



Universidad Autónoma  
de Madrid



FACULTAD DE  
**CIENCIAS**

**Trabajo Fin de Máster. Curso 2024 – 2025**

**Master Erasmus Mundus in**

Theoretical Chemistry and Computational  
Modelling

**Modelling the mode of action of UV solar  
filters based on natural products**

**Fernando Mata Rabadán**

Directores: Inés Corral Pérez y Jesús González-Vázquez  
Universidad Autónoma de Madrid, Facultad de Ciencias,  
Departamento de Química.



# Contents

<b>List of abbreviations</b>	<b>iii</b>
<b>Abstract</b>	<b>v</b>
<b>Resumen</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Skin and Ultra-Violet radiation . . . . .	1
1.2 Solar filters . . . . .	2
1.3 Organic absorption and Alkaloids . . . . .	3
<b>2 Objectives</b>	<b>6</b>
<b>3 Theoretical Methods</b>	<b>7</b>
3.1 The Schrödinger equation . . . . .	7
3.1.1 The Born-Oppenheimer approximation . . . . .	8
3.1.2 Antisymmetry of the wave function and Exchange . . . . .	9
3.1.3 The Hartree-Fock method . . . . .	10
3.1.4 Electron correlation . . . . .	12
3.2 Density functional theory . . . . .	13
3.2.1 Hohenberg-Kohn theorems . . . . .	13
3.2.2 Kohn-Sham formalism . . . . .	16
3.2.3 Exchange-Correlation functionals . . . . .	17
3.3 Time-Dependent Density Functional Theory . . . . .	23
3.3.1 Runge-Gross Theorem . . . . .	23
3.3.2 Linear Response Time-Dependent Density Functional Theory . . . . .	24
3.3.3 Casida Equations and Tamm-Dancoff Approximation . . . . .	26
3.3.4 Introduction of a basis set . . . . .	28
3.4 Absorption, emission and photostability . . . . .	30
3.5 Non-Adiabatic Chemistry . . . . .	32
3.6 Emission Spectra and Constant Rates . . . . .	35
<b>4 Computational Methods</b>	<b>37</b>
4.1 Conical intersection optimization . . . . .	37

4.2 Conical intersection characterization . . . . .	39
<b>5 Computational details</b>	<b>43</b>
<b>6 Results</b>	<b>44</b>
6.1 Substituent effect . . . . .	44
6.1.1 Xanthine . . . . .	44
6.1.2 Caffeine . . . . .	48
6.2 Beta-carboline . . . . .	52
6.3 Tryptamine . . . . .	55
6.4 Nicotine . . . . .	59
6.5 Cytisine . . . . .	63
<b>7 Conclusions</b>	<b>67</b>
<b>A Appendix: Computational implementation</b>	<b>73</b>
<b>B Appendix: Additional Figures and Tables</b>	<b>80</b>
B.1 Xanthine . . . . .	80
B.2 Caffeine . . . . .	81
B.3 Beta-Carboline . . . . .	88
B.4 Tryptamine . . . . .	91
B.5 Nicotine . . . . .	94
B.6 Cytisine . . . . .	96

# List of abbreviations

**CDV** Coupling Derivative Vector.

**CoIn** Conical intersection.

**DFT** Density Functional Theory.

**DGV** Difference Gradient Vector.

**FC** Franck-Condon.

**GGA** Generalized gradient approximation.

**LDA** Local density approximation.

**MECI** Minimum energy conical intersection.

**MEP** Minimum Energy Path.

**NAC** Non-Adiabatic Coupling.

**NEB** Nudged Elastic Band.

**PES** Potential Energy Surface.

**TD-DFT** Time-Dependent Density Functional Theory.

**TDA** Tamm-Danoff Approximation.

**TISE** Time-Independent Schrödinger equation.

**UBP** Updated Branching Plane.

**UV** Ultra-Violet.



# Abstract

Six alkaloids (xanthine, caffeine,  $\beta$ -carboline, tryptamine, nicotine, and cytisine) were investigated to elucidate their photophysical deactivation mechanisms and assess their potential as UV-protective agents. Fluorescence quantum yields and internal conversion rates were determined via the adiabatic Hessian approach revealing low fluorescence quantum yields ( $\leq 0.2$ ) for xanthine,  $\beta$ -carboline, nicotine, and cytisine. Semiclassical absorption spectra were generated from Wigner distributions to refine spectral band profiles and compared with available experimental data revealing generally good agreement in band structure. A program based in the Updated Branching Plane (UBP) algorithm for systematic localization of minimum-energy conical intersections (MECIs) was developed. A tool for the  $\mathcal{P}/\mathcal{B}$  characterisation of MECIs was developed to identify the topology around these critical points. All systems exhibited bifurcating MECIs. MECIs were energetically inaccessible in  $\beta$ -carboline, cytisine, caffeine, and tryptamine due to either large energetic differences between minima and MECIs or barriers in the  $S_1$  surface. Nicotine presented an accessible MECI 0.64 eV under its  $S_1$  minimum, with a 0.17 eV barrier while xanthine displayed an accessible MECI 0.16 eV below an  $S_1$  minimum with a barrier of 0.37 eV. In both cases, the total barrier energy is lower than the absorption energy, indicating efficient non-radiative decay. Possible photoproduct formation was studied by following bifurcation coordinates and re-optimizing on the  $S_0$ . Xanthine can undergo a five-membered ring opening leading to a photoproduct while caffeine can produce a similar species under UVC conditions. Nicotine and potentially xanthine are promising as natural UV filters, combining low fluorescence yields and accessible non-radiative decay pathways. The critical role of substituents suggests that functionalization is key to tune optical and topological properties.

## Resumen

Se estudiaron seis alcaloides (xantina, cafeína,  $\beta$ -carbolina, triptamina, nicotina y citisina) con el objetivo de comprender sus mecanismos de desactivación fotofísica y evaluar su potencial como agentes fotoprotectores frente a la radiación UV. Se calcularon los rendimientos cuánticos de fluorescencia y las constantes de conversión interna mediante el método del hessiano adiabático, obteniéndose valores bajos ( $\leq 0.2$ ) en xantina,  $\beta$ -carbolina, nicotina y citisina. Se generaron espectros de absorción semicásicos a partir de distribuciones de Wigner, los cuales mostraron buena concordancia con los datos experimentales disponibles. Se desarrolló un programa basado en el algoritmo "Updated Branching Plane" (UBP) para localizar sistemáticamente intersecciones cónicas de mínima energía (MECIs), junto con una herramienta para caracterizar su topología  $\mathcal{P}/\mathcal{B}$ . Todos los compuestos presentaron MECIs bifurcadas. Las MECIs fueron inaccesibles energéticamente en  $\beta$ -carbolina, citisina, cafeína y triptamina debido a grandes diferencias de energía o barreras en la superficie  $S_1$ . En cambio, la nicotina mostró una MECI accesible 0.64 eV por debajo de su mínimo en la superficie  $S_1$  con una barrera de 0.17 eV, y la xantina presentó una MECI accesible 0.16 eV debajo del mínimo  $S_1$ . Esto indica una vía eficiente de desactivación no radiativa en estos sistemas. Además, se exploró la posible formación de fotoproductos. La xantina puede experimentar una apertura del anillo de cinco miembros, mientras que la cafeína puede generar un producto similar bajo radiación UVC. Estos resultados posicionan a la nicotina y la xantina como candidatos prometedores para filtros UV naturales. El papel clave de los sustituyentes sugiere que la funcionalización es fundamental para ajustar las propiedades ópticas y topológicas de estos compuestos.

# Acknowledgements

I would like to thank the computational resources provided by the Red Española de Supercomputación (RES) and the Centro de Computación Científica (CCC) of the Universidad Autónoma de Madrid, which have been crucial for the development of this thesis.

I also want to thank Ignacio Fernandez Galván for providing me with aditional data of ethylene to benchmark my results.

I would like to thank my supervisors Inés and Jesús, who have introduced me these years to the world of computational chemistry. Jesús for the long hours he spent repeating same things (such as explaining non-adiabatic coupling for the thirteenth time) to me until I understood them, and Inés encouraging me to do my best and supporting me since meeting her at Physical Chemistry IV three years ago.

I am also deeply grateful to Daniel Escudero and the Escudero group in KU Leuven for welcoming me to their group and making me feel home away from home. Specially to Stef 1, Stef 2, and Ula, for the great times teaching mus, tasting all fries' sauces, showing me lots of places, and making sure I know that Linkin Park is back.

A heartfelt thank you to Cris, Alejandro, Aurora, Bea and Dani for being great friends and teaching me so much, putting up with my La La Land references, and helping me make the best of any situation. I would also like to thank my friends from the UAM (and Lucía), for making this journey so much more fun throughout Oviedo, Zaragoza, and all the good times in Aspet.

I am immensely thankful to my parents for all the support they have given me throughout my life, and for always believing in me. For teaching me the value of hard work, companionship, and the importance of enjoying what you do.

Finally, I would like to thank my cats Sara, Bro, and Caribe for joining me (sleeping) throughout the writing of this thesis, and for being the best company I could ask for.



# 1 Introduction

## 1.1 Skin and Ultra-Violet radiation

The skin is divided into two primary layers: the epidermis and the dermis. The epidermis is the outermost layer and serves as a barrier against external agents such as pathogens, chemical substances, and UV radiation. Within the epidermis, keratinocytes are the most abundant cell type. These cells synthesize cytokeratins, structural proteins that contribute to the skin's integrity and the tight arrangement of keratinocytes [1].

Keratinocytes also play a crucial role in the skin's response to Ultra-Violet (UV) radiation by accumulating melanin, a pigment that partially blocks the penetration of UV radiation through the skin [2]. Melanin absorbs incoming UV radiation and dissipates the energy as heat. Melanin synthesis is carried out by melanocytes, the second most abundant cell type in the epidermis. Melanin is stored in organelles called melanosomes, which are transferred from melanocytes to keratinocytes [3].

Melanocytes produce two types of melanin pigments. Eumelanin exhibits a black to brown colour, while pheomelanin is yellow to reddish brown (Figure 1). Both are derived from the amino acid tyrosine but follow different biosynthetic routes [4]. Fair-skinned individuals have less epidermal eumelanin than darker-skinned individuals. However, pheomelanin levels are similar across skin types, making eumelanin the primary determinant of skin complexion and UV protection [2].

Melosome transfer occurs within a structure known as the epidermal melanin unit, consisting of several keratinocytes and a single melanocyte. The melanocyte extends dendrites to deliver melanosomes to the surrounding keratinocytes [5]. Once received, keratinocytes incorporate the melanin granules and aggregate them over the nucleus, forming a protective barrier against UV radiation. The processes of incorporation, aggregation, and degradation of melanin by keratinocytes vary among individuals and are responsible for differences in skin colouring [6].

Excess UV exposure can produce immediate or delayed effects. Immediate effects include sunburn, pigmentation of the skin, an increase in epidermal thickness, and blister formation. Repeated injury to the skin can result in chronic conditions such as photoaging, immunosuppression, and photocarcinogenesis [7]. In order to prevent subsequent UV damage, increased amounts of melanin are synthesised and accumulated in the skin. Defects in this synthesis pathway are linked to cancer susceptibility [2].

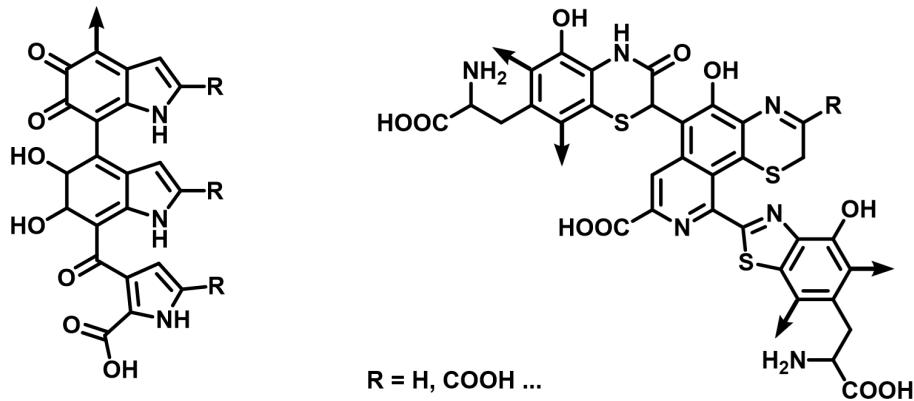


Figure 1: Molecular structure of eumelanin (left) and pheomelanin (right).

In order to understand the effect of UV radiation on the skin, it is useful to divide the UV radiation in three regions: UVA (320–400 nm or 3.10–3.94 eV), UVB (280–320 nm or 3.94–4.43 eV), and UVC (100–280 nm or 4.43–12.4 eV). While UVC is mostly absorbed by the ozone layer and is not present in Earth’s surface, UVA and UVB radiation reaches humans, and exposure to them can cause pathologies. The wavelength of UV radiation determines its penetrating power, with longer wavelengths penetrating further into the skin. Therefore, there are different mechanisms responsible for UV-induced DNA damage.

On one hand, UVA can damage DNA through an indirect mechanism by generating reactive oxygen species that react with DNA. On the other hand, UVB has been shown to induce direct damage to DNA, for example, by stimulating a photoreaction resulting in the formation of cyclobutane-pyrimidine dimers [7]. Failure to mitigate these effects can disrupt base pairing, leading to the development of skin cancer [7].

## 1.2 Solar filters

Photoprotectors play a key role in preventing health hazards resulting from UV exposure. There is a wide range of sunscreens in the cosmetic industry that block UVA and UVB radiation, but in recent years concern has risen regarding the environmental and health impact of these photoprotectors. The effectiveness of a sunscreen is measured by its Sun Protection Factor (SPF). SPF indicates how long skin covered with sunscreen can be exposed to sunlight before showing signs of damage, compared to unprotected skin. The most common commercial sunscreens have an SPF between 15 and 50, with higher SPFs providing more protection [8].

Commercial sunscreens are broadly categorised as inorganic or organic filters that reflect or absorb the incoming UV radiation. Inorganic sunscreens reflect, scatter, and can even slightly absorb UV radiation. Some of these are titanium dioxide or zinc oxide, but their use is not widespread due to the colouring they show when applied to the skin. Organic filters are widely used and are composed of molecules that absorb UV radiation, with a rapid relaxation that dissipates the energy as heat.

In recent years, there has been a large amount of research focused on the effect that these commercial sunscreens may have on human health and the environment. Accumulation of organic sunscreens such as oxybenzone has been reported in human serum, breast milk, and urine samples. Benzophenone derivatives, such as benzophenone-4, have been detected in human placental tissue as well [8]. Animal studies have reported estrogenic properties of common solar filters resulting in a significant increase in uterine weight in rats [9]. In vitro studies show that UV filters such as benzophenone-3, 4-methylbenzylidene, octyl-methoxycinnamate, and octyl-dimethyl-PABA have endocrine-disrupting capacities [10]. Inorganic solar filters have also been reported to induce airway disease and respiratory allergies in mice [11].

In addition to the health concerns, multiple studies have reported a significant environmental impact of commercial sunscreens. UV filters can enter aquatic environments through excretion from the body or washing off during swimming or bathing. Wastewater treatment plants are often unable to remove these filters from treated water, leading to their leakage into water sources, where they spread via oceanic currents [8]. The exposure of aquatic organisms to these filters has been linked to negative effects on development and reproduction [10]. An alarming result of solar filter presence is the bleaching of coral. The exposure of coral reefs to solar filters, even in small quantities, can generate bleaching and death of coral, endangering the whole reef ecosystem [12].

Due to the crucial aspect of health and environmental impact, research groups have started the search for photoprotectors based on natural molecules or inspired on compounds present in nature. Some examples include microbe-derived agents or plant-based compounds [13]. Unfortunately, studies show that sunscreens based only on natural compounds may provide lower protection than the minimum commercial sunscreens (SPF 15) [14]. Therefore, the search for new effective solar filters based on natural compounds is still ongoing.

### 1.3 Organic absorption and Alkaloids

Regarding the absorption, relaxation, and environmental requirements, the ideal solar filter should possess the following properties:

- **Absorption in the UV-visible range:** It should be able to absorb light in the UV-visible

range, specifically in the UVA and UVB regions.

- **Efficient non-radiative deactivation:** It should deactivate quickly through a process that emits no radiation back to the skin.
- **Photostability:** The solar filter should be able to retain its original properties, allowing repeated absorption of UV radiation and deactivation cycles without degradation or loss of effectiveness.
- **Environmental friendliness:** The solar filter should not have negative effects on the environment or human health. Ideally, it should be biodegradable and already present in nature.

Aromatic hydrocarbons are able to absorb in the UV as the  $\pi \rightarrow \pi^*$  and  $n \rightarrow \pi^*$  transitions are usually in this energy range. The absorption properties of aromatic rings can be modified by the introduction of substituents. The substituent effect can generate a displacement in the absorption wavelength towards higher energies (hypsochromic shift) or to lower energies (bathochromic shift). The most common type of substituents are *inductive* or *mesomeric* substituents. This classification is based on Hückel Molecular Orbital theory, where inductive refers to substituents that change the  $\alpha$  integral (stabilization due to atom), while mesomeric refers to those that modify the  $\beta$  integral (stabilization due to neighbouring atoms) [15].

Inductive substituents are substitutions in aromatic rings that alter orbital energies through inductive effect. Mesomeric substituents possess  $\pi$  orbitals, and thus can extend the aromatic system [15]. Mesomeric substituents are generally classified as  $\pi$  donors ( $+M$ ) or  $\pi$  acceptors ( $-M$ ). This extension of the pi system affects the stabilization of the molecular orbitals, altering their energy differences and thus generating a shift in absorption energies. Substitution can affect the orbital energy gap differently depending on the substitution position, even for the same type of substituent.

It is possible to appreciate in Figure 1 that melanin possesses the previously discussed heterocyclic aromatic groups and inductive and mesomeric substituents that grant it the desired absorption properties. Considering the previous properties of aromatic hydrocarbons and substituents, and the environmental effect of actual solar filters, the focus of this work is the study of compounds of a family of natural products named alkaloids.

Alkaloids are nitrogen-containing compounds with heterocyclic aromatic rings, allowing them to absorb in the UV energy range. Alkaloids are present mainly in plants, but can also be present in microorganisms and animals. There is a wide range of alkaloids, with over 21,000 found in plants. Due to the nitrogen content, they generally possess a slight basic character, facilitating their extraction and isolation [16].

In particular, this work focuses on six alkaloid compounds of different sizes, but with aromatic cores. These compounds are beta-carboline, caffeine, xanthine, tryptamine, cytisine, and nicotine (Figure 2). Caffeine and xanthine are both purine derivatives, with caffeine being a methylated derivative of xanthine. The study of these systems with the same aromatic core will be performed with the idea of gaining insight into the effect of methylation on the xanthine core. Beta-carboline and tryptamine possess a similar, but not identical, indole structure, with the latter missing a carbon atom and thus losing the pyridine aromatic cycle. Finally, cytisine and nicotine are pyridine derivatives. All of these possess different substituents, but can be considered as organic cores, due to the available positions for functionalization via electrophilic aromatic substitution [17].

This work sets up a computational study of the optical and photophysical properties of these compounds in their natural form, with the aim of determining their potential as solar filters. For this, the topology of the different pathways of de-excitation will be studied, in order to determine their properties and their stability.

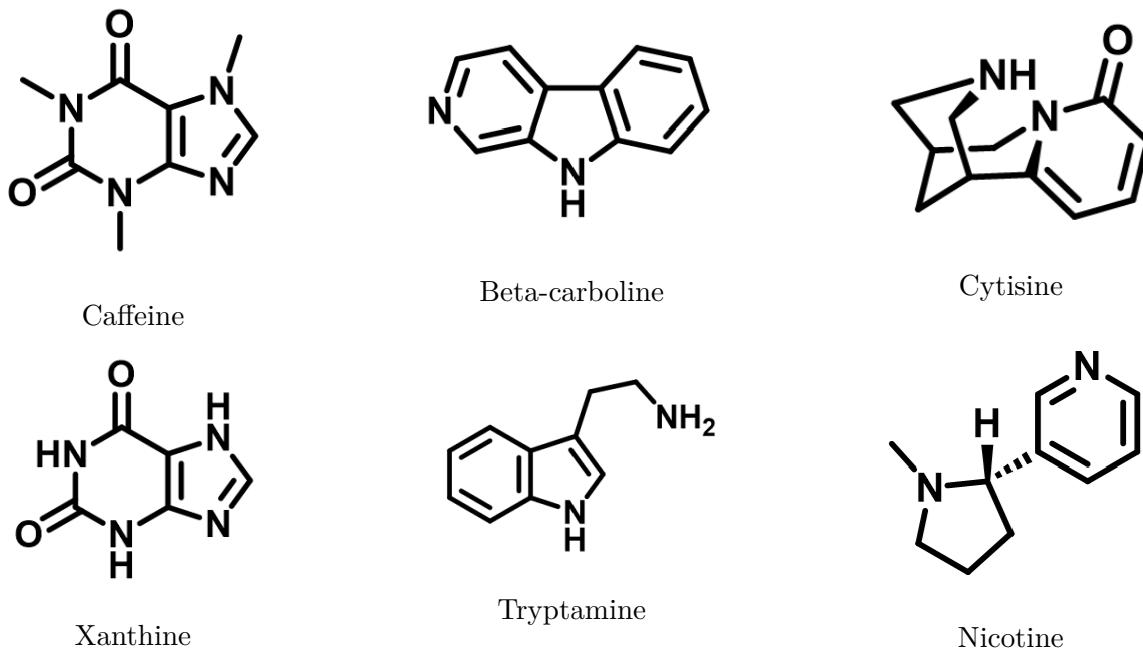


Figure 2: Structures of studied alkaloids in this work: caffeine, beta-carboline, cytisine, xanthine, tryptamine, and nicotine.

## 2 Objectives

The present work studies the excited state Potential Energy Surface (PES) of six alkaloids, with the goal of determining their potential as solar filters. The systems studied were beta-carboline, caffeine, xanthine, tryptamine, cytisine, and nicotine. The specific objectives were:

1. Study of the absorption phenomena and calculation of radiative and non-radiative constant rates to determine the efficiency of the deactivation pathways.
2. Implementation of a conical intersection optimization algorithm to study the topology of the excited state potential energy surfaces.
3. Implementation of a conical intersection characterization algorithm to study the conical intersection properties and benchmark against available literature.
4. Study of the lowest-lying excited state PES topology of the six alkaloids, searching for conical intersections and deactivation pathways. Connect the energy-accessible excited state minima with conical intersections to determine energetic barriers.
5. Study the possible formation of photoproducts.

## 3 Theoretical Methods

In the following sections, the theoretical background of the methods used in this work is presented.

Section 3.1 describes the Schrödinger equation and the Born-Oppenheimer approximation, which are the basis of quantum chemistry. This section is heavily based in the books of Szabo and Ostlund [18], Cohen-Tannoudji [19], and Jensen [20].

Section 3.2 introduces density functional theory, a formalism that allows the calculation of the ground state properties of many-electron systems using electron density instead of wave functions. This section is based on the seminal work of Hohenberg and Kohn [21], Kohn and Sham [22], and the book of Koch and Holthausen [23].

Section 3.3 presents time-dependent density functional theory, which allows the study of excited states using a density functional theory based approach and is based on the work of Runge and Gross [24], Casida [25], and the book of Ulrich [26].

Finally, Section 3.5 discusses non-adiabatic effects and rate constant calculations, and is based on the work of Fdez. Galván et al. [27], do Casal et. al. perspective [28], and Matsika's review [29].

### 3.1 The Schrödinger equation

The description of a quantum-mechanical system is governed by the Schrödinger equation:

$$\hat{H}(\mathbf{R}, \mathbf{x}, t) |\Psi(\mathbf{R}, \mathbf{x}, t)\rangle = i\hbar \frac{\partial}{\partial t} |\Psi(\mathbf{R}, \mathbf{x}, t)\rangle \quad (3.1)$$

Here,  $|\Psi(\mathbf{R}, \mathbf{x}, t)\rangle$  is the wave function of the system,  $\mathbf{R}$  the nuclear coordinates, and  $\mathbf{x}$  the spin-spatial coordinates of the electrons of the system. In the Schrödinger equation,  $\hat{H}(\mathbf{r}, t)$  is the Hamiltonian operator, which describes the total energy of the system and is defined using atomic units as:

$$\hat{H}(\mathbf{R}, \mathbf{x}, t) = \hat{T}(\mathbf{R}, \mathbf{x}) + \hat{V}(\mathbf{R}, \mathbf{x}, t) \quad (3.2)$$

here,  $\hat{T}(\mathbf{R}, \mathbf{x})$  and  $\hat{V}(\mathbf{R}, \mathbf{x}, t)$  are kinetic energy and potential energy *operators*. Linear operators are mathematical objects that act on the wave function to yield new wave functions or observables such as momentum, position, or dipole moment by applying them to the wave function:

$$\langle \hat{O} \rangle = \langle \Psi(\mathbf{R}, \mathbf{x}, t) | \hat{O}(\mathbf{R}, \mathbf{x}, t) | \Psi(\mathbf{R}, \mathbf{x}, t) \rangle \quad (3.3)$$

where  $\langle \hat{O} \rangle$  is the mean or expected value of the operator  $\hat{O}$ . Therefore, all the information of the system can be extracted from the wave function.

In the case of bound systems where the Hamiltonian does not depend on time, the solution of the Schrödinger equation is equivalent up to a phase factor. Therefore, the time dependence of the Schrödinger equation can be discarded, resulting in the Time-Independent Schrödinger equation (TISE):

$$\hat{H} |\Psi(\mathbf{R}, \mathbf{x})\rangle = E |\Psi(\mathbf{R}, \mathbf{x})\rangle \quad (3.4)$$

which is the typical eigenvalue problem in linear algebra. By solving the TISE, it is possible to obtain the stationary state (or eigenstate) wave functions. The eigenvalues associated with each eigenstate are the energies of each stationary state.

A very important theorem, the *variational principle*, states that the energy of any trial wave function  $\Psi'$  is always greater than or equal to the true ground state energy  $E_0$ :

$$E_0 \leq \frac{\langle \Psi' | \hat{H} | \Psi' \rangle}{\langle \Psi' | \Psi' \rangle} \quad (3.5)$$

This principle is a powerful tool for estimating the ground state energy. By choosing a trial wave function that depends on a set of parameters, it is possible to minimise the energy with respect to these parameters, obtaining the closest value to the true ground state energy.

### 3.1.1 The Born-Oppenheimer approximation

The Born-Oppenheimer approximation simplifies the treatment of molecular systems by decoupling the motion of nuclei and electrons. Since nuclei are much heavier than electrons, they move more slowly, allowing the nuclei to be treated as fixed points in space while solving the electronic problem. This approximation neglects non-adiabatic coupling terms that arise from the interaction between nuclear and electronic motion.

Under this approximation, the Hamiltonian is defined as:

$$\hat{H} = \sum_N T_N + \sum_e T_e + \sum_N \sum_{N'} V_{NN'} + \sum_e \sum_N V_{Ne} + \sum_e \sum_e V_{ee} \quad (3.6)$$

In the previous equation, the last four terms compose an effective Hamiltonian usually referred to as *electronic Hamiltonian*. Inside the BO approximation, the total wave function can be described as a product of the so-called electronic and nuclear wave functions, where the electronic wave function is defined for every nuclei configuration:

$$|\Psi(\mathbf{R}, \mathbf{x})\rangle_{tot} = |\Phi(\mathbf{R})\rangle_{nuc} |\psi(\mathbf{R}, \mathbf{x})\rangle_{el} \quad (3.7)$$

By treating the nuclei as fixed, the electronic wave function can be expressed parametrically with respect to the nuclear positions:

$$|\psi(\mathbf{r})\rangle_{el} = |\psi(\mathbf{R}; \mathbf{r})\rangle_{el} \quad (3.8)$$

Solving the electronic part of the TISE for each nuclear configuration yields the electronic energy as a function of nuclear coordinates, a Potential Energy Surface (PES). This surface governs the nuclear motion and provides a framework for understanding chemical processes. Since the problem to solve within the Born-Oppenheimer approximation is the electronic part, in the following sections the wave function will refer to electronic wave function.

### 3.1.2 Antisymmetry of the wave function and Exchange

The wave function of an  $N$ -particle system can be defined as the product of individual single-particle wave functions, named *orbitals*. These orbitals depend individually on the spatial and spin coordinates of each electron,  $\mathbf{x}_N \equiv \{\mathbf{r}, \sigma\}$ .

Due to electrons being fermions, the wave function must fulfil Pauli's antisymmetry principle. This principle states that two electrons can not occupy the same quantum state simultaneously. Therefore, the wave function of a system of electrons must be antisymmetric with respect to the exchange of two electrons:

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\rangle = -|\Psi(\mathbf{x}_2, \mathbf{x}_1, \dots, \mathbf{x}_N)\rangle \quad (3.9)$$

To introduce the antisymmetry of the wave function and the exchange correlation effects, a mathematical construct named the Slater determinant is used. It is defined as:

$$|\Psi_{SD}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \dots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \dots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \dots & \chi_k(\mathbf{x}_N) \end{vmatrix} \quad (3.10)$$

The antisymmetry of the wave function is guaranteed by the properties of determinants, which change sign when two rows or columns are exchanged. For the Slater determinant to be properly normalised, the orbitals must be orthonormal, i.e., the inner product of two different orbitals must be zero, and the inner product of an orbital with itself must be one:

$$\langle \chi_i | \chi_j \rangle = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (3.11)$$

### 3.1.3 The Hartree-Fock method

Considering the Slater determinant definition of the wave function, the simplest method to solve the TISE is the Hartree-Fock method. The Hartree-Fock method assumes that the electronic wave function can be described with a single determinant and relies in the variational principle (sec. 3.1) to obtain the orbitals that compose it. The so-called Hartree-Fock energy is calculated as:

$$E_{HF} = \sum_i \langle \chi_i | \hat{H}_{core} | \chi_i \rangle + \frac{1}{2} \left[ \sum_{ij} \langle \chi_i | \hat{J}_j | \chi_i \rangle - \langle \chi_i | \hat{K}_j | \chi_j \rangle \right] \quad (3.12)$$

here,  $\hat{H}_{core}$  is the core Hamiltonian that includes the kinetic energy of the orbital and the electron-nuclei attraction potential, and  $\hat{J}_{ij}$  and  $\hat{K}_{ij}$  are the Coulomb and exchange operators, respectively, defined as:

$$\hat{J}_i(\mathbf{x}_1) \chi_j(\mathbf{x}_1) = \left[ \int \chi_i^*(\mathbf{x}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \chi_i(\mathbf{x}_2) d\mathbf{x}_2 \right] \chi_j(\mathbf{x}_1) \quad (3.13)$$

and the exchange operator is defined as:

$$\hat{K}_i(\mathbf{x}_1) \chi_j(\mathbf{x}_1) = \left[ \int \chi_i^*(\mathbf{x}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \chi_j(\mathbf{x}_2) d\mathbf{x}_2 \right] \chi_i(\mathbf{x}_1) \quad (3.14)$$

Exchange is a non-local interaction between electrons, depending on the integral over all points in

space. As the orbitals are spin-dependent, and because of the orthogonality of spin functions, this interaction vanishes between electrons of different spins. The exchange interaction is responsible for some stabilisation of the system, as this contribution is negative and decreases the energy.

From this definition of the exchange and Coulomb interactions, after performing one of the integrations, both  $\mathbf{J}_{ij}$  and  $\mathbf{K}_{ij}$  depend only on one set of coordinates, as the other has been averaged out. The inner integrals define an effective potential acting on orbital  $\chi_i(\mathbf{x}_1)$ . This is the *mean-field approximation*, which implies that the Coulomb and exchange energies are defined as effective one-electron properties interacting with a mean field of an electron cloud.

From the variational principle, the  $E_{HF}$  of any trial set of orbitals is always greater than or equal to the true ground state energy. Due to the construction of the wave function as a Slater determinant, it is necessary that orbitals remain orthonormal in order for the total wave function to be normalised. To perform the minimisation of the energy with respect to the orbitals, while maintaining the orthogonality constraint, Lagrange multipliers are used.

The Lagrangian is defined as:

$$\mathcal{L}[\{\chi\}] = E_{HF}[\{\chi\}] - \sum_{ij} \lambda_{ij} (\langle \chi_i | \chi_j \rangle - \delta_{ij}) \quad (3.15)$$

where the second term is the orthonormality constraint and  $\lambda_{ij}$  are the undefined Lagrange multipliers. The minimum of this Lagrangian function must satisfy:

$$\frac{\delta \mathcal{L}[\{\chi\}]}{\delta \chi_i^*} = 0 \quad (3.16)$$

Taking the derivative with respect to  $\chi_i^*$  and equating to zero yields:

$$\left[ \hat{H}_{core} + \sum_j \hat{J}_j - \hat{K}_j \right] |\chi_i\rangle = \lambda_i |\chi_i\rangle \quad (3.17)$$

where the term in brackets is the Fock operator  $\hat{F}_j$ , which is a one-electron operator, and the Lagrange multiplier  $\lambda_i$  is the energy associated with orbital  $\chi_i$  by solving this equation. This equation appears to be a typical eigenvalue equation, but it is actually a pseudo-eigenvalue equation as the operators  $\hat{J}_j$  and  $\hat{K}_j$  depend on the orbitals  $\{\chi\}$ .

The solution of the Hartree-Fock equations is an iterative process known as the *self-consistent field*

(SCF) method. Starting with an initial guess for the wave function, the Fock operator is constructed, and the Hartree-Fock equations are solved to generate a new set of orbitals. This process is repeated until the orbitals used to define the Fock operator are the same as those obtained in the current step, achieving *self-consistency*.

### 3.1.4 Electron correlation

The simplest way to solve the TISE is the previously presented Hartree-Fock method. The description of electron-electron interaction was performed in a mean-field way, where each electron interacts with the average field of all other electrons. While this is a good approximation for weakly correlated systems i.e. systems where the electron-electron interaction is weak, it fails to describe strongly correlated systems where the electron-electron interaction is significant, and the mean-field approximation breaks down. Additionally, the Hartree-Fock method neglects instantaneous electron-electron interactions, which are crucial in many systems. This results in an incomplete description of the wave function and, consequently, an incorrect description of the system's energy.

The missing contribution from these instantaneous interactions is referred to as *electron correlation*. The energy associated with electron correlation effects is called the *correlation energy*, and it is defined as:

$$E_{corr} = E_0 - E_{HF} \quad (3.18)$$

where  $E_0$  is the exact ground state energy of the system, and  $E_{HF}$  is the Hartree-Fock energy. Although the correlation energy is typically small compared to the total Hartree-Fock energy, it often conveys essential information about the electronic structure of the system, which is necessary for accurately describing chemical processes.

To include electron correlation effects, it is necessary to employ either perturbation theory or more advanced "post Hartree-Fock" wave-function based methods, such as coupled cluster, configuration interaction, or multireference approaches. Wave-function based methods describe the total wave function as a linear combination of multiple Slater determinants, optimizing the weights of these determinants variationally to achieve the lowest possible energy. By including additional determinants, these methods provide a more precise description of electron correlation. However, this increased accuracy comes at a cost, as the complexity of the wave function grows significantly, leading to computational demands that often scale beyond practical limits for time and resources.

This practical limitation is what leads to the use of density functional theory-based approaches,

where the treatment of electron correlation can be described exactly (in theory) for any system, while dealing only with the three-dimensional electron density rather than the  $4N$ -dimensional wave function.

### 3.2 Density functional theory

The core idea of Density Functional Theory (DFT) is that if a relation between the electron density and the wave function could be established, and the electron density could be obtained exactly, then the wave function would not need to be calculated explicitly. This idea eliminates the use of wave functions that depend on 4 variables per electron, in favour of a three-variable object, the electron density  $\rho(\mathbf{r})$ . Even without the explicit description of the wave function, all properties of the system could be formally and exactly obtained from the electron density. This reduction in complexity would, in theory, allow any system to be treated exactly, as any property depends only on the density.

Before describing the fundamental theorems of DFT, it is important to define the electron density itself and its properties. The electron density  $\rho(\mathbf{r})$  is a function that depends on the position in space and yields a scalar value representing the probability of finding an electron at those spatial coordinates. It is obtained from the wave function by integrating all spin and spatial coordinates except one:

$$\rho(\mathbf{r}) = N \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\sigma_1 d\mathbf{x}_2 \dots d\mathbf{x}_N \quad (3.19)$$

The electron density must follow some constraints. It is positive-valued and vanishes at infinity. Its integral over all space is normalised and yields the number of electrons  $N$ .

#### 3.2.1 Hohenberg-Kohn theorems

Hohenberg and Kohn set the foundation of DFT in 1964 by proving two theorems that establish the relationship between the electron density and the ground state energy of a many-electron system [21]. The connection between the wave function and the electron density can be demonstrated in the following way. Recalling the definition of the TISE:

$$\hat{H} |\Psi(\mathbf{r})\rangle = E |\Psi(\mathbf{r})\rangle \quad \text{where} \quad \hat{H} = \hat{T} + \hat{V} + \hat{U} \quad (3.20)$$

where the term  $\hat{T}$  is the electron kinetic energy,  $\hat{V}$  the electron-electron interaction and  $\hat{U}$  the

electron-electron exchange and correlation interactions. Spin dependency has been dropped, as the spin-dependent formalism will be introduced later and is not necessary for the following explanations.

The last two terms in Equation 3.20 conform an effective potential usually referred as *external potential*. The ground state wave function depends on this external potential introduced in the Hamiltonian. This dependence is a *functional* dependence, i.e., the potential shapes the wave function. Similarly, the electron density is an observable of the wave function and thus depends functionally on the wave function:

$$v(\mathbf{r}) \rightarrow \Psi(\mathbf{r}) \rightarrow \rho(\mathbf{r}) \quad (3.21)$$

The first Hohenberg-Kohn theorem presents the existence of a unique linear map from the external potential  $v(\mathbf{r})$  to the density of the system  $\rho(\mathbf{r})$  (up to an additive constant). It is proven by contradiction, assuming that there are two different Hamiltonians  $\hat{H}$  and  $\hat{H}'$  that yield the same ground state density  $\rho(\mathbf{r})$ . These Hamiltonians differ in the definition of the external potentials  $\hat{V}$  and  $\hat{V}'$  and thus the ground state wave functions resulting from solving the TISE,  $\Psi$  and  $\Psi'$ , must be different. From the variational principle, if the energy is calculated for both trial wave functions with the Hamiltonian  $\hat{H}$ :

$$E_0 = \langle \Psi | \hat{H} | \Psi \rangle < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' + \hat{V}' - \hat{V} | \Psi' \rangle \quad (3.22)$$

As the Hamiltonians differ only in the potential term, the difference in energies must be:

$$E_0 < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' | \Psi' \rangle + \int \rho(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r} \quad (3.23)$$

But doing the same with the opposite Hamiltonian results in the previous term but with the opposite primed terms. This leads to the contradiction that  $E_0 + E'_0 < E_0 + E'_0$ . Therefore, both potentials must be equal (up to an additive constant). This implies that the map from the potential to the electron density is bijective, i.e., there is a unique potential that maps to each electron density, and at the same time the electron density maps uniquely to an external potential:

$$v(\mathbf{r}) \leftrightarrow \Psi(\mathbf{r}) \leftrightarrow \rho(\mathbf{r}) \quad (3.24)$$

The second Hohenberg-Kohn theorem states that the ground state energy of the system is a func-

tional of the ground state density. The proof is based on the variational principle. Considering that the wave function is a functional of the density, and that the energy is a functional of the wave function itself, it is possible to define the energy as a functional of the density. The following reasoning can be applied to any Hamiltonian, but in this case the focus will be the Born-Oppenheimer Hamiltonian (eq. 3.6). This Hamiltonian can be written as a sum of terms that depend on the electronic density as:

$$E[\rho] = \langle \Psi | \hat{H} | \Psi \rangle = T_N[\rho] + V_{NN}[\rho] + V_{Ne}[\rho] + T_e[\rho] + V_{ee}[\rho] \quad (3.25)$$

Due to the Born-Oppenheimer approximation, the nuclear kinetic energy  $T_N$  and nuclei-nuclei potential  $V_{NN}$  are treated as constants. The nuclei-electron attraction potential can be simply defined as:

$$V_{Ne}[\rho] = \int v(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} \quad (3.26)$$

Finally, the last electron kinetic energy  $T_e$  and electron-electron potential  $V_{ee}$  can be grouped in a universal functional of the density  $F[\rho]$  that includes the kinetic energy, correlation and exchange effects:

$$F[\rho(\mathbf{r})] = \langle \Psi | \hat{T} + \hat{U} | \Psi \rangle = T_e[\rho] + V_{ee}[\rho] \quad (3.27)$$

This universal functional is valid for any number of particles and any external potential and defines *exactly* electron behaviour and energy. If this functional is known, the ground state energy of the system can be defined from the ground state density  $\rho_0(\mathbf{r})$ . From the variational principle, the real ground state energy is lower than the energy of any trial wave function. In the same way, having established the relation between the wave function and the density  $\rho'(\mathbf{r})$ , any trial density will yield an energy higher than the real ground state energy:

$$E_0[\rho_0] = \int v(\mathbf{r})\rho_0(\mathbf{r}) d\mathbf{r} + F[\rho_0] < E[\rho'] = \int v(\mathbf{r})\rho'(\mathbf{r}) d\mathbf{r} + F[\rho'] \quad (3.28)$$

Therefore, in order to obtain the ground state energy, it is necessary to minimise the energy functional with respect to the density variationally.

### 3.2.2 Kohn-Sham formalism

In 1965, Kohn and Sham, building on the work of Hohenberg and Kohn, introduced a formalism to describe the universal functional  $F[\rho]$  defined in equation 3.27. They proposed a set of equations analogous to the Hartree-Fock equations, incorporating electron correlation for inhomogeneous systems [22].

The key idea is that the exact ground-state density  $\rho_0(\mathbf{r})$  of a real interacting system can be reproduced by a fictitious system of non-interacting electrons. In this system, there are no explicit two-electron terms in the Hamiltonian; instead, the interactions are included in an effective potential  $v_{xc}(\mathbf{r})$ , which accounts for electron-electron interactions and exchange-correlation effects. The wave function of the fictitious system is described as a single Slater determinant of single-particle orbitals  $\chi_i(\mathbf{r})$ , subject to the orthonormality condition. The density is defined as:

$$\rho(\mathbf{r}) = \sum_i^N |\chi_i(\mathbf{r})|^2 \quad (3.29)$$

The universal functional  $F[\rho(\mathbf{r})]$  is expressed as:

$$F[\rho] = T_s[\rho] + E_{xc}[\rho] \quad (3.30)$$

where  $T_s[\rho]$  is the kinetic energy functional of the non-interacting system, and  $E_{xc}[\rho]$  is the exchange-correlation energy functional. The total energy of the system is then given by:

$$E[\rho] = \int v_{ext}(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} + J[\rho] + T_s[\rho] + E_{xc}[\rho] \quad (3.31)$$

As established in the Hohenberg-Kohn theorems (sec. 3.2.1), the electron density  $\rho(\mathbf{r})$  uniquely determines the external potential  $v(\mathbf{r})$ .

To minimise the energy functional with respect to the density, the Kohn-Sham equations are derived analogously to the Hartree-Fock equations (sec. 3.1.3) by introducing the constraint that the integral of the density equals the number of electrons:

$$\mathcal{L}[\rho] = E[\rho] - \mu \left( \int \rho(\mathbf{r}) d\mathbf{r} - N \right) \quad (3.32)$$

The resulting restricted Kohn-Sham equations are:

$$\left[ v_{ext}(\mathbf{r}) + v_J(\mathbf{r}) - \frac{1}{2} \nabla^2 + v_{xc}(\mathbf{r}) \right] \chi_i(\mathbf{r}) = \varepsilon_i \chi_i(\mathbf{r}) \quad (3.33)$$

In spin-polarised systems, the Kohn-Sham equations can be extended to include spin-dependent terms. This allows the treatment of systems where the spin-up and spin-down electron densities differ, such as magnetic materials. The unrestricted (spin-dependent) Kohn-Sham equations are written as:

$$\left[ v_{ext}(\mathbf{r}) + v_J(\mathbf{r}) - \frac{1}{2} \nabla^2 + v_{xc,\sigma}(\mathbf{r}) \right] \chi_{i,\sigma}(\mathbf{r}) = \varepsilon_{i,\sigma} \chi_{i,\sigma}(\mathbf{r}) \quad (3.34)$$

Here,  $\sigma$  denotes the spin index ( $\uparrow$  or  $\downarrow$ ), and  $v_{xc,\sigma}(\mathbf{r})$  is the spin-dependent exchange-correlation potential. The total electron density is obtained by summing over the spin components:

$$\rho(\mathbf{r}) = \rho_\uparrow(\mathbf{r}) + \rho_\downarrow(\mathbf{r}) \quad (3.35)$$

The Kohn-Sham equations are solved iteratively using a self-consistent process similar to the Hartree-Fock method (sec. 3.1.3). Starting with an initial guess for the density, the equations are solved iteratively to update the effective potential and calculate a new density. This process is repeated until convergence is reached. At convergence (assuming the exact functional), the energy obtained is the same as the real energy of the system, allowing all exact properties to be derived from this density.

### 3.2.3 Exchange-Correlation functionals

In the energy functional described for the Kohn-Sham equations (eq. 3.31), all terms are known except the exchange-correlation energy functional  $E_{xc}[\rho]$ . While the Kohn-Sham formalism provides an exact way to obtain the real energy and density of the system, it is necessary for the exact description of the exchange-correlation energy functional to be known. Unfortunately, this functional is not known, and it is impossible to derive it from first principles. Therefore, it is necessary to approximate this functional in order to obtain a practical solution to the Kohn-Sham equations.

To understand the different approaches to defining the exchange-correlation energy functional and their constraints, it is necessary to understand the nature of the exchange and correlation effects. For this, it is useful to define a function that describes the electron-electron interaction, the pair density:

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\sigma_1 d\sigma_2 d\mathbf{x}_3 \dots d\mathbf{x}_N \quad (3.36)$$

The pair density  $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$  establishes the probability of finding one electron at position  $\mathbf{r}_1$  and another electron at position  $\mathbf{r}_2$ . This pair density will be governed by two effects: the Coulomb repulsion between electrons and the antisymmetry of the wave function. As a consequence of this, two key effects can be identified: the *Coulomb hole* and the *Fermi hole*.

The Coulomb hole is a direct consequence of the Coulomb repulsion between electrons. Since electrons repel each other, the electron density tends to decrease in regions where the density is high. Conversely, the density increases in regions where it is lower, as electrons are less likely to encounter strong repulsion in such areas. The Fermi hole, on the other hand, arises from the antisymmetry of the wave function, as dictated by the Pauli exclusion principle. As a result, the probability of finding two electrons with the same spin in the same region of space is zero.

Considering the Coulomb and Fermi holes, it is possible to define an exchange-correlation hole function:

$$h_{xc}(\mathbf{r}_1; \mathbf{r}_2) = h_c^{\sigma_1, \sigma_2}(\mathbf{r}_1; \mathbf{r}_2) + h_x^{\sigma_1 = \sigma_2}(\mathbf{r}_1; \mathbf{r}_2) \quad (3.37)$$

here, the parametrical dependence arises due to the hole functions being conditioned probabilities, i.e., the hole probability at position  $\mathbf{x}_2$  depends on an electron located at  $\mathbf{x}_1$ . A constraint of the hole functions, later used to define exchange-correlation functionals is that they must fulfil the sum rule:

$$\int h_x(\mathbf{r}_1; \mathbf{r}_2) d\mathbf{r}_2 = -1 \quad \int h_c(\mathbf{r}_1; \mathbf{r}_2) d\mathbf{r}_2 = 0 \quad (3.38)$$

Because the Coulomb and exchange holes represent distinct physical effects, it is standard practice to split the exchange-correlation energy functional into two separate terms:

$$E_{xc}[\rho] = E_x[\rho] + E_c[\rho] \quad (3.39)$$

Coulomb and exchange hole functions must satisfy strict physical constraints such as normalisation, negativity, proper cusp behaviour, the sum rule, and decay at large distances. These requirements ensure a physically accurate description of the energy, as depicted in the definition of the exchange-

correlation functional.

Modern DFT codes offer a wide selection of approximate exchange-correlation functionals, each with its own strengths and limitations. Although no single functional is universally best, their diversity allows for the selection of one that best fits the system under study. These functionals have been built progressively, starting from simple models and incorporating increasing levels of physical detail.

Many exchange-correlation functionals are constructed by modelling the exchange-correlation energy as an integral over an energy per particle  $\varepsilon_{xc}$ :

$$E_{xc}[\rho] = \int \rho(\mathbf{r}) \varepsilon_{xc}(\rho(\mathbf{r})) d\mathbf{r} \quad (3.40)$$

This energy per particle is approximated using the density and related quantities, including parameters for fine-tuning. The parameters in these models are typically fitted to experimental data or results from high-level electronic structure methods, or defined non-empirically in order to satisfy the proper constraints of the hole functions.

As a result, exchange-correlation functionals are commonly classified by the increasing complexity and accuracy of their approximations. There are two main categories: local and semi-local functionals, and hybrid functionals. The first category of functionals is based purely on the local density or its properties and includes the local density approximation, the generalised gradient approximation, and the meta generalised gradient approximation. The second category includes exact exchange or correlation effects through a non-local term and comprises hybrid and double hybrid functionals.

[30]

### **Local Density Approximation**

The Local density approximation (LDA) was already introduced in the publication of Kohn and Sham in 1965 [22]. The LDA assumes that the exchange-correlation energy functional can be approximated by the exchange-correlation energy of a uniform electron gas. This is simple for non-polarized systems, but for different spin densities, the local spin density approximation (LSDA) is used.

The LSDA functional is defined as:

$$E_{xc}^{LSDA}[\rho_\uparrow, \rho_\downarrow] = \int \rho(\mathbf{r}) \varepsilon_{xc}^{uniform}(\rho(\mathbf{r})_\uparrow, \rho(\mathbf{r})_\downarrow) d\mathbf{r} \quad (3.41)$$

where  $\varepsilon_{xc}^{uniform}$  is the exchange-correlation energy per particle of a uniform electron gas. This approximation is valid in the limit where the density is constant throughout space or varies slowly. This is specially accurate in materials such as metals, where the density is high and relatively constant. The LDA satisfies the sum rule of the exchange-correlation hole function which must integrate to  $-1$  (eq. 3.38). The LDA also fulfils the spin-rescaling condition:

$$E_{xc}^{LSDA}[\rho_\uparrow, \rho_\downarrow] = \frac{1}{2}E_x[2\rho_\uparrow] + \frac{1}{2}E_x[2\rho_\downarrow] \quad (3.42)$$

The LDA is computationally efficient and provides reasonable results for many systems, particularly those with slowly varying densities. However, it tends to underestimate the exchange-correlation energy in systems with rapidly varying densities or strong electron correlations.

This approximation was expected to be valid in two limiting cases: when the density is slowly varying and in the high electron density regime. This definition fails in the case of non-homogeneous systems such as molecules or surfaces. In these cases, the exchange energy can be underestimated by  $\approx 10\%$ . Also, electron correlation is greatly overestimated, leading to a poor description of the energy, bonds, and molecular properties.

### Generalised Gradient Approximation

In order to improve upon LDA functionals, it is necessary to accurately define the behaviour of the exchange-correlation functional for non-uniform systems. One way of doing this is not only to consider the local density, but also its derivative  $\nabla\rho(\mathbf{r})$ . This is the basis of the Generalized gradient approximation (GGA) functionals:

$$E_{xc}^{LSDA}[\rho_\uparrow, \rho_\downarrow] = \int \rho(\mathbf{r})\varepsilon_{xc}^{GGA}(\rho(\mathbf{r})_\uparrow, \rho(\mathbf{r})_\downarrow, \nabla\rho(\mathbf{r})_\uparrow, \nabla\rho(\mathbf{r})_\downarrow) d\mathbf{r} \quad (3.43)$$

Initially, some constraints of the exchange-correlation hole function were not satisfied, such as the negativity of the hole function. This resulted in very negative correlation energies, and demonstrated the need for not only the inclusion of the derivative, but also a proper definition of the hole functions.

Some famous GGA functionals include the one proposed by Becke (B88) [31], which reproduces the asymptotic behaviour of the exchange energy. Another popular functional is the one proposed by Lee, Yang and Parr (LYP) [32], where the parametrisation is rather complex and results in a good description of correlation energies in noble gases. Both functionals were designed by empirically fitting the parameters to experimental data. It is common to see them combined in the BLYP

functional, where the exchange functional is defined by B88 and the correlation is given by LYP.

On the other hand, the Perdew-Burke-Ernzerhof (PBE) functional [33] was developed without fitting to any experimental data, but all parameters were fundamental constants. This functional was built around the satisfaction of several exact constraints such as the sum rule (eq. 3.38). It conserves well-defined properties of the LDA approach, but with a significant improvement in the description of the exchange-correlation energy. This is why this functional is widely used in DFT calculations that involve materials and solids.

### Meta-GGA

Meta-GGA functionals consider the orbital kinetic energy density  $\tau(\mathbf{r})$ . It is the highest semi-local family of functionals that avoids the explicit calculation of non-local terms, remaining relatively computationally efficient. The meta-GGA functionals can be defined as:

$$E_{xc}^{meta-GGA}[\rho_{\uparrow}, \rho_{\downarrow}] = \int \rho(\mathbf{r}) \varepsilon_{xc}^{meta-GGA}(\rho(\mathbf{r})_{\uparrow}, \rho(\mathbf{r})_{\downarrow}, \nabla \rho(\mathbf{r})_{\uparrow}, \nabla \rho(\mathbf{r})_{\downarrow}, \tau_{\uparrow}(\mathbf{r}), \tau_{\downarrow}(\mathbf{r})) d\mathbf{r} \quad (3.44)$$

By using the orbital kinetic energy density, meta-GGA can recognise regions of space with varying electron density, even when the gradient of the density is small, and treat them accordingly. This allows differentiation, for example, between a slowly varying density and a one- or two-electron ground-state density.

The addition of  $\tau$  also allows the cancellation of correlation energy in a one-electron atom ( $E_c[\rho, 0] = 0$ ), and makes possible the exact calculation of exchange energy for some particular densities [34].

One of the most popular meta-GGA functionals is the TPSS functional [35], which presents good performance in a wide range of systems such as molecules, hydrogen-bonded complexes, and ionic solids.

### Hybrid and Double Hybrid Functionals

The Limitation of LDA, GGA, and meta-GGA functionals is that they are (semi)local. Therefore, they cannot describe non-local exchange as a non-local function needs to be included in their definition. Hybrid functionals were therefore developed to include the exact exchange component of the Kohn-Sham orbitals obtained from Hartree-Fock theory. This term is added in the definition of the exchange-correlation energy functional. Taking the famous B3LYP functional (B88 + LYP + HF exchange) as an example, which combines Becke's three-parameter exchange functional [36] with the LYP functional [32], the exact exchange can be included as:

$$E_{xc}^{B3LYP} = (1 - a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^{B88} + (1 - c)E_c^{LSDA} + cE_c^{LYP} \quad (3.45)$$

where  $a$ ,  $b$ , and  $c$  are semi-empirical parameters, hence the 3 in B3LYP. Another very popular hybrid functional is the hybrid version of PBE, PBE0 [37], presenting much more accurate results in molecules and solid lattices than its non-hybrid counterpart.

Finally, at the top of the hierarchy of exchange-correlation functionals, we find the double hybrid functionals. These functionals include a non-local exchange term and a second-order perturbation correction to the correlation energy. In this way, virtual orbitals are included in the functional definition, generating a more accurate description of electron correlation effects. An example of these types of functionals is the B2PLYP functional [38], which combines the B88 exchange functional with a second-order perturbation correction to the correlation energy:

$$E_{xc}^{DH} = (1 - \alpha)E_x^{DFT} + \alpha E_x^{HF} + (1 - \beta)E_c^{DFT} + \beta E_c^{PT2} \quad (3.46)$$

The inclusion of a more accurate Hartree-Fock exchange and PT2-like term provides a good description of electron correlation, but also increases the computational cost significantly. This is why the application of these functionals to large systems is limited, especially double hybrids, as the computational cost is on the order of the PT2 method itself.

### Range Separation

A problem reported with exchange-correlation functionals is the description of long-range interactions. This is particularly relevant in systems with weak interactions such as van der Waals forces, hydrogen bonds, and dispersion interactions. These interactions are often underestimated by standard functionals, leading to inaccurate predictions of molecular geometries and binding energies.

In order to tackle this problem, range-separated hybrid functionals were developed. These functionals separate the exchange energy into short-range and long-range components, allowing for a more accurate treatment of long-range interactions. In this scheme, short-range exchange is described by the DFT functional, while the long-range orbital-orbital exchange is described through the exact Hartree-Fock integral [39]. The range separation is usually treated by an Ewald split:

$$\frac{1}{r_{12}} = \frac{1}{r_{12}^{SR}} + \frac{1}{r_{12}^{LR}} = \frac{1 - \text{erf}(\mu r_{12})}{r_{12}} + \frac{\text{erf}(\mu r_{12})}{r_{12}} \quad (3.47)$$

where  $\text{erf}(x)$  is the Gaussian error function, which in the positive value range varies between 0 and 1, and  $\mu$  is a range-separation parameter (also named  $\omega$  elsewhere) that determines the separation between short-range and long-range interactions. The Gaussian error function has the properties  $\text{erf}(0) = 0$  and  $\lim_{x \rightarrow \infty} \text{erf}(x) = 1$ . With this definition, the short-range contribution is attenuated as distance increases while the long-range contribution increases. These range-separated parts act as operators and substitute the  $r_{12}^{-1}$  operators commonly found in Coulomb and exchange integrals as:

$$E_x^{\text{LR-HF}} = -\frac{1}{2} \sum_{i,j}^{\text{occ.}} \left\langle ij \left| \frac{\text{erf}(\mu r_{12})}{r_{12}} \right| ij \right\rangle \quad (3.48)$$

Some prominently used functionals are the "Coulomb Attenuating Method" B3LYP (CAM-B3LYP) [39], and the  $\omega$ B97X functional [40]. The first adds extra parameters to the Ewald split, while the second optimises the  $\mu$  parameter. Both functionals have shown improved results compared to their non-separated counterparts in systems with non-covalent interactions.

### 3.3 Time-Dependent Density Functional Theory

Hohenberg-Kohn-Sham DFT is time-independent and thus cannot treat problems involving time-dependent potentials. These types of potentials are fundamental when treating excitations and spectroscopic properties. Therefore, it is necessary to generalise time-independent DFT to a time-dependent case.

#### 3.3.1 Runge-Gross Theorem

The extension to the time-dependent regime of ground-state density-based definition of the wave function was shown by Runge and Gross [24]. Consider a time-dependent Hamiltonian generated by introducing an external potential that varies with time:

$$\hat{H}(t) = -\frac{1}{2} \sum_i^N \nabla_i^2 + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \hat{V}(t) \quad (3.49)$$

where the first term is the electron kinetic energy, the second the electron-electron repulsion, and  $\hat{V}(t)$  the external potential:

$$\hat{V}(t) = \sum_i^N v(\mathbf{r}_i, t) \quad (3.50)$$

The Runge–Gross theorem proves that two external potentials that differ by more than a function of time itself cannot produce the same density starting from the same initial state. The basis of this proof is the assumption that the external potential is Taylor-expandable, and, using the continuity equation relating the current density and the density, it is possible to show that the density evolution under the action of two different potentials diverges at infinitesimal times after the application of the potential if they are not the same.

Therefore, the Runge–Gross theorem sets the foundation of Time-Dependent Density Functional Theory (TD-DFT) by demonstrating that the map between the time-dependent potential and time-dependent density is bijective, i.e., there is a unique potential that maps to a given density. This result is analogous to the Hohenberg–Kohn theorem for ground-state DFT, but in the time-dependent regime. It provides the conceptual basis for describing excited-state and dynamical phenomena in quantum systems using the time-dependent density alone, avoiding the need of the full many-body wave function.

### 3.3.2 Linear Response Time-Dependent Density Functional Theory

While the Runge–Gross theorem establishes a theoretical foundation for TD-DFT, in practice, evolving the full time-dependent Kohn–Sham system is often computationally expensive and unnecessary. For many systems and properties of interest, especially in spectroscopy and excited-state calculations, a full time-propagation is not required.

Instead, linear response (LR) theory provides a more efficient alternative. In this approach, the system is initially in its ground state and is subject to a small, time-dependent perturbation. By analysing how the density responds linearly to this perturbation, it is possible to extract valuable information about the system’s excited states and properties, such as excitation energies, oscillator strengths, polarisabilities, and absorption spectra.

The LR-TD-DFT formalism is based on the derivation from the interaction picture in quantum mechanics, where the evolution of the system can be separated into a time-independent Hamiltonian and a time-dependent perturbation. Only the main equations of linear response theory and the interaction picture in quantum mechanics will be presented, while the full derivation can be found elsewhere [26, 41].

The main focus of linear response theory (in this context) is determining the evolution of an operator  $\hat{\alpha}$  with time when perturbed by a Hamiltonian of the form:

$$\hat{V}_{ext}(t) = \begin{cases} 0 & \text{if } t \leq t_0 \\ F(t)\hat{\beta} & \text{if } t > t_0 \end{cases} \quad (3.51)$$

where the perturbing field is activated at a time  $t_0$ . It is possible to define the *response* function of the operator  $\alpha$  and expand it as:

$$\alpha(t) - \alpha_0 = \alpha_1(t) + \alpha_2(t) + \alpha_3(t) + \dots \quad (3.52)$$

From the previous expansion, following the necessary algebra [26], the linear term  $\alpha_1(t)$  is expressed as:

$$\alpha_1(t) = \int_{-\infty}^{\infty} dt' \chi_{\alpha\beta}(t-t')F(t') \quad (3.53)$$

where  $\chi_{\alpha\beta}(t-t')$  is the retarded response function, defined as:

$$\chi_{\alpha\beta}(t-t') = -i\theta(t-t') \langle \Psi_0 | [\hat{\alpha}(t-t'), \hat{\beta}] | \Psi_0 \rangle \quad (3.54)$$

The response function  $\chi_{\alpha\beta}(t-t')$  is central to linear response theory and demonstrates causality due to the presence of the Heaviside step function  $\theta(t-t')$ : the response of the system at time  $t$  depends only on perturbations applied at earlier times  $t'$ . In this context,  $t'$  represents the time at which the perturbation  $F(t')$  is applied, while  $t$  is the time at which the response  $\alpha_1(t)$  is measured.

The response function also contains information on the evolution of the system due to the commutator  $[\hat{\alpha}(t-t'), \hat{\beta}]$ . This ensures that the response function is sensitive to the specific operators involved in the perturbation and the observable being measured.

In the context of TD-DFT, the observable of interest is the density, and thus the density-density response function  $\chi_{\rho\rho}$ . It is common to work in Fourier space by Fourier transforming the response function:

$$\alpha_1(\omega) = \chi_{\alpha\beta}(\omega)F(\omega) \quad (3.55)$$

where:

$$\chi_{\alpha\beta}(\omega) = -i \int_{-\infty}^{\infty} d\tau \theta(\tau) \langle \Psi_0 | [\hat{\alpha}(\tau), \hat{\beta}] | \Psi_0 \rangle e^{i\omega\tau} \quad (3.56)$$

In the context of the density-density response function, the commutator  $[\hat{\alpha}(\tau), \hat{\beta}]$  can be expanded, and by introducing the completeness relation  $\sum_n |\Psi_n\rangle \langle \Psi_n| = \mathbb{I}$  and the integral form of the step function  $\theta(\tau)$ , this leads to the following expression:

$$\chi_{\rho\rho}(\mathbf{r}, \mathbf{r}'; \omega) = \lim_{\eta \rightarrow 0^+} \sum_{n=1}^{\infty} \left[ \frac{\langle \Psi_0 | \hat{\rho}(\mathbf{r}) | \Psi_n \rangle \langle \Psi_n | \hat{\rho}(\mathbf{r}') | \Psi_0 \rangle}{\omega - \Omega_n + i\eta} - \frac{\langle \Psi_0 | \hat{\rho}(\mathbf{r}') | \Psi_n \rangle \langle \Psi_n | \hat{\rho}(\mathbf{r}) | \Psi_0 \rangle}{\omega + \Omega_n + i\eta} \right] \quad (3.57)$$

Here,  $\Omega_n$  represents the  $n$ -th excitation frequency of the system. This equation, known as the *Lehmann representation* of the linear response function, encapsulates all the information about the excitation spectrum of the system.

From equation 3.57, the excitation energies of the system can be identified. When the frequency  $\omega$  approaches an excitation energy  $\Omega_n$ , the response function exhibits a resonance. The imaginary term  $i\eta$  prevents divergence by shifting the pole into the complex plane. This not only avoids singularities but also enables the determining the excitation spectrum by examining the imaginary part of the response function, which behaves like a Dirac delta function:

$$\lim_{\eta \rightarrow 0^+} \text{Im} \frac{1}{\omega - \omega' + i\eta} = -i\pi\delta(\omega - \omega') \quad (3.58)$$

Not only the spectrum can be obtained, but other properties can be determined from the response function. For example, the trace of the dynamic polarizability response function defines the mean polarizability, and the residue at each pole  $\text{Res}[\chi_{zz}, \Omega_n] = 2\Omega_n/3 \sum_{\mu}^3 |\langle \Psi_n | \hat{r}_{\mu}(\mathbf{r}) | \Psi_0 \rangle|^2$  measures the transition strength (or oscillator strength) from the ground to the  $n$ -th excited state.

### 3.3.3 Casida Equations and Tamm-Dancoff Approximation

The Lehmann representation provides a theoretical foundation for understanding the excitation spectrum of a system. However, practical calculations require a computationally efficient approach. In the context of LR-TD-DFT, this is achieved by reformulating the problem in terms of the Kohn-Sham orbitals and their response to the perturbation. The formulation of this problem in a matrix form was introduced by Casida [25] and is the basis of many TD-DFT implementations.

Equation 3.57 can be expanded, and by introducing the Kohn-Sham framework, potential, orbitals, and exchange-correlation functionals, we arrive at the equation that forms the basis of Casida's

formalism:

$$\sum_{\sigma'} \sum_{j'k'} [\delta_{ij'} \delta_{ak'} \delta_{\sigma\sigma'} \omega_{j'k'\sigma'} + \alpha_{j'k'\sigma'} K_{ia\sigma, j'k'\sigma'}(\Omega)] \beta_{j'k'\sigma'}(\Omega) = \Omega \beta_{ia\sigma}(\Omega) \quad (3.59)$$

The matrix elements  $\alpha_{j'k'\sigma'}$  represent the differences in occupation numbers between occupied and unoccupied orbitals, while  $\omega_{j'k'\sigma'}$  represents the energy differences between these orbitals. The matrix elements  $K_{ia\sigma, j'k'\sigma'}(\Omega)$  are the elements of the interaction kernel, which describes the coupling between single-particle excitations and includes the exchange-correlation kernel of TD-DFT. Finally,  $\beta_{ia\sigma}(\Omega)$  are the excitation amplitudes, which represent the change in the electron density due to the perturbation, and  $\Omega$  is the excitation frequency associated.

Equation 3.59 forms a non-Hermitian eigenvalue problem that allows the calculation of the excitation energies  $\Omega$  and the excitation amplitudes  $\beta$  from a Kohn-Sham reference system. It is common to use  $X_{ia\sigma} = -\beta_{ia\sigma}$  and  $Y_{ia\sigma} = \beta_{ia\sigma}$  to denote the excitation amplitudes, where  $X_{ia\sigma}$  represents the excitation amplitudes while  $Y_{ia\sigma}$  represents the de-excitation amplitudes. WIth this definition of excitation and de-excitation amplitudes, the Casida equation is defined as:

$$\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{B} & \mathbb{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \Omega \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \quad (3.60)$$

where the matrix elements are defined as:

$$A_{ia\sigma, i'a'\sigma'} = \delta_{ii'} \delta_{aa'} \delta_{\sigma\sigma'} \omega_{a'i'\sigma'} + K_{ia\sigma, i'a'\sigma'}(\Omega) \quad (3.61)$$

$$b_{ia\sigma, i'a'\sigma'} = K_{ia\sigma, i'a'\sigma'}(\Omega) \quad (3.62)$$

In practice, this system of equations can be further simplified by exploiting the properties of  $(\mathbb{A} - \mathbb{B})$ , which is diagonal when using real orbitals, yielding a symmetric eigenvalue equation:

$$\mathbb{C}\mathbf{Z} = \Omega^2 \mathbf{Z} \quad (3.63)$$

where:

$$\mathbb{C} = (\mathbb{A} - \mathbb{B})^{1/2} (\mathbb{A} + \mathbb{B}) (\mathbb{A} - \mathbb{B})^{1/2} \quad (3.64)$$

and:

$$\mathbf{Z} = (\mathbb{A} - \mathbb{B})^{-1/2}(\mathbf{X} + \mathbf{Y}). \quad (3.65)$$

This transformed eigenvalue equation is numerically advantageous, as it reduces the original non-Hermitian problem to a symmetric one. The excitation energies  $\Omega$  are obtained from the square roots of the eigenvalues of  $\mathbb{C}$ , while the corresponding eigenvectors determine the excitation character.

Different approximations can be made to further reduce the computational cost of the calculations. The most common one is the Tamm-Dancoff Approximation (TDA), which neglects the de-excitation processes. In practice, this is done by removing the  $\mathbb{B}$  matrix from the equations, leading to a simplified eigenvalue equation:

$$\mathbb{A}\mathbf{X} = \Omega\mathbf{X} \quad (3.66)$$

This is formally identical to the configuration interaction singles (CIS) method, and thus all the results obtained from the TDA can be interpreted as CIS results. Also, TDA has been shown to perform better in some cases, such as open-shell systems away from the ground-state equilibrium geometry [26].

### 3.3.4 Introduction of a basis set

Throughout the previous sections, the equations shown implicitly assume the exact definition of the molecular orbitals  $\chi_i$  in the Hartree-Fock, Kohn-Sham, and Casida formalisms. In practice, this is not possible, and thus a finite *basis set* must be employed. The molecular orbitals are defined as a linear expansion of *basis* functions:

$$\chi_i(\mathbf{r}) = \sum_{\mu}^{\infty} C_{i\mu} \phi_{\mu}(\mathbf{r}) \quad (3.67)$$

Limiting the basis set implies an approximation in the definition of the molecular orbitals, and thus in the accuracy of the calculations. Therefore, expanding the basis set leads to the *basis set limit*, a limit in which the results converge to the exact solution of the problem within the theoretical framework, for example in the case of Hartree-Fock to the exact HF energy (that can still be improved by adding correlation effects).

There are many basis set functions, but the most common ones used in molecular systems are

Slater-type orbitals and Gaussian-type orbitals. Slater-type orbitals are defined as an exponential ( $-\zeta r$ ), and mimic the exact solution of the hydrogen atom:

$$\phi^{\text{STO}}(\mathbf{r}) = Ne^{-\zeta r} Y_{lm}(\theta, \phi) \quad (3.68)$$

where  $N$  is a normalisation constant,  $\zeta$  is the orbital exponent,  $Y_{lm}(\theta, \phi)$  are the spherical harmonics, and  $r$  is the distance to the nucleus. Slater-type orbitals are particularly useful for describing the behaviour of electrons in atoms and molecules, as they can accurately represent the radial part of the wave function and reproduce the discontinuity at the nucleus.

Conversely, Gaussian-type orbitals are defined as a Gaussian function ( $-\alpha r^2$ ) multiplied by a spherical harmonic:

$$\phi^{\text{GTO}}(\mathbf{r}) = Nx^l y^m z^n e^{-\alpha r^2} \quad (3.69)$$

This definition is not able to reproduce the discontinuity of the Slater-type orbitals at the nucleus, but it is computationally more efficient due to the properties of Gaussian functions, such as the product of two Gaussian functions being another Gaussian function, and the integral of a Gaussian function being analytically solvable. This makes them particularly suitable for numerical calculations, as they allow for efficient evaluation of integrals and other operations.

Gaussian-type orbitals are usually not described as a single Gaussian function, but are instead *contracted*, i.e., defined as a linear combination of several Gaussian functions:

$$\phi^{\text{CGTO}}(\mathbf{r}) = \sum_{p=1}^{N_c} d_p x^l y^m z^n e^{-\alpha_p r^2} \quad (3.70)$$

where  $N_c$  is the number of *primitive* Gaussian functions,  $d_p$  are contraction coefficients, and  $\alpha_p$  are the exponents of each primitive. Contracted Gaussian-type orbitals combine the computational efficiency of Gaussian-type orbitals with improved flexibility to approximate atomic orbitals towards the exact hydrogen atom solutions.

The parametrisation of  $d_p$  and  $\alpha_p$  allows fine-tuning of the basis set to a specific system or methodology. As the variational principle states, the parameters that yield the lowest energy will produce more accurate results. Therefore, there are numerous families of CGTOs, depending on the intended application.

Throughout this work, the Dunning correlation-consistent *cc-pVDZ* basis set was used [42]. The cc-pVDZ is a *double-zeta* basis set, which means that it contains two sets of Gaussian-type orbitals for each atom, allowing for a better description of the electron density and wave function. It also includes polarisation functions, functions with higher angular momentum than the valence orbitals, which allow for a more accurate description of the molecular orbitals.

### 3.4 Absorption, emission and photostability

When light interacts with matter, it can produce an excitation in the rotational, vibrational, or electronic levels of the system. When the energy of the photon matches the energy difference between electronic levels, Ultra-Violet (UV) radiation can promote a molecule from the ground electronic state to an excited electronic state. A photon with energy on the UV range is able to promote electrons from occupied to virtual orbitals, for example from the Highest Occupied Molecular Orbital (HOMO) to the Lowest Unoccupied Molecular Orbital (LUMO), generating an excited species. Since it is very unlikely that the minima of the ground and excited electronic states lie on the same geometry, the electronic excitation is usually accompanied by a vibrational excitation following the Franck-Condon (FC) principle [43]. This principle states that electronic transitions take place instantly, without a geometrical rearrangement, making transitions vertical. The transition intensity is governed by the overlap between vibrational wave functions. As a consequence, transitions take place more likely to excited vibrational states, where the overlap is maximum.

After the excitation, the excess electronic and vibrational energy on the excited species is dissipated by decaying back to the ground state through radiative or non-radiative pathways. If the relaxation process is non-radiative, a change in the final chemical structure of the system can happen, creating a new species. This formation of a new product is known as photochemistry, and it can take place via two different mechanisms. The change of the system's structure can during the relaxation process on an excited-state surface. Another possibility is that the excess of energy can be enough to overpass energetic barriers on the ground state, allowing access to new products. If the vibrational relaxation does not result in any meaningful change of the system, due to vibrational energy being dissipated very fast, or it not being enough to overcome energetic barriers, this process is known as photophysics.

Radiative processes such as fluorescence or phosphorescence emit light spontaneously. Non-radiative pathways are vibrational relaxation, transition between states of the same multiplicity (internal conversion), or between states of different multiplicity (intersystem crossing). These processes are shown in Figure 3, a Jablonski diagram, which is a simplified graphical representation of the

processes involved in vibronic excitation and deactivation obviating the change of coordinates of the system following relaxation.

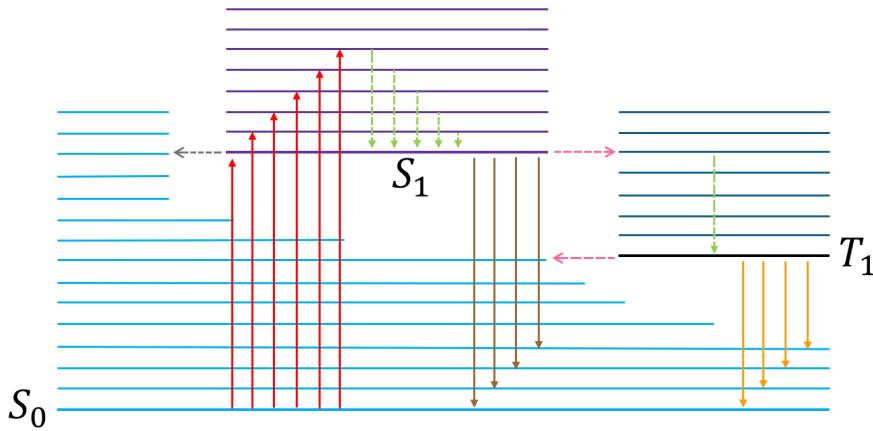


Figure 3: Jablonski diagram illustrating the main processes in a vibronic excitation process. Radiative transitions pictured in solid lines: absorption (red), fluorescence (brown), and phosphorescence (orange). In dashed lines, non-radiative transitions: vibrational relaxation (green), internal conversion (grey), and intersystem crossing (pink). Bold lines represent vibrational state  $v_0$  of each electronic state. Thin lines represent excited vibrational states.

The different processes illustrated in Figure 3 take place on different time scales and compete as deactivation pathways. Absorption of a photon takes place in the range of  $10^{-15}$  to  $10^{-12}$  seconds, while vibrational relaxation to the lowest vibrational state occurs in the range  $10^{-14}$  to  $10^{-12}$  seconds [44]. After reaching the lowest vibrational state, the molecule must decay through radiative or non-radiative processes.

Fluorescence is a spontaneous emission and its efficiency is dictated by Fermi's Golden Rule (FGR) (section 3.6). Fermi's golden rule states that the rate of emission is proportional to the transition dipole moment between the two states. Internal conversion depends on the overlap of the vibrational wave functions of the upper and lower electronic states. Intersystem crossing requires a transition between states of different multiplicity, which is forbidden. Therefore, the process takes longer and is slower than fluorescence and internal conversion.

Due to the nature of the previous processes, the type of decay greatly depends on the system and

its properties. Therefore, some systems will favour radiative deactivation, while others will favour non-radiative pathways.

A complete relaxation can result in two outcomes: the formation of photoproducts or the return to the original ground state. The outcome of the relaxation is determined by the efficiency of the different competing deactivation pathways and thus by the potential energy landscape of the system. Therefore, in order to determine the fate of excited population, it is necessary to study the efficiency of the different possible pathways.

### 3.5 Non-Adiabatic Chemistry

Even though the Born-Oppenheimer approximation works well for many systems, it usually breaks down when different electronic states are involved in a process. As the potential energy surfaces get closer, the coupling terms neglected in the Born-Oppenheimer approximation become significantly large. The processes that involve these couplings between different electronic states are called *non-adiabatic processes*. These processes are especially important in photochemistry, where non-radiative relaxation of excited systems occurs through the coupling of electronic states.

These processes happen because the variation of the electronic wave function with the nuclear coordinates is not smooth. To quantify this behaviour, the first-order and second-order non-adiabatic coupling terms between two electronic states  $A$  and  $B$  are evaluated as:

$$h_{\alpha}^{AB}(\mathbf{R}) = \langle \psi_A(\mathbf{R}; \mathbf{r}) | \nabla_{\alpha} | \psi_B(\mathbf{R}; \mathbf{r}) \rangle \quad (3.71)$$

$$K^{AB}(\mathbf{R}) = \langle \psi_A(\mathbf{R}; \mathbf{r}) | \nabla_{\alpha}^2 | \psi_B(\mathbf{R}; \mathbf{r}) \rangle \quad (3.72)$$

where  $\alpha$  is the nuclear coordinate,  $\mathbf{R}$  is the nuclear coordinate vector, and  $\mathbf{r}$  is the electronic coordinate vector. The first-order non-adiabatic coupling term  $h_{\alpha}^{AB}(\mathbf{R})$  is the largest, and drives the non-adiabatic transitions. The direction of the coupling can be defined as an N-dimensional vector, called the Coupling Derivative Vector (CDV), that can be obtained between two eigenfunctions of the Hamiltonian as::

$$\mathbf{h}^{AB} = \frac{\langle \psi_A | \nabla \hat{H}_{el} | \psi_B \rangle}{E_B - E_A} \quad (3.73)$$

where  $E_A$  and  $E_B$  are the energies of the states  $A$  and  $B$ , respectively and  $\hat{H}_{el}$  is the electronic

Hamiltonian introduced in Section 3.1.1. The CDV is a vector that points in the direction of the non-adiabatic coupling, and its magnitude is proportional to the strength of the coupling. From this definition, it follows that it vanishes at large energy differences and diverges when states are degenerate.

The efficiency of non-adiabatic transitions between two states depends on the derivative coupling vector between the states [29]. At the degeneracy between states, the coupling is maximum, allowing an efficient transition between them. The degeneracy point is called a Conical intersection (CoIn), as the intersection between two electronic states presents a double-cone shape. CoIns play a fundamental role in photochemistry and photophysics, as they provide ultrafast and efficient pathways for non-radiative transitions between electronic states. This makes CoIns essential for understanding complete non-radiative decay mechanisms, as these processes depend on internal conversion and intersystem crossings.

CoIns are not isolated points on the PES, but exist in an  $N - 2$  dimensional *hypersurface* or *seam*, where  $N$  is the number of degrees of freedom [45]. Along the  $N - 2$  dimensional seam, there exists a single intersection that presents the lowest energy of the seam. This point is called a Minimum energy conical intersection (MECI). The MECI is the most efficient point for non-adiabatic transitions, as relaxation along the CoIn hypersurface leads to this point (or to a local minimum).

To understand the nature of the intersection between two electronic states, it is useful to introduce the *branching plane*. The branching plane is a two-dimensional subspace of nuclear coordinates that locally lifts the degeneracy between the intersecting states. It is defined by two vectors: the Coupling Derivative Vector (CDV),  $\mathbf{h}^{AB}$ , and the Difference Gradient Vector (DGV),  $\mathbf{g}^{AB}$  [46]. The difference gradient vector is defined as:

$$\mathbf{g}^{AB} = \frac{1}{2}(\mathbf{g}^B - \mathbf{g}^A) \quad (3.74)$$

where  $\mathbf{g}^A$  and  $\mathbf{g}^B$  are the gradients of the states  $A$  and  $B$ , respectively. Displacements along  $\mathbf{h}^{AB}$  and  $\mathbf{g}^{AB}$  remove the degeneracy, while displacements orthogonal to the branching plane (along the seam) preserve the degeneracy while altering the energy. Therefore, as will be explained in 4.1, it is possible to find the MECI by searching for the minimum energy along the seam, while keeping the branching plane fixed.

As the branching plane is defined uniquely by the CDV and the gradients of each state, it is possible to define two magnitudes,  $\mathcal{P}$  and  $\mathcal{B}$ , that characterise the type of intersection. The derivation of

these magnitudes is explained in Section 4.2.

The  $\mathcal{P}$  parameter refers to the slope of the CoIn, i.e., the tilt of the conical intersection, while the  $\mathcal{B}$  parameter refers to the bifurcation of the CoIn. A value of  $\mathcal{P} < 1$  indicates a peaked CoIn and  $\mathcal{P} > 1$  a sloped CoIn, whereas  $\mathcal{B} < 1$  indicates a single-path CoIn and  $\mathcal{B} > 1$  a bifurcating intersection. These are shown in Figure 4.

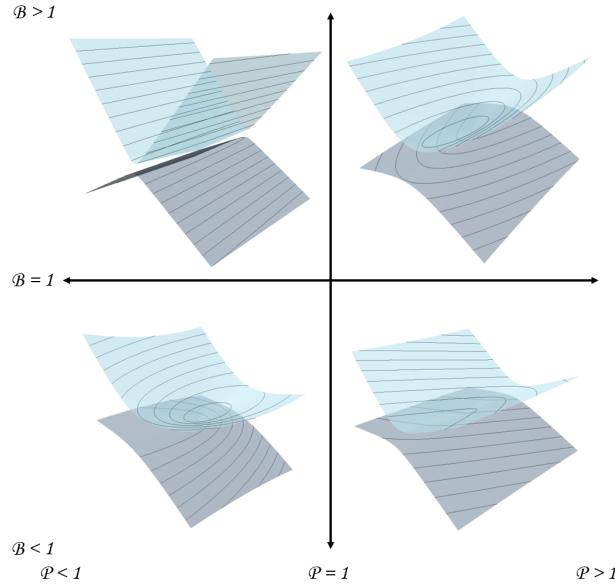


Figure 4: Topologies of minimum energy conical intersections (MECIs) classified by  $\mathcal{P}$  and  $\mathcal{B}$  parameters. Upper left: Peaked Bifurcating. Lower left: Peaked Single-Path. Upper right: Sloped Bifurcating. Lower right: Sloped Single-Path.

Characterisation by means of  $\mathcal{P}$  and  $\mathcal{B}$  can be intuitively understood by looking at the graphical representation of the CoIn. Smaller values of  $\mathcal{P}$  indicate that the upper surface cone is centred around the CoIn geometry, while larger values indicate that the upper cone is tilted. Conversely, a smaller value of  $\mathcal{B}$  indicates that energy decreases quickly in a single direction from the CoIn geometry, while larger values indicate that the energy decreases in two directions.

The difference between peaked and tilted CoIns can play a significant role in the efficiency of the CoIn. A peaked CoIn facilitates the transfer of population to the lower energy state. In contrast, a tilted intersection requires the wavepacket to cross over the conical intersection point multiple times, transferring smaller wavepacket fractions[47].

Single-path CoIns funnel the mechanism towards the formation of a single photoproduct, and are indicative of photostable molecules. Bifurcating conical intersections are characterised by the formation of multiple photoproducts, and some typical examples include cis-trans photoisomerisation

[47].

### 3.6 Emission Spectra and Constant Rates

The rate of absorption, spontaneous emission, internal conversion, and intersystem crossing can be calculated by the Fermi Golden Rule:

$$k_{if} = \frac{2\pi}{\hbar} \left| \langle f | \hat{H}' | i \rangle \right|^2 \rho(E_f) \quad (3.75)$$

where  $i$  is the initial state,  $f$  the final state,  $\rho(E_f)$  the density of states at the energy of the final state, and  $\hat{H}'$  the matrix elements of the perturbed Hamiltonian. The perturbed Hamiltonian varies depending on the nature of the transition that takes place. In the case of absorption or emission, the transition dipole moment  $\mu$  dictates the transition, while it is the Non-Adiabatic Coupling (NAC) and Spin-Orbit Coupling (SOC) in the case of internal conversion and intersystem crossing, respectively. The previous equation can be recast when the Born-Oppenheimer approximation is valid as:

$$k_{if} = \frac{2\pi}{\hbar} |V_{if}|^2 \sum_{v_i, v_f} P_{i, v_i}(T) \left| \langle \Theta_{i, v_i} | \Theta_{f, v_f} \rangle \right|^2 \delta(E_{f, v_f}^f - E_{i, v_i}) \quad (3.76)$$

Here,  $V_{if}$  is the electronic coupling element between the initial and final states,  $P_{i, v_i}(T)$  is the Boltzmann population of the initial vibrational state at temperature  $T$ , and  $|\Theta_{i, v_i}\rangle$ ,  $|\Theta_{f, v_f}\rangle$  are the vibrational wave functions of the initial and final states, respectively.

The lack of dependence of the electronic coupling term on the vibrational degrees of freedom, allowing the separation of electronic and vibrational contributions, is the Condon approximation.

The exact definition of the vibronic wave function would require a complete mapping of the PES, which is not feasible for most systems. Therefore, the harmonic approximation is introduced, where the vibrational wave functions are written as a product of uncoupled harmonic oscillators. Due to the introduction of the harmonic approximation, it is necessary to define the harmonic oscillators that will be used for the calculation of rate constants. This results in three different models: the adiabatic Hessian, vertical Hessian, and vertical gradient.

The vertical Hessian and adiabatic Hessian models both take into account the Hessians of the existing and final electronic states, but differ in how these are evaluated. In the vertical Hessian model, the Hessians for both the upper and lower states are calculated at the same geometry. In

the adiabatic Hessian model, each state's Hessian is calculated at its own equilibrium geometry. Conversely, the vertical gradient model uses a single Hessian for both states. Because the potential energy surfaces are anharmonic, the vertical and adiabatic Hessian approaches can yield different results. Generally, vertical models provide a more accurate description of the Franck-Condon region, while adiabatic models are more accurate for transitions near the 0 – 0 point [48].

With the chosen definition of the vibrational wave functions, the emission spectrum can be calculated in the frequency domain as:

$$\sigma(\omega, T) = \frac{4\omega^3}{3c^3} \sum_{v_i, v_f} P_{i, v_i}(T) |\langle \Psi_v | \vec{\mu} | \Psi_i \rangle|^2 \delta(E_{fi} + E_{i, v_i} - E_{f, v_f} - \hbar\omega) \quad (3.77)$$

where  $\vec{\mu}$  is the transition dipole moment operator, and the integral over  $\omega$  yields the radiative rate constant  $k_r$ .

# 4 Computational Methods

In this section a brief overview of the implemented computational methods will be presented. The methods will be divided in two sections: the first one will focus on the optimization of Minimum energy conical intersection (MECI)s, while the second in their characterization.

## 4.1 Conical intersection optimization

Throughout this work, the exploration of the different Potential Energy Surface (PES) regarding the search of Conical intersection (CoIn)s was done with a homemade program. This program is an implementation of the Updated Branching Plane (UBP) method [46], which combines an iterative estimation of the Coupling Derivative Vector (CDV) vector with the Gradient Projection method [49]. The program interfaces with an electronic structure code in order to request the required properties for the optimization. In this section, a brief description of the UBP algorithm will be presented.

Geometry optimization is often one of the most fundamental processes in quantum chemistry. Force calculation is usually fairly simple, as most of the time the forces of interest are the gradient of the energy with respect to the coordinates. Optimizing algorithms manage the optimization process, including the caveats that can arise, such as gradient rescaling, convergence criteria, etc. Throughout this work two different optimizers have been used: the Broyden-Fletcher-Goldfarb-Shanno (BFGS) [50] as implemented in ORCA [51], and the Fast Inertial Relaxation Engine (FIRE) [52] optimizer as implemented in ASE [53].

The BFGS algorithm belongs to the family of Quasi-Newton optimization algorithms, based in the estimation of the inverse Hessian matrix, and an iterative update of it. Using the inverse Hessian and the gradient, the algorithm estimates the new position of the atoms by updating the position of the previous step as  $x_{k+1} = x_k + \alpha_k p_k$ , where  $\alpha_k$  is a factor ensuring correct scaling of the step size and  $p_k$  is the search direction defined as  $p_k = -H_k^{-1}\nabla f(x_k)$ . On the other hand, the FIRE algorithm is based the gradient and a velocity term instead of the Hessian. The velocity term is updated with the forces, and the position of the atoms is propagated with a molecular-dynamics integrator such as the velocity Verlet. While the BFGS algorithm is more efficient when the PES has a smooth topology, the FIRE algorithm has no memory of previous steps giving better results when the topology of the PES can change abruptly (often the case close to CoIns).

As described in section 3.5, the two conditions around a MECI is that  $\Delta E_{ij} = 0$  and that the energy of the upper state is the lowest possible while maintaining degeneracy. These two conditions can

be considered defining the effective force as a combination of two components:

$$\mathbf{g}_{\text{eff}} = \mathbf{g}^{AB} + \mathbf{P}_{\text{GDV},\text{CDV}} \times \mathbf{g}^{AB} \quad (4.1)$$

The inclusion of the Difference Gradient Vector (DGV) ensures degeneracy as  $\mathbf{g}_{\text{eff}}$  approaches 0. The second term is the component that minimizes the mean gradient  $\mathbf{g}^{AB}$ , resulting in a descent of the CoIn hyperseam all the way to the lowest (local) energy point, the MECI.

In order to minimize the energy without breaking degeneracy, it is required that the energy is minimized in all coordinates except the ones that form the branching plane ( $\mathbf{g}^{AB}$  and  $\mathbf{h}^{AB}$ ), as a change in these coordinates lifts degeneracy. This is why the projector matrix  $\mathbf{P}_{\text{GDV},\text{CDV}}$  is included:

$$\mathbf{P}_{\text{GDV},\text{CDV}} = \mathbb{I} - \mathbf{x}\mathbf{x}^T - \mathbf{y}\mathbf{y}^T \quad (4.2)$$

where  $\mathbb{I}$  is the identity matrix,  $\mathbf{x}$  is the normalized DGV, and  $\mathbf{y}$  is the normalized CDV. This projector discards these directions from the mean gradient, leaving them unaltered. With this definition of the effective force, it is possible to use it as the force with any optimizer.

This method is often used without including the CDV term in cases where it is not available. This neglection can lead to a CoIn, but may not be the minimum energy one. To improve upon this the UBP method [46] was developed, that iteratively estimates the Non-Adiabatic Coupling (NAC) vector with almost no extra cost using the following equation:

$$\mathbf{y}_k = \frac{(\mathbf{y}_{k-1} \cdot \mathbf{x}_k)\mathbf{x}_{k-1} - (\mathbf{x}_{k-1} \cdot \mathbf{x}_k)\mathbf{y}_{k-1}}{\sqrt{(\mathbf{y}_{k-1} \cdot \mathbf{x}_k)^2 + (\mathbf{x}_{k-1} \cdot \mathbf{x}_k)^2}} \quad (4.3)$$

where once again  $\mathbf{x}$  is the normalized DGV and  $\mathbf{y}$  is the normalized CDV. The CDV in the first iteration is guessed as a vector perpendicular to the DGV, and it has been shown to converge to the actual CDV in the vicinity of a MECI [46].

The interface program written follows the previous algorithm as represented in Figure 5. Some example code such as the input file or the UBP and effective force calculation are shown in appendix A.

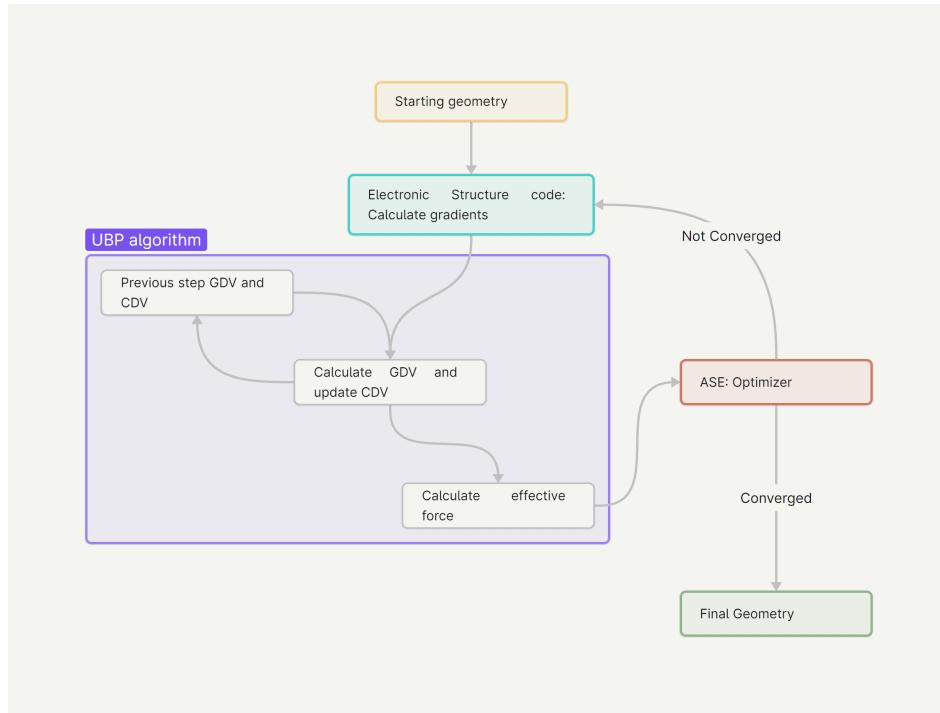


Figure 5: Workflow of the conical intersection optimization algorithm as implemented.

## 4.2 Conical intersection characterization

As the CDV  $\mathbf{y}_k$  at the end of a MECI optimization was available from the UBP implementation, the conical intersection characterization algorithm defined by Fdez. Galván et. al. [27] was implemented. In this section, a brief overview of the algorithm and will be presented.

The energies near the CoIn can be approximated by using 3 vectors (or the combination of them). These are the energy gradients of each state  $\mathbf{g}^A$  and  $\mathbf{g}^B$ , and the non-adiabatic coupling vector  $\mathbf{h}^{AB}$ . Considering the mean gradient and mean difference gradient vectors:

The branching plane is defined then by the vectors  $\mathbf{g}^{AB}$  (equation 3.74) and  $\mathbf{h}^{AB}$  (equation 3.73). Even though these vectors are uniquely defined, a linear combination of them would suffice to define the branching plane correctly. Therefore, it is convenient to perform a rotation with respect to an angle  $\beta$  such that these vectors are orthogonal:

$$\tilde{\mathbf{g}} = \mathbf{g}^{AB} \cdot \cos(\beta) + \mathbf{h}^{AB} \cdot \sin(\beta) \quad \text{and} \quad \tilde{\mathbf{h}} = \mathbf{g}^{AB} \cdot \cos(\beta) - \mathbf{h}^{AB} \cdot \sin(\beta) \quad (4.4)$$

The value of the angle of rotation  $\beta$  can then be obtained via:

$$\tan(2\beta) = \frac{2\mathbf{g}^{AB} \cdot \mathbf{h}^{AB}}{\mathbf{g}^{AB} \cdot \mathbf{g}^{AB} - \mathbf{h}^{AB} \cdot \mathbf{h}^{AB}} \quad (4.5)$$

There are 4 possibilities of  $\beta$  that satisfy this orthogonality condition and correspond to  $\frac{\pi}{2}$  rotations of the  $\tilde{\mathbf{g}}$  and  $\tilde{\mathbf{h}}$  vectors. By normalizing these orthogonal vectors, it is possible to define the branching plane with the unitary vectors:

$$\hat{\mathbf{x}} = \frac{\tilde{\mathbf{g}}}{\sqrt{\tilde{\mathbf{g}} \cdot \tilde{\mathbf{g}}}} \quad \text{and} \quad \hat{\mathbf{y}} = \frac{\tilde{\mathbf{h}}}{\sqrt{\tilde{\mathbf{h}} \cdot \tilde{\mathbf{h}}}} \quad (4.6)$$

This way, it is possible to define a 2-dimensional projection in space over these two directions that defines the geometries in the vicinity of the conical intersection:

$$\mathbf{R}(x, y) = \mathbf{R}^\times + x\hat{\mathbf{x}} + y\hat{\mathbf{y}} \quad (4.7)$$

where  $\mathbf{R}^\times$  is the geometry at the conical intersection and  $\mathbf{R}(x, y)$  represents the coordinates along the unitary vectors that define the branching plane. Similarly, the energy difference in the branching plane can be estimated with respect to the  $x$  and  $y$  coordinates:

$$|\Delta E^{AB}(x, y)| = 2\sqrt{x^2 \cdot \tilde{\mathbf{g}} \cdot \tilde{\mathbf{g}} + y^2 \cdot \tilde{\mathbf{h}} \cdot \tilde{\mathbf{h}}} \quad (4.8)$$

Instead of using this definition, it is simpler to define the pitch  $\delta_{gh}$  and the asymmetry  $\Delta_{gh}$  as:

$$\delta_{gh} = \sqrt{\frac{1}{2}(\tilde{\mathbf{g}} \cdot \tilde{\mathbf{g}} + \tilde{\mathbf{h}} \cdot \tilde{\mathbf{h}})} \quad \text{and} \quad \Delta_{gh} = \frac{\tilde{\mathbf{g}} \cdot \tilde{\mathbf{g}} - \tilde{\mathbf{h}} \cdot \tilde{\mathbf{h}}}{\tilde{\mathbf{g}} \cdot \tilde{\mathbf{g}} + \tilde{\mathbf{h}} \cdot \tilde{\mathbf{h}}} \quad (4.9)$$

So that the energy difference is obtained directly via:

$$|\Delta E^{AB}(x, y)| = 2\delta_{gh}\sqrt{(x^2 + y^2) + \Delta_{gh}(x^2 - y^2)} \quad (4.10)$$

The energy of the upper and lower states in the vicinity of the conical intersection can then be estimated from the mean energy plane and the energy difference as:

$$E(x, y) = \bar{E}^{AB}(x, y) \pm |\Delta E^{AB}(x, y)| \quad (4.11)$$

where the average energy in the branching plane  $\bar{E}^{AB}$  is defined as:

$$\bar{E}^{AB}(x, y) = E^\times + x(\mathbf{s}^{AB} \cdot \hat{\mathbf{x}}) + y(\mathbf{s}^{AB} \cdot \hat{\mathbf{y}}) \quad (4.12)$$

Being  $E^\times$  is the energy at the conical intersection.

In the MECI, there is a direction in which the degeneracy is broken the fastest. This direction is aligned with the mean gradient vector  $\mathbf{s}^{AB}$  of the two states and is defined as:

$$\mathbf{s}^{AB} = \frac{1}{2}(\mathbf{g}^A + \mathbf{g}^B) \quad (4.13)$$

At a MECI,  $\mathbf{s}^{AB}$  is completely contained in the branching plane [27].

Only the projection of the  $\mathbf{s}^{AB}$  vector (equation 4.13) is necessary in order to define the mean energy plane. Therefore, it is useful to define a relative tilt  $\sigma$  that indicates magnitude of this projection and is scaled with the pitch:

$$\sigma = \sqrt{s_x^2 + s_y^2} \quad \text{where} \quad s_x = \frac{\mathbf{s}^{AB} \cdot \hat{\mathbf{x}}}{\delta_{gh}} \quad \text{and} \quad s_y = \frac{\mathbf{s}^{AB} \cdot \hat{\mathbf{y}}}{\delta_{gh}} \quad (4.14)$$

Finally, the angle  $\theta_s$  is the angle formed by the  $\hat{\mathbf{x}}$  vector and the projection  $\mathbf{s}_{ab}$  in the branching plane.

As mentioned earlier, there are four possible solutions of  $\beta$  that define the  $\tilde{\mathbf{g}}$  and  $\tilde{\mathbf{h}}$  vectors. By convention, usually the angle  $\beta$  is rotated arbitrarily such as the asymmetry  $\Delta_{gh}$  is positive, and the projection of  $\mathbf{s}^{AB}$  is contained within the first quadrant (implying that  $[0 < \theta_s < \frac{\pi}{2}]$ ). A failure to fulfil this requirement could indicate that the definition of the branching plane vectors is not correct, or that the CoIn analysed has not converged to a MECI.

With the previous definitions of  $\delta_{gh}$ ,  $\Delta_{gh}$ ,  $\sigma$ , and  $\theta_s$  it is possible to characterize the conical intersections according to the magnitudes  $\mathcal{P}$  and  $\mathcal{B}$ . The magnitude  $\mathcal{P}$  is defined as:

$$\mathcal{P} = \frac{\sigma^2}{1 - \Delta_{gh}^2} (1 - \Delta_{gh} \cdot \cos(2 \cdot \theta_s)) \quad (4.15)$$

And  $\mathcal{B}$  is defined as:

$$\mathcal{B} = \sqrt[3]{\frac{\sigma^2}{4\Delta_{gh}^2}} \times \left( \sqrt[3]{(1 + \Delta_{gh}) \cos^2 \theta_s} + \sqrt[3]{(1 - \Delta_{gh}) \sin^2 \theta_s} \right) \quad (4.16)$$

As mentioned in section 3.5 the  $\mathcal{P}$  parameter refers to the slope of the CoIn, i.e., the tilt of the conical intersection while the  $\mathcal{B}$  parameter refers to the bifurcation of the CoIn. A  $\mathcal{P} < 1$  indicates a peaked CoIn and  $\mathcal{P} > 1$  a sloped CoIn otherwise. Where  $\mathcal{B} < 1$  indicates a single path CoIn and  $\mathcal{B} > 1$  a bifurcating intersection.

## 5 Computational details

Ground state optimisations and vertical absorption spectra calculations were performed with the ORCA [51] quantum chemistry package. All electronic structure calculations throughout this work were performed with the CAM-B3LYP [39] functional and the correlation-consistent cc-pVDZ [42] basis set.

The semiclassical spectra were calculated either with NEWTON-X [54] interfaced with the GAUSSIAN 16 [55] quantum chemistry package, or by overlapping the vertical absorption spectra calculated in ORCA. In both cases the initial conditions are sampled from a Wigner quasiprobability distribution. In both cases, 200 initial conditions were generated for the spectra.

Excited state optimisations were performed on the first excited states using ORCA, following the character of the root. An additional  $S_1$  optimisation in each system was performed in GAUSSIAN.

Constant rates and emission spectra were calculated using the FCClasses [48] program, employing the GAUSSIAN  $S_1$ -optimised minima. The  $S_1/S_0$  non-adiabatic coupling matrix elements for the internal conversion rates were calculated with Q-CHEM [56] and ORCA. All rate constants were calculated using the adiabatic Hessian model.

Minimum energy conical intersection (MECI)s were obtained with the Updated Branching Plane (UBP) implementation, which interfaces the atomistic software environment [53] with ORCA. The  $\mathcal{P}$  and  $\mathcal{B}$  parameters were also obtained, using as Coupling Derivative Vector (CDV) the last step's  $\mathbf{y}_k$  vector.

The JMOL [57] program was used for the representation of molecular structures throughout this work.

# 6 Results

## 6.1 Substituent effect

In this section, we present the results of the study of xanthine and its derivative caffeine. The aim of this section is to compare the photophysical and photochemical properties of both systems, and to explore the effect of methylation on the xanthine system.

### 6.1.1 Xanthine

The vertical and semiclassical absorption spectra were calculated using CAM-B3LYP/cc-pVDZ in its time dependent variant and within the Tamm-Dancoff Approximation (TDA). The resulting absorption spectra are shown in Figure 6.

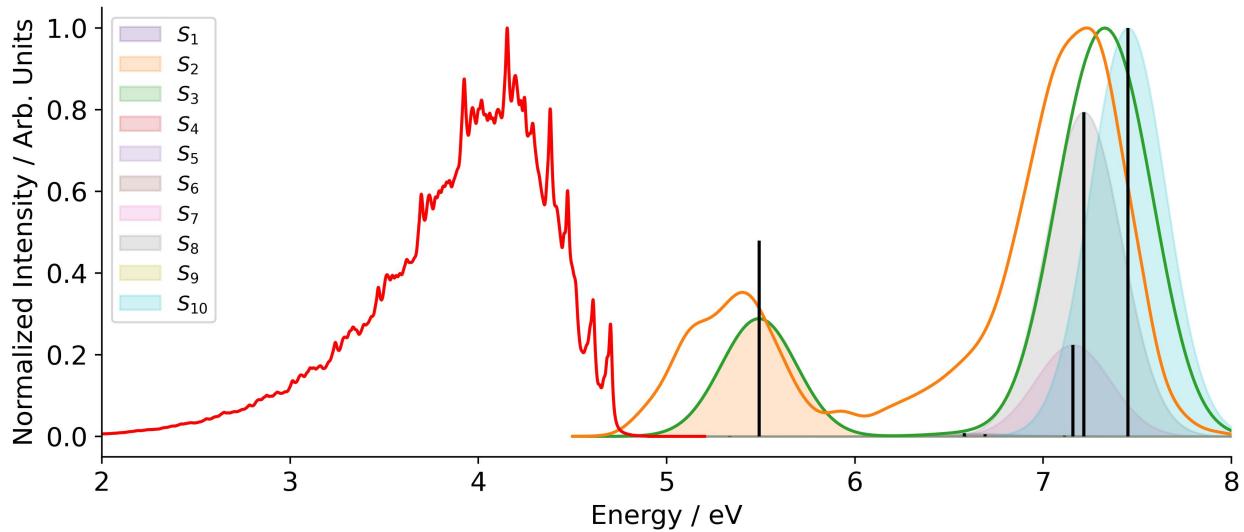


Figure 6: Absorption spectra of xanthine. In orange, semiclassical absorption spectrum. In green, convoluted vertical absorption spectrum. Black lines indicate the positions and relative oscillator strength of vertical excitations. In red, calculated emission spectrum.

Characterization of the excited states character was performed via analysis of Natural Transition Orbitals (NTO) (Appendix B.1). The first absorption in the vertical spectrum is located at 5.34 eV, corresponding to an  $n_O \rightarrow \pi^*$  transition, and its oscillator strength is small. The most intense absorption peaks of the vertical spectrum are associated with transitions to the  $S_2$  (5.49 eV),  $S_7$  (7.16 eV),  $S_8$  (7.22 eV), and  $S_{10}$  (7.45 eV). The states  $S_2$ ,  $S_7$ ,  $S_8$ , and  $S_{10}$  all present a  ${}^1(\pi\pi^*)$  character involving different  $\pi$  orbitals.

The semiclassical absorption spectrum shows a red shift of 0.10 eV in the main absorption band

compared to the vertical spectrum. It presents a similar band structure, with the main absorption located at 6.32 eV, and the first peak observed in the vertical spectrum is replaced by a band with notable absorption between 4.6–5.8 eV.

Relaxation of the molecular geometry from the Franck-Condon (FC) structure in each of the first ten excited states converged to four distinct geometries on the  $S_1$  surface (Appendix B.1). From these geometries, two minima were identified in the  $S_1$  manifold, as shown in Figure 7. The first presents a  ${}^1(n_O\pi^*)$  character, while the second minimum presents an elongation of the C–O bond and a  ${}^1(\pi\pi^*)$  character. An additional minimum on the  $S_3$  surface was found at 6.00 eV above the reference energy. Figure 8 shows a schematic representation of the excited Potential Energy Surface (PES) of xanthine.

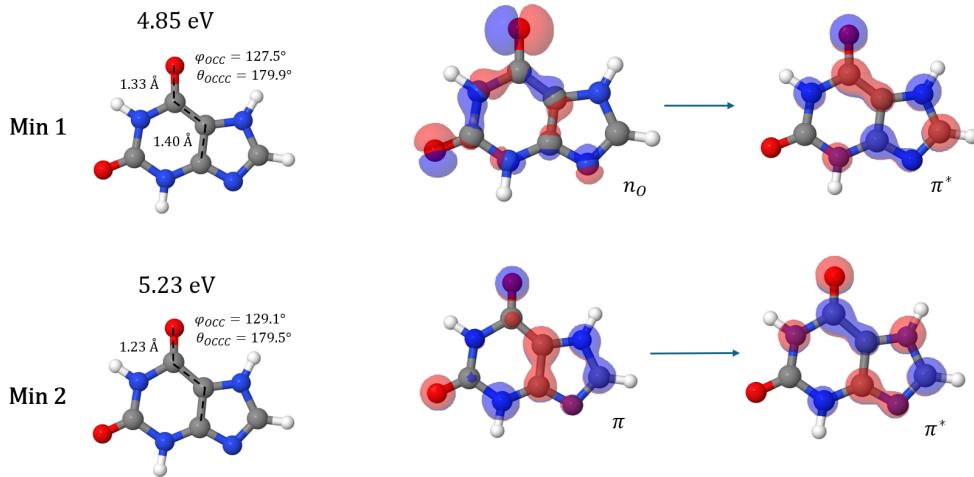


Figure 7: Xanthine optimised  $S_1$  excited state minima electronic state character.

The emission spectrum and rate constants were calculated for the lowest-energy excited state minimum (Figure 6). The rate constants and quantum yield are presented in Table 1. The radiative rate constants are smaller than the calculated internal conversion rate constants, resulting in a fluorescence quantum yield of 0.14. Therefore, non-radiative relaxation may be the dominant deactivation pathway.

$k_r / s^{-1}$	$k_{IC} / s^{-1}$	$\Phi_f = \frac{k_r}{k_r+k_{ic}}$
$1.0917 \times 10^{08}$	$6.519 \times 10^{08}$	0.1434

Table 1: Xanthine  $S_1$  rate constants and fluorescence quantum yield.

Starting from the two excited state minima (and the additional  $S_1$  geometries), Minimum energy conical intersection (MECI) optimizations were performed. Only a single conical intersection close in

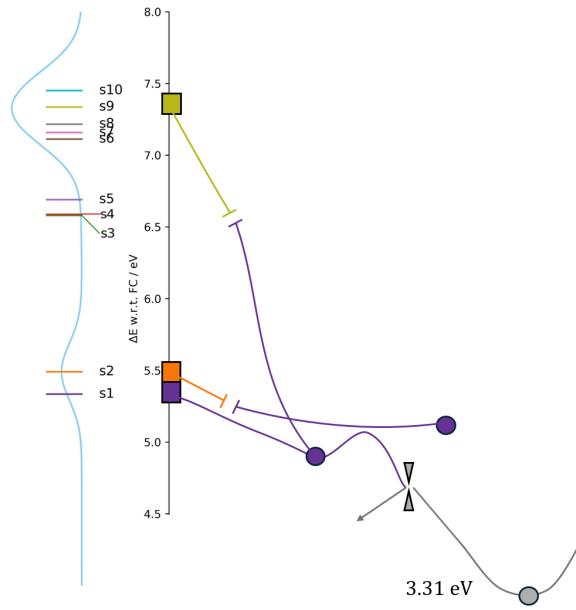


Figure 8: Schematic representation of the excited-state PES of xanthine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Gaps indicate change in energy surface maintaining state character. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows indicate relaxation to the FC geometry. Energies relative to the FC energy.

energy to the  $S_1$  minima was found (Table 2), while the rest of the MECIs located were energetically unaccessible (Appendix B.1). This MECI is characterised by a ring opening of the imidazole-like ring, elongating the  $C-N$  bond to 2.87 Å and a  ${}^1(\pi\pi_{CN}^*)$   $S_1$  character (Figure 9). This MECI is characterised as bifurcating (Table 2).

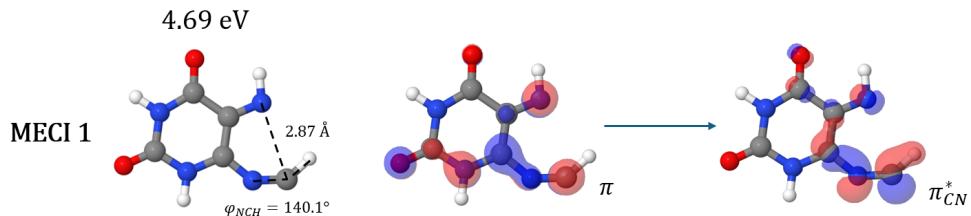


Figure 9: Xanthine optimised  $S_1/S_0$  MECI geometry and electronic state character.

Min	$\Delta E$ min-FC	$\Delta E$ MECI-FC	$\Delta E$ min-MECI	$\mathcal{P}$	$\mathcal{B}$	Type
1	4.85	4.69	-0.16	0.1278	0.1012	Peaked Bifurcating

Table 2: Xanthine MECI  $\mathcal{P}$  and  $\mathcal{B}$  parameters and classification. Energies are given in eV.

The Minimum Energy Path (MEP) between the first  $S_1$  minimum and the located MECI was calculated as to determine the accessibility of this  $S_1/S_0$  conical intersection (Appendix B.1). The pathway connecting these two points was obtained using a Nudged Elastic Band (NEB) in the  $S_1$  electronic state, requiring the breaking of a  $C-N$  bond, with an energy barrier of 0.37 eV above the excited state minimum, totalling 5.27 eV above the FC energy. The vertical excitation energy to the  $S_1$  is 5.33 eV, and thus this energy barrier can be surpassed, rendering the MECI accessible.

To determine the possible paths that could be followed after the  $S_1/S_0$  intersection, the geometry was displaced in the direction of minimum energy on both sides, and were optimised in the  $S_0$  state. This resulted to ring closure in one minimum energy direction, but the other direction converges to a ground state minimum following the ring opening, 3.31 eV above the FC energy (Figure 10).

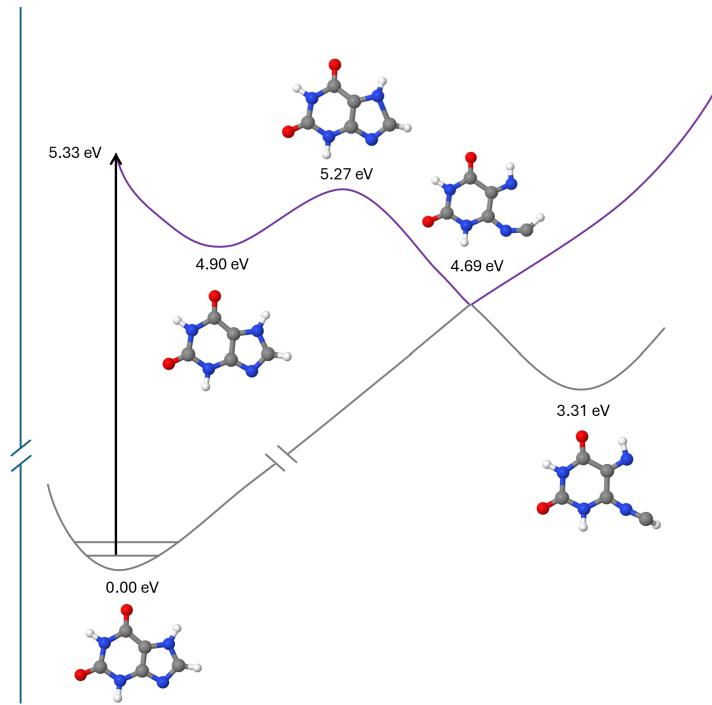


Figure 10: Schematic representation of the PES along the first  $S_1$  minimum of xanthine.

Considering the low radiative rate constant and the accessibility of an  $S_1/S_0$  Conical intersection (CoIn), non-radiative decay is possible in this system, either towards the FC geometry or towards ring opening. However, to determine the exact deactivation pathway, further dynamical calculations are required. Nevertheless, this system could potentially act as an effective solar filter.

### 6.1.2 Caffeine

Caffeine is the tri-methylated derivative of xanthine, and thus comparison between the systems can provide information regarding the effect of these substituents. The vertical and semiclassical absorption spectra were calculated using CAM-B3LYP/cc-pVDZ in its time dependent variant and within the TDA. The resulting absorption spectra are shown in Figure 11, and compared with the experimental reference [58].

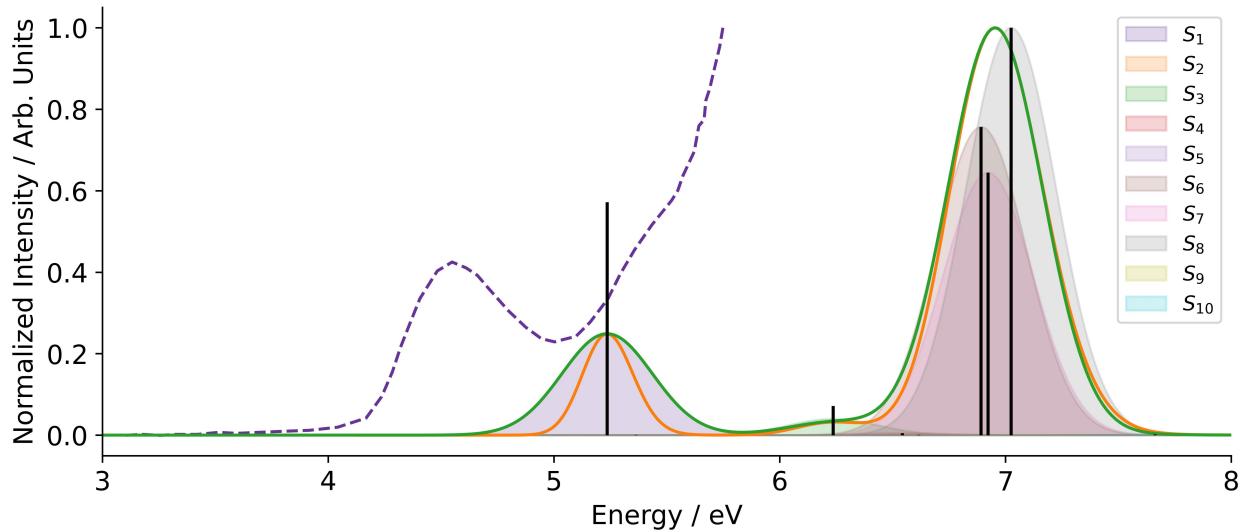


Figure 11: Absorption spectra of caffeine. In purple: experimental absorption spectrum digitalised from reference [58]. In orange: semiclassical absorption spectrum. In green: convoluted vertical absorption spectrum. Black lines indicate the positions and relative oscillator strength of vertical excitations.

The vertical and semiclassical absorption spectra show two distinct absorption bands. The first band is located at 5.24 eV and corresponds to the  $S_0 \rightarrow S_1$  transition, which has a large oscillator strength. This transition is characterised by a  $\pi \rightarrow \pi^*$  transition (Appendix B.2) in the aromatic system. The second absorption band appears due to the  $S_0 \rightarrow S_5$ ,  $S_0 \rightarrow S_6$ , and  $S_0 \rightarrow S_7$  transitions, presenting  ${}^1(n_O\pi^*)$ ,  ${}^1(\pi\pi^*)$ , and  ${}^1(\pi\pi^*)$  character respectively.

Comparing the semiclassical and vertical absorption spectra, it is possible to observe that there is no significant differences in band structure except for a narrowing of the first band. Both calculated spectra overestimate the first band energy by 0.70 eV. Nevertheless, the band structure is in similar to experimental results, including the small shoulder present in the most intense band.

A direct comparison between the absorption spectra of caffeine (Figure 11) and xanthine (Figure 6) reveals that there is no significant effect of methyl substitution on the band structure. Both systems exhibit a similar band structure, with intense bands corresponding to  $\pi \rightarrow \pi^*$  transitions. However,

it is possible to appreciate a slight variation in band structure between the two systems. The first band in caffeine is narrower and is not shifted towards lower energies as in the case of xanthine.

This phenomenon is likely due to the population of methyl rotation vibrational modes in caffeine, which do not alter the aromatic system and thus do not modify absorption properties. Conversely, xanthine vibrational modes generate elongation of bonds and displacement of atoms in the aromatic system, excited state energies. Vertical absorption spectra of both systems exhibit quite similar shapes, with a slight red shift of the most intense band of  $\approx 0.10$  eV in the caffeine system. This is an indicator that methylation decreases the orbital energy gap by inductive effect (Section 1.3).

Relaxation on the first ten excited states converged on four  $S_1$  geometries (Appendix B.2). Two distinct  $S_1$  minima were located with similar geometries, but different electronic state character (Figure 12). The first minimum presents a  ${}^1(n'_O\pi^*)$  character while the second one presents a  ${}^1(\pi'\pi^*)$  character. Geometrically, the most notable differences between these minima is the rotation of the methyl substituents. Two additional minima on the  $S_5$  surface were located for caffeine, at an energy of  $\approx 6.5$  eV (Appendix B.2). Notably, caffeine and xanthine  $S_1$  minima are contained within the same energy range of 4.9–5.2 eV with respect to their reference. Figure 13 shows a schematic representation of the excited PES of caffeine.

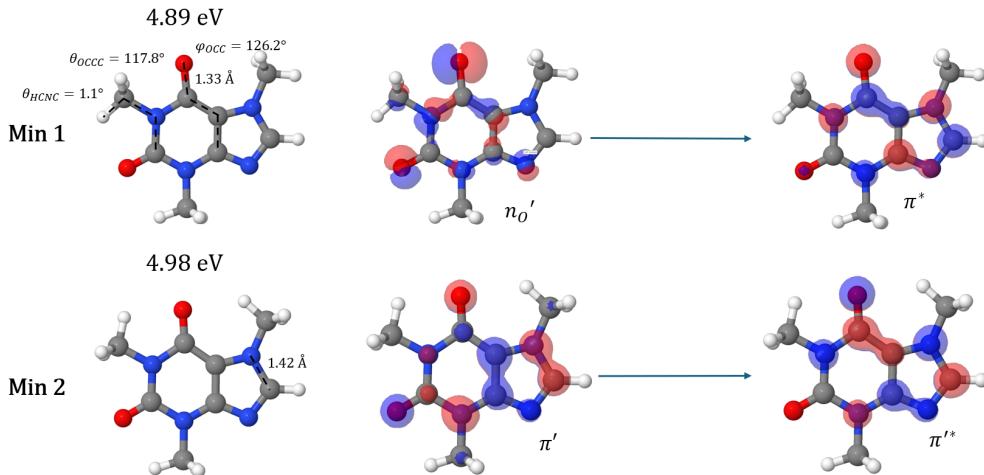


Figure 12: Optimised caffeine  $S_1$  excited state minima and electronic state character.

The photoemission spectrum and rate constants were calculated at the lowest-energy  $S_1$  minimum. The resulting photoemission rate constant was  $\approx 0$ . This would imply that the non-radiative deactivation pathway is the dominant one. However, a comparison between the FC geometry and the  $S_1$  geometry used to calculate the Hessian in the adiabatic Hessian model presented rotation in two of the three methyl groups. As a consequence, there is no vibrational overlap between the FC

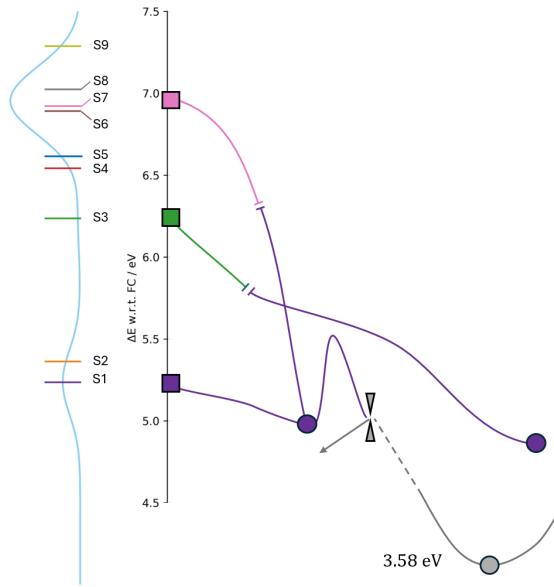


Figure 13: Schematic representation of the excited-state PES of caffeine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Gaps indicate change in energy surface maintaining excited state character. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows indicate a return to FC geometry. Energies relative to the FC energy.

and  $S_1$  vibrational wave functions resulting in the zero constant rate.

Starting from the  $S_1$  minima (and the additional  $S_1$  geometries),  $S_1/S_0$  MECI optimizations were performed. As in the xanthine system, only a single conical intersection close in energy to the  $S_1$  minima was found (Table 3), while the rest of the MECIs located were energetically unaccessible (Appendix B.2). This MECI is again characterised by a ring opening of the imidazole-like ring, elongating the  $C-N$  bond to 2.73 Å and a  ${}^1(\pi\pi_{CN}^*)$   $S_1$  character (Figure 14). This MECI presents a Peaked bifurcating topology (Table 3).

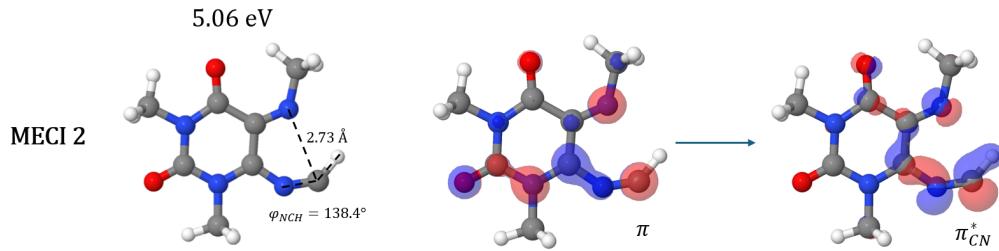


Figure 14: Caffeine optimised  $S_1/S_0$  MECI geometry and electronic state character.

The minimum energy path between the  $S_1$  minima 2 the obtained MECI was calculated (Appendix B.2). The path leading to the MECI presented an energy barrier of 0.54 eV in due to the

Min	$\Delta E$ min-FC	$\Delta E$ MECI-FC	$\Delta E$ min-MECI	$\mathcal{P}$	$\mathcal{B}$	Type
2	4.98	5.0	+0.08	0.1381	0.1704	Peaked Bifurcating

Table 3: Caffeine MECI  $\mathcal{P}$  and  $\mathcal{B}$  parameters and classification. Energies are given in eV.

*C–N* bond stretching. The total adiabatic energy of the barriers amounts to 5.52 eV above the FC energy for. Considering that the vertical  $S_1$  excitation energy is 5.24 eV, the path leading to the MECI may be inaccessible.

After the *C–N* bond dissociation barrier, a second smaller barrier appeared in the calculation of the NEB, due to the rearrangement of the *N–C–H* angle. This barrier corresponds to an in-plane angle change of *N–C–H*, which is not expected to be present in the real path, as an out of plane rotation would take place. Therefore, this barrier is overestimated and is not expected along the real path.

The possible formation of photoproducts from caffeine was explored by displacing the geometry in both directions of the bifurcating MECI and reoptimising in the  $S_0$  state. In this case, the bifurcation leads both to the formation of a photoproduct and the return to the FC geometry (Figure 15). The formed product is a minimum in the  $S_0$  surface located at 3.58 eV above the FC reference, with a ring opening conformation similar to the one in the xanthine system.

In summary, the photophysical and photochemical properties of caffeine reveal differences when compared to xanthine. An increase in the energy barrier between the excited state minimum and the connected MECI was observed. As a consequence, due to the high minimum–MECI barrier, it is unlikely that neither the MECIs nor the photoproduct will be accessible in a real dynamic process under UVA–UVB conditions. Therefore, the efficient non-radiative deactivation of the excited state is not expected, and the system will likely return to the ground state radiatively.

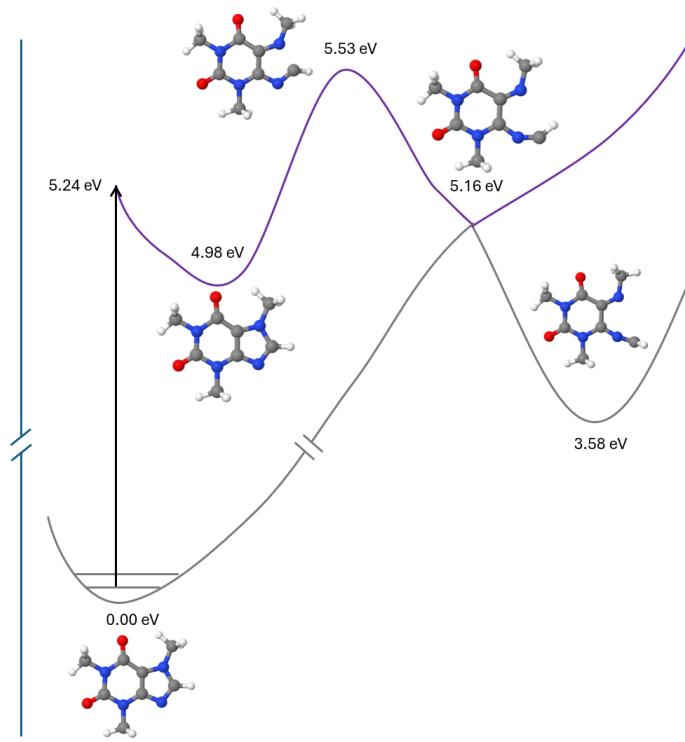


Figure 15: Schematic representation of the PES along the third  $S_1$  minimum of caffeine.

## 6.2 Beta-carboline

The vertical and semiclassical absorption spectra were calculated using CAM-B3LYP/cc-pVDZ in its time dependent variant and within the TDA. Figure 16 shows the calculated absorption spectra and the experimental absorption spectrum [59].

The first vertical absorption of beta-carboline is located at 4.49 eV, corresponding to a  $\pi \rightarrow \pi^*$  transition (Appendix B.3), and presents a small oscillator strength. The first transition with notable oscillator strength corresponds to the  $S_0 \rightarrow S_3$  transition at 5.08 eV which presents a  ${}^1(\pi\pi^*)$  character. The most intense peaks are located at 6.25 and 6.90 eV, corresponding to the  $S_0 \rightarrow S_7$  and ( $S_0 \rightarrow S_{10}$ ) transitions, presenting a  ${}^1(\pi\sigma^*)$  and  ${}^1(\pi\sigma_{NH}^*)$  character respectively.

The semiclassical absorption spectrum presents a red shift of 0.26 eV compared with the vertical spectrum. As illustrated in Figure 16, the semiclassical absorption spectrum aligns more closely with the experimental band positions than the vertical spectrum, but lacks a completely correct description. While the overall band pattern is better captured by the semiclassical approach, both computational methods underestimate the intensity of the first absorption band relative to experiment.

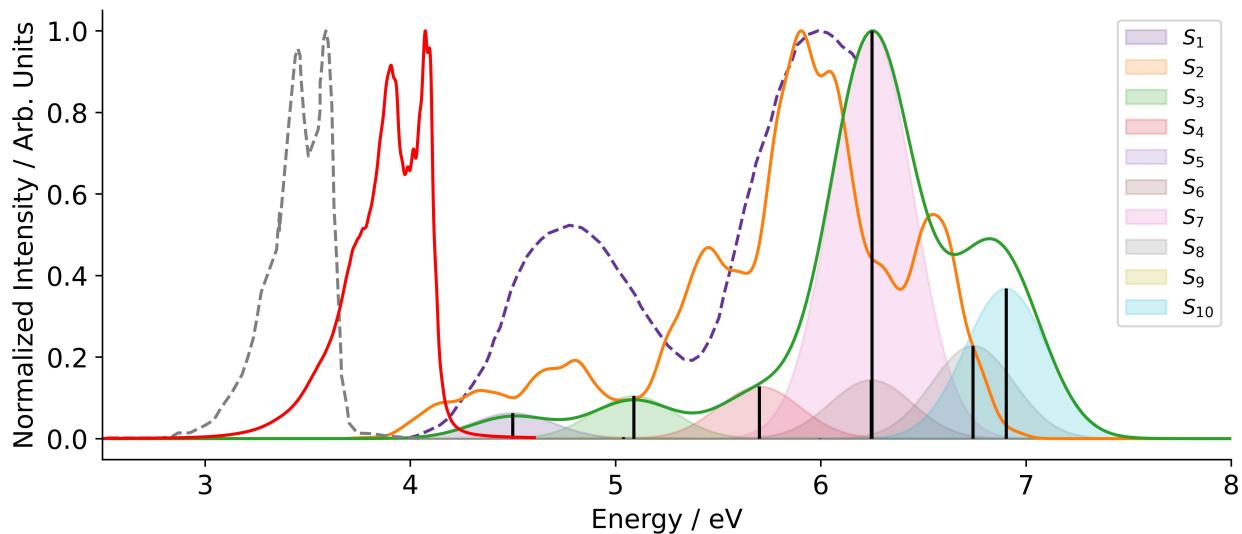


Figure 16: Absorption and emission spectra of beta-carboline. In orange, semiclassical absorption spectrum. In purple dashed lines: experimental absorption spectrum digitalised from reference [59]. In green, convoluted vertical absorption spectrum. Black lines indicate the positions and relative oscillator strength of vertical excitations. In red: calculated photoemission spectrum. In grey dashed lines: experimental emission spectrum digitalised from reference [60].

Relaxation of the molecular geometry on the first ten excited states converged on three  $S_1$  geometries (Appendix B.3). Two distinct minima were found in the  $S_1$  surface at energies of 4.37 eV and 4.49 eV over the FC energy (Figure 17). The first excited state minimum presents a very similar geometry to the FC structure and a  ${}^1(\pi\pi^*)$  character. The second minimum presents a slight distortion of the geometry of the nitrogen atom of the 6-membered ring towards the ring centre and presents a  ${}^1(n_N\pi^*)$  character. Three other distinct minima were found in the  $S_4$ ,  $S_8$ , and  $S_{10}$  states respectively, all of them at energies over 5.6 eV. Therefore, these were not considered relevant in the absorption energy range of interest. Figure 18 shows a schematic representation of the excited PES of beta-carboline.

The photoemission spectrum (Figure 16) and rate constants (Table 4) were calculated for the lowest-energy excited state minimum. The calculated photoemission spectrum reproduces the band pattern of the experimental emission spectrum almost perfectly, with a shift of 0.49 eV towards higher energies. The fluorescence quantum yield is slightly underestimated, with a value of 0.2062 compared to the experimental value of 0.30 [60].

Despite the small calculated fluorescence quantum yield, which could indicate a possible efficient deactivation pathway, previous experimental studies [60] have reported a phosphorescence quantum yield of 0.40. This indicates a total radiative quantum yield of  $\Phi_r = 0.70$ , indicating that this

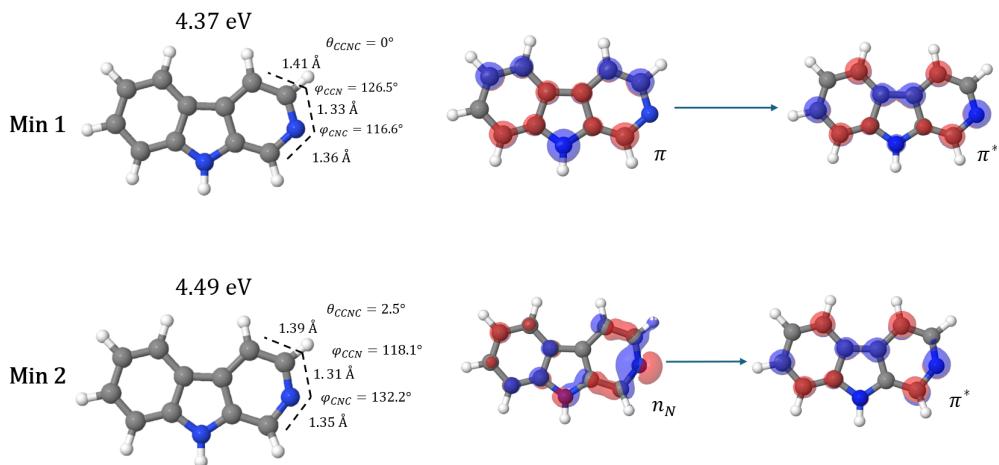


Figure 17: Optimised beta-carboline  $S_1$  excited state minima and electronic state character.

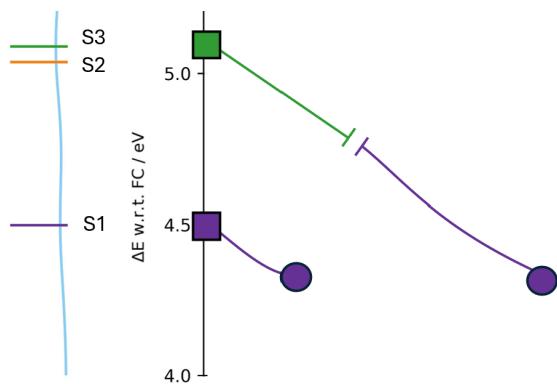


Figure 18: Schematic representation of the excited-state PES of beta-carboline. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Gaps indicate change in energy surface maintaining excited state character. Minima are marked as points. Energies relative to the FC energy.

molecule deactivates efficiently in a radiative manner.

The phosphorescence quantum yield value was calculated using the FC minimum and the  $T_1$  minimum, in order to compare with the experimental result. The calculated phosphorescence rate is  $1.72 \times 10^{-3} \text{ s}^{-1}$ . This rate results in a phosphorescence quantum yield of  $\Phi_p \approx 0$ , which indicates that the phosphorescence rates cannot be modelled appropriately this way for beta-carboline.

$k_r / \text{s}^{-1}$	$k_{IC} / \text{s}^{-1}$	$\Phi_f = \frac{k_r}{k_r + k_{ic}}$	Reference [60]
$4.4883 \times 10^7$	$1.727 \times 10^8$	0.2062	0.30

Table 4: Beta-carboline  $S_1$  rate constants and fluorescence quantum yield.

Starting from the excited state minima, an optimisation to locate MECIs was performed. However, no relevant conical intersections were found, as all the found MECIs were energetically unaccessible ( $\Delta E \geq 0.78 \text{ eV}$ ) from any minima.

The high experimental fluorescence and phosphorescence quantum yield, combined with the energetic inaccessibility of decay pathways via conical intersections, indicates that beta-carboline cannot act as a photoprotector itself. Nevertheless, further research on the effect of functionalization in this molecule should be considered. If substituent modifications could facilitate access to new energetically favourable the MECIs, beta-carboline could potentially act as a building block for a more sophisticated solar filter.

### 6.3 Tryptamine

The vertical and semiclassical absorption spectra were calculated using CAM-B3LYP/cc-pVDZ in its time dependent variant and within the TDA (Figure 19).

Tryptamine presents its first absorption at 5.14 eV, exhibiting a  ${}^1(\pi\pi^*)$  character (Appendix B.4). This transition, alongside the  $S_0 \rightarrow S_2$  transition at 5.21 eV ( ${}^1(\pi\pi^*)$ ), form a first band in the absorption spectrum. The second most intense band is located at 6.80 eV, where the contributions of the  $S_0 \rightarrow S_4$  ( ${}^1(\pi\pi^*)$ ) and  $S_0 \rightarrow S_6$  ( ${}^1(\pi\pi^*)$ ) transitions predominate. The semiclassical absorption spectrum presents a red shift of 0.43 eV in the most intense band while preserving the shape of the convoluted vertical absorption spectrum. The first absorption band in the semiclassical spectrum is located at 4.82 eV, 0.50 eV higher than the experimentally reported 4.32 eV peak [61].

Optimization of the ten first excited states of tryptamine resulted in four geometries on the  $S_1$  PES (Appendix B.4). Considering the geometries and character, two distinct minima in the  $S_1$  state were identified at 4.87 and 4.97 eV above the FC reference geometry. (Figure 20). These two minima

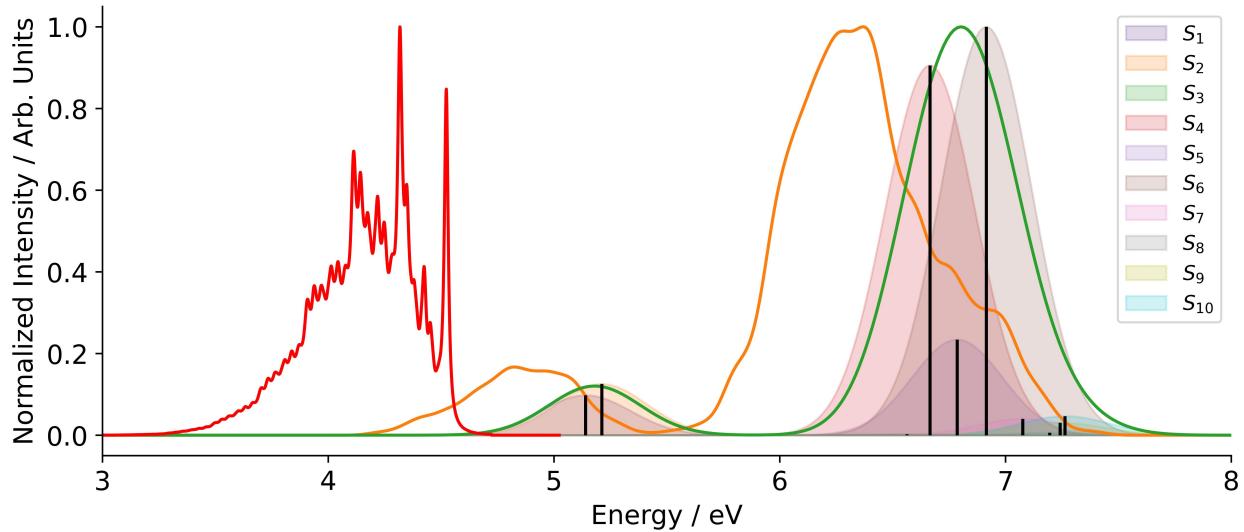


Figure 19: Absorption spectra of tryptamine. In orange: semiclassical absorption spectrum. In green: convoluted vertical absorption spectrum. In red: calculated photoemission spectrum. The vertical lines indicate the positions and intensities of vertical excitations.

present a similar geometry in the aromatic core to the FC geometry and a  ${}^1(\pi'\pi'^*)$  character, but show conformational differences in the side chain. Two additional minima in the  $S_5$  and  $S_7$  states were located, but they are not relevant for the low-lying excited state deactivation pathways. Figure 21 shows a schematic representation of the excited PES of tryptamine.

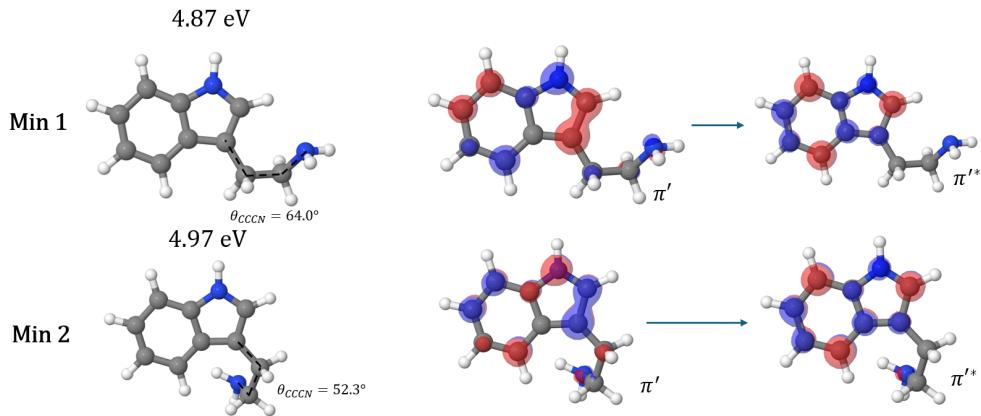


Figure 20: Optimized tryptamine  $S_1$  excited state minima and electronic state character.

The photoemission spectrum and rate constants were calculated for the lowest-energy excited state minimum. The resulting spectrum is shown in Figure 19. The calculated photoemission rate constant is  $k_r = 6.255 \times 10^7 \text{ s}^{-1}$  for the  $S_1$  minimum. Conversely, the calculated internal conversion rate is approximately zero. This indicates that the radiative deactivation of this system may be

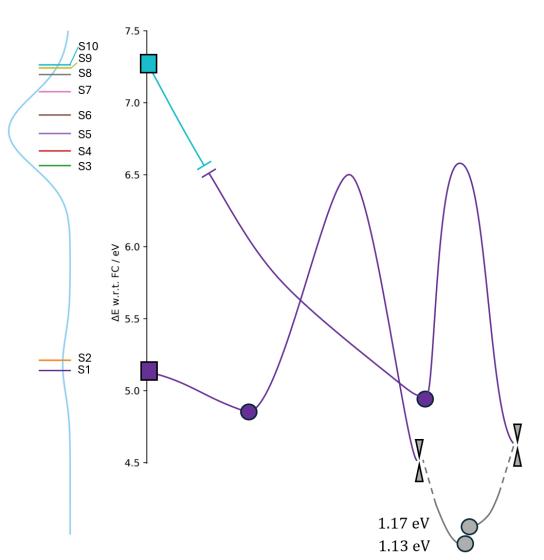


Figure 21: Schematic representation of the excited-state PES of tryptamine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones.

favoured.

Further optimization of the  $S_1$  PES resulted in the location of two  $S_1/S_0$  MECIs (Table 5). The MECI geometries are shown in Figure 22. The obtained MECIs present a notable difference with respect to the FC geometry, as a proton from the side chain is transferred to the aromatic core. However, both MECIs present a similar geometry, where the only notable difference is the rotation of the side chain.

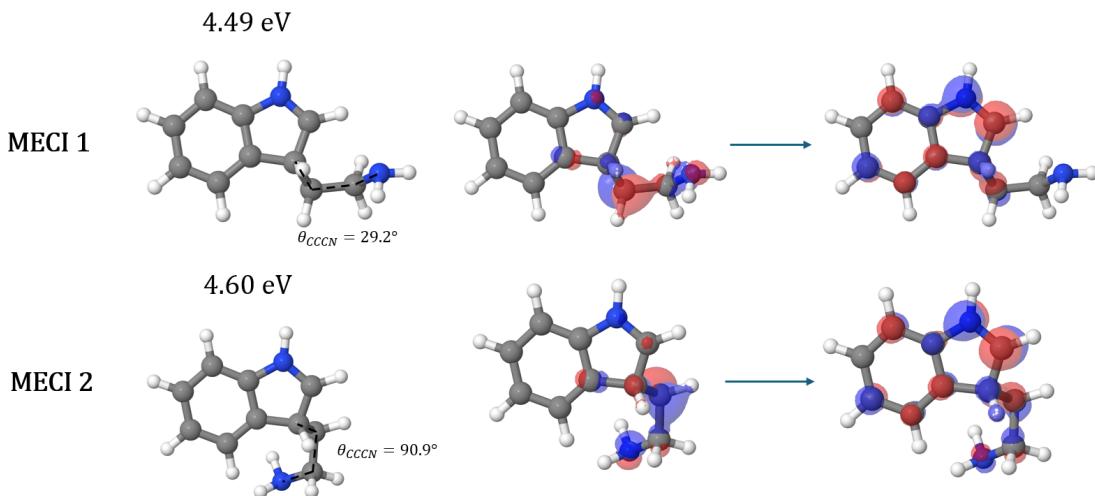


Figure 22: Caffeine optimised  $S_1/S_0$  MECI geometry and electronic state character.

Min	$\Delta E$ min-FC	$\Delta E$ MECI-FC	$\Delta E$ min-MECI	$\mathcal{P}$	$\mathcal{B}$	Type
1	4.87	4.49	-0.38	1.0939	0.0640	Sloped Bifurcating
2	4.97	4.60	-0.37	1.0495	0.2472	Sloped Bifurcating

Table 5: Tryptamine MECI  $\mathcal{P}$  and  $\mathcal{B}$  parameters and classification. Energies are given in eV.

One of the most interesting findings in this system is that the obtained MECIs not only involve a proton transfer from the side chain to the indole core, but are also lower in energy than the excited state minima identified. This strongly suggests the possibility of the formation of a photoproduct named 2-(Indolin-3-yl)ethanamine (Figure 23) along these pathways. Therefore, further research of the properties of this 2-(Indolin-3-yl)ethanamine, in order to determine its stability and absorption, is required.

The path between these minima and their associated MECI was determined through NEB calculations. Even though the MECIs are lower in energy, barriers in the  $S_1$  state of 1.64 eV and 1.75 eV are present. The MEP connecting minimum 1 and MECI 1 is represented in Figure B.4 while the MEP between minimum 2 and MECI 2 is omitted due to similarity (Appendix B.4). These barriers are associated with the increase in energy during proton transfer. The high energy of this barrier creates difficulties for the formation of 2-(Indolin-3-yl)ethanamine, as the system must overcome total energy barriers of 6.51 and 6.72 eV over the FC energy respectively, while the first absorptions are located at a lower energy. Therefore, the formation of 2-(Indolin-3-yl)ethanamine is expected to occur at higher energies than those considered for a solar filter.

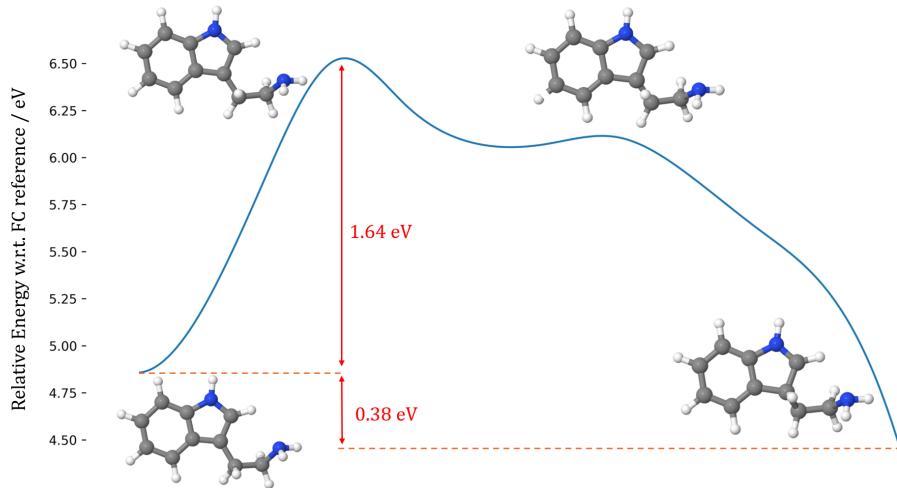


Figure 23: MEP pathway of minimum / MECI 1.

The obtained MECIs presented a bifurcating character. Therefore, the system directions after the intersection were analysed by displacing the geometries in the corresponding directions and relaxing the geometry in the  $S_0$  state. This relaxation results in two conformers of 2-(Indolin-3-yl)ethanamine, located 1.13 and 1.17 eV higher in energy than the original tryptamine molecule (Figure 21). This increase in photoproduct energy is likely due to the loss of stabilization granted by the aromaticity of the indole core due to the proton transfer.

The topology of the tryptamine potential energy landscape presents  $S_1$  excited state minima and low-energy CoIns, leading to deactivation pathway that involves a proton transfer indicating a possible route for 2-(Indolin-3-yl)ethanamine formation. However, the high energy barriers in these pathways suggest that 2-(Indolin-3-yl)ethanamine formation is unlikely under typical Ultra-Violet (UV) excitation conditions.

Future studies on tryptamine should include an investigation of the stability, absorption properties, and photoreactivity of 2-(Indolin-3-yl)ethanamine. The functionalisation of the tryptamine molecule may also have a major influence on the energies, barriers, and deactivation pathways, allowing for the tuning of its photophysical properties.

#### 6.4 Nicotine

The vertical and semiclassical absorption spectra were calculated using CAM-B3LYP/cc-pVDZ in its time dependent variant and within the TDA. Figure 24 compares the calculated semiclassical absorption spectrum with the experimental reference spectrum [62].

The vertical absorption spectrum presents a first weak absorption at 5.15 eV, presenting to a  $^1(n_N\pi^*)$  character (Appendix B.5). There are two weak absorption bands corresponding to transitions to the  $S_3$  and  $S_5$  states located at 5.65 and 6.63 eV, with  $^1(\pi\pi^*)$  and  $^1(n_N\pi^*)$  character respectively. The most intense absorption peak is associated with the  $S_0 \rightarrow S_9$  transition located at 7.91 eV.

Figure 24 shows that the semiclassical spectrum presents a red shift of 1.04 eV compared in the most intense band compared to the vertical spectrum. Although the semiclassical spectrum delivers a better agreement with the experimental spectrum, the main peak is still overestimated by 0.63 eV. Also, the band structure is not fully accurate, as the semiclassical spectrum displays a single broad first band, whereas the experimental data shows two distinct absorption bands.

Optimization of the ten first excited states of tryptamine resulted in four geometries in the  $S_1$  PES (Appendix B.5). Three distinct minima were found in the  $S_1$  manifold at 4.63, 4.68, and 4.69 eV above the FC reference geometry. All of these minima present a  $^1(n_N\pi^*)$  character, but the first two

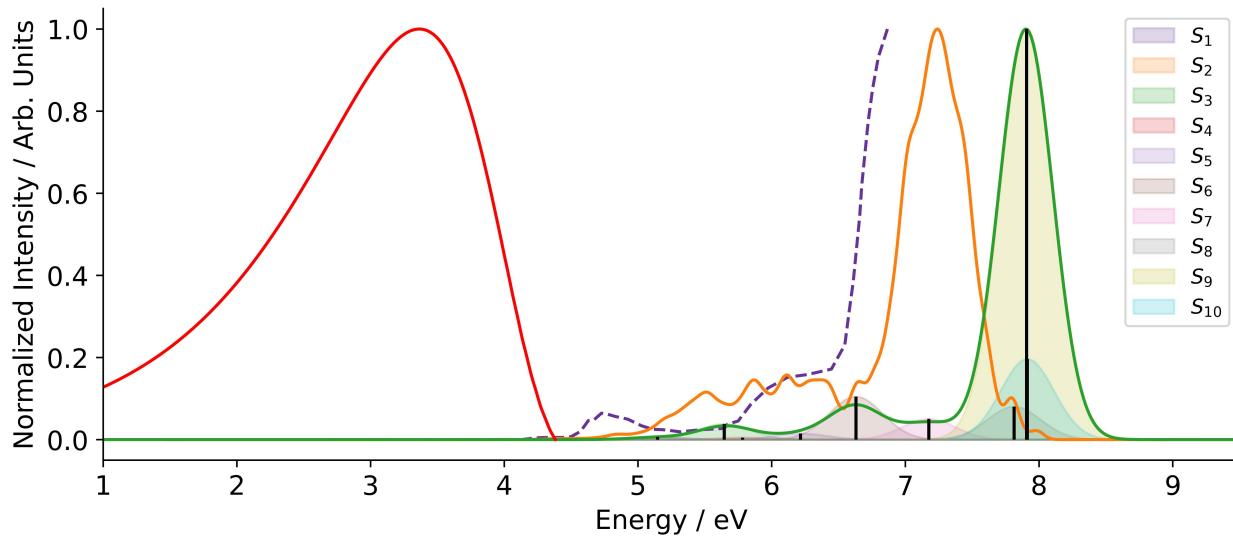


Figure 24: Absorption and emission spectra of nicotine. In orange, semiclassical absorption spectrum. In purple dashed lines: experimental absorption spectrum digitalised from reference [62]. In green, convoluted vertical absorption spectrum. Black lines indicate the positions and relative oscillator strength of vertical excitations. In red: calculated photoemission spectrum.

correspond to an excitation of the lone pair in the pyridine ring, while the third to an excitation of the lone pair in the five-member ring. It is possible to appreciate that the first two minima present a slight loss of planarity of the pyridine ring, and that the lone pair excitations occur to different  $\pi^*$  orbitals. Also, a rotation of the methyl group can be appreciated in the third minima, that doesn't take place in the first two minima. Two additional minima were found on the  $S_6$  and  $S_8$  surfaces with energies above 6 eV relative to the  $S_0$  FC reference (Appendix B.5). Figure 26 shows a schematic representation of the excited PES of nicotine.

The photoemission spectrum and rate constants were calculated for the lowest-energy excited state minimum. The resulting spectrum is shown in Figure 24, while the rate constants and quantum yield are presented in Table 6. The radiative rate constant is three orders of magnitude lower than the internal conversion rate constant, resulting in a quantum yield close to zero. This indicates that non-radiative decay is the dominant deactivation pathway in this system.

$k_r / s^{-1}$	$k_{IC} / s^{-1}$	$\Phi_f = \frac{k_r}{k_r+k_{ic}}$
$1.706 \times 10^{06}$	$3.568 \times 10^{09}$	0.00048

Table 6: Nicotine  $S_1$  rate constants and fluorescence quantum yield.

Starting from the excited state minima, optimisations were carried out to locate the corresponding MECIs. This procedure a single accessible  $S_1/S_0$  MECI (Table 7). The geometry and character

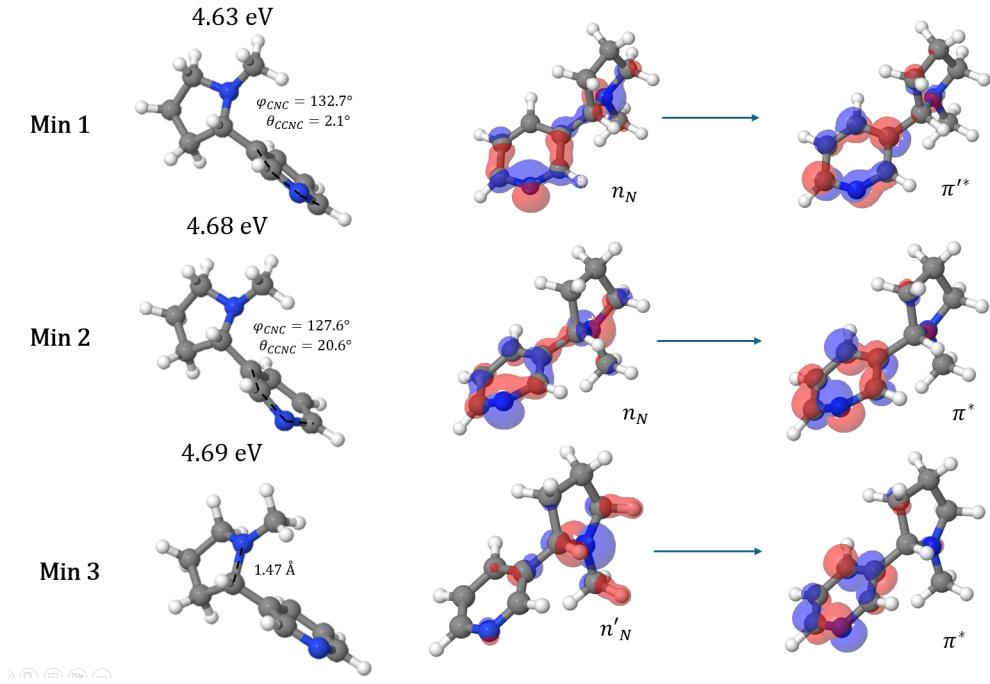


Figure 25: Optimised nicotine  $S_1$  excited state minima and electronic state character.

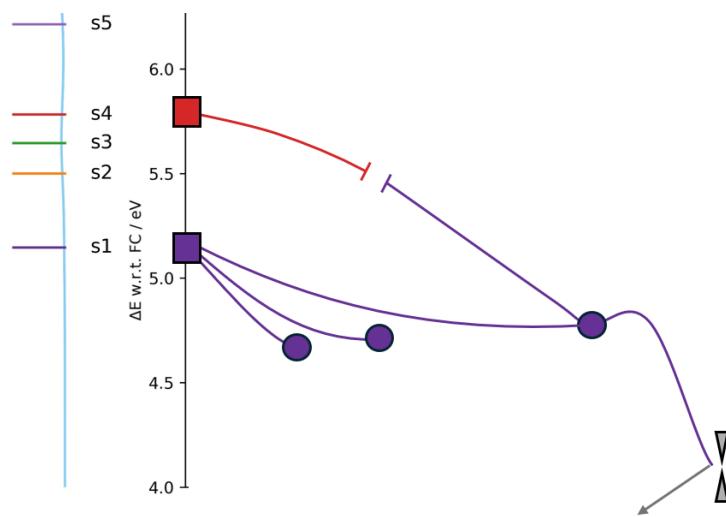


Figure 26: Schematic representation of the excited-state PES of nicotine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surfaces. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows at  $S_0/S_1$  MECIs indicate a return to FC geometry.

is shown in Figure 27. This MECI is characterised by the elongation of the  $C-N$  bond in the five-membered ring to 2.47 Å and a  ${}^1(n_N\pi^*)$ .

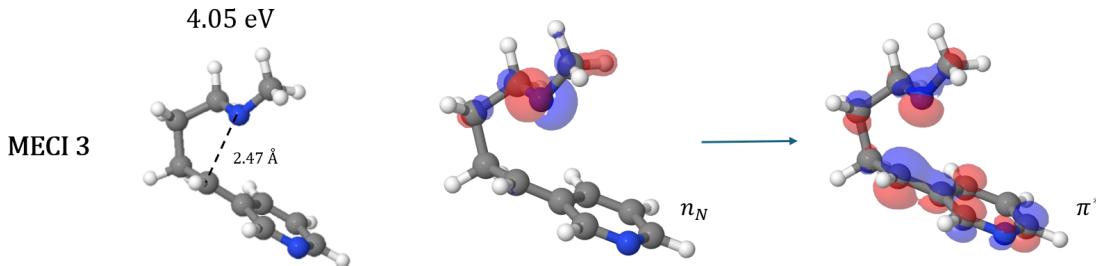


Figure 27: Nicotine optimised  $S_1/S_0$  MECI geometry and electronic state character.

Min	$\Delta E$ min-FC	$\Delta E$ MECI-FC	$\Delta E$ min-MECI	$\mathcal{P}$	$\mathcal{B}$	Type
3	4.69	4.05	-0.64	0.3399	0.0625	Peaked Bifurcating

Table 7: Tryptamine MECI  $\mathcal{P}$  and  $\mathcal{B}$  parameters and classification. Energies are given in eV.

The complete deactivation pathway was characterised by connecting the third minimum and the obtained MECI. There is an energetic barrier between the minimum and the MECI of 0.17 eV associated with the stretching of the  $C-N$  bond in the five-membered ring. However, the  $S_1$  absorption energy (5.15 eV) is higher than the total 4.86 eV barrier energy. As a consequence, this MECIs is expected to be accessible.

The obtained MECI presents a bifurcating topology. To determine the possible formation of photoproducts, the geometry was displaced in each minimum energy direction and optimised on the  $S_0$  surface. In both directions, the optimisation resulted in the return to the FC geometry, without any photoproduct formation.

Considering the previous results: the calculated rate constants, low quantum yield, an energetically accessible  $S_1/S_0$  intersection, and lack of formation of photoproducts, this suggests that nicotine may act as a photoprotector, as the non-radiative deactivation pathway is favoured and returns the molecule to its original structure.

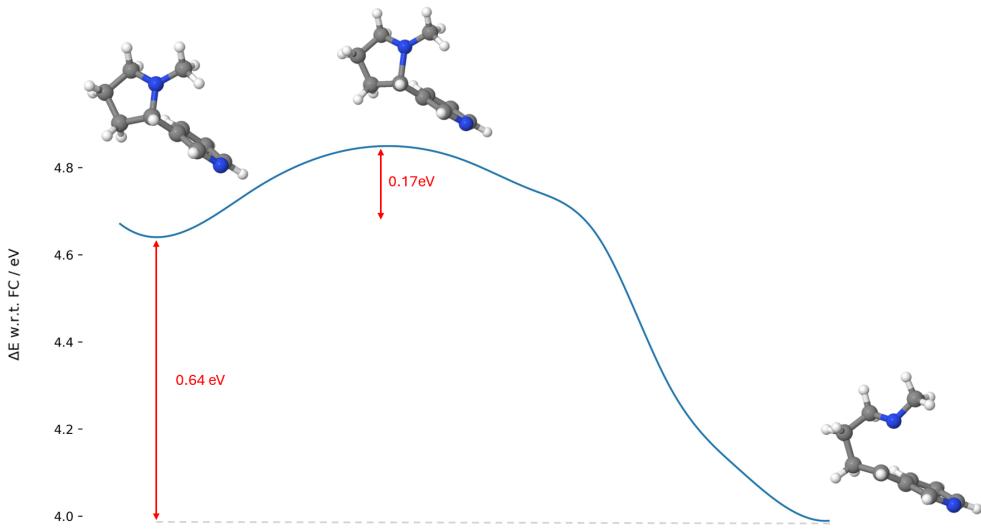


Figure 28: MEP pathway of minimum /MECI 3. (Placeholder, i will make the molecules consistent with the rest when i can)

## 6.5 Cytisine

The vertical and semiclassical absorption spectra were calculated using CAM-B3LYP/cc-pVDZ in its time dependent variant and within the TDA. Figure 29 shows the calculated spectra and the experimental absorption spectrum [63].

The first and most intense absorption is located at 4.71 eV and is associated with the  $S_0 \rightarrow S_1$  transition, which presents a  ${}^1(\pi\pi^*)$  character (Appendix B.6). There are two other less intense peaks corresponding to transitions to the  $S_5$  and  $S_{10}$  states located at 6.22 eV and 7.43 eV respectively, presenting a  ${}^1(n_O\pi^*)$  character.

The calculated semiclassical absorption spectrum presents a 0.50 eV red shift in the first band when compared with the vertical absorption spectrum. More notably, the band intensities change drastically, with the third band being the most intense. The semiclassical absorption spectrum overestimates the energy of the first and second bands by 0.16 eV when compared to the experimental spectrum [63]. While the semiclassical absorption spectrum presents two differentiated bands as in the experimental spectrum, the relative band intensities are not correctly described, underestimating the second band's intensity.

Optimisation of the FC structure in the first 10 excited states resulted in the location of a single  $S_1$  minimum shown in Figure 30. This minimum presents very similar geometry to the FC geometry and is located at 4.37 eV above the FC energy, presenting a  ${}^1(\pi\pi^*)$  character. In addition, four

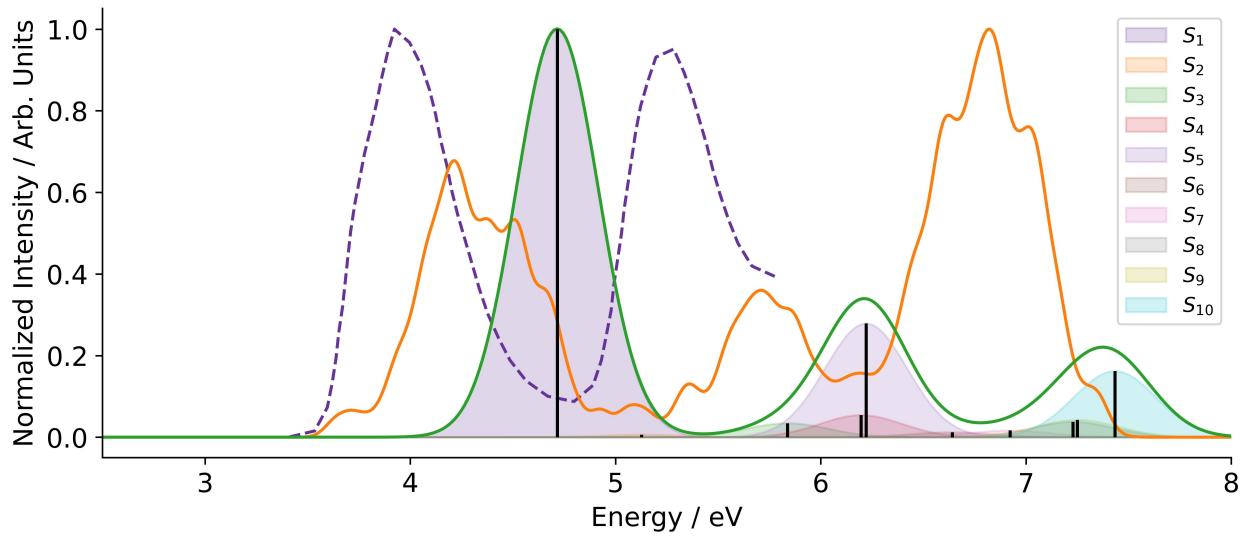


Figure 29: Absorption and emission spectra of cytisine. Left panel (a): in orange, semiclassical absorption spectrum. In purple: experimental data digitalised from [63]. In green, convoluted vertical absorption spectrum. Black lines indicate the positions and relative oscillator strengths of vertical excitations. Left panel (b): in red: calculated photoemission spectrum.

further excited state minima were found (Figure 31): one in the  $S_7$  PES (6.70 eV), two in the  $S_8$  PES (6.91 and 6.99 eV), and one in the  $S_9$  PES (7.11 eV). Due to the energy of the excitations being over 6.5 eV, these minima were not considered in the present work. Figure 31 shows a schematic representation of the excited PES of cytisine.

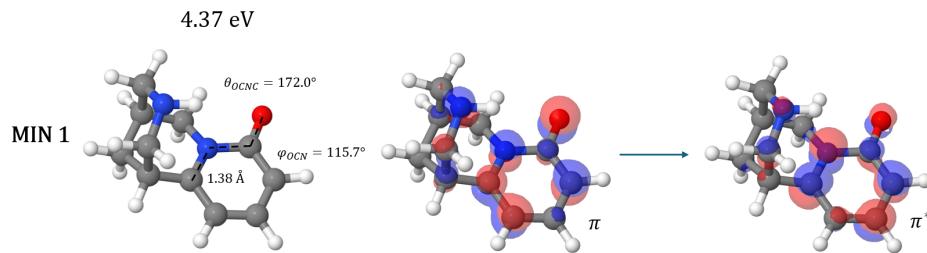


Figure 30: Cytisine optimized  $S_1$  excited state minimum and electronic state character.

Fluorescence and internal conversion rate constants were calculated using the  $S_1$  minimum conformation with the adiabatic Hessian model. Results are shown in Table 8. While the obtained results were calculated in the gas phase, there is experimental reference data for the fluorescence rate constant in acetonitrile [63], which is used here for comparison. The obtained rate constants are similar to experimental values in acetonitrile. The calculated fluorescence quantum yield in the gas phase is 0.195, which is larger than the  $7 \cdot 10^{-3}$  value in acetonitrile [63].

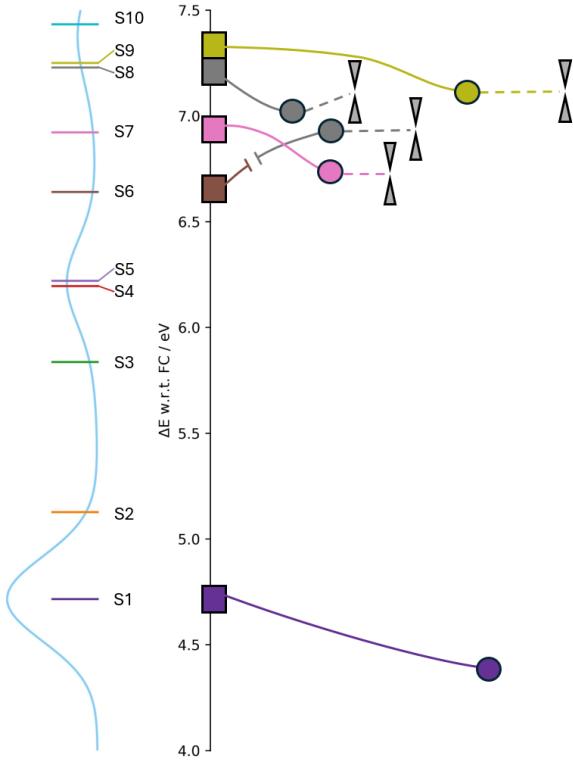


Figure 31: Schematic representation of the excited-state PES of cytisine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones.

$k_r / s^{-1}$	$k_{IC} / s^{-1}$	$\Phi_f = \frac{k_r}{k_r+k_{ic}}$	$k_r^{\text{Reference}} / s^{-1}$ (Acetonitrile) [63]
$7.0456 \cdot 10^7$	$2.901 \cdot 10^8$	0.195	$6.6 \cdot 10^7$

Table 8: Cytisine  $S_1$  rate constants and fluorescence quantum yield.

Starting from the  $S_1$  minimum, a MECI search was performed. However, no energetically accessible MECIs were located (4.37 eV over the  $S_1$  minimum energy).

Cytisine presents a low calculated fluorescence quantum yield, in agreement with previous experimental results (in acetonitrile), indicating that non-radiative decay should be favoured. However, no energetically accessible  $S_1/S_0$  funnels were found. Therefore, further research on this system is required to understand the deactivation mechanisms present in cytisine.

## 7 Conclusions

Six natural product molecules belonging to the alkaloid family were studied to determine their suitability as solar filters. These compounds possess aromatic systems and various substituents. Specifically, the studied compounds were caffeine, xanthine, beta-carboline, tryptamine, nicotine, and cytisine.

The first absorption bands of all the studied systems were located in the 4.49 eV (beta-carboline) to 5.34 eV (xanthine) range in gas phase. This energy range corresponds to 276–232 nm, at the lower energy end of UVC region. The absorption bands of all the compounds were attributed to  $\pi \rightarrow \pi^*$  transitions localized in the aromatic cores. Semiclassical absorption spectra were calculated, improving the resolution of band structure and energy range, showing generally good agreement with available experimental results, while still slightly overestimating absorption energies.

Fluorescence rate constants between an  $S_1$  excited-state minimum and the ground-state geometry were calculated in all the systems using the adiabatic Hessian model. The obtained fluorescence quantum yields were  $\leq 0.2$  in xanthine, beta-carboline, nicotine, and cytisine, indicating that these compounds favour non-radiative deactivation processes. However, the limitations of the adiabatic Hessian model should be considered, as conformational changes in the excited state geometry may lead to an underestimation of the fluorescence quantum yield.

To explore the excited-state potential energy surfaces, a tool was developed to optimise geometries towards minimum energy conical intersections, based on the Updated Branching Plane (UBP) algorithm. This tool enabled the identification of several Conical intersection (CoIn)s in the studied systems. These intersections generally exhibited very low energy differences between states of  $<0.005$  eV, reaching the CoIn seam, and further relaxing the energy along the seam towards Minimum energy conical intersection (MECI)s.

The CoIn characterisation algorithm based on the  $\mathcal{P}$  and  $\mathcal{B}$  parameters was also successfully implemented. A benchmark carried out on the ethylene system yielded exactly the expected parameters (up to numerical error) compared with the reference [27]. The implementation was also tested against OPENMOLCAS' [64] implementation, yielding identical results.

Considering the previous point, it is notable that all the CoIns found in the studied systems were bifurcating conical intersections, even when no photoproduct formation is expected. Furthermore, the characterisation of the CoIns resulted in some cases in intersections which could not be rotated in such a way that  $0 < \Delta_{gh}$  and  $[0 < \theta_s < \frac{\pi}{2}]$  were fulfilled (Section 4.2), which is expected. As the

characterisation implementation matches previous literature, there are two main possibilities for this overestimation of bifurcating intersections based on the approximations made. A possibility is that the approximation of the Coupling Derivative Vector (CDV) is not accurate enough, leading the geometry towards degeneracy, but not to the MECI. Another possibility is that Time-Dependent Density Functional Theory (TD-DFT) could define inaccurately the branching plane vectors in  $S_0/S_1$  intersections, leading to wrong characterisation parameters. However, the CoIns obtained present degeneracy, and their energy was decreased until the imposed effective gradient convergence threshold. Therefore, these can be considered a good approximation to the real MECI, and as a starting point for further studies with more accurate electronic-structure methods.

In various cases, the energy of the converged conical intersections was much higher than the absorption energy, such as in beta-carboline and cytisine. The rest of the systems presented conical intersections close in energy to the  $S_1$  excited-state minima. The critical points in the excited-state topology of the systems were connected when the energy differences were in this range, presenting energy barriers between the minima and conical intersections. Also, a study of possible photoproduct formation was performed in all the systems. In caffeine, xanthine, and tryptamine photoproduct formation is possible if the energetic barriers separating the  $S_1$  excited-state minima and the conical intersections can be overcome. However, these barriers present total energies over the UVA and UVB energy range.

Xanthine presented a low calculated fluorescence quantum yield, and an energetically accessible conical intersection, with a small barrier of 0.37 eV to access the funnel from the  $S_1$  minimum. The conical intersection was characterised as bifurcating, and indicates the possibility of formation of a photoproduct involving the opening of the five-membered ring. Further investigation of this photoproduct should be considered. If it demonstrates the expected solar filter properties, xanthine could be a promising candidate for use as a solar filter, as it offers an accessible non-radiative deactivation pathway.

The effect of xanthine methylation was studied by analysing the caffeine molecule. Caffeine presented slightly lower absorption energies in the semiclassical spectrum than xanthine ( $\approx 0.1$  eV). The conical intersection found was located at 0.08 eV over the connected  $S_1$  excited-state minimum. It presented a higher energy barrier of 0.54 eV, larger than the one of the accessible xanthine conical intersection. The total energy of this barrier surpasses the first absorption energies, which indicates that this conical intersection is not accessible, and therefore caffeine is not suitable as a solar filter.

Tryptamine presents a near-unity fluorescence quantum yield. This is consistent with the large barriers of 1.64 and 1.75 eV over their respective  $S_1$  minima. This deems the conical intersections

inaccessible, even if they present lower energies than the excited state minima, indicating that tryptamine is not suitable as a solar filter. Still, photoproduct formation is possible at high energies, leading to the formation of 2-(Indolin-3-yl)ethanamine, where a proton is transferred from the side chain to the aromatic core, breaking the aromaticity of the indole system.

Nicotine presented a low fluorescence quantum yield, and an accessible conical intersection with a barrier of 0.17 eV above the excited state minimum. The total energy barrier is smaller than the first absorption energy. Also, no formation of photoproducts is expected, returning to the ground state geometry. This indicates that nicotine can be a good candidate as a solar filter, as it presents an accessible non-radiative deactivation pathway.

Beta-carboline presented a low calculated fluorescence quantum yield. However, previous literature indicates that beta-carboline has a high phosphorescence quantum yield. This fact is consistent with the lack of energy accessible MECIs found. This indicates that beta-carboline is not suitable as a solar filter, as it does not present an accessible non-radiative deactivation pathway.

Finally, even though cytisine presented a low fluorescence quantum yield, the no accessible MECIs were located. Therefore, further research is needed to understand the underlying deactivation processes and the suitability of cytisine as a solar filter.

From the studied systems, xanthine and nicotine are the best candidates as solar filters, as they present low fluorescence quantum yields and accessible non-radiative deactivation pathways. Caffeine, beta-carboline, tryptamine and cytisine are not suitable as solar filters, as they do not present accessible non-radiative deactivation pathways. However, as observed in the caffeine system, functionalisation can significantly tune the optical properties and excited state topology, and further study of the functionalisation of these systems could lead to new candidates as solar filters. Also, the study of the properties of potential photoproducts could reveal their potential as solar filters.

A limitation in the current work is the lack of symmetry-breaking of the planar structures in these systems. Currently, a tool to analyse the internal coordinates of a system and displace only ring coplanar atoms is in development. Therefore, further work considering deplanarization should be performed, as it could lead to the discovery of new conical intersections.

The results obtained in this work provide a foundation for further exploration of natural alkaloids and their derivatives as potential solar filters. Future research could include dynamical studies to explore the non-radiative deactivation pathways in more detail, shading light into potential competing deactivation mechanisms. A higher level of theory could also be employed to obtain a more accurate description of the excited-state potential energy surfaces and conical intersections.

## References

- (1) Nordlund, J. J. *Dermatologic Clinics* **2007**, *25*, Pigmentary Disorders, 271–281.
- (2) D’Orazio, J.; Jarrett, S.; Amaro-Ortiz, A.; Scott, T. *International Journal of Molecular Sciences* **2013**, *14*, 12222–12248.
- (3) Cardinali, G.; Bolasco, G.; Aspite, N.; Lucania, G.; Lotti, L. V.; Torrisi, M. R.; Picardo, M. *Journal of Investigative Dermatology* **2008**, *128*, 558–567.
- (4) ITO, S.; WAKAMATSU, K.; OZEKI, H. *Pigment Cell Research* **2000**, *13*, 103–109.
- (5) Nordlund, J. J. *Dermatologic Clinics* **2007**, *25*, 271–281.
- (6) Boissy, R. E. *Experimental Dermatology* **2003**, *12*, 5–12.
- (7) Sklar, L. R.; Almutawa, F.; Lim, H. W.; Hamzavi, I. *Photochemical and Photobiological Sciences* **2013**, *12*, 54–64.
- (8) Shetty, N.; Schalka, S.; Lim, H. W.; Mohammad, T. F. *Photochemical and Photobiological Sciences* **2023**, *22*, 2463–2471.
- (9) Schlumpf, M.; Cotton, B.; Conscience, M.; Haller, V.; Steinmann, B.; Lichtensteiger, W. *Environmental Health Perspectives* **2001**, *109*, 239–244.
- (10) Kim, S.; Choi, K. *Environment International* **2014**, *70*, 143–157.
- (11) Jonasson, S.; Gustafsson, Å.; Koch, B.; and, A. B. *Inhalation Toxicology* **2013**, *25*, PMID: 23480194, 179–191.
- (12) Danovaro, R.; Bongiorni, L.; Corinaldesi, C.; Giovannelli, D.; Damiani, E.; Astolfi, P.; Greci, L.; Pusceddu, A. *Environmental Health Perspectives* **2008**, *116*, 441–447.
- (13) Woolley, J. M.; Losantos, R.; Sampedro, D.; Stavros, V. G. *Physical Chemistry Chemical Physics* **2020**, *22*, 25390–25395.
- (14) Widsten, P.; Tamminen, T.; Liitiä, T. *ACS Omega* **2020**, *5*, 13438–13446.
- (15) Klessinger, M.; Michl, J., *Excited States and Photochemistry of Organic Molecules*; VCH: 1995.
- (16) In *Medicinal Natural Products*; John Wiley and Sons, Ltd: 2009; Chapter 6, pp 311–420.
- (17) Carey, F.; Sundberg, R., *Advanced Organic Chemistry: Part A: Structure and Mechanisms*; Advanced Organic Chemistry; Springer US: 2007.
- (18) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, First; Dover Publications, Inc.: Mineola, 1996.
- (19) Cohen-Tannoudji, C.; Diu, B.; Laloë, F., *Quantum mechanics; 1st ed.* Trans. of : Mécanique quantique. Paris : Hermann, 1973; Wiley: New York, NY, 1977.

- (20) Jensen, F., *Introduction to Computational Chemistry*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006.
- (21) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.
- (22) Kohn, W.; Sham, L. J. *Self-Consistent Equations Including Exchange and Correlation Effects*; tech. rep.; 1965, A1133–A1138.
- (23) Holthausen, M. C.; Koch, W., *A Chemist's Guide to Density Functional Theory*; John Wiley & Sons, Inc.: 2001.
- (24) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.
- (25) Casida, M. E., *Time-Dependent Density Functional Response Theory for Molecules*.
- (26) Huix-Rotllant, M.; Ferré, N.; Barbatti, M., *Time-Dependent Density Functional Theory*; John Wiley and Sons, Ltd: 2020.
- (27) Fdez. Galván, I.; Delcey, M. G.; Pedersen, T. B.; Aquilante, F.; Lindh, R. *Journal of Chemical Theory and Computation* **2016**, *12*, PMID: 27327873, 3636–3653.
- (28) Do Casal, M. T.; Veys, K.; Bousquet, M. H. E.; Escudero, D.; Jacquemin, D. *The Journal of Physical Chemistry A* **2023**, *127*, PMID: 37988002, 10033–10053.
- (29) Matsika, S. *Chemical Reviews* **2021**, *121*, 9407–9449.
- (30) Perdew, J. P.; Schmidt, K. *AIP Conference Proceedings* **2001**, *577*, 1–20.
- (31) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (32) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (33) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (34) Perdew, J. P. *MRS Bulletin* **2013**, *38*, 743–750.
- (35) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (36) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.
- (37) Adamo, C.; Barone, V. *The Journal of Chemical Physics* **1999**, *110*, 6158–6170.
- (38) Grimme, S. *The Journal of Chemical Physics* **2006**, *124*, 034108.
- (39) Yanai, T.; Tew, D. P.; Handy, N. C. *Chemical Physics Letters* **2004**, *393*, 51–57.
- (40) Chai, J.-D.; Head-Gordon, M. *The Journal of Chemical Physics* **2008**, *128*, 084106.
- (41) Sakurai, J. J.; Napolitano, J., *Modern Quantum Mechanics*, 3rd ed.; Cambridge University Press: 2020.
- (42) Dunning Thom H., J. *The Journal of Chemical Physics* **1989**, *90*, 1007–1023.
- (43) Engel, T.; Reid, P., *Physical Chemistry: Pearson New International Edition PDF eBook*; Pearson Education: 2013.
- (44) Omary, M. A.; Patterson, H. H. In *Encyclopedia of Spectroscopy and Spectrometry*, Lindon, J. C., Ed.; Elsevier: Oxford, 1999, pp 1186–1207.

- (45) Sicilia, F.; Blancafort, L.; Bearpark, M. J.; Robb, M. A. *Journal of Chemical Theory and Computation* **2008**, *4*, 257–266.
- (46) Maeda, S.; Ohno, K.; Morokuma, K. *Journal of Chemical Theory and Computation* **2010**, *6*, 1538–1545.
- (47) Farfan, C. A.; Turner, D. B. *Phys. Chem. Chem. Phys.* **2020**, *22*, 20265–20283.
- (48) Cerezo, J.; Santoro, F. *Journal of Computational Chemistry* **2023**, *44*, 626–643.
- (49) Bearpark, M. J.; , M. A. R.; Schlegel, H. B. *Chemical Physics Letters* **1994**, *223*, 269–274.
- (50) Shanno, D. F. **1970**, *24*.
- (51) Neese, F. *WIREs Comput. Molec. Sci.* **2012**, *2*, 73–78.
- (52) Bitzek, E.; Koskinen, P.; Gähler, F.; Moseler, M.; Gumbsch, P. *Physical Review Letters* **2006**, *97*, DOI: 10.1103/PhysRevLett.97.170201.
- (53) Hjorth Larsen, A. et al. *Journal of Physics: Condensed Matter* **2017**, *29*, 273002.
- (54) Barbatti, M. et al. *Journal of Chemical Theory and Computation* **2022**, *18*, PMID: 36194696, 6851–6865.
- (55) Frisch, M. J. et al. Gaussian 16, Gaussian Inc. Wallingford CT, 2016.
- (56) Shao, Y. et al. *Molecular Physics* **2015**, *113*, 184–215.
- (57) Hanson, R. M. *Journal of Applied Crystallography* **2010**, *43*, 1250–1260.
- (58) Navarra, G.; Moschetti, M.; Guerrasi, V.; Mangione, M. R.; Militello, V.; Leone, M. *Journal of Chemistry* **2017**, *2017*, 6435086.
- (59) Reyman, D.; Pardo, A.; Poyato, J. M. L. *The Journal of Physical Chemistry* **1994**, *98*, 10408–10411.
- (60) Varela, A. P.; Burrows, H. D.; Douglas, P.; da Graça Miguel, M. *Journal of Photochemistry and Photobiology A: Chemistry* **2001**, *146*, 29–36.
- (61) Carney, J. R.; Zwier, T. S. *The Journal of Physical Chemistry A* **2000**, *104*, 8677–8688.
- (62) Clayton, P. M.; Vas, C. A.; Bui, T. T.; Drake, A. F.; McAdam, K. *Chirality* **2013**, *25*, 288–293.
- (63) Krystkowiak, E.; Przybył, A. K.; Bayda-Smykaj, M.; Koput, J.; Maciejewski, A. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2018**, *203*, 375–382.
- (64) Li Manni, G. et al. *Journal of Chemical Theory and Computation* **2023**, *19*, PMID: 37216210, 6933–6991.

# A Appendix: Computational implementation

All the code of the homemade CoIn optimization and characterisation tool can be accessed in this GitHub repository.

## Input file

```

1 geom= cytisine_min_1_1_1.xyz
2
3 mode = ci
4 profile = ubp
5 calc_nacme = False
6
7 functional = CAM-B3LYP
8 basis = cc-pvdz
9 iroot = 0
10 jroot = 1
11 nroots = 10
12
13 charge = 0
14 mult = 1
15 program = 'ORCA'
16 label = None
17 convergence = 'natural'
18
19 slow_start = True

```

## (Simplified) Updated Branching Plane implementation

```

1 def ubp_forces(self):
2     iteration = int(np.loadtxt('iteration.dat'))
3     if iteration == 1:
4         en1, en2, _ = np.loadtxt("energies.dat")
5     else:
6         enerline = np.loadtxt("energies.dat")[-1]
7         en1, en2, _ = enerline
8
9     # management of the gradients
10    grad1 = np.loadtxt("engrad0.dat").reshape(-1)
11    grad2 = np.loadtxt("engrad1.dat").reshape(-1)

```

```

13     # check if previous iteration values are in memory
14
15     try:
16
16         x_minus_one = np.loadtxt('x_minus_one.dat')
17         y_minus_one = np.loadtxt('y_minus_one.dat')
18
19         first_iter = False
20
21     except:
22
23         first_iter = True
24
25
26     # x is the unit vector parallel to the difference gradient vector
27     x = (grad1-grad2).reshape(-1)
28     x /= np.linalg.norm(x)
29
30
31     # the definition of y can be: first step, updated, exact
32     if first_iter:
33
34         y_space = scypi.linalg.null_space(x)
35
36         y = y_space[:, 0]
37
38
39     elif not self.calc_nacme:
40
41         y = (np.dot(y_minus_one, x) * x_minus_one - np.dot(x_minus_one, x) *
42             y_minus_one) / (np.dot(y_minus_one, x)**2 + np.dot(x_minus_one, x)**2)**0.5
43
44         y /= np.linalg.norm(y)
45
46
47     else:
48
49         y = np.loadtxt('nacme.dat').reshape(-1)
50
51         y /= np.linalg.norm(y)
52
53
54     # generate the effective gradients
55     g_diff = 2 * (en1 - en2) * x
56     mean_grad = (grad1 + grad2)/2
57
58
59     # Generate the projector
60     P = np.identity(len(grad1.reshape(-1))) - np.outer(x,x) - np.outer(y,y)
61
62
63     # calculate the total effective gradient with the projector
64     total_gradient = iter_weight * g_diff.reshape(-1) + P @ mean_grad.reshape(-1)
65     total_gradient = total_gradient.reshape([-1,3])
66
67
68     # feed the optimizer with energies and gradients
69     print(total_gradient)
70
71     self.results["energy"] = en2
72
73     self.results["forces"] = - total_gradient * (ase.units.Hartree / ase.units.Bohr
74 )

```

## (Simplified) Conical intersection characterization

```
1 class CiCharacterizer:
2
3     def __init__(self,
4                  ga_filename: str,
5                  gb_filename: str,
6                  hab_filename: str,
7                  ci_energy: float = 0.0
8 ) -> None:
9
10        self._ga_filename = ga_filename
11        self._gb_filename = gb_filename
12        self._hab_filename = hab_filename
13        self._ci_energy = ci_energy
14        self.rotation_counter = 0
15
16    self._init()
17
18    def __init__(self) -> None:
19        self._load_vectors()
20
21    for i in range(0, 4):
22        if self.is_rotation_needed(self.asymmetry, self.theta_s):
23            self._rotate_for_beta()
24        else:
25            break
26
27    @property
28    def ci_energy(self) -> float:
29        return float(self._ci_energy)
30
31    def _load_vectors(self) -> None:
32        self._ga = np.loadtxt(self._ga_filename).reshape(-1)
33        self._gb = np.loadtxt(self._gb_filename).reshape(-1)
34        self._h_ab = np.loadtxt(self._hab_filename).reshape(-1)
35
36    @cached_property
37    def g_ab(self) -> np.ndarray:
38        return 0.5 * np.copy(self._gb - self._ga)
39
40
41    @cached_property
```

```
42     def s_ab(self) -> np.ndarray:
43         return 0.5 * np.copy(self._gb + self._ga)
44
45     @cached_property
46     def h_ab(self) -> np.ndarray:
47         return np.copy(self._h_ab)
48
49     @cached_property
50     def _pre_beta(self) -> np.array:
51         beta = 0.5 * np.arctan2(2 * np.dot(self.g_ab, self.h_ab), (np.dot(self.g_ab,
52             , self.g_ab) - np.dot(self.h_ab, self.h_ab)))
53
53         beta_2 = np.arctan2(
54             2 * np.dot(self.g_ab, self.h_ab),
55             np.dot(self.g_ab, self.g_ab) - np.dot(self.h_ab, self.h_ab)
56         )
57
58         beta = beta_2/2
59
60         return [beta + i*np.pi/2 for i in range(0, 4)]
61
62     def _rotate_for_beta(self):
63         self.rotation_counter += 1
64
65     @property
66     def beta(self) -> float:
67         return self._pre_beta[self.rotation_counter % 4]
68
69     def is_rotation_needed(self, asymmetry:float, theta_s:float) -> bool:
70         return not (asymmetry > 0 and theta_s > -10**-12 and theta_s < np.pi/2)
71
72     @property
73     def _g_tilde(self) -> np.ndarray:
74         return self.g_ab * np.cos(self.beta) + self.h_ab * np.sin(self.beta)
75
76     @property
77     def _h_tilde(self) -> np.ndarray:
78         return self.h_ab * np.cos(self.beta) - self.g_ab * np.sin(self.beta) # original
79
80     @property
81     def x(self) -> np.ndarray:
```

```
82     return np.copy(self._g_tilde / np.linalg.norm(self._g_tilde))
83
84     @property
85     def y(self) -> np.ndarray:
86         return np.copy(self._h_tilde / np.linalg.norm(self._h_tilde))
87
88     @property
89     def pitch(self) -> float:
90         ''' Pitch \delta_{gh}.'''
91
92         return np.sqrt(1/2 * (np.dot(self._g_tilde, self._g_tilde) + np.dot(self._h_tilde, self._h_tilde)))
93
94     @property
95     def asymmetry(self) -> float:
96         ''' Asymmetry \Delta_{gh}.'''
97
98         asym = (np.dot(self._g_tilde, self._g_tilde) - np.dot(self._h_tilde, self._h_tilde)) / (np.dot(self._g_tilde, self._g_tilde) + np.dot(self._h_tilde, self._h_tilde))
99
100        return float(asym) # original
101
102    def energy_difference(self, x:float, y:float) -> float:
103        return 2 * self.pitch * np.sqrt((x**2 + y**2) + self.asymmetry * (x**2 - y**2))
104
105    def average_energy(self, x:float, y:float) -> float:
106        return self.ci_energy + x * np.dot(self.s_ab, self.x) + y * np.dot(self.s_ab, self.y)
107
108    def E_A(self, x:float, y:float) -> float:
109        s_x = np.dot(self.s_ab, self.x) / self.pitch
110        s_y = np.dot(self.s_ab, self.y) / self.pitch
111        # return x + y
112        return self.ci_energy + self.pitch * (x * s_x + y * s_y - ((x**2 + y**2) + self.asymmetry * (x**2 - y**2))**0.5)
113
114    def E_B(self, x:float, y:float) -> float:
115        diff = self.energy_difference(x, y)
116        aver = self.average_energy(x, y)
117        return aver + diff
```

```
118
119     @property
120     def theta_s(self, n_points:int = 1800, radius:float=1):
121
122         angles = np.linspace(0, 2 * np.pi, n_points, endpoint=False)
123         x_points = radius * np.cos(angles)
124         y_points = radius * np.sin(angles)
125
126         pairs = [[x_points[i], y_points[i]] for i in range(n_points)]
127
128         samples = []
129
130         for pair in pairs:
131             x, y = pair
132             sample = [x, y, self.average_energy(x, y)]
133             samples.append(sample)
134
135         x,y,z = max(samples, key=lambda item: item[2])
136
137         vec = np.array([x,y])
138         theta = np.arctan2(y,x)
139
140         return theta
141
142     @property
143     def sigma(self):
144         s_x = np.dot(self.s_ab, self.x) / self.pitch
145         s_y = np.dot(self.s_ab, self.y) / self.pitch
146
147         return np.sqrt(s_x**2 + s_y**2)
148
149     @property
150     def p(self) -> Tuple[float, str]:
151
152         p = self.sigma**2 / (1 - self.asymmetry**2) * (1 - self.asymmetry * np.cos(2*self.theta_s))
153
154         p_type = 'Peaked' if p < 1 else 'Sloped'
155
156         return (p, p_type)
157
158     @property
```

```
159     def b(self) -> Tuple[float, str]:  
160  
161         b = (self.sigma**2/(4*self.asymmetry**2)) **(1/3) * (((1+self.asymmetry)*np  
162             .cos(self.theta_s)**2)**(1/3) + ((1-self.asymmetry)*np.sin(self.theta_s)**2)  
163             **(1/3))  
164  
165         b_type = 'Bifurcating' if b < 1 else 'Single Path'  
166  
167         return (b, b_type)
```

## B Appendix: Additional Figures and Tables

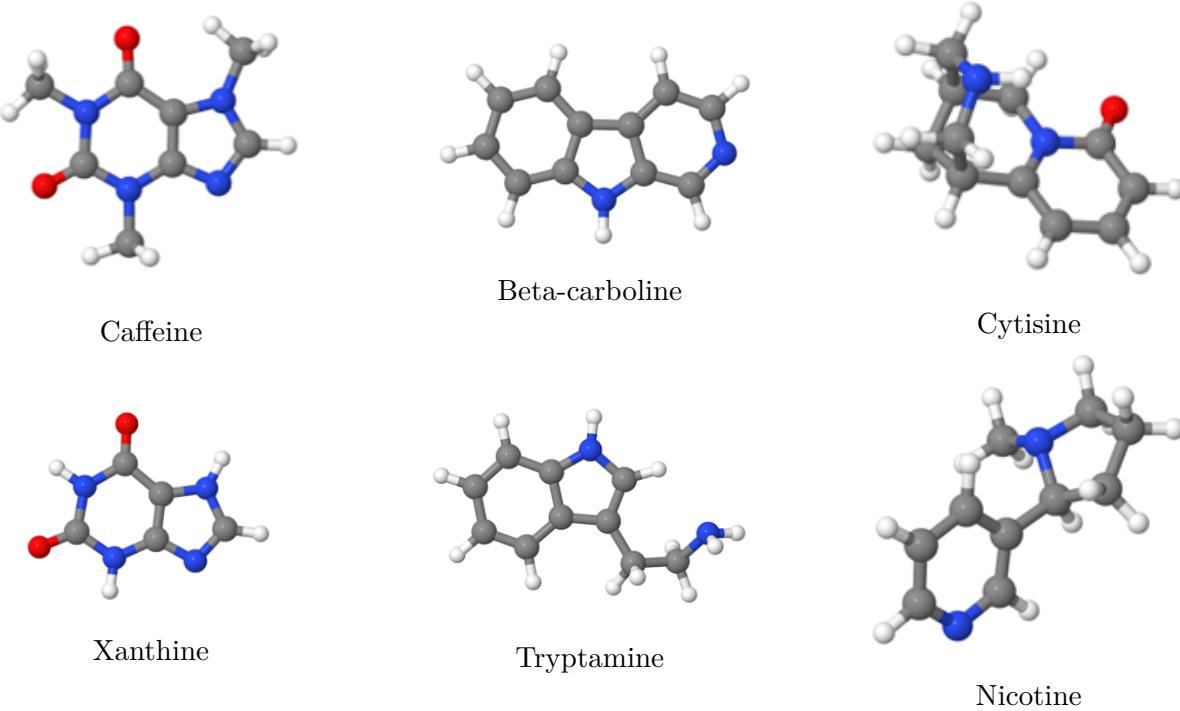


Figure 32: Optimized ground state structures of caffeine, beta-carboline, cytisine, xanthine, tryptamine, and nicotine.

### B.1 Xanthine

**FC geometry NTOS**

**Minima excited state character**

**MECI excited state character**

Starting root	Minimum	$\Delta E$ min-FC / eV	$\Delta E$ MECI-FC / eV	$\Delta E$ min-MECI / eV
1	1	4.85	7.18	+2.33
9	2	4.86	7.89	+3.03
1	3	4.90	4.69	-0.21
2	4	5.23	7.96	+2.73

Table 9: Relative energies of additional xanthine  $S_1$  geometries and optimized Minimum energy conical intersection (MECI)s with respect to the FC reference energy.

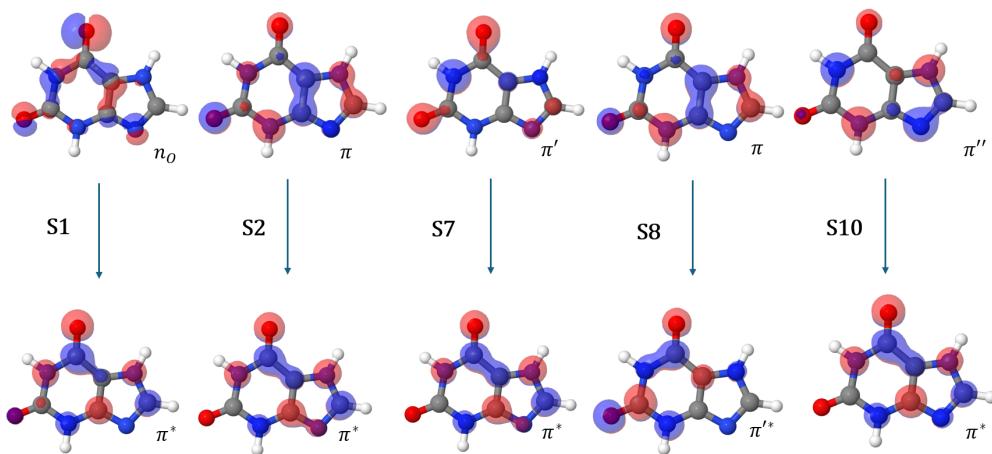


Figure 33: Natural Transition Orbitals (NTOs) of Xanthine at the FC geometry.

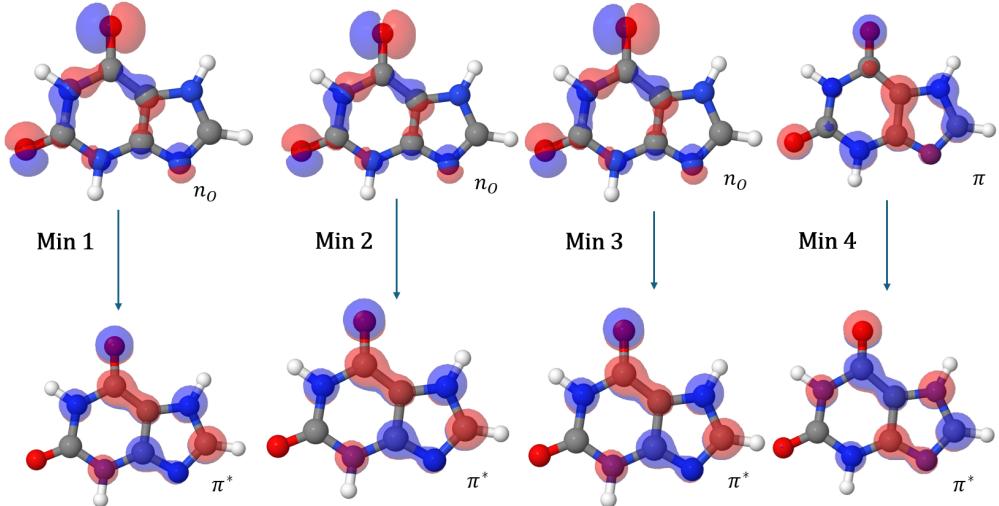


Figure 34: Transition orbitals of Xanthine at the optimised excited-state geometries.

## B.2 Caffeine

### FC geometry NTOs

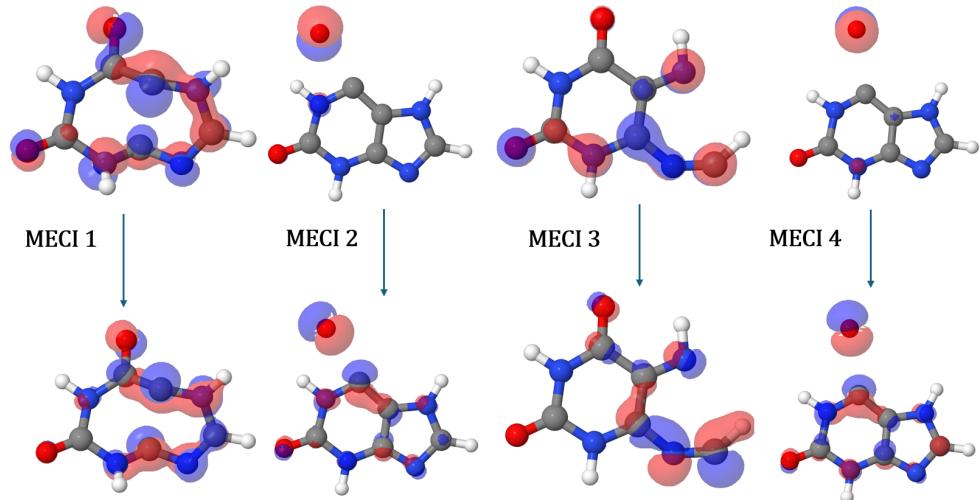
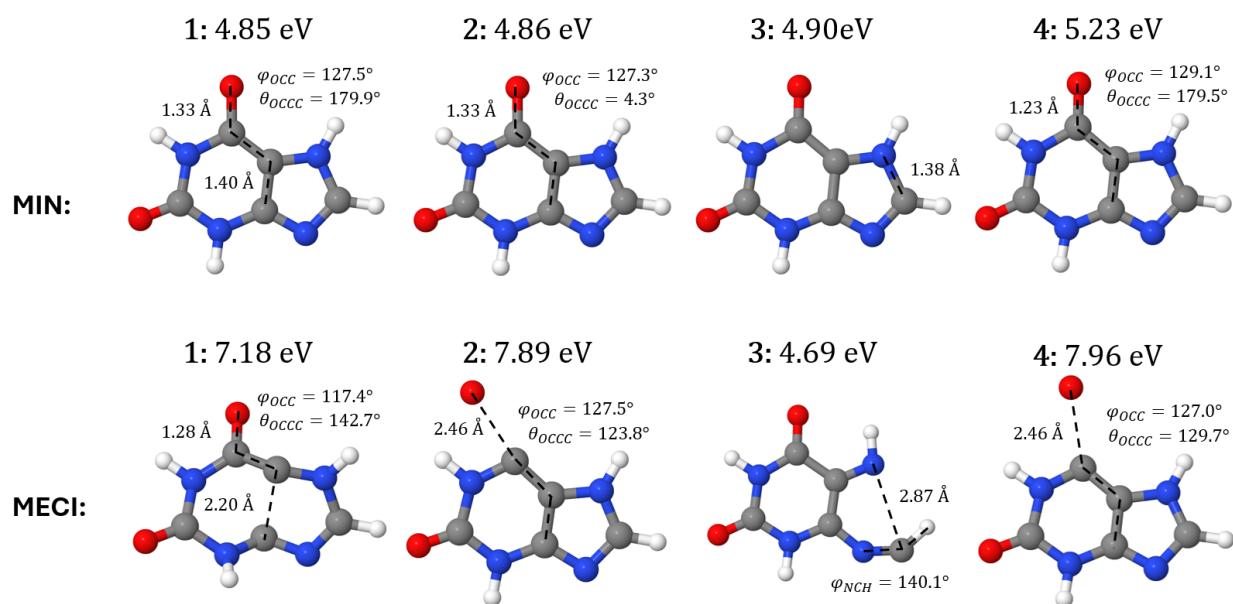


Figure 35: Transition orbitals of Xanthine at the optimised MECI.

Figure 36: Optimised  $S_1$  excited state geometries and  $S_1/S_0$  conical intersection pairs.

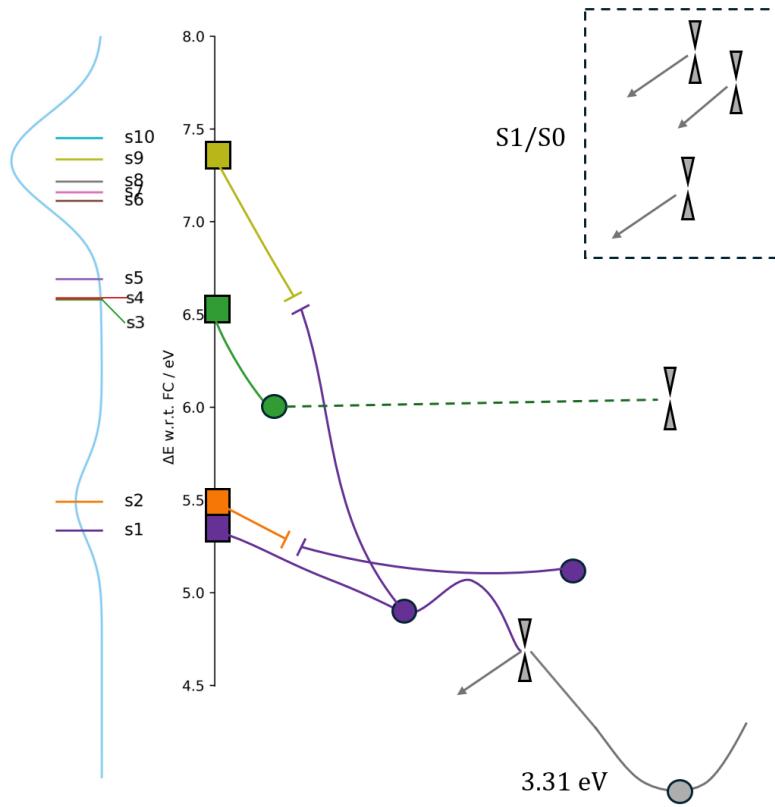


Figure 37: Extended schematic representation of the excited-state Potential Energy Surface (PES) of xanthine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Gaps indicate change in energy surface maintaining excited state character. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows indicate a return to Franck-Condon (FC) geometry. Energies relative to the FC energy.

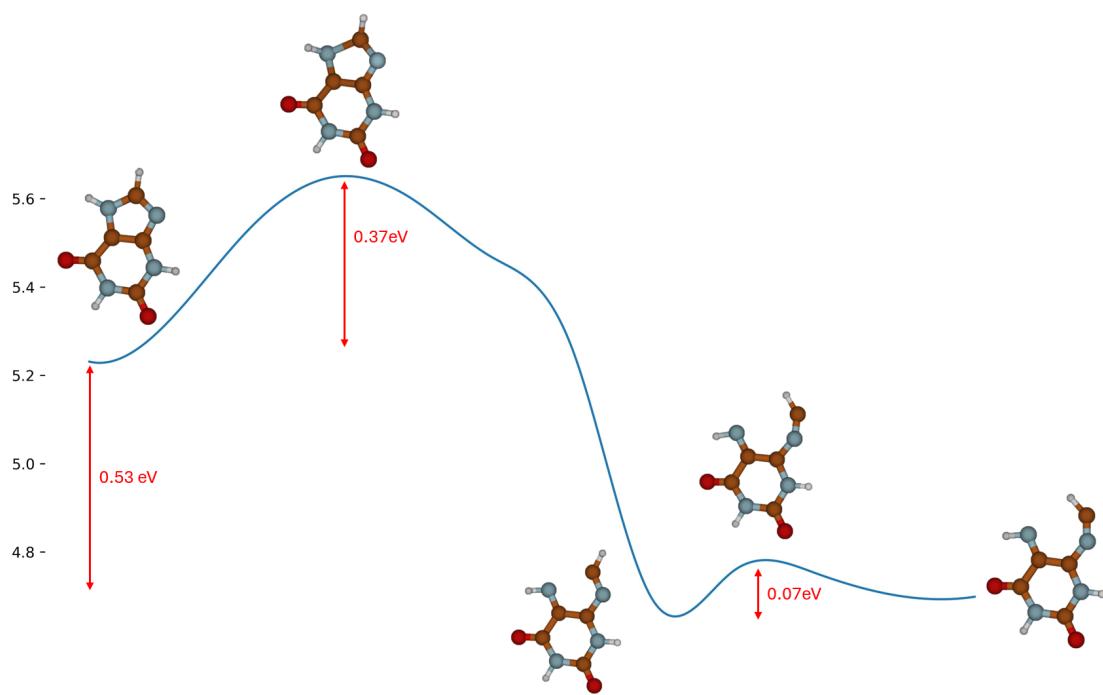


Figure 38: Xanthine minimum 1 MEP.

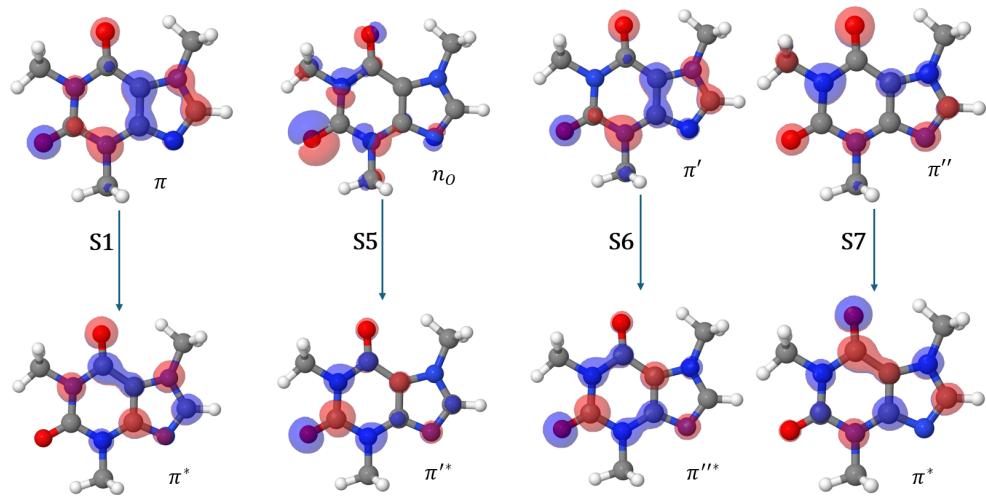


Figure 39: Natural Transition Orbitals (NTOs) of Caffeine at the FC geometry.

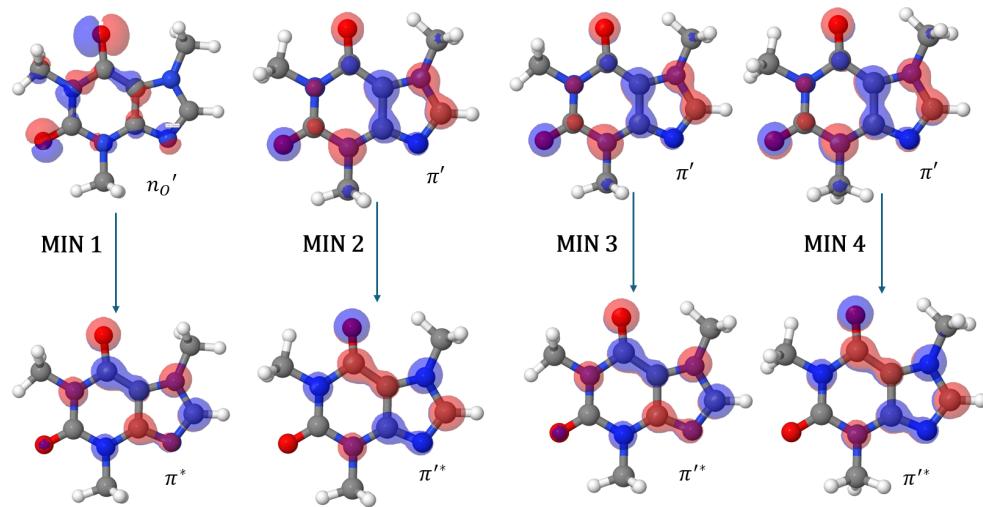
**Minima excited state character**

Figure 40: Transition orbitals of Caffeine at the optimised excited-state geometries.

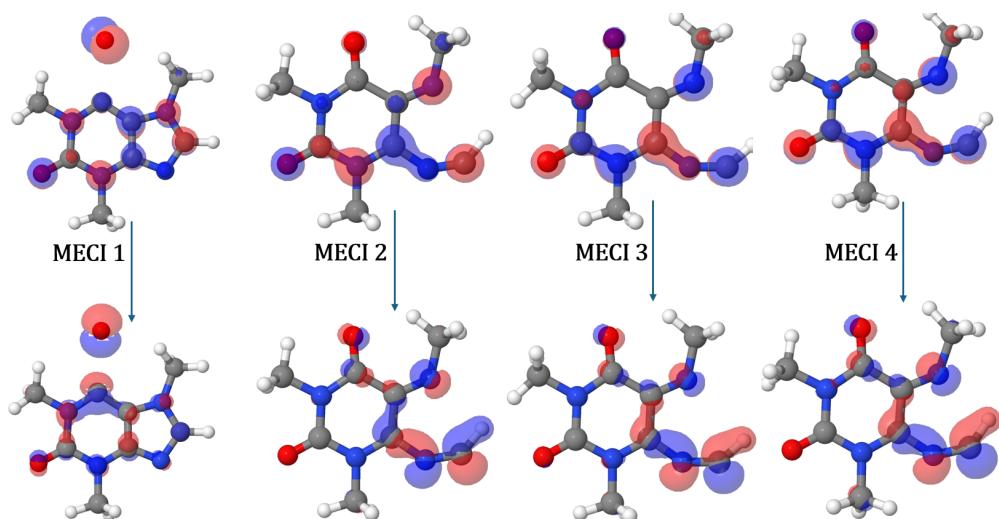
**MECI excited state character**

Figure 41: Transition orbitals of Caffeine at the optimised MECI.

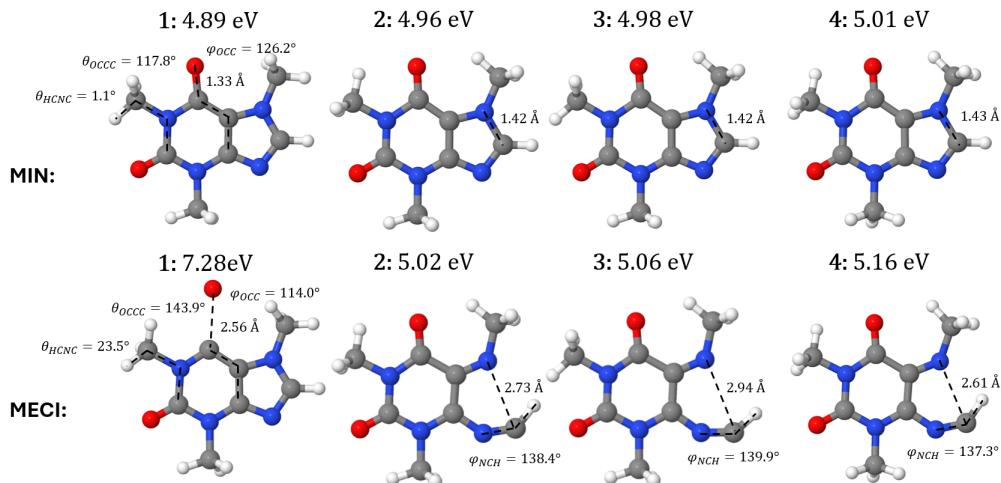


Figure 42: Optimised caffeine  $S_1$  excited state geometries and conical intersection pairs.

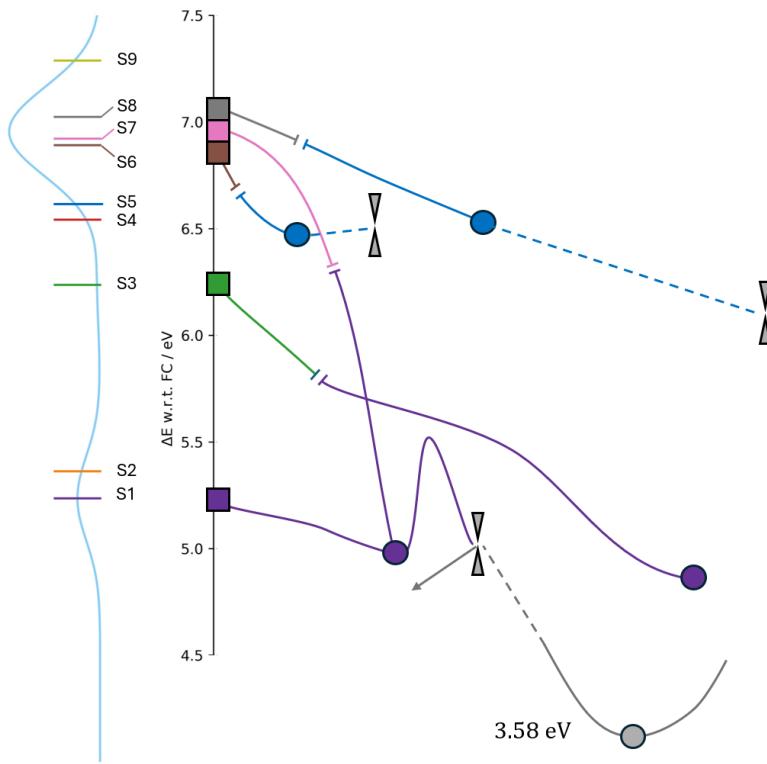


Figure 43: Extended schematic representation of the excited-state PES of caffeine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Gaps indicate change in energy surface maintaining excited state character. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows indicate a return to FC geometry. Energies relative to the FC energy.

Starting root	Minimum	$\Delta E$ min-FC / eV	$\Delta E$ MECI-FC / eV	$\Delta E$ min-MECI / eV
3	1	4.89	7.28	+2.39
7	2	4.96	5.02	+0.06
1	3	4.98	5.06	+0.08
1	4	5.01	5.16	+0.15

Table 10: Relative energies of additional caffeine  $S_1$  geometries and optimized MECIs with respect to the FC reference energy.

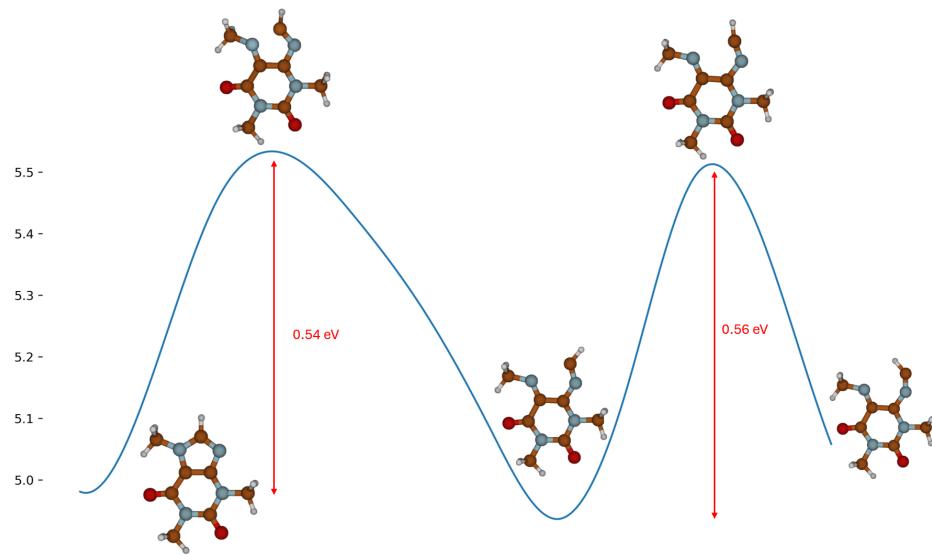


Figure 44: Caffeine minimum 2 MEP.

### B.3 Beta-Carboline

#### FC geometry NTOS

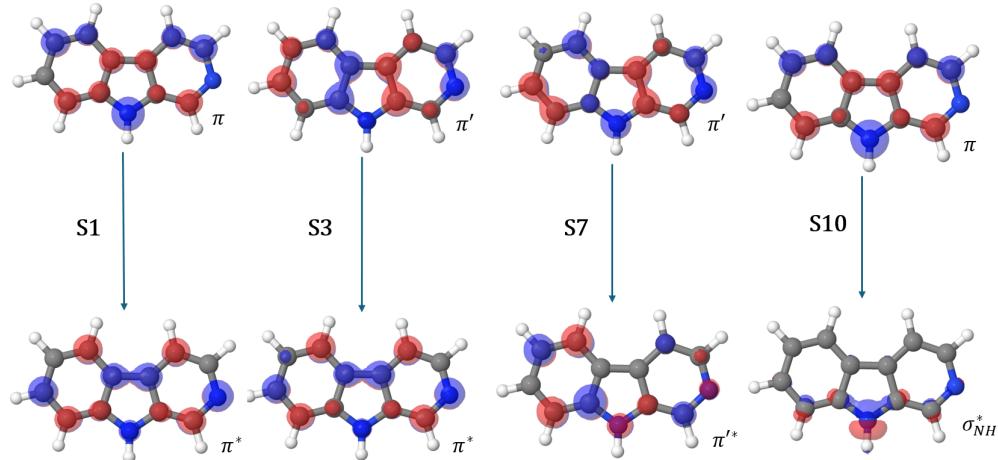


Figure 45: Natural Transition Orbitals (NTOs) of beta-carboline at the FC geometry.

#### Minima excited state character

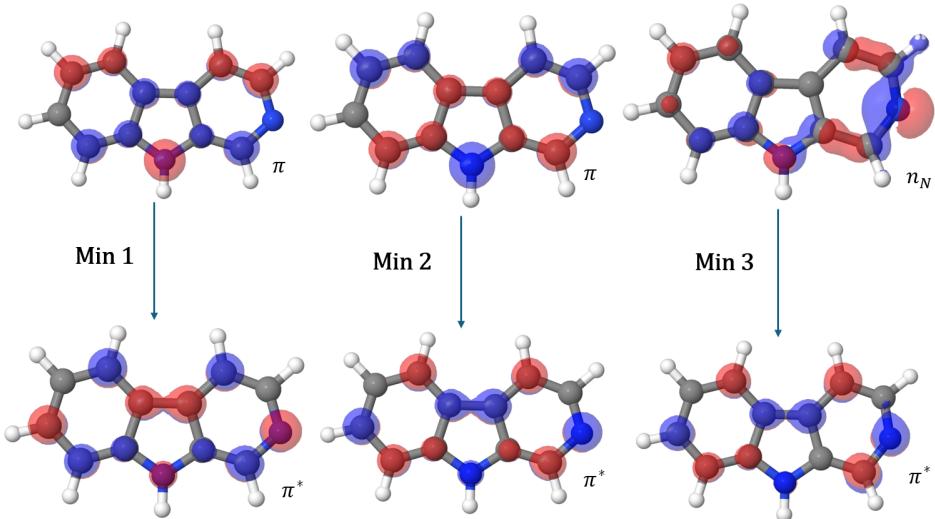


Figure 46: Transition orbitals of beta-carboline at the optimised excited-state geometries.

#### MECI excited state character

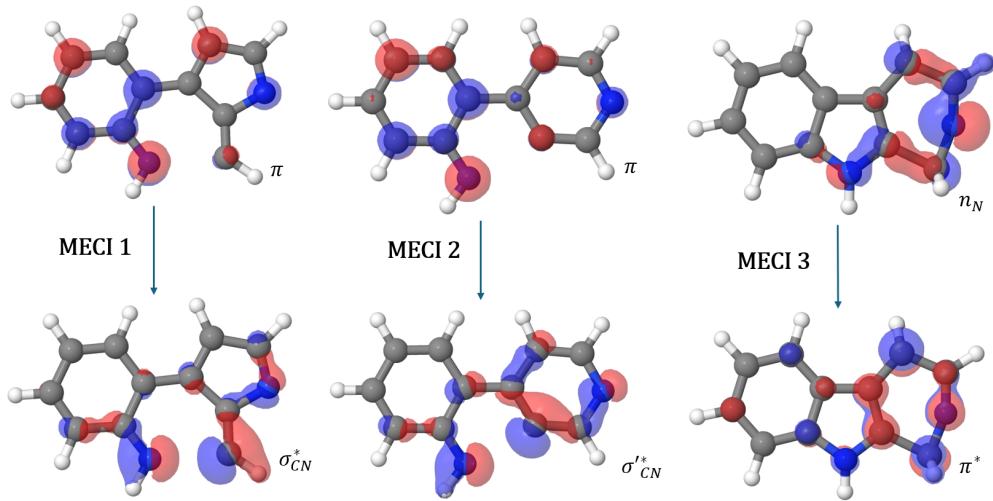
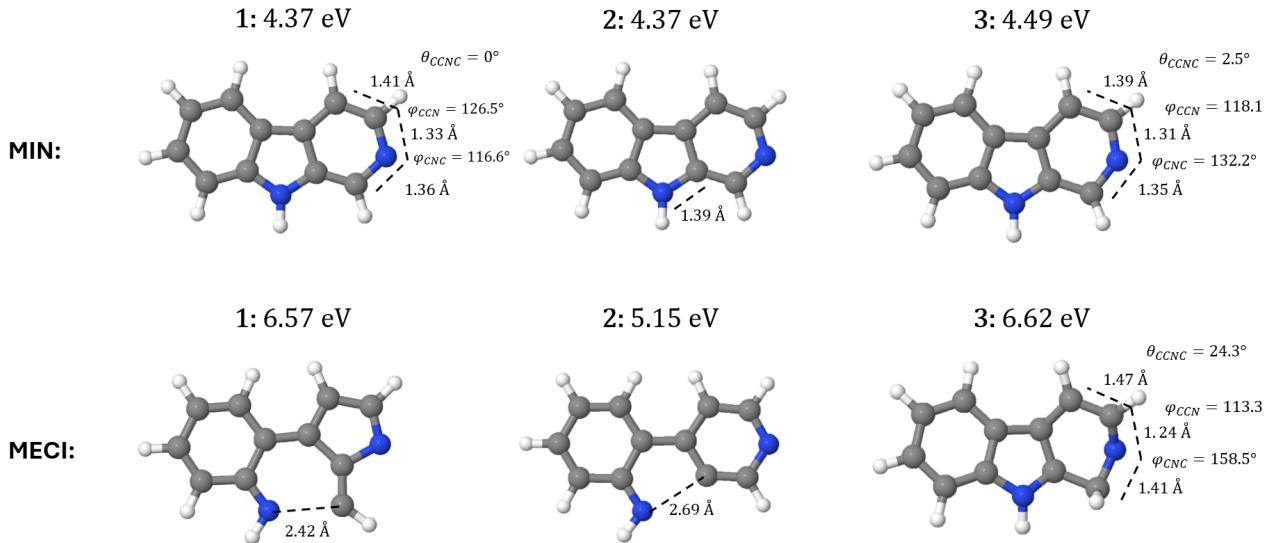


Figure 47: Transition orbitals of beta-carboline at the optimised MECI.

Figure 48: Optimized beta-carboline  $S_1$  excited state minima and conical intersection pairs.

Starting root	Minimum	$\Delta E$ min-FC / eV	$\Delta E$ MECI-FC / eV	$\Delta E$ min-MECI / eV
1	1	4.37	6.57	+2.20
2	2	4.37	5.15	+0.78
3	3	4.49	6.72	+1.67

Table 11: Relative energies of additional beta-carboline  $S_1$  geometries and optimized MECIs with respect to the FC reference energy.

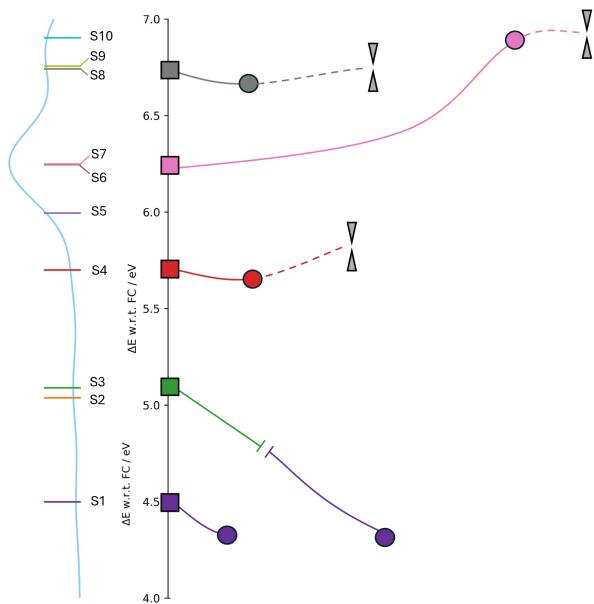


Figure 49: Extended schematic representation of the excited-state PES of beta-carboline. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows at  $S_0/S_1$  MECIs indicate a return to FC geometry.

## B.4 Tryptamine

### FC geometry NTOS

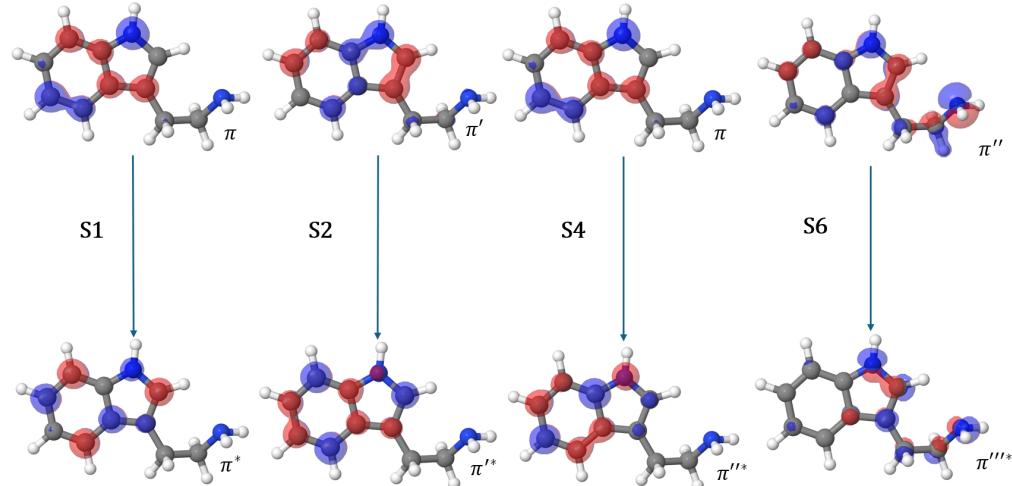


Figure 50: Natural Transition Orbitals (NTOs) of Tryptamine at the FC geometry.

### Minima excited state character

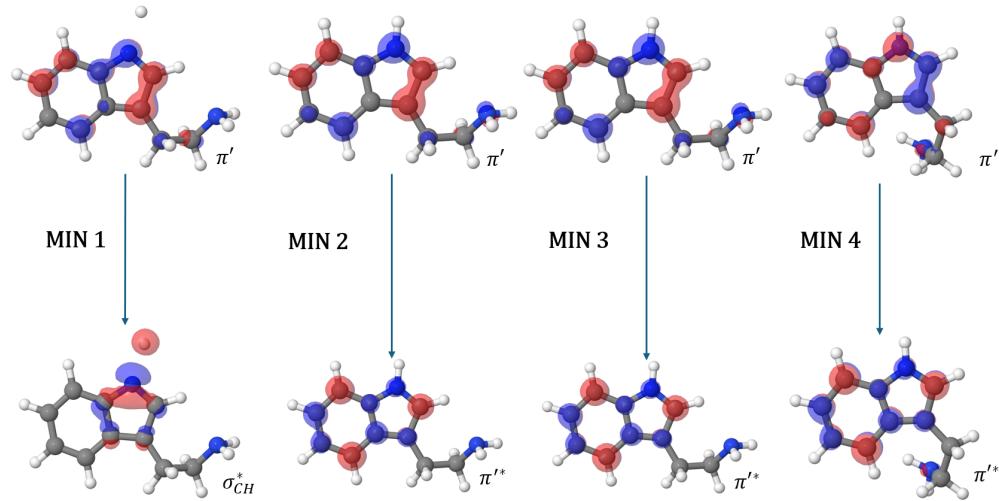


Figure 51: Transition orbitals of Tryptamine at the optimised excited-state geometries.

### MECI excited state character

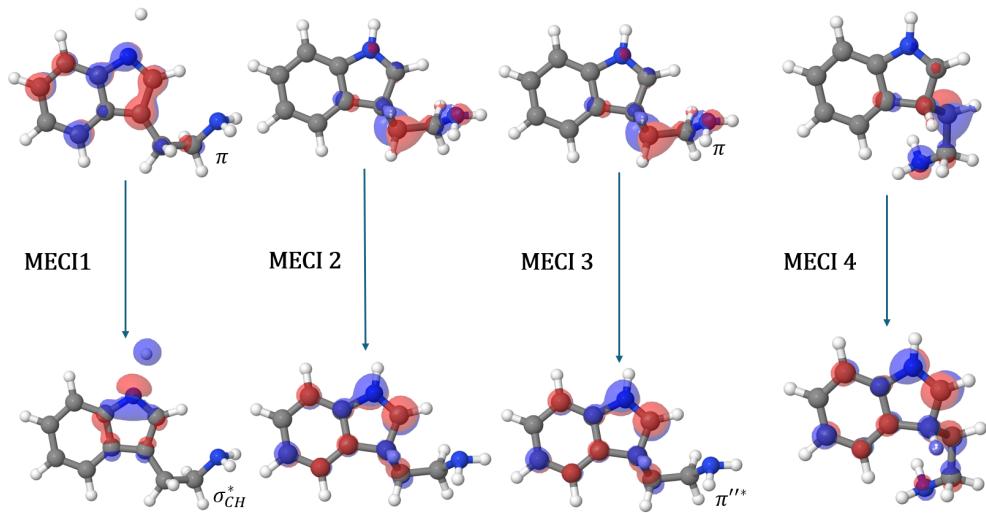
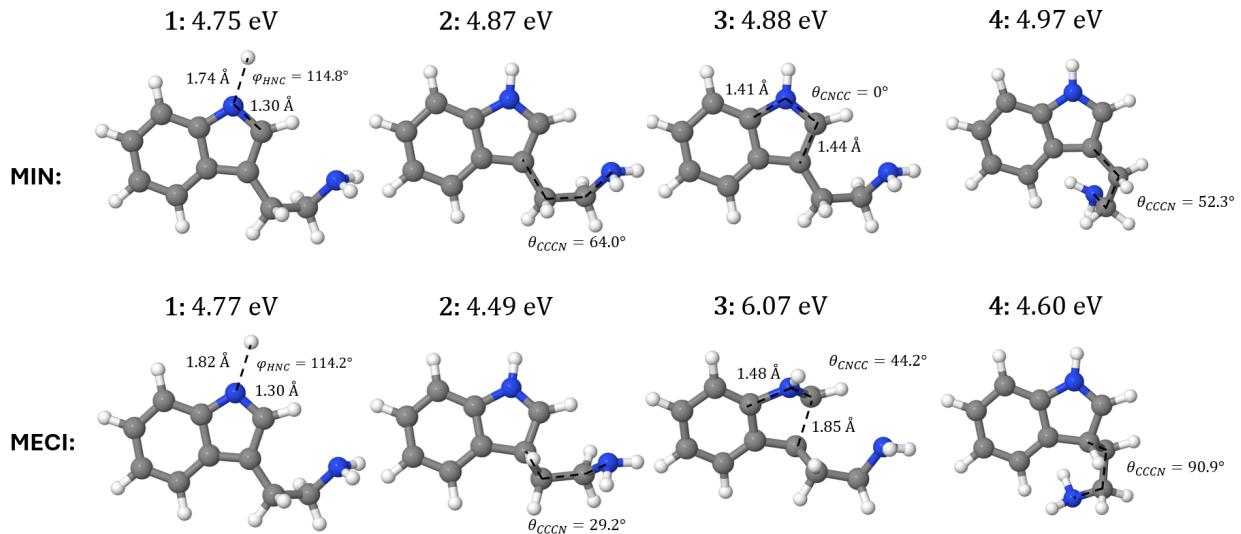


Figure 52: Transition orbitals of Tryptamine at the optimised MECI.

Figure 53: Optimized tryptamine  $S_1$  excited state minima and conical intersection pairs.

Starting root	Minimum	$\Delta E$ min-FC / eV	$\Delta E$ MECI-FC / eV	$\Delta E$ min-MECI / eV
1	1	4.87	4.49	-0.38
1	2	4.88	6.07	+1.20
10	3	4.97	4.60	-0.37

Table 12: Relative energies of additional tryptamine  $S_1$  geometries and optimized MECIs with respect to the FC reference energy.

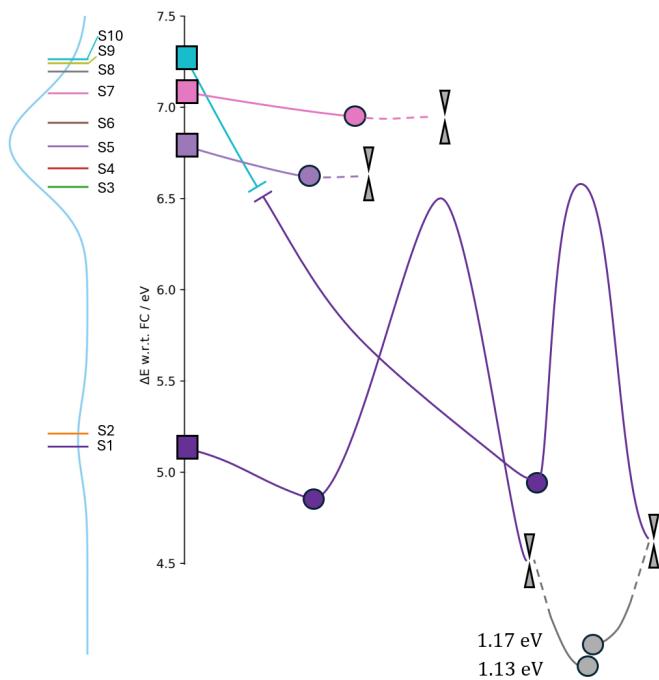


Figure 54: Extended schematic representation of the excited-state PES of tryptamine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surface. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows at  $S_0/S_1$  MECIs indicate a return to FC geometry.

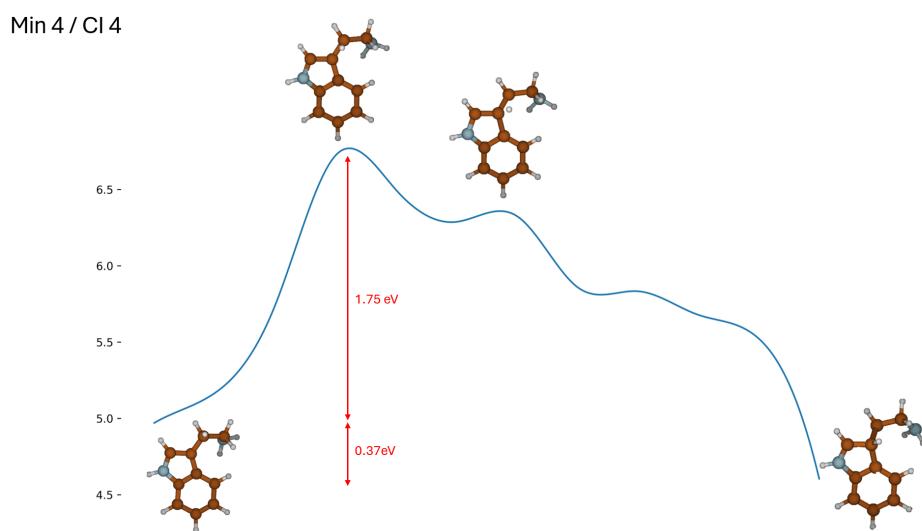


Figure 55: Minimum Energy Path (MEP) pathway of minimum / MECI 2.

## B.5 Nicotine

### FC geometry NTOS

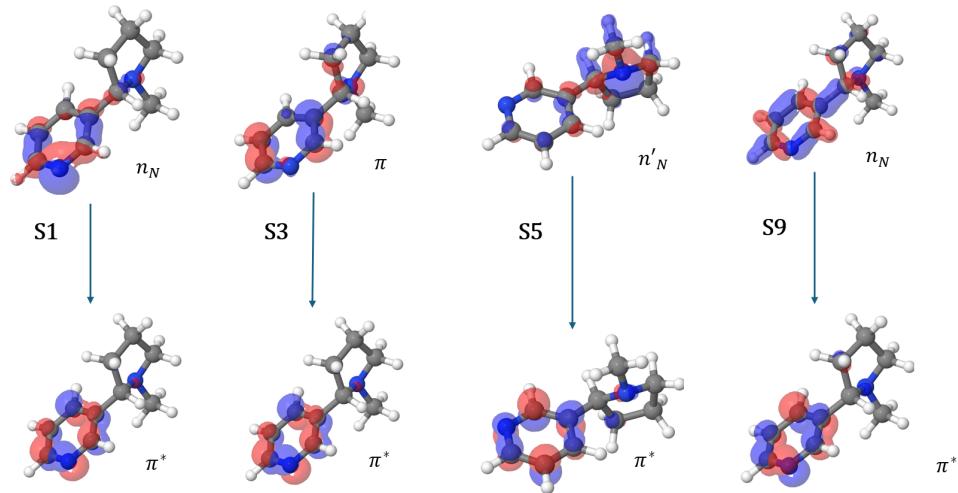


Figure 56: Natural Transition Orbitals (NTOs) of nicotine at the FC geometry.

### Minima excited state character

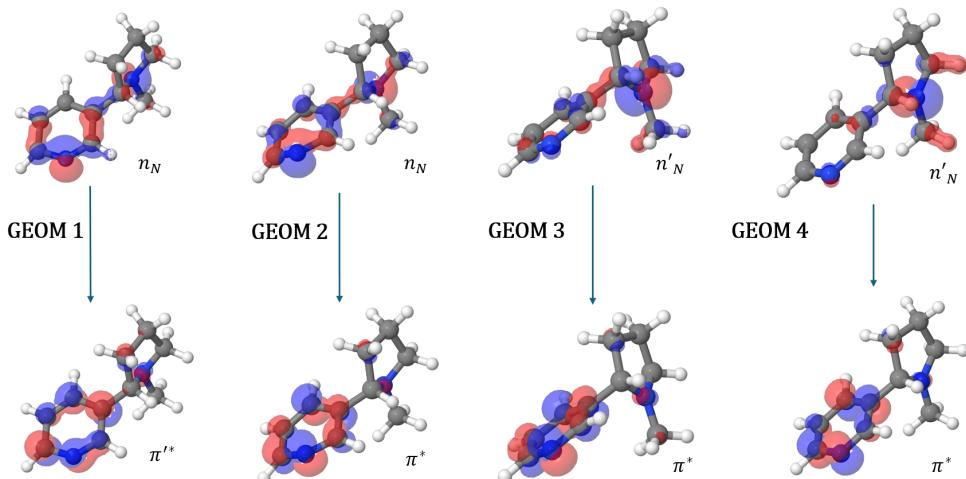


Figure 57: Transition orbitals of nicotine at the optimised excited-state geometries.

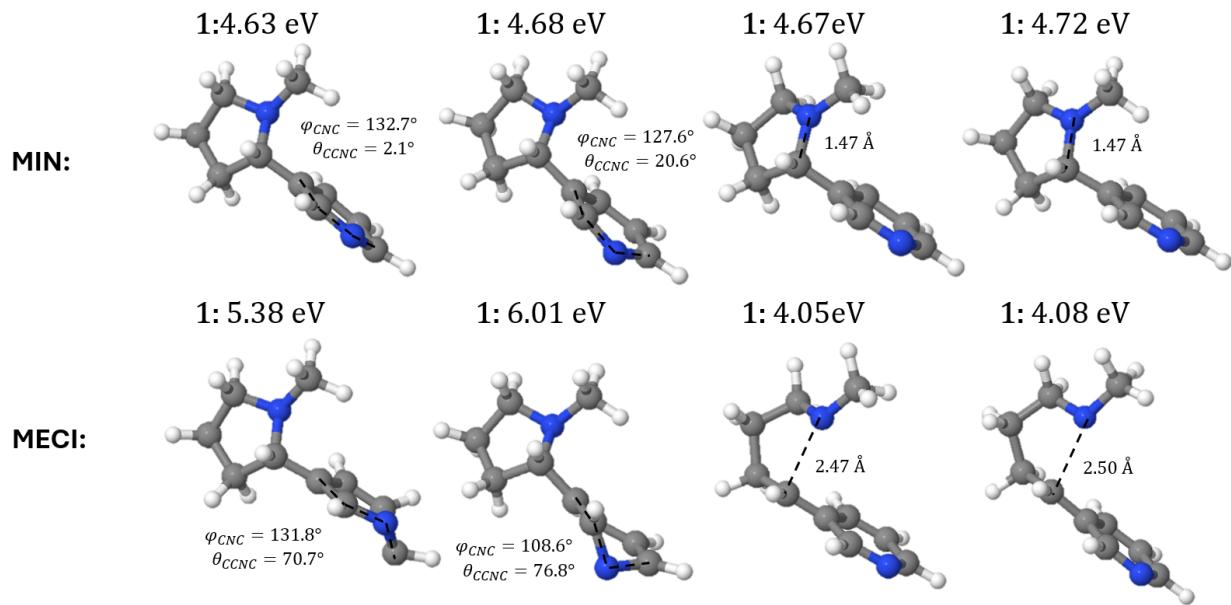


Figure 58: Optimised  $S_1$  excited state geometries and conical intersection pairs.

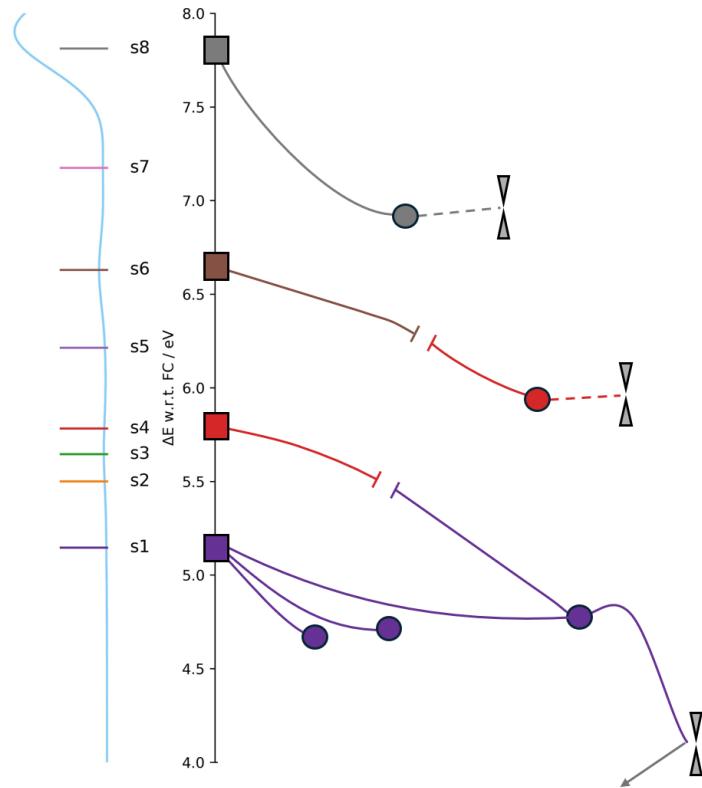


Figure 59: Schematic representation of the excited-state PES of nicotine. The vertical absorption spectrum is shown on the left. Colours indicate excited-state surfaces. Minima are marked as points, and conical intersections between surfaces and the next lower state are depicted as grey double cones. Left grey arrows at  $S_0/S_1$  MECIs indicate a return to FC geometry.

Starting root	Minimum	$\Delta E$ min-FC / eV	$\Delta E$ MECI-FC / eV	$\Delta E$ min-MECI / eV
1	1	4.68	5.38	+0.70
1	2	4.63	6.01	+1.38
1	3	4.69	4.05	-0.64
4	4	4.72	4.08	-0.64

Table 13: Relative energies of the four nicotine  $S_1$  geometries and their associated MECI with respect to the FC reference energy.

## B.6 Cytisine

### FC geometry NTOS

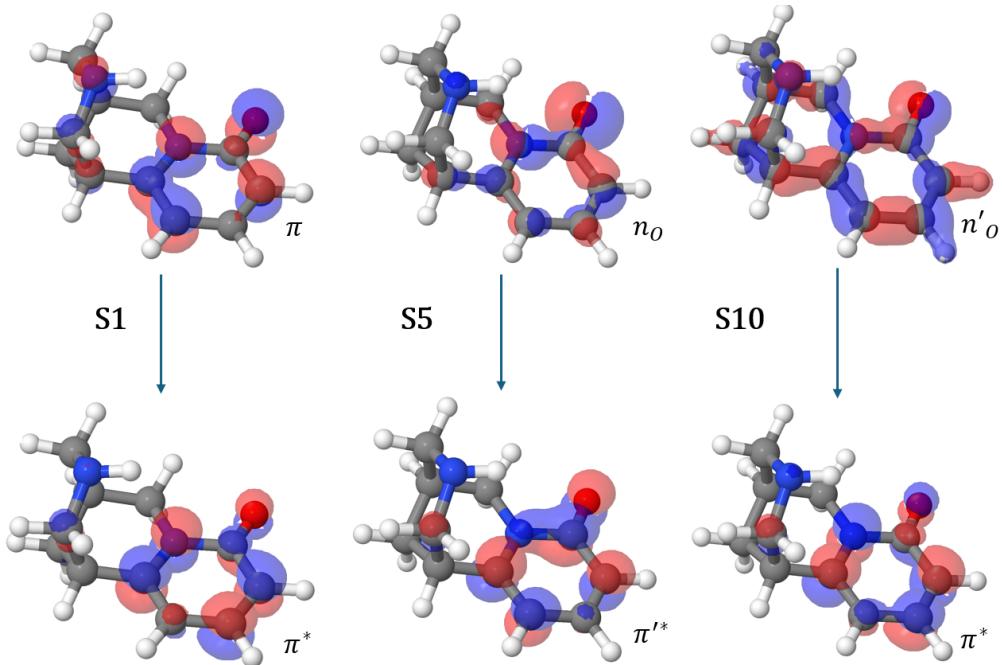


Figure 60: Natural Transition Orbitals (NTOs) of cytisine at the FC geometry.

Starting root	Minimum	$\Delta E$ min-FC / eV	$\Delta E$ MECI-FC / eV	$\Delta E$ min-MECI / eV
1	1	4.37	8.74	+4.37
1	2	4.37	8.94	+4.56

Table 14: Relative energies of the two cytisine  $S_1$  minima and their associated MECI with respect to the FC reference energy.

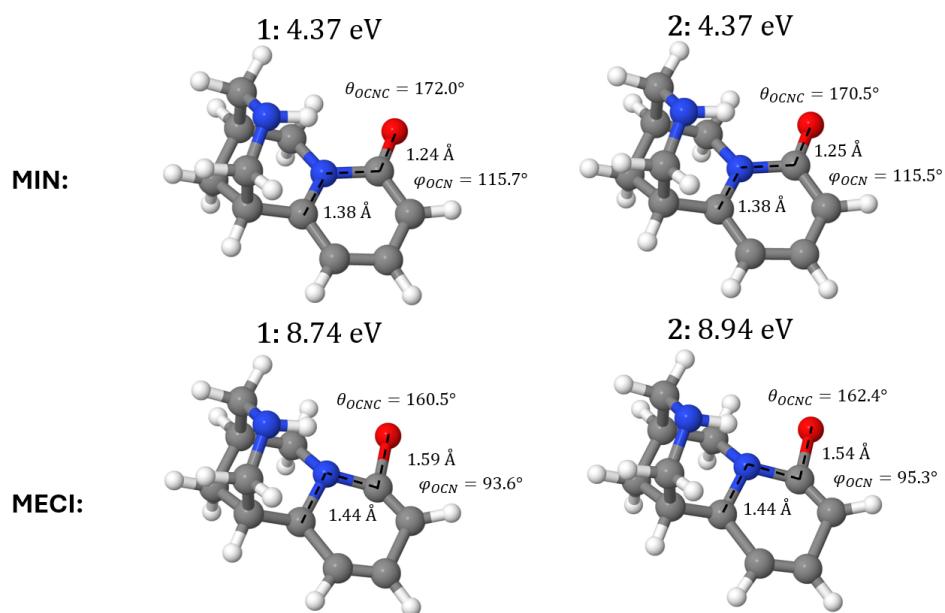


Figure 61: Cytisine optimized  $S_1$  excited state minimum and electronic state character.