Low-Lying $\pi\pi^*$ Excited States in Five-Membered Ring Heterocycles: A Continuing Challenge

Chaiyaporn Lakmuang,[†,‡] Thierry Tran,[¶] and Antonio Prlj[*,†]

†*Division of Physical Chemistry, Ruđer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia*

‡*Department of Chemistry, Faculty of Science, University of Zagreb, Horvatovac 102a, 10000 Zagreb, Croatia*

¶*Faculty of Chemistry, Wrocław University of Science and Technology, PL-50370, Wrocław, Poland*

E-mail: antonio.prlj@irb.hr

Abstract

The lowest electronically excited $\pi\pi^*$ states of conjugated hydrocarbons, as well as related heterocyclic compounds, pose a significant challenge for quantum chemical calculations. The difficulty arises from the need to achieve a balanced description of excited states with markedly different electronic characters, such as bright *vs.* dark, ionic *vs.* covalent, and singly *vs.* doubly excited states. Here, we focus on the low-lying $\pi\pi^*$ states of five-member-ring heterocycles, such as furan, pyrrole and thiophene. We analyze the origins of the diverse excited-state characters from both molecular orbital and valence bond theory perspectives, employing pseudosymmetry arguments to bridge the two viewpoints and drawing comparisons with the $\pi\pi^*$ excited states of the *cis*-butadiene and benzene model systems. We further discuss the critical role of electronic correlation in obtaining a balanced description of excited states and explain

why it is difficult, or sometimes impossible, to achieve high accuracy with commonly used electronic-structure approximations. As an illustration, we focus on results from standard time-dependent density functional theory approximations, as well as a series of algebraic diagrammatic construction methods with systematically improvable correlation levels through perturbative truncation schemes. Finally, we address the origins and implications of valence–Rydberg mixing, which arguably remains an open question in excited-state quantum chemistry.

1 Introduction

Five-membered ring heterocycles (Figure 1), such as furan, pyrrole and thiophene, are essential building blocks of biomolecules and systems relevant to optoelectronic applications.[1,2] It is not surprising that these systems, and in particular their electronically excited states, have attracted tremendous research interest over the years. Along with many experimental studies (see Palmer *et al.*[3–5] and the references therein), numerous computational studies have focused on excited states of furan,[3,6–18] pyrrole[4,6,11,13,14,16,17,19–23] and thiophene,[5,16,24–33] employing methods such as time-dependent density-functional theory (TDDFT),[11,12,16,21,29] multi-reference configuration interaction (MRCI),[3–5,21] complete active space second-order perturbation theory (CASPT2),[6,8,19,20,24,25] coupled cluster (CC),[8–10,13,27] algebraic diagrammatic construction (ADC)[15,18,23,27,29] *etc.* Pyrrole and furan are part of the Thiel benchmark set,[34] which is commonly used for benchmarking (new) excited-state methods. Due to their nontrivial photophysics and photochemistry, featuring the interplay of many electronic states and competing relaxation pathways, heterocycles appear as ideal model systems for testing new theoretical methods for nonadiabatic molecular dynamics. Consequently, extensive research has been done on the nonadiabatic dynamics of furan,[35–40] thiophene,[37,41–48] and (most frequently) pyrrole,[20,37,49–63] mostly dating from the 2000s onward. Thiophene's selenium analogue, selenophene, has been used as a model system for intersystem crossing induced by heavy-atom effects.[45] Less common heterocycles from the 15th and 16th columns

of the periodic table, such as phosphole, arsole, and tellurophene, are predominantly discussed in the experimental literature,[64–67] while theoretical analyses have mainly focused on their ground-state aromaticity.[68–70] Interestingly, weakly aromatic phosphole and arsole are non-planar (Figure 1), as the energy stabilization from aromaticity is insufficient to overcome distortion through pyramidalization.[70]

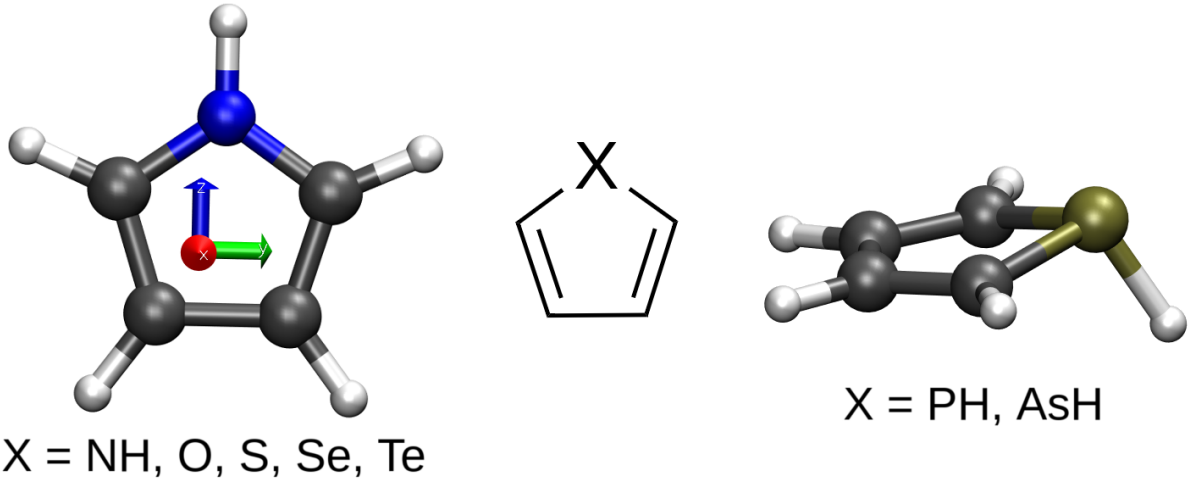


Figure 1: Five-membered ring heterocycles. Following recommendation by Mulliken,[71] planar C_{2v} molecules are placed in yz plane, where z direction coincides with the C_2 rotation axis, and x axis is perpendicular to the molecular plane. With this orientation, the lowest $\pi\pi^*$ excited states have A_1 or B_2 symmetry. For non-planar phosphole and arsole $\pi\pi^*$ states have A' and A'' irreducible representations (irreps) within C_s point group symmetry. Insets: pyrrole (left) and phosphole (right).

The reason behind the large (and still growing) number of theoretical studies on the excited states of five-membered ring heterocycles is their challenging electronic structure. Excited-state dynamics strongly depends on achieving a proper balance between $\pi\pi^*$, $n\pi^*$, $\pi\sigma^*$ or Rydberg excitations, which is challenging even when high-level wavefunction-based methods are employed. As a consequence, it is not uncommon for the same nonadiabatic dynamics method to yield very different results depending on the electronic structure method employed.

The theoretical description of $\pi\pi^*$ excited states *per se* may be perceived as less challenging, as only π-excitations are involved, but this is somewhat misleading. Earlier anal-

yses of thiophene and its fused derivatives (thienoacenes),[29] have shown that two low-lying $\pi\pi^*$ singlet states were highly unbalanced when TDDFT was employed, whereas correlated wavefunction-based methods (ADC, CC, CASPT2) provided a picture more consistent with experiment. In the case of thiophene, standard TDDFT approximations predict an incorrect ordering of the B_2 (dominated by the $HOMO \rightarrow LUMO$ excitation) and A_1 (dominated by the $HOMO-1 \rightarrow LUMO$ excitation) $\pi\pi^*$ states, akin to configuration interaction singles (CIS). The inclusion of double excitations, even at the perturbative level, largely resolves this problem, even though both states remain dominated by single excitations. This issue was compared to the unbalanced description of the L_a and L_b states in linear acenes (and other polycyclic aromatic hydrocarbons) at the TDDFT level, a well-known problem that has been thoroughly investigated[72–81] since its first report in 2003.[82] The short-axis polarized L_a state of acenes ($HOMO \rightarrow LUMO$) is significantly underestimated by local and semilocal functionals, whereas range-separated hybrid functionals tend to provide a good description. The long-axis polarized L_b state consists of equal (or roughly equal) contributions from the $HOMO-1 \rightarrow LUMO$ and $HOMO \rightarrow LUMO+1$ configurations, but it is also sensitive to the inclusion of differential correlation effects (*i.e.*, double excitation corrections).[75,83] Unlike L_a, L_b excitation energies are well predicted by global hybrid functionals such as B3LYP, while range-separated hybrids tend to overestimate L_b. The different behavior of the L_a and L_b states was connected to their ionic and covalent valence bond (VB) characters, respectively.[82] The issue of the unbalanced description of L_a and L_b would likely not have such an impact if L_a-like and L_b-like states did not play a significant role in a large number of fused-ring hydrocarbons and their heteroaromatic analogues. L_a and L_b type states dominate the photochemistry and photophysics of naphthols,[84] indoles,[85,86] azulene,[87] aminoacids (adenine, guanine),[88] some nucleobases (tryptophan),[89] porphyrins,[90] and many acene-like systems.[83] Since in heteroaromatic molecules L_a and L_b states lose some of their main features (*e.g.*, short- and long-axis polarization), it was recently proposed to rename them as L_w (weakly correlated) and L_s (strongly correlated), respectively.[91] Overall,

L_a and L_b type states certainly play an important role for a large scope of molecules. When it comes to excited states of five-membered ring heterocycles, is it worth asking whether L_a/L_b nomenclature should also be used unequivocally?

Perhaps the oldest and most well-known puzzle concerning the elusive positions of the $\pi\pi^*$ states is that of the $1B_u$ and $2A_g$ states in *trans*-polyenes. The $1B_u$ state of polyenes is dominated by the single $HOMO \rightarrow LUMO$ excitation, whereas the $2A_g$ state is typically referred to as a "doubly excited" state, despite being a mixture of single and double excitations.[92] Just like the L_a and L_b states of acenes, the $1B_u$ is an ionic state in VB theory, while the $2A_g$ state is covalent.[93,94] Historically, the relevant question of the correct energy ordering of $1B_u$ and $2A_g$ was both an experimental and a theoretical issue - early developments from both perspectives were summarized in a review by Hudson, Kohler, and Schulten.[95] Excited states of polyenes have remained a topic of interest and debate up to the present.[96–111] Polyenes are used as fundamental model systems for testing new excited-state methods,[34] while the excited states of the smallest representatives, such as butadiene, are now being calculated within chemical accuracy using methods such as full configuration interaction - quantum Monte Carlo (FCI-QMC),[112] CC up to connected quadruples (CCSDTQ)[113] and semistochastic heat-bath configuration interaction (SHCI).[114] Some controversy arose regarding the TDDFT performance for polyenes - it was reported that standard TDDFT (using the Tamm–Dancoff approximation, TDA) combined with GGA functionals provided reasonable excitation energies for the $2A_g$ states.[115] This was especially surprising, considering that standard TDDFT within the adiabatic approximation is a single-excitation method that cannot describe doubly excited states.[116] The "good" performance of adiabatic TDDFT was later rationalized as a result of a cancellation of errors,[117] while efforts to develop TDDFT beyond the adiabatic approximation and explicitly include doubly excited configurations (*e.g.* dressed TDDFT approach[118,119]) have yielded promising results for compounds such as butadiene and hexatriene. Finally, to connect with five-membered ring heterocycles, we note that heterocycles have a "hidden" *cis*-butadiene structure in their ring. Low-lying $\pi\pi^*$

states of heterocycles have A_1 and B_2 symmetry labels, the same as in *cis*-butadiene, which are equivalents of A_g and B_u in *trans*-butadiene, respectively. It is reasonable to pose the question: how closely do the $\pi\pi^*$ excited states of five-membered ring heterocycles resemble those of polyenes?

In the present contribution we take a deeper look into the excited $\pi\pi^*$ states of five-membered ring heterocycles and their connection with excited states of polyenes and acenes, focusing on three key compounds: furan, pyrrole and thiophene. To better understand the trends among the congener compounds, some comparisons with selenophene, tellurophene, phosphole, and arsole are also made (see Supporting Information, SI). For all systems, the emphasis is placed on the two lowest $\pi\pi^*$ states, which remain challenging to describe in a balanced manner. We look into different perspectives to analyze the distinct characters of excited states, using arguments from molecular orbital (MO) and VB theories, as well as symmetry and pseudosymmetry considerations. We try to rationalize the following questions: how double excitations affect the low-lying states, why some states appear optically bright or dark, when and how Rydberg states start to mix with low-lying $\pi\pi^*$ states, and how much triplet states differ from singlets? Finally, we scrutinize the performance of selected excited-state methods. Among the wide range of wavefunction-based methods, we inspect the series of ADC methods[120] with different orders in perturbation theory, namely ADC(1), an equivalent of CIS, the popular ADC(2), and the higher-level ADC(3) method, which can describe excited states with partial double-excitation character, although not necessarily very accurately.[121] The comparison of ADC(1), ADC(2), and ADC(3) illustrates the importance of increasingly accounting for correlation effects in $\pi\pi^*$ states. Coupled-cluster singles, doubles, and triples (CC3),[122] which has emerged as the state-of-the-art method for excited states of medium-sized molecules,[123] was employed as a reference. We also analyze the performance of adiabatic TDDFT using a series of standard underlying density functional approximations to gain a better understanding of intrinsic limitations and the improvements necessary within the TDDFT framework.

2 Theoretical analysis

2.1 Molecular orbital theory

The character of the lowest $\pi\pi^*$ singlet excited states in polyenes, acenes, and heteroaromatic rings can be rationalized by a simple model of four electrons in four orbitals. In their seminal study on the electronic structure of simple polyenes, Schulten and Karplus[93] employed MO theory to analyze the character of singly excited B_u state and the "doubly excited" A_g state in *trans*-polyenes. It was sufficient to consider four π orbitals, namely HOMO-1, HOMO, LUMO, and LUMO+1. The lowest-energy singly excited configuration can be formed by a $HOMO \rightarrow LUMO$ orbital excitation (hereafter denoted as $|H \rightarrow L\rangle$), which belongs to the B_u irrep and gives rise to the B_u excited state. Likewise, the lowest A_g excited state is composed of singly excited configurations belonging to the A_g irrep, namely $|H-1 \rightarrow L\rangle$ and $|H \rightarrow L+1\rangle$. In the idealized model these two configurations are energetically degenerate and interact as a plus and minus linear combination. The configuration interaction approach gives rise to two excited states, A_g^+ and A_g^-. This mixing of configurations is shown schematically in the upper part of Figure 2. Karplus and Schulten observed that the A_g^+ and A_g^- states have similar excitation energies at the semiempirical CIS level, but a dramatic change occurs as soon as doubly excited configurations are taken into account. In particular, only the A_g^- state interacts with double excitations, which substantially lowers its excitation energy. Note that configuration interaction is a generic term for the linear combination of multiple electronic configurations, but the implications of the term "interaction" should not be overlooked. The double excitation character in A_g^- primarily corresponds to the $|H, H \rightarrow L, L\rangle$ configuration, which belongs to the totally symmetric A_g irrep. As a consequence of the interaction between different configurations, the multiconfigurational A_g^- state drops into the same energy range as B_u, or even below. Since the low-lying A_g^- and B_u excited states possess very different characters, it is unsurprising that the exact order of these two states was a matter of debate for decades. Many electronic structure

methods were unable to properly describe excited states with multiconfigurational character (including double excitation), while the optically dark A_g^-, sometimes called a "phantom" state, could only be measured by two-photon spectroscopy. Needless to say, absorption from the totally symmetric ground state (A_g) to both A_g^- and A_g^+ is one-photon forbidden, *i.e.*, *gerade-gerade* transitions are forbidden by Laporte symmetry rules. However, this applies only to centrosymmetric *trans*-polyenes.

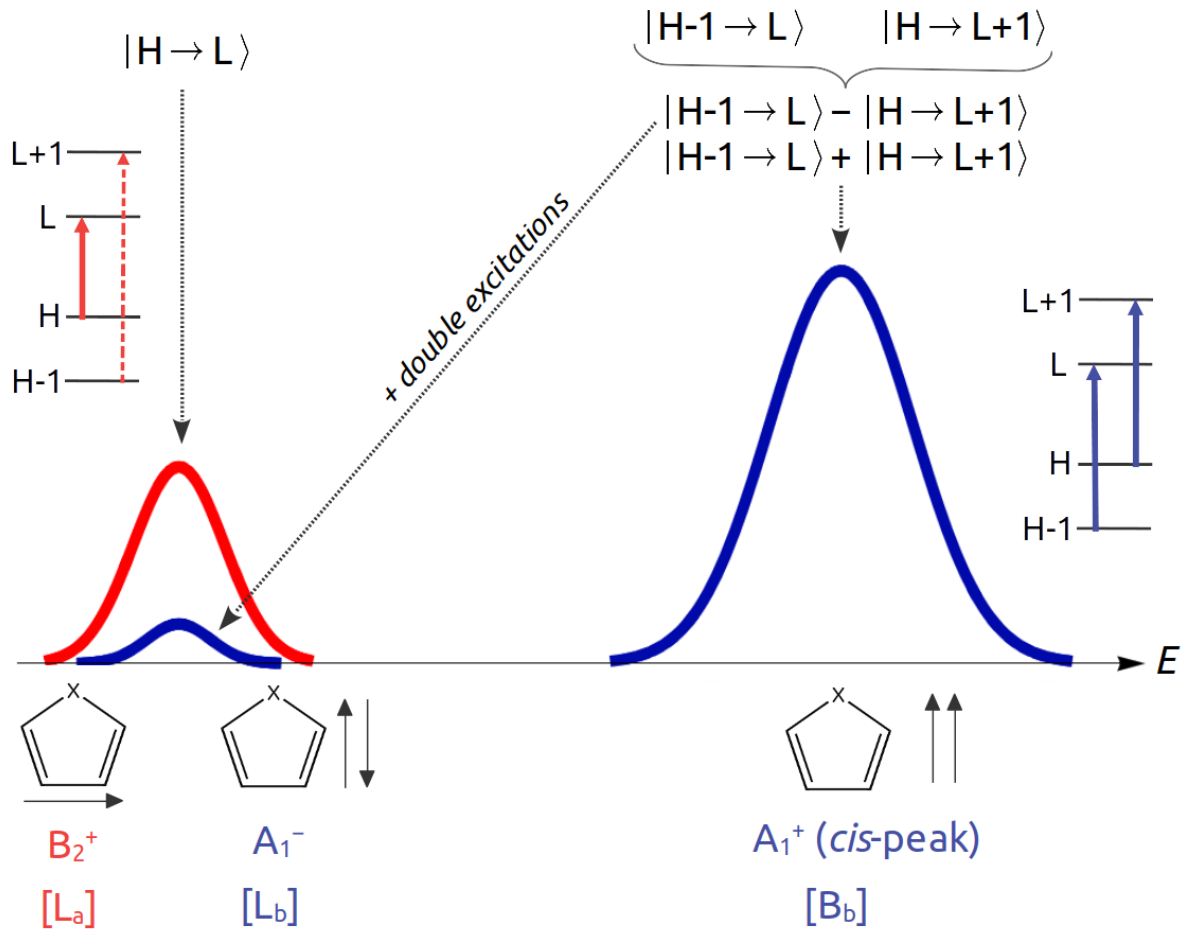


Figure 2: MO configurations forming $\pi\pi^*$ singlet excited states. The scheme of the absorption maxima approximately describes the spectrum of *cis*-polyenes and five-member-ring heterocycles (states labeled as A_1/B_2), as well as some acenes (Platt's nomenclature[124] is in square brackets). Small black arrows next to the heterocycle structures represent transition dipole moment as vectors which rationalize the overall absorption peak intensities.

In *cis*-polyenes, inversion symmetry is broken, yielding the dark A_1^- and bright A_1^+ excited states. The latter, also known as *cis*-peak,[93,125] owes its large absorption intensity to

the constructive addition of transition dipole moments corresponding to two singly excited configurations, as opposed to the destructive interference in dark A_1^-. The addition and cancellation of transition dipole moments can be deduced from a simple mathematical model for configuration interaction of $X \pm Y$ type.[126] Note that pure double excitation does not contribute to one-photon absorption intensity. On the other hand, the bright B_u state of *trans*-polyene converts to bright B_2 state in *cis*-polyene. The schematic pattern of the photoabsorption spectrum for *cis*-butadiene (and related compounds) is depicted in Figure 2. Importantly, *cis*-butadiene serves as a model for interpreting the spectra of five-membered ring heterocycles, which exhibit two lowest $\pi\pi^*$ states of B_2 and A_1 symmetry, whereas their higher lying A_1 excited state is typically the most intense state in the entire absorption spectrum below the ionization threshold. As a matter of fact, Figure 2 also schematizes the absorption spectrum of acenes and many acene-like organic molecules, yet the state symmetries are different. In acenes, the bright $HOMO \rightarrow LUMO$ excitation is known as L_a state according to Platt's nomenclature[124,127] (or the p-band in Clar's nomenclature[127]). The dark state, which is formed by destructive interference of the $|H-1 \rightarrow L\rangle$ and $|H \rightarrow L+1\rangle$ configurations, is known as L_b (α-band in Clar's nomenclature). This state also has a sizable contribution from double excitations. The exact order of the L_a and L_b states is system-dependent – in the experimental spectrum, the dark L_b is often hidden behind the bright L_a. The dark L_b (α-band) is paired with the higher-lying bright B_b (β-band), which originates from the constructive interference of $|H-1 \rightarrow L\rangle$ and $|H \rightarrow L+1\rangle$ configurations, and the latter state is not very sensitive to double excitations. In linear acenes (*i.e.*, naphthalene and larger), the bright L_a state has B_{2u} symmetry, while the dark L_b state (and its bright "antipode", B_b, at higher energies) corresponds to B_{3u} irrep. Note that the L_b state cannot contain contributions from the $|H, H \rightarrow L, L\rangle$ double excitation, which necessarily belongs to the totally symmetric irrep (A_g) – the L_b state can only mix with double excitations belonging to B_{3u} irrep.

Due to its higher symmetry, benzene is an outlier in the acenes family. In benzene, the L_a

state corresponds to the B_{1u}, and L_b to the B_{2u} irrep. Importantly, the L_a state of benzene is dark, just like L_b, which is an unusual property. This is a consequence of the orbital degeneracies in benzene, *i.e.*, the HOMO is degenerate with HOMO-1 and the LUMO is degenerate with LUMO+1. Unlike in other compounds where L_a is strongly dominated by the $|H \rightarrow L\rangle$ configuration, the L_a of benzene is an equal mixture of the $|H \rightarrow L\rangle$ and $|H-1 \rightarrow L+1\rangle$ configurations. These two configurations have antiparallel transition dipole moments, the addition of which leads to a net transition dipole moment close to zero and a vanishing oscillator strength. It is worth mentioning that in other compounds (including polyenes and heterocycles), the $|H \rightarrow L\rangle$ will also mix with $|H-1 \rightarrow L+1\rangle$, but since the latter configuration is energetically less favorable, its weight in the wavefunction of the low-lying $\pi\pi^*$ state will normally not exceed several percent. The $|H-1 \rightarrow L+1\rangle$ configuration will be dominant in a higher-lying state of the same symmetry.

2.2 Alternancy symmetry

Another useful perspective is Pariser's classification of the excited states of alternant π-conjugated hydrocarbons (polyenes, acenes) into *plus* and *minus* parities.[128] Pariser's alternancy symmetry arises in the Hückel[129] and Pariser-Parr-Pople[130–132] methods, which assume zero differential overlap, and it is exact within the framework of these methods. Yet, alternancy symmetry can still be used as a qualitative tool to interpret data from more complex excited-state theories and applies to conjugated systems that are not strictly hydrocarbons (*e.g.*, heterocycles).[17] Thus, in realistic applications, alternancy symmetry effectively reduces to a pseudosymmetry consideration. The application of alternancy symmetry in π-conjugated molecules was standardized and particularly advocated by Hirao and coworkers.[17,94,133,134] According to the alternancy symmetry rules, the ground state of alternant hydrocarbons is characterized as a *minus* state, while the excited state of $HOMO \rightarrow LUMO$ character transforms as a *plus* state. Two excited states of multiconfigurational nature corresponding to the plus and minus linear combinations of $|H-1 \rightarrow L\rangle$ and $|H \rightarrow L+1\rangle$

are classified as *plus* and *minus* states, respectively. The state labeled *minus* always has a lower energy than its corresponding *plus* counterpart. Note that the coincidence between the *plus-minus* parity (hereby marked in italic font) and the plus-minus sign in configuration interaction (roman font) is a matter of convention – flipping the orbital phase may eventually change the sign in configuration interaction, but not the alternancy symmetry label of the state. Finally, double excitations of the $A, A \rightarrow B, B$ type are classified as *minus* and mix only with other *minus* states.

Minus and *plus* states behave very differently in terms of electron correlation. While *minus* states may have important double excitation character, *plus* states exhibit large dynamic $\sigma - \pi$ polarization effects (see Section 4.2). The *plus* and *minus* pseudoparity labels are also highlighted for all excited states in Figure 2. The transition dipole moment between states of the same pseudoparity is zero in the idealized case, while the transitions between *minus* and *plus* states are optically pseudoparity-allowed. Thus, excited states of *minus* character are dark with respect to excitation from the *minus* ground state, while *plus* excited states are allowed by pseudosymmetry rules. Note that alternancy symmetry appears in addition to spatial symmetry. Pseudosymmetry rules for absorption intensities can only apply if transitions are allowed by symmetry selection rules for electronic absorption (*e.g.*, not in inversion-symmetry-forbidden *trans*-polyene A_g states). Furthermore, pseudosymmetry rules are slightly alleviated when it comes to realistic computations, but the general trends are well predicted, in accord with the transition dipole moment analysis of the previous subsection.

2.3 Valence bond theory

Unlike the MO theory, which invokes orbital excitations to describe electronically excited states, in VB theory all electronic states (including the ground state) are described as linear combinations of VB structures (*e.g.*, resonance structures). A well-known example is the ground state of benzene, which is approximately described as a plus linear combination of

two Kekuleé structures, along with other corrections (*e.g.*, Dewar structures). Alternancy symmetry provides a trivial relationship between the MO and VB pictures in all π-conjugated hydrocarbons. *Minus* states are covalent in VB theory, while *plus* states are ionic.[94,133,135–137] Explicit VB computations were performed by Hirao *et al.*[135,136] using the CASVB method, a VB equivalent of CASSCF, showing that *minus* states are mainly composed of neutral covalent resonance structures, while the *plus* states consist of charge-separated ionic resonance structures. For instance, the L_b state of benzene is described as a minus linear combination of two Kekulé structures in the VB framework, and therefore it is dominantly a covalent state, just like the A_g/A_1 excited state of *trans/cis*-butadiene, which mostly consists of Dewar-like diradical structures in the VB description. These dominantly covalent states also contain minor contributions from the ionic structures. On the other hand, the L_a in benzene and B_u/B_2 state in *trans/cis*-butadiene are composed of charge-separated ionic resonance structures, mainly singly and doubly polar, without contributions from pure covalent structures. Indeed, the covalent nature of the acene L_b and the polyene A_g^- state, as well as the ionic nature of the acene L_a state and the polyene B_u^+ state, are well known in the literature.[82,93] The reference to alternancy symmetry makes this characterization straightforward without even performing explicit VB computations.

2.4 Molecular plasmonics

Plasmons are collective excitations in periodic materials that play an important role in various optoelectronic applications. There has been a significant effort to determine what the analogues of plasmons are in small molecules, *i.e.*, the plasmon-like molecular excitations.[138,139] The term molecular plasmon has been used interchangeably in somewhat different contexts, since there is no exact definition or general agreement on what the term should be used for. Here, we focus entirely on two studies: the work of Aikens *et al.*[126] on plasmon resonance in acenes, and the study of Krauter *et al.*[140] on the identification of plasmons in *trans*-polyenes.

Due to their similarity to a periodic graphene sheet, acenes are logical target molecules to look for molecular plasmons. Aikens *et al.*[126] analyzed $\pi\pi^*$ excited states in small acenes from naphthalene to hexacene, and found that a plasmon occurs when all MO configurations interact constructively. The constructive interaction was found in case of the high-energy β-band, while its low-intensity counterpart, the α-band (L_b), is a result of destructive interaction. Therefore, following Aikens *et al.*, the β-band is an example of a molecular plasmon. The optically intense p-peak (L_a) was not described as a plasmon but as a single-particle excitation.

On the other hand, Krauter *et al.*[140] identified plasmons based on the so-called scaling approach,[141] which relies on the fact that plasmonic excitation energies strongly depend on the electron-electron interaction, as opposed to single-particle excitations. The characterization was further confirmed by inspection of transition densities and other quantities. Plasmon-like states are strongly affected by non-dynamical long-range correlation (*i.e.*, the exchange part of electron-electron correlation), while dynamical correlation effects, which are typically introduced *via* higher orders of perturbation theory are more important for so called single-particle excitations.[140] According to this analysis, the first plasmon in *trans*-polyene is B_u^+ (one configuration), the second plasmon is A_g^+ (two configurations), the third plasmon has three configurations, *etc.* In contrast, A_g^- is characterized as a single-particle excitation. It was found that single-particle excitations have strong contributions from doubly-excited configurations, as opposed to plasmons which are composed of singly-excited configurations. We note that plasmon-like excitations, as defined by Krauter *et al.* coincide with ionic *plus* states, while single-particle excitations coincide with covalent *minus* excited states. Importantly, unlike the analysis of Aikens *et al.*,[126] Krauter *et al.*[140] do not classify plasmons according to their large absorption intensity, as some of the plasmon-type states are optically dark – *e.g.*, plasmonic A_g^+ of *trans*-polyene is symmetry-forbidden. Nevertheless, the "plasmonic potential" of A_g^+ becomes evident when the geometry flips from *trans* to *cis* form – the dark A_g^+ becomes the strongly allowed A_1^+, *i.e.*, the *cis*-peak shown in Figure 2.

Note that the same conclusion – that plasmonic excitations correspond to ionic states in VB theory – was also indicated by Kimber and Plasser.[142]

2.5 Summary

The different ways to look at the two lowest $\pi\pi^*$ excited states in conjugated molecules (polyenes, heterocycles, and acenes) are summarized in Table 1. These general properties could certainly be extrapolated to other conjugated molecules, but one has to be careful when doing so. For instance, Schulten *et al.*[143] argued that all $\pi\pi^*$ excited states in carbonium ions, cyanine dyes, and protonated Schiff base polyenes are, to a great extent, of covalent character, and are poorly described by singly excited configurations alone. Excited states of these compounds could not fit within the dichotomy shown in Table 1.

Table 1: Contrasting the two low-lying $\pi\pi^*$ singlet excited states in common π-conjugated molecules ("heterocycles" refer to the five-membered ring heterocycles studied herein).

B_u (*trans*-polyenes)	A_g (*trans*-polyenes)
B_2 (*cis*-polyenes)	A_1 (*cis*-polyenes)
B_2 (heterocycles)	A_1 (heterocycles)
L_a (acenes)	L_b (acenes)
L_w	L_s
ionic	covalent
plus	*minus*
single excitations	singles & doubles
$\|H \rightarrow L\rangle$ configuration	multiconfigurational
plasmonic	not plasmonic
bright	dark

3 Computational details

The ground-state geometries of furan, pyrrole, thiophene, and benzene were taken from Ref. 123, where they were optimized using the CC3/aug-cc-pVTZ method. The ground-state structures of selenophene, tellurophene, phosphole, arsole, and *cis*-butadiene were optimized with the MP2/aug-cc-pVTZ method (the structures and additional details are available in

the SI). Note that for *cis*-butadiene there has been a debate about whether the ground-state minimum is a planar C_{2v} geometry or a non-planar *gauche* geometry with C_2 symmetry.[144,145] MCSCF computations predicted a planar geometry,[144,146] and many excited-state studies on *cis*-butadiene were performed for a planar system.[98,101,135,144,146,147] Nevertheless, MP2 predicts a *gauche* rotamer,[148,149] in accord with the high-level coupled-cluster computations[145] (but also the MCSCF with a large active space[150]). Since we use *cis*-butadiene only as a model system to compare with planar heterocycles, we opt here for a planar geometry, *i.e.*, the planarity was imposed during the optimization.

Excited-state computations (ADC(1), ADC(2), ADC(3), TDDFT) and analyses were performed with the Q-Chem 6.3.1 electronic-structure package.[151] We employ a strict variant of ADC(2), where single excitations are treated to second order in perturbation theory, the coupling between the singles and doubles manifolds in first order of perturbation theory, while double excitations are computed only to zeroth order (*i.e.*, only as orbital energy differences).[120] In ADC(3), all perturbation orders are consistently one order higher than in strict ADC(2); ADC(3) does not explicitly treat triple excitations, which are, for instance, included in CC3.[120] Note that triple excitations are still implicitly present in ADC(2)/ADC(3) excited-state wavefunctions and energies through higher-order terms in perturbation theory.[117] In this study, we employed conventional TDDFT within the linear-response formalism and the adiabatic approximation. Excited-state characters were determined using natural transition orbital (NTO) analysis in Q-Chem.[152,153] Since NTOs are computed for each excited state separately, the downside of this approach is that the same types of orbitals (*e.g.* HOMOs) are not completely identical in different states (which can be circumvented by using state-averaged NTOs[152]). For raw data, additional details, and supplementary computations (including TDDFT with other functionals and CASSCF/CASPT2 calculations) that are not presented in the main text, see the SI.

CC3/aug-cc-pVTZ excitation energies of Loos *et al.*[123] were taken as reference values for furan, pyrrole, thiophene, and benzene. The reference CC3 energies were nearly converged

with the aug-cc-pVTZ basis set, and the deviation from CC3/aug-cc-pVQZ was between 0.00 and 0.02 eV. CC3/aug-cc-pVTZ singlet excited states of *cis*-butadiene were computed using the eT package.[154]

4 Results and discussion

4.1 Model systems

The similarity between the $\pi\pi^*$ excited states of polyenes, acenes, and five-membered ring heterocycles does not seem accidental if we compare the frontier π-orbitals of the simplest model systems. Figure 3 schematically depicts the frontier π-orbitals of benzene, *cis*-butadiene, and a heterocycle – showing that the orbitals exhibit a very similar arrangement of nodes and lobes. In particular, the frontier orbitals of a heterocycle are somewhere in between those of benzene and *cis*-butadiene. Note that in benzene, HOMO and HOMO-1 are degenerate, just like LUMO and LUMO+1. However, this is specific to benzene and generally does not hold for other compounds, including the larger acenes.

We now take a closer look at the excited-state characters of three model compounds: benzene, planar *cis*-butadiene, and furan. For the sake of analysis, we opt for the ADC(3) electronic-structure method, although other methods (*e.g.*, MCSCF, CC, TDDFT) could also serve well. It is advantageous to employ a method that can capture the double-excitation character of the target excited states. The choice of ADC(3) is also partly motivated by the implementation of wavefunction analysis tools in the Q-Chem package.[152,153] ADC(3) computations typically rely on canonical Hartree-Fock (HF) orbitals, which are not suitable to rationalize the excited-state characters. This is especially true when diffuse basis sets are used (see the discussion in Ref. 155). As an alternative to the HF orbitals, we visualize the excited states using NTOs, which provide a compact description of orbitals involved in single excitations,[142] and have more physical meaning than canonical orbitals.[156] NTOs are obtained by singular value decomposition of the one-particle transition-density matrix

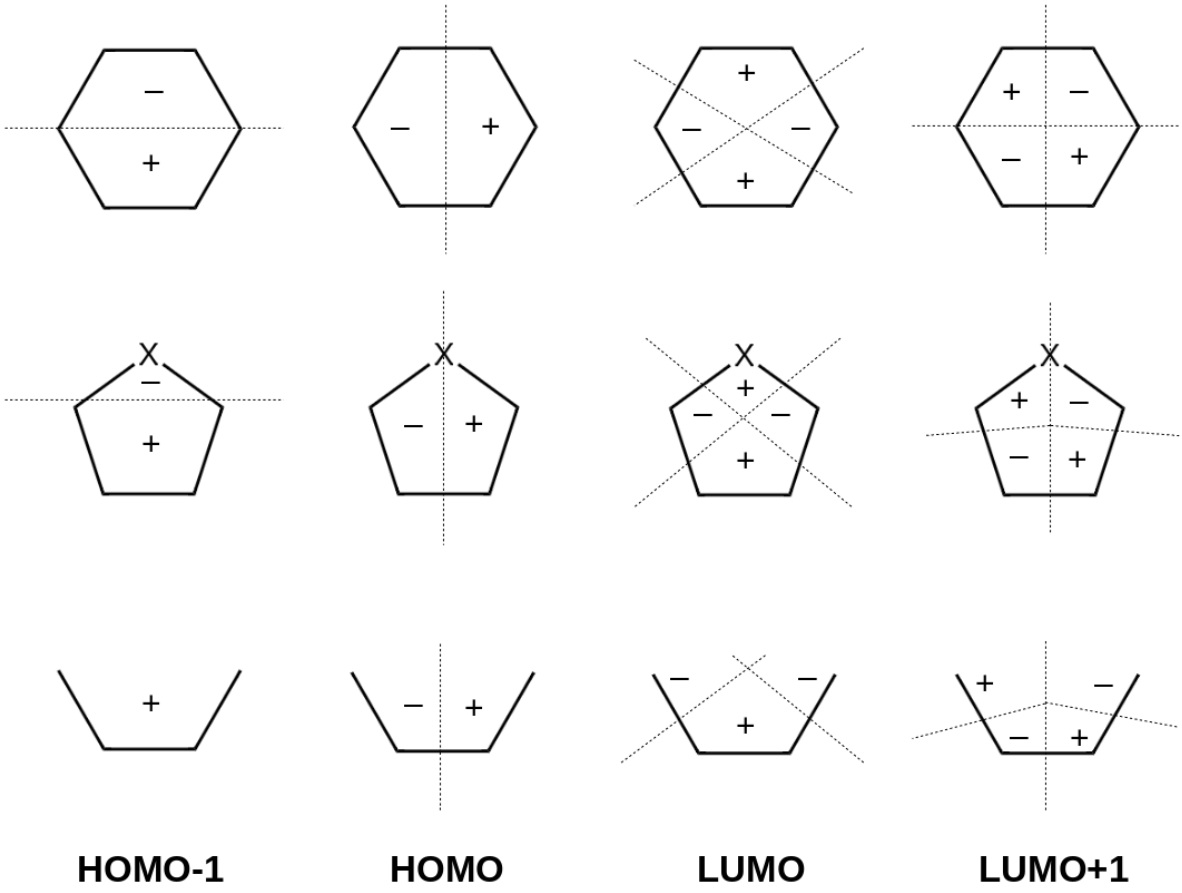


Figure 3: Depiction of the frontier π-orbitals in benzene (top), a five-membered ring heterocycle (middle), and *cis*-butadiene (bottom). Note that in benzene, the two HOMOs and the two LUMOs are nearly degenerate. $\pm$ sign designates the π orbital phase, which changes upon reflection through the molecular plane. Dotted lines represent approximate and simplistic borders between orbital lobes.

(1TDM), which can be interpreted as an effective exciton wavefunction.[142,157] Each pair of hole and particle NTOs has a corresponding weight. The sum of all NTO weights is equal to Ω, which is the squared Frobenius norm of the 1TDM. In practice, Ω represents the total weight of all single excitations, whereas the value of $1-\Omega$ is suggested as an estimate of the double-excitation character, or more precisely the amount of all non-singles.[152,153] Proper analysis of the double excitations requires considering a two-particle transition-density matrix,[153] but this is beyond the scope of this work.

First, we investigate the lowest covalent *minus* singlet excited states of three model compounds (Figure 4). All orbital labels (H, L, $H-1$, $L+1$) are assigned according to Figure 3, considering the NTO shapes and their nodal structure. From the NTO analysis, it is evident that these excited states are multiconfigurational even at the level of singles. The $1-\Omega$ values differ for the three compounds, showing that the double-excitation character is relatively small in benzene, somewhat larger in furan, and dominant in butadiene. Note that the ADC(3) method is not very accurate for polyenes and overestimates their double-excitation character.[a] Multireference methods predict approximately 50% of the double-excitation character in A_1 of *cis*-butadiene[24] (see also Table S11 for our CASSCF computations). On the other hand, the L_b state of benzene is dominantly singly excited. Hence, the L_b states of acenes were traditionally not brought into relation with "doubly excited" states of polyenes. Notable exceptions exist, for instance, in the SI of the Ref. 34 the L_a/L_b notation was also used for excited states of polyenes, which was certainly motivated by the similar orbital excitations involved in acenes and polyenes. Nevertheless, Platt's L_a/L_b notation is hardly acceptable in polyenes if we consider the original definition by Platt, which is based on the perimeter model – *i.e.*, it applies to cyclic molecules. For heterocycles, the L_a/L_b labels could be used. These compounds are somewhere in between acenes and polyenes: heterocycles resemble benzene by similar orbital excitations of the singles and the relatively low double-excitation character, while on the other hand they share the same symmetry

[a]Polyenes are notoriously difficult to treat – the quantitative description of electronic states of polyenes with the CI method requires single, double, triple, and quadruple excitations.[105,158]

of $\pi\pi^*$ excited states (A_1/B_2) as a planar *cis*-butadiene model. The symmetry of excited states has some further repercussions. The double-excitation character in the A_1^- state of a five-membered ring heterocycle and *cis*-butadiene mainly corresponds to the $|H, H \rightarrow L, L\rangle$ configuration, which is the energetically most favorable configuration that belongs to the totally symmetric A_1 irrep (all orbital levels doubly filled). In acenes, a configuration of the $|H, H \rightarrow L, L\rangle$ type will also belong to the totally symmetric irrep A_g (A_{1g} in benzene), and therefore it will contribute to a completely different excited state, which will not mix with L_b. From that point of view, the A_1^- state of heterocycles should not be labeled as L_b, but only as an L_b-like state.

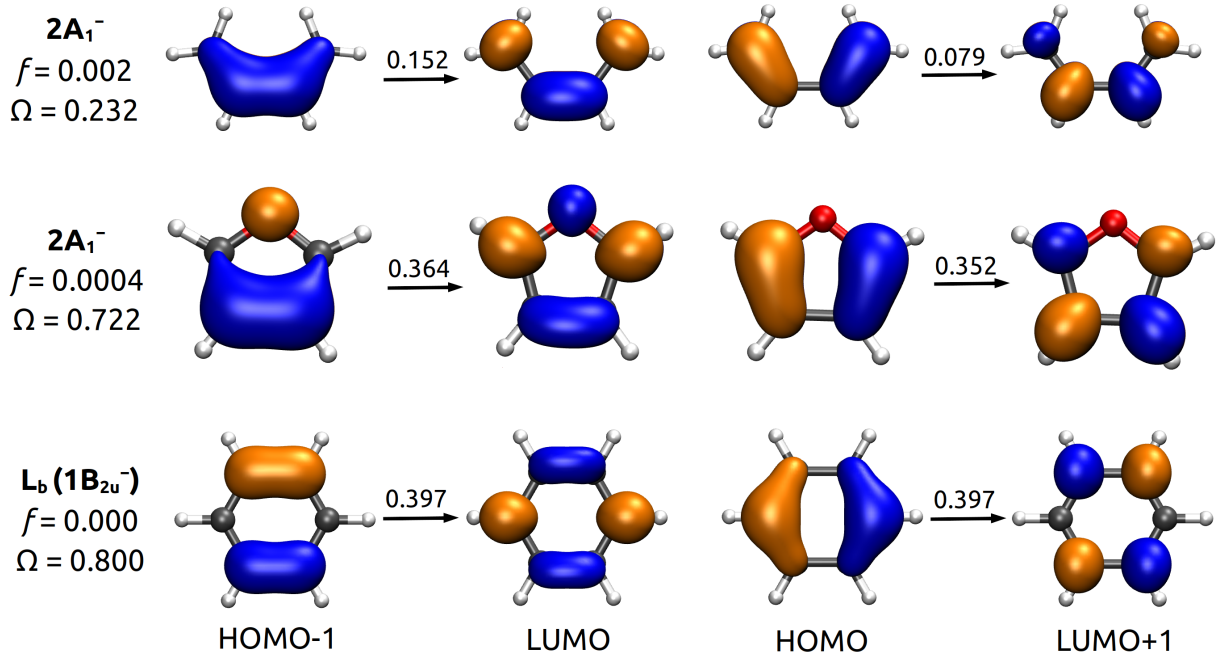


Figure 4: Dominant hole/particle NTOs of the low-lying covalent *minus* singlet excited states of *cis*-butadiene, furan, and benzene. Excited-state symmetry labels are listed along with oscillator strengths (f) and Ω values. NTO weights are given for each hole–particle pair. ADC(3)/aug-cc-pVTZ level; isovalue = 0.06.

The oscillator strengths of covalent *minus* states are small, as predicted by the alternancy symmetry rules. The states are not forbidden, but the transition dipole moments of the singly excited configurations largely cancel each other, while double excitations do not contribute to one-photon absorption. The oscillator strength is proportional to the squared value of the

net transition dipole moments. Note that in *cis*-butadiene, two singly excited configurations do not have perfectly balanced weights, which is also the case in the CASSCF wavefunction.[24] This can be explained by orbital level shifts, likely due to larger intramolecular interactions compared to the *trans*-isomer.[b]

As shown in Figure 2, the low-lying covalent *minus* state is paired with a high-lying ionic *plus* state of the same symmetry (*e.g.*, A_1^+ in furan and *cis*-butadiene). These high-lying states are optically very intense (*e.g.*, the *cis*-peak in Figure 2) and tend to exhibit a smaller double-excitation character. Their oscillator strengths at the ADC(3) level amount to 0.278, 0.671, and 0.336 for *cis*-butadiene, benzene, and furan, respectively (Table S2). Note that in the case of benzene the large oscillator strength corresponds to a doubly degenerate state (see below). The $1-\Omega$ values are 0.23, 0.13, and 0.22 for *cis*-butadiene, benzene, and furan, respectively (Table S7). These values are not as small as one would expect from the idealized picture (*i.e.*, no mixing with double excitations), especially in furan and *cis*-butadiene. This can be explained by the breakdown of alternancy symmetry, which has the consequence that double excitations may lose their *minus* features and start to mix with *plus* states as well.[17]

The NTOs for the lowest ionic *plus* singlet excited states are shown in Figure 5. The amount of higher-than-singles character in these excited states is small. Ω values are greater than 0.85, which implies that these states are dominantly singly excited. The remaining corrections of up to 15% can be considered as usual differential correlation effects. Therefore, single-excitation methods that cover basic correlation effects, such as CC2, ADC(2), or even LR-TDDFT, should be able to properly describe such states.

From the present analysis we see that low-lying B_2^+ states of furan and *cis*-butadiene are strongly dominated by the $|H \rightarrow L\rangle$ configuration and have a large oscillator strength. Regrettably, the L_a state of benzene is an outlier due to the orbital degeneracies of the

[b]Hirao *et al.* argue that the B_2^+ state in *cis*-butadiene is stabilized compared to the corresponding B_u^+ state in *trans*-butadiene because the distance between opposite charges (terminal C atoms) in the dominant VB ionic structures is greatly reduced in the *cis*-isomer.[135] In MO language, this may be interpreted as HOMO–LUMO gap reduction due to orbital energy shifts, which should also affect the composition of the A_1^- state.

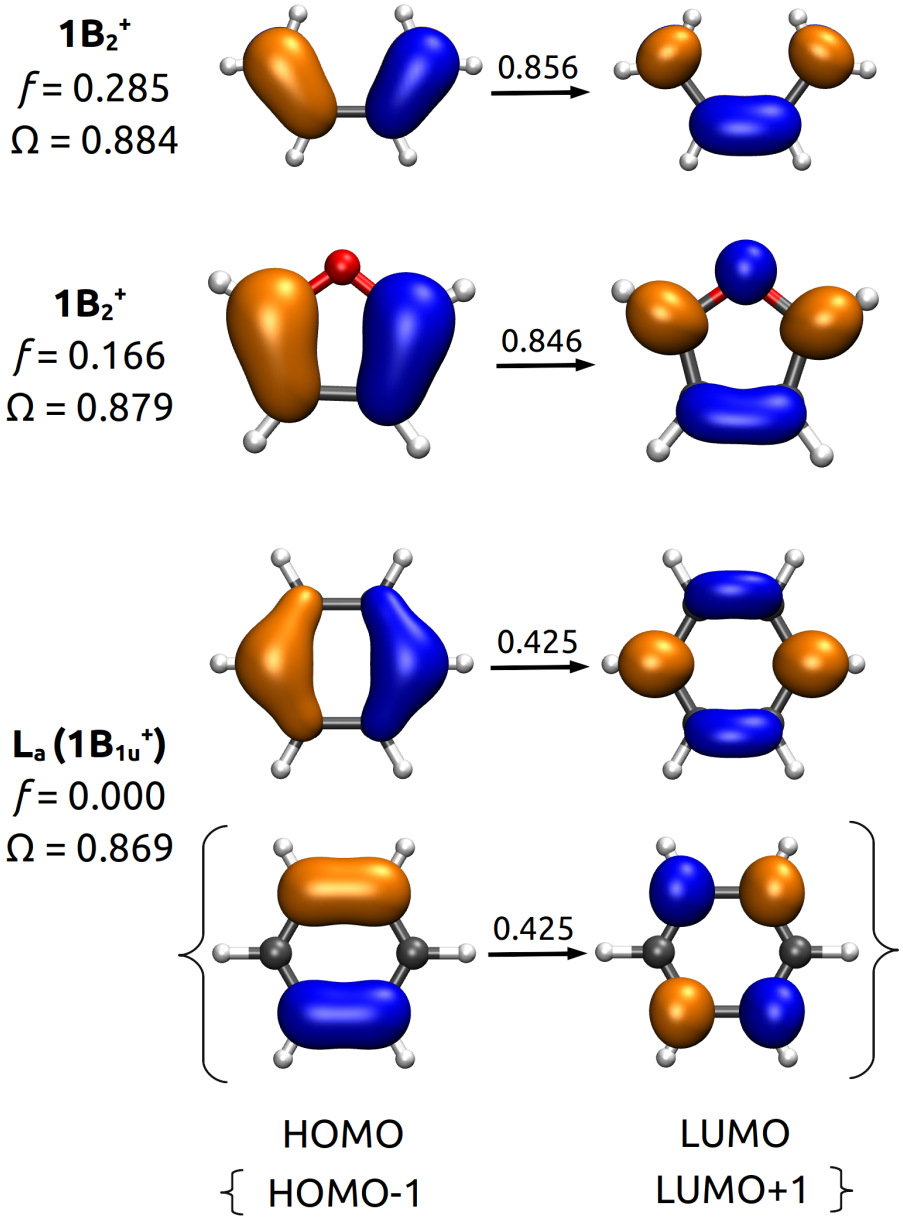


Figure 5: Dominant hole/particle NTOs of the low-lying ionic *plus* singlet excited states of *cis*-butadiene, furan, and benzene. Excited-state symmetry labels are listed along with oscillator strengths (f) and Ω values. NTO weights are given for each hole–particle pair. ADC(3)/aug-cc-pVTZ level; isovalue = 0.06.

HOMOs and LUMOs. Benzenes' L_a is composed of equal amounts of $|H \rightarrow L\rangle$ and $|H-1 \rightarrow L+1\rangle$. As noted earlier, the two configurations possess antiparallel transition dipole moments which cancel each other, hence, the L_a state of benzene is dark, just like L_b.[c]

In larger linear acenes (*e.g.*, naphthalene, anthracene, *etc.*) the orbital degeneracy is split, L_a state is strongly dominated by $|H \rightarrow L\rangle$, and the excited state has a large oscillator strength.[133,134] For these compounds, the second configuration ($|H-1 \rightarrow L+1\rangle$) is energetically less favorable, so its contribution to the L_a state is small. Accordingly, in the lowest B_2^+ state of furan and *cis*-butadiene, the amount of $|H-1 \rightarrow L+1\rangle$ character is smaller than 1% (Figure 6). Evidently, a high-lying B_2 state also exists (which is also ionic *plus*![17]), and is dominated by $|H-1 \rightarrow L+1\rangle$ with a small mixture of $|H \rightarrow L\rangle$. This excited state appears in the vicinity of the A_1^+ state (*cis*-peak), it is bright, and typically shows as a knee next to the *cis*-peak (for simplicity not shown in Figure 2). The same principle holds for linear acenes – there is a high-lying ionic bright state of the same symmetry as L_a, but dominated by $|H-1 \rightarrow L+1\rangle$.[133,134] In this respect, benzene is again very peculiar.[133] The dark L_a of benzene (B_{1u}^+) is paired with the higher-lying bright ionic state (B_{1u}^+), which also consists of equal fractions of $|H \rightarrow L\rangle$ and $|H-1 \rightarrow L+1\rangle$, but with a different CI sign. It gets even more complex as this high-lying B_{1u}^+ is degenerate with B_{2u}^+, which is a high-lying ionic pair of L_b. The high-lying B_{1u}^+ and B_{2u}^+ together form a doubly degenerate E_{1u}^+ state, which has the net oscillator strength from both of its two components.

In the following sections, we will set aside the discussion about high-lying states and focus on two low-lying $\pi\pi^*$ states of heterocycles that play a vital role in their nonadiabatic photochemistry and photophysics.

[c]The behavior is similar in porphyrin free base which has nearly degenerate HOMOs and LUMOs. L_a and L_b states are both dark and together they form a Q band. In their FCI-QMC study, Alavi *et al.*[159] noticed that "dynamical correlation is more important for Q_y band then for Q_x band", posing the question: "why would two qualitatively similar states behave so differently in terms of correlation effects?" It is due to the different nature of L_a and L_b states.[90]

4.2 Characters of low-lying excited states: $^1B_2^+$ vs $^1A_1^-$

We analyze the characters of the two lowest $\pi\pi^*$ singlet states, B_2^+ and A_1^-, in furan, pyrrole, and thiophene, as well as the equivalent states in *cis*-butadiene and benzene for comparison. Complementary data for selenophene, tellurophene, phosphole, and arsole are presented in the SI. While we acknowledge that ADC(3) results are not necessarily of benchmark quality, particularly with respect to accurately capturing the fraction of double-excitation character, the primary objective here is to revisit the concepts introduced in the Section 2 using real compounds.

The B_2^+ state of all heterocycles and butadiene is mostly composed of the $|H \rightarrow L\rangle$ configuration (Figure 6); the contribution from the second configuration, $|H-1 \rightarrow L+1\rangle$, is small but still detectable. In the L_a state of benzene, the two contributions are equal due to orbital degeneracies. The estimated amount of non-singles is almost constant among the compounds, and it falls within the margins characteristic of dominantly singly excited states.

It is interesting to take a closer look at the other singles configurations (dark green bars in Figure 6) that appear as corrections to the dominant $|H \rightarrow L\rangle$. These are configurations of $n\sigma^*$ and $\sigma\sigma^*$ character. Their contribution is generally larger for heavier heteroatoms (*e.g.*, up to 8% of "other singles" in tellurophene B_2, see Table S6). Most of these additional configurations can be rationalized as dynamic σ-polarization, which is known to be important in ionic states.[160] Note that a small fraction of configurations involving σ orbitals also appears in covalent A_1 states, but it is at least twice as small as in ionic B_2 states (Tables S6 and S7). This indirectly indicates that covalent states also have a certain amount of ionic character (*e.g.*, alternancy symmetry is only approximately satisfied), which is also valid for covalent states of hydrocarbons, as well as covalent ground states.[135,136] Covalent states generally exhibit some degree of ionic character.

The σ-contributions to $\pi\pi^*$ states can be rationalized as follows. The ionic π-system, which is described in VB theory by charge-separated resonance structures, polarizes the σ-

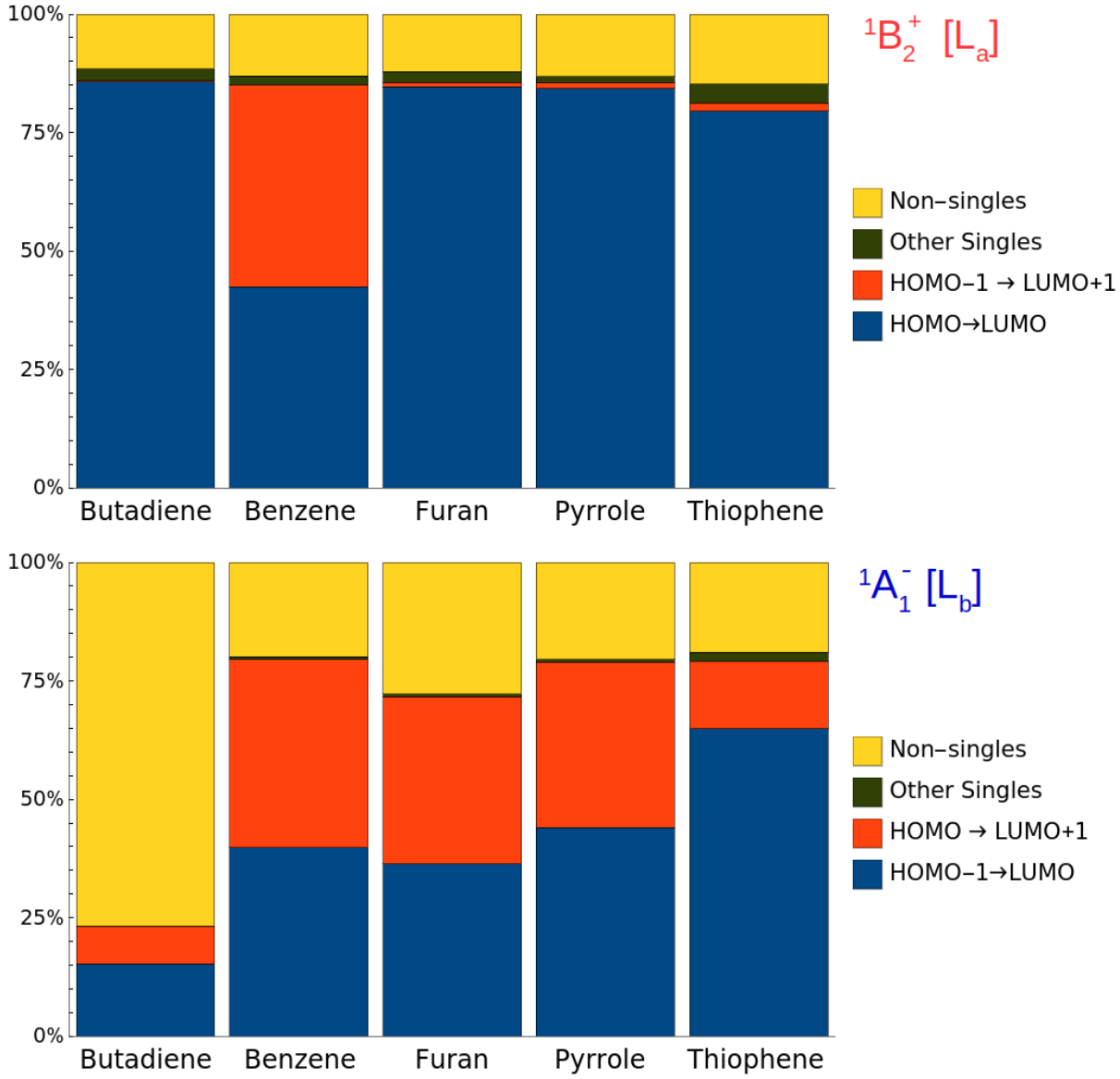


Figure 6: The weights of different configurations in the $^1B_2^+$ and $^1A_1^-$ singlet excited states of *cis*-butadiene, furan, pyrrole, and thiophene. In the case of benzene, we refer to the L_a and L_b labels in square brackets. The amount of non-singles is estimated as $1 - \Omega$, while the amount of other singles is estimated by subtracting the NTO weights of the dominant configurations from the Ω value. ADC(3)/aug-cc-pVTZ level.

skeleton of the molecules. The response of the σ-electrons is, in MO language, equivalent to adding configurations that involve σ-orbitals (both single and double excitations[160,161]). Dynamic σ-polarization is followed by p-contraction, a secondary effect related to the reduction of ionicity due to σ polarization. These two effects together are called the $\sigma - \pi$ correlation.[160] The importance of $\sigma - \pi$ correlation is well established within MCSCF theory. The CASSCF method is known to severely overestimate excitation energies of ionic $\pi\pi^*$ states (often by more than 1 eV) if the active space is composed of π-orbitals only (when σ orbitals are inactive and treated only at the mean-field level).[96,133] Certain tricks have been proposed to improve excitation energies at the MCSCF level.[146,147,160] However, $\sigma - \pi$ correlation is only described sufficiently well when dynamical correlation is properly taken into account, for instance via perturbation theory (*e.g.*, in MS-CASPT2[162]).

Kimber and Plasser[142] argued that σ-contributions mixed into the $\pi\pi^*$ wavefunction, even though they are small, greatly reduce the transition dipole moments and oscillator strengths of the bright states. Our results agree with this observation – pyrrole has the lowest amount of "other singles" (1%) and the largest B_2 oscillator strength of 0.171. The oscillator strength of the furan B_2 state is only slightly smaller (0.166), while the amount of "other singles" is around 2%. However, the oscillator strength is greatly reduced for the thiophene B_2 state (0.085), which has around 4% of "other singles" configurations mixed in. This trend continues for the B_2 states of selenophene and tellurophene (see Tables S2 and S6, also including arsole and phosphole).

As opposed to the B_2^+ state, which is dominated by one configuration, the A_1^- state is intrinsically multiconfigurational (Figure 6, lower panel). Two singly excited configurations dominate, namely $|H-1 \rightarrow L\rangle$ and $|H \rightarrow L+1\rangle$. The relative weight of these two configurations, which combine through CI with a minus sign, helps rationalize the oscillator strengths of A_1^- states in heterocycles. The intensity of A_1^- in furan and pyrrole is close to zero, as expected from the alternancy symmetry rules. The two dominant single excitations are expected to interact destructively, and the state appears optically dark. However, in the

A_1^- state of thiophene, the $|H-1 \rightarrow L\rangle$ configuration dominates. This suggests that the cancellation of transition dipole moments is incomplete. In thiophene, the A_1^- state is bright, in fact as bright as B_2^+, suggesting a breakdown of the alternancy symmetry prediction. The $|H-1 \rightarrow L\rangle$ configuration becomes even more dominant in selenophene and tellurophene, which have even larger oscillator strengths, and the same holds for phosphole and arsole (which have a bright A′ state – equivalent of A_1). The reason behind the prevalence of $|H-1 \rightarrow L\rangle$ over $|H \rightarrow L+1\rangle$ is the shift in orbital energy levels. Roos *et al.*[24] attributed the change in orbital energies to the larger quasi-aromatic character in thiophene. However, this explanation does not support the cases of arsole and phosphole, which are very weakly aromatic.

Orbital shifts cannot be directly observed with NTOs (which do not have associated orbital energies), but they are already apparent when considering the DFT Kohn-Sham orbitals. Figure 7 compares the orbital levels of furan and thiophene computed with B3LYP/cc-pVTZ. In thiophene, the LUMO is stabilized, while the HOMO–1 is destabilized as compared to furan, which overall makes the $|H-1 \rightarrow L\rangle$ configuration energetically more favorable than $|H \rightarrow L+1\rangle$. The MO shifts are due to the interactions of π orbital lobes of the conjugated carbon system and the valence p orbitals of the heteroatom (interaction with d orbitals is also possible but is likely small). HOMO and LUMO+1 orbitals have a vertical nodal plane passing through the heteroatom, *i.e.*, they are not significantly affected by the heteroatom orbitals. The interaction affects HOMO–1 and LUMO, which have larger lobes on the sulphur heteroatom. The MO energy shifts become more pronounced for heavier heteroatoms, which have more spacious valence atomic orbitals and thus larger overlap with the π orbitals of the conjugated carbon skeleton. Therefore, $|H-1 \rightarrow L\rangle$ becomes even more dominant down the group, and its weight in A_1^- is the largest in tellurophene (approximately 20 times greater than the weight of the second configuration). In phosphole and arsole, the overlaps are more complex due to the loss of planarity, but the trend (with respect to pyrrole) is the same as with chalcogens – going down the group, the A′ state gains a dominant

$|H-1 \rightarrow L\rangle$ character and a sizable oscillator strength (Table S2). The stabilization of the LUMO has an effect on the B_2 state as well – its excitation energy decreases following the smaller HOMO–LUMO gaps. This happens for all heteroatoms below the 2nd row in the periodic table.

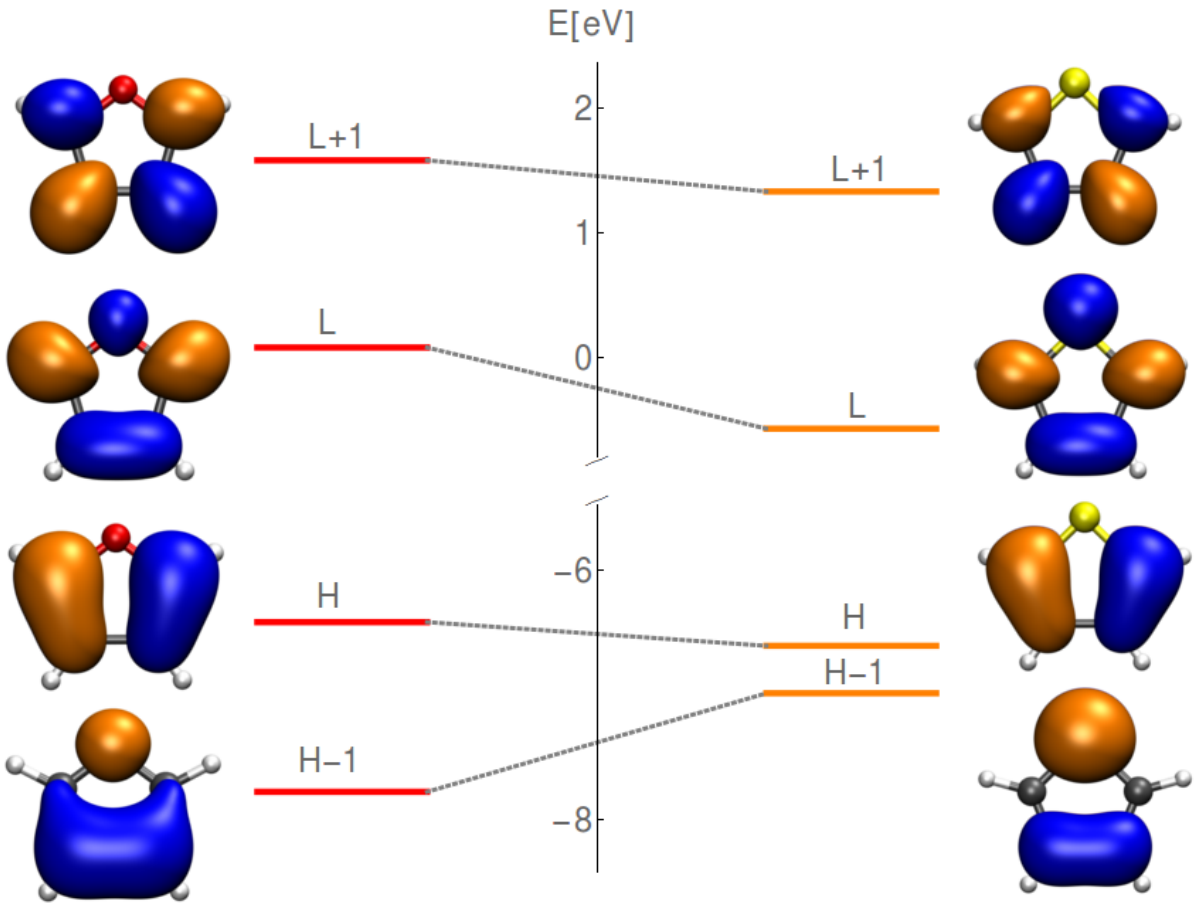


Figure 7: Kohn-Sham π orbital energy levels of furan (left) and thiophene (right). B3LYP/cc-pVTZ level; isovalue = 0.04.

Another interesting issue is the amount of non-singles character in A_1^- (Figure 6). As noted earlier, this partly corresponds to $|H, H \rightarrow L, L\rangle$, which is the lowest-energy double excitation within the A_1 irrep, but the other configurations may mix as well (*e.g.*, $|H-1, H \rightarrow L, L+1\rangle$,[17] *etc.*). The non-singles character is the largest in furan (27.8%), followed by pyrrole (20.4%) and thiophene (19.0%). For selenophene, phosphole, arsole, and tellurophene, the amount of non-singles is relatively small and becomes similar to that of

the B_2^+ state (Tables S6 and S7). Double excitations are likely overestimated at the ADC(3) level, which is the case in polyenes, but even larger double-excitation weights were reported in CASSCF wavefunctions. For comparison, at the CASSCF level the total double-excitation (non-singles) characters were 43.7 (46.9)%, 25.3 (26.8)%, and 15.5 (20.7)% for furan, pyrrole, and thiophene, respectively.[24] How can the amount of double excitations in the A_1^- states of different compounds be rationalized?

Dreuw *et al.*[117] examined why the low-lying A_g^- state of linear polyenes contains a large amount of double-excitation character. From a naive perspective, one might expect that an excited state of doubly excited character should have roughly twice the excitation energy of the corresponding singly excited state. In fact, a simple mathematical treatment demonstrates that the "pure" $\pi\pi^*$ doubly excited state should have an excitation energy equal to the sum of the lowest singlet $\pi\pi^*$ and the lowest triplet $\pi\pi^*$ single excitations. However, the double excitation in many polyenes appears in the lowest (S_1) state in the spectrum, even below the "$HOMO \rightarrow LUMO$" excitation, which is intuitively expected to be the S_1 state. Dreuw *et al.* postulated two conditions that allow the appearance of double excitations among low-energy excited states. The first is configuration interaction with other single configurations, which can be energetically favorable. The second necessary (though not sufficient) condition is a large singlet–triplet (S–T) gap, which pushes the "$HOMO \rightarrow LUMO$" excitation higher in energy. Polyenes have notoriously large S-T gaps (*i.e.*, $E(^1B_u) - E(^3B_u)$). For *cis*-butadiene we estimated the gap around 3 eV at the ADC(3) level. In heterocycles, S–T gaps are smaller than in polyenes, being 2.25 eV, 2.14 eV, and 1.86 eV for furan, thiophene, and pyrrole, respectively (ADC(3) results in Table S2). Nevertheless, the S–T gap values cannot be trivially related to the fractions of double excitations in the lowest *minus* state – for instance, in polyenes the double-excitation character of $2A_g^-$ is rather constant from butadiene to decapentaene, while the S–T gaps become smaller.[94] Large S–T gaps primarily matter as a condition for state inversion but do not explain the efficiency of double-excitation mixing in a *minus* state.

Overall, the smaller double-excitation character in A_1^- states of heterocyclic compounds, as compared to polyenes, can be qualitatively understood by their greater similarity to the L_b state of benzene. Alternative explanations, such as the breakdown of alternancy symmetry properties in heterocycles, are also possible. The efficiency of mixing single excitations in the A_1^- state somehow also coincides with the amount of double-excitation character. In the A_1^- state of furan, two singly excited configurations have roughly equal weights, and the fraction of double-excitation character is the largest among the heterocycles. Pyrrole is very similar, but the imbalance of singles is larger and the double-excitation character drops as compared to furan. Pyrrole is followed by thiophene, with a similar trend. For the larger heteroatoms, one configuration becomes dominant ($|H-1 \rightarrow L\rangle$), with an ever-decreasing double-excitation character (smallest in tellurophene and arsole) – the A_1 state becomes bright and, to a large extent, monoconfigurational, just like the B_2 state.

4.3 Benchmark of selected excited state methods

We assess the performance of several excited-state methods by comparing them with the CC3 reference (for discussion and comparison with the available experimental results, see Section S-II in the SI). The importance of correlation effects is examined through a series of ADC methods, from uncorrelated ADC(1), which is equivalent to CIS, through ADC(2), a method that is appropriate for dominantly singly excited states, to ADC(3), which can capture double excitations. For standard TDDFT, we test functionals from the generalized gradient approximation (BLYP),[163,164] the global hybrid family (B3LYP),[164,165] and the range-separated hybrid family (CAM-B3LYP).[166] Other authors used a similar choice of functionals[80,167] By opting for these functionals, we examine the role of exact exchange (0% in BLYP, 20% in B3LYP, 19% in the short range and 65% in the long range in CAM-B3LYP) in predicting excitation energies. In fact, the results obtained with these functionals are to the large extent representative of their families. Similar results are obtained with other generalized gradient approximation functionals (PBE), global hybrid functionals (PBE0), and

range-separated functionals (ωB97xD), for which the data are listed in the SI (Table S8 and Figure S2). The meta-hybrid functional with a large portion of exact exchange (M06-2X) tends to perform similarly to the range-separated functionals.

CC3 provides benchmark-level values for excited states that do not have a strong double-excitation character – the exception, in our case, is the A_1 state in *cis*-butadiene. Double excitations mixed with single excitations are certainly less of an issue than "genuine" double excitations. For *trans*-butadiene A_g, it was reported that CC3 values differ from the theoretical best estimate by less than 0.2 eV.[110] Nevertheless, our primary goal here is not high accuracy *per se*, but rather to identify common patterns in the behavior of the methods, or a sort of fingerprint of the excited-state energy variations among methods.

Figure 8 compares the relative errors in excitation energies of furan, pyrrole, and thiophene, obtained with different electronic structure methods. When comparing ADC(1) and ADC(2), the most striking is the large stabilization of the A_1 excited state as soon as double excitations are taken into account, even at the lowest order. This is a common feature for all the molecules, including selenophene, tellurophene, phosphole and arsole (Table S2), even though their A_1 states do not have large double-excitation characters, and even seem predominantly mono-configurational. ADC(1) severely overshoots the A_1 state excitation energy by more than 1 eV – the largest deviation being around 1.4 eV for furan. On the other hand, ADC(1) provides an excellent estimate of the B_2 excitation energy, though, the error near zero is undoubtedly accidental. ADC(2) provides a balanced description of both states with relatively small errors of up to +0.2 eV. The contrast between ADC(1) and ADC(2) is equivalent to the difference between CIS and CIS(D), which was proposed as a quick test to differentiate between L_a- and L_b-like states in acene-like molecules.[83] The solid performance of ADC(2) for the A_1 state of all heterocycles is puzzling considering the variable double-excitation characters – for instance, the error in excitation energy is not very different in furan and thiophene, although the A_1 state of furan is expected to have more double excitations. The fact that ADC(2) treats the coupling between singly and doubly excited

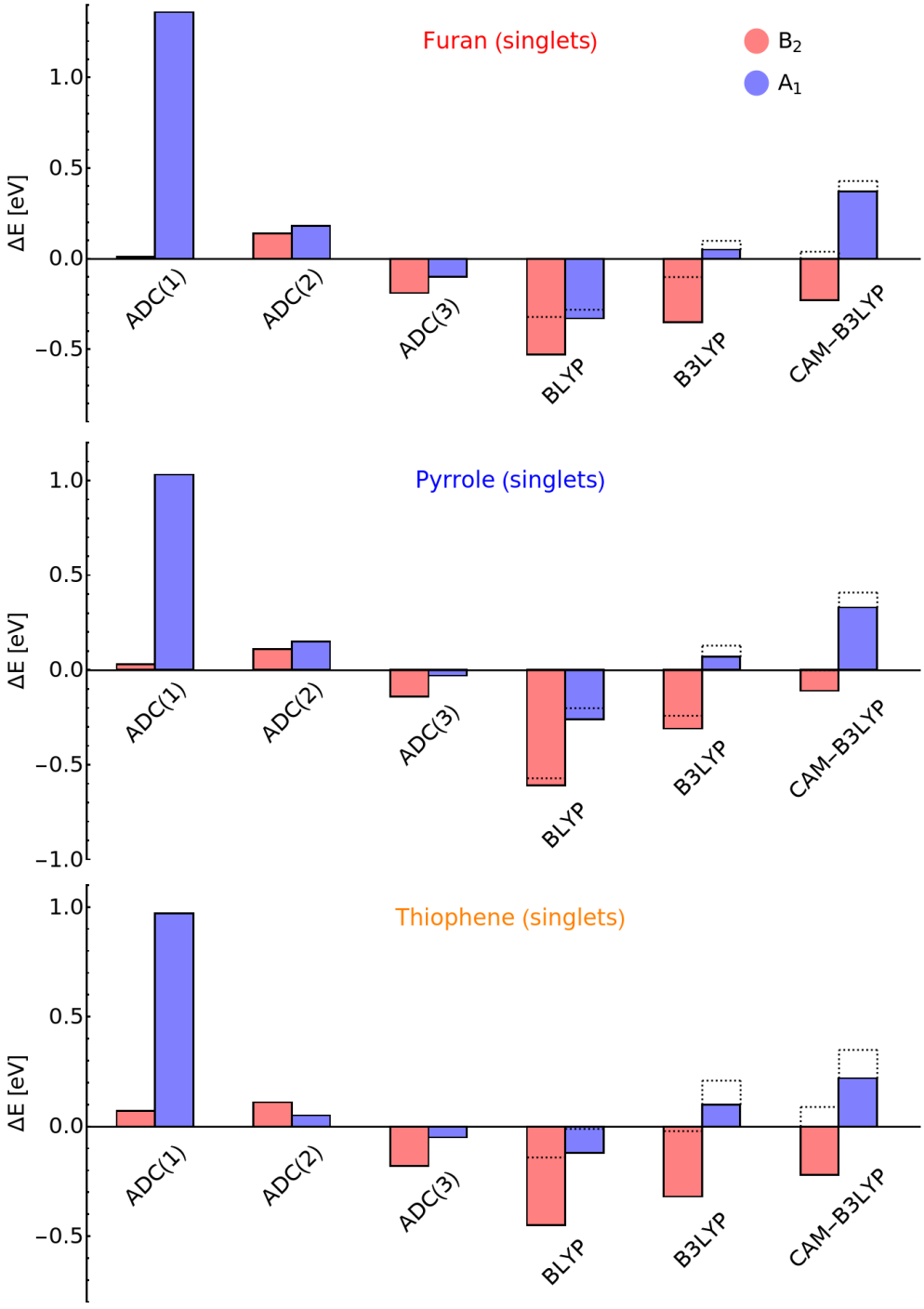


Figure 8: Deviation of excitation energies of the low-lying singlet excited states, $^1A_1$ and $^1B_2$, from the CC3 reference. Bars with dotted lines correspond to the TDA. All data were obtained with the aug-cc-pVTZ basis set.

configurations (treated at first order) might be more important for excitation energies than the proper treatment of double excitations (treated at zeroth order). It follows that the A_1 state might not be well described for compounds with a larger amount of double-excitation character, such as furan, but this is not necessarily reflected in the excitation energy. This deficiency could potentially be reflected in other properties (*e.g.*, quality of excited-state surfaces).

ADC(3) improves A_1 excitation energies, which is expected considering the partial double-excitation character, but overcorrects B_2, which is underestimated by around 0.2 eV. The performance of ADC(3) is somewhat disappointing considering the significant computational effort over ADC(2), which itself performs equally well, or even better than ADC(3). Loos *et al.* proposed ADC(2.5), *i.e.* the arithmetic average of ADC(2) and ADC(3), as a good estimate of excitation energies based on a set of organic molecules.[121]

Pure doubly excited states are completely absent from the adiabatic TDDFT spectrum, but this does not hold for multiconfigurational states with both singly and doubly excited configurations – for these states only single excitation character is captured.[16,116] Tozer *et al.*[16] pointed out the surprising fact that there is no clear correlation between the error in the computed TDDFT excitation energy and the doubly excited character of the state (but this statement is, to some extent, true also for ADC(2)). TDDFT results for the A_1 and B_2 states of heterocycles are shown in Figure 8. The performance of three exemplary functionals by all means resembles the earlier benchmarks on L_a and L_b states.[83] The L_a-like excitation, B_2, is sensitive to the treatment of exact exchange. Note that the L_a state was called a charge-transfer-like state without net charge transfer,[73] which is in principle equivalent to an ionic state. Therefore, the B_2 trends resemble the well-known TDDFT problem with CT states. GGA and hybrid functionals severely underestimate the excitation energies of B_2, while range-separated functionals perform fairly well, with very small errors. However, A_1, *i.e.* the L_b-resembling state, is also sensitive to exact exchange, but in the sense that a larger amount of exact exchange upshifts it way too high. Range-separated functionals

severely overestimate the excitation energies of A_1, which is perhaps a good sign, since standard TDDFT does not treat double excitations and should not predict the excitation energies very accurately. The good performance of B3LYP is likely due to a fortuitous cancellation of errors. Although standard TDDFT only treats single excitations, it contains approximate correlation, so the results are significantly better than CIS, which contains no correlation whatsoever. The fact that TDDFT cannot provide a balanced description of two low-lying $\pi\pi^*$ states, B_2 and A_1 (but also L_a and L_b), whatever functional is employed, points to the inherent deficiency of standard TDDFT approximations, presumably its adiabatic approximation framework. Other TDDFT-based approaches that go beyond or alleviate the adiabatic approximation framework, such as double-hybrid and constricted-variational DFT (CV-DFT), provide a much better description of L_a and L_b states in acenes and acene-like systems.[75,77] Methods such as mixed-reference spin-flip TDDFT (MRSF-TDDFT)[168] have shown good performance even when a large double-excitation character plays a role, enabling a balanced description of the $2A_g$ and $1B_u$ states of *trans*-polyenes.[169]

It is also instructive to look at the performance of the TDA.[170] In fact, TDA is considered a first-order approximation in CV-DFT (which is fully converged with infinite orders). Again, the TDA results (dotted bars in Figure 8) are similar to those from an earlier study on L_a and L_b states.[83] TDA upshifts B_2 by roughly 0.2 eV and A_1 by about 0.1 eV, regardless of the functional. This slightly improves interstate gaps with respect to standard TDDFT, but it can hardly be stated that TDA results are superior to TDDFT. Regarding the size of the shift, one exception is the B_2 state of pyrrole, where the energy correction is much smaller than 0.2 eV. This is probably explained by the sizable valence–Rydberg (VR) mixing in the B_2 state of pyrrole (see Section 4.4).

Finally, charts similar to Figure 8 are also produced for benzene and *cis*–butadiene (Figure S1). While benzene shows a similar pattern to heterocycles, this is not entirely true for butadiene, where a critically large double-excitation character induces much bigger errors in ADC(2) and ADC(3), confirming that polyenes are more challenging systems to describe.

It is somewhat surprising that the results of standard TDDFT are not exceptionally poor, but the "good performance" (e.g., when B3LYP is used) is very likely an accidental match. Nevertheless, even CC3 does not constitute a high-level reference because of the strong double-excitation character of the $2A_1$ state of *cis*-butadiene, which requires at least CC4 calculations to achieve benchmark-level accuracy.[110]

4.4 Valence-Rydberg mixing

The mixing of valence and Rydberg excitations is a well known observation in excited states of organic molecules,[171] for instance in H_2O molecule.[172] Despite the evidence of VR mixing in many molecules, *e.g.* occurring through the Rydbergization process described by Mulliken,[173] the mixing of Rydberg and $\pi\pi^*$ states specifically was often considered unphysical, being a result of an insufficient treatment of electron correlation. CASSCF wavefunctions are prone to erroneous VR mixing; however, with MS-CASPT2 this effect is largely or completely removed.[19,162] Too strong VR mixing might also occur in standard TDDFT.[16] It is known that ionic $\pi\pi^*$ states are more likely to undergo VR mixing,[160] for example in the MCSCF treatment of the ionic $^1B_u^+$ state of *trans*-butadiene[96] (in the *cis*-isomer the mixing is supposed to be slightly smaller[98,144]). Lischka *et al.*[104] performed high-level MRCI and MRCC computations on *trans*-butadiene, showing the clear separation of compact valence and diffuse Rydberg B_u states. Yet, Truhlar *et al.*[109] stated that the valence ^1B_u state has a "certain degree of Rydberg character that is not artificial." The most systematic account of VR mixing is the study of Angeli[160] on the ionic V state of ethylene (*i.e.*, HOMO $\rightarrow$ LUMO), demonstrating that no mixing occurs if correlation effects are treated properly.

The ionic $^1B_2^+$ state of pyrrole has been deemed a notorious case. The state can mix with an energetically close πp_x^* Rydberg state of B_2 symmetry, and depending on the degree of mixing, the two states exchange their characters while the intensity spreads among two transitions. Roos *et al.*[19] reported that some degree of VR mixing appears even at MS-CASPT2. Došlić *et al.*[58] simulated nonadiabatic excited-state dynamics with ADC(2) and

TDDFT, showing that large VR mixing occurs when diffuse basis sets are employed, which leads to the nonradiative decay (upon 200 nm excitation) that is three times slower than the experimental estimate. Conversely, a basis set without diffuse p_x-functions prevents the deleterious VR mixing,[58,59] and good agreement with experiment is observed. On the other hand, Barbatti *et al.*[51] reinterpreted the same experimental results and showed that ADC(2)/aug-cc-pVDZ provides results that agree well with experiments, implying that there is nothing in principle wrong with the electronic structure exhibiting VR mixing. Therefore, the case of VR mixing in pyrrole remains puzzling.

We examine the extent of VR mixing in heterocycles by comparing methods with different degrees of electron correlation: ADC(1), ADC(2), and ADC(3). The expectation value of x^2, *i.e.* $\langle x^2 \rangle$, where x is the axis perpendicular to the molecular plane, is used to measure the state diffuseness; this quantity is often referred to as the second moment.[14,172,174] Alternatively, the measure of electron size[175] is more appropriate for non-planar systems (*e.g.*, phosphole, arsole) and relevant values for ADC and TDDFT computations are given in the Table S4.

Figure 9 shows $\langle x^2 \rangle$ values of the lowest valence and Rydberg states of furan, pyrrole, and thiophene, computed at the ADC(3)/aug-cc-pVTZ level. It is evident that the two B_2 states are largely mixed in pyrrole, which is not the case for furan and thiophene. In furan and thiophene, there is a clear distinction between the valence and the Rydberg state. The B_2 $\pi\pi^*$ state of furan appears slightly more diffuse than the ground state, while thiophene shows no sign of VR mixing. Accordingly, no VR mixing occurs in selenophene, tellurophene, arsole, and phosphole (see Table S4). This is easy to understand following a simple argument. Rydberg state energies are related to the molecular ionization potential, and the lowest Rydberg states in standard organic molecules typically appear around an excitation energy of 6 eV (200 nm) and higher.[92] Rydberg states may mix with the valence states that appear in their energy range, but they have a limited impact on the excited states that are substantially lower. In thiophene and other compounds below the 2nd row of the periodic table, the LUMO

orbital is stabilized with respect to the HOMO, and the excitation energies of B_2 are shifted lower, which prevents mixing with Rydberg states. On the other hand, we find evidence that the high-lying $\pi\pi^*$ states, which are submerged within a number of Rydberg excitations, indeed undergo VR mixing (all values marked with a star in Tables S2-S4 likely exhibit some level of VR mixing.). It is a matter of coincidence that, due to the proximity of a Rydberg state, the large VR mixing occurs for the lowest ionic $\pi\pi^*$ state of pyrrole. Notice also that, due to the lower first ionization potential in pyrrole as compared to furan, the Rydberg states are expected to appear in the lower-energy region.[14]

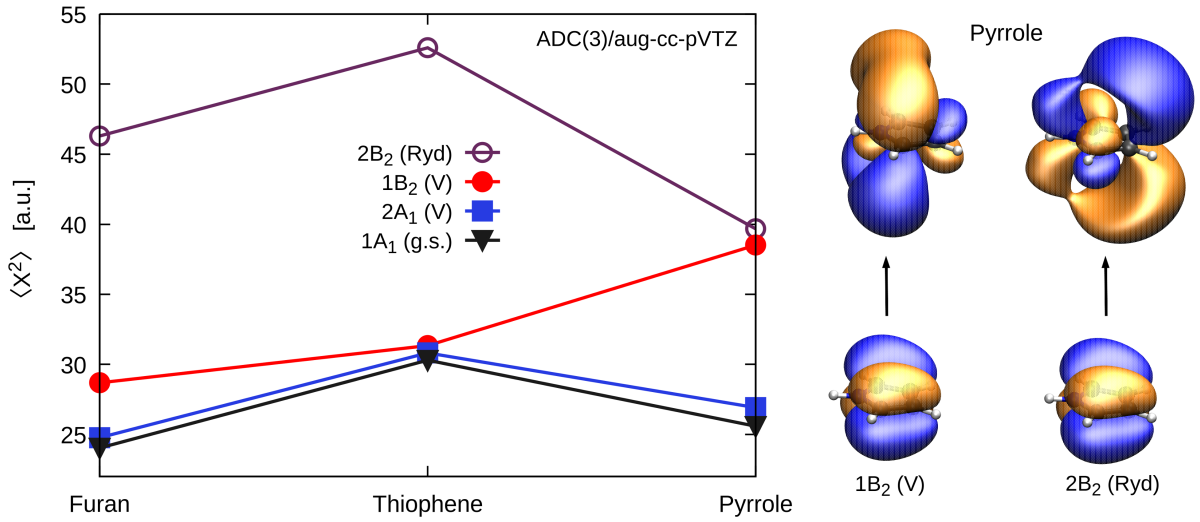


Figure 9: Diffuseness of different singlet electronic states in three compounds, including the ground state (g.s.), two valence states (V), and a Rydberg state (Ryd), is shown on the left. On the right, NTOs of the two B_2 states in pyrrole are shown. Computed at the ADC(3)/aug-cc-pVTZ level; ground state values are based on MP2/aug-cc-pVTZ. Isovalue = 0.02.

In Figure 10, we take a closer look at how $\langle x^2 \rangle$ values of B_2 change within the ADC series. $\langle x^2 \rangle$ in furan and thiophene decreases when going from ADC(1) to ADC(3). This should perhaps not be confused with the elimination of VR mixing, but rather with the effect of *p*-contraction upon introducing dynamic correlation, as described by Angeli.[160] The $\langle x^2 \rangle$ values show convergent behavior within the ADC series. On the other hand, the $\pi\pi^*$ wavefunction is undoubtedly highly mixed in pyrrole for all levels of theory. The mixing always occurs at the level of virtual orbitals – *i.e.*, it does not involve introducing multiconfigurational character.

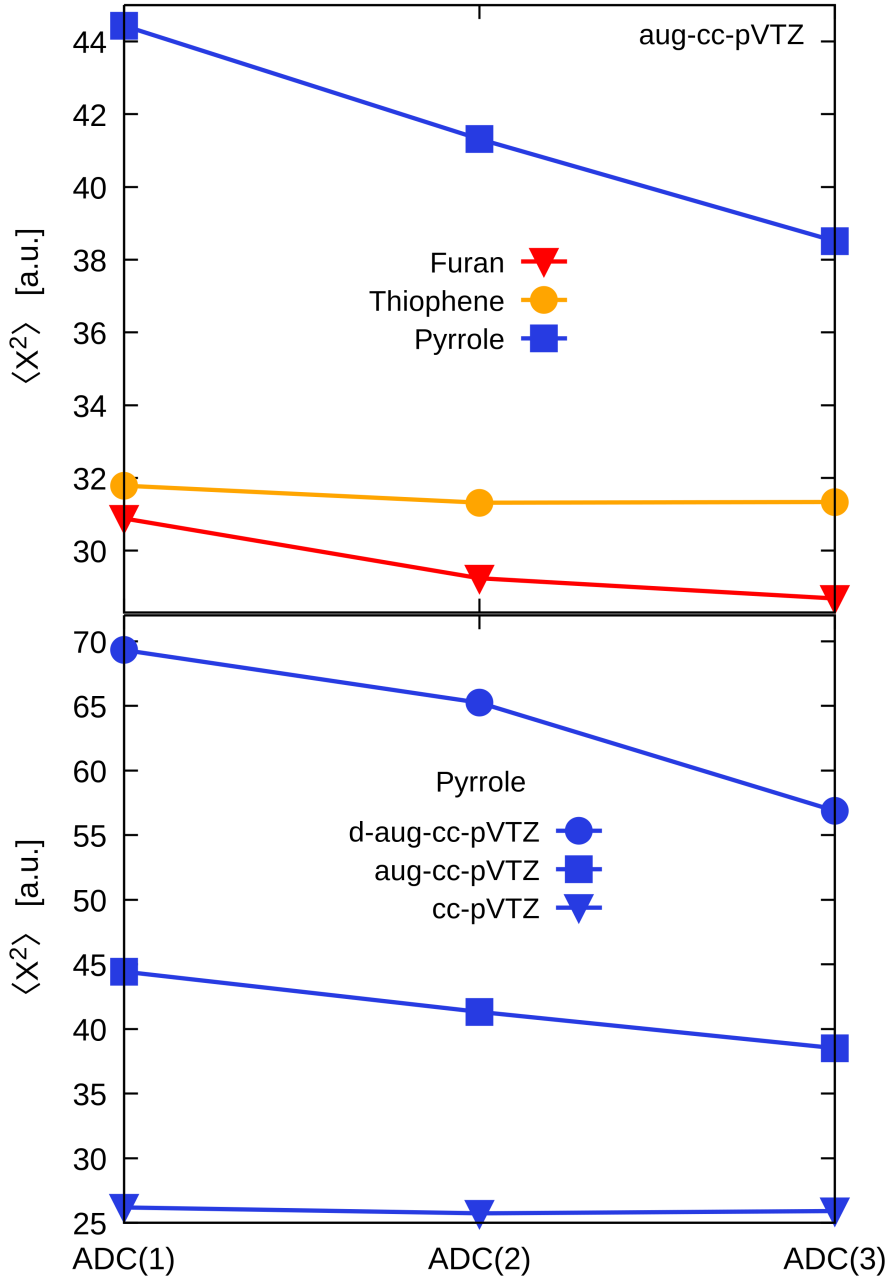


Figure 10: Diffuseness of the lowest valence-like (*i.e.*, bright) B_2 singlet excited states of three compounds, computed with the ADC series and the aug-cc-pVTZ basis set (upper panel). The lower panel shows the diffuseness of the bright B_2 state of pyrrole with different ADC methods and different basis sets.

At the ADC(2) level, the valence-like state (based on NTO shape and large intensity) is even more diffuse than the Rydberg-like state. Note that Barbatti *et al.*[51] assigned the lower B_2 state as a Rydberg state (πp_x^*), while Došlić *et al.*[58] argued that this should be a valence state. Already in 2002, Roos *et al.*[19] stated that "one can find in the theoretical literature as many arguments in favor of the Rydberg character of the band as about its valence nature." Two decades later, this issue is not fully resolved.

Unlike furan and thiophene, the trend from ADC(1) to ADC(3) in pyrrole B_2 does not show any sign of convergence. It is possible that at higher ADC levels (*i.e.*, ADC(> 3)), the VR mixing becomes completely eliminated, given that ADC(n) converges to full-CI with increasing n.[120] We recomputed the pyrrole trend with several different basis sets (Figure 10, lower panel). Obviously, there is no mixing when the non-diffuse cc-pVTZ basis set is used. However, the use of such a basis set compromises the energy convergence of both $\pi\pi^*$ and Rydberg states. Diffuse basis sets show different degrees of VR mixing but share the decreasing $\langle x^2 \rangle$ trend following the increase of electron correlation. In addition to that, we also note that the decreasing trend in $\langle x^2 \rangle$ does not correlate with the energy gaps between the B_2 valence-like and Rydberg-like states (see Table S5; although one should arguably consider diabatic rather than adiabatic states). This further points to electron correlation as the primary underlying factor behind VR mixing. Nevertheless, the case of VR mixing in pyrrole remains moot and may perhaps be resolved by future FCI computations. Alavi, Filippi *et al.*[112] already conducted MQC-FCI computations on *trans*-butadiene B_u, implying the elimination of VR mixing, although no measure was presented in this respect.

Not surprisingly, large mixing in pyrrole B_2 also occurs with TDDFT/TDA, where the two excitations exhibit similar electron sizes which are larger than typical valence and smaller than typical Rydberg state values (see Table S4). TDA generally predicts slightly greater diffuseness than TDDFT, even when no VR mixing is present, likely due to the neglect of de-excitations. As for the covalent $2A_1^-$ states of heterocycles, the $\langle x^2 \rangle$ diffuseness values are similar to those of the ground state ($1A_1$), indicating that there is no significant VR

mixing for these states with any of the electronic structure methods considered here. It was previously argued that the TDDFT description of A_g excited states in polyenes is biased toward Rydberg character,[118,176] which could be a mechanism by which adiabatic TDDFT compensates for the missing double-excitation character. However, we could not find clear evidence of this effect in the A_1 states of heterocycles.

Based on all cases of VR mixing that we were able to pinpoint (including the high-lying states, see Table S4), we find that ionic states are more prone to mixing as opposed to covalent states (one exception was the A_1^- state of *cis*-butadiene, which was substantially mixed at the ADC(1) level). While for these non-pyrrole cases we did not always find a "linear" decrease of diffuseness within the ADC series, in every single example ADC(3) (i.e., the method with the highest level of correlation) predicted the smallest diffuseness.

4.5 Triplets

We demonstrated that the lowest two $\pi\pi^*$ singlet states of five-membered ring heterocycles represent a challenge for electronic structure theory. But what about the corresponding triplets? According to the alternancy symmetry rules provided by Hirao *et al.*, the excited state of $HOMO \rightarrow LUMO$ character is ionic (*plus*) in singlets, but covalent (*minus*) in triplets.[94,133] On the other hand, the A_1 state that is composed of minus CI of $|H-1 \rightarrow L\rangle$ and $|H \rightarrow L+1\rangle$ is covalent (*minus*) for both singlet and triplet multiplicity. Therefore, in the triplet manifold we are dealing with two low-lying covalent *minus* states, and we expect that they will be described in a more balanced manner. This does not mean that no triplets can have ionic character – *e.g.*, higher triplet states such as A_1^+ are ionic.

Figure 11 shows the characters of the low-lying triplet states for furan, pyrrole, and thiophene, along with benzene and *cis*-butadiene for comparison. The B_2 triplet is composed of $|H \rightarrow L\rangle$, with a small contribution from other configurations for all compounds (except benzene – due to orbital degeneracies). "Other singles" mainly comprise $\pi\pi^*$ deexcitations, while the amount of non-singles is small. In $^3A_1$, the ratio of single excitation configurations

is similar to that in the singlet, but the major difference is the small amount of non-singles in the triplet, even in notorious *cis*-butadiene. This can be explained by the fact that configurations of the $|H, H \rightarrow L, L\rangle$ type are forbidden by Hund's rule. That could be the reason for the small amount of double excitation character in covalent $^3B_2$ as well, as many higher excitations in which two spin-parallel electrons end up in the same orbital are spin-forbidden (*e.g.* $|H-1, H \rightarrow L, L\rangle$). Therefore, not only are the two lowest triplet states both covalent, but their double excitation character is also small. Another mitigating factor is the absence of VR mixing in the low-lying triplet states (Table S4). Triplet $\pi\pi^*$ states tend to appear at lower energies than their singlet counterparts. On the other hand, triplet Rydbergs have similar energies as singlet Rydbergs, which can be understood by the same arguments as for diminishing S-T gaps in charge-transfer states, which is due to the small exchange integral (*i.e.* Rydberg states represent a special case of charge separation). Therefore, the lowest $\pi\pi^*$ triplet states are pushed away from Rydberg states, which prevents possible VR mixing.

Figure 12 shows the performance of ADC methods and three representative functionals within standard TDDFT/TDA (TDA results are indicated with dotted bars) with respect to the CC3 reference. The performance of ADC(1) is poor and is characterized by a strong underestimation of two states. Interestingly, the trend is completely opposite to the overestimation of the covalent state in singlets. ADC(2) results are balanced but slightly overestimating excitation energies, just like in singlets. ADC(3) again overcorrects ADC(2) – the results are somewhat underestimated and not fully balanced for two states. The performance of TDDFT is intriguing. It is clear that, with standard TDDFT, results are underestimated, and including more exact exchange significantly downshifts B_2, deteriorating the excitation energies. This issue is known as triplet instability.[167] It seems to significantly affect only one state. The problem is largely solved with TDA (as suggested by Peach *et al.*[167]), which gives smaller errors and a balanced performance for all the functionals. Nevertheless, functionals with a very large amount of exact exchange are not recommended for triplets. In

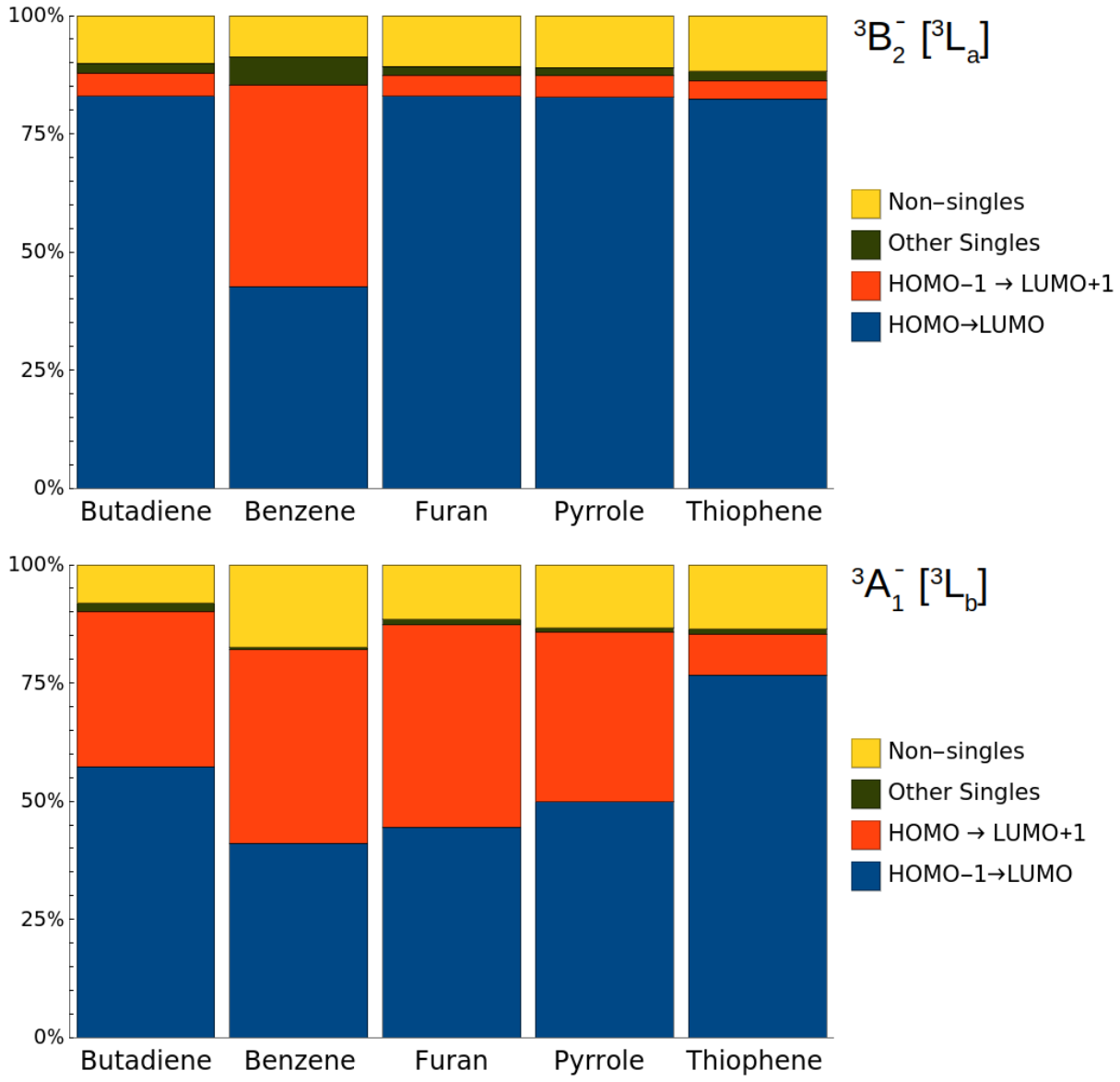


Figure 11: The weights of different configurations in the $^3B_2^-$ and $^3A_1^-$ triplet excited states of *cis*-butadiene, furan, pyrrole, and thiophene. In the case of benzene, we refer to the ^3L_a and ^3L_b labels in square brackets, The amount of non-singles is estimated as $1 - \Omega$, while the amount of other singles is estimated by subtracting the NTO weights of the dominant configurations from the Ω value. ADC(3)/aug-cc-pVTZ level.

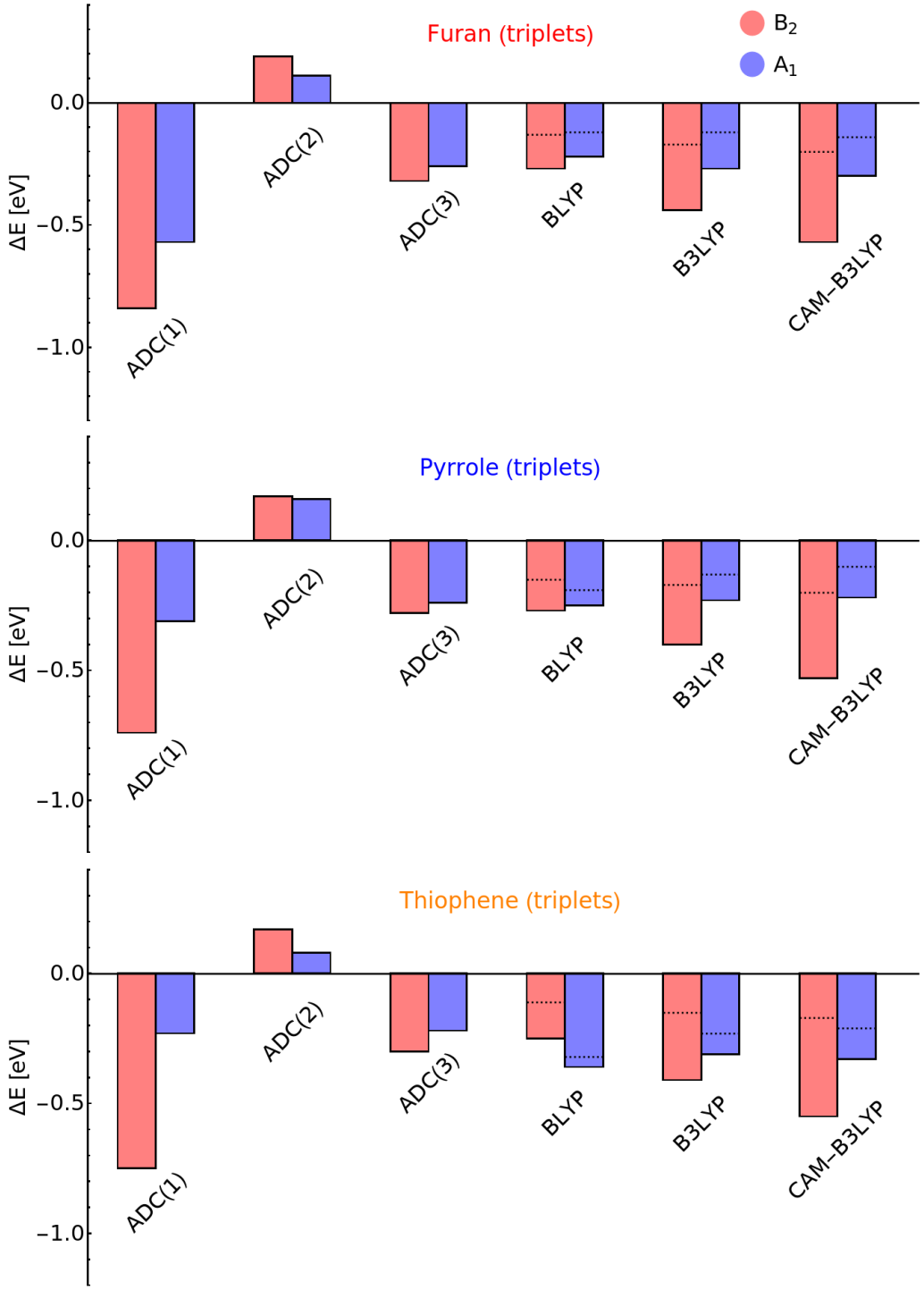


Figure 12: Deviation of excitation energies of the low-lying triplet excited states, $^3A_1$ and $^3B_2$, from the CC3 reference. Bars with dotted lines correspond to the TDA. All data were obtained with the aug-cc-pVTZ basis set.

fact, ADC(1) (*i.e.*, the CIS) is a Tamm-Dancoff approximation of TDHF, and it can be interpreted as a functional with 100% exact exchange and no correlation. The latter is also expected to suffer from a serious triplet instability problem, which also explains the poor performance of ADC(1).

Peach *et al.*[167] argued that the triplet instability problem in TDDFT is similar to the triplet instability in HF theory, where the spin-restricted HF ground state becomes too high in energy due to unphysical ionic components in the wavefunctions (as exemplified by the dissociation of H_2). They introduced a stability measure for each excited state and showed that TDA leads to significant improvements for problematic triplet states. Test computations were performed on butadiene and naphthalene, identifying the triplet excitations for which the presence of exact exchange is undesirable. The two lowest triplets of butadiene, ^3B_u and ^3A_g, behave similarly to the triplet states of heterocycles – the ^3B_u excitation energies decrease dramatically with increasing the fraction of exact exchange in the functional, while the decreasing trend is more moderate for ^3A_g. Interestingly, in naphthalene, only $^3B_{2u}$ (a triplet equivalent of the L_a singlet) experiences the instability, while $^3B_{3u}$ (a triplet equivalent of the L_b singlet) even shows a small increase in excitation energies with a larger fraction of exact exchange. From the perspective of this particular aspect, the electronic structure of heterocycles appears more similar to that of butadiene than to that of acenes. It remains to be examined whether only covalent triplet states can potentially suffer from the instability problem. Note that in the VB description of singlet states, ionic components in the covalent L_b state of benzene are significantly larger than those in the covalent ^1A_g state of butadiene,[135,136] but unfortunately no data are available for the corresponding triplet states.

5 Conclusion and outlook

The problems of the unbalanced description of L_a and L_b states in acenes and related compounds, as well as the lowest B_u and A_g singlet excited states in polyenes, have attracted

large research interest over the decades. Here, we attempt to generalize the problem by examining the common features among the excited states of these compounds, as well as those of closely related five-member-ring heterocycles. To achieve this, we assembled numerous ideas scattered throughout the extensive literature on excited states of conjugated molecules, pointing out the concepts that arguably deserve to be part of standard textbook knowledge. The two lowest excited $\pi\pi^*$ singlets can be described as covalent *minus* and ionic *plus* states, which stem from their description within VB and alternancy symmetry theories. *Minus* states typically involve both single and double excitations, whereas *plus* states are dominated by single excitations (being $HOMO \rightarrow LUMO$ for the lowest state). Within this framework, we analyzed excited states of heterocycles with more scrutiny. The lowest $\pi\pi^*$ states in heterocycles exhibit inherent problems that may challenge the performance of many approximate electronic structure methods. The covalent A_1 state contains an important (but not necessarily large) double excitation character. The ionic B_2 state is dominated by one configuration and it may occasionally mix with Rydberg excitation (*e.g.* in pyrrole), but it remains unclear whether this is an artifact of the lack of correlation effects in approximate electronic structure methods. In terms of the method performance, the difference between ADC(1) and ADC(2) serves well to discriminate between the two low-lying excited states, which react differently to the inclusion of double excitations even at the lowest orders of perturbation theory. When compared to the CC3 reference, ADC(2) tends to slightly overestimate excitation energies, whereas ADC(3) somewhat underestimates them. Standard TDDFT within the adiabatic approximation appears to have an inherent difficulty in describing both A_1 and B_2 states in a balanced manner. Triplet states are somewhat less challenging compared to singlets, the reason being that both $\pi\pi^*$ triplet states (A_1 and B_2) have a covalent *minus* character with little double excitation character and no VR mixing. Nevertheless, standard TDDFT suffers from a triplet instability problem, which is largely solved by employing TDA.

Many properties of the low-lying $\pi\pi^*$ states can be deduced from their symmetry and

pseudosymmetry. Yet, perturbations to symmetric molecular structures should not significantly alter their electronic structure. However, for larger geometry distortions, one should examine the excited-state potential energy surfaces of these compounds, including the excited-state minima and the conical intersections, which should be the next step following this work. While a lot of work on excited-state dynamics beyond the Franck–Condon region has already been discussed for furan, pyrrole, and thiophene, future work should deliver more systematic evaluations of their potential energy surfaces with different electronic structure methods. Comparisons with the potential energy surfaces of *cis*-butadiene, benzene, and other related systems will also be beneficial. Only when problems related to the electronic structure description of these compounds are fully resolved can these molecules become true model systems for nonadiabatic dynamics simulations. Other properties, such as intersystem crossing, emission, and vibronic structure, or two-photon absorption cross-sections, also deserve a more stringent investigation and comparison among congener compounds. Future developments in electronic structure methods (both density-based and wavefunction-based) should revisit the issues highlighted in the present account, as excited states of heterocycles are expected to remain intriguing to both theorists and spectroscopists even in the forthcoming decades. We hope that, in the near future, FCI reference computations with large basis sets will be available, but the development of corresponding wavefunction analysis tools will also be vital.

Acknowledgement

The authors thank Prof Clemence Corminboeuf, Prof Basile Curchod, and Prof Kun-Han Lin for helpful discussions and valuable input during the early stages of the research. Prof Ewa Pastorczak and Dr Jack Taylor are acknowledged for their insightful suggestions regarding the manuscript. This research was performed using the Advanced computing service provided by University of Zagreb – University Computing Centre (SRCE).

Associated Content

Data and Software Availability

The data underlying this study are available in the Supporting Information. Calculations were carried out using commercial and open-source software packages referenced in the manuscript. No custom software was developed as part of this work.

Supporting Information

The Supporting Information is available free of charge at `https://pubs.acs.org/`.

Contents: Additional details on the experimental values, tables with raw data, supporting calculations using multireference methods and other density functionals, and molecular structures in XYZ format.

References

(1) Alvarez-Builla, J.; Vaquero, J. J.; Barluenga, J. *Modern Heterocyclic Chemistry, 4 Volume Set*; John Wiley & Sons, 2011; Vol. 2.

(2) Joule, J. A.; Mills, K. *Heterocyclic chemistry at a glance*; John Wiley & Sons, 2012.

(3) Palmer, M. H.; Walker, I. C.; Ballard, C. C.; Guest, M. F. The electronic states of furan studied by VUV absorption, near-threshold electron energy-loss spectroscopy and ab initio multi-reference configuration interaction calculations. *Chem. Phys.* **1995**, *192*, 111–125.

(4) Palmer, M. H.; Walker, I. C.; Guest, M. F. The electronic states of pyrrole studied by optical (VUV) absorption, near-threshold electron energy-loss (EEL) spectroscopy and ab initio multi-reference configuration interaction calculations. *Chem. Phys.* **1998**, *238*, 179–199.

(5) Palmer, M. H.; Walker, I. C.; Guest, M. F. The electronic states of thiophene studied by optical (VUV) absorption, near-threshold electron energy-loss (EEL) spectroscopy and ab initio multi-reference configuration interaction calculations. *Chem. Phys.* **1999**, *241*, 275–296.

(6) Serrano-Andres, L.; Merchan, M.; Nebot-Gil, I.; Roos, B. O.; Fulscher, M. Theoretical study of the electronic spectra of cyclopentadiene, pyrrole, and furan. *J. Am. Chem. Soc.* **1993**, *115*, 6184–6197.

(7) Pastore, M.; Angeli, C.; Cimiraglia, R. An application of second and third-order n-electron valence state perturbation theory to the calculation of the vertical electronic spectrum of furan. *Chem. Phys. Lett.* **2006**, *426*, 445–451.

(8) Stenrup, M.; Larson, Å. A computational study of radiationless deactivation mechanisms of furan. *Chem. Phys.* **2011**, *379*, 6–12.

(9) Christiansen, O.; Jørgensen, P. The electronic spectrum of furan. *J. Am. Chem. Soc.* **1998**, *120*, 3423–3430.

(10) Gromov, E. V.; Trofimov, A. B.; Vitkovskaya, N. M.; Schirmer, J.; Köppel, H. Theoretical study of the low-lying excited singlet states of furan. *J. Chem. Phys.* **2003**, *119*, 737–753.

(11) Burcl, R.; Amos, R. D.; Handy, N. C. Study of excited states of furan and pyrrole by time-dependent density functional theory. *Chem. Phys. Lett.* **2002**, *355*, 8–18.

(12) Crespo-Otero, R.; Barbatti, M. Spectrum simulation and decomposition with nuclear ensemble: formal derivation and application to benzene, furan and 2-phenylfuran. *Theor. Chem. Acc.* **2012**, *131*, 1237.

(13) Li, X.; Paldus, J. A multireference coupled-cluster study of electronic excitations in furan and pyrrole. *J. Phys. Chem. A* **2010**, *114*, 8591–8600.

(14) Wan, J.; Meller, J.; Hada, M.; Ehara, M.; Nakatsuji, H. Electronic excitation spectra of furan and pyrrole: Revisited by the symmetry adapted cluster–configuration interaction method. *J. Chem. Phys.* **2000**, *113*, 7853–7866.

(15) Holland, D. M. P.; Seddon, E. A.; Trofimov, A. B.; Gromov, E. V.; Wormit, M.; Dreuw, A.; Korona, T.; De Oliveira, N.; Archer, L. E.; Joyeux, D. A study of the excited electronic states of normal and fully deuterated furan by photoabsorption spectroscopy and high-level ab initio calculations. *J. Mol. Spectrosc.* **2015**, *315*, 184–195.

(16) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andres, L. Does density functional theory contribute to the understanding of excited states of unsaturated organic compounds? *Mol. Phys.* **1999**, *97*, 859–868.

(17) Nakano, H.; Tsuneda, T.; Hashimoto, T.; Hirao, K. Theoretical study of the excitation spectra of five-membered ring compounds: Cyclopentadiene, furan, and pyrrole. *J. Chem. Phys.* **1996**, *104*, 2312–2320.

(18) Trofimov, A. B.; Schirmer, J. Polarization propagator study of electronic excitation in key heterocyclic molecules II. Furan. *Chem. Phys.* **1997**, *224*, 175–190.

(19) Roos, B. O.; Malmqvist, P.-Å.; Molina, V.; Serrano-Andrés, L.; Merchán, M. Theoretical characterization of the lowest-energy absorption band of pyrrole. *J. Chem. Phys.* **2002**, *116*, 7526–7536.

(20) Neville, S. P.; Worth, G. A. A reinterpretation of the electronic spectrum of pyrrole: A quantum dynamics study. *J. Chem. Phys.* **2014**, *140*, 034317.

(21) Palmer, M. H.; Wilson, P. J. The singlet electronic states of pyrrole: a theoretical study by both ab initio multi-reference configuration interaction methods and time-dependent density functional theory and a reconsideration of the experimental VUV spectral data. *Mol. Phys.* **2003**, *101*, 2391–2408.

(22) Pastore, M.; Angeli, C.; Cimiraglia, R. The vertical electronic spectrum of pyrrole: A second and third order n-electron valence state perturbation theory study. *Chem. Phys. Lett.* **2006**, *422*, 522–528.

(23) Trofimov, A. B.; Schirmer, J. Polarization propagator study of electronic excitation in key heterocyclic molecules I. Pyrrole. *Chem. Phys.* **1997**, *214*, 153–170.

(24) Serrano-Andrés, L.; Merchán, M.; Fülscher, M.; Roos, B. O. A theoretical study of the electronic spectrum of thiophene. *Chem. Phys. Lett.* **1993**, *211*, 125–134.

(25) Stenrup, M. Theoretical study of the radiationless deactivation mechanisms of photo-excited thiophene. *Chem. Phys.* **2012**, *397*, 18–25.

(26) Pastore, M.; Angeli, C.; Cimiraglia, R. A multireference perturbation theory study on the vertical electronic spectrum of thiophene. *Theor. Chem. Acc.* **2007**, *118*, 35–46.

(27) Holland, D. M. P.; Trofimov, A. B.; Seddon, E. A.; Gromov, E. V.; Korona, T.; De Oliveira, N.; Archer, L. E.; Joyeux, D.; Nahon, L. Excited electronic states of thiophene: high resolution photoabsorption Fourier transform spectroscopy and ab initio calculations. *Phys. Chem. Chem. Phys.* **2014**, *16*, 21629–21644.

(28) Wan, J.; Hada, M.; Ehara, M.; Nakatsuji, H. Electronic excitation spectrum of thiophene studied by symmetry-adapted cluster configuration interaction method. *J. Chem. Phys.* **2001**, *114*, 842–850.

(29) Prlj, A.; Curchod, B. F. E.; Fabrizio, A.; Floryan, L.; Corminboeuf, C. Qualitatively incorrect features in the TDDFT spectrum of thiophene-based compounds. *J. Phys. Chem. Lett.* **2015**, *6*, 13–21.

(30) Salzmann, S.; Kleinschmidt, M.; Tatchen, J.; Weinkauf, R.; Marian, C. M. Excited states of thiophene: ring opening as deactivation mechanism. *Phys. Chem. Chem. Phys.* **2008**, *10*, 380–392.

(31) Weinkauf, R.; Lehr, L.; Schlag, E. W.; Salzmann, S.; Marian, C. M. Ultrafast dynamics in thiophene investigated by femtosecond pump probe photoelectron spectroscopy and theory. *Phys. Chem. Chem. Phys.* **2008**, *10*, 393–404.

(32) Wu, X.-F.; Zheng, X.; Wang, H.-G.; Zhao, Y.-Y.; Guan, X.; Phillips, D. L.; Chen, X.; Fang, W. A resonance Raman spectroscopic and CASSCF investigation of the Franck–Condon region structural dynamics and conical intersections of thiophene. *J. Chem. Phys.* **2010**, *133*, 134507.

(33) Dash, M.; Moroni, S.; Filippi, C.; Scemama, A. Tailoring CIPSI expansions for QMC calculations of electronic excitations: the case study of thiophene. *J. Chem. Theory Comput.* **2021**, *17*, 3426–3434.

(34) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* **2008**, *128*, 134110.

(35) Fuji, T.; Suzuki, Y.-I.; Horio, T.; Suzuki, T.; Mitrić, R.; Werner, U.; Bonačić-Koutecký, V. Ultrafast photodynamics of furan. *J. Chem. Phys.* **2010**, *133*, 234303.

(36) Oesterling, S.; Schalk, O.; Geng, T.; Thomas, R. D.; Hansson, T.; de Vivie-Riedle, R. Substituent effects on the relaxation dynamics of furan, furfural and β-furfural: a combined theoretical and experimental approach. *Phys. Chem. Chem. Phys.* **2017**, *19*, 2025–2035.

(37) Köppel, H.; Gromov, E. V.; Trofimov, A. B. Multi-mode–multi-state quantum dynamics of key five-membered heterocycles: spectroscopy and ultrafast internal conversion. *Chem. Phys.* **2004**, *304*, 35–49.

(38) Gromov, E. V.; Reddy, V. S.; Gatti, F.; Köppel, H. Reaction surface approach to multimode vibronic coupling problems: General framework and application to furan. *J. Chem. Phys.* **2013**, *139*, 234306.

(39) Díaz Mirón, G.; Lien-Medrano, C. R.; Banerjee, D.; Monti, M.; Aradi, B.; Sentef, M. A.; Niehaus, T. A.; Hassanali, A. Non-adiabatic couplings in surface hopping with tight binding density functional theory: The case of molecular motors. *J. Chem. Theory Comput.* **2024**, *20*, 10602–10614.

(40) Tiefenbacher, M. X.; Bachmair, B.; Chen, C. G.; Westermayr, J.; Marquetand, P.; Dietschreit, J. C. B.; González, L. Excited-state nonadiabatic dynamics in explicit solvent using machine learned interatomic potentials. *Digit. Discov.* **2025**, *4*, 1478–1491.

(41) Fazzi, D.; Barbatti, M.; Thiel, W. Modeling ultrafast exciton deactivation in oligothiophenes via nonadiabatic dynamics. *Phys. Chem. Chem. Phys.* **2015**, *17*, 7787–7799.

(42) Prlj, A.; Curchod, B. F. E.; Corminboeuf, C. Excited state dynamics of thiophene and bithiophene: new insights into theoretically challenging systems. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14719–14730.

(43) Schnappinger, T.; Kölle, P.; Marazzi, M.; Monari, A.; González, L.; de Vivie-Riedle, R. Ab initio molecular dynamics of thiophene: The interplay of internal conversion and intersystem crossing. *Phys. Chem. Chem. Phys.* **2017**, *19*, 25662–25670.

(44) Cui, G.; Fang, W. Ab Initio trajectory surface-hopping study on ultrafast deactivation process of thiophene. *J. Phys. Chem. A* **2011**, *115*, 11544–11550.

(45) Pederzoli, M.; Pittner, J. A new approach to molecular dynamics with non-adiabatic and spin-orbit effects with applications to QM/MM simulations of thiophene and selenophene. *J. Chem. Phys.* **2017**, *146*, 114101.

(46) Suchan, J.; Janoš, J.; Slavicek, P. Pragmatic Approach to Photodynamics: Mixed Landau-Zener Surface Hopping with Inter-system Crossing. *J. Chem. Theory Comput.* **2020**, *16*, 5809–5820.

(47) Heller, E. R.; Joswig, J.-O.; Seifert, G. Exploring the effects of quantum decoherence on the excited-state dynamics of molecular systems. *Theor. Chem. Acc.* **2021**, *140*, 42.

(48) Parkes, M. A.; Worth, G. A. The "simple" photochemistry of thiophene. *J. Chem. Phys.* **2024**, *161*, 114305.

(49) Vazdar, M.; Eckert-Maksić, M.; Barbatti, M.; Lischka, H. Excited-state non-adiabatic dynamics simulations of pyrrole. *Mol. Phys.* **2009**, *107*, 845–854.

(50) Barbatti, M.; Pittner, J.; Pederzoli, M.; Werner, U.; Mitrić, R.; Bonačić-Kouteckỳ, V.; Lischka, H. Non-adiabatic dynamics of pyrrole: Dependence of deactivation mechanisms on the excitation energy. *Chem. Phys.* **2010**, *375*, 26–34.

(51) Barbatti, M.; Sen, K. Effects of different initial condition samplings on photodynamics and spectrum of pyrrole. *Int. J. Quantum Chem.* **2016**, *116*, 762–771.

(52) Faraji, S.; Vazdar, M.; Reddy, V. S.; Eckert-Maksic, M.; Lischka, H.; Koeppel, H. Ab initio quantum dynamical study of the multi-state nonadiabatic photodissociation of pyrrole. *J. Chem. Phys.* **2011**, *135*, 154310.

(53) Makhov, D. V.; Saita, K.; Martinez, T. J.; Shalashilin, D. V. Ab initio multiple cloning simulations of pyrrole photodissociation: TKER spectra and velocity map imaging. *Phys. Chem. Chem. Phys.* **2015**, *17*, 3316–3325.

(54) Saita, K.; Nix, M. G.; Shalashilin, D. V. Simulation of ultrafast photodynamics of pyrrole with a multiconfigurational Ehrenfest method. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16227–16235.

(55) Lan, Z.; Dupays, A.; Vallet, V.; Mahapatra, S.; Domcke, W. Photoinduced multimode quantum dynamics of pyrrole at the $^1\pi\sigma^*$–S_0 conical intersections. *J. Photochem. Photobiol. A* **2007**, *190*, 177–189.

(56) Vallet, V.; Lan, Z.; Mahapatra, S.; Sobolewski, A. L.; Domcke, W. Photochemistry of pyrrole: Time-dependent quantum wave-packet description of the dynamics at the $^1\pi\sigma^*$–S_0 conical intersections. *J. Chem. Phys.* **2005**, *123*, 144307.

(57) Wu, G.; Neville, S. P.; Schalk, O.; Sekikawa, T.; Ashfold, M. N.; Worth, G. A.; Stolow, A. Excited state non-adiabatic dynamics of pyrrole: A time-resolved photoelectron spectroscopy and quantum dynamics study. *J. Chem. Phys.* **2015**, *142*, 074302.

(58) Sapunar, M.; Ponzi, A.; Chaiwongwattana, S.; Mališ, M.; Prlj, A.; Decleva, P.; Došlić, N. Timescales of N–H bond dissociation in pyrrole: a nonadiabatic dynamics study. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19012–19020.

(59) Heindl, M.; González, L. A XMS-CASPT2 non-adiabatic dynamics study on pyrrole. *Comput. Theor. Chem.* **2019**, *1155*, 38–46.

(60) Lan, Z.; Domcke, W. Role of vibrational energy relaxation in the photoinduced nonadiabatic dynamics of pyrrole at the $^1\pi\sigma^*$–S_0 conical intersection. *Chem. Phys.* **2008**, *350*, 125–138.

(61) Frank, I.; Damianos, K. Restricted open-shell Kohn-Sham theory: Simulation of the pyrrole photodissociation. *J. Chem. Phys.* **2007**, *126*, 125105.

(62) Gudem, M.; Kowalewski, M. Controlling the photostability of pyrrole with optical nanocavities. *J. Phys. Chem. A* **2021**, *125*, 1142–1151.

(63) Makhov, D. V.; Adeyemi, S.; Cowperthwaite, M.; Shalashilin, D. V. Simulation of the dynamics of vibrationally mediated photodissociation for deuterated pyrrole. *J. Phys. Commun.* **2022**, *6*, 025001.

(64) Mathey, F. The organic chemistry of phospholes. *Chem. Rev.* **1988**, *88*, 429–453.

(65) Imoto, H.; Naka, K. The Dawn of Functional Organoarsenic Chemistry. *Chem. Eur. J.* **2019**, *25*, 1883–1894.

(66) Janosik, T.; Bergman, J. *Progress in Heterocyclic Chemistry*; Elsevier, 2009; Vol. 21; pp 115–144.

(67) Nordén, B.; Håkansson, R.; Pedersen, P. B.; Thulstrup, E. W. The magnetic circular dichroism of five-membered ring heterocycles. *Chem. Phys.* **1978**, *33*, 355–366.

(68) Epiotis, N. D.; Cherry, W. On the aromaticity of phospholes and arsoles. *J. Am. Chem. Soc.* **1976**, *98*, 4365–4370.

(69) Dransfeld, A.; Nyulászi, L.; Schleyer, P. v. R. The aromaticity of polyphosphaphospholes decreases with the pyramidality of the tricoordinate phosphorus. *Inorg. Chem.* **1998**, *37*, 4413–4420.

(70) Chesnut, D. B.; Quin, L. D. The important role of the phosphorus lone pair in phosphole aromaticity. *Heteroat. Chem.* **2007**, *18*, 754–758.

(71) Mulliken, R. S. Report on notation for the spectra of polyatomic molecules. *J. Chem. Phys.* **1955**, *23*, 1997–2011.

(72) Richard, R. M.; Herbert, J. M. Time-dependent density-functional description of the ^1L_a state in polycyclic aromatic hydrocarbons: Charge-transfer character in disguise? *J. Chem. Theory Comput.* **2011**, *7*, 1296–1306.

(73) Kuritz, N.; Stein, T.; Baer, R.; Kronik, L. Charge-transfer-like π-π^* excitations in time-dependent density functional theory: A conundrum and its solution. *J. Chem. Theory Comput.* **2011**, *7*, 2408–2415.

(74) Wong, B. M.; Hsieh, T. H. Optoelectronic and excitonic properties of oligoacenes: Substantial improvements from range-separated time-dependent density functional theory. *J. Chem. Theory Comput.* **2010**, *6*, 3704–3712.

(75) Goerigk, L.; Grimme, S. Double-hybrid density functionals provide a balanced description of excited ^1L_a and ^1L_b states in polycyclic aromatic hydrocarbons. *J. Chem. Theory Comput.* **2011**, *7*, 3272–3277.

(76) Lopata, K.; Reslan, R.; Kowalska, M.; Neuhauser, D.; Govind, N.; Kowalski, K. Excited-state studies of polyacenes: a comparative picture using EOMCCSD, CR-EOMCCSD(T), range-separated (LR/RT)-TDDFT, TD-PM3, and TD-ZINDO. *J. Chem. Theory Comput.* **2011**, *7*, 3686–3693.

(77) Krykunov, M.; Grimme, S.; Ziegler, T. Accurate theoretical description of the ^1L_a and ^1L_b excited states in acenes using the all order constricted variational density functional theory method and the local density approximation. *J. Chem. Theory Comput.* **2012**, *8*, 4434–4440.

(78) Senn, F.; Krykunov, M. Excited State Studies of Polyacenes Using the All-Order Constricted Variational Density Functional Theory with Orbital Relaxation. *J. Phys. Chem. A* **2015**, *119*, 10575–10581.

(79) Moore, B.; Sun, H.; Govind, N.; Kowalski, K.; Autschbach, J. Charge-Transfer Versus Charge-Transfer-Like Excitations Revisited. *J. Chem. Theory Comput.* **2015**, *11*, 3305–3320.

(80) Mewes, S. A.; Plasser, F.; Dreuw, A. Communication: Exciton analysis in time-dependent density functional theory: How functionals shape excited-state characters. *J. Chem. Phys.* **2015**, *143*, 171101.

(81) Peng, W.-T.; Chai, J.-D. Assessment of asymptotically corrected model potentials for charge-transfer-like excitations in oligoacenes. *Phys. Chem. Chem. Phys.* **2014**, *16*, 21564–21569.

(82) Grimme, S.; Parac, M. Substantial errors from time-dependent density functional

theory for the calculation of excited states of large π systems. *ChemPhysChem* **2003**, *4*, 292–295.

(83) Prlj, A.; Sandoval-Salinas, M. E.; Casanova, D.; Jacquemin, D.; Corminboeuf, C. Low-lying $\pi\pi^*$ states of heteroaromatic molecules: a challenge for excited state methods. *J. Chem. Theory Comput.* **2016**, *12*, 2652–2660.

(84) Novak, J.; Prlj, A.; Basarić, N.; Corminboeuf, C.; Došlić, N. Photochemistry of 1-and 2-Naphthols and Their Water Clusters: The Role of $^1\pi\pi^*(L_a)$ Mediated Hydrogen Transfer to Carbon Atoms. *Chem. Eur. J.* **2017**, *23*, 8244–8251.

(85) Brisker-Klaiman, D.; Dreuw, A. Explaining level inversion of the L_a and L_b states of indole and indole derivatives in polar solvents. *ChemPhysChem* **2015**, *16*, 1695–1702.

(86) Arulmozhiraja, S.; Coote, M. L. ^1L_a and ^1L_b states of indole and azaindole: is density functional theory inadequate? *J. Chem. Theory Comput.* **2012**, *8*, 575–584.

(87) Prlj, A.; Begušić, T.; Zhang, Z. T.; Fish, G. C.; Wehrle, M.; Zimmermann, T.; Choi, S.; Roulet, J.; Moser, J.-E.; Vaníček, J. Semiclassical Approach to Photophysics Beyond Kasha's Rule and Vibronic Spectroscopy Beyond the Condon Approximation. The Case of Azulene. *J. Chem. Theory Comput.* **2020**, *16*, 2617–2626.

(88) Santoro, F.; Improta, R.; Fahleson, T.; Kauczor, J.; Norman, P.; Coriani, S. Relative stability of the L_a and L_b excited states in adenine and guanine: Direct evidence from TD-DFT calculations of MCD spectra. *J. Phys. Chem. Lett.* **2014**, *5*, 1806–1811.

(89) Nolting, D.; Marian, C.; Weinkauf, R. Protonation effect on the electronic spectrum of tryptophan in the gas phase. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2633–2640.

(90) Hashimoto, T.; Choe, Y.-K.; Nakano, H.; Hirao, K. Theoretical study of the Q and B bands of free-base, magnesium, and zinc porphyrins, and their derivatives. *J. Phys. Chem. A* **1999**, *103*, 1894–1904.

(91) Hoffmann, M.; Mewes, S. A.; Wieland, S.; Popp, C.; Dreuw, A. Electron–hole correlation as unambiguous and universal classification for the nature of low-lying $\pi\pi^*$ states of nitrogen heterocycles. *J. Phys. Chem. Lett.* **2019**, *10*, 6112–6117.

(92) Michl, J.; Bonacic-Koutecky, V. *Electronic aspects of organic photochemistry*; Wiley, 1990.

(93) Schulten, K.; Karplus, M. On the origin of a low-lying forbidden transition in polyenes and related molecules. *Chem. Phys. Lett.* **1972**, *14*, 305–309.

(94) Nakayama, K.; Nakano, H.; Hirao, K. Theoretical study of the $\pi \rightarrow \pi^*$ excited states of linear polyenes: The energy gap between $1^1B_u^+$ and $2^1A_g^-$ states and their character. *Int. J. Quantum Chem.* **1998**, *66*, 157–175.

(95) Hudson, B. S.; Kohler, B. E.; Schulten, K. *Excited states*; Elsevier, 1982; Vol. 6; pp 1–95.

(96) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Lindh, R.; Roos, B. O. Towards an accurate molecular orbital theory for excited states: Ethene, butadiene, and hexatriene. *J. Chem. Phys.* **1993**, *98*, 3151–3162.

(97) Orlandi, G.; Zerbetto, F.; Zgierski, M. Z. Theoretical analysis of spectra of short polyenes. *Chem. Rev.* **1991**, *91*, 867–891.

(98) Cave, R. J.; Davidson, E. R. A theoretical investigation of some low-lying singlet states of 1,3-butadiene. *J. Phys. Chem.* **1987**, *91*, 4481–4490.

(99) Cave, R. J. Size-inconsistency effects in molecular properties for states with valence-Rydberg mixing: The low-lying $\pi \rightarrow \pi^*$ states of ethylene and butadiene. *J. Chem. Phys.* **1990**, *92*, 2450–2456.

(100) Lappe, J.; Cave, R. J. On the Vertical and Adiabatic Excitation Energies of the $2A_g$ State of trans-1,3-Butadiene. *J. Phys. Chem. A* **2000**, *104*, 2294–2300.

(101) Kitao, O.; Nakatsuji, H. Cluster expansion of the wavefunction: Valence and Rydberg excitations of trans-and cis-butadiene. *Chem. Phys. Lett.* **1988**, *143*, 528–534.

(102) Saha, B.; Ehara, M.; Nakatsuji, H. Singly and doubly excited states of butadiene, acrolein, and glyoxal: Geometries and electronic spectra. *J. Chem. Phys.* **2006**, *125*, 014316.

(103) Schmidt, M.; Tavan, P. Electronic excitations in long polyenes revisited. *J. Chem. Phys.* **2012**, *136*, 124309.

(104) Dallos, M.; Lischka, H. A systematic theoretical investigation of the lowest valence- and Rydberg-excited singlet states of trans-butadiene. The character of the 1^1B_u (V) state revisited. *Theor. Chem. Acc.* **2004**, *112*, 16–26.

(105) Tavan, P.; Schulten, K. The low-lying electronic excitations in long polyenes: A PPP-MRD-CI study. *J. Chem. Phys.* **1986**, *85*, 6602–6609.

(106) Dokukina, I.; Marian, C. M.; Weingart, O. New Perspectives on an Old Issue: A Comparative MS-CASPT2 and OM2-MRCI Study of Polyenes and Protonated Schiff Bases. *Photochem. Photobiol.* **2017**, *93*, 1345–1355.

(107) Glover, W. J.; Mori, T.; Schuurman, M. S.; Boguslavskiy, A. E.; Schalk, O.; Stolow, A.; Martínez, T. J. Excited state non-adiabatic dynamics of the smallest polyene, trans 1,3-butadiene. II. Ab initio multiple spawning simulations. *J. Chem. Phys.* **2018**, *148*, 164303.

(108) Shu, Y.; Truhlar, D. G. Doubly excited character or static correlation of the reference state in the controversial 2^1A_g state of trans-butadiene? *J. Am. Chem. Soc.* **2017**, *139*, 13770–13778.

(109) Dong, S. S.; Gagliardi, L.; Truhlar, D. G. Nature of the 1^1B_u and 2^1A_g excited states

of butadiene and the Goldilocks principle of basis set diffuseness. *J. Chem. Theory Comput.* **2019**, *15*, 4591–4601.

(110) Kossoski, F.; Boggio-Pasqua, M.; Loos, P.-F.; Jacquemin, D. Reference energies for double excitations: Improvement and extension. *J. Chem. Theory Comput.* **2024**, *20*, 5655–5678.

(111) Baerends, E. J.; Amati, M.; Stoia, S. Improved TDDFT Excitation Energies with an Accurate Kohn–Sham Potential: Reassessment of the Double-Excitation Character of the Low Lying 2^1A_g Excited State of s-Trans-1,3-Butadiene. *J. Chem. Theory Comput.* **2025**, *21*, 12755–12772.

(112) Daday, C.; Smart, S.; Booth, G. H.; Alavi, A.; Filippi, C. Full configuration interaction excitations of ethene and butadiene: Resolution of an ancient question. *J. Chem. Theory Comput.* **2012**, *8*, 4441–4451.

(113) Watson, M. A.; Chan, G. K.-L. Excited states of butadiene to chemical accuracy: Reconciling theory and experiment. *J. Chem. Theory Comput.* **2012**, *8*, 4013–4018.

(114) Chien, A. D.; Holmes, A. A.; Otten, M.; Umrigar, C. J.; Sharma, S.; Zimmerman, P. M. Excited states of methylene, polyenes, and ozone from heat-bath configuration interaction. *J. Phys. Chem. A* **2018**, *122*, 2714–2722.

(115) Hsu, C.-P.; Hirata, S.; Head-Gordon, M. Excitation energies from time-dependent density functional theory for linear polyene oligomers: butadiene to decapentaene. *J. Phys. Chem. A* **2001**, *105*, 451–458.

(116) Ullrich, C. A. *Time-dependent density-functional theory: concepts and applications*; Oxford University Press, 2011.

(117) Starcke, J. H.; Wormit, M.; Schirmer, J.; Dreuw, A. How much double excitation

character do the lowest excited states of linear polyenes have? *Chem. Phys.* **2006**, *329*, 39–49.

(118) Cave, R. J.; Zhang, F.; Maitra, N. T.; Burke, K. A dressed TDDFT treatment of the 2^1A_g states of butadiene and hexatriene. *Chem. Phys. Lett.* **2004**, *389*, 39–42.

(119) Dar, D. B.; Maitra, N. T. Capturing the elusive curve-crossing in low-lying states of butadiene with dressed TDDFT. *J. Phys. Chem. Lett.* **2025**, *16*, 703–709.

(120) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2015**, *5*, 82–95.

(121) Loos, P.-F.; Jacquemin, D. Is ADC(3) as Accurate as CC3 for Valence and Rydberg Transition Energies? *J. Phys. Chem. Lett.* **2020**, *11*, 974–980.

(122) Koch, H.; Christiansen, O.; Jørgensen, P.; Sanchez de Merás, A. M.; Helgaker, T. The CC3 model: An iterative coupled cluster approach including connected triples. *J. Chem. Phys.* **1997**, *106*, 1808–1818.

(123) Loos, P.-F.; Lipparini, F.; Boggio-Pasqua, M.; Scemama, A.; Jacquemin, D. A Mountaineering Strategy to Excited States: Highly Accurate Energies and Benchmarks for Medium Sized Molecules. *J. Chem. Theory Comput.* **2020**, *16*, 1711–1741.

(124) Platt, J. R. Classification of spectra of cata-condensed hydrocarbons. *J. Chem. Phys.* **1949**, *17*, 484–495.

(125) Tavan, P.; Schulten, K. The "phantom" photochemical singlet state of stilbene and its diphenylpolyene relative. *Chem. Phys. Lett.* **1978**, *56*, 200–204.

(126) Guidez, E. B.; Aikens, C. M. Origin and TDDFT benchmarking of the plasmon resonance in acenes. *J. Phys. Chem. C* **2013**, *117*, 21466–21475.

(127) Klessinger, M.; Michl, J. *Excited states and photochemistry of organic molecules*; Wiley-VCH, 1995.

(128) Pariser, R. Theory of the electronic spectra and structure of the polyacenes and of alternant hydrocarbons. *J. Chem. Phys.* **1956**, *24*, 250–268.

(129) Hückel, E. Quantentheoretische beiträge zum benzolproblem. *Z. Phys.* **1931**, *70*, 204–286.

(130) Pariser, R.; Parr, R. G. A Semi-Empirical Theory of the Electronic Spectra and Electronic Structure of Complex Unsaturated Molecules. I. *J. Chem. Phys.* **1953**, *21*, 466–471.

(131) Pariser, R.; Parr, R. G. A semi-empirical theory of the electronic spectra and electronic structure of complex unsaturated molecules. II. *J. Chem. Phys.* **1953**, *21*, 767–776.

(132) Pople, J. A. Electron interaction in unsaturated hydrocarbons. *Trans. Faraday Soc.* **1953**, *49*, 1375–1385.

(133) Hashimoto, T.; Nakano, H.; Hirao, K. Theoretical study of the valence $\pi \rightarrow \pi^*$ excited states of polyacenes: Benzene and naphthalene. *J. Chem. Phys.* **1996**, *104*, 6244–6258.

(134) Kawashima, Y.; Hashimoto, T.; Nakano, H.; Hirao, K. Theoretical study of the valence $\pi \rightarrow \pi^*$ excited states of polyacenes: anthracene and naphthacene. *Theor. Chem. Acc.* **1999**, *102*, 49–64.

(135) Hirao, K.; Nakano, H.; Nakayama, K.; Dupuis, M. A complete active space valence bond (CASVB) method. *J. Chem. Phys.* **1996**, *105*, 9227–9239.

(136) Hirao, K.; Nakano, H.; Nakayama, K. A complete active space valence bond method with nonorthogonal orbitals. *J. Chem. Phys.* **1997**, *107*, 9966–9974.

(137) Plasser, F. Visualisation of Electronic Excited-State Correlation in Real Space. *ChemPhotoChem* **2019**, *3*, 702–706.

(138) Manjavacas, A.; Marchesin, F.; Thongrattanasiri, S.; Koval, P.; Nordlander, P.; Sanchez-Portal, D.; Garcia de Abajo, F. J. Tunable molecular plasmons in polycyclic aromatic hydrocarbons. *ACS Nano* **2013**, *7*, 3635–3643.

(139) Lauchner, A.; Schlather, A. E.; Manjavacas, A.; Cui, Y.; McClain, M. J.; Stec, G. J.; Garcia de Abajo, F. J.; Nordlander, P.; Halas, N. J. Molecular plasmonics. *Nano Lett.* **2015**, *15*, 6208–6214.

(140) Krauter, C. M.; Bernadotte, S.; Jacob, C. R.; Pernpointner, M.; Dreuw, A. Identification of plasmons in molecules with scaled ab initio approaches. *J. Phys. Chem. C* **2015**, *119*, 24564–24573.

(141) Bernadotte, S.; Evers, F.; Jacob, C. R. Plasmons in molecules. *J. Phys. Chem. C* **2013**, *117*, 1863–1878.

(142) Kimber, P.; Plasser, F. Toward an understanding of electronic excitation energies beyond the molecular orbital picture. *Phys. Chem. Chem. Phys.* **2020**, *22*, 6058–6080.

(143) Schulten, K.; Dinur, U.; Honig, B. The spectra of carbonium ions, cyanine dyes, and protonated Schiff base polyenes. *J. Chem. Phys.* **1980**, *73*, 3927–3935.

(144) Serrano-Andrés, L.; Roos, B. O.; Merchán, M. Theoretical study of the electronic spectra of cis-1,3,5-hexatriene and cis-1,3-butadiene. *Theor. Chim. Acta* **1994**, *87*, 387–402.

(145) Feller, D.; Craig, N. C. High level ab initio energies and structures for the rotamers of 1,3-butadiene. *J. Phys. Chem. A* **2009**, *113*, 1601–1607.

(146) Santolini, V.; Malhado, J. P.; Robb, M. A.; Garavelli, M.; Bearpark, M. J. Photochemical reaction paths of cis-dienes studied with RASSCF: the changing balance between ionic and covalent excited states. *Mol. Phys.* **2015**, *113*, 1978–1990.

(147) Boggio-Pasqua, M.; Bearpark, M. J.; Klene, M.; Robb, M. A. A computational strategy for geometry optimization of ionic and covalent excited states, applied to butadiene and hexatriene. *J. Chem. Phys.* **2004**, *120*, 7849–7860.

(148) Wiberg, K. B.; Rosenberg, R. E. Butadiene. 1. A normal coordinate analysis and infrared intensities. Structure of the second rotamer. *J. Am. Chem. Soc.* **1990**, *112*, 1509–1519.

(149) Guo, H.; Karplus, M. Ab initio studies of polyenes. I. 1,3-butadiene. *J. Chem. Phys.* **1991**, *94*, 3679–3699.

(150) Nguyen, K. A.; Gordon, M. S. Isomerization of bicyclo[1.1.0]butane to butadiene. *J. Am. Chem. Soc.* **1995**, *117*, 3835–3847.

(151) Epifanovsky, E.; Gilbert, A. T. B.; Feng, X.; Lee, J.; Mao, Y.; Mardirossian, N.; Pokhilko, P.; White, A. F.; Coons, M. P.; Dempwolff, A. L.; Gan, Z.; Hait, D.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Kussmann, J.; Lange, A. W.; Lao, K. U.; Levine, D. S.; Liu, J.; McKenzie, S. C.; Morrison, A. F.; Nanda, K. D.; Plasser, F.; Rehn, D. R.; Vidal, M. L.; You, Z.-Q.; Zhu, Y.; Alam, B.; Albrecht, B. J.; Aldossary, A.; Alguire, E.; Andersen, J. H.; Athavale, V.; Barton, D.; Begam, K.; Behn, A.; Bellonzi, N.; Bernard, Y. A.; Berquist, E. J.; Burton, H. G. A.; Carreras, A.; Carter-Fenk, K.; Chakraborty, R.; Chien, A. D.; Closser, K. D.; Cofer-Shabica, V.; Dasgupta, S.; de Wergifosse, M.; Deng, J.; Diedenhofen, M.; Do, H.; Ehlert, S.; Fang, P.-T.; Fatehi, S.; Feng, Q.; Friedhoff, T.; Gayvert, J.; Ge, Q.; Gidofalvi, G.; Goldey, M.; Gomes, J.; González-Espinoza, C. E.; Gulania, S.; Gunina, A. O.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A.; Herbst, M. F.; Hernández Vera, M.; Hodecker, M.; Holden, Z. C.; Houck, S.; Huang, X.; Hui, K.; Huynh, B. C.; Ivanov, M.; Jász, Á.; Ji, H.; Jiang, H.; Kaduk, B.; Kähler, S.; Khistyaev, K.; Kim, J.; Kis, G.; Klunzinger, P.; Koczor-Benda, Z.; Koh, J. H.; Kosenkov, D.; Koulias, L.; Kowalczyk, T.; Krauter, C. M.; Kue, K.; Kunitsa, A.; Kus, T.; Ladjánszki, I.; Lan-

dau, A.; Lawler, K. V.; Lefrancois, D.; Lehtola, S.; Li, R. R.; Li, Y.-P.; Liang, J.; Liebenthal, M.; Lin, H.-H.; Lin, Y.-S.; Liu, F.; Liu, K.-Y.; Loipersberger, M.; Luenser, A.; Manjanath, A.; Manohar, P.; Mansoor, E.; Manzer, S. F.; Mao, S.-P.; Marenich, A. V.; Markovich, T.; Mason, S.; Maurer, S. A.; McLaughlin, P. F.; Menger, M. F. S. J.; Mewes, J.-M.; Mewes, S. A.; Morgante, P.; Mullinax, J. W.; Oosterbaan, K. J.; Paran, G.; Paul, A. C.; Paul, S. K.; Pavošević, F.; Pei, Z.; Prager, S.; Proynov, E. I.; Rák, Á.; Ramos-Cordoba, E.; Rana, B.; Rask, A. E.; Rettig, A.; Richard, R. M.; Rob, F.; Rossomme, E.; Scheele, T.; Scheurer, M.; Schneider, M.; Sergueev, N.; Sharada, S. M.; Skomorowski, W.; Small, D. W.; Stein, C. J.; Su, Y.-C.; Sundstrom, E. J.; Tao, Z.; Thirman, J.; Tornai, G. J.; Tsuchimochi, T.; Tubman, N. M.; Veccham, S. P.; Vydrov, O.; Wenzel, J.; Witte, J.; Yamada, A.; Yao, K.; Yeganeh, S.; Yost, S. R.; Zech, A.; Zhang, I. Y.; Zhang, X.; Zhang, Y.; Zuev, D.; Aspuru-Guzik, A.; Bell, A. T.; Besley, N. A.; Bravaya, K. B.; Brooks, B. R.; Casanova, D.; Chai, J.-D.; Coriani, S.; Cramer, C. J.; Cserey, G.; DePrince, I., A. Eugene; DiStasio, J., Robert A.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Goddard, I., William A.; Hammes-Schiffer, S.; Head-Gordon, T.; Hehre, W. J.; Hsu, C.-P.; Jagau, T.-C.; Jung, Y.; Klamt, A.; Kong, J.; Lambrecht, D. S.; Liang, W.; Mayhall, N. J.; McCurdy, C. W.; Neaton, J. B.; Ochsenfeld, C.; Parkhill, J. A.; Peverati, R.; Rassolov, V. A.; Shao, Y.; Slipchenko, L. V.; Stauch, T.; Steele, R. P.; Subotnik, J. E.; Thom, A. J. W.; Tkatchenko, A.; Truhlar, D. G.; Van Voorhis, T.; Wesolowski, T. A.; Whaley, K. B.; Woodcock, I., H. Lee; Zimmerman, P. M.; Faraji, S.; Gill, P. M. W.; Head-Gordon, M.; Herbert, J. M.; Krylov, A. I. Software for the frontiers of quantum chemistry: An overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **2021**, *155*, 084801.

(152) Plasser, F.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. I. Formalism. *J. Chem. Phys.* **2014**, *141*, 024106.

(153) Plasser, F.; Bäppler, S. A.; Wormit, M.; Dreuw, A. New tools for the systematic analysis and visualization of electronic excitations. II. Applications. *J. Chem. Phys.* **2014**, *141*, 024107.

(154) Folkestad, S. D.; Kjønstad, E. F.; Myhre, R. H.; Andersen, J. H.; Balbi, A.; Coriani, S.; Giovannini, T.; Goletto, L.; Haugland, T. S.; Hutcheson, A.; others eT 1.0: An open source electronic structure program with emphasis on coupled cluster and multilevel methods. *J. Chem. Phys.* **2020**, *152*, 184103.

(155) Baerends, E. J.; Gritsenko, O. V.; Van Meer, R. The Kohn–Sham gap, the fundamental gap and the optical gap: the physical meaning of occupied and virtual Kohn–Sham orbital energies. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16408–16425.

(156) Krylov, A. I. From orbitals to observables and back. *J. Chem. Phys.* **2020**, *153*, 080901.

(157) Bäppler, S. A.; Plasser, F.; Wormit, M.; Dreuw, A. Exciton analysis of many-body wave functions: Bridging the gap between the quasiparticle and molecular orbital pictures. *Phys. Rev. A* **2014**, *90*, 052521.

(158) Schulten, K.; Ohmine, I.; Karplus, M. Correlation effects in the spectra of polyenes. *J. Chem. Phys.* **1976**, *64*, 4422–4441.

(159) Li Manni, G.; Smart, S. D.; Alavi, A. Combining the complete active space self-consistent field method and the full configuration interaction quantum Monte Carlo within a super-CI framework, with application to challenging metal-porphyrins. *J. Chem. Theory Comput.* **2016**, *12*, 1245–1258.

(160) Angeli, C. On the nature of the $\pi \rightarrow \pi^*$ ionic excited states: The V state of ethene as a prototype. *J. Comput. Chem.* **2009**, *30*, 1319–1333.

(161) Malrieu, J.-P.; Guihéry, N.; Calzado, C. J.; Angeli, C. Bond electron pair: Its relevance

and analysis from the quantum chemistry point of view. *J. Comput. Chem.* **2007**, *28*, 35–50.

(162) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chem. Phys. Lett.* **1998**, *288*, 299–306.

(163) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098.

(164) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785.

(165) Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(166) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(167) Peach, M. J. G.; Williamson, M. J.; Tozer, D. J. Influence of triplet instabilities in TDDFT. *J. Chem. Theory Comput.* **2011**, *7*, 3578–3585.

(168) Lee, S.; Filatov, M.; Lee, S.; Choi, C. H. Eliminating spin-contamination of spin-flip time dependent density functional theory within linear response formalism by the use of zeroth-order mixed-reference (MR) reduced density matrix. *J. Chem. Phys.* **2018**, *149*, 104101.

(169) Horbatenko, Y.; Lee, S.; Filatov, M.; Choi, C. H. How beneficial is the explicit account of doubly-excited configurations in linear response theory? *J. Chem. Theory Comput.* **2021**, *17*, 975–984.

(170) Hirata, S.; Head-Gordon, M. Time-dependent density functional theory within the Tamm–Dancoff approximation. *Chem. Phys. Lett.* **1999**, *314*, 291–299.

(171) Reisler, H.; Krylov, A. I. Interacting Rydberg and valence states in radicals and molecules: experimental and theoretical studies. *Int. Rev. Phys. Chem.* **2009**, *28*, 267–308.

(172) Rubio, M.; Serrano-Andrés, L.; Merchán, M. Excited states of the water molecule: Analysis of the valence and Rydberg character. *J. Chem. Phys.* **2008**, *128*, 104305.

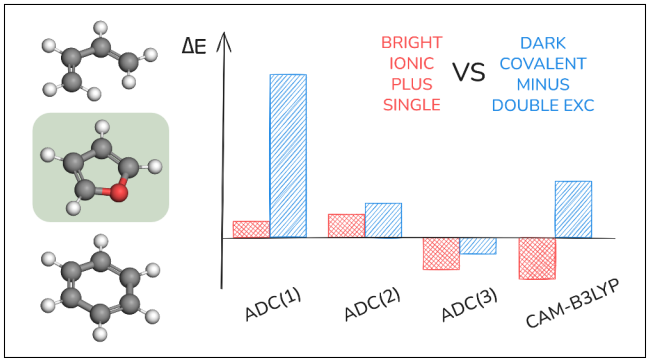
(173) Mulliken, R. S. Rydberg states and Rydbergization. *Acc. Chem. Res.* **1976**, *9*, 7–12.

(174) Lindh, R.; Roos, B. O. A theoretical study of the diffuseness of the V ($^1B_{1u}$) state of planar ethylene. *Int. J. Quantum Chem.* **1989**, *35*, 813–825.

(175) Mewes, S. A.; Plasser, F.; Krylov, A.; Dreuw, A. Benchmarking excited-state calculations using exciton properties. *J. Chem. Theory Comput.* **2018**, *14*, 710–725.

(176) Levine, B. G.; Ko, C.; Quenneville, J.; Martínez, T. J. Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.* **2006**, *104*, 1039–1051.

TOC Graphic


ΔE
BRIGHT
IONIC
PLUS
SINGLE
VS
DARK
COVALENT
MINUS
DOUBLE EXC
ADC(1)
ADC(2)
ADC(3)
CAM-B3LYP