Toolkits for simulation and interpretation of photoinduced processes: a mixed classical-quantum approach

Sapunar, Marin

Doctoral thesis / Disertacija

2021

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:767444

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**



Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb







University of Zagreb
FACULTY OF SCIENCE

Marin Sapunar

TOOLKITS FOR SIMULATION AND INTERPRETATION OF PHOTOINDUCED PROCESSES: A MIXED CLASSICAL-QUANTUM APPROACH

DOCTORAL DISSERTATION

Supervisor:
Dr. Nađa Došlić

Zagreb, 2021



Sveučilište u Zagrebu
PRIRODOSLOVNO-MATEMATIČKI FAKULTET

Marin Sapunar

ALATI ZA SIMULACIJU I INTERPRETACIJU FOTOINDUCIRANIH PROCESA – MIJEŠANI KLASIČNO-KVANTNI PRISTUP

DOKTORSKI RAD

Mentor:
dr. sc. Nađa Došlić

Zagreb, 2021.

Zahvale/Thank you..

… mentorici dr. sc. Nađi Došlić na svemu što sam od nje naučio, na svim savjetima, pomoći i na vrlo ugodnom boravku u grupi kroz sve ove godine. Hvala na poticaju da slijedim svoju znatiželju čak i kada me vodila vijugavim putem do cilja.

… Momiru, Aurori i Tomislavu na pomoći u ovom i ostalim znanstvenim radovima te na brojnim diskusijama, znanstvenim i ostalim.

… Marku, Ivanu, Darku, Davoru, Jurici, Sermsiri, Tareku, Mihaelu, Jovanu i Rajku na znanstvenim raspravama te na ugodnoj radnoj atmosferi.

… to prof. Piero Decleva and prof. Wolfgang Domcke for welcoming me to their groups and giving me the opportunity to learn from them. Also to Nicola, Xiang and Sebastian for making my stays in Trieste and Munich enjoyable.

… Grgi i Kiki koji su uvijek pri ruci za podijeliti i dobre i loše i potpuno nevažne trenutke.

… mojoj obitelji, mami, tati, sestri i baki koji su mi od rođenja potpora i stabilnost u svemu, velikom i malom. Didi i nonu koji su znali da ću uspjeti sve što započnem. Vesni, Toniju i Karli koji su me u potpunosti prihvatili u svoju obitelj.

… najviše od svega, hvala Antei, Marti i uskoro Luni na beskonačnoj ljubavi, sreći i potpori u ovom periodu.

Sadržaj

SAŽETAK ..... **XI**
ABSTRACT ..... **XIII**
PROŠIRENI SAŽETAK ..... **XV**
§ 1. INTRODUCTION **1**
§ 2. LITERATURE OVERVIEW **4**
2.1. Nonadiabatic dynamics **4**
2.2. Mixed quantum-classical dynamics **8**
2.2.1. Trajectory based methods *8*
2.2.2. Fewest-switching surface hopping *10*
2.2.3. Couplings and wave function overlaps *14*
2.2.4. Landau-Zener surface hopping *24*
2.2.5. Nuclear phase space distributions *28*
2.3. Photochemical problems **33**
2.3.1. Pyrrole *34*
2.3.2. Pyrazine *37*
§ 3. COMPUTATIONAL METHODS **40**
3.1. Electronic structure **40**
3.1.1. Algebraic diagrammatic construction *42*
3.1.2. Time-dependent density functional theory *44*
3.1.3. Complete active space methods *45*
3.1.4. Other excited state electronic structure methods *46*
3.2. Implementation of surface hopping algorithms **47**
3.2.1. Landau-Zener surface hopping *52*
3.2.2. Molecular orbital overlap integrals *54*
3.3. Photoionization observables **58**
§ 4. RESULTS AND DISCUSSION **61**
4.1. Evaluation of SH algorithms **61**
4.1.1. Pyrazine *62*
4.1.2. Pyrole *66*
4.1.1. Numerical stability of LZSH *68*
4.2. Wave function overlap algorithms **70**

4.2.1. Performance *74*
4.2.1. Approximations *77*
4.3. Decomposition of spectra **82**
4.3.1. Absorption spectra of nucleobases *84*
4.4. Pyrrole **90**
4.4.1. UV absorption spectrum *93*
4.4.2. FSSH simulations *95*
4.4.3. Photoionization *101*
4.5. Pyrazine **103**
4.5.1. Potential energy surfaces *104*
4.5.2. FSSH simulations *106*
§ 5. CONCLUSION **111**
§ 6. LIST OF ABBREVIATIONS **115**
§ 7. REFERENCES **116**
§ 8. CURRICULUM VITAE ..... **XXXIII**

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Kemijski odsjek

Doktorska disertacija

SAŽETAK

ALATI ZA SIMULACIJU I INTERPRETACIJU FOTOINDUCIRANIH PROCESA – MIJEŠANI KLASIČNO-KVANTNI PRISTUP

Marin Sapunar
Zavod za fizičku kemiju, Institut Ruđer Bošković, Bijenička 54, 10000 Zagreb, Hrvatska

Fotokemijske reakcije uključuju više elektronskih stanja koja međusobno interagiraju zbog čega se moraju proučavati metodama koje idu dalje od Born-Oppenheimerove aproksimacije. Cilj ovog rada je doprinijeti miješanim klasično-kvantnim metodama za proučavanje takvih reakcija. Razvijeni su vrlo efikasni numerički algoritmi za računanje integrala preklapanja između elektronskih valnih funkcija na različitim geometrijama molekula. Ti integrali nužni su za simulacije neadijabatske dinamike na miješanom klasično-kvantnom nivou, ali ovdje pokazujemo da su i iznimno korisni za analizu elektronskih stanja u velikim ansamblima nuklearnih konfiguracija. Takva analiza je zatim iskorištena za ponovno sagledavanje dvije široko proučavane reakcije u fotokemiji, disocijaciju vodika u pirolu i internu konverziju u pirazinu. Kod pirola se proučava utjecaj miješanja svijetlog valentnog $\pi\pi^*$ stanja s Rydbergovim stanjima na vremensku skalu reakcije, dok se kod pirazina daje dokaz vrlo ranog populiranja dva nπ^* stanja.

(130 stranica, 36 slika, 6 tablica, 286 literaturnih navoda, jezik izvornika: engleski)

Rad je pohranjen u Središnjoj kemijskoj knjižnici, Horvatovac 102a, Zagreb i Nacionalnoj i sveučilišnoj knjižnici, Hrvatske bratske zajednice 4, Zagreb.

Ključne riječi: metoda nuklearnog ansambla / neadijabatska dinamika / pirazin / pirol / pobuđena stanja / preklapanja valnih funkcija

Mentor: dr. sc. Nađa Došlić, znanstveni savjetnik u trajnom zvanju

Rad prihvaćen: 2. lipnja 2021.

Ocjenitelji:
1. prof. dr. sc. Tomica Hrenar
2. dr. sc. Marko Cvitaš, zn. sur.
3. dr. sc. Janez Mavri, zn. savj.



University of Zagreb
Faculty of Science
Department of Chemistry

Doctoral Thesis

ABSTRACT

TOOLKITS FOR SIMULATION AND INTERPRETATION OF PHOTOINDUCED PROCESSES: A MIXED CLASSICAL-QUANTUM APPROACH

Marin Sapunar
Department of Physical Chemistry, Ruđer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia

Photochemical reactions involve multiple interacting electronic states which means they need to be treated using methods beyond the Born-Oppenheimer approximation. In the current work, contributions are made to mixed classical-quantum methods for studying these reactions. First, highly efficient numerical algorithms were developed to calculate overlap integrals between electronic wave functions at different nuclear configurations. These integrals are essential for nonadiabatic dynamics simulations at the mixed classical-quantum level but are also shown here to be very useful for facilitating the analysis of the electronic states in large ensembles of nuclear configurations. This type of analysis is then used to revisit two prototypical reactions in photochemistry, hydrogen dissociation in pyrrole and internal conversion in pyrazine. In pyrrole, the effect of Rydberg-valence mixing in the bright $\pi\pi^*$ state on the time scale of the reaction is discussed. In pyrazine, the very early population of two nπ^* states is conclusively proven.

(130 pages, 36 figures, 6 tables, 286 references, original in English)

Thesis deposited in Central Chemical Library, Horvatovac 102A, Zagreb, Croatia and National and University Library, Hrvatske bratske zajednice 4, Zagreb, Croatia.

Keywords: excited states / nonadiabatic dynamics / nuclear ensemble approach / pyrazine / pyrrole / wave function overlaps

Supervisor: Dr. Nađa Došlić, Scientific advisor

Thesis accepted: July 4th 2021

Reviewers:
Dr. Tomica Hrenar, Professor
Dr. Marko Cvitaš, Research Associate
Dr. Janez Mavri, Senior Scientist

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Kemijski odsjek

Doktorska disertacija

PROŠIRENI SAŽETAK

Apsorpcija elektromagnetskog zračenja omogućava brojne kemijske reakcije koje se u osnovnom stanju ne mogu odvijati ili imaju vrlo nizak prinos. Mehanizmi tih fotokemijskih reakcija odvijaju se preko pobuđenih stanja putevima koji nisu termalno dostupni te je njihovo razumijevanje predmet temeljnih istraživanja u kemiji. S teorijske strane takva istraživanja su otežana činjenicom da fotokemijske reakcije, po definiciji, uključuju više od jednog elektronskog stanja. To znači da tijekom reakcije dolazi do sloma Born-Oppenheimerove aproksimacije na kojoj je bazirana većina polja kvantne kemije.

Blizu koničnih presjeka (regija degeneracije ploha potencijalne energije) male promjene u valnom paketu jezgara uzrokuju velike promjene u elektronskoj strukturi zbog čega se ne mogu potpuno odvojiti elektronski i nuklearni stupnjevi slobode u Schrödingerovoj jednadžbi. Za teorijski opis dinamike kroz takve presjeke potrebno je uzeti u obzir neadijabatska sprezanja među elektronskim stanjima. Metode kojima se taj problem može tretirati uključuju metode kvantne dinamike te miješane klasično kvantne metode u kojima se jezgre razmatraju klasično (kao roj trajektorija), ali se neadijabatski efekti uključuju kao modifikacija plohe potencijalne energije. Među metodama kvantne dinamike najpopularnija je multi-konfiguracijska vremenski ovisna Hartree metoda (eng. *Multi-Configuration Time-Dependent Hartree*, MCTDH) i njene varijante.[1,2] Među miješanim klasično kvantnim metodama najpopularnija je Tullyeva metoda skokova među plohama s minimalno obrata (eng. *Fewest-Switches Surface Hopping*, FSSH)[3,4] u kojoj se ploha potencijalne energije mijenja „skokovima“ u različita elektronska stanja ovisno o neadijabatskim sprezanjima među stanjima. Uz navedene metode, često se koristi i *ab initio* metoda višestrukog stvaranja (eng. *Ab Initio Multiple Spawning*, AIMS)[5,6] koja se prema razini aproksimacija nalazi između kvantnog i SH pristupa neadijabatskoj dinamici.[7]

Kod kvantnih metoda glavni problemi su njihovo skaliranje s brojem stupnjeva slobode te potreba za prethodnim generiranjem globalnih dijabatskih ploha potencijalne energije. Miješane klasično kvantne metode same po sebi nisu računalno zahtjevne (gotovo cijela cijena ukupnog računa su izračuni elektronske strukture) te imaju veliku prednost u činjenici da se mogu računati „u letu“. S druge strane, SH metode bazirane su na većem broju aproksimacija koje mogu značajno utjecati na kvalitetu dobivenih rezultata.[13,14] Zbog rada s nezavisnim trajektorijama, SH metode ne uključuju dekoherenciju i kvantne efekte jezgara (eng. *Nuclear Quantum Effects*, NQE) kao što su energija nulte točka i tuneliranje. Zbog navedenih nedostataka, aktivno se radi na unapređivanju obiju vrsta metoda za neadijabatsku dinamiku. Tako su razvijene aproksimativne kvantne metode kojima se mogu tretirati veći sustavi (ML-MCTDH)[8–10] ili se mogu računati „u letu“ (odnosno bez prethodnog računanja ploha), npr. varijacijski multi-konfiguracijski Gaussijani (vMCG).[11,12] FSSH i slični algoritmi unaprijeđeni su: efikasnijim metodama za računanje neadijabatskih sprezanja,[15–17] metodama za rješavanje numeričkih problema kod naglih križanja stanja,[18–22] uključivanjem dekoherencije[23,24] ili djelomičnim uključivanjem NQE. Među potonjima je obećavajuća kombinacija SH metoda s dinamikom polimera prstena (eng. *Ring Polymer Molecular Dynamics*, RPMD).[25–29]

Također, važan je i razvoj metode nuklearnog ansambla za dobivanje opservabli iz roja trajektorija gdje je cilj iz što manje količine podataka (dakle s što manje računa elektronske strukture) dobiti što preciznije opservable.[30–33] Dok se ovom metodom mogu jednostavno dobiti tražene opservable, razumijevanje ovih rezultata u terminima elektronskih svojstava sustava obično nije jednostavan. Razlog tome je da rad u adijabatskoj bazi koji omogućava računanje „u letu“ istovremeno otežava kasniju interpretaciju rezultata u terminima dijabatskih elektronskih stanja koja je lakše izravno povezati s promjenama elektronskih svojstava sustava i s eksperimentalnim zapažanjima.[14]

Cilj ovog rada je evaluacija i unaprjeđenje postojećih miješanih klasično-kvantnih metoda. U tu svrhu implementirane su FSSH i Landau-Zener metoda skokova među plohama (LZSH) te su detaljno ispitane kroz usporedbu s metodama kvantne dinamike.[34] Također, razvijeni su efikasni algoritmi za računanje integrala preklapanja valnih funkcija[35] koji su potrebni u FSSH algoritmu, ali i u novo razvijenoj metodi za analizu nuklearnog ansambla u terminima dijabatskih stanja.[35,36] U konačnici, razvijene metode korištene su za dobivanje novih spoznaja o ključnim primjerima fotokemijskih procesa.[37]

Metode skokova među plohama

Za propagaciju koordinata jezgara u klasičnoj adijabatskoj molekulskoj dinamici potrebni su samo energija i gradijent elektronskog stanja. Osim tih vrijednosti, za FSSH dinamiku su potrebna sprezanja među stanjima koja se mogu dobiti u obliku vektora neadijabatskog sprezanja ili u obliku skalarnih vremenski deriviranih sprezanja (eng. *Time-Derivative Coupling*, TDC). Ova dva oblika međusobno su povezana preko vektora brzine jezgara te u praksi daju gotovo jednake rezultate.[15,38] Vektori neadijabatskog sprezanja mogu se izračunati za određenu geometriju te je njihov račun implementiran u raznim programima za elektronsku strukturu.[39–42] S druge strane, TDC se moraju računati duž trajektorije zbog eksplicitne ovisnosti o vremenu, najčešće numerički metodom konačnih razlika koristeći integrale preklapanja valne funkcije u uzastopnim koracima.[4,15,16]

Neovisno o vrsti sprezanja koja se koristi, jedan od ključnih problema FSSH metode je prepoznavanje i pravilno tretiranje brzih promjena neadijabatskih sprezanja. Energije i gradijenti elektronskih stanja mijenjaju se glatko između vremenskih koraka u dinamici te se bez straha od velikih numeričkih grešaka može koristi vremenski korak od 0.5 do 1.0 fs. S druge strane, neadijabatska sprezanja mogu se mijenjati na istoj vremenskoj skali, ali i mnogo brže. Čak i uz vrlo kratki vremenski korak, neadijabatsko sprezanje može unutar koraka značajno narasti i ponovno se smanjiti; kada je interakcija među stanjima zanemariva sprezanje među njima može nalikovati delta funkciji. Takve nagle promjene sprezanja nazivaju se trivijalni prijelazi. Do sada su razvijene brojne metode temeljene na računu preklapanja valnih funkcija koje dijelom ili potpuno rješavaju taj problem.[18–22] Kako bi se te metode mogle ispravno koristiti potrebno je i pravilno pratiti fazu valnih funkcija između koraka u dinamici što u generalnom slučaju također nije jednostavan problem.[43]

Za male sustave račun preklapanja valnih funkcija je značajno brži od samog računa elektronske strukture, ali skaliranje računa je lošije te bez aproksimacija ili efikasne implementacije može postati najskuplji dio ukupnog računa. Razlog za loše skaliranje je računanje integrala između funkcija koje su razapete različitim baznim skupovima (pomakom jezgara pomiču se i bazne funkcije), čije funkcije nisu međusobno ortogonalne niti pokrivaju isti prostor. Zbog toga je nužno uzeti u obzir valne funkcije u njihovoj potpunoj složenosti te eksplicitno računati preklapanja Slaterovih determinanti (SD) od kojih su valne funkcije građene.[15,16] Računi takvih integrala nisu standardni dio alata u polju računa elektronske

strukture te ih je potrebno odvojeno implementirati (zajedno sa sučeljima za željene metode za elektronsku strukturu) u programima za neadijabatsku dinamiku.

Ovaj korak je najsloženiji dio implementacije FSSH algoritma za nove metode i programe za račun elektronske strukture. Konkretno, neadijabatska sprezanja rijetko su dostupna u ranim implementacijama modernih metoda, a valne funkcije često nisu dostupne u obliku iz kojeg se mogu lako izračunati integrali preklapanja. Dok postoje pokušaji da se potrebne vrijednosti naprave lakše dostupnima iz samih programa za elektronsku strukturu,[44–47] u praksi nedostatak standarda za ispis atomskih/molekulskih orbitala te koeficijenata valnih funkcija predstavlja značajnu prepreku u izradi sučelja između FSSH algoritma i različitih programa za elektronsku strukturu.

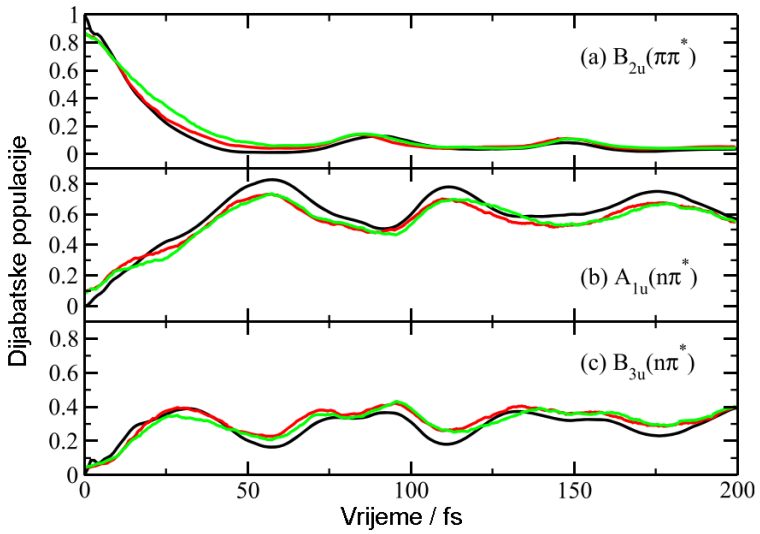
S druge strane, trivijalno je napraviti sučelje između bilo kojeg programa za elektronsku strukturu i metode kojoj nisu potrebna neadijabatska sprezanja u bilo kakvom obliku. U novije vrijeme razvijene su metode za neadijabatsku dinamiku koje koriste Landau-Zener (LZ)[48–50] ili Zhu-Nakamura formule[51,52] za čiju su upotrebu potrebne samo energije i gradijenti elektronskih stanja. Ove metode već se dugo koriste u proučavanju sudarnih procesa i jednostavnih modelnih sustava, ali interes za njih je obnovljen kada je LZ formula preformulirana u terminima adijabatskih ploha potencijalne energije čime je omogućeno njeno korištenje u jednostavnoj implementaciji neadijabatske dinamike „u letu“.

Dok je jednodimenzionalni model s dva stanja iz kojeg je izvedena LZ formula ekstremno jednostavan, aproksimacije tog modela (uska regija interakcije između samo dva stanja, linearna promjena dijabatskih energija u području križanja, stalna brzina tijekom prolaska kroz križanje) su iznenađujuće dobro zadovoljene u velikom broju stvarnih sustava zbog uske lokaliziranosti koničnih presjeka. Metoda je metoda pokazala obećavajuće rezultate na modelnim sustavima, no još uvijek nedostaje detaljnih ispitivanja metode na realnim sustavima u punoj dimenzionalnosti.

Evaluacija SH metoda

Kako bi se potvrdila prikladnost miješanih klasično-kvantnih metoda za sustave koji će se proučavati, prvo je napravljena usporedba s kvantnom dinamikom na MCTDH razini za najbolji dostupni model ploha pirazina koji se može računati kvantnom dinamikom.[34] Ovaj modelni potencijal uključuje tri najniža pobuđena stanja ($B_{3u}(n\pi^*)$, $A_{1u} * (n\pi^*)$ i $B_{2u}(\pi\pi^*)$) i devet najrelevantnijih vibracijskih stupnjeva slobode.[53,54]

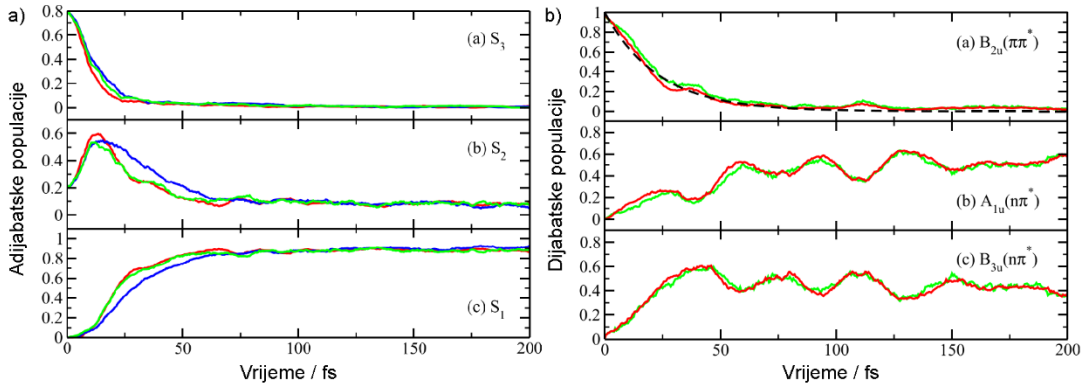
Slika I prikazuje populacije stanja dobivene različitim metodama. Rezultati dobiveni FSSH i LZSH metodama međusobno se gotovo uopće ne razlikuju te obje metode kvalitativno odgovaraju referentnim rezultatima dobivenim MCTDH metodom (uz blago podcjenjivanje populacije $A_{1u}(n\pi^*)$ stanja). Nakon početnog prijenosa populacije u $n\pi^*$ stanja, vidljive su oscilacije populacije između tih stanja. Klasično-kvantne metode dobro reproduciraju i ove oscilacije.



Slika I: Dijabatske populacije najniža tri pobuđena stanja pirazina dobivene MCTDH (crna), FSSH (zelena) i LZSH (crvena) metodama na modelnom potencijalu s tri stanja i devet nuklearnih stupnjeva slobode.

Time je pokazano da, za ovu reakciju, LZSH dinamika daje jednako dobre rezultate kao i FSSH. Međutim, ovi početni rezultati dobiveni su za niskodimenzionalni modelni sustav koji je nastao regresijom na jednostavni polinomni oblik. Pravi ispit LZSH metode su računi u punoj dimenzionalnosti na realnim plohama potencijalne energije. U tu svrhu provedeni su računi neadijabatske dinamike pirazina na ADC(2)/aug-cc-pVDZ razini teorije FSSH i LZSH metodama. U ovom slučaju usporedba je rađena s dvije različite verzije FSSH algoritma, uobičajenom adijabatskom formulacijom (A-FSSH) te u formalizmu lokalne diabatizacije (LD-FSSH) prilikom propagacije elektronske valne funkcije (Slika II).

Kao i kod modelnog sustava, vidljivo je gotovo potpuno slaganje između LD-FSSH i LZSH metoda. A-FSSH blago odstupa od ostalih metoda u tome što je potrebno nešto više vremena da bi se populiralo S_1 stanje. Slaganje je vidljivo i u dijabatskoj bazi (u ovom slučaju određenoj na temelju velike razlike u oscilatorskoj snazi među stanjima) u kojoj su jasno vidljive i oscilacije u populacijama.



Slika II: (a) Adijabatske i (b) dijabatske populacije tri najniža pobuđena stanja pirazina dobivene A-FFSH (plava), LD-FSSH (zelena) i LZSH (crvena) metodama na ADC(2)/aug-cc-pVDZ razini teorije.

Iako su populacije gotovo identične, razlika između LD-FSSH i LZSH metoda vidljiva je kada se promatraju sami skokovi. U LZSH dinamici se skokovi u prosjeku događaju kod manje razlike energija između stanja i gotovo nikada ne dolazi do skokova većih od 0,5 eV, dok u LD-FSSH dinamici (rijetko) dolazi do skokova i kada su stanja udaljena za 1 eV. Ova razlika je i očekivana, LZ vjerojatnost evaluira se samo u minimumu razlike energija stanja, dok je u FSSH dinamici neadijabatsko sprezanje različito od nule u širem području oko koničnog presjeka radi čega do skokova u prosjeku dolazi pri nešto višim energijama.

Ovi rezultati daju obećavajuću sliku za LZSH dinamiku čime bi se omogućila istraživanja dinamike u pobuđenim stanjima uz račun samo energija i gradijenata elektronskih stanja. U tom slučaju, pojednostavljeno je korištenje modernih metoda za elektronsku strukturu gdje nisu razvijeni algoritmi za analitički račun neadijabatskih sprezanja te često nije jednostavno izraziti valnu funkciju u obliku iz kojeg bi se jednostavno izračunala preklapanja među susjednim točkama u dinamici. Međutim, slično ispitivanje za složeniji fotokemijski proces dalo je veće razlike između metoda što ukazuje da se LZSH metoda mora pažljivo primjenjivati na sustave u kojima ne dolazi do interakcije većeg broja stanja u uskoj regiji ploha potencijalne energije.

Račun integrala preklapanja valnih funkcija

Ključan dio ovog rada je vrlo efikasna implementacija[35] računa integrala preklapanja pobuđenih valnih funkcija za konkretan slučaj valnih funkcija oblika konfiguracijske interakcije s jednostrukim pobudama (engl. *Configuration Interaction Singles*; CIS) što uključuje pomoćne valne funkcije dobivene vremenski-ovisnom teorijom funkcionala gustoće (TDDFT).[55,56] Isti algoritam može se, uz zanemarivanje dvostrukih pobuda koje manje

doprinose valnoj funkciji, koristiti i za metodu algebarske diagramatske konstrukcije drugog reda (ADC(2))[57–59] čime su pokrivene najčešće metode zasnovane na jednoj referentnoj konfiguraciji koje se koriste u računima neadijabatske dinamike. Ovi integrali imaju oblik

$$S_{AB} = \langle \Psi_A | \Psi'_B \rangle = \sum_o \sum_p d_o^A d'^B_p \langle \Phi_o | \Phi'_p \rangle$$

gdje indeksi A i B označavaju elektronska stanja s valnim funkcijama Ψ izgrađenim kao linearne kombinacije Slaterovih determinanti Φ_o s koeficijentima d_o^A. Apostrofom su označene sve vrijednosti vezane za stanje B koje je definirano na različitoj geometriji od A. Za CIS valne funkcije indeksi o i p označavaju sve moguće kombinacije pobuda iz okupiranih i virtualnih molekulskih orbitala za svako od stanja. Dakle, skaliranje ovog računa je $O(n_{occ}^5 n_{virt}^2)$ gdje n_{occ}^3 dolazi iz računa integrala $\langle \Phi_o | \Phi'_p \rangle$.

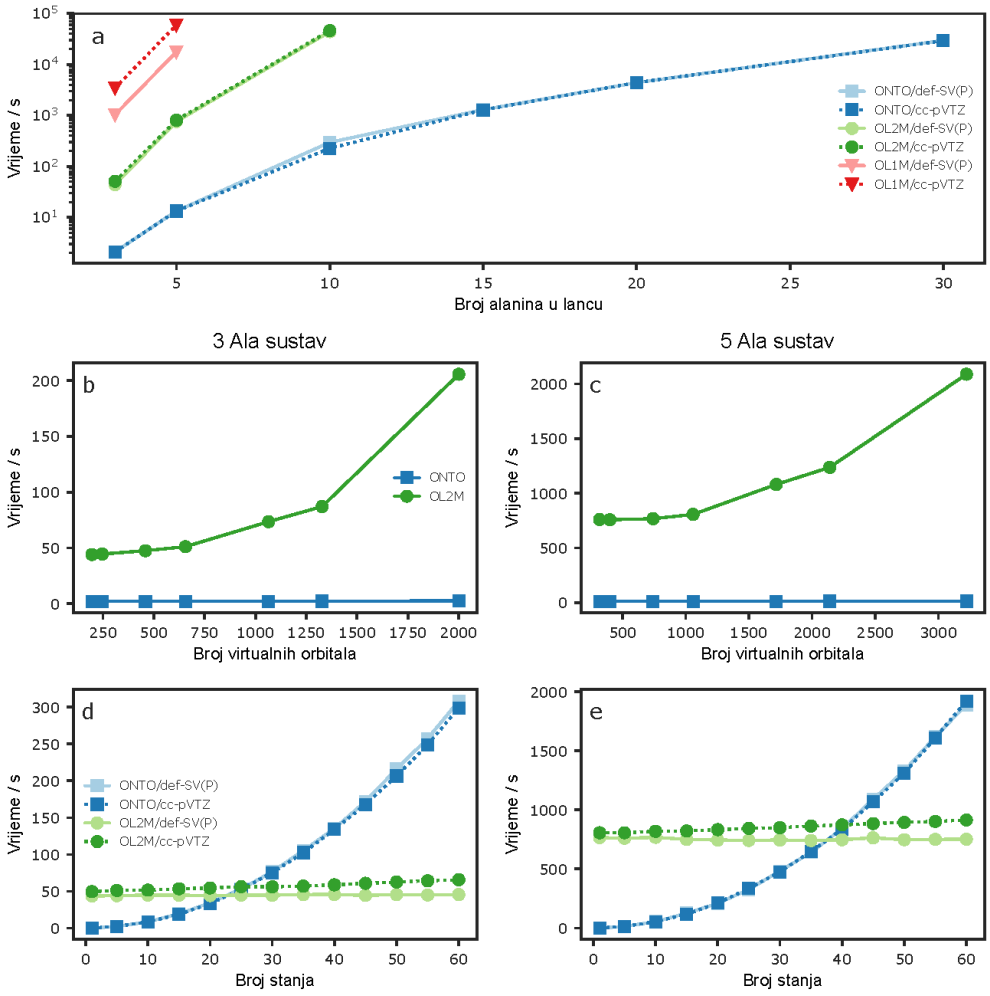
Prvi efikasniji pristup ovoj jednadžbi zasnovan je na činjenici da se većina Slaterovih determinanti koje se pojavljuju u gornjoj sumi razlikuje u samo jednom ili dva reda/stupca. Ta činjenica se može iskoristiti raspisivanjem determinanti u minore drugog stupnja te predračunavanjem tih minora u računu integrala (ovaj algoritam biti će označen s OL2M). Drugi algoritam baziran je na prikazu valnih funkcija u kompaktnoj bazi prirodnih prijelaznih orbitala (engl. *Natural Transition Orbitals*, NTO). U toj bazi CIS valne funkcije imaju oblik

$$|\Psi_A\rangle = \sum_o^{n_{occ}} \lambda_o^A |\Theta_o^A\rangle$$

gdje suma ide samo po okupiranim orbitalama zbog toga što prijelazne orbitale dolaze u parovima tako da je pobuda iz okupirane prijelazne orbitale o moguća samo u virtualnu prijelaznu orbitalu o (algoritam koji koristi ovu bazu biti će označen s ONTO). Time cijena računa jednog integrala preklapanja pada na $O(n_{occ}^5)$.

Slika III prikazuje značajno ubrzanje računa preklapanja valnih funkcija razvijenim algoritmima. Dok neoptimiziranim algoritmi bez dodatnih aproksimacija postaju skuplji od TDDFT pa čak i ADC(2) računa elektronske strukture već za sustave s 20-30 atoma, ONTO algoritam, zahvaljujući manjem (n^5) skaliranju, može se koristiti i za sustave s 300 atoma odnosno 600 okupiranih orbitala prije nego njegova cijena postane usporediva sa standardnim TDDFT računom. Osim za velike sustave, oba razvijena algoritma pogodna su i za račune s velikim baznim skupovima. U tom pogledu, ONTO gotovo da uopće nema ovisnost o baznom skupu, dok OL2M tek za vrlo velike bazne skupove počinje pokazivati kvadratnu ovisnost o broju virtualnih orbitala (kada korak računa preklapanja iz minora postane skuplji od samog

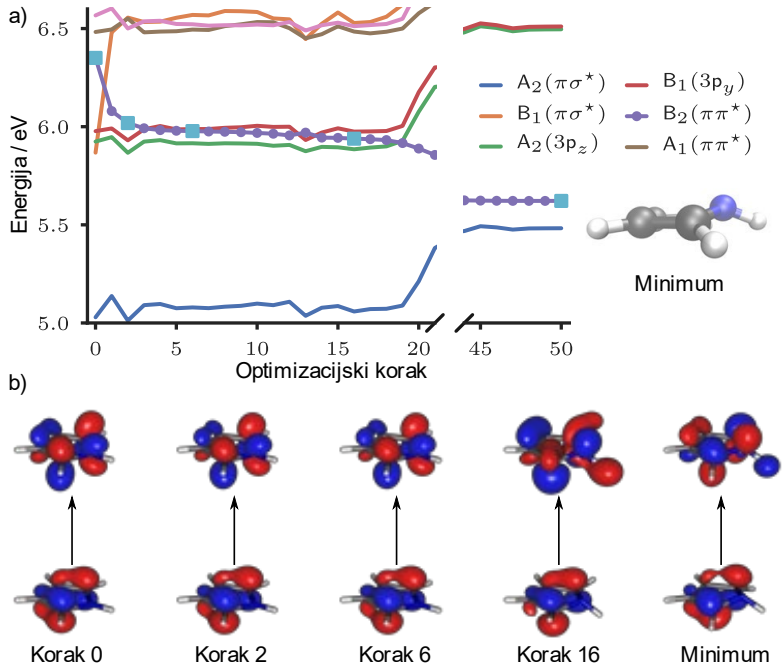
računa minora). Na kraju se još ispituje i skaliranje algoritama s brojem stanja (dimenzijom matrice preklapanja). Kao što je očekivano, ONTO algoritam u ovom slučaju pokazuje kvadratnu ovisnost zbog toga što svako elektronsko stanje ima različite NTO. S druge strane, OL2M ima vrlo slabu ovisnost o broju stanja te se pokazuje kao brži algoritam kada je broj stanja dovoljno velik. To znači da se preporuča korištenje OL2M algoritma za sustave s velikom gustoćom stanja ili sustave koji se pobuđuju u visokoenergetska stanja. Vrijedi spomenuti i da se OL2M algoritam može dodatno optimizirati efikasnim ažuriranjem intermedijera tijekom računa determinanti minora. Nakon implementacije takvog ažuriranja[60] dobiva se otprilike deseterostruku ubrzanje algoritma čime isplativiji od ONTO već za značajno manji broj stanja.



Slika III: Usporedba cijene izvođenja OL1M (crvena), OL2M (zelena) i ONTO (plava) algoritama. (a) Skaliranje s veličinom sustava (brojem peptide u lancu) za pet pobuđenih stanja. Dodatni testovi izvedeni su za sustav građen od 3 alanina (b, d) i 5 alanina (c, e). Vremena izvođenja su prikazana za pet pobuđenih stanja i različitim baznim skupovima od def-SV(P) pa sve do aug-cc-pVQZ (b, c) te s fiksnim baznim skupom i različitim brojem stanja (d, e).

Razvijeni algoritmi mogu se koristiti, uz zanemarivu cijenu, pri računima neadijabatske dinamike za bilo koji od sustava koji se danas proučavaju na TDDFT razini. Međutim, aktivno se razvijaju nove i efikasnije implementacije računa za elektronsku strukturu te se brzo pomiče granica veličine sustava koji se mogu proučavati.[61–63] U slučaju kada račun preklapanja ponovno postane usko grlo u računu dinamike, mogu se bez problema uvesti aproksimacije. Jednostavan pristup je samo „skratiti" valnu funkciju tako da se uzimaju najveći koeficijenti dok njena norma ne postane veća od zadanog praga. Ispitivanjem moguće greške u izračunatim neadijabatskim sprezanjima tijekom nasumično odabranih FSSH trajektorija pirola na ADC(2)/aug-cc-pVDZ i B3LYP/def2-TZVPD razinama, pokazano je da je srednja greška u izračunatim preklapanjima vrlo mala, ali da su u rijetkim slučajevima moguće i velike greške (reda $2\sqrt{1-t}$, gdje je t norma odrezane valne funkcije). Unatoč tome, uz odabir dovoljno velikog praga za normu ($0{,}999$ ili više), mogu se dobiti precizni rezultati uz smanjenje broja determinanti koje se moraju izračunati za nekoliko redova veličine. Ovakva aproksimacija posebno je prikladna uz ONTO algoritam koji koristi kompaktni zapis valnih funkcija u terminima NTO gdje uvijek dominira mali broj prijelaza.

Osim u neadijabatskoj dinamici, dobiveni integrali preklapanja mogu se koristiti općenitije u bilo kojem slučaju kada je potrebno pratiti pojedino elektronsko stanje kroz promjene nuklearnih koordinata, na primjer za asignaciju spektara generiranih metodom nuklearnog ansambla u slučajevima kada se više elektronskih stanja nalazi u istom energetskom rasponu. Također, praćenje stanja je potrebno tijekom optimizacije viših pobuđenih elektronskih stanja čiji minimumi se vrlo često nalaze ispod koničnih presjeka s jednim ili više pobuđenih stanja koja se nalaze na nižoj energiji u FC regiji. Kako bi se takve optimizacije uspješno izvršavale, modificiran je standardni optimizacijski algoritam. Tijekom optimizacije se nakon svakog koraka računa preklapanje između valne funkcije ciljanog stanja u početnoj točci i svih stanja u trenutnom koraku te se optimizacija nastavlja uvijek prateći gradijent onog stanja čija je valna funkcija najbliža onoj traženog stanja. Slika IV prikazuje optimizaciju svijetlog $B_2(\pi\pi^*)$ stanja pirola koje je peto pobuđeno stanje prilikom vertikalne pobude, ali drugo pobuđeno stanje u minimumu. Koristeći modificirani algoritam, optimizacija bez problema u svakom trenutku prati traženo stanje te uspješno pronalazi njegov minimum bez potrebe za ručnim mijenjanjem stanja.



Slika IV: (a) Optimizacija $B_2(\pi\pi^*)$ stanja pirola na B3LYP/def2-TZVPD razini teorije s prijelazom kroz tri konična presjeka tijekom prvih 20 koraka. (b) Dominantne NTO kontribucije trenutnom ciljanom stanju duž optimizacijskog puta (u točkama označenim svijetlo plavim kvadratima).

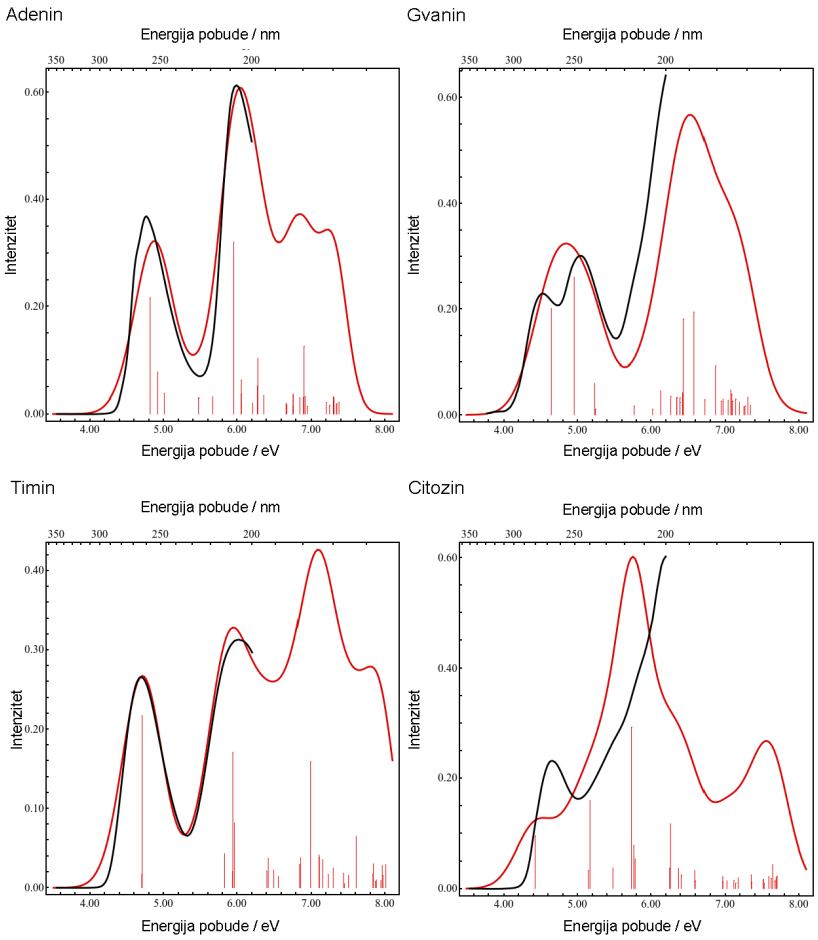
Dekompozicija izračunatih spektara

Kao što je spomenuto u uvodu, FSSH i ostale metode koje računaju neadijabatsku dinamiku „u letu“ ne daju jednostavan način za promatranje rezultata u terminima dijabatskih stanja. Uz dostupne integrale preklapanja valnih funkcija na različitim geometrijama, ta se činjenica donekle može promijeniti. Glavna karakteristika dijabatskih stanja je da se ne mijenjaju (odnosno da se sporo mijenjaju) s promjenom nuklearne konfiguracije, a preklapanja valnih funkcija nam daju kvantitativnu mjeru te promjene. Stoga se, uz poznavanje stanja na jednoj geometriji stanja, na svim ostalim geometrijama mogu raspisati kao linearne kombinacije tih stanja što se može koristiti kao prva aproksimacija prijelaza iz adijabatske u dijabatsku bazu. Na temelju toga, definiraju se dvije metode za dekompoziciju opservabli dobivenih metodom nuklearnog ansambla. U prvoj, kao „dijabatska“ stanja uzimaju se samo adijabatska stanja na pojedinoj geometriji koja imaju najveće preklapanje sa stanjima na referentnoj geometriji koja su uzeta kao definicija dijabatskih stanja. U drugoj metodi, svako stanje na svakoj geometriji $|\Psi_F(\mathbf{R}_i)\rangle$ doprinosi svim „dijabatskim“ stanjima, ali uteženo koeficijentom tog adijabatskog stanja u bazi referentnih stanja $c_{AF}(\mathbf{R}_i) = \langle\Psi_A(\mathbf{R}_0)|\Psi_F(\mathbf{R}_i)\rangle$.

Ove metode iskorištene su za promatranje UV apsorpcijskih spektara nukleobaza DNK (adenina, timina, citozina i gvanina) u vodi. Konkretno, omogućena je analiza utjecaja okoline

na pojedina elektronska stanja sistematskim dodavanjem sve složenijih efekata u račun nuklearnih ansambala i energija pobude.

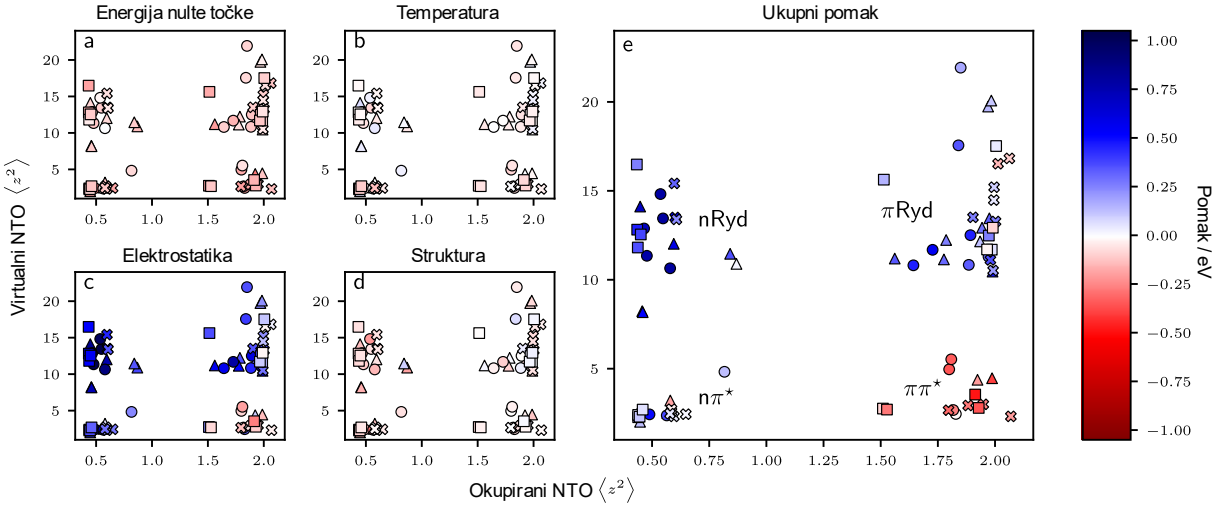
Slika V prikazuje dobivene spektre koji su u vrlo dobrom slaganju s eksperimentalnim spektrima vodenih otopina tih molekula, pogotovo u slučaju adenina i timina. Prikazani linijski spektar sadrži informacije o prosječnim energijama i oscilatorskim snagama pojedinih stanja. Koristeći podatke o dominantnim NTO parovima, jednostavno je asignirati sve linije unutar spektara. Kod citozina primjećujemo najveća odstupanja od eksperimentalnih mjerenja, ali nešto bolji rezultati dobivaju se nakon optimizacije citozina u COSMO okolini u kojoj NH_2 grupa prelazi iz piramidalne u gotovo planarnu konformaciju.



Slika V: Eksperimentalni (crna linija) i simulirani (crvena linija, ADC(2)/aug-cc-pVDZ razina) spektri nukleobaza DNK u vodi. Linijski spektar prikazuje prosječne energije pojedinih stanja unutar ansambla.

Zanimljivo je i proučiti kako se ovi spektri, koji uzimaju u obzir gibanje jezgara oko minimuma te utjecaj otapala, razlikuju od energija vertikalnih prijelaza koje se dobivaju iz jednostavnog računa pobuđenih stanja iz geometrije minimuma u vakuumu. U tu svrhu izračunati su spektri nukleobaza koji postepeno uključuju sve složenije efekte okoline (Slika

VI) te su praćenjem preklapanja valnih funkcija asignirana stanja svake geometrije svakog ansambla. Kao prvi korak, prijelaz s geometrije minimuma na nuklearni ansambl pri 0K (uključenje energije nulte točke) pomiče sva stanja prema crvenom dijelu spektra, u prosjeku za ≈ 0,1 eV. Takav učinak je očekivan, pomaci iz minimuma osnovnog stanja u prosjeku osnovno stanje destabiliziraju više nego pobuđena stanja. Zatim je promatrana promjena temperature s 0 K na 298 K koja ima vrlo mali učinak na same spektre.



Slika VI: Utjecaj temperature i okoline na energije pobude različitih pobuđenih stanja. Prikazani su pomaci od vertikalne energije pobude uz postepeno dodavanje (a) energije nulte točke, (b) temperature, (c) elektrostatskog učinka otapala, (d) efekta otapala na strukturu molekule te konačno (e) svih efekata zajedno. Za razlikovanje stanja različitog tipa, kao osi su korištene očekivane vrijednosti $\langle z^2 \rangle$ dominantnih ispražnjenih i ispunjenih oribtala prema kojima se stanja jasno grupiraju u $n\pi^*$, $\pi\pi^*$, πRyd i nRyd tipove. Boje točaka označavaju pomak pojedinog stanja (u eV) u odnosu na odgovarajuće stanje u vakuumu.

S druge strane, elektrostatski efekt otapala je vrlo značajan i nehomogen (Slika VIc). Kada se energije pobude iz ansambla na 298 K računaju u COSMO okolini dolazi do pomaka nπ*, nRyd i πRyd stanja prema plavom dijelu spektra za 0,38, 0,69 i 0,34 eV dok s ππ* stanja blago pomiču (0.03 eV). Ovi pomaci najviše ovise o promjeni dipolnog momenta prilikom prelaska iz osnovnog u pobuđeno stanje. Veće promjene dipolnog momenta nisu povoljne u otapalu te se takva stanja više pomiču prema plavom. Uz elektrostatski doprinos, otapalo utječe i na ravnotežnu geometriju molekule. Taj efekt proučen je optimizacijom osnovnog stanja u COSMO okolini te generiranjem novog ansambla koristeći taj minimum (Slika VId), čime se prosječne energije pobude nisu značajno promijenile.

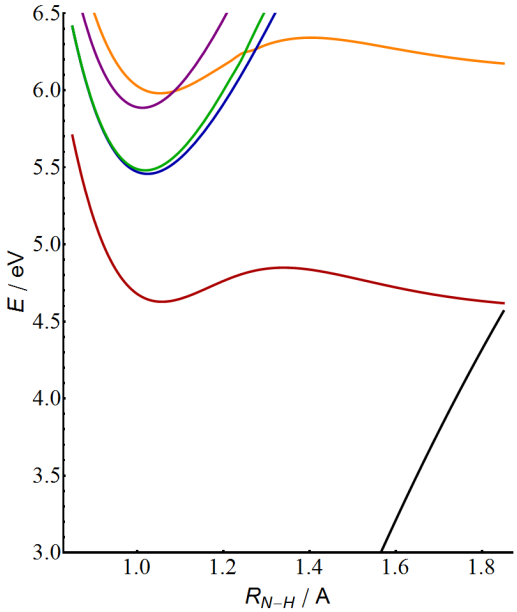
Promatrani pomaci u spektru kvalitativno su jednaki za sve nukleobaze te se zaključci mogu primijeniti i na druge molekule sa istim tipovima pobuđenih stanja. Osim nabrojenog, iznimno važan utjecaj na spektar mogu imati i vodikove veze između molekule i otapala. Energije i oscilatorske snage stanja s pobudama blizu atoma koji tvore vodikove veze mogu

se vrlo značajno promijeniti. Međutim, za promatranje tog učinka potrebno je eksplicitno uključiti molekule otapala tijekom generiranja nuklearnog ansambla i računa elektronske strukture, što je van opsega trenutnog rada.

Pirol

Uz razvijene metode moguće je vratiti se na neke od osnovnih fotokemijskih reakcija i sagledati ih iz nove perspektive. Jedna od tih reakcija je disocijacija vodika putem $\pi\sigma^*$ stanja nakon pobude pirola UV zračenjem.[64,65] Unatoč velikom broju radova koji su proučavali ovaj problem, dugo vremena je nedostajao kvantitativan opis ovisnosti brzine ove reakcije o valnoj duljini zračenja kojim je molekula pobuđena.[66]

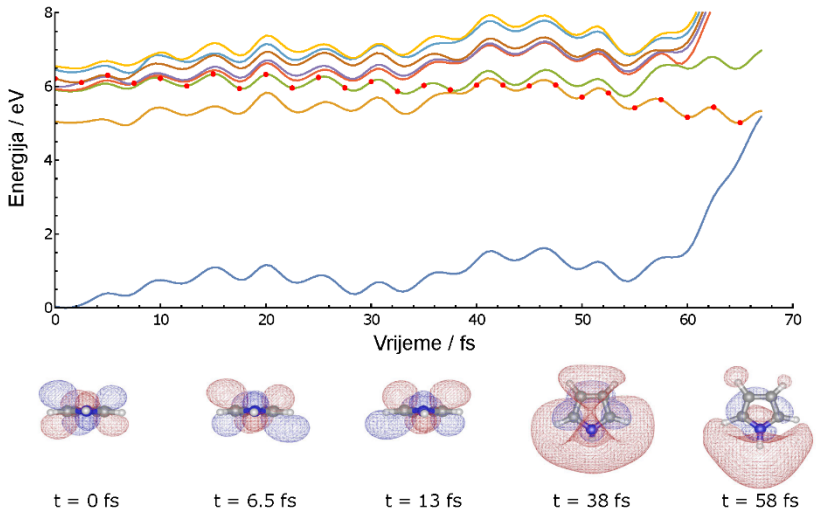
Ova reakcija istražena je simulacijama neadijabatske dinamike u širokom rasponu valnih duljina $(200-250\ \text{nm})$[37] koristeći kombinaciju B3LYP/def2-TZVPD i ADC(2)/aug-cc-pVDZ razina teorije. Za opis reakcije nakon pobude na 250 nm ključan je opis minimuma najnižeg stanja, $A_2(\pi\sigma^*)$, u koji se izravno pobuđuje molekula pri ovoj valnoj duljini. To stanje u regiji oko minimuma ima $\pi 3p_z$ karakter, ali izduživanjem N-H veze prelazi u disocijativno $\pi\sigma^*$ stanje (Slika VII). Pokazano je da TDDFT (koristeći razne funkcionale) ne može dobro opisati ovaj minimum, zbog čega u simulacijama dinamike uvijek daje previše brzu disocijaciju nakon pobude.



Slika VII: Presjek ploha potencijalne energije duž koordinate N-H istezanja pri čemu su sve ostale koordinate optimizirane u S_1 stanju na ADC(2)/aug-cc-pVDZ razini teorije.

S druge strane, na ADC(2) razini zbog prevelikog miješanja $\pi\pi^*$ (valentnog) i $\pi 3p_x$ (Rydberg) karaktera nije dobro opisana dinamika nakon pobude u svijetlo B_2 čiji je minimum

bolje opisan na TDDFT razini. Ovo stanje je peto pobuđeno stanje u geometriji minimuma osnovnog stanja, ali na TDDFT razini se može vidjeti da gradijent vodi ovo stanje mimo ostalih stanja u toj regiji dok ne postane S_2 stanje, poviše samo $A_2(\pi\sigma^*)$ stanja. Od tu je moguća vrlo brza interna konverzija izravno iz $B_2(\pi\pi^*)$ u $A_2(\pi\sigma^*)$ stanje. Taj mehanizam vidljiv je u dinamici promatranjem elektronskih stanja duž reprezentativnih trajektorija (Slika VIIIb).

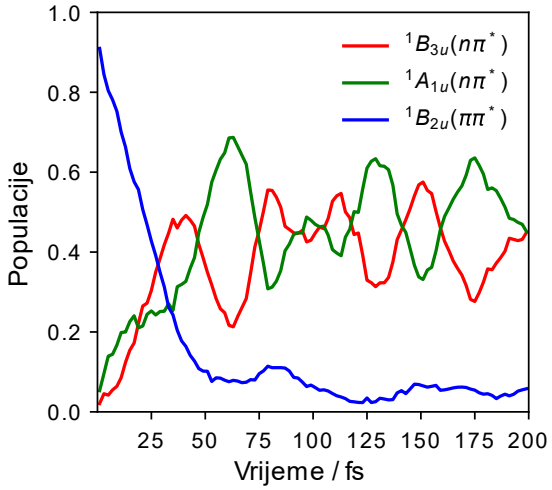


Slika VIII: Potencijalne energije osnovnog stanja i devet najnižih pobuđenih stanja pirola duž reprezentativne TDDFT/B3LYP/def2-TZVPD trajektorije. Trenutno populirano stanje prikazano je crvenim točkama, a dominantne virtualne NTO tog stanja u različitim trenutcima prikazane su ispod.

Pirazin

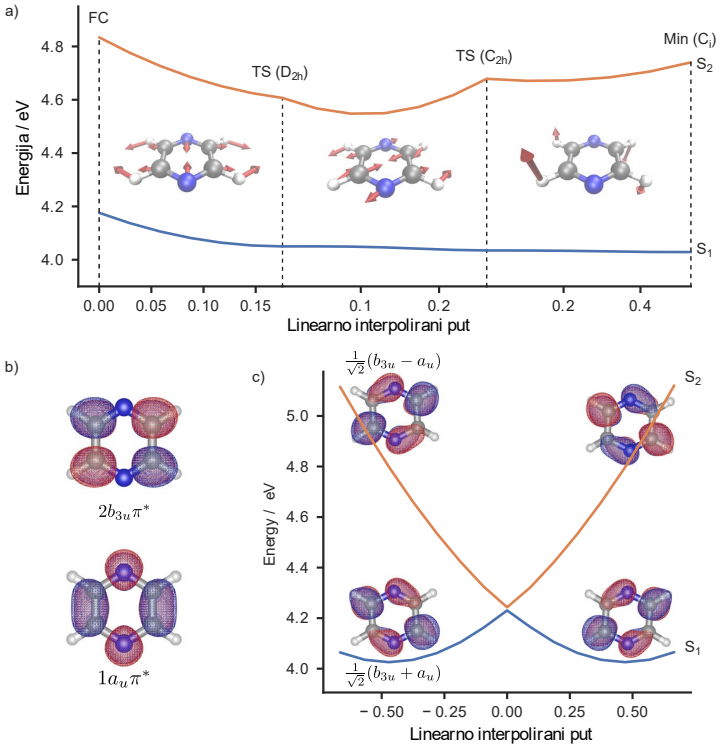
Iduća fotokemijska reakcija koja se promatra je interna konverzija u nakon pobude pirazina u $B_{2u}(\pi\pi^*)$ stanje.[67–70] Za ovu konverziju dugo se smatralo da ide direktno u najniže stanje, $B_{3u}(n\pi^*)$, te je zbog svoje jednostavnosti često korištena za testiranje novih metoda za kvantnu dinamiku. Međutim, u novije vrijeme otvorena je rasprava o sudjelovanju još jednog stanja u fotokemijskom procesu, $A_u(n\pi^*)$ stanja koje dio teorijskih metoda smješta između ova dva dobro istražena stanja.[53,54,71–73] Sudjelovanje ovog stanja do sada nije promatrano u punoj dimenzionalnosti, jer predstavlja problem za koji je potrebno dobiti dijabatske populacije iz simulacija neadijabatske dinamike u adijabatskoj bazi.

Stoga je napravljena dinamika za ovaj sustav na ADC(2)/aug-cc-pVDZ razini te su izračunate „dijabatske“ populacije stanja gore opisanom metodom (Slika IX). Iz ovoga se jasno vidi da je $A_u(n\pi^*)$ stanje populirano vrlo brzo tijekom procesa i to u jednakoj razini kao i $B_{3u}(n\pi^*)$ stanje.



Slika IX: Dijabatske populacije pirazina duž ansambla FSSH trajektorija na ADC(2)/aug-cc-pVDZ razini.

Uz činjenicu da su oba stanja populirana, primijećeno je da su koeficijenti ovih stanja na različitim geometrijama slični. To se može objasniti promatranjem minimuma S_1 stanja (Slika X) u kojem su ova dva stanja potpuno pomiješana. Do ovog miješanja dolazi zbog promjene energije orbitala prilikom izduživanja dvije nasuprotne C-N veze u minimumu čime se stabiliziraju π^* orbitale s čvorovima na tim vezama.



Slika X: a) Energije S_1 i S_2 stanja na putu od minimuma osnovnog stanja pirazina D_{2h} simetrije do minimuma S_1 stanja C_i simetrije. Gibanja duž puta prikazana su unutar slike. b) Virtualne NTO S_1 i S_2 u geometriji minimuma osnovnog stanja. c) Put između dva ekvivalentna minimuma S_1 stanja s prikazanim virtualnim NTO S_1 i S_2 stanja koja se mogu dobiti miješanjem orbitala iz minimuma osnovnog stanja.

Naravno, uvijek je moguće da su ova zapažanja rezultat krivog opisa ploha potencijalne energije na odabranoj razini teorije. Kako bi se potvrdilo da to nije slučaj, potrebna je usporedba s eksperimentalnim zapažanjima.

Zaključak

Cilj ovog rada bio je unaprijediti miješane klasično-kvantne metode za proučavanje fotokemijskih procesa.

Napravljena je modularna i fleksibilna implementacija LZSH i FSSH algoritama te su metode evaluirane kroz usporedbu s kvantnom dinamikom i međusobno. Unatoč jednostavnosti LZSH metode, pokazano je da za slučaj interne konverzije u pirazinu daje rezultate koji se odlično slažu s FSSH i MCTDH metodama. Međutim, kod pirola LZSH predviđa značajno brže vrijeme reakcije od FSSH. U konačnici, može se zaključiti da LZSH daje barem kvalitativno točan opis mehanizama u fotokemiji. Time ova metoda postaje važan alat za simulaciju neadijabatske dinamike, jer vrlo često faktor koji ograničava preciznost računa nije metoda za dinamiku već opis plohe potencijalne energije. U takvim slučajevima, LZSH se može lako primijeniti uz naprednije metode za elektronsku strukturu i u slučajevima kada neadijabatska sprezanja ili valne funkcije nisu dostupne.

Također, razvijeni su algoritmi za račun integrala preklapanja valnih funkcija dobivenih TDDFT ili ADC(2) metodama za elektronsku strukturu. Uz ove algoritme potpuno je uklonjen jedini dio FSSH algoritma čija je cijena usporediva s računom elektronske strukture za male i srednje velike molekule.

Osim bržeg računa neadijabatskih sprezanja, razvijeni kod omogućava i lakšu integraciju integrala preklapanja u ostale dijelove teorijskih istraživanja pobuđenih stanja. Tako su predloženi algoritmi za optimizaciju pobuđenih stanja, automatsku asignaciju orbitala i elektronskih stanja te za dekompoziciju spektara izračunatih metodom nuklearnog ansambla na njihove dijabatske kontribucije.

Ove metode iskorištene su za analizu utjecaja okoline na pojedina elektronska stanja DNK nukleobaza. Sistematskim dodavanjem sve složenijih efekata u račun nuklearnih ansambala i energija pobude može se doći do generalnog zaključka da promjena energije nekog stanja u otopini najviše ovisi o vektoru promjene dipolnog momenta između tog stanja i osnovnog stanja.

U konačnici, proučavane su dva prototipa reakcija u fotokemiji, interna konverzija u pirazinu te disocijacija vodika u pirolu. U pirolu je objašnjena ovisnost brzine reakcije o

energiji pobude te je pokazano kako je dobar opis miješanja između valentnih i Rydberg-ovih stanja ključan za kvantitativno točne simulacije. U slučaju interne konverzije pirazina, po prvi put je za dinamiku u punoj dimenzionalnosti pokazano da $A_{1u}(n\pi^*)$ stanje nedvojbeno sudjeluje u ovom procesu. Kako bi se odredile dijabatske populacije stanja te kontribucije pojedinih stanja fotoelektronskom spektru pirazina u različitim trenutcima nakon pobude, korištene su razvijene metode za asignaciju stanja i dekompoziciju spektara. Time je pokazano da je opis dinamike u kojoj sudjeluju tri stanja u potpunosti konzistentan s eksperimentalnim zapažanjima.

§1. INTRODUCTION

When molecules absorb a photon in the ultraviolet (UV) or visible region they are promoted to excited electronic states. In these states they can undergo reactions which are otherwise not possible or unlikely. The mechanisms of these photochemical reaction involve pathways through excited states which are not thermally accessible. Examples of important photochemical reactions can be found in all areas of chemistry. In nature, photosynthesis is used to convert carbon dioxide and water into glucose and oxygen; photochemical reactions such as the isomerization of retinal are the starting point for vision; in humans Vitamin D is formed only through exposure to sunlight. These types of reactions are also used in medicine for photodynamic therapy and in industry for production of various compounds such as benzyl chloride or in various polymerization reactions. Many photochemical reactions have negative effects we wish to avoid such as photodegradation of many materials or damage of DNA by UV radiation. Understanding these reactions at a detailed mechanistic level is an area of significant importance in chemistry.

From the theoretical side, studies of these reactions are made difficult by the fact that, by definition, photochemical reactions involve multiple electronic states which means that a breakdown of the Born-Oppenheimer approximation occurs at some point during the reaction. In the vicinity of conical intersections (CIs, regions where the potential energy surfaces of two or more state become degenerate), small changes in the nuclear wave packet cause large changes in the electronic structure of the system. In such cases, it becomes impossible to fully separate the electronic and nuclear degrees of freedom in the Schrödinger equation. This problem can be treated through quantum dynamics methods such as the multi-configuration time-dependent Hartree (MCTDH) method and its variants[1,2] or mixed classical-quantum methods which treat the nuclei classically and electrons quantum mechanically, but include nonadiabatic effects through interactions between the two. Among these methods the most popular are surface hopping (SH) methods and more specifically Tully's fewest-switches surface hopping (FSSH) method[3,4] in which nuclei move on a single potential energy surface (PES) at any given moment, but "hops" between different surfaces are possible depending on the nonadiabatic coupling between the states.

Each of these classes of methods has distinct advantages and disadvantages. Quantum dynamics can be incredibly accurate, but are very expensive without significant approximations.

They also require the prior construction of global diabatic potential energy surfaces which is another difficult task for which there is no general solution. On the other hand, mixed classical-quantum methods are usually simple to perform "on-the-fly" (without precomputed surfaces) which makes them especially useful when there is no prior knowledge of the possible mechanism. However, they also require very large numbers of evaluations of PESs and do not include nuclear quantum effects (NQE).

Both types of methods are still in active development. From the quantum dynamics side, developments include approximate methods capable of treating very large systems, such as the multi-level MCTDH method[8–10] or "on-the-fly" methods which do not require precomputed surfaces, such as the variational multiconfiguration Gaussian (vMCG).[11,12] From the mixed classical-quantum side, various expansions and corrections to the FSSH algorithm have been developed. These include various improvements to the calculation of the required time-derivative couplings (TDCs),[15–17] solve numerical issues at sudden surface crossings,[18–22] schemes to correct for the lack of decoherence in the FSSH algorithm,[23,24] and many other areas. One promising path to fix a basic weakness of FSSH involves adding some NQEs through connecting the FSSH method with ring-polymer molecular dynamics (RPMD).[25–29]

Another area of development is the nuclear ensemble approach (NEA) through which observables are calculated in mixed classical-quantum simulations. This approach allows calculating observables by averaging over values calculated at points representing the nuclear phase space distribution of the system and is closely related to the overall cost of the calculation since the goal is always to obtain accurate results while running as few calculations as possible.[30–33] However, while we can easily calculate observables, extracting physical insight on the electronic properties of a system from mixed classical-quantum simulations is not often straightforward. The only values connected to the electronic degrees of freedom actively used by the FSSH algorithm are the adiabatic populations and nonadiabatic couplings which, by themselves, give little meaningful insight into the electronic properties of the system. For this reason, even though photochemical reactions are intrinsically driven by changes in the electronic structure of the system upon photoexcitation, studies based on SH dynamics calculations are often focused primarily on the changes in the nuclear degrees of freedom which can be analyzed more directly.

In general, it is possible to obtain a qualitative diabatic population by performing a series of transformations along SH trajectories.[75] However, this approach fails for more gradual changes

of electronic character along trajectories and especially when new states drop into the space being considered. An alternate to this approach are tools for quantitative analysis of excited state wave functions such as those developed by Plasser and coworkers.[76–81] These tools are more robust and informative, but have so far not been employed for analyzing the electronic properties within large nuclear ensembles generated by SH dynamics simulations apart from specific problems such as the analysis of charge transfer in transition metal complexes.[82]

The goal of the current work is to improve the efficiency of mixed classical-quantum dynamics simulations and simplify the analysis of the results we obtain through these simulations. First, in section 4.2 we will present algorithms for highly efficient algorithms for calculating wave function overlaps which is the only part of the FSSH algorithm with a cost comparable to, or higher than, electronic structure calculations. Then, in section 4.3 we will show how these overlaps can be used for characterization of electronic states across nuclear ensembles. This approach finds middle ground between the simple and general, but path-dependent approach suggested by Mai *et al.*[75] and the robust, quantitative but less general methods suggested by Plasser and coworkers.[76–81] Lastly in sections 4.4 and 4.5, we will apply the developed methods to reexamine two prototypical problems in photochemistry, the hydrogen detachment in pyrrole and internal conversion in pyrazine, with a focus on an analysis of the excited states involved in the processes.

§2. LITERATURE OVERVIEW

2.1 Nonadiabatic dynamics

We will begin with the basics of quantum chemistry needed for understanding the concepts in the following section. These topics are covered in detail by many books[83–85] and reviews[6,86] so here we will only give a brief overview following the notation suggested by Malhado.[86] For chemical phenomena where relativistic effects do not play a role, the dynamics of electrons and nuclei can generally be described using the time-dependent Schrödinger equation (TDSE)

$$\hat{H}\,|\Psi\rangle = i\frac{\mathrm{d}}{\mathrm{d}t}\,|\Psi\rangle\,. \tag{2.1}$$

In the above equation, and throughout the text, atomic units are used. For a system consisting of N_n nuclei with coordinates, masses and charges $\{\mathbf{R}_a\}$, $\{M_a\}$ and $\{Z_a\}$, respectively, and N_e electrons with coordinates $\{\mathbf{r}_i\}$, the Hamiltonian operator $\hat{H}$ is

$$\hat{H} = -\frac{1}{2}\sum_a^{N_n}\frac{1}{M_a}\nabla_a^2 - \frac{1}{2}\sum_i^{N_e}\nabla_i^2 + \hat{V} = \hat{T}_N + \hat{T}_e + \hat{V} = \hat{T}_N + \hat{H}_e\,, \tag{2.2}$$

where the potential energy $\hat{V}$ is given by

$$\hat{V} = -\sum_a^{N_n}\sum_i^{N_e}\frac{Z_a}{|\mathbf{R}_a - \mathbf{r}_i|} + \sum_i^{N_e}\sum_j^{N_e}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_a^{N_N}\sum_b^{N_N}\frac{Z_a Z_b}{|\mathbf{R}_a - \mathbf{R}_b|}\,. \tag{2.3}$$

It is not possible to solve equation 2.1 analytically for any molecular system and even numerical solutions for systems with more than a few atoms/electrons are possible only with some approximations.[87] The most common path for solving this problem is to start from the Born-Oppenheimer approximation.[88,89] The starting point for this approximation is the observation that the mass of an electron is over three orders of magnitude smaller than the mass of a proton which results in a significantly faster time scale for the motion of electrons. Taking this into account, it makes sense to separate the "slow" nuclear degrees of freedom and the "fast" electronic degrees of freedom, viewing the full Hilbert space $\mathscr{H}$ as a tensor product between the two subspaces

$$\mathscr{H} = \mathscr{H}_N \otimes \mathscr{H}_e\,, \tag{2.4}$$

and to consider the effect of nuclei on the electrons as slowly changing "external" conditions.

In the limit of frozen (clamped) nuclei, the Schrödinger equation of the electronic Hamiltoninan is

$$\hat{H}_e(\mathbf{R})\,|\psi_n;\mathbf{R}\rangle = E_n(\mathbf{R})\,|\psi_n;\mathbf{R}\rangle\ , \tag{2.5}$$

where $\hat{H}_e$ depends parametrically on the positions of the nuclei $\mathbf{R}$. In this case, its eigenstates $|\psi_n;\mathbf{R}\rangle$ also have a parametric dependence on these coordinates and the eigenvalues $E_n(\mathbf{R})$ are a function of the nuclear positions. These eigenstates form a complete basis for the Hilbert space $\mathscr{H}_e$, and thus, an arbitrary state of the full system can be expanded in a basis formed from products of the $|\psi_n;\mathbf{R}\rangle$ and basis states of the nuclear degrees of freedom

$$|\Psi\rangle = |\mathbf{R},\psi_n;\mathbf{R}\rangle = |\mathbf{R}\rangle\otimes|\psi_n;\mathbf{R}\rangle = \sum_n \int |\mathbf{R},\psi_n;\mathbf{R}\rangle\,\langle\mathbf{R},\psi_n;\mathbf{R}|\Psi\rangle\,d\mathbf{R}\,. \tag{2.6}$$

In the coordinate representation this can be written as

$$\begin{aligned}\Psi(\mathbf{R},\mathbf{r}) &= \sum_n \int d\mathbf{R}'\,\langle\mathbf{R},\mathbf{r}|\mathbf{R}',\psi_n;\mathbf{R}'\rangle\,\langle\mathbf{R}',\psi_n;\mathbf{R}'|\Psi\rangle \\ &= \sum_n \langle\mathbf{r}|\psi_n;\mathbf{R}\rangle\,\langle\mathbf{R},\psi_n;\mathbf{R}|\Psi\rangle \\ &= \sum_n \psi_n(\mathbf{r};\mathbf{R})\chi_n(\mathbf{R})\,.\end{aligned} \tag{2.7}$$

The expansion in equation 2.7 is known as the Born-Oppenheimer or Born-Huang *ansatz*. In practice, it is not possible to obtain all of the solutions $|\psi_n;\mathbf{R}\rangle$ of equation 2.5. However, formally these solutions exist and the expansion given above is exact. Inserting this into the TDSE from equation 2.1 and dropping the explicit parametric dependence of $|\psi_n;\mathbf{R}\rangle$ on $\mathbf{R}$ for simplicity, one obtains

$$i\frac{\mathrm{d}}{\mathrm{d}t}\sum_i |\psi_i\rangle\,|\chi_i\rangle = \sum_i \left(-\sum_a^{N_n}\frac{1}{2M_a}\nabla^2_{R_a} + \hat{H}_e\right)|\psi_i\rangle\,|\chi_i\rangle\ . \tag{2.8}$$

which, after multiplying from the left with $\langle \psi_j |$, becomes

$$\begin{aligned} i\frac{\mathrm{d}}{\mathrm{d}t}|\chi_j\rangle = &-\sum_a^{N_n}\frac{1}{2M_a}\nabla_{R_a}^2|\chi_j\rangle + E_j|\chi_j\rangle \\ &-\sum_a^{N_n}\frac{1}{2M_a}\sum_i(2\langle\psi_j|\nabla_{R_a}|\psi_i\rangle\nabla_{R_a} + \langle\psi_j|\nabla_{R_a}^2|\psi_i\rangle)|\chi_i\rangle\ , \end{aligned} \tag{2.9}$$

where the terms that couple the coefficients of the total wave function through the motion of the nuclei are known, respectively, as the derivative and scalar

$$\mathbf{F}_{ji} = \langle\psi_j|\nabla_R|\psi_i\rangle \tag{2.10}$$

$$G_{ji} = \langle\psi_j|\nabla_R^2|\psi_i\rangle \tag{2.11}$$

nonadiabatic couplings.

Equation 2.9 is still formally exact. An obvious approximation from here is to neglect all nonadiabatic couplings. This is called the Born-Oppenheimer approximation and results in uncoupled equations

$$i\frac{\mathrm{d}}{\mathrm{d}t}|\chi_j\rangle = -\sum_a^{N_n}\frac{1}{2M_a}\nabla_{R_a}^2|\chi_j\rangle + E_j|\chi_j\rangle\ . \tag{2.12}$$

The electronic states still depend on the positions of the nuclei, but are always relaxed with respect to the motion of the nuclei. This corresponds to the adiabatic theorem which states that a physical system remains in its instantaneous eigenstate if a perturbation is acting on it slowly enough and if its eigenvalue is well separated from the rest. For this reason the Born-Oppenheimer approximation is also sometimes called the adiabatic approximation (but this term is also used for the closely related Born–Huang approximation obtained when the diagonal nonadiabatic terms of equation 2.9 are retained). When a system is initially in a given electronic state $|\psi_n\rangle$, it will always remain in the given state and the motion of the nuclei is simply motion on a potential energy surface (PES) given by $E_n(\mathbf{R})$ generated by that electronic state. The concept of nuclei moving on a PES is central to most dynamical considerations in chemistry. The "slowly acting perturbation" requirement of the adiabatic theorem is covered by the initial reasoning for separating the total Hamiltonian based on fast electronic degrees of freedom and slow nuclear degrees of freedom (the perturbation). The second requirement of well separated eigenstates is usually true for the electronic ground state in molecules. How-

ever, for photochemical reactions, which involve excited electronic states, this requirement is almost never satisfied and a breakdown of the Born-Oppenheimer approximation occurs. In this case there is need to consider the nonadiabatic couplings which were neglected to arrive at the adiabatic single state view.

To better understand the reason why the Born-Oppenheimer approximation breaks down when electronic states are nearly degenerate, it is helpful to take a closer look at the nonadiabatic couplings. By applying the nuclear gradient to the electronic Schrödinger equation (2.5) and inserting the definition of the derivative couplings (2.10), one obtains an alternative expression for the couplings

$$\mathbf{F}_{ij} = \frac{\langle \psi_i | \nabla \hat{H}_e | \psi_j \rangle}{E_i - E_j} \,. \tag{2.13}$$

The denominator in equation 2.13 explicitly shows that the derivative couplings tend to become large when two states approach each other. Importantly, at points where two states are degenerate the couplings diverge and the adiabatic approximation breaks down completely.

On the other hand, equation 2.13 also confirms that couplings become smaller for states that are further apart. This allows us to make a new approximation, the group Born-Oppenheimer or group adiabatic approximation. Assuming a group of g electronic states are well separated energetically from all other states, the Born-Huang ansatz (2.7) can be truncated to include only a subset $\{g\}$ of electronic states

$$\Psi(\mathbf{R},\mathbf{r}) = \sum_{n\in\{g\}} \psi_n(\mathbf{r};\mathbf{R})\chi_n(\mathbf{R}) \,, \tag{2.14}$$

where the subset $\{g\}$ usually involves the lowest electronic states of the system in the relevant range of nuclear coordinates. This ansatz results in the same TDSE as the full Born-Huang ansatz, only with electronic states limited to those in $\{g\}$

$$\begin{aligned} i\frac{\mathrm{d}}{\mathrm{d}t}|\chi_j\rangle = &-\sum_a^{N_n} \frac{1}{2M_a} \nabla_{R_a}^2 |\chi_j\rangle + E_j |\chi_j\rangle \\ &-\sum_a^{N_n} \frac{1}{2M_a} \sum_{i\in\{g\}} \left(2\langle \psi_j | \nabla_{R_a} | \psi_i \rangle \nabla_{R_a} + \langle \psi_j | \nabla_{R_a}^2 | \psi_i \rangle\right) |\chi_i\rangle \,. \end{aligned} \tag{2.15}$$

Obviously, the Born-Oppenheimer approximation is a special case with only a single state in $\{g\}$. Since it is never possible to treat all electronic states, equation 2.15 is the starting

point for almost all dynamical calculations beyond the Born-Oppenheimer approximation in quantum chemistry. This approximation is often taken as a given, and in practice it is not even necessary to ensure that the selected electronic states are well separated from higher states. When dynamics are started with the initial population in some state $i < g$, the higher lying states often never acquire a significant population during a propagation. In this case, mixing (large coupling) between the highest states in $\{g\}$ and the states above them which are not included in the calculation has a negligible effect on the overall dynamics.

Methods that can in principle give numerically exact solution to equation 2.15 include grid based quantum dynamics (QD) approaches, the multiconfiguration time-dependent Hartree (MCTDH) method which introduces time-dependent basis functions[1,2] and trajectory based methods where the basis functions explicitly follow the nuclear wave functions. The latter include the variational multiconfiguration Gaussian (vMCG),[11] full multiple spawning (FMS)[90] and multiconfiguration Ehrenfest (MCE)[91,92] methods.

All of these methods are extremely computationally demanding, scaling (unless major approximations are introduced) exponentially with the number of dimensions of the system being studied. Since these methods are limited to systems composed of several atoms, there is a need for more approximate methods which can include some nonadiabatic effects beyond the Born-Oppenheimer approximation without the cost associated with a fully quantum treatment of the problem.

2.2 Mixed quantum-classical dynamics

2.2.1 Trajectory based methods

To arrive at a more tractable algorithm for considering nonadiabatic dynamics, we start with the basic idea of treating the nuclei as classical particles. This is a step further than the separation of the slow and fast degrees of freedom in the Born-Oppenheimer approximation, but the goal is to start with this larger approximation and later reintroduce the nonadiabatic effects into the simplified description of the system. Classical particles can be described by a point in nuclear phase space $(\mathbf{R}, \mathbf{v})$. However, for a description closer to the nuclear wave function $|\chi\rangle$ a statistical treatment is needed. For this reason, the basic classical construct to describe the nuclei is their phase space distribution $\rho(\mathbf{R}, \mathbf{v})$. The expectation value of any observable μ can

be obtained by integrating its value weighted by the probability distribution $\rho(\mathbf{R},\mathbf{v})$ over the full phase space

$$\langle\mu\rangle = \int\int \mu(\mathbf{R},\mathbf{v})\rho(\mathbf{R},\mathbf{v})d\mathbf{R}d\mathbf{v} = \int \mu(\mathbf{R})\rho(\mathbf{R})d\mathbf{R}\,, \tag{2.16}$$

where we assume for simplicity that μ, as will always be the case further in the text, is a function of the coordinates only and that the velocity can safely be integrated out of the equation giving $\rho(\mathbf{R})$ which describes the coordinate distribution.

To study dynamical processes, a time evolution $\rho(\mathbf{R},\mathbf{v};t)$ of the phase space distribution is needed. In purely classical dynamics, this evolution is given by the Liouville equation, but in practice the most common way for propagating this equation is to represent the distribution using a set of discreet points $\{(\mathbf{R_i},\mathbf{v_i})\}$. The full distribution can then be propagated through time by propagating the individual points using Newton's equation of motion

$$\mathbf{F}_i = m\frac{\mathrm{d}^2\mathbf{R}_i}{\mathrm{d}t^2}\,, \tag{2.17}$$

where the force in the Born-Oppenheimer picture is governed by the electronic Schrödinger equation as $\mathbf{F}_i = -\nabla E_0(\mathbf{R}_i)$ with $E_0(\mathbf{R}_i)$ being the PES of the ground state upon which the nuclei are propagated. This gives a set of trajectories $\{\mathbf{R}_i\}$ averaging over which at any moment t is equivalent to integrating over the phase space distribution so eq. 2.16 becomes

$$\langle\mu(t)\rangle = \sum_i^{N_p} w_i\mu(\mathbf{R}_i(t))\,, \tag{2.18}$$

where w_i is the weight of each trajectory and N_p the number of trajectories. Usually the initial points are sampled directly from $\rho(\mathbf{R},\mathbf{v};0)$ so each trajectory has the same weight, $\frac{1}{N_p}$.

In mixed quantum-classical dynamics the quantum and classical degrees of freedom interact so the description of the time evolution given above is not enough. Formally the evolution of the system in such a case is given by the mixed quantum-classical Liouville equation (QCLE) where the quantum particles are coupled to a bath of classical particles. This equation is also usually propagated using trajectories,[93] but the equations of motion are more complex and the trajectories are coupled. Numerically stable and efficient algorithms for solving these equations are still an open problem.[94–96]

On the other hand, the most studied and widely used mixed quantum-clasical dynamics method is the fewest-switches surface hopping (FSSH) algorithm developed by Tully.[3,4] In principle FSSH can be derived starting from the QCLE[13] or from (also formally exact) nonadiabatic Bohmian dynamics[97,98] and making specific approximations. However, FSSH was introduced as an *ad hoc* correction to include nonadiabatic effects into molecular dynamics (MD) simulations[99] and is usually still regarded as such. The approach is conceptually simple, intuitively captures the effects of coupling between electronic states and has given reasonable results for a wide range of problems in photochemistry.

2.2.2 Fewest-switches surface hopping

As stated above, in this section we will start from the simple fully classical trajectories given by eq. 2.17, and attempt to reintroduce as much of the nonadiabatic effects while keeping the method as simple and efficient as possible.

While the definition of the force acting on the nuclei $\mathbf{F}_i$ in the adiabatic approximation is clear since there is only one electronic state, it becomes less obvious when multiple electronic states $\psi_A(\mathbf{r};\mathbf{R})$ need to be taken into account. Going beyond the BO approximation, each of these states has an accompanying nuclear wave function $|\chi_A\rangle$ in the Born-Huang ansatz of equation 2.14. Above, we replaced a single nuclear wave function by a phase space distribution $\rho(\mathbf{R},\mathbf{v})$. If we hope to retain a classical view of the nuclei when multiple states are involved we can imagine that each state will need its own phase space distribution $\rho_A(\mathbf{R},\mathbf{v})$. Then the key question for developing trajectory based nonadiabatic dynamics methods is how to connect the propagation of trajectories representing $\{\rho_A(\mathbf{R},\mathbf{v})\}$ with the propagation and transfer of population between electronic states $\{|\psi_A(\mathbf{r};\mathbf{R}(t))\rangle\}$.

In the surface hopping view, each trajectory is assigned to a particular electronic state. Representing the flow of population between the different states at a given time is accomplished by changing this assignment (hops between surfaces). This means that each trajectory is described by the coordinates and current state as a function of time, $(\mathbf{R}_i(t),\lambda_i(t))$. The force acting on the nuclei is then defined simply by the gradient of the current electronic state.

The key question to answer is how to determine the population flow between the electronic states. In FSSH this is done by propagating an electronic wave function $|\Psi(\mathbf{r},t;\mathbf{R})\rangle$ in conjunc-

tion with each trajectory. The ansatz used is similar to the Born-Huang expansion

$$|\Psi(\mathbf{r},t;\mathbf{R})\rangle = \sum_A c_A(t)\,|\psi_A(\mathbf{r};\mathbf{R}(t))\rangle\,. \tag{2.19}$$

with the nuclear wave functions $|\chi_A\rangle$ replaced by time-dependent coefficients $c_A(t)$ which do not explicitly contain information on the motion of the nuclei. Inserting 2.19 into the TDSE and multiplying from the left with $|\psi_B\rangle$ we obtain an equation for the time evolution of the wave function coefficients

$$\frac{\mathrm{d}c_B}{\mathrm{d}t} = \sum_A c_A(-\langle\psi_B|\hat{H}_e|\psi_A\rangle + i\,\langle\psi_B|\frac{\mathrm{d}}{\mathrm{d}t}|\psi_A\rangle)\,. \tag{2.20}$$

In the basis of eigenstates of the electronic Hamiltonian equation 2.20 becomes

$$\frac{\mathrm{d}c_B}{\mathrm{d}t} = -c_B E_B + \sum_A c_A\,\langle\psi_B|\frac{\mathrm{d}}{\mathrm{d}t}|\psi_A\rangle\,, \tag{2.21}$$

where we see that the electronic states are coupled through the time-derivative couplings $\langle\psi_B|\frac{\mathrm{d}}{\mathrm{d}t}|\psi_A\rangle$ (TDC). These couplings are related to the derivative nonadiabatic couplings 2.10 and the nuclear velocity using the chain rule:

$$\sigma_{BA} \equiv \langle\psi_B|\frac{\mathrm{d}}{\mathrm{d}t}|\psi_A\rangle = \langle\psi_B|\nabla_R|\psi_A\rangle\,\frac{\mathrm{d}\mathbf{R}}{\mathrm{d}t}\,. \tag{2.22}$$

Changes in the coefficients $\{c_A\}$ along a trajectory are used to determine when a hop (change of state) should occur. At each step during the trajectory propagation, the hopping probability from the current state B to any other state A is equal to the change in the population of state B ($c_B^*(t)c_B(t)$) due to flux to A during the time step Δt. Using equation 2.21, we can derive that this probability should be

$$P_{B\to A} = \max(0, -\frac{2\Delta t}{c_B^* c_B}\,\mathrm{Re}(\sigma_{BA}c_A c_B^*))\,, \tag{2.23}$$

where max() is used because the population can flow from A to the current state B, in which case the flux is negative. However, a negative flux just means that hops would instead occur in trajectories where the current state is B. Hops are then decided by sampling a random number

$r \in [0,1]$ and evaluating

$$\sum_{C}^{A-1} P_{B\to k} \leq r \leq \sum_{C}^{A} P_{B\to k} \tag{2.24}$$

for each state A. If the condition is true and if energy conservation can be ensured after a hop, the propagation is continued on that state. For now we assumed that the time step Δt is infinitesimal and we will deal with the implications of discretized time steps in section 2.2.3.

This type of sudden switch from one PES to another is not physical. However, when attempting to describe quantum effects using classical trajectories, it is in general more meaningful to consider the full swarm of trajectories instead of individual trajectories. When repeating the above procedure for many trajectories the goal is to ensure that the proportion of trajectories in a given state will be equal to the average population for each state A,

$$\left\langle |c_A(t)|^2 \right\rangle = \sum_i \delta_{\lambda_i(t)A} w_i \tag{2.25}$$

The reason for using the population flux instead of the populations themselves to determine hopping probabilities is that this minimizes the number of hops needed to satisfy the requirement given by equation 2.25 at any moment (this is why the method is called "fewest-switches"). The importance of this point is easily illustrated by imagining an algorithm with population based hopping probabilities going through a region of heavy mixing. Equation 2.25 would still be satisfied, but at each step the gradient followed by many of the trajectories would change. This would result in the motion of the individual trajectories, but also of the whole ensemble, being similar to motion on an averaged PES instead of two distinct types of motion on one state or the other.

The general procedure described above is the basis of FSSH. Once the initial conditions are defined, the only values needed to propagate a trajectory are the energies of the electronic states, the gradient of the current electronic states and the couplings between all electronic states. Computing energies and gradients is the most standard task of electronic structure calculations, and analytical nonadiabatic coupling vectors (related to TDCs through equation 2.22) have been implemented for various electronic structure methods.

However, for a robust algorithm we also need to consider many additional details. One active field of research is how to generate the initial conditions for trajectories[31,100–105] and how to obtain observables based on an ensemble of nuclear trajectories[31–33,106] which will be covered

in section 2.2.5. Another active field deals with evaluating the nonadiabatic couplings[4,15,16] and integrating the TDSE while taking into account possible sharp peaks in the couplings[19–21,107] which will be covered in section 2.2.3. Two other important question which we will only briefly touch on here are how exactly to enforce energy conservation after hops[108–111] and how to correct for the lack of decoherence in SH.[13,112–115]

Energy conservation: For a hop to occur it is not necessary for the two states to be degenerate and it is possible (if unlikely) for the energy separation to be substantial. This creates an energy conservation problem during FSSH trajectories which is solved by scaling the velocity whenever a hop occurs. This should be done along the direction of the nonadiabatic coupling vector $\mathbf{F}_{ij}$,[108,109] but is also sometimes (when coupling vectors are not available) done along the direction of the momentum or uniformly along all directions. If energy conservation is not possible due to lack of kinetic energy along the chosen direction, the hop is rejected. This situation is referred to as a frustrated or forbidden hop and is handled either by reversing the velocity along the chosen direction (as in elastic scattering) or by ignoring the hop continuing the trajectory on the initial state with no changes. The former approach is considered more justified and gives better results in most (but not all[116]) cases.[110,111]

Correctly achieving energy conservation during dynamics is important from a theoretical perspective, but usually less so from a practical perspective. For real systems with many degrees of freedom these corrections are usually minor since hops mostly occur when the gap between two states is much smaller than the total kinetic energy of the system. However, in some cases the choice regarding the treatment of momentum rescaling and frustrated hops can have a significant impact on the overall dynamics even for molecular systems.[117]

Decoherence: Overcoherence in the standard FSSH algorithm is an additional problem that needs to be resolved. The coupling at a given moment in each trajectory depends only on the current nuclear position which is determined solely by motion on the current surface. For a given trajectory, amplitudes on all states other than the current one represent "virtual" trajectories moving alongside the "real" trajectory.[112] In effect, this means that after bifurcating at a crossing, two parts of a nuclear wave packet on different surfaces do not move apart. Because of this, they also do not experience the same level of decoherence as they would in global methods.

An ideal solution to this problem is not possible inside the independent trajectory approximation. However, a number of approximate recipes have been given based on damping the amplitudes of all states other than the currently populated one. Of these we will only mention commonly used procedure suggested by Granucci and Persico[114] based on the work by Zhu and coworkers[113,118] referred to as energy based decoherence or as non-linear decay of mixing. Here the populations of all states other than the currently populated one are reduced after each time step based on their energy difference with the the current state:

$$\begin{aligned} C'_A &= C_A e^{-\frac{\Delta t}{2\tau_{AB}}}, \quad \forall A \neq B \\ C'_B &= C_B \left(\frac{1 - \sum_{A \neq B} C'_A}{|C_B|^2} \right)^{1/2} \\ \tau_{AB} &= \frac{\hbar}{|E_A - E_B|} \left(1 + \frac{\alpha}{E_{\text{kin}}} \right), \end{aligned} \tag{2.26}$$

where E_{kin} is the nuclear kinetic energy and α an empirical parameter determining the strength of the decoherence correction. Based on some initial numerical tests, small changes in the empirical parameter were shown to have little impact on the overall dynamics.[113] A default value of 0.1 Hartree was suggested[114] and was used in almost all subsequent calculations.

Like the question of energy conservation, decoherence is more often studied from the theoretical side than for the practical effect it has on simulations of actual molecular systems. Mostly it is considered satisfactory to have at least some form of correction, which is why the simple algorithm presented above is the most common choice. On the other hand, if a system is expected to pass through the same interaction region many times during a photochemical reaction, a correct treatment of decoherence is critical.[13] This is an active area of research which is beyond the scope of this overview. For a more detailed theoretical consideration of this problem, the reader is referred to the work of Subotnik and coworkers[13] and the references within.

2.2.3 Couplings and wave function overlaps

The key quantity that differentiates FSSH molecular dynamics and most other nonadiabatic dynamics algorithms from ground state Born-Oppenheimer molecular dynamics is the coupling between electronic states. As stated by equation 2.22 time-derivative couplings σ_{BA} can be easily obtained using the nuclear velocity vector and nonadiabatic coupling vectors

$\langle \psi_j | \nabla_R | \psi_i \rangle$. Analytical algorithms for obtaining the latter have been developed for numerous electronic structure methods. This includes couplings for TDDFT[119] with linear response TDDFT[41,56,120,121] and within the Tamm-Dancoff approximation,[42,122] equation-of-motion CC methods,[40,123] MCSCF,[124] state averaged CASSCF,[125] MRCI,[126,127] and CASPT2[128] which covers many of the most commonly used electronic structure methods for excited electronic states.

While this is in theory all that is needed for FSSH calculations, there is one physical and two practical problems with using TDCs obtained in this way. The physical problem is that TDCs can change quickly and are often not well represented by their values at the start and end of a nuclear time step (we will get back to this problem shortly). The practical problems are that 1) analytical NACV calculations add a significant extra step to the electronic structure calculations which are already the bottleneck of FSSH calculations and 2) implementing this type of calculation is not trivial and is usually done only for methods that are well established which prevents the use of cutting edge electronic structure methods for FSSH.

These practical problems have been noticed very early in the development of FSSH and already in 1994 Hammes-Schiffer and Tully[4] proposed calculating the TDCs by numerical differentiation of the electronic wave functions along the trajectory using finite differences

$$\sigma_{AB}(t+\frac{\Delta t}{2}) \approx \frac{1}{2\Delta t}(\langle \psi_A(t) | \psi_B(t+\Delta t) \rangle - \langle \psi_B(t+\Delta t) | \psi_A(t) \rangle) . \tag{2.27}$$

Calculation of couplings based on this equation is usually called the HST method. The accuracy of this method for calculating couplings was confirmed in several studies.[15,38]

Using this equation, it is possible to obtain the TDCs for any wave function for which we can compute overlap matrix elements of the form $S_{AB} = \langle \psi_A | \psi'_B \rangle$. Here we use the prime to signify that the two wave functions correspond to different time steps. It is important to stress that this does not just mean that the two states are not orthogonal (due to being solutions of the electronic Schrödinger equation at different geometries), but also usually means that they are built using different basis sets since atomic basis functions explicitly depend on the nuclear coordinates. This complicates the calculation since it requires taking into account the Slater determinant (SD) form of the wave functions.

In general we can consider the problem of calculating the matrix for two sets of states $\{|\psi_A\rangle\}$

and $\left\{\left|\psi'_B\right\rangle\right\}$

$$|\psi_A\rangle = \sum_a^{n_{det}} d_a^A \, |\Phi_a\rangle \tag{2.28}$$

and

$$\left|\psi'_B\right\rangle = \sum_b^{n'_{det}} d_b'^B \left|\Phi'_b\right\rangle \tag{2.29}$$

where a and b are collective indices running over all combinations of orbitals from which SDs $|\Phi_a\rangle = \left|\phi_{a(1)\alpha}\cdots\phi_{a(n_\alpha)\alpha}\phi_{a(n_\alpha+1)\beta}\cdots\phi_{a(n_\alpha+n_\beta)\beta}\right|$ are constructed. We use $\{\phi_{i\alpha}\}$ and $\left\{\phi_{i\beta}\right\}$ to denote the set of all α and β spin orbitals, respectively. Inserting into the expression for overlap matrix elements, we obtain

$$S_{AB} = \sum_a^{n_{det}} \sum_b^{n'_{det}} d_a^A d_b'^B \left\langle\Phi_a\middle|\Phi'_b\right\rangle . \tag{2.30}$$

Elements $\left\langle\Phi_a\middle|\Phi'_b\right\rangle$ can be calculated as the determinant of the overlaps of individual orbitals from which the SDs are built.[16,129] The orbital overlap matrix is block diagonal so the final expression can be divided into separate spin α and β blocks. The overlap between the two SDs can then be written as

$$\begin{aligned}\left\langle\Phi_a\middle|\Phi'_b\right\rangle = &\begin{vmatrix} \left\langle\phi_{a(1)\alpha}\middle|\phi'_{b(1)\alpha}\right\rangle & \cdots & \left\langle\phi_{a(1)\alpha}\middle|\phi'_{b(n_\alpha)\alpha}\right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle\phi_{a(n_\alpha)\alpha}\middle|\phi'_{b(1)\alpha}\right\rangle & \cdots & \left\langle\phi_{a(n_\alpha)\alpha}\middle|\phi'_{b(n_\alpha)\alpha}\right\rangle \end{vmatrix} \times \\ &\begin{vmatrix} \left\langle\phi_{a(n_\alpha+1)\beta}\middle|\phi'_{b(n_\alpha+1)\beta}\right\rangle & \cdots & \left\langle\phi_{a(n_\alpha+1)\beta}\middle|\phi'_{b(n_\alpha+n_\beta)\beta}\right\rangle \\ \vdots & \ddots & \vdots \\ \left\langle\phi_{a(n_\alpha+n_\beta)\beta}\middle|\phi'_{b(n_\alpha+1)\beta}\right\rangle & \cdots & \left\langle\phi_{a(n_\alpha+n_\beta)\beta}\middle|\phi'_{b(n_\alpha+n_\beta)\beta}\right\rangle \end{vmatrix} \end{aligned} \tag{2.31}$$

$$= \left\langle\Phi_a\middle|\Phi'_b\right\rangle_\alpha \left\langle\Phi_a\middle|\Phi'_b\right\rangle_\beta , \tag{2.32}$$

where $\left\langle\Phi_a\middle|\Phi'_b\right\rangle_\sigma$ was defined as the determinant of the spin σ block of the overlap.

The calculation of a single determinant is usually performed using LU factorization and scales as $O(n^3)$, where n is the size of the matrix. This means that equation 2.30 scales as $O(n_{det}n'_{det}n^3)$. This scaling is always worse than the underlying electronic structure calculation used to compute the wave function meaning that this type of calculation can quickly become the bottleneck of FSSH calculations. However, due to a much smaller prefactor, even without any optimizations or approximations the evaluation of eq. 2.30 is significantly faster than electronic

structure calculations while the number of determinants is sufficiently low.

In order to evaluate the overlap matrix for larger systems, active spaces or basis sets, more efficient algorithms are needed. For this, one should also take into account the form of the wave function in the electronic structure method. For single reference methods, expansions 2.28 and 2.29 will contain many determinants which differ in only a single or a few orbitals. For multireference methods based on the complete active space expansion, the wave functions will contain many different determinants, but determinants with a given arrangement of orbitals of a given spin will repeat in determinants with many different arrangements of orbitals of the opposite spin. These properties were used by Plasser *et al.*[16] to develop a fast general algorithm for such calculations. Their algorithm is based on sorting and reusing the spin blocks appearing many times in CASSCF wave function expansions and on expanding the determinants in terms of minors and reusing these minors to compute multiple determinants when enough of them differ by only a single orbital.

This approach significantly increases the efficiency of overlap calculations, but it only pushes the point at which these calculations become the bottleneck to medium sized systems or active spaces. An alternative approach is to introduce approximations to the overlap calculation. Pittner *et al.*[15] introduced two thresholds to their implementation of eq. 2.30, a mutual excitation threshold t^{me} (determinants with a significantly different number of excitations from the reference orbital are assumed to have negligible overlap) a coefficient product threshold t^{cp} (when the product of the CI coefficients of two determinants is negligible in all states the corresponding matrix element is not calculated) so the overlap between two SDs a and b are calculated only when conditions

$$t_{ab}^{\mathrm{me}} \geq \left| \sum_i \delta_{a(i)i} - \sum_i \delta_{b(i)i} \right| \tag{2.33}$$

$$t_{ab}^{\mathrm{cp}} \geq \sum_A \sum_B \left| d_a^A d_b'^B \right| \tag{2.34}$$

are met. Another approximation scheme was offered by Plasser *et al.*,[16] here the approximation is made at the wave function level introducing a threshold for the norm of the wave function t^{norm}, the CI expansions 2.28 and 2.29 are sorted and truncated to keep the smallest number of

determinants satisfying the condition

$$t_A^{\text{norm}} \leq \sum_a \left| d_a^A \right|^2 \tag{2.35}$$

to form truncated wave functions $|\tilde{\psi}_A\rangle$ which are then used in the overlap calculation. If we assume that these approximations have negligible effect on the angle between the wave functions, errors due to the reduction of their norms can be canceled out simply by re-normalizing (or orthogonalizing) the final overlap matrix. Another method for speeding up these calculations along with a more detailed analysis of the effects of approximating the wave functions will be given in 4.2.

A completely different approach for obtaining TDCs was developed by Ryabinkin and coworkers[17] based on orbital derivatives. In this approach TDCs are also evaluated by numerical differentiation, but the SDs are expanded before differentiation resulting in an expression where the finite difference scheme can be applied to the orbitals and coefficients themselves instead of the determinants. This approach is always significantly faster, scaling with the third power of the size of the system for a CIS type wave function. However, this approach does not compute the overlap matrix between wave functions at the two time steps which is itself a key quantity for FSSH calculations. This is especially true for the problem of trivial crossings which will be described in detail in the following section.

Trivial crossings: In classical molecular dynamics, motion is governed by a single potential energy surface which is a smooth functions of nuclear coordinates and the equation of motion 2.17 can be integrated using standard methods such as the velocity Verlet algorithm.[130] A time step on the order of 1 fs is usually small enough to adequately cover the high frequency motions in molecular systems and ensure a stable propagation.

The situation in mixed quantum-classical dynamics simulations is significantly more complicated because, in the vicinity of conical intersections, nonadiabatic couplings are highly peaked functions. There is no limit to how narrow or high these peaks can be in the case of weakly avoided crossings. In the case of unavoided (trivial) crossings, couplings are delta functions. In practice, this means that couplings can be nearly zero at the start and end of a time-step, but increase to very high values during the time step. This is illustrated in Fig 2.1

where we see two intersections of the same diabatic states with different coupling strengths. In the first case, the two states mix over four nuclear time steps and the nonadiabatic coupling is reasonably described by the values calculated at discreet time steps. On the other hand, in Fig. 2.1a, the two states are very weakly coupled which leads to a highly peaked nonadiabatic coupling which occurs in between two time steps. In this case, a nonadiabatic transition should almost definitely occur, but likely won't if only the couplings at the discreet time steps are used to calculate the probability.

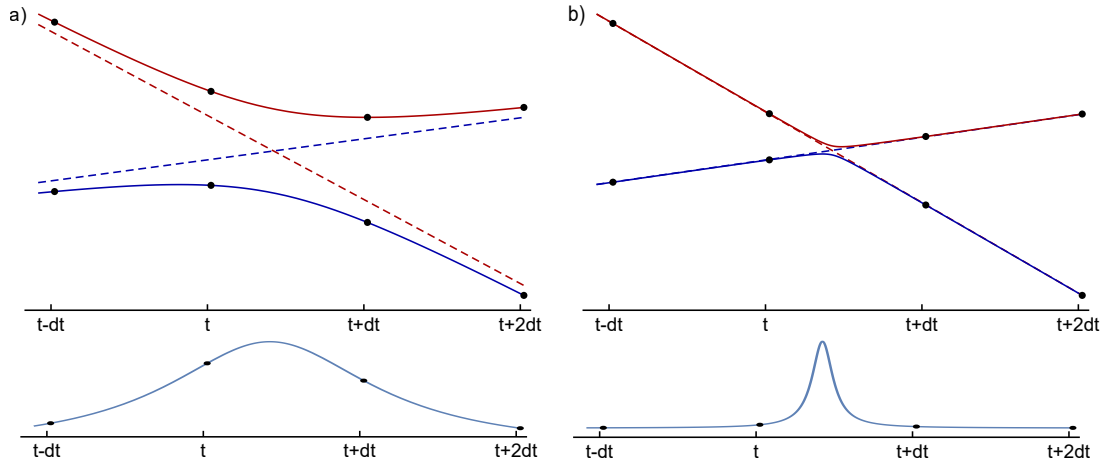


Figure 2.1: Adiabatic (full lines) and diabatic (dashed lines) states and nonadiabatic couplings (blue lines, bottom panels) for the case of strong (a) or weak (b) coupling between diabatic states. Values calculated at discreet nuclear time steps are marked by black dots. In the case of weak coupling the peak in the nonadiabatic coupling is too narrow to be captured by discreet time steps.

The problem of trivial crossings is perhaps most easily seen for systems with a large physical separation between excited states. When simulating a system composed of two subsystems separated by hundreds of ångström analytic TDCs between two states on different subsystems are likely to always be almost exactly zero, but if two states cross during a time step it is obvious that the same diabatic state should be followed.[19,107] The job of any practical FSSH algorithm is then to be able to recognize and handle this type of situation, but also any borderline situation. This also needs to be done without resorting to arbitrarily small nuclear time steps, ideally without any additional electronic structure evaluations.

Since they can vary massively during a time step, direct interpolation of TDCs is not an option. On the other hand, information on any sudden peaks in the coupling during the time step will be preserved in the wave functions at the end of a time step. The $\sigma_{AB}(t + \frac{\Delta t}{2})$ on

the left-hand side of equation 2.27 is actually a measure of the change in the wave functions during the time step, an approximation of the *average* σ_{AB} during the full time step. Due to this property, TDCs obtained from wave function overlaps are significantly more numerically stable than analytical calculations.

However, equation 2.27 by itself is not enough for a correct evaluation of hopping probabilities, especially in the case of trivial or nearly trivial crossings. The issue that needs to be addressed is what assumptions to make about the changes of the electronic states during the time step while integrating the TDSE. This can be split into two connected but partially distinct problems; 1) how to interpolate the wave functions or couplings during the step while integrating the TDSE and 2) how to match the phases of the wave functions at the beginning and end of the time step.

Interpolation and integration: Approaches to integrating the TDSE given in equation 2.20 can in general be separated into two caregories, adiabatic and diabatic. As mentioned above, if the equation is integrated in its original adiabatic form 2.20 care needs to be taken to correctly interpolate the possibly highly peaked couplings. On the other hand, the equation can be transformed into a basis of (quasi)diabatic states between which the couplings are (almost) zero. In this case the main problem is finding an adequate diabatic basis.

In the adiabatic representation, the averaged couplings evaluated at $t + \frac{\Delta t}{2}$ using eq. 2.27 are a better approximation than those evaluated analytically at t and $t + \Delta t$ since they are actually a measure of the change in the wave functions during the time step. However, this procedure still causes large numerical errors such as probabilities larger than unity. These problems can be alleviated by evaluating the hopping probability at substeps during the integration of the TDSE instead of once for each nuclear time step:

$$P_{B\to A} = \max(0, -\frac{2\Delta\tau}{c_B(\tau)^* c_B(\tau)} \mathrm{Re}(\sigma_{BA}(\tau) c_A(\tau) c_B(\tau)^*)), \tag{2.36}$$

where τ is the time during the current integration substep and $\Delta\tau$ is the substep duration. This probability can then be integrated during the full nuclear time step to determine whether a hop should occur at the end of the step[56] or hops can be allowed during each substep.[131] In either case, $\sigma_{BA}(\tau)$ needs to be evaluated in some way. At first this was done by linear interpolation/extrapolation of the TDCs evaluated at $t + \frac{\Delta t}{2}$ using eq. 2.27.[56,131]

However, this type of interpolation of the individual elements of the TDC matrix does not correspond to physically meaningful changes of the electronic wave functions along the substeps. Meek and Levine[20] suggested a method called norm-preserving interpolation (NPI) deriving $\sigma_{BA}(\tau)$ by assuming that the adiabatic wave functions during the time step change via a unitary transformation

$$|\psi_A(\tau)\rangle = \mathbf{U}(\tau)\,|\psi_A(t)\rangle \tag{2.37}$$

where $\mathbf{U}(\tau)$ is a unitary matrix corresponding to pairwise rotations connecting the initial wave functions $\{|\psi_A(t)\rangle\}$ with the final wave functions $\{|\psi_A(t+\Delta t)\rangle\}$. The TDCs can then be calculated from $\{|\psi_A\tau\rangle\}$ and even analytically integrated over the time step

$$\sigma_{BA}(t+\frac{\Delta t}{2}) = \frac{1}{\Delta t}\int_t^{t+\Delta t} \langle\psi_B(t)|\mathbf{U}^\dagger(\tau)\frac{\partial}{\partial\tau}\mathbf{U}(\tau)|\psi_A(t)\rangle\, d\tau \tag{2.38}$$

to give a better approximation of the average TDC than the finite difference approach.

The key advantage of this procedure is that it maintains orthonormality between the states at all times. This approach was improved by Jain *et al.*[24] when they derived a general expression for σ which does not rely on pairwise rotations. Starting from the time derivative of the overlap matrix elements

$$\frac{\partial S_{BA}}{\partial\tau} = \langle\psi_B(t)|\frac{\partial}{\partial\tau}|\psi_B(t+\tau)\rangle \tag{2.39}$$

$$= \sum_M \langle\psi_B(t)|\psi_M(t+\tau)\rangle \left\langle \psi_M(t+\tau)\middle|\frac{\partial\psi_B(t+\tau)}{\partial\tau}\right\rangle \tag{2.40}$$

$$= \sum_M S_{BM}\sigma_{MA} \tag{2.41}$$

leads to the differential equation in matrix notation

$$\frac{\partial\mathbf{S}}{\partial\tau} = \mathbf{S}\sigma \tag{2.42}$$

Which can be used to evaluate σ. Averaged over the nuclear time step this gives

$$\sigma(t+\frac{\Delta t}{2}) \approx \frac{1}{\Delta t}\log\mathbf{S}(t,t+\Delta t)\,, \tag{2.43}$$

where $\mathbf{S}(t,t+\Delta t)$ is the overlap matrix between adiabatic wave functions at two nuclear time

steps. In order to evaluate the equation above, **S** needs to be a unitary matrix corresponding to a proper rotation. This can be achieved by orthogonalizing the matrix and correctly assigning the phases of the wave functions (see below).

If the dynamics are based on propagating the TDSE in the diabatic basis, the most commonly used approach is the local diabatization (LD) formalism introduced by Granucci and coworkers.[18] In this approach, instead of constructing global diabatic surfaces, the diabatization is performed on-the-fly along the nuclear trajectory so that couplings are negligible only along the path. The adiabatic to diabatic transformation matrix is defined simply as the orthogonalized overlap matrix $\mathbf{T} = \mathbf{S}[\mathbf{S}^T\mathbf{S}]^{-1/2}$. This matrix is used to transform the Hamiltonian into the diabatic basis

$$\mathbf{H}(t+\Delta t) = \mathbf{T}(t+\Delta t)\mathbf{E}(t+\Delta t)\mathbf{T}^T(t+\Delta t)\,, \tag{2.44}$$

where **E** is the Hamiltonian in the adiabatic representation (containing the adiabatic energies on the diagonal). Since the diabatic states are constructed to change as little as possible along the time step, this matrix can be safely interpolated by linear interpolation

$$\mathbf{H}(\tau) \approx \mathbf{E}(t) + [\mathbf{H}(t+\Delta t) - \mathbf{E}(t)]\frac{\tau}{\Delta t}\,. \tag{2.45}$$

Within this approximation, the TDSE is easily integrated analytically and the hopping probabilities are defined using the changes of adiabatic populations along the time step.

Since large errors are expected only in case of trivial crossings, a third option is to recognize this situation and treat it as a special case. This is done by Wang and coworkers[21,22] by detecting crossings based on the overlap matrix and using a different formula for calculating the probability of hopping to the state with the maximum overlap with the current state. Fernandez-Alberti and coworkers[19,107] assigned the states by solving the assignment problem for the overlap matrix. If the current state is assigned to a different state of the next step and if the corresponding overlap matrix element is greater than some threshold, the crossing is treated as a trivial crossing and the hop is made without evaluating the probability.

Phase matching: Electronic wave functions are eigenvectors of the electronic Schrödinger equation. This means that multiplying them by -1 (for real wave functions) results in a wave function that satisfies the same equation. Normally a change of sign makes no difference since

all relevant expectation values are insensitive to this phase factor. However, when integrating the TDSE during FSSH simulations, we assume that the wave functions change smoothly during the time step and if the phase factor is different between the two time steps this assumption cannot hold.

Tracking and correcting these phases is critical for the correct evaluation of hopping probabilities at trivial crossings when propagating in the adiabatic basis. This can be illustrated for the simple toy model of a trivial crossing between two electronic states. Four possible overlap matrices can be obtained in such a case:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2.46}$$

depending on the (basically random) phase factor of each wave function returned by the electronic structure calculation. While each of these matrices is unitary, the first two are not rotation matrices (their determinants are -1 instead of 1). Attempting to use these matrices with the HST or NPI methods would result in all TDCs being exactly zero and the logarithm of the matrix required by the method of Jain *et al.*[24] would be a complex number matrix.

In practice elements of overlap matrices are basically never exactly zero so the simplest way to ensure that the overlap matrix corresponds to a rotation is to multiply the columns of the matrix so that all diagonal elements are positive.[15] However, simply ensuring that the matrix is a rotation matrix is not enough to ensure that the path followed during interpolation will be the best path, especially where there are multiple state crossings within the time step. In such cases the choice of phase factors will determine which states will mix during the interpolation.

A sensible solution is usually obtained by assigning the states at the beginning of the time step to those at the end of the step. This is most simply done by placing the elements with the maximum value on the diagonal before correcting the phases.[132] However, in some cases, this method still leaves some ambiguity regarding the relative phases of the states. Zhou *et al.*[43] attempt to solve this ambiguity by suggesting that the optimal selection of phase factors is the one for which the path covered by the wave functions during the rotation is minimal. This is obtained by minimizing $\mathrm{Tr}\left(|\log \mathbf{U}|^2\right)$ for which they give an approximate numerical algorithm.

2.2.4 Landau-Zener surface hopping

Trajectory surface hopping methods differ from standard adiabatic molecular dynamics by answering one additional question: what is the probability of transitions between different electronic states? It is therefore instructive to look at a model system where this question can be answered analytically, the Landau-Zener model.[133–135] The solution of this apparently simple model can give useful insight into the parameters which have an effect on the transition probability. The equation can also be used as the basis of an approximate algorithm for much more complex situations. While this algorithm is not expected to perform ideally in all cases, it has the advantage of bypassing the most difficult part of the FSSH algorithm - the calculation of nonadiabatic couplings.

The assumptions underpinning the Landau-Zener model are: (1) only two-states are interacting in the relevant region of the nuclear coordinate space, (2) the two states are interacting in a limited region, (3) the system can be described by diabatic states which coincide with the adiabatic states far from the interaction region, (4) all elements of the diabatic Hamiltonian have at most a linear dependence on the nuclear coordinates and (5) the velocity along the trajectory is constant. In general, we can write the Hamiltonian in the diabatic or adiabatic basis as

$$\hat{H}_d = \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{12}(z) & H_{22}(z) \end{bmatrix}, \quad \hat{H}_a = \begin{bmatrix} V_+(z) & 0 \\ 0 & V_-(z) \end{bmatrix}, \tag{2.47}$$

where, $z(t)$ is a general nuclear coordinate passing through an interaction region where the diabatic surfaces cross at $z(t_c) = z_c$. We will denote the diabatic states $\{|\phi_1\rangle, |\phi_2\rangle\}$ and the adiabatic states $\{|\phi_-\rangle, |\phi_+\rangle\}$. The second assumption above is ensured by the condition

$$\lim_{|z|\to\infty} \frac{H_{12}(z)}{|H_{11}(z) - H_{22}(z)|} = 0, \tag{2.48}$$

which in combination with assumption (4) means that the off-diagonal elements cannot depend on z. This gives the form for the Hamiltonian of the Landau-Zener model as

$$\hat{H}_d = \begin{bmatrix} h_{11} + s_1 z & h_{12} \\ h_{12} & h_{22} + s_2 z \end{bmatrix}, \tag{2.49}$$

which can further be simplified without loss of generality by choosing the coordinates so that $t_c = 0$ and $z_c = 0$

$$\hat{H}_d = \begin{bmatrix} s_1 z & h_{12} \\ h_{12} & s_2 z \end{bmatrix}. \tag{2.50}$$

For this model, the adiabatic energies (eigenvalues of $\hat{H}_d$) are

$$V_{\pm}(z) = \frac{1}{2}(s_1 z + s_2 z \pm \sqrt{4h_{12}^2 + (s_1 - s_2)^2 z^2}) \tag{2.51}$$

and the gap between the states is

$$g(z) = |V_+(z) - V_-(z)| = \sqrt{4h_{12}^2 + (s_1 - s_2)^2 z^2}, \tag{2.52}$$

which is equal to $2|h_{12}|$ at the gap minimum. So while the gap depends on the difference between the slopes of the two states $\Delta s \equiv s_1 - s_2$, how close they can approach depends only on the coupling between the states. For the Landau-Zener formula (see below) to be exact, the slopes of the two states need to have the same sign. The model is illustrated in Fig. 2.2. The diabatic surfaces are straight lines intersecting at t_c while the adiabatic surfaces exhibit and avoided crossing in the region around the diabatic crossing point. Because of this, the lower surface V_- corresponds to diabatic state H_{11} in the limit as z goes to $-\infty$ and to H_{22} as z goes to $+\infty$.

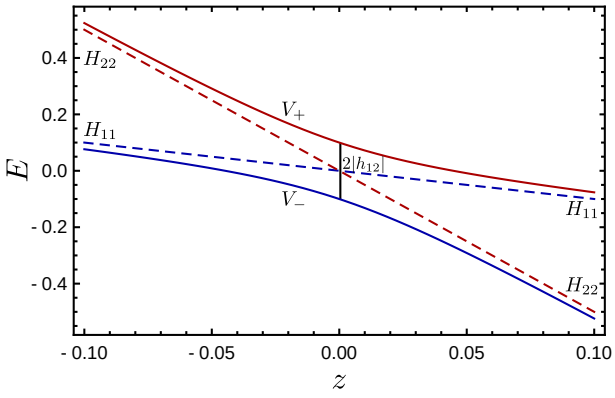


Figure 2.2: Adiabatic (full lines) and diabatic (dashed lines) potentials in the Landau-Zener model. The diabatic states cross at $z_c = 0$ while the adiabatic states exhibit an avoided crossing.

The electronic wave function of the system can be written in terms of adiabatic states as $|\Psi(z,t)\rangle = c_-(t)\,|\phi_-(z)\rangle + c_+(t)\,|\phi_+(z)\rangle$. Assuming the full population is initially in one of the

states (i.e. $|\Psi(z,t_0)\rangle = |\phi_+\rangle$), the probability that a non-adiabatic transition occurred up to some point t is simply the population of the other state $|c_+(t)|^2$. Alternately, we can write the wave function in terms of diabatic states $|\Psi(z,t)\rangle = c_1(t)\,|\phi_1(z)\rangle + c_2(t)\,|\phi_2(z)\rangle$. Now we can insert the diabatic Hamiltonian into the time-dependent Schrödinger equation to obtain a system of equations

$$\begin{aligned} i\hbar\frac{\mathrm{d}c_1}{\mathrm{d}t} &= h_{12}c_2\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t}\Delta s z(t)dt\right) = h_{12}c_2\exp\left(-\frac{iv}{2\hbar}\Delta s(t^2-t_0^2)\right) \\ i\hbar\frac{\mathrm{d}c_2}{\mathrm{d}t} &= h_{12}c_1\exp\left(-\frac{i}{\hbar}\int_{t_0}^{t}\Delta s z(t)dt\right) = h_{12}c_1\exp\left(-\frac{iv}{2\hbar}\Delta s(t^2-t_0^2)\right) \end{aligned}, \tag{2.53}$$

where we have used the assumption of constant velocity to expand $z(t)$ as vt and solve the integrals in the exponent. In the limit of $t_0 \ll t_c \ll t$ this system of equations can be solved analytically. Since the adiabatic and diabatic states can be equated in the same limit so $c_+(-\infty) = c_2(-\infty)$ and $c_-(\infty) = c_2(\infty)$, this gives the Landau-Zener formula for the nonadiabatic transition probability

$$P = \exp\left(-\frac{2\pi h_{12}^2}{\hbar v\Delta s}\right). \tag{2.54}$$

Equivalently, the equation above can be written in terms of the adiabatic energies[48,49] as

$$P = \exp\left(-\frac{\pi}{2\hbar}\sqrt{\frac{g^3(t_c)}{\ddot{g}(t_c)}}\right), \tag{2.55}$$

where $\ddot{g}(t_c)$ is the second-order time derivative of the energy gap. These equations show that the transition probability increases exponentially when the minimum of the gap between two states decreases, when the difference between the slopes of the states increases, or when the velocity with which the crossing region is approached increases.

Now that we have detailed the approximations of the LZ model we are ready to consider how and when the equations derived for this simple model system are applicable to more realistic systems. Historically, the diabatic LZ formula was used to study atomic or molecular collisions where the interaction is dominated by a single coordinate (distance between the subsystems) and is large only in a small region when the two subsystems are near, simplifying the construction of a diabatic basis. However, classical trajectories are one dimensional objects in phase space regardless of the number of dimensions of the system so the trajectory can always be written in

some general form $z(t)$. Construction of a diabatic basis is also not needed if diabatization can be performed locally along the trajectory in a way similar to the locally diabatic formalism of FSSH[18] or if the recently derived adiabatic version of the LZ formula[48] is used.

In terms of studying photochemical reactions, nonadiabatic interactions between electronic states in molecules are known to be sharply peaked satisfying the requirement for a localized interaction region. On the other hand, many photochemical reactions of interest involve more than two states. This is a problem for the LZ model unless each interaction region involves only two states and is well separated from other states. The topography of the potential energy surfaces is also surely more complicated than the linear dependence in the LZ model, but a linear model might still be adequate in the narrow intersection region.

Interest has recently grown in using the Landau-Zener formula for a surface hopping algorithm. This was done first using by constructing a locally diabatic basis and applying the diabatic LZ formula[136] and a few years later the adiabatic version was derived[48,49] and used to study low dimensional models of molecular systems.[50,137] Algorithms were also developed (also with constructing a locally diabatic basis) using the related, but more general Zhu-Nakamura formula and were shown to give a very good agreement with FSSH for molecular systems.[51,52] The LZ formula was also used recently to treat intersystem crossing both in the context of AIMS[138] and surface hopping.[139]

The primary motivation for all of these studies is due to the fact neither formula for hopping probability requires the knowledge of nonadiabatic couplings. This is a very attractive property mainly because it almost completely trivializes interfacing the surface hopping algorithm with any electronic structure method/code which can compute analytical gradients of excited states, but also because the coupling calculation is the most computationally expensive and hardest to implement part of the FSSH algorithm. Algorithms based on the Zhu-Nakamura formulas were also used in recent attempts to use potential energy surfaces based on machine learning models in nonadiabatic dynamics simulations since the highly peaked nonadiabatic couplings are hard to learn using machine learning algorithms.[140,141]

Section 3.2.1 details a practical implementation of a surface hopping algorithm based on equation 2.55 with care taken to assure the numerical stability of the calculated probabilities. In section 4.1 the LZ algorithm is compared with FSSH for two prototype molecules with a significantly different photochemistry, pyrazine and pyrrole.

2.2.5 Nuclear phase space distributions

In 2.2.1 we decided to represent the nuclei by phase space distributions $\{\rho_A(\mathbf{R},\mathbf{v},t)\}$ which are in turn represented by a set of trajectories $\{(\mathbf{R}_i(t),\lambda_i(t))\}$. So far we have focused exclusively on how to propagate these trajectories, but here we will focus on how to generate the initial distributions from which trajectories are sampled.

Typical photochemical studies are concerned with processes which follow a path where: 1) the system is in equilibrium in the ground state, 2) a photon is absorbed exciting the system to some upper state and 3) the excess energy is either redistributed through nonradiative processes or is lost through luminescence. Surface hopping algorithms are at their best when modelling nonradiatve processes which means that steps 1) and 2) usually need to be taken into account before starting the SH calculation. This involves specifying a ground state phase space distribution and a way to "lift" it to the excited states to model the absorption process.

Ground state ensemble: Two different approaches to describing the ground state phase space distribution are common. These approaches differ based on whether the quantum nature of the nuclei (mainly the zero-point energy, ZPE) is taken into account or not.

For "classical" sampling, it is usual to run long Born-Oppenheimer ground state dynamics trajectories with a thermostat employed to ensure a canonical ensemble. After the trajectories are fully equilibrated and if they are long enough, the set of geometries and velocities from these trajectories represent a classical phase space distribution. Initial conditions are then obtained by randomly sampling snapshots from these trajectories. Under this procedure, the system is imparted with a thermal energy of $\varepsilon_T = kT$.

The standard "quantum" sampling procedure is to construct the Wigner quasiprobability distribution for the system which is the quantum analogue of a classical probability distribution in phase space

$$W(q,p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} \langle q+q'|\hat{\rho}|q-q'\rangle\, e^{-2ipq'/\hbar} dq' \tag{2.56}$$

where $\hat{\rho}$ is the density matrix of the system and $\langle q|\psi\rangle = \psi(q)$ is the wave function in position representation. For a harmonic oscillator in the ground state, this expression can be evaluated analytically,

$$W(q,p) = \frac{1}{\pi\hbar}\exp\left(\frac{m\omega}{\hbar}q^2 + \frac{1}{m\omega\hbar}p^2\right), \tag{2.57}$$

where m and ω are the mass and frequency of the oscillator. Analytical expressions are available also for excited states of the harmonic oscillator. Temperature effects can be included by sampling from each state based on the Boltzmann distribution,[104] or by directly inserting a canonical ensemble density matrix into equation 2.56 which gives

$$W(q,p) = 2\tanh\left(\frac{\hbar\omega}{2kT}\right)\exp\left(-\tanh\left(\frac{\hbar\omega}{2kT}\right)\left(\frac{m\omega}{\hbar}q^2+\frac{1}{m\omega\hbar}p^2\right)\right). \tag{2.58}$$

Equations 2.57 and 2.58 are exact for the harmonic oscillator, but they can also be used to approximate the quantum distributions of positions and momenta for real systems. In this case, each normal mode is represented as a harmonic oscillator with mass m_j and frequency ω_j and no interaction with the other modes so the full distribution can be written as a product of one-dimensional distributions

$$W(\mathbf{q},\mathbf{p}) = \prod_j W_j(q_j,p_j) \tag{2.59}$$

and each mode can be sampled individually. Another property of equations 2.57 and 2.58 to notice is that the position and momentum are uncorrelated which means they can be sampled separately (which is not always the case).

At room temperature the ZPE is significantly higher than the thermal energy of the system. This means that a simulation with initial conditions sampled from the Wigner distribution has significantly more vibrational energy available and, thus, can explore a wider region of phase space. Comparing the amplitudes $Q^a = (2\varepsilon/m\omega^2)^{1/2}$ of harmonic oscillators with energy ε_T and ε_Q:[101]

$$\frac{Q^{aT}}{Q^{aQ}} = \frac{2kT}{\hbar\omega}^{1/2}, \tag{2.60}$$

we see that the difference is greater for high energy modes.

Equation 2.56 which includes nuclear quantum effects is obviously a more realistic description of the quantum systems being studied than a purely classical view of the nuclei. Sampling using equations 2.57 or 2.58 and 2.59 is also much less computationally demanding than a classical trajectory since it requires only a single evaluation of the Hessian of the system. On the other hand, representing the ground state of a molecule as a set of independent harmonic oscillators is a very large approximation which can severely limit the range of sampled conformations (for example when more than one minimum is accessible in the ground state) but

also extend it where the energy is underestimated by the harmonic approximation (particularly for low-frequency, large-amplitude motions[102]). For select cases this approximation can be improved, such as by sampling randomly from two or more Wigner distributions when the molecule has a few well separated minima,[31] but there is no way to systematically improve the distribution. Special care also has to be taken when an explicit environment is included in the calculations.[102,142]

Another quantum sampling scheme that is gaining traction is the of path integral molecular dynamics (PIMD). In this method, nuclear quantum effects are included by a classical evolution of an extended system composed of multiple copies ("beads") of the real system connected by harmonic potentials. The computational cost of this procedure is higher than classical MD since it involves propagating multiple beads simultaneously. This cost has steadily been decreasing with the development of quantum thermostats with the approach based on the generalized Langevin equation (GLE) with only 4-10 beads required to obtain a converged distribution.[143] This thermostat approach converges to the exact quantum distribution in the case of a harmonic potential, but unlike the analytical Wigner formula also provides a reasonable estimate for highly anharmonic cases.[143] The development of path integral based methods is an active field of research, for an overview of the topic the reader is referred to the recent review by Markland and Ceriotti and references therein.[144]

This procedure has been used by several groups for sampling geometries to simulate absorption spectra.[32,145–148] Momentum distributions can also be obtained using PIMD in conjunction with the GLE thermostat which means it can be used to sample initial conditions for SH simulations.[149,150] Several promising early attempts have also been made to directly join PIMD with SH approaches to include nuclear quantum effects in trajectory based nonadiabatic dynamics simulations.[25–27,29]

Initial excited state distribution: Several attempts have been made to include the electric field in photochemical studies to directly model the absorption process.[151–154] However, the more usual approach is to model the absorption separately and start the dynamics simulation with a wave packet (phase space distribution) already in the excited state.

Within the electric dipole approximation the semi-classical expression for the photoabsorp-

tion cross-section is

$$\sigma(E) = \frac{\pi}{3\hbar c\varepsilon_0 E} \sum_F \int \rho_0 |\mu_{0F}|^2 \delta(E - E_{0F}) d\mathbf{R} \tag{2.61}$$

where the sum goes over all possible final states F and it is assumed that initially the full population is in the ground electronic state, $E_{0F} = E_F(\mathbf{R}) - E_0(\mathbf{R})$ is the vertical excitation energy and μ_{0A} is the electronic transition dipole moment

$$\mu_{0A} = \langle \psi_0(\mathbf{r};\mathbf{R}) | \mu_e(\mathbf{r};\mathbf{R}) | \psi_A(\mathbf{r};\mathbf{R}) \rangle \,, \tag{2.62}$$

which can be expressed in terms of oscillator strength f_{0F} as

$$\mu_{0A}^2 = \frac{3\hbar^2 e^2}{2m_e E_{0F}} f_{0F} \,. \tag{2.63}$$

Equation 2.61 tells us that the probability of absorbing a photon of a given energy depends for a given geometry is possible only for $E = \Delta E_{0F}$ and depends linearly on the oscillator strength.

The only other consideration to take into account is the source of photons. The most common approach is to assume that the excitation is performed by a weak laser pulse (small perturbation) with an extremely narrow time envelope. This translates to a very wide energy envelope meaning that all photon energies in a relevant energy window can be assumed to be equally probable. In practice, this means sampling points from the ground state distribution weighted only by their oscillator strength and using the same initial time for all selected trajectories.

Other starting assumptions can be that the system is excited by a continuum wave laser field[150] or by black-body radiation.[105] In the former case the photon energy distribution is narrow, if the energy of the laser is at the tail of the spectrum it might be difficult to sample enough points in the selected energy window. The solution offered in this case is to add a Lagrange multiplier ensuring the resonance condition during the ground state dynamics so that only points in the correct energy window are sampled.[150] In the case of black-body radiation, the radiation can be treated as an ensemble of coherent short pulses[155] and averaging over time by displacing the initial time of each trajectory can correct for the fact that not all molecules in the ensemble are excited at the same moment.[105]

Nuclear ensemble approach: Because of the delta function in equation 2.61, the photoabsorption cross-section (and many other spectroscopic observables) cannot be practically calculated by equation 2.18 since it would give a line spectrum regardless of the number of points sampled. Instead, the delta functions are replaced by finite width line shape functions centered at $E = \Delta E_{0F}$ and the absorption spectrum is calculated as

$$\sigma(E) = \frac{\pi e^2 \hbar}{2mc\varepsilon_0 E} \sum_F \sum_i^{N_p} w_i E_{0F}(\mathbf{R}_i) f_{0F}(\mathbf{R}_i) k\left(\frac{E - E_{0F}(\mathbf{R}_i)}{\delta_F}\right), \tag{2.64}$$

where $k(\frac{E-E_{0F}(\mathbf{R}_i)}{\delta_F})$ is the normalized line shape and δ_F a parameter (bandwidth) which determines the width of the function. This equation corresponds to the statistical method called kernel density estimation (KDE) whose goal is to estimate an underlying distribution from a set of points sampled from the distribution[156,157] (in this method k is called the kernel). In the current context, this approach is usually called the nuclear ensemble approach and replacing the delta functions with line shapes can easily be justified by invoking the uncertainty principle.

Crespo-Otero and Barbatti[106] present a formal derivation of the equation above starting from the expression for the cross-sections within the BO approximation and employing a Monte Carlo procedure for solving integrals containing the nuclear wave functions of the initial and final vibrational state. Depending on the assumptions made for the nuclear wave functions, the line shape is either a Lorentzian

$$k\left(\frac{E - E_{0F}(\mathbf{R}_i)}{\delta_F}\right) = \frac{1}{\pi\delta_F} \frac{\delta_F^2}{\delta_F^2 + (E - E_{0F})^2}, \tag{2.65}$$

or a Gaussian

$$k\left(\frac{E - E_{0F}(\mathbf{R}_i)}{\delta_F}\right) = \frac{1}{\delta_F\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{E - E_{0F}(\mathbf{R_i})}{\delta_F}\right)^2\right\}. \tag{2.66}$$

The former are used to describe homogeneous broadening while the latter are associated with broadening caused by collisions. The broadening parameter in this view is associated with the lifetime of the excited state and is usually determined empirically to obtain the best fit with experimental data. On the other hand, Sršeň and co-workers offer a statistical view of the method and its limitations.[32] They use Silverman's rule of thumb to set the broadening

parameter and bootstrap resampling to set confidence intervals for the calculated observable.

The computational cost of photochemical studies is often directly related to the number of sampled points needed to obtain converged observables so the goal is always to minimize this number. Much work has been devoted to this goal and the suggested solutions depend on what is being studied. In the previous section we mentioned the possibility of adding a constraint to ground state dynamics to more efficiently sample the tail end of the spectrum.[150] For sampling rare events in the surface hopping view there is the army ants algorith.[158] When the calculation of the observable is more expensive than sampling the phase space distribution, machine learning can be used to build a model for quickly evaluating (approximate) observables for many points.[33]

Another useful approach is importance sampling[31] for situations when the observables need to be calculated for multiple similar distributions (for example at different temperatures). The expectation value for some distribution ρ' can simply be expressed in terms of another known distribution ρ if the ratio between the distributions is known

$$\langle \mu \rangle = \int \mu(\mathbf{R}) \rho'(\mathbf{R}) d\mathbf{R} = \int \mu(\mathbf{R}) \frac{\rho'(\mathbf{R})}{\rho(\mathbf{R})} \rho(\mathbf{R}) d\mathbf{R}. \tag{2.67}$$

In this case, if a collection of points have been sampled from ρ, one only needs to adjust their weights $w(\mathbf{R_i}) = \rho'(\mathbf{R_i})/\rho(\mathbf{R_i})$ to obtain results for ρ'.

2.3 Photochemical problems

In the following section we will take a look at the work done so far on a few of the prototypical problems in photochemistry. The systems studied are pyrrole and pyrazine (Figure 2.3). The photochemical processes in these systems are very different, but both systems are relatively small and the studied processes occur on very fast (femtosecond) time scales which allows them to be studied using a variety of advanced methods both experimental and theoretical. Despite this fact, some basic aspects of these reactions, such as which electronic states are relevant and to what extent, are still under debate.

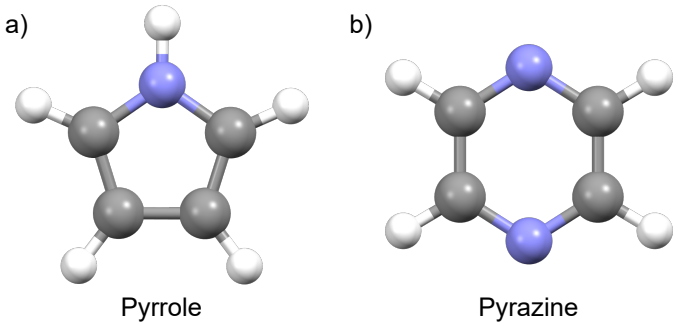


Figure 2.3: The prototypical systems under investigation, a) pyrrole and b) pyrazine.

2.3.1 Pyrrole

Pyrrole is a simple five-membered heterocycle commonly found as a building block of biologically important molecules such as proline, indole and porphyrins of which the best known is probably heme.

From the perspective of photochemistry, it is interesting due to being, along with phenol, imidazole and indole, a prototypical example of excited state hydrogen detachment driven by dissociative $^1\pi\sigma^*$ states as noted by Domcke and coworkers.[66,67] This process usually involves excitation to a $^1\pi\pi^*$ state, followed by internal conversion to the (usually dark) $^1\pi\sigma^*$ which leads to a conical intersection with the ground state along the hydrogen detachment coordinate. The process is governed by the relative energy of the two excited states and the position of the CI between them. The molecules mentioned above exhibit a complex photochemistry while their size and prevalence as building blocks of larger chromophores make them ideal targets for a bottom-up approach to photochemical studies. Over the past twenty years, this process in pyrrole has been widely studied from both the theoretical[159–164] and experimental[68,165–173] side.

In terms of spectroscopy, despite numerous studies of the vertical excitation energies of pyrrole,[174–180] there are still some disagreements about the character and relative positions of the lowest excited states. There are two $^1\pi\sigma^*$ states, $A_2(\pi\sigma^*)$ and $B_1(\pi\sigma^*)$, corresponding to excitations from either the HOMO or HOMO-1 π orbitals into an orbital with significant antibonding σ^* character. While these states are often denoted σ^*, at the ground state minimum geometry the character of the particle orbital is of Rydberg $3s$ type.[180,181] The σ^* character becomes dominant only as the N-H bond is extended. Unlike in phenol and imidazole, these

states are the lowest excited states of pyrrole. Both are dark states, but the $A_2(\pi\sigma^*)$ state is possible to detect as a very weak tail of the absorption spectrum up to 250 nm where the 0-0 transition lies. On the other hand, the $B_1(\pi\sigma^*)$ state is not experimentally verified. There are also two relevant $^1\pi\pi^*$ states. One of these is the bright $B_2(\pi\pi^*)$ state which is agreed upon as the main state contributing to the lowest band in the UV spectrum of pyrrole between 5.6 and 6.2 eV (220 to 200 nm).[182] The other is the $A_1(\pi\pi^*)$ state which is approximately an order of magnitude darker which has sometimes been placed below and sometimes above $B_2(\pi\pi^*)$. Between the $^1\pi\sigma^*$ and $^1\pi\pi^*$ theoretical studies usually predict the presence of two Rydberg states, $A_2(\pi 3p_z)$ and $B_1(\pi 3p_y)$. The third p type Rydberg state ($3p_x$), is usually placed at significantly higher energies. An overview of the vertical excitation energies at various levels of theory reported in the literature is given in table 4.3 in section 4.4.

The absorption spectrum was also simulated at different levels of theory,[101,180,183] with the best agreement with experiment achieved by Neville *et al.*[180] based on MCTDH calculations by simulating individual spectra obtained from vertical excitations to $A_1(\pi\pi^*)$, $B_1(\pi 3p_y)$ and $B_2(\pi\pi^*)$. A model comprised of six states and ten modes based on energies obtained at the CASPT2(8,8) level using the aug-cc-pVDZ basis with additional diffuse s and p functions was used for $A_1(\pi\pi^*)$ and $B_2(\pi\pi^*)$ states. The $B_1(\pi 3p_y)$ was described using a single state model with ten harmonic modes. It was found that the sharp peak at 5.85 eV comes from the $B_1(\pi 3p_y)$ state, in agreement with previous assignments.[176,182,183] While it was confirmed that most of the intensity of the lowest peak comes from excitation to the $B_2(\pi\pi^*)$ state, it was also found that coupling with the $A_2(\pi 3p_z)$ and $B_1(\pi\sigma^*)$ states significantly influences the shape and position of the peak.

The dynamics following excitation of pyrrole to the $A_2(\pi\sigma^*)$ state has also been studied by a number of groups.[159,160,162,164,173,184,185] The primary relaxation mechanism is agreed to be a fast ejection of H-atoms. The time scale of this process was always reported below 60 fs, but it has been found that this mechanism depends on the energy imparted along the N-H mode.[159] Wu *et al.*[173] also performed dynamics from the $B_1(\pi\sigma^*)$ states based on the model Hamiltonian developed by Neville *et al.*[180] This resulted in two competing processes, direct dissociation from $B_1(\pi\sigma^*)$ and to internal conversion to $A_2(\pi\sigma^*)$ followed by dissociation with time scales of 19 and 61 fs, respectively.

A few dynamics studies have also been performed starting from higher excitation energies,

mostly on the $B_2(\pi\pi^*)$ state.[161–163] At these energies other deactivation mechanisms, mainly ring puckering and ring opening, start becoming relevant. These mechanisms were first investigated by Barbatti *et al.* who investigated different CIs of the molecule at the MRCI level.[181] In further studies, FSSH dynamics simulations were performed at the MRCI[161] and TDDFT[162] level. In the former study, 80% of the trajectories deactivated through N-H stretching and 13% deactivatd throguh ring puckering or opening mechanisms. A deactivation time constant of 139 fs was obtained. In the latter study, three excitation windows centered at 5.0, 5.85 and 6.43 eV were explored. The two higher energy windows populated mostly the $B_2(\pi\pi^*)$ state and led to relaxation through N-H stretching either on the S_1 or on the S_2 surface in $> 90\%$ of the trajectories. At this level of theory the ring opening mechanism was not noticed and the ring puckering mechanism was present in only 4%-7% of the trajectories. The reported time scales were 20 fs, 166 fs and 184 fs for the three time windows. A somewhat faster deactivation ($\tau \approx 100$ fs) after excitation to the $B_2(\pi\pi^*)$ state was obtained by Faraji and coworkers.[163] In this study, a model Hamiltonian was built based on MRCI computations and including five electronic states (ground state, the two $^1\pi\sigma^*$ and the two $^1\pi\pi^*$ states) and 24 modes (all vibrational degrees of freedom of the molecule). Anharmonic treatment was adopted only for the N-H stretching mode so this was the only possible relaxation mechanism.

From the experimental side, the time scale of H-elimination in pyrrole following excitation to the $A_2(\pi\sigma^*)$ state was first determined by Lippert *et al.*[167] who reported two time constants ($\tau_1 = 110 \pm 80$ fs and $\tau_2 = 1.1 \pm 0.5$ ps) for NH fission after excitation at $\lambda = 250$ nm. The former was assigned to direct hydrogen detachment along the $A_2(\pi\sigma^*)$ surface, while the latter was attributed to detachment from the vibrationally excited ground state. The shorter time scale was confirmed by Roberts *et al.* using time-resolved ion yield and velocity map imaging at multiple excitation wavelengths. The excitation wavelength dependent time constants for N-H bond fission have been obtained from the kinetic fit of the normalized H^+ signal transient (REMPI probed H-atoms) as a function of pump-probe delay time. The exponential rise of the H^+ signal at positive pump-probe delay times revealed single time constants of $\tau = 126 \pm 28$ fs and $\tau = 46 \pm 22$ fs for H-elimination at 250 nm (band origin) and 238 nm, respectively.[68] The authors also investigated the dynamics of monodeuterated pyrrole (pyrrole-d_1) and reported a time constant of 1.4 ± 0.3 ps for N-D dissociation at 250 nm giving rise to a kinetic isotope effect of $KIE \approx 11$. These results are consistent with the existence of a small barrier on the S_1

surface as predicted by theory.[159] On the other hand, Wu *et al.* used pump pulses at 242 and 236 nm and reported an ultrafast decay of the system from the ionization window on a single timescale of less than 20 fs for both wavelengths.[173] Similar results were obtained by Kirkby *et al.*[186] who reported a timescale of < 50 fs following excitation at 249.5, 245 and 240 nm.

At shorter wavelengths other states, primarily the $B_2(\pi\pi^*)$ state, are accessible. Roberts *et al.*[68] reported two time constants of $\tau_1 = 52 \pm 12$ fs and $\tau_2 = 1.0 \pm 0.4$ ns following excitation at 200 nm the first of which they assigned to the $^1\pi\sigma^*$ mediated process while the second was assigned to H-elimination from vibrationally excited ground state species. At the same wavelength Kirkby *et al.*[186] reported a timescale of < 50 fs. Wu *et al.* excited at 217 nm and reported two time constants of $\tau_1 = 13 \pm 10$ and $\tau_2 = 29 \pm 10$ fs which are in reasonable agreement with the theoretically predicted time scales reported in the same study.

While a general picture of the photodynamics of pyrrole is known, some details are still not understood. We will attempt to address these in section 4.4. The main goal of that section will be to eliminate the discrepancy between the computed and measured time constants for pyrrole excited to higher lying states. In addition to this, significant attention will be given to the exact nature of the $B_2(\pi\pi^*)$ state and the interaction between this state and the *p*-type Rydberg state which are near in energy.

2.3.2 Pyrazine

Pyrazine is a prototypical system for photochemical studies as a standard example of vibronic coupling.[69] A large part of the reason for this is due to the convenience of studying the system thanks to its high symmetry and the fact that the coupling between the two most relevant states of the system can be described reasonably using a two dimensional model. These are the $B_{2u}(\pi\pi^*)$ and $B_{3u}(n\pi^*)$ states visible in the absorption spectrum at, respectively, 4.81 and 3.83 eV. Due to symmetry considerations, these can only be coupled in first order through normal modes of b_{1g} symmetry, of which there is only one in the molecule, the Q_{10a} normal mode (the normal mode notation used here, as is standard in the literature, is based on the normal modes of benzene.[69,187])

Over the years, the photochemistry of the system has been widely studied. From the theoretical side, a series of papers by Domcke and coworkers[70–72,188] investigated the process using two-state vibronic coupling models containing three to four modes. These models captured the

physics responsible for internal conversion in pyrazine and were used to calculate various spectra of pyrazine in excellent agreement with experimental ones. Similar results were obtained by others using different electronic structure and quantum dynamics methods.[189] Over time full dimensional models were developed, first with additional modes included heuristically[190–192] and finally a quadratic vibronic coupling model encompassing all 24 degrees of freedom was built.[193] The developed two state pyrazine models have often been used to test new numerical methods in quantum dynamics.[10,194–199]

From the experimental side, Stert *et al.*[200] measured the lifetime of the $B_{2u}(\pi\pi^*)$ state using time-resolved photoelectron spectroscopy (TRPES) to approximately 20 fs, in excellent agreement with theoretical predictions. Recent advances in the time-resolution of the experiments finally provided conclusive direct evidence of the ultrafast internal conversion in pyrazine with a reported lifetime of around 22 fs.[201,202] On significantly longer time scales, both internal conversion to the ground state and intersystem crossing to the triplet manifold are seen with timescales in the range of tens or hundreds of ps.[202] To our knowledge, direct spectroscopic evidence of the $A_u(n\pi^*)$ state has not been reported.

From the mixed quantum-classical perspective, Werner *et al.*[73] performed SH dynamics on the system at the TDDFT/B3LYP/TZVP level. They reported significant populations of $A_u(n\pi^*)$ and $B_{2g}(n\pi^*)$ in the early moments of the dynamics calculations. However, this was done without a diabatization procedure, simply equating the states with the S_2 and S_4 adiabatic states, respectively. In a subsequent study, Tomasello *et al.*[74] calculated the TRPES at the same level of theory to obtain a reasonable agreement with experiment. Based on static calculations and potential energy scans, Lin and co-workers[203] also suggested that the $A_u(n\pi^*)$ state should be important in the ultrafast relaxation mechanism of the $B_{2u}(\pi\pi^*)$ state. To see whether these states were actually populated during the dynamics, Sala *et al.*[53] built two, three and four state models spanning ten relevant normal modes based on electronic structure calculations at the XMCQDPT2(10, 8) level. Through these calculations they confirmed that the $A_u(n\pi^*)$ state is indeed significant in the early dynamics while the $B_{2g}(n\pi^*)$ state was never populated. They also found[54] that the $A_u(n\pi^*)$ state was responsible for internal conversion to the ground state at a longer time scale. Sun and co-workers used the same two and three state model to compute multiple spectral signals and found that including the $A_u(n\pi^*)$ state has a significant effect on the calculated signals.

On the other hand, Kanno and co-workers performed calculations at the MRCISD(10,8) also including four electronic states, from quantum dynamics propagated in different subspaces spanned by two modes[204] they concluded that the $A_u(n\pi^*)$ state is not populated in the early moments after excitation. In a continuation of this work, Mignolet and coworkers[7] simulated the vacuum ultraviolet photoelectron spectrum of pyrazine at the same level of theory. They computed signals of the three states at their potential minima along the Q_{6a} normal mode. When computed in this way, the main signal coming from the ionization of $A_u(n\pi^*)$ is centered at a photoelectron kinetic energy of 4.90 eV, where no peak is visible in the time resolved photoelectron imaging experiment.[202]

Most of the work on the photochemistry of pyrazine has focused on these lowest excited states. Some work has been done to study higher lying states,[205,206] but these states are not relevant in the current context. In section 4.5 we will take another look at the internal conversion in pyrazine from the mixed quantum-classical perspective. We will employ a higher level of theory than previous SH results and specifically look for signs of the $A_u(n\pi^*)$ state whose significance in the dynamics is still disputed.[7,204]

§3. COMPUTATIONAL METHODS

At the mixed classical-quantum level we usually need to employ three distinct sets of computational methods to obtain a full picture of a photochemical problem. First, we need a method for treating the electronic structure of the system. Next, we need a method for treating the nuclear degrees of freedom: generating an initial ensemble and propagating it. Finally, we need a method for simulating observables which can be directly verified by experiments. Each of these steps is important and the computational methods chosen for each are described in this section.

3.1 Electronic structure

To explain a photochemical reaction means to understand the interplay between the relevant electronic states and nuclear motions. This poses a significant challenge for electronic structure methods since it requires accurately describing states with significantly different electronic characters at an equal level and at a wide range of nuclear configurations. This includes both bound and dissociative valence state, diffuse Rydberg states and highly delocalized or charge-transfer states. This problem cannot be solved using blunt force since any method which would be accurate in all cases would also be prohibitively expensive for practical calculations. Because of this, the choice of the level of theory for the electronic structure calculation is the basis for any successful photochemical study.

For dynamics simulations, this choice is limited by the need for thousands of evaluations of energies, excited state gradients and nonadiabatic couplings along the ensemble of trajectories. The need for excited state gradients specifically means that most state of the art methods for excitation energies cannot be used since efficient analytical gradients are difficult to implement for most of these methods. However, these methods are often useful for benchmarking less accurate methods at selected relevant geometries.

In the studies and tests reported in chapter 4 multiple electronic structure methods and basis sets will be used. Mostly, we will be using the ADC(2) method (see below) which has proven to provide a good balance between cost and accuracy for the molecules with ten to twenty atoms studied here. For these calculations, the efficient implementation in the Turbomole program package[207,208] based on the resolution of the identity (RI) approximation[209–211] was used. In

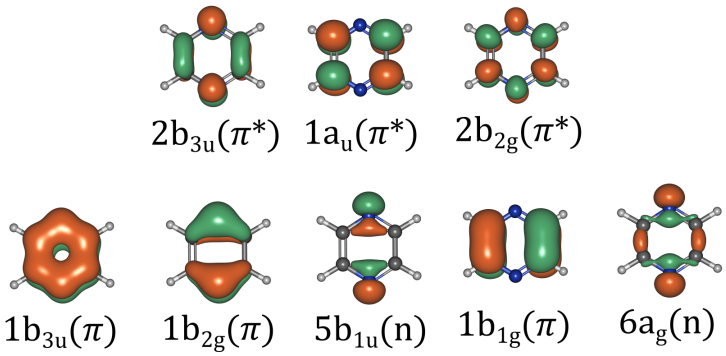


Figure 3.1: Orbitals included in the active space for XMS-CASPT2 calculations of pyrazine.

all cases where ADC(2) was used, the core electrons were frozen. All ADC(2) calculations were performed using the cc-pVNZ[212] and aug-cc-pVNZ[213] family of basis sets. In the case of pyrrole, calculations FSSH dynamics simulations were also repeated using the spin-component scaled version of ADC(2).[214]

Additional calculations for pyrrole were performed at the TDDFT level using the B3LYP[215,216] functional, also using the Turbomole program package.[217] The def2-TZVPD basis set was used in this case.[218] As will be discussed in detail in section 4.4, this method gives a significantly different description of the bright $B_2(\pi\pi^*)$ state of pyrrole. For pyrazine, additional calculations were performed at the XMS-CASPT2/aug-cc-pVDZ using the BAGEL program package.[39] The five lowest states of the neutral molecule and the seven lowest states of the cation were calculated with an active space consisting of 10 electrons (9 for the cation) in 8 orbitals: $\pi(1b_{3u})$, $\pi(1b_{2g})$, $n(5b_{1u})$, $\pi(1b_{1g})$, $n(6a_g)$, $\pi^*(2b_{3u})$, $\pi^*(1a_u)$ and $\pi^*(2b_{2g})$ (Figure 3.1). These calculations were performed for a subset of geometries sampled from the dynamics calculations to simulate the photoelectron spectra since this level of theory provided significantly better ionization potentials than ADC(2).

In section 4.1 we will also compare FSSH and LZSH algorithms to quantum dynamics studies. For this we will employ model potentials developed by Sala *et al.*[53] based on electronic structure calculations at the extended multi-configuration quasi-degenerate second-order perturbation theory (XMCQDPT2)[219] level using an active space of ten electrons in eight orbitals and the aug-cc-pVDZ basis set.[212] In the two state model the normal modes which give rise to non-vanishing first-order couplings between the $B_{3u}(n\pi^*)$ and $B_{2u}(\pi\pi^*)$ states are included, these are the totally symmetric modes Q_{6a}, Q_1, Q_{9a} and Q_{8a} and the B_{1g} mode Q_{10a}. The three state model also includes the $A_u(n\pi^*)$ state which couples to the $B_{3u}(n\pi^*)$ and $B_{2u}(\pi\pi^*)$ states

through B_{2g} and B_{3g} modes, respectively. Of these, the Q_4, Q_5, Q_3 and Q_{8b} modes are included in the model. Benchmark quantum dynamics calculations for these systems were performed using the MCTDH method, for the details of these calculations the reader is referred to the original publication.[34]

3.1.1 Algebraic diagrammatic construction

The algebraic diagrammatic construction scheme is based on the theory of the polarisation propagator.[57] This propagator has the form

$$\Pi(\omega) = \mathbf{x}^{\dagger}(\omega - \Omega)^{-1}\mathbf{x}, \tag{3.1}$$

where $\mathbf{x} = \langle \psi_n | c_r^{\dagger} c_s | \psi_0 \rangle$ is a matrix of transition amplitudes and Ω is a diagonal matrix containing the vertical excitation energies ω_n. The procedure for obtaining these excitation energies involves writing the operator in a non-diagonal form based on a representation of the effective Hamiltonian after which searching for the poles of the propagator (excitation energies where eq. 3.1 goes to infinity) can be reduced to a diagonalization of a Hermitian matrix. This method was later reformulated in the so-called intermediate-state representation.[58] In this representation, excited states are built starting from a (correlated) ground state wave function $|\psi_0\rangle$ and consecutively acting on it with classes of standard configuration interaction (CI) excitation operators

$$\mathbf{C}_I = \left\{ c_o^{\dagger} c_a; \quad c_o^{\dagger} c_p^{\dagger} c_a c_b; \quad c_o^{\dagger} c_p^{\dagger} c_q^{\dagger} c_a c_b c_c; \quad \ldots \right\}, \tag{3.2}$$

where indices $a, b, \ldots$ denote occupied orbitals and $o, p, \ldots$ denote virtual orbitals. When the ground state is composed of multiple determinants, the states obtained in this way are not orthogonal. Because of this, an orthogonalization procedure also needs to be employed after acting with each class (single excitations, double excitations, ...) of operators. For ADC methods the ground state is the Møller-Plesset (MP) wave function and orthogonalization is done using a combination of the Gram-Schmidt process (between excitation classes) and symmetric orthogonalization (within the same excitation class).[220,221] A series of ADC(n) methods is available by limiting the order of the MP ground state and the level of CI excitation operators used to n-th order. Coupled cluster (CC) methods can be derived in a similar way with a different ground state and orthogonalization procedure.

Alternately, ADC(2) can be considered[222] as a further approximation of the second-order approximate coupled cluster singles and doubles model (CC2).[223] However, the symmetrisation of the Jacobian in ADC(n) methods leads to some key advantages over CCn methods. They are fully size-intensive and (starting from the third order) have better scaling with the number of basis functions. Most importantly, ADC methods involve diagonalization of a Hermitian matrix. This means that they are more stable close to conical intersections which is critical for nonadiabatic dynamics simulations.[59] Another advantage of this is that excited state properties are twice as fast to calculate since the left and right eigenvectors are the same. On the other hand, the CCn methods include a description of the ground state which is superior to the MPn methods. This means that ADC(2) is less reliable when the ground state has significant multi-reference character which results in an unbalanced description of the various excitations of the system. However, in such cases CCn also quickly becomes unreliable since it is still a single reference method.[224]

A more detailed overview of ADC(n) methods is presented in the recent review by Drew *et al.*[225] and the references within. The accuracy of these methods has been assessed in a number of benchmarks.[226–228] Mean errors for ADC(2) have been reported as 0.22 ± 0.38 eV for various singlet states[226] and 0.01 ± 0.16 eV for charge transfer states,[228] results similar to CC2. Even better results are obtained using the ADC(3) method. This method is less accurate than CC3,[229] but has the advantage of $O(n^6)$ scaling which makes it significantly faster than CC3 for larger systems. Efficient ADC(3) analytical gradients have only recently become available[230] which makes the method a likely candidate for future benchmark nonadiabatic dynamics studies.

One additional alternative to mention are the spin-component scaled (SCS) and spin-opposite scaled (SOS) variants of MP, ADC and CC methods.[214,231,232] The former is an empirical correction based on scaling the same-spin and opposite-spin integrals appearing in the equations. In the latter, the same-spin integrals are completely neglected and opposite-spin integrals are scaled to compensate. The SOS-ADC(2) method scales approximately as $O(n^4)$ making it significantly faster for larger systems than ADC(2). For our purposes however, the main advantage of these methods is that they have been shown to alleviate the problems of these methods in the treatment of Rydberg states both in terms of vertical excitation energies and shapes of potential energy surfaces.[227,233]

3.1.2 Time-dependent density functional theory

Density functional theory (DFT) is an approach to electronic structure calculations based on the Hohenberg-Kohn[234] theorems. The first of these theorems states that the total energy for a system of electrons moving under the influence of an external potential $v_{ext}(\mathbf{r})$ (usually the electron-nuclei interaction) is uniquely determined by its electron density $n(\mathbf{r})$ or in other words, that the total energy is a functional of the electron density

$$E[n(\mathbf{r})] = \int v_{ext}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + F[n(\mathbf{r})]\,, \tag{3.3}$$

where $F[n(\mathbf{r})]$ is a universal functional completely defined by the number of electrons. The second Hohenberg-Kohn theorem states that the density which minimizes the total energy is the exact ground state density n_0 which allows for practical applications of DFT. The problem is that a part of the universal functional is not fully known. In the Kohn-Sham approach[235] a practical approach to DFT is given by replacing the interacting system with a fictitious system of non-interacting particles which generate the same density. For the non-interacting system, the single-particle orbitals (densities) can be obtained through an eigenvalue equation with an effective potential. In this approach, the unknown part of the energy functional is contained in the exchange-correlation energy $E_{xc}[n(\delta r)]$ which enters the effective potential as $\frac{\partial E_{xc}[n(\delta r)]}{\partial n(\delta r)}$. Since this part of the effective potential depends on the density itself, the equation is solved in an iterative manner improving the density and effective potential until self-consistent solutions are obtained. The key problem of DFT is finding the best possible approximation for the exchange-correlation functional $E_{xc}[n(\delta r)]$.

In 1984, Runge and Gross[236] extended DFT to time-dependent systems where they have shown that there exists a one-to-one mapping between a time-dependent potential in which the system evolves and the density of the system. This theorem provides the foundation of time-dependent density functional theory (TDDFT) which can be used to extend the applicability of DFT to excited states. From a practical perspective, the most widely used approach for excited state calculations is linear response TDDFT. In this approach, the variation of the system following a small perturbation depends only on the ground state density. This means that the excited states can be obtained directly from Kohn-Sham orbitals from a non-hermitian eigenvalue

equation of the form

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix}, \tag{3.4}$$

where the matrices **A** and **B** involve four-center integrals over pairs of occupied and virtual Kohn-Sham orbitals so have dimensions equal to those in the configuration interaction singles (CIS) approach. In fact, in the Tamm-Dancoff approximation the off diagonal elements **B** are assumed to be zero and the problem computationally reduces to the same form as CIS.

TDDFT is likely the most widely used method for simulating excited states as it provides reasonable accuracy in many cases at a very low computational cost. Mean errors for vertical excitation energies computed by TDDFT are usually reported in the 0.2-0.4 eV range.[237,238] However, it also fails qualitatively for certain types of excited states, mainly those with charge-transfer or double excitation character. Another issue is that there is no way to systematically improve TDDFT results which has resulted in an ever growing number of exchange-correlation functionals with no clear hierarchy and no sure way to decide which functional is best suited for a particular application.[238,239]

As long as it accurately describes the electronic states of the system under study, TDDFT is also mostly well suited for SH calculations. Multiple methods for calculating the couplings between TDDFT excited states are available.[41,42,56,65,120–122] The method is also quite numerically stable and retains its accuracy in wide areas of the potential energy surface. However, while it can predict the location of CIs, the exact shape of the intersections can be incorrect. This is especially true for intersections between the S_0 and S_1 states which are qualitatively incorrect.[240] This means that, like single reference wave function methods, TDDFT is not suited for following the dynamics of a system through type of intersection.

3.1.3 Complete active space methods

There are many situations where the ground state of a system cannot be adequately described by a single SD. In these cases the Hartree-Fock (HF) method fails qualitatively in a way that is difficult to overcome for single-reference methods built upon it and reference states consisting of more SDs are needed. These are usually generated by multi-configurational self-consistent field (MC-SCF) in which the wave function is a linear combination of SDs or configuration state functions calculated by simultaneous variation of both the linear combination coefficients and

orbitals from which the determinants are built. The complete active space self-consistent field (CASSCF)[241,242] method is the most commonly used variant of MC-SCF. In this method, an active space of relevant orbitals and electrons is selected and a full CI expansion of this active space is included in the MC-SCF wave function ansatz with all other orbitals always either doubly occupied or empty.

Due to the factorial scaling of the method with the size of the active space only a small number of electrons/orbitals can be treated in this way. Typically, this means that CASSCF calculations are limited to ten to twenty electrons in as many orbitals. Since excitations to/from most orbitals are not included at all in the CASSCF wave functions, most of the dynamical electronic correlation is not captured by this method so the errors of the method are quite large. This can be remedied by the complete active space second-order perturbation theory (CASPT2)[243,244] method which corrects the CASSCF wave functions using perturbation theory similar to the MP2 correction to HF wave functions. Depending on the details of how the perturbation is applied, single-state, multi-state[245] and extended multi-state (XMS-CASPT2)[246] variants of CASPT2 are available. The latter, based on quasi-degenerate perturbation theory was developed specifically to treat numerical problems arising in MS-CASPT2 when state mixing is strong.

Analytical gradients for CASPT2 have been derived in 2003.[247] However, the method has been employed for FSSH simulations only recently,[248] following the very efficient implementations of CASPT2 gradients[249] and nonadiabatic coupling vectors[128] in the BAGEL electronic structure code.[39] With its inclusion of static correlation and, consequently, ability to treat CIs involving the ground state, CASPT2 has very clear advantages over ADC(2) and similar single-reference methods with otherwise similar accuracy.[238] On the other hand, the method is still more difficult to use correctly than most single-reference methods and more prone to errors due to problems such as intruder states.

3.1.4 Other excited state electronic structure methods

The accurate evaluation of excited state energies and properties is a very active field of interest in chemistry and one that is still far from solved.[238] In the previous subsections, we have briefly covered the electronic structure methods used in the current work. These are also the methods most commonly used for on-the-fly nonadiabatic dynamics simulations.

Among single-reference methods, the coupled cluster family of methods, either in the linear response or equation of motion formulation, is extremely important for excited state studies. These methods can be quite accurate for a wide variety of excited states and range from CC2 which has a similar cost and accuracy to ADC(2) through CCSD to CC3, CCSDT and CCSDTQ. The latter three are benchmark methods for small to very small molecules with accuracy in the range of a few hundredths of an eV, but a scaling between $O(n^7)$ for CC3 to $O(n^{10})$ for CCSDTQ. These methods are widely used with efficient implementations available in many electronic structure codes along with analytic gradients for excited states.[209,211,250,251] Nonadiabatic coupling vectors have also been implemented,[40] but as mentioned in the ADC section, the problem of CC methods for on-the-fly dynamics is their inability to correctly treat conical intersections[59] so they are still mostly used in static calculations.

Among multi-reference methods, we can mention the density matrix renormalization group (DMRG),[252] *n*-electron valance state perturbation method (NEVPT2)[253] and various multireference CI methods.[254,255] DMRG is a method for optimization of wave function coefficients in terms of so-called matrix product states which can effectively expand the number of active orbitals for which a full CI treatment can be applied to around one hundred. Like CASSCF, by itself this method does not include enough dynamical correlation, but can be built upon and progress in this direction is ongoing.[256] NEVPT2 is a method related to CASPT2. It is size consistent and free of the intruder state problem,[253] but lacks the efficient analytical gradient implementations available for CASPT2 making it unsuitable for dynamics simulations. MR-CI is another way to add dynamical correlation to the CASSCF (or other MRSCF) wave functions by including (usually single and double) excitations out of all reference determinants. Like CASPT2, this method can provide quite accurate results for different kinds of excited states, but is also difficult to use and very expensive for large reference wave functions.

For a more detailed recent overview of the methods mentioned here see the book edited by Gonzales and Lindh.[257] For a discussion on the accuracy of the methods, the reader is referred to the recent review of benchmarks by Loos and coworkers.[238]

3.2 Implementation of surface hopping algorithms

The general scheme for a nonadiabatic dynamics simulation based on mixed quantum-classical methods is shown in figure 3.2. We can roughly separate the procedure into four consecutive,

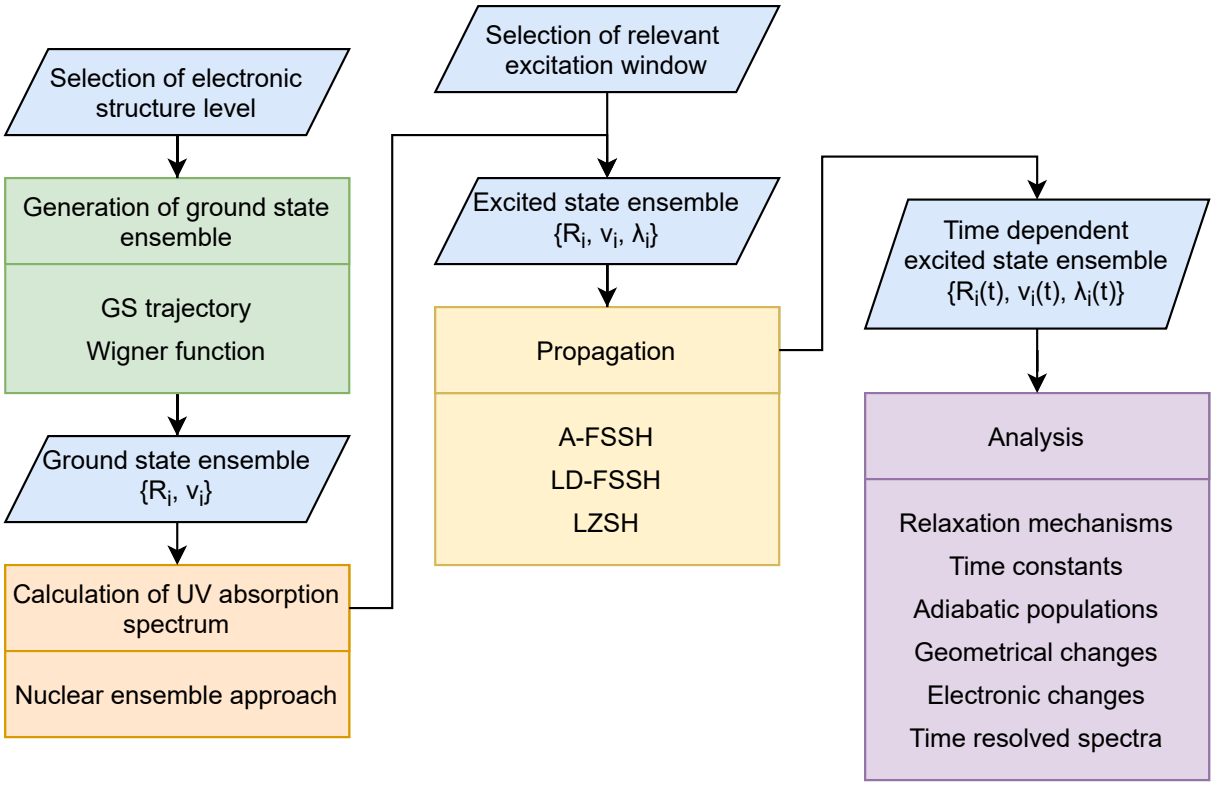


Figure 3.2: Workflow of a typical mixed quantum-classical nonadiabatic dynamics simulation. It consists of generating an initial ensemble in the ground state (green), simulating the photoexcitation (orange), propagating the ensemble of trajectories (yellow) and finally analysing the results (purple). The output of each step along this process is the input for the next step.

partially independent parts:

1. sampling the ground state ensemble,
2. simulating the excitation process,
3. propagating the chosen initial conditions,
4. analysing the results.

1) Established procedures for sampling the ground state phase space distribution $\rho(\mathbf{R},\mathbf{v})$, along with their strengths and weaknesses, were covered in section 2.2.5. In the present study, initial conditions $\{\mathbf{R}_i,\mathbf{v}_i\}$ will be sampled from the thermal Wigner distribution given by eq. 2.58. As already mentioned, the distribution given by this method includes the ZPE, but can be far from the actual distribution in the presence of normal modes which are highly anharmonic and in particular low-frequency torsional modes. To ensure that the harmonic approximation adequately describes the ground state PES of the systems under study, the energy of each point sampled based on the harmonic approximation E_i^h will be compared to the actual energy E_i^{qm}

calculated at that geometry. In cases where this difference is larger than a selected threshold t_{anh} the sampled point is discarded. For rigid systems with a deep ground state minimum, such as DNA nucleobases, pyrrole and pyrazine, the harmonic approximation is quite accurate so sampled points are rarely discarded.

2) A part of the ground state ensemble needs to be lifted to the excited states. To do this, the excited state energies and oscillator strengths $\{E_A, f_{0A}\}$ at each geometry $\{\mathbf{R}_i\}$ need to be calculated. From these, the UV absorption spectrum can easily be calculated using the nuclear ensemble approach. This is an important step since it provides a point early in the study where the results can directly be compared with the experiment to judge whether the chosen electronic structure level is adequate to describe at least the bright excited states of the molecule in the region close to the ground state minimum. After this, a subset of points sampled from the ground state ensemble needs to be chosen to represent the initial wave packet in the excited state. The choice of initial excited states is done based on oscillator strength. In addition to this, we also always have to limit our exploration to a specific energy window we are interested in. The probability to sample state A at geometry $\mathbf{R}_i$ is then calculated as

$$P_{iA} \propto f_{OA}(\mathbf{R}_i)\mathscr{E}(\omega - E_A(\mathbf{R}_i))\,, \tag{3.5}$$

where $\mathscr{E}(\omega - E_A(\mathbf{R}_i))$ is the energy window centered at ω limiting the dynamics exploration to the relevant energy range. In the case of pyrrole, three rectangular energy windows are selected centered at 200, 238 and 250 nm with widths of 2 nm. In the case of pyrazine, the energy window is a Gaussian function centered at $\omega = 4.7eV$ with a full width at half maximum of 0.2 eV chosen to best represent the pump pulse which was used by Horio *et al.*[202] to obtain the experimental time-resolved photoelectron spectrum.

3) The key step is the propagation of trajectories. While trajectory surface hopping algorithms are simple in principle, there are many different variants and corrections which might give different results depending on the situation. For this reason, one of the main goals of our implementation was for it to be as modular as possible to allow for freely choosing the best approach for a particular problem. Figure 3.3 shows the work flow during the propagation of a single trajectory. The core loop involves multiple steps. First the nuclear degrees of freedom are propagated. Since electronic structure calculations are the computational bottleneck of

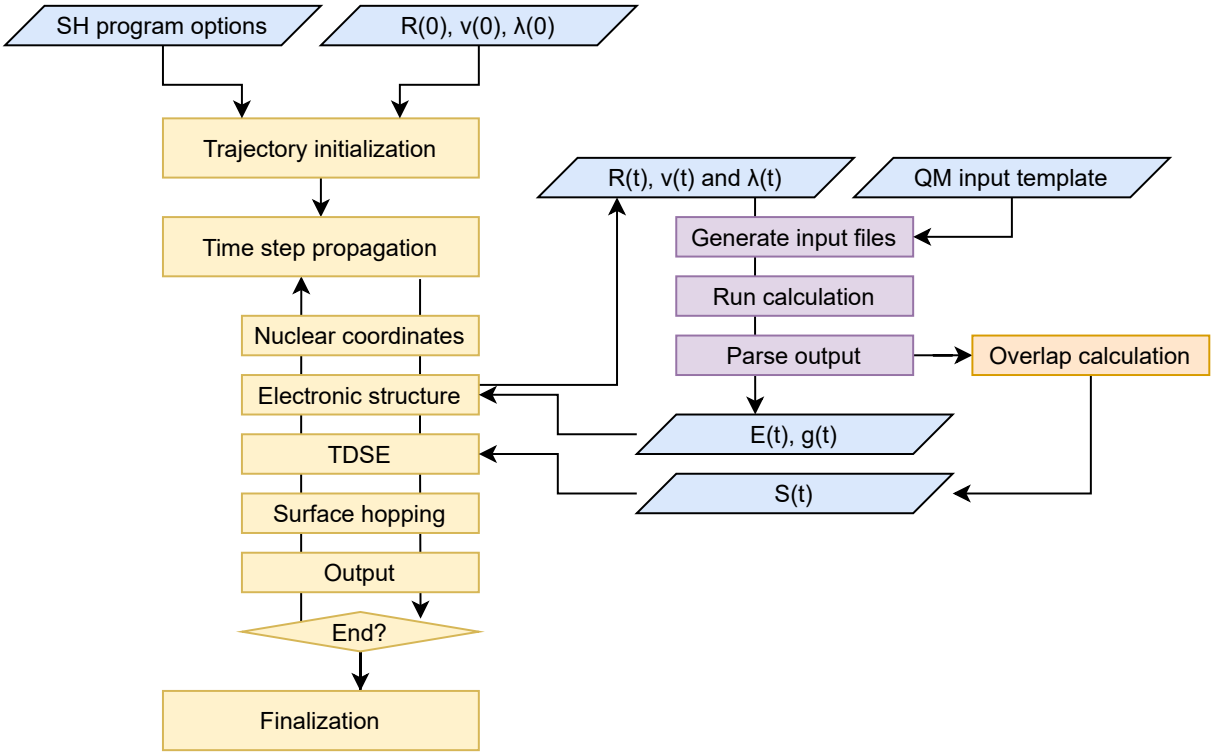


Figure 3.3: Propagation of a trajectory using the FSSH algorithm. The main FSSH code (yellow) handles the propagation of the nuclear coordinates and the total electronic wave function. When energies and gradients are required a script (purple) is called to run a chosen electronic structure code and parse the output. The overlap calculation is handled by a standalone program also interfaced to the electronic structure code.

these calculations, the velocity Verlet algorithm[130] is always used. For model systems where additional evaluations of the PES are not computationally expensive one could use higher order propagators. It is possible to set constraints during this propagation[258] or to project rotation and translation from the motion at each step to keep the molecule in the Eckart frame.[259] The next step is to calculate the energies and gradient at the new geometry which is handled by an outside script which communicates with the main code through simple files containing the necessary inputs/outputs.

For FSSH algorithms this is followed by calculating the overlap matrix between the current and previous step, propagating the TDSE and determining the hopping probabilities. In the present work we have developed new algorithms for calculating the overlap matrix which are described in detail in section 4.2. To propagate the TDSE and obtain accurate hopping probabilities the many details explained in section 2.2.3 need to be taken into account. Most of the methods mentioned in that section have been implemented as options in the SH code. In the

studies reported here, the phases of the overlap matrix were corrected by transforming the matrix so that the largest elements are on the diagonal, multiplying the columns so these elements are all positive and transforming back to the original basis. When propagating in the adiabatic basis (A-FSSH) the nuclear time step was split into 10000 substeps[131] during which the energies were linearly interpolated and the overlap matrix was interpolated based on the NPI scheme.[20] Hops were allowed at each substep. Alternately, propagation was performed based on the locally diabatic formalism (LD-FSSH).[18] The energy based decoherence scheme of Granucci and Persico was employed with the default decoherence constant of 0.1 E_h.[114] Energy conservation after hops was ensured by uniformly rescaling the nuclear momenta. For LZSH these steps are skipped. Instead, the hopping probability is evaluated only if a gap minimum in the previous step is detected. To ensure the numerical accuracy of the evaluated probability, the trajectory is returned to a step before the gap minimum and the minimum and second derivative of the energy are evaluated more accurately using a numerical procedure described in section 3.2.1.

Since the overlap code has been fully separated from the dynamics code, the trajectories never need information on the basis functions, molecular orbitals or wave functions which makes them completely indifferent to the chosen level of theory. The only information required to be passed from the dynamics to the ES program is the geometry and current state, while the only information returned are the state energies and gradient of the current state (and nonadiabatic coupling vectors if they are used). This makes writing the interface script trivial for any electronic structure code. For LZSH this is all that is needed, for FSSH an interface for the overlap code also has to be written. This part is more difficult since it requires working with the basis set, molecular orbitals and wave function coefficients. However, maximum attention has been devoted when writing the overlap code to simplify this procedure. Details concerning the interface between electronic structure codes and the wave function overlap calculation are covered in section 3.2.2.

4) After a set of trajectories is propagated, we are left with a huge amount of data from which physical insight needs to be extracted. The most basic information are the relaxation pathway for individual trajectories and the adiabatic populations of the ensemble. These can be used to extract (through fitting) time constants of the studied reactions which can be directly compared to information available through time resolved experiments. Diabatic populations are not available from on-the-fly SH calculations, but we can estimate them by using wave function

overlaps to compare the electronic wave functions at different geometries in the ensemble. This procedure is explained in section 4.3. Lastly, observables such as time-resolved spectra can be calculated using the nuclear ensemble approach for each time step.

3.2.1 Landau-Zener surface hopping

The adiabatic form of the Landau Zener equation for the transition probability given in section 2.2.4,

$$P_{ij} = \exp\left(-\frac{\pi}{2\hbar}\sqrt{\frac{g^3(t_c)}{\ddot{g}(t_c)}}\right), \tag{3.6}$$

does not contain the TDCs which allows us to bypass the overlap calculation. However, overlaps are used in all algorithms for handling trivial crossings so in our implementation of LZSH we take care to not "reinvent" this problem by inaccurately treating such intersections where we expect the gap minimum to become very small during the time step.

While the FSSH algorithm gives a "local" transition probability (actually transition rate) at each step of the dynamics, equation 3.6 gives the "global" transition probability for the intersection region. This means that the probability should be evaluated only when a minimum of the energy gap between two states $g(t) = E_i(t) - E_j(t)$ is found along a trajectory. After evaluating the probability at this minimum, the trajectory either resumes its evolution on the previously populated state or is transferred to the other state and evolution is resumed from t_c.

The second derivative of the energy gap is calculated when the gap minimum is reached ($t = t_c$) using a three point finite difference

$$\ddot{g}(t_c) = \frac{g(t-\delta) + g(t+\delta) - 2g(t)}{\delta^2} = \frac{2d}{\delta^2}, \tag{3.7}$$

where δ is the time-step and we defined $d \equiv ((g(t-\delta) - g(t)) + (g(t+\delta) - g(t)))/2$ as the average change in the gap between the time steps around the minimum. Inserting into the LZ probability formula, we obtain

$$P_{ij} = \exp\left(-\frac{\pi}{2}\sqrt{\frac{g^3(t)\delta^2}{2d}}\right). \tag{3.8}$$

In this form, we see that we need only the state energies of three consecutive time steps to

evaluate the LZ probability. This means that, at each step, LZ surface hopping requires no information in addition to that required for classical molecular dynamics, the energies of the electronic states and the gradient on the current state. This makes LZ surface hopping incredibly easy to interface with any electronic structure program even when nonadiabatic couplings are not available and without the need to implement the calculation of wave function overlaps which is typically the most labour intensive part of implementing Tully's fewest-switches surface hopping algorithm.

However, there are two possible sources of error in the formula above. The first is the assumption that $g(t_c) = g(t)$. In fact we only know that $t - \delta < t_c < t + \delta$. This is important because $g(t_c)$ appears in the exponent so small changes in the value of the minimum can have a large impact on the calculated probability. The second source of error is the precision σ with which state energies are calculated as set during the electronic structure calculation. This is more important when the time step is small and the second derivative is calculated by taking a difference between numbers of similar magnitude.

The former source of error is basically a minimization problem where we are finding the local minimum of $g(t)$. This is most simply done by reducing the time steps when a gap minimum is found (for example using the bisection method). However, since each additional evaluation of the electronic energies is costly, it is useful to only do this when necessary. Since d is the change in the gap around the minimum, we can assume that $g(t) - d \leq g(t_c) \leq g(t)$. Inserting these upper and lower limits into equation 3.8, we obtain an error interval for the hopping probability $P_{min} \leq P \leq P_{max}$. This allows us to set a precision threshold directly on the probability and only reduce the time step when needed. The same procedure can be done for the error due to σ by assuming all energies entering equation 3.8 are only correct up to $\pm\sigma$.

Due to the exponential form of equation 3.6, there is a wide range of possible intersections characterized by $(g(t_c), \ddot{g}(t_c))$ combinations which give hopping probabilities close to zero or one and are not sensitive to small deviations in either value. On the other hand, when the term in the exponent is close to unity (i.e. the hopping probability is close to 50%), small errors in the calculated values can have a large impact on the probability so additional calculations will be required. By always calculating error estimates, we hope to minimize the number of additional potential energy evaluations without sacrificing accuracy.

3.2.2 Molecular orbital overlap integrals

We mentioned above how interfacing electronic structure codes with a wave function overlap calculation is significantly more complex than interfacing with the molecular dynamics portion of the code. In this section we will explain the reason for this and describe how we attempt to minimize this problem through the implementation of the code. The two inputs needed for any wave function overlap calculation are the wave function coefficients and the matrix of overlaps between molecular orbitals used to generate the two sets of wave functions. While the wave function coefficients can usually be read directly from a file, molecular orbital overlaps between different geometries are rarely available from electronic structure programs and have to be calculated from the molecular orbital coefficients and information on the basis functions.

The most common type of basis functions used in quantum chemical programs are Gaussian type orbitals (GTOs) which allow for very efficient evaluations of integrals needed for electronic structure calculations. Two varieties of these basis functions are used,[260] Cartesian GTOs are defined as

$$g_{l_x l_y l_z}(\mathbf{R};\mathbf{R_0},\xi) = N_g x^{l_x} y^{l_y} z^{l_z} e^{\xi(x^2+y^2+z^2)}, \tag{3.9}$$

where $q \equiv R_q - R_{0q}$ for $q \in \{x,y,z\}$, while spherical GTOs are defined as

$$\tilde{g}_{nlm}(\mathbf{R};\mathbf{R_0},\xi) = N_{\tilde{g}} Y_m^l(\theta,\phi) r^n e^{-\xi r^2}, \tag{3.10}$$

where r, θ and ϕ are spherical coordinates centered at $\mathbf{R_0}$ and $Y_m^l(\theta,\phi)$ are the spherical harmonic functions. The two types of GTOs are normalized as

$$N_g = \left(\frac{2^{\frac{3}{2}+l} \xi^{\frac{3}{2}+l}}{\Gamma(\frac{1}{2}+l_x)\Gamma(\frac{1}{2}+l_y)\Gamma(\frac{1}{2}+l_z)} \right)^{\frac{1}{2}}, \tag{3.11}$$

and

$$N_{\tilde{g}} = \left(\frac{2^{2n+3}(n+1)!\xi^{(n+\frac{3}{2})}}{(2n+2)!\pi^{\frac{1}{2}}} \right)^{\frac{1}{2}}, \tag{3.12}$$

respectively. Typically, only spherical Gaussians with $l = n$ are used and are defined in groups (subshells) with the same l and all possible m values. For a given angular momentum $l = l_x + l_y + l_z$, the number of Cartesian functions is greater ($(l+1)(l+2)/2$) than the number of

spherical functions $(2l+1)$. The spherical functions with $l = n$ can be expanded as a linear combination of the appropriate Cartesian functions[260]

$$\tilde{g}_{lm}(\mathbf{R};\mathbf{R_0},\xi) = \sum_{l_x,l_y,l_z} T^{lm}_{l_x l_y l_z} g_{l_x l_y l_z}(\mathbf{R};\mathbf{R_0},\xi)\,. \tag{3.13}$$

General analytical expressions for the transformation coefficients $T^{lm}_{l_x l_y l_z}$ have been derived and, since they depend only on the quantum numbers, can easily be tabulated.[260]

While these basis functions are efficient, they do not have the correct behavior for $r = 0$ (where atomic orbitals should have a cusp) or for large r (where they should fall of as e^{-r}). Because of this, the actual basis functions used in most electronic structure programs are contracted Gaussian type orbital (cGTO)

$$\begin{aligned} G_{l_x l_y l_z}(\mathbf{R};\mathbf{R_0},\xi,\mathbf{b}) &= N_G \sum_i b_i g_{l_x l_y l_z}(\mathbf{R};\mathbf{R_0},\xi_i) \\ &= N_G x^{l_x} y^{l_y} z^{l_z} \sum_i N_{g_i} b_i e^{\xi_i (x^2+y^2+z^2)}\,, \end{aligned} \tag{3.14}$$

which are linear combinations of GTOs with the same angular part built to better reproduce a single Slater type orbital (STO). The coefficients b_i are known as contraction coefficients. The full cGTO can be normalized as

$$N_G = \Gamma_{xyz} \sum_i \sum_j b_i b_j N_{g_i} N_{g_j} (\xi_i + \xi_j)^{-\frac{3}{2}-l}\,, \tag{3.15}$$

where we have defined $\Gamma_{xyz} \equiv \Gamma(\frac{1}{2}+l_x)\Gamma(\frac{1}{2}+l_y)\Gamma(\frac{1}{2}+l_z)$.

Once the basis set is defined, it is straightforward to calculate matrix elements in the space of the basis functions. The elements of the atomic orbital overlap matrix $\mathbf{S}$ are calculated as

$$S_{ab} = \langle G_a | G_b \rangle = N_{G_a} N_{G_b} \sum_i \sum_j b_i b_j \left\langle g_{i,a} \middle| g_{j,b} \right\rangle . \tag{3.16}$$

The integrals over Cartesian functions $\left\langle g_{i,a} \middle| g_{j,b} \right\rangle$ can be calculated analytically.[261] In the basis of spherical functions, operators can be obtained either by numerically computing the integrals

$\langle \tilde{g}_{i,a} | \tilde{g}_{j,b} \rangle$ or by transforming the Cartesian matrices

$$\tilde{\mathbf{S}} = \mathbf{T}_{\text{cs}}^{\mathsf{T}} \mathbf{S}^{\text{C}} \mathbf{T}_{\text{cs}} \,, \tag{3.17}$$

where $\mathbf{T}_{\text{cs}}$ is a block diagonal matrix containing the transformation coefficients between Cartesian and spherical functions $T^{lm}_{l_x l_y l_z}$ for each subshell. Operators in the basis of molecular orbitals

$$\phi_i(\mathbf{R}) = \sum_a C_{i,a} G_a \quad \text{or} \quad \phi_i(\mathbf{R}) = \sum_a \tilde{C}_{i,a} \tilde{G}_a \tag{3.18}$$

are easily obtained by matrix multiplication from the corresponding atomic orbital operators

$$\mathbf{S}_{\mathsf{MO}} = \mathbf{C}^{\mathsf{T}} \mathbf{S} \mathbf{C} = \tilde{\mathbf{C}}^{\mathsf{T}} \tilde{\mathbf{S}} \tilde{\mathbf{C}} . \tag{3.19}$$

From the practical side, to calculate the required values, one needs to know the molecular orbital coefficients and the basis functions. To fully define each basis function, we need to know its origin $\mathbf{R_0}$ (almost always the coordinates of an atom), exponents ξ, contraction coefficients $\mathbf{b}$ and angular momentum quantum numbers (l_x, l_y, l_z) or (l, m). The former quantities are the same inside a subshell regardless of the choice between spherical and Cartesian basis functions, so the basis set is most usually given as a list of $\{l, \xi, \mathbf{b}\}$.

While almost all quantum chemistry packages dealing with isolated molecules use contracted Gaussian basis sets, how these are expanded into actual basis functions is often different from program to program. Even after defining the order of the basis functions within subshells, some ambiguities (such as a different sign for some of the spherical harmonics, or a different normalization procedure) remain due to possible non-standard definitions used in certain packages. Because of this, we add an additional step in the calculation of the MO overlap matrix $\mathbf{S}_{\mathsf{MO}}$. First, the AO overlap matrix is calculated according to equation 3.16 or 3.17 based only on the basis set information and assuming standard ordering of basis functions within subshells. Before using equation 3.19, the AO overlap matrix (either $\mathbf{S}$ or $\tilde{\mathbf{S}}$) is further transformed as

$$\mathbf{S}' = \mathbf{T}_{\text{prog}}^{\mathsf{T}} \mathbf{S} \mathbf{T}_{\text{prog}} \,, \tag{3.20}$$

where $\mathbf{T}_{\text{prog}}$ is a permutation and scaling matrix depending on the electronic structure program

Table 3.1: Order of spherical and Cartesian basis functions expected for molecular orbitals given in the Molden format.[264]

l	m order / $l_x l_y l_z$ order
2 (d)	0, 1, -1, 2, -2 / 200, 020, 002, 110, 101, 011
3 (f)	0, 1, -1, 2, -2, 3, -3 / 300, 030, 003, 120, 210, 201, 102, 012, 021, 111
4 (g)	0, 1, -1, 2, -2, 3, -3, 4, -4 / 400, 040, 004, 310, 301, 130, 031, 103, 013, 220, 202, 022, 211, 121, 112

Table 3.2: Changes required to standardize Molden files generated by different quantum chemistry packages. The fixes were tested to give correct values for MO overlap matrix elements for the given versions of the programs and with basis sets containing functions with angular momenta up to l_{max}.

Package	Version	l_{max}	Type	Change
Bagel	11. 5. 2020	h	Sphe	None
Molcas	19.11	g	Sphe	None
Molpro	2018.1	f	Cart	None
CFour	2.1	f	Cart	Scale all coefficients by $\sqrt{(2l_x-1)!!(2l_y-1)!!(2l_z-1)!!}$
Turbomole	7.0.1	g	Cart	Scale all coefficients by $\sqrt{2l-1}$
Orca	4.2.1	h	Sphe	Printed contraction coefficients are already multiplied by N_g.[a] Coefficients corresponding to AOs with $2 < \|m\| < 5$ scaled by -1.

[a] Fixed by dividing contraction coefficients by N_g when reading them instead of by $\mathbf{T}_{\text{prog}}$.

used. In this way, creating an interface to a different electronic structure program requires only knowing the ordering of functions within subshells and possible scaling factors. These are readily found without any outside information by comparing the standard AO overlap matrix $\mathbf{S}$ for the basis set and a known final matrix $\mathbf{S}'$. The latter can be obtained for an orthonormal set of MOs where only the MO coefficients are needed to calculate the AO overlap matrix

$$\mathbf{S}_{\mathsf{MO}} = \mathbf{I} = \mathbf{C}^{\mathsf{T}}\mathbf{S}'\mathbf{C} \quad \Rightarrow \quad \mathbf{S}' = \mathbf{C}^{\mathsf{T},-1}\mathbf{C}^{-1}. \tag{3.21}$$

The closest to a standard format for writing atomic basis sets and molecular orbitals is the Molden format[262,263] since most packages can print MOs in this format. It supports both spherical and Cartesian functions and has a published ordering[264] for expanding the basis functions within each subshell up to g type orbitals (Table 3.1). As noted above, this by itself is not enough and "fixes" are often needed to handle Molden files generated by different quantum chemistry packages. Examples of such changes are given in Table 3.2.

A further example is given for MOs written in the native format for Turbomole. Turbomole uses spherical basis functions and it uses the same ordering of basis functions as Molden, but one f-type basis function ($\tilde{g}_{3,-3}$) and two g-type basis functions ($\tilde{g}_{3,-2}$ and $\tilde{g}_{4,-3}$) are defined with a different sign from those in equation 3.10. Listing 1 gives the two functions needed to automatically generate the transformation matrix from this format to any other format for which functions like this are also defined.

```
def turbo_m_order(l):
    ''' Order of m for subshell with given l.'''
    if l == 1:
        return [1, -1, 0]
    m = [0]
    for i in range(1, l+1):
        m.append(i)
        m.append(-i)
    return m

def turbo_lm_scale(l, m):
    ''' Scaling factor for bf with quantum numbers l, m.'''
    if l == 3 and m == -3:
        return -1
    if l == 4 and (m == -3 or m == -2):
        return -1
    return 1
```

Listing 1: Functions needed to generate $\mathbf{T}_{\text{prog}}$ for transforming MO coefficients to and from the Turbomole native format.

3.3 Photoionization observables

For the photochemical processes studied in later sections, some the most informative experiments might be those based on time-resolved photoelectron spectroscopy.[173,186,202] In these experiments, the pump pulse excites the system to start the reaction and the probe pulse ionizes the system. Then the kinetic energy (and angular distribution) of the photoelectron can be measured to obtain information on the state of the system from which the electron was ejected.

At a particular molecular geometry, the likelihood of photoionization is connected to the

dipole transition moment between the initial ($\left|\Psi_I^N\right\rangle$) and final ($\left|\Phi_{Fk}^N\right\rangle$) wave functions involved in the transition. The partial differential cross-section in the molecular frame is

$$\frac{d\sigma_{IFk}}{d\mathbf{k}} = 4\pi^2\alpha\omega\left|\left\langle\Psi_I^N\middle|\mu\middle|\Phi_{Fk}^N\right\rangle\right|^2, \tag{3.22}$$

where $\mathbf{k}$ is the momentum of the photoelectron, α is the fine structure constant, ω is the photon energy and $\left\langle\Psi_I^N\middle|\mu\middle|\Phi_{Fk}^N\right\rangle = \mu_{IFk}$ is the dipole transition moment. This value is most commonly evaluated in the single channel approximation where the final state is expressed as an antisymmetrized product of the wave function of the ejected photoelectron φ_k with kinetic energy $E_k = \frac{\hbar^2k^2}{2m_e}$ and of the remaining bound ionized system Ψ_F^{N-1}.

For single photon photoionization with linearly polarized light, the differential cross-section will have the following form

$$\frac{d\sigma_{IFk}}{d\mathbf{k}} = \frac{\sigma_{IFk}}{4\pi}[1+\beta_{IFk}P_2(\cos\theta)], \tag{3.23}$$

where P_2 is the second-order Legendre polynomial, θ is the angle between the electron emission direction $\mathbf{k}$ and the polarization vector of the incident light, and β is an asymmetry parameter (ranging between -1 and 2) fully describes the angular distribution of the photoelectron.

The calculation of the dipole transition moment can be simplified to a single electron picture by integrating out the $N-1$ electrons of the bound system

$$\mu_{IFk} = \left\langle\phi_{IF}^D\middle|\mu\middle|\phi_k\right\rangle, \tag{3.24}$$

where $\left|\phi_{IF}^D\right\rangle$ is the Dyson orbital

$$\left|\phi_{IF}^D\right\rangle = \sqrt{N}\int\Psi_F^{N-1}(x_1,...,x_{N-1})^*\Psi_I^N(x_1,x_2,...,x_N)dx_1...dx_{N-1} \tag{3.25}$$

and an additional so-called conjugate contribution is usually neglected.

Humeniuk *et al.* derived expressions for Dyson orbitals within linear response TDDFT.[265] The wave function is represented by the TDDFT linear response eigenvectors as a linear combination of single-excitations from the Kohn-Sham reference Slater determinant. After expanding the SDs of the neutral system into minors along the last row to give a sum over SDs of dimension

$N-1$, the Dyson orbital can be written as

$$\left|\phi_{IF}^{D}\right\rangle = \sum_{a,i}\sum_{b,j} c_{aF}^{*N-1} c_{bI}^{N} \sum_{k} \left\langle \Phi_{aF}^{iN-1} \middle| a_k \middle| \Phi_{bI}^{jN} \right\rangle |\phi_k\rangle \,, \tag{3.26}$$

where the first two sums go over all combinations of occupied orbitals a, b and virtual orbitals i, j of the final (cation) and initial (neutral) state, respectively, and operator a_k removes an electron from orbital k of the neutral wave function.

The norm of a Dyson orbital can itself be used as a qualitative approximation of the photoionization cross-section. For a more accurate treatment, or for obtaining the photoelectron angular distribution, we need to solve eq. 3.24. This will be done at the static-exchange DFT level[266] with the DFT ground state density used to build the Hamiltonian matrix. Both the Hamiltonian matrix and the previously computed Dyson orbitals are expanded in a new basis set consisting of an expansion of radial B-spline functions multipled by real spherical harmonics. This basis set consists of a set of basis functions positioned at the center of mass of the molecule and expanding to large distances (needed to accurately treat the continuum states) and basis functions positioned at the nuclei (needed to accurately treat the bound states). The continuum states are obtained through the Galerkin approach.[267] and are further normalized to K-matrix asymptotic boundary conditions.

For the photoionization observable calculations performed in the study of pyrrole in section 4.4, the ground state density is calculated at the DFT/LB94/DZP level[268] using the ADF program.[269] This functional was previously shown to give accurate results for photoionization observables.[270] For the B-spline basis set, a radial grid was placed at the center of mass of the molecule with nodes extending up to 25 a_0 with a step size of 0.2 a_0 and combined with spherical harmonics with angular momentum up to 15. Additional expansions were placed at the nuclei, associated with spheres of radius 1.2, 1.0 and 0.7 a_0 and angular momenta up to 2, 2 and 1 for N, C and H atoms, respectively.

§4. RESULTS AND DISCUSSION

4.1 Evaluation of SH algorithms

As noted in section 2.2.4, significant interest in recent years has been devoted to developing a surface hopping algorithm where the energies and gradients of the electronic states would be the only required input during the propagation of the trajectories.[48–52,136,137,271]

Most of these studies have focused on model systems when comparing the LZSH type algorithms to other methods. An exception to this is the work of Zhu and coworkers[52] who used the photoisomerization reaction in azobenzene to compare surface hopping procedures based on the Zhu-Nakamura formula and on the fewest-switching algorithm. They found very good agreement between the two methods both for the full population and for individual trajectories and hopping events. However, this reaction usually involves only a single hopping event for each trajectory at an S_0/S_1 conical intersection well separated from all other states so agreement here cannot be generalized for more complex reactions. On the other hand, Xie and Domcke[137] compared both FSSH and LZSH algorithms with exact quantum dynamics simulations for a three-state two mode model of photoinduced hydrogen detachment in phenol. This reaction involves a transfer of population from the $\pi\pi^*$ state into a dissociative $\pi\sigma^*$ state which leads to an intersection with the ground state along the hydrogen detachment coordinate. In this case, both SH algorithms were in agreement, but both failed to quantitatively reproduce the exact results. However, it was difficult to judge how much of this failure was due to the approximate treatment of nonadiabatic effects and how much was due to system-specific nuclear quantum effects since the reaction involves a significant degree of H-atom tunneling.

In the present work, we will look at two prototypical problems in photochemistry, the first is internal conversion in pyrazine after excitation to the $B_{2u}(\pi\pi^*)$ and the second is hydrogen detachment in pyrrole after excitation to the $B_2(\pi\pi^*)$ state. The former involves a fast transfer of population from the $B_{2u}(\pi\pi^*)$ state to the lower lying $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ states. In pyrrole, the $B_2(\pi\pi^*)$ is the adiabatic S_5 state. Hydrogen detachment proceeds mostly from the $A_2(\pi\sigma^*)$ state which is the adiabatic S_1 state so, at least, the $\pi\sigma^*$ and two Rydberg states between these two states need to be included in the consideration. These reactions will be covered in significantly more detailed in sections 4.5 and 4.4, respectively. Here, we will just

comment on the reactions in terms of their treatment using FSSH and LZSH. Nonadiabatic dynamics were performed using SH algorithms in the adiabatic basis using the NPI formalism to ensure numerical stability of the calculated TDCs (A-FSSH), in the locally diabatic basis (LD-FSSH), or based on the Landau-Zener formula (LZSH).

4.1.1 Pyrazine

The dynamics of pyrazine were considered at three distinct levels in order to obtain the best possible assessment of the methods for the problem and also to consider the effect of the number of dimensions included in the model on the performance of the methods. The first two are the two and three state linear vibronic coupling model Hamiltonians develepoed by Sala *et al.*[53] described in section 3.1 and the third is the ADC(2)/aug-cc-pVDZ full dimensional potential.

Figure 4.1 shows the population of the $B_{2u}(\pi\pi^*)$ state in the diabatic basis and of S_2 in the adiabatic basis for dynamics performed using the two state model. The population decays quickly within the first 45 fs to approximately 0.15 and then exhibits recurrences around 80 fs and 140 fs after the start of the dynamics. Overall, the exact results (black curve) are accurately reproduced by both LZSH and by LD-FSSH with only minor differences visible. The deviations between the FSSH and LZSH results are very minor. The population recurrences are also mostly reproduced by both SH methods.

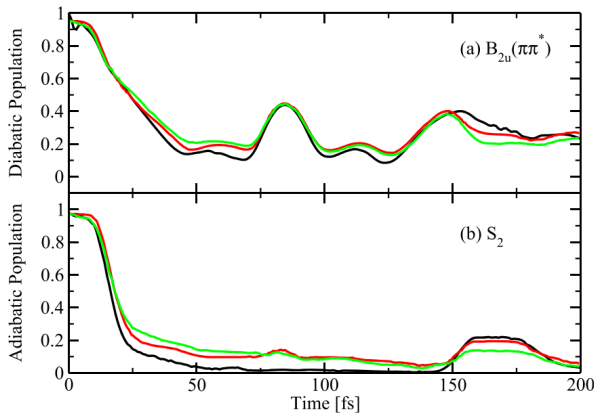


Figure 4.1: Diabatic population of the $B_{2u}(\pi\pi^*)$ state (upper panel) and adiabatic population of the S_2 state for the two state model of pyrazine during MCTDH (black), LD-FSSH (green) and LZSH (red) dynamics. Figure originally published in Ref. (*34*).

Figure 4.2 shows the populations of the diabatic states for dynamics performed using the three state model. The decay to the $B_{2u}(\pi\pi^*)$ state is just as fast as in the two state model, but

unlike in that model decays almost completely to zero in the same time frame. There are also almost no recurrences of the population of the $B_{2u}(\pi\pi^*)$ state, instead the population oscillates between the $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ states. As with the two state model, the agreement between both of the mixed classical-quantum methods and MCTDH is very good, with only a slight ($\approx$ 0.05) underestimation of the population $A_u(n\pi^*)$. A similar level of agreement between LZSH and LD-FSSH can be seen when looking at the populations of the adiabatic states in Figure 4.3. The decay of the S_3 population is faster than that of $B_{2u}(\pi\pi^*)$. After approximately 60 fs, the population is split between the S_1 and S_2 states at approximately 80% to 20% and does not significantly change until the end of the simulation.

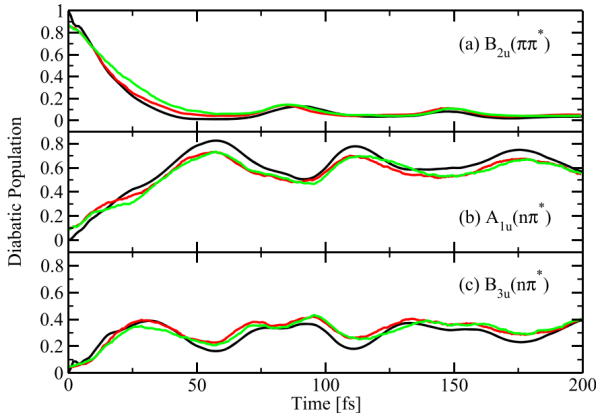


Figure 4.2: Diabatic populations of the $B_{2u}(\pi\pi^*)$ (a), $A_u(n\pi^*)$ (b) and $B_{3u}(n\pi^*)$ (c) states for the three state model of pyrazine during MCTDH (black), LD-FSSH (green) and LZSH (red) dynamics. Figure originally published in Ref. (*34*).

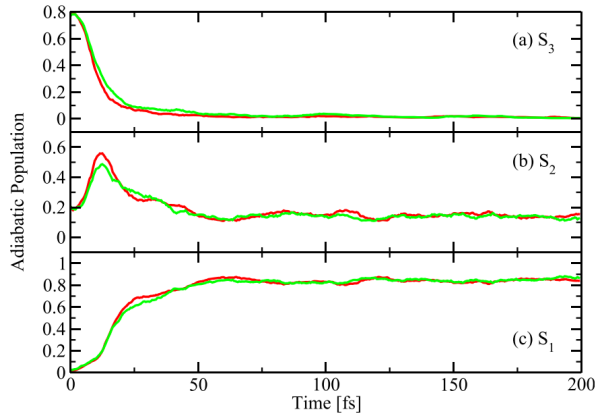


Figure 4.3: Adiabatic populations of the S_3 (a), S_2 (b) and S_1 (c) states for the three state model of pyrazine during MCTDH (black), LD-FSSH (green) and LZSH (red) dynamics. Figure originally published in Ref. (*34*).

Finally, *ab initio* on-the-fly simulations were performed at the ADC(2)/aug-cc-pVDZ level. In these simulations, initial conditions were sampled from the Wigner distribution and an instantaneous excitation to the bright $B_{2u}(\pi\pi^*)$ state was assumed. Five states (ground state and lowest four excited states) were included in the calculation. However, the 200 fs simulations were too short for any decay to the ground state and the S_4 state was also never significantly (< 0.03) populated during the dynamics.

Figure 4.4 shows the populations of the adiabatic states for dynamics performed using the full system at the ADC(2) level. There are no qualitative differences between the adiabatic populations in simulations using the full-dimensional simulations and the three state model. As before, we see an excellent agreement between the LD-FSSH and LZSH methods. On the other hand, the A-FSSH underestimates the decay rate of the intermediate S_2 population. At longer time scales, the populations given by the three methods converge to the same value.

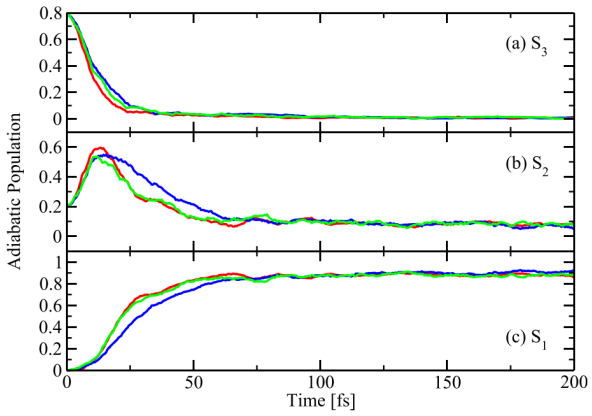


Figure 4.4: Adiabatic populations of the S_3 (a), S_2 (b) and S_1 (c) states of pyrazine during A-FSSH (blue), LD-FSSH (green) and LZSH (red) dynamics calculated at the ADC(2)/aug-cc-pVDZ level of theory. Figure originally published in Ref. (*34*).

The dynamics on the fully dimensional system were performed in the adiabatic basis and no diabatization procedure to generate fully coupled diabatic states for a system of this size is available. Instead, the diabatic populations are approximated using two simple procedures based on the properties of the states. In the first attempt, the states are simply arranged according to their oscillator strengths. At the ground state minimum geometry, the $B_{2u}(\pi\pi^*)$ state is relatively bright ($\sim 5 \times 10^{-2}$ oscillator strength), the $B_{3u}(n\pi^*)$ state is dark ($\sim 5 \times 10^{-3}$) while the transition to $A_u(n\pi^*)$ is forbidden (0). These large differences allow separating the states based on the oscillator strengths, but are not reliable at different geometries. The diabatic popula-

tions obtained from oscillator strengths are shown for LZSH and FSSH dynamics in Figure 4.5. Whereas the adiabatic populations were smooth, the diabatic populations of the $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ states exhibit oscillations with the two methods giving almost completely identical results. However, these oscillations have a significantly shorter period ($\sim$ 30 fs) compared to those of the model system ($\sim$ 60 fs).

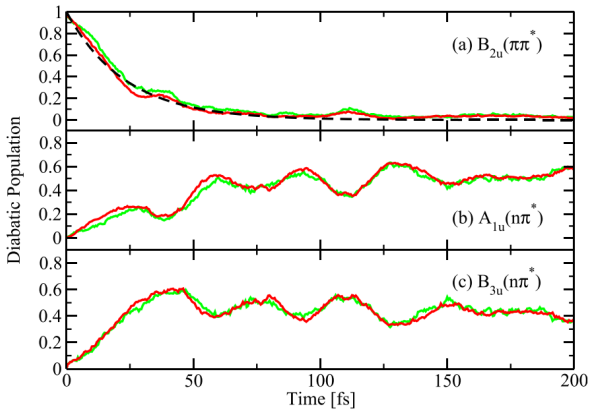


Figure 4.5: Approximate diabatic populations of the $B_{2u}(\pi\pi^*)$ (a), $A_u(n\pi^*)$ (b) and $B_{3u}(n\pi^*)$ (c) states of pyrazine during A-FSSH (blue), LD-FSSH (green) and LZSH (red) dynamics calculated at the ADC(2)/aug-cc-pVDZ level of theory. Figure originally published in Ref. (*34*).

One area where the two methods differ is the actual points where the hops occur. Figure 4.6 shows the distribution of gaps at which the hops occur for each of the methods tested. The total number of hops is similar in each simulation, but we see that significantly more hops occur at low gaps ($\sim$ 0.1 eV) in the LZSH simulation compared to either FSSH simulation. This observation is expected since hopping probabilities in LZSH are evaluated only in the gap minima. On the other hand, hops in FSSH can occur at any point regardless of the gap between states so hops at larger gap values are possible.

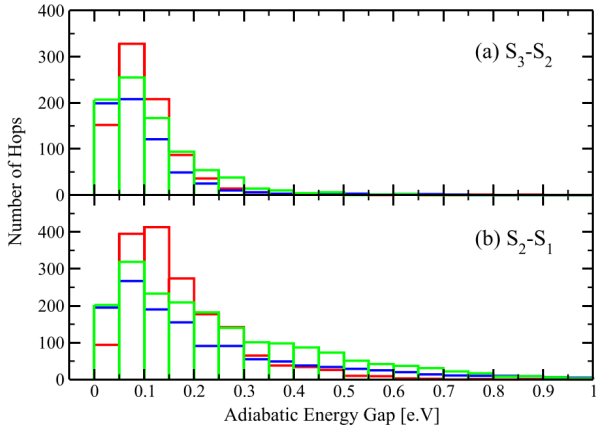


Figure 4.6: Distribution of the instantaneous energy gaps for successful hops between (a) S_3 and S_2 and (b) between S_2 and S_1 in the full-dimensional ADC(2) simulations. The red, green, and blue distributions correspond to LZSH, LD-FSSH and A-FSSH methods, respectively. Figure originally published in Ref. (*34*).

4.1.2 Pyrrole

For pyrrole, sets of trajectories were propagated using each of the SH algorithms (A-FSSH, LD-FSSH and LZSH) using three different electronic structure methods, ADC(2)/aug-cc-pVDZ, SCS-ADC(2)/aug-cc-pVDZ and B3LYP/def2-TZVPD. For each method a total of 100 trajectories were sampled (weighted by oscillator strengths) from the ground state Wigner distribution. In all of the selected trajectories, the bright $B_2(\pi\pi^*)$ state from which the dynamics were started was the S_5 state.

Table 4.1 shows the time scales for relaxation to the ground state for each dynamics simulation. Unlike the situation in pyrazine, the quantitative picture for the reaction in pyrrole is significantly different based on both the electronic structure and the dynamics level of theory. The reaction is significantly slower at the ADC(2) level than TDDFT and significantly faster with LZSH than LD-FSSH. Differences between ADC(2) and SCS-ADC(2) are negligible. In terms of reaction mechanisms, most trajectories ended through dissociation of the N-H bond, and a smaller fraction (similar across all methods) through the ring-puckering mechanism. Some of the N-H dissociation trajectories remain in the S2 state when the S0/S1 CI is reached, but at these points the S2 state is always within 1 eV of the CI, suggesting the S2/S0 CI would be reached quickly from there.

Populations of all adiabatic states are shown in Figure 4.7. Qualitatively the results LD-FSSH/LZSH results are very similar. The main visible difference is in the populations of states

Table 4.1: Time constants (in fs) for relaxation to the ground state for different combinations of surface hopping algorithm and electronic structure method.

	LD-FSSH	LZSH
ADC(2)/aug-cc-pVDZ	185	128
SCS-ADC(2)/aug-cc-pVDZ	189	122
B3LYP/def2-TZVPD	95	52

S_4 (purple), S_3 (red) and S_2 (green) in the early steps of the B3LYP dynamics which are populated for a shorter amount of time in the LZSH dynamics. An initial comparison of the populated diabatic states at 10 and 20 fs also did not show qualitative differences between LZSH and LD-FSSH dynamics.

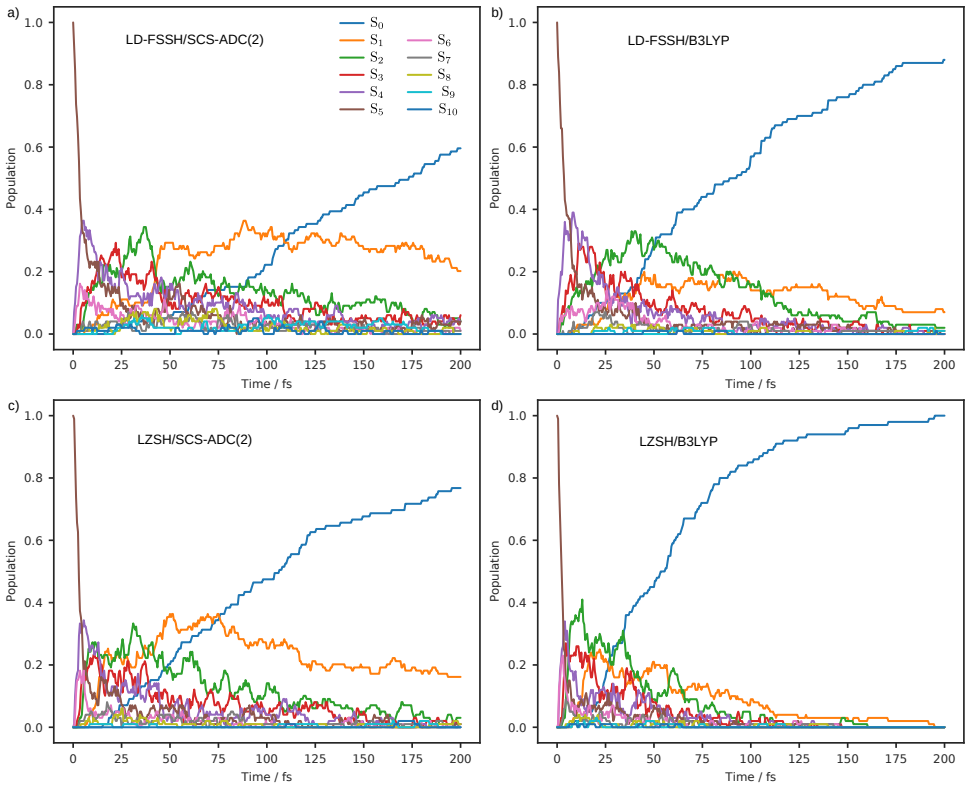


Figure 4.7: Populations of the ground state and ten lowest excited states during LD-FSSH and LZSH dynamics at the SCS-ADC(2)/aug-cc-pVDZ (a, c) and B3LYP/def2-TZVPD (b, d) levels of theory.

The distribution of hops with respect to the gap minimum is quite similar to the one shown for pyrazine in Figure 4.6 with the LZSH distribution more highly peaked at low gap values. Also like in pyrazine, the overall number of hops is similar for both algorithms algorithms. However, in pyrrole the two LZSH simulations end faster which means that the same number of hops occurs in less time. Indeed, we see that the key difference between the two methods

in pyrrole is in the very early stages of the simulation (first 5 fs) where significantly more hops occur in the LZSH simulations. This causes the LZ trajectories to more quickly pass through the initial region where the states S2-S6 are all within a very narrow energy window.

4.1.3 Numerical stability of LZSH

Continuing with the pyrrole B3LYP dynamics as our test case, we will now take a look at the numerical stability of the LZSH algorithm with respect to the possible errors caused by discreet time steps and finite precision of the electronic structure calculations as discussed in section 3.2.1. The LZSH results shown above are for trajectories where the time step was bisected four times before evaluating the hopping probability at each gap minimum. The relaxation time scale is almost identical (within 2 fs) if only a single bisection is performed, but slight differences could be seen for the populations of the S_2 and S_1 states. Due to the frequency of gap minima encountered during the trajectory, the average number of evaluations of the PES is $\sim 20\%$ higher when four bisections are performed at each hop than when a single bisection is performed (244 and 203, respectively).

Figure 4.8 shows the $(\log(g), \log(d))$ distributions from the pyrrole B3LYP trajectories with different time steps for the evaluation of hopping probabilities. Here, as defined in section 3.2.1, g is the energy gap and $d = ((g(t-\delta) - g(t)) + (g(t+\delta) - g(t)))/2$ is the change of the gap value between time steps. Also shown are regions where the errors in the calculated probability (assuming energies are calulated with a precision $\sigma = 10^{-6}$) are greater than 5% (blue contours), 10% (orange), 20% (green) or 50% (red) for the given combination of g and d and a time step of 0.5 (a), 0.25 (b) 0.125 (c) and 0.03125 fs. The area with large possible errors in the upper part of the figures is due to uncertainty in the position of t_c, while the one in the lower part of the figures is due to σ. We see that a time step of 0.5 fs is often enough, but a not insignificant portion of the hops are in the upper "error" area, suggesting that this time step is sometimes too large. On the other hand, a time step of 0.03125 fs is often too small, since a significant portion of the determined gaps could be affected by the accuracy of the electronic structure calculation.

So decreasing the time step increases the precision with which the gap energy is calculated, but introduces larger errors in the second derivative. From these results, it appears that using a time step around 0.25 fs should give satisfactory accuracy for almost all hops in this system. However, in around 80% of the cases a step of 0.5 fs is already enough to be certain that the

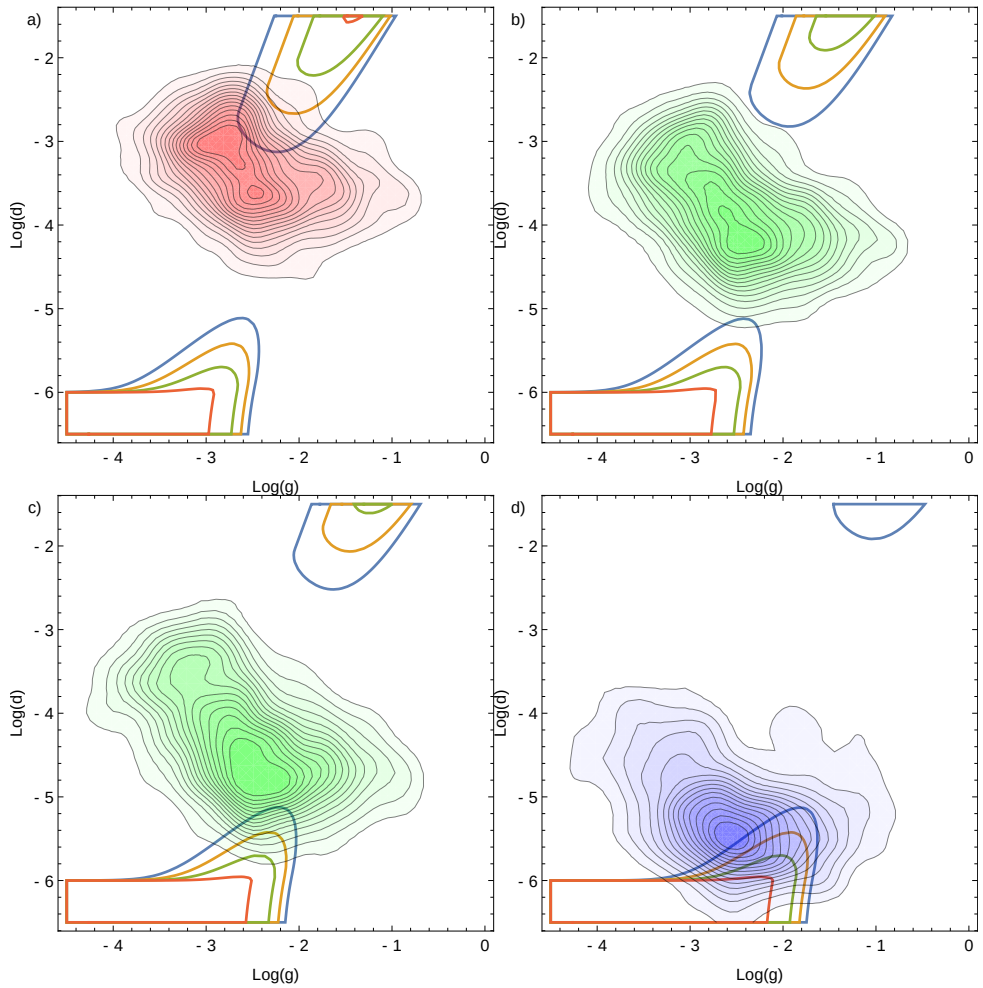


Figure 4.8: Error estimates for eq. 2.55 with $\sigma = 10^{-6}$ and a time step of 0.5 (a), 0.25 (b) 0.125 (c) and 0.03125 fs. Combinations of g and d for which errors greater than 5%, 10%, 20% and 50% are estimated are within the blue, orange, green and red contours, respectively. Distributions $(\log(g), \log(d))$ from the pyrrole B3LYP trajectories with the corresponding time step are shown by the shaded areas.

probability was calculated within 5% accuracy. The average pyrrole B3LYP trajectory encountered 18 gap minima. This means that using the developed algorithm for adapting the time step based on error estimates often reduces the overall cost of the simulation by ~ 21 evaluations per trajectory or 15% overall with no trade-off in the accuracy.

An analysis of the actual differences in hopping probabilities evaluated at the same gap minima over the entire simulation was performed to see whether the error estimates used to determine the time step are realistic. On average, the computed probabilities were quite accurate regardless of the time step. However, outliers where the calculated probability is different by 10 percentage points depending on the time step were not uncommon. In these cases, the estimated

errors were similarly high which means the adaptive algorithm would automatically modify the time step to accurately calculate the probability. With these results, we can be confident in the numerical stability of the developed LZSH algorithm. However, from the almost perfect agreement seen in pyrazine and significant differences seen in pyrrole, we can also note that increasing the density of states has a significant impact on the accuracy of LZSH.

4.2 Wave function overlap algorithms

In this section we are interested in the calculation of the overlap matrix between a set of N_A electronic states $\{|\Psi_A\rangle\}$ and N_B states $\{|\Psi'_B\rangle\}$ where the states are expanded in terms of Slater determinants (SDs)

$$S_{AB} = \langle \Psi_A | \Psi'_B \rangle = \sum_{o}^{n_{det}} \sum_{p}^{n'_{det}} d_o^A d_p'^B \langle \Phi_o | \Phi'_p \rangle \,, \tag{4.1}$$

where n_{det} is the number of SDs in the expansion of A, Φ_o are the SDs and d_o^A are their coefficients and all primed values refer to states $\{|\Psi'_B\rangle\}$. Here, and in the rest of the section, $o, p, ...$ are general indices, while $a, b, ...$, $i, j, ...$ and σ, τ will be used for occupied orbitals, virtual orbitals and spin functions, respectively. The only requirement for calculating S_{AB} is that the two sets of states share the same number of occupied orbitals (electrons). Apart from this, they can be expanded in terms of basis sets defined at different geometries (which is the usual case when they are used in FSSH calculations), but also using different basis sets altogether.

As stated in section 2.2.3, equation 4.1 requires the evaluation of a huge number of determinants and has a formal scaling of $O(n_{det} n'_{det} (n_alpha^3 + n_beta^3))$. Instead of trying to solve the general problem, we limit our scope to states which can be described by a CIS type wave function

$$|\Psi_A\rangle = \sum_{\sigma}^{\{\alpha,\beta\}} \sum_{a}^{n_{o,\sigma}} \sum_{i}^{n_{v,\sigma}} d_{ai,\sigma}^A \left| \Phi_{a,\sigma}^i \right\rangle \tag{4.2}$$

$$|\Psi'_B\rangle = \sum_{\tau}^{\{\alpha,\beta\}} \sum_{b}^{n_{o,\tau}} \sum_{j}^{n'_{v,\tau}} d_{bj,\tau}'^B \left| \Phi_{b,\tau}'^j \right\rangle \,, \tag{4.3}$$

where $\left| \Phi_{a,\sigma}^i \right\rangle$ are SDs with one electron moved from orbital a to orbital i.

Inserting the CIS ansatz into equation 4.1 we obtain

$$\left\langle \Psi_A \middle| \Psi'_B \right\rangle = \sum_{\sigma}^{\{\alpha,\beta\}} \sum_{\tau}^{\{\alpha,\beta\}} \sum_{a}^{n_{o,\sigma}} \sum_{b}^{n_{o,\tau}} \sum_{i}^{n_{v,\sigma}} \sum_{j}^{n'_{v,\tau}} d^{A}_{ai,\sigma} d'^{B}_{bj,\tau} \left\langle \Phi^{i}_{a,\sigma} \middle| \Phi'^{j}_{b,\tau} \right\rangle \tag{4.4}$$

or, after splitting into blocks as in equation 2.31 and rearranging the sums

$$\begin{aligned}
\left\langle \Psi_A \middle| \Psi'_B \right\rangle = \left\langle \Phi_0 \middle| \Phi'_0 \right\rangle_\beta & \left(\sum_{a}^{n_{o,\alpha}} \sum_{b}^{n_{o,\alpha}} \sum_{i}^{n_{v,\alpha}} \sum_{j}^{n'_{v,\alpha}} d^{A}_{ai,\alpha} d'^{B}_{bj,\alpha} \left\langle \Phi^{i}_{a} \middle| \Phi'^{j}_{b} \right\rangle_\alpha \right) \\
& + \left(\sum_{a}^{n_{o,\alpha}} \sum_{i}^{n_{v,\alpha}} d^{A}_{ai,\alpha} \left\langle \Phi^{i}_{a} \middle| \Phi'_{0} \right\rangle_\alpha \right) \left(\sum_{b}^{n_{o,\beta}} \sum_{j}^{n'_{v,\beta}} d'^{B}_{bj,\beta} \left\langle \Phi_{0} \middle| \Phi'^{j}_{b} \right\rangle_\beta \right) \\
& + \left(\sum_{a}^{n_{o,\beta}} \sum_{i}^{n_{v,\beta}} d^{A}_{ai,\beta} \left\langle \Phi^{i}_{a} \middle| \Phi'_{0} \right\rangle_\beta \right) \left(\sum_{b}^{n_{o,\alpha}} \sum_{j}^{n'_{v,\alpha}} d'^{B}_{bj,\alpha} \left\langle \Phi_{0} \middle| \Phi'^{j}_{b} \right\rangle_\alpha \right) \\
& + \left\langle \Phi_0 \middle| \Phi'_0 \right\rangle_\alpha \left(\sum_{a}^{n_{o,\beta}} \sum_{b}^{n_{o,\beta}} \sum_{i}^{n_{v,\beta}} \sum_{j}^{n'_{v,\beta}} d^{A}_{ai,\beta} d'^{B}_{bj,\beta} \left\langle \Phi^{i}_{a} \middle| \Phi'^{j}_{b} \right\rangle_\beta \right).
\end{aligned} \tag{4.5}$$

The above equation consists of eight blocks (four for each spin) which can be calculated separately. For convenience, since they are not important for further considerations, we will drop the spin indices going forward and also assume $n'_v = n_v$ and $N = N_A = N_B$. The most expensive step is the calculation of $n_o^2 n_v^2$ determinants of the form $\left\langle \Phi^{i}_{a} \middle| \Phi'^{j}_{b} \right\rangle$ in the quadruple sums. Apart from the determinant calculation, the sums themselves are costly because the CI coefficients are state dependent so the sums have to be computed for each pair of states separately, when implemented using matrix multiplications this adds an additional $O(n_o^2 n_v^2 N + n_o^2 n_v N^2)$ step resulting in a scaling of $O(N_A N_B n_o^2 n_v^2)$ for a straightforward implementation.

One approach to reducing the cost of such calculations is to notice that most matrices whose determinants are needed for overlap calculations are very similar. Specifically for CIS wave functions, in equation 4.5 we can see that the marices differ by at most two rows and two columns. This property was used previously by Plasser and coworkers[16] to significantly speed up calculations for MR-CIS expansion by using a single step Laplace's recursive formula during the determinant calculations. Expanding the determinants into minors along a column allows one to reuse the minors for all determinants which are different from the starting determinant by only that column. For the case of elements $\left\langle \Phi^{i}_{a} \middle| \Phi'^{j}_{b} \right\rangle$ appearing in CIS expansions, expanding

along column j we arrive at

$$\left\langle \Phi_a^i \middle| \Phi_b'^j \right\rangle = \sum_{c\neq a}^{n_o} o_{cj}(-1)^{c+b} \left\langle \Phi_{a,c}^i \middle| \Phi_b' \right\rangle + o_{ij}(-1)^{a+b} \left\langle \Phi_a \middle| \Phi_b' \right\rangle , \tag{4.6}$$

where o_{op} denotes the overlap between molecular orbitals o and p. Here, none of the determinants have a dependence on index j, meaning that the determinant calculations can be taken out of the sum over n_v' virtual orbitals (at a cost of an additional sum over $n_o - 1$ orbitals).

This approach can be furthered by expanding the level 1 minors (L1M) again using Laplace's recursive formula and writing $\left\langle \Phi_a^i \middle| \Phi_b'^j \right\rangle_\sigma$ in terms of level 2 minors (L2M)

$$\begin{aligned} \left\langle \Phi_a^i \middle| \Phi_b'^j \right\rangle = &\sum_{c\neq a}^{n_o} \sum_{d\neq b}^{n_o} o_{cj} o_{id} \operatorname{sgn}(b-d) \operatorname{sgn}(c-a)(-1)^{a+b+c+d} \left\langle \Phi_{a,c} \middle| \Phi_{b,d}' \right\rangle \\ &+ o_{ij}(-1)^{a+b} \left\langle \Phi_a \middle| \Phi_b' \right\rangle , \end{aligned} \tag{4.7}$$

where $\operatorname{sgn}(o-p)$ are used to change sign beyond the row/column removed from the L1Ms where the indices are reduced by one compared to the original matrix. Both sums in equation 4.7 are over occupied orbitals and the elements $\left\langle \Phi_{a,c} \middle| \Phi_{b,d}' \right\rangle$ also do not depend on i or j. Since interchanging the indices a and c or b and d has no effect on the values, the total number of unique determinants to calculate is reduced to $\frac{1}{4} n_o^2 (n_o - 1)^2$. This effectively removes the dependence on the number of virtual orbitals from the cost of calculating determinants changing the scaling to $O(n_o^7)$. An additional step is needed to reconstruct the full determinants which can be done efficiently using matrix-matrix multiplications with a scaling of $O(n_o^4 n_v + n_o^3 n_v^2)$, and the final sum still remains the same.

Another approach is based on calculating the natural transition orbitals (NTOs) for the CIS wave functions.[60,61,272] In the case of CIS wave functions, NTOs can be calculated by separate unitary transformations in the subspaces of occupied and virtual orbitals which is conveniently done by a singular value decomposition (SVD) of the matrix containing the CIS coefficients d_{ai}^A

$$\mathbf{D}^A = \mathbf{O}\boldsymbol{\Lambda}\mathbf{V}^\mathsf{T} . \tag{4.8}$$

This results in an $n_o \times n_o$ matrix $\mathbf{O}$ containing the hole NTOs in the basis of occupied orbitals, an $n_v \times n_v$ matrix $\mathbf{V}$ containing the particle NTOs in the basis of virtual orbitals and $n_o \times n_v$

matrix Λ containing the n_o (assuming $n_o < n_v$, as is always the case in practice) singular values λ_k^A on the diagonal and zeroes everywhere else.

The CIS wave functions of equations 4.2 and 4.3 can be written in terms of the NTOs as

$$|\Psi_A\rangle = \sum_{\sigma}^{\{\alpha,\beta\}} \sum_{o}^{n_\sigma} \lambda_{o,\sigma}^A \left|\Theta_{o,\sigma}^A\right\rangle \tag{4.9}$$

$$|\Psi_B\rangle = \sum_{\tau}^{\{\alpha,\beta\}} \sum_{p}^{n_\sigma} \lambda_{p,\sigma}^B \left|\Theta_{p,\sigma}^B\right\rangle, \tag{4.10}$$

where $\left|\Theta_{o,\sigma}^A\right\rangle$ are singly excited SDs with excitations from the oth occupied to oth virtual natural orbital. In this form, the full wave function is a linear combination of only $n_\alpha + n_\beta$ SDs. Inserting into Eq. 4.1 we get

$$\begin{aligned}
\left\langle \Psi_A \middle| \Psi'_B \right\rangle = &\left\langle \Theta_0^A \middle| \Theta_0'^B \right\rangle_\beta \left(\sum_o^{n_\alpha} \sum_p^{n_\alpha} \lambda_{o,\alpha}^A \lambda_{p,\alpha}'^B \left\langle \Theta_o^A \middle| \Theta_p'^B \right\rangle_\alpha \right) \\
&+ \left(\sum_o^{n_\alpha} \lambda_{o,\alpha}^A \left\langle \Theta_o^A \middle| \Theta_0'^B \right\rangle_\alpha \right) \left(\sum_p^{n_\beta} \lambda_{p,\beta}'^B \left\langle \Theta_0^A \middle| \Theta_p'^B \right\rangle_\beta \right) \\
&+ \left(\sum_o^{n_\beta} \lambda_{o,\beta}^A \left\langle \Theta_o^A \middle| \Theta_0'^B \right\rangle_\beta \right) \left(\sum_p^{n_\alpha} \lambda_{p,\alpha}'^B \left\langle \Theta_0^A \middle| \Theta_o'^B \right\rangle_\alpha \right) \\
&+ \left\langle \Theta_0^A \middle| \Theta_0'^B \right\rangle_\alpha \left(\sum_o^{n_\beta} \sum_p^{n_\beta} \lambda_{o,\beta}^A \lambda_{p,\beta}'^B \left\langle \Theta_o^A \middle| \Theta_p'^B \right\rangle_\beta \right).
\end{aligned}$$

Only double sums over n_o remain in equation 4.11 and only n_o^2 determinants need to be calculated for each spin. The trade-off here is that the orbitals are now state-dependent so the determinants $\left\langle \Theta_o^A \middle| \Theta_p'^B \right\rangle$ need to be calculated separately for each pair of states, resulting in an overall scaling of $O(N^2 n_o^5)$. Since only a small number of states is included in a typical calculation, it can be assumed that $N^2 \ll n_o^2 \ll n_v^2$. This would mean that calculations based on equation 4.11 (denoted ONTO further in the text) are expected to significantly outperform those based on equation 4.7 (denoted OL2M) which, in turn, are faster than those based on equation 4.5 (denoted OCIS).

We should note here that this type of calculation can trivially be extended to the calculation of Dyson orbitals using equation 3.25. This requires an additional loop to cycle over the orbitals from which an electron can be removed when expanding the N electron wave func-

tion into minors following which $N-1$ overlap determinants are calculated. Due to this extra loop, calculating Dyson orbitals in this manner has even worse scaling than regular overlap calculations and thus performance improvements are even more crucial.

4.2.1 Performance

In section 4.2 we presented two efficient algorithms for calculating overlaps of wave functions expanded in terms of Slater determinants built from different sets of basis functions. We have shown that the developed algorithms should have significantly better scaling than a trivial implementation of the underlying equations. In this section, we will test the actual performance of these algorithms.

The cost of a wave function overlap calculation depends on three parameters: the size of the system, the size of the basis set and the number of states for which overlaps are required. The size of the system (number of electrons) fully determines the number of occupied orbitals, while the number of virtual orbitals depends both on the number of atoms of each type and the number of basis functions added for each such atom. The number of states included depends on the problem being studied, but in general larger systems have a denser manifold of states which means that they usually require the consideration of more states.

A series of poly-alanine peptides was chosen for numerical tests of the performance of the algorithms. The systems ranged from three to thirty residues. This corresponds to systems with 62 to 575 occupied orbitals. All calculations were performed for two basis sets of significantly different size, the small def-SV(P) and the larger cc-pVTZ basis set. The number of virtual orbitals went from 196 to 1843 for the smaller set and from 656 to 6083 for the larger one. All calculations were performed on a single CPU core (Intel(R) Xeon(R) CPU E3-1245 v5 @ 3.50GHz) and the timings are shown in Figure 4.9.

The trivial implementation of the calculation is significantly slower than all of the results shown so comparisons are made only between the new algorithms (OL2M and ONTO) and the previously available algorithm based on level 1 minors (OL1M algorithm).[16] Both algorithms improve upon the OL1M algorithm, but the ONTO algorithm also outperforms the OL2M algorithm by an order of magnitude even for the smallest test case. The differences are more pronounced for larger systems. Going beyond ≈ 200 occupied orbitals, the OL2M algorithm quickly becomes too expensive while the ONTO algorithm still takes only eight hours on a

single processor at 575 occupied orbitals.

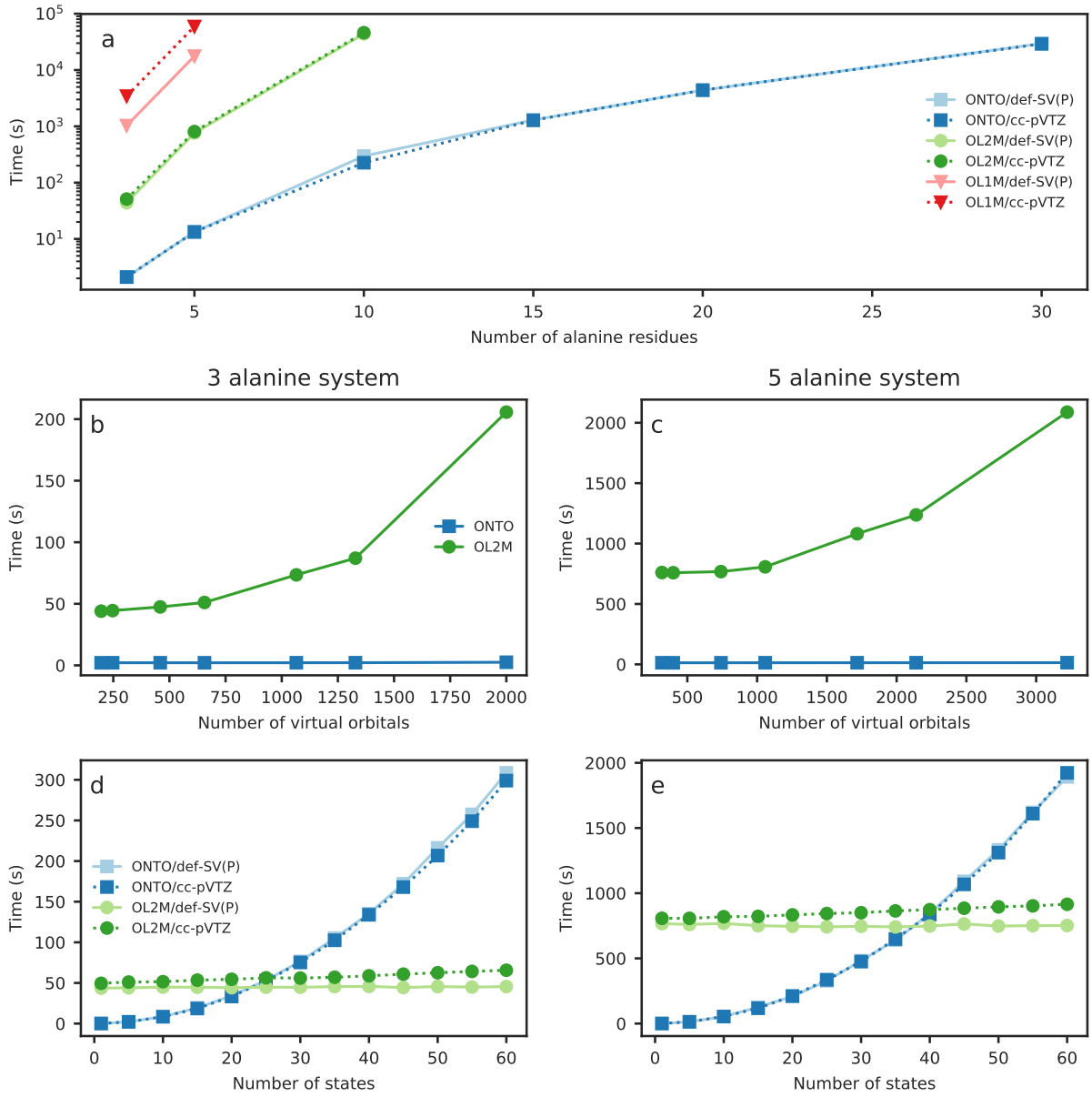


Figure 4.9: Comparison of the performance of the OL1M (red), OL2M (green) and ONTO (blue) algorithms. (a) Scaling with the system size (number of residues) with 5 excited states. Additional tests were performed on the 3 alanine (b, d) and 5 alanine (c, e) systems. (b) 3-Ala and (c) 5-Ala scaling with basis set size for 5 excited states going from the def-SV(P) to the aug-cc-pVQZ basis set. (d) 3-Ala and (e) 5-Ala scaling with the number of states included in the calculation. Figure originally published in Ref. (*35*).

We see in Figure 4.9a that the curves for the two basis sets almost perfectly overlap for both of the new algorithms. This shows that the benefits of these algorithms increases with the size of the basis set. To further test this, in Figure 4.9b and c we look at a wider range of basis sets when applied to the two smaller systems with 3 and 5 alanine residues, respectively.

The series of basis sets included was def-SV(P) < cc-pVDZ < aug-cc-pVDZ < cc-pVTZ < aug-cc-pVTZ < cc-pVQZ < aug-cc-pVQZ. As expected the ONTO algorithm still shows no dependence on the number of virtual orbitals since all matrices entering the algorithm after the initial SVD decomposition to generate the NTOs (which has a negligible cost even for the largest calculations) have dimensions proportional only to the number of occupied orbitals. On the other hand, the cost of the calculation with the OL2M algorithm initially also does not depend on the basis set size, but as we begin to use augmented triple zeta or quadruple zeta basis sets a quadratic dependence emerges. This is because of the $O(n_o^3 n_v^2)$ scaling step of reconstructing the full determinants from the minors which becomes the bottleneck of the algorithm for large enough n_v.

The last variable when looking at the scaling of the different algorithms is the size of the overlap matrix being calculated. In Figure 4.9d and 4.9e we see that the OL2M algorithm eventually becomes preferable to the NTO algorithm. This is expected since the dominant steps for large systems in the OL2M and ONTO algorithms are always $O(n^7)$ and $O(n^5 N_A N_B)$, respectively. From this, we expect that the point at which the OL2M algorithm starts outperforming the ONTO algorithm is when $N_A N_B / n^2 > c$, where c is some implementation dependent constant. From the data in Figure 4.9 we see that $c \approx \frac{1}{6}$ in the current code.

Here we note that both algorithms are amenable to parallelization since most of the time is spent in the loops calculating determinants which do not depend on each other. While the results shown above were performed on a single processor, the code itself is parallelized and shows a nearly linear speedup with the number of cores. In addition to this, both algorithms still consist of calculations of many similar determinants. We can take further advantage of this fact by re-utilizing and updating the LU factorization while calculating the determinants. This has been implemented for the OL2M algorithm resulting in a 10 to 100 times faster calculation of the minors.[62] This implementation fills the gap where calculations of overlaps could become the bottleneck for medium sized systems when a large number of states needed to be considered. It also opens the possibility of extending the procedure to compute the overlaps between more complex wave function types such as MR-CIS without making approximations.

4.2.2 Approximations

Based on the results of section 4.2.1 we can conclude that overlaps can be calculated at virtually no cost compared to the electronic structure calculation for most systems typically studied using FSSH. For ADC(2) this is certainly true, but TDDFT scales more favorably and can be used to treat much larger systems at a low cost, especially with the use of graphical processing units (GPUs).[63–65] When the cost of the overlap calculation is not negligible approximations should be considered.

In section 2.2.3 we mentioned two approximation schemes for calculating TDCs. Here we will take a more detailed look at these schemes in the context of our developed algorithms and comment on the viability of their implementation and possible errors. From the approximation scheme suggested by Pittner *et al.*,[15] the mutual excitation threshold is not relevant for CIS type wave functions since all determinants are singly excited. On the other hand, the coefficient product threshold t^{cp} of Pittner and the wave function norm threshold t^{norm} of Plasser *et al.*[16] are both applicable to these types of wave functions.

The efficiency of the OL2M algorithm stems from the ability to efficiently calculate overlaps between many Slater determinants at once. Dropping a small number of SDs from the wave function expansion or SD overlaps from eq. 4.5 would have negligible effect on the overall cost of the calculation. With larger numbers of dropped terms (higher level of approximation), one would have to select which minors are worth precomputing and when it is more efficient to directly calculate the overlap determinant. The more terms are dropped, the less minors would be worth calculating and the difference between the OL2M and direct algorithms would be smaller. For this reason, the OL2M algorithm is not particularly amenable to this type of approximation.

On the other hand, the ONTO algorithm is based on reducing the number of determinants in the expansions and reducing this number even further directly reduces the cost of the calculation. Wave functions expanded in terms of NTO excitations are usually dominated by a small number of determinants[272] which makes them ideal for approximations based on neglecting determinants with small coefficients. Both the t^{cp} and t^{norm} approximations are trivial to implement in this algorithm, with the condition for t^{cp} changed slightly from eq. 2.33 to reflect that the orbitals are different for each state so coefficients for each pair of states are checked

separately

$$t^{\mathrm{cp}}_{ab,AB} \geq \left| d^A_a d'^B_b \right| . \tag{4.11}$$

The two approximation schemes are quite similar. For the purposes of analyzing the possible errors, the t^{norm} scheme is more practical so we will use it in this section. However, similar conclusions can be made for the scheme based on t^{cp}. We start by truncating the wave functions to keep only the k_t largest coefficients

$$\left|\tilde{\Psi}_A\right\rangle = \sum_o^{k_t} d^A_o \left|\Phi^A_o\right\rangle , \tag{4.12}$$

so that the norm of each truncated wave function $\left|\tilde{\Psi}_A\right\rangle$ is larger than a selected threshold parameter t^{norm}

$$\left\langle\tilde{\Psi}_A\middle|\tilde{\Psi}_A\right\rangle = \sum_o^{k_t} (d^A_o)^2 \geq t . \tag{4.13}$$

This truncation obviously reduces the norm of the wave functions causing a systematic underestimation of the calculated overlaps but it can also affect the angles between them. The norm can easily be fixed by renormalizing either the truncated states before calculating the overlaps or the overlap matrix itself. In the cases where angles are not changed by the truncation, the exact overlaps are recovered in this way and initial tests using this methods have resulted in very accurate overlaps for threshold values above 0.95.[16]

The maximum possible overestimation error introduced in this way is $1-t$ which is acceptable (and expected). However, it is also possible for the overlaps to be underestimated by as much as $2\sqrt{t}\sqrt{1-t}$. This can be simply illustrated with an overlap of two states dominated by the same two excitations

$$|\Psi_A\rangle = \sqrt{t}\,|\Theta_1\rangle + \sqrt{1-t}\,|\Theta_2\rangle \quad \Rightarrow \quad \left|\tilde{\Psi}_A\right\rangle = \sqrt{t}\,|\Theta_1\rangle \tag{4.14}$$

$$|\Psi_B\rangle = \sqrt{t}\,|\Theta_2\rangle + \sqrt{1-t}\,|\Theta_1\rangle \quad \Rightarrow \quad \left|\tilde{\Psi}_B\right\rangle = \sqrt{t}\,|\Theta_2\rangle . \tag{4.15}$$

The overlap of these two states is equal to $2\sqrt{t}\sqrt{1-t}$ when the full wave functions are considered. If the states are truncated, only the dominant excitations are kept and the overlap is 0. For $t = 0.9$, this produces an error of 0.6, and even for $t = 0.99$ the error is still 0.2 and these errors are not reduced by renormalizing or orthogonalizing the overlaps. While this situation might

appear extreme at first glance, it is the type of behaviour expected near conical intersections or avoided crossings where significant mixing between electronic states can occur for small changes in nuclear coordinates.

There is also reason to suspect that the ONTO algorithm will be more vulnerable to these types of errors since NTO expansions are dominated by a small number of determinants. For this reason detailed numerical tests of the approximation scheme based on the threshold t^{norm} for individual wave functions were performed. To ensure that the approximation is tested thoroughly, we repeat all overlap calculations along FSSH trajectories using different values of the t^{norm} truncation threshold parameter. For the test case, two trajectories (one TDDFT/B3LYP/def2-TZVPD and one ADC(2)/aug-cc-pVDZ trajectory randomly selected from the ensemble) simulating the nonadiabatic dynamics of pyrrole excited to $B_2(\pi\pi^\star)$ state[37] were chosen. These dynamics are covered in detail in section 4.4 and were originally performed without truncation since the molecule is small so the overlap calculation is fast. The reason for choosing this system is that the manifold of excited states is quite dense for such a small molecule which results in many state crossings during the dynamics. As noted in section 4.2.2, such state crossings are regions where the suggested truncation scheme might result in large errors.

Both trajectories consist of 500 steps (250 fs with a 0.5 fs time step) with 10 excited states resulting in a a total of 50 000 overlap matrix elements between excited states. The basis set for the TDDFT calculations included 18 occupied and 212 virtual orbitals, while the basis set for the ADC(2) calculation included 13 occupied (the five core orbitals were frozen) and 142 virtual orbitals. The total number of determinants computed along the trajectory with different settings for t^{norm} is shown in Figure 4.10.

The efficiency of the truncation scheme is determined by the number of excitations with significant coefficients in the wave function expansion. In the TDDFT calculations, the auxilliary wave functions, written in terms of excitations from Kohn-Sham orbitals, are dominated by only a few excitations. Because of this, for small values of t^{norm} the possibility of reusing determinants for multiple states in the OCIS algorithm slightly outweighs the compactness of the wave functions in the ONTO algorithm. However, as the threshold is increased above $t^{\text{norm}} = 0.98$, the cost of the OCIS algorithm quickly overtakes the ONTO algorithm. In the ADC(2) calculation, a larger number of singly excited determinants has significant contributions and the ONTO algorithm requires less determinants for $t^{\text{norm}} = 0.99$ than the OCIS algorithm for $t^{\text{norm}} = 0.9$.

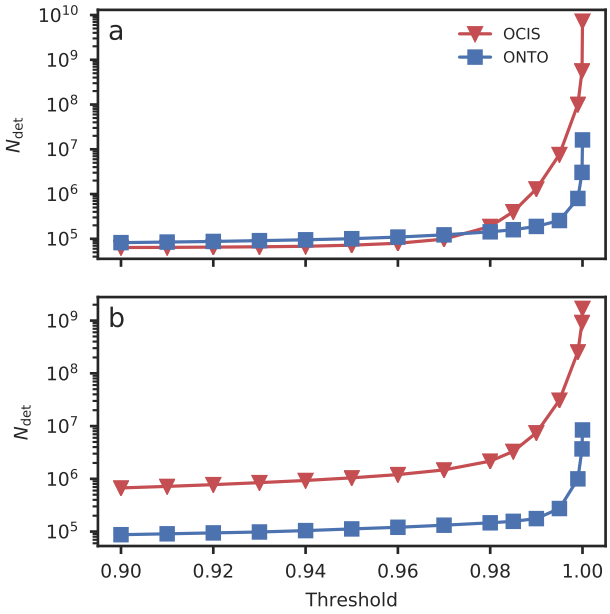


Figure 4.10: Comparison of the total number of determinants computed along a (a) TDDFT(B3LYP)/def2-TZVPD and (b) ADC(2)/aug-cc-pVDZ (b) FSSH trajectory of pyrrole for different values of the CI vector truncation threshold t^{norm} using the ONTO and OCIS algorithms. Figure originally published in Ref. (*35*).

As expected, for both trajectories the total number of determinants which needed to be computed was lower with the ONTO algorithm by more than two orders of magnitude for higher values of t^{norm}. For larger systems or smaller numbers of states, these differences are expected to be even more pronounced in favour of the ONTO algorithm.

The more important information, the errors introduced by the approximation, is shown in Figure 4.11. All overlap matrix elements were compared to the corresponding elements calculated using the full wave functions. The mean error of the approximation scheme is small for both the OCIS and ONTO algorithm and is reduced further by orthogonalization of the overlap matrix. However, looking at mean errors is not sufficient since a large error in a single matrix element can cause an unphysical hop which affects the remainder of the trajectory. As stated previously, we expected that wave functions with large contributions from a few determinants would result in larger errors after the approximation. However, this was only partially the case. The observed maximum errors are approximately $\sqrt{1-t^{\text{norm}}}$ for low threshold values (half of the worst-case scenario error). For higher threshold levels ($t^{\text{norm}} >= 0.999$), the observed maximum errors separate further from the worst-case scenario, but always remain significantly higher than $1-t^{\text{norm}}$.

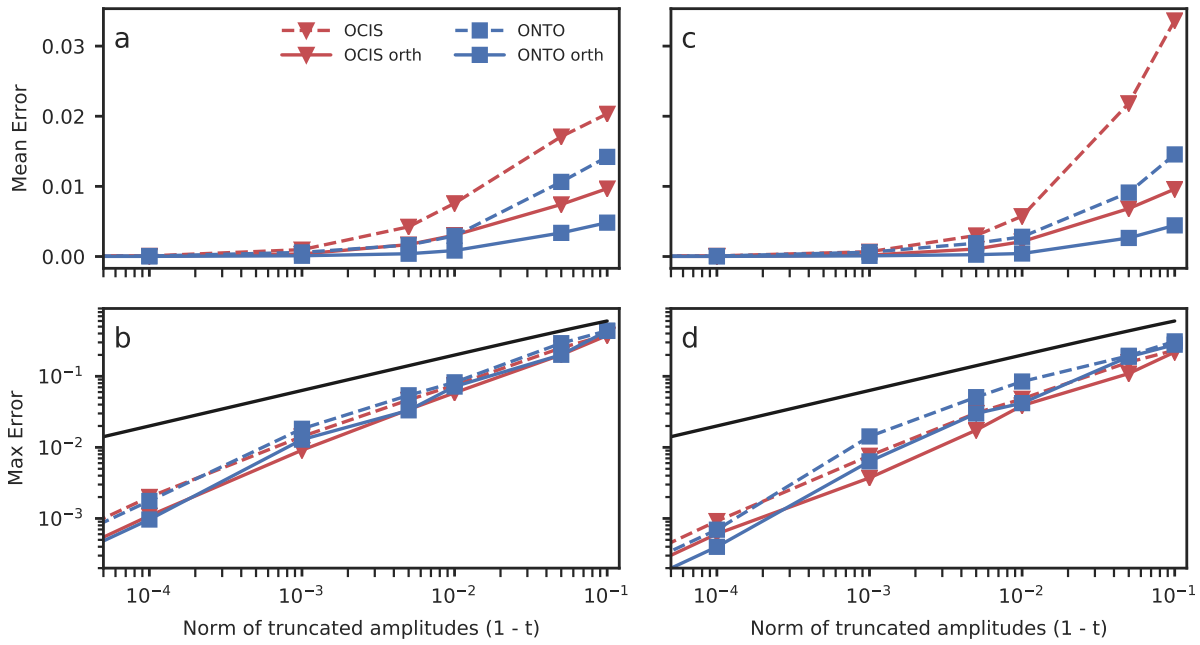


Figure 4.11: Mean (a, c) and maximum (b, d) errors for overlap elements computed with different threshold values along a (a, b) TDDFT(B3LYP)/def2-TZVPD and (c, d) ADC(2)/aug-cc-pVDZ FSSH trajectory using the ONTO and OCIS algorithms. Errors for the raw overlap matrix elements (full lines), errors after orthogonalization (dashed lines) and the maximum error expected for the approximation scheme ($2\sqrt{1-t}\sqrt{t}$, black line) are shown. Figure originally published in Ref. (*35*).

In the case of the TDDFT trajectory, both the CIS and NTO expansions are quite compact and the maximum errors are approximately the same with both algorithms. ADC(2) wave functions usually have contributions from a larger number of excitations. Still, the maximum errors of the OCIS algorithm are only slightly smaller than those of the ONTO algorithm. Orthogonalization has almost no effect on the maximum errors since they are not caused by the loss of norm of the wave functions. In any case, it is important to remember that, at least for TDDFT and ADC(2) wave functions, errors of the order $\sqrt{1-t^{\text{norm}}}$ are possible with this approximation scheme and it should be avoided or used with the largest possible values of the truncation threshold when quantitative results are required. It is worth noting here that approximate overlaps used for CC2/ADC(2) calculations are actually always calculated by truncating the wave functions since only their singles amplitudes are taken to construct a formal CIS wave function.[59] This is different from the approximation mentioned above since none of the neglected determinants are singly excited determinants and, assuming that the reference determinant doesn't change significantly, will have negligible overlaps with the non-neglected singly excited determinants.

The maximum error in this case should be approximately equal to the norm of the neglected amplitudes which is on average equal to 0.076 along the tested trajectory.

4.3 Decomposition of spectra

One outstanding problem of surface hopping methods is the interpretation of obtained results in terms of changes in the electronic structure of the system. In this regard, the main advantage of surface hopping methods is also its problem. Simulating dynamics on-the-fly in the basis of adiabatic states is the reason why this method is cheap and simple to use compared to methods which require constructing global diabatic potential energy surfaces. On the other hand, diabatic surfaces are useful when interpreting the results because electronic properties change smoothly over the surface. While we can compute reasonably accurate estimates of many observables using the nuclear ensemble approach described in section 2.2.5, we often want to know how the different electronic states contribute to these observables.

In this section, we will look at the absorption spectra and how we can assign them within the nuclear ensemble method. It is important to note that we are not aiming for a perfect quantitative method, but rather a qualitative way to show how much a particular "diabatic" state contributes to a particular region of the absorption spectrum. Wave function overlaps are a natural choice for this task since they are the most direct comparison of the electronic structure of two states at different nuclear geometries. The procedure is simple. 1) Define a reference geometry, this is usually the ground state minimum geometry. 2) Characterize the states at this geometry (this step will be detailed later). 3) Calculate the overlap matrix between this geometry and all other points of the ensemble. 4) Assign the states at all other points using those at the reference geometry based on the overlaps, this is done by solving the assignment problem using the Hungarian algorithm.[273,274] 5) Calculate averaged values for each "diabatic" state by summing over the states assigned to each state of the reference geometry.

Upon assigning the states at all points of the ensemble, the spectra are calculated as

$$\sigma(E) \propto \sum_{A}^{N_s} \frac{\bar{f}_{0A}}{\sqrt{2\pi}\gamma_A} \exp\left(-\frac{(E-\bar{E}_{0A})^2}{2\gamma_A^2} \right), \tag{4.16}$$

where $\bar{E}_{0A}$ and $\bar{f}_{0A}$ are, respectively, the mean excitation energy and oscillator strength of each of the thirty states and γ_i is the standard deviation of the excitation energy. A graphical summary

of this procedure is shown in Figure 4.12.

These spectra are different from the standard nuclear ensemble approach where each point is broadened using the same kernel function. In our approach, the spectra use state-specific widths based on the standard deviations γ_i. However, in general the two approaches give a very similar overall shape of the spectrum with the main difference being the easy assignment of the contributions from individual states using the current approach.

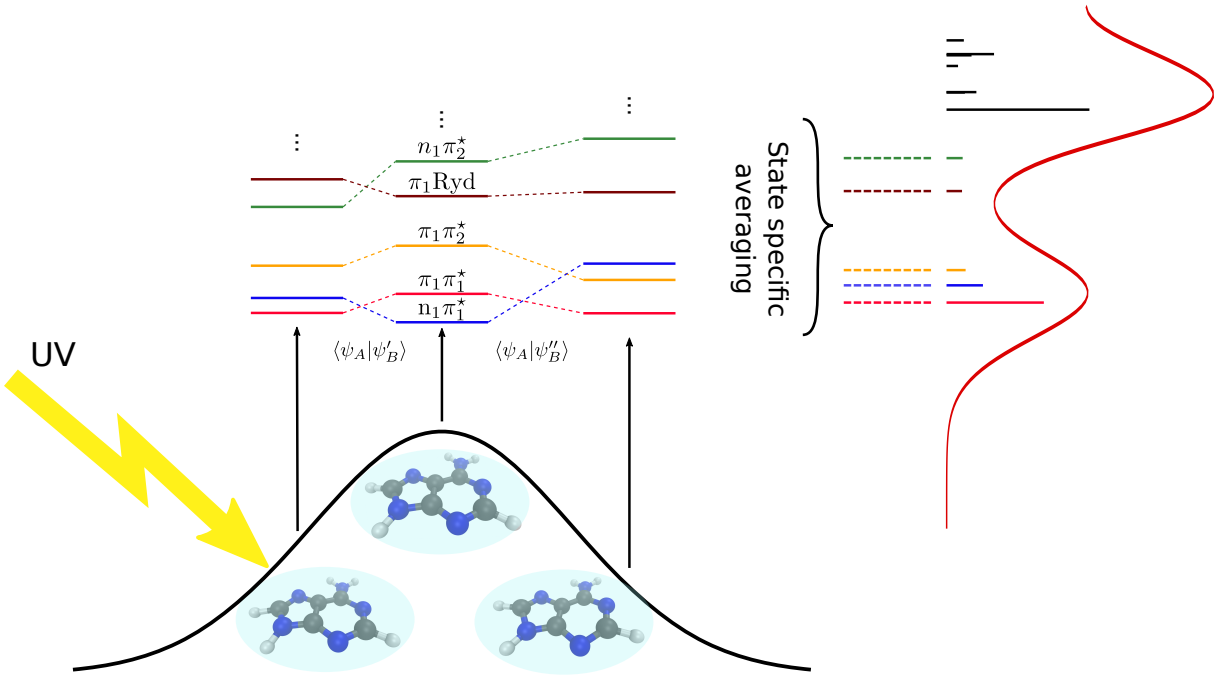


Figure 4.12: Schematic representation of the simulation of state-averaged absorption spectra.

It is clear that this is by no means a real diabatization procedure. When a state is a nearly equal mixture between two states of the reference geometry it will be assigned fully to one of those states. However, if the reference states are chosen well strong mixing should be limited to regions near conical intersections and most entries of the overlap matrices will be close to zero or one. If strong mixing between states is present at a significant portion of the geometries, an alternative algorithm can be used. In this algorithm, we replace steps 4) and 5) above and instead spectra for each "diabatic" state are obtained by summing over all states at each geometry weighted by the square of their overlap with the reference state. This is equivalent to writing the states at each geometry in the basis of reference states and using the squares of the coefficients

$c_{AF}(\mathbf{R_i}) = \langle \psi_A(\mathbf{R_0}) | \psi_F(\mathbf{R_i}) \rangle$ as additional weights in the nuclear ensemble approach:

$$\sigma_A(E) \propto \sum_F \sum_i^{N_p} w_i E_{0F}(\mathbf{R}_i) |c_{AF}(\mathbf{R_i})|^2 f_{0F}(\mathbf{R}_i) k\left(\frac{E - E_{0F}(\mathbf{R}_i)}{\delta_F}\right). \tag{4.17}$$

This alternate procedure looses the simplicity of 1:1 correspondence between the states at each geometry and the easy definition of state-specific broadening factors $\{\gamma_A\}$. On the other hand, it is less vulnerable to errors due to mixing between the states.

4.3.1 Absorption spectra of nucleobases

The procedure described above was used to study the spectra of DNA nucleobases in a wide range of excitation energies (350-190 nm). In this study we were interested in characterizing the absorption spectra and also in the effect of the zero point energy, temperature and solvation on the positions of the peaks of individual states in the spectra. The four nucleobases studied are adenine, thymine, guanine and cytosine (Figure 4.13). For each of these, three phase space distributions were sampled using eq. 2.58. These were the gas phase distribution at $T = 0$ K (100 points for each molecule), the gas phase distribution at $T = 298$ K (100 points) and a distribution at $T = 298$ where the minimum and normal modes were obtained with the COSMO model.

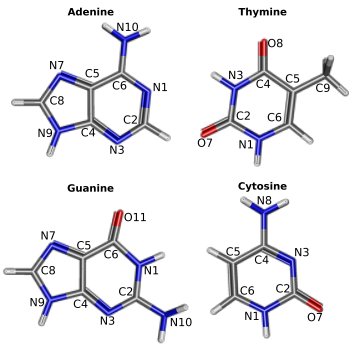


Figure 4.13: MP2/aug-cc-pVDZ optimized ground state geometries of the four nucleobases and atom labeling. Figure originally published in Ref.(*36*).

The electronic states at the reference geometry of each of the four bases (the gas phase ground state minima) were assigned based on their dominant hole and particle transition orbitals. For all four nucleobases all states have either an n or π hole orbital and either a π^* or a Rydberg particle orbital so we can roughly categorize them into four types: nπ^*, nRyd, $\pi\pi^*$ and

πRyd. The states are mutually orthogonal, but often share the same dominant hole or particle orbital. This can be automatically determined by computing the overlaps between the individual hole/particle orbitals of each state, mostly resulting in values close to zero or one. For example, the lowest 14 states of adenine can all be described in terms of four hole and seven particle NTOs (Table 4.2).

In addition to overlaps, we compute the second moments along the z-axis (out-of-plane axis) of the orbitals to aid in the state assignment (and later visualization). Since the nucleobases are planar or almost planar, the four types of relevant orbitals can be very clearly separated based on these moments. Among the hole orbitals, n type orbitals are mostly confined to the plane so their $\left\langle z^2 \right\rangle$ values are small ($0.4-0.9$ a_0^2) while π orbitals extend further out of plane ($1.5-1.8$ a_0^2). The π^* orbitals are more diffuse ($2.0-5.0$ a_0^2) while the defining characteristic of Rydberg orbitals is their very diffuse character ($8.0-32$ a_0^2). Using these moments to determine the type of each orbital and overlaps to detect orbitals repeating in multiple states, an automatic procedure for assigning all states can be easily devised. This type of assignment is shown in Table 4.2 for adenine, while additional information on the states of all four nucleobases both in gas phase and in solution is presented in the original article[36] with a more detailed discussion of specific relevant states and comparisons with previously reported findings.

Table 4.2: ADC(2) excitation energies (in eV) and oscillator strengths of 9H-adenine computed at the optimized ground state geometry. All transition below 6.53 eV (190 nm) and selected transitions with large oscillator strengths at higher energies are included. States are characterized by their dominant NTO pairs contributing. Table adapted from Ref. (*36*).

S	NTO character	E^{ref}	f^{ref}
S_1	$n_1\pi_1^*$	4.99	0.013
S_2	$\pi_1\pi_1^*$	5.10	0.251
S_3	$\pi_1\pi_2^*$+ π_2 pi_1^*	5.12	0.036
S_4	$\pi_1 Ryd_1$	5.38	0.010
S_5	$n_1\pi_2^*$	5.63	0.002
S_6	$\pi_1 Ryd_2$	5.69	0.004
S_7	$n_1 Ryd_1$	5.96	0.022
S_8	$n_2\pi_2^*$	6.03	0.001
S_9	$\pi_1 Ryd_3$	6.17	0.008
S_{10}	$\pi_2\pi_1^*$ + $\pi_1\pi_2^*$ +$n_1 Ryd_2$	6.20	0.425
S_{11}	$n_1 Ryd_2$+ $\pi_1\pi_2^*$+ $\pi_2\pi_1^*$	6.40	0.0975
S_{12}	$\pi_1 Ryd_4 + \pi_2 Ryd_1$	6.42	0.0319
S_{13}	$\pi_1 Ryd_5$	6.46	0.0056
S_{14}	$\pi_2 Ryd_1$	6.50	0.0019
S_{16}	$\pi_3\pi_1^*$ + $n_1 Ryd_3$+ $\pi_2\pi_2^*$	6.59	0.039
S_{27}	$\pi_2\pi_2^*$ + $n_1 Ryd_4$+ $\pi_3\pi_1^*$	7.17	0.163

Spectra obtained using Eq. 4.16 are shown in Figure 4.14 along with the experimental absorption spectra. Mean excitation energies and oscillator strengths are shown as a red stick spectrum. Spectra were normalized to scale with the experiment, but no shift in the energy domain was used. We see that there is good agreement between the experiment and calculated spectra. In the case of adenine and thymine this agreement is almost perfect. For guanine, the peaks corresponding to the L_a and L_b are further apart in the experiment than in the simulated spectrum where they appear merged as a single band. This difference is acceptable, especially in light of the fact that the distance and relative intensity of these two bands depends on the exact experimental procedure.[275,276] It is also possible that the discrepancy is due to the 7H form of guanine which was not accounted for in the theoretical spectrum but might have a significant contribution to the experimental spectrum. For cytosine we found that the full simulated spectrum is redshifted with respect to the experimental one and the height of the first band is underestimated.

With the ability to assign states across different geometries, we can analyze how the average

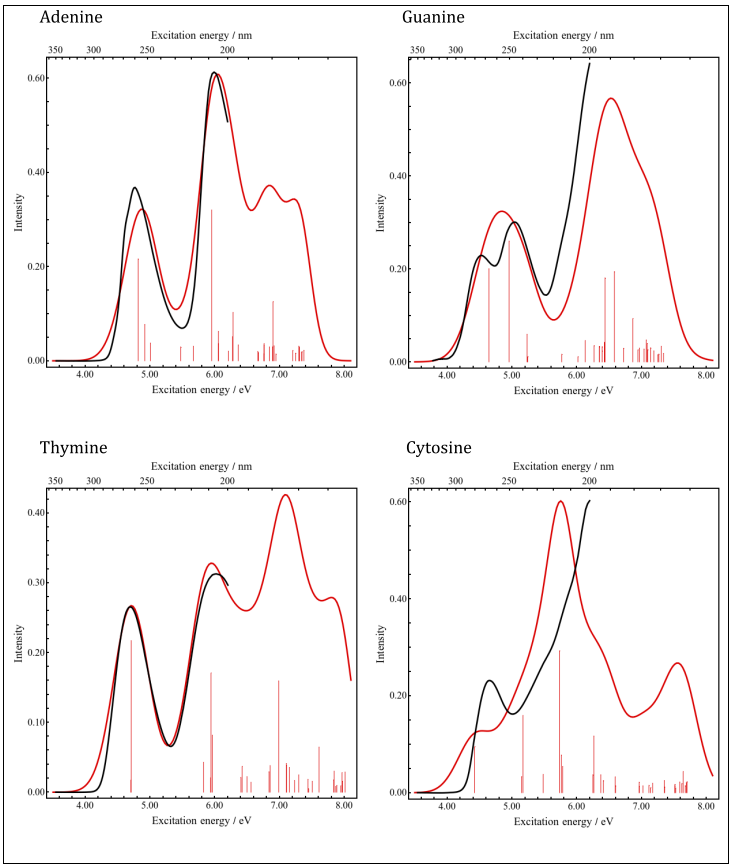


Figure 4.14: Theoretical (state-averaged, red line) and experimental (black line) UV absorption spectra of DNA nucleobases in water. Sticks indicate the mean excitation energies and oscillator strengths of the lowest 30 transitions. Figure originally published in Ref. (*36*).

energies of these states are shifted across the Wigner distribution around the ground state minimum. In addition to this, we can compare the states between different ensembles, for example at different temperatures or with a different environment. These effects are studied individually and all together in Figure 4.15. Each state is represented by a single marker (crosses, triangles, squares and circles for the states of adenine, guanine, thymine and cytosine, respectively) whose color indicates the shift of the mean energy of the state with respect to the vertical excitation energy. As mentioned above, $\left\langle z^2 \right\rangle$ of the hole and particle orbitals can be used to differentiate orbitals of different type which is used here to group the states by their character.

First we consider what we gain by considering an ensemble of geometries instead of vertical excitation energies. Compared to the ground state minimum, the gap between the ground state and any other state should be larger at some geometries and smaller at other geometries. However, on average the gap is smaller. Among the electronic states shown here, the π^* states are redshifted by 0.15 eV while the Rydberg states are shifted by 0.09 eV on average. This is

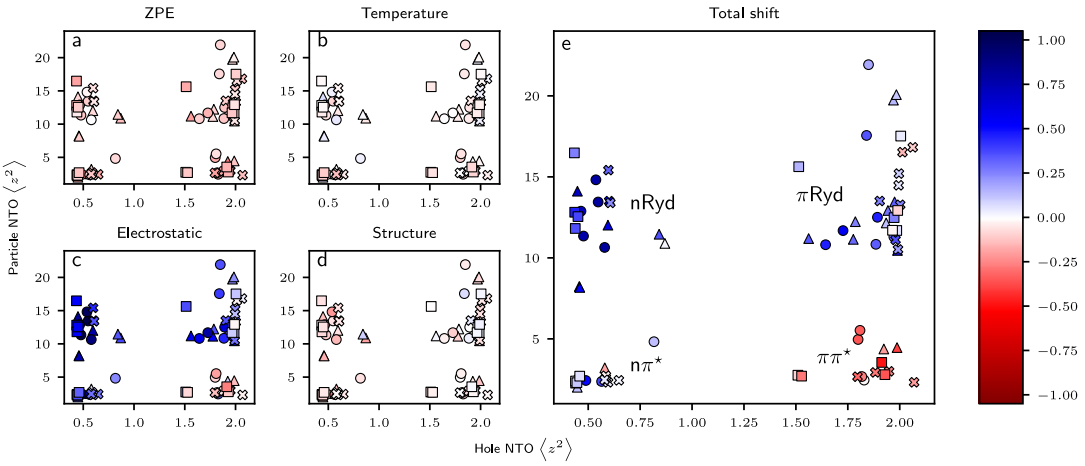


Figure 4.15: Solvent induced shifts for different types of electronic transitions. The states are characterized by their dominant hole and particle NTOs which are recognized and grouped based on their second moments. States of adenine, guanine, thymine and cytosine are denoted with crosses, triangles, squares and circles, respectively. Figure originally published in Ref. (*36*).

consistent with results from Bai and coworkers[277] who calculated an average shift of 0.11 eV for the states of a set of 28 organic molecules. Based on simple analytic models they show that the main cause of this shift is that the ground state is more tightly bound (and is thus more destabilized by geometry changes) than the excited states. The difference between the shifts of the π^* and Rydberg states are also explained by similar considerations. The main difference here is the significant softening of out-of-plane modes due to the weakening of π-conjugation in π^* states which occurs to a much smaller degree in Rydberg states. In Figure 4.15b we see a redshift of the π^* and Rydberg states by 0.04 and 0.01 eV, respectively, due to a change in temperature from 0 K to 298 K. This effect is smaller than the zero-point energy since it involves significantly less energy. In general it is explained by similar considerations of the frequencies of the normal modes, but at the temperature considered only the low-frequency normal modes are noticeably excited.

Next, we look at the effect of changing the environment on the excitation energies. In Figure 4.15c we look only at the electrostatic effects by computing the energies within the COSMO environment. These effects are quite large and lead to a blueshift of the nπ^*, nRyd and πRyd states by 0.38, 0.69 and 0.34 eV. On the other hand, $\pi\pi^*$ are redshifted by 0.03 eV. Figure 4.15d shows the contribution of structural effects due to the difference between the ground state minimum geometries of the molecules in vacuum and with a COSMO environment.

This causes a further redshift of the different states by 0.03 to 0.10 on average. This contribution seems to depend more on the individual molecules than type of excited state. The average shift is roughly proportional to the change in geometry between the vacuum and COSMO optimized structures. The root mean square deviations are larger for guanine (0.34 Å amu$^{1/2}$) and cytosine (0.53 Å amu$^{1/2}$) than for adenine and thymine, (<0.1Å amu$^{1/2}$).

We saw above that the largest effects are caused by electrostatic effects. To further explore this component of solvent shift, in Figure 4.16 we show how it relates to the magnitude of the vector of the difference of static dipole moment vectors in the ground and excited states at reference geometries, $\Delta\mu^{ref} = |\mu_n^{ref} - \mu_0^{ref}|$. Here we see that a sudden change in the direction of the static dipole moment is energetically unfavorable in solution. In other words, states whose electron densities are very different from those in the ground state are strongly blueshifted in solution. This is most clearly the case for nRyd states where an electron is transferred from a quite localized n-orbital to a highly delocalized Rydberg orbital.

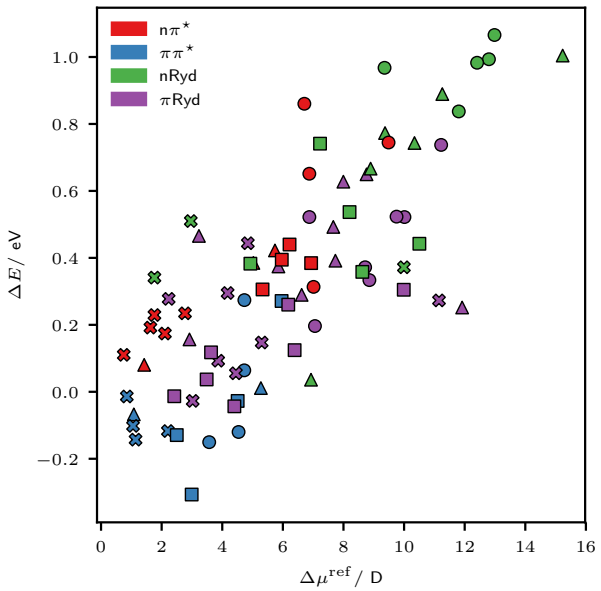


Figure 4.16: Dependence of the electrostatic component of the solvent shift (ΔE) on the magnitude of the difference of static dipole moment vectors in the ground and excited states at reference geometries, $\Delta\mu^{ref} = |\mu_n^{ref} - \mu_0^{ref}|$. The lowest 20 electronic states of each base are shown. Adenine, guanine, thymine and cytosine are denoted with crosses, triangles, squares and circles, respectively. Figure originally published in Ref. (*36*).

In the above section, we have shown how an extremely simple state tracking algorithm based on overlap matrices can be used to extract useful information about the diabatic states of a

molecule from an ensemble of nuclear geometries. For the molecules in question, each overlap calculation takes only a few seconds so their cost is completely negligible compared to the electronic structure calculations. On the other hand, in the next section we will see the limits of the information we can obtain using these simple assignment (Eq. 4.16) or weighting (Eq. 4.17) procedures when analyzing the absorption spectrum of pyrrole. In that case, these procedures can point to the presence of intensity borrowing between the states, but a full diabatization procedure is required for a quantitative interpretation of the spectrum.

4.4 Pyrrole

In this section we will explore the nuances of the photochemistry of pyrrole, primarily the wavelength dependent time scale of hydrogen detachment seen by Roberts and coworkers.[68]

We begin by choosing an appropriate level of theory for the electronic structure calculations. The two methods selected are ADC(2)/aug-cc-pVDZ and TDDFT/B3LYP/def2-TZVPD. Both methods are in good agreement with benchmark electronic structure calculations for vertical excitation energies (Table 4.3). We see that all states at both levels of theory are within a few tenths of an eV of the reference values. Oscillator strengths are also in agreement with MS-CASPT2 values reported by Roos *et al.* with the exception of the $A_1(\pi\pi^*)$ state which is an order of magnitude darker using both methods. Specifically looking at the $B_2(\pi\pi^*)$ state, we see that the vertical excitation energy obtained at both current levels of theory is in good agreement with that reported by Neville *et al.*[180] who have shown that this energy can be significantly higher than expected based on the experiment since the actual absorption spectrum is shifted to lower energies due to coupling (intensity borrowing) with the other states. Looking at the main orbital contributions for this excited state, we see a considerable difference between the two levels of theory. At the TDDFT/B3LYP/def2-TZVPD level the state is a pure $\pi\pi^*$ state. On the other hand, at the ADC(2)/aug-cc-pVDZ level the state has a mixed valence/Rydberg character with significant $3p_x$ character for the dominant particle orbital. This difference is due to the employment of the def2-TZVPD basis set and not the B3LYP functional that was used in the calculations. The augmented def2 series of basis sets contains diffuse s and d functions, but lacks diffuse p functions. The aug-cc-pVNZ (N = D, T, Q) basis sets do contain diffuse p functions, which enhance mixing of the $B_2(\pi\pi^*)$ and $B_2(\pi 3p_x)$ states. Let us note that when the ADC(2) method is used with the def2 basis sets the mixing is removed, but the vertical

excitation energies are severely over-estimated. On the other hand, B3LYP used with the aug-cc-pVTZ basis sets yields mixed states and underestimated vertical excitation energies.

Table 4.3: Vertical excitation energies (in eV) computed with the TDDFT/B3LYP/def2-TZVPD and ADC(2)/aug-cc-pVDZ methods compared with previous theoretical results and experimental data. Computed oscillator strengths are given in parentheses where available.

	$A_2(\pi\sigma^*)$	$B_1(\pi\sigma^*)$	$A_2(\pi 3p_z)$	$B_1(\pi 3p_y)$	$B_2(\pi\pi^*)$	$A_1(\pi\pi^*)$
Experiment[177,182]	5.22[a]	-	-	5.85[b]	5.90[b]	-
B3LYP	4.99	5.87	5.89	5.95	6.32	6.47
	(0.000)	(0.013)	(0.000)	(0.026)	(0.174)	(0.001)
ADC(2)	5.13	5.75	5.86	5.89	6.35	6.49
	(0.000)	(0.013)	(0.000)	(0.031)	(0.214)	(0.000)
GMS CCSD[179]	5.10	5.79	5.81	5.96	5.96	6.53
MS-CASPT2[176]	5.22	-	5.97	5.87	5.87	5.82
	(0.000)	-	(0.000)	(0.026)	(0.209)	(0.036)
CASPT2/CCSD[180]	5.06	5.86	5.87	6.00	6.24	6.01

a Band origin. *b* Band maximum.

Going beyond the vertical excitation, we look at the minima of the key states governing the photochemistry, the $^1\pi\sigma^*$ states and the $B_2(\pi\pi^*)$ state. The former states are known to be quasi-bound with respect to the N-H stretching coordinate, with a low barrier leading to the dissociation pathway.[68,159] In agreement with previous the TDDFT/PBE0 study by Barbatti *et al.*,[162] we found that TDDFT fails to accurately reproduce this feature, giving a purely dissociative PES along the N-H stretching coordinate with a mostly flat region where the minimum should be present. On the other hand, the surface along this coordinate is accurately reproduced at the ADC(2)/aug-cc-pVDZ level, with a minimum and a barrier of $E_0 = 1780$ cm^{-1} in good agreement with previous CASSCF, MRCI and CASPT2 calculations yielding barriers of 2090, 1935 and 1615 cm^{-1}, respectively.[159,161,172] Figure 4.17 shows a relaxed scan of the PES for the ground and five lowest singlet excited states. The scan was obtained by constrained optimizations of the S_1 state along the N-H stretching coordinate. The two $^1\pi\sigma^*$ states, $A_2(\pi\sigma^*)$ and $B_1(\pi\sigma^*)$ (which corresponds to S_5 at the $A_2(\pi\sigma^*)$ minimum) are dissociative with respect to the relevant coordinate while the other states are bound and strongly destabilized by the elongation of the N-H bond. Until bond lengths close to the CI are reached, all states have mostly ($> 90\%$) single excitation character and the ground state has very little multireference character ($D1 < 0.04$).

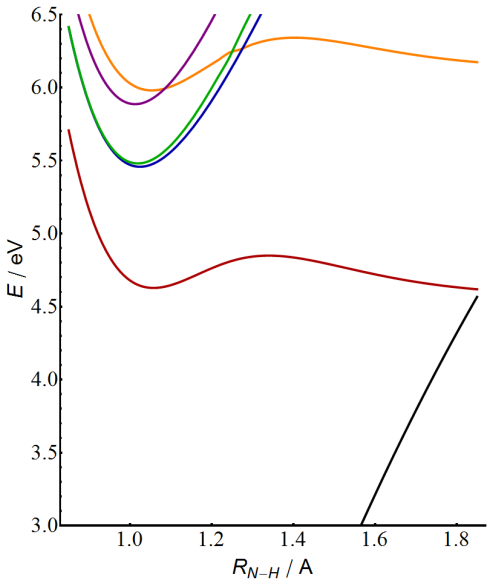


Figure 4.17: Relaxed scan of the PES along the N-H stretching coordinate optimized in the $A_2(\pi\sigma^*)$ state (red) at the ADC(2)/aug-cc-pVDZ level. The states shown are the ground state (black), $A_2(\pi\sigma^*)$ (red), $A_2(\pi 3p_z)$ (blue), $B_1(\pi 3p_y)$ (green), $B_2(\pi\pi^*)$ (purple) and $B_1(\pi\sigma^*)$ (orange). All energies are in eV and with respect to the ground state minimum energy. Figure originally published in Ref. (*37*).

Next, we focus on the geometry optimization of the bright $B_2(\pi\pi^*)$ state. At the ground state minimum geometry this is the S_5 state and optimization immediately leads to a conical intersection with the $B_1(\pi\sigma^*)$ state which is heavily destabilized as the $B_2(\pi\pi^*)$ state is relaxed. In order to simplify the procedure of finding the minimum "diabatic" $B_2(\pi\pi^*)$ state, we modify the optimization procedure to compute the wave function overlap between the $B_2(\pi\pi^*)$ state at the ground state minimum geometry and the states at each subsequent step of the optimization. Thus, at each step we continue the optimization on the adiabatic state which is characterized as the $B_2(\pi\pi^*)$ state at the current geometry. In this way, when a CI is reached during optimization, the optimization always follows the target state, unlike the standard adiabatic geometry optimization which usually reaches the first CI and fails to converge. In the present example at the TDDFT/B3LYP/def2-TZVPD level (Figure 4.18), three such CIs (with the $B_1(\pi\sigma^*)$, $B_1(\pi 3p_y)$ and $A_2(\pi 3p_z)$) states) are crossed and the algorithm converges to a minimum where $B_2(\pi\pi^*)$ state is the S_2 adiabatic state. From the dominant NTO pairs along the optimization path (Figure 4.18b), we see that the character of the state has remained virtually unchanged during the whole optimization. This optimization procedure is conceptually similar to the one developed concurrently by Sanz García *et al.*[278,279] based on tracking the target state through

the dominant NTO pair only.

The converged geometry, with the N atom and corresponding H atom out of the molecular plane in opposite directions, is in good agreement with the previously reported $B_2(\pi\pi^*)$ state minimum[247] and is located 0.7 eV below the vertical excitation energy. On the other hand, at the ADC(2)/aug-cc-pVDZ level the $B_2(\pi\pi^*)$ state has a more shallow planar minimum closer to the FC geometry at which the target state is still the S_4 state. This minimum is 0.34 eV below the vertical excitation energy and still above the p-type Rydberg states. As we will see shortly, this difference has a substantial effect on the dynamics of the system.

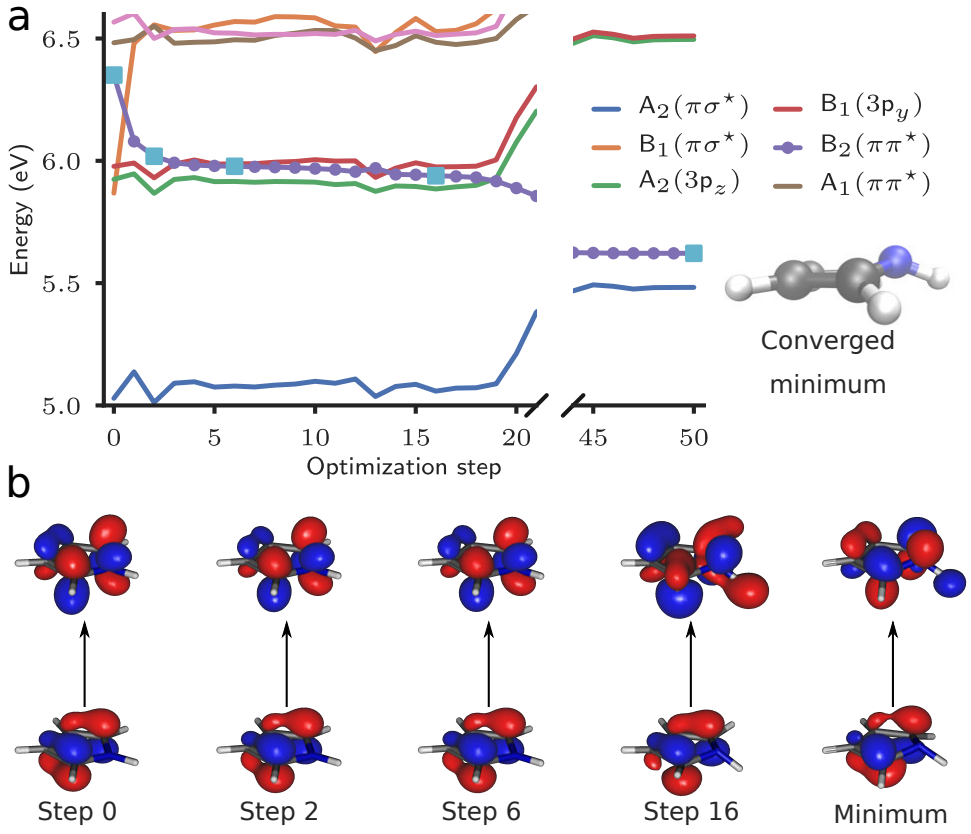


Figure 4.18: (a) Optimization of the $B_2(\pi\pi^*)$ state of pyrrole at the TDDFT/B3LYP/def2-TZVPD level of theory with three CI crossings occurring during the first 20 steps. (b) The dominant natural transition orbital pairs for the target state along the optimization path (at points marked by light blue squares). Figure originally published in Ref. (*35*).

4.4.1 UV absorption spectrum

The nuclear ensemble approach was used to calculate the UV absorption spectrum at both levels of theory. Equation 2.58 was used to generate 8000 geometries and velocities based on the har-

monic normal modes of the ground state minima at $T = 293$ K. Pyrrole is an aromatic system with a deep ground state minimum with no normal modes below 500 cm^{-1} and no highly anharmonic modes. Due to these facts the harmonic approximation is sufficiently accurate. This can be seen by comparing the distribution of potential energies in the harmonic approximation and the actual energies calculated at the generated geometries (Figure 4.19). At both levels of theory we see that the energy distribution in the harmonic approximation is virtually identical to the actual distribution of the energies for the sampled geometries. However, some outliers (less than 1% of all geometries) are present in both cases where the error of the harmonic approximation is larger than 0.5 eV. These outliers have been removed from further consideration.

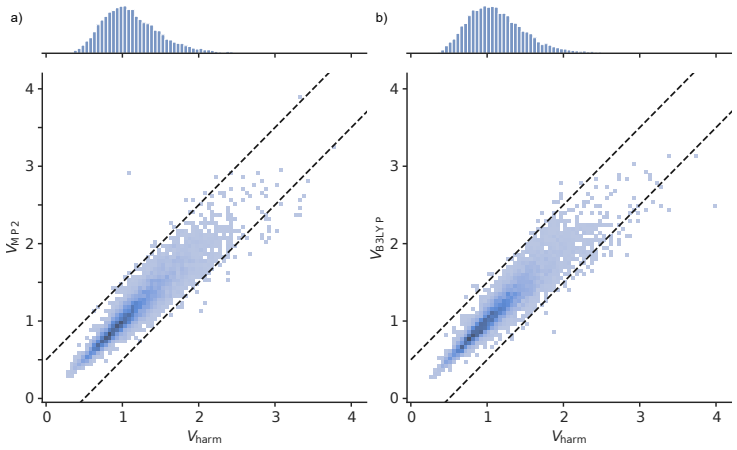


Figure 4.19: Ground state potential energies of pyrrole in the harmonic approximation and at the (a) MP2/aug-cc-pVDZ and (b) TDDFT/B3LYP/def2-TZVPD level for geometries sampled from the thermal Wigner distribution at 293 K based on normal modes computed at the respective level of theory.

At both levels of theory the UV spectrum was simulated from the generated geometries based on the energies and oscillator strengths of the ten lowest excited states and broadened by a Lorentzian function with a width of 0.1 eV. Both methods give a similar absorption spectrum, with the band center at ≈ 6.05 eV, which is slightly blueshifted with respect to the experimental spectrum. Additionally, the red tail of the calculated spectra, especially at the ADC(2) level is wider than the experimental spectrum. As is expected for the nuclear ensemble approach, the fine structure of the spectrum, including the intense and very narrow peak usually assigned to the $B_1(\pi 3p_y)$ state, was not reproduced.

By looking at the wave functions of the electronic states at each geometry we can estimate the contributions to the spectrum from each diabatic state. For this goal we use Eq. 4.17

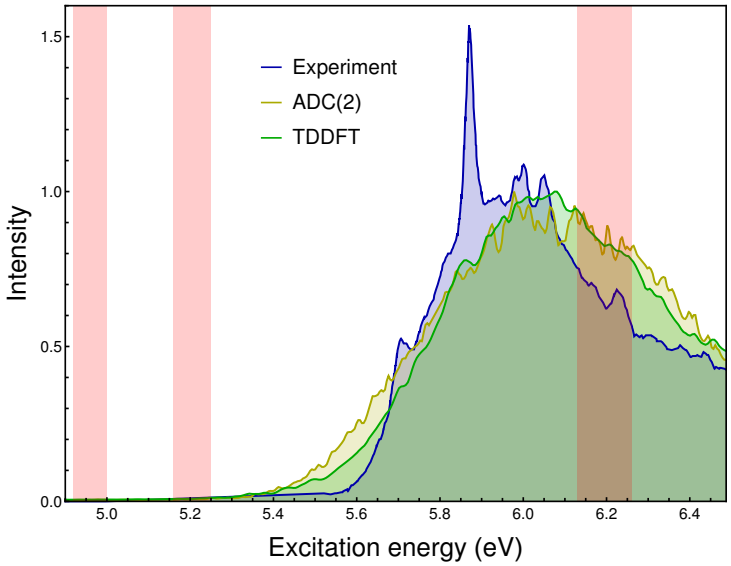


Figure 4.20: Comparison of the first absorption band of the experimental[182] UV spectrum (blue) with those obtained using the nuclear ensemble approach based ten excited states calculated at the ADC(2)/aug-cc-pVDZ (yellow) and TDDFT/B3LYP/def2-TZVPD (green) levels of theory for 8000 geometries. The shaded red areas centered at 4.96, 5.21 and 6.2 eV (250, 238 and 200 nm) with a width of 4 nm indicate windows from which initial conditions were sampled for the nonadiabatic dynamics simulations.

as described in section 4.3 since mixing between states is common. From Figure 4.21a, we can confirm that the main contribution to the first band in the spectrum of pyrrole is from the $B_2(\pi\pi^*)$ state. However, especially at lower energies, we also see significant contributions to the spectrum from the states assigned as $B_1(\pi\sigma^*)$, $A_2(\pi 3p_z)$ and $B_1(\pi 3p_y)$ which have negligible oscillator strengths at the ground state minimum geometry. Looking at these states at other geometries (Figure 4.21b and c), it can be seen that their oscillator strengths are proportional to their overlaps with the $B_2(\pi\pi^*)$ state at the ground state geometry. This indicates that their contributions to the spectrum are due to their mixing with the bright $B_2(\pi\pi^*)$ state. This is in agreement with the observation by Neville *et al.* based on quantum dynamics calculations that intensity borrowing by these states plays a significant role in shifting the first absorption band to lower energies. While only the TDDFT/B3LYP/def2-TZVPD spectrum is shown in Figure 4.21, very similar results can be seen at the ADC(2)/aug-cc-pVDZ level as well.

4.4.2 FSSH simulations

Finally, we can look at the dynamics of the system. We are interested in three excitation windows shown in Figure 4.20. The three spectral windows centered at 250, 238 and 200 nm (4.96,

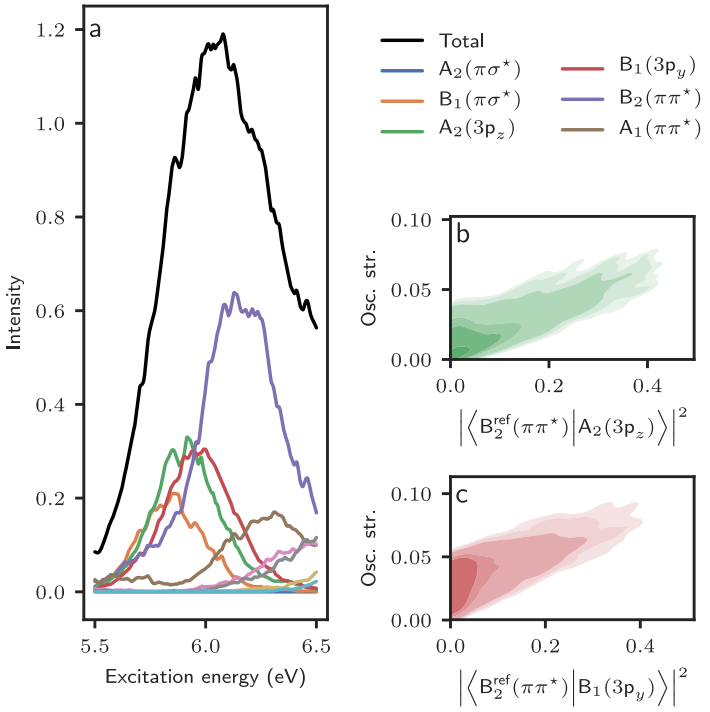


Figure 4.21: (a) Spectrum of pyrrole computed using the nuclear ensemble method and decomposed into contributions by state based on overlaps with the reference states defined at the ground state minimum geometry. (b) $A_2(\pi 3p_z)$ and (c) $B_1(\pi 3p_y)$ state distributions of oscillator strengths and their overlaps with the $B_2(\pi\pi^*)$ state at the ground state minimum geometry reference state. Figure originally published in Ref. (*35*).

5.21 and 6.2 eV) with a width of 4 nm roughly correspond to the wavelengths of the pump pulses used in the experiment of Roberts *et al.*[68] From 2000 geometries sampled from the thermal Wigner distribution, for each energy window a total of 100 initial conditions was selected based on oscillator strengths.

The first excitation window is in the region of the 0-0 transition of the $A_2(\pi\sigma^*)$ state. As stated above, TDDFT fails to reproduce this region of the PES so only results at the ADC(2)/aug-cc-pVDZ are reported for this window. All trajectories within this energy window start with very low potential energy and in order to relax to the ground state they need to cross the 1780 cm^{-1} barrier along the N-H stretching coordinate. Indeed, only 10% of the trajectories reached the ground electronic state within the simulation time of 500 fs. This confirms that relaxation in this region is primarily driven by tunneling of the hydrogen. Since nuclear quantum effects are not included in the SH approach, the relaxed PES scan along the N-H stretching coordinate (Figure 4.17) was used as a 1D model to estimate the hydrogen tunneling dynamics based on the standard WKB approach. Assuming the reduced masses of the N–H and N–D fragments

as effective mass parameters, the eigenstates of the potential were computed by diagonalizing the Hamiltonian with the Lanczos–Arnoldi integration scheme implemented in the MCTDH program package.[2] Quasi-bound states of pyrrole and pyrrole-d_1 were found at 897.9 and 795.4 cm^{-1}. The tunnelling probabilities were computed as $P = e^{2S}$ where S is the classical action integral through the barrier. The KIE was obtained as the ratio of hydrogen vs. deuterium tunnelling probabilities. The semiclassical solution yields an isotope effect of 9.9, in very good agreement with the experimentally determined value of 11.[68]

The second excitation window corresponds to excitation with ≈ 0.25 eV higher energy than the first window. This is not high enough to excite the molecule to S_2 or above, but is sufficient to cross the barrier along the N-H stretching coordinate in most cases, as seen by a dissociation yield of 86% within 500 fs. The time dependent populations of the $A_2(\pi\sigma^*)$ state are shown in Figure 4.22. The calculated time constant for the depopulation of the S_1 state of 28 fs is in good agreement with the experimentally reported value of 46 ± 22. Since a ADC(2) is a single reference method, we cannot track the dynamics beyond the S_0/S_1 CI and FSSH simulations are stopped when the energy gap between S_0 and S_1 is 0.1 eV or lower. However, based on the high velocity of the hydrogen atom in the final moments of all trajectories, we can assume that little to no bifurcation occurs at this CI and that the main products are a ground state pyrrolyl radical and a hydrogen atom. In this wavelength range, this has been demonstrated through the total kinetic energy release spectra, which do not show a signal corresponding to the formation of an excited state radical, and time-resolved photoelectron spectra, both revealing only one time constant.[68,169,173]

The third excitation window falls in the higher energy region of the first absorption band, where the $B_2(\pi\pi^*)$ state dominates the absorption. In this region FSSH simulations were performed at both the ADC(2)/aug-cc-pVDZ and TDDFT/B3LYP/def2-TZVPD levels of theory. At the former level a total relaxation yield of 82% was found within the simulation time of 500 fs. Among the trajectories which reached the ground state, 70% deactivated through the N-H bond dissociation channel while the remaining 30% deactivated through the ring deformation channel. None of the trajectories exhibited a ring opening or other pathway. The time constants obtained both for the N-H dissociation and the ring deformation pathways are higher than 150 fs. This is three times longer than the experimental value of 52 ± 12 fs, but in agreement with previous simulations in this energy region performed at the MR-CISD+Q(6,5)[161]

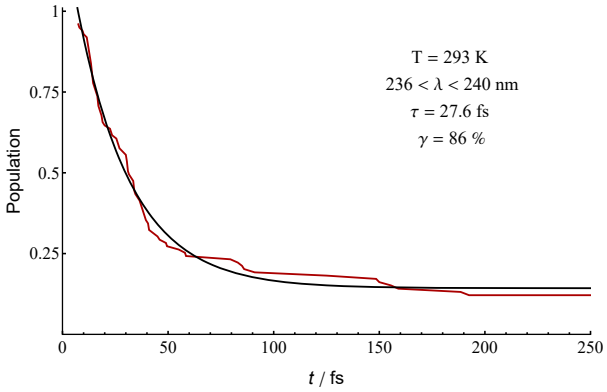


Figure 4.22: Population of the S_1 state of pyrrole during FSSH simulations at the ADC(2)/aug-cc-pVDZ level for initial conditions sampled at $T = 293$ K in the wavelength range $236 < \lambda < 240$. The black line represents a mono-exponential fit for the decay of the S1 state. Figure originally published in Ref. (*37*).

and TDDFT/PBE0[162] levels reporting time constants of 140 and 180 fs, respectively. Additional simulations were performed by considering a higher excitation window (196 – 198 nm), a larger basis set (cc-pVTZ with diffuse functions of p, d, and f type added to the center of mass of the molecule to accurately describe the Rydberg states[280]), or by employing the spin-component scaled ADC(2) (which has been reported to improve the description of Rydberg states[227]). None of these were found to have a significant effect on the results. At the TDDFT/B3LYP/def2-TZVPD level N–H bond dissociation was again found to be the dominant deactivation mechanism, while ring puckering occurred only in 6% of the cases. The time constant for the relaxation to the ground state via the N-H dissociation pathway was 48 fs, in excellent agreement with the measurment of Roberts *et al.*,[68] and in sharp contrast to the ADC(2)/aug-cc-pVDZ results. The total relaxation yield was also higher, at 98%.

The differences between the two sets of dynamics simulations from the $B_2(\pi\pi^*)$ state can be directly connected to the topography of the $B_2(\pi\pi^*)$ state at the respective electronic structure levels. A typical B3LYP/def2-TZVPD trajectory deactivating through N–H bond dissociation is shown in Figure 4.23. The simulation started from the $B_2(\pi\pi^*)$ state which is the adiabatic S_5 state. This state is intercalated in a dense manifold of excited states and two Rydberg states are found lower in energy. However, the valence and Rydberg states are separated and the initial excitation is of clear $\pi\pi^*$ character as shown by the leftmost NTO shown in Figure 4.23. During the dynamics, the system evolves through out-of-plane motions towards the minimum of the state and $B_2(\pi\pi^*)$ becomes the S_2 state. Since there is little mixing between the $B_2(\pi\pi^*)$

state and the Rydberg states no hop occurs to these states as the system evolves past them. After some time in the S_2 state, an S_1/S_2 CI is reached and a change of character occurs in a region where the $B_2(\pi\pi^*)$ and $A_2(\pi\sigma^*)$ states are heavily mixed. Once on the $A_2(\pi\sigma^*)$ surface, the S_0/S_1 CI is quickly reached through N-H bond elongation.

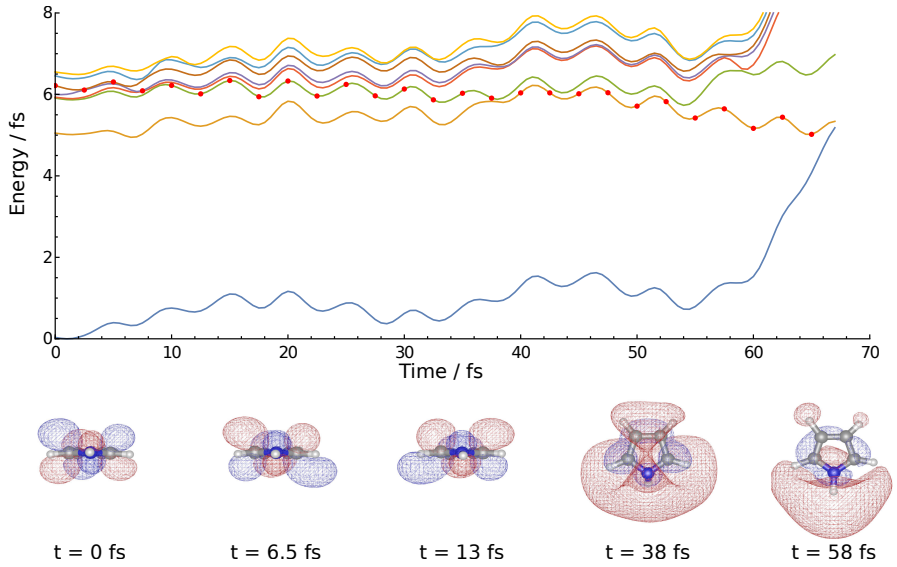


Figure 4.23: Potential energy of the ground state and nine lowest excited electronic states along a representative TDDFT/B3LYP/def2-TZVPD trajectory. Dominant particle NTOs of the currently populated state at points during the trajectory are shown below. Figure adapted from Ref. (*37*).

On the other hand, at the ADC(2)/aug-cc-pVDZ level the minimum of the $B_2(\pi\pi^*)$ state is planar and higher in energy than the two Rydberg states. A representative trajectory at this level is shown in Figure 4.24. Because of this, the system spends more time in the region where the three states are close in energy thus having more opportunities to hop between the different states. In addition to the system spending more time in the region where the states are close in energy, the couplings between the states are higher than at the TDDFT/B3LYP/def2-TZVPD level due to the high level of Rydberg-valence mixing. This can be seen in the trajectory by frequent changes of electronic character. Since there is no clear motion bringing the populated state closer to $A_2(\pi\sigma^*)$, it takes a longer time to reach the S_1/S_2 CI. However, once the CI is finally reached the relaxation mechanism is the same as in TDDFT. Of course, this picture only shows one representative trajectory from each ensemble. Some ADC(2) trajectories quickly reach the S_1/S_2 CI without hops to the Rydberg p_z and p_y state while some TDDFT trajectories do spend time in the Rydberg states.

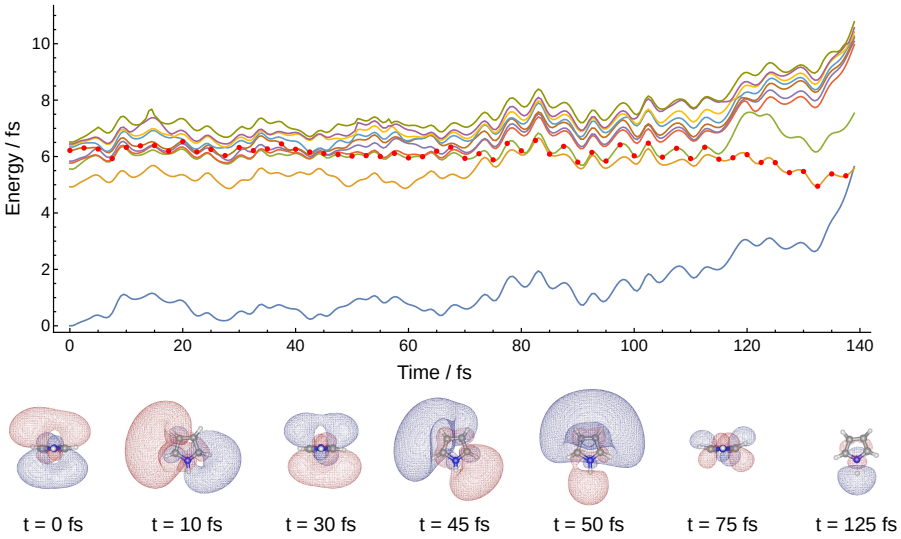


Figure 4.24: Potential energy of the ground state and nine lowest excited electronic states along a representative ADC(2)/aug-cc-pVDZ trajectory. Dominant particle NTOs of the currently populated state at points during the trajectory are shown below. Figure adapted from Ref. (*37*).

To further test the effect of the different description of Rydberg-valence mixing, FSSH dynamics simulations were performed using two additional basis sets at the TDDFT level. The first was the def2-QZVP basis set with the f and g functions removed. This basis set provides a similar description of valence states to def2-TZVPD, but due to the lack of diffuse functions, Rydberg states are destabilized so they are higher in energy than the $B_2(\pi\pi^*)$ state. Despite having two states fewer between the initially excited state and the dissociative S_1 state, the effect on the dynamics is negligible, showing that these states are not important in the overall process. On the other hand, when the aug-cc-pVTZ basis set is used mixing between valence and Rydberg states is present, similar to that at the ADC(2)/aug-cc-pVDZ. This is reflected in the dynamics, increasing the time constant of the process beyond 200 fs.

Based on these results and previous literature, we can conclude that the dynamics of pyrrole in the second excitation window at 238 nm is straightforward and basically solved at this point. The current results for the excitation at 250 nm are also in excellent agreement with the experimental findings of Roberts *et al.* providing a satisfying picture of the dynamics in this region. However we should note that in the experiment of Kirkby *et al.*[186] following excitation at 249.5 nm a sub-50 fs time scale was obtained so no trapping by the low barrier was found. This indicates that it is extremely easy to overcome the barrier and benchmark quantum dynamics calculations for both pyrrole and pyrrole-d_1 are needed to obtain a definitive quantitative picture

of the dynamics in this region.

The dynamics in the third excitation region around 200 nm is the most complex. Our results suggest that the picture given by the TDDFT/B3LYP/def2-TZVPD is correct both in terms of the topography of the $B_2(\pi\pi^*)$ state and the dynamics and that a lack of Rydberg-valence mixing is the key to this. Since these dynamics simulations were first published, more studies have looked at the dynamics of pyrrole starting from the $B_2(\pi\pi^*)$ state.[186,281,282] Kirkby *et al.*[186] explored the dynamics using the model Hamiltonian developed in ref.[180] and, in contrast to the current results, obtained a significant transient population of the $A_2(\pi 3p_z)$ state with a maximum at 35 fs, but also noted that the dynamics are highly sensitive to the addition of additional modes, namely the nitrogen out-of-plane mode. Geng *et al.* specifically looked at the effect of the Rydberg states using a quadratic vibronic coupling Hamiltonian and ML-MCTDH based on ADC(2)/aug-cc-pVDZ electronic structure calculations and also found a transient population of the $A_2(\pi 3p_z)$ state.[281] On the other hand, Heindl and González[282] performed calculations at the XMS-CASPT2(8,8)/ANO-L+ level to simulate the dynamics starting from an energy window also centered at 200 nm. By excluding the $3p_x$ orbital from their active space they ensured that the $B_2(\pi\pi^*)$ state is of pure $\pi\pi^*$ character. In their simulations 91% of the trajectories resulted in N-H dissociation (with a time constant of 64 ± 13 fs) and 9% in ring-puckering, results in very good agreement to those reported here. Due to the use of a multireference method, they could follow the dynamics through S_0/S_1 CI and calculate the total kinetic energy spectrum in which the main feature of the spectrum was in qualitative agreement with the experiment.

4.4.3 Photoionization

In a continuation of the work attempting to clarify the nature of the $B_2(\pi\pi^*)$ state, we have calculated the photoionization observables based on a Dyson orbitals approach used together with an accurate representation of the molecular continuum based on an expansion in terms of B-spline basis functions. These observables were calculated at the ground state minimum geometry at the TDDFT/B3LYP and ADC(2) levels employing both the aug-cc-pVDZ and def2-TZVPD basis sets in order to see whether the different description of the $B_2(\pi\pi^*)$ state leaves a footprint in the observable photoelectron spectrum.

At the ground state minimum geometry, the Dyson obital corresponding to ionization from the $B_2(\pi\pi^*)$ state to the ground state of the cation at the ADC(2)/aug-cc-pVDZ level is com-

posed mainly of the $3b_1(3p_x)$, $4b_1(3p_x)$ and $5b_1(\pi^*)$ virtual MOs with contributions of 79%, 7% and 5%. On the other hand, at the B3LYP/def2-TZVPD level the Dyson orbital is basically exclusively $3b_1(\pi^*)$ since the Rydberg orbitals are at higher energies. This results in an overlap of only 0.65 between the two Dyson orbitals. For comparison, the overlap between the Dyson orbitals from the $A_2(\pi\sigma^*)$ excited state of the molecule calculated using the same methods is 0.95.

The calculated cross-sections and asymmetry parameters for the $B_2(\pi\pi^*)$ excited state are presented in Figure 4.25. One can observe differences in magnitude and shape in their profiles starting from the threshold region up to about 20 eV. Cross-section profiles calculated with both electronic structure methods and the aug-cc-pVDZ basis set reach a value of about 35 Mb rapidly decreasing to the value of 2 Mb within 10 eV. On the other hand, cross-section profiles computed with TDDFT and ADC(2) and the def2-TZVPD basis set reach their maximum value at around 20 Mb with a much slower decrease to the plateau value of 2 Mb at 20 eV. This difference is expected as the slower decrease is a typical signature of a spatially more contracted orbital. For the asymmetry parameters, the most significant difference between the calculated profiles can be observed within the 15 eV.

These results in addition to the molecular frame photoelectron angular distributions which were also calculated in the same work,[283] suggest that a carefully planned photoelectron spectroscopy experiment should be able to provide definitive insight into the nature of the $B_2(\pi\pi^*)$ excitation and the degree to which valence-Rydberg mixing plays a role in it. From the theoretical side, for converged and persuasive results one would need both very large basis sets and a treatment of correlation significantly beyond the ADC(2) level.

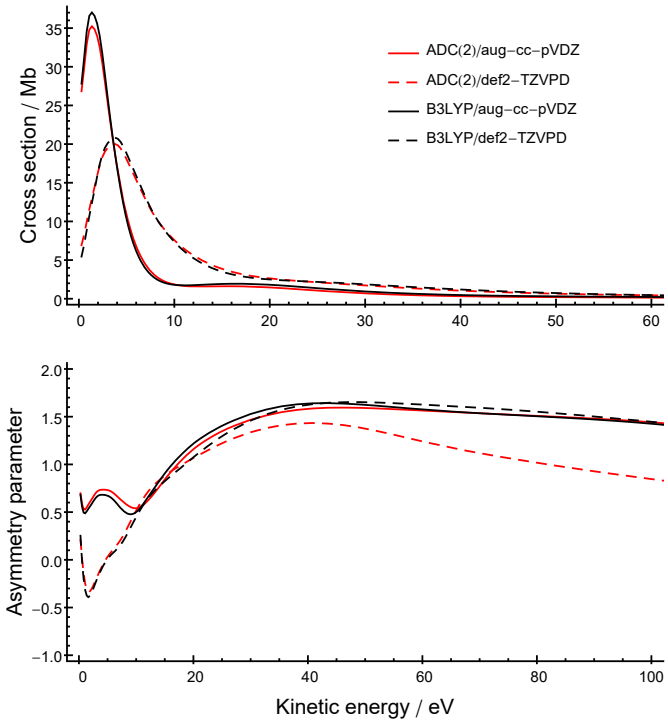


Figure 4.25: Computed partial cross-sections and asymmetry parameters for the photoionization from the first B2 excited state of neutral pyrrole to the ground state of the cation. Figure originally published in Ref. (*283*).

4.5 Pyrazine

As covered in section 2.3.2, the internal conversion process of pyrazine following excitation to the $B_{2u}(\pi\pi^*)$ state has been used for the past three decades as a testing ground for quantum dynamics methods. However, prompted by FSSH calculations of Werner *et al.*,[73,74] recent studies have shown that the process is more complex than the well-established two-state picture would suggest. This motivates us to reexamine pyrazine from a mixed quantum-classical perspective where we can take into account the full complexity of the potential energy surfaces while treating all excited states in the relevant energy region in an unbiased way.

The ADC(2)/aug-cc-pVDZ level of theory was chosen as it was shown to provide a reasonably accurate description of the excited states at a low cost. Table 4.4 shows a comparison of the vertical excitation energies (and 0-0 transitions where possible) with previously reported theoretical and experimental data. We can see that all states are within 0.3 eV of experimental values with errors comparable to those of higher level methods. This is in contrast to B3LYP/TZVP calculations used by Werner *et al.*[73] which give a significantly higher excitation energy for the

$B_{2u}(\pi\pi^*)$ state. The position of the $A_u(n\pi^*)$ state has not been conclusively verified by experiment, but most high level electronic structure calculations place it either slightly below the bright $B_{2u}(\pi\pi^*)$ or the two states are essentially degenerate.

Table 4.4: Vertical excitation energies (in eV) of pyrazine computed with the ADC(2)/aug-cc-pVDZ method compared with previous theoretical results and experimental data. 0-0 transitions are given in parentheses.

	$B_{3u}(n\pi^*)$	$A_u(n\pi^*)$	$B_{2u}(\pi\pi^*)$	$B_{2g}(n\pi^*)$
Experiment[284]	- (3.83)	-	4.81[a] (4.69)	-
ADC(2)[34]	4.18 (4.05)	4.83	5.08 (4.81)	5.85
XMCQDPT2[53]	3.93	4.45	4.79	5.38
CASPT2[285]	4.02	4.75	4.80	5.56
CC2[286]	4.26	4.95	5.13	5.92
CC3[286]	4.24	5.05	5.02	5.74
MRCISD[204]	4.55	5.16	5.52	5.91
TDDFT/B3LYP[73]	3.96	4.6	5.46	6.3

a Band maximum.

4.5.1 Potential energy surfaces

To further test the electronic structure method, one-dimensional cuts of the diabatic PESs of the three lowest excited states along the four most relevant dimensionless tuning modes were calculated and compared with the more accurate but computationally expensive multi-reference XMCQDPT2 method as a benchmark (Figure 4.26). Along each of the scans the two electronic structure methods give almost parallel potential energies with crossings at roughly the same positions. The well studied CI between the $B_{2u}(\pi\pi^*)$ and $B_{3u}(n\pi^*)$ states is reached along the Q_{6a} mode. In each of the scans we also reach a CI between the $B_{2u}(\pi\pi^*)$ and $A_u(n\pi^*)$ states, with the CI along the Q_{8a} mode closest to the equilibrium structure. Additionally, CIs between the $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ can be reached along the Q_{9a} and Q_{8a} modes. Recent quantum dynamics calculations have shown that the latter CI is responsible for the oscillations of population between the two states.[53]

Next, we will look at the shape of the excited state PESs in their full dimensionality. At the ground state minimum geometry (S_0^{min}) the $B_{3u}(n\pi^*)$ state is 0.65 eV below the other states. Optimizing the S_1 state without symmetry constraints leads to a minimum of C_i symmetry

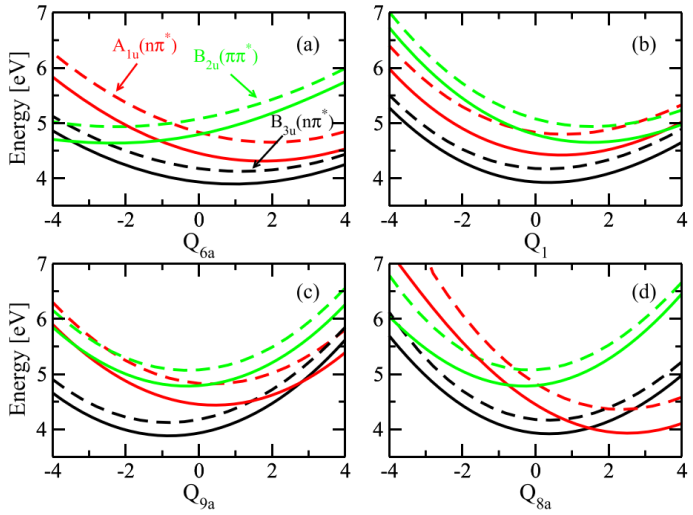


Figure 4.26: One-dimensional cuts of the potential energy surfaces of the $B_{3u}(n\pi^*)$ (black), $A_u(n\pi^*)$ (red), and $B_{2u}(\pi\pi^*)$ (green) states along the four most important totally symmetric dimensionless normal modes, Q_{6a} (a), Q_1 (b), Q_{9a} (c), Q_{8a} (d). The full and dashed lines represent the XMCQDPT2/aug-cc-pVDZ and ADC(2)/aug-cc-pVDZ results, respectively. Figure originally published in Ref. (*34*).

(S_1^{min}) 4.02 eV above the ground state minimum. There are in fact four symmetrically equivalent minima, breaking the symmetry at a D_{2h} transition state along the imaginary Q_{8b} normal mode leads to one of two C_{2h} transition states where a negative frequency for an out of plane hydrogen motion leads to the C_i symmetry minima. The path from S_0^{min} through the two transition states to one of the S_1 minima is shown in Figure 4.27a, with the corresponding motions shown in the inset. At the minimum, the S_2 state is 0.77 eV above S_1. However, when we compare the NTOs of the two states to those of $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ at the FC geometry, we see that they are significantly different (Figure 4.27b and c). The particle NTOs at S_1^{min} are rotated with respect to those at S_0^{min}. This is due to the fact that the Q_{8b} mode, along which the C_{2h} transition state was found, is of B_{3g} symmetry so it can directly couple the $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ states. We obtain a more quantiative picture by calculating the overlaps between the states at S_0^{min} and S_1^{min}. In the orthogonalized overlap matrix, we see that S_1 and S_2 are evenly mixed combinations of $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ with overlaps equal almost exactly to $\pm\frac{1}{\sqrt{2}}$.

This mixing between the two states along the Q_{8b} mode is caused by a stabilization of the π^* orbital with a node on the stretched C-N bond with respect to the π^* orbital with a node on the contracted C-N bond. This is in contrast to the situation in D_{2h} symmetry where the $1a_u$ orbital, with nodes on all four C-N bonds, is lower in energy (in benzene the two orbitals

are degenerate). The CI between S_1 and S_2 can easily be found along a path connecting two symmetrically equivalent S_1^{min} (Figure 4.27c). Since the system will naturally move toward these minima and oscillate around them during the dynamics, this heavy mixing implies that a significant population in the $A_u(n\pi^*)$ state should be expected.

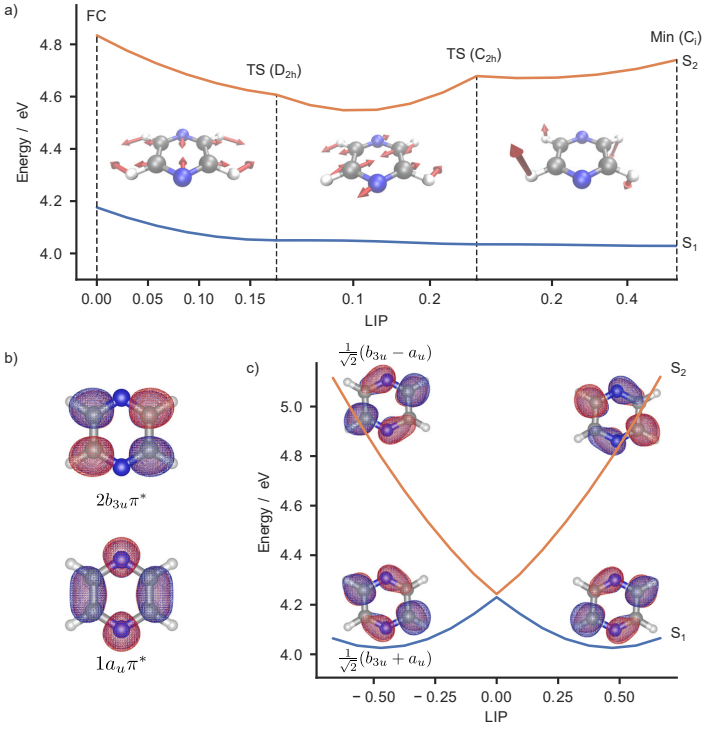


Figure 4.27: a) Energies of the S_1 and S_2 states along a path from the FC geometry (left) through transition states of D_{2h} and C_{2h} symmetry to the S_1 minimum (right). The motion corresponding to each part of the path is shown in the inset. b) Dominant particle NTOs of the $B_{3u}(n\pi^*)$ and $A_u(n\pi^*)$ states at the FC geometry. c) Linearly interpolated path between two symmetrically equivalent minima on the S_1 surface with the particle NTOs of the S_1 and S_2 states shown above the corresponding minima.

4.5.2 FSSH simulations

In section 4.1 we have already shown that the different SH algorithms perform quite well for this system when compared against exact quantum dynamics calculations on the model systems of Sala *et al.*. In addition to this, we have also shown that A-FSSH, LD-FSSH and LZSH are in quite good agreement with each other both for the model systems and for the full dimensional system. With method testing out of the way, in this section we will take the LD-FSSH results and focus on their interpretation. Because we will be comparing our dynamics to the experi-

mental results of Hori *et al.*,[202] we have further limited our selection of trajectories based on the energy shape of the pump pulse used in that experiment. From the initial set of 500 trajectories sampling was performed using a Gaussian function centered at 4.7 eV with a FWHM of 0.2 eV. All further analysis was performed for a subset of 86 trajectories chosen in this way.

We have already shown both the adiabatic and diabatic populations obtained through nonadiabatic dynamics simulations in Figure 4.4 and 4.5. However, for the diabatic picture we have used only oscillator strengths to obtain the adiabatic to diabatic transformation. A more reliable, albeit still not perfect, way to obtain diabatic populations is based on comparing the wave functions themselves. At each geometry, overlaps with the reference states at the ground state minimum geometry are calculated. This overlap matrix is then used to transform the current states into the basis of reference states. Populations obtained in this way are shown in Figure 4.28a. Qualitatively, the picture we obtain is the same as the one based on overlaps, with the $B_{2u}(\pi\pi^*)$ state quickly depopulated in favor of both $A_u(n\pi^*)$ and $B_{3u}(n\pi^*)$. After approximately 50 fs, most of the population is in these states and oscillations are seen with a period of approximately 35 fs. Looking at the actual coefficients of the currently populated state in the basis of reference states (Figure 4.28b), we see this same oscillatory behavior. However, we also see that at all time steps, states at most geometries in the ensemble are not clearly described as either $A_u(n\pi^*)$ or $B_{3u}(n\pi^*)$ state. Instead, the coefficients mostly show a mixture of the two states. This is in line with the expectation that the system is exploring the region around the S_1 minimum.

We have also calculated the photoelectron spectrum based on the nuclear ensemble approach at $t = 0$ fs (Figure 4.29) and $t = 50$ fs (Figure 4.30) at the XMS-CASPT2/aug-cc-pVDZ level. Photoelectron kinetic energies were calculated based on a probe pulse of 9.3 eV, corresponding to the pulse used by Horio *et al.*.[202] In the two figures, the upper panels show the spectra calculated from all neutral states weighted according to the coefficients of the $B_{3u}(n\pi^*)$ (a), $A_u(n\pi^*)$ (b) and $B_{2u}(\pi\pi^*)$ (c) reference states. The bottom panels (d) show the actual photoelectron spectrum as calculated based only on the currently populated state of each trajectory. In Figure 4.29d we clearly see three peaks, centered at 0.5, 2.5 and 4 eV, in very good agreement with the experiment.[202] These peaks all originate from the bright $B_{2u}(\pi\pi^*)$ state which is the only one populated at $t = 0$. The peaks at 4 and 2.5 eV correspond to ionization to D_1 and D_3 while ionization to D_0 is not allowed. On the other hand, D_0 is accessible from the $B_{3u}(n\pi^*)$

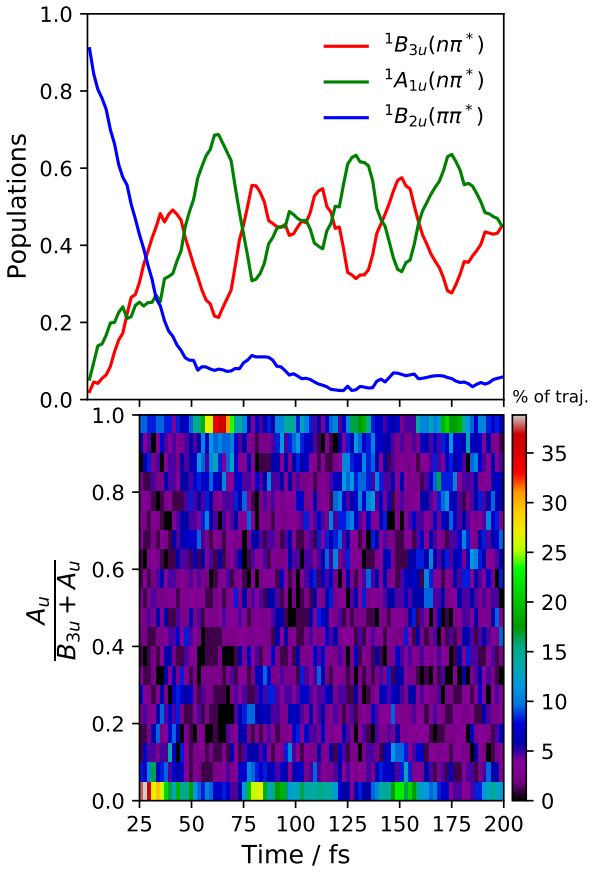


Figure 4.28: Time-dependent populations of the diabatic $B_{3u}(n\pi^*)$, $A_u(n\pi^*)$ and $B_{2u}(\pi\pi^*)$ states obtained from ADC(2)/aug-cc-pVDZ nonadiabatic dynamics simulations from a sample of 86 trajectories that satisfy the experimental excitation conditions (upper panel). Distribution of the relative $A_u(n\pi^*)$ character of the currently-populated state for steps where $n\pi^*$ character is dominant.

and $A_u(n\pi^*)$ states with peaks centered around 3.8 eV and 4.4 eV, respectively. Since neither of these states have allowed transitions to D_1-D_4, they have no signals in the 1-3.5 eV region of the photoelectron spectrum which means that the population of the $B_{2u}(\pi\pi^*)$ state can be unambiguously connected to the peak at 2.5 eV.

A significantly different picture can be seen after 50 fs when most of the population has transfered from $B_{2u}(\pi\pi^*)$ to the $n\pi^*$ states. In Figure 4.27d we see that the peak at 2.5 eV has disappeared and in the 3-5 eV region the spectrum has a single peak centered around 3.6 eV. This is again in good agreement with the experiment,[202] also reproducing the shift of the second peak to lower energies. We can also see that this spectrum can not be assigned to either the $A_u(n\pi^*)$ or the $B_{3u}(n\pi^*)$ state. The spectra corresponding to these two states are virtually identical, composed of two distinct peaks in the 3-5 eV region. This difference between the

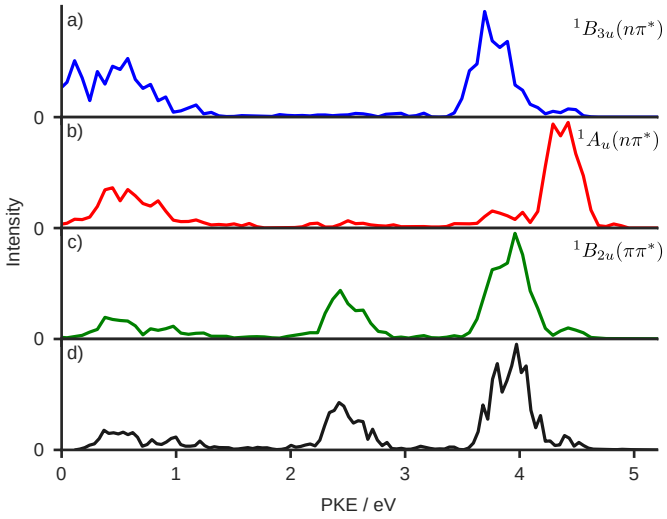


Figure 4.29: Photoionization cross-sections for the nuclear ensemble at $t = 0$ fs with (a) $B_{3u}(n\pi^*)$, (b) $A_u(n\pi^*)$, (c) $B_{2u}(\pi\pi^*)$ and (d) the currently populated state as the initial state.

two pictures can be easily explained if we refer back to Figure 4.27. Once on the S_1 surface, motion towards the minimum mixes the $A_u(n\pi^*)$ and $B_{3u}(n\pi^*)$ states so the signals seen in Figure 4.27a and b correspond to the mixed S_1 (3-4 eV region) and S_2 (4-5 eV region) states. Since almost all trajectories are moving on the S_1 surface, the peak at higher kinetic energies is never seen.

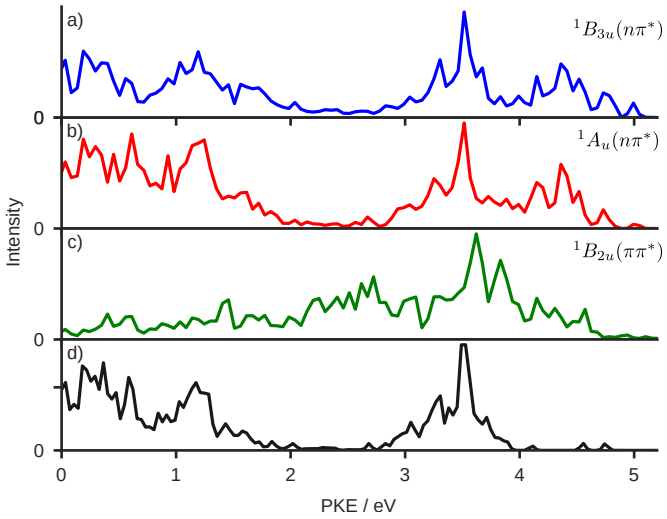


Figure 4.30: Photoionization cross-sections for the nuclear ensemble at $t = 50$ fs with (a) $B_{3u}(n\pi^*)$, (b) $A_u(n\pi^*)$, (c) $B_{2u}(\pi\pi^*)$ and (d) the currently populated state as the initial state.

The lack of any peak in the experimental spectrum above 4 eV was taken as evidence that the $A_u(n\pi^*)$ state is never populated by Mignolet *et al.*[7] since they estimated that the ionization from this state to D_0 should leave a signal around 5 eV. However, as we have shown above,

this peak would only be present before any relaxation of the nuclear ensemble (wave packet) has occurred. Upon relaxation, mixing of the two states ensures that it is no longer possible to assign the population to just one diabatic state and the experimentally observed spectrum is fully consistent with this picture.

§5. CONCLUSION

The focus of this work is the potential of the mixed classical-quantum approach for the study of photoinduced processes. The basic tools in this approach are the trajectory surface hopping algorithm for simulating the dynamics of systems in the excited state and the nuclear ensemble method for calculating observables over ensembles of nuclear configurations.

Our work began with creating a flexible and modular implementation of FSSH and LZSH algorithms including recent improvements of the algorithm for the correct treatment the trivial crossings problem. We have tested these algorithms against quantum dynamics simulations on model systems and also for the first time performed a detailed comparison of LZSH and FSSH without employing analytical models. Here we saw that the two methods are in very good agreement for simple internal conversion in pyrazine, but for the more complex reaction in pyrrole predict a time scale which is faster by a factor between 1.5 and 2. Nevertheless, we can conclude that LZSH can at least give qualitatively correct descriptions of reaction mechanisms in photochemistry. This makes it an important tool since the factor limiting the precision of nonadiabatic dynamics simulations is often the underlying electronic structure method and not the method of propagation itself. In these cases, LZSH can be used alongside more advanced electronic structure methods for which nonadiabatic couplings or wave function overlaps are not yet available or too expensive.

Next, we developed and implemented two new algorithms for the calculation of wave function overlap integrals for use alongside TDDFT and ADC(2) electronic structure calculations. Without these algorithms, these calculations would become extremely expensive for larger systems due to their high scaling. The highly unfavorable $O(n_v^2 n_o^5)$ scaling of a regular wave function overlap algorithm was reduced significantly to $O(n_o^7)$ in the OL2M or $O(N^2 n_o^5)$ in the ONTO algorithm. With these changes, the overall cost of these calculations is now negligible compared to electronic structure calculations for all systems studied with the currently available methods.

In addition to the performance benefit afforded by the new algorithms, we have focused on usability and flexibility in the implementation. This has allowed for easier interfacing with different electronic structure codes, but also for easier integration of the wave function overlap calculation into the work flow of photochemical studies. The code was connected to an opti-

mization procedure to find the minimum of a particular state of pyrrole while passing through multiple conical intersections. It was also used to find combinations of hole and particle orbitals to simplify the assignment of excited states in a wide energy window for four DNA nucleobases.

The code can also calculate related integrals. In the study of DNA nucleobases, it was used to calculate $\left\langle z^2 \right\rangle$ values of the hole and particle NTOs of various states to automatically determine their character (n, π, π^* and Rydberg). Another important application is the calculation of Dyson orbitals which are closely related to wave function overlaps (after expanding the SDs of an N electron wave function into minors the calculation reduces to a series of overlap calculations between $N-1$ electron determinants). These orbitals were used both in the pyrrole and in the pyrazine studies to compare results with the experiment.

Most importantly, wave function overlaps were used to track the electronic character of states across large ensembles of nuclear configurations simplifying the assignment of absorption and photoelectron spectra, allowing us to comment on the diabatic populations based on FSSH calculations in the adiabatic basis and to detect the level of mixing between states of different character. These tools have allowed us to reexamine some of the prototypical problems in photochemistry through a new lens.

For our study of the absorption spectra of DNA nucleobases, we started from a large number of excited states at the GS minimum geometry and tracked these states in ensembles including zero-point energy, temperature effects and solvation effects. In this way, we could directly see the small redshift of all states due to the zero-point energy and the large shifts in different directions for different types of excited states when the electrostatic effects of the solution are introduced. We found that nπ^*, nRyd and πRyd states are always blueshifted by the polar environment while $\pi\pi^*$ states are slightly redshifted with respect to their counterparts in the gas phase. These effects are quite general and connected to the magnitude of the difference of static dipole moment vectors between the ground and excited states.

In our original study of pyrrole, we also looked at the simulation and decomposition of absorption spectra using the nuclear ensemble method and wave function overlaps and were able to reproduce the general shape of the spectrum and show signs of intensity borrowing from the $B_2(\pi\pi^*)$ state to the lower lying Rydberg states as seen in benchmark quantum dynamics simulations.[180] Mainly, we were interested in how hydrogen detachment is affected by the excitation wavelength. The lowest energy window studied, at 250 nm, involves tunneling of the hydrogen

atom and is thus an example where the SH algorithm cannot provide an accurate description of the dynamics. On the other hand, the window at 238 nm excites the $A_2(\pi\sigma^*)$ state in a region significantly above the barrier and results in a very fast and direct dissociation.

The third excitation window involves excitation to the $B_2(\pi\pi^*)$ state and passage through multiple conical intersections before relaxation to the ground state is possible. In this case, the dynamics are highly sensitive to the shape of the $B_2(\pi\pi^*)$ PES at the underlying level of theory. We have performed simulations at multiple levels of theory showing that the relaxation mechanism is correctly reproduced in all cases, but the time scale of the reaction depends heavily on the level of Rydberg-valence mixing. At the TDDFT/B3LYP/def2-TZVPD level and other methods where this mixing is not present our results are in very good agreement with the experiment.[68] These results were later corroborated by FSSH simulations of Heindl and González[282] at the XMS-CASPT2(8,8)/ANO-L+ level where the $3p_x$ orbital was excluded from the active space to ensure that the $B_2(\pi\pi^*)$ state is of pure $\pi\pi^*$ character. However, recent quantum dynamics studies have suggested a significant transient population of the $A_2(\pi 3p_z)$[186,281] which is more in line with our simulations at the ADC(2)/aug-cc-pVDZ level.

When different levels of theory give a conflicting description of a system it can be difficult to decide with confidence which description is correct without direct confirmation from experiment. For this reason, we have attempted to calculate observables which would give direct insight into the nature of the $B_2(\pi\pi^*)$ state. With a high degree of Rydberg-valence mixing the state obtains a significantly more diffuse character which should be visible in photoelectron spectroscopy. Indeed, cross-section profiles calculated for the $B_2(\pi\pi^*)$ state with a significant contribution from the Rydberg $3p_x$ orbital decrease in value much faster than for the pure $\pi\pi^*$ state. These results suggest that it should be possible to experimentally determine the nature of the $B_2(\pi\pi^*)$ excitation and the degree to which valence-Rydberg mixing plays a role in it.

Pyrazine is an interesting example because it was considered for a very long time as a two state problem based on quantum dynamics simulations.[70–72,188] This view was first challenged by FSSH simulations at the TDDFT level.[73,74] However, these studies presented the populations of the S_1, S_2 and S_3 adiabatic states as directly equal to diabatic populations of the $B_{3u}(n\pi^*)$, $A_u(n\pi^*)$ and $B_{2u}(\pi\pi^*)$. In this picture, the $A_u(n\pi^*)$ state appeared as a transiently populated state before most of the population was transferred to $B_{2u}(\pi\pi^*)$. This is in contrast to the quantum dynamics studies of Sala *et al.*[53,54] in which reduced dimensionality models were

designed specifically to address the lack of $A_u(n\pi^*)$ state in the standard models of pyrazine used in quantum dynamics simulations. In our first study on pyrazine,[34] we have shown that both FSSH and LZSH simulations on the model systems of pyrazine give populations which agree quite well with those obtained through MCTDH. Additionally, we have performed SH calculations on the full dimensional system at the ADC(2)/aug-cc-pVDZ level and calculated actual diabatic populations where we have seen that the $A_u(n\pi^*)$ state is indeed populated on a longer time scale than suggested by Mitrić and coworkers.[73,74]

This difference between the adiabatic and diabatic picture was explained through an analysis of the S_1 potential energy surface showing a strong mixing between the $A_u(n\pi^*)$ and $B_{2u}(\pi\pi^*)$ states in a region around the S_1 minimum. This mixing was also found in the nuclear ensemble generated through SH calculations. Other recent studies have argued that the $A_u(n\pi^*)$ state is not populated based on the lack of a peak corresponding to this state in the experimental photoelectron spectrum.[7,204] By directly simulating slices of the time-resolved photoelectron spectra reported Horio *et al.*[202] and decomposing the spectrum based on contributions of the different states, we could show that the peak in the 3.5 eV PKE region arises both from the $B_{3u}(n\pi^*)$ state (at early times) and the mixed $B_{2u}(\pi\pi^*)$ and $A_u(n\pi^*)$ states (at later times). This mixing is only seen after relaxation of the molecule along multiple normal modes towards the S_1 minimum which shifts the peak of the $A_u(n\pi^*)$ state from higher PKEs seen in the FC region to the position seen in the experiment. In this we see a clear advantage of mixed quantum-classical methods being able to treat all degrees of freedom of the molecule on an equal footing compared to low dimensional models in which this sort of relaxation and shift of the peak is not seen.

§6. List of abbreviations

ADC	Algebraic diagrammatic construction
AIMS	Ab inition multiple spawning
CASSCF	Complete active space self consistent field
CASPT2	Complete active space second-order perturbation theory
CI	Conical intersection / Configuration interaction
CIS	Configuration interaction singles
DFT	Density functional theory
DMRG	Density matrix renormalization group
FMS	Full multiple spawning
GTO	Gaussian type orbital
HST	Hammes-Schiffer and Tully method for calculating TDCs
L1M	Level 1 minor
L2M	Level 2 minor
MCE	Multiconfiguration Ehrenfest
MCTDH	Multiconfiguration time-dependent Hartree
MD	Molecular dynamics
MPn	Møller–Plesset perturbation theory to n-th order
MRCI	Multireference configuration interaction
MRSCF	Multireference self-consistent field
NEA	Nuclear ensemble approach
NTO	Natural transition orbital
NQE	Nuclear quantum effects
NEVPT2	n-electron valence second order perturbation theory
OCIS	Overlap calculation using trivial alrgorithm
OL2M	Overlap calculation using L2M based algorithm
ONTO	Overlap calculation using NTO based algorithm
PES	Potential energy surface
PIMD	Path integral molecular dynamics
QCLE	Mixed quantum-classical Liouville equation
QD	Quantum dynamics
SD	Slater determinant
STO	Slater type orbital
SVD	Singular value decomposition
TDC	Time-derivative coupling
TDDFT	Time-dependent density functional theory
TDSE	Time-dependent Schrödinger equation
TRPES	Time-resolved photoelectron spectroscopy
UV	Ultraviolet
vMCG	Variational multiconfiguration Gaussian
XMS-CASPT2	Extended multi-state CASPT2
ZPE	Zero-point energy

§7. REFERENCES

1. H.-D. Meyer, U. Manthe, and L. S. Cederbaum, *Chem. Phys. Lett.* **165** (1990) 73–78.
2. M. Beck, *Phys. Rep.* **324** (2000) 1–105.
3. J. C. Tully, *J. Chem. Phys.* **93** (1990) 1061–1071.
4. S. Hammes-Schiffer and J. C. Tully, *J. Chem. Phys.* **101** (1994) 4657–4667.
5. M. Ben-Nun, J. Quenneville, and T. J. Martínez, *J. Phys. Chem. A* **104** (2000) 5161–5175.
6. B. F. E. Curchod and T. J. Martínez, *Chem. Rev.* **118** (2018) 3305–3336.
7. B. Mignolet, M. Kanno, N. Shimakura, S. Koseki, F. Remacle, H. Kono, and Y. Fujimura, *Chem. Phys.* **515** (2018) 704–709.
8. H. Wang and M. Thoss, *J. Chem. Phys.* **119** (2003) 1289–1299.
9. U. Manthe, *J. Chem. Phys.* **128** (2008) 164116.
10. O. Vendrell and H.-D. Meyer, *J. Chem. Phys.* **134** (2011) 044135.
11. G. A. Worth, M. A. Robb, and I. Burghardt, *Faraday Discuss.* **127** (2004) 307.
12. G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne, *Int. Rev. Phys. Chem.* **34** (2015) 269–308.
13. J. E. Subotnik, W. Ouyang, and B. R. Landry, *J. Chem. Phys.* **139** (2013) 214107.
14. J. E. Subotnik, A. Jain, B. Landry, A. Petit, W. Ouyang, and N. Bellonzi, *Annu. Rev. Phys. Chem.* **67** (2016) 387–417.
15. J. Pittner, H. Lischka, and M. Barbatti, *Chem. Phys.* **356** (2009) 147–152.
16. F. Plasser, M. Ruckenbauer, S. Mai, M. Oppel, P. Marquetand, and L. González, *J. Chem. Theory Comput.* **12** (2016) 1207–1219.
17. I. G. Ryabinkin, J. Nagesh, and A. F. Izmaylov, *J. Phys. Chem. Lett.* **6** (2015) 4200–4203.
18. G. Granucci, M. Persico, and A. Toniolo, *J. Chem. Phys.* **114** (2001) 10608–10615.
19. S. Fernandez-Alberti, A. E. Roitberg, T. Nelson, and S. Tretiak, *J. Chem. Phys.* **137** (2012) 014512.

20. G. A. Meek and B. G. Levine, *J. Phys. Chem. Lett.* **5** (2014) 2351–2356.

21. L. Wang and O. V. Prezhdo, *J. Phys. Chem. Lett.* **5** (2014) 713–719.

22. J. Qiu, X. Bai, and L. Wang, *J. Phys. Chem. Lett.* **9** (2018) 4319–4325.

23. G. Granucci, M. Persico, and A. Zoccante, *J. Chem. Phys.* **133** (2010) 134111.

24. A. Jain, E. Alguire, and J. E. Subotnik, *J. Chem. Theory Comput.* **12** (2016) 5256–5268.

25. P. Shushkov, R. Li, and J. C. Tully, *J. Chem. Phys.* **137** (2012) 22A549.

26. J. O. Richardson and M. Thoss, *J. Chem. Phys.* **139** (2013) 031102.

27. F. A. Shakib and P. Huo, *J. Phys. Chem. Lett.* **8** (2017) 3073–3080.

28. J. O. Richardson, P. Meyer, M.-O. Pleinert, and M. Thoss, *Chem. Phys.* **482** (2017) 124–134.

29. S. Ghosh, S. Giannini, K. Lively, and J. Blumberger, *Faraday Discuss.* **221** (2020) 501–525.

30. W. Arbelo-González, R. Crespo-Otero, and M. Barbatti, *J. Chem. Theory Comput.* **12** (2016) 5037–5049.

31. F. Kossoski and M. Barbatti, *J. Chem. Theory Comput.* **14** (2018) 3173–3183.

32. Š. Sršeň, J. Sita, P. Slavíček, V. Ladányi, and D. Heger, *J. Chem. Theory Comput.* **16** (2020) 6428–6438.

33. B.-X. Xue, M. Barbatti, and P. O. Dral, *J. Phys. Chem. A* **124** (2020) 7199–7210.

34. W. Xie, M. Sapunar, N. Došlić, M. Sala, and W. Domcke, *J. Chem. Phys.* **150** (2019) 154119.

35. M. Sapunar, T. Piteša, D. Davidović, and N. Došlić, *J. Chem. Theory Comput.* **15** (2019) 3461–3469.

36. M. Sapunar, W. Domcke, and N. Došlić, *Phys. Chem. Chem. Phys.* **21** (2019) 22782–22793.

37. M. Sapunar, A. Ponzi, S. Chaiwongwattana, M. Mališ, A. Prlj, P. Decleva, and N. Došlić, *Phys. Chem. Chem. Phys.* **17** (2015) 19012–19020.

38. E. Fabiano, G. Groenhof, and W. Thiel, *Chem. Phys.* **351** (2008) 111–116.

39. T. Shiozaki, *WIREs Comput. Mol. Sci.* **8** (2018) e1331.

40. A. Tajti and P. G. Szalay, *J. Chem. Phys.* **131** (2009) 124104.

41. I. Tavernelli, E. Tapavicza, and U. Rothlisberger, *J. Chem. Phys.* **130** (2009) 124107.

42. I. Tavernelli, B. F. E. Curchod, A. Laktionov, and U. Rothlisberger, *J. Chem. Phys.* **133** (2010) 194104.

43. Z. Zhou, Z. Jin, T. Qiu, A. M. Rappe, and J. E. Subotnik, *J. Chem. Theory Comput.* **16** (2020) 835–846.

44. F. Peng, M.-S. Wu, M. Sosonkina, T. Windus, J. Bentz, M. Gordon, J. Kenny, and C. Janssen, Tackling component interoperability in quantum chemistry software, 2007 Symposium on Component and framework technology in high-performance and scientific computing, New York, 2007, Conference Proceedings pp. 101–110.

45. R. M. Parrish *et al.*, *J. Chem. Theory Comput.* **13** (2017) 3185–3197.

46. D. G. A. Smith *et al.*, *J. Chem. Theory Comput.* **14** (2018) 3504–3511.

47. J. M. H. Olsen *et al.*, *J. Chem. Phys.* **152** (2020) 214115.

48. A. K. Belyaev and O. V. Lebedev, *Phys. Rev. A* **84** (2011) 014701.

49. A. K. Belyaev, C. Lasser, and G. Trigila, *J. Chem. Phys.* **140** (2014) 224108.

50. A. K. Belyaev, W. Domcke, C. Lasser, and G. Trigila, *J. Chem. Phys.* **142** (2015) 104307.

51. L. Yu, C. Xu, Y. Lei, C. Zhu, and Z. Wen, *Phys. Chem. Chem. Phys.* **16** (2014) 25883–25895.

52. L. Yue, L. Yu, C. Xu, Y. Lei, Y. Liu, and C. Zhu, *ChemPhysChem* **18** (2017) 1274–1287.

53. M. Sala, B. Lasorne, F. Gatti, and S. Guérin, *Phys. Chem. Chem. Phys.* **16** (2014) 15957.

54. M. Sala, S. Guérin, and F. Gatti, *Phys. Chem. Chem. Phys.* **17** (2015) 29518–29530.

55. M. E. Casida, in, *Recent Adv. Density Funct. Methods*, World Scientific, Singapore, 1995, pp. 155–192.

56. E. Tapavicza, I. Tavernelli, and U. Rothlisberger, *Phys. Rev. Lett.* **98** (2007) 023001.

57. J. Schirmer, *Phys. Rev. A* **26** (1982) 2395–2416.

58. J. Schirmer, *Phys. Rev. A* **43** (1991) 4647–4659.

59. F. Plasser, R. Crespo-Otero, M. Pederzoli, J. Pittner, H. Lischka, and M. Barbatti, *J. Chem. Theory Comput.* **10** (2014) 1395–1405.

60. I. Mayer, *Chem. Phys. Lett.* **437** (2007) 284–286.

61. P. R. Surján, *Chem. Phys. Lett.* **439** (2007) 393–394.

62. P. Alonso-Jordá, D. Davidović, M. Sapunar, J. R. Herrero, and E. S. Quintana-Ortí, *Comput. Phys. Commun.* **258** (2021) 107521.

63. C. M. Isborn, N. Luehr, I. S. Ufimtsev, and T. J. Martínez, *J. Chem. Theory Comput.* **7** (2011) 1814–1823.

64. L. D. Peters, J. Kussmann, and C. Ochsenfeld, *J. Chem. Theory Comput.* **15** (2019) 6647–6659.

65. H. Uratani and H. Nakai, *J. Chem. Phys.* **152** (2020) 224109.

66. A. L. Sobolewski and W. Domcke, *Chem. Phys.* **259** (2000) 181–191.

67. A. L. Sobolewski, W. Domcke, C. Dedonder-Lardeux, and C. Jouvet, *Phys. Chem. Chem. Phys.* **4** (2002) 1093–1100.

68. G. M. Roberts, C. A. Williams, H. Yu, A. S. Chatterley, J. D. Young, S. Ullrich, and V. G. Stavros, *Faraday Discuss.* **163** (2013) 95.

69. K. K. Innes, I. G. Ross, and W. R. Moomaw, *J. Mol. Spectrosc.* **132** (1988) 492–544.

70. R. Schneider and W. Domcke, *Chem. Phys. Lett.* **159** (1989) 61–65.

71. L. Seidner, G. Stock, A. L. Sobolewski, and W. Domcke, *J. Chem. Phys.* **96** (1992) 5298–5309.

72. C. Woywod, W. Domcke, A. L. Sobolewski, and H.-J. Werner, *J. Chem. Phys.* **100** (1994) 1400–1413.

73. U. Werner, R. Mitrić, T. Suzuki, and V. Bonačić-Koutecký, *Chem. Phys.* **349** (2008) 319–324.

74. G. Tomasello, A. Humeniuk, and R. Mitrić, *J. Phys. Chem. A* **118** (2014) 8437–8445.

75. S. Mai, P. Marquetand, and L. González, *J. Chem. Phys.* **140** (2014) 204302.

76. F. Plasser and H. Lischka, *J. Chem. Theory Comput.* **8** (2012) 2777–2789.

77. F. Plasser, M. Wormit, and A. Dreuw, *J. Chem. Phys.* **141** (2014) 024106.

78. F. Plasser, S. A. Bäppler, M. Wormit, and A. Dreuw, *J. Chem. Phys.* **141** (2014) 024107.

79. F. Plasser, B. Thomitzni, S. A. Bäppler, J. Wenzel, D. R. Rehn, M. Wormit, and A. Dreuw, *J. Comput. Chem.* **36** (2015) 1609–1620.

80. P. Kimber and F. Plasser, *Phys. Chem. Chem. Phys.* **22** (2020) 6058–6080.

81. F. Plasser, *J. Chem. Phys.* **152** (2020) 084108.

82. S. Mai, F. Plasser, J. Dorn, M. Fumanal, C. Daniel, and L. González, *Coord. Chem. Rev.* **361** (2018) 74–97.

83. A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, 1989.

84. D. J. Tannor, *Introduction to quantum mechanics: a time-dependent perspective*, University Science Books, Sausalito, CA, 2007.

85. W. Domcke, D. R. Yarkony, and H. Köppel, *Conical Intersections*, World Scientific, 2004, vol. 15 p. 838.

86. J. P. Malhado, M. J. Bearpark, and J. T. Hynes, *Front. Chem.* **2** (2014) 97.

87. A. Palacios, H. Bachau, and F. Martín, *Phys. Rev. A* **75** (2007) 013408.

88. M. Born and R. Oppenheimer, *Ann. Phys.* **389** (1927) 457–484.

89. M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, Clarendon Press, 1954.

90. T. J. Martínez, M. Ben-Nun, and R. D. Levine, *J. Phys. Chem.* **100** (1996) 7884–7895.

91. D. V. Shalashilin, *J. Chem. Phys.* **130** (2009) 244101.

92. D. V. Makhov, C. Symonds, S. Fernandez-Alberti, and D. V. Shalashilin, *Chem. Phys.* **493** (2017) 200–218.

93. R. Kapral and G. Ciccotti, *J. Chem. Phys.* **110** (1999) 8919–8929.

94. A. Kelly and T. E. Markland, *J. Chem. Phys.* **139** (2013) 014104.

95. A. A. Kananenka, C.-Y. Hsieh, J. Cao, and E. Geva, *J. Phys. Chem. Lett.* **7** (2016) 4809–4814.

96. J. Liu and G. Hanna, *J. Phys. Chem. Lett.* **9** (2018) 3928–3933.

97. F. de Carvalho, M. Bouduban, B. F. E. Curchod, and I. Tavernelli, *Entropy* **16** (2013) 62–85.

98. B. F. E. Curchod and I. Tavernelli, *J. Chem. Phys.* **138** (2013) 184112.

99. J. C. Tully and R. K. Preston, *J. Chem. Phys.* **55** (1971) 562–572.

100. N. Klaffki, O. Weingart, M. Garavelli, and E. Spohr, *Phys. Chem. Chem. Phys.* **14** (2012) 14299.

101. M. Barbatti and K. Sen, *Int. J. Quantum Chem.* **116** (2016) 762–771.

102. S. Mai, H. Gattuso, A. Monari, and L. González, *Front. Chem.* **6** (2018).

103. J. P. Zobel, M. Heindl, J. J. Nogueira, and L. González, *J. Chem. Theory Comput.* **14** (2018) 3205–3217.

104. J. P. Zobel, J. J. Nogueira, and L. González, *Phys. Chem. Chem. Phys.* **21** (2019) 13906–13915.

105. M. Barbatti, *J. Chem. Theory Comput.* **16** (2020) 4849–4856.

106. R. Crespo-Otero and M. Barbatti, *Theor. Chem. Acc.* **131** (2012) 1237.

107. T. Nelson, S. Fernandez-Alberti, A. E. Roitberg, and S. Tretiak, *Chem. Phys. Lett.* **590** (2013) 208–213.

108. M. F. Herman, *J. Chem. Phys.* **81** (1984) 754–763.

109. D. F. Coker and L. Xiao, *J. Chem. Phys.* **102** (1995) 496–510.

110. M. D. Hack, A. W. Jasper, Y. L. Volobuev, D. W. Schwenke, and D. G. Truhlar, *J. Phys. Chem. A* **103** (1999) 6309–6326.

111. A. Carof, S. Giannini, and J. Blumberger, *J. Chem. Phys.* **147** (2017) 214113.

112. M. Thachuk, M. Y. Ivanov, and D. M. Wardlaw, *J. Chem. Phys.* **109** (1998) 5747–5760.

113. C. Zhu, S. Nangia, A. W. Jasper, and D. G. Truhlar, *J. Chem. Phys.* **121** (2004) 7658.

114. G. Granucci and M. Persico, *J. Chem. Phys.* **126** (2007) 134114.

115. J. E. Subotnik and N. Shenvi, *J. Chem. Phys.* **134** (2011) 024105.

116. A. W. Jasper, M. D. Hack, and D. G. Truhlar, *J. Chem. Phys.* **115** (2001) 1804–1816.

117. F. Plasser, S. Mai, M. Fumanal, E. Gindensperger, C. Daniel, and L. González, *J. Chem. Theory Comput.* **15** (2019) 5031–5045.

118. C. Zhu, A. W. Jasper, and D. G. Truhlar, *J. Chem. Theory Comput.* **1** (2005) 527–540.

119. V. Chernyak and S. Mukamel, *J. Chem. Phys.* **112** (2000) 3572–3579.

120. I. Tavernelli, B. F. E. Curchod, and U. Rothlisberger, *J. Chem. Phys.* **131** (2009) 196101.

121. Q. Ou, G. D. Bellchambers, F. Furche, and J. E. Subotnik, *J. Chem. Phys.* **142** (2015) 064114.

122. Q. Ou, S. Fatehi, E. Alguire, Y. Shao, and J. E. Subotnik, *J. Chem. Phys.* **141** (2014) 024114.

123. O. Christiansen, *J. Chem. Phys.* **110** (1999) 711–723.

124. K. L. Bak, P. Jo/rgensen, H. J. A. Jensen, J. Olsen, and T. Helgaker, *J. Chem. Phys.* **97** (1992) 7573–7584.

125. I. Fdez. Galván, M. G. Delcey, T. B. Pedersen, F. Aquilante, and R. Lindh, *J. Chem. Theory Comput.* **12** (2016) 3636–3653.

126. B. H. Lengsfield, P. Saxe, and D. R. Yarkony, *J. Chem. Phys.* **81** (1984) 4549–4553.

127. H. Lischka, M. Dallos, P. G. Szalay, D. R. Yarkony, and R. Shepard, *J. Chem. Phys.* **120** (2004) 7322–7329.

128. J. W. Park and T. Shiozaki, *J. Chem. Theory Comput.* **13** (2017) 2561–2570.

129. P.-O. Löwdin, *Phys. Rev.* **97** (1955) 1474–1489.

130. W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.* **76** (1982) 637–649.

131. R. Mitrić, U. Werner, and V. Bonačić-Koutecký, *J. Chem. Phys.* **129** (2008) 164118.

132. E. Fabiano, T. W. Keal, and W. Thiel, *Chem. Phys.* **349** (2008) 334–347.

133. L. D. Landau, *Phys. Z. Sowijet.* **1** (1932) 51–88.

134. L. D. Landau, *Phys. Z. Sowijet.* **2** (1932) 46–51.

135. C. Zener, *Proc. R. Soc. London. Ser. A, Contain. Pap. a Math. Phys. Character* **137** (1932) 696–702.

136. C. Fermanian Kammerer and C. Lasser, *J. Chem. Phys.* **128** (2008) 1–10.

137. W. Xie and W. Domcke, *J. Chem. Phys.* **147** (2017) 184114.

138. D. A. Fedorov, A. O. Lykhin, and S. A. Varganov, *J. Phys. Chem. A* **122** (2018) 3480–3488.

139. J. Suchan, J. Janoš, and P. Slavíček, *J. Chem. Theory Comput.* **16** (2020) 5809–5820.

140. D. Hu, Y. Xie, X. Li, L. Li, and Z. Lan, *J. Phys. Chem. Lett.* **9** (2018) 2725–2732.

141. W.-K. Chen, X.-Y. Liu, W.-H. Fang, P. O. Dral, and G. Cui, *J. Phys. Chem. Lett.* **9** (2018) 6702–6708.

142. M. De Vetta, O. Baig, D. Steen, J. Nogueira, and L. González, *Molecules* **23** (2018) 2932.

143. M. Ceriotti, M. Parrinello, T. E. Markland, and D. E. Manolopoulos, *J. Chem. Phys.* **133** (2010) 124104.

144. T. E. Markland and M. Ceriotti, *Nat. Rev. Chem.* **2** (2018) 0109.

145. M. Ončák, L. Šištík, and P. Slavíček, *J. Chem. Phys.* **133** (2010) 174303.

146. O. Svoboda, M. Ončák, and P. Slavíček, *J. Chem. Phys.* **135** (2011) 1–17.

147. Y. K. Law and A. A. Hassanali, *J. Chem. Phys.* **148** (2018) 102331.

148. T. J. Zuehlsdorff, J. A. Napoli, J. M. Milanese, T. E. Markland, and C. M. Isborn, *J. Chem. Phys.* **149** (2018) 024107.

149. D. Hollas, O. Svoboda, and P. Slavíček, *Chem. Phys. Lett.* **622** (2015) 80–85.

150. J. Suchan, D. Hollas, B. F. E. Curchod, and P. Slavíček, *Faraday Discuss.* **212** (2018) 307–330.

151. G. A. Jones, A. Acocella, and F. Zerbetto, *J. Phys. Chem. A* **112** (2008) 9650–9656.

152. R. Mitrić, J. Petersen, and V. Bonačić-Koutecký, *Phys. Rev. A - At. Mol. Opt. Phys.* **79** (2009) 1–6.

153. I. Tavernelli, B. F. E. Curchod, and U. Rothlisberger, *Phys. Rev. A - At. Mol. Opt. Phys.* **81** (2010) 1–13.

154. M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, *J. Chem. Theory Comput.* **7** (2011) 1253–1258.

155. A. Chenu and P. Brumer, *J. Chem. Phys.* **144** (2016) 044103.

156. M. Rosenblatt, *Ann. Math. Stat.* **27** (1956) 832–837.

157. E. Parzen, *Ann. Math. Stat.* **33** (1962) 1065–1076.

158. S. Nangia, A. W. Jasper, T. F. Miller, and D. G. Truhlar, *J. Chem. Phys.* **120** (2004) 3586–3597.

159. V. Vallet, Z. Lan, S. Mahapatra, A. L. Sobolewski, and W. Domcke, *Faraday Discuss.* **127** (2004) 283–293.

160. V. Vallet, Z. Lan, S. Mahapatra, A. L. Sobolewski, and W. Domcke, *J. Chem. Phys.* **123** (2005) 144307.

161. M. Vazdar, M. Eckert-Maksić, M. Barbatti, and H. Lischka, *Mol. Phys.* **107** (2009) 845–854.

162. M. Barbatti, J. Pittner, M. Pederzoli, U. Werner, R. Mitrić, V. Bonačić-Koutecký, and H. Lischka, *Chem. Phys.* **375** (2010) 26–34.

163. S. Faraji, M. Vazdar, V. S. Reddy, M. Eckert-Maksic, H. Lischka, and H. Köppel, *J. Chem. Phys.* **135** (2011) 154310.

164. K. Saita, M. G. D. Nix, and D. V. Shalashilin, *Phys. Chem. Chem. Phys.* **15** (2013) 16227.

165. B. Cronin, A. L. Devine, M. G. D. Nix, and M. N. R. Ashfold, *Phys. Chem. Chem. Phys.* **6** (2004) 5031.

166. J. Wei, J. Riedel, A. Kuczmann, F. Renth, and F. Temps, *Faraday Discuss.* **127** (2004) 267.

167. H. Lippert, H.-H. Ritze, I. V. Hertel, and W. Radloff, *ChemPhysChem* **5** (2004) 1423–1427.

168. A. J. van den Brom, M. Kapelios, Un, T. N. Kitsopoulos, N. H. Nahler, B. Cronin, and M. N. R. Ashfold, *Phys. Chem. Chem. Phys.* **7** (2005) 892.

169. B. Cronin, A. L. Devine, M. G. D. Nix, and M. N. R. Ashfold, *Phys. Chem. Chem. Phys.* **8** (2006) 3440.

170. R. Montero, Á. Peralta Conde, V. Ovejas, M. Fernández-Fernández, F. Castaño, J. R. Vázquez de Aldana, and A. Longarte, *J. Chem. Phys.* **137** (2012) 064317.

171. R. Montero, V. Ovejas, M. Fernández-Fernández, Á. Peralta Conde, and A. Longarte, *J. Chem. Phys.* **141** (2014) 014303.

172. T. N. V. Karsili, B. Marchetti, R. Moca, and M. N. R. Ashfold, *J. Phys. Chem. A* **117** (2013) 12067–12074.

173. G. Wu, S. P. Neville, O. Schalk, T. Sekikawa, M. N. R. Ashfold, G. A. Worth, and A. Stolow, *J. Chem. Phys.* **142** (2015) 074302.

174. L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, B. O. Roos, and M. Fulscher, *J. Am. Chem. Soc.* **115** (1993) 6184–6197.

175. O. Christiansen, J. Gauss, J. F. Stanton, and P. Jo/rgensen, *J. Chem. Phys.* **111** (1999) 525–537.

176. B. O. Roos, P.-Å. Malmqvist, V. Molina, L. Serrano-Andrés, and M. Merchán, *J. Chem. Phys.* **116** (2002) 7526–7536.

177. M. H. Palmer and P. J. Wilson, *Mol. Phys.* **101** (2003) 2391–2408.

178. M. Pastore, C. Angeli, and R. Cimiraglia, *Chem. Phys. Lett.* **422** (2006) 522–528.

179. X. Li and J. Paldus, *J. Phys. Chem. A* **114** (2010) 8591–8600.

180. S. P. Neville and G. A. Worth, *J. Chem. Phys.* **140** (2014) 034317.

181. M. Barbatti, M. Vazdar, A. J. a. Aquino, M. Eckert-Maksić, and H. Lischka, *J. Chem. Phys.* **125** (2006) 164323.

182. M. H. Palmer, I. C. Walker, and M. F. Guest, *Chem. Phys.* **238** (1998) 179–199.

183. A. B. Trofimov and J. Schirmer, *Chem. Phys.* **214** (1997) 153–170.

184. I. Frank and K. Damianos, *J. Chem. Phys.* **126** (2007) 125105.

185. D. V. Makhov, K. Saita, T. J. Martínez, and D. V. Shalashilin, *Phys. Chem. Chem. Phys.* **17** (2015) 3316–3325.

186. O. M. Kirkby, M. A. Parkes, S. P. Neville, G. A. Worth, and H. H. Fielding, *Chem. Phys. Lett.* **683** (2017) 179–185.

187. R. C. Lord, A. L. Marston, and F. A. Miller, *Spectrochim. Acta* **9** (1957) 113–125.

188. G. Stock, C. Woywod, W. Domcke, T. Swinney, and B. S. Hudson, *J. Chem. Phys.* **103** (1995) 6851–6860.

189. U. Manthe and H. Köppel, *J. Chem. Phys.* **93** (1990) 1658–1669.

190. S. Krempl, M. Winterstetter, H. Plöhn, and W. Domcke, *J. Chem. Phys.* **100** (1994) 926–937.

191. G. A. Worth, H.-D. Meyer, and L. S. Cederbaum, *J. Chem. Phys.* **105** (1996) 4412–4426.

192. G. A. Worth, H.-D. Meyer, and L. S. Cederbaum, *J. Chem. Phys.* **109** (1998) 3518–3529.

193. A. Raab, G. A. Worth, H.-D. Meyer, and L. S. Cederbaum, *J. Chem. Phys.* **110** (1999) 936–946.

194. M. Thoss, W. H. Miller, and G. Stock, *J. Chem. Phys.* **112** (2000) 10282–10292.

195. P. Puzari, B. Sarkar, and S. Adhikari, *J. Chem. Phys.* **125** (2006) 194316.

196. I. Burghardt, K. Giri, and G. A. Worth, *J. Chem. Phys.* **129** (2008) 174104.

197. D. V. Shalashilin, *J. Chem. Phys.* **132** (2010) 244111.

198. I. Thanopulos, P. Brumer, and M. Shapiro, *J. Chem. Phys.* **133** (2010) 154111.

199. L. Chen, M. F. Gelin, and W. Domcke, *J. Chem. Phys.* **150** (2019) 024101.

200. V. Stert, P. Farmanara, and W. Radloff, *J. Chem. Phys.* **112** (2000) 4460–4464.

201. Y.-I. Suzuki, T. Fuji, T. Horio, and T. Suzuki, *J. Chem. Phys.* **132** (2010) 174302.

202. T. Horio, R. Spesyvtsev, K. Nagashima, R. A. Ingle, Y.-i. Suzuki, and T. Suzuki, *J. Chem. Phys.* **145** (2016) 044306.

203. C. K. Lin, Y. Niu, C. Zhu, Z. Shuai, and S. H. Lin, *Chem. - An Asian J.* **6** (2011) 2977–2985.

204. M. Kanno, Y. Ito, N. Shimakura, S. Koseki, H. Kono, and Y. Fujimura, *Phys. Chem. Chem. Phys.* **17** (2015) 2012–2024.

205. C. Woywod, A. Papp, G. J. Halász, and Á. Vibók, *Theor. Chem. Acc.* **125** (2010) 521–533.

206. T. Horio, Y. I. Suzuki, and T. Suzuki, *J. Chem. Phys.* **145** (2016).

207. *TURBOMOLE V7.0 2015, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, 2007.*

208. F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, and F. Weigend, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4** (2014) 91–100.

209. C. Hättig and F. Weigend, *J. Chem. Phys.* **113** (2000) 5154.

210. A. Köhn and C. Hättig, *J. Chem. Phys.* **119** (2003) 5021.

211. C. Hättig, in, *Adv. Quantum Chem.* Vol. 50 pp. 37–60.

212. T. H. Dunning, *J. Chem. Phys.* **90** (1989) 1007–1023.

213. R. A. Kendall, T. H. Dunning, and R. J. Harrison, *J. Chem. Phys.* **96** (1992) 6796–6806.

214. A. Hellweg, S. A. Grün, and C. Hättig, *Phys. Chem. Chem. Phys.* **10** (2008) 4119–4127.

215. A. D. Becke, *J. Chem. Phys.* **98** (1993) 5648–5652.

216. P. J. Stephens, F. J. Devlin, C. S. Ashvar, C. F. Chabalowski, and M. J. Frisch, *Faraday Discuss.* **99** (1994) 103.

217. R. Bauernschmitt and R. Ahlrichs, *Chem. Phys. Lett.* **256** (1996) 454–464.

218. D. Rappoport and F. Furche, *J. Chem. Phys.* **133** (2010) 134105.

219. A. A. Granovsky, *J. Chem. Phys.* **134** (2011) 214113.

220. F. Mertins and J. Schirmer, *Phys. Rev. A* **53** (1996) 2140–2152.

221. F. Mertins, J. Schirmer, and A. Tarantelli, *Phys. Rev. A* **53** (1996) 2153–2168.

222. R. Izsák, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **10** (2020) 1–29.

223. O. Christiansen, H. Koch, and P. Jørgensen, *Chem. Phys. Lett.* **243** (1995) 409–418.

224. D. Tuna, D. Lefrancois, Ł. Wolański, S. Gozem, I. Schapiro, T. Andruniów, A. Dreuw, and M. Olivucci, *J. Chem. Theory Comput.* **11** (2015) 5758–5781.

225. A. Dreuw and M. Wormit, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **5** (2015) 82–95.

226. P. H. P. Harbach, M. Wormit, and A. Dreuw, *J. Chem. Phys.* **141** (2014) 064113.

227. A. Tajti, L. Tulipán, and P. G. Szalay, *J. Chem. Theory Comput.* **16** (2020) 468–474.

228. B. Kozma, A. Tajti, B. Demoulin, R. Izsák, M. Nooijen, and P. G. Szalay, *J. Chem. Theory Comput.* **16** (2020) 4213–4225.

229. P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin, *J. Chem. Theory Comput.* **14** (2018) 4360–4379.

230. D. R. Rehn and A. Dreuw, *J. Chem. Phys.* **150** (2019) 174110.

231. N. O. C. Winter and C. Hättig, *J. Chem. Phys.* **134** (2011) 184101.

232. C. M. Krauter, M. Pernpointner, and A. Dreuw, *J. Chem. Phys.* **138** (2013) 044107.

233. A. Tajti and P. G. Szalay, *J. Chem. Theory Comput.* **15** (2019) 5523–5531.

234. P. Hohenberg and W. Kohn, *Phys. Rev.* **136** (1964) B864–B871.

235. W. Kohn and L. J. Sham, *Phys. Rev.* **140** (1965) A1133–A1138.

236. E. Runge and E. K. U. Gross, *Phys. Rev. Lett.* **52** (1984) 997–1000.

237. A. D. Laurent and D. Jacquemin, *Int. J. Quantum Chem.* **113** (2013) 2019–2039.

238. P.-F. Loos, A. Scemama, and D. Jacquemin, *J. Phys. Chem. Lett.* **11** (2020) 2374–2383.

239. H. S. Yu, S. L. Li, and D. G. Truhlar, *J. Chem. Phys.* **145** (2016) 130901.

240. B. G. Levine, C. Ko, J. Quenneville, and T. J. MartÍnez, *Mol. Phys.* **104** (2006) 1039–1051.

241. B. O. Roos, P. R. Taylor, and P. E. Sigbahn, *Chem. Phys.* **48** (1980) 157–173.

242. P. E. M. Siegbahn, J. Almlöf, A. Heiberg, and B. O. Roos, *J. Chem. Phys.* **74** (1981) 2384–2396.

243. K. Andersson, P. Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, *J. Phys. Chem.* **94** (1990) 5483–5488.

244. K. Andersson, P. Å. Malmqvist, and B. O. Roos, *J. Chem. Phys.* **96** (1992) 1218–1226.

245. J. Finley, P. Å. Malmqvist, B. O. Roos, and L. Serrano-Andrés, *Chem. Phys. Lett.* **288** (1998) 299–306.

246. T. Shiozaki, W. Gyroffy, P. Celani, and H.-J. Werner, *J. Chem. Phys.* **135** (2011).

247. P. Celani and H.-J. Werner, *J. Chem. Phys.* **119** (2003) 5044–5057.

248. J. W. Park and T. Shiozaki, *J. Chem. Theory Comput.* **13** (2017) 3676–3683.

249. W. Győrffy, T. Shiozaki, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **138** (2013) 104104.

250. Y. Shao *et al.*, *Mol. Phys.* **113** (2015) 184–215.

251. D. A. Matthews, L. Cheng, M. E. Harding, F. Lipparini, S. Stopkowicz, T.-c. Jagau, P. G. Szalay, J. Gauss, and J. F. Stanton, *J. Chem. Phys.* **152** (2020) 214108.

252. S. R. White, *Phys. Rev. Lett.* **69** (1992) 2863–2866.

253. C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J. P. Malrieu, *J. Chem. Phys.* **114** (2001) 10252.

254. P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka, and R. Shepard, *Chem. Rev.* **112** (2012) 108–181.

255. H. Lischka, D. Nachtigallová, A. J. A. Aquino, P. G. Szalay, F. Plasser, F. B. C. Machado, and M. Barbatti, *Chem. Rev.* **118** (2018) 7293–7361.

256. A. Baiardi and M. Reiher, *J. Chem. Phys.* **152** (2020) 040903.

257. L. González and R. Lindh, *Quantum Chemistry and Dynamics of Excited States*, Wiley, 2020.

258. H. C. Andersen, *J. Comput. Phys.* **52** (1983) 24–34.

259. A. Y. Dymarsky and K. N. Kudin, *J. Chem. Phys.* **122** (2005) 39–41.

260. H. B. Schlegel and M. J. Frisch, *Int. J. Quantum Chem.* **54** (1995) 83–87.

261. S. Obara and A. Saika, *J. Chem. Phys.* **84** (1986) 3963–3974.

262. G. Schaftenaar and J. H. Noordik, *J. Comput. Aided. Mol. Des.* **14** (2000) 123–134.

263. G. Schaftenaar, E. Vlieg, and G. Vriend, *J. Comput. Aided. Mol. Des.* **31** (2017) 789–800.

264. http://cheminf.cmbi.ru.nl/molden/molden_format.html (accessed on 06/25/2020).

265. A. Humeniuk, M. Wohlgemuth, T. Suzuki, and R. Mitrić, *J. Chem. Phys.* **139** (2013) 134104.

266. D. Toffoli, M. Stener, G. Fronzoni, and P. Decleva, *Chem. Phys.* **276** (2002) 25–43.

267. M. Brosolo, P. Decleva, and A. Lisini, *Comput. Phys. Commun.* **71** (1992) 207–214.

268. R. van Leeuwen and E. J. Baerends, *Phys. Rev. A* **49** (1994) 2421–2431.

269. G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, *J. Comput. Chem.* **22** (2001) 931–967.

270. M. Stener, S. Furlan, and P. Decleva, *J. Phys. B At. Mol. Opt. Phys.* **33** (2000) 1081–1102.

271. L. Yu, C. Xu, and C. Zhu, *Phys. Chem. Chem. Phys.* **17** (2015) 17646–17660.

272. R. L. Martin, *J. Chem. Phys.* **118** (2003) 4775–4777.

273. H. W. Kuhn, *Nav. Res. Logist. Q.* **2** (1955) 83–97.

274. J. Munkres, *J. Soc. Ind. Appl. Math.* **5** (1957) 32–38.

275. M. Barbatti, A. J. A. Aquino, and H. Lischka, *Phys. Chem. Chem. Phys.* **12** (2010) 4959.

276. C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández, and B. Kohler, *Annu. Rev. Phys. Chem.* **60** (2009) 217–239.

277. S. Bai, R. Mansour, L. Stojanović, J. M. Toldo, and M. Barbatti, *J. Mol. Model.* **26** (2020) 107.

278. J. Sanz García, M. Boggio-Pasqua, I. Ciofini, and M. Campetella, *J. Comput. Chem.* **40** (2019) 1420–1428.

279. M. Campetella and J. Sanz García, *J. Comput. Chem.* **41** (2020) 1156–1164.

280. K. Kaufmann, W. Baumeister, and M. Jungen, *J. Phys. B At. Mol. Opt. Phys.* **22** (1999) 2223–2240.

281. T. Geng, O. Schalk, S. P. Neville, T. Hansson, and R. D. Thomas, *J. Chem. Phys.* **146** (2017) 144307.

282. M. Heindl and L. González, *Comput. Theor. Chem.* **1155** (2019) 38–46.

283. A. Ponzi, M. Sapunar, N. Došlić, and P. Decleva, *Theor. Chem. Acc.* **139** (2020) 137.

284. M. Stener, P. Decleva, D. M. P. Holland, and D. A. Shaw, *J. Phys. B At. Mol. Opt. Phys.* **44** (2011) 075203.

285. M. R. Silva-Junior, M. Schreiber, S. P. A. Sauer, and W. Thiel, *J. Chem. Phys.* **133** (2010) 174318.

286. M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, *J. Chem. Phys.* **128** (2008) 134110.

§ 8. CURRICULUM VITAE

Marin Sapunar

Born: May 30th, 1990 in Split, Croatia

Education:

2012 univ. bacc. chem., Undergraduate Studies in Chemistry, Faculty of Science, University of Zagreb

2014 mag. chem., Graduate Studies in Chemistry (Inorganic and Physical Chemistry), Faculty of Science, University of Zagreb

Employment:

2014-2015 Research assistant at Ruđer Bošković Institute, UKF project "Time resolved photoelectron spectroscopy as a probe for ultrafast dynamics"

2015-2016 Research assistant at Ruđer Bošković Institute, ESF project "Networks for professional training of young scientists in interdisciplinary research of innovative surfaces and materials (MIPoMat)"

2016- Assistant, Division of Physical Chemistry, Ruđer Bošković Institute

Teaching:

2014-2018 Teaching assistant, Quantum Chemistry, Faculty of Science, University of Zagreb

2016-2017 Teaching assistant, Advanced Physical Chemistry Laboratory, Faculty of Science, University of Zagreb

Exchange visits and collaboration:

April 2016 Short stay at the International Centre for Theoretical Physics in Trieste working with Dr. Ali Hassanali.

May 2016 Short term scientific mission (through COST project XLIC) at the University of Trieste working with Prof. Piero Decleva.

April/May 2021 Short stay at the Technical University of Munich working with Prof. Wolfgang Domcke.

Publications:

Number of articles: **15**; Number of citations (Web of Science): **173**; *h*-index: **8**

(1) Kallay, N.; Preočanin, T.; Sapunar, M.; Namjesnik, D. Common Surface Potential of Two Different Crystal Planes. *Surf. Innov.* **2014**, *2* (3), 142–150.

(2) Chaiwongwattana, S.; Sapunar, M.; Ponzi, A.; Decleva, P.; Došlić, N. Exploration of Excited State Deactivation Pathways of Adenine Monohydrates. *J. Phys. Chem. A* **2015**, *119* (43), 10637–10644.

(3) Sapunar, M.; Ponzi, A.; Chaiwongwattana, S.; Mališ, M.; Prlj, A.; Decleva, P.; Došlić, N. Timescales of N–H Bond Dissociation in Pyrrole: A Nonadiabatic Dynamics Study. *Phys. Chem. Chem. Phys.* **2015**, *17* (29), 19012–19020.

(4) Thisuwan, J.; Chaiwongwattana, S.; Sapunar, M.; Sagarik, K.; Došlić, N. Photochemical Deactivation Pathways of Microsolvated Hydroxylamine. *J. Photochem. Photobiol. A Chem.* **2016**, *328*, 10–15.

(5) Ponzi, A.; Sapunar, M.; Angeli, C.; Cimiraglia, R.; Došlić, N.; Decleva, P. Photoionization of Furan from the Ground and Excited Electronic States. *J. Chem. Phys.* **2016**, *144* (8), 084307.

(6) Knežević, A.; Dokli, I.; Sapunar, M.; Šegota, S.; Baumeister, U.; Lesac, A. Induced Smectic Phase in Binary Mixtures of Twist-Bend Nematogens. *Beilstein J. Nanotechnol.* **2018**, *9* (1), 1297–1307.

(7) Knežević, A.; Sapunar, M.; Buljan, A.; Dokli, I.; Hameršak, Z.; Kontrec, D.; Lesac, A. Fine-Tuning the Effect of π–π Interactions on the Stability of the N TB Phase. *Soft Matter* **2018**, *14* (42), 8466–8474.

(8) Sapunar, M.; Ayari, T.; Došlić, N. Comparative Study of the Photodynamics of Malonaldehyde and Acetylacetone. *Chem. Phys.* **2018**, *515*, 622–627.

(9) Squibb, R. J.; Sapunar, M.; Ponzi, A.; Richter, R.; Kivimäki, A.; Plekan, O.; Finetti, P.; Sisourat, N.; Zhaunerchyk, V.; Marchenko, T.; Journel, L.; Guillemin, R.; Cucini, R.; Coreno, M.; Grazioli, C.; Di Fraia, M.; Callegari, C.; Prince, K. C.; Decleva, P.; Simon, M.; Eland, J. H. D.; Došlić, N.; Feifel, R.; Piancastelli, M. N. Acetylacetone Photodynamics at a Seeded Free-Electron Laser. *Nat. Commun.* **2018**, *9* (1), 63.

(10) Sapunar, M.; Domcke, W.; Došlić, N. UV Absorption Spectra of DNA Bases in the 350–190 Nm Range: Assignment and State Specific Analysis of Solvation Effects. *Phys. Chem. Chem. Phys.* **2019**, *21* (41), 22782–22793.

(11) Sapunar, M.; Piteša, T.; Davidović, D.; Došlić, N. Highly Efficient Algorithms for CIS Type Excited State Wave Function Overlaps. *J. Chem. Theory Comput.* **2019**, *15* (6), 3461–3469.

(12) Xie, W.; Sapunar, M.; Došlić, N.; Sala, M.; Domcke, W. Assessing the Performance of Trajectory Surface Hopping Methods: Ultrafast Internal Conversion in Pyrazine. *J. Chem. Phys.* **2019**, *150* (15), 154119.

(13) Grisanti, L.; Sapunar, M.; Hassanali, A.; Došlić, N. Toward Understanding Optical Properties of Amyloids: A Reaction Path and Nonadiabatic Dynamics Study. *J. Am. Chem. Soc.* **2020**, *142* (42), 18042–18049.

(14) Ponzi, A.; Sapunar, M.; Došlić, N.; Decleva, P. Photoionization of Pyrrole from the $$B_2$$ State: A Computational Study on the Effects of Rydberg–Valence Mixing. *Theor. Chem. Acc.* **2020**, *139* (8), 137.

(15) Alonso-Jordá, P.; Davidović, D.; Sapunar, M.; Herrero, J. R.; Quintana-Ortí, E. S. Efficient Update of Determinants for Many-Electron Wave Function Overlaps. *Comput. Phys. Commun.* **2021**, *258*, 107521.

Selected Conference Presentations:

- "Ultrafast photodynamics of pyrrole: a comparative TDDFT and ADC(2) study", Poster, Control of Chemical Reactivity: Working Group 3 Meeting of COST Action CM1204, Birmingham, United Kingdom, 14. 4. 2014.
- "Computation of many-electron wave function overlaps for photochemical studies", Lecture, Central European Symposium on Theoretical Chemistry, Srni, Czech Republic, 9. 9. 2018.
- "Tracking excited electronic states in nuclear coordinate space", Lecture, Computational Chemistry Day, Zagreb, Croatia, 11.5.2019.
- "Wave function overlap based assignment of UV absorption spectra and state specific analysis of solvation effects", Poster, Central European Symposium on Theoretical Chemistry, Burg Schlaining, 10. 9. 2019.
- "Analysis of excited electronic states at many nuclear geometries" Poster, 55th Symposium on Theoretical Chemistry, Rostock, Germany, 24. 9. 2019.

- “Excited states characterization using natural transition orbitals and wave function overlaps”, Poster, Modelling Photoinduced Processes in Molecular systems, London, United Kingdom, 6. 2. 2020.