CASSCF can capture the multiconfigurational character of the system, although second-order perturbation theory (CASPT2) or other post-SCF methods are necessary to recover the missing dynamic correlation.

Unfortunately, CASSCF calculations with active spaces larger than sixteen electrons in sixteen orbitals, i.e., about 35 million configuration state functions (CSFs), are too large to be performed in current standard software packages.[53] Larger active spaces can be achieved by excluding additional configurations via the restricted active space self-consistent field (RASSCF)[7] method. In RASSCF, instead of having one active space, there are three subspaces, RAS1, RAS2, and RAS3. RAS1 orbitals are doubly

occupied, but with a user-defined number of excitations allowed into RAS2 or RAS3. RAS3 orbitals are unoccupied, but with a user-defined number of excitations allowed from RAS1 or RAS2. RAS2 orbitals can have any occupation. While this approach eliminates many of the possible configurations, the remaining number of configurations must still remain within the practical limit. Dynamic correlation can be added on top of the RAS wave function using RASPT2.[10]

Our active space notation is as follows: For CASSCF, we use (n,N), where n is the number of active electrons and N is the number of active orbitals. For RASSCF, we use (n,h,p;N_1,N_2,N_3), where n is the total number of active electrons, h is the maximum number of holes permitted in RAS1, p is the maximum number of particles permitted in RAS3, and N_n is the number of orbitals in the RASn subspace.

The simplest approach to active space selection is to try to include all valence orbitals, but this is unaffordable for almost any system of interest. However, many of the valence orbitals are sufficiently low or high in energy that they remain doubly occupied or unoccupied, respectively, in all of the configurations needed to describe the multireference character of the system. As systems become more complex, and especially as they become less symmetric, degeneracy is often broken and only a few orbitals may be needed in the active space. The specific orbitals involved in each active space are given in the Supporting Information, but we provide here some comments on the general approach to selecting the active space orbitals for the systems we study here and discuss the most challenging or unusual cases.

The active space of the isolated O_2 includes all 2p orbitals, namely the σ, σ*, π, and π* formed from the 2p orbitals, but excluding the σ, σ* orbitals formed by the 2s orbitals, as suggested in Ref. [47]. In isolated O_2, 2p σ and σ* orbitals exhibit the same degree of multireference character as the π and π*: the occupation numbers of the π and

π* natural orbitals are 1.96 and 1.04, respectively, and the occupation numbers of the σ and σ* orbitals are 1.96 and 0.04, respectively. In the cat-M-O_2 supersystems, the σ and σ* natural orbitals retain similar occupation numbers (1.93-1.97 and 0.03-0.07, respectively) and do not interact with the metal orbitals. The π and π* orbitals, however, are more complicated. While both π and both π* orbitals are degenerate to each other in isolated O_2, and therefore they all have to be active, that degeneracy is broken in the cat-M-O_2 supersystem. The O_2 invariably takes a side-on binding orientation to the metal, and so one of the π/π* orbital pairs is in a δ orientation to the $3d_{yz}$ of the metal, while the other is in a π orientation to either the $3d_{xy}$ or $3d_{xz}$, depending on whether the O_2 is in the plane or perpendicular to the plane of the catecholate. In the supersystems, one of the O_2 π/π* pairs is very low-lying in energy and cannot be made active; the orbitals rotate out during the optimization procedure, leaving both orbitals doubly occupied in the inactive space. The other π/π* pair needs to be active; the π* orbital is either singly occupied (with Mg, Sc, Co, Cu, and Zn) or forms an interacting pair with a singly occupied 3*d* metal orbital (with Ti, V, Cr, Mn, Fe, and Ni). The interacting pairs' natural orbitals occupation numbers vary widely, from 1.94 and 0.09 (Ti) to 1.22 and 0.81 (Cr). The π natural orbital is doubly occupied (1.97-1.98), but it is still correlated to the π* orbital and leaving it out of the active space lowers the PT2 energies by as much as 9.8 kcal/mol (4.7 kcal/mol on average), and raises the CASSCF energies by as much as 32.6 kcal (20.7 kcal/mol on average).

Isolated N_2 is simpler than O_2, as the 2*p* σ/σ* orbitals do not exhibit as much multireference character as in the O_2 case, and only the π and π* orbitals need to be active (an active space of 4 electrons in 4 orbitals). The N_2 2*p* σ and σ* occupation numbers are 1.98 and 0.02, respectively and including them in the isolated N_2 active space (6,6) lowers the PT2 energy by only 1.3 kcal/mol. Including them in the cat-Mg-N_2 active space lowers the PT2 energy by only 1.4 kcal/mol, and the effect on the binding energy is

less than 0.1 kcal/mol. The N_2 π and π* orbitals generally remain two degenerate pairs in the cat-M-N_2 supersystems, with the π natural orbitals being doubly occupied (1.93-1.96) and the π* unoccupied (0.04-0.07). The only exceptions are Sc, in which one of the N_2 π* orbitals is singly occupied (1.00), and Ti, in which one of the N_2 π* orbitals forms a π/π* interacting pair with a 3*d* orbital of the metal (with occupation numbers of 1.83 and 0.20, respectively). However, excluding the doubly occupied π orbitals and unoccupied π* orbitals from the cat-M-N_2 active spaces increases the PT2 energies by an average of 8.4 kcal/mol and the CASSCF energies by an average of 53.6 kcal/mol, and so they must remain active as well.

The isolated cat-M active spaces include a minimum of the six π orbitals on the carbon ring, as well as any singly occupied orbitals on the metal. The lowest two π natural orbitals are always close to doubly occupied (with occupation numbers 1.94-1.98 and 1.90-1.96), and the highest three π natural orbitals are always close to unoccupied (0.03-0.09, 0.02-0.10, and 0.00-0.03). The third π orbital is doubly occupied (1.90-1.92) with most metals, but it is singly occupied (0.99-1.01) with Ni, Cu, and Zn. The missing electron is located on the metal center instead. The occupancy of the third π orbital also changes in selected cases when guests are bound to the metal, which is addressed below. There are two additional π orbitals primarily located on the oxygens and their adjacent carbons, but they are low-lying in energy and remain doubly occupied (1.98) even when active with no significant impact on the binding energy. For example, adding them to the cat-Mg and cat-Mg-N_2 active spaces decreases the binding energy by only 0.2 kcal/mol.

The 3*d* orbitals of the metal do not always need to be included in the metal-catecholate active spaces, although that is a complex matter that varies from metal to metal. The variability of these systems is driven primarily by the differences in 3*d* occupations in different metals. The 3*d* orbitals need to be active when they are singly

occupied or when they are in partially occupied bonding/antibonding pairs with the catecholate or guest orbitals. Additional orbitals need to be active for the metals Fe, Co, Ni, and Cu due to the "double-shell effect".[534–536] For first-row transition metals with more than half-filled 3*d* subshells, it has been shown[10,202,325,536] that an additional set of unoccupied 3*d* ' orbitals need to be included in the active space, especially when changes in 3*d* occupancy are involved.[202,325] The double-shell effect has been shown to be less pronounced for earlier transition metals,[325] and our results are in agreement with that conclusion. For cat-Fe, including the double-shell effect decreases the binding energy by only 3.2 and 2.3 kcal/mol for N_2 and O_2, respectively, with similarly small changes for Mn and Cr. It proved impossible to even get active spaces with the double-shell effect for the earlier transition metals, as most of the 3*d* orbitals are unoccupied and often rotate out of the active space during optimization. However, for the later transition metals, the double-shell effect was sometimes quite dramatic. Of the cases studied here, cat-Co-N_2 and cat-Cu-O_2 especially highlight the importance of the double-shell effect when there are changes in 3*d* occupancy.

In cat-Co, three 3*d* orbitals are singly occupied ($3d_{xy}$, $3d_{xz}$, and $3d_{yz}$) and two are doubly occupied ($3d_{x^2-y^2}$ and $3d_{z^2}$). In cat-Co-N_2, however, only two are singly occupied ($3d_{xz}$ and $3d_{x^2-y^2}$), while three are doubly occupied ($3d_{xy}$, $3d_{xz}$, and $3d_{z^2}$). The additional 3*d* electron comes from the third catecholate π orbital, which is singly occupied in cat-Co-N_2. Using only the singly occupied 3*d* orbitals in each active space results in an N_2 binding energy of -47.7 kcal/mol, which is implausibly high in comparison to the other N_2 binding energies. However, including the full double-shell set of orbitals results in a binding energy of -19.9 kcal/mol, similar to the N_2 binding energies with most of the other metals.

Cat-Cu-O_2 has one fewer doubly occupied 3*d* orbital in the supersystem than isolated cat-Cu. Isolated cat-Cu has a full 3*d* subshell due to reduction of the metal center by the catecholate. In the cat-Cu-O_2 supersystem, however, the Cu $3d_{xy}$ orbital is singly occupied, with the other electron being in a singly occupied O_2 π^* orbital. Omitting doubly occupied Cu 3*d* orbitals leads to an implausible prediction of a positive 19.0 kcal/mol binding energy of O_2 to Cu (effectively predicting no binding). Conversely, making all 3*d* and 3*d* ' orbitals active in both cat-Cu and cat-Cu-O_2 results in a binding energy of -9.5 kcal/mol. While this energy is still considerably lower than the O_2 binding energies to other cat-M systems, we assign this difference to the fact that the full 3*d* subshell is broken up when Cu reduces O_2.

The 4*s* orbital of the metal is generally unoccupied and does not need to be active, but it is singly occupied in isolated cat-Sc, cat-Ti, cat-Ni, and cat-Zn. Unlike the 3*d*, the 4*s* orbital does not necessarily need to be active in the supersystem, even when it is singly occupied in the isolated cat-M. While the five 3*d* orbitals have the potential to be degenerate to each other, the 4*s* orbital is the only one in its subshell, and omitting it when unoccupied does not lead to an imbalanced treatment. For example, including the 4*s* in the cat-Zn active space but not in the cat-Zn-N_2 and cat-Zn-O_2 active spaces changes the binding energies by less than 1.4 kcal/mol.

As we hope we have made clear, choosing an active space is not a black-box procedure, but based on our findings we offer a general recipe for the metal-catecholate active spaces:

- The six π orbitals of the catecholate carbon ring.
- Any singly occupied metal 3*d* or 4*s* orbitals.
    - For first-row transition metals with 3*d* occupancies of 6, 7, 8, or 9, all 3*d* and “double-shell” 3*d* ' orbitals should be included.

- The π and π* orbitals of N_2.
- The 2p σ and 2p σ* orbitals of O_2, and the higher-lying π and π* orbitals. (However, in isolated O_2, the π orbitals are degenerate, as are the π*, so all four of them must be active.).
- Orbitals formed from interactions of the above guest orbitals and metal orbitals (e.g., a bonding/antibonding pair formed from a metal 3*d* and an O_2 π*).

To resolve uncertainty as to whether orbitals need to be active, the best approach is to compare energy changes with and without the orbital(s) in question, both at the CASSCF and CASPT2 level. Since CASSCF is variational, the CASSCF energies are expected to lower for larger active space. CASPT2, on the other hand is not variational, so the best one can hope for is that the CASPT2 energy converges with the active space. If it is necessary to use RAS or some other variant, it is important to use the same level of theory for all active spaces being compared.

The above recipe was used for the results presented here. Often, the CASSCF and RASSCF calculations along with their corresponding PT2 results indicate multiple near-degenerate spin states. For all spin states close to the ground state, absolute energies and the weight of the dominant configurations are presented in the Supporting Information.

4.3.4 Results and Discussion

4.3.4.1 O_2 v. N_2 Selectivity

The calculated electronic binding energies (ΔE_b) of O2 and N_2 to cat-M systems are presented in Figure 33. The DFT N_2 results were also published in our previous work.[461] The trends across the period for both guests are generally consistent between DFT functionals, as is the trend across the period for O_2/N_2 selectivity. However, there are significant differences between functionals when we compare individual binding energies. The hybrid functionals, M06 and PBE0, usually yield smaller ΔE_b values than the local functionals, M06-L and PBE. The differences between hybrid and local functionals are larger for PBE0 and PBE than for M06 and M06-L, and vary more widely (but are generally larger) for O_2 than for N_2. The range of disagreement between functionals is 4 to 13 kcal/mol for N_2, but for O_2 it is between 11 and 26 kcal/mol for all cases other than Mg and Zn. The largest differences are most often between PBE and PBE0. While all functionals indicate that most of these systems show promise for O_2/N_2 selectivity, it is difficult to identify the degree of selectivity based on DFT results alone.

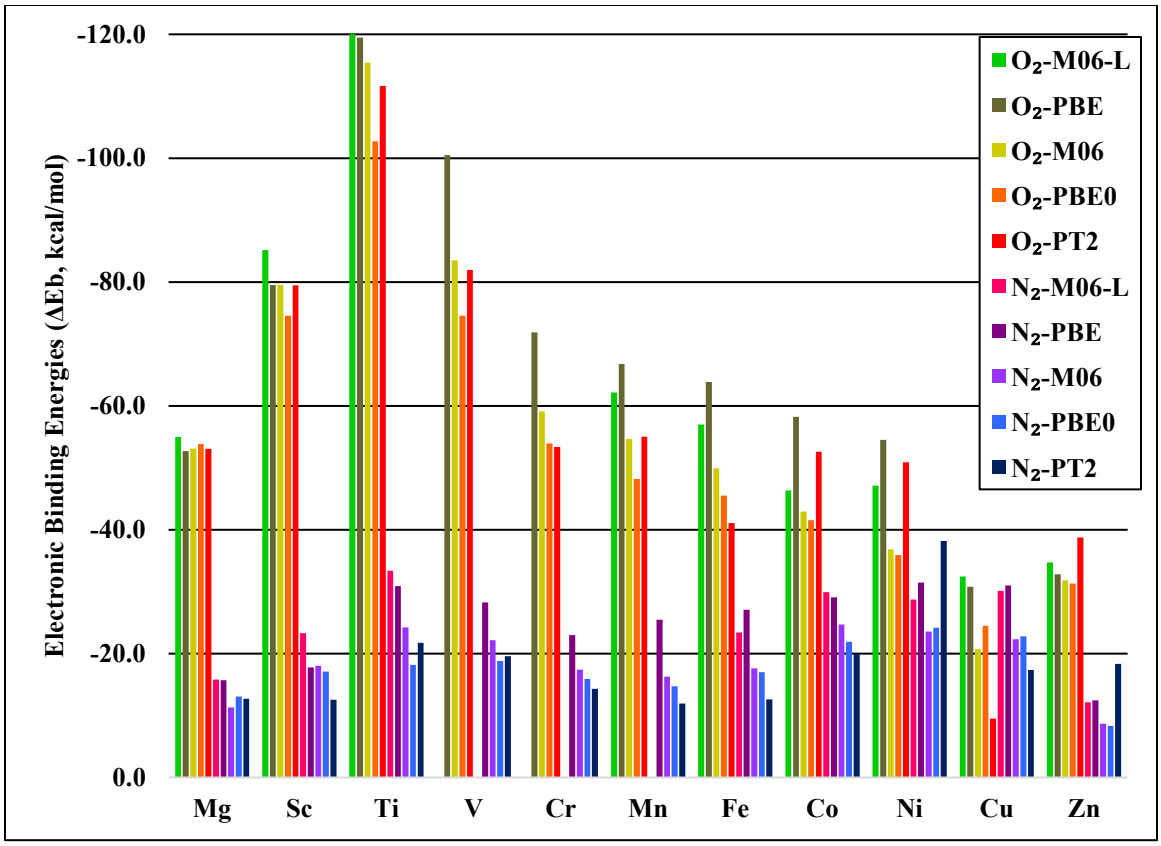


Figure 33. DFT and PT2 electronic bonding energies (ΔE_b, kcal/mol) of O_2 and N_2 to cat-M^{2+} systems. DFT N_2 results are from previous work.[461] M06-L results are absent for Mn-N_2, V, and Cr due to convergence failures.

In addition to the differences between the functionals, the DFT results had significant spin contamination with Cr for N_2 and Cr, Mn, Fe, and Co for O_2. (Calculated spin expectation values for all cases are presented in the Supporting Information.) When transition metals and open-shell systems are involved, as is the case here, it is not uncommon to have multireference character due to the potential for several near-degenerate configurations. Significant spin contamination often occurs when the multireference character of a system does not allow a single reference method, such as DFT, to obtain good energetic results while maintaining symmetry constraints. Lifting the constraints by performing spin-unrestricted calculations, as we did here, may yield

better energetics, but the physical interpretation of other properties such as the spin density may be less meaningful.[271] Accordingly, we employed the multireference wave function methods CASPT2 and RASPT2 (collectively referred to as "PT2") to gain additional insight into these systems. In addition to being able to capture the multireference character of these systems, the PT2 methods rely on CASSCF and RASSCF wave functions, which are spin eigenfunctions and by construction do not suffer from spin contamination.

The PT2 ΔE_b values of O_2 are greater than those of N_2 by at least 12 kcal/mol for all metals studied except Cu. N_2 ΔE_b values are near to or smaller than -20 kcal/mol in almost every case, while O_2 ΔE_b values range from -38.8 to -111.7 kcal/mol (excepting Cu, which is addressed below). PT2 and DFT results are in general agreement with regards to the trends in binding and selectivity, but no one functional has consistent agreement with PT2 in terms of individual binding energies. The O_2 binding energy mean unsigned errors (MUEs) with respect to PT2 are 8.0, 11.0, 5.7. and 7.5 kcal/mol for M06-L, PBE-D3(BJ), M06, and PBE0-D3(BJ), respectively, while for N_2 the respective MUEs are 9.3, 8.9, 5.3, and 4.5. (M06-L MUEs exclude cases that did not converge.) On average, the hybrid functionals (and M06 in particular) are in better agreement with PT2 than the local functionals, and all functionals are in better agreement with PT2 for N_2 than for O_2.

Table 30. CASSCF/RASSCF Maximum configuration weights for cat-M and cat-M-guest systems. Smaller values indicate greater multireference character.

Metal	cat-M	cat-M-O_2	cat-M-N_2
Mg	0.89	0.88	0.84
Sc	0.89	0.86	0.85
Ti	0.89	0.83	0.78
V	0.90	0.79	0.84
Cr	0.90	0.17	0.84
Mn	0.90	0.29	0.84
Fe	0.91	0.60	0.88
Co	0.90	0.38	0.86
Ni	0.60	0.41	0.88
Cu	0.93	0.92	0.90
Zn	0.92	0.88	0.84

Notice that we take the PT2 values as references, but there are no experimental values available against which the PT2 values can be verified. However, the multireference character of these systems is confirmed by the CASSCF/RASSCF configuration weights (Table 30). The configuration weight can be thought of as the degree of contribution a particular electronic configuration makes to the final wave function. A maximum configuration weight of 1.00 would indicate a purely single-reference case, while maximum weights much smaller than 1.00 indicate multireference character. For these systems, cat-M-O_2 cases are especially multireference, with several systems having a maximum configuration weight of less than 0.50, indicating that no single configuration makes up even half of the final wave function.

An analysis of the CASSCF/RASSCF orbital occupations helps us understand the physical causes of the trends in the binding energy. A key difference between the guest molecules is that the two π^* antibonding orbitals are unoccupied in isolated N_2, but they are singly occupied in isolated O_2. While the π^* orbitals of N_2 generally remain

unoccupied, with the exceptions of Sc, Ti, and Co, O_2 is always reduced by the cat-M system. The greater binding of O_2 than N_2 is primarily assigned to this redox activity between O_2 and the cat-M systems.

The differences in the binding energies of different metals to O_2 can be understood as differences in the redox chemistry. In the supersystems, the degeneracy of the O_2 π^* orbitals is broken, resulting in one doubly occupied orbital and what would be one unoccupied O_2 π^* orbital. However, in all cases studied this higher-lying π^* orbital receives an electron from either the catecholate or the metal center. (Figure 34) The strongest binding occurs with Sc^{2+}, Ti^{2+}, and V^{2+}, when O_2 is reduced by the metal center (Figure 34a). These metals are typically found in higher oxidation states than 2+, so it is energetically favorable for them to be oxidized to 3+. However, Ti should realistically be at 4+, so that particular case may not be as physically feasible as the others studied. We limited ourselves to M^{2+} for this study to improve our ability to observe trends across the period, but in future work we intend to consider additional oxidation states, especially if they are of experimental interest.

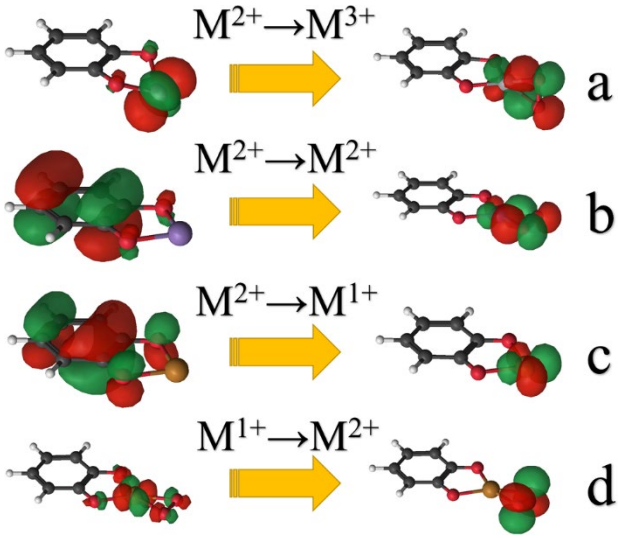


Figure 34. Electron transfer to O_2. (a) In early transition metals, the electron comes from a singly occupied orbital on the metal. (b) In middle transition metals and in Mg, a catecholate π orbital is the source of the electron. (c) Late transition metals are reduced by the catecholate before guests have been added, and (d) the metal in turn reduces the O_2.

In contrast to Sc, Ti, and V, where the metal reduces the O_2, the metal remains in a 2+ state for Mg, Cr, Mn, Fe, and Co (Figure 34b). In these cases, the O_2 is reduced by the catecholate instead. Ni, Cu, and Zn are especially interesting cases, as the catecholate reduces the metal in the cat-M subsystem (Figure 34c), resulting in 1+ oxidation states, but then the metals reduce the oxygen in the supersystem and thus return to 2+ (Figure 34d). The change in metal oxidation state is especially important in understanding the low binding of O_2 to Cu. In isolated cat-Cu, the 3*d* subshell is full, as the Cu is reduced by the catecholate. However, the 3*d* subshell is broken up when Cu reduces O_2, inducing an energy penalty that reduces the magnitude of the binding energy. N_2 does not get reduced by Cu, and so the 3*d* subshell remains full with N_2 and N_2 does not see the energy penalty that O_2 does. This breaking of the 3*d* subshell is why O_2 binding is predicted to be weaker than N_2 for cat-Cu, in contrast to the other metals that do not break up a full 3*d* subshell.

For the purpose of O_2/N_2 separation, it is important to have a significant difference in the interaction energies of O_2 and N_2 and to have both interaction energies low enough that binding is readily reversible. All the metals studied except Cu^{2+} offer large differences in the O_2 and N_2 interaction energies across all methods. In many cases, however, the O_2 binding is so strong that it would be impractical to reverse the binding, which would make it impossible to use a material with the given catecholate-metal as part of an ongoing industrial process. Of the metals studied here, the early transition metals are clearly unsuitable, at least in the admittedly unlikely 2+ oxidation state. The middle and late transition metals have somewhat weaker O_2 binding, but it is not clear whether it would be weak enough to make reversibility easily accessible. The best candidates out of the metals studied are Fe^{2+} and Zn^{2+}, as they have the lowest O_2 binding energies that are still over 20 kcal/mol greater than N_2 binding energies. However, it may be possible to tune the redox activity of the catecholate using additional functional groups such as an additional aromatic ring[531] on the opposite side of the catecholate carbon ring, and therefore tune the O_2 binding for other metals. Additional metals and oxidation states may also offer additional candidates. We plan to investigate these possibilities in future studies, as well as the effects of multiple guests (including solvent molecules) on a given metal center.

4.3.4.2 NO Binding and Selectivity

Our previous work[461] focused on using metal-catecholates in an improved gas mask filter to bind NO and other toxic gases selectively against benign gases that would be expected to be present. While these systems showed significant promise for selectivity of NO vs. N_2, CO_2, and H_2O, O_2 was not addressed. We now add consideration of O_2 to the previous data (Figure 35). We present M06 results here, as it was the functional with

the best overall agreement with PT2 for N_2 and O_2 (as discussed in section 4.3.4.1). Results for other functionals for all guests other than O_2 may be found in our previous work.[461] O_2 binds more strongly than the toxic gases (NO, CO, and NH_3) for most metals, but NO binding is stronger than O_2 by over 8 kcal/mol for Co^{2+}, Ni^{2+}, and Cu^{2+}, and all three toxic gases have stronger binding than O_2 for Cu^{2+}. For toxic gas filtration in atmospheric conditions, Cu^{2+} is the most promising candidate, although Co^{2+} and Ni^{2+} also show promise for NO specifically.

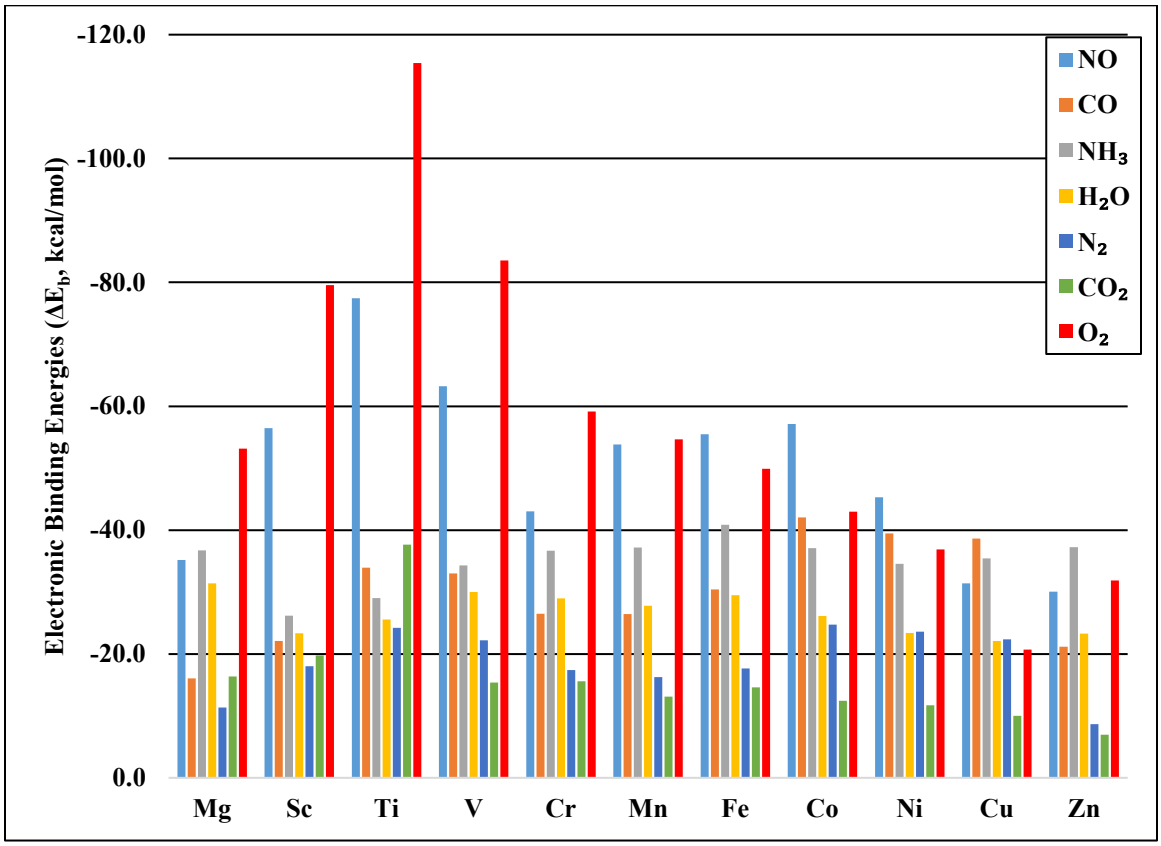


Figure 35. M06 free energies of adsorption (ΔE_b, kcal/mol) of various gases to cat-M^{2+} systems. All results other than O_2 are from previous work.[461]

4.3.5 Conclusions

In the interest of finding new materials for air separation, the binding of O_2 and N_2 to catecholates with 2+ transition metals was investigated with DFT and multireference PT2 wave function methods. The multireference nature of many of these systems makes them challenging, but there is good agreement among the different methods with regards to the overall trends. In addition to confirming the general trends, the multireference methods provide physical insight and enable us to explain the trends in O_2 binding in terms of redox activity between the catecholate, the metal, and the O_2 guest.

Some guidance is offered for choosing active spaces for these and similar systems: We found that it is sufficient to include only the six π orbitals on the carbon ring of the catecholate, a limited subset of the guest N_2 or O_2 valence orbitals, and any singly occupied 3*d* or 4*s* orbitals of the metal. For the later first-row transition metals, additional *d* orbitals are needed to account for the "double-shell" effect.

All metals studied, with the exception of Cu^{2+}, are predicted to have much stronger O_2 binding than N_2, and the weak binding of O_2 to Cu^{2+} suggests that such systems may be useful for other applications in which O_2 binding is not desired, such as toxic gas filtration. For air separation, Fe^{2+} and Zn^{2+} are the most promising out of those considered, as they are the most likely to have O_2 binding weak enough to enable reversibility while still having good O_2/N_2 separation. Further work is planned to consider tuning O_2 binding with functional groups on the catecholate ring, additional metals and oxidation states, and a more detailed consideration of realistic coordination environments on the metal center. Additionally, larger-scale simulations are planned for specific MOF structures modified with the metal-catecholates studied here. Based on the

results we have found so far, metalated catecholates in MOFs have the potential to offer excellent selectivity for air separation and are worthy of considerable ongoing study.

4.3.6 Additional Information

Supporting Information: Discussion of active space selection for CASSCF and RASSCF, along with absolute energies, dominant configuration weights, orbital occupation numbers, and visualized orbitals. CASPT2 and RASPT2 electronic binding energies, absolute energies, and reference weights. DFT optimized geometries, calculated spin expectation values, electronic binding energies and free energies of adsorption. All DFT absolute energies used in calculating interaction energies, including electronic energies with and without zero-point correction, enthalpies, and free energies. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.8b03599.

This research is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-17ER16362. The authors acknowledge the Minnesota Supercomputing Institute (MSI) for providing computing resources.

4.4. Metal-Organic Frameworks with Metal Catecholates for O_2/N_2 Separation

This section describes the outcome of a collaborative research project carried out by Hakan Demir, Samuel J. Stoneburner, WooSeok Jeong, Debmalya Ray, and Xuan Zhang, (and advised by Omar K. Farha, Christopher J. Cramer, J. Ilja Siepmann, and Laura Gagliardi). A report on this research project has been published.[537]

Samuel J. Stoneburner conceived the original idea, performed the cluster DFT calculations, contributed to the structural analysis and screening, and contributed to the manuscript.

4.4.1 Overview

Oxygen and nitrogen are widely produced feedstocks with diverse fields of applications, but are primarily obtained via the energy-intensive cryogenic distillation of air. More energy-efficient processes are desirable, and materials such as zeolites and metal-organic frameworks (MOFs) have been studied for air separation. Inspired by recent theoretical work identifying metal-catecholates for enhancement of O_2 selectivity MOFs, the computation-ready experimental (CoRE) database of MOF structures was screened to identify promising candidates for incorporation of metal catecholates. Based on structural requirements, preliminary Grand-Canonical Monte Carlo simulations, and further constraints to ensure the computational feasibility, over 5,000 structures were eliminated and four MOFs (UiO-66(Zr), Ce-UiO-66, MOF-5, and IRMOF-14) were treated with periodic density functional theory (DFT). Metal catecholates (Mg, Co, Ni, Zn, and Cd) were selected based on cluster DFT calculations and were added to the shortlisted MOFs. Periodic DFT was used to compute O_2 and N_2 binding energies near metal catecholates. We find that the binding energies are primarily dependent on the metals in the metal catecholates, all of which bind O_2 quite strongly (80-258 kJ/mol) and have weaker binding for N_2 (3-148 kJ/mol). Of those studied here, Cd-catecholated MOFs are identified as the most promising.

4.4.2 Introduction

The production of oxygen gas is at enormous levels (>100 million tons/year) and oxygen gas is one of the most essential chemicals due to its various uses in medicine, chemical manufacturing, coal gasification, wastewater treatment, fuel cells, and the paper industry.[109,538,539] High-purity oxygen (>99%) is crucial for a variety of areas, such as medical,[539] military and aerospace,[540] semiconductor,[541] cylinder filling,[542] ozone

generation,[542] plasma chemistry,[542] and oxy-fuel combustion[516] applications. For instance, for surgeries in the U.S., the minimum oxygen purity is 99%, and it is even higher in Japan (99.6%).[540] In semiconductor, military, and aerospace applications similar concentrations of 99.8%, 99.5%, and 99.5%, respectively, are required.[540] Thus, effective separation of oxygen and nitrogen from air can provide large sources of commodity gas that would lower the cost of the aforementioned applications.

Separation of O_2 and N_2 has been regarded as one of the most challenging separations due to their similar molecular sizes.[543] Cryogenic distillation has been used to produce O_2 at industrial scales since the 1920s.[544] However, the boiling points of O_2 and N_2 at 1 atm are -183 and -196 °C, respectively, and so a great deal of energy is required to condense the gases at very low temperatures.[545] Although cryogenic distillation is a proven technology for large quantities of air separation (>200 tons/day), for small and medium scale production more economical alternatives around room temperature are possible, such as membrane separation, vacuum swing adsorption (VSA), and pressure swing adsorption (PSA).[546–548] It has been demonstrated that the PSA technique with current materials can provide oxygen at concentrations up to 95% for small and medium scale needs with much less energy than required for cryogenic distillation.[540,545]

Zeolites are traditionally used as adsorbents for many applications, including O_2/N_2 separation. Many zeolites have shown preference for N_2 over O_2 such as 4A, 5A, LiAgX, LiLSX, and Linde 10X zeolite, which are molecular sieves hosting different pore aperture sizes with disparate chemical contents in terms of silica amount and the absence/presence of various cations (Li, Ag).[549,550] The more favorable interactions for N_2 compared to O_2 have been suggested to arise from the stronger interactions of the larger quadrupole moment of N_2 (-1.4 Debye.Å for N_2 vs. -0.4 Debye.Å for O_2)[439,551] with the electric field in the framework.[550,552] For obtaining O_2 with air separation,

however, the adsorbent should selectively capture O_2 over N_2 around room temperature. Capturing O_2 would bring great economic benefit compared to zeolites since capturing lean gas with porous materials means smaller volumes of gas need to be processed. Metal-organic frameworks (MOFs)[94,96,553,554] are promising candidates for selective O_2 adsorption. MOFs can be efficient O_2 selective materials since they have much wider chemical diversity than traditional adsorbents as well as highly porous and functionalizable structures.

The basic concept underlying the construction of the MOFs is the linking of two building blocks (i.e., metal nodes and organic linkers), and so almost infinitely many MOFs can be constructed by combining different building blocks.[395] Recently, the number of experimental and theoretical MOFs has risen swiftly and several collections of structures involving MOFs such as the Cambridge Structural Database (CSD)[428], hypothetical MOFs (hMOFs)[395], computation-ready experimental MOFs (CoRE MOFs[131]), and ToBaCCo[555] have been reported.

Although high numbers of MOFs might initially appear to be an advantage, performance testing each MOF experimentally would be daunting in terms of time and financial resources. Computational tools can be highly beneficial in accelerating the experimental efforts to find the best material for a particular application. Grand-Canonical Monte Carlo (GCMC) is a molecular simulation method routinely used to investigate the adsorption thermodynamics in nanoporous materials having a wide range of structural diversity.[556–558] For instance, Moghadam et al.[559] performed GCMC simulations for more than 2900 MOFs in the CoRE MOF database to acquire adsorption loadings and deliverable O_2 capacity, both at 298 K. Having identified top MOFs from GCMC simulations, they synthesized UMCM-152 and achieved the highest deliverable oxygen capacity (22.5% higher than the second top material in the literature) and

matching loadings with those in simulations. DeCoste et al.[560] made a similar analysis for 10,000 MOFs and identified NU-125 as a better O_2 storage material than NaX and Norit activated carbon (237%, and 98%, respectively) in terms of excess capacity. GCMC simulations can be categorized into two types: force-field based and ab-initio based simulations. The accuracy of the former heavily relies on the force field parameters which determine the intramolecular/intermolecular interactions.[561] Many examples have been demonstrated where generic force fields can make predictions close to experimental adsorption values for sorbates such as CH_4 and H_2.[562–564] However, when Zeitler et al.[565] studied 98 materials for O_2 adsorption using a generic force field, UFF,[440] their comparisons with experimental data demonstrated that UFF is incapable of describing O_2-open metal site interactions accurately. Density functional theory (DFT) is a relatively affordable quantum chemical method which can be used to derive specialized force fields that could significantly improve generic force field predictions and reproduce experimental data,[105,421,566] but the use of DFT-based force fields is generally limited to the specific systems for which they were designed. DFT can also be used to study adsorption in systems for which a force field is unavailable.[567] Wang et al.[568] studied O_2 adsorption in $M_3(BTC)_2$ (M= Cr, Mn, Fe, Co, Ni and Cu) with PBE-D2 and found $Ni_3(BTC)_2$ to be a potentially useful oxygen adsorbent that could favorably bind O_2 over N_2. Parkes et al.[513] screened the binding energy of O_2 and N_2 in M_2(dobdc) and $M_3(BTC)_2$ with 14 different metals with PBE-D2. The MOFs substituted with early transition metals were recommended as the best materials for selectively capturing O_2 over N_2 because both MOFs show higher O_2 binding energy with early transition metals than with late transition metals. Similar studies have been conducted by Gallis et al.[521], Verma et al.[522] and Xiao et al.[525] on M-BTC (M = Mn, Fe, Co, Cu), Fe_2(dobdc), and Co-BTTri and Co-BDTriP, respectively.

Open-metal MOFs that have been synthesized and studied experimentally regarding O_2 adsorption/separation typically have had critical performance failures such as poor cyclability and/or limited separation performance except at relatively low temperatures. For example, $Cr_3(BTC)_2$ exhibits much more favorable O_2 adsorption interaction compared to the physisorption of N_2, driven by a partial electron transfer from Cr^{2+} to the bound O_2 sorbate, but it shows a steady decrease in O_2 uptake for repeated cycles.[569] A Co(II) carborane-based MOF shows an O_2/N_2 selectivity of 6.5 at low pressure, but it quickly diminishes to ~2 at higher pressures.[570] Bloch et al.[109] have demonstrated that Fe_2(dobdc) is an O_2 selective MOF, but with irreversible O_2 binding above 226 K. They found that charge transfer from Fe(II) to O_2 changes from partial transfer at low temperature into complete transfer at room temperature. Later, Bloch et al.[571] reported that Cr-BTT exhibits rapid O_2 adsorption/desorption kinetics with good O_2/N_2 selectivity. These studies illustrated the importance of the presence of redox-active metal in MOFs for preferential O_2 uptake over N_2. However, these open metal MOFs have been known to lose crystallinity and O_2 adsorption performance when adsorption conditions are not well controlled.[525] In contrast, Co-BTTri and Co-BDTriP showed good recyclability up to 10 cycles, but they exhibited high selectivities only at low temperatures (13 and 40 at 243 K for Co-BTTri and Co-BDTriP, respectively).[525] Low O_2/N_2 selectivity at room temperature is a common problem for MOFs. For example, it has been shown that MOF-177 possesses an O_2/N_2 selectivity of 1.8 at 298 K, 1 atm.[572] Likewise, UMCM-1 exhibits an O_2/N_2 selectivity of 1.64 at similar conditions (298 K, 0.96 bar).[573] MOFs demonstrating high O_2/N_2 selectivity around room temperature have not been reported yet to the best of our knowledge, which is a gap in material space and is one of the motives of this work.

Metalated catecholate linkers have brought new opportunities in adsorption due to the strong interactions between the open metal site in the metal catecholate with multiple

sorbates (e.g., Weston et al.[497]), and thus could be used to obtain higher O_2/N_2 selectivity and O_2 uptake. A previous computational investigation[326] predicted that metal catecholates (Mg, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, Cu, and Zn) could be beneficial for air separation. Except for Cu catecholate, all metal catecholates were shown to interact stronger with O_2 than N_2. Fe and Zn catecholates were suggested as the most appropriate choices due to relatively weak affinity towards O_2, which is expected to result in a relatively easy desorption and improve the regenerability of the system. In addition, generation of strong adsorption sites at the metal-catecholate rather than at the nodes could be a good way to avoid structural stability issues. Conventional open metal sites in metal nodes are closely related to structural integrity, since the weakest bond in MOFs is the metal-ligand bond.[574,575] In contrast, the metal sites in metalated catecholate linkers are not directly bonded to the pristine MOF structure and are not expected to play a significant role in the general structural stability. While there are additional synthetic and reusability challenges that arise from the reactivity of the catecholates and the undercoordination of the metal, metal-catecholates and related moieties have already been synthesized in several MOFs[494–496,529,576] and porous organic polymers[129,497,498,530], with applications in gas separations[129], gas storage[497], and catalysis[494,495,498,529,530]. For example, Fei et al.[495] reported the synthesis of UiO-66(Zr) with Cr-catecholates for oxidation catalysis with no loss of Cr and with good stability with respect to temperature and aqueous solvent. Similarly, Huang et al.[529] reported the synthesis and hydroboration of carbonyls on ANL1-Ti$(OiPr)_2$ with no significant loss of the alkoxide-supported Ti. Tanabe et al.[498] found that functionalizing a porous organic polymer with a Ta^V trialkyl resulted in no loss of thermal and structural robustness.

In this study, we aim to develop highly O_2 selective MOFs for O_2/N_2 separation by inserting metal-catecholates into experimentally known MOF structures. Our objective can be divided into two targets: to find parent MOF structures that can be good platforms

for metalated structures, and then to develop metalated MOF structures from the parent ones. The scheme of this multi-stage work is depicted below in Figure 36. In the first stage several structural criteria, which are indicated in the Computational Methods section, are applied to CoRE MOFs to identify MOFs with appropriate pore space where metal catecholates could be incorporated. In the second stage the shortlisted MOFs are used, without any modification, in binary GCMC simulations (O_2/N_2 = 20:80) to predict O_2/N_2 selectivity at 1 bar, 298 K. MOFs with O_2/N_2 selectivity larger than 1 are considered to have good potential for selective O_2 separation, inasmuch as they are not N_2-selective and will therefore not have structural effects counterproductive to our goal of O_2 selectivity. In the third stage metal-catecholate functionalized MOF structures are generated for the screened pristine MOFs, with the metals selected based on previous work[326] and new cluster DFT calculations. Duplicate metalation sites are excluded using an in-house developed code in Python. In the last stage a subset of the metalated MOF structures are chosen based on computational feasibility of structural optimization and the binding energies of O_2 and N_2 are calculated with periodic DFT.

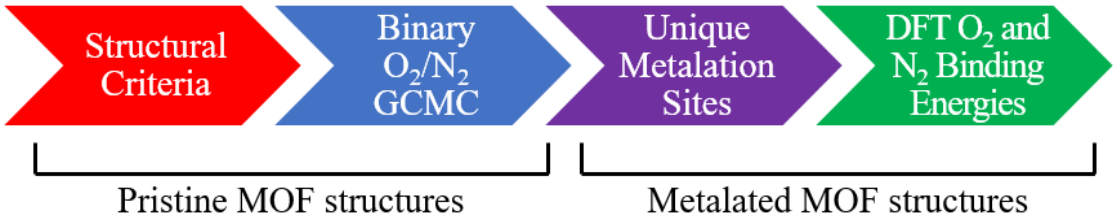


Figure 36. Multi-stage screening approach.

4.4.3 Computational Methods

4.4.3.1 Overview of the Screening Procedure

Our work consisted of multiple steps, summarized in Figure 37. The CoRE MOF database includes 5,109 experimentally reported MOF structures. The initial 5,109 structures were reduced to 2,867 by requiring 6-membered rings consisting of C and/or N atoms and requiring that at least two adjacent ring atoms have one H atom attached to each. These requirements ensured that the given structures have site(s) for inserting metalated catecholates, resulting in the selection of MOFs that have organic ligands containing benzene(-like) moiety such as 1,4-benzenedicarboxylate (BDC).

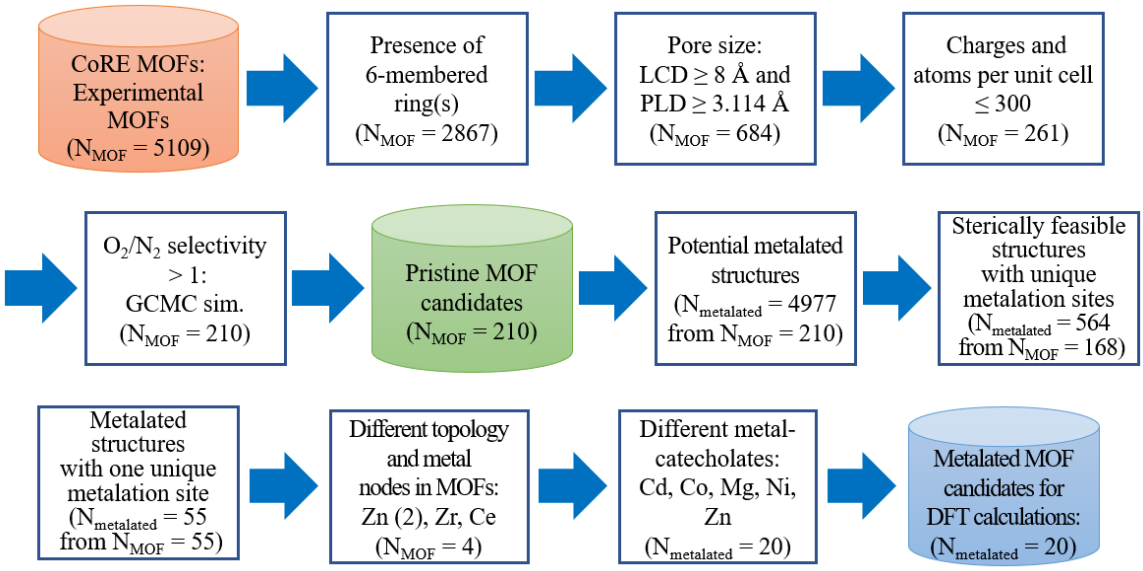


Figure 37. Screening process for selecting metalated MOF structures for DFT calculations.

Next, MOFs were required to have pore limiting diameters (PLD) larger than or equal to 3.114 Å and largest cavity diameters (LCD) larger than or equal to 8.000 Å. The pore limiting diameter (i.e., pore aperture diameter) minimum was chosen as 90% of the

O_2 kinetic diameter, which filtered out most of the MOFs that would experience high diffusional constraints. The largest cavity diameter (i.e., largest pore diameter) was chosen so that there would be enough space available in the pore space to incorporate a metal catecholate. The pore specifications reduced the number of MOF structures to 684.

In anticipation of GCMC calculations and recognizing the need for relatively reliable charges for the framework atoms, MOFs were omitted that involve atoms for which there were not sufficient electron affinity and ionization potential data in the EQeq code,[460] mostly lanthanides and actinides. To reduce computational cost for subsequent periodic DFT calculations, MOFs were also required to have less than or equal to 300 atoms per unit cell. Thus, 261 MOF structures out of the original 5,109 survived to GCMC calculations. Note that omissions for computational feasibility took place only after the structural screening was complete, and so only 423 structures (i.e., less than 8.3% of the original 5,109) were excluded from GCMC calculations for the technical reasons described here. See the Supporting Information for further details on the size-excluded structures.

4.4.3.2 Structural Analysis

MOF structures were obtained from the CoRE MOF database, which includes experimentally reported structures only.[131] Potential sites for metal-catecholates, i.e., 6-membered rings of carbon and/or nitrogen atoms, of which two adjacent ring atoms have one H atom each attached, were identified using the geometry analysis tools of Platon.[577,578] Geometrical analysis on pore sizes and the number of open metal MOFs was conducted with Zeo++.[579] Additional structural analysis for metalated MOF structures was performed with custom in-house code as described in subsequent sections.

4.4.3.3 GCMC

Binary GCMC simulations for binary gas mixtures (20% O_2, 80% N_2) were carried out using the RASPA[438] code to obtain O_2 and N_2 loadings and O_2/N_2 selectivity data at p = 1 bar and T = 298 K as follows

$$S_{O_2/N_2} = \frac{N_{O_2}/N_{N_2}}{x_{O_2}/x_{N_2}} \tag{79}$$

where N is the gas uptake (in units of molecules per simulation cell) and x is the mole fraction of this component in the feed mixture. Adsorbate-adsorbate and adsorbate-MOF interactions were modeled through a combination of Lennard-Jones and Coulomb potentials. The force field parameters for O_2[551]and N_2[439] (3-site models) were obtained from the TraPPE force field, while UFF Lennard-Jones parameters were used for the framework atoms. The framework atoms were assigned charges based on the EQeq[460] (extended charge equilibration) method. Although the framework flexibility may play an important role in some adsorption cases,[580–583] in this study the atomic positions were kept frozen during the simulations due to the unavailability of generic flexible force fields. The Lennard-Jones interactions were truncated at 12 Å without analytical tail correction, as is common in simulations for MOFs.[584,585] (Larger truncation distances do not lead to any significant change in selectivity, as shown in the Supporting Information.) The Ewald summation technique[586] was used with a relative precision of 10^{-6} for electrostatic interaction calculations. The GCMC simulations involved 50,000 cycles in total with equal equilibration and production cycles, where the allowed GCMC moves were translation, rotation, reinsertion,[587] molecular identity change, and random insertion/deletion with equal probability. Here, a Monte Carlo cycle is defined as max(20,N) Monte Carlo steps, in which N denotes the total number of O_2 and N_2 guest molecules in the simulation box. The gas loadings were computed in 5 blocks (i.e, for 1-10000, 10001-20000, …, 40001-50000 cycles). The ratios of gas uptake in the first and

last simulation block over the average gas loading are plotted in Figure S8, which shows the insignificant deviations along the simulation. As a supplement, the average and variance of these ratios are presented in Table S11 where it can be seen that the averages over all materials are very similar in the first and last simulation blocks for both O_2 and N_2 and the spread of these ratios is narrow, which implies that 50,000 cycles were sufficient to get converged GCMC results.

4.4.3.4 Generation of Metalated Catecholate Functionalized Structures

Metalated MOF structures were created from the prescreened pristine MOFs (EDUVOO [IRMOF-14], RUBTAK02 [UiO-66(Zr)], SAHYIK [MOF-5], and Ce-UiO-66) by inserting only one metal-catecholate moiety in each potential metalation site (defined above under "Structural Analysis"). Ce-UiO-66 is not included in the CoRE MOF database, however, it has been added to our list of materials to study the effect of metal type in MOF nodes on binding energies, which was previously shown to be large for the binding energy of H_2O.[588] In this work, we report both the refcode (i.e., a 6-character alphanumeric reference code) and common name for a MOF structure. Note that there might be multiple refcodes for a specific MOF depending on experimental conditions such as temperature, the presence of solvent molecules, etc. The geometry of the added metal-catecholate moiety was obtained from the DFT optimized geometry of N_2 bound to Zn-catecholate as published in previous work.[326] The MOF structures with metalated catecholates were checked for steric hindrances and symmetrical redundancy using an in-house code developed in Python, the procedure of which is detailed further below.

In the environmental analysis implemented in the code, a unit cell was expanded to a supercell with consideration of an environment checking radius of 15 Å for periodic

boundary conditions. Next, within the checking radius centered on the added metal atom, up to 50 nearest framework atoms or more when distances of the atoms are too close (<0.05 Å) were selected to produce environment information for the specific metalation site. Four different characteristics were calculated as the environment information: atom type, distance between the added metal atom and the framework atom (r_{ij}), angle between the vector from the center of a ring to the added metal atom and the vector from the added metal atom to the framework atom (θ_1), and angle between the vector from one oxygen atom to the other oxygen atom in the catecholate ligand and the vector from the added metal atom to the framework atom (θ_2). Finally, based on the environment information, duplicate metalation sites were excluded (Figure 38). In addition, sterically unphysical metalation sites, which were too close to other framework atoms (<2.5 Å), were also filtered out.

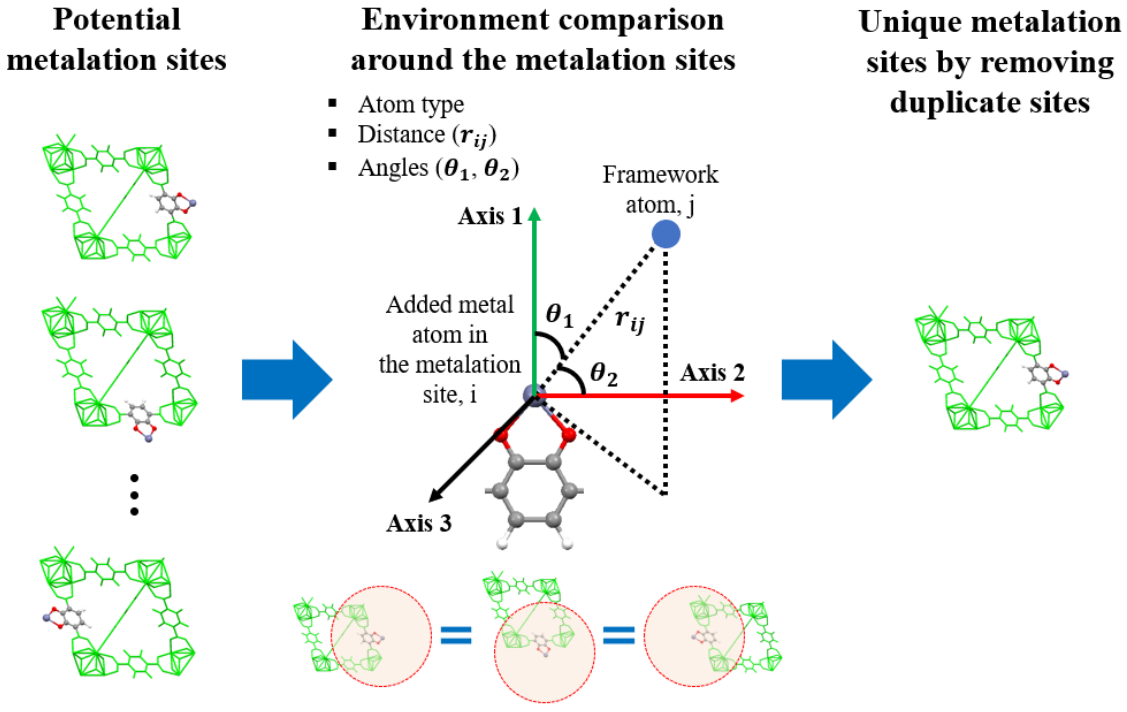


Figure 38. Environment comparison for identifying unique metalation sites among potential metalation sites.

4.4.3.5 Cluster DFT

Many different metals could be considered for the metal-catecholates, and treating each of them in each shortlisted MOF structure with periodic DFT would have required more calculations than could readily be performed. Instead, we chose a selected list of specific metals based on cluster calculations in previous work[326] and additional cluster DFT calculations using PBE-D3(BJ),[82,414,589] which allowed for direct comparison to periodic DFT results. In keeping with previous work,[326] the binding energies of guests (O_2 and N_2) to a given metal-catecholate were calculated as the energies of the isolated guest and the isolated metal-catecholate (Figure 39) subtracted from the energy of the catecholate-metal-guest supersystem (Equation (80)).

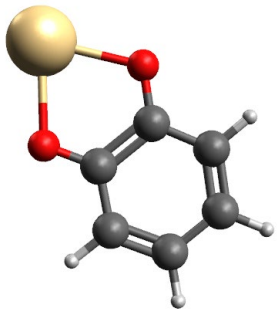


Figure 39. Cluster model of metal-catecholate system.

$$E_{bind} = E_{complex} - E_{metal-catecholate} - E_{adsorbate} \quad (80)$$

All new cluster DFT calculations were performed with Amsterdam Density Functional (ADF) 2017[590–592] using the all-electron TZ2P basis set[593] and the zeroth-order regular approximation (ZORA) for scalar relativity corrections.[594–596] All spin states were considered. The calculated Gibbs free energies are based on a standard state of ideal gas at 1 atm. For the sake of simplicity in the periodic DFT calculations, only 2+ oxidation states were considered for the data presented here. Metals were selected based on

estimated availability of precursors and the likelihood of obtaining a 2+ oxidation state, specifically Be^{2+}, Mg^{2+}, Ca^{2+}, Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Zn^{2+}, Sr^{2+}, Pd^{2+}, Cd^{2+}, Ba^{2+}, Pt^{2+}, and Pb^{2+}.

4.4.3.6 Periodic DFT

The periodic spin-polarized DFT binding energies of O_2 and N_2 were calculated using the Vienna Ab Initio Simulation Package (VASP)[432,597] with PBE-D3(BJ).[82,414,589] In these calculations, both the MOF unit cell information and the MOF atomic coordinates were used as reported in the CoRE MOF database. A metal catecholate was incorporated into the viable MOFs (see Results and Discussion for details) and then the adsorbate molecules were initially placed near the metal catecholate in three orthogonal configurations and were then fully optimized. For Co and Ni in the metal-catecholates, high spin states were assumed for the metals with ferromagnetically coupled O_2 in the framework. All periodic DFT calculations were performed using Γ-point sampling and a 500 eV kinetic energy cutoff. The energy and force convergence criteria were 10^{-5} eV and 0.01 eV/Å, respectively. Missing hydrogens in the RUBTAK02 structure were added as in Yang et al.[598] in order to have a neutral structure. The proton topology assigned for UiO-66(Zr) follows that of the most stable configuration of NU-1000 as demonstrated by Planas et al.[599] For periodic DFT calculations, the electronic binding energy is defined as:

$$E_{bind} = E_{complex} - E_{framework} - E_{adsorbate} \quad (81)$$

4.4.4 Results and Discussion

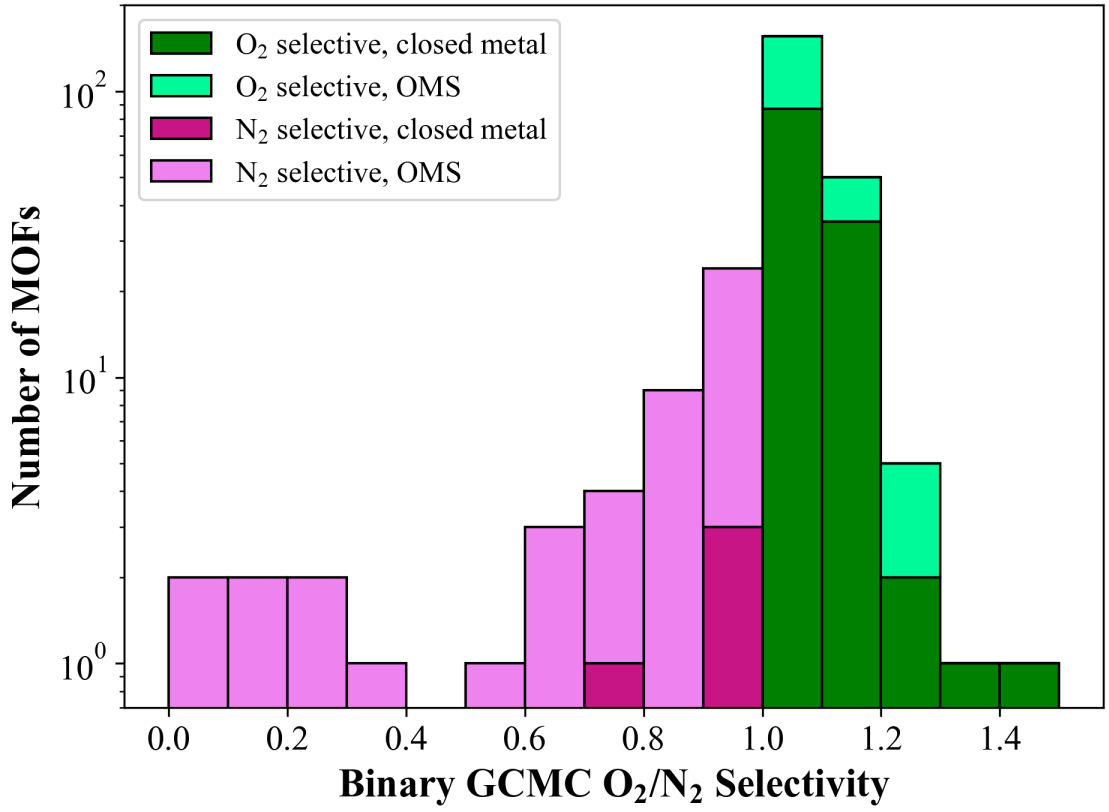


Figure 40. Binary GCMC O_2/N_2 (20%/80%) selectivity for 261 pristine MOFs obtained at 1 bar, 298 K.

In Figure 40, a histogram of the O_2/N_2 selectivities obtained from the binary GCMC simulations is shown for the 261 pristine MOFs. The selectivity varies between 0 and ~1.5, but most materials perform near a selectivity of 1. Unexpectedly, there are seven MOFs that show highly N_2 selective behavior ($0 < S_{O_2/N_2} < 0.4$), but all of these contain open metal sites. A more detailed analysis for NEXXEV, the most N_2 selective material, indicates that (after removal of solvent molecules) the structure contains rows of free (essentially uncoordinated) Li cations that cause erroneous partial charge

assignments with the EQeq method and lead to unreasonably strong interactions with N_2 molecules (see Supporting Information). Detailed GCMC results can also be found in the Supporting Information.

Our screening analysis focuses on the materials that are O_2 selective ($S_{O_2/N_2} > 1$) based on the GCMC results. These materials have structural properties that allow preferential adsorption of O_2 over N_2, taking into account dispersion and electrostatic interactions only. However, it should be noted that the generic force fields (i.e. UFF) may fail to accurately account for the interactions between the sorbates and the adsorbent[561] for structures both with open metal sites (e.g., CO_2 adsorption in Mg-MOF-74[105]) and without open metal sites (e.g., Ar adsorption in ZIF-8[566]). Thus, in reality, these O_2 selective materials could exhibit higher selectivity than predicted by GCMC simulations. For example, the experimental O_2/N_2 selectivity is calculated as ~1.7, and ~1.7-1.8 for UMCM-1, and MOF-177, respectively, by taking the ratio of pure O_2 and N_2 adsorption loadings at 298 K, 1 bar.[600] To test the proximity of UFF predictions with respect to experiments, we performed pure O_2 and N_2 GCMC simulations and the O_2/N_2 selectivity is predicted to be 1.05 and 1.02 for UMCM-1 and MOF-177, respectively. Our conclusions match those of Zeitler et al.,[565] who found that UFF failed to account for O_2-metal interactions in a study comparing the experimental O_2/N_2 selectivities of several MOFs to those obtained by GCMC simulations . Thus, the calculated GCMC O_2/N_2 selectivities in this work likely underpredict O_2/N_2 selectivity that would be obtained in experimental conditions. Accordingly, it should not be assumed that the GCMC step of the screening identifies that the best candidates without further screening being necessary. On the contrary, the GCMC step only eliminates the cases that would be N_2-selective, which would be counterproductive for O_2-selective modifications.

To obtain a more accurate description of O_2 and N_2 adsorption, periodic DFT calculations were performed on the metal-catecholated versions of a subset of these materials. Considering only cases for which $S_{O_2/N_2} > 1$ left 210 surviving MOF structures out of the original 5,109 for possible metalation. From these 210 structures, a total of 4,977 potential metalated structures can be generated by assuming every two adjacent H atoms in the rings can be transformed to a metal-catecholate complex, and by including only one metalation site per metalated structure. However, to perform DFT calculations for all the generated potential metalated structures would be a daunting task, especially as there would need to be at least 7 calculations per metalated structure: with and without two different gas molecules, N_2 and O_2, and with 3 different initial configurations for each (see Figure 41).

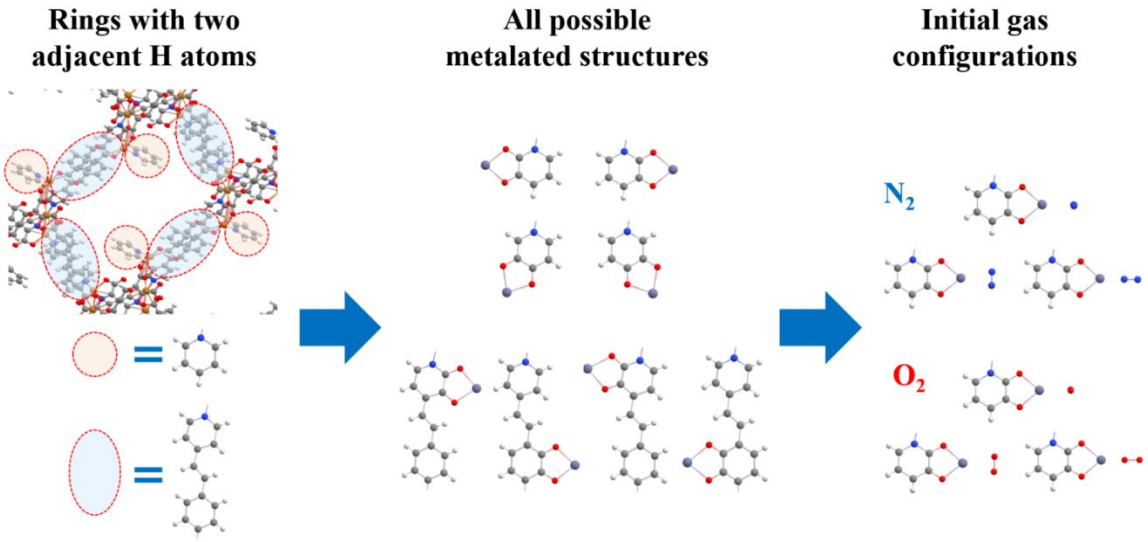


Figure 41. Generation of metalated structures for the periodic DFT calculations. Color code: oxygen atoms, red; carbon, gray; hydrogen, white; nitrogen, blue; metal, blue-gray.

To narrow down the number of potential metalated structures for further DFT calculations, metalated structures with duplicate or sterically unrealistic metalation sites were excluded via environment comparisons (Figure 38) using in-house developed code. Unrealistic metalation sites could be created when the added metalation site is in a small pore, or the direction of the metal-catecholate complex (i.e., direction from the center of a ring to the added metal atom) faces toward the pore wall (see Supporting Information for details). Note that this set of exclusions is therefore physically motivated, not merely a concession to computational limits.

Following the environment comparisons, 564 unique metalated structures from 168 parent MOF structures were obtained. To simplify the analysis, MOF structures that possess more than one unique metalation site were ruled out. MOFs with multiple unique metalation sites would likely have an unpredictable mixture of metalated sites in experiment, making it more difficult to make comparisons between experiment and theory. While it is possible that some of the excluded structures could yield even better air separation capabilities than those we consider for the remainder of this study, our focus on this work is identifying several promising candidates in such a way that the computational results can be experimentally tested.

Of the 564 possible metalated structures, only 55 have only one unique metalation site. In these structures, the structure nodes feature different metals: Cd (1), Cu (8), Mn (2), Sc (1), Zn (41), and Zr (2), where the number in the parentheses is the number of structures. 55 structures would still have led to an overabundance of calculations, so we attempted to select a diverse sample by considering at least one MOF with each metal in the node for DFT optimization after the metal catecholate incorporation. However, we observe that for many structures incorporation of the metal catecholate led to large structural changes during the optimization (e.g., for UKIQIP-Zn-cat the Zn of the

catecholate unit is found to approach closely the C atoms of the nearest benzene ring), and we surmise that these structures would not be stable in their metal catecholate form. In general, we have observed that converging to a physically reliable structure (i.e. no undesired bond breakage/formation) becomes harder as the metalated linker points more towards the MOF scaffold rather than the open porous space. In those cases, the strong confinement effect around the metalated linkers may be responsible for the significant distortion of the structure. We have not investigated the effect of the orientation of the metalated linker on the optimization of the structures, however, as it is beyond the scope of our screening study. For simplicity, our metalated structure optimizations are initialized with the same orientation of the linkers as in the pristine MOFs. Only SAHYIK (Zn)-M-cat, EDUVOO (Zn)-M-cat, and RUBTAK02 (Zr)-M-cat resulted in functional structures after optimization (SAHYIK: $Zn_4O(BDC)_3$, where BDC = 1,4-benzene-dicarboxylate, EDUVOO: $Zn_4O(PDC)_3$ where PDC = pyrene-2,7-dicarboxylate, RUBTAK02 (H atoms added): $Zr_6O_4(OH)_4(BDC)_6$) where the interatomic distances are reliable with no undesired bond formation/breakage. The vibrational frequency calculations of the optimized structures are not performed due to the high computational cost. It is notable that these MOFs have large pores where the metalated linkers are oriented towards the pore space. The structures identified using the screening methodology are listed with their structural information in Table 31 for parent and metal-catecholated MOFs with 5 different metals in catecholates that will be discussed in the next paragraph. Note that UiO-66(Zr) (here as RUBTAK02) has been successfully used before as a support for metalated catecholates.[495,496] As expected, with the addition of metal catecholates, there are slight decreases in void fractions. However, in terms of pore sizes, the incorporation of metal catecholate can cause an increase or decrease depending on the rotation of the linker in the optimized structures.

Table 31. Final MOF candidates studied for O_2/N_2 separation.

MOF	Metal node	Metal-catecholate	PLD (Å)	LCD (Å)	Void fraction
SAHYIK (MOF-5)	Zn	-	7.8	14.9	0.77
		Cd	8.0	15.1	0.76
		Co	8.0	15.2	0.76
		Mg	8.0	15.1	0.76
		Ni	8.0	15.2	0.76
		Zn	8.0	15.2	0.76
EDUVOO (IRMOF-14)	Zn	-	10.6	20.9	0.84
		Cd	10.4	21.0	0.83
		Co	10.4	21.0	0.83
		Mg	10.4	21.0	0.83
		Ni	10.5	21.0	0.83
		Zn	10.4	21.0	0.83
RUBTAK02 (UiO-66)	Zr	-	3.9	8.5	0.47
		Cd	4.2	7.7	0.44
		Co	4.2	7.7	0.43
		Mg	4.2	7.7	0.44
		Ni	4.2	7.7	0.45
		Zn	4.2	7.7	0.45
UiO-66	Ce	-	4.4	9.1	0.50
		Cd	4.5	8.1	0.48
		Co	4.5	8.1	0.47
		Mg	4.5	8.1	0.48
		Ni	4.5	8.1	0.48
		Zn	4.5	8.1	0.49

In selecting the candidate metals for the metal-catecholates, it is important to strike a balance between the absolute O_2 binding energy, which is important for reversibility, and the difference in O_2 and N_2 binding energies ($E_{sep} = E_{b,\,O2} - E_{b,\,N2}$), which speaks to potential for separation. In the previous cluster calculations[326] all metals

other than Cu (i.e, Mg, Sc, Ti, V, Cr, Mn, Fe, Co, Ni, and Zn) were predicted to be O_2 selective and had high values for E_{sep}, but were also predicted to have extremely high absolute O_2 binding that would likely be irreversible. In order to see whether the magnitude of the O_2 binding might decrease at the periodic level, we selected the cases with the (relatively) weakest O_2 binding among the investigated metals above: Co, Mg, Ni, and Zn. However, when their periodic DFT binding energies were calculated in this work, they were found to be similar to the cluster DFT binding energies and therefore still have very strong absolute O_2 binding energies (in the Supporting Information, Table S3). Therefore, for this study we performed new cluster calculations on a larger set of metals in order to identify more favorable metal candidates. As explained in the Computational Methods section, these metals were selected for their expected experimental feasibility based on the availability of synthetic precursors and accessibility of the 2+ oxidation state. The results of the new cluster calculations are presented in Figure 42, and as before, most metals studied are predicted to have overly strong binding for O_2. There is not necessarily a known specific value for the electronic binding energy below which O_2 binding will be reversible, in part because it would depend on the thermal stability of the specific MOF support. However, we expect that ideally the absolute O_2 binding energies would need to be within ~20-40 kJ/mol, and therefore in order to obtain good O_2/N_2 separation the absolute N_2 binding energies would ideally be near 0 kJ/mol so as to maintain a maximum differential. Of the metals studied, only Cd and Pb offer sufficiently low absolute N_2 binding energies. In Cd-catecholate, the difference between O_2 and N_2 binding energy is substantial, which suggests N_2 may not compete with O_2. This competitive advantage is not expected to hold for Pb-catecholate due to the smaller difference between O_2 and N_2 binding energies. In light of these factors, Pb catecholates are not considered further, and the final list of metals for the metal-catecholates is Cd, Co, Mg, Ni, and Zn. We acknowledge that all of these metals

feature absolute O_2 binding energies larger than 40 kJ/mol, but as we discuss further down we believe there are mitigating factors that keep Cd within the realm of practical possibilities. While Co, Mg, Ni, and Zn all have much larger absolute O_2 binding energies and are not expected to allow for reversible binding, we include their periodic DFT results for the sake of analyzing whether binding trends are exclusively due to the metals in the metal-catecholates or whether the support structures also play a role.

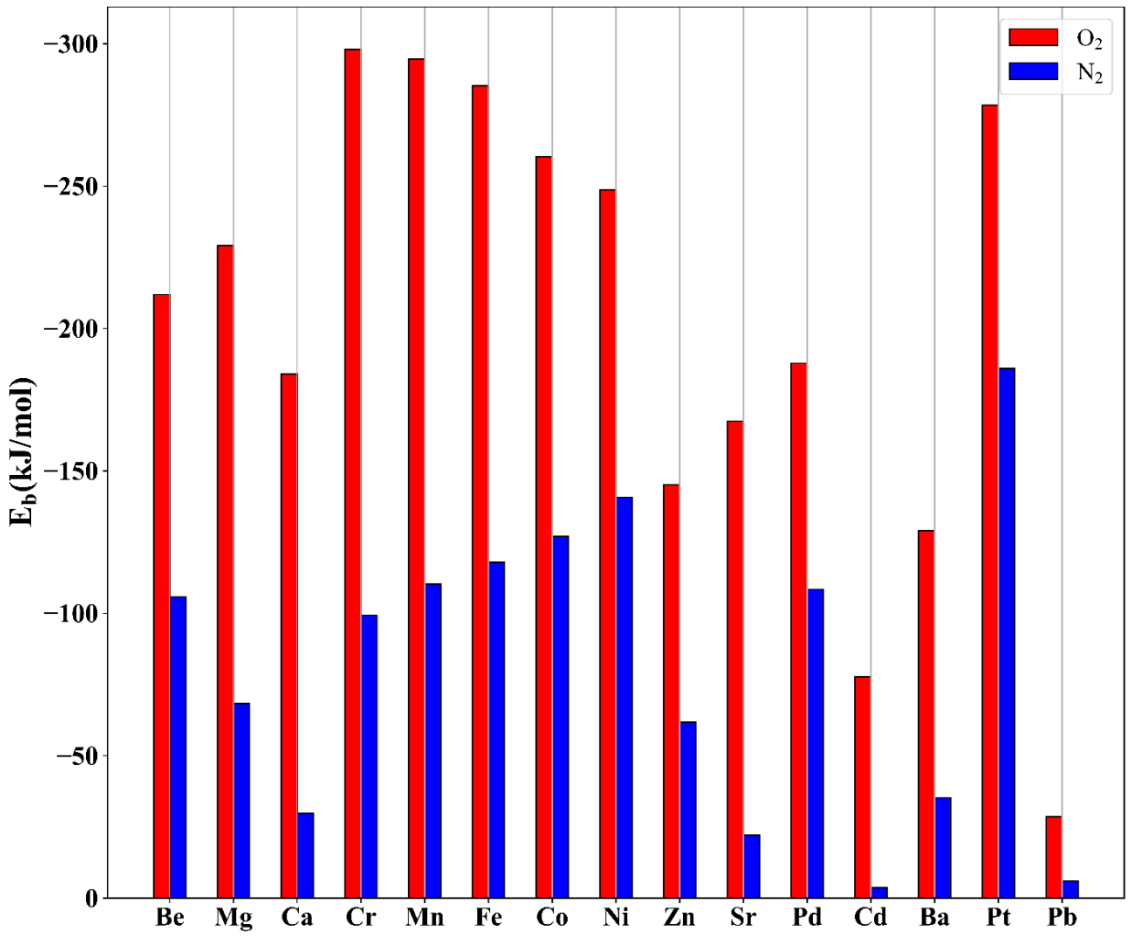


Figure 42. PBE-D3(BJ) binding energies (E_b) of O_2 and N_2 in cluster models. Results presented here use the lowest energy spin state for each system.

The selected metal-catecholates are included in four MOFs to obtain the binding energies using periodic DFT. Those binding energies are then compared with those from cluster models. Figure 43 presents O_2 and N_2 binding energies for each material, where the type of metal-catecholate is denoted with a suffix after the MOF name (i.e. MOF Name-Zn-cat). The binding energies shown in Figure 43 are calculated with the lowest energy configurations optimized from the three orthogonal initial positions. Note that the MOF structure (i.e. the environment of the metal catecholate) has limited effect on the strength of O_2 and N_2 adsorption, as can be seen when comparing data with the same metal-catecholates in different MOFs. However, it should be understood that a different observation is possible for other MOFs, especially MOFs having smaller pores that create larger confinement effects around the adsorbate. Such an effect does not exist in the studied structures due to the distance of the atoms (except the metal catecholate) around the adsorbates. Besides the full framework effect, it should also be noted that the binding energies of O_2 and N_2 change very little when comparing between UiO-66(Zr) and its Ce analogue, implying that the effect of the metal in the node is minimal when the binding is taking place at a metal-catecholate inserted into a linker. The nature of the support structure, however, may remain relevant in that some degree of thermal stability will be required in order to maintain structural integrity during the adsorption and release of O_2. The strength of O_2 adsorption follows the order of Mg-cat > Co-cat > Ni-cat ≈ Zn-cat > Cd-cat while N_2 adsorption strength is in the order of Ni-cat ≈ Co-cat > Mg-cat ≈ Zn-cat > Cd-cat. To estimate the O_2/N_2 selectivity in periodic systems, the following relation is used:

$$S_{O2/N2,DFT} = e^{-(\Delta G_{O2} - \Delta G_{N2})/RT} \quad (82)$$

Here, $\Delta G = \Delta E + \Delta G_{corr}$ where ΔE is the electronic binding energy obtained in the periodic model and ΔG_{corr} is the difference between ΔG and ΔE obtained from the cluster

calculations. As can be seen in Table S10, $S_{O2/N2,DFT}$ follows the order: Mg-cat > Zn-cat > Cd-cat > Co-cat > Ni-cat for all MOF types. It should be noted that $S_{O2/N2,DFT}$ is only a qualitative value that is used to give an idea about the selectivity trends at low pressure. It should not be considered a quantitative value for selectivity (i.e. $S_{O2/N2,DFT}$ value of 0.3 for SAHYIK-Ni-cat does not imply that structure is N_2 selective, it rather demonstrates that, among SAHYIK materials, SAHYIK-Ni-cat is predicted to be the least selective since others possess higher selectivity estimates.)

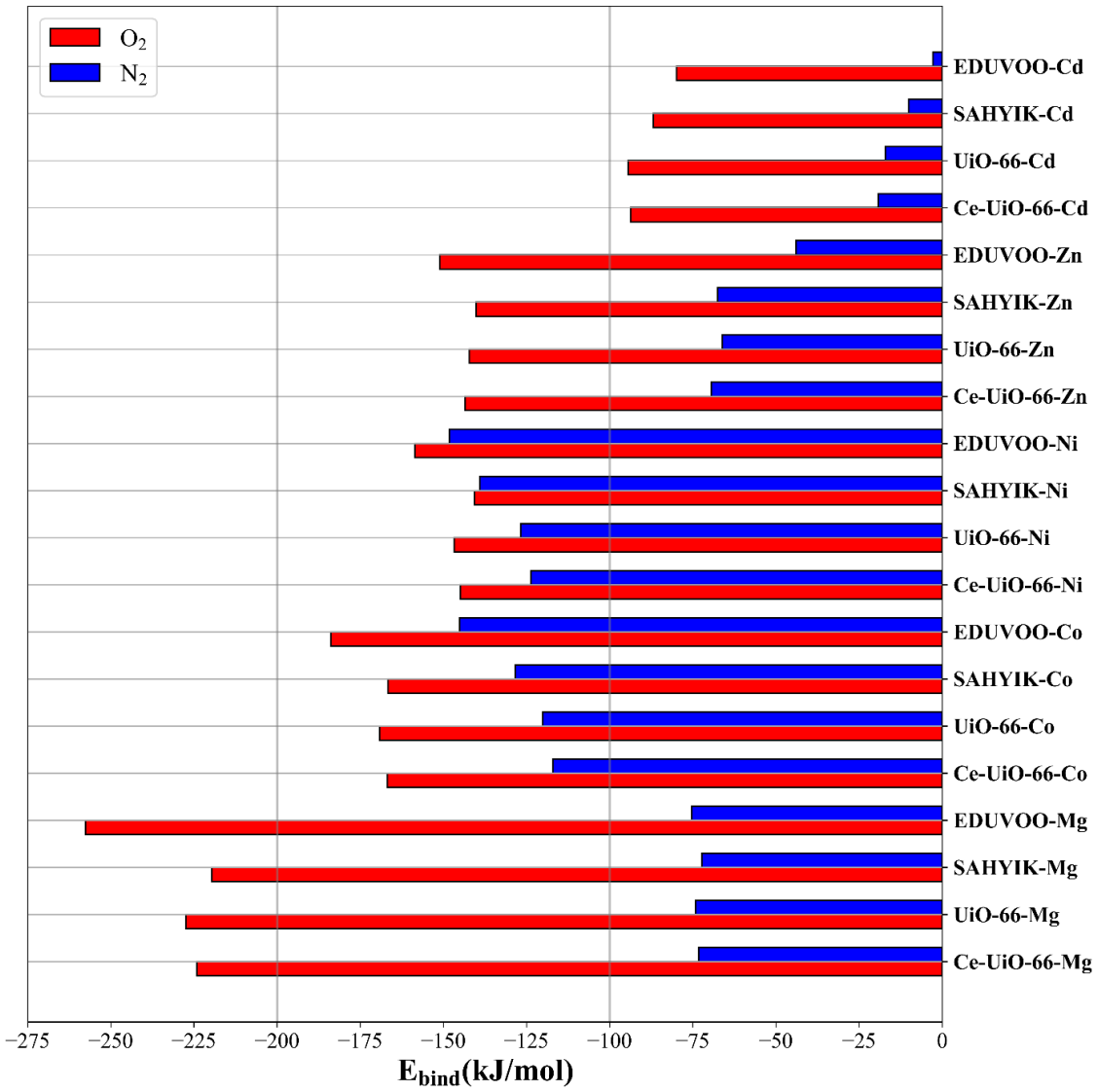


Figure 43. PBE-D3(BJ) binding energies (E_b) of O_2 and N_2 in RUBTAK02 (UiO-66(Zr)), Ce-UiO-66, SAHYIK (MOF-5), and EDUVOO (IRMOF-14) incorporated with metal-catecholates (Mg, Co, Ni, Zn, and Cd).

In order to rationalize the binding energy trends, we further analyzed our periodic results for EDUVOO MOF and calculated Quantum Theory of Atoms in Molecules (QTAIM) charges using the Bader charge analysis code developed by the Henkelman Group at the University of Texas – Austin[601–605] on the catecholate metal center and on

the bonded O_2/N_2 molecules. We do not see a significant difference in the QTAIM charges of metal centers when comparing between the three different guest conformations, so the QTAIM charges of the metal centers are reported for the most stable conformations only (in the Supporting Information, Table S9). We note that while charge is an important electrostatic descriptor, there is another key difference between these metals, namely that the Co and Ni metals have partially filled *d* orbital shells while Mg, Zn, and Cd are characterized by all shells of a given angular momentum being fully filled. We found that for Mg, Zn, and Cd, the QTAIM charge on the metal center follows the same order as that of the binding energy (i.e. Mg > Zn > Cd). For Co and Ni, we observed that for Co-catechol the QTAIM charge on the Co center is slightly higher than that of the Ni in the O_2-bound catechol, suggesting that O_2 binds slightly stronger to the Co-catechol compared to the Ni-catechol because it oxidizes the metal to a greater extent. The charges on the Co and Ni centers are very similar in the N_2 bounded structures and thus their N_2 binding energies are very similar as well.

The periodic results mirror those reported for cluster calculations in previous work[326] because the large pores result in there not being a significant adsorption site near the adsorbate other than the metal catecholate. Zn-cat had lower absolute O_2 and N_2 binding energies than most of the other metals studied due to a full 3*d* subshell that minimizes covalency in O_2 or N_2 binding. In our current work, we see that Cd has even lower absolute O_2 and N_2 binding energies than Zn; this is consistent with the more diffuse character of the Cd 5*s* orbital compared to the 4*s* for Zn, which leads to the guests binding at greater distances and with reduced electrostatic interactions as well as reduced overlap for any covalent contributions to binding.

While a significant difference between O_2 and N_2 binding energies in favor of O_2 adsorption would be beneficial for efficient O_2/N_2 separation, absolute O_2 binding

energies cannot be too large without causing a high energy requirement for regeneration. O_2 and N_2 binding energies in MOFs with Mg-catecholate suggest that, in spite of the large E_{sep}, they are not the ideal materials due to having overly high absolute O_2 binding energies (over 220 kJ/mol). MOFs with other metal-catecholates also exhibit high absolute O_2 adsorption energies (~80-180 kJ/mol) despite being lower than those in MOFs with Mg-catecholates. Within the five metal catecholates studied with periodic DFT, Cd-catecholates lead to the lowest absolute O_2 adsorption energies in MOFs (~80-95 kJ/mol) together with low absolute N_2 adsorption energies (~3-20 kJ/mol). Even at 80 kJ/mol, however, it would likely be difficult to evacuate O_2 from the material at low pressures, which would lead to a high regeneration cost when using these MOFs to separate an O_2/N_2 mixture. That being said, there are several possible solutions. In this work, the adsorption energies of only the first molecules adsorbed are investigated in detail. However, it is likely that many of the subsequent molecules would adsorb with less energy, which can render a material partially regenerable. This hypothesis is best tested using MOFs with Cd-catecholates, as their O_2 adsorption is expected to be weaker than the other materials. To illustrate this idea, DFT calculations employing two O_2 molecules near the Cd-catecholate of EDUVOO have been performed with different initial positions selected for O_2 molecules. The lowest energy configuration shows a combined binding energy for two O_2 molecules to be -110.4 kJ/mol, or an average of -65.2 kJ/mol. As the lowest binding energy for the first O_2 molecule in EDUVOO-Cd-Cat is -79.9 kJ/mol, this implies that a second O_2 molecule will have weaker binding than the first one. Furthermore, it may be possible to weaken the intensity of the O_2 binding by further functionalization of the catecholate carbon ring. Zhang et al. have found that the redox capabilities of a metal-catecholate can be tuned by adding functional groups to the opposite side of the 6-membered ring.[531] While their study was for catalysis, it is possible that a similar strategy could be used for gas separation. Because tuning the redox activity

would likely proceed by adding functional groups to the 6-membered ring according to established principles of organic chemistry, our current work should serve as a useful starting point for any such future studies.

4.4.5 Conclusions

Air separation has long been an active research topic because oxygen and nitrogen are highly desired gases for many industries and applications. In contrast to earlier N_2-selective zeolite studies, this work focuses on the potentially O_2-selective MOFs. O_2-selective materials are more economically desirable for high-purity and high-volume oxygen due to requiring lower volumes of processed gas in a real application. The MOFs that are investigated in this work (RUBTAK02 (UiO-66(Zr)), SAHYIK (MOF-5), and EDUVOO (IRMOF-14)) are shortlisted from the CoRE MOF database using structural criteria and binary GCMC O_2/N_2 selectivity data. Ce-UiO-66 is also added to this list despite not being a constituent of the CoRE MOF database in order to study the effect of metals in UiO-66 nodes on adsorption energies. These MOFs are modified with metal catecholates (Mg, Co, Ni, Zn, and Cd) to investigate O_2 and N_2 binding affinity using DFT. In general, the change in O_2 and N_2 binding energies across different MOFs is small, implying the dominant factor to be the metal type in the metal catecholate and that the MOF structure has only a secondary effect on the binding energies. This suggests that, in most of the cases, the interaction of sorbates near metal catecholates is dominant over other moieties in the structure. Since overly strong adsorption is not favored for regenerability, materials that have relatively low absolute O_2 binding energies are sought, along with low absolute N_2 binding energies for high O_2/N_2 selectivity. Out of five metal catecholates, Cd-catecholates show the least intense O_2 adsorption (~80-95 kJ/mol), and with much less favorable N_2 adsorption (~3-20 kJ/mol). The big difference between O_2

and N_2 binding energies could enable high O_2/N_2 selectivity, however, due to the relatively high O_2 binding energies, MOFs with Cd-catecholates likely would experience regenerability problems. It is possible, however, that in practice O_2 binding would be weaker than predicted due to anions or solvents such as water binding to the metal, increasing the coordination number. It also may be possible to tune the binding energies by modifying the catecholate carbon ring. Although it is not studied in this work, another approach could be using these MOFs for catalysis, where these materials might provide a suitable medium for oxidation of molecules such as methane or ethane.

4.4.6 Additional Information

Supporting Information: Breakdown of numbers of atoms in size-eliminated MOFs; detailed GCMC simulation results; neutrality check for MOFs; metalation code details; DFT binding energies and associated absolute energies in cluster and periodic models; analysis of the most N_2 selective MOF; analysis of the sensitivity of GCMC O_2/N_2 selectivities to Lennard-Jones cutoff distances; periodic DFT energies for two O_2 molecules in EDUVOO-Cat-Cd; gas loadings across the GCMC simulation blocks. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.9b02848.

This research is supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-17ER16362. Computer resources were provided by the Minnesota Supercomputing Institute at the University of Minnesota.

References

(1) Gaggioli, C. A.; Stoneburner, S. J.; Cramer, C. J.; Gagliardi, L. Beyond Density Functional Theory: The Multiconfigurational Approach to Model Heterogeneous Catalysis. *ACS Catal.* **2019**, submitted for publication.

(2) Ma, D.; Li Manni, G.; Gagliardi, L. The Generalized Active Space Concept in Multiconfigurational Self-Consistent Field Methods. *J. Chem. Phys.* **2011**, *135*, 044128.

(3) Li Manni, G.; Ma, D.; Aquilante, F.; Olsen, J.; Gagliardi, L. SplitGAS Method for Strong Correlation and the Challenging Case of Cr_2. *J. Chem. Theory Comput.* **2013**, *9*, 3375–3384.

(4) Zhao, Y.; Tishchenko, O.; Gour, J. R.; Li, W.; Lutz, J. J.; Piecuch, P.; Truhlar, D. G. Thermochemical Kinetics for Multireference Systems: Addition Reactions of Ozone. *J. Phys. Chem. A* **2009**, *113*, 5786–5799.

(5) Roos, B. O. The Complete Active Space Self-Consistent Field Method and Its Applications in Electronic Structure Calculations. In *Advances in Chemical Physics, v. 69: ab initio Methods in Quantum Chemistry, part II*; Lawley, K. P., Ed.; Wiley: New York, 1987; pp 399–445.

(6) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A Complete Active Space SCF Method (CASSCF) Using a Density Matrix Formulated Super-CI Approach. *Chem. Phys.* **1980**, *48*, 157–173.

(7) Malmqvist, P.-Å.; Rendell, A.; Roos, B. O. The Restricted Active Space Self-Consistent-Field Method, Implemented with a Split Graph Unitary Group Approach. *J. Phys. Chem.* **1990**, *94*, 5477–5482.

(8) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-Order Perturbation Theory with a CASSCF Reference Function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(9) Andersson, K.; Malmqvist, P.-A.; Roos, B. O. Second-Order Perturbation Theory with a Complete Active Space Self-Consistent Field Reference Function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

(10) Malmqvist, P. Å.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The Restricted Active Space Followed by Second-Order Perturbation Theory Method: Theory and Application to the Study of CuO_2 and Cu_2O_2 Systems. *J. Chem. Phys.* **2008**, *128*, 204109.

(11) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2014**, *10*, 3669−3680.

(12) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. Erratum: Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2016**, *12*, 458.

(13) Verma, P.; Truhlar, D. G. HLE16: A Local Kohn-Sham Gradient Approximation with Good Performance for Semiconductor Band Gaps and Molecular Excitation

Energies. *J. Phys. Chem. Lett.* **2017**, *8*, 380–387.

(14) Schrödinger, E. Quantisierung Als Eigenwertproblem. *Ann. Phys.* **1926**, *384*, 361–376.

(15) Szabo, A.; Ostlund, N. S. Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory; McGraw-Hill: New York, 1982; pp 1–137, 231–272, 321–326.

(16) Löwdin, P. O. Correlation Problem in Many-Electron Quantum Mechanics: I. Review of Different Approaches and Discussion of Some Current Ideas. In *Advances in Chemical Physics, v. 2*; Interscience Publishers, Inc.: New York, 1959; pp 207–322.

(17) Tew, D. P.; Klopper, W.; Helgaker, T. Electron Correlation: The Many-Body Problem at the Heart of Chemistry. *J. Comput. Chem.* **2007**, *28*, 1307–1320.

(18) Born, M.; Oppenheimer, R. Zur Quantentheorie Der Molekeln. *Ann. Phys.* **1927**, *20* (84), 457–484.

(19) Slater, J. C. The Theory of Complex Spectra. *Phys. Rev.* **1929**, *34*, 1293–1322.

(20) Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*, 89–110.

(21) Fock, V. Näherungsmethode Zur Lösung Des Quantenmechanischen Mehrkörperproblems. *J Zeitschrift für Phys.* **1930**, *61*, 126–148.

(22) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.

(23) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(24) Peach, M. J. G.; Tozer, D. J.; Handy, N. C. Exchange and Correlation in Density Functional Theory and Quantum Chemistry. *Int. J. Quantum Chem.* **2011**, *111*, 563–569.

(25) Tishchenko, O.; Zheng, J.; Truhlar, D. G. Multireference Model Chemistries for Thermochemical Kinetics. *J. Chem. Theory Comput.* **2008**, *4*, 1208–1219.

(26) Ruedenberg, K.; Schmidt, M. W. Physical Understanding through Variational Reasoning: Electron Sharing and Covalent Bonding. *J. Phys. Chem. A* **2009**, *113*, 1954–1968.

(27) Koch, W.; Holthausen, M. C. A Chemist's Guide to Density Functional Theory; Wiley-VCH: Weinheim, 2000; pp 1–91.

(28) Odoh, S. O.; Cramer, C. J.; Truhlar, D. G.; Gagliardi, L. Quantum-Chemical Characterization of the Properties and Reactivities of Metal–Organic Frameworks. *Chem. Rev.* **2015**, *115*, 6051–6111.

(29) Helgaker, T.; Jørgensen, P.; Olsen, J. Molecular Electronic-Structure Theory; John Wiley & Sons Ltd: Chichester, 2000; pp 1-200,433-647,724-816.

(30) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*,

B864–B871.

(31) Schuch, N.; Verstraete, F. Computational Complexity of Interacting Electrons and Fundamental Limitations of Density Functional Theory. *Nat. Phys.* **2009**, *5*, 732–735.

(32) Alipour, M. Making a Happy Match between Orbital-Free Density Functional Theory and Information Energy Density. *Chem. Phys. Lett.* **2015**, *635*, 210–212.

(33) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

(34) von Lilienfeld, O. A.; Tuckerman, M. E. Molecular Grand-Canonical Ensemble Density Functional Theory and Exploration of Chemical Space. *J. Chem. Phys.* **2006**, *125*, 154104.

(35) Matuszak, D.; Aranovich, G. L.; Donohue, M. D. Lattice Density Functional Theory of Molecular Diffusion. *J. Chem. Phys.* **2004**, *121*, 426–435.

(36) Cembran, A.; Song, L.; Mo, Y.; Gao, J. Block-Localized Density Functional Theory (BLDFT), Diabatic Coupling, and Their Use in Valence Bond Theory for Representing Reactive Potential Energy Surfaces. *J. Chem. Theory Comput.* **2009**, *5*, 2702–2716.

(37) Peverati, R.; Truhlar, D. G. Quest for a Universal Density Functional: The Accuracy of Density Functionals across a Broad Spectrum of Databases in Chemistry and Physics. *Phil. Trans. R. Soc A.* **2014**, *372*, 20120476.

(38) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Doubly Hybrid Meta DFT: New Multi-Coefficient Correlation and Density Functional Methods for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A* **2004**, *108*, 4786–4791.

(39) Karton, A.; Tarnopolsky, A.; Lamère, J. F.; Schatz, G. C.; Martin, J. M. L. Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical . *J. Phys. Chem. A* **2008**, *112*, 12868–12886.

(40) Carlson, R. K.; Li Manni, G.; Sonnenberger, A. L.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory: Barrier Heights and Main Group and Transition Metal Energetics. *J. Chem. Theory Comput.* **2015**, *11*, 82–90.

(41) Goerigk, L.; Grimme, S. A Thorough Benchmark of Density Functional Methods for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.

(42) Becke, A. D. Perspective: Fifty Years of Density-Functional Theory in Chemical Physics. *J. Chem. Phys.* **2014**, *140*, 18A301.

(43) Burke, K. Perspective on Density Functional Theory. *J. Chem. Phys.* **2012**, *136*, 150901.

(44) Kedziora, G. S.; Barr, S. A.; Berry, R.; Moller, J. C.; Breitzman, T. D. Bond Breaking in Stretched Molecules: Multi-Reference Methods versus Density

Functional Theory. *Theor. Chem. Acc.* **2016**, *135*, 79.

(45) Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration Self Consistent Field and Multireference Configuration Interaction Methods and Applications. *Chem. Rev.* **2012**, *112*, 108–181.

(46) Shavitt, I. Matrix Element Evaluation in the Unitary Group Approach to the Electron Correlation Problem. *Int. J. Quantum Chem. Symp.* **1978**, *12*, 5–32.

(47) Veryazov, V.; Malmqvist, P. Å.; Roos, B. O. How to Select Active Space for Multiconfigurational Quantum Chemistry? *Int. J. Quantum Chem.* **2011**, *111*, 3329–3338.

(48) Davidson, E. The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices. *J. Comput. Phys.* **1975**, *17*, 87–94.

(49) Davidson, E. R. Super-Matrix Methods. *Comput. Phys. Commun.* **1989**, *53*, 49–60.

(50) Davidson, E. R. Monster Matrices: Their Eigenvalues and Eigenvectors. *Comput. Phys.* **1993**, *7*, 519–522.

(51) Li Manni, G. New Methods to Treat Strongly Correlated Systems, Thesis, Université de Genève, 2013.

(52) Schmidt, M. W.; Gordon, M. S. The Construction and Interpretation of MCSCF Wavefunctions. *Annu. Rev. Phys. Chem.* **1998**, *49*, 233–266.

(53) Aquilante, F.; Autschbach, J.; Carlson, R. K.; Chibotaru, L. F.; Delcey, M. G.; De Vico, L.; Fdez. Galván, I.; Ferré, N.; Frutos, L. M.; Gagliardi, L.; Garavelli, M.; Giussani, A.; Hoyer, C. E.; Li Manni, G.; Lischka, H.; Ma, D.; Malmqvist, P. Å.; Müller, T.; Nenov, A.; Olivucci, M.; Pedersen, T. B.; Peng, D.; Plasser, F.; Pritchard, B.; Reiher, M.; Rivalta, I.; Schapiro, I.; Segarra-Martí, J.; Stenrup, M.; Truhlar, D. G.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Vysotskiy, V. P.; Weingart, O.; Zapata, F.; Lindh, R. Molcas 8: New Capabilities for Multiconfigurational Quantum Chemical Calculations across the Periodic Table. *J. Comput. Chem.* **2016**, *37*, 506–541.

(54) Vogiatzis, K. D.; Ma, D.; Olsen, J.; Gagliardi, L.; De Jong, W. A. Pushing Configuration-Interaction to the Limit: Towards Massively Parallel MCSCF Calculations. *J. Chem. Phys.* **2017**, *147*, 184111.

(55) Ivanic, J.; Ruedenberg, K. Identification of Deadwood in Configuration Spaces through General Direct Configuration Interaction. *Theor. Chem. Acc.* **2001**, *106*, 339–351.

(56) Ivanic, J.; Ruedenberg, K. Deadwood in Configuration Spaces. II. Singles + Doubles and Singles + Doubles + Triples + Quadruples Spaces. *Theor. Chem. Acc.* **2002**, *107*, 220–228.

(57) Bytautas, L.; Ruedenberg, K. A Priori Identification of Configurational Deadwood. *Chem. Phys.* **2009**, *356*, 64–75.

(58) White, S. R. Density Matrix Formulation for Quantum Renormalization Groups. *Phys. Rev. Lett.* **1992**, *69*, 2863–2866.

(59) White, S. R. Density-Matrix Algorithms for Quantum Renormalization Groups. *Phys. Rev. B* **1993**, *48* (14), 10345–10356.

(60) White, S. R.; Martin, R. L. Ab Initio Quantum Chemistry Using the Density Matrix Renormalization Group. *J. Chem. Phys.* **1999**, *110*, 4127–4130.

(61) Mizukami, W.; Kurashige, Y.; Yanai, T. More π Electrons Make a Difference: Emergence of Many Radicals on Graphene Nanoribbons Studied by *Ab Initio* DMRG Theory. *J. Chem. Theory Comput.* **2013**, *9*, 401–407.

(62) Kurashige, Y. Multireference Electron Correlation Methods with Density Matrix Renormalisation Group Reference Functions. *Mol. Phys.* **2014**, *112* (11), 1485–1494.

(63) Pulay, P. A Perspective on the CASPT2 Method. *Int. J. Quantum Chem.* **2011**, *111*, 3273–3279.

(64) Zobel, J. P.; Nogueira, J. J.; González, L. The IPEA Dilemma in CASPT2. *Chem. Sci.* **2016**, *8*, 1482–1499.

(65) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. A Modified Definition of the Zeroth-Order Hamiltonian in Multiconfigurational Perturbation Theory (CASPT2). *Chem. Phys. Lett.* **2004**, *396*, 142–149.

(66) Pierloot, K.; Phung, Q. M.; Domingo, A. Spin State Energetics in First-Row Transition Metal Complexes: Contribution of (3*s*3*p*) Correlation and Its Description by Second-Order Perturbation Theory. *J. Chem. Theory Comput.* **2017**, *13*, 537–553.

(67) Ruipérez, F.; Aquilante, F.; Ugalde, J. M.; Infante, I. Complete vs Restricted Active Space Perturbation Theory Calculation of the Cr_2 Potential Energy Surface. *J. Chem. Theory Comput.* **2011**, *7*, 1640–1646.

(68) Forsberg, N.; Malmqvist, P.-Å. Multiconfiguration Perturbation Theory with Imaginary Level Shift. *Chem. Phys. Lett.* **1997**, *274*, 196–204.

(69) Ma, D.; Li Manni, G.; Olsen, J.; Gagliardi, L. Second Order Perturbation Theory for Generalized Active Space Self-Consistent-Field Wave Functions. *J. Chem. Theory Comput.* **2016**, *12*, 3208–3213.

(70) Kurashige, Y.; Yanai, T. Second-Order Perturbation Theory with a Density Matrix Renormalization Group Self-Consistent Field Reference Function: Theory and Application to the Study of Chromium Dimer. *J. Chem. Phys.* **2011**, *135*, 094104.

(71) Sand, A. M.; Truhlar, D. G.; Gagliardi, L. Efficient Algorithm for Multiconfiguration Pair-Density Functional Theory with Application to the Heterolytic Dissociation Energy of Ferrocene. *J. Chem. Phys.* **2017**, *146*, 034101.

(72) Sand, A. M.; Truhlar, D. G.; Gagliardi, L. Erratum: "Efficient Algorithm for Multiconfiguration Pair-Density Functional Theory with Application to the Heterolytic Dissociation Energy of Ferrocene" [J. Chem. Phys. 146, 034101 (2017)]. *J. Chem. Phys.* **2017**, *146*, 189901.

(73) Vogiatzis, K. D.; Li Manni, G.; Stoneburner, S. J.; Ma, D.; Gagliardi, L. Systematic Expansion of Active Spaces beyond the CASSCF Limit: A

GASSCF/SplitGAS Benchmark Study. *J. Chem. Theory Comput.* **2015**, *11*, 3010–3021.

(74) Löwdin, P.-O. A Note on the Quantum-Mechanical Perturbation Theory. *J. Chem. Phys.* **1951**, *19*, 1396–1401.

(75) Sharma, P.; Truhlar, D. G.; Gagliardi, L. Active Space Dependence in Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.* **2018**, *14*, 660–669.

(76) Perdew, J. P.; Ernzerhof, M.; Burke, K.; Savin, A. On-Top Pair-Density Interpretation of Spin Density Functional Theory, with Applications to Magnetism. *Int. J. Quantum Chem.* **1997**, *61*, 197–205.

(77) Perdew, J. P.; Savin, A.; Burke, K. Escaping the Symmetry Dilemma through a Pair-Density Interpretation of Spin-Density Functional Theory. *Phys. Rev. A* **1995**, *51*, 4531–4541.

(78) Carlson, R. K.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory: A Fully Translated Gradient Approximation and Its Performance for Transition Metal Dimers and the Spectroscopy of $Re_2Cl_8^{2-}$. *J. Chem. Theory Comput.* **2015**, *11*, 4077–4085.

(79) Becke, A. D.; Savin, A.; Stoll, H. Extension of the Local-Spin-Density Exchange-Correlation Approximation to Multiplet States. *Theor. Chim. Acta* **1995**, *91*, 147–156.

(80) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(81) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(82) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(83) Zhang, Y.; Yang, W. Comment on "Generalized Gradient Approximation Made Simple." *Phys. Rev. Lett.* **1998**, *80*, 890–890.

(84) Sharma, P.; Bernales, V.; Knecht, S.; Truhlar, D. G.; Gagliardi, L. Density Matrix Renormalization Group Pair-Density Functional Theory (DMRG-PDFT): Singlet-Triplet Gaps in Polyacenes and Polyacetylenes. *Chem. Sci.* **2019**, *10*, 1716–1723.

(85) Zhou, C.; Gagliardi, L.; Truhlar, D. G. Multiconfiguration Pair-Density Functional Theory for Iron Porphyrin with CAS, RAS, and DMRG Active Spaces. *J. Phys. Chem. A* **2019**, *123*, 3389–3394.

(86) Sayfutyarova, E. R.; Sun, Q.; Chan, G. K.-L.; Knizia, G. Automated Construction of Molecular Active Spaces from Atomic Valence Orbitals. *J. Chem. Theory Comput.* **2017**, *13*, 4063–4078.

(87) Ryde, U.; Mata, R. A.; Grimme, S. Does DFT-D Estimate Accurate Energies for the Binding of Ligands to Metal Complexes? *Dalt. Trans.* **2011**, *40*, 11176–11183.

(88) Stein, C. J.; Reiher, M. Automated Selection of Active Orbital Spaces. *J. Chem.*

Theory Comput. **2016**, *12*, 1760−1771.

(89) Stein, C. J.; Von Burg, V.; Reiher, M. The Delicate Balance of Static and Dynamic Electron Correlation. *J. Chem. Theory Comput.* **2016**, *12*, 3764–3773.

(90) Keller, S.; Boguslawski, K.; Janowski, T.; Reiher, M.; Pulay, P. Selection of Active Spaces for Multiconfigurational Wavefunctions. *J. Chem. Phys.* **2015**, *142*, 244104.

(91) Bao, J. J.; Dong, S. S.; Gagliardi, L.; Truhlar, D. G. Automatic Selection of an Active Space for Calculating Electronic Excitation Spectra by MS-CASPT2 or MC-PDFT. *J. Chem. Theory Comput.* **2018**, *14*, 2017–2025.

(92) Stein, C. J.; Reiher, M. Automated Identification of Relevant Frontier Orbitals for Chemical Compounds and Processes. **2017**, *71*, 170–176.

(93) Yaghi, M.; Li, H. Hydrothermal Synthesis of a Metal-Organic Framework Containing Large Rectangular Channels. *J. Am. Chem. Soc.* **1995**, *117*, 10401–10402.

(94) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. Reticular Synthesis and the Design of New Materials. *Nature* **2003**, *423*, 705–714.

(95) Eddaoudi, M.; Moler, D. B.; Li, H.; Chen, B.; Reineke, T. M.; O'Keeffe, M.; Yaghi, O. M. Modular Chemistry: Secondary Building Units as a Basis for the Design of Highly Porous and Robust Metal-Organic Carboxylate Frameworks. *Acc. Chem. Res.* **2001**, *34*, 319–330.

(96) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The Chemistry and Applications of Metal-Organic Frameworks. *Science.* **2013**, *341*, 1230444.

(97) Leus, K.; Liu, Y.-Y.; Van Der Voort, P. Metal-Organic Frameworks as Selective or Chiral Oxidation Catalysts. *Cat. Rev. - Sci. Eng.* **2014**, *56*, 1–56.

(98) Dhakshinamoorthy, A.; Opanasenko, M.; Cejka, J.; Garcia, H. Metal Organic Frameworks as Heterogeneous Catalysts for the Production of Fine Chemicals. *Catal. Sci. Technol.* **2013**, *3*, 2509–2540.

(99) Lee, J.; Farha, O. K.; Roberts, J.; Scheidt, K. A.; Nguyen, S. T.; Hupp, J. T. Metal-Organic Framework Materials as Catalysts. *Chem. Soc. Rev.* **2009**, *38*, 1450–1459.

(100) Xiao, D. J.; Bloch, E. D.; Mason, J. A.; Queen, W. L.; Hudson, M. R.; Planas, N.; Borycz, J.; Dzubak, A. L.; Verma, P.; Lee, K.; Bonino, F.; Crocellà, V.; Yano, J.; Bordiga, S.; Truhlar, D. G.; Gagliardi, L.; Brown, C. M.; Long, J. R. Oxidation of Ethane to Ethanol by N_2O in a Metal-Organic Framework with Coordinatively Unsaturated Iron(II) Sites. *Nat. Chem.* **2014**, *6*, 590–595.

(101) Jiao, L.; Wang, Y.; Jiang, H.-L.; Xu, Q. Metal–Organic Frameworks as Platforms for Catalytic Applications. *Adv. Mater.* **2018**, *30*, 1703663.

(102) Yang, D.; Gates, B. C. Catalysis by Metal Organic Frameworks: Perspective and Suggestions for Future Research. *ACS Catal.* **2019**, *9*, 1779–1798.

(103) Trickett, C. A.; Helal, A.; Al-Maythalony, B. A.; Yamani, Z. H.; Cordova, K. E.; Yaghi, O. M. The Chemistry of Metal-Organic Frameworks for CO_2 Capture, Regeneration and Conversion. *Nat. Rev. Mater.* **2017**, *2*, 17045.

(104) Kim, K. C.; Moghadam, P. Z.; Fairen-Jimenez, D.; Snurr, R. Q. Computational Screening of Metal Catecholates for Ammonia Capture in Metal–Organic Frameworks. *Ind. Eng. Chem. Res.* **2015**, *54*, 3257–3267.

(105) Dzubak, A. L.; Lin, L.-C.; Kim, J.; Swisher, J. A.; Poloni, R.; Maximoff, S. N.; Smit, B.; Gagliardi, L. *Ab Initio* Carbon Capture in Open-Site Metal–Organic Frameworks. *Nat. Chem.* **2012**, *4*, 810–816.

(106) Lee, K.; Isley, W. C.; Dzubak, A. L.; Verma, P.; Stoneburner, S. J.; Lin, L.-C. C.; Howe, J. D.; Bloch, E. D.; Reed, D. A.; Hudson, M. R.; Brown, C. M.; Long, J. R.; Neaton, J. B.; Smit, B.; Cramer, C. J.; Truhlar, D. G.; Gagliardi, L. Design of a Metal-Organic Framework with Enhanced Back Bonding for Separation of N_2 and CH_4. *J. Am. Chem. Soc.* **2014**, *136*, 698–704.

(107) Liu, J.; Thallapally, P. K.; McGrail, B. P.; Brown, D. R.; Liu, J. Progress in Adsorption-Based CO_2 Capture by Metal–Organic Frameworks. *Chem. Soc. Rev.* **2012**, *41*, 2308–2322.

(108) Yang, Q.; Liu, D.; Zhong, C.; Li, J.-R. Development of Computational Methodologies for Metal-Organic Frameworks and Their Application in Gas Separations. *Chem. Rev.* **2013**, *113*, 8261–8323.

(109) Bloch, E. D.; Murray, L. J.; Queen, W. L.; Chavan, S.; Maximoff, S. N.; Bigi, J. P.; Krishna, R.; Peterson, V. K.; Grandjean, F.; Long, G. J.; Smit, B.; Bordiga, S.; Brown, C. M.; Long, J. R. Selective Binding of O_2 over N_2 in a Redox-Active Metal-Organic Framework with Open Iron(II) Coordination Sites. *J. Am. Chem. Soc.* **2011**, *133*, 14814–14822.

(110) DeCoste, J. B.; Peterson, G. W. Metal–Organic Frameworks for Air Purification of Toxic Chemicals. *Chem. Rev.* **2014**, *114*, 5695–5727.

(111) Barea, E.; Montoro, C.; Navarro, J. A. R. Toxic Gas Removal-Metal-Organic Frameworks for the Capture and Degradation of Toxic Gases and Vapours. *Chem. Soc. Rev.* **2014**, *43*, 5419–5430.

(112) Zhao, X.; Wang, Y.; Li, D. S.; Bu, X.; Feng, P. Metal–Organic Frameworks for Separation. *Adv. Mater.* **2018**, *30*, 1705189.

(113) Tsivion, E.; Long, J. R.; Head-Gordon, M. Hydrogen Physisorption on Metal − Organic Framework Linkers and Metalated Linkers: A Computational Study of the Factors That Control Binding Strength. *J. Am. Chem. Soc.* **2014**, *136*, 17827–17835.

(114) Getman, R. B.; Bae, Y.; Wilmer, C. E.; Snurr, R. Q. Review and Analysis of Molecular Simulations of Methane , Hydrogen , and Acetylene Storage in Metal-Organic Frameworks Supporting Information. *Chem. Rev* **2012**, *112*, 703–723.

(115) Bloch, E. D.; Queen, W. L.; Chavan, S.; Wheatley, P. S.; Zadrozny, J. M.; Morris, R.; Brown, C. M.; Lamberti, C.; Bordiga, S.; Long, J. R. Gradual Release of Strongly Bound Nitric Oxide from $Fe_2(NO)_2$(Dobdc). *J. Am. Chem. Soc.* **2015**, *137*, 3466–3469.

(116) Hu, Q.; Yu, J.; Liu, M.; Liu, A.; Dou, Z.; Yang, Y. A Low Cytotoxic Cationic Metal-Organic Framework Carrier for Controllable Drug Release. *J. Med. Chem.*

2014, *57*, 5679–5685.

(117) Chen, B.; Xiang, S.; Qian, G. Metal-Organic Frameworks with Functional Pores for Recognition of Small Molecules. *Acc. Chem. Res.* **2010**, *43*, 1115–1124.

(118) Kreno, L. E.; Leong, K.; Farha, O. K.; Allendorf, M.; Van Richard P., D.; Hupp, J. T. Metal-Organic Framework Materials as Chemical Sensors. *Chem. Rev.* **2012**, *112*, 1105–1125.

(119) Meek, S. T.; Greathouse, J. A.; Allendorf, M. D. Metal-Organic Frameworks: A Rapidly Growing Class of Versatile Nanoporous Materials. *Adv. Mater.* **2011**, *23*, 249–267.

(120) Liu, L.; Zhou, Y.; Liu, S.; Xu, M. The Applications of Metal−Organic Frameworks in Electrochemical Sensors. *ChemElectroChem* **2018**, *5*, 6–19.

(121) Campbell, M. G.; Dincă, M. Metal–Organic Frameworks as Active Materials in Electronic Sensor Devices. *Sensors (Switzerland)* **2017**, *17*, 1108.

(122) Evans, J. D.; Sumby, C. J.; Doonan, C. J. Post-Synthetic Metalation of Metal–Organic Frameworks. *Chem. Soc. Rev.* **2014**, *43*, 5933–5951.

(123) Islamoglu, T.; Goswami, S.; Li, Z.; Howarth, A. J.; Farha, O. K.; Hupp, J. T. Postsynthetic Tuning of Metal-Organic Frameworks for Targeted Applications. *Acc. Chem. Res.* **2017**, *50*, 805–813.

(124) Silva, P.; Vilela, S. M. F.; Tomé, J. P. C.; Almeida Paz, F. A. Multifunctional Metal-Organic Frameworks: From Academia to Industrial Applications. *Chem. Soc. Rev.* **2015**, *44*, 6774–6803.

(125) Colón, Y. J.; Snurr, R. Q. High-Throughput Computational Screening of Metal-Organic Frameworks. *Chem. Soc. Rev.* **2014**, *43*, 5735–5749.

(126) Fei, H.; Cohen, S. M. Metalation of a Thiocatechol-Functionalized Zr(IV)-Based Metal–Organic Framework for Selective C–H Functionalization. *J. Am. Chem. Soc.* **2015**, *137*, 2191–2194.

(127) Mondloch, J. E.; Bury, W.; Fairen-Jimenez, D.; Kwon, S.; DeMarco, E. J.; Weston, M. H.; Sarjeant, A. A.; Nguyen, S. T.; Stair, P. C.; Snurr, R. Q.; Farha, O. K.; Hupp, J. T. Vapor-Phase Metalation by Atomic Layer Deposition in a Metal--Organic Framework. *J. Am. Chem. Soc.* **2013**, *135*, 10294–10297.

(128) Nguyen, H. G. T.; Weston, M. H.; Farha, O. K.; Hupp, J. T.; Nguyen, S. T. A Catalytically Active Vanadyl(Catecholate)-Decorated Metal Organic Framework via Post-Synthesis Modifications. *CrystEngComm* **2012**, *14*, 4115–4118.

(129) Weston, M. H.; Peterson, G. W.; Browe, M. A.; Jones, P.; Farha, O. K.; Hupp, J. T.; Nguyen, S. T. Removal of Airborne Toxic Chemicals by Porous Organic Polymers Containing Metal–Catecholates. *Chem. Commun.* **2013**, *49*, 2995–2997.

(130) Hashim, S. S.; Mohamed, A. R.; Bhatia, S. Oxygen Separation from Air Using Ceramic-Based Membrane Technology for Sustainable Fuel Production and Power Generation. *Renew. Sustain. Energy Rev.* **2011**, *15*, 1284–1293.

(131) Chung, Y. G.; Camp, J.; Haranczyk, M.; Sikora, B. J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O. K.; Sholl, D. S.; Snurr, R. Q. Computation-Ready,

Experimental Metal–Organic Frameworks: A Tool To Enable High-Throughput Screening of Nanoporous Crystals. *Chem. Mater.* **2014**, *26*, 6185–6192.

(132) Löwdin, P. O. Quantum Theory of Many-Particle Systems. III. Extension of the Hartree-Fock Scheme to Include Degenerate Systems and Correlation Effects. *Phys. Rev.* **1955**, *97*, 1509–1520.

(133) Cremer, D. Møller-Plesset Perturbation Theory: From Small Molecule Methods to Methods for Thousands of Atoms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 509–530.

(134) Bartlett, R. J.; Musiał, M. Coupled-Cluster Theory in Quantum Chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

(135) González, L.; Escudero, D.; Serrano-Andrés, L. Progress and Challenges in the Calculation of Electronic Excited States. *ChemPhysChem* **2012**, *13*, 28–51.

(136) Malrieu, J. P.; Caballol, R.; Calzado, C. J.; de Graaf, C.; Guihéry, N. Magnetic Interactions in Molecules and Highly Correlated Materials: Physical Content, Analytical Derivation, and Rigorous Extraction of Magnetic Hamiltonians. *Chem. Rev.* **2014**, *114*, 429–492.

(137) Hicks, J.; Hoyer, C. E.; Moubaraki, B.; Manni, G. L.; Carter, E.; Murphy, D. M.; Murray, K. S.; Gagliardi, L.; Jones, C. A Two-Coordinate Manganese(0) Complex with an Unsupported Mn-Mg Bond: Allowing Access to Low Coordinate Homo- and Heterobimetallic Compounds. *J. Am. Chem. Soc.* **2014**, *136*, 5283–5286.

(138) Brogden, D. W.; Turov, Y.; Nippe, M.; Li Manni, G.; Hillard, E. A.; Clérac, R.; Gagliardi, L.; Berry, J. F. Oxidative Stretching of Metal-Metal Bonds to Their Limits. *Inorg. Chem.* **2014**, *53*, 4777–4790.

(139) Buenker, R. J.; Peyerimhoff, S. D.; Butscher, W. Applicability of the Multi-Reference Double-Excitation CI (MRD-CI) Method to the Calculation of Electronic Wavefunctions and Comparison with Related Techniques. *Mol. Phys.* **1978**, *35*, 771–791.

(140) Huron, B.; Malrieu, J. P.; Rancurel, P. Iterative Perturbation Calculations of Ground and Excited State Energies from Multiconfigurational Zeroth-Order Wave Functions. *J. Chem. Phys.* **1973**, *58*, 5745–5759.

(141) Cave, R. J.; Xantheas, S. S.; Feller, D. Exploiting Regularity in Systematic Sequences of Wavefunctions Which Approach the Full CI Limit. *Theor. Chim. Acta* **1992**, *83*, 31–55.

(142) Hanrath, M.; Engels, B. New Algorithms for an Individually Selecting MR-CI Program. *Chem. Phys.* **1997**, *225*, 197–202.

(143) Olsen, J.; Roos, B. O.; Jørgensen, P.; Jensen, H. J. A. Determinant Based Configuration Interaction Algorithms for Complete and Restricted Configuration Interaction Spaces. *J. Chem. Phys.* **1988**, *89*, 2185–2192.

(144) Rolik, Z.; Szabados, Á.; Surján, Ṕter R. A Sparse Matrix Based Full-Configuration Interaction Algorithm. *J. Chem. Phys.* **2008**, *128*, 144101.

(145) Fleig, T.; Olsen, J.; Visscher, L. The Generalized Active Space Concept for the

Relativistic Treatment of Electron Correlation. II. Large-Scale Configuration Interaction Implementation Based on Relativistic 2- and 4-Spinors and Its Application. *J. Chem. Phys.* **2003**, *119*, 2963–2971.

(146) Ivanic, J. Direct Configuration Interaction and Multiconfigurational Self-Consistent-Field Method for Multiple Active Spaces with Variable Occupations . I . Method. *J. Chem. Phys.* **2003**, *119*, 9364–9376.

(147) Li Manni, G.; Aquilante, F.; Gagliardi, L. Strong Correlation Treated via Effective Hamiltonians and Perturbation Theory. *J. Chem. Phys.* **2011**, *134*, 034114.

(148) Nitzsche, L. E.; Davidson, E. R. A Perturbation Theory Calculation on the $1\pi\pi*$ State of Formamide. *J. Chem. Phys.* **1978**, *68*, 3103–3109.

(149) Staroverov, V. N.; Davidson, E. R. The Reduced Model Space Method in Multireference Second-Order Perturbation Theory. *Chem. Phys. Lett.* **1998**, *296*, 435–444.

(150) Shavitt, I. The A_k and B_k Approximate CI Methods. Comment on a Paper by Maynau and Heully. *Chem. Phys. Lett.* **1992**, *192*, 135–137.

(151) Mayhall, N. J.; Goldey, M.; Head-Gordon, M. A Quasidegenerate Second-Order Perturbation Theory Approximation to RAS-NSF for Excited States and Strong Correlations. *J. Chem. Theory Comput.* **2014**, *10*, 589–599.

(152) Mayhall, N. J.; Head-Gordon, M. Computational Quantum Chemistry for Single Heisenberg Spin Couplings Made Simple: Just One Spin Flip Required. *J. Chem. Phys.* **2014**, *141*, 134111.

(153) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Neogrády, P.; Pedersen, T. B.; Pitoňák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. MOLCAS 7: The Next Generation. *J. Comput. Chem.* **2010**, *31*, 224–247.

(154) Olsen, J.; Jørgensen, P.; Simons, J. Passing the One-Billion Limit in Full Configuration-Interaction (FCI) Calculations. *Chem. Phys. Lett.* **1990**, *169*, 463–472.

(155) Werner, H.; Knowles, P. J. An Efficient Internally Contracted Multiconfiguration–Reference Configuration Interaction Method. *J. Chem. Phys.* **1988**, *89*, 5803–5814.

(156) Knowles, P. J.; Werner, H. J. An Efficient Method for the Evaluation of Coupling Coefficients in Configuration Interaction Calculations. *Chem. Phys. Lett.* **1988**, *145*, 514–522.

(157) Werner, H. J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: A General-Purpose Quantum Chemistry Program Package. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 242–253.

(158) Roos, B. O.; Lindh, R.; Malmqvist, P. Å.; Veryazov, V.; Widmark, P. O. Main Group Atoms and Dimers Studied with a New Relativistic ANO Basis Set. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.

(159) Roos, B. O.; Lindh, R.; Malmqvist, P. Å.; Veryazov, V.; Widmark, P. O. New Relativistic ANO Basis Sets for Transition Metal Atoms. *J. Phys. Chem. A* **2005**,

109, 6575–6579.

(160) Brand, J. C. D.; Cross, K. J.; Hoy, A. R. The Huggins Bands of Ozone. *Can. J. Phys.* **1978**, *56*, 327–333.

(161) Swanson, N.; Celotta, R. J. Observation of Excited States in Ozone near the Dissociation Limit. *Phys. Rev. Lett.* **1975**, *35* (12), 783–785.

(162) VonRosenberg, C. W.; Trainor, D. W. Excitation of Ozone Formed by Recombination. II. *J. Chem. Phys.* **1975**, *63* (12), 5348–5353.

(163) McGrath, W. D.; Maguire, J. M.; Thompson, A.; Trocha-Grimshaw, J. The Production of Electronically Excited Ozone by Flash Irradiation. *Chem. Phys. Lett.* **1983**, *102* (1), 59–63.

(164) Anderson, S. M.; Morton, J.; Mauersberger, K. Near-Infrared Absorption Spectra of $^{16}O_3$ and $^{18}O_3$: Adiabatic Energy of the $^1A_2$ State? *J. Chem. Phys.* **1990**, *93*, 3826–3832.

(165) Arnold, D. W.; Xu, C.; Kim, E. H.; Neumark, D. M. Study of Low-Lying Electronic States of Ozone by Anion Photoelectron Spectroscopy of O^-_3. *J. Chem. Phys.* **1994**, *101*, 912–922.

(166) Günther, J.; Anderson, S. M.; Hilpert, G.; Mauersberger, K. Rotational Structure in the Absorption Spectra of $^{18}O_3$ and $^{16}O_3$ near 1 mm: A Comparative Study of the $^3A_2$ and $^3B_2$ States. *J. Chem. Phys.* **1997**, *108*, 5449–5457.

(167) Steinfeld, J. I.; Adler-Golden, S. M.; Gallagher, J. W. Critical Survey of Data on the Spectroscopy and Kinetics of Ozone in the Mesosphere and Thermosphere. *J. Phys. Chem. Ref. Data* **1987**, *16*, 911–951.

(168) Thunemann, K. H.; Peyerimhoff, S. D.; Buenker, R. J. Configuration Interaction Calculations for the Ground and Excited States of Ozone and Its Positive Ion: Energy Locations and Transition Probabilities. *J. Mol. Spectrosc.* **1978**, *70*, 432–448.

(169) Borowski, P.; Fiilscher, M.; Malmqvist, P.-Å.; Roos, B. O. A Theoretical Study of the Low-Lying Excited States of Ozone. *Chem. Phys. Lett.* **1995**, *237*, 195–203.

(170) Szalay, P. G.; Bartlett, R. J. Analytic Energy Gradients for the Two-Determinant Coupled Cluster Method with Application to Singlet Excited States of Butadiene and Ozone. *J. Chem. Phys.* **1994**, *101*, 4936–4944.

(171) Banichevich, A.; Peyerimhoff, S. D. Theoretical Study of the Ground and Excited States of Ozone in Its Symmetric Nuclear Arrangement. *Chem. Phys.* **1993**, *174*, 93–109.

(172) Musiał, M.; Kucharski, S. A.; Zerzucha, P.; Kuś, T.; Bartlett, R. J. Excited and Ionized States of the Ozone Molecule with Full Triples Coupled Cluster Methods. *J. Chem. Phys.* **2009**, *131*, 194104.

(173) Miliordos, E.; Xantheas, S. S. On the Bonding Nature of Ozone (O_3) and Its Sulfur-Substituted Analogues SO_2, OS_2, and S_3: Correlation between Their Biradical Character and Molecular Properties. *J. Am. Chem. Soc.* **2014**, *136*, 2808–2817.

(174) Schmelz, T.; Chambaud, G.; Rosmus, P.; Köppel, H.; Cederbaum, L.; Werner, H.-J. Electronic States of the O^{3+} Radical Cation. *Chem. Phys. Lett.* **1991**, *183*, 209–216.

(175) Koch, W.; Frenking, G.; Steffen, G.; Reinen, D.; Jansen, M.; Assenmacher, W. The Low Lying Electronic States of O^-_3. *J. Chem. Phys* **1993**, *99*, 1271–1277.

(176) Palmer, M. H.; Nelson, A. D. An Ab Initio Molecular Orbital Study of the Electronically Excited and Cationic States of the Ozone Molecule and a Comparison with Spectral Data. *Mol. Phys.* **2002**, *100*, 3601–3614.

(177) Elliott, R.; Compton, R.; Levis, R.; Matsika, S. Excited Electronic States of the Cyclic Isomers of O_3 and SO_2. *J. Phys. Chem. A* **2005**, *109*, 11304–11311.

(178) Bag, A.; Manohar, P. U.; Vaval, N.; Pal, S. First- and Second-Order Electrical Properties Computed at the FSMRCCSD Level for Excited States of Closed-Shell Molecules Using the Constrained-Variational Approach. *J. Chem. Phys.* **2009**, *131*, 024102.

(179) Barysz, M.; Rittby, M.; Bartlett, R. J. Fock Space Multi-Reference Coupled-Cluster Study of Excitation Energies and Dipole Oscillator Strengths of Ozone. *Chem. Phys. Lett.* **1992**, *193* (5), 373–379.

(180) Vaval, N.; Pal, S. Adiabatic States of Ozone Using Fock Space Multireference Coupled Cluster Method. *J. Chem. Phys.* **1999**, *111*, 4051–4055.

(181) Kowalski, K.; Piecuch, P. New Coupled-Cluster Methods with Singles, Doubles, and Noniterative Triples for High Accuracy Calculations of Excited Electronic States. *J. Chem. Phys.* **2004**, *120*, 1715–1738.

(182) Tsuneda, T.; Nakano, H.; Hirao, K. Study of Low-Lying Electronic States of Ozone by Multireference Møller-Plesset Perturbation Method. *J. Chem. Phys.* **1995**, *103*, 6520–6528.

(183) Decleva, P.; De Alti, G.; Lisini, A. Theoretical Study of the Valence Photoelectron Spectrum of Ozone: An Analysis of Correlation Effects and Configuration Interaction (CI) Model Spaces. *J. Chem. Phys.* **1988**, *89*, 367–373.

(184) Lias, S. G.; Bartmess, J. E.; Liebman, J. F.; Holmes, J. L.; Levin, R. D.; Mallard, W. G. *Gas-Phase Ion and Neutral Thermochemistry*; Published by the American Chemical Society and the American Institute of Physics for the National Bureau of Standards, 1988.

(185) Aquilante, F.; Bondo Pedersen, T.; Sánchez de Merás, A.; Koch, H. Fast Noniterative Orbital Localization for Large Molecules. *J. Chem. Phys.* **2006**, *125*, 174101.

(186) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Roos, B. O.; Fülscher, M. Theoretical Study of the Electronic Spectra of Cyclopentadiene, Pyrrole, and Furan. *J. Am. Chem. Soc.* **1993**, *115*, 6184–6197.

(187) Flicker, W. M.; Mosher, O. A.; Kuppermann, A. Electron Impact Investigation of Electronic Excitations in Furan, Thiophene, and Pyrrole. *J. Chem. Phys.* **1976**, *64*, 1315–1321.

(188) Fülscher, M. P.; Roos, B. O. Theoretical Study of the Electronic Spectrum of Cytosine. *J. Am. Chem. Soc.* **1995**, *117*, 2089–2095.

(189) Serrano-Andrés, L.; Roos, B. O. Theoretical Study of the Absorption and Emission Spectra of Indole in the Gas Phase and in a Solvent. *J. Am. Chem. Soc.* **1996**, *118*, 185–195.

(190) Hübner, O.; Himmel, H.-J. Cyclic and Linear NiO_2: A Multireference Configuration Interaction Study. *J. Phys. Chem. A* **2012**, *116*, 9181–9188.

(191) Citra, A.; Chertihin, G. V.; Andrews, L.; Neurock, M. Reactions of Laser-Ablated Nickel Atoms with Dioxygen. Infrared Spectra and Density Functional Calculations of Nickel Oxides NiO, ONiO, $Ni2O_2$, and Ni_2O_3, Superoxide NiOO, Peroxide $Ni(O_2)$, and Higher C. *J. Phys. Chem. A* **1997**, *101*, 3109–3118.

(192) Allouti, F.; Manceron, L.; Alikhani, M. E. The Ni + O_2 Reaction: A Combined IR Matrix Isolation and Theoretical Study of the Formation and Structure of NiO_2. *Phys. Chem. Chem. Phys.* **2006**, *8*, 448–455.

(193) Wu, H.; Wang, L. S. A Study of Nickel Monoxide (NiO), Nickel Dioxide (ONiO), and $Ni(O_2)$ Complex by Anion Photoelectron Spectroscopy. *J. Chem. Phys.* **1997**, *107*, 16–21.

(194) Ramond, T. M.; Davico, G. E.; Hellberg, F.; Svedberg, F.; Salén, P.; Söderqvist, P.; Lineberger, W. C. Photoelectron Spectroscopy of Nickel, Palladium, and Platinum Oxide Anions. *J. Mol. Spectrosc.* **2002**, *216*, 1–14.

(195) Blomberg, M. R. A.; Siegbahn, P. E. M.; Strich, A. A Theoretical Study of the Interaction between Nickel and Oxygen. *Chem. Phys.* **1985**, *97*, 287–301.

(196) Bauschlicher, C. W. What Is the Ground State of $Ni(O_2)$? *J. Phys. Chem. A* **2004**, *108*, 2871–2873.

(197) Song, J.; Aprá, E.; Khait, Y. G.; Hoffmann, M. R.; Kowalski, K. High-Level Ab Initio Calculations on the NiO_2 System. *Chem. Phys. Lett.* **2006**, *428*, 277–282.

(198) Gutsev, G. L.; Rao, B. K.; Jena, P. Systematic Study of Oxo, Peroxo, and Superoxo Isomers of 3*d*-Metal Dioxides and Their Anions. *J. Phys. Chem. A* **2000**, *104*, 11961–11971.

(199) Pierloot, K. The CASPT2 Method in Inorganic Electronic Spectroscopy: From Ionic Transition Metal to Covalent Actinide Complexes. *Mol. Phys.* **2003**, *101*, 2083–2094.

(200) Deng, K.; Yang, J.; Zhu, Q. A Theoretical Study of the NiO2 Species. *J. Chem. Phys.* **2003**, *118*, 6868–6873.

(201) Yancoillie, S.; Pierloot, K. Multiconfigurational g Tensor Calculations as a Probe for the Covalency of the Copper-Ligand Bonds in Copper(II) Complexes: $[CuCl_4]^{2-}$, $[Cu(NH_3)_4]^{2+}$, and Plastocyanin. *J. Phys. Chem. A* **2008**, *112*, 4011–4019.

(202) Sauri, V.; Serrano-Andrés, L.; Shahi, A. R. M.; Gagliardi, L.; Vancoillie, S.; Pierloot, K. Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2) Method for Electronic Excited States: A Benchmark

Study. *J. Chem. Theory Comput.* **2011**, *7*, 153–168.

(203) Vancoillie, S.; Malmqvist, P.-Å.; Pierloot, K. Calculation of EPR g Tensors for Transition-Metal Complexes Based on Multiconfigurational Perturbation Theory (CASPT2). *ChemPhysChem* **2007**, *8*, 1803–1815.

(204) Chow, C.; Chang, K.; Willett, R. D. Electron Spin Resonance Spectra and Covalent Bonding in the Square-Planar $CuCl_4^{2-}$ and $CuBr_4^{2-}$ Ions. *J. Chem. Phys.* **1973**, *59*, 2629–2640.

(205) Szilagyi, R. K.; Metz, M.; Solomon, E. I. Spectroscopic Calibration of Modern Density Functional Methods Using $[CuCl_4]^{2-}$. *J. Phys. Chem. A* **2002**, *106*, 2994–3007.

(206) Stoneburner, S. J.; Shen, J.; Ajala, A. O.; Piecuch, P.; Truhlar, D. G.; Gagliardi, L. Systematic Design of Active Spaces for Multi-Reference Calculations of Singlet-Triplet Gaps of Organic Diradicals, with Benchmarks against Doubly Electron-Attached Coupled-Cluster Data. *J. Chem. Phys.* **2017**, *147*, 164120.

(207) Pedersen, S.; Herek, J. L.; Zewail, A. H. The Validity of the "Diradical" Hypothesis: Direct Femtoscond Studies of the Transition-State Structures. *Science.* **1994**, *266* (5189), 1359–1364.

(208) Ko, C.-C.; Yam, V. W.-W. Photochromic Transitional Metal Complexes for Photosensitization. In *Photochromic Materials: Preparation, Properties and Applications*; Tian, H., Zhang, J., Eds.; Wiley: Hoboken, NJ, 2016; pp 47–70.

(209) Cho, D.; Ko, K. C.; Lee, J. Y. Quantum Chemical Approaches for Controlling and Evaluating Intramolecular Magnetic Interactions in Organic Diradicals. *Int. J. Quantum Chem.* **2016**, *116*, 578–597.

(210) Davis, R. M.; Sowers, A. L.; Degraff, W.; Bernardo, M.; Thetford, A.; Krishna, M. C.; Mitchell, J. B. A Novel Nitroxide Is an Effective Brain Redox Imaging Contrast Agent and in Vivo Radioprotector. *Free Radic. Biol. Med.* **2011**, *51*, 780–790.

(211) Sugawara, T.; Komatsu, H.; Suzuki, K. Interplay between Magnetism and Conductivity Derived from Spin-Polarized Donor Radicals. *Chem. Soc. Rev.* **2011**, *40*, 3105–3118.

(212) Sanvito, S. Molecular Spintronics. *Chem. Soc. Rev.* **2011**, *40*, 3336–3355.

(213) Nakano, M.; Champagne, B. Nonlinear Optical Properties in Open-Shell Molecular Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2016**, *6*, 198–210.

(214) Minami, T.; Nakano, M. Diradical Character View of Singlet Fission. *J. Phys. Chem. Lett.* **2012**, *3*, 145–150.

(215) Melo, J. S. S. de; Burrows, H. D.; Pina, J. Triplet States on π-Conjugated Polymers, Oligomers and Related Materials. In *Photochemistry: Volume 43*; Fasani, E., Albini, A., Eds.; Royal Society of Chemistry: Cambridge, 2016; pp 83–102.

(216) Li, X.; Tang, M. L. Triplet Transport in Thin Films: Fundamentals and Applications. *Chem. Commun.* **2017**, *53*, 4429–4440.

(217) Hedley, G. J.; Ruseckas, A.; Samuel, I. D. W. Light Harvesting for Organic Photovoltaics. *Chem. Rev.* **2017**, *117*, 796–837.

(218) Niklas, J.; Poluektov, O. G. Charge Transfer Processes in OPV Materials as Revealed by EPR Spectroscopy. *Adv. Energy Mater.* **2017**, *7*, 1602226.

(219) Gallagher, N. M.; Olankitwanit, A.; Rajca, A. High-Spin Organic Molecules. *J. Org. Chem.* **2015**, *80*, 1291–1298.

(220) Ibeji, C. U.; Ghosh, D. Singlet-Triplet Gaps in Polyacenes: A Delicate Balance between Dynamic and Static Correlations Investigated by Spin-Flip Methods. *Phys. Chem. Chem. Phys.* **2015**, *17*, 9849–9856.

(221) Ess, D. H.; Johnson, E. R.; Hu, X.; Yang, W. Singlet-Triplet Energy Gaps for Diradicals from Fractional-Spin Density-Functional Theory. *J. Phys. Chem. A* **2011**, *115*, 76–83.

(222) Garza, A. J.; Jiménez-Hoyos, C. A.; Scuseria, G. E. Electronic Correlation without Double Counting via a Combination of Spin Projected Hartree-Fock and Density Functional Theories. *J. Chem. Phys.* **2014**, *140*, 244102.

(223) Demel, O.; Shamasundar, K. R.; Kong, L.; Nooijen, M. Application of Double Ionization State-Specific Equation of Motion Coupled Cluster Method to Organic Diradicals. *J. Phys. Chem. A* **2008**, *112*, 11895–11902.

(224) Abe, M. Diradicals. *Chem. Rev.* **2013**, *113*, 7011–7088.

(225) Shen, J.; Piecuch, P. Merging Active-Space and Renormalized Coupled-Cluster Methods via the CC(P;Q) Formalism, with Benchmark Calculations for Singlet-Triplet Gaps in Biradical Systems. *J. Chem. Theory Comput.* **2012**, *8*, 4968–4988.

(226) Zeng, Z.; Shi, X.; Chi, C.; López Navarrete, J. T.; Casado, J.; Wu, J. Pro-Aromatic and Anti-Aromatic π-Conjugated Molecules: An Irresistible Wish to Be Diradicals. *Chem. Soc. Rev.* **2015**, *44*, 6578–6596.

(227) Varga, Z.; Truhlar, D. G. B2N2O4: Prediction of a Magnetic Ground State for a Light Main-Group Molecule. *Inorg. Chem.* **2015**, *54*, 8552–8559.

(228) Varga, Z.; Truhlar, D. G. Singlet-Triplet Competition in the Low-Lying Energy States of $C_4O_{4-n}S_n$ (n = 1-3) Molecules. *Struct. Chem.* **2015**, *26*, 1229–1240.

(229) Sancho-García, J. C.; Pittner, J.; Čársky, P.; Hubač, I. Multireference Coupled-Cluster Calculations on the Energy of Activation in the Automerization of Cyclobutadiene: Assessment of the State-Specific Multireference Brillouin–Wigner Theory. *J. Chem. Phys.* **2000**, *112*, 8785.

(230) Coester, F. Bound States of a Many-Particle System. *Nucl. Phys.* **1958**, *7*, 421–424.

(231) Coester, F.; Kümmel, H. Short-Range Correlations in Nuclear Wave Functions. *Nucl. Phys.* **1960**, *17*, 477–485.

(232) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256.

(233) Čížek, J. On the Use of the Cluster Expansion and the Technique of Diagrams in

Calculations of Correlation Effects in Atoms and Molecules. In *Advances in Chemical Physics, v. 14*; LeFebvre, R., Moser, C., Eds.; John Wiley & Sons, Inc., 1969; pp 35–89.

(234) Čížek, J.; Paldus, J. Correlation Problems in Atomic and Molecular Systems III. Rederivation of the Coupled-Pair Many-Electron Theory Using the Traditional Quantum Chemical Methodst. *Int. J. Quantum Chem.* **1971**, *5*, 359–379.

(235) Paldus, J.; Čížek, J.; Shavitt, I. Correlation Problems in Atomic and Molecular Systems. IV. Extended Coupled-Pair Many-Electron Theory and Its Application to the BH_3 Molecule. *Phys. Rev. A* **1972**, *5*, 50–67.

(236) Purvis, G. D.; Bartlett, R. J. A Full Coupled-Cluster Singles and Doubles Model: The Inclusion of Disconnected Triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(237) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(238) Ghosh, S.; Cramer, C. J.; Truhlar, D. G.; Gagliardi, L. Generalized-Active-Space Pair-Density Functional Theory: An Efficient Method to Study Large, Strongly Correlated, Conjugated Systems. *Chem. Sci.* **2017**, *8*, 2741–2750.

(239) Saito, T.; Nishihara, S.; Yamanaka, S.; Kitagawa, Y.; Kawakami, T.; Yamada, S.; Isobe, H.; Okumura, M.; Yamaguchi, K. Symmetry and Broken Symmetry in Molecular Orbital Description of Unstable Molecules IV: Comparison between Single- and Multi-Reference Computational Results for Antiaromtic Molecules. *Theor. Chem. Acc.* **2011**, *130*, 749–763.

(240) Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. Are Atoms Intrinsic to Molecular Electronic Wavefunctions? I. The FORS Model. *Chem. Phys.* **1982**, *71*, 41–49.

(241) Gagliardi, L.; Truhlar, D. G.; Manni, G. L.; Carlson, R. K.; Hoyer, C. E.; Bao, J. L. Multiconfiguration Pair-Density Functional Theory: A New Way To Treat Strongly Correlated Systems. *Acc. Chem. Res.* **2017**, *50*, 66–73.

(242) Bao, J. L.; Sand, A.; Gagliardi, L.; Truhlar, D. G. Correlated-Participating-Orbitals Pair-Density Functional Method and Application to Multiplet Energy Splittings of Main-Group Divalent Radicals. *J. Chem. Theory Comput.* **2016**, *12*, 4274–4283.

(243) Bao, J. L.; Odoh, S. O.; Gagliardi, L.; Truhlar, D. G. Predicting Bond Dissociation Energies of Transition-Metal Compounds by Multiconfiguration Pair-Density Functional Theory and Second-Order Perturbation Theory Based on Correlated Participating Orbitals and Separated Pairs. *J. Chem. Theory Comput.* **2017**, *13*, 616–626.

(244) Mahapatra, U. S.; Datta, B.; Mukherjee, D. A Size-Consistent State-Specific Multireference Coupled Cluster Theory: Formal Developments and Molecular Applications. *J. Chem. Phys.* **1999**, *110*, 6171–6188.

(245) Balková, A.; Bartlett, R. J. A Multireference Coupled-Cluster Study of the Ground State and Lowest Excited States of Cyclobutadiene. *J. Chem. Phys.* **1994**, *101*, 8972–8987.

(246) Eckert-Maksić, M.; Vazdar, M.; Barbatti, M.; Lischka, H.; Maksić, Z. B. Automerization Reaction of Cyclobutadiene and Its Barrier Height: An Ab Initio Benchmark Multireference Average-Quadratic Coupled Cluster Study. *J. Chem. Phys.* **2006**, *125*, 064310.

(247) Vazdar, M.; Eckert-Maksić, M.; Lischka, H. The Antiferromagnetic Spin Coupling in Non-Kekulé Acenes—Impressive Polyradical Character Revealed by High-Level Multireference Methods. *ChemPhysChem* **2016**, 2013–2021.

(248) Musiał, M.; Kucharski, S. A.; Bartlett, R. J. Multireference Double Electron Attached Coupled Cluster Method with Full Inclusion of the Connected Triple Excitations: MR-DA-CCSDT. *J. Chem. Theory Comput.* **2011**, *7*, 3088–3096.

(249) Musiał, M. Multireference Fock Space Coupled Cluster Method in the Effective and Intermediate Hamiltonian Formulation for the (2,0) Sector. *J. Chem. Phys.* **2012**, *136*, 134111.

(250) Musiał, M.; Kowalska-Szojda, K.; Lyakh, D. I.; Bartlett, R. J. Potential Energy Curves via Double Electron-Attachment Calculations: Dissociation of Alkali Metal Dimers. *J. Chem. Phys.* **2013**, *138*, 194103.

(251) Shen, J.; Piecuch, P. Doubly Electron-Attached and Doubly Ionized Equation-of-Motion Coupled-Cluster Methods with 4-Particle-2-Hole and 4-Hole-2-Particle Excitations and Their Active-Space Extensions. *J. Chem. Phys.* **2013**, *138*, 194102.

(252) Shen, J.; Piecuch, P. Doubly Electron-Attached and Doubly Ionised Equation-of-Motion Coupled-Cluster Methods with Full and Active-Space Treatments of 4-Particle–2-Hole and 4-Hole–2-Particle Excitations: The Role of Orbital Choices. *Mol. Phys.* **2014**, *112*, 868–885.

(253) Ajala, A. O.; Shen, J.; Piecuch, P. Economical Doubly Electron-Attached Equation-of-Motion Coupled- Cluster Methods with an Active-Space Treatment of Three-Particle− One-Hole and Four-Particle−Two-Hole Excitations. *J. Phys. Chem. A* **2017**, *121*, 3469−3485.

(254) Piecuch, P.; Bartlett, R. J. EOMXCC: A New Coupled-Cluster Method for Electronic Excited States. In *Advances in Quantum Chemistry*; 1999; Vol. 34, pp 295–380.

(255) Levchenko, S. V.; Krylov, A. I. Equation-of-Motion Spin-Flip Coupled-Cluster Model with Single and Double Substitutions: Theory and Application to Cyclobutadiene. *J. Chem. Phys.* **2004**, *120*, 175–185.

(256) Piecuch, P. Active-Space Coupled-Cluster Methods. *Mol. Phys.* **2010**, *108*, 2987–3015.

(257) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(258) Papajak, E.; Truhlar, D. G. Efficient Diffuse Basis Sets for Density Functional Theory. *J. Chem. Theory Comput.* **2010**, *6*, 597–601.

(259) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(260) Piecuch, P.; Kucharski, S. A.; Kowalski, K.; Musiał, M. Efficient Computer Implementation of the Renormalized Coupled-Cluster Methods: The R-CCSD[T], R-CCSD(T), CR-CCSD[T], and CR-CCSD(T) Approaches. *Comput. Phys. Commun.* **2002**, *149*, 71–96.

(261) Włoch, M.; Gour, J. R.; Kowalski, K.; Piecuch, P. Extension of Renormalized Coupled-Cluster Methods Including Triple Excitations to Excited Electronic States of Open-Shell Molecules. *J. Chem. Phys.* **2005**, *122*, 214107.

(262) Piecuch, P.; Gour, J. R.; Włoch, M. Left-Eigenstate Completely Renormalized Equation-of-Motion Coupled-Cluster Methods: Review of Key Concepts, Extension to Excited States of Open-Shell Systems, and Comparison with Electron-Attached and Ionized Approaches. *Int. J. Quantum Chem.* **2009**, *109*, 3268–3304.

(263) Papajak, E.; Leverentz, H. R.; Zheng, J.; Truhlar, D. G. Efficient Diffuse Basis Sets: Cc-PVxZ+ and Maug-Cc-PVxZ. *J. Chem. Theory Comput.* **2009**, *5*, 1197–1202.

(264) Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. Density Matrix Averaged Atomic Natural Orbital (ANO) Basis Sets for Correlated Molecular Wave Functions - I. First Row Atoms. *Theor. Chim. Acta* **1990**, *77*, 291–306.

(265) Veryazov, V.; Widmark, P.-O.; Serrano-Andrés, L.; Lindh, R.; Roos, B. O. 2MOLCAS as a Development Platform for Quantum Chemistry Software. *Int. J. Quantum Chem.* **2004**, *100*, 626–635.

(266) Kovačević, G.; Veryazov, V. Luscus: Molecular Viewer and Editor for MOLCAS. *J. Cheminform.* **2015**, *7*, 16.

(267) Raghavachari, K.; Pople, J. A.; Replogle, E. S.; Head-Gordon, M.; Handy, N. C. Size-Consistent Brueckner Theory Limited to Double and Triple Substitutions. *Chem. Phys. Lett.* **1990**, *167*, 115–121.

(268) Szalay, P. G.; Bartlett, R. J. Multi-Reference Averaged Quadratic Coupled-Cluster Method: A Size-Extensive Modification of Multi-Reference CI. *Chem. Phys. Lett.* **1993**, *214*, 481–488.

(269) Stoneburner, S. J.; Truhlar, D. G.; Gagliardi, L. MC-PDFT Can Calculate Singlet-Triplet Splittings of Organic Diradicals. *J. Chem. Phys.* **2018**, *148*, 064108.

(270) Pople, J. A.; Gill, P. M. W.; Handy, N. C. Spin-Unrestricted Character of Kohn-Sham Orbitals for Open-Shell Systems. *Int. J. Quantum Chem.* **1995**, *56*, 303–305.

(271) Jacob, C. R.; Reiher, M. Spin in Density-Functional Theory. *Int. J. Quantum Chem.* **2012**, *112*, 3661–3684.

(272) Yamaguchi, K.; Jensen, F.; Dorigo, A.; Houk, K. N. A Spin Correction Procedure for Unrestricted Hartree-Fock and Møller-Plesset Wavefunctions for Singlet

Diradicals and Polyradicals. *Chem. Phys. Lett.* **1988**, *149*, 537–542.

(273) Yamaguchi, K.; Takahara, Y.; Fueno, T.; Houk, K. N. Extended Hartree-Fock (EHF) Theory in Chemical Reactions III. Projected Moller-Plesset (PMP) Perturbation Wavefunctions for Transition Structures of Organic Reactions. *Theor. Chim. Acta* **1988**, *73*, 337–364.

(274) Saito, T.; Thiel, W. Analytical Gradients for Density Functional Calculations with Approximate Spin Projection. *J. Phys. Chem. A* **2012**, *116*, 10864–10869.

(275) Hratchian, H. P. Communication: An Efficient Analytic Gradient Theory for Approximate Spin Projection Methods. *J. Chem. Phys.* **2013**, *138*, 101101.

(276) Saito, T.; Nishihara, S.; Kataoka, Y.; Nakanishi, Y.; Kitagawa, Y.; Kawakami, T.; Yamanaka, S.; Okumura, M.; Yamaguchi, K. Reinvestigation of the Reaction of Ethylene and Singlet Oxygen by the Approximate Spin Projection Method. Comparison with Multireference Coupled-Cluster Calculations. *J. Phys. Chem. A* **2010**, *114*, 7967–7974.

(277) Saito, T.; Nishihara, S.; Kataoka, Y.; Nakanishi, Y.; Matsui, T.; Kitagawa, Y.; Kawakami, T.; Okumura, M.; Yamaguchi, K. Transition State Optimization Based on Approximate Spin-Projection (AP) Method. *Chem. Phys. Lett.* **2009**, *483*, 168–171.

(278) Wang, J.; Becke, A. D.; Smith, V. H. Evaluation of $\langle S^2 \rangle$ in Restricted, Unrestricted Hartree–Fock, and Density Functional Based Theories. *J. Chem. Phys.* **1995**, *102*, 3477–3480.

(279) Wittbrodt, J. M.; Schlegel, H. B. Some Reasons Not to Use Spin Projected Density Functional Theory. *J. Chem. Phys.* **1996**, *105*, 6574–6577.

(280) Costa, R.; Valero, R.; Mañeru, D. R.; Moreira, I. D. P. R.; Illas, F. Spin Adapted versus Broken Symmetry Approaches in the Description of Magnetic Coupling in Heterodinuclear Complexes. *J. Chem. Theory Comput.* **2015**, *11*, 1006–1019.

(281) Hoyer, C. E.; Ghosh, S.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory Is as Accurate as CASPT2 for Electronic Excitation. *J. Phys. Chem. Lett.* **2016**, *7*, 586–591.

(282) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.

Gaussian 09. Gaussian, Inc.: Wallingford CT 2009.

(283) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *J. Chem. Phys.* **1996**, *105*, 9982–9985.

(284) Yamaguchi, K.; Takahara, Y.; Fueno, T. Ab-Initio Molecular Orbital Studies of Structure and Reactivity of Transition Metal-OXO Compounds. In *Applied Quantum Chemistry*; Smith, V. H., Schaefer, H. F., Morokuma, K., Eds.; Springer: Dordrecht, 1986; pp 155–184.

(285) Caballol, R.; Castell, O.; Illas, F.; Moreira, P. R.; Malrieu, J. P. Remarks on the Proper Use of the Broken Symmetry Approach to Magnetic Coupling. *J. Phys. Chem. A* **1997**, *101*, 7860–7866.

(286) Shoji, M.; Koizumi, K.; Kitagawa, Y.; Kawakami, T.; Yamanaka, S.; Okumura, M.; Yamaguchi, K. A General Algorithm for Calculation of Heisenberg Exchange Integrals *J* in Multispin Systems. *Chem. Phys. Lett.* **2006**, *432*, 343–347.

(287) Presti, D.; Stoneburner, S. J.; Truhlar, D. G.; Gagliardi, L. Full Correlation in a Multiconfigurational Study of Bimetallic Clusters : Restricted Active Space Pair-Density Functional Theory Study of [2Fe-2S] Systems. *J. Phys. Chem. C* **2019**, *123*, 11899–11907.

(288) Beinert, H.; Holm, R. H.; Münck, E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science.* **1997**, *277*, 653–659.

(289) Österberg, R. Origins of Metal Ions in Biology. *Nature* **1974**, *249*, 382–383.

(290) Gillum, W. O.; Frankel, R. B.; Foner, S.; Holm, R. H. Synthetic Analogues of the Active Sites of Iron-Sulfur Proteins. XIII. Further Electronic Structural Relationships between the Analogues $[Fe_2S_2(SR)_4]^{2-}$ and the Active Sites of Oxidized 2Fe-2S* Proteins. *Inorg. Chem.* **1976**, *15*, 1095–1100.

(291) Scintilla, S.; Bonfio, C.; Belmonte, L.; Forlin, M.; Rossetto, D.; Li, J.; Cowan, J. A.; Galliani, A.; Arnesano, F.; Assfalg, M.; Mansy, S. S. Duplications of an Iron–Sulphur Tripeptide Leads to the Formation of a Protoferredoxin. *Chem. Commun.* **2016**, *52*, 13456–13459.

(292) Qi, W.; Cowan, J. A. Structural, Mechanistic and Coordination Chemistry of Relevance to the Biosynthesis of Iron-Sulfur and Related Iron Cofactors. *Coord. Chem. Rev.* **2011**, *255*, 688–699.

(293) Blondin, G.; Girerd, J. J. Interplay of Electron Exchange and Electron Transfer in Metal Polynuclear Complexes in Proteins or Chemical Models. *Chem. Rev.* **1990**, *90* (8), 1359–1376.

(294) Dutta, S. K.; Ensling, J.; Werner, R.; Flörke, U.; Haase, W.; Gütlich, P.; Nag, K. Valence-Delocalized and Valence-Trapped $Fe^{II}Fe^{III}$ Complexes: Drastic Influence of the Ligands. *Angew. Chemie Int. Ed. English* **1997**, *36*, 152–155.

(295) Wächtershäuser, G. Groundworks for an Evolutionary Biochemistry: The Iron-Sulphur World. *Prog. Biophys. Mol. Biol.* **1992**, *58*, 85–201.

(296) Wächtershäuser, G. Life as We Don't Know It. *Science.* **2000**, *289*, 1307–1308.

(297) Wächtershäuser, G. Before Enzymes and Templates: Theory of Surface

Metabolism. *Microbiol. Rev.* **1988**, *52*, 452–484.

(298) Wächtershäuser, G. Evolution of the First Metabolic Cycles. *Proc. Natl. Acad. Sci.* **1990**, *87*, 200–204.

(299) Bonfio, C.; Valer, L.; Scintilla, S.; Shah, S.; Evans, D. J.; Jin, L.; Szostak, J. W.; Sasselov, D. D.; Sutherland, J. D.; Mansy, S. S. UV-Light-Driven Prebiotic Synthesis of Iron–Sulfur Clusters. *Nat. Chem.* **2017**, *9*, 1229–1234.

(300) Venkateswara Rao, P.; Holm, R. H. Synthetic Analogues of the Active Sites of Iron−Sulfur Proteins. *Chem. Rev.* **2004**, *104*, 527–560.

(301) Pandelia, M. E.; Lanz, N. D.; Booker, S. J.; Krebs, C. Mössbauer Spectroscopy of Fe/S Proteins. *Biochim. Biophys. Acta - Mol. Cell Res.* **2015**, *1853*, 1395–1405.

(302) Koch, F.; Berkefeld, A.; Speiser, B.; Schubert, H. Mechanistic Aspects of Redox-Induced Assembly and Disassembly of S-Bridged [2M-2S] Structures. *Chem. – A Eur. J.* **2017**, *23*, 16681–16690.

(303) Sharma, S.; Sivalingam, K.; Neese, F.; Chan, G. K.-L. Low-Energy Spectrum of Iron-Sulfur Clusters Directly from Many-Particle Quantum Mechanics. *Nat. Chem.* **2014**, *6*, 927–933.

(304) Noodleman, L.; Davidson, E. R. Ligand Spin Polarization and Antiferromagnetic Coupling in Transition Metal Dimers. *Chem. Phys.* **1986**, *109*, 131–143.

(305) Dong, G.; Phung, Q. M.; Hallaert, S. D.; Pierloot, K.; Ryde, U. H_2 Binding to the Active Site of [NiFe] Hydrogenase Studied by Multiconfigurational and Coupled-Cluster Methods. *Phys. Chem. Chem. Phys.* **2017**, *19*, 10590–10601.

(306) Noodleman, L.; Peng, C. Y.; Case, D. A.; Mouesca, J.-M. Orbital Interactions, Electron Delocalization and Spin Coupling in Iron-Sulfur Clusters. *Coord. Chem. Rev.* **1995**, *144*, 199–244.

(307) Ruiz, E.; Cano, J.; Alvarez, S.; Alemany, P. Broken Symmetry Approach to Calculation of Exchange Coupling Constants for Homobinuclear and Heterobinuclear Transition Metal Complexes. *J. Comput. Chem.* **1999**, *20*, 1391–1400.

(308) Raghu, C.; Rudra, I.; Sen, D.; Ramasesha, S. Properties of Low-Lying States in Some High-Nuclearity Mn, Fe, and V Clusters: Exact Studies of Heisenberg Models. *Phys. Rev. B* **2001**, *64*, 064419.

(309) Hübner, O.; Sauer, J. Structure and Thermochemistry of $Fe_2S_2^{-/0/+}$ Gas Phase Clusters and Their Fragments. B3LYP Calculations. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5234–5243.

(310) Hübner, O.; Sauer, J. The Electronic States of $Fe_2S_2^{-/0/+/2+}$. *J. Chem. Phys.* **2001**, *116*, 617–628.

(311) Shoji, M.; Koizumi, K.; Taniguchi, T.; Kitagawa, Y.; Yamanaka, S.; Okumura, M.; Yamaguchi, K. Theory of Chemical Bonds in Metalloenzymes III: Full Geometry Optimization and Vibration Analysis of Ferredoxin-Type [2Fe–2S] Cluster. *Int. J. Quantum Chem.* **2007**, *107*, 116–133.

(312) Chang, C. H.; Kim, K. Density Functional Theory Calculation of Bonding and

Charge Parameters for Molecular Dynamics Studies on [FeFe] Hydrogenases. *J. Chem. Theory Comput.* **2009**, *5*, 1137–1145.

(313) Li, Z.; Chan, G. K.-L. Spin-Projected Matrix Product States: Versatile Tool for Strongly Correlated Systems. *J. Chem. Theory Comput.* **2017**, *13*, 2681–2695.

(314) Chu, S.; Bovi, D.; Cappelluti, F.; Orellana, A. G.; Martin, H.; Guidoni, L. Effects of Static Correlation between Spin Centers in Multicenter Transition Metal Complexes. *J. Chem. Theory Comput.* **2017**, *13*, 4675–4683.

(315) Barborini, M.; Guidoni, L. Geometries of Low Spin States of Multi-Centre Transition Metal Complexes through Extended Broken Symmetry Variational Monte Carlo. *J. Chem. Phys.* **2016**, *145*, 124107.

(316) Ghosh, S.; Sonnenberger, A. L.; Hoyer, C. E.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory Outperforms Kohn-Sham Density Functional Theory and Multireference Perturbation Theory for Ground-State and Excited-State Charge Transfer. *J. Chem. Theory Comput.* **2015**, *11*, 3643–3649.

(317) Ghosh, S.; Verma, P.; Cramer, C. J.; Gagliardi, L.; Truhlar, D. G. Combining Wave Function Methods with Density Functional Theory for Excited States. *Chem. Rev.* **2018**, *118*, 7249–7292.

(318) Presti, D.; Truhlar, D. G.; Gagliardi, L. Intramolecular Charge Transfer and Local Excitation in Organic Fluorescent Photoredox Catalysts Explained by RASCI-PDFT. *J. Phys. Chem. C* **2018**, *122*, 12061–12070.

(319) Douglas, M.; Kroll, N. M. Quantum Electrodynamical Corrections to the Fine Structure of Helium. *Ann. Phys. (N. Y).* **1974**, *155*, 89–155.

(320) Hess, B. A. Relativistic Electronic-Structure Calculations Employing a Two-Component No-Pair Formalism with External-Field Projection Operators. *Phys. Rev. A* **1986**, *33*, 3742–3748.

(321) Wolf, A.; Reiher, M.; Hess, B. A. The Generalized Douglas–Kroll Transformation. *J. Chem. Phys.* **2002**, *117*, 9215–9226.

(322) Stålring, J.; Bernhardssonn, A.; Lindh, R. Analytical Gradients of a State Average MCSCF State and a State Average Diagnostic. *Mol. Phys.* **2001**, *99*, 103–114.

(323) Momma, K.; Izumi, F. VESTA 3 for Three-Dimensional Visualization of Crystal, Volumetric and Morphology Data. *J. Appl. Crystallogr.* **2011**, *44*, 1272–1276.

(324) Mayerle, J. J.; Denmark, S. E.; Depamphilis, B. V.; Holm, R. H.; Ibers, J. A.; Holm, R. H. Synthetic Analogs of the Active Sites of Iron-Sulfur Proteins. XI. Synthesis and Properties of Complexes Containing the Iron Sulfide (Fe2S2) Core and the Structures of Bis[o-Xylyl-α,α'-Dithiolato-μ-Sulfido-Ferrate(III)] and Bis. *J. Am. Chem. Soc.* **1975**, *97*, 1032–1045.

(325) Pierloot, K. Nondynamic Correlation Effects in Transition Metal Coordination Compounds. In *Computational Organometallic Chemistry*; Cundari, T. R., Ed.; Marcel Dekker: New York, 2001; pp 123–158.

(326) Stoneburner, S. J.; Gagliardi, L. Air Separation by Catechol-Ligated Transition

Metals: A Quantum Chemical Screening. *J. Phys. Chem. C* **2018**, *122*, 22345–22351.

(327) Crouse, B. R.; Johnson, M. K.; Meyer, J. Spectroscopic Evidence for a Reduced Fe2S2 Cluster with a S = $^9/_2$ Ground State in Mutant Forms of *Clostridium Pasteurianum* 2Fe Ferredoxin. *J. Am. Chem. Soc.* **1995**, *117*, 9612–9613.

(328) Achim, C.; Golinelli, M.-P.; Bominaar, E. L.; Meyer, J.; Münck, E. Mössbauer Study of Cys56Ser Mutant 2Fe Ferredoxin from *Clostridium Pasteurianum*: Evidence for Double Exchange in an $[Fe_2S_2]^+$ Cluster. *J. Am. Chem. Soc.* **2002**, *118*, 8168–8169.

(329) Subramanian, S.; Duin, E. C.; Fawcett, S. E. J. J.; Armstrong, F. A.; Meyer, J.; Johnson, M. K. Spectroscopic and Redox Studies of Valence-Delocalized [Fe2S2]+ Centers in Thioredoxin-like Ferredoxins. *J. Am. Chem. Soc.* **2015**, *137*, 4567–4580.

(330) Stoneburner, S. J.; Truhlar, D. G.; Gagliardi, L. Affordable and Accurate Transition Metal Spin-State Energetics via MC-PDFT Using TPBE with High Local Exchange. **2019**, to be submitted for publication.

(331) Bauer, I.; Knölker, H.-J. Iron Catalysis in Organic Synthesis. *Chem. Rev.* **2015**, *115*, 3170–3387.

(332) Harvey, J. N.; Poli, R.; Smith, K. M. Understanding the Reactivity of Transition Metal Complexes Involving Multiple Spin States. *Coord. Chem. Rev.* **2003**, *238–239*, 347–361.

(333) Meunier, B.; de Visser, S. P.; Shaik, S. Mechanism of Oxidation Reactions Catalyzed by Cytochrome P450 Enzymes. *Chem. Rev.* **2004**, *104*, 3947–3980.

(334) Goff, H.; Mar, G. N. L.; Reed, C. A. Nuclear Magnetic Resonance Investigation of Magnetic and Electronic Properties of "Intermediate Spin" Ferrous Porphyrin Complexes. *J. Am. Chem. Soc.* **1977**, *99*, 3641–3646.

(335) Stavretis, S. E.; Atanasov, M.; Podlesnyak, A. A.; Hunter, S. C.; Neese, F.; Xue, Z. L. Magnetic Transitions in Iron Porphyrin Halides by Inelastic Neutron Scattering and Ab Initio Studies of Zero-Field Splittings. *Inorg. Chem.* **2015**, *54*, 9790–9801.

(336) Phung, Q. M.; Pierloot, K. The Dioxygen Adducts of Iron and Manganese Porphyrins: Electronic Structure and Binding Energy. *Phys. Chem. Chem. Phys.* **2018**, *20*, 17009–17019.

(337) Vancoillie, S.; Zhao, H.; Radoń, M.; Pierloot, K. Performance of CASPT2 and DFT for Relative Spin-State Energetics of Heme Models. *J. Chem. Theory Comput.* **2010**, *6*, 576–582.

(338) Ghosh, A. Transition Metal Spin State Energetics and Noninnocent Systems: Challenges for DFT in the Bioinorganic Arena. *J. Biol. Inorg. Chem.* **2006**, *11*, 712–724.

(339) Pierloot, K.; Vancoillie, S. Relative Energy of the High-($^5T_{2g}$) and Low-($^1A_{1g}$) Spin States of the Ferrous Complexes [Fe(L)(NHS_4)]: CASPT2 versus Density Functional Theory. *J. Chem. Phys.* **2008**, *128*, 034104.

(340) Radoń, M. Spin-State Energetics of Heme-Related Models from DFT and Coupled Cluster Calculations. *J. Chem. Theory Comput.* **2014**, *10*, 2306–2321.

(341) Phung, Q. M.; Pierloot, K. Low-Lying Electromeric States in Chloro-Ligated Iron(IV)-Oxo Porphyrin as a Model for Compound I, Studied with Second-Order Perturbation Theory Based on Density Matrix Renormalization Group. *J. Chem. Theory Comput.* **2019**, [Early online access] DOI:10.1021/acs.jctc.9b00166.

(342) Kupper, C.; Mondal, B.; Serrano-Plana, J.; Klawitter, I.; Neese, F.; Costas, M.; Ye, S.; Meyer, F. Nonclassical Single-State Reactivity of an Oxo-Iron(IV) Complex Confined to Triplet Pathways. *J. Am. Chem. Soc.* **2017**, *139*, 8939–8949.

(343) Pinter, B.; Al-Saadon, R.; Chen, Z.; Yang, W. Spin-State Energetics of Iron(II) Porphyrin from the Particle-Particle Random Phase Approximation. *Eur. Phys. J. B* **2018**, *91*, 270.

(344) Nachtigallová, D.; Antalík, A.; Lo, R.; Sedlák, R.; Manna, D.; Tuček, J.; Ugolotti, J.; Veis, L.; Legeza, Ö.; Pittner, J.; Zbořil, R.; Hobza, P. An Isolated Molecule of Iron(II) Phthalocyanin Exhibits Quintet Ground-State: A Nexus between Theory and Experiment. *Chem. - A Eur. J.* **2018**, *24*, 13413–13417.

(345) Guo, M.; Källman, E.; Pinjari, R. V.; Couto, R. C.; Kragh Sørensen, L.; Lindh, R.; Pierloot, K.; Lundberg, M. Fingerprinting Electronic Structure of Heme Iron by Ab Initio Modeling of Metal L-Edge X-Ray Absorption Spectra. *J. Chem. Theory Comput.* **2019**, *15*, 477–489.

(346) Li Manni, G.; Alavi, A. Understanding the Mechanism Stabilizing Intermediate Spin States in Fe(II)-Porphyrin. *J. Phys. Chem. A* **2018**, *122*, 4935–4947.

(347) Ali, M. E.; Sanyal, B.; Oppeneer, P. M. Electronic Structure, Spin-States, and Spin-Crossover Reaction of Heme-Related Fe-Porphyrins: A Theoretical Perspective. *J. Phys. Chem. B* **2012**, *116*, 5849–5859.

(348) Radoń, M.; Broclawik, E.; Pierloot, K. DFT and Ab Initio Study of Iron-Oxo Porphyrins: May They Have a Low-Lying Iron(V)-Oxo Electromer? *J. Chem. Theory Comput.* **2011**, *7*, 898–908.

(349) Radoń, M.; Pierloot, K. Binding of CO, NO, and O_2 to Heme by Density Functional and Multireference Ab Initio Calculations. *J. Phys. Chem. A* **2008**, *112*, 11824–11832.

(350) Radoń, M.; Pierloot, K. Correction to "Binding of CO, NO, and O_2 to Heme by Density Functional and Multireference Ab Initio Calculations". *J. Phys. Chem. A* **2011**, *115*, 7871–7871.

(351) Groenhof, A. R.; Swart, M.; Ehlers, A. W.; Lammertsma, K. Electronic Ground States of Iron Porphyrin and of the First Species in the Catalytic Reaction Cycle of Cytochrome P450s. *J. Phys. Chem. A* **2005**, *109*, 3411–3417.

(352) Liao, M. S.; Scheiner, S. Electronic Structure and Bonding in Metal Porphyrins, Metal = Fe, Co, Ni, Cu, Zn. *J. Chem. Phys.* **2002**, *117*, 205–219.

(353) Schöneboom, J. C.; Lin, H.; Reuter, N.; Thiel, W.; Cohen, S.; Ogliaro, F.; Shaik, S. The Elusive Oxidant Species of Cytochrome P450 Enzymes: Characterization

by Combined Quantum Mechanical/Molecular Mechanical (QM/MM) Calculations. *J. Am. Chem. Soc.* **2002**, *124*, 8142–8151.

(354) Phung, Q. M.; Wouters, S.; Pierloot, K. Cumulant Approximated Second-Order Perturbation Theory Based on the Density Matrix Renormalization Group for Transition Metal Complexes: A Benchmark Study. *J. Chem. Theory Comput.* **2016**, *12*, 4352–4361.

(355) Vancoillie, S.; Zhao, H.; Tran, V. T.; Hendrickx, M. F. A.; Pierloot, K. Multiconfigurational Second-Order Perturbation Theory Restricted Active Space (RASPT2) Studies on Mononuclear First-Row Transition-Metal Systems. *J. Chem. Theory Comput.* **2011**, *7*, 3961–3977.

(356) Phung, Q. M.; Feldt, M.; Harvey, J. N.; Pierloot, K. Toward Highly Accurate Spin State Energetics in First-Row Transition Metal Complexes: A Combined CASPT2/CC Approach. *J. Chem. Theory Comput.* **2018**, *14*, 2446–2455.

(357) Li Manni, G.; Kats, D.; Tew, D. P.; Alavi, A. Role of Valence and Semicore Electron Correlation on Spin Gaps in Fe(II)-Porphyrins. *J. Chem. Theory Comput.* **2019**, *15*, 1492–1497.

(358) Smith, J. E. T.; Mussard, B.; Holmes, A. A.; Sharma, S. Cheap and Near Exact CASSCF with Large Active Spaces. *J. Chem. Theory Comput.* **2017**, *13*, 5468–5478.

(359) Saitow, M.; Kurashige, Y.; Yanai, T. Fully Internally Contracted Multireference Configuration Interaction Theory Using Density Matrix Renormalization Group: A Reduced-Scaling Implementation Derived by Computer-Aided Tensor Factorization. *J. Chem. Theory Comput.* **2015**, *11*, 5120–5131.

(360) Li Manni, G.; Smart, S. D.; Alavi, A. Combining the Complete Active Space Self-Consistent Field Method and the Full Configuration Interaction Quantum Monte Carlo within a Super-CI Framework, with Application to Challenging Metal-Porphyrins. *J. Chem. Theory Comput.* **2016**, *12*, 1245–1258.

(361) Munro, O. Q.; Wet, M. De; Pollak, H.; van Wyk, J.; Marques, H. M. Haempeptide Models for Haemoproteins Part 3: N-Acetylmicroperoxidase-8 : EPR, Mössbauer and Magnetic Susceptibility Studies on an Iron(III) Porphyrin in Thermal Equilibrium between S= 3/2, 5/2 and S= 1/2 States. *Faraday Trans.* **1998**, *94*, 1743–1752.

(362) Strauss, S. H.; Silver, M. E.; Long, K. M.; Thompson, R. G.; Hudgens, R. A.; Spartalian, K.; Ibers, J. A. Comparison of the Molecular and Electronic Structures of (2,3,7,8, 12,13,17,18-Octaethylporphyrinato)Iron(II) and (Trans-7,8-Dihydro-2,3,7,8,12,13,17,18-Octaethyl-Porphyrinato)Iron(II). *J. Am. Chem. Soc.* **1985**, *107*, 4207–4215.

(363) Mispelter, J.; Momenteau, M.; Lhoste, J. M. Proton Magnetic Resonance Characterization of the Intermediate (S=1) Spin State of Ferrous Porphyrins. *J. Chem. Phys.* **1980**, *72*, 1003–1012.

(364) Boyd, P. D. W.; Buckingham, D. A.; McMeeking, R. F.; Mitra, S. Paramagnetic Anisotropy, Average Magnetic Susceptibility, and Electronic Structure of

Intermediate-Spin S = 1 (5,10,15,20-Tetraphenylporphyrin)Iron(II). *Inorg. Chem.* **1979**, *18*, 3585–3591.

(365) Lang, G.; Spartalian, K.; Reed, C. A.; Collman, J. P. Mössbauer Effect Study of the Magnetic Properties of S = 1 Ferrous Tetraphenylporphyrin. *J. Chem. Phys.* **1978**, *69*, 5424–5427.

(366) Kitagawa, T.; Teraoka, J. The Resonance Raman Spectra of Intermediate-Spin Ferrous Porphyrin. *Chem. Phys. Lett.* **1979**, *63*, 443–446.

(367) Dolphin, D.; Sams, J. R.; Tsin, T. B.; Wong, K. L. Synthesis and Mossbauer Spectra of Octaethylporphyrin Ferrous Complexes. *J. Am. Chem. Soc.* **1976**, *98*, 6970–6975.

(368) Collman, J. P.; Hoard, J. L.; Kim, N.; Lang, G.; Reed, C. A. Synthesis, Stereochemistry, and Structure-Related Properties of α,β,γ,δ-Tetraphenylporphinatoiron(II). *J. Am. Chem. Soc.* **1975**, *97*, 2676–2681.

(369) Sligar, S. G. Coupling of Spin, Substrate, and Redox Equilibriums in Cytochrome P450. *Biochemistry* **1976**, *15*, 5399–5406.

(370) Kepp, K. P. Consistent Descriptions of Metal–Ligand Bonds and Spin-Crossover in Inorganic Chemistry. *Coord. Chem. Rev.* **2013**, *257*, 196–209.

(371) Swart, M.; Groenhof, A. R.; Ehlers, A. W.; Lammertsma, K. Validation of Exchange-Correlation Functional for Spin States of Iron Complexes. *J. Phys. Chem. A* **2004**, *108*, 5479–5483.

(372) Ganzenmüller, G.; Berkaïne, N.; Fouqueau, A.; Casida, M. E.; Reiher, M. Comparison of Density Functionals for Differences between the High- ($^5T_{2g}$) and Low- ($^1A_{1g}$) Spin States of Iron(II) Compounds. IV. Results for the Ferrous Complexes [Fe(L)('NHS_4')]. *J. Chem. Phys.* **2005**, *122*, 234321.

(373) Reiher, M. Theoretical Study of the $Fe(Phen)_2(NCS)_2$ Spin-Crossover Complex with Reparametrized Density Functionals. *Inorg. Chem.* **2002**, *41*, 6928–6935.

(374) Deeth, R. J.; Fey, N. The Performance of Nonhybrid Density Functionals for Calculating the Structures and Spin States of Fe(II) and Fe(III) Complexes. *J. Comput. Chem.* **2004**, *25*, 1840–1848.

(375) Pierloot, K.; Vancoillie, S. Relative Energy of the High-($^5T_{2g}$) and Low-($^1A_{1g}$) Spin States of $[Fe(H_2O)_6]^{2+}$, $[Fe(NH_3)_6]^{2+}$, and $[Fe(Bpy)_3]^{2+}$: CASPT2 Versu. *J. Chem. Phys.* **2006**, *125*, 124303.

(376) Radón, M.; Broclawik, E.; Pierloot, K. Electronic Structure of Selected $\{FeNO\}^7$ Complexes in Heme and Non-Heme Architectures: A Density Functional and Multireference Ab Initio Study. *J. Phys. Chem. B* **2010**, *114*, 1518–1528.

(377) Wilbraham, L.; Verma, P.; Truhlar, D. G.; Gagliardi, L.; Ciofini, I. Multiconfiguration Pair-Density Functional Theory Predicts Spin-State Ordering in Iron Complexes with the Same Accuracy as Complete Active Space Second-Order Perturbation Theory at a Significantly Reduced Computational Cost. *J. Phys. Chem. Lett.* **2017**, *8*, 2026–2030.

(378) Verma, P.; Truhlar, D. G. HLE17: An Improved Local Exchange-Correlation

Functional for Computing Semiconductor Band Gaps and Molecular Excitation Energies. *J. Phys. Chem. C* **2017**, *121*, 7144–7154.

(379) Feldt, M.; Phung, Q. M.; Pierloot, K.; Mata, R. A.; Harvey, J. N. Limits of Coupled-Cluster Calculations for Non-Heme Iron Complexes. *J. Chem. Theory Comput.* **2019**, *15*, 922–937.

(380) Radoń, M. Benchmarking Quantum Chemistry Methods for Spin-State Energetics of Iron Complexes against Quantitative Experimental Data. *Phys. Chem. Chem. Phys.* **2019**, *21*, 4854–4870.

(381) Sharma, P.; Truhlar, D. G.; Gagliardi, L. Multiconfiguration Pair-Density Functional Theory Investigation of the Electronic Spectrum of MnO_4^-. *J. Chem. Phys.* **2018**, *148*, 124305.

(382) Carlson, R. K.; Li Manni, G.; Sonnenberger, A. L.; Truhlar, D. G.; Gagliardi, L. Correction to Multiconfiguration Pair-Density Functional Theory: Barrier Heights and Main Group and Transition Metal Energetics. *J. Chem. Theory Comput.* **2016**, *12*, 457.

(383) Sharkas, K.; Gagliardi, L.; Truhlar, D. G. Multiconfiguration Pair-Density Functional Theory and Complete Active Space Second Order Perturbation Theory. Bond Dissociation Energies of FeC, NiC, FeS, NiS, FeSe, and NiSe. *J. Phys. Chem. A* **2017**, *121*, 9392–9400.

(384) Sharma, P.; Bernales, V.; Truhlar, D. G.; Gagliardi, L. Valence $\pi\pi*$ Excitations in Benzene Studied by Multiconfiguration Pair-Density Functional Theory. *J. Phys. Chem. Lett.* **2019**, *10*, 75–81.

(385) Chen, H.; Lai, W.; Shaik, S. Exchange-Enhanced H-Abstraction Reactivity of High-Valent Nonheme Iron(IV)-Oxo from Coupled Cluster and Density Functional Theories. *J. Phys. Chem. Lett.* **2010**, *1*, 1533–1540.

(386) Kirkland, J. K.; Khan, S. N.; Casale, B.; Miliordos, E.; Vogiatzis, K. D. Ligand Field Effects on the Ground and Excited States of Reactive FeO^{2+} Species. *Phys. Chem. Chem. Phys.* **2018**, *20*, 28786–28795.

(387) Balabanov, N. B.; Peterson, K. A. Systematically Convergent Basis Sets for Transition Metals. I. All-Electron Correlation Consistent Basis Sets for the 3*d* Elements Sc-Zn. *J. Chem. Phys.* **2005**, *123*, 064107.

(388) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(389) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(390) Reiher, M.; Wolf, A. Exact Decoupling of the Dirac Hamiltonian. I. General Theory. *J. Chem. Phys.* **2004**, *121*, 2037–2047.

(391) Reiher, M.; Wolf, A. Exact Decoupling of the Dirac Hamlltonian. II. The Generalized Douglas-Kroll-Hess Transformation up to Arbitrary Order. *J. Chem.*

Phys. **2004**, *121*, 10945–10956.

(392) Howes, B. D.; Feis, A.; Indiani, C.; Marzocchi, M. P.; Smulevich, G. Formation of Two Types of Low-Spin Heme in Horseradish Peroxidase Isoenzyme A2 at Low Temperature. *J. Biol. Inorg. Chem.* **2000**, *5*, 227–235.

(393) Decker, A.; Rohde, J.-U.; Klinker, E. J.; Wong, S. D.; Que, L.; Solomon, E. I. Spectroscopic and Quantum Chemical Studies on Low-Spin Fe^{IV}=O Complexes: Fe-O Bonding and Its Contributions to Reactivity. *J. Am. Chem. Soc.* **2007**, *129*, 15983–15996.

(394) Ongari, D.; Tiana, D.; Stoneburner, S. J.; Gagliardi, L.; Smit, B. Origin of the Strong Interaction between Polar Molecules and Copper(II) Paddle-Wheels in Metal Organic Frameworks. *J. Phys. Chem. C* **2017**, *121*, 15135–15144.

(395) Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. Large-Scale Screening of Hypothetical Metal–Organic Frameworks. *Nat. Chem.* **2012**, *4*, 83–89.

(396) Morris, R. E.; Wheatley, P. S. Gas Storage in Nanoporous Materials. *Angew. Chemie - Int. Ed.* **2008**, *47*, 4966–4981.

(397) Li, J.-R.; Kuppler, R. J.; Zhou, H.-C. Selective Gas Adsorption and Separation in Metal-Organic Frameworks. *Chem. Soc. Rev.* **2009**, *38*, 1477–1504.

(398) Zhang, T.; Lin, W. Metal-Organic Frameworks for Artificial Photosynthesis and Photocatalysis. *Chem. Soc. Rev.* **2014**, *43*, 5982–5993.

(399) Liu, J.; Chen, L.; Cui, H.; Zhang, J.; Zhang, L.; Su, C.-Y. Applications of Metal-Organic Frameworks in Heterogeneous Supramolecular Catalysis. *Chem. Soc. Rev.* **2014**, *43*, 6011–6061.

(400) Poloni, R.; Lee, K.; Berger, R. F.; Smit, B.; Neaton, J. B. Understanding Trends in CO_2 Adsorption in Metal-Organic Frameworks with Open-Metal Sites. *J. Phys. Chem. Lett.* **2014**, *5*, 861–865.

(401) Evans, J. D.; Fraux, G.; Gaillac, R.; Kohen, D.; Trousselet, F.; Vanson, J. M.; Coudert, F.-X. Computational Chemistry Methods for Nanoporous Materials. *Chem. Mater.* **2017**, *29*, 199–212.

(402) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.

(403) Chui, S. S.; Lo, S. M.; Charmant, J. P. H.; Orpen, A. G.; Williams, I. D. A Chemically Functionalizable Nanoporous Material $[Cu_3(TMA)_2(H_2O)_3]N$. *Science.* **1999**, *283*, 1148–1151.

(404) Mason, J. A.; Veenstra, M.; Long, J. R. Evaluating Metal–Organic Frameworks for Natural Gas Storage. *Chem. Sci.* **2014**, *5*, 32–51.

(405) Min Wang, Q.; Shen, D.; Bülow, M.; Ling Lau, M.; Deng, S.; Fitch, F. R.; Lemcoff, N. O.; Semanscin, J. Metallo-Organic Molecular Sieve for Gas Separation and Purification. *Microporous Mesoporous Mater.* **2002**, *55*, 217–230.

(406) Hamon, L.; Jolimaître, E.; Pirngruber, G. D. CO_2 and CH_4 Separation by Adsorption Using Cu-BTC Metal−Organic Framework. *Ind. Eng. Chem. Res.*

2010, *49*, 7497–7503.

(407) Ge, L.; Zhou, W.; Rudolph, V.; Zhu, Z. Mixed Matrix Membranes Incorporated with Size-Reduced Cu-BTC for Improved Gas Separation. *J. Mater. Chem. A* **2013**, *1*, 6350–6358.

(408) Wee, L. H.; Janssens, N.; Bajpe, S. R.; Kirschhock, C. E. A.; Martens, J. A. Heteropolyacid Encapsulated in $Cu_3(BTC)_2$ Nanocrystals: An Effective Esterification Catalyst. *Catal. Today* **2011**, *171*, 275– 280.

(409) Pérez-Mayoral, E.; Čejka, J. [$Cu_3(BTC)_2$]: A Metal-Organic Framework Catalyst for the Friedländer Reaction. *ChemCatChem* **2011**, *3*, 157–159.

(410) Ye, J.; Liu, C. $Cu_3(BTC)_2$: CO Oxidation over MOF Based Catalysts. *Chem. Commun.* **2011**, *47*, 2167–2169.

(411) Wu, H.; Simmons, J. M.; Srinivas, G.; Zhou, W.; Yildirim, T. Adsorption Sites and Binding Nature of CO_2 in Prototypical Metal-Organic Frameworks: A Combined Neutron Diffraction and First-Principles Study. *J. Phys. Chem. Lett.* **2010**, *1*, 1946–1951.

(412) Grajciar, L.; Nachtigall, P.; Bludský, O.; Rubeš, M. Accurate Ab Initio Description of Adsorption on Coordinatively Unsaturated Cu^{2+} and Fe^{3+} Sites in MOFs. *J. Chem. Theory Comput.* **2015**, *11*, 230–238.

(413) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(414) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(415) Dion, M.; Rydberg, H.; Schroder, E.; Langreth, D. C.; Lundqvist, B. I. Van Der Waals Density Functional for General Geometries. *Phys. Rev. Lett.* **2004**, *92*, 246401.

(416) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. Higher-Accuracy van Der Waals Density Functional. *Phys. Rev. B* **2010**, *82*, 081101.

(417) Lee, K.; Howe, J. D.; Lin, L. C.; Smit, B.; Neaton, J. B. Small-Molecule Adsorption in Open-Site Metal-Organic Frameworks: A Systematic Density Functional Theory Study for Rational Design. *Chem. Mater.* **2015**, *27*, 668–678.

(418) Poloni, R.; Smit, B.; Neaton, J. B. CO_2 Capture by Metal-Organic Frameworks with van Der Waals Density Functionals. *J. Phys. Chem. A* **2012**, *116*, 4957–4964.

(419) Poloni, R.; Smit, B.; Neaton, J. B. Ligand-Assisted Enhancement of CO_2 Capture in Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2012**, *134*, 6714–6719.

(420) Lin, L.-C.; Lee, K.; Gagliardi, L.; Neaton, J. B.; Smit, B. Force-Field Development from Electronic Structure Calculations with Periodic Boundary Conditions: Applications to Gaseous Adsorption and Transport in Metal-Organic Frameworks. *J. Chem. Theory Comput.* **2014**, *10*, 1477–1488.

(421) Mercado, R.; Vlaisavljevich, B.; Lin, L.-C.; Lee, K.; Lee, Y.; Mason, J. A.; Xiao, D. J.; Gonzalez, M. I.; Kapelewski, M. T.; Neaton, J. B.; Smit, B. Force Field

Development from Periodic Density Functional Theory Calculations for Gas Separation Applications Using Metal-Organic Frameworks. *J. Phys. Chem. C* **2016**, *120*, 12590−12604.

(422) Grajciar, L.; Wiersum, A. D.; Llewellyn, P. L.; Chang, J. S.; Nachtigall, P. Understanding CO_2 Adsorption in CuBTC MOF: Comparing Combined DFT-Ab Initio Calculations with Microcalorimetry Experiments. *J. Phys. Chem. C* **2011**, *115*, 17925–17933.

(423) Tiana, D.; Hendon, C. H.; Walsh, A. Ligand Design for Long-Range Magnetic Order in Metal-Organic Frameworks. *Chem. Commun.* **2014**, *50*, 13990–13993.

(424) Grajciar, L.; Bludský, O.; Nachtigall, P. Water Adsorption on Coordinatively Unsaturated Sites in CuBTC MOF. *J. Phys. Chem. Lett.* **2010**, *1*, 3354–3359.

(425) Vogiatzis, K. D.; Mavrandonakis, A.; Klopper, W.; Froudakis, G. E. Ab Initio Study of the Interactions between CO_2 and N-Containing Organic Heterocycles. *ChemPhysChem* **2009**, *10*, 374–383.

(426) Vogiatzis, K. D.; Klopper, W.; Mavrandonakis, A.; Fink, K. Magnetic Properties of Paddlewheels and Trinuclear Clusters with Exposed Metal Sites. *ChemPhysChem* **2011**, *12*, 3307–3319.

(427) Boyd, P. G.; Moosavi, S. M.; Witman, M.; Smit, B. Force-Field Prediction of Materials Properties in Metal-Organic Frameworks. *J. Phys. Chem. Lett.* **2017**, *8*, 357–363.

(428) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72*, 171–179.

(429) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces. **2008**, *100*, 136406.

(430) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; De Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M. QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys. Condens. Matter* **2009**, *21*, 395502.

(431) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59*, 1758–1775.

(432) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(433) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120*,

215–241.

(434) Knowles, P. J.; Andrews, J. S.; Amos, R. D.; Handy, N. C.; Pople, J. A. Restricted Møller-Plesset Theory for Open-Shell Molecules. *Chem. Phys. Lett.* **1991**, *186*, 130–136.

(435) Davidson, E. R. Comment on "Comment on Dunning's Correlation-Consistent Basis Sets." *Chem. Phys. Lett.* **1996**, *260*, 514–518.

(436) Boys, S.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553–566.

(437) Aquilante, F.; Gagliardi, L.; Pedersen, T. B.; Lindh, R. Atomic Cholesky Decompositions: A Route to Unbiased Auxiliary Basis Sets for Density Fitting Approximation with Tunable Accuracy and Efficiency. *J. Chem. Phys.* **2009**, *130*, 154107.

(438) Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. RASPA: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials. *Mol. Simul.* **2016**, *42*, 81–101.

(439) Potoff, J. J.; Siepmann, J. I. Vapor–Liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen. *AIChE J.* **2001**, *47*, 1676–1682.

(440) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

(441) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: A Generic Force Field for Molecular Simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

(442) Campañá, C.; Mussard, B.; Woo, T. K. Electrostatic Potential Derived Atomic Charges for Periodic Systems Using a Modified Error Functional. *J. Chem. Theory Comput.* **2009**, *5*, 2866–2878.

(443) Otero-De-La-Roza, A.; Johnson, E. R.; Luaña, V. Critic2: A Program for Real-Space Analysis of Quantum Chemical Interactions in Solids. *Comput. Phys. Commun.* **2014**, *185*, 1007–1018.

(444) Fischer, M.; Gomes, J. R. B.; Jorge, M. Computational Approaches to Study Adsorption in MOFs with Unsaturated Metal Sites. *Molecular Simulation*. Taylor & Francis 2014, pp 537–556.

(445) Liechtenstein, A. I.; Asimonov, V.; Zaneen, J. DFT and Strong Interations: Orbital Ordering in Mott-Hubbard Insulator. *Phys. Rev. B, rapid comm* **1995**, *52*, 5.

(446) Mann, G. W.; Lee, K.; Cococcioni, M.; Smit, B.; Neaton, J. B. First-Principles Hubbard U Approach for Small Molecule Binding in Metal-Organic Frameworks. *J. Chem. Phys.* **2016**, *144*, 174104.

(447) Wang, L.; Maxisch, T.; Ceder, G. Oxidation Energies of Transition Metal Oxides within the GGA+U Framework. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2006**, *73*, 195107.

(448) Thonhauser, T.; Zuluaga, S.; Arter, C. A.; Berland, K.; Schröder, E.; Hyldgaard, P.

Spin Signature of Nonlocal Correlation Binding in Metal-Organic Frameworks. *Phys. Rev. Lett.* **2015**, *115*, 136402.

(449) Hamada, I. Van Der Waals Density Functional Made Accurate. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2014**, *89*, 121103(R).

(450) Klimeš, J.; Michaelides, A. Perspective: Advances and Challenges in Treating van Der Waals Dispersion Forces in Density Functional Theory. *J. Chem. Phys.* **2012**, *137*, 120901.

(451) Řezáč, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and Their Applications. *Chem. Rev.* **2016**, *116*, 5038–5071.

(452) Chen, B.; Eddaoudi, M.; Reineke, T. M.; Kampf, J. W.; O'Keeffe, M.; Yaghi, O. M. $Cu_2(ATC)\cdot 6H_2O$: Design of Open Metal Sites in Porous Metal-Organic Crystals (ATC: 1,3,5,7-Adamantane Tetracarboxylate) [27]. *J. Am. Chem. Soc.* **2000**, *122*, 11559–11560.

(453) Maurice, R.; Sivalingam, K.; Ganyushin, D.; Guihéry, N.; De Graaf, C.; Neese, F. Theoretical Determination of the Zero-Field Splitting in Copper Acetate Monohydrate. *Inorg. Chem.* **2011**, *50*, 6229–6236.

(454) Hu, J.; Liu, J.; Liu, Y.; Yang, X. Improving Carbon Dioxide Storage Capacity of Metal Organic Frameworks by Lithium Alkoxide Functionalization: A Molecular Simulation Study. *J. Phys. Chem. C* **2016**, *120*, 10311–10319.

(455) Aprea, P.; Caputo, D.; Gargiulo, N.; Iucolano, F.; Pepe, F. Modeling Carbon Dioxide Adsorption on Microporous Substrates: Comparison between Cu-BTC Metal-Organic Framework and 13X Zeolitic Molecular Sieve. *J. Chem. Eng. Data* **2010**, *55*, 3655–3661.

(456) Chowdhury, P.; Mekala, S.; Dreisbach, F.; Gumma, S. Adsorption of CO, CO_2 and CH_4 on Cu-BTC and MIL-101 Metal Organic Frameworks: Effect of Open Metal Sites and Adsorbate Polarity. *Microporous Mesoporous Mater.* **2012**, *152*, 246–252.

(457) Krishna, R. Adsorptive Separation of $CO_2/CH_4/CO$ Gas Mixtures at High Pressures. *Microporous Mesoporous Mater.* **2012**, *156*, 217–223.

(458) Li, B.; Zhang, Z.; Li, Y.; Yao, K.; Zhu, Y.; Deng, Z.; Yang, F.; Zhou, X.; Li, G.; Wu, H.; Nijem, N.; Chabal, Y. J.; Lai, Z.; Han, Y.; Shi, Z.; Feng, S.; Li, J. Enhanced Binding Affinity, Remarkable Selectivity, and High Capacity of CO2 by Dual Functionalization of a Rht-Type Metal-Organic Framework. *Angew. Chemie - Int. Ed.* **2012**, *51*, 1412–1415.

(459) Spanopoulos, I.; Bratsos, I.; Tampaxis, C.; Vourloumis, D.; Klontzas, E.; Froudakis, G. E.; Charalambopoulou, G.; Steriotis, T. A.; Trikalitis, P. N. Exceptional Gravimetric and Volumetric CO_2 Uptake in a Palladated NbO-Type MOF Utilizing Cooperative Acidic and Basic, Metal-CO_2 Interactions. *Chem. Commun.* **2016**, *52*, 10559–10562.

(460) Wilmer, C. E.; Kim, K. C.; Snurr, R. Q. An Extended Charge Equilibration Method. *J. Phys. Chem. Lett.* **2012**, *3*, 2506–2511.

(461) Stoneburner, S. J.; Livermore, V.; McGreal, M. E.; Yu, D.; Vogiatzis, K. D.; Snurr, R. Q.; Gagliardi, L. Catechol-Ligated Transition Metals: A Quantum Chemical Study on a Promising System for Gas Separation. *J. Phys. Chem. C* **2017**, *121*, 10463–10469.

(462) Stoneburner, S. J.; Livermore, V.; McGreal, M. E.; Yu, D.; Vogiatzis, K. D.; Snurr, R. Q.; Gagliardi, L. Correction to "Catechol-Ligated Transition Metals: A Quantum Chemical Study on a Promising System for Gas Separation." *J. Phys. Chem. C* **2017**, *121*, 20553.

(463) Moncada, S.; Palmer, R. M.; Higgs, E. A. Nitric Oxide: Physiology, Pathophysiology, and Pharmacology. *Pharmacol. Rev.* **1991**, *43*, 109–142.

(464) Weinberger, B.; Laskin, D. L.; Heck, D. E.; Laskin, J. D. The Toxicology of Inhaled Nitric Oxide. *Toxicol. Sci.* **2001**, *59* (1), 5–16.

(465) Szinicz, L. History of Chemical and Biological Warfare Agents. *Toxicology* **2005**, *214*, 167–181.

(466) Guo, R.-T.; Hao, J.-K.; Pan, W.-G.; Yu, Y.-L. Liquid Phase Oxidation and Absorption of NO from Flue Gas: A Review. *Sep. Sci. Technol.* **2015**, *50*, 310–321.

(467) Long, X. L.; Xiao, W. D.; Yuan, W. K. Removal of Sulfur Dioxide and Nitric Oxide Using Cobalt Ethylenediamine Solution. *Ind. Eng. Chem. Res.* **2005**, *44*, 686–691.

(468) McKinlay, A. C.; Eubank, J. F.; Wuttke, S.; Xiao, B.; Wheatley, P. S.; Bazin, P.; Lavalley, J. C.; Daturi, M.; Vimont, A.; De Weireld, G.; Horcajada, P.; Serre, C.; Morris, R. E. Nitric Oxide Adsorption and Delivery in Flexible MIL-88(Fe) Metal-Organic Frameworks. *Chem. Mater.* **2013**, *25*, 1592–1599.

(469) Fioretos, K. A.; Psofogiannakis, G. M.; Froudakis, G. E. Ab-Initio Study of the Adsorption and Separation of NO_x and SO_x Gases in Functionalized IRMOF Ligands. *J. Phys. Chem. C* **2011**, *115*, 24906–24914.

(470) Yu, K.; Kiesling, K.; Schmidt, J. R. Trace Flue Gas Contaminants Poison Coordinatively Unsaturated Metal-Organic Frameworks: Implications for CO_2 Adsorption and Separation. *J. Phys. Chem. C* **2012**, *116*, 20480–20488.

(471) Wang, L.; Wang, L.; Zhao, J.; Yan, T. Adsorption of Selected Gases on Metal-Organic Frameworks and Covalent Organic Frameworks: A Comparative Grand Canonical Monte Carlo Simulation. *J. Appl. Phys.* **2012**, *111*, 112628.

(472) Tan, K.; Zuluaga, S.; Gong, Q.; Gao, Y.; Nijem, N.; Li, J.; Thonhauser, T.; Chabal, Y. J. Competitive Coadsorption of CO_2 with H_2O, NH_3, SO_2, NO, NO_2, N_2, O_2, and CH_4 in M-MOF-74 (M = Mg, Co, Ni): The Role of Hydrogen Bonding. *Chem. Mater.* **2015**, *27*, 2203–2217.

(473) Ding, L.; Yazaydin, A. Ö. How Well Do Metal-Organic Frameworks Tolerate Flue Gas Impurities? *J. Phys. Chem. C* **2012**, *116*, 22987–22991.

(474) Brozek, C. K.; Miller, J. T.; Stoian, S. a; Dinca, M. NO Disproportionation at a Mononuclear Site-Isolated Fe^{2+} Center in Fe^{2+}-MOF-5. *J. Am. Chem. Soc.* **2015**,

137, 7495–7501.

(475) Xiao, B.; Wheatley, P. S.; Zhao, X.; Fletcher, A. J.; Fox, S.; Rossi, A. G.; Megson, I. L.; Bordiga, S.; Regli, L.; Thomas, K. M.; Morris, R. E. High-Capacity Hydrogen and Nitric Oxide Adsorption and Storage in a Metal--Organic Framework. *J. Am. Chem. Soc.* **2007**, *129*, 1203–1209.

(476) Bordiga, S.; Regli, L.; Bonino, F.; Groppo, E.; Lamberti, C.; Xiao, B.; Wheatley, P. S.; Morris, R. E.; Zecchina, A. Adsorption Properties of HKUST-1 toward Hydrogen and Other Small Molecules Monitored by IR. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2676.

(477) Barth, B.; Mendt, M.; Pöppl, A.; Hartmann, M. Adsorption of Nitric Oxide in Metal-Organic Frameworks: Low Temperature IR and EPR Spectroscopic Evaluation of the Role of Open Metal Sites. *Microporous Mesoporous Mater.* **2015**, *216*, 97–110.

(478) Bonino, F.; Chavan, S.; Vitillo, J. G.; Groppo, E.; Agostini, G.; Lamberti, C.; Dietzel, P. D. C.; Prestipino, C.; Bordiga, S. Local Structure of CPO-27-Ni Metallorganic Framework upon Dehydration and Coordination of NO. *Chem. Mater.* **2008**, *20*), 4957–4968.

(479) Nguyen, J. G.; Tanabe, K. K.; Cohen, S. M. Postsynthetic Diazeniumdiolate Formation and NO Release from MOFs. *CrystEngComm* **2010**, *12*, 2335–2338.

(480) Ingleson, M. J.; Heck, R.; Gould, J. a.; Rosseinsky, M. J. Nitric Oxide Chemisorption in a Postsynthetically Modified Metal−Organic Framework. *Inorg. Chem.* **2009**, *48*, 9986–9988.

(481) Shen, J.; Liu, G.; Huang, K.; Li, Q.; Guan, K.; Li, Y.; Jin, W. UiO-66-Polyether Block Amide Mixed Matrix Membranes for CO_2 Separation. *J. Memb. Sci.* **2016**, *513*, 155–165.

(482) Bloch, W. M.; Babarao, R.; Hill, M. R.; Doonan, C. J.; Sumby, C. J. Post-Synthetic Structural Processing in a Metal-Organic Framework Material as a Mechanism for Exceptional CO_2/N_2 Selectivity. **2013**.

(483) Li, T.; Chen, D.-L.; Sullivan, J. E. E.; Kozlowski, M. T. T.; Johnson, J. K. K.; Rosi, N. L. L. Systematic Modulation and Enhancement of CO_2 : N_2 Selectivity and Water Stability in an Isoreticular Series of Bio-MOF-11 Analogues. *Chem. Sci.* **2013**, *4*, 1746–1755.

(484) Couck, S.; Gobechiya, E.; Kirschhock, C. E. A.; Serra-Crespo, P.; Juan-Alcañiz, J.; Martinez Joaristi, A.; Stavitski, E.; Gascon, J.; Kapteijn, F.; Baron, G. V.; Denayer, J. F. M. Adsorption and Separation of Light Gases on an Amino-Functionalized Metal-Organic Framework: An Adsorption and in Situ XRD Study. *ChemSusChem* **2012**, *5*, 740–750.

(485) Liu, H.; Zhao, Y.; Zhang, Z.; Nijem, N.; Chabal, Y. J.; Peng, X.; Zeng, H.; Li, J. Ligand Functionalization and Its Effect on CO2 Adsorption in Microporous Metal-Organic Frameworks. *Chem. - An Asian J.* **2013**, *8*, 778–785.

(486) Tan, Y. X.; He, Y. P.; Zhang, J. Temperature-/Pressure-Dependent Selective Separation of CO_2 or Benzene in a Chiral Metal-Organic Framework Material.

ChemSusChem **2012**, *5*, 1597–1601.

(487) Li, H.; Shi, W.; Zhao, K.; Niu, Z.; Chen, X.; Cheng, P. A Robust Porous Metal-Organic Framework with a New Topology That Demonstrates Pronounced Porosity and High-Efficiency Sorption/Selectivity Properties of Small Molecules. *Chem. - A Eur. J.* **2012**, *18*), 5715–5723.

(488) Liu, H.; Zhao, Y.; Zhang, Z.; Nijem, N.; Chabal, Y. J.; Zeng, H.; Li, J. The Effect of Methyl Functionalization on Microporous Metal-Organic Frameworks' Capacity and Binding Energy for Carbon Dioxide Adsorption. *Adv. Funct. Mater.* **2011**, *21*, 4754–4762.

(489) Lin, X. M.; Li, T. T.; Wang, Y. W.; Zhang, L.; Su, C. Y. Two Zn^{II} Metal-Organic Frameworks with Coordinatively Unsaturated Metal Sites: Structures, Adsorption, and Catalysis. *Chem. - An Asian J.* **2012**, *7*, 2796–2804.

(490) Babarao, R.; Rubio-Martinez, M.; Hill, M. R.; Thornton, A. W. Interpenetrated Zirconium-Organic Frameworks: Small Cavities versus Functionalization for CO_2 Capture. *J. Phys. Chem. C* **2016**, *120*, 13013–13023.

(491) Zhang, Y.; Su, W.; Sun, Y.; Liu, J.; Liu, X.; Wang, X. Adsorption Equilibrium of N_2, CH_4, and CO_2 on MIL-101. *J. Chem. Eng. Data* **2015**, *60*, 2951–2957.

(492) Tanaka, H.; Nishibayashi, Y.; Yoshizawa, K. Interplay between Theory and Experiment for Ammonia Synthesis Catalyzed by Transition Metal Complexes. *Acc. Chem. Res.* **2016**, *49*, 987–995.

(493) Guru, M. M.; Shima, T.; Hou, Z. Conversion of Dinitrogen to Nitriles at a Multinuclear Titanium Framework. *Angew. Chemie Int. Ed.* **2016**, *55*, 12316–12320.

(494) Nguyen, H. G. T.; Weston, M. H.; Sarjeant, A. A.; Gardner, D. M.; An, Z.; Carmieli, R.; Wasielewski, M. R.; Farha, O. K.; Hupp, J. T.; Nguyen, S. T. Design, Synthesis, Characterization, and Catalytic Properties of a Large-Pore Metal-Organic Framework Possessing Single-Site Vanadyl(Monocatecholate) Moieties. *Cryst. Growth Des.* **2013**, *13*, 3528–3534.

(495) Fei, H.; Shin, J.; Meng, Y. S.; Adelhardt, M.; Sutter, J.; Meyer, K.; Cohen, S. M. Reusable Oxidation Catalysis Using Metal-Monocatecholato Species in a Robust Metal−Organic Framework. *J. Am. Chem. Soc.* **2014**, *136*, 4965–4973.

(496) Nguyen, H. G. T.; Mao, L.; Peters, A. W.; Audu, C. O.; Brown, Z. J.; Farha, O. K.; Hupp, J. T.; Nguyen, S. T. Comparative Study of Titanium-Functionalized UiO-66: Support Effect on the Oxidation of Cyclohexene Using Hydrogen Peroxide. *Catal. Sci. Technol.* **2015**, *5*, 4444–4451.

(497) Weston, M. H.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Nguyen, S. T. Synthesis and Metalation of Catechol-Functionalized Porous Organic Polymers. *Chem. Mater.* **2012**, *24*, 1292–1296.

(498) Tanabe, K. K.; Siladke, N. A.; Broderick, E. M.; Kobayashi, T.; Goldston, J. F.; Weston, M. H.; Farha, O. K.; Hupp, J. T.; Pruski, M.; Mader, E. A.; Johnson, M. J. A.; Nguyen, S. T. Stabilizing Unstable Species through Single-Site Isolation: A Catalytically Active Ta^{V} Trialkyl in a Porous Organic Polymer. *Chem. Sci.* **2013**,

4, 2483–2489.

(499) Stergiannakos, T.; Tylianakis, E.; Klontzas, E.; Froudakis, G. E. Enhancement of Hydrogen Adsorption in Metal−Organic Frameworks by Mg^{2+} Functionalization: A Multiscale Computational Study. *J. Phys. Chem. C* **2010**, *114*, 16855–16858.

(500) Getman, R. B.; Miller, J. H.; Wang, K.; Snurr, R. Q. Metal Alkoxide Functionalization in Metal-Organic Frameworks for Enhanced Ambient-Temperature Hydrogen Storage. *J. Phys. Chem. C* **2011**, *115*, 2066–2075.

(501) Tsivion, E.; Mason, J. A.; Gonzalez, M. I.; Long, F. R.; Head-Gordan, M. A Computational Study of CH_4 Storage in Porous Framework Materials with Metalated Linkers: Connecting the Atomistic Character of CH_4 Binding Sites to Usable Capacity. *Chem. Sci.* **2016**, *7*, 4503–4518.

(502) Ilyakina, E. V.; Poddel'sky, A. I.; Cherkasov, V. K.; Abakumov, G. A. Binding of NO by Nontransition Metal Complexes Ekaterina. *Mendeleev Commun.* **2012**, *22*, 208–210.

(503) Maihom, T.; Wannakao, S.; Boekfa, B.; Limtrakul, J. Production of Formic Acid via Hydrogenation of CO_2 over a Copper-Alkoxide-Functionalized MOF: A Mechanistic Study. *J. Phys. Chem. C* **2013**, *117*, 17650–17658.

(504) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

(505) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(506) Vogiatzis, K. D.; Barnes, E. C.; Klopper, W. Interference-Corrected Explicitly-Correlated Second-Order Perturbation Theory. *Chem. Phys. Lett.* **2011**, *503*, 157–161.

(507) Vogiatzis, K. D.; Haunschild, R.; Klopper, W. Accurate Atomization Energies from Combining Coupled-Cluster Computations with Interference-Corrected Explicitly Correlated Second-Order Perturbation Theory. *Theor. Chem. Acc.* **2014**, *133*, 1–12.

(508) Mavrandonakis, A.; Vogiatzis, K. D.; Boese, A. D.; Fink, K.; Heine, T.; Klopper, W. Ab Initio Study of the Adsorption of Small Molecules on Metal–Organic Frameworks with Oxo-Centered Trimetallic Building Units: The Role of the Undercoordinated Metal Ion. *Inorg. Chem.* **2015**, *54*, 8251–8263.

(509) Aquilante, F.; Pedersen, T. B.; Veryazov, V.; Lindh, R. MOLCAS-a Software for Multiconfigurational Quantum Chemistry Calculations. *WIREs Comput Mol Sci* **2013**, *3*, 143–149.

(510) Greenwood, N. N.; Earnshaw, A. Nitrogen. In *Chemistry of the Elements*; Butterworth-Heinemann: Oxford, 1997; pp 406–472.

(511) Greenwood, N. N.; Earnshaw, A. Oxygen. In *Chemistry of the Elements*; Butterworth-Heinemann: Oxford, 1997; pp 600–644.

(512) Zhang, X.; Chen, J.; Yao, L.; Huang, Y.; Zhang, X.; Qiu, L. Research and Development of Large-Scale Cryogenic Air Separation in China. *J. Zhejiang Univ. Sci. A* **2014**, *15*, 309–322.

(513) Parkes, M. V.; Sava Gallis, D. F.; Greathouse, J. A.; Nenoff, T. M. Effect of Metal in $M_3(Btc)_2$ and $M_2(Dobdc)$ MOFs for O_2/N_2 Separations: A Combined Density Functional Theory and Experimental Study. *J. Phys. Chem. C* **2015**, *119*, 6556–6567.

(514) Kneer, R.; Toporov, D.; Förster, M.; Christ, D.; Broeckmann, C.; Pfaff, E.; Zwick, M.; Engels, S.; Modigell, M. OXYCOAL-AC: Towards an Integrated Coal-Fired Power Plant Process with Ion Transportmembrane-Based Oxygen Supply. *Energy Environ. Sci.* **2010**, *3*, 198–207.

(515) Markewitz, P.; Kuckshinrichs, W.; Leitner, W.; Linssen, J.; Zapp, P.; Bongartz, R.; Schreiber, A.; Müller, T. E. Worldwide Innovations in the Development of Carbon Capture Technologies and the Utilization of CO_2. *Energy Environ. Sci.* **2012**, *5*, 7281–7305.

(516) Kather, A.; Scheffknecht, G. The Oxycoal Process with Cryogenic Oxygen Supply. *Naturwissenschaften* **2009**, *96*, 993–1010.

(517) Yu, J.; Xie, L.-H.; Li, J.-R.; Ma, Y.; Seminario, J. M.; Balbuena, P. B. CO_2 Capture and Separations Using MOFs: Computational and Experimental Studies. *Chem. Rev.* **2017**, *117*, 9674–9754.

(518) Scheffknecht, G.; Al-Makhadmeh, L.; Schnell, U.; Maier, J. Oxy-Fuel Coal Combustion-A Review of the Current State-of-the-Art. *Int. J. Greenh. Gas Control* **2011**, *5S*, S16–S35.

(519) Smith, A. R.; Klosek, J. A Review of Air Separation Technologies and Their Integration with Energy Conversion Processes. *Fuel Process. Technol.* **2001**, *70*, 115–134.

(520) Sava Gallis, D. F.; Chapman, K. W.; Rodriguez, M. A.; Greathouse, J. A.; Parkes, M. V.; Nenoff, T. M. Selective O_2 Sorption at Ambient Temperatures via Node Distortions in Sc-MIL-100. *Chem. Mater.* **2016**, *28*, 3327–3336.

(521) Sava Gallis, D. F.; Parkes, M. V.; Greathouse, J. A.; Zhang, X.; Nenoff, T. M. Enhanced O_2 Selectivity versus N_2 by Partial Metal Substitution in Cu-BTC. *Chem. Mater.* **2015**, *27*, 2018–2025.

(522) Verma, P.; Maurice, R.; Truhlar, D. G. Identifying the Interactions That Allow Separation of O_2 from N_2 on the Open Iron Sites of $Fe_2(Dobdc)$. *J. Phys. Chem. C* **2015**, *119*, 28499–28511.

(523) Parkes, M. V.; Greathouse, J. A.; Hart, D. B.; Sava Gallis, D. F.; Nenoff, T. M. *Ab Initio* Molecular Dynamics Determination of Competitive O_2 vs. N_2 Adsorption at Open Metal Sites of $M_2(Dobdc)$. *Phys. Chem. Chem. Phys.* **2016**, *18*, 11528–11538.

(524) Zhang, W.; Banerjee, D.; Liu, J.; Schaef, H. T.; Crum, J. V.; Fernandez, C. A.; Kukkadapu, R. K.; Nie, Z.; Nune, S. K.; Motkuri, R. K.; Chapman, K. W.; Engelhard, M. H.; Hayes, J. C.; Silvers, K. L.; Krishna, R.; McGrail, B. P.; Liu, J.;

Thallapally, P. K. Redox-Active Metal-Organic Composites for Highly Selective Oxygen Separation Applications. *Adv. Mater.* **2016**, *28*, 3572–3577.

(525) Xiao, D. J.; Gonzalez, M. I.; Darago, L. E.; Vogiatzis, K. D.; Haldoupis, E.; Gagliardi, L.; Long, J. R. Selective, Tunable O_2 Binding in Cobalt(II)-Triazolate/Pyrazolate Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2016**, *138*, 7161–7170.

(526) Lorzing, G. R.; Trump, B. A.; Brown, C. M.; Bloch, E. D. Selective Gas Adsorption in Highly Porous Chromium(II)-Based Metal-Organic Polyhedra. *Chem. Mater.* **2017**, *29*, 8583−8587.

(527) Ilyakina, E. V.; Poddel'sky, A. I.; Cherkasov, V. K.; Abakumov, G. A. Binding of NO by Nontransition Metal Complexes. *Mendeleev Commun.* **2012**, *22*, 208–210.

(528) Bobbitt, N. S.; Snurr, R. Q. *Ab Initio* Screening of Metal Catecholates for Adsorption of Toxic Pnictogen Hydride Gases. *Ind. Eng. Chem. Res.* **2017**, *56*, 14324–14336.

(529) Huang, Z.; Liu, D.; Camacho-Bunquin, J.; Zhang, G.; Yang, D.; López-Encarnación, J. M.; Xu, Y.; Ferrandon, M. S.; Niklas, J.; Poluektov, O. G.; Jellinek, J.; Lei, A.; Bunel, E. E.; Delferro, M. Supported Single-Site Ti(IV) on a Metal–Organic Framework for the Hydroboration of Carbonyl Compounds. *Organometallics* **2017**, *36*, 3921−3930.

(530) Camacho-Bunquin, J.; Ferrandon, M.; Das, U.; Dogan, F.; Liu, C.; Larsen, C.; Platero-Prats, A. E.; Curtiss, L. A.; Hock, A. S.; Miller, J. T.; Nguyen, S. T.; Marshall, C. L.; Delferro, M.; Stair, P. C. Supported Aluminum Catalysts for Olefin Hydrogenation. *ACS Catal.* **2017**, *7*, 689–694.

(531) Zhang, X.; Vermeulen, N. A.; Huang, Z.; Cui, Y.; Liu, J.; Krzyaniak, M. D.; Li, Z.; Noh, H.; Wasielewski, M. R.; Delferro, M.; Farha, O. K. Effect of Redox "Non-Innocent" Linker on the Catalytic Activity of Copper-Catecholate-Decorated Metal–Organic Frameworks. *ACS Appl. Mater. Interfaces* **2018**, *10*, 635–641.

(532) Karlström, G.; Lindh, R.; Malmqvist, P. Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P. O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. MOLCAS: A Program Package for Computational Chemistry. *Comput. Mater. Sci.* **2003**, *28*, 222–239.

(533) Beebe, N. H. F.; Linderberg, J. Simplifications in the Generation and Transformation of Two-Electron Integrals in Molecular Calculations. *Int. J. Quant. Chem.* **1977**, *12*, 683–705.

(534) Dunning, T. H.; Botch, B. H.; Harrison, J. F. On the Orbital Description of the $4s3d^{N+1}$ States of the Transition Metal Atoms. *J. Chem. Phys.* **1980**, *72*, 3419–3420.

(535) Botch, B. H.; Dunning, T. H.; Harrison, J. F. Valence Correlation in the s^2d^n, Sd^{N+1}, and d^{N+2} States of the First-row Transition Metal Atoms. *J. Chem. Phys.* **1981**, *75*, 3466–3476.

(536) Andersson, K.; Roos, B. O. Excitation Energies in the Nickel Atom Studied with the Complete Active Space SCF Method and Second-Order Perturbation Theory.

Chem. Phys. Lett. **1992**, *191*, 507–514.

(537) Demir, H.; Stoneburner, S. J.; Jeong, W.; Ray, D.; Zhang, X.; Farha, O. K.; Cramer, C. J.; Siepmann, J. I.; Gagliardi, L. Metal-Organic Frameworks with Metal Catecholates for O2/N2 Separation. *J. Phys. Chem. C* Article ASAP, DOI: 10.1021/acs.jpcc.9b02848.

(538) Fernández-Barquín, A.; Casado-Coterillo, C.; Valencia, S.; Irabien, A. Mixed Matrix Membranes for O_2/N_2 Separation: The Influence of Temperature. *Membranes.* **2016**, *6*, 28.

(539) Santos, J. C.; Cruz, P.; Regala, T.; Magalhães, F. D.; Mendes, A. High-Purity Oxygen Production by Pressure Swing Adsorption. *Ind. Eng. Chem. Res.* **2007**, *46*, 591–599.

(540) Ferreira, D.; Boaventura, M.; Bárcia, P.; Whitley, R. D.; Mendes, A. Two-Stage Vacuum Pressure Swing Adsorption Using AgLiLSX Zeolite for Producing 99.5+% Oxygen from Air. *Ind. Eng. Chem. Res.* **2016**, *55*, 722–736.

(541) Kim, M. B.; Jee, J. G.; Bae, Y. S.; Lee, C. H. Parametric Study of Pressure Swing Adsorption Process to Purify Oxygen Using Carbon Molecular Sieve. *Ind. Eng. Chem. Res.* **2005**, *44*, 7208–7217.

(542) Jee, J.-G.; Lee, S.-J.; Kim, M.-B.; Lee, C.-H. Three-Bed PVSA Process for High-Purity O_2 Generation from Ambient Air. *AIChE J.* **2005**, *51* (11), 2988–2999.

(543) Adil, K.; Belmabkhout, Y.; Pillai, R. S.; Cadiau, A.; Bhatt, P. M.; Assen, A. H.; Maurin, G.; Eddaoudi, M. Gas/Vapour Separation Using Ultra-Microporous Metal-Organic Frameworks: Insights into the Structure/Separation Relationship. *Chem. Soc. Rev.* **2017**, *46*, 3402–3430.

(544) Rege, S. U.; Yang, R. T. Limits for Air Separation by Adsorption with LiX Zeolite. *Ind. Eng. Chem. Res.* **1997**, *36*, 5358–5365.

(545) Baksh, M. S. A.; Kikkinides, E. S.; Yang, R. T. Lithium Type X Zeolite as a Superior Sorbent for Air Separation. *Sep. Sci. Technol.* **1992**, *27*, 277–294.

(546) Zanota, M.-L.; Heymans, N.; Gilles, F.; Su, B.-L.; De Weireld, G. Thermodynamic Study of LiNaKLSX Zeolites with Different Li Exchange Rate for N_2/O_2 Separation Process. *Microporous Mesoporous Mater.* **2011**, *143*, 302–310.

(547) Jayaraman, A.; Yang, R. T. Stable Oxygen-Selective Sorbents for Air Separation. *Chem. Eng. Sci.* **2005**, *60*, 625–634.

(548) Lee, S.-J.; Jung, J.-H.; Moon, J.-H.; Jee, J.-G.; Lee, C.-H. Parametric Study of the Three-Bed Pressure-Vacuum Swing Adsorption Process for High Purity O_2 Generation from Ambient Air. *Ind. Eng. Chem. Res.* **2007**, *46*, 3720–3728.

(549) Nolan, J. T.; McKeehan, T. W.; Danner, R. P. Equilibrium Adsorption of Oxygen, Nitrogen, Carbon Monoxide, and Their Binary Mixtures on Molecular Sieve Type 10X. *J. Chem. Eng. Data* **1981**, *26*, 112–115.

(550) Kim, Y. H.; Lee, D. G.; Moon, D. K.; Byeon, S. H.; Ahn, H. W.; Lee, C. H. Effect of Bed Void Volume on Pressure Vacuum Swing Adsorption for Air Separation. *Korean J. Chem. Eng.* **2014**, *31* (1), 132–141.

(551) Zhang, L.; Siepmann, J. I. Direct Calculation of Henry's Law Constants from Gibbs Ensemble Monte Carlo Simulations: Nitrogen, Oxygen, Carbon Dioxide and Methane in Ethanol. *Theor. Chem. Acc.* **2006**, *115*, 391–397.

(552) Richards, A. J.; Watanabe, K.; Austin, N.; Stapleton, M. R. Computer Simulation of the Gas Separation Properties of Zeolite Li-X. *J. Porous Mater.* **1995**, *2*, 43–49.

(553) Kitagawa, S.; Kitaura, R.; Noro, S. I. Functional Porous Coordination Polymers. *Angew. Chemie - Int. Ed.* **2004**, *43*, 2334–2375.

(554) Férey, G. Hybrid Porous Solids: Past, Present, Future. *Chem. Soc. Rev.* **2008**, *37*, 191–214.

(555) Colón, Y. J.; Gómez-Gualdrón, D. A.; Snurr, R. Q. Topologically Guided, Automated Construction of Metal-Organic Frameworks and Their Evaluation for Energy-Related Applications. *Cryst. Growth Des.* **2017**, *17*, 5801–5810.

(556) Smit, B.; Maesen, T. L. M. Molecular Simulations of Zeolites: Adsorption, Diffusion, and Shape Selectivity. *Chem. Rev.* **2008**, *108*, 4125–4184.

(557) Watanabe, T.; Sholl, D. S. Accelerating Applications of Metal-Organic Frameworks for Gas Adsorption and Separation by Computational Screening of Materials. *Langmuir* **2012**, *28*, 14114–14128.

(558) Ryan, P.; Farha, O. K.; Broadbelt, L. J.; Snurr, R. Q. Computational Screening of Metal-Organic Frameworks for Xenon/Krypton Separation. *AIChE J.* **2011**, *57*, 1759–1766.

(559) Moghadam, P. Z.; Islamoglu, T.; Goswami, S.; Exley, J.; Fantham, M.; Kaminski, C. F.; Snurr, R. Q.; Farha, O. K.; Fairen-Jimenez, D. Computer-Aided Discovery of a Metal-Organic Framework with Superior Oxygen Uptake. *Nat. Commun.* **2018**, *9*, 1378.

(560) DeCoste, J. B.; Weston, M. H.; Fuller, P. E.; Tovar, T. M.; Peterson, G. W.; LeVan, M. D.; Farha, O. K. Metal-Organic Frameworks for Oxygen Storage. *Angew. Chemie - Int. Ed.* **2014**, *53* (51), 14092–14095.

(561) Fang, H.; Demir, H.; Kamakoti, P.; Sholl, D. S. Recent Developments in First-Principles Force Fields for Molecules in Nanoporous Materials. *J. Mater. Chem. A* **2014**, *2*, 274–291.

(562) Basdogan, Y.; Keskin, S. Simulation and Modelling of MOFs for Hydrogen Storage. *CrystEngComm* **2015**, *17*, 261–275.

(563) Assfour, B.; Seifert, G. Hydrogen Storage in 1D Nanotube-like Channels Metal-Organic Frameworks: Effects of Free Volume and Heat of Adsorption on Hydrogen Uptake. *Int. J. Hydrogen Energy* **2009**, *34*, 8135–8143.

(564) Lennox, M. J.; Bound, M.; Henley, A.; Besley, E. The Right Isotherms for the Right Reasons? Validation of Generic Force Fields for Prediction of Methane Adsorption in Metal-Organic Frameworks. *Mol. Simul.* **2017**, *43*, 828–837.

(565) Zeitler, T. R.; Van Heest, T.; Sholl, D. S.; Allendorf, M. D.; Greathouse, J. A. Predicting Low-Pressure O_2 Adsorption in Nanoporous Framework Materials for Sensing Applications. *ChemPhysChem* **2013**, *14*, 3740–3750.

(566) Demir, H.; Greathouse, J. A.; Staiger, C. L.; Perry IV, J. J.; Allendorf, M. D.; Sholl, D. S. DFT-Based Force Field Development for Noble Gas Adsorption in Metal Organic Frameworks. *J. Mater. Chem. A* **2015**, *3*, 23539–23548.

(567) Coudert, F.-X.; Fuchs, A. H. Computational Characterization and Prediction of Metal-Organic Framework Properties. *Coord. Chem. Rev.* **2016**, *307*, 211–236.

(568) Wang, Y.; Yang, J.; Li, Z.; Zhang, Z.; Li, J.; Yang, Q.; Zhong, C. Computational Study of Oxygen Adsorption in Metal-Organic Frameworks with Exposed Cation Sites: Effect of Framework Metal Ions. *RSC Adv.* **2015**, *5*, 33432–33437.

(569) Murray, L. J.; Dinca, M.; Yano, J.; Chavan, S.; Bordiga, S.; Brown, C. M.; Long, J. R. Highly-Selective and Reversible O_2 Binding in Cr_3(1,3,5-Benzenetricarboxylate)$_2$. *J. Am. Chem. Soc.* **2010**, *132*, 7856–7857.

(570) Bae, Y.-S.; Spokoyny, A. M.; Farha, O. K.; Snurr, R. Q.; Hupp, J. T.; Mirkin, C. A. Separation of Gas Mixtures Using Co(II) Carborane-Based Porous Coordination Polymers. *Chem. Commun.* **2010**, *46*, 3478–3480.

(571) Bloch, E. D.; Queen, W. L.; Hudson, M. R.; Mason, J. A.; Xiao, D. J.; Murray, L. J.; Flacau, R.; Brown, C. M.; Long, J. R. Hydrogen Storage and Selective, Reversible O_2 Adsorption in a Metal–Organic Framework with Open Chromium(II) Sites. *Angew. Chemie - Int. Ed.* **2016**, *55*, 8605–8609.

(572) Li, Y.; Yang, R. T. Gas Adsorption and Storage in Metal-Organic Framework MOF-177. *Langmuir* **2007**, *23*, 12937–12944.

(573) Mu, B.; Schoenecker, P. M.; Walton, K. S. Gas Adsorption Study on Mesoporous Metal-Organic Framework UMCM-1. *J. Phys. Chem. C* **2010**, *114*, 6464–6471.

(574) Yuan, S.; Feng, L.; Wang, K.; Pang, J.; Bosch, M.; Lollar, C.; Sun, Y.; Qin, J.; Yang, X.; Zhang, P.; Wang, Q.; Zou, L.; Zhang, Y.; Zhang, L.; Fang, Y.; Li, J.; Zhou, H.-C. Stable Metal–Organic Frameworks: Design, Synthesis, and Applications. *Adv. Mater.* **2018**, *30*, 1704303.

(575) Howarth, A. J.; Liu, Y.; Li, P.; Li, Z.; Wang, T. C.; Hupp, J. T.; Farha, O. K. Chemical, Thermal and Mechanical Stabilities of Metal–Organic Frameworks. *Nat. Rev. Mater.* **2016**, *1*, 15018.

(576) Zhang, X.; Vermeulen, N. A.; Huang, Z.; Cui, Y.; Liu, J.; Krzyaniak, M. D.; Li, Z.; Noh, H.; Wasielewski, M. R.; Delferro, M.; Farha, O. K. Effect of Redox "Non-Innocent" Linker on the Catalytic Activity of Copper-Catecholate-Decorated Metal–Organic Frameworks. *ACS Appl. Mater. Interfaces* **2018**, *10*, 635–641.

(577) Spek, A. L. Single-Crystal Structure Validation with the Program PLATON. *J. Appl. Crystallogr.* **2003**, *36*, 7–13.

(578) Spek, A. L. Structure Validation in Chemical Crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2009**, *65*, 148–155.

(579) Willems, T. F.; Rycroft, C. H.; Kazi, M.; Meza, J. C.; Haranczyk, M. Algorithms and Tools for High-Throughput Geometry-Based Analysis of Crystalline Porous Materials. *Microporous Mesoporous Mater.* **2012**, *149*, 134–141.

(580) Witman, M.; Ling, S.; Jawahery, S.; Boyd, P. G.; Haranczyk, M.; Slater, B.; Smit,

B. The Influence of Intrinsic Framework Flexibility on Adsorption in Nanoporous Materials. *J. Am. Chem. Soc.* **2017**, *139*, 5547–5557.

(581) Gee, J. A.; Sholl, D. S. Effect of Framework Flexibility on C_8 Aromatic Adsorption at High Loadings in Metal–Organic Frameworks. *J. Phys. Chem. C* **2016**, *120* (1), 370–376.

(582) Wang, X.; Krishna, R.; Li, L.; Wang, B.; He, T.; Zhang, Y.-Z.; Li, J.-R.; Li, J. Guest-Dependent Pressure Induced Gate-Opening Effect Enables Effective Separation of Propene and Propane in a Flexible MOF. *Chem. Eng. J.* **2018**, *346*, 489–496.

(583) Bakhshian, S.; Sahimi, M. Theoretical Model and Numerical Simulation of Adsorption and Deformation in Flexible Metal-Organic Frameworks. *J. Phys. Chem. C* **2018**, *122*, 9465–9473.

(584) Moghadam, P. Z.; Fairen-Jimenez, D.; Snurr, R. Q. Efficient Identification of Hydrophobic MOFs: Application in the Capture of Toxic Industrial Chemicals. *J. Mater. Chem. A* **2016**, *4*, 529–536.

(585) Borycz, J.; Lin, L.-C.; Bloch, E. D.; Kim, J.; Dzubak, A. L.; Maurice, R.; Semrouni, D.; Lee, K.; Smit, B.; Gagliardi, L. CO_2 Adsorption in Fe_2(Dobdc): A Classical Force Field Parameterized from Quantum Mechanical Calculations. *J. Phys. Chem. C* **2014**, *118*, 12230–12240.

(586) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithm to Applications*; Academic Press: San Diego, CA, 2002.

(587) Dubbeldam, D.; Torres-Knoop, A.; Walton, K. S. On the Inner Workings of Monte Carlo Codes. *Mol. Simul.* **2013**, *39*, 1253–1292.

(588) Islamoglu, T.; Ray, D.; Li, P.; Majewski, M. B.; Akpinar, I.; Zhang, X.; Cramer, C. J.; Gagliardi, L.; Farha, O. K. From Transition Metals to Lanthanides to Actinides: Metal-Mediated Tuning of Electronic Properties of Isostructural Metal–Organic Frameworks. *Inorg. Chem.* **2018**, *57*, 13246–13251.

(589) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

(590) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

(591) Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.

(592) Baerends, E. J.; Ziegler, T.; Atkins, A. J.; Autschbach, J.; Bashford, D.; Baseggio, O.; Bérces, A.; Bickelhaupt, F. M.; Bo, C.; Boerritger, P. M.; Cavallo, L.; Daul, C.; Chong, D. P.; Chulhai, D. V; Deng, L.; Dickson, R. M.; Dieterich, J. M.; Ellis, D. E.; van Faassen, M.; Ghysels, A.; Giammona, A.; van Gisbergen, S. J. A.; Goez, A.; Götz, A. W.; Gusarov, S.; Harris, F. E.; van den Hoek, P.; Hu, Z.; Jacob, C. R.; Jacobsen, H.; Jensen, L.; Joubert, L.; Kaminski, J. W.; van Kessel, G.; König, C.; Kootstra, F.; Kovalenko, A.; Krykunov, M.; van Lenthe, E.; McCormack, D. A.; Michalak, A.; Mitoraj, M.; Morton, S. M.; Neugebauer, J.;

Nicu, V. P.; Noodleman, L.; Osinga, V. P.; Patchkovskii, S.; Pavanello, M.; Peeples, C. A.; Philipsen, P. H. T.; Post, D.; Pye, C. C.; Ramanantoanina, H.; Ramos, P.; Ravenek, W.; Rodríguez, J. I.; Ros, P.; Rüger, R.; Schipper, P. R. T.; Schlüns, D.; van Schoot, H.; Schreckenbach, G.; Seldenthuis, J. S.; Seth, M.; Snijders, J. G.; Solà, M.; M., S.; Swart, M.; Swerhone, D.; te Velde, G.; Tognetti, V.; Vernooijs, P.; Versluis, L.; Visscher, L.; Visser, O.; Wang, F.; Wesolowski, T. A.; van Wezenbeek, E. M.; Wiesenekker, G.; Wolff, S. K.; Woo, T. K.; Yakovlev, A. L. ADF2017, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, Https://Www.Scm.Com.

(593) van Lenthe, E.; Baerends, E. J. Optimized Slater-Type Basis Sets for the Elements 1-118. *J. Comput. Chem.* **2003**, *24*, 1142–1156.

(594) Van Lenthe, E.; Van Leeuwen, R.; Baerends, E. J.; Snijders, J. G. Relativistic Regular Two-Component Hamiltonians. *Int. J. Quantum Chem.* **1996**, *57*), 281–293.

(595) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic Total Energy Using Regular Approximations. *J. Chem. Phys.* **1994**, *101*, 9783–9792.

(596) van Lenthe, E.; Ehlers, A.; Baerends, E.-J. Geometry Optimizations in the Zero Order Regular Approximation for Relativistic Effects. *J. Chem. Phys.* **1999**, *110*, 8943–8953.

(597) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for *Ab Initio* Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(598) Yang, D.; Momeni, M. R.; Demir, H.; Pahls, D. R.; Rimoldi, M.; Wang, T. C.; Farha, O. K.; Hupp, J. T.; Cramer, C. J.; Gates, B. C.; Gagliardi, L. Tuning the Properties of Metal-Organic Framework Nodes as Supports of Single-Site Iridium Catalysts: Node Modification by Atomic Layer Deposition of Aluminium. *Faraday Discuss.* **2017**, *201*, 195–206.

(599) Planas, N.; Mondloch, J. E.; Tussupbayev, S.; Borycz, J.; Gagliardi, L.; Hupp, J. T.; Farha, O. K.; Cramer, C. J. Defining the Proton Topology of the Zr_6-Based Metal-Organic Framework NU-1000. *J. Phys. Chem. Lett.* **2014**, *5*, 3716–3723.

(600) Mu, B.; Schoenecker, P. M.; Walton, K. S. Gas Adsorption Study on Mesoporous Metal−Organic Framework UMCM-1. *J. Phys. Chem. C* **2010**, *114*, 6464–6471.

(601) Tang, W.; Sanville, E.; Henkelman, G. A Grid-Based Bader Analysis Algorithm without Lattice Bias. *J. Phys. Condens. Matter* **2009**, *21*, 084204.

(602) Sanville, E.; Kenny, S. D.; Smith, R.; Henkelman, G. Improved Grid-Based Algorithm for Bader Charge Allocation. *J Comput Chem* **2007**, *28*, 899–908.

(603) Henkelman, G.; Arnaldsson, A.; Jónsson, H. A Fast and Robust Algorithm for Bader Decomposition of Charge Density. *Comput. Mater. Sci.* **2006**, *36*, 354–360.

(604) Yu, M.; Trinkle, D. R. Accurate and Efficient Algorithm for Bader Charge Integration. *J. Chem. Phys.* **2011**, *134*, 064111.

(605) Bader, R. F. W. *Atoms in Molecules. A Quantum Theory.*; Oxford University Press: New York, 1990.