Modeling the Energetic Landscape of Transition Metal Complexes via Electronic Structure Theory and Chemically-Informed Artificial Intelligence Methodologies

Alexis W. Mills

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Xiaosong Li, Chair

Brandi Cossairt

Ting Cao

Program Authorized to Offer Degree:
Chemistry

University of Washington

Abstract

Modeling the Energetic Landscape of Transition Metal Complexes via Electronic Structure Theory and Chemically-Informed Artificial Intelligence Methodologies

Alexis W. Mills

Chair of the Supervisory Committee:
Professor Xiaosong Li
Department of Chemistry

Methods used to computationally study the electronic structures of complex chemical systems are ever-evolving to address the desire for increased accuracy and reduced computational cost. Modern advancements have given rise to sophisticated methodologies, such as Density Functional Theory (DFT), and offer a means in which to evaluate systems ranging from simple organic materials to transition metal nanoparticles. Part I of this thesis will employ DFT in order to simulate spectroscopic signatures of bimetallic platinum(II) complexes and analyze the electronic structure as a function of the complex ligand. Through this study, a general trend can be extracted to define a set of design rules for building Pt(II) dimer complexes with desirable electron transfer behavior. Part II of this work will introduce a physics-informed reinforcement machine learning (RL) algorithm that has been designed to learn from physically-motivated actions and seek to address the cost challenges of studying large metal-hydride systems by adapting the RL algorithm to simulate the electronic landscape. Electronic structure theory will supplement the RL algorithm, in the metal-hydride material application, to improve the quality of the simulated physical properties and aid in training the model.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY

AC: Actor-critic

AI: Artificial intelligence

AO: Atomic orbital

BL: Bridging ligand

BOMD: Born-Oppenheimer molecular dynamics

BPY: 7,8-benzoquinoline

CML: Cyclometalating ligand

CT: Charge transfer

DDPG: Deep Deterministic Policy Gradient

DFT: Density Functional Theory

DP: Dynamic programming

HDDPG: hybrid-DDPG

HK: Hohenberg-Kohn

HOMO: Highest occupied molecular orbital

ISC: Intersystem crossing

KS: Kohn-Sham

LC: Ligand centered

LDA: Local density approximation

LUMO: Lowest unoccupied molecular orbital

MB: Müller-Brown

MDP: Markov decision process

MG: Markov game

ML: Machine learning

MLCT: Metal-to-ligand charge transfer

MMLCT: Metal-metal-to-ligand charge transfer

NN: Neural network

NTO: Natural transition orbitals

OU: Ornstein-Uhlenbeck

PES: Potential energy surface

PPY: 2-phenylpyridine

RL: Reinforcement learning

SCF: Self consistent field

SE: Schrödinger Equation

TDDFT: Time-Dependent Density Functional Theory

X2C: Exact-two-component

XC: Exchange-Correlation

ACKNOWLEDGMENTS

Graduate school has been one of the most valuable and humbling experiences I have had the opportunity of enduring. I have learned a great deal from both my colleagues and research that I will carry with me throughout my life.

I would like to begin by expressing my deepest appreciation to Prof. Xiaosong Li. This endeavor would not have been possible without his endless patience, encouragement, and incredible support in the form of both guidance and challenges. In addition to granting me the opportunity to completely change disciplines from physical organic chemistry to theoretical chemistry as I transitioned into graduate school, he continued to open countless opportunities for me to grow and establish myself as a professional. Through his support, I was able to participate in the DIRECT program and quickly discover my passion for artificial intelligence research.

Next, I would like to credit all of the post-docs I have had the pleasure of working with over these past few years. First, a huge thanks is owed to Dr. Andrew Valentine for challenging me and regularly providing constructive criticism. These exercises and feedback have instilled a more critical perspective and have encouraged me to continually ask questions. I would like to extend my gratitude and utmost appreciation to Dr. Joshua Goings, who offered both invaluable insights into machine learning and practical suggestions in response to my proposals. Finally, I would like to thank Dr. Chad Hoyer, Dr. Andrew Wildman, and Dr. Ernesto Baez-Martinez for all the Gloomhaven board game nights fueled by decaffeinated coffee and good company.

I would be remiss if I did not acknowledge the mentor who inspired an endless desire for learning and discovery, Prof. Michael Walter, my undergraduate PI at the

University of North Carolina at Charlotte. His infectious enthusiasm and curiosity inspired me to pursue a higher education and motivates me to continue learning each and every day.

I am extremely grateful for my family who has shown constant support and interest in my studies over these past few years. They have been my rock from the very beginning and have continually encouraged me to show up in the face of adversity and do my best. I am so very thankful for all of the lessons and guidance they have taught and provided me throughout my life. The experiences they have created for me have been instrumental in my success.

Finally, I would like to end with a special thanks to the most important person in my life, my husband. I truly believe none of this would have been possible if not for my husband's unparalleled patience and unwavering support. The past few years have been periods of immense growth personally and professionally for the both of us. I am incredibly proud of everything that we have accomplished together and all that he has achieved in industry as a professional chemist. I am forever thankful for his willingness to reestablish himself from an organic chemist to a biochemist in order to support us as we built our life in Seattle. I am eternally grateful for all that he has done and continues to do for us.

DEDICATION

to my loving husband

Part I

ELECTRONIC STRUCTURE THEORY

Chapter 1

THEORETICAL BACKGROUND

1.1 Density Functional Theory

Prior to discussing the results presented in this thesis, it is worthwhile introducing the methods used to computationally study the systems investigated. This discussion will begin with a brief introduction to density functional theory (DFT), which is a standard methodology that allows one to quantum mechanically treat a system and characterize the electronic structure of matter.[88, 23] This is done by evaluating the interactions between the electrons within the system and the electrons with an external potential; where the external potentials simulate the electron's attraction to the nuclei of the system. When deriving DFT, it is most natural to begin with the time-independent Schrödinger eigenequation

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{1.1}$$

where $\hat{H}$ is the Hamiltonian operator that acts on Ψ, the eigenvector, to find the scalar eigenvalue, E. Unless otherwise stated, the Hamiltonian considered in this work will be limited to the nonrelativistic variety, meaning interactions such as spin-orbit coupling are neglected for simplicity. Additionally, the eigenvector, Ψ, in this context will refer to the wave function of the system that holds all relevant information about the system's current quantum state.

It is useful to decompose $\hat{H}$ into the electronic kinetic ($\hat{T}_e$), nuclear kinetic ($\hat{T}_n$), electron-electron ($\hat{V}_{ee}$), electron-nuclear ($\hat{V}_{en}$), and nuclear-nuclear ($\hat{V}_{nn}$) interaction terms such that the Hamiltonian becomes, $\hat{H} = \hat{T}_e + \hat{T}_n + \hat{V}_{ee} + \hat{V}_{en} + \hat{V}_{nn}$. Consid-

ering the various interaction terms and recognizing the inherent complexity that accompanies systems of high dimensionality, it is apparent that analytically solving the Schrödinger equation for larger systems quickly becomes prohibitive.[46] To address this, a number of approximations are assumed in an effort to ease the computational challenge; the most fundamental being the Born-Oppenheimer approximation. This approximation postulates that the electrons move considerably faster than their nuclei counterparts thus allowing the nuclei to be treated as stationary particles.[196] As such, terms exclusively pertaining to the nuclei are fixed leaving any variability to the electronic terms. These terms are defined below and can be used to construct the electronic Hamiltonian:

$$\hat{T} = -\frac{1}{2}\sum_{i} \nabla_i^2 \tag{1.2}$$

$$\hat{V}_{ee} = \sum_{i<j} \frac{1}{r_i - r_j} \tag{1.3}$$

$$\hat{V}_{en} = -\sum_{i,I} \frac{Z_I}{r_i - R_I} \tag{1.4}$$

where lowercase letters pertain to electrons and uppercase letters represent properties of the nuclei. The indices of the corresponding particles are represented by, $i, j, ..., etc.$, and the positions of these particles are held within the r vector. Additionally, Z is the charge of the corresponding nucleus. Provided the Born-Oppenheimer approximation is enforced, the V_{en} term can be redefined as V_{ext}:

$$\hat{V}_{ext} = \sum_{i} \nu(r_i) \tag{1.5}$$

where ν is the external potential acting only on the electrons of the system. Note the $\hat{T}$ and Coulomb ($\hat{V}_{ee}$) terms are both considered universal and therefore the same for all N-electron systems, while the $\hat{V}_{en}$ term is system-dependent as positively charged particles become relevant.[23] Once the latter term has been defined for the

system, Ψ can be determined via **Equation 1.1**, then the observables can be evaluated by calculating the expectation values for the corresponding operators of interest. The wave function, however, is quite abstract compared to experimentally observable properties, such as the electron density, imploring one to identify a more approachable description of the system. This realization lays the groundwork for the development of DFT.

1.1.1 Hohenberg-Kohn Theorems

Prior to 1964, DFT was non-existent and the Thomas-Fermi model[1] was utilized to approximate the electronic distribution of a system. The foundation for DFT was made possible once Hohenberg and Kohn (HK) introduced two theorems. The first stating the ground-state energy of a system is related to the electron density and is thus unique to that system and the second stating the ground-state energy can be determined by minimizing the energy via the electron density.[204, 6]

HK Theorem I

A *reductio ad absurdum* argument will be pursued to demonstrate this first theorem. To begin, a system with a nondegenerate ground-state is considered and the following relationship is declared

$$\int \rho(r)dr - N_e = 0 \tag{1.6}$$

which indicates the number of electrons in the system, N_e, is determined by the electron density, ρ. Next, two potentials are introduced, $v(r)$ and $v'(r)$, that have ground-states, Ψ and Ψ', respectively, and correspond to the same electronic density. As such, Ψ' will be used as a trial function for H. Considering the variational

[1] L. H. Thomas and E. Fermi independently devised an approximation in which to model the electron distribution of a system and conveniently published their models in the same year, 1927.[175]

principle[2], the following can be defined[204]

$$E < \langle \Psi' | H | \Psi' \rangle = \langle \Psi' | H' | \Psi' \rangle + \langle \Psi' | H - H' | \Psi' \rangle \tag{1.7}$$

$$E < E' + \int \rho(r) \{ \nu(r) - \nu'(r) \} dr \tag{1.8}$$

and conversely

$$E' < \langle \Psi | H' | \Psi \rangle = \langle \Psi | H | \Psi \rangle + \langle \Psi | H' - H | \Psi \rangle \tag{1.9}$$

$$E' < E - \int \rho(r) \{ \nu(r) - \nu'(r) \} dr \tag{1.10}$$

such that the summation of **Equation 1.8** and **Equation 1.10** become

$$E + E' < E' + E \tag{1.11}$$

which is not possible and therefore shows the ground-state density must be related to a unique external potential.

HK Theorem II

The second theorem proposes that once the ground-state electron density has been determined, then all information about the molecule can be known. Their work proved a universal functional of the electron density, $F[\rho]$, exists and gives the ground-state energy, provided the appropriate ν is identified.[76, 23]

$$E_0 = \min_{\Psi \to \rho_0} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle + \int d^3 r \rho(r) \nu(r) =: F[\rho] + V[\rho] \tag{1.12}$$

where $F[\rho]$ is simply the sum of both the kinetic and Coulomb repulsion energies. The above implies that if one knows the ground-state density of the system, then the

[2]The variational method is simply a means by which a trial wave function is adjusted, via variational parameters, to minimize the energy. In short, it allows one to approximate an impossible problem by evaluating a solvable problem.

unique external potential of the system can be determined and the ground-state Ψ can be found through a functional relationship. Assuming this holds true, then ground-state observables are also functionals of ρ_0 giving rise to the general relationship in **Equation 1.13**.[8]

$$O[\rho] = \langle \Psi_0[\rho] | \hat{O} | \Psi_0[\rho] \rangle \tag{1.13}$$

1.1.2 The Kohn-Sham Equations: Exchange-Correlation Functionals

In 1967 Kohn and Sham revisited the relationship defined in **Equation 1.12** and devised the following relationship

$$E = \int \nu(r)\rho(r)dr + \frac{1}{2}\int\int \frac{\rho(r)\rho(r')}{|r-r'|}drdr' + T_s[\rho(r)] + E_{xc}[\rho(r)] \tag{1.14}$$

where $T_s[\rho(r)]$ is the kinetic energy and $E_{xc}[\rho(r)]$ is the exchange-correlation (XC) energy of the system.[89] The former considers non-interacting electrons while the latter considers interacting electrons. The authors then declare if the density of the system is changing sufficiently slowly, thus abiding by the local density approximation (LDA), then the following can be said about the exchange-correlation term

$$E_{xc}[\rho] = \int \rho(r)\epsilon_{xc}(\rho(r))dr \tag{1.15}$$

where ϵ_{xc} is the XC energy of each electron.[89, 88] Plenty of ongoing work aims at developing suitable XC functionals beyond those constrained to the just LDA. Additional classes of XC functionals include:

1. *Gradient-Corrected Functionals*: dependent on the density and gradient of the density
2. *Meta-Generalized Gradient Approximation (Meta-GGA)*: utilizes density, density gradient, and higher order derivatives of the density
3. *Hybrid functionals*: based on the previous two functionals and adds an additional term with some portion of Hartree-Fock exchange energy contribution

1.2 Linear Response Function

Deriving the linear response function aids in a deeper understanding of how various properties, namely excitation energies, are calculated allowing for a better understanding of the system's response to external forces. It is assumed the forces acting on the system are sufficiently small such that they can be described reasonably well mathematically. The two main components that must be considered when discussing time-dependent (TD) DFT include the Runge-Gross Theorem and the Kohn-Sham (KS) formalism. The former allows a fixed initial state such that a relationship between ρ and an external potential can be made while the latter maintains the HK theorem holds true and a non-interacting reference exists for the system investigated.[89, 46, 159]

1.2.1 Runge-Gross Theorem

The Runge-Gross theorem expands on previous work in the field and extends to the TD domain. The theorem states that a direct mapping exists between the potential a system evolves under and the density of that system.[159] Just as before, the brief derivation will begin with the Schrödinger equation, however, this time the focus will be on the time-dependent alternative since a "response" is sought after, implying an evolution with respect to time.

$$\hat{H}\Psi = i\frac{\partial \Psi}{\partial t} \tag{1.16}$$

where

$$\hat{H} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ext} \tag{1.17}$$

and the $\hat{V}_{ext}$ term becomes

$$\hat{V}_{ext} = \sum_{i}^{N} \nu(r_i, t) \tag{1.18}$$

where N is the number of electrons and ν is the single-particle potential; note the potential is now dependent on time in addition to the position. Having redefined the above equations to accommodate the TD case, a brief proof of the Runge-Gross theorem may proceed to show that two densities arising from a common ground state as a result of two different potentials will always be unique if, and only if, the potentials are sufficiently different.[46] That is to say

$$\nu^A(r,t) \neq \nu^B(r,t) + c(t) \tag{1.19}$$

where c is a constant at time t. It is assumed that these potentials are expandable, via a Taylor series, about t_0. Considering this and recalling that the potentials must be sufficiently different, it can be said that a small non-negative integer, k, can be identified such that [46]

$$v_k^A(r) = \frac{\partial^k}{\partial t^k}\nu^A(r,t)|_{t=t_0} \tag{1.20}$$

and similarly,

$$v_k^B(r) = \frac{\partial^k}{\partial t^k}\nu^B(r,t)|_{t=t_0} \tag{1.21}$$

differ by more than just a constant, thus

$$\nu_k^A(r) - \nu_k^B(r) \neq c \tag{1.22}$$

It is understood that the current densities, $j(r,t)$, corresponding to the above potentials are different and therefore have different electron densities. The general mathematic representation of the current density is as follows: [46]

$$j(r,t) = \frac{1}{2i}[\Psi^*\nabla\Psi - \nabla\Psi^*\Psi] \tag{1.23}$$

where the wave functions are dependent on both the electron positions and the time in

this case. It should be noted that the current density and the electron density abide by the continuity equation, which says the change in electron density is related to the flux of $j(r,t)$.[46] Through additional algebra and subtracting the two different densities, a direct relationship between the external potentials and the current densities can be made such that

$$\frac{\partial^2}{\partial t^2}[\rho^A - \rho^B] = -\nabla\frac{\partial}{\partial t}[j^A - j^B] = \nabla[\rho_0\nabla(\nu^A - \nu^B)] \tag{1.24}$$

which gives rise to a relationship between the time-dependent electron densities on the left-most side and the time-dependent external potentials on the right-most side of the equation. This eventually leads to the mentioned one-to-one mapping stated at the beginning of this section.[46]

$$\rho \leftrightarrow \nu[\rho] + c \leftrightarrow \Psi[\rho]e^{-i\alpha} \tag{1.25}$$

1.2.2 Kohn-Sham Formalism

Prior to moving forward, it is assumed that a reference state exists and is represented by the following relationship

$$\rho = \sum_i^N |\phi_i|^2 \tag{1.26}$$

where ϕ is the single-electron orbital that represents the i^{th} electron out of N total electrons and is dependent on both r and t. The single-electron orbitals can be represented as

$$\phi_i = \sum_j^M c_{ij}\chi_j \tag{1.27}$$

where χ is the time-independent single-particle ψ in basis M. These single-electron orbitals make up the single Slater determinant, Φ. It is further assumed that this non-interacting reference has an external one-particle potential, ν_s; as such, the time-

dependent Schrödinger equation can be rewritten as [46]

$$i\frac{\partial}{\partial t}\phi_i = \left(-\frac{1}{2}\nabla_i^2 + \nu_s\right)\phi_i \tag{1.28}$$

The universal function, and subsequently the action functionals, can be used to derive the exchange-correlation, A_{xc}, term which adjusts **Equation 1.28** to the following

$$i\frac{\partial}{\partial t}\phi_i = \left(-\frac{1}{2}\nabla_i^2 + \nu + \int d^3r' \frac{\rho(r',t)}{|r-r'|} + \frac{\delta A_{xc}[\rho]}{\delta\rho(r,t)}\right)\phi_i = \hat{F}^{KS}\phi_i \tag{1.29}$$

It should be noted that the exchange-correlation action functional, or the xc kernel, is unknown and thus requires one to make use of approximations. The most commonly used approximation is the adiabatic approximation, which posits that the system moves sufficiently slow as time progresses.[190]

1.2.3 Linear-Response TDDFT

Having presented a brief overview of the previous theorems, the foundation for the succeeding derivation is set and has been well defined. To begin, the notation will shift to density matrix notation and the starting state will be the ground-state as no electric field has been applied yet.

$$i\frac{\partial}{\partial t}P^0_{pr} = \sum_q [F^0_{pq}P^0_{qr} - P^0_{pq}F^0_{qr}] = 0 \tag{1.30}$$

where F_{pq} and P_{pq} correspond to the KS Hamiltonian and density matrix of the ground-state, respectively, and p, q, r refer to orbital indices. Additionally, the idempotency condition holds, meaning the sum of $P^0_{pq}P^0_{qr}$ over q is equal to P^0_{pr}.[46]

Next, an external force is applied in order to observe the first-order linear response, given by:

$$i\frac{\partial}{\partial t}P^1_{pr} = \sum_q [F^0_{pq}P^1_{qr} - P^1_{pq}F^0_{qr} + F^1_{pq}P^0_{qr} - P^0_{pq}F^1_{qr}] \tag{1.31}$$

The KS Hamiltonian is adapted to accommodate the first-order change by including the actual electric field, g_{pq}, and the change in density. The former is represented as

$$g_{pq} = \frac{1}{2}[f_{pq}e^{-i\omega t} + f^*_{qp}e^{i\omega t}] \tag{1.32}$$

while the latter is

$$\Delta F^0_{pq} = \sum_{st} \frac{\partial F^0_{pq}}{\partial P_{st}} P^1_{st} \tag{1.33}$$

where ω is the frequency and f is the one-electron operator. These two definitions come together to give the updated KS Hamiltonian

$$F^1_{pq} = g_{pq} + \Delta F^0_{pq} \tag{1.34}$$

Alternatively, the density matrix is updated such that the perturbation densities, d, are considered

$$P^1_{pq} = \frac{1}{2}[d_{pq}e^{-i\omega t} + d^*_{qp}e^{i\omega t}] \tag{1.35}$$

Through substitution, a bit of algebra, and maintaing the idempotency condition, the following is found, which forces the matrix blocks relating occupied-to-occupied and virtual-to-virtual orbitals to be zero.

$$\sum_q [P^0_{pq}P^1_{qr} + P^1_{pq}P^0_{qr}] = P^1_{pr} \tag{1.36}$$

Finally, the following equations can be derived

$$F^0_{aa}d_{ai} - d_{ai}F^0_{ii} + \left(f_{ai} + \sum_{bj} \frac{\partial F_{ai}}{\partial P_{bj}} x_{bj} + \frac{\partial F_{ai}}{\partial P_{jb}} y_{bj}\right) P^0_{ii} = \omega d_{ai} \tag{1.37}$$

$$F^0_{ii}d_{ia} - d_{ia}F^0_{aa} - P^0_{ii}\left(f_{ia} + \sum_{bj} \frac{\partial F_{ia}}{\partial P_{bj}} x_{bj} + \frac{\partial F_{ia}}{\partial P_{jb}} y_{bj}\right) P^0_{ii} = \omega d_{ia} \tag{1.38}$$

such that the following non-Hermitian eigenvalue equation can be found when a small perturbation is applied to the system [46]

$$\begin{bmatrix} A & B \\ B^* & A^* \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = \sigma \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} \tag{1.39}$$

where σ is the energy, X is the excitation from an occupied to virtual orbital while Y is the de-excitation from a virtual to an occupied orbital. The A and B matrix elements are as follows

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + (ia|jb) + (ia|f_{xc}|jb) \tag{1.40}$$

$$B_{ia,jb} = (ia|bj) + (ia|f_{xc}|bj) \tag{1.41}$$

with i, j and a, b corresponding to occupied and virtual orbitals, respectively, and ϵ is the orbital energy.

Plenty of information can be gained from **Equation 1.39** including the charge transfer (CT) numbers, which indicate the likelihood of finding a particle or hole on a given fragment of a molecule.[186] The CT numbers can be determined by first defining the right and left eigenvectors to find the transition density matrices.

$$D_{pq}^{01} = \left\langle \Psi_0 | \hat{a}_p^\dagger \hat{a}_p | \Psi_1 \right\rangle = \begin{bmatrix} 0 & X \\ Y & 0 \end{bmatrix}_{pq} \tag{1.42}$$

$$D_{pq}^{10} = \left\langle \Psi_1 | \hat{a}_p^\dagger \hat{a}_p | \Psi_0 \right\rangle = \begin{bmatrix} 0 & -Y^\dagger \\ X^\dagger & 0 \end{bmatrix}_{pq} \tag{1.43}$$

where $\hat{a}_p^\dagger$ and $\hat{a}_q$ are the creation and annihilation operators, respectively. Through rotating the matrices into the atomic orbital (AO) basis, the charge transfer numbers

can be determined

$$\Omega_{AB} = \frac{1}{4} \sum_{\alpha \in A} \sum_{\beta \in B} \left[D^{10}_{\beta\alpha} (SD^{01}S)_{\alpha\beta} + (SD^{10})_{\beta\alpha} (SD^{01})_{\alpha\beta} (D^{10}S)_{\beta\alpha} (D^{01}S)_{\alpha\beta} + (SD^{10}S)_{\beta\alpha} D^{01}_{\alpha\beta} \right] \tag{1.44}$$

where β and α index AO fragments of components A and B, and S is the atomic orbital overlap matrix.

Chapter 2

EVALUATING THE ELECTRONIC STRUCTURE OF BIMETALLIC PLATINUM(II) COMPLEXES VIA DENSITY FUNCTIONAL THEORY

Using the methods introduced in the previous chapter, a series of bimetallic platinum complexes are evaluated to understand the architectural influence on the electronic structure. Specifically, the ligand composition is investigated to identify the dominating factors contributing to the experimentally observed platinum-platinum contraction and ultimately the energetic landscape lending to ultrafast dynamics. Reprinted with permission from Mills, A., Valentine, A., Hoang, K., Roy, S., Castellano, F., Chen, L., and Li, X., 2021. General design rules for bimetallic platinum(II) complexes. The Journal of Physical Chemistry A, 125, 43, pp.9438-9449. Copyright 2021, American Chemical Society.

2.1 Introduction

Square-planar transition metal complexes are often studied because of their attractive photophysical and photochemical properties making them excellent candidates for advancing molecular materials.[81, 28, 201] Platinum(II) complexes of this design are of particular interest because of their rich excited-state properties extending from long lifetimes and versatile photoluminescence to easily accessible redox profiles.[40, 35, 24, 106, 34] Notable applications capitalizing on these properties include organic light emitting diodes[69, 194, 118, 200, 25], bio-sensors[97], and catalysis[18].

Binuclear platinum(II) complexes, in particular, have gained traction because of their increased tunability accessed through modifying both the conjugation and bulk-

iness of the attached ligands. The choice of ligand strongly influences whether the complex behaves as a mononuclear system or takes advantage of the binuclear nature of the complex. A large array of platinum(II) complexes have been synthesized with attachments ranging from simple monodentate ligands[81] to large chelating cyclometalating ligands[28, 118, 206, 24] to fine-tune and manipulate the electronic structure and geometric configuration for the aforementioned applications.[193, 80, 79]

The primary use of these platinum dimer complexes has been to induce charge transfer from the metal centers to adjacent electron-accepting ligands. Experimentally, transient absorption spectroscopy is used to observe large structural rearrangements[28] and intersystem crossing (ISC) to long-lived triplet states[85, 86, 19]; future studies will explore the possibility of further transferring the charge-separated electron coherently to additional ligands to drive additional chemical processes. To that end, a series of platinum dimers were synthesized according to a butterfly motif.[69, 71] Bridging ligands connecting the platinum atoms rigidly enforce a given metal-metal separation in both the ground and excited states, with direct implications for the nature and ordering of the electronic states. Shorter Pt-Pt separations destabilize the occupied $5d_{z^2}$-$5d_{z^2}$ σ^* orbital, lowering the energy of metal-metal-to-ligand charge-transfer (MMLCT) states relative to more local ligand-centered (LC) or monometallic metal-to-ligand-charge-transfer (MLCT) states. The former deplete an antibonding orbital between the metal centers, increasing the bond order and inducing large structural changes.[25, 28, 83, 199, 24, 118, 172] Cyclometalating ligands then function either as the location of local π-π^* excitations or as electron-accepting ligands in the case of (M)MLCT, with π-stacking interactions further influencing the shape of the complex. The butterfly design allows for both a fixed metal-metal separation and flexible motion as the complex relaxes in the excited state.

To gauge the effects of various input factors, the suite of molecules studied in this paper systematically varies the constituent ligands. Smaller, pyridyl derivative, bridging ligands enforce a narrower Pt-Pt separation and architecturally encourage

MMLCT transitions, while bulkier fused rings hold the Pt atoms farther apart and lead to more low-lying local LC and MLCT states. Cyclometalating ligands considered in this study are planar aromatic hydrocarbons that form suitable electron acceptors. Compared to the 2-phenylpyridine cyclometalting ligand, 7,8-benzoquinoline ligands have a smaller π-π^* splitting and therefore lower the energy of the lowest-unoccupied molecular orbital. These tunable elements vary the HOMO-LUMO energy gap, the degree of flexibility in the molecule, and perhaps most importantly the relative energetic ordering of the low-lying excited states. As fast ISC is desired, it is important to have triplet states with local (LC) character[186] of lower energy than the bright ^{1}MMLCT state, as ISC is fastest between states of different orbital character.[53]

Recent theoretical work by the authors have shown that square planar bimetallic platinum(II) complexes are capable of supporting long-lived electronic coherence if the interplatinum distance is reasonably small.[106] Coherent electronic behavior is of great fundamental interest because of its influence over excited-state processes prevalent in nature such as efficient light harvesting.[33] It is, however, recognized that the anticipated coherence in molecules of this design quickly dissipates once molecular vibrational motions are considered as coupling to these modes begins to dominate and encourage decoherence.[154] To compliment experimental results suggesting ultrafast photochemical processes, a theoretical study was conducted to support rapid intersystem crossing, which is expected to take place within tens of femtoseconds.[34] Those results demonstrated that in the presence of strong spin-orbit coupling, ultrafast intersystem crossing can be anticipated.[186] These findings prompted further investigation on the complex's architectural influence over the electronic structure and excited-state behavior.

Herein, we investigate a collection of platinum dimers to devise a set of molecular design guidelines to fine-tune the electronic structure and charge transfer character of binuclear platinum(II) complexes. This study focuses on the effects of the steric bulkiness and conjugation of the bridging and cyclometalating ligands, respectively.

Through these modifications greater control over the excited-states can be realized and the long-lived electronic or vibrational coherences can be exploited to drive valuable excited-state processes.

2.2 Methods

To study the geometric features of these platinum dimers, optimized geometries are calculated in GAUSSIAN16 using spin-restricted density functional theory (DFT).[57] Linear response time-dependent (TD) DFT is utilized for all excited-state calculations. The charge transfer (CT) character for each complex is determined quantitatively via linear-response TDDFT in order to assign CT character to select excited states that have the potential to participate in the relaxation process. Charge transfer numbers are calculated using a locally modified developmental version of GAUSSIAN16 following the methods outlined in previous work.[115, 147, 148] Unless otherwise stated, all calculations are completed using the ωB97X-D[26] functional with the LanL2DZ effective core potential[191] and associated double-ζ basis set on the platinum atoms and the 6-31g(d) basis set on all other atoms. Finally, all calculations are completed in the gas phase as environmental effects (e.g. solvent) are expected to have a uniform effect on the systems studied given their structural similarities.

2.3 Results & Discussion

A series of bimetallic platinum(II) complexes were synthesized by Castellano and coworkers with varying cyclometalating (CML) and bridging ligands (BL) to modulate the geometric and excited-state properties.[28] In this report, we discuss the effects of varying ligand compositions and devise design rules that can be derived from these observations. The CMLs for all dimers in this study are either 2-phenylpyridine (ppy) or 7,8-benzoquinoline (bpy). The bridging ligands, on the other hand, vary by both core architecture and substituent composition. Structures **1-9** feature a pyridine derivative as the bridging ligand with either an oxygen or sulfur heteroatom atom

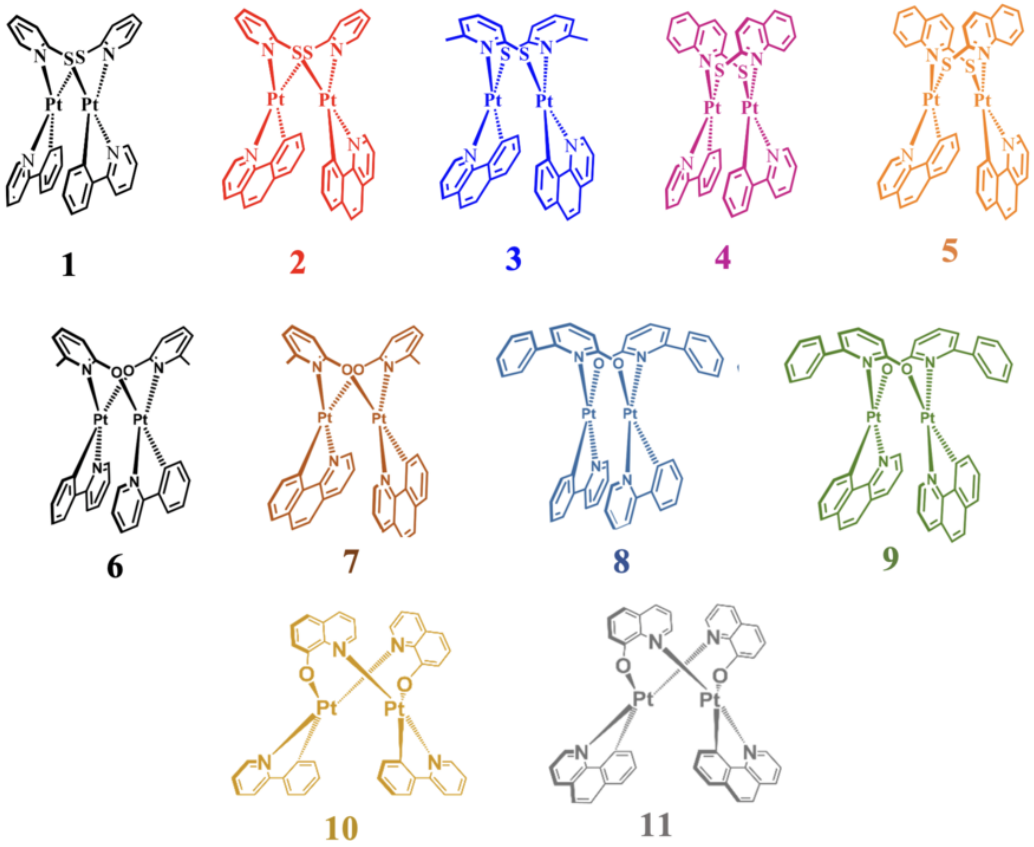


Figure 2.1. Structures of bimetallic platinum(II) dimers

attached alpha to the nitrogen (see **Fig. 2.1**). Additionally, dimers **3** and **6-9** have either a methyl or phenyl substituent in the other alpha position (6-substitution) to increase the bulkiness of the ligand and ideally to modulate the platinum-platinum distance. The influence of this coordinate is further explored in structures **4-5** and **10-11** by exchanging the pyridine group for a quinoline ligand that has a sulfur attached in the second position or an oxygen attached in the eighth position. It should be noted that the nitrogens of the CMLs in the sulfur series (**1-5**) have a *cis* configuration while the nitrogens of the oxygen series (**6-11**) have been arranged in

a *trans* orientation; this is likely a consequence of "trans effects" often observed in square-planar complexes.[36, 4]

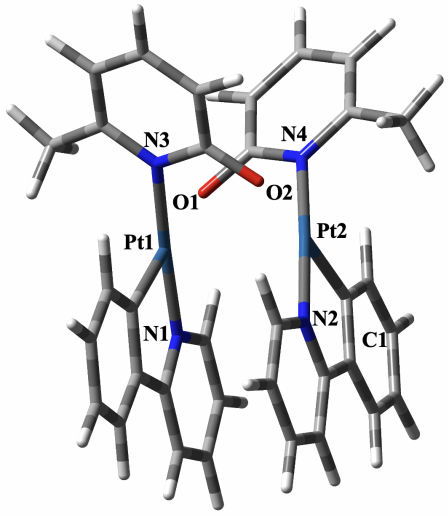


Figure 2.2. General numbering scheme for all dimers in this study. Note that N1 of the sulfur series will be on the opposite ring of the corresponding CML.

To identify the optimum basis set/functional combination, benchmarking is completed on complexes **6-9**. Pairings include B3LYP with 6-31g(d)/LanL2DZ, CAM-B3LYP with 6-31g(d)/LanL2DZ, ωB97X-D with 6-31g(d)/LanL2DZ, and ωB97X-D with def2-TZVP/SDD. Qualitatively, a substantial difference is observed in the orientation of the CM ligands when moving from B3LYP and CAM-B3LYP to ωB97X-D, as seen in **Figure 2.3**. While B3LYP and CAM-B3LYP find an A-frame orientation of the molecule, the dispersion corrections in ωB97X-D induce an increased π-π interaction between the CML ligands, leading a more H-frame configuration that closely matches the experimental crystal structure.[28] Complex **7** is further analyzed in **Table 2.1** to quantitatively demonstrate the increased agreement between the calculated structure and the experimentally reported geometry. In addition to using long-range and dispersion correction, the data presented support the use of a smaller

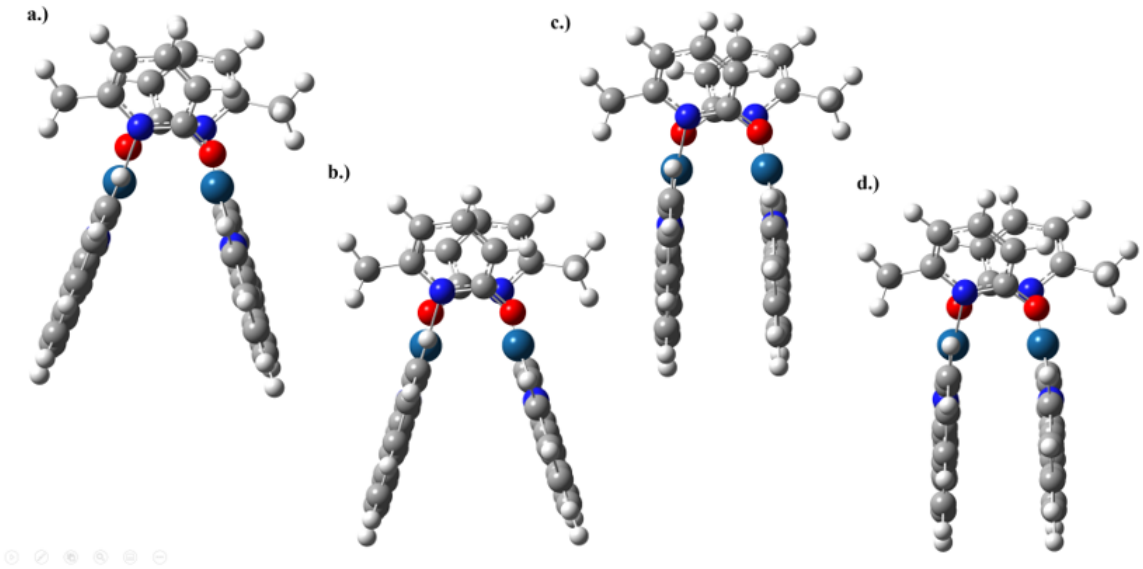


Figure 2.3. Optimized geometries of complex **7** using a.) B3LYP and 6-31g(d)/LanL2DZ, b.) CAM-B3LYP and 6-31g(d)/LanL2DZ, c.) ωB97X-D and 6-31g(d)/LanL2DZ, and d.) ωB97X-D and def2-TZVP/SDD, where the first basis set listed is applied to all light atoms and the second basis set is used exclusively on the Pt atoms.

basis set as the larger basis set shows minute improvement, if any, when comparing the equilibrium geometries calculated at the same level of theory.

Additionally, the static electronic spectra are calculated at the same level of theory used to find the corresponding equilibrium geometries. The resulting spectra show the line shape remains relatively consistent across all basis sets and functionals for the lowest energy transition. Large differences begin to show up in the local character of the second, higher energy peak of the spectra simulated via CAM-B3LYP. Notable differences between these simulated spectra and the experimentally observed spectra are found, specifically, in the line shapes of complexes **7** and **9**, which display multiple peaks in the region just under 350 nm. Interestingly, better agreement between sim-

ulated and experimental spectra is seen when using B3LYP, which finds the correct lower energy peak ordering across the complexes considered. The relative ordering between **7** and **8** appears to have changed once long-range correction is included, which is demonstrated by both CAM-B3LYP and ωB97X-D. However, given the increased agreement between the calculated ground-state geometry using the latter functional and the experimentally observed crystal structure, this study employs ωB97X-D for the remaining calculations. Although not demonstrated here, the authors found, in a previous study, that including range correction is significant as it finds the correct ordering of the triplet states, further supporting the use of ωB97X-D.[186] Finally, a larger basis set is used with ωB97X-D to simulate the static electronic spectra in appendix A. The only notable difference seen when increasing the basis set is a slight red-shift in the transition energies, justifying the use of a smaller basis.

Ground State Structural Characterization

The effects of the various ligands on the geometric characteristics at the ground-state equilibrium geometry are presented in **Tab. 2.2**. At the ground-state geometry, complexes **1-9** appear to have very similar Pt-Pt distances at around ∼2.9 Å. Unsurprisingly, moving the interacting atoms of the BL from the second to the eighth position substantially increases the distance between the platinum atoms. This is demonstrated in complexes **10** and **11** where the distance between the two transition metals was calculated to be roughly 0.3 Å greater than the previous nine complexes.

Basis set	6-31g(d)/LanL2DZ	6-31g(d)/LanL2DZ	6-31g(d)/LanL2DZ	def2-TZVP/SDD	-
Functional	B3LYP	CAM-B3LYP	ωB97X-D	ωB97X-D	Exp.[28]
Pt1-Pt2	3.071	3.023	2.911	2.913	2.816
Pt1-N1	2.057	2.045	2.048	2.042	2.017
Pt1-N3	2.081	2.066	2.068	2.065	2.041
Pt1-Pt2-N2	98.69	97.57	90.49	90.81	91.25
N1-Pt1-Pt2-N2	63.13	62.32	64.21	64.69	61.37

Table 2.1. Selected bond lengths, angles, and dihedrals of **7** at the ground-state geometry in units of angstroms and degrees (see **Figure 2.2** for atom reference).

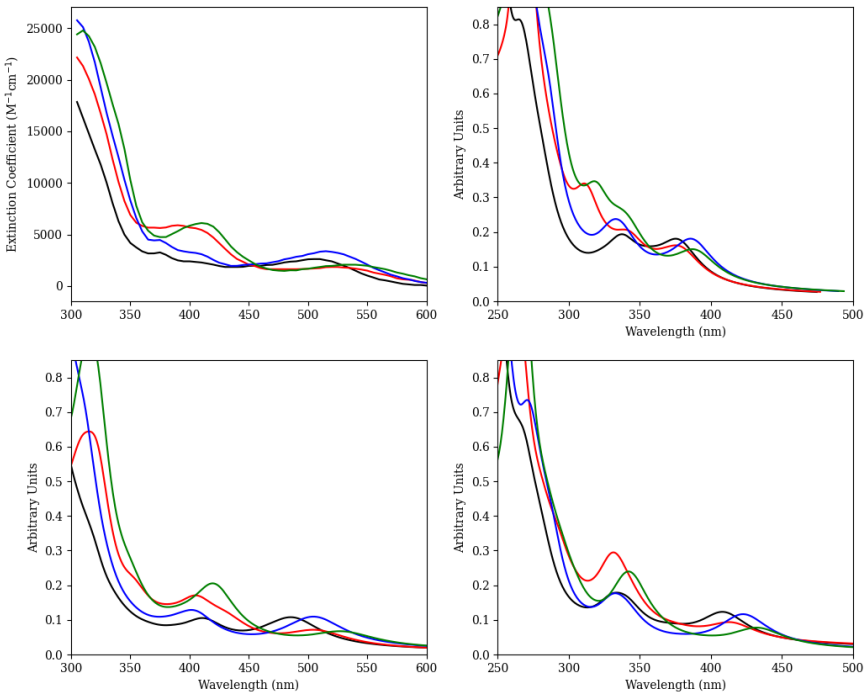


Figure 2.4. Collected absorption spectra for complexes **6** (black), **7** (red), **8** (blue), and **9** (green). The spectra presented were found experimentally[28] (top left) and computationally using the following functional and light/heavy atom basis sets: CAM-B3LYP with 6-31g(d)/LanL2DZ (top right), B3LYP with 6-31g(d)/LanL2DZ (lower left), and ωB97X-D with 6-31g(d)/LanL2DZ (lower right).

This is an immediate reflection of the increased role played by the BL in mediating the position of the transition metals.

As expected, the choice of chalcogen greatly influences the distance between the Pt atom and the BL. By simply increasing the atomic size, the bridging ligand is forced further away from the heart of the structure. The distance between these two BLs, calculated via the center of mass point of pyridyl ring, seems to be unaffected by the chalcogen and instead appears to be largely dependent on the CML archi-

tecture. Generally speaking, as the CML becomes more rigid, and the flexibility of the molecule becomes hindered, the BLs are forced further from one another. This effect can be countered by incorporating a bulky, free-rotating, substituent on the BL. As demonstrated by the sulfur series through the BL-Pt1-BL coordinate, the methyl group introduced in **3** shows some influence over the BL-BL positions by slightly decreasing the angle, compared to **2**, and forcing the BLs closer than they would be otherwise.

The impact of the free-rotating substituent becomes more pronounced in the oxygen series when comparing **7** to **9**. The bulkiness of the phenyl group has a much greater influence over the BL-BL distance than the comparatively smaller methyl group. The effects of the larger hanging substituent are further explored by comparing **8** to **9** where the CML is the only differing feature between the complexes. Based on the previous observations, one might expect **8** to have slightly closer BLs since the complex has not assumed a locked configuration induced by the rigid CML; this is, however, not the case. In fact, the BL-Pt1-BL angle and the distance between the BLs is substantially greater in **8** than **9**. This is a consequence of the phenyl group substituents of the BL, which have the freedom to orient in space in various ways and subsequently influence the overall geometry. The influence of the CML can, however,

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Pt1-Pt2	2.945	2.946	2.950	2.938	2.934	2.911	2.911	2.940	2.912	3.257	3.258
Pt2-N4	2.222	2.217	2.237	2.239	2.231	2.073	2.064	2.071	2.075	2.091	2.087
Pt1-O1/S1	2.345	2.339	2.371	2.343	2.336	2.185	2.177	2.203	2.183	2.175	2.165
BL-BL	4.379	4.494	4.504	4.615	4.840	4.668	4.845	5.147	4.342	3.641	3.624
CML-CML	3.582	3.579	3.646	3.641	3.609	3.751	3.723	3.717	3.748	4.416	4.425
BL-Pt1-BL	64.632	66.668	65.861	67.354	71.168	72.253	75.391	79.973	67.161	66.374	67.362
Pt-Pt-C1	89.180	88.182	91.543	93.582	92.220	103.932	103.245	105.799	101.035	115.660	114.951

Table 2.2. Selected bond lengths, angles, and dihedrals of **1-11** at the ground-state geometry in units of angstroms and degrees (see **Figure 2.2** for atom reference). Calculations including BLs and CMLs were completed by identifying the center of mass points on the respective ligands. For uniformity purposes, only the pyridyl ring of the BL was considered when finding the center of mass point.

be preserved as the BL increases in size if the BL maintains reduced mobility. This is supported by comparing **4** with **5** and **10** with **11**, which both have BLs composed of fused rings.

Static Electronic Spectra

The calculated absorption spectra for all complexes in this study are presented in **Figure A.1** and **Figure 2.5**. The lower energy peaks of the absorption spectra, for complexes **1-9**, are assigned as having metal-metal-to-ligand charge-transfer (MMLCT) character, which is consistent with findings reported in the literature.[28, 37, 94]

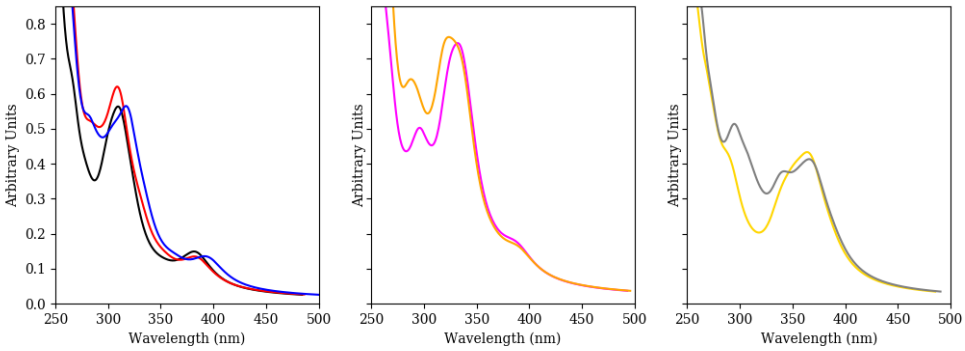


Figure 2.5. Calculated absorption spectra for complexes **1-5** and **10-11**. The far left panel holds **1** (black), **2** (red), and **3** (blue). The center panel shows both **4** (magenta) and **5** (orange) and the far right panel displays both **10** (gold) and **11** (grey).

This MMLCT character is not observed in structures **10-11** of the oxygen series. Instead **10-11** yield lowest excited states that are generally centered on the bridging ligand. The absence of MMLCT character in these structures is anticipated considering the MMLCT state arises when the d_{z^2} orbitals of the two Pt atoms are sufficiently close; thus forming a noticeable gap between both the $d\sigma$ and $d\sigma^*$ orbitals and promoting the latter to the HOMO. The excitation out of the $d\sigma^*$ orbital into the π^* orbital is what characterizes the MMLCT transition and in turn increases the bond order between the Pt atoms upon excitation. Recalling the Pt-Pt distance in

complexes **10** and **11**, it is recognized that their natural positioning prevents them from interacting.

Stabilization of the MMLCT State

The internal heavy metal effect can be considered for these complexes as they hold two platinum atoms at the heart of their structure. The heavy metal effect is simply the increase in spin-orbit coupling between the singlets and triplets thus increasing the probability of intersystem crossing (ISC).[137, 31] Considering this and experimental evidence of ISC, additional electronic transitions to the triplet excited states, at the ground-state equilibrium geometry, are analyzed via charge transfer numbers and natural transition orbitals (NTOs).[28]

To emphasize the influence of the functional on the characterization of the electronic structure and the energetic landscape, the charge transfer (CT) numbers were calculated using different functionals. **Figure 2.6** demonstrates the change in the energetic landscape as the functional, and subsequently the geometry, changes for **6**; the chosen representative complex for this analysis. Although the electronic character of the first singlet excited state of complex **6** is consistently assigned to ^{1}MM(CM)LCT character, with an average charge-transfer ratio of 0.849, the energy of this state is observed to be quite sensitive to the functional. This CT state undergoes considerable destabilization ($\sim$1 eV) when moving from B3LYP to CAM-B3LYP. This energy sensitivity carries over to the triplets as well and is most notable in states exhibiting strong charge transfer character. The change in CT state ordering can be attributed to the inclusion of range-correction, which has been shown to better predict charge-transfer excitation energies.[184] This correction has a greater impact on CT states as contributions from both the exchange functional and Hartree-Fock exchange integral are modified as a function of distance, thus giving rise to a better description of the long-range orbital-orbital interactions.[182] Provided this, states with local transitions are expected to be less affected by this correction, which is generally what is observed in **Figure 2.6** considering the yellow states vary by only about 0.2 eV. The

CT numbers were also calculated for complexes **7-9** and can be found in the appendix B showing similar trends to those observed here.

The effects of both the geometry and functional can be observed by comparing panels **II-IV** of **Figure 2.6**. Like CAM-B3LYP, ωB97X-D is a range-corrected functional, but it also directly includes dispersion corrections. To tease apart the two effects, **III** gives the excited-state spectrum calculated using ωB97X-D at the same molecular geometry as in **II**: the state orderings and characters are nearly identical between the two functionals, with a slight energetic shift. However, in allowing the molecule to relax to the ωB97X-D optimized geometry in **IV**, additional changes are observed. The CT states are observed to stabilize once the CM ligands are brought in close proximity via dispersion-corrections. Additionally, this change in geometry allows states that were previously strongly CT in character to mix with states that had predominately local character; giving rise to "mixed" states that have large contributions from both (CM)LC and MM(CM)LCT electronic character. This is more quantitatively represented in **Table 2.3** where the elements of Ω are as follows:

$$\Omega = \begin{pmatrix} MC & MM(CM)LCT \\ (CM)MCT & (CM)LC \end{pmatrix} \tag{2.1}$$

the *MC* notation indicates metal-centered, *MM(CM)LCT* is metal-metal-to-ligand CT, *(CM)MCT* is ligand-to-metal CT, and *(CM)LC* is ligand centered. All of the elements within this matrix correspond to either the platinum atoms, the CM ligand, or both.

The energetic landscape strongly impacts the dynamic processes that can be observed theoretically, which makes selecting the appropriate functional of utmost importance. Given prior studies have demonstrated that these systems undergo fast ISC, there should be low-lying triplet states with different character from the first singlet state; per El-Sayed's rules.[28, 186, 86, 53] This leaves CAM-B3LYP and ωB97X-D as the remaining viable options, from this analysis, for properly predicting the electronic

structure. Recalling the increased agreement between the equilibrium geometry found via ωB97X-D, calculations going forward will utilize this functional.

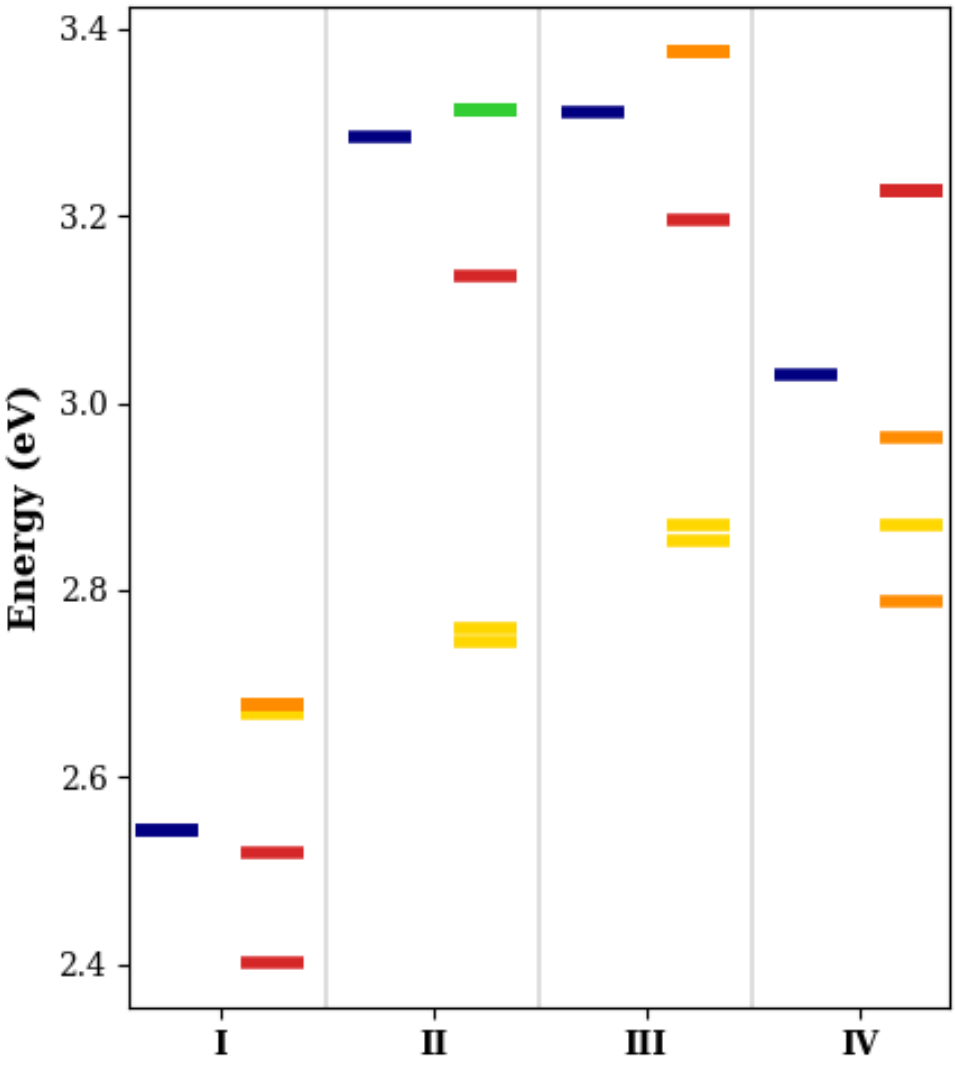


Figure 2.6. Jablonski diagram of complex **6** using I) B3LYP with 6-31g(d)/LanL2DZ, II) CAM-B3LYP with 6-31g(d)/LanL2DZ, III) ω97X-D with 6-31g(d)/LanL2DZ at the CAM-B3LYP optimized geometry, and IV.)ω97X-D with 6-31g(d)/LanL2DZ at the ω97X-D optimized geometry. The states considered in this analysis are: ^{1}MM(CM)LCT (blue), ^{3}MM(CM)LCT (red), ^{3}(CM)LC (yellow), ^{3}Mixed (orange), and ^{3}(BD)LC (green).

	State	Energy (eV)	Ω	Charge-Transfer Ratio	Designation
I.	1	2.402	$\begin{pmatrix} 0.05 & 0.68 \\ 0.01 & 0.14 \end{pmatrix}$	0.804	^{3}MM(CM)LCT
	2	2.519	$\begin{pmatrix} 0.03 & 0.62 \\ 0.01 & 0.26 \end{pmatrix}$	0.713	^{3}MM(CM)LCT
	3	2.543	$\begin{pmatrix} 0.04 & 0.78 \\ 0.01 & 0.06 \end{pmatrix}$	0.906	^{1}MM(CM)LCT
	5	2.669	$\begin{pmatrix} 0.03 & 0.34 \\ 0.03 & 0.53 \end{pmatrix}$	0.438	^{3}Mixed
	6	2.677	$\begin{pmatrix} 0.03 & 0.43 \\ 0.02 & 0.43 \end{pmatrix}$	0.540	^{3}Mixed
II.	1	2.745	$\begin{pmatrix} 0.02 & 0.12 \\ 0.03 & 0.82 \end{pmatrix}$	0.162	^{3}(CM)LC
	2	2.759	$\begin{pmatrix} 0.01 & 0.10 \\ 0.03 & 0.85 \end{pmatrix}$	0.138	^{3}(CM)LC
	3	3.137	$\begin{pmatrix} 0.11 & 0.61 \\ 0.01 & 0.17 \end{pmatrix}$	0.719	^{3}MM(CM)LCT
	4	3.285	$\begin{pmatrix} 0.06 & 0.70 \\ 0.01 & 0.12 \end{pmatrix}$	0.816	^{1}MM(CM)LCT
	5	3.314	$\begin{pmatrix} 0.01 & 0.01 \\ 0.00 & 0.00 \end{pmatrix}$	0.067	^{3}(B)LC
III.	1	2.852	$\begin{pmatrix} 0.02 & 0.16 \\ 0.03 & 0.77 \end{pmatrix}$	0.210	^{3}(CM)LC
	2	2.869	$\begin{pmatrix} 0.01 & 0.13 \\ 0.03 & 0.81 \end{pmatrix}$	0.173	^{3}(CM)LC
	3	3.196	$\begin{pmatrix} 0.10 & 0.61 \\ 0.01 & 0.18 \end{pmatrix}$	0.718	^{3}MM(CM)LCT
	4	3.312	$\begin{pmatrix} 0.06 & 0.71 \\ 0.01 & 0.12 \end{pmatrix}$	0.814	^{1}MM(CM)LCT
	5	3.375	$\begin{pmatrix} 0.25 & 0.17 \\ 0.01 & 0.01 \end{pmatrix}$	0.400	^{3}Mixed*
IV.	1	2.788	$\begin{pmatrix} 0.05 & 0.38 \\ 0.02 & 0.48 \end{pmatrix}$	0.469	^{3}Mixed
	2	2.869	$\begin{pmatrix} 0.02 & 0.14 \\ 0.03 & 0.79 \end{pmatrix}$	0.192	^{3}(CM)LC
	3	2.963	$\begin{pmatrix} 0.06 & 0.41 \\ 0.02 & 0.46 \end{pmatrix}$	0.485	^{3}Mixed
	4	3.03	$\begin{pmatrix} 0.06 & 0.75 \\ 0.01 & 0.07 \end{pmatrix}$	0.861	^{1}MM(CM)LCT
	5	3.228	$\begin{pmatrix} 0.04 & 0.74 \\ 0.01 & 0.14 \end{pmatrix}$	0.818	^{3}MM(CM)LCT

Table 2.3. CT numbers of **6** corresponding to **Figure 2.6** where group I) uses B3LYP with 6-31g(d)/LanL2DZ, II) uses CAM-B3LYP with 6-31g(d)/LanL2DZ, III) uses ω97X-D with 6-31g(d)/LanL2DZ at the CAM-B3LYP optimized geometry, and IV.) uses ω97X-D with 6-31g(d)/LanL2DZ at the ω97X-D optimized geometry. * Mixed state with ^{3}MC and ^{3}MM(CM)LCT contributions

Complex	MM(CM)LCT	MM(B)LCT	(CM)LC	(B)LC	Misc.
3	0.7155	0.0205	0.0956	0.0063	0.1623
4	0.6517	0.0274	0.0875	0.0263	0.2072
6	0.7499	0.0085	0.1075	0.0019	0.1322
7	0.7719	0.0056	0.1167	0.0012	0.1046
8	0.7467	0.0083	0.0974	0.0020	0.1455
10	0.0038	0.0874	0.0018	0.8539	0.0531

Table 2.4. CT numbers for select complexes at the ground-state geometry for the first singlet excited state where MMLCT is the metal-metal-to-ligand CT to the specified ligand in parentheses and LC is ligand centered about the ligand denoted in parentheses.

Table 2.4 outlines the charge transfer numbers for the first singlet excited state of select complexes extending from sulfur derivatives to the larger oxygen derivatives. These calculations show the S_1 state having ^{1}MMLCT character for complexes with a Pt-Pt distance < 3 Å and ^{1}bridging-ligand centered (BLC) character for systems with a Pt-Pt distance > 3 Å. This assignment is determined considering the contribution of charge transfer from the metal to the *cis*-CML and the metal to the *trans*-CML. If a noticeable contribution is observed to come from the M(*trans*-CM)LCT, in addition to the M(*cis*-CM)LCT, then the metal atoms are considered close enough to facilitate MMLCT. This assumption implies the leaving orbital, $d\sigma*$, is shared among the platinum atoms, which is visualized in **Figure 2.7**

The dramatic difference in CT character of both **10** and **11** compared to all other complexes in this study has largely to do with the bridging ligand composition, which inhibits many of the sought-after features in binuclear transition metal systems. This ligand not only forces the platinum atoms to maintain a great distance from one another, it increases the conjugation of the overall complex allowing greater dispersion of electron density about the BL. The increased delocalization in turn lowers the energy of the accepting orbitals making transitions more energetically competitive.

Selected NTOs are displayed in **Figure 2.7** to both support the CT assignment

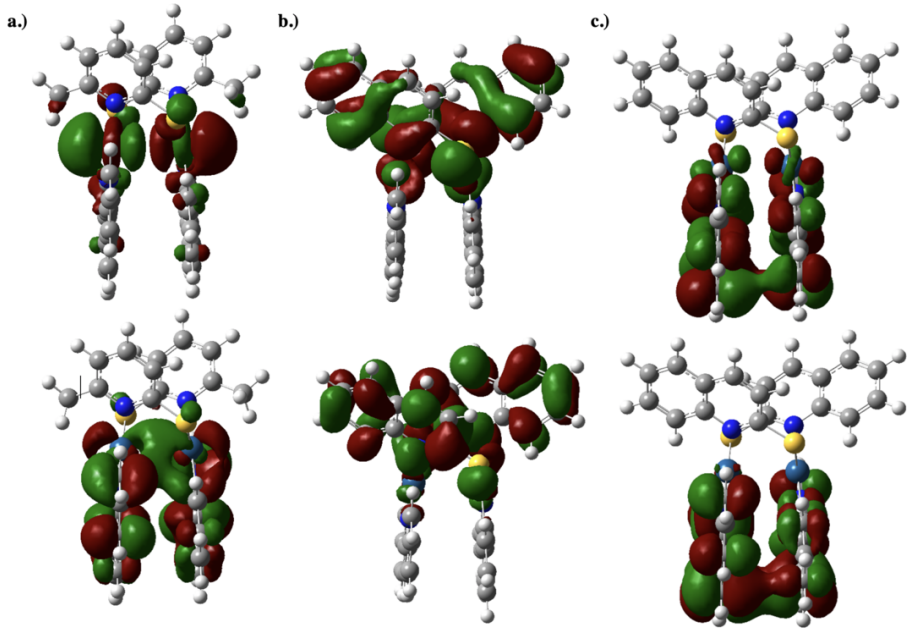


Figure 2.7. Representative NTOs of excited states are shown here where a.) is ^{1}MMLCT of **3**, b.) is ^{3}BLC of **4**, and c.) is ^{3}CMLC of **5** at the ground-state geometry. The hole orbital is found in the top row and the particle orbital can be found on the bottom row.

and visually demonstrate the electronic character of various excited states. As mentioned previously, complexes **1-9** support ^{1}MMLCT where electrons are promoted out of the $d\sigma^{*}$ orbital onto the ligand π^{*} orbital (**Figure 2.7a**). Conversely, **10-11** experience bridging-ligand-centered electronic character as a consequence of their inherent architecture. Similar character is seen in the low-lying triplet of **4**, which is shown in **Figure 2.7b**. Surprisingly, the ^{3}BLC appears in **5** at higher energy triplets as opposed to the low-lying triplet. The low-lying triplet of this complex is consistent with complexes **1-9**, which all have cyclometalating ligand centered (CMLC) character. Perhaps this is a result of the differing conjugation between the BL and CMLs.

Bringing attention to **Figure 2.8**, it seems the increased conjugation of the CML plays a significant role in stabilizing the triplet states while increasing the bulkiness of

the BL has a minor stabilizing effect on the singlet state, in most cases. This can be qualitatively analyzed by considering the differences between the singlet and triplet states in the Jablonski diagram.

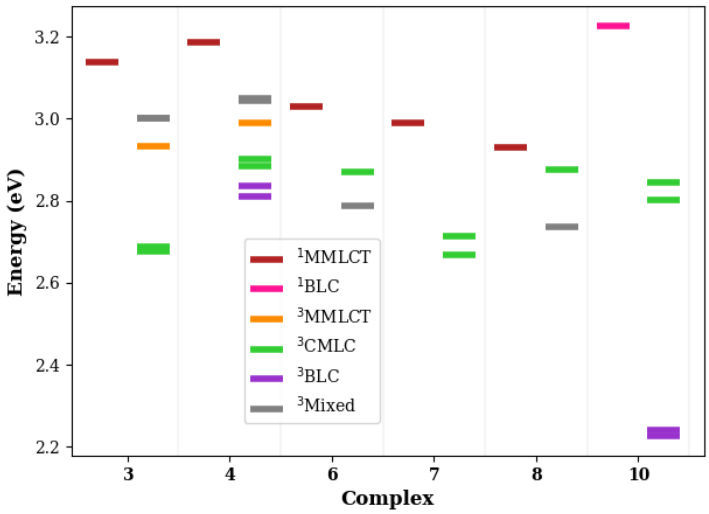


Figure 2.8. Jablonski diagram for complexes **3-4**, **6-8**, and **10**.

The most direct comparison of complexes across the oxygen and sulfur series is between **3** and **7**, which differ only by the choice of chalcogen. When comparing these two complexes it becomes obvious that introducing a larger atom in the second position of the BL causes increased electronic destabilization, which is reflected in the higher energy of the ^{1}MMLCT state of **3**. This destabilization is further emphasized in both **4-5** and **10-11** by incorporating sizable fused rings in close proximity. Though the BL composition clearly plays a role in the ^{1}MMLCT state, the positioning of the CMLs does show some influence over the ^{1}MMLCT energy. Considering only complexes with a Pt-Pt bond length < 3.0 Å, there appears to be a linear relationship between the ^{1}MMLCT energy and the distance between the center of mass point of the CMLs. An increase in distance of $\sim$0.025 Å can stabilize the ^{1}MMLCT state by

nearly 0.1 eV. The Pearson correlation coefficient was found to be −0.92 indicating a strong negative correlation between this particular geometric feature and energy (appendix C). This suggests that through controlling the position of the CML the energy of the ^{1}MMLCT can be fine-tuned to create a more accessible state. Other explored geometric features do not show a significant correlation with the energy. Complexes **10-11** are neglected in this calculation because of the significant architectural difference in the BL.

Additionally, it is recognized that the energy of the ^{3}CMLC state is minimally changed when comparing **3** and **7**, indicating the BL has little to no influence over the triplet states. Similarly, looking explicitly at the oxygen series, the relative positioning of the ^{3}CMLC state for **6** and **8** are reasonably consistent with one another while **7** shows noticeably different behavior. Complex **7** shows decent stabilization of the ^{3}CMLC states compared to **6** as a result of well separated conjugated CMLs. As noted, the triplet states of **8** show no such stabilization with the addition of the phenyl group on the bridging ligand compared to **6**. Complex **8** does, however, show a slight stabilization of the ^{1}MMLCT state as a result of the modified bridging ligand. Considering these observations, it comes as no surprise that both ^{1}MMLCT and ^{3}CMLC states are stabilized by the bulkier BL and conjugated CMLs of **9** (appendix D). The stabilization of the triplet states, as the CML increases in conjugation, is observed for all complex pairs in this study as a consequence of increased electron delocalization.[68]

Charge Transfer Modulation

The extent to which the Pt-Pt distance influences the charge transfer character is investigated by manually modulating the positions of the transition metals in complexes **10** and **11**. A series of geometries are interpolated between the ground-state equilibrium and an optimized geometry that has a fixed Pt-Pt distance of ∼2.6 Å. Single-point energy calculations are then completed to evaluate the energy of the first singlet excited state at each geometry to build potential energy surfaces (PESs) for

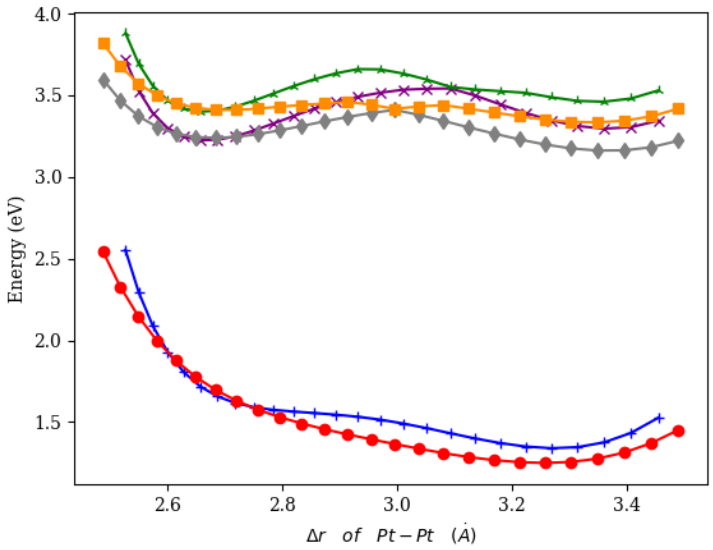


Figure 2.9. Potential energy surfaces of complexes **10** and **11** where the ground-states are blue and red, respectively. The first two singlet excited-states for **10** are represented by the purple (S_1) and green (S_2) surfaces while the grey (S_1) and orange (S_2) are **11** surfaces. It should be noted the ground-state surfaces have been shifted up by 1.25 eV for plotting purposes.

further analysis (**Figure 2.9**).

In decreasing the distance between the platinum atoms, there is a noticeable rise and subsequent fall in S_1 energy as the complex proceeds toward a shorter bond length. This rise in the energy can be attributed to the geometric rearrangement forcing the complex into a less than optimal configuration. In directing the complex toward a geometry with a decreased distance between the Pt atoms, metal-metal interactions can be accessed allowing the system to experience ^{1}MMLCT, which is expected considering the results presented earlier. Through this experiment, the exact distance required to facilitate ^{1}MMLCT can be identified via NTO analysis.

To determine the minimum distance needed to encourage a shift in electronic character, geometries about the maximum point of the S_1 surface are sampled and

their NTOs are collected. A noticeable change in the CT character is observed in increments of just 0.008 Å, as seen in **Figure 2.10**. The LC character about the bridging ligand is maintained until the Pt-Pt distance reached 2.996 Å, at which point the CT transitions to mixed character. Among the mixed character observed at this geometry, some contribution is found to come from ^{1}MMLCT; though this does not become the dominate contributing character until a further 0.008 Å decrease is experienced. This same trend is seen in complex **10**, however, the mixed character is observed at a slightly longer Pt-Pt distance (3.102 Å). Similarly, by decreasing the distance only 0.008 Å, the dominating charge transfer character is observed to be ^{1}MMLCT for the first singlet excited state.

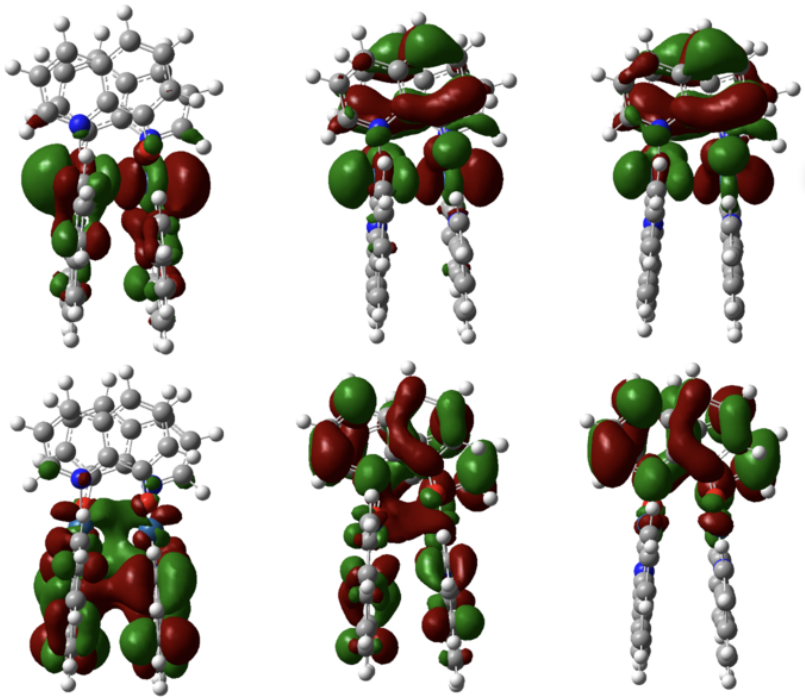


Figure 2.10. NTOs of complex **11** with a Pt-Pt bond length (from right to left) of 2.987, 2.996, and 3.004 Å and electronic character MMLCT, mixed, and LC, respectively. The hole orbital is in the top row and the particle orbital is in the bottom row.

Alternatively, the Pt-Pt bond length of complex **7** is increased to study the necessary bond length required to stabilize the ^{1}MMLCT state and destabilize the LC state in systems with a significantly smaller bridging ligand. Comparing the relevant surfaces, there is a substantially steeper increase in energy as the system is forced out of the equilibrium geometry and finds a metal separation of > 3 Å (see appendix E). Though this increase is expected, it is noticeably more dramatic than the change in energy observed for complexes **10** and **11**. This is likely due to the increased strain forced on the smaller bridging ligand in order to obtain such separation. As before, NTOs for geometries about the "hill-like" region are sampled to identify the transition between ^{1}MMLCT and the new electronic character. Complex **7** is observed to host ^{1}LC character about the CML, which is unsurprising given the CML architecture is well suited for greater electronic delocalization compared to the smaller BL.

Excited State Geometries

The first singlet excited-state equilibrium geometries are calculated to build a better understanding of the structural characteristics post-photoexcitation and beyond the Franck-Condon point. The most notable and consistent geometric changes, across all isomers in this study, are those seen in both the distance between the Pt atoms and the shifting of the CMLs relative to one another. **Figure 2.11** compares the geometries of both **3** and **7** to analyze the impact of modifying the chalcogen of the BL. Interestingly, this did not appear to have a substantial influence over the Pt-Pt contraction from the S_0 to the S_1 geometry. In fact, complexes **1-9** all experience a Pt-Pt contraction of just over 0.2 Å. Systems **10** and **11**, on the other hand, undergo a slight increase in distance; though this is perhaps unsurprising considering the bridging ligand composition.

The change in bridging ligand position between the S_0 and S_1 geometries is, however, inconsistent across the complexes. As before, the position of the BLs were evaluated by determining the center of mass of the pydridyl rings then calculating the distance between the two points. It is observed that the BLs in the complexes

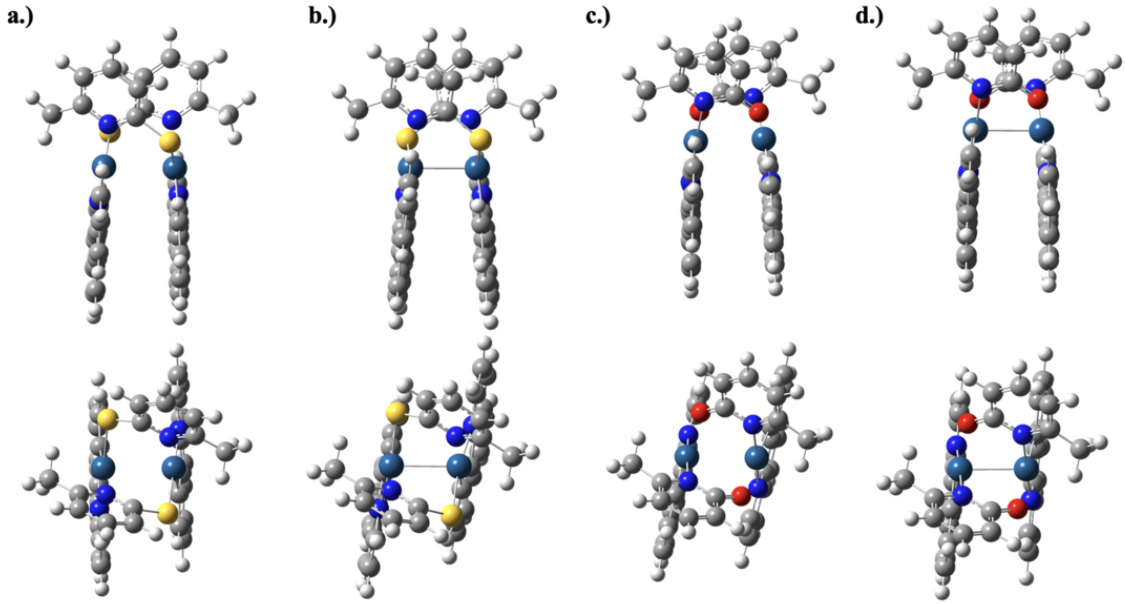


Figure 2.11. Optimized geometries for the S_0 and S_1 states of both **3** (a and b, respectively) and **7** (c and d, respectively). The platinum, sulfur/oxygen, nitrogen, carbon, and hydrogen atoms are represented by the cyan, yellow/red, blue, grey, and white spheres.

belonging to the sulfur series undergo a noticeable shift towards one another from S_0 to S_1. Complexes **10-11**, of the oxygen series, also experience a shift in BLs toward one another with a comparable magnitude of ~0.4-0.5 Å as seen in **4-5**. This rearrangement is allowed as a result of the CMLs distancing themselves from each other through a twisting, divergent-type motion. Conversely, complexes **6-9** generally see a slight repulsion between the bridging ligands when going from the ground-state geometry to the first singlet excited-state geometry. As a consequence, systems **6-9** experience the opposite distortion in their CMLs and instead see them shifting toward one another.

Perhaps the most interesting geometric change seen in these structures is that of the angle formed by atoms Pt1-Pt2-C1. The difference in this angle for the sulfur series was far more dramatic than that observed for the oxygen series. All sulfur

complexes undergo a change of $\sim$ 10 degrees while the oxygen complexes have a modest change of $\sim$ 4.5 degrees. This large difference may be attributed to the difference in size between the oxygen and sulfur atoms forcing a greater shift in the CMLs of the sulfur complexes to accommodate the structural limitations.

2.4 Conclusions

Although DFT does an exceptional job at simulating experimental results, it should be emphasized that (TD)DFT is a non-relativistic, single reference method. Although relativistic effects, namely spin-orbit coupling, are expected to be strong in these bimetallic systems, they are largely neglected in this study. Instead, scalar relativistic effects are indirectly considered via effective core potentials. Remaining within the non-relativistic scheme allows a more direct analysis of the electronic structure by representing states as spin-pure; thus excluding state splitting and additional state mixing. A more rigorous treatment of relativistic effects is not expected to have a significant impact on the calculated equilibrium geometry or simulated spectra presented in this study. Though the density of states will increase, given additional transitions would become allowed under the relativistic regime, explicit inclusion of relativistic effects is not expected to substantially improve the intense spectral signals.[38] Additionally, since exchange-correlation kernels are fundamentally frequency-independent, the calculations presented in this study are confined to the adiabatic approximation. Finally, some evidence of double excitations, in the form of satellite peaks, can be seen in the experimental spectra, the agreement between the observed and simulated spectra suggests single excitations are dominant in these systems. Despite the limitations noted, DFT is able to adequately describe these systems and provide results with excellent agreement to experimental findings.

In conclusion, quantum mechanical calculations on the ground and excited-state structures have been carried out for a series of square-planar platinum(II) dimers with a diverse selection of bridging and cyclometalating ligands. These calculations aim at

giving insight to the effects of bridging ligand bulkiness and cyclometalating ligand conjugation. Through this study, it was further confirmed the bridging ligand architecture and composition is largely responsible for modulating the platinum-platinum distance. Additionally, it was observed that the position of the bridging ligands relative to one an other were somewhat influenced by the rigidity of the CML.

As anticipated, the geometric changes were accompanied by electronic structure differences. The π-conjugation within the CML was found to have a noticeable stabilizing effect on the triplet states of the bimetallic complexes while the bridging ligand composition appeared to have a greater, albeit less dramatic, influence over the energy of the singlet state. Though the composition of the CMLs also seemed to slightly stabilize the singlet states, the lowering of energy in the singlet states became more pronounced when the CML positions changed. This stabilization can be attributed to both the increase of delocalization, as a result of increased conjugated structures, and the system reorienting in space to accommodate structural features. These results reveal individual effects of ligand replacements as well as their mutual influence in both electronic and nuclear structures in both the ground and excited states, which serves as guidance in future molecular design and chemical tuning of transition metal complexes for various applications.

Acknowledgments

This work was supported by the Ultrafast Initiative of the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, through Argonne National Laboratory under Contract No. DE-AC02-06CH11357. The development of time-dependent electronic structure methods is supported by the National Science Foundation (CHE-1856210 to X.L.). Computations were facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system at the University of Washington, funded by the Student Technology Fee. L.X.C. are grateful for the support from the National Science Foundation

(CHE-1955806). F.N.C. was supported by the National Science Foundation (CHE-1955795).

Chapter 3

MODELING THE ULTRAFAST INTERSYSTEM CROSSING DYNAMICS IN PLATINUM DIMER COMPLEXES

The previous chapter proposed a compelling energetic landscape suitable for fascinating intersystem crossing dynamics, which have been evidenced experimentally on an ultrafast timescale.[87] This chapter introduces a method in which the origins of ultrafast intersystem crossing can be investigated via variational relativistic means. The work presented has been adapted with permission from Valentine, A., Radler, J., Mills, A., Kim, P., Castellano, F., Chen, L., and Li, X., 2019. Resolving the ultrafast intersystem crossing in a bimetallic platinum complex. The Journal of Chemical Physics, 151, 114303. Copyright 2019, American Institue of Physics.

3.1 Introduction

Platinum (II) complexes have been well-studied due to their luminescent properties. As a result of their strong spin-orbit coupling, such complexes phosphoresce efficiently and are good candidates for organic light-emitting diodes (OLEDs).[194, 75] Platinum dimer complexes such as tetrakis(pyrophosphito)diplatinate(II) ($[Pt_2(pop)_4]^{4-}$), being smaller than large supramolecular systems and thus easier to control, nevertheless exhibit many of the same photophysical characteristics and have been extensively studied experimentally [188, 187, 156] and theoretically.[130, 95] A series of binuclear platinum complexes with different bridging and cyclometalating ligands have been synthesized, forming a full library of bichromophoric metal complexes with subtly different electronic structure.[118, 117, 29] This new generation of butterfly-shaped

platinum dimers, such as the [Pt(ppy)(μ-^{t}Bu$_2$pz)]$_2$ complex (ppy=phenylpyridine, pz=pyrazolate) shown in Fig. 3.1, experimentally evinces rapid intersystem crossing (ISC) and long-lasting vibrational coherence after photoexcitation.[114, 34, 19, 85]

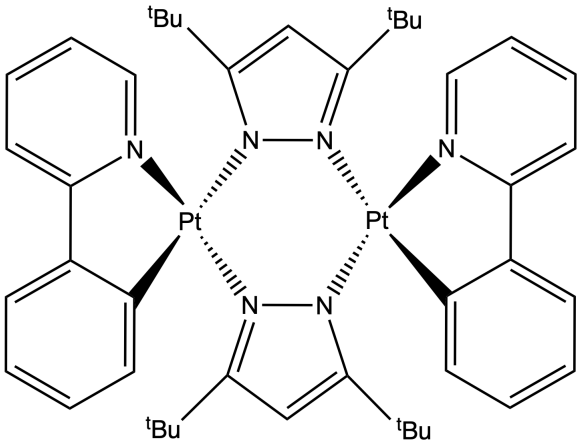


Figure 3.1. Molecular structure for the ^{t}Bu-substituted pyrazolate-cyclometalated diplatinum complex.

Recent theoretical work from the authors[106] demonstrated that the long-lived (∼20 fs) electronic coherence in the singlet metal-metal to ligand charge transfer (^{1}MMLCT) states of binuclear platinum complexes can be achieved through structural modifications of the bridging pyrazolate ligands that sterically enforce certain inter-platinum distances. However, pure electronic S_1-S_2 coherence quickly collapses to the S_1 electronic state within 20 fs due to the coupling with molecular vibrations.[154] The next stage of the ultrafast photochemical process in binuclear platinum complexes is the intersystem crossing to a triplet manifold, experimentally reported to take place before 150 fs after photoexcitation.[34] In this paper, we continue to investigate the cascading excited-state pathway in photoexcited [Pt(ppy)(μ-^{t}Bu$_2$pz)]$_2$ by studying the energetic landscape that supports the intersystem crossing dynamics from S_1 to lower-lying triplet states.

3.2 Methods

3.2.1 Theory

Intersystem crossing, the transition between states of different spin multiplicities that is formally spin-forbidden within non-relativistic quantum mechanics, plays an important role in photochemistry. This change of spin state is only possible through spin-orbit coupling, a fundamentally relativistic phenomenon.[120, 138] We have recently developed a non-perturbative, variational approach based on relativistic theory to compute the spin-orbit coupling strength between spin-pure states.[185] We only present a brief review herein and refer readers to Ref. [185] for theoretical details.

To observe a change in population between different spin states, a basis of states must first be selected, and then couplings between those states must be evaluated. As states calculated variationally in the presence of spin-orbit coupling will no longer belong to a pure spin state, but rather a superposition of several spin states, it is natural to perform calculations in a spin-pure basis where singlets, triplets, *etc.*, can be easily identified. As such, we define spin-pure states to be our spin-diabatic basis, and spin-orbit-coupled states to be our spin-adiabats. In order to generate these, we apply a spin separation technique so that the Dirac Hamiltonian can be written as,[49, 155]

$$\mathbf{H} = \begin{pmatrix} V & \hat{T} \\ \hat{T} & \frac{\mathbf{p}V\cdot\mathbf{p}}{4m^2c^2} - \hat{T} \end{pmatrix} + \begin{pmatrix} \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \frac{i\boldsymbol{\sigma}\cdot\mathbf{p}V\times\mathbf{p}}{4m^2c^2} \end{pmatrix} \tag{3.1}$$

where V is the scalar potential, $\hat{T}$ the kinetic energy operator, c the speed of light, and m the electron mass. Spin and orbital angular momenta are coupled through the $\sigma \cdot \mathbf{p}V \times \mathbf{p}$ term: the vector σ contains the Pauli spin matrices, and $\mathbf{p}$ is the linear momentum operator. The first term in Eq. (3.1) is the spin-pure portion of the Dirac Hamiltonian, which contains scalar relativistic effects, while the second term gives

rise to spin-couplings.

Variationally solving the Dirac equation without the spin-orbit-coupling term will produce spin-diabatic states, *e.g.,* singlets, triplets, *etc.*, denoted as $\{\widetilde{\psi}_K\}$. In contrast, when spin-couplings are included in the Dirac Hamiltonian, the resulting eigenstates, $\{\psi_L\}$, correspond to spin-adiabatic states. In order to calculate the state-to-state couplings between states of different multiplicities, we search for a unitary transformation matrix, $\mathbf{T}$, relating $\{\psi_L\}$ and $\{\widetilde{\psi}_K\}$ (see Ref. [185] for details). Finally, taking the diagonal spin-adiabatic Hamiltonian $\mathbf{H}$

$$\mathbf{H} = \begin{bmatrix} E_1 & & \\ & \ddots & \\ & & E_M \end{bmatrix} \tag{3.2}$$

and rotating it into the spin-diabatic basis

$$\widetilde{\mathbf{H}} = \mathbf{T}\mathbf{H}\mathbf{T}^\dagger = \begin{bmatrix} \widetilde{E}_1 & V_{12} & \cdots \\ V_{12}^* & \widetilde{E}_2 & \\ \vdots & & \ddots \end{bmatrix} \tag{3.3}$$

yields a diabatic Hamiltonian in a spin-pure basis with off-diagonal spin-orbit couplings that drive the intersystem crossing event. For the purpose of this work, nonadiabatic couplings between triplet states are effectively neglected, as we are principally interested in the transition from singlet to triplet. Once the spin-orbit couplings can be evaluated, what remains is to sample the nuclear configurations that result during the photoexcitation process; we elect to perform a single Born-Oppenheimer molecular dynamics (BOMD) trajectory on the first singlet excited state, which will produce representative nuclear geometries for the early times of the photoexcited dynamics.

In this work, we employ the exact-two-component (X2C)[93, 113, 139, 78, 111, 112, 160, 102, 140, 51, 62, 91, 52] transformed relativistic time-dependent density-

functional theory (TDDFT)[52, 84, 145] method to compute spin-diabatic and spin-adiabatic states, using the one-electron Dirac Hamiltonian without and with the spin-orbit coupling term. To account partially for two-electron relativistic effects, the spin-orbit operator within X2C is scaled by a semiempirical fudge factor,[15] reminiscent of the use of effective nuclear charges to account for nuclear screening of valence electrons by core electrons.

3.3 Computational Details

All calculations were performed using a locally modified copy of the development version of GAUSSIAN16.[57] Geometries for the platinum dimer were optimized[98] using the CAM-B3LYP functional,[202] the 6-31g(d) basis on light atoms, and the LanL2DZ effective core potential and double-ζ basis set on the Pt atoms.[73] The range-separated functional was employed in order to better describe the excited states of this species, many of which are charge transfer in character; a comparison of different functionals is presented in appendix F.[5] Born-Oppenheimer molecular dynamics of the first singlet excited state were performed at the same level of theory, beginning at the ground-state geometry with zero initial momentum, evaluated at 0.5 fs time steps over 200 fs. Relativistic X2C-TDDFT calculations were performed with the same functional and basis on light atoms, with the relativistic Sapporo double-ζ basis applied to Pt.[129] All non-relativistic density functional theory calculations were performed on a pruned (75, 302) grid (Grid=UltraFine in GAUSSIAN16), while the convergence of relativistic calculations was improved by switching to a pruned (175, 974) grid for light atoms and an unpruned (250, 974) grid for Pt (Grid=SuperFine).

3.4 Results & Discussion

Before discussing the dynamics, much information can be gleaned from static calculations at the ground- and excited-state geometries of the $[\mathrm{Pt(ppy)}(\mu\text{-}^{t}\mathrm{Bu_2pz})]_2$ dimer complex, shown in Fig. 3.2 and given in full in appendix F. Relative to the

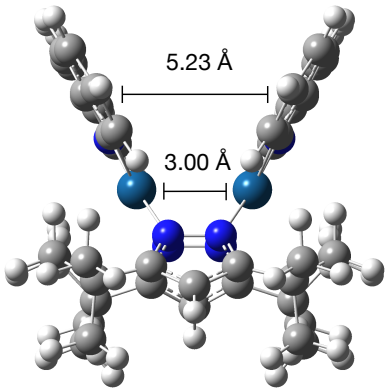


Figure 3.2.1. S_0

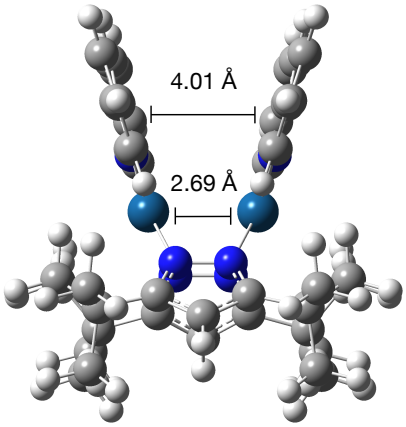


Figure 3.2.2. S_1

Figure 3.2. Optimized geometries for the ground (a) and first singlet excited (b) states of the platinum dimer. The dimer contracts upon photoexcitation.

ground-state, the Pt-Pt bond distance in S_1 contracts from 3.0 Å to 2.69 Å, and the phenylpyridine ligands fold toward one another, with their separation decreasing from 5.23 Å to 4.01 Å. This behavior can be explained by examining the natural transition orbitals (NTOs)[121] that correspond to an electronic transition from the ground-state to the state of interest, plotted in Fig. 3.3. The S_1 state, which is a metal-metal-to-ligand charge-transfer (MMLCT) state, promotes an electron from the antibonding d_{z^2} orbital between the two Pt atoms to a bonding combination of π^* orbitals on the ppy ligands. This increases the bond order both between the Pt atoms and between the ppy ligands, resulting in shorter separations between each. At the ground-state geometry, there are three triplet states below S_1 in energy. T_1 and T_2 are ligand-centered (LC) states, while T_3 is also MMLCT. Though not shown, at the S_1 geometry, S_1 remains MMLCT and T_2 remains LC, while T_1 and T_3 switch in character. For a description of how excited-state characters are assigned,[116, 147, 148]

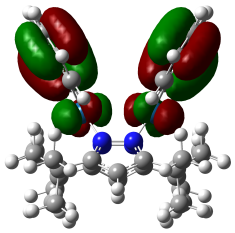
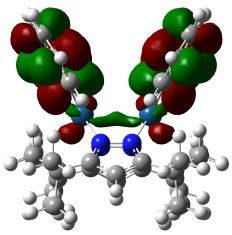


Figure 3.3.1. T_1



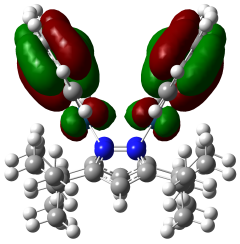
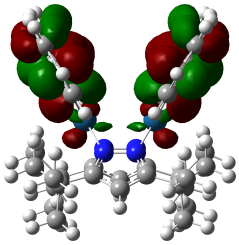


Figure 3.3.2. T_2



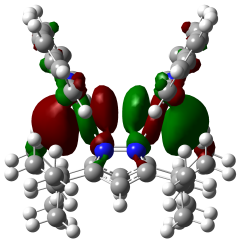
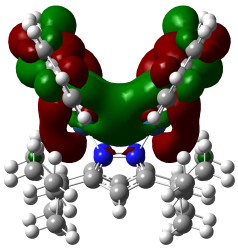


Figure 3.3.3. T_3



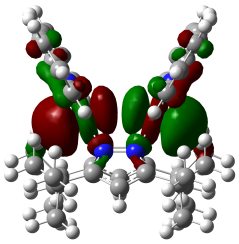


Figure 3.3.4. S_1

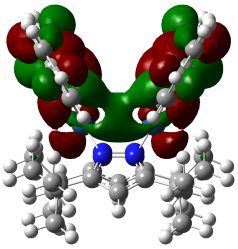


Figure 3.3. (a)-(d): Natural transition orbitals (NTOs) of the first four excited states at the ground-state geometry, corresponding to a vertical excitation. For each state, the hole orbital is on the left and the particle orbital is on the right. T_1 and T_2 are LC, while T_3 and S_1 are MMLCT.

	T_1	T_2	T_3	T_4
V_{ST} (meV)	28	84	40	31
Γ_{ST}^{-1} (fs)	134	15	66	110

Table 3.1. Spin-orbit coupling matrix elements V_{ST} of the four lowest-energy triplets to S_1 at the ground-state geometry, and the corresponding intersystem crossing rates Γ_{ST}^{-1} estimated from Fermi's Golden Rule. Intersystem crossing is fastest to T_2, with non-negligible transition probabilities to the other triplets as well.

see appendix F.

El-Sayed's rules [53] maintain that ISC is fastest between singlets and triplets that differ in electronic character. T_2 is the only triplet that is LC at both geometries, while S_1 is MMLCT. In addition, T_2 is lower in energy than S_1 at the ground-state geometry, while it is higher in energy at the S_1 minimum, implying there is a crossing between the two states somewhere between the two geometries. Those two facts allow us to postulate that T_2 is the most likely candidate for ultrafast ISC. However, using the procedure outlined above, it is possible to be more quantitative. The spin-orbit couplings at the ground-state geometry between the four lowest-energy triplets and S_1 are shown in Tab. 3.1. As expected, the coupling to T_2 is quite strong at 84 meV (680 cm^{-1}), which is easily strong enough to allow for ultrafast intersystem crossing. A crude estimate for the ISC rate can be obtained by inserting these coupling values into Fermi's Golden Rule,

$$\Gamma_{if} = \frac{2\pi}{\hbar}|\langle f|\hat{H}|i\rangle|^2\rho(E_f), \tag{3.4}$$

and neglecting the density-of-states by setting $\rho(E_f) = 1$ eV^{-1}. Shown also in Tab. 3.1, the ISC rate is fastest between the S_1 and T_2 states, at a rate of 15 fs. The ISC rates to the other triplets are slower, but are nevertheless still considered to be ultrafast events, with rates ranging from 66 fs to 134 fs—all faster than the experimental resolution of 150 fs.

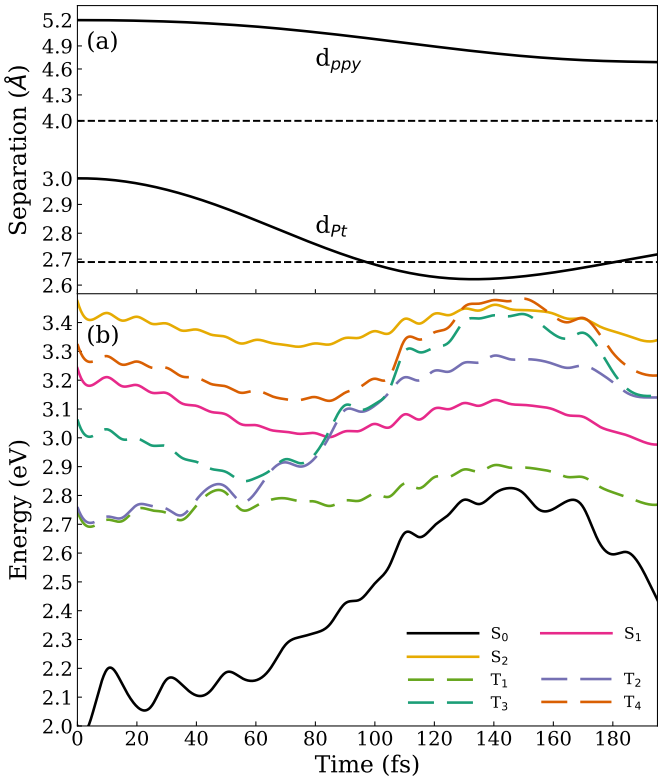


Figure 3.4. Non-relativistic Born-Oppenheimer molecular dynamics simulation along the S_1 surface of the platinum dimer. **(a)** Geometric parameters as a function of time. The distance between the platinum atoms and between the ppy ligands both decrease over the early times of the simulation. Dashed lines indicate the value of these parameters at the S_1 equilibrium geometry. Note the different scales on the y-axis for the two parameters. **(b)** The potential energy surface (PES) of the low-lying singlet and triplet states relative to the ground-state energy at equilibrium (S_0 shifted 2.0 eV for clarity). S_1 reaches a relative minimum, and intersects with T_2 and T_3, at 85 fs with a Pt-Pt bond distance of 2.72 Å.

While these static data are compelling evidence of ultrafast intersystem crossing to the second triplet excited state, ISC is an inherently dynamic process.[138] In order to qualitatively incorporate some of the photoexcited dynamics into the study of this molecule, an *ab initio* direct Born-Oppenheimer molecular dynamics (BOMD)[99, 100] trajectory was performed for 200 fs on the S_1 surface, at the same TD-CAM-B3LYP/LanL2DZ level of theory as above. The simulation began from the Frank-Condon point with zero initial momentum. The full potential energy surface (PES) at this level of theory is shown in Fig. 3.4b. At this level of theory, S_1 proceeds directly to the vicinity of the S_1 minimum, reaching a local energy minimum around $t = 85$ fs. In these first 85 fs, the Pt-Pt bond length (shown in Fig. 3.4a) decreases rapidly, from 3.0 Å to 2.72 Å, near the S_1 minimum value of 2.69 Å. The energy rises somewhat as the Pt-Pt bond length decreases past its equilibrium value to a minimum of 2.63 Å, after which it begins to turn back toward equilibrium. The ppy ligands, whose orientation is the second major change between the ground- and excited-state geometries (also shown in Fig. 3.4a), are much slower to react, steadily approaching one another over the course of the simulation but remaining far from the equilibrium separation of 4.01 Å.

Interesting dynamics are also visible in the triplet states over this range of geometries. As expected, S_1 intersects with both T_2 and T_3, around the same time as it reaches its local energy minimum. The change in electronic character of these states over the course of the simulation is evident. S_1 remains MMLCT over the entire trajectory, and T_2 remains LC. However, T_1 is LC at short times, becoming MMLCT by $t = 85$ fs, while T_3 follows the inverse pattern. This behavior is manifested within the PES: T_1 and T_2 track one another closely for the first 50 fs, after which they diverge and T_2 and T_3 then rise in energy together. From these dynamics, it is clear that the photodynamics are strongly downhill for the first 80 fs, where we therefore consider the BOMD results to be representative of configurations likely sampled at short times after photoexcitation. After that point, particularly if ISC is fast enough

that lower-lying triplet states might be meaningfully populated, it is less clear that the BOMD simulation remains physically meaningful, and consequently we restrict ourselves to considering the first 60 fs of the simulation.

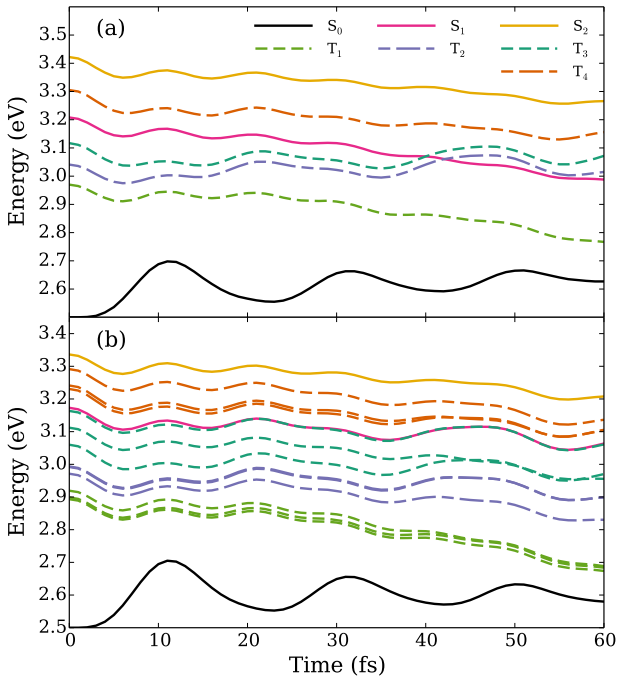


Figure 3.5. Relativistic PES of the Pt dimer computed with X2C-TDDFT using geometries sampled from Born-Oppenheimer molecular dynamics simulations along the S_1 surface (S_0 shifted 2.5 eV for clarity). **(a)** Spin-diabatic PES predicts S_1 crossing the triplets T_2 and T_3 at an earlier time of 40 fs. **(b)** Spin-adiabatic PES shows that strong spin-orbit coupling dramatically changes the potential energy landscape, splitting triplets in energy by up to 0.1 eV and preventing the intersection of surfaces.

With nuclear configurations now sampled, what remains is to evaluate the spin-orbit coupling between excited states at these geometries. To that end, single-point calculations were performed every 1 fs over the first 60 fs using X2C-TDDFT with and without spin-orbit coupling, with the resultant surfaces shown in Fig. 3.5. The PESs of spin-diabatic X2C results (Fig. 3.5a) look much the same as those from the effective core potential. However, even in the absence of spin-orbit coupling,

the explicit inclusion of scalar relativistic effects leads to an even earlier singlet-triplet crossing, occurring around $t = 40$ fs. Each triplet state in the spin-free X2C calculations correctly exhibits a three-fold degeneracy; for clarity, only one of each triplet state is shown in Fig. 3.5a. In contrast, a pronounced change occurs upon the inclusion of spin-orbit coupling (Fig. 3.5b). States that belong to the same triplet manifold are no longer degenerate and split in energy, by up to 0.1 eV in the case of T_3. The spin-orbit-coupled state corresponding to S_1 at short times is nearly flat over the trajectory, unlike its spin-diabatic counterpart, which steadily decreases in energy. While any degree of coupling between singlets and triplets would prevent true state crossings, one might expect to see several avoided crossings in regions where S_1 crossed triplet states in the spin-free calculations. Instead, there is very little change in relative energies along the trajectory, suggesting that spin-orbit coupling between states is strong relative to energetic differences. The dramatic change induced by spin-orbit coupling is evidence of very strong singlet-triplet interactions in the platinum dimer that lead to rapid ISC.

Using the procedure outlined above and detailed in Ref. [185], the spin-adiabatic states are rotated into the spin-diabatic basis, with the results plotted in Fig. 3.6a. Remarkably, the qualitative character of the spin-pure results is almost completely recovered; S_1 again decreases in energy, triplets nearly regain their degeneracy, and there is a clear state crossing between S_1 and the triplets T_2 and T_3. The diabatic couplings to S_1 from the various triplet states are also plotted in Fig. 3.6b. Singlet-triplet spin-orbit coupling is very strong in this species. As expected, T_2 couples most strongly to S_1, by nearly 90 meV (725 cm^{-1}) over the first 40 fs. Coupling to T_3 and T_4 is also significant, varying from 30 meV to 65 meV, while the S_1-T_1 coupling strength averages only 17 meV. However, this does not rule out the importance of the T_1 state in the photochemical process because it can couple to other triplet states after the ISC occurs.

Finally, the effect of spin-orbit coupling on singlet and triplet populations imme-

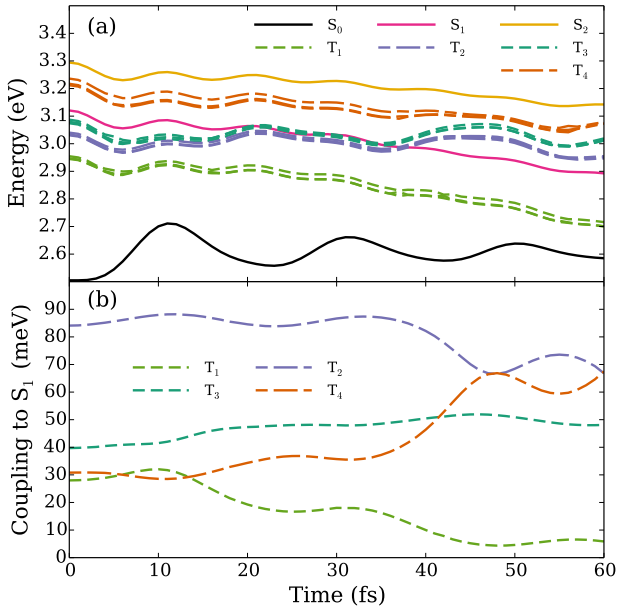


Figure 3.6. **(a)** Energies of spin-adiabatic states rotated into the spin-diabatic basis. The diabatic transformation of the spin-adiabatic states recovers the qualitative character of the spin-diabatic states: the triplet states regain their degeneracies, and S_1 decreases in energy over time and crosses with T_2 and T_3. **(b)** Diabatic coupling of triplet states to S_1. The norm of each set of singlet-triplet couplings is presented. T_2 couples most strongly, by nearly 90 meV over the first 40 fs.

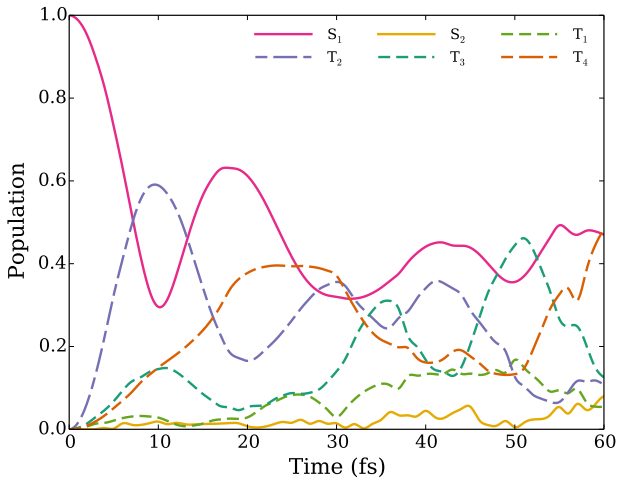


Figure 3.7. Population dynamics driven by spin-orbit coupling over the first 60 fs. The S_1 population drops by 70% over the first 10 fs, followed by oscillatory behavior.

diately following photoexcitation can be explored by directly propagating the populations according to

$$c(t + \Delta t) = \exp\left[\frac{-i\widetilde{H}\Delta t}{\hbar}\right] c(t), \tag{3.5}$$

where $c_i(t)$ is the coefficient of state $\widetilde{\psi}_i$ at time t, and where non-adiabatic coupling has been neglected. The resulting populations, equal to $|c_i(t)|^2$, are shown in Fig. 3.7 for the first 60 fs of the simulation, beginning with the total population in S_1 at time $t = 0$. A rapid depletion of S_1 is observed, dropping by 70% in just 10 fs. As expected, much of the population is transferred into T_2, reaching an initial maximum of 60% at time $t = 10$ fs. After 10 fs, other triplets gain appreciable magnitude and oscillatory populations are observed. That the population of S_1 oscillates over time rather than steadily decreases is expected, as a consequence of a lack of coupling between triplet states. In particular, though the T_1 state is not efficiently populated in this simulation, it is lowest in energy and does not couple efficiently to S_1; were the triplet populations to funnel into T_1, the reverse ISC reaction seen in the S_1

oscillations would be prevented. Transfer to T_1 can only be achieved through two factors neglected in this study: nonadiabatic coupling and vibronic effects. The former allows for direct population transfer between triplet states, while the latter can lead to broken symmetries and an increased internal conversion rate. Were these effects included, the intersystem crossing event would likely be rendered irreversible, and from these results, we predict the ISC rate to be under 15 fs.

3.5 Conclusion

In this work, we have demonstrated that strong spin-orbit coupling effects lead to ultrafast intersystem crossing in the $[Pt(ppy)(\mu\text{-}^t Bu_2pz)]_2$ dimer upon photoexcitation to the first singlet excited state. State-to-state spin-orbit couplings were evaluated using a variational approach, applied to structures sampled from the *ab initio* excited-state molecular dynamics. The rate of ISC is evaluated using both Fermi's Golden Rule and time-propagation of state coefficients. Due to its LC electronic character, its low energy separation from S_1, and its strong spin-orbit coupling to S_1, the second triplet excited state T_2 is the most likely candidate for the initial ISC event. While the experimental resolution of intersystem crossing in this species is 150 fs, this work suggests the intersystem crossing event is much more likely to occur within the first 15 fs, dramatically narrowing the window during which a change in spin state is likely to occur.

In the absence of explicitly evaluating nonadiabatic coupling between triplet states, and without allowing nuclear relaxation along triplet trajectories, nothing in the Hamiltonian in Eq. (3.3) will prevent the reverse ISC reaction. Development of analytical nuclear gradients, valuable in both nuclear dynamics and the evaluation of nonadiabatic coupling, remains therefore an important goal for variational relativistic methods. Future work will also consider the possibility of implicitly incorporating nonadiabatic coupling via a local diabatic transformation:[66, 146] once triplet states can efficiently couple to T_1, which does not meaningfully participate in the ISC, it is

expected that the intersystem crossing will no longer appear reversible.

Finally, this work is evidence of the importance of choosing an appropriate basis for describing dynamical processes. Fully variational relativistic calculations, though more accurate in their depiction of spin-orbit coupling, render identification of the relevant states much more difficult, as it is no longer meaningful to speak of them as singlets or triplets. In contrast, after rotating into a spin-pure basis as above, states fully recover their multiplicities and are clearly identified. This procedure, in effect a form of spin-purification, is an important step in the use of variational relativistic calculations in the simulation of intersystem crossing events.

Acknowledgments

The development of two-component electronic structure method is funded by the US Department of Energy (DE-SC0006863). The development of non-perturbative spectroscopic methods is supported by the National Science Foundation (CHE-1856210). Computational intersystem crossing is partially supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences, through Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Computations were facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system at the University of Washington, funded by the Student Technology Fee and the National Science Foundation (MRI-1624430).

Part II

CHEMICALLY-INFORMED ARTIFICIAL INTELLIGENCE APPROACH

Chapter 4

THE FOUNDATIONAL FRAMEWORK OF AI ALGORITHMS

"Artificial intelligence" is the broad term the encapsulates all forms of machine learning and is generally divided into three primary categories: supervised, unsupervised, and reinforcement machine learning. These three methodologies all seek the same primary goal, which is to identify the underlying patterns of known information such that future predictions are as accurate as possible. Typical problems addressed by these forms of analysis include classification, regression, clustering, and detection.[101, 63] The work in part II of this thesis will focus on building and implementing reinforcement learning (RL) algorithms to detect optimal states and ultimately solve problems of chemical relevance.

4.1 Elements of Reinforcement Machine Learning

Prior to discussing the more technical aspects of RL, it is worthwhile introducing the universal elements of the basic framework to better appreciate the architecture and novel functionality. As mentioned earlier, reinforcement machine learning differs from supervised and unsupervised methods in a number of ways, the most notable being RL does not require a pre-trained model and therefore handles input in a unique manner, relative to the more classical techniques. The key elements contributing to the unique functionality of RL include the following:

1. *State(s)*: the current description of the system at a given time step. This should hold very specific information, unique from other states, allowing the algorithm to learn the consequences/advantages of being in a given state. If one

considers building a molecule from scratch using RL, the state of this problem may hold information such as the set of atoms included in the molecule, the positions of each atom in Cartesian space, and even the multiplicity of atoms considered.[171]

2. *Action(s)*: a change to the current state determined by the RL algorithm; or more specifically, the agent of the RL algorithm. Actions allow for the collection of more data points in order to build a reasonable data set in which to train the model. In the early phases of model building during RL, the actions tend to be more stochastic; however, once sufficient data has been collected, the actions are more logically selected in order to achieve the predefined objective.
3. *Agent*: the decision-making component that selects which actions to take when in a given state. Initially, the agent is completely unaware of how the system behaves in the environment and must collect data (states) by taking actions. As new states are collected, the agent periodically undergoes training to understand the effects of its actions. After sufficient data collection and training, the agent is capable of detecting patterns and thus making informed decision allowing optimal states to be reached more efficiently. The agent will be represented by a neural network in this work.
4. *Policy*: the set of guidelines/rules the agent devises in response to training. In short, as the policy advances, it is better equipped to guide the agent and encourage the selection of better actions to find desirable states.
5. *Environment*: the interacting portion of the problem that receives the new state, which is determined by applying the selected actions on the current state. The environment evaluates the new state and a reward is granted that rates the quality of the actions. A potential energy surface (PES) may serve as a suitable environment for one who is looking to locate signature states of a molecule (i.e. global minimum or saddle points) via RL. If one were to consider a PES as the environment, information that can be extracted to evaluate a given state include the energy and subsequently the gradient.

6. *Reward*: a scalar value that rates the quality of the new state once it has been subjected to the environment. The reward function is a predefined component that should reflect the objective and encourage the agent to proceed toward said objective by awarding greater values when seemingly good actions have been selected. This is one of the most complicated components to build and should be strategically designed to keep the agent from converging to suboptimal states or succumbing to the cobra effect[1].

4.2 Dynamic Programming

Reinforcement machine learning operates under the dynamic programming (DP) regime and involves solving control problems via sequential problem solving and splitting the problem into sub-problems.[14] The objective is to identify an optimal policy such that the reward is maximized. In order to identify said policy, DP assumes a Markov decision process (MDP) model, which provides information about the states, actions, dynamics, and rewards. An analytic representation of the dynamics is unnecessary, instead a transition probability is generated to predict a mapping from the current state to the next state.[21]

Dynamic programming follows the principle of optimality, which states that all remaining decisions, after the initial state and decision, make up the highest performing policy.[13] The principle of optimality is significant because it gives rise to the functional equation and thus a means in which the optimal policy, π, can be identified. An effective optimal policy is one that is capable of maximizing the value function, $v(s)$, which is the expectation of the return function given the state. Simply put, it dictates how the agent will respond in a given state such that the maximum possible

[1]The cobra effect is an example of perverse incentive, which is when the reward function encourages unintentional outcomes. The experienced outcomes are generally negative and unconducive to solving the problem of interest.

reward is achieved.

$$v_\pi(s) = \mathbf{E}_\pi[G_t|S_t = s] \tag{4.1}$$

where v is the state-value function that follows policy π, s is the state at time t, and G_t is the return function. The return function is defined as the summation of all future rewards at a discounted value leading the former equation to become:

$$v_\pi(s) = \mathbf{E}_\pi[\sum_{i=0}^{\infty} \gamma^i R_{t+i+1}|S_t = s] \tag{4.2}$$

where γ is the discount factor, ranging anywhere from 0 to 1. The return function represents the quality of the current state by estimating whether or not the current state has the potential to give rise to higher rewarding future states; as such, a higher return value is desirable and indicative of a good state.[177, 203] The discount factor is important because it ensures convergence of the value function and dictates the extent to which the agent is concerned with future rewards. Finally, the expectation indicates what the reward is expected to be should the agent continue following π from state s.

4.3 The Value Function Bellman Equations

In an effort to derive the the state-value Bellman equation, one may rewrite **Equation 4.2** as:

$$v_\pi(s) = \mathbf{E}_\pi[R_{t+1} + \gamma G_{t+1}|S_t = s] \tag{4.3}$$

in order to separate the initial state from all subsequent states and remain compliant with the principle of optimality. Provided the algorithm progresses dynamically, a transition probability distribution must be defined

$$p(s', r|s, a) = \mathbf{P}(S_t = s', R_t = r|S_{t-1} = s, A_{t-1} = a) \tag{4.4}$$

which states that given state s and taking action a, the new state will be s' and the agent will be awarded r. The sum of the probability distribution for selecting a in s will sum to 1.

$$\sum_{s' \in S} \sum_{r \in R} p(s', r|s, a) = 1 \tag{4.5}$$

Continuing with the derivation, **Equation 4.3** can be split into two parts:

$$\mathbf{E}_\pi[R_{t+1}|S_t = s] = \sum_a \pi(s, a) \sum_{s'} p(s', r|s, a)r \tag{4.6}$$

and

$$\mathbf{E}_\pi[\gamma G_{t+1}|S_t = s] = \mathbf{E}_\pi[\gamma \sum_{i=0}^{\infty} \gamma^i r_{t+i+2}|S_t = s] \tag{4.7}$$

Focusing on **Equation 4.7**, the discount factor can be moved such that the equation becomes:

$$\mathbf{E}_\pi[\gamma G_{t+1}|S_t = s] = \gamma \mathbf{E}_\pi[\sum_{i=0}^{\infty} \gamma^i r_{t+i+2}|S_t = s] \tag{4.8}$$

and the expectation can be rewritten as:

$$\mathbf{E}_\pi[\sum_{i=0}^{\infty} \gamma^i r_{t+i+2}|S_t = s] = \sum_a \pi(s, a) \sum_{s'} p(s', r|s, a) \mathbf{E}_\pi[\sum_{i=0}^{\infty} \gamma^i r_{t+i+2}|S_{t+1} = s'] \tag{4.9}$$

Adapting the state-value function from the previous section, one finds:

$$v_\pi(s') = \mathbf{E}_\pi[\sum_{i=0}^{\infty} \gamma^i r_{t+i+2}|S_{t+1} = s'] \tag{4.10}$$

$$\mathbf{E}_\pi[\sum_{i=0}^{\infty} \gamma^i r_{t+i+2}|S_t = s] = \sum_a \pi(s, a) \sum_{s'} p(s', r|s, a) v_\pi(s') \tag{4.11}$$

The two parts of the initial equation can be recombined to find the Bellman

equation of the state-value function, which is a recursive relationship:

$$v_\pi(s') = \sum_a \pi(s,a) \sum_{s'} p(s', r|s, a)[r + \gamma v_\pi(s')] \tag{4.12}$$

Furthermore, the Bellman Optimality equation can be defined as:

$$v_{\pi^*}(s) = max_\pi v_\pi(s) \tag{4.13}$$

which indicates the best possible policy will give the greatest possible reward. Therefore the Bellman Optimiality equation of **Equation 4.12** is:[177]

$$v_{\pi^*}(s') = max_{a \in A} \sum p(s', r|s, a)[r + \gamma v_{\pi^*}(s')] \tag{4.14}$$

Through a similar derivation, the Bellman Optimality equation of the state-action value function is found to be:

$$q_{\pi^*}(s, a) = \mathbf{E}_{\pi^*}[R_{t+1} + \gamma max_{a \in A} q_{\pi^*}(S_{t+1}, a')|S_t = s, A_t = a] \tag{4.15}$$

4.4 Neural Network Architecture

The value functions discussed in the previous section are of utmost importance in RL as they connect all of the key elements of the architecture together. From the previous section, it is recognized that the equations derived collect all data in a single function and use the devised function to assess future states and actions in order to achieve the main goal.[161] Because of the inherent complexity of these functions, it is most appropriate to use neural networks in order to adequately represent them in complex problems. Neural networks are a suitable choice because of their advanced computational structures that have been modeled after biological network architectures.[60]

The general architecture of neural networks (NN) includes a series of layers and a number of neurons per layer. The depth of the NN and the number of neurons in

each layer is largely dependent on the objective and complexity of the problem. A standard NN consists of an input layer, hidden layer(s), and an output layer. The input layer will have a number of neurons equal to the number of input values, or features of the input, while the output layer will have a number of neurons equal to the number of expected output values. The hidden layer(s), positioned between the input and output layers, will have a predetermined number of neurons, which is given as a hyperparameter prior to training and should be determined based on the complexity of the problem investigated; the number of hidden layers can also be adjusted to accommodate the complexity of the problem. For increasingly complex problems, deep NN are implemented to capitalize on the increased nonlinear mapping capabilities offered under these architectures.

Figure 4.1 demonstrates the general structure of a simple NN where the input layer takes a single input value and the expected output from the NN has a size of two. Additionally, the sample NN includes two hidden layers each with 4 neurons. The dotted lines indicate the connections between each layer and are assigned unique weights, w_{ij}, which are updated throughout training.

Once a suitable architectures has been devised, one may define an appropriate loss function to minimize such that the weights of the network optimize via backpropagation during training. RL seeks to specifically optimize the value function, therefore the loss function in this work will be a modified state-action value function.

4.5 Training Neural Networks via Backpropagation

Backpropagation is a means by which one can train neural networks and minimize the loss function; this method works by continually updating the weights, which connect the layers of the NN.[158, 197] In other words, backpropagation ultimately seeks to find weights that allow a network to output results that are sufficiently close to the true results. It is useful to develop a deeper understanding of this process before moving forward, to better appreciate the power of this method and its impact on

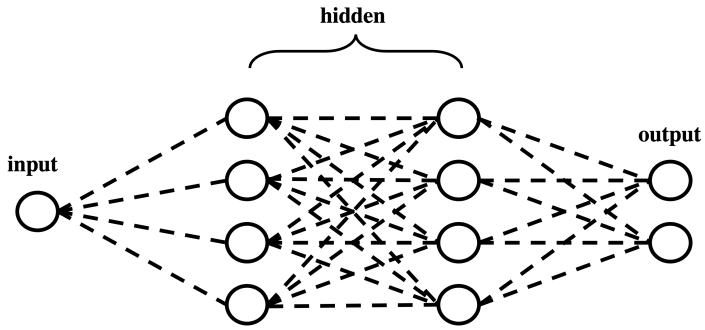


Figure 4.1. Schematic of a deep neural network where the circles indicate a neuron and the dotted lines are each assigned unique weights that neurons of one layer to another. The neural network pictured here takes a single input and maps it to the first hidden layer, which is composed of 4 neurons. Each neuron of the first hidden layer then maps the output to the second hidden layer. Finally, the second hidden layer maps the output to an output layer with two units.

developing successful machine learning models. To begin, the following relationship, connecting the input to the output, is defined:

$$x_j = \sum_i y_i w_{ji} \tag{4.16}$$

where x is the input of neuron j, the weight connecting j to neuron i is w_{ji}, and the output of i is y. Each layer is typically assigned an activation function that provides a real-value output value; this derivation will use the sigmoid activation function[2], which is:

$$y_j = \frac{1}{1 + e^{-x_j}} \tag{4.17}$$

[2]Other popular activation functions include the hyperbolic tangent (tanh) function, the rectified linear unit (ReLU) function, and the softmax function. These functions are imperative for incorporating non-linearity into the NN.

Next, the loss function will be defined as follows for simplicity:

$$L = \frac{1}{2}\sum_{j}(y_j - t_j)^2 \tag{4.18}$$

where t is the true output value. Having defined the above equation, the partial derivative of the loss function with respect to the output can be represented as:

$$\frac{\partial L}{\partial y_j} = y_j - t_j \tag{4.19}$$

which gives rise to the relationship

$$\frac{\partial L}{\partial x_j} = \frac{\partial L}{\partial y_j}\frac{\partial y_j}{\partial x_j} \tag{4.20}$$

through the chain rule. By differentiating the sigmoid activation function with respect to x and substituting the resulting equation into the above, one finds:

$$\frac{\partial L}{\partial x_j} = \frac{\partial L}{\partial y_j} y_j(1 - y_j) \tag{4.21}$$

The chain rule can again be utilized to write:

$$\frac{\partial L}{\partial w_{ji}} = \frac{\partial L}{\partial x_j}\frac{\partial x_j}{\partial w_{ji}} = \frac{\partial L}{\partial x_j} y_i \tag{4.22}$$

and

$$\frac{\partial L}{\partial y_i} = \sum_{j}\frac{\partial L}{\partial x_j} w_{ji} \tag{4.23}$$

At this point, the partial derivative of the loss function with respect to y can be determined for any neuron in the second to last layer when the $\frac{\partial L}{\partial y}$ is known.[158] This process is repeated for each earlier layer in order to find $\frac{\partial L}{\partial w}$. The latter term is

then used to update the weights via gradient descent.

$$\Delta w = -\alpha \frac{\partial L}{\partial w} \tag{4.24}$$

where α is the learning rate limiting the weight update.

Chapter 5

DEVELOPING A PHYSICALLY-MOTIVATED DEEP REINFORCEMENT MACHINE LEARNING ALGORITHM

Considering the high computational cost of evaluating the energetic landscape of systems using traditional quantum methods (i.e. DFT), a deep machine learning approach is explored in this chapter to build a foundation for an alternative analysis method. A physics-supplemented deep deterministic policy gradient algorithm is extended upon and implemented in a model system to demonstrate the capability and versatility of AI methods in solving problems of chemical interest. The work presented in this chapter has been adapted with permission from Mills, A., Goings, J., Beck, D., Yang, C., Li, X. Exploring potential energy surfaces using reinforcement machine learning. The Journal of Chemical Information and Modeling, 62 (13), 3169-3179.

5.1 Introduction

Computationally modeling chemical systems has been a valuable tool for many years as theoretical methods allow a deeper look at the systems under investigation. In order to extract meaningful data from theoretical simulations and closely reproduce experimental results, the systems investigated must first be optimized to find the equilibrium geometry allowing for more accurate characterization. Optimization techniques have been around for decades and range from the rational function method to the direct inversion in the iterative subspace method.[162, 176, 9, 39] Traditional optimization methods can be improved upon by including regularly updated second-

order couplings between coordinates via the Hessian matrix.[50] This, however, has the potential to substantially increase the computational cost; especially in cases with high-dimensional spaces.[163, 2]

High computational costs have been combated with quasi-Newton methods that circumvent calculating the full Hessian and instead find a suitable approximation.[20, 168] Further advancements have led to the development of first-order direct inversion of the iterative subspace methods; which has demonstrated similar, if not better, performance compared to quasi-Newton methods.[55, 149, 70, 98, 122, 61] These algorithms, however, have their faults, which include premature convergence, oscillatory behavior, and taking steps in the wrong direction consequently leading to unreliable solutions.[55, 98] Additionally, most optimization algorithms rely on the input structure being reasonably close to the equilibrium geometry, through the aid of molecular visualization programs, which is not always the case; especially as the structures become larger and increasingly more complex.

In an effort to improve the performance of optimization methods, while maintaining good quality energy predictions, machine learning algorithms have recently been explored. Notable advancements in the past decade have utilized supervised learning methods to build DFT-quality potential energy surfaces (PES).[173, 12] However, methods relying on training data often suffer from substantial overhead, that comes with pre-training models, and insufficient sampling variability, which can be a consequence of specialized datasets. One might consider overcoming such limitations through the implementation of reinforcement machine learning techniques, which is a subcategory of machine learning that does not require datasets for initialization and is capable of learning on-line.

Recently, RL has been applied, in concert with classical methods, to molecular optimization problems with the objective of improving convergence.[2] RL works by optimizing a policy through a developer designed reward system that evaluates the quality of the algorithm's predictions. This method has been implemented in the

work of Ahuja et al. where they devise a method using RL to find an appropriate correction term for steps within the BFGS algorithm. However, this work implements a preconditioned policy thus forgoing one of the greatest benefits of RL, the ability of learning without initializing a pre-trained model.

The work presented in this study demonstrates the remarkable problem solving capabilities of an RL agent as it navigates through various PES to identify the global minimum and locate local minima points. The surfaces included in this study were explicitly selected as model environments because of their multiple minima regions making the global minimum less obvious to the agent; thus significantly increasing the difficulty of the problem. The algorithm implemented herein extends on the deep deterministic policy gradient (DDPG) method, introduced in the work of Lilicrap et al., by incorporating an optimizing target state and utilizing gradient vector information indirectly provided by the state defined in this problem.[105] Though the DDPG framework has been shown to work well in high-dimensional spaces, the experiments herein will be limited to two degrees of freedom for ease of visualization and interpretability.

5.2 Theory

At the most fundamental level, an RL algorithm consists of an agent that interacts with a predefined environment to collect the highest possible rewards. Through trial-and-error and accumulated rewards, the agent uses the acquired experiences to optimize a policy (*i.e.*, function approximator) in order to solve a given problem. As more experiences are collected and the policy matures, the agent begins to make more informed decisions and subsequently completes tasks with increased efficiency.

5.2.1 The Markov Decision Process

The optimization problem, within the RL regime, will be framed as a Markov decision process (MDP); which is the process of selecting an action to achieve the highest

rewarding future state, given the transition probabilities.[150, 3] An MDP is preferred because of its simplicity and superior performance in stochastic environments.[10, 179] Within this formalism, the states are assumed to obey the Markov property, which is to say that all pertinent information for future states is held in the current state. States of this nature are considered "memory-less" since past states are deemed irrelevant and are promptly ignored.[77] Mathematically, the transition probability of a Markov state can be represented by the following relationship:[177]

$$\begin{aligned} p(s', r|s, a) &= Pr[R_{t+1} = r, S_{t+1} = s'|S_0, A_0, R_1, ..., S_{t-1}, A_{t-1}, R_t, S_t, A_t] \\ &= Pr[R_{t+1} = r, S_{t+1} = s'|S_t, A_t] \end{aligned} \tag{5.1}$$

where s and s' are the current and next states, respectively, r is the reward, a is the action, and t is the time. Since the problem is framed as a process, the dynamics of the environment are evaluated via the transition probabilities.[178] Because of their dynamic nature, MDP models require the following information: the state space, action space, reward function, and transition probability function; generally collected in a tuple, $(\mathcal{S}, \mathcal{A}, \mathcal{R}, p)$.[108] Problems within the MDP framework can be approximately solved by learning an optimal policy through reward updates as high probability trajectories in the environment are traversed.[131]

5.2.2 Policy Gradient Theorem

Policy gradient methods are commonly used to solve MDP problems and will be the class implemented in this study. These particular methods are appealing because of their increased convergence arising from parameterizing the policy such that actions are selected in the direction of the greatest possible reward via gradient ascent.[170, 143, 144] Methods of this class are optimal for real world problems as they are able to handle continuous action spaces by representing their policy as a

probability distribution; generally written as $\mu_\theta(a;s) = P[a;s,\theta]$.[143, 82] This indicates the policy, μ, follows parameter vector θ to select a_t, given the current state, s_t. The parameter vector, in this context, will be composed of weights that are extracted from a function approximator.

5.2.3 Actor-Critic Method

The actor-critic (AC) method is a sub-model of the policy gradient theorem class that combines formerly unrelated models to capitalize on their individual benefits and make a, single, superior model. The AC model consists of two parts:[170, 90]

I.) an actor that optimizes the parameters of a stochastic policy via gradient ascent

II.) a critic that estimates the Q-value function via temporal-difference learning

The two components work in tandem to find an optimal policy function through parameter updates directed by the learned value function.[90] The latter component learns through the recursive relationship:

$$Q(s_t, a_t) = \mathbb{E}[r(s_t, a_t) + \gamma\, \mathbb{E}[Q(s_{t+1}, a_{t+1})]] \tag{5.2}$$

where $\mathbb{E}$ is the expectation and γ is the discount factor, which is a predetermined hyperparameter that indicates how heavily future rewards are weighted. **Equation 5.2** is commonly known as the Bellman equation and is of great significance in RL as it relates the state and the action to one another. The critic is then updated by minimizing the following loss function:

$$L(\theta^Q) = \mathbb{E}[(Q(s_t, a_t; \theta^Q) - y_t)^2] \tag{5.3}$$

where y_t is

$$y_t = r(s_t, a_t) + \gamma Q(s_{t+1}, \mu(s_{t+1}); \theta^Q) \tag{5.4}$$

The critic network is then used to update the actor network through gradient ascent

$$\nabla_{\theta^\mu} J \approx \mathbb{E}[\nabla_{\mu(s_t)} Q(s, \mu(s_t); \theta^Q) \nabla_{\theta^\mu} \mu(s_t; \theta^\mu)] \tag{5.5}$$

The AC model combines Q-learning-type methods with policy gradient methods to improve convergence issues common in the former by including gradient-based optimization via the latter. Additionally, large variances, often seen in actor-only models, are controlled by incorporating low-variance knowledge of the dynamics via the critic model; which has the added advantage of increasing the agent's rate of learning.[67]

5.2.4 Deep Deterministic Policy Gradient

The DDPG method expands on the AC architecture by implementing target networks and including a memory of past experiences, which work together to improve both the stability and sampling efficiency, respectively.[105] The target networks, denoted by a prime symbol, are slowly updated throughout the training process, relative to the online actor and critic networks, using the following soft update implementation

$$\theta \leftarrow \tau\theta^x + (1 - \tau)\theta^{x'} \tag{5.6}$$

where τ is a very small constant and x can be either μ or Q for the actor or critic, respectively.

5.3 Methods

Herein, a hybrid-DDPG (hDDPG) algorithm is implemented that utilizes the gradient vector, calculated from the potential energy and position information, in an effort to help direct the agent toward the global minimum of the PES. This takes inspiration

from classical methods that rely solely on optimizing systems via gradient-descent and novel techniques that have shown improved optimization convergence. Additionally, an updating target state is implemented by first introducing a recorded target value and then periodically replacing this value as a higher rewarded state is identified; this is incorporated in the reward function to aid in finding the optimal solution. The general process is outlined in **Algorithm 1**.

Algorithm 1 hDDPG

Initialize the environment and agent's memory
for episode=1,2,...,M **do**
    Reset the initial state
    if *done* is not True **then** ▷ termination condition has not been triggered
        Select a_t using the following conditions:
        if energy increased 3 consecutive steps **then**
            Use gradient vector forces to select actions: $a_t = -\nabla E$
        else
            Use $\mu_\theta(a; s)$ to select a_t
        end if
        Apply a_t and evaluate the environment's response to s_t via the reward function
        Consult the termination conditions to update *done* if necessary
        Update memory and and train the agent
    end if
end for

The agent is composed of an AC architecture with complimentary target actor and target critic networks as well as an experience memory. The target networks are additional neural networks used to ensure stability as the agent advances; they are updated during training following **Equation 5.6**.[170]

To encourage exploration, this work implements both Ornstein-Uhlenbeck (OU) noise and randomized initial states, within preset environmental bounds, for each episode. OU noise has been included as it satisfies the Markov property and the OU process demonstrates temporal homogeneity.[45] Additionally, randomized initial states allow the agent to build a better understanding of the environment as more of

the area is stochastically surveyed.

The Müller-Brown PES was selected as the initial model environment for this study because of its increased complexity arising from the numerous minima energy regions. This surface is modeled by the following equation, which is simply the sum of four Gaussians:[16]

$$\lambda(x,y) = \sum_{i=1}^{4} A_i exp[a_i(x-\bar{x}_i)^2 + b_i(x-\bar{x}_i)(y-\bar{y}_i) + c_i(y-\bar{y}_i)^2] \tag{5.7}$$

where

$$\begin{aligned} A = (-200,-100,-170,15); a = (-1,-1,-6.5,0.7); \\ b = (0,0,11,0.6); c = (-10,-10,-6.5,0.7); \\ \bar{x} = (1,0,-0.5,-1); \bar{y} = (0,0.5,1.5,1) \end{aligned} \tag{5.8}$$

An observation space was set by limiting the x value between [−2,1.25] and the y between [−0.5,2.25], which can be visualized in **Figure 5.1**. Actions selected via the function approximator are screened and clipped prior to passing through the step function to ensure s_t remains within the aforementioned limits. This keeps the agent from saturating the memory with useless experiences that may hinder learning or force exploding/vanishing gradient problems.[47]

Considering the stochastic nature of RL and the complexity of PES, a number of termination conditions are defined to prevent nonphysical states from dominating the replay buffer. The agent is instructed to abandon an episode if the calculated energy increases a total of 25 times, a reward of zero is granted 5 times, the allotted steps have been reached, or there is a notable increase in energy; this last condition is case specific. For example, in the case of Müller-Brown surface, if the following relationship is met, the episode will promptly terminate: $\lambda_t > 2 * |\lambda_i|$, where λ_t is the current state's energy and λ_i is the initial state's energy of the episode. This particular condition enforces tighter exploration constraints when the initial state is in an energetically positive or null region. Alternatively, if the initial state is in a more

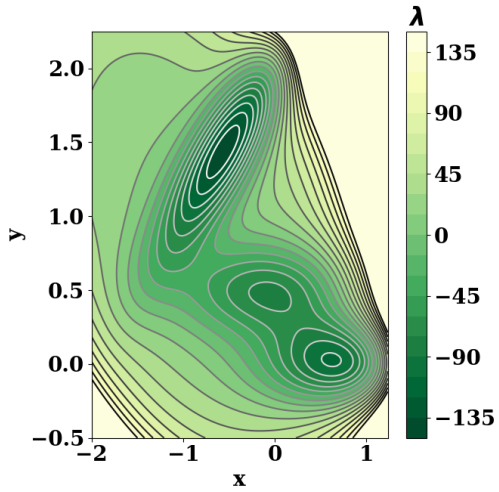


Figure 5.1. Contour of the Müller-Brown potential energy surface.

energetically negative region, the agent is granted more freedom to explore the area per this relationship. It should be reiterated that this final condition is well suited for the Müller-Brown potential energy surface and may not perform well for surfaces with different energetic ranges.

5.4 Results & Discussion

5.4.1 Improving the Critic with an Advanced Reward Function

The critic is immediately dependent on the quality of the reward granted (**Equation 5.2**), therefore devising a strategic reward function is instrumental in the agent's ability to learn and solve the problem of interest. As such, the reward function is scenario specific and must incorporate relevant properties of the environment to ensure the highest quality actions are identified and, subsequently, utilized.[132, 127, 128] Two reward functions are compared in this section to illustrate the significance of a well defined function that compresses multiple aspects of the PES into a single, unitless, value. The first reward function implemented is simply the difference between

λ of the previous state and current state:

$$r_t = \lambda_{t-1} - \lambda_t \tag{5.9}$$

Algorithm 2 Moving Target Reward Function

```
Calculate Euclidean distance, ϵ_t, between [x_low, y_low] and [x_t, y_t]
if λ_t > λ_low then
    if ϵ_t = ϵ_{t−1} then
        r_t = 0
    else if ϵ_t > ϵ_{t−1} then
        r_t = −1 ∗ ϵ_t
    else
        r_t = clip(1/ϵ_t, 0, φ)
    end if
else if λ_t = λ_low then
    r_t = ϵ_t
else
    r_t = clip(1/ϵ_t, 0, φ) ∗φ
    Update λ_low with new lowest observed energy
end if
```

The second reward function evaluated can be found in **Algorithm 2**. The very first step of this function is to extract the positional components of the state, $[x, y]$ and calculate the Euclidean distance, ϵ, between the current coordinates and the coordinates with the lowest recorded potential energy, λ_{low}. Next, the λ_t is compared to λ_{low} to determine if the agent has found a lower energy state. If the agent's most recent state is found to have a more positive energy, the first sub-route is chosen and assigns a reward based on the calculated Euclidean distance. Notice the only way to obtain a positive reward, via this route, is to have found a smaller ϵ indicating the agent is indeed moving toward the target state. However, if a new target has not been found, the maximum positive reward will be limited to ϕ, where ϕ is set to 10 for all PES explored in this work, via the *clip* function; which ensures the value does not exceed ϕ or go below 0. Alternatively, if the agent decides to remain

stationary, a reward of zero is assigned as nothing is learned therefore nothing is gained. Finally, if the agent identifies a lower energy state than λ_{low}, the target state is updated and a maximum reward of $\phi * \varphi$ can be collected, where φ is set to 5 in this study thus limiting the maximum possible reward to 50. It is worth mentioning that reward clipping seen in lines 8 and 13 of **Algorithm 2** is necessary to avoid exploding/vanishing gradient issues that accompany states with extraordinarily large value disparities. The update feature within the reward function will be termed the *moving target* going forward as this recorded state is frequently updated as the agent learns more about the environment. It should be recognized that the recorded state represents the best possible solution to the problem, with respect to the acquired experiences.

Utilizing the standard DDPG implementation, the agent trained on the moving target reward function (**Algorithm 2**) was evaluated against the agent trained via the simple reward function (**Equation 5.9**). Both agents were allotted 750 episodes to optimize a policy where the first 90% of episodes randomly initialized a starting state, which can be thought of as the training set, while the last 10% of episodes used a fixed initial state that was deemed sufficiently far from the optimal point to evaluate what the agent has learned. These final 10% of episodes are intended to act loosely as a "test" set to evaluate the quality of the policy. It should be noted that the agent never stops learning and will still collect the experiences from these last 10% of episodes to train the networks.

Learning the Müller-Brown Potential Energy Surface

The initial states recorded during training, using both the simple and moving target reward functions, have been plotted in **Figure 5.2** to demonstrate the extent in which exploration is practiced. **Figure 5.2a** suggests that both agents have seemingly equal opportunities to traverse the surface and build a reasonable picture of the landscape as the area is well surveyed in both simulations. The corresponding surface, in **Figure**

5.2b, however, plots the recorded terminal states and illustrates the learning variance between the two agents. The agent learning from the simple reward function falls short of locating the global minimum and instead fixates on a region approaching the nearest local minima. The agent trained on the moving reward function appears to frequent this region initially, but eventually learns of a better solution and begins to find terminal states localized in the global minimum of the surface, which is likely a consequence of having a well defined target to aid the agent's decision-making.

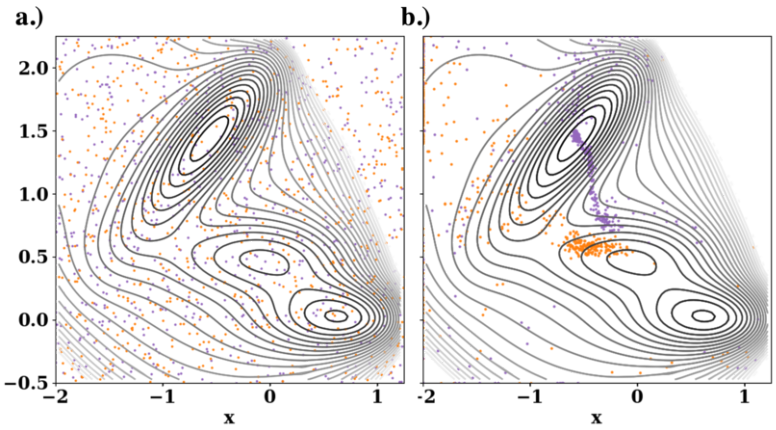


Figure 5.2. a.) Initial and b.) terminal states of all episodes for the agent trained on the simple reward function (orange) and the agent trained on the moving target reward function defined by **Algorithm 2** (purple) atop the true Müller-Brown potential energy surface.

Demonstrating Knowledge of the Surface

To evaluate the quality of training completed prior, the last 10% of episodes will begin with a fixed initial state. Of these last 10% of episodes, the final 50 episodes are plotted in the form of trajectories in **Figure 5.3** to visually represent the improved convergence when an updating target is included in training. Lacking any formal direction via a target, the orange trajectories are observed to localize at a suboptimal

area between two minima regions, which reflects the trend observed during training. This is likely a consequence of receiving comparable positive rewards from the surrounding minima regions rendering the agent unsure of which is the better option. From this analysis, it becomes clear that **Equation 5.9** is not suitable for efficiently traversing complicated environments in the allotted episodes. Perhaps given more time to learn and explore via randomized state initialization, the agent would eventually converge to the global minimum; however, from the collected data, it seems the agent has settled on converging to a rather poor solution. Although granting the agent more time to collect experiences is a viable option, it is certainly not the most efficient; instead, modifying the reward function to indirectly hold more information about the surface will likely lend to increased learning and thus convergence to a more suitable solution.

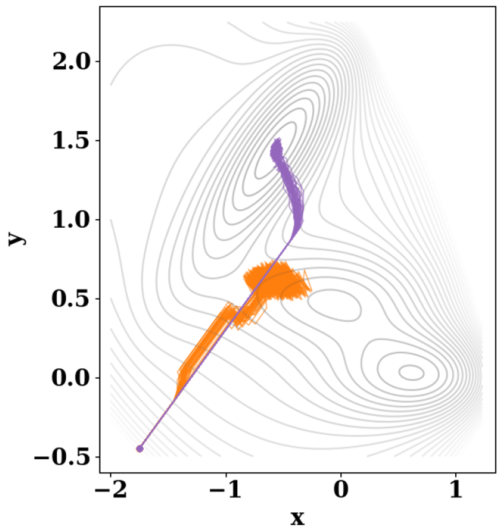


Figure 5.3. Trajectories for episodes 700-750, which have a fixed starting point $(-1.75, -0.45)$, using a simple reward function (orange) and the moving target reward function (purple) on the Müller-Brown potential energy surface.

The purple trajectories from **Figure 5.3** show increasingly better solutions than their counterpart orange trajectories. Every trajectory the agent pursues in these final episodes immediately fall down the surface until a slight increase in λ_t is found, in which case the agent redirects and proceeds toward the global minimum region. Though impressive, this is perhaps unsurprising as the reward function with the optimizing target is explicitly designed to give a better description of the environment by considering both λ_t and ϵ to relate s_{t-1}, s_t and s_{low} to one another. That is to say, the geography of the environment is now considered in addition to the state's potential energy, which has a noticeable impact on the quality of the policy. Having motivated the importance of a descriptive reward function, moving forward, an optimizing target is implemented and the moving target reward function, outlined in **Algorithm 2**, will be utilized.

5.4.2 Guiding the Actor with Physics

To further improve the agent's learning rate, it is expected that ensuring valuable experiences are collected in the memory will allow for an increased rate of convergence to optimal solutions. To address this, the hDDPG algorithm is developed where the actor is designed to incorporate physically meaningful experiences, via gradient vector information, whilst learning the surface. This vector is calculated intermittently, based off of a predefined criterion, and used in place of policy selected actions. The gradient is utilized to help direct the agent specifically when it demonstrates a lack of environmental awareness through consecutively collecting negative rewards. Additionally, the inclusion of gradient vector information broadens the applicability of this framework and allows for additional stable states to be easily located; namely, potential reactants and intermediates of the investigated system.

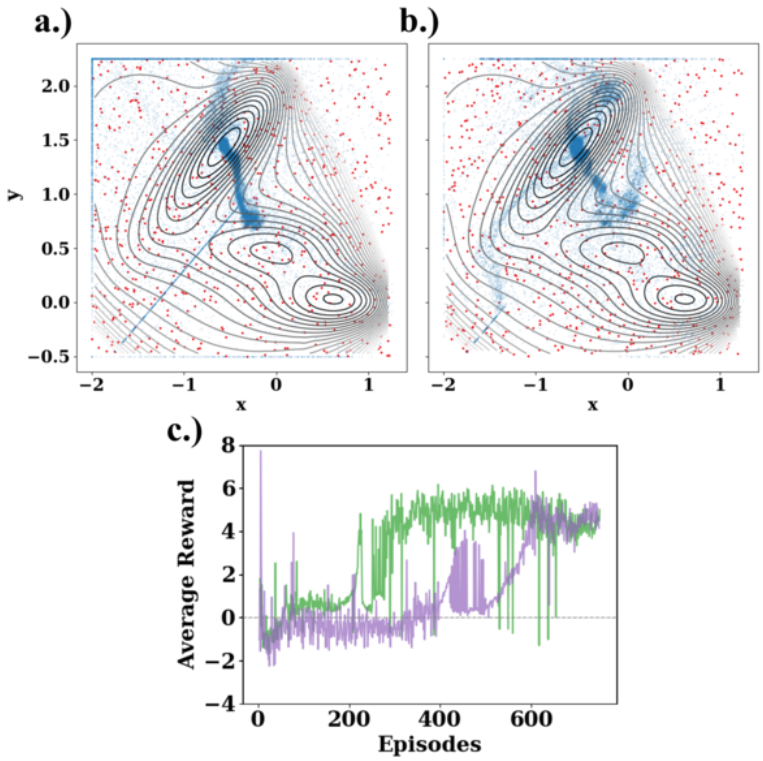


Figure 5.4. Contour plot of Müller-Brown potential energy surface built using every sampled point during exploration via a.) DDPG and b.) hDDPG where red points are initial states for each episode and the blue points are sampled states during the episodes. The average reward (c.) for each episode is recorded for standard DDPG (purple) and the hDDPG (green) algorithms.

Learning with Physics

A series of training runs are analyzed to illustrate the differences in the agent's ability to locate the global minima region using the DDPG and hDDPG algorithm. A side-by-side comparison of representative training runs, from each algorithm, is presented in **Figure 5.4**. The red points represent the initial state of each episode and the blue points are all other states visited by the agent at some point during the episode. The initial states are visually found to be well dispersed about the environment allowing

for increased exploration and a diverse set of states to learn from. The increased exploration significantly improves the agent's ability to describe the environment, which is demonstrated by the underlying contour plot that is interpolated using strictly the states visited by the agent of the respective algorithm. Additionally, the blue points emphasize the regions the agent found to be of great interest through clustering, indicated by the dense dark blue regions. The higher density cloud about coordinates $(-0.75,1.5)$ in **Figure 5.4b** suggests the hDDPG algorithm does a better job at locating the global minimum region and continually visits that area when compared to DDPG. The DDPG agent, on the other hand, appears to have an affinity for the space between the global minimum and a local minimum well.

The quality of the policy and rate of learning is evaluated through analyzing the average collected reward, which is seen in **Figure 5.4c** where the average rewards are plotted for the DDPG and hDDPG algorithms. The glaring difference between the two plots is the rapid increase in average reward allotted for the latter compared to the former, suggesting a significant difference in the learning rates. An average reward of $\sim$5 is granted by episode 300 for the hDDPG agent, which is nearly 300 episodes sooner than convergence to a comparable average reward for the DDPG agent. This indicates the hDDPG agent is able to quickly locate a satisfactory solution and continually select subsequent actions that would grant generally higher rewards. A closer look at the two plots show the DDPG agent struggles to maintain an average positive reward for nearly the first half of training, unlike the hDDPG agent who maintains a positive reward, consistently, after about 100 episodes-worth of experiences. This difference can be attributed to the use of physically motivated actions that help direct the hDDPG agent toward better solutions and thus higher rewards.

Quality experiences collected through improved guidance via gradient vector actions has a substantial impact on the agent's ability to learn and make improved decision. It should be noted that the hDDPG actor used the gradient to direct the trajectory for only 2% of the total steps during training. Additionally, these physically

motivated actions were less frequently utilized in later episodes indicating the agent gained enough information to continue independent of the gradient vector. These superior experiences contribute to the collection of more physically meaningful states, which strongly enhance the agent's visualization of the environment. An additional effect of finding more physically meaningful states and successfully progressing toward a defined target is avoiding premature termination. Unlike the hDDPG training, the standard DDPG training was found to have terminated many of the earlier episodes rather quickly, depriving the agent of sufficient data/experiences to learn from. This very likely contributes to the difference in learning rate between the DDPG and hDDPG agents.

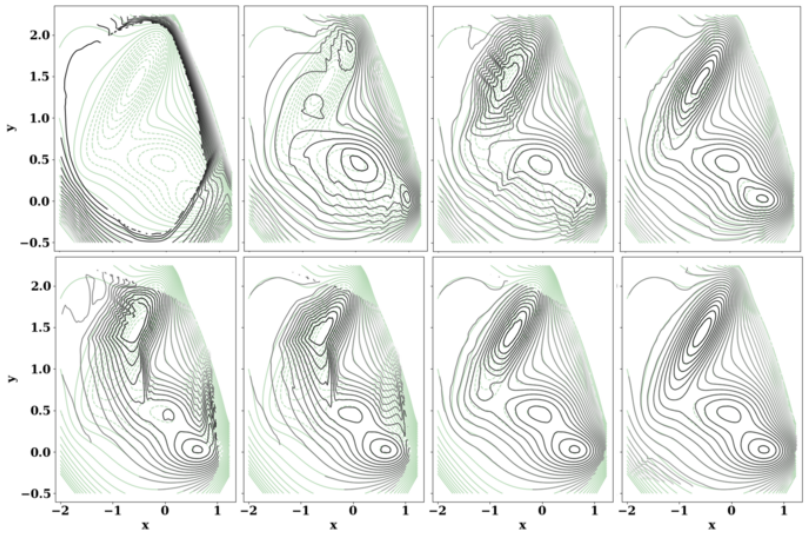


Figure 5.5. The interpolated surface (black) compared to the true Müller-Brown surface (green) for the DDPG agent (top row) and the hDDPG agent (bottom row) after 5, 10, 25, and 50 episodes (from left to right).

A visualization of the environment through the agent's perspective is shown in **Figure 5.5** to demonstrate how the collected experiences, after a set number of episodes, influence the agent's perception of the surface. The plots represented by

the black curves are interpolated from the limited experiences collected after the indicated number of episodes. These surfaces are plotted above the true Müller-Brown surface to appreciate the agent's ability to build the environment purely based on collected states. Remarkably, after only 5 episodes, the hDDPG agent is capable of building a surface that resembles the Müller-Brown surface and by 50 episodes, the agent's idea of the environment is nearly identical to the true surface. Alternatively, the DDPG agent is unable to initially form a reasonable image of the surface as the selected actions during exploration have led to states populating the outer regions of the environment. Furthermore, because of the poor actions, the first couple of episodes during training for the DDPG agent were terminated prematurely to keep from saturating the memory with states along the environment's boundaries and subsequently collecting an abundance of negative rewards. This lack of data contributes strongly to the agent's inability to demonstrate a satisfactory understanding of the environment in early episodes. However, through additional training, the standard DDPG agent begins to form a better picture of the surface. In fact, by 50 episodes, both agents seem to have collected sufficient data to build a surface that closely resembles the expected Müller-Brown PES suggesting the observation space has been well explored.

Though the agent has built a sensible interpretation of the environment by episode 50, that does not necessarily indicate the policy has been optimized to solve the problem under investigation. This is further emphasized by revisiting the average reward plots (**Figure 5.4c**), which show the soonest an agent is capable of making consistently high rewarding actions is just before episode 300. Although the first 50 episodes acquired rather low rewards, these episodes are still valuable as they allow the agent to explore critical point regions as it searches for the global minimum. The RL algorithm has been devised such that states resembling critical points are recorded during training if they satisfy the following conditions: i.) the magnitude of the gradient is approaching zero and ii.) the state of interest is sufficiently far

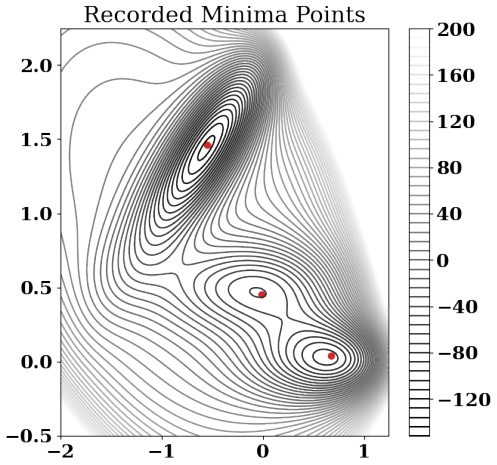


Figure 5.6. Recorded minima points (red) collected during 50 episodes training and visualized on the true Müller-Brown surface.

from all previously recorded points. Condition *i* is set to confirm the current state of interest does not reside on a slope of the surface while condition *ii* restricts the number of points that can be recorded within a given proximity. As training progresses and additional states are identified as potential critical states, they are compared to previously recorded states, via the Euclidean distance, and are either added to the list of recorded minima points or they replace a nearby existing point. A state can only replace an existing state if the observed energy of the new state is lower than the existing state's energy.

Condition *i* is recognized to have the potential of recording problematic regions of the surface such as saddle points while the *replace* action, applied during the list update, can force an existing state out of this region and onto a slope of the surface. To remove such states from the minima list a final check is completed by first ensuring conditions *i* and *ii* are still satisfied and then evaluating the quality of the recorded states relative to the interpolated PES built using all states visited by the agent.

Each state from the list is evaluated on the interpolated surface by sampling a small grid-like region about the recorded state and comparing the energies of each point. If any of the sampled states are observed to decrease in energy by 10% of the recorded state's energy, then the state is decidedly not a local minima and is removed from the list.

The collected minima points of the Müller-Brown surface from representative simulations are recorded in **Table 5.1** to show that although a suitable policy is not yet optimized for either DDPG or hDDPG agent, they are both capable of exploring the surface and identifying points in local minima regions. It is recognized that the DDPG agent has not been allotted sufficient training time to capture all minima regions as indicated by the missing data in the table. The limited training time emphasizes the impact that intelligent actions have on efficient data collection. The recorded minima points collected by the hDDPG agent are presented in **Figure 5.6** to further visualize the collected states. The resolution has been increased, in the form of contour lines, to illustrate the accuracy of the recorded states relative to the expected minima point of the true surface. Remarkably, all local mininma regions are successfully found and recorded during training regardless of well depth or location. This is largely due to the incorporated physics, which encourage directed learning and guides the agent along the surface towards lower energy regions of the environment. Though the physics-informed actions appear to play a significant role in locating minima regions early on in training, the importance of exploration should not be overlooked as this also has a

DDPG			**hDDPG**		
x	y	λ	x	y	λ
-0.0965	0.4572	-80.3822	-0.0154	0.4536	-80.5743
-	-	-	0.6689	0.0417	-107.1873
-0.5207	1.3663	-134.1448	-0.5498	1.4637	-146.4170

Table 5.1. Collected minima points recorded by both the DDPG and hDDPG agents during the first 50 episodes of training.

substantial impact on the agent's ability to identify and collect representative points of these regions.

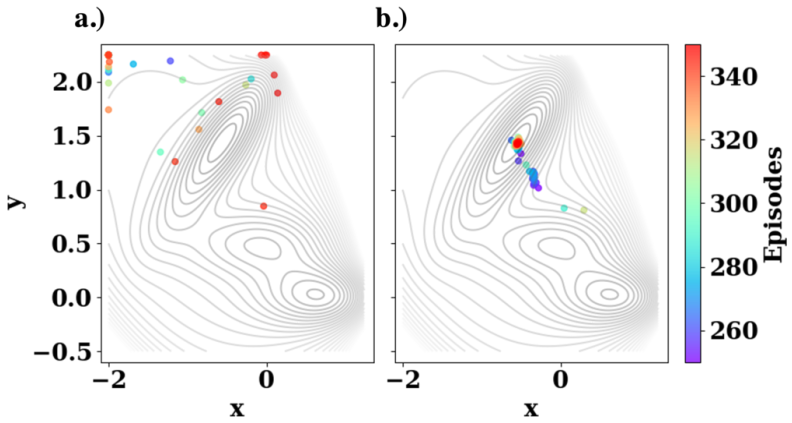


Figure 5.7. Terminal states for episodes 250-350 atop the true Müller-Brown potential energy surface for a.) DDPG and b.) hDDPG where the color bar indicates the episode number and corresponding color.

As stated previously and observed in **Figure 5.4c**, higher rewards are not accumulated until roughly episode 300, it is at this point that an optimized policy begins to take form. Additional information, giving insight to the agent's progression, can be extracted from the location of the terminal states for each episode. **Figure 5.7** showcases the agent's ability to identify and proceed toward the global minimum between episodes 250 and 350 by plotting the terminal state for each episode in the given range. This range was selected as it covers the transition from lower to higher average rewards in the case of the hDDPG agent. The agent trained via DDPG displays a chaotic arrangement of terminal states with hardly any comprehensible trend. Conversely, the hDDPG agent appears to be honing in on the global minimum, indicated by the color gradient approaching red as the deepest region of the well is reached. The trend in terminal state location differs greatly between the two algorithms and demonstrates the significance of learning from physically meaningful states, which are

objectively states of higher quality and therefore greater value.

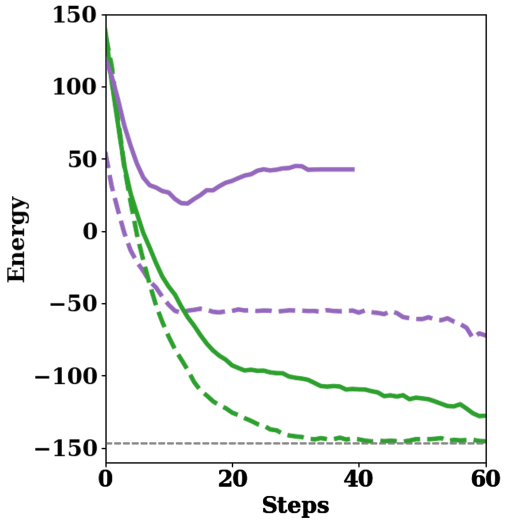


Figure 5.8. The average energies at each step for episode ranges 250-350 (solid lines), corresponding to **Figure 5.7** episodes, and 500-600 (dotted lines) for the standard DDPG algorithm (purple) and the hDDPG algorithm (green). The dotted grey line indicates the true global minimum of the Müller-Brown surface.

The average energies at each step between episodes 250-350 are plotted in **Figure 5.8** as solid lines to show the agent's progression as each step passes. These results clearly show the hDDPG agent has a far superior understanding of the environment and is capable of nearing the global minimum much sooner than the DDPG agent. The standard DDPG agent, on the other hand, is still finding it quite difficult to satisfactorily solve the problem and instead converges to an energy of ~ 50, well off of the anticipated ~ -146 value. Additionally, the standard DDPG agent terminates, on average, after roughly 40 steps thus halting any further data collection. After additional training, episodes 500-600 show more promising results, at least for the hybrid agent, which appears to nearly reach the global minimum in the later steps of the episode. Though the standard agent is further from the global minimum,

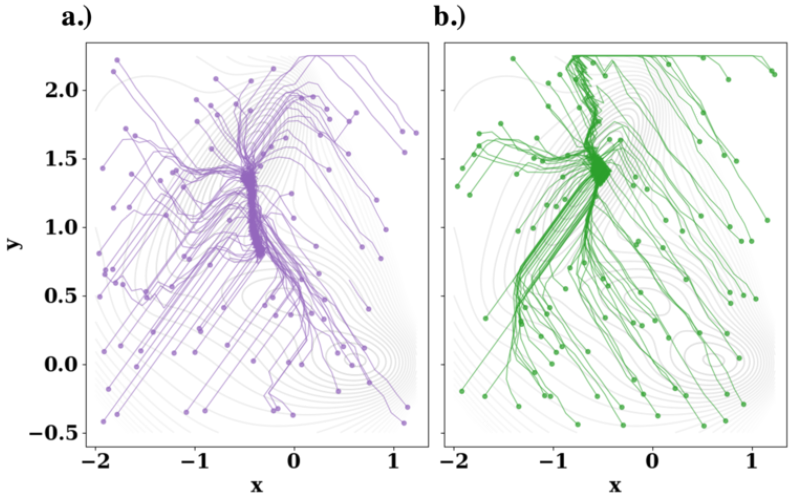


Figure 5.9. Trajectories for episodes 500-600 for the a.) standard DDPG algorithm and b.) hDDPG algorithm where the solid circle is the starting state of the episode.

it is still showing significant improvement compared to earlier episodes; however, it has yet to show improved performance over the hybrid agent, which exhibited better decision-making after only ∼300 episodes.

Once the agent has been allotted a generous period of training it begins to show improved decision making and consistently reaches positions reasonably close to the global minimum as shown in **Figure 5.9**. This is particularly exciting as a number of starting points force the agent through local minima regions. The agents, however, do not succumb to the false positive reinforcement awarded when s_t has a lower energy than s_{t-1}. This can be attributed to the advanced reward function and more specifically the implemented moving target, which aids in directing the state to the position with the lowest recorded energy.

The target's location for the DDPG and hDDPG algorithms is recorded in **Figure 5.10** to show the progression as training advances. At first glance, it is apparent that the target is infrequently updated considering the tens of thousands of steps taken

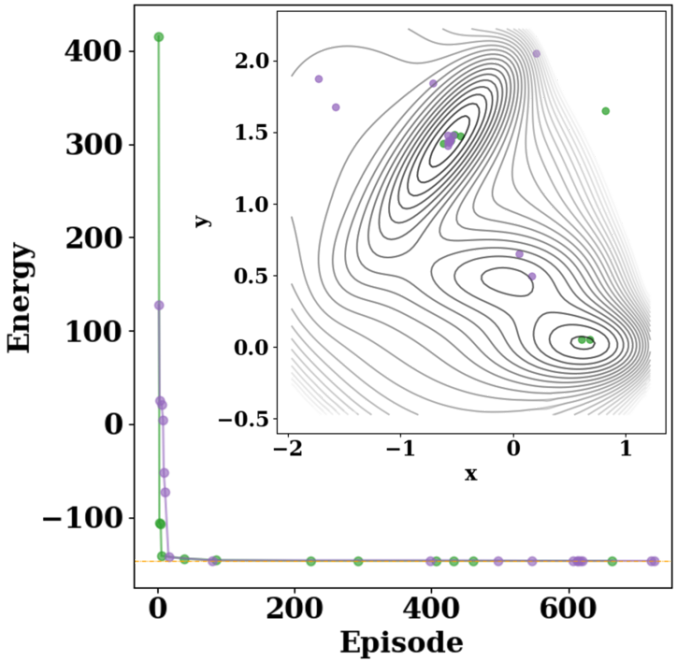


Figure 5.10. Energy and position (inset plot) of the recorded targets as DDPG (purple) and hDDPG (green) training progressed.

during training. This suggests the optimal target is located rather quickly, which is supported by the rapid decrease in energy of the target state. In fact, the energy of the target begins to converge by episode 100 in both cases. Target updates made later in training are recognized to have rather small improvements in energy and remain in the vicinity of global minimum well.

As previously observed, through convergence to a lower energy seen in **Figure 5.8**, the hDDPG agent is found to more rapidly learn the minimum region which is likely due to its ability to quickly identify a low energy target (**Figure 5.10**). Additionally, fewer suboptimal targets were identified, compared to the DDPG algorithm, which supports the collection of more valuable experiences and thus leads to overall improved training. In fact, it only took the hybrid agent 0.03% of the total steps to find

the minimum point, which is half of what the standard agent required to find a comparable target. The standard agent required more steps because the initial poor quality targets, located about the surrounding areas of the optimal minima region, led to the collection of poor quality experiences. As reflected in the data presented, there are consequences to learning with a suboptimal target and they manifest in the agent's learning rate. These targets are, however, quickly replaced by better states as additional exploration is completed allowing the agent to optimize the policy. This shows both the adaptability of the policy and the importance of identifying a good target state early on to ensure subsequent experiences collected are of higher quality. Since a target near the global minimum is found earlier in the hDDPG method, through the aid of gradient vector information, the agent is capable of selecting actions that move the state toward the optimal target sooner in training. This allows the memory to fill with more valuable experiences that are then used to train the neural network and more rapidly build an improved policy.

Evaluating the Physics-Informed Actor

The quality of the trained policy can be evaluated in the last few episodes, where an initial state has been fixed, to analyze the agent's trajectory of choice based on the experiences collected. The trajectories in **Figure 5.11** show both agents are able to identify the global minimum, however, the paths they pursue are quite different. The hDDPG agent finds trajectories that generally follow the gradient of the surface while the DDPG agent continues along the surface until there is a rise in energy and then proceeds toward the global minimum.

5.4.3 Beyond the Müller-Brown Surface

Additional PES have been developed, by modifying the Müller-Brown surface, to evaluate the hDDPG agent's ability to learn in different environments. Surface **I**, in **Figure 5.12**, modifies the variables defined in **Equation G.3** (see appendix G)

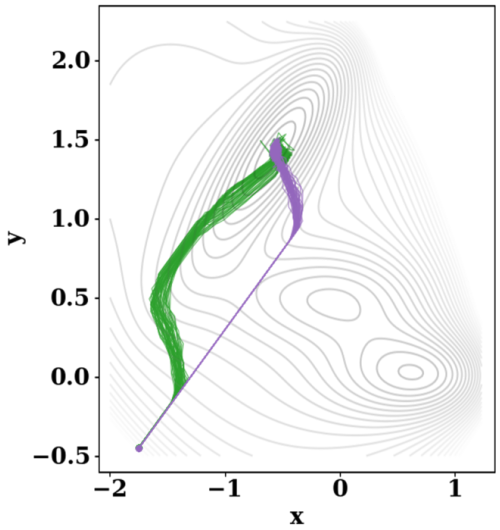


Figure 5.11. The final 50 trajectories from the DDPG (purple) and hDDPG (green) simulations.

such that a new global minimum is introduced in the lower left region. Alternatively, surface **II** is developed to hold a series of large, energetically competitive, regions in which the agent has the potential of erroneously identifying as the global minimum. The true global minimum, located in the lower right region, has been designed to take up the least area on the surface and make locating the solution more challenging for the agent.

In changing only the landscape and maintaining all previously used initial conditions, the hDDPG agent is capable of successfully locating the global minimum region in both environments. The average reward collected per episode is plotted in **Figure 5.13a** and **5.13b** to visualize the agent's ability to rapidly converge to a solution on surfaces **I** and **II**, respectively. In both instances, the agent only requires roughly 200 episodes of learning before identifying a suitable target state and consistently finding the global minimum region of the respective surface. The terminal states have been

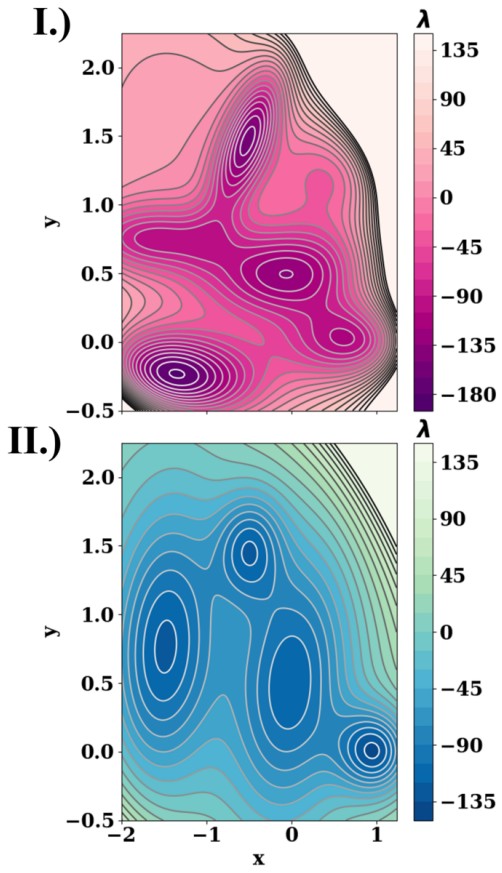


Figure 5.12. Modified Müller-Brown potential energy surfaces.

plotted in **Figure 5.13** to further support the identification of the global minimum region as the recorded terminal states begin to cluster about the minimum energy region.

5.5 Conclusions

The research presented in this study has demonstrated that implementing an advanced reward function, which compresses multiple environmental properties (i.e. the energy and corresponding position), has the potential to substantially improve the quality

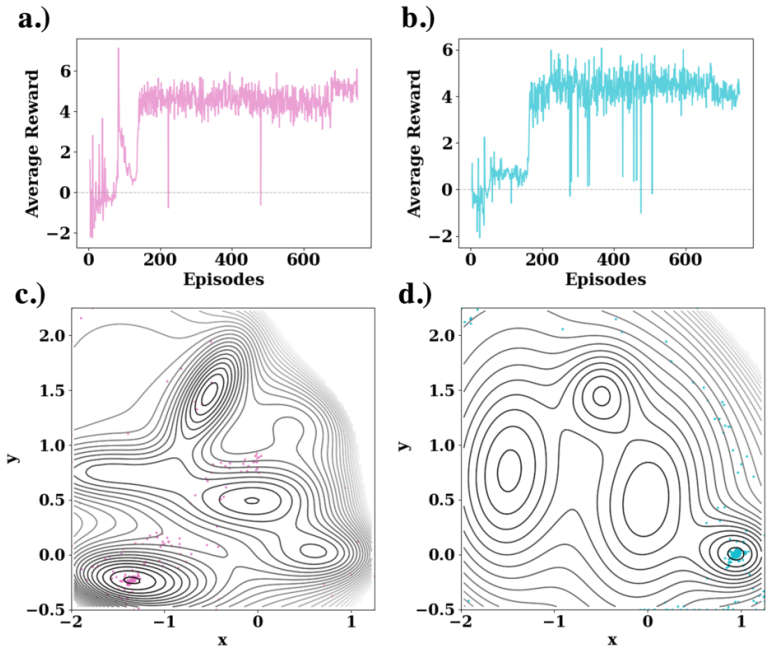


Figure 5.13. Average reward per episode in environment **I** (a) and **II** (b) and the recorded terminal states of episodes 1-750 on surfaces **I** (c) and **II** (d).

of the policy. By offering a more complete picture of the environment and defining a set target point in a single value, the agent receives a fine-tuned reward that directly relates the quality of s_t to the optimal state identified. This provides better direction to the agent and noticeably improves not only the rate of convergence, but the rate of convergence to the optimal solution. Policy optimization is further improved by incorporating physically motivated actions to guide the agent through the environment. Incorporating meaningful actions allows for valuable positive experiences to be collected in the memory during training, which ultimately increases the probability of selecting good actions in the future. When trained on positive, physically sound, experiences, the agent has a better chance of developing a more accurate picture of the environment and subsequently converging to a more optimal solution.

It is recognized that the action space explored in this work was confined to two

dimensions for demonstrative purposes, which is rather limited. This model is, however, expected to extend well to problems with increased dimensionality and greater action spaces. It should be noted that with an increased action space, the agent will require more exploration in order to devise a suitable policy; which will lead to an increase in computational cost. Additionally, a slight increase in cost is expected to come from training a more complicated neural network, however, this is anticipated to scale linearly with the system complexity.

Through this research, we hope to have shed light on the robustness of reinforcement machine learning and its applicability to complicated environments. Reinforcement machine learning has been realized as a powerful tool for solving traditionally challenging problems in all fields and most recently in the field of physical chemistry. It is our belief that these methods have the potential to dramatically accelerate research discoveries in chemistry and allow previously inaccessible research avenues to finally be pursued.

Acknowledgments

This work was partially supported by the Ultrafast Initiative of the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, through Argonne National Laboratory under Contract No. DE-AC02-06CH11357. Y.C. and X.L. acknowledge the support to develop reduced scaling computational methods from the Scientific Discovery through Advanced Computing (SciDAC) program sponsored by the Offices of Advanced Scientific Computing Research (ASCR) and Basic Energy Sciences (BES) of the U.S. Department of Energy (#DE-SC0022263).

Chapter 6

APPLICATION: SOLVING HYDRIDE POSITIONS IN COPPER-HYDRIDE NANOMATERIALS

Work from the previous chapter will be adapted in this chapter to accommodate more complex systems with increasing challenges and advanced metrics by which to learn from. A series of experimentally known copper-hydride nanomaterials will be used as benchmarking tools to evaluate the performance of the RL algorithm and illustrate the agent's versatility.

6.1 Introduction

Metal-hydride nanoclusters are useful materials for a number of different applications ranging from CO_2 reduction to hydrogen storage.[181, 126, 174] These materials have been recognized as candidates for the aforementioned applications because of their unique reactivity arising from the transition-metal hydride character.[135] Though these materials have desirable properties, oftentimes, the exact structures are inferred from spectra (i.e. X-Ray crystallography and nuclear magnetic resonance) as a consequence of spectroscopic limitations leaving the atomically-precise structure a mystery.[17, 96, 103] In order to utilize more accurate methods, such as neutron diffraction, and obtain resolvable spectra, one must grow large pure crystals, which is often quite challenging in and of itself.[44] Provided these known limitations, artificial intelligence algorithms have been explored in an effort to resolve nanocluster structures at a fraction of the cost and with comparable precision to advanced spectroscopic techniques.

Artificial intelligence methods have gained increased traction in the past few

decades because of their advanced utility and robust pattern recognition. Such methods have found applications in the field of chemistry with developments such as SchNet or the CGCNN model, both of which implement convolution neural networks (CNN) to predict physical properties.[166, 198] Recently, three-dimensional CNNs were devised and used to predict the hydride sites within copper nanoclusters with impressive accuracy.[192] The previously mentioned algorithms can be categorized under the classical supervised machine learning method, which requires prior model training and, consequently, access to sizable data sets. The work presented in this chapter seeks to forgo such overhead by adapting the recently developed hybrid-deep deterministic policy gradient (hDDPG) method to optimize hydride positions within copper nanoclusters. The hDDPG algorithm differs from the previous methods as it falls under the reinforcement machine learning regime.

Reinforcement learning algorithms offer powerful optimizing potential and incredible versatility resulting from their advanced architecture and increased model adaptability. A key component of the hDDPG architecture is the inclusion of deep neural networks, which are utilized because of their ability to solve complex, non-trivial, problems. The adaptability of RL algorithms, in general, can be attributed to the fact these models are not initially trained on a specific set of data, but rather learn from experiences within a defined environment. The environments in this work will be the experimentally observed X-ray crystal structures, supplemented by electronic structure theory, and the signals will be determined through theoretical calculations using the PM6 semi-empirical method, which was selected for computational cost considerations.

6.2 Methods

The hDDPG framework is implemented in this study and extended to nanomaterial systems where the hydride is controlled by the agent and all other atoms of the complex are held stationary, thus abiding by the approximations made in Markov

decision processes; which belong to the class of stochastic sequential processes.[150] These particular processes are unique in that a decision-maker exists to determine which actions to take in a given state. Upon selecting actions, a reward signal is received that rates the decision and a probability distribution is specified for the next state.[150]

Prior to running a simulation, the hydride-free environment is initialized using the structure obtained from X-ray crystallography, allowing the exact positions of all the heavy atoms to be known. Next, if available, the spectroscopic signatures from NMR are evaluated to determine the minimum hydride bonding order; which is then used as a hyperparameter for the RL algorithm. This experimental information is used during each step to evaluate whether the proposed structure is chemically valid by confirming the current hydride state satisfies the minimum inferred bonding order and the proposed hydride position is sufficiently far from all surrounding atoms. Should the hydride state pass these validation checks, the current state is evaluated via the GAUSSIAN16 software package using the PM6 semi-empirical method. This method provides the environment's signals using single point energy calculations that evaluate both the total energy of the current state and the corresponding gradient; the latter being contingent on the convergence of the SCF calculation. The calculated total energy is then utilized in the reward function outlined in **Algorithm 3**, which has been adapted from the Müller-Brown potential energy surface application to accommodate this problem.[124] The most notable modification made is the ability to handle states that have been deemed chemically invalid.

At the topmost level, the reward function searches for the total energy, λ_t, of the current state, then passes through the appropriate loop. If the current state has been evaluated and the total energy has been determined, the second level compares the current energy to the lowest recorded energy, λ_{low}. If the current state is found to be lower in energy, then the state is recorded as the *new* lowest energy state and a unitless reward is calculated based on the Euclidean distance. It should be noted

Algorithm 3 Reward Function

Calculate Euclidean distance, ϵ_t, between s_{low} and s_t
if λ_t exists **then**
    if $\lambda_t > \lambda_{low}$ **then**
        if $\epsilon_t = \epsilon_{t-1}$ **then**
            $r_t = 0$
        else if $\epsilon_t > \epsilon_{t-1}$ **then**
            $r_t = -1 * \epsilon_t$
        else
            $r_t = \text{clip}(1/\epsilon_t, 0, \phi)$
        end if
    else if $\lambda_t = \lambda_{low}$ **then**
        $r_t = \epsilon_t$
    else
        $r_t = \text{clip}(1/\epsilon_t, 0, \phi) * \varphi$
        Update λ_{low} with new lowest observed energy
    end if
else ▷ s_t is chemically invalid
    if λ_{low} exists **then**
        if $\epsilon_t = \epsilon_{t-1}$ **then**
            $r_t = 0$
        else if $\epsilon_t > \epsilon_{t-1}$ **then**
            $r_t = -2 * \epsilon_t$
        else
            $r_t = \text{clip}(1/\epsilon_t, 0, \varphi)$
        end if
    else
        r_t = -1 * φ
    end if
end if

that ϕ and φ are defined as 10 and 5, respectively, in this application. As such, the highest possible reward at any given state is 50. Conversely, if the state is found to be chemically invalid by the prior checks, a total energy is not calculated and the reward function proceeds through the second route where the Euclidean distances of the previous and current states are compared. This is done to evaluate whether the agent is moving toward the lowest energy state or away. If the state is found to be closer, though chemically invalid, then a positive reward should be granted to encourage the agent to continue moving in the direction that would approach the known lower energy state. Notice the increased negative reward granted, compared to the comparable condition from the first loop, when the system is found to be both invalid and moving away from the lowest energy state. This is done to discourage such trajectories since they provide little information in that they are unreasonable states and they appear to be moving in a direction that is not anticipated to net any meaningful knowledge. Finally, if the state is invalid and insufficient exploration has been experienced, up to the current time step evaluated, then the function is lacking physically meaningful information and is left granting a fixed reward.

In addition to evaluating the quality of the current structure, via the energy and ultimately the reward function, states of desirable properties are recorded in a list and returned upon completion of the training simulation. The conditions to update the list are evaluated during each step and consider both the energy and the force of the current state, if available. As a first condition, if the current state's energy is found to be lower than any of the prior list entries, then the state is considered for recording. Alternatively, if the forces are available, then the force magnitude of the hydride is evaluate and considered for recording if it falls below a preset threshold, which was set to 0.09 in this work. Once a state has been marked as a state of interest, it may be added by appending to the list or replacing an existing entry. The act of replacing a state within the list is allowed if the energy of the new state is within 0.0001 Hartrees of an existing state's energy. If this condition is met, then the lower

energy state replaces the existing state in the list. Conversely, if this criteria is not met, the state is simply appended to the list. Additionally, the act of appending may occur by considering the forces alone, should they exist. This has been explicitly allowed in order to capture states that may be considered "good" states, though fail to reflect that in the calculated total energy; which is a recognized consequence of theoretical methods. This list will be referred to as the candidate list going forward.

To ensure adequate exploration is pursued by the agent during training, a Voronoi diagram of the nanocluster is created where each atom serves as a base point. From the diagram, the vertices with nearby transition metal atoms are randomly selected and used as the initial state for the hydride to incorporate an artificial element of exploration and aid in developing a more versatile problem-solving agent. This is done for the first 80% of episodes, the last 20% of episodes randomly select a state from the existing candidate list to use as the initial state. By initializing states using candidate list entries, the agent has the opportunity to fine-tune the entries of the list and find more suitable hydride positions within the general region.

In an effort to reduce the number of seemingly unphysical states and keep the episodes from running for extended periods of time, a series of termination conditions are defined to end an episode. The conditions set are as follows:

1. The maximum number of steps have been reached
2. 60% of the maximum steps where found to be chemically invalid
3. 60% of the maximum steps failed to converge during the SCF procedure
4. The cumulative reward for the episode exceeds -75

6.3 Results & Discussion

A collection of experimentally synthesized and structurally known copper clusters are investigated in this study to demonstrate the precision that can be achieved using reinforcement machine learning.[104, 110, 43] Each structure features a single hydride and a copper core, which can be seen in **Figure 6.1**. The ligands about the core vary

in size and bulkiness to provide a variety of environment types.

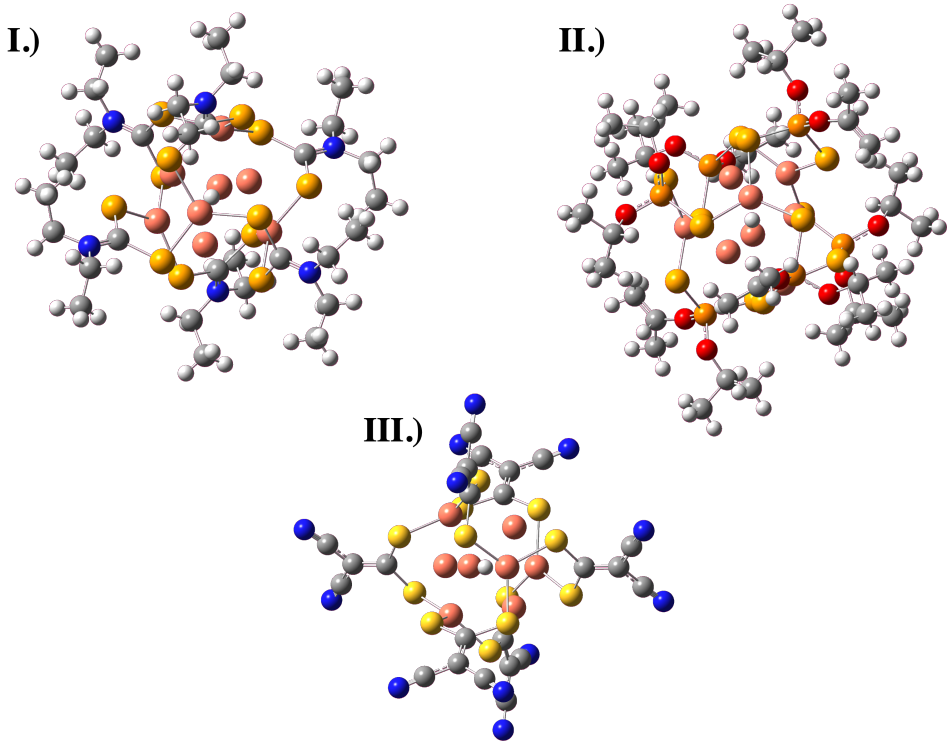


Figure 6.1. Copper nanoclusters studied in this work where white, grey, blue, yellow, pink, orange, dark orange, and red correspond to hydrogen, carbon, nitrogen, sulfur, copper, selenium, phosphorous, and oxygen, respectively.

6.3.1 Evaluation via Cumulative Rewards

The most straightforward method in which to gauge the quality of a trained agent and subsequent policy is to evaluate the accumulated rewards as a function of the episode. The total reward for each episode is presented in **Figure 6.2** to visualize the agent's progress throughout the simulation. To begin, the left most plot of **Figure 6.2** indicates the agent requires just under 200 episodes of exploration before developing a fundamental understanding of the nanomaterial. The negative rewards

in the first 100 episodes reflect the agent's poor understanding of the environment and lack of basic chemical knowledge. These episodes generally conclude early as a consequence of termination condition 2 in that many of the proposed structures find hydride locations far too close to neighboring atoms or colliding with stationary atoms of the environment. Though the states collected in the first one hundred episodes are poorly rewarded, they ultimately contribute to the steep learning curve that follows in the next 200 episodes.

As mentioned, the newfound understanding of the environment is reflected in the positive rewards collected after episode 200 and can be attributed to the increased exploration arising from randomized initial starting states. This increased exploration allows for more diverse experiences and enforces more informed learning, which encourages better decision-making. As the agent learns to take better actions, which manifests in chemically valid structures, the number of steps the episodes experience are observed to increase. This, in turn, provides more opportunities for the agent to collect rewards and increase the cumulative reward upon completion of an episode. In fact, the termination condition generally triggered from episode 200 onward is condition 1.

After the dramatic learning curve experienced between episodes 200 and 400, the remaining episodes demonstrate generally better action selection; however, a temporary decrease in reward is observed just before episode 600. This originates from a number of chemically invalid structure proposals, that arise from a desire to continue exploration. It should be noted that although these episodes netted a lower overall reward, the trajectories were not entirely ludicrous. These episodes simply experimented with different pathways in order to find a comparable terminal state. Soon after this brief detour, the plot is observed to converge to a reward of 140 indicating a suitable policy has been identified. Similarly, convergence to a reasonably high reward is found to occur in case **II**. Case **II**, however, could have benefited from additional training as the final episodes do not appear to converge but rather steadily increase.

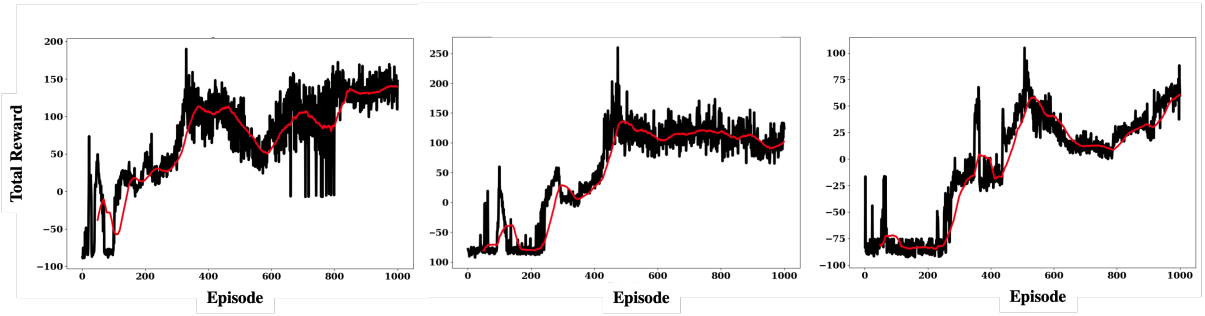


Figure 6.2. Total reward accumulated at the end of each episode for clusters **I**, **II**, and **III** (from left to right) where the red trace indicates the rolling average of the reward with a window length of 50.

6.3.2 Episodic Evaluation

To better visualize the agent's progression during training in case **I**, one can investigate the trajectories explored during episodes at varying points of the simulation. **Figure 6.3** presents the trajectories at the beginning of the simulation (left) and toward the middle of the simulation (right), which covers the region of the reward plot where the most dramatic change is observed. This is done to illustrate the improvement of the policy as more experiences are collected. It is clear that during the earlier episodes a suitable target state has yet to be identified rendering exploration seemingly unguided, as evidenced by the vast area covered and chaotic trajectories. Remarkably, by episode 300, a general point of interest appears to have been identified and is consistently pursued in the subsequent ten episodes. In the ten consecutive trajectories illustrated, the agent appears to terminate in a common area despite initializing in a number of completely unrelated states. This can be attributed to exceptional exploration that has allowed the agent to develop a practical policy and, subsequently, a good understanding of the nanocluster structure.

Though the policy appears to be optimizing, given the convergence of the final state in trajectories 300-310, it is unclear as to whether the policy is optimizing to

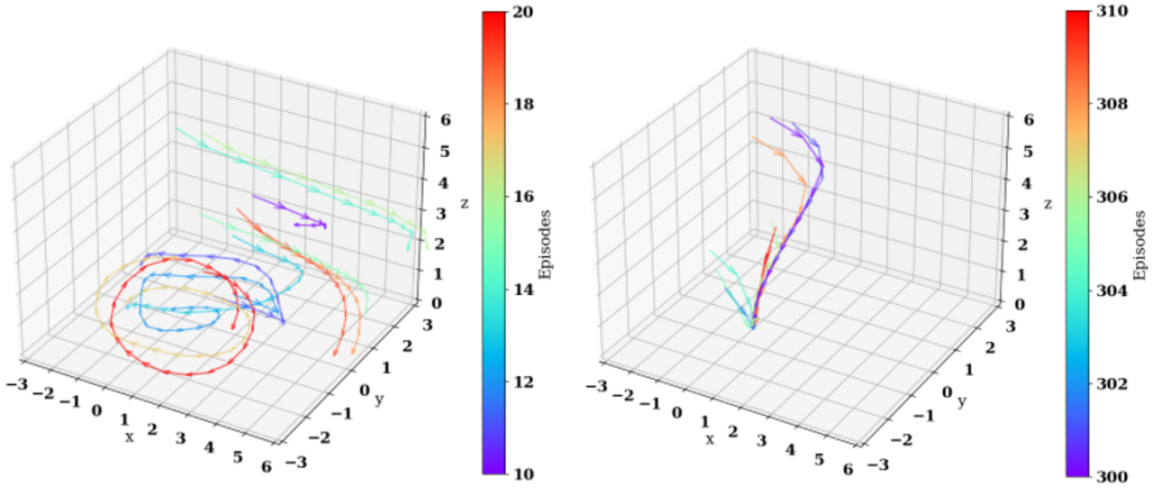


Figure 6.3. Hydride trajectory for case **I** during episodes 10 to 20 (left) and 300 to 310 (right).

the experimentally observed hydride position. To confirm this, a series of trajectories throughout the simulation can be extracted and the final states can be compared to the known hydride position. For the most straightforward analysis, a handful of episodes, with the same initial state, were analyzed and are presented in **Table 6.1**.

At the very beginning of the simulation, the agent demonstrates understandably poor decision-making as a consequence of an underdeveloped policy. Without any formal guidelines in which to follow, the agent is incapable of proposing a single valid state at this point and thus receives a large negative reward as the final state finds a hydride position nearly 7 angstroms away from the true position. This behavior is quickly corrected by episode 41 where 21 of the states visited qualified as chemically valid and the final position of the trajectory was found to be just over 0.5 angstroms away from the experimentally reported position. Little improvement is observed over the next 100 episodes in terms of distance, however, the total energy is reported to have undergone a notable decrease. Unsurprisingly, as training progresses the distance

Episode	Complete Episode			Final Step		
	Valid States	Total Reward	Episode Duration (s)	Δr (Å)	Total Energy (Hartree)	Force Mag.
2	0	-89.086	10	6.824	—	—
41	21	9.037	303	0.662	15.894	0.148
151	25	25.531	329	0.542	15.858	0.034
286	25	47.020	284	0.253	15.847	0.021
443	25	145.231	252	0.139	15.829	0.014
641	25	123.241	237	0.142	15.829	0.014
798	15	98.220	175	0.113	15.843	0.011

Table 6.1. Analysis of episodes that had the same initial state, which was recorded to be 2.766 angstroms from the experimentally observed hydride position. The distances calculated indicate the distance between the final state of the episode and the position of the experimentally observed hydride position. Finally, the force magnitude on the hydride was evaluated and the total energies for the respective states were calculated.

and energy decrease and the reward begins to significantly increase. The increase in cumulative reward indicates the ability of the agent to find regions very near the lowest recorded point and, presumably, the optimal position.

The method used for calculating the total energy, however, is the limiting factor and will ultimately dictate the quality of the agent's prediction. Such limitations are directly reflected in the results from episode 443 compared to episode 798. The former has a lower energy, but is observed to reside further from the true position when compared to the latter. It is, however, noted that the force magnitude calculated on the hydride atom of episode 798 is lower than all other reported states, despite the energy not being the minimum recorded energy. This limitation is acknowledged and addressed by updating the candidate list based on both an energy criteria and a force criteria such that states of this nature are not overlooked and excluded from post-simulation data analysis.

6.3.3 Post-Simulation Analysis

Absence of Minimum Energy State Biasing

Once the simulation has completed and training concludes, the candidate list is evaluated to identify the best positions found by the agent. Additional analysis is required to vet the resulting list and extract notable points. As a first step, the data is clustered based on their position in space in order to group proposals that fall within close proximity. From each cluster, the lowest energy state is extracted and saved. The results, after the first 80% of episodes, are recorded in **Table 6.2**, which shows the ten lowest energy states as well as the corresponding episode and step they were found. Only the first 80% are considered in this analysis to demonstrate the quality of the policy prior to biasing good initial states for the final 20% of episodes.

Episode	Step	Energy	Force Mag.	Δr
797	3	15.8277	0.0092	0.1128
325	4	15.8299	0.0158	0.2829
323	3	15.8301	0.0158	0.2792
732	4	15.8316	0.0258	0.1587
753	8	15.8359	0.0309	0.2848
282	21	15.8369	0.0315	0.3731
194	5	15.8402	0.0435	0.5469
768	5	15.8425	0.0068	0.0447
726	5	15.8427	0.0075	0.1014
689	4	15.8428	0.0075	0.1005

Table 6.2. Proposed states from the candidate list after clustering into 15 group and extracting the lowest energy state from each cluster. The energy is reported in Hartrees and Δr, reported in angstroms, is calculated between the predicted hydride and the experimentally observed hydride position.

These data further demonstrate that the level of theory used to evaluate the total energy of the structure will strongly influence the agent's ability to identify an experimentally accurate target state. Considering this, the forces are evaluated and

the force magnitude on the hydride of interest is calculated as an additional metric. This proves fruitful as the candidate list is updated considering both energy and force magnitude and therefore captures a remarkably accurate state in episode 768. Although the energy is not the lowest in the list, the force magnitude is closest to zero and the position is calculated to be 0.04 angstroms away from total agreement with the crystal structure. This builds a strong argument for using the forces as the driving metric, however, without a successfully converged SCF calculation the forces cannot be evaluated rendering an entirely force-dependent algorithm unfeasible. Instead, the forces are used as a secondary metric for determining the quality of a state, provided it has been evaluated during the semi-empirical calculation.

Though the gradient is not always readily available, it does prove highly useful when it has been evaluated. In addition to considering the forces in the reward function, they are consulted when the agent continually selects poor actions. More specifically, the gradient is used in place of NN selected actions to encourage the system to proceed in a direction that minimizes the total energy. This periodically aids in training as the forces are used for 3% of the total steps. It should be acknowledged that a large contributing factor to the minimal use of the forces during training may be due to the absence of their availability as a limitation of the method and failure to converge during the SCF procedure. Though limited, they are highly influential and ultimately support the optimization process.

Minimum Energy State Biasing

The final 20% of episodes select initial states from the candidate list in order to bias ideal regions of the PES where the equilibrium geometry is expected. In doing this and taking advantage of the agent's accumulated knowledge, one can expect to find a number of proposed structures reasonably close to the true value. **Table 6.3** reports the the difference between the proposed structures and the crystal structure hydride positions. The first few guesses for case **I** and **II** are remarkably close to the

experimentally observed structure and locate positions within 0.2 Å. The inability of case **III**, to find a comparably close position, could be a consequence of insufficient training, which is implied by the failure to converge the cumulative reward. Additionally, it was observed that although a number states were deemed chemically valid, the SCF procedures were unable to converge ultimately depriving the agent from utilizing gradient directed actions. This withheld crucial information from the agent and likely contributed to the subpar candidate structures.

	I	II	III
1	0.1105	0.1755	0.2594
2	0.1823	0.1857	0.3143
3	0.1485	0.1600	0.3221
4	0.1592	0.2716	0.3509
5	0.2445	0.4143	0.4419

Table 6.3. Proposed states from the candidate list after clustering into 15 group and extracting the 5 lowest energy state from each cluster. The proposed positions are then sorted by energy and the distances, reported in units of angstroms, are calculated relative to the respective crystal structure.

6.3.4 AI versus DFT Optimization

To illustrate the power of a well trained AI algorithm, compared to more traditional optimization methods, the initial state of episode 798, from section 6.3.2, was optimized using DFT with the CAM-B3LYP functional and basis set 6-31g(d). The optimization concluded after 10 hours and 30 minutes, however, it failed to meet the minimum SCF convergence criteria. Conversely, the agent, trained on 797 episodes, was capable of completing episode 798 in just under 3 minutes and found a reasonably accurate hydride position. The failure of DFT is likely due to the complex potential energy surface and the presences of intense energetic barriers between the initial and true equilibrium state making it difficult for DFT to locate the optimal geometry when

given a poor initial structure. This highlights the shortcomings of standard methods in that they often require user intervention and expertise when initial guesses for such calculations are exceptionally far from the optimal states. This further supports the increased robustness of the reinforcement learning methodology for solving complicated environments and identifying trends that traditional algorithms are incapable of navigating independently.

Though DFT struggles to find an optimal hydride position given the same initial state as episode 798, without additional user intervention, the final state of this episode, from the RL model, is expected to provide a suitable initial structure for optimization via DFT. After roughly 4 hours, the DFT optimization does indeed complete and identifies a position 0.102 angstroms away from the known hydride location. This improves the RL guess by $\sim$0.010 angstroms, which is a surprisingly small improvement. Perhaps additional functional tuning is required to find better agreement between the simulated and crystal structures, but that is beyond the scope of this study.

Considering the previous results, it is proposed that rather than using the RL model and DFT methods competitively, they can be used in tandem. The prior results emphasize the importance of providing an initial state near the optimal structure in order for DFT to be successful, so the lowest energy structure from the candidate list can be selected as the initial guess for DFT optimization in order to maximize the quality of the predicted state.

6.4 Conclusion

The study presented herein demonstrates the impressive problem-solving ability of strategically designed artificial intelligence algorithms. It is shown that without sufficient knowledge of where the hydride resides within the nanocluster, the RL agent is able to successfully learn the general structure, develop a basic understanding of chemistry, and outperform traditional DFT calculations when given initial guesses

sufficiently far from the equilibrium geometry. Additionally, a well trained model is found to provide comparable results to DFT calculations in a fraction of the time when given objectively good initial states to optimize.

Though the agent is capable of finding a number of excellent candidate state, it must be recognized that the algorithm is only as accurate as the level of theory used. Additionally, candidate states will tend toward states with minimum energies defined by the semi-empirical method (used in this study), which has the tendency to provide positions that disagree with experiment; as shown in section 6.3.2. Improvements on this can be done by implementing strict chemistry rules without sacrificing the collection of experiences and improving the reward and update functions to prioritize the force magnitudes when they exist. Finally, this method is capable of handling a single hydride, which is very limiting in that many systems of interest are designed to hold a higher concentration of hydrides. As such, further method development is necessary to accommodate cases of this motif.

To conclude, this work has aimed to design and support a foundation for further research and advancement in AI development for chemistry applications. The methods introduced here are intended to provide a means in which to identify structurally sound hydride positions within large materials that are otherwise unapproachable to standard workflows as a consequence of their immense size and need for an extraordinary amount of computation resources.

Chapter 7

DEVELOPING A MULTI-AGENT FRAMEWORK FOR OPTIMIZATION METAL-HYDRIDE COMPLEXES

The previous chapter laid the groundwork for the methodology used in this chapter, which is intended to optimize materials with more than one unknown hydride position. As before, the algorithm used here will build on that which was introduced in chapter 5, however, instead of using a single actor, multiple actors will be initialized to solve the problem of interest. The data presented herein are preliminary results and demonstrate the progress of ongoing work.

7.1 Motivation

As mentioned in the previous chapter, copper metal-hydride materials have various application including the reduction of CO_2 via catalysis mechanisms, which is significant for emission control and reducing unnecessary atmospheric damage.[41] Based on reported DFT calculations, the primary contributors to the reduction process were found to be both the surface metals as well as the hydride ligands of the system.[181] Considering this, in order to develop efficient catalysts the materials should hold an increased number of sites in which this mechanism can proceed; thus requiring materials with both various surface metal sites and hydride ligands available for interactions. The work in this chapter will fixate on the desire to increase the number of hydrides within the system and focus on a means in which to optimize the overall structures of the materials. Specifically, the algorithm from chapter 6 will be adapted to accommodate systems with more than one hydride.

7.2 Theory

By considering systems with more than one moving component, the observation space is naturally expected to increase, which consequently increases the computational cost. A multi-agent framework is proposed to adjust the computational load of each agent, however, this introduces new challenges that were largely absent in the previous single-agent application; one of which being the element of non-stationarity. Non-stationarity arises from multiple moving parts within a single environment rendering the traditional MDP invalid for this problem; instead, the problem must be reframed as a Markov (stochastic) game.

7.2.1 Markov Games

To properly accommodate a multi-agent framework, one must first redefine the problem as a Markov game (MG), which is merely an extension of the classic Markov decision process. This is necessary as the actions selected from all agents in the system strongly contribute to the state dynamics.[22] As such, the environment from each agent's perspective becomes non-stationary. At each step the MG considers the complete set of actions and rewards from all agent giving rise to the modified tuple:

$$MG = \langle N, S, A, R, p \rangle \tag{7.1}$$

where the only new component of the tuple is N, which represents the set of agents. Additionally, A is redefined to be a set of joint actions, $\Pi_{i=1}^{n} A_i$, thus becoming a joint action space where n is the total number of agents.[119, 22] This requires a modification to R and p, as they both rely on the joint-actions. The reward function becomes, $S \times A \rightarrow R^n$, and the transition probability becomes, $p(s'|s, \overrightarrow{a})$, where $\overrightarrow{a}$ is the joint-action.[119, 107].

7.3 Methods

As before, the problem investigated will operate under the model-free regime for ease of implementation and to avoid unintentional biases from inadequately defining the model. This renders the algorithm both goal-oriented and increasingly more versatile as the former point allows for adaptability to different environments.[22]

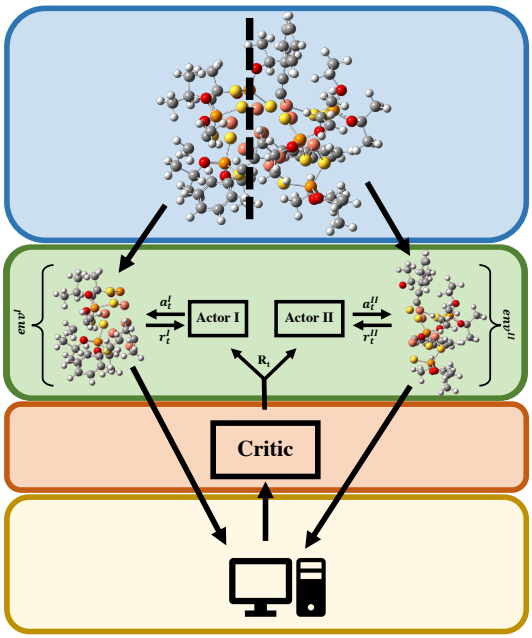


Figure 7.1. General schematic of the multi-agent framework where a_t^x and r_t^x represent the action and reward at step t for the respective agent, denoted by the superscript x. Additionally, R_t, indicates the global reward determined after the complete system has been evaluated via GUASSIAN16.

To mitigate the inevitable increase in computational cost, the environment is segmented via K-means clustering where k is set equal to the number of hydrides that are expected to reside within the nanocluster and the atoms of the cluster are partitioned into k clusters. For instance, a system with two hydrides will be split into two cluster, or sub-environments, as indicated by the dotted line in the blue portion highlighted of **Figure 7.1**. Each resulting cluster will have a designated actor that

selects actions independent of one another, which is illustrated in the green highlighted section of **Figure 7.1**. Additionally, each actor will receive a minor reward, r_t, at this phase that rates the action based on crude chemical criteria. The local reward is based on whether or not the local environment, with the proposed hydride position, meets the following conditions: a.) the hydride is not colliding with surrounding atoms and b.) a transition metal is within close proximity. The hydride bonding order is of little use at this point because the complete environment is not considered during this check. If both agents have proposed seemingly valid states, the sub-environments are joined and a second check is completed to confirm the chemical validity of the structure. This evaluation extends on the previous conditions and asserts the hydrides are not colliding and they meet the minimum bond order define by the set hyperparameter (derived from NMR spectra). Should the state prove adequate and the second check indicates the structure has maintained chemical validity upon combination, a single point energy calculation is completed to determined a global reward using the function outlined in **Algorithm 3** of the previous chapter. To summarize, the actors will work in a cooperative capacity, as opposed to competitive, and seek to collectively optimize the total reward. Each actor will execute actions locally and receive indirect information of the global environment through the single critic that communicates to all actors.

7.4 Results & Discussion

The preceding discussion analyzes two different simulations, the first utilizes a purely NN approach whilst the other consults the gradient when the agent-selected actions increase the energy of the system for three or more consecutive steps. The quality of these actions are illustrated in **Figure 7.2** by plotting the number of chemically valid states relative to the total number of steps taken per episode. It should be noted that the first simulation was allotted 1500 episodes while the physics-informed simulation had only 1000 episodes to train. Remarkably, both simulations rapidly find trajec-

tories that encourage the hyrides toward the center of the Cu cluster, which is the general region they are expected to reside. However, it is surprising the agent devoid of physics appeared to find more chemically valid structures throughout the simulation and concludes with a series of episodes that find a number of viable structures.

Alternatively, the physics-informed agent demonstrates poor understanding and a lack of learning within the given timeframe. Upon further investigation, it was found that this agent rarely utilized the gradient due to its limited availability as a consequence of either chemically invalid structures or failure to converge during the SCF procedure. Considering the lack of gradient accessibility, it is reasonable to consider this simulation largely NN driven, which explains the resemblance between the results of **Figure 7.2a** and **b** for the comparable regions.

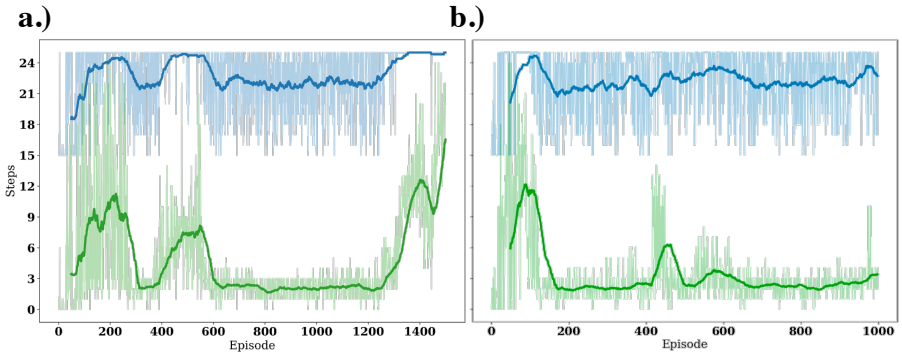


Figure 7.2. Total steps (blue) and number of chemically valid structures (green) are illustrated in the plots above for a simulation run in the absence of physically-motivated actions (a) and with physically-motivated actions (b). The bold solid lines indicate the rolling mean value using a window size of 50.

Although the plot indicates a large number of proposed structures are chemically invalid based on the criteria defined in the "Methods" section, that is not to say the structures are completely meaningless. A more in depth analysis shows the actors are capable of quickly identifying the central region of the cluster as an ideal location for

the hydrides and continually seek this area in subsequent episodes. In fact, it seems that many of these structures fail the second validity check because the hydrides are colliding with one another. This is due to the overlapping area between the sub-environments that both actors have access to. This inevitably leads to orientation complications as both actors can direct their respective hydrides to identical locations. Provided this, extended periods of training are necessary allowing the actors to realize this obstacle and identify a solution. The pure NN simulation appears to have had sufficient training allowing the agent to address this challenge and identify sensible structures during the later episodes.

Toward the end of the NN pure simulation, the individual actors are observed to work together much more effectively and demonstrate suitable spatial awareness, which is manifested in the high volume of chemically valid structures. At this point in training, the proposed hydride positions are sufficiently far apart and generally localized in the proper structural region allowing the actors to fine-tune their actions. The increase in valid structures after episode 1200 is in part due to the fact the initial states are sampled from the candidate list. In doing this, regions in which the hydrides are expected to reside are biased to improve the quality of the agent-proposed structures.

Episode	Step	Energy	H_1 Δr	H_2 Δr
1353	3	4.5792	0.5204	0.7802
208	9	4.5869	0.7824	0.2366
498	20	4.5874	0.7883	0.8726
1499	10	4.5876	1.1424	0.8783
522	17	4.5888	0.8949	1.1627

Table 7.1. The first five lowest energy states, from the proposed structures, extracted during post-simulation data analysis via K-means clustering where k is set equal to 15. The energies are reported in Hartrees and the distances, Δr (recorded in angstroms), are calculated between the predicted hydride positions and the hydride positions from the reported crystal structures.

Upon completion of the simulation, it is observed that after 1500 episodes the lowest recorded energy is found to be 4.5792 Hartrees. Following the same post-simulation data analysis practiced in chapter 6, the 5 lowest energy states are recorded in **Table 7.1**. Naturally the lowest energy state was found later in the simulation as these episodes initialized with guesses near the anticipated equilibrium state. This allows for a new, lower energy, state to be more easily identified.

7.5 Conclusion

It should be reiterated that this work is in the early phases and has been shown here to illustrate the progress made by advancing the algorithm from a single-agent to a multi-agent framework. Further research is required to optimize this methodology and encourage increased agent performance. For instance, designing a more elegant local reward function is expected to significantly improve the decision maker's ability to select sensible actions by providing a more informative rating.

Preliminary results have been presented in this chapter to introduce ongoing work focused around optimizing complex materials through advanced AI algorithms. Though in the early phases, the data presented shows encouraging results and has the potential of providing alternative methods for resolving nanomaterial structures. Through this research, we hope to support experimentalist in designing and analyzing novel materials for practical applications.

BIBLIOGRAPHY

[1] Antoni Aguilar-Mogas, Xavier Giménez, and Josep Maria Bofill. Finding reaction paths using the potential energy as reaction coordinate. 128(10):104102, 2008.

[2] Kabir Ahuja, William H. Green, and Yi-Pei Li. Learning to Optimize Molecular Geometries Using Reinforcement Learning. 17(2):818–825, 2021.

[3] Oguzhan Alagoz, Heather Hsu, Andrew J Schaefer, and Mark S Roberts. Markov Decision Processes: A Tool for Sequential Decision Making Under Uncertainty. 30(4):474–483, 2010.

[4] Kirsty M. Anderson and A. Guy Orpen. On the relative magnitudes of cis and trans influences in metal complexes. pages 2682–2683, 2001.

[5] Andrew J. Atkins and Leticia González. Trajectory Surface-Hopping Dynamics Including Intersystem Crossing in $[Ru(bpy)_3]^{2+}$. 8(16):3840–3845, aug 2017.

[6] Peter Atkins and Ronald Friedman. *Molecular Quantum Mechanics*. Oxford University Press, 2011.

[7] K. Azizzadenesheli, E. Brunskill, and A. Anandkumar. Efficient exploration through bayesian deep q-networks. In *2018 Information Theory and Applications Workshop (ITA)*, pages 1–9, 2018.

[8] E. J. Baerends and O. V. Gritsenko. A Quantum Chemical View of Density Functional Theory. *J. Phys. Chem. A*, 101(30):5383–5403, 1997.

[9] Ajit Banerjee, Noah Adams, Jack Simons, and Ron Shepard. Search for Stationary Points on Surfaces. 89(1):52–57, 1985.

[10] Andrew G. Barto, Steven J. Bradtke, and Satinder P. Singh. Learning to Act Using Real-Time Dynamic Programming. 72(1):81–138, 1995.

[11] Axel D. Becke. Density-Functional Thermochemistry. III. The Role of Exact Exchange. 98(7):5648–5652, apr 1993.

[12] Jörg Behler and Michele Parrinello. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. 98:146401, Apr 2007.

[13] Richard Bellman. *Dyanmic Programming*. Princeton University Press, 1957. https://www.gwern.net/docs/statistics/decision/1957-bellman-dynamicprogramming.pdf.

[14] Dimitri P. Bertsekas. *Reinforcement Learning and Optimal Control*. Athena Scientific, 2019. https://web.mit.edu/dimitrib/www/$RL_1 - SHORT - INTERNET - POSTED.pdf$.

[15] J. C. Boettger. Approximate Two-Electron Spin-Orbit Coupling Term For Density-Functional-Theory DFT Calculations Using The Douglas-Kroll-Hess Transformation. 62:7809–7815, 2000.

[16] Silvia Bonfanti and Walter Kob. Methods to Locate Saddle Points in Complex Landscapes. 147(20):204104, 2017.

[17] Megalamane S. Bootharaju, Raju Dey, Lieven E. Gevers, Mohamed N. Hedhili, Jean-Marie Basset, and Osman M. Bakr. A New Class of Atomically Precise, Hydride-Rich Silver Nanoclusters Co-Protected by Phosphines. 138(42):13770–13773, 2016.

[18] Delphine Brissy, Myriem Skander, Pascal Retailleau, Gilles Frison, and Angela Marinetti. Platinum(ii) complexes featuring chiral diphosphines and n-heterocyclic carbene ligands: Synthesis and evaluation as cycloisomerization catalysts. 28:140–151, 2009.

[19] Samantha E. Brown-Xu, Matthew S. J. Kelley, Kelly A. Fransted, Arnab Chakraborty, George C. Schatz, Felix N. Castellano, and Lin X. Chen. Tunable Excited-State Properties and Dynamics as a Function of Pt–Pt Distance in Pyrazolate-Bridged Pt(II) Dimers. 120:543–550, 2016.

[20] Broyden, C. G. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA J. Appl. Math.*, 6(1):76–90, 03 1970.

[21] Lucian Busoniu, Robert Babuska, Bart De Schutter, and Damien Ernst. *Reinforcement Learning and Dynamic Programming Using Function Approximators*. CRC press, 2017.

[22] Lorenzo Canese, Gian Carlo Cardarilli, Luca Di Nunzio, Rocco Fazzolari, Daniele Giardino, Marco Re, and Sergio Spanò. Multi-Agent Reinforcement

Learning: A Review of Challenges and Applications. *Applied Sciences*, 11(11), 2021.

[23] Klaus Capelle. A Bird's-Eye View of Density-Functional Theory. 2002.

[24] Maya Chaaban, Yu-Chieh Chi, Michael Worku, Chenkun Zhou, Haoran Lin, Sujin Lee, Azza Ben-Akacha, Xinsong Lin, Chen Huang, and Biwu Ma. Thiazol-2-Thiolate-Bridged Binuclear Platinum(II) Complexes with High Photoluminescence Quantum Efficiencies of up to Near Unity. 59:13109–13116, 2020.

[25] Maya Chaaban, Chenkun Zhou, Haoran Lin, Brandon Chyi, and Biwu Ma. Platinum(II) Binuclear Complexes: Molecular Structures, Photophysical Properties, and Applications. 7:5910, 2019.

[26] Jeng-Da Chai and Martin Head-Gordon. Long-Range Corrected Hybrid Density Functionals with Damped Atom–Atom Dispersion Corrections. 10(44):6615–6620, 2008.

[27] Arnab Chakraborty, Joseph C. Deaton, Alexandre Haefele, and Felix N. Castellano. Charge-Transfer and Ligand-Localized Photophysics in Luminescent Cyclometalated Pyrazolate-Bridged Dinuclear Platinum(II) Complexes. 32:3819–3829, 2013.

[28] Arnab Chakraborty, James E. Yarnell, Roger D. Sommer, Subhangi Roy, and Felix N. Castellano. Excited-State Processes of Cyclometalated Platinum(II) Charge-Transfer Dimers Bridged by Hydroxypyridines. 57:1298–1310, 2018.

[29] Arnab Chakraborty, James E. Yarnell, Roger D. Sommer, Subhangi Roy, and Felix N. Castellano. Excited-State Processes of Cyclometalated Platinum(II) Charge-Transfer Dimers Bridged by Hydroxypyridines. 0(57):1298–1310, 2018.

[30] Alan Kwun-Wa Chan and Vivian Wing-Wah Yam. Precise Modulation of Molecular Building Blocks from Tweezers to Rectangles for Recognition and Stimuli-Responsive Processes. 51(12):3041–3051, 2018.

[31] Hsiu-Yi Chao, Wei Lu, Yanqin Li, Michael C. W. Chan, Chi-Ming Che, Kung-Kai Cheung, and Nianyong Zhu. Organic Triplet Emissions of Arylacetylide Moieties Harnessed through Coordination to [Au(PCy3)]+. Effect of Molecular Structure upon Photoluminescent Properties. 124(49):14696–14706, 2002. PMID: 12465981.

[32] Wei-Chih Chen, Pi-Tai Chou, and Yuan-Chung Cheng. Low Internal Reorganization Energy of the Metal–Metal-to-Ligand Charge Transfer Emission in Dimeric Pt(II) Complexes. 123(16):10225–10236, 2019.

[33] Aurélia Chenu and Gregory D. Scholes. Coherence in energy transfer and photosynthesis. 66(1):69–96, 2015. PMID: 25493715.

[34] Sung Cho, Michael W. Mara, Xianghuai Wang, Jenny V. Lockard, Aaron A. Rachford, Felix N. Castellano, and Lin X. Chen. Coherence in MetalMetal-to-Ligand-Charge-Transfer Excited States of a Dimetallic Complex Investigated by Ultrafast Transient Absorption Anisotropy. 115:3990–3996, 2011.

[35] Yang-Jin Cho, So-Yoen Kim, Chang Min Choi, Nam Joon Kim, Chul Hoon Kim, Dae Won Cho, Ho-Jin Son, Chyongjin Pac, and Sang Ook Kang. Photophysics and Excited-State Properties of Cyclometalated Iridium(III)Platinum(II) and Iridium(III)Iridium(III) Bimetallic Complexes Bridged by Dipyridylpyrazine. 56:5305–5315, 2017.

[36] Zdenek Chval, Miroslav Sip, and Jaroslav V. Burda. The trans effect in square-planar platinum(II) complexes—A density functional study. 29(14):2370–2381, 2008.

[37] Leander M. Cinninger, Lyndon D. Bastatas, Yulong Shen, Bradley J. Holliday, and Jason D. Slinker. Luminescent Properties of a 3,5-Diphenylpyrazole Bridged Pt(II) Dimer. 48:9684–9691, 2019.

[38] Joel Creutzberg and Erik Donovan Hedegård. Investigating the Influence of Relativistic Effects on Absorption Spectra for Platinum Complexes with Light-Activated Activity Against Cancer Cells. 22:27013–27023, 2020.

[39] Pál Császár and Péter Pulay. Geometry Optimization by Direct Inversion in the Iterative Subspace. 114:31–34, 1984.

[40] Scott D. Cummings and Richard Eisenberg. Tuning the Excited-State Properties of Platinum(II) Diimine Dithiolate Complexes. 118:1949–1960, 1996.

[41] Steven J. Davis, Ken Caldeira, and H. Damon Matthews. Future CO_2 Emissions and Climate Change from Existing Energy Infrastructure. *Science*, 329(5997):1330–1333, 2010.

[42] Marc Deisenroth and Carl Rasmussen. Pilco: A model-based and data-efficient approach to policy search. pages 465–472, 01 2011.

[43] Rajendra S Dhayal, Jian-Hong Liao, Hsing-Nan Hou, Ria Ervilita, Ping-Kuei Liao, and CW Liu. Copper (I) Diselenocarbamate Clusters: Synthesis, Structures and Single-Source Precursors for Cu and Se Composite Materials. *Dalton Transactions*, 44(12):5898–5908, 2015.

[44] Rajendra S. Dhayal, Jian-Hong Liao, Yan-Ru Lin, Ping-Kuei Liao, Samia Kahlal, Jean-Yves Saillard, and C. W. Liu. A Nanospheric Polyhydrido Copper Cluster of Elongated Triangular Orthobicupola Array: Liberation of H2 from Solar Energy. *JACS*, 135(12):4704–4707, 2013.

[45] J. L. Doob. The Brownian Movement and Stochastic Equations. *Ann. Math.*, 43(2):351–369, 1942.

[46] Andreas Dreuw and Martin Head-Gordon. Single-Reference Ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chemical reviews*, 105(11):4009–4037, 2005.

[47] Jingliang Duan, Yang Guan, Shengbo Eben Li, Yangang Ren, Qi Sun, and Bo Cheng. Distributional Soft Actor-Critic: Off-Policy Reinforcement Learning for Addressing Value Estimation Errors. pages 1–15, 2021.

[48] K. G. Dyall and K. Faegri Jr. *Introduction to Relativistic Quantum Chemistry*. Oxford University Press, 2007.

[49] Kenneth G. Dyall and Knut Fægri, Jr. *Introduction to Relativistic Quantum Chemistry*. Oxford University Press, 2007.

[50] Frank Eckert, Peter Pulay, and Hans-Joachim Werner. Ab Initio Geometry Optimization for Large Molecules. 18(12):1473–1483, 1997.

[51] Franco Egidi, Joshua J. Goings, Michael J. Frisch, and Xiaosong Li. Direct Atomic-Orbital-Based Relativistic Two-Component Linear Response Method for Calculating Excited-State Fine Structures. 12(8):3711–3718, aug 2016.

[52] Franco Egidi, Shichao Sun, Joshua J. Goings, Giovanni Scalmani, Michael J. Frisch, and Xiaosong Li. Two-Component Non-Collinear Time-Dependent Spin Density Functional Theory for Excited State Calculations. 13(6):2591–2603, 2017.

[53] M. A. El-Sayed. Spin—Orbit Coupling and the Radiationless Processes in Nitrogen Heterocyclics. 38(12):2834–2838, 1963.

[54] Logan Engstrom, Andrew Ilyas, Shibani Santurkar, Dimitris Tsipras, Firdaus Janoos, Larry Rudolph, and Aleksander Madry. Implementation matters in deep rl: A case study on ppo and trpo. In *International Conference on Learning Representations*, 2020.

[55] Ödön Farkas and H Bernhard Schlegel. Methods for Optimizing Large Molecules. Part III. An Improved Algorithm for Geometry Optimization Using Direct Inversion in the Iterative Subspace (GDIIS). 4(1):11–15, 2002.

[56] F. Fischer, M. Rovatsos, and G. Weiss. Hierarchical reinforcement learning in communication-mediated multiagent coordination. pages 1334–1335, 2004.

[57] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian˜16 Revision C.01, 2016. Gaussian Inc. Wallingford CT.

[58] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox. Gaussian Development Version Revision I.11+.

[59] MJ Frisch, GW Trucks, HB Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, GA Petersson, Hiroshi Nakatsuji, et al. Gaussian 16, 2016.

[60] Garrett B. Goh, Nathan O. Hodas, and Abhinav Vishnu. Deep Learning for Computational Chemistry. *J. Comp. Chem.*, 38(16):1291–1307, 2017.

[61] Joshua J. Goings, Feizhi Ding, and Xiaosong Li. Self-consistent field using direct inversion in iterative subspace method and quasi-newton vectors. 68:77–86, 2014.

[62] Joshua J. Goings, Joseph M. Kasper, Franco Egidi, Shichao Sun, and Xiaosong Li. Real Time Propagation of the Exact Two Component Time-Dependent Density Functional Theory. 145(10):104107, 2016.

[63] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[64] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

[65] Abhijit Gosavi. Reinforcement learning: A tutorial survey and recent advances. *INFORMS J. Comput.*, 21(2):178–192, 2009.

[66] G. Granucci, M. Persico, and A. Toniolo. Direct Semiclassical Simulation of Photochemical Processes with Semiempirical Wave Functions. 114(24):10608–10615, jun 2001.

[67] Ivo Grondman, Lucian Busoniu, Gabriel A. D. Lopes, and Robert Babuska. A Survey of Actor-Critic Reinforcement Learning: Standard and Natural Policy Gradients. 42(6):1291–1307, 2012.

[68] Fabien Gutierrez, Christine Tedeschi, Laurent Maron, Jean-Pierre Daudey, Romuald Poteau, Joëlle Azema, Pierre Tisnès, and Claude Picard. Quantum Chemistry-Based Interpretations on the Lowest Triplet State of Luminescent Lanthanides Complexes. Part 1. Relation Between the Triplet State Energy of Hydroxamate Complexes and Their Luminescence Properties. pages 1334–1347, 2004.

[69] Kristoffer Haldrup, Asmus O. Dohn, Megan L. Shelby, Michael W. Mara, Andrew B. Stickrath, Michael R. Harpham, Jier Huang, Xiaoyi Zhang, Klaus B. Møller, Arnab Chakraborty, Felix N. Castellano, David M. Tiede, and Lin X.

Chen. Butterfly Deformation Modes in a Photoexcited Pyrazolate-Bridged Pt Complex Measured by Time-Resolved X-Ray Scattering in Solution. 120:7475–7483, 2016.

[70] Tracy P. Hamilton and Péter Pulay. Direct Inversion in the Iterative Subspace (DIIS) Optimization of Open-Shell, Excited-State, and Small Multiconfiguration SCF Wave Functions. 84:5728, 1986.

[71] Mingu Han, Yu Tian, Zhao Yuan, Lei Zhu, and Biwu Ma. A Phosphorescent Molecular "Butterfly" that undergoes a Photoinduced Structural Change allowing Temperature Sensing and White Emission. 53(41):10908–10912, 2014.

[72] Mance E Harmon and Stephanie S Harmon. Reinforcement learning: A tutorial. 1997.

[73] P. Jeffrey Hay and Willard R. Wadt. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for Potassium to Gold Including the Outermost Core Orbitals. *J. Chem. Phys.*, 82:299–310, 1985.

[74] Nicolas Heess, Dhruva TB, Srinivasan Sriram, Jay Lemmon, Josh Merel, Greg Wayne, Yuval Tassa, Tom Erez, Ziyu Wang, S. M. Ali Eslami, Martin A. Riedmiller, and David Silver. Emergence of locomotion behaviours in rich environments. *CoRR*, abs/1707.02286, 2017.

[75] Cheuk-Lam Ho, Wai-Yeung Wong, Bing Yao, Zhiyuen Xie, Lixiang Wang, and Zhenyang Lin. Synthesis, Characterization, Photophysics and Electrophosphorescent Applications of Phosphorescent Platinum Cyclometalated Complexes with 9-arylcarbazole Moieties. 694(17):2735–2749, 2009.

[76] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.

[77] Brooke E. Husic and Vijay S. Pande. Markov State Models: From an Art to a Science. 140(7):2386–2396, 2018.

[78] Miroslaŵ Ilias and Trond Saue. An Infinite-Order Relativistic Hamiltonian by a Simple One-Step Transformation. 126:064102, 2007.

[79] Hershel Jude, Jeanette A. Krause Bauer, and William B. Connick. Synthesis, Structures, and Emissive Properties of Platinum(II) Complexes with a Cyclometallating Aryldiamine Ligand. 41:2275–2281, 2002.

[80] Hershel Jude, Jeanette A. Krause Bauer, and William B. Connick. Tuning the Electronic Structures of Platinum(II) Complexes with a Cyclometalating Aryldiamine Ligand. 43:725–733, 2004.

[81] Hershel Jude, Jeanette A. Krause Bauer, and William B. Connick. Luminescent Platinum(II) Dimers with a Cyclometallating Aryldiamine Ligand. 44:1211–1220, 2005.

[82] Sham M. Kakade. A Natural Policy Gradient. 14, 2001.

[83] Jan Kalinowski, Valeria Fattori, Massimo Cocchi, and J. A. Gareth Williams. Light-Emitting Devices Based on Organometallic Platinum Complexes as Emitters. 255:2401–2425, 2011.

[84] Joseph M. Kasper, Patrick J. Lestrange, Torin F. Stetina, and Xiaosong Li. Modeling $L_{2,3}$-Edge X-ray Absorption Spectroscopy with Real-Time Exact Two-Component Relativistic Time-Dependent Density Functional Theory. 14(4):1998–2006, 2018.

[85] Pyosang Kim, Matthew S. Kelley, Arnab Chakraborty, Nolan L. Wong, Richard P. Van Duyne, George C. Schatz, Felix N. Castellano, and Lin X. Chen. Coherent Vibrational Wavepacket Dynamics in Platinum(II) Dimers and Their Implications. 122:14195–14204, 2018.

[86] Pyosang Kim, Andrew J. S. Valentine, Subhangi Roy, Alexis W. Mills, Arnab Chakraborty, Felix N. Castellano, Xiaosong Li, and Lin X. Chen. Ultrafast Excited-State Dynamics of Photoluminescent Pt(II) Dimers Probed by a Coherent Vibrational Wavepacket. 12(29):6794–6803, 2021.

[87] Pyosang Kim, Andrew J. S. Valentine, Subhangi Roy, Alexis W. Mills, Arnab Chakraborty, Felix N. Castellano, Xiaosong Li, and Lin X. Chen. Ultrafast Excited-State Dynamics of Photoluminescent Pt(II) Dimers Probed by a Coherent Vibrational Wavepacket. 12(29):6794–6803, 2021.

[88] W. Kohn, A. D. Becke, and R. G. Parr. Density Functional Theory of Electronic Structure. *J. Phys. Chem.*, 100(31):12974–12980, 1996.

[89] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[90] Vijay R Konda and John N Tsitsiklis. Actor-Critic Algorithms. *Adv. Neural Inf. Process. Syst.*, 12, 1999.

[91] Lukas Konecny, Marius Kadek, Stanislav Komorovsky, Olga L. Malkina, Kenneth Ruud, and Michal Repisky. Acceleration of Relativistic Electron Dynamics by Means of X2C Transformation: Application to the Calculation of Nonlinear Optical Properties. 12(12):5823–5833, Dec 2016.

[92] Tamami Koshiyama, Ai Omura, and Masako Kato. Redox-Controlled Luminescence of a Cyclometalated Dinuclear Platinum Complex Bridged with Pyridine-2-Thiolate Ions. 33(10):1386–1387, 2004.

[93] Werner Kutzelnigg and Wenjian Liu. Quasirelativistic Theory Equivalent to Fully Relativistic Theory. 123(24):241102, dec 2005.

[94] Siu-Wai Lai, Michael Chi-Wang Chan, Tsz-Chun Cheung, Shie-Ming Peng, and Chi-Ming Che. Probing d8d8 Interactions in Luminescent Mono- and Binuclear Cyclometalated Platinum(II) Complexes of 6-Phenyl-2,2‘-bipyridines. 38(18):4046–4055, 1999.

[95] Gianluca Levi, Mátyás Pápai, Niels E. Henriksen, Asmus O. Dohn, and Klaus B. Møller. Solution Structure and Ultrafast Vibrational Relaxation of the Pt-POP Complex Revealed by ΔSCF-QM/MM Direct Dynamics Simulations. 122(13):7100–7119, apr 2018.

[96] Jiaye Li, Jonathan M. White, Roger J. Mulder, Gavin E. Reid, Paul S. Donnelly, and Richard A. J. O'Hair. Synthesis, Structural Characterization, and Gas-Phase Unimolecular Reactivity of Bis(diphenylphosphino)amino Copper Hydride Nanoclusters $[Cu_3(X)(\mu_3 - H)((PPh_2)_2NH)_3](BF_4), WhereX = \mu_2 - Clandμ_3 - BH_4$. *Inorg, Chem,*, 55(19):9858–9868, 2016.

[97] Kai Li, Glenna So Ming Tong, Qingyun Wan, Gang Cheng, Wai-Yip Tong, Wai-Hung Ang, Wai-Lun Kwong, and Chi-Ming Che. Highly Phosphorescent Platinum(II) Emitters: Photophysics, Materials and Biological Applications. 7:1653–1673, 2016.

[98] Xiaosong Li and Michael J. Frisch. Energy-Represented Direct Inversion in the Iterative Subspace within a Hybrid Geometry Optimization Method. 2(3):835–839, 2006.

[99] Xiaosong Li, John M. Millam, and H. Bernhard Schlegel. Ab Initio Molecular Dynamics Studies of the Photodissociation of Formaldehyde, $H_2CO \rightarrow H_2$ + CO: Direct Classical Trajectory Calculations by MP2 and Density Functional Theory. 113:10062–10067, 2000.

[100] Xiaosong Li, John M. Millam, and H. Bernhard Schlegel. Glyoxal Photodissociation. An Ab Initio Direct Classical Trajectory Study of $C_2H_2O_2 \rightarrow H_2 + 2CO$. 114:8897–8904, 2001.

[101] Yuxi Li. Deep Reinforcement Learning: An Overview. *ArXiv.org*, 2017.

[102] Zhendong Li, Yunlong Xiao, and Wenjian Liu. On the Spin Separation of Algebraic Two-Component Relativistic Hamiltonians. 137(15):154114, 2012.

[103] Jian-Hong Liao, Rajendra Singh Dhayal, Xiaoping Wang, Samia Kahlal, Jean-Yves Saillard, and CW Liu. Neutron Diffraction Studies of a Four-Coordinated Hydride in Near Square-Planar Geometry. *Inorg. Chem.*, 53(20):11140–11145, 2014.

[104] Ping-Kuei Liao, Kuan-Guan Liu, Ching-Shiang Fang, C. W. Liu, John P. Fackler, and Ying-Yann Wu. A Copper(I) Homocubane Collapses to a Tetracapped Tetrahedron Upon Hydride Insertion. *Inorg. Chem.*, 50(17):8410–8417, 2011.

[105] Timothy P. Lillicrap, Jonathan J. Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *ArXiv.org*, 2019.

[106] David B. Lingerfelt, Patrick J. Lestrange, Joseph J. Radler, Samantha E. Brown-Xu, Pyosang Kim, Felix N. Castellano, Lin X. Chen, and Xiaosong Li. Can Excited State Electronic Coherence Be Tuned via Molecular Structural Modification? A First-Principles Quantum Electronic Dynamics Study of Pyrazolate-Bridged Pt(II) Dimers. 121:1932–1939, 2017.

[107] Michael L. Littman. *Markov games as a framework for multi-agent reinforcement learning*. Morgan Kaufmann, 1994.

[108] Michael L Littman, Thomas L Dean, and Leslie Pack Kaelbling. On the Complexity of Solving Markov Decision Problems. *ArXiv.org*, 2013.

[109] Michael L Littman and Csaba Szepesvári. A generalized reinforcement-learning model: Convergence and applications. In *ICML*, volume 96, pages 310–318. Citeseer, 1996.

[110] C. W. Liu, Bijay Sarkar, Yao-Jheng Huang, Ping-Kuei Liao, Ju-Chun Wang, Jean-Yves Saillard, and Samia Kahlal. Octanuclear Copper(I) Clusters Inscribed in a Se12 Icosahedron: Anion-Induced Modulation of the Core Size and Symmetry. page 11222, 2009.

[111] W. Liu and D. Peng. Exact Two-Component Hamiltonians Revisited. 131(3):031104, 2009.

[112] Wenjian Liu. Ideas of Relativistic Quantum Chemistry. 108:1679–1706, 2010.

[113] Wenjian Liu and Daoling Peng. Infinite-Order Quasirelativistic Density Functional Method Based on the Exact Matrix Quasirelativistic Theory. 125:044102, 2006.

[114] Jenny V. Lockard, Aaron A. Rachford, Grigory Smolentsev, Andrew B. Stickrath, Xianghuai Wang, Xiaoyi Zhang, Klaus Atenkoffer, Guy Jennings, Alexander Soldatov, Arnold L. Rheingold, Felix N. Castellano, and Lin X. Chen. Triplet Excited State Distortions in a Pyrazolate Bridged Platinum Dimer Measured by X-ray Transient Absorption Spectroscopy. 114(48):12780–12787, 2010.

[115] A. V. Luzanov and O. A. Zhikol. Electron Invariants and Excited State Structural Analysis for Electronic Transitions Within CIS, RPA, and TDDFT Models. 110(4):902–924, 2010.

[116] A. V. Luzanov and O. A. Zhikol. Electron Invariants and Excited State Structural Analysis for Electronic Transitions within CIS, RPA, and TDDFT Models. 110(4):902–924, 2010.

[117] B Ma, P I Djurovich, S Garon, B Alleyne, and M E Thompson. Platinum Binuclear Complexes as Phosphorescent Dopants for Monochromatic and White Organic Light-Emitting Diodes. 16(18):2438–2446, 2006.

[118] Biwu Ma, Jian Li, Peter I. Djurovich, Muhammed Yousufuddin, Robert Bau, and Mark E. Thompson. Synthetic Control of PtPt Separation and Photophysics of Binuclear Platinum Complexes. 127:28–29, 2004.

[119] Liam MacDermed, Charles Isbell, and Lora Weiss. Markov Games of Incomplete Information for Multi-Agent Reinforcement Learning. In *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*, 2011.

[120] Christel M. Marian. Spin-Orbit Coupling and Intersystem Crossing in Molecules. 2(2):187–203, mar 2012.

[121] R. L. Martin. Natural Transition Orbitals. *J. Chem. Phys.*, 118(11):4775–4777, 2003.

[122] Joseph W. May, Jeremy D. Lehner, Michael J. Frisch, and Xiaosong Li. Transition state search using a guided direct inversion in the iterative subspace method. 8:5175–5179, 2012.

[123] David R McMillin and Jeffrey J Moore. Luminescence that Lasts From Pt(trpy)Cl+ Derivatives (trpy=2,2;6,2-terpyridine). 229(1):113–121, 2002.

[124] Alexis W Mills, Joshua J Goings, David Beck, Chao Yang, and Xiaosong Li. Exploring Potential Energy Surfaces Using Reinforcement Machine Learning. *Journal of Chemical Information and Modeling*, 2022.

[125] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. 518(7540):529–533, 2015.

[126] R. Mohtadi and S. Orimo. The Renaissance of Hydries as Energy Materials. *Nat. Rev. Mater.*, 2:16091, 2017.

[127] Andrew Y Ng, Daishi Harada, and Stuart Russell. Policy Invariance Under Reward Transformations: Theory and Application to Reward Shaping. *Icml*, 99:278–287, 1999.

[128] Scott Niekum, Andrew G. Barto, and Lee Spector. Genetic Programming for Reward Function Search. *IEEE Trans. Auton. Ment. Dev.*, 2(2):83–90, 2010.

[129] Takeshi Noro, Masahiro Sekiya, and Toshikatsu Koga. Segmented Contracted Basis Sets for Atoms H through Xe: Sapporo-(DK)-nZP Sets (n = D, T, Q). 131(2):1124, feb 2012.

[130] Irina V. Novozhilova, Anatoliy V. Volkov, and Philip Coppens. Theoretical Analysis of the Triplet Excited State of the $[Pt_2(H_2P_2O_5)_4]^{4-}$ Ion and Comparison with Time-Resolved X-ray and Spectroscopic Results. 125:1079–1087, 2003.

[131] Maximilian E Ororbia and Gordon P Warn. Design Synthesis Through a Markov Decision Process and Reinforcement Learning Framework. 22(2):021002, 2021.

[132] Xinyan Ou, Qing Chang, and Nilanjan Chakraborty. Simulation Study on Reward Function of Reinforcement Learning in Gantry Work Cell Scheduling. 50:1–8, 2019.

[133] Qing-Jiang Pan, Hong-Gang Fu, Hai-Tao Yu, and Hong-Xing Zhang. Theoretical Insight into Electronic Structures and Spectroscopic Properties of $[Pt_2(pop)_4]^{4-}$, $[Pt_2(pcp)_4]^{4-}$, and Related Derivatives (pop = $P_2O_5H_2$ $^{2-}$ and pcp = $P_2O_4CH_4$ $^{2-}$). 45:8729–8735, 2006.

[134] Ralph G. Pearson. Hard and Soft Acids and Bases. 85:3533–3539, 1963.

[135] Ralph G Pearson. The Transition-Metal-Hydrogen Bond. *Chemical Reviews*, 85(1):41–49, 1985.

[136] Ralph G. Pearson. Hard and Soft Acids and Bases — The Evolution of a Chemical Concept. 100:403–425, 1990.

[137] Robert Pellow and Martin Vala. The External Heavy Atom Effect: Theory of Spin-Orbit Coupling of Alkali and Noble Metals in Rare Gas Matrices. 90:5612, 1989.

[138] Thomas J. Penfold, Etienne Gindensperger, Chantal Daniel, and Christel M. Marian. Spin-Vibronic Mechanism for Intersystem Crossing. 118(15):6975–7025, aug 2018.

[139] Daoling Peng, Wenjian Liu, Yunlong Xiao, and Lan Cheng. Making Four- and Two-Component Relativistic Density Functional Methods Fully Equivalent Based on the Idea of 'From Atoms to Molecule. 127:104106, 2007.

[140] Daoling Peng, Nils Middendorf, Florian Weigend, and Markus Reiher. An Efficient Implementation of Two-Component Relativistic Exact-Decoupling Methods for Large Molecules, journal = jcp, volume =.

[141] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. 77(18):3865–3868, oct 1996.

[142] Diego Perez-Liebana, Katja Hofmann, Sharada Prasanna Mohanty, Noburu Kuno, Andre Kramer, Sam Devlin, Raluca D Gaina, and Daniel Ionita. The Multi-Agent Reinforcement Learning in Malm\" O (MARL\" O) Competition. *arXiv preprint arXiv:1901.08129*, 2019.

[143] Jan Peters and Stefan Schaal. Policy Gradient Methods for Robotics. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2219–2225, 2006.

[144] Jan Peters, Sethu Vijayakumar, and Stefan Schaal. Natural Actor-Critic. In *European Conference on Machine Learning*, pages 280–291. Springer, 2005.

[145] Alessio Petrone, David B. Williams-Young, Shichao Sun, Torin F. Stetina, and Xiaosong Li. An Efficient Implementation of Two-Component Relativistic Density Functional Theory with Torque-Free Auxiliary Variables. 91(7):169, 2018.

[146] Felix Plasser, Giovanni Granucci, Jiri Pittner, Mario Barbatti, Maurizio Persico, and Hans Lischka. Surface Hopping Dynamics using a Locally Diabatic Formalism: Charge Transfer in the Ethylene Dimer Cation and Excited State Dynamics in the 2-pyridone Dimer. 137(22):22A514, dec 2012.

[147] Felix Plasser and Hans Lischka. Analysis of Excitonic and Charge Transfer Interactions from Quantum Chemical Calculations. 8(8):2777–2789, aug 2012.

[148] Felix Plasser, Michael Wormit, and Andreas Dreuw. New Tools for the Systematic Analysis and Visualization of Electronic Excitations. I. Formalism. 141(2):024106, jul 2014.

[149] Péter Pulay. Convergence Acceleration of Iterative Aequences. The Case of SCF Iteration. 73(2):393–398, 1980.

[150] Martin L Puterman. Markov decision processes. *Int. Ser. Oper. Res. Manag. Sci.*, 2:331–434, 1990.

[151] Emma V. Puttock, Melissa T. Walden, and J.A. Gareth Williams. The Luminescence Properties of Multinuclear Platinum Complexes. 367:127–162, 2018.

[152] Xiangfei Qi, P.I. Djurovich, N.C. Giebink, M.E. Thompson, and S.R. Forrest. Evidence for Enhanced Dipolar Interactions Between Pt Centers in Binuclear Phosphorescent Complexes. 458(4):323–328, 2008.

[153] Dongfang Qiu, Jiang Wu, Zhiyuan Xie, Yanxiang Cheng, and Lixiang Wang. Synthesis, Photophysical and Electrophosphorescent Properties of Mononuclear Pt(II) Complexes with Arylamine Functionalized Cyclometalating Ligands. 694(5):737–746, 2009.

[154] Joseph J. Radler, David B. Lingerfelt, Felix N. Castellano, Lin X. Chen, and Xiaosong Li. Role of vibrational dynamics on excited-state electronic coherence in a binuclear platinum complex. 122(23):5071–5077, 2018. PMID: 29652504.

[155] Markus Reiher and Alexander Wolf. Wiley-VCH, second edition, 2015.

[156] D. Max Roundhill, Harry B. Gray, and Chi Ming Che. Pyrophosphito-Bridged Diplatinum Chemistry. 22(2):55–61, feb 1989.

[157] D. Max Roundhill, Harry B. Gray, and Chi Ming Che. Pyrophosphito-Bridged Diplatinum Chemistry. 22(2):55–61, 1989.

[158] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning Representations by Back-Propagating Errors. *Nature*, (323):533–536, 1986.

[159] Erich Runge and E. K. U. Gross. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.*, 52:997–1000, Mar 1984.

[160] Trond Saue. Relativistic Hamiltonians for Chemistry: A Primer. *ChemPhysChem*, 12:3077–3094, 2011.

[161] Tom Schaul, Daniel Horgan, Karol Gregor, and David Silver. Universal value function approximators. *J. Machine Learning Research*, 37:1312–1320, 2015.

[162] H. Bernhard Schlegel. Optimization of equilibrium geometries and transition structures. 3(2):214–218, 1982.

[163] H. Bernhard Schlegel. Geometry optimization. 1(5):790–809, 2011.

[164] John Schulman, Philipp Moritz, Sergey Levine, Michael Jordan, and Pieter Abbeel. High-dimensional continuous control using generalized advantage estimation, 2018.

[165] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *CoRR*, abs/1707.06347, 2017.

[166] Kristof Schütt, Pieter-Jan Kindermans, Huziel Enoc Sauceda Felix, Stefan Chmiela, Alexandre Tkatchenko, and Klaus-Robert Müller. Schnet: A Continuous-Filter Convolutional Neural Network for Modeling Quantum Interactions. *Adv. Neural Inf. Process. Syst.*, 30, 2017.

[167] Marsel Z. Shafikov, Dmitry N. Kozhevnikov, Michael Bodensteiner, Fabian Brandl, and Rafał Czerwieniec. Modulation of Intersystem Crossing Rate by Minor Ligand Modifications in Cyclometalated Platinum(II) Complexes. 55:7457–7466, 2016.

[168] David F Shanno. Conditioning of quasi-Newton Methods for Function Minimization. 24(111):647–656, 1970.

[169] Violeta Sicilia, Pilar Borja, and Antonio Martín. Half-Lantern Pt(II) and Pt(III) Complexes. New Cyclometalated Platinum Derivatives. 2(3):508–523, 2014.

[170] David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller. Deterministic Policy Gradient Algorithms. In Eric P. Xing and Tony Jebara, editors, *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 387–395, Bejing, China, 22–24 Jun 2014. PMLR.

[171] Gregor Simm, Robert Pinsler, and Jose Miguel Hernandez-Lobato. Reinforcement Learning for Molecular Design Guided by Quantum Mechanics. *J. Machine Learning Research*, 119:8959–8969, 2020.

[172] Vasily Sivchik, Aleksandr Kochetov, Toni Eskelinen, Kristina S. Kisel, Anastasia I. Solomatina, Elena V. Grachova, Sergey P. Tunik, Pipsa Hirva, and Igor O. Koshevoy. Modulation of Metallophilic and – Interactions in Platinum Cyclometalated Luminophores with Halogen Bonding. 27(5):1787–1794, 2021.

[173] Justin S Smith, Olexandr Isayev, and Adrian E Roitberg. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. 8(4):3192–3203, 2017.

[174] H Smithson, CA Marianetti, D Morgan, A Van der Ven, A Predith, and G Ceder. First-Principles Study of the Stability and Electronic Structure of Metal Hydrides. *Phys. Rev. B*, 66(14):144107, 2002.

[175] Larry Spruch. Pedagogic notes on Thomas-Fermi theory (and on some improvements): atoms, stars, and the stability of bulk matter. *Reviews of Modern Physics*, 63(1):151, 1991.

[176] Alain St-Amant and Dennis R. Salahub. New Algorithm for the Optimization of Geometries in Local Density Functional Theory. 169(5):387–392, 1990.

[177] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, London, 2018.

[178] Richard S Sutton, David McAllester, Satinder Singh, and Yishay Mansour. Policy Gradient Methods for Reinforcement Learning with Function Approximation. *Adv. Neural Inf. Process. Syst.*, 12, 1999.

[179] Richard S Sutton, Doina Precup, and Satinder Singh. Between MDPs and Semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning. 112(1-2):181–211, 1999.

[180] Ming Tan. Multi-Agent Reinforcement Learning: Independent vs. Cooperative Agents. *In Proceedings of the Tenth International Conference on Machine Learning*, pages 330–337, 1993.

[181] Qing Tang, Yongjin Lee, Dai-Ying Li, Woojun Choi, C. W. Liu, Dongil Lee, and De-en Jiang. Lattice-Hydride Mechanism in Electrocatalytic CO2 Reduction by Structurally Precise Copper-Hydride Nanoclusters. 139(28):9728–9736, 2017.

[182] Yoshihiro Tawada, Takao Tsuneda, Susumu Yanagisawa, Takeshi Yanai, and Kimihiko Hirao. A Long-Range-Corrected Time-Dependent Density Functional Theory. *The Journal of Chemical Physics*, 120(18):8425–8433, 2004.

[183] Alexander Tronnier and Thomas Strassner. (C^C*) Cyclometalated Binuclear N-heterocyclic Biscarbene Platinum(II) Complexes – Highly Emissive Phosphorescent Emitters. 42:9847, 2013.

[184] Takao Tsuneda and Kimihiko Hirao. Long-range Correction for Density Functional Theory. *WIREs Computational Molecular Science*, 4(4):375–390, 2014.

[185] A.J.S Valentine and Xiaosong Li. Toward the Evaluation of Intersystem Crossing Rates with Variational Relativistic Methods. 151:084107, 2019.

[186] Andrew J. S. Valentine, Joseph J. Radler, Alexis Mills, Pyosang Kim, Felix N. Castellano, Lin X. Chen, and Xiaosong Li. Resolving the Ultrafast Intersystem Crossing in a Bimetallic Platinum Complex. 151:114303, 2019.

[187] Renske M. van der Veen, Andrea Cannizzo, Frank van Mourik, Antonin Vlček, and Majed Chergui. Vibrational Relaxation and Intersystem Crossing of Binuclear Metal Complexes in Solution. 133(2):305–315, jan 2011.

[188] Renske M. van der Veen, Chris J. Milne, Amal El Nahhas, Frederico A. Lima, Van-Thai Pham, Jonathan Best, Julia A. Weinstein, Camelia N. Borca, Rafael Abela, Christian Bressler, and Majed Chergui. Structural Determination of a Photochemically Active Diplatinum Molecule by Time-Resolved EXAFS Spectroscopy. 48(15):2711–2714, mar 2009.

[189] H. van Hasselt and M. A. Wiering. Reinforcement learning in continuous action spaces. In *2007 IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning*, pages 272–279, 2007.

[190] G. Vignale, C. A. Ullrich, and S. Conti. Time-Dependent Density Functional Theory Beyond the Adiabatic Local Density Approximation. *Phys. Rev. Lett.*, 79:4878–4881, Dec 1997.

[191] Willard R. Wadt and P. Jeffrey Hay. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for Main Group Elements Na to Bi. 82(1):284–298, 1985.

[192] Song Wang, Zili Wu, Sheng Dai, and De-en Jiang. Deep learning accelerated determination of hydride locations in metal nanoclusters. *Angewandte Chemie International Edition*, 60(22):12289–12292, 2021.

[193] Julia A. Weinstein, Mark T. Tierney, E. Stephen Davies, Karel Base, Anthony A. Robeiro, and Mark W. Grinstaff. Probing the Electronic Structure of Platinum(II) Chromophores: Crystal Structures, NMR Structures, and Photophysical Properties of Six New Bis- and Di- Phenolate/Thiolate Pt(II)Diimine Chromophores. 45:4544–4555, 2006.

[194] J. A. Gareth Williams, Stephanie Develay, David L. Rochester, and Lisa Murphy. Optimising the Luminescence of Platinum(II) Complexes and Their Application in Organic Light Emitting Devices(OLEDs). 252:2596–2611, 2008.

[195] Keith Man-Chung Wong and Vivian Wing-Wah Yam. Self-Assembly of Luminescent Alkynylplatinum(II) Terpyridyl Complexes: Modulation of Photophysical Properties through Aggregation Behavior. 44(6):424–434, 2011.

[196] R.G. Woolley and B.T. Sutcliffe. Molecular Structure and the Born—Oppenheimer Approximation. *Chemical Physics Letters*, 45(2):393–398, 1977.

[197] Barry J. Wythoff. Backpropagation Neural Networks: A tutorial. *Chemom. Intell. Lab. Syst.*, 18(2):115–155, 1993.

[198] Tian Xie and Jeffrey C Grossman. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.*, 120(14):145301, 2018.

[199] Vivian Wing-Wah Yam, Vonika Ka-Man Au, and Sammual Yu-Lut Leung. Light-Emitting Self-Assembled Materials Based on d8 and d10 Transition Metal Complexes. 115(15):7589–7728, 2015. PMID: 26158432.

[200] Vivian Wing-Wah Yam, Alan Kwun-Wa Chan, and Eugene Yau-Hin Hong. Charge-Transfer Processes in Metal Complexes Enable Luminescence and Memory Functions. 4:528–541, 2020.

[201] Vivian Wing-Wah Yam and Angela Sin-Yee Law. Luminescent d8 Metal Complexes of Platinum(II) and Gold(III): From Photophysics to Photofunctional Materials and Probes. 414:213298, 2020.

[202] Takeshi Yanai, David P Tew, and Nicholas C Handy. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). 393(1-3):51–57, July 2004.

[203] Tengteng Zhang and Hongwei Mo. Reinforcement Learning for Robot Research: A Comprehensive Review and Open Issues. *Int. J. Adv. Robot. Syst.*, 18(3):17298814211007305, 2021.

[204] Di Zhou. An introduction of density functional theory and its application. 2008.

[205] Chao Zou, Jinqiang Lin, Sa Suo, Mo Xie, Xiaoyong Chang, and Wei Lu. Palladium(ii) N-Heterocyclic Allenylidene Complexes with Extended Intercationic PdPd Interactions and MMLCT Phosphorescence. 54:5319–5322, 2018.

[206] Juan A. Zuleta, Joanne M. Bevilacqua, Davide M. Proserpio, Pierre D. Harvey, and Richard Eisenberg. Spectroscopic and Theoretical Studies on the Excited State in Diimine Dithiolate Complexes of Platinum(II). 31:2396–2404, 1992.

Appendix A

STATIC ELECTRONIC SPECTRA WITH ωB97X-D AND DEF2TZVP SDD

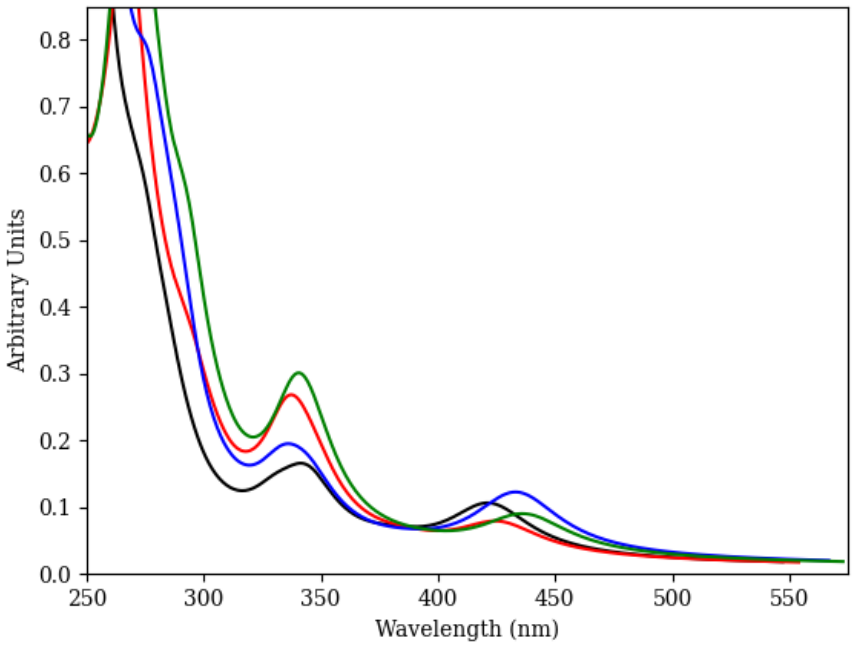


Figure A.1. Collected absorption spectra for complexes **6** (black), **7** (red), **8** (blue), and **9** (green) using ωB97X-D with def2-TZVP/SDD.

Appendix B

CHARGE TRANSFER NUMBERS OF BIMETALLIC PLATINUM DIMERS

The following CT tables demonstrate the remaining oxygen derivatives, with shorter Pt-Pt distances, experience the same trends observed in the main text for complex **6**. The ^{1}MM(CM)LCT state is found to be quite sensitive to the functional used; specifically, the inclusion of range-correction is shown to largely influence the electronic landscape by shifting the CT states. This energetic shift carries over to the triplet states exhibiting strong CT character as well and can lead to state mixing.

	State	Energy (eV)	Ω	Charge-Transfer Ratio	Designation
B3LYP	1	2.402	$\begin{pmatrix} 0.05 & 0.68 \\ 0.01 & 0.14 \end{pmatrix}$	0.804	^{3}MM(CM)LCT
	2	2.519	$\begin{pmatrix} 0.03 & 0.62 \\ 0.01 & 0.26 \end{pmatrix}$	0.713	^{3}MM(CM)LCT
	3	2.543	$\begin{pmatrix} 0.04 & 0.78 \\ 0.01 & 0.06 \end{pmatrix}$	0.906	^{1}MM(CM)LCT
	5	2.669	$\begin{pmatrix} 0.03 & 0.34 \\ 0.03 & 0.53 \end{pmatrix}$	0.438	^{3}Mixed
	6	2.677	$\begin{pmatrix} 0.03 & 0.43 \\ 0.02 & 0.43 \end{pmatrix}$	0.54	^{3}Mixed
CAM-B3LYP	1	2.745	$\begin{pmatrix} 0.02 & 0.12 \\ 0.03 & 0.82 \end{pmatrix}$	0.162	^{3}(CM)LC
	2	2.759	$\begin{pmatrix} 0.01 & 0.10 \\ 0.03 & 0.85 \end{pmatrix}$	0.138	^{3}(CM)LC
	3	3.137	$\begin{pmatrix} 0.11 & 0.61 \\ 0.01 & 0.17 \end{pmatrix}$	0.719	^{3}MM(CM)LCT
	4	3.285	$\begin{pmatrix} 0.06 & 0.70 \\ 0.01 & 0.12 \end{pmatrix}$	0.816	^{1}MM(CM)LCT
	5	3.314	$\begin{pmatrix} 0.01 & 0.01 \\ 0.00 & 0.00 \end{pmatrix}$	0.067	^{3}Mixed
ωB97X-D	1	2.788	$\begin{pmatrix} 0.05 & 0.38 \\ 0.02 & 0.48 \end{pmatrix}$	0.469	^{3}Mixed
	2	2.869	$\begin{pmatrix} 0.02 & 0.14 \\ 0.03 & 0.79 \end{pmatrix}$	0.192	^{3}(CM)LC
	3	2.963	$\begin{pmatrix} 0.06 & 0.41 \\ 0.02 & 0.46 \end{pmatrix}$	0.485	^{3}Mixed
	4	3.03	$\begin{pmatrix} 0.06 & 0.75 \\ 0.01 & 0.07 \end{pmatrix}$	0.861	^{1}MM(CM)LCT
	5	3.228	$\begin{pmatrix} 0.04 & 0.74 \\ 0.01 & 0.14 \end{pmatrix}$	0.818	^{3}MM(CM)LCT

Table B.1. CT numbers of complex **6** using basis sets 6-31g(d) on light atoms and LanL2DZ on Pt

	State	Energy (eV)	Ω	Charge-Transfer Ratio	Designation
B3LYP	1	2.321	$\begin{pmatrix} 0.03 & 0.68 \\ 0.01 & 0.17 \end{pmatrix}$	0.795	^{3}MM(CM)LCT
	2	2.368	$\begin{pmatrix} 0.02 & 0.65 \\ 0.01 & 0.24 \end{pmatrix}$	0.741	^{3}MM(CM)LCT
	3	2.432	$\begin{pmatrix} 0.02 & 0.80 \\ 0.01 & 0.06 \end{pmatrix}$	0.919	^{1}MM(CM)LCT
	5	2.589	$\begin{pmatrix} 0.01 & 0.20 \\ 0.01 & 0.73 \end{pmatrix}$	0.253	^{3}(CM)LC
	6	2.623	$\begin{pmatrix} 0.01 & 0.24 \\ 0.01 & 0.71 \end{pmatrix}$	0.282	^{3}(CM)LC
CAM-B3LYP	1	2.552	$\begin{pmatrix} 0.00 & 0.05 \\ 0.01 & 0.93 \end{pmatrix}$	0.066	^{3}(CM)LC
	2	2.556	$\begin{pmatrix} 0.00 & 0.05 \\ 0.01 & 0.93 \end{pmatrix}$	0.066	^{3}(CM)LC
	3	2.983	$\begin{pmatrix} 0.04 & 0.34 \\ 0.03 & 0.54 \end{pmatrix}$	0.415	^{3}Mixed
	4	3.033	$\begin{pmatrix} 0.01 & 0.24 \\ 0.02 & 0.69 \end{pmatrix}$	0.293	^{3}(CM)LC
	6	3.271	$\begin{pmatrix} 0.05 & 0.70 \\ 0.01 & 0.14 \end{pmatrix}$	0.809	^{1}MM(CM)LCT
ωB97X-D	1	2.67	$\begin{pmatrix} 0.00 & 0.09 \\ 0.01 & 0.87 \end{pmatrix}$	0.13	^{3}(CM)LC
	2	2.715	$\begin{pmatrix} 0.01 & 0.06 \\ 0.01 & 0.91 \end{pmatrix}$	0.086	^{3}(CM)LC
	3	2.846	$\begin{pmatrix} 0.07 & 0.60 \\ 0.02 & 0.22 \end{pmatrix}$	0.715	^{3}MM(CM)LCT
	4	2.989	$\begin{pmatrix} 0.05 & 0.77 \\ 0.01 & 0.07 \end{pmatrix}$	0.881	^{1}MM(CM)LCT
	5	3.095	$\begin{pmatrix} 0.02 & 0.36 \\ 0.02 & 0.57 \end{pmatrix}$	0.417	^{3}Mixed

Table B.2. CT numbers of complex **7** using basis sets 6-31g(d) on light atoms and LanL2DZ on Pt

	State	Energy (eV)	Ω	Charge-Transfer Ratio	Designation
B3LYP	1	2.33	$\begin{pmatrix} 0.05 & 0.70 \\ 0.01 & 0.10 \end{pmatrix}$	0.85	^{3}MM(CM)LCT
	2	2.451	$\begin{pmatrix} 0.03 & 0.71 \\ 0.01 & 0.14 \end{pmatrix}$	0.832	^{3}MM(CM)LCT
	3	2.451	$\begin{pmatrix} 0.03 & 0.75 \\ 0.01 & 0.05 \end{pmatrix}$	0.915	^{1}MM(CM)LCT
	5	2.672	$\begin{pmatrix} 0.02 & 0.32 \\ 0.03 & 0.57 \end{pmatrix}$	0.41	^{3}Mixed
	6	2.675	$\begin{pmatrix} 0.02 & 0.29 \\ 0.03 & 0.61 \end{pmatrix}$	0.365	^{3}(CM)LC
CAM-B3LYP	1	2.744	$\begin{pmatrix} 0.02 & 0.13 \\ 0.03 & 0.81 \end{pmatrix}$	0.173	^{3}(CM)LC
	2	2.762	$\begin{pmatrix} 0.01 & 0.10 \\ 0.03 & 0.85 \end{pmatrix}$	0.136	^{3}(CM)LC
	3	3.076	$\begin{pmatrix} 0.09 & 0.52 \\ 0.01 & 0.14 \end{pmatrix}$	0.635	^{3}MM(CM)LCT
	4	3.095	$\begin{pmatrix} 0.01 & 0.09 \\ 0.00 & 0.02 \end{pmatrix}$	0.143	^{3}Mixed
	6	3.203	$\begin{pmatrix} 0.06 & 0.71 \\ 0.01 & 0.09 \end{pmatrix}$	0.844	^{1}MM(CM)LCT
ωB97X-D	1	2.738	$\begin{pmatrix} 0.07 & 0.53 \\ 0.02 & 0.28 \end{pmatrix}$	0.647	^{3}Mixed
	2	2.876	$\begin{pmatrix} 0.01 & 0.15 \\ 0.03 & 0.79 \end{pmatrix}$	0.197	^{3}(CM)LC
	3	2.93	$\begin{pmatrix} 0.06 & 0.74 \\ 0.01 & 0.07 \end{pmatrix}$	0.874	^{1}MM(CM)LCT
	4	2.933	$\begin{pmatrix} 0.03 & 0.26 \\ 0.02 & 0.64 \end{pmatrix}$	0.331	^{3}(CM)LC
	5	3.104	$\begin{pmatrix} 0.37 & 0.26 \\ 0.02 & 0.01 \end{pmatrix}$	0.582	^{3}Mixed

Table B.3. CT numbers of complex **8** using basis sets 6-31g(d) on light atoms and LanL2DZ on Pt

	State	Energy (eV)	Ω	Charge-Transfer Ratio	Designation
B3LYP	1	2.252	$\begin{pmatrix} 0.03 & 0.71 \\ 0.01 & 0.11 \end{pmatrix}$	0.856	^{3}MM(CM)LCT
	2	2.297	$\begin{pmatrix} 0.02 & 0.71 \\ 0.00 & 0.14 \end{pmatrix}$	0.84	^{3}MM(CM)LCT
	3	2.339	$\begin{pmatrix} 0.02 & 0.77 \\ 0.00 & 0.05 \end{pmatrix}$	0.927	^{1}MM(CM)LCT
	5	2.573	$\begin{pmatrix} 0.01 & 0.18 \\ 0.01 & 0.77 \end{pmatrix}$	0.224	^{3}(CM)LC
	6	2.599	$\begin{pmatrix} 0.01 & 0.18 \\ 0.01 & 0.77 \end{pmatrix}$	0.226	^{3}(CM)LC
CAM-B3LYP	1	2.551	$\begin{pmatrix} 0.00 & 0.05 \\ 0.01 & 0.93 \end{pmatrix}$	0.066	^{3}(CM)LC
	2	2.557	$\begin{pmatrix} 0.00 & 0.05 \\ 0.01 & 0.93 \end{pmatrix}$	0.064	^{3}(CM)LC
	3	2.961	$\begin{pmatrix} 0.05 & 0.42 \\ 0.02 & 0.43 \end{pmatrix}$	0.516	^{3}Mixed
	4	3.025	$\begin{pmatrix} 0.01 & 0.29 \\ 0.02 & 0.63 \end{pmatrix}$	0.35	^{3}(CM)LC
	7	3.184	$\begin{pmatrix} 0.05 & 0.72 \\ 0.01 & 0.10 \end{pmatrix}$	0.85	^{1}MM(CM)LCT
ωB97X-D	1	2.665	$\begin{pmatrix} 0.01 & 0.18 \\ 0.01 & 0.76 \end{pmatrix}$	0.23	^{3}(CM)LC
	2	2.708	$\begin{pmatrix} 0.01 & 0.06 \\ 0.01 & 0.91 \end{pmatrix}$	0.089	^{3}(CM)LC
	3	2.757	$\begin{pmatrix} 0.06 & 0.57 \\ 0.01 & 0.26 \end{pmatrix}$	0.679	^{3}MM(CM)LCT
	4	2.861	$\begin{pmatrix} 0.04 & 0.79 \\ 0.01 & 0.05 \end{pmatrix}$	0.905	^{1}MM(CM)LCT
	5	2.999	$\begin{pmatrix} 0.39 & 0.26 \\ 0.02 & 0.01 \end{pmatrix}$	0.597	^{3}Mixed

Table B.4. CT numbers of complex **9** using basis sets 6-31g(d) on light atoms and LanL2DZ on Pt

Appendix C

PEARSON CORRELATION COEFFICIENT

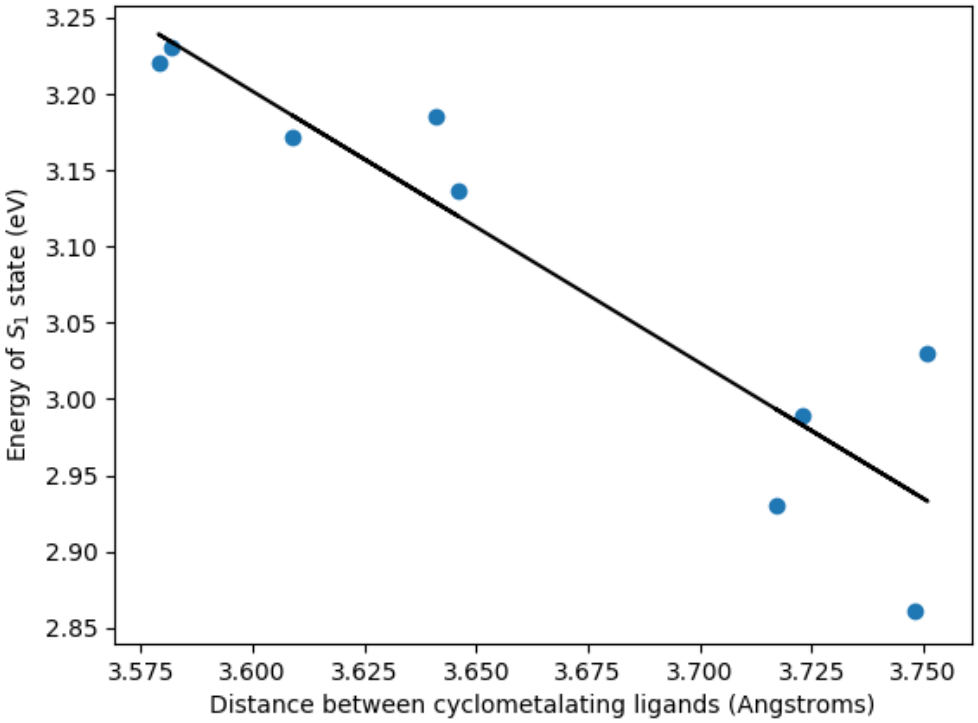


Figure C.1. Plot of S_1 state energy as a function of distance between the cyclometalating ligands at the center of mass point. The slope was calculated to be -1.781 and the Pearson coefficient was calculated to be -0.92.

Appendix D

JABLONSKI DIAGRAM OF ALL PLATINUM COMPLEXES

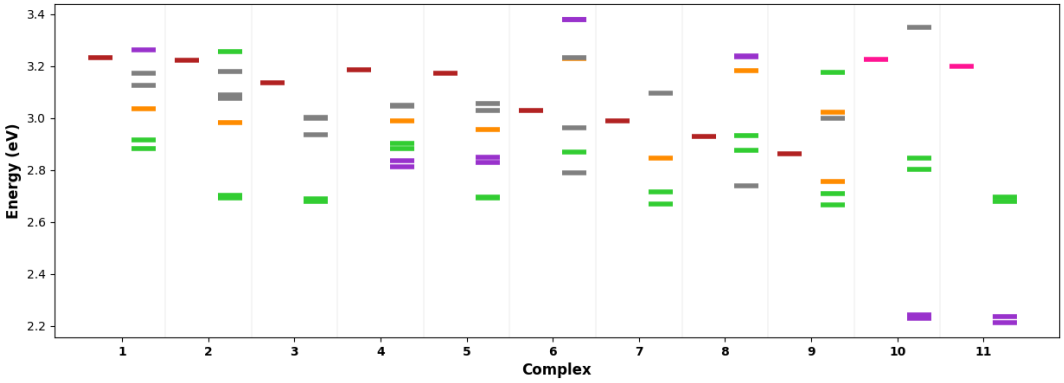


Figure D.1. Jablonski diagram for complexes **1-11**. The colors of each level are as follows: ^{1}MMLCT (red), ^{1}BLC (pink), ^{3}MMLCT (orange), ^{3}CMLC (green), ^{3}BLC (purple), and ^{3}Mixed (grey).

Appendix E

POTENTIAL ENERGY AS A FUNCTION OF PLATINUM-PLATINUM BOND LENGTH

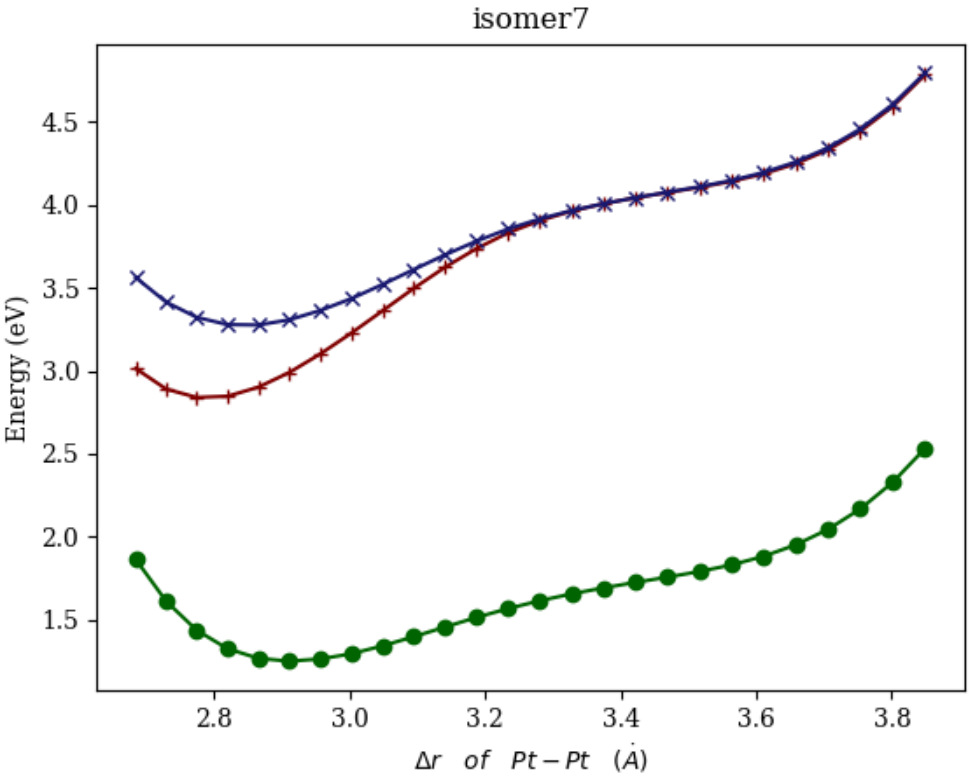


Figure E.1. Potential energy surface of **7** where the ground-state is green, the first singlet excited-state surface is red, and the second singlet excited-state is blue.

Appendix F

FUNCTIONAL DEPENDENCE

To analyze the excited-state character of the platinum dimer complex, and to examine the choice of density functional, the charge-transfer numbers Ω_{AB} are evaluated at the ground-state geometry (see below) for the first six excited states using the functionals CAM-B3LYP,[202] PBE,[141] B3LYP,[11] and ωB97X-D[26], presented in the following tables. The 6-31g(d) basis was used on light atoms, and the LanL2DZ effective core potential and double-ζ basis set on the Pt atoms.[73] The fragment A was chosen to be the π-accepting ligands and the fragment B was chosen to be the platinum centers. Consequently, the elements of Ω correspond to following state designations

$$\Omega = \begin{pmatrix} \mathrm{LC} & \mathrm{LMCT} \\ \mathrm{MMLCT} & \mathrm{MC} \end{pmatrix}, \tag{F.1}$$

where LC and MC mean ligand- and metal-centered, respectively, while LMCT and MMLCT are ligand-to-metal and metal-metal-to-ligand charge transfer, respectively.

First, the range-separated hybrid functional CAM-B3LYP, which was used in the text, is presented in Table C.1. The first two triplet states are clearly ligand-centered, followed by a number of MMLCT states. In sharp contrast, the pure functional PBE, which lacks any exact Hartree-Fock exchange, finds only charge-transfer states, all much lower in energy than the states captured by CAM-B3LYP. This is consistent with the tendency of pure functionals to dramatically overstabilize charge-transfer states. Next, the hybrid functional B3LYP (Table C.3), which contains some degree of exact exchange, does not stabilize the charge-transfer states as radically, but still finds a number of MMLCT states below the first LC states. Interestingly, the LC

Table F.1. Charge-transfer numbers from TD-CAM-B3LYP.

State	Energy (eV)	Ω	Designation
1	2.745	$\begin{pmatrix} 0.85 & 0.03 \\ 0.10 & 0.01 \end{pmatrix}$	^{3}LC
2	2.760	$\begin{pmatrix} 0.87 & 0.03 \\ 0.08 & 0.01 \end{pmatrix}$	^{3}LC
3	3.064	$\begin{pmatrix} 0.12 & 0.01 \\ 0.67 & 0.14 \end{pmatrix}$	^{3}MMLCT
4	3.242	$\begin{pmatrix} 0.09 & 0.01 \\ 0.75 & 0.09 \end{pmatrix}$	^{1}MMLCT
5	3.326	$\begin{pmatrix} 0.21 & 0.01 \\ 0.68 & 0.05 \end{pmatrix}$	^{3}MMLCT
6	3.475	$\begin{pmatrix} 0.09 & 0.00 \\ 0.80 & 0.04 \end{pmatrix}$	^{1}MMLCT

Table F.2. Charge-transfer numbers from TD-PBE.

State	Energy (eV)	Ω	Designation
1	1.633	$\begin{pmatrix} 0.06 & 0.00 \\ 0.79 & 0.07 \end{pmatrix}$	^{3}MMLCT
2	1.754	$\begin{pmatrix} 0.05 & 0.00 \\ 0.81 & 0.05 \end{pmatrix}$	^{1}MMLCT
3	1.789	$\begin{pmatrix} 0.06 & 0.00 \\ 0.81 & 0.05 \end{pmatrix}$	^{3}MMLCT
4	1.883	$\begin{pmatrix} 0.05 & 0.00 \\ 0.82 & 0.05 \end{pmatrix}$	^{1}MMLCT
5	2.160	$\begin{pmatrix} 0.06 & 0.00 \\ 0.86 & 0.01 \end{pmatrix}$	^{3}MMLCT
6	2.175	$\begin{pmatrix} 0.06 & 0.00 \\ 0.86 & 0.01 \end{pmatrix}$	^{1}MMLCT

Table F.3. Charge-transfer numbers from TD-B3LYP.

State	Energy (eV)	Ω	Designation
1	2.308	$\begin{pmatrix} 0.08 & 0.01 \\ 0.76 & 0.09 \end{pmatrix}$	^{3}MMLCT
2	2.435	$\begin{pmatrix} 0.07 & 0.00 \\ 0.80 & 0.06 \end{pmatrix}$	^{1}MMLCT
3	2.496	$\begin{pmatrix} 0.11 & 0.01 \\ 0.78 & 0.05 \end{pmatrix}$	^{3}MMLCT
4	2.596	$\begin{pmatrix} 0.06 & 0.00 \\ 0.83 & 0.04 \end{pmatrix}$	^{1}MMLCT
5	2.704	$\begin{pmatrix} 0.66 & 0.03 \\ 0.25 & 0.02 \end{pmatrix}$	^{3}LC
6	2.734	$\begin{pmatrix} 0.68 & 0.03 \\ 0.24 & 0.02 \end{pmatrix}$	^{3}LC

Table F.4. Charge-transfer numbers from ωB97X-D.

State	Energy (eV)	Ω	Designation
1	2.856	$\begin{pmatrix} 0.81 & 0.03 \\ 0.13 & 0.02 \end{pmatrix}$	^{3}LC
2	2.874	$\begin{pmatrix} 0.85 & 0.03 \\ 0.10 & 0.01 \end{pmatrix}$	^{3}LC
3	3.118	$\begin{pmatrix} 0.13 & 0.01 \\ 0.67 & 0.14 \end{pmatrix}$	^{3}MMLCT
4	3.273	$\begin{pmatrix} 0.09 & 0.01 \\ 0.75 & 0.09 \end{pmatrix}$	^{1}MMLCT
5	3.3372	$\begin{pmatrix} 0.18 & 0.01 \\ 0.71 & 0.06 \end{pmatrix}$	^{3}MMLCT
6	3.494	$\begin{pmatrix} 0.02 & 0.02 \\ 0.27 & 0.52 \end{pmatrix}$	^{3}MC

states from B3LYP have nearly the same energy as those from CAM-B3LYP, suggesting that both functionals treat LC states similarly, but B3LYP overstabilizes the charge-transfer states. Finally, the range-separated and dispersion-corrected ωB97X-D functional (Table C.4) yields very similar results for the first five excited states. These results are consistent with the intuition that range-separated functionals should be used when examining systems featuring both local and charge-transfer excitations in order to treat both on an equal footing, and for that reason CAM-B3LYP was used in this study.

Appendix G

BEYOND THE MÜLLER-BROWN SURFACE

The additional surfaces were devised using the following equation:[16]

$$\lambda(x,y) = \sum_{i=1}^{4} A_i exp[a_i(x-\bar{x}_i)^2 + b_i(x-\bar{x}_i)(y-\bar{y}_i) + c_i(y-\bar{y}_i)^2] \tag{G.1}$$

where surface **I** assigns the following values to the corresponding variables

$$\begin{aligned} A = (-150,-120,-250,-200,-150,-170,15); a = (-1,-1,-1,-1,-1,-10,0.7); \\ b = (0,0,-1,0,1,11,0.6); c = (-5,-15,-10,-10,-10,-6.5,0.7); \\ \bar{x} = (1,,1.7,-1.5,1,0,-0.5,-1); \bar{y} = (1.25,0.75,-0.25,0,0.5,1.5,1) \end{aligned} \tag{G.2}$$

and surface **II** uses the following values

$$\begin{aligned} A = (-130,-150,-125,-100,5); a = (-2.5,-5,-2,-6.5,0.5); \\ b = (0.5,0,0.5,0.75,0.6); c = (-1,-10,-1,-6.5,0.7); \\ \bar{x} = (-1.5,1,0,-0.5,-1); \bar{y} = (0.75,0,0.5,1.5,1) \end{aligned} \tag{G.3}$$