THESIS
submitted to the

UNIVERSITY OF BORDEAUX
DOCTORAL SCHOOL OF CHEMICAL SCIENCES

by
Carmelo NAIM

in fulfillment for the Degree of
DOCTOR OF PHILOSOPHY

in PHYSICAL CHEMISTRY

DETERMINATION OF NONLINEAR OPTICAL PROPERTIES WITH QUANTUM CHEMISTRY: FROM BENCHMARKS TO EXPERIMENTAL SYSTEMS

Thesis Directors: Prof. Frédéric CASTET and Dr. Eduard MATITO

Defense date: 02 of December, 2022

Review committee members:

Prof.	**Juan-Carlos SANCHO-GARCIA**	University of Alicante	President
Prof.	**Denis JACQUEMIN**	University of Nantes	Reviewer
Dr.	**Marc de WERGIFOSSE**	Catholic University of Louvain	Reviewer
Prof.	**Jesus UGALDE**	University of the Basque Country	Examiner
Dr.	**Lionel TRUFLANDIER**	University of Bordeaux	Examiner

Abstract

This thesis addresses different aspects related to the computation of nonlinear optical (NLO) properties, from the investigation of reference benchmark systems using highly accurate wavefunction methods to that of experimental systems by means of Density Functional Theory (DFT) approaches.

In a first part, we address the performances of several acceleration techniques to the calculation of static NLO properties of small reference systems. The first family of methods tested is based on the Resolution of Identity approximation, which allows reducing the computational cost by reducing the four-index integrals to three-index integrals by a density-fitting procedure. The other family of methods is based on the so-called domain-localized based approximation, which exploits the local nature of dynamic correlation by employing different localization schemes for orbitals. These methods simplify the wavefunction through a judicious employment of several thresholds, cutoffs and parameters which heavily cut-down the computational cost. Two families of methods which aims at enhancing the performances of canonical methods without increasing their computational cost have been studied. The first one is the spin-component scaled Møller-Plesset second-order Perturbation theory (SCS-MP2) methods, which consists in decomposing the MP2 correlation energy in spin components and scaling them in different manners. The last method tested is the Møller-Plesset third-order Perturbation theory Kohn-Sham method (MP3:KS), which uses canonical KS reference orbitals in place of standard Hartree-Fock orbitals.

In a second part, wavefunction-based methods and DFT are employed to decipher the impact of Van der Waals interactions on the structures and NLO properties of a series of azobenzene molecules symmetrically substituted in meta-position with functional groups of different bulkiness. We assess the performance of a large set of density functional approximations in reproducing the geometry, the relative energy and first hyperpolarizability of *E* and *Z* azobenzene isomers in comparison with calculations using MP2 and Coupled Cluster approaches. Moreover, we analyze the individual contribution of the substituents on the NLO response of this series of compounds, giving insights into the precise role of the functional groups responsible for dispersion interactions.

Finally, we report a joint theoretical and experimental investigation of the second-order nonlinear optical properties of four series of amphiphilic cationic chromophores, which involve different push–pull extremities and increasingly large polyenic bridges. These systems have an enhanced second-harmonic response and are of interest for use as probes in biological systems such as lipid membranes. Experimental Electric Field Induced Second Harmonic Generation (EFISHG) measures are made possible for these cationic chromophores by using a solvent with a low relative permittivity, which induces the formation of neutral ion pairs including the positively charged dye and its iodine counterion. The theoretical methodology that has been employed combines classical molecular dynamics and DFT calculations, describing the effects of structural fluctuations on the EFISHG properties of the complexes formed by the dye and its iodine counterion, and providing a rationale to EFISHG experimental measurements.

Contents

Introduction

Nonlinear optics begins in the 1960s with the invention of the Ruby laser by Mainan [1] and the subsequent measurement of the Second Harmonic Generation (SHG) by Franken et al. in 1961 [2]. Since then, this field has continuously grown, bringing a huge impact on the development of photonic technologies. Among all the possible applications, nonlinear optical (NLO) processes are widely employed in dynamic image processing [3], data storage [4], sensing [5] and microscopic imaging [6, 7].

NLO phenomena occur when systems subjected to electromagnetic fields respond non linearly with respect to the field strength [8]. Among all the NLO processes, the already mentioned SHG is one of the most exploited. SHG occurs when two incident photons interact through a medium and combine into one single photon with twice the frequency of the incident ones. This process takes place in specific range of energies that do not match resonant frequencies of molecules that build the medium. Therefore, differently from an absorption/emission process, it takes place in an ultra-short time scale and does not involve energy transfer between the photons and the NLO medium. Nowadays, SHG finds a wide range of applications in optoelectronics [4] but also in bioimaging and therapy [9, 10]. The SHG response is deeply connected to the structural and electronic characteristics of the NLO material and of its constitutive components. Therefore, a proper characterization of a SHG signal provided by a system requires a deep knowledge of its microscopic properties, its response in bulk, and its interaction with the external environment.

The rational design of efficient materials with large SHG response combines different disciplines: chemical synthesis, spectroscopic characterization, and computational modelling. The latter has become essential since it allows the interpretation of structure/properties relationships that can not be fully understood through experimental measurements. For this purpose, identifying an appropriate computational protocol able to model the target phenomena, and that guarantees an acceptable compromise between accuracy and computational cost, is of paramount importance.

At the microscopic level, if a molecule is irradiated with an external electric field $\mathbf{F}$, it experiences an induced polarization or, in other words, a change in its dipole moment

$\boldsymbol{\mu}$, which can be expressed as a Taylor expansion with respect to the strength of $\mathbf{F}$:

$$\mu_i = \mu_i^0 + \sum_j \alpha_{ij} F_j + \sum_{j,k} \beta_{ijk} F_j F_k + \sum_{j,k,l} \gamma_{ijl} F_j F_k F_l + ..., \tag{1}$$

where the expansion coefficients are the components of the optical response tensors, namely the permanent dipole moment μ_i^0, the linear polarizability α_{ij}, and the nonlinear first hyperpolarizabilty β_{ijk} and second hyperpolarizability γ_{ijkl}. If the applied electric fields are oscillating, the optical responses of the molecule are dynamic and depend on the frequency of the incident fields, while static electric fields generate static NLO properties.

The proper modeling of molecular NLO properties relies on quantum mechanical (QM) simulations. However, their practical computation can be particularly challenging. In fact, in order to accurately describe NLO phenomena, a computational method should account for electron correlation effects, employ extended basis sets to make the electronic cloud flexible enough, use tight numerical thresholds criteria and, for embedded systems, model properly the external environment. Therefore, extensive computational resources are needed, which limits the size of the systems that can be investigated and/or the quality of the QM approximations that can be used. A smart selection of the appropriate method according to the scale of the system investigated is thus fundamental.

Generally, the class of methods that provides the best accuracy are wavefunction-based methods (WFN). In particular, post-HF schemes such as Coupled Cluster (CC) or Møller Plesset Perturbation theory (MPn) are considered as the most effective techniques for computing NLO properties. However, because of computational limitations, they can be applied only to small size systems [11–13], and are mainly employed for benchmark purposes on isolated chromophores. On the other hand, Density Functional Theory (DFT) is considered as an excellent alternative to wavefunction methods from the computational point of view and it is widely used in the field of NLO simulations. However, a large number of density functional approximations (DFAs) are not able to quantitatively reproduce the absolute magnitude of NLO responses. These failures are typically attributed to the so-called delocalization error which, in standard DFAs, leads to an over delocalization of the electrons [13–18]. Range-separated DFAs are usually able to reduce the delocalization error, but they usually do not reach the accuracy of wavefunction methods [16, 17]. Nevertheless, provided a proper choice of the exchange-correlation functional, DFT calculations are usually suitable for reproducing experimental trends and provide qualitative structure/property relationships.

In this thesis, different aspects related to the computation of NLO properties have been addressed, from the investigation of reference benchmark systems using highly accurate wavefunction methods to that of experimental systems by means of DFT approaches.

The manuscript starts by a general overview of nonlinear optical phenomena and provides a physical and mathematical description of their microscopical origin. Focusing on

SHG, we show the connection between chemical structure and NLO response, identifying the important structural features for designing molecules able to deliver large SHG signal. Moreover, we present the most common experimental techniques for measuring SHG properties of chromophores in solution.

In Chapter 2, the principle and basic equations of quantum chemical methods employed throughout the thesis are presented, with an highlight on their advantages, limitations, and computational performances. Acceleration techniques based on the Resolution of Identity (RI) approximation or on the use of localized orbitals, whose performance for computing static first and second hyperpolarizabilities are addressed in Chapter 4, are also presented.

Chapter 3 focuses more specifically on the methods employed for computing NLO properties. Three classes of methods are presented with their adavantages and limitations: 1) Sum-Over-State (SOS) approaches based on time-dependent perturbation theory [19], 2) Finite Field methods (FF), which allow to straightforwardly determine the static NLO properties through numerical derivatives of perturbed ground-state energies [20–23], and 3) response theory that allows, for variational wavefunctions, to determine static and dynamic response properties through time averaged derivatives of the so-called “quasi-energy” with respect to the electric field [24, 25].

In Chapters 4 and 5, we address the performances of some state-of-the-art wavefunction methods for the FF calculation of static NLO properties of small benchmark systems of different nature. These methods aim to improve the cost-efficiency of standard post-HF methods by following different strategies. They can be classified into two groups: the acceleration techniques, which aim at reducing the computational cost of standard WFN approximations without affecting their accuracy, and enhancement techniques which aim at improving the accuracy without increasing the computational effort. Although the performances of accelerated and enhanced methods for calculations of thermodynamic, kinetics, and molecular properties have been already assessed [26–36], a systematic study of their performances on NLO properties has been overlooked so far in literature.

Chapter 4 reports the performances of several acceleration techniques applied to MP2 and CC methods. The first family of methods tested is based on the RI approximation [26–28], which allows reducing the computational cost by reducing the four-index integrals to three-index integrals by a density-fitting procedure. Another family of methods is based on the so-called domain-localized based approximation, which exploits the local nature of dynamic correlation by employing different localization schemes for orbitals [29–34]. These methods simplify the wavefunction through the employment of several thresholds and parameters which heavily cut-down the computational cost. The last family of methods tested in this chapter is the spin-component scaled MP2 methods, which consist in decomposing the MP2 correlation energy in same spin and opposite spin components and scaling them in different manners [37, 38]. These last methods, although they do not explicitly accelerate canonical MP2, they are supposed to improve its performances.

In Chapter 5, the performance of the so-called MP3 Kohn-Sham method (MP3:KS) [35, 36] are evaluated. This method improves standard MP3 (referred to as MP3:HF) by using canonical KS reference orbitals in place of standard HF orbitals. MP3:KS was shown to improve MP3:HF in the calculation of thermochemical properties, barrier heights and dipole moments, being comparable to CCSD with a fraction of its computational cost [36]. The performance of MP3:KS was addressed for several DFAs including different percentages of explicit HF exchange, highlighting the importance of the delocalization error in DFT.

In Chapter 6, we employed WFN and DFT methods to decipher the impact of Van der Waals (vdW) interactions on the structures and NLO properties of a series of azobenzene molecules symmetrically substituted in *meta*-position with functional groups of different bulkiness. It has been already observed how vdW interactions impact the thermal stability and the photochemistry of molecular switches [39, 40], however, their impact on NLO has been overlooked so far. We assess the performance of a large set of DFAs in reproducing the geometry, the relative energy and static first hyperpolarizability of E and Z azobenzene isomers in comparison with MP2 and Couple Cluster singels and doubles (CC2) calculations. Moreover, we analyze the individual contribution of the substituents on the NLO response of this series of compounds, giving insights into the precise role of the functional groups responsible for dispersion interactions.

Finally, in Chapter 7, the second-order nonlinear optical properties of four series of amphiphilic cationic chromophores involving different push–pull extremities and increasingly large polyenic bridges have been investigated both theoretically and experimentally. These systems have an enhanced second-harmonic response and are of interest for use as probes in biological systems such as lipid membranes [4]. Experimental Electric Field Induced Second Harmonic Generation (EFISHG) measures are made possible for these cationic chromophores by using a solvent with a low relative permittivity, which induces the formation of neutral ion pairs including the positively-charged dye and its iodine counterion [41]. The computational approach combines sequentially classical molecular dynamics (MD) and DFT calculations. This theoretical methodology has been employed to describe the effects of structural fluctuations on the EFISHG properties of the complexes formed by the dye and its iodine counterion, and provides a rationale to EFISHG experimental measurements.

Bibliography

[1] Maiman, T. Stimulated Optical Radiation in Ruby. *Nature* **1960**, 493–494.

[2] Franken, P. A.; Hill, A. E.; Peters, C. W.; Weinreich, G. Generation of Optical Harmonics. *Phys. Rev. Lett.* **1961**, *7*, 118–119.

[3] Van Steenwinckel, D.; Hendrickx, E.; Persoons, A.; Samyn, C. Large Dynamic Ranges in Photorefractive NLO Polymers and NLO-Polymer-Dispersed Liquid Crystals Using a Bifunctional Chromophore as a Charge Transporter. *Chem. Mater.* **2001**, *13*, 1230–1237.

[4] Dalton, L. R.; Sullivan, P. A.; Bale, D. H. Electric field poled organic electro-optic materials: state of the art and future prospects. *Chem. Rev.* **2010**, *110*, 25–55.

[5] Ray, P. C. Size and Shape Dependent Second Order Nonlinear Optical Properties of Nanomaterials and Their Application in Biological and Chemical Sensing. *Chem. Rev.* **2010**, *110*, 5332–5365.

[6] Huff, T. B.; Shi, Y.; Fu, Y.; Wang, H.; Cheng, J.-X. Multimodal Nonlinear Optical Microscopy and Applications to Central Nervous System Imaging. *IEEE J. Sel. Top. Quantum Electron.* **2008**, *14*, 4–9.

[7] Wang, H.-W.; Langohr, I. M.; Sturek, M.; Cheng, J.-X. Imaging and Quantitative Analysis of Atherosclerotic Lesions by CARS-Based Multimodal Nonlinear Optical Microscopy. *Arterioscler. Thromb. Vasc. Biol.* **2009**, *29*, 1342–1348.

[8] Boyd, R. W. *Nonlinear Optics (Fourth Edition)*, fourth edition ed.; Academic Press, 2020; pp 1–64.

[9] Campagnola, P. J.; Wei, M. D.; Loew, L. M. High-Resolution Nonlinear Optical Imaging of Live Cells by Second Harmonic Generation. *Biophys. J.* **1999**, *77*, 3341–3349.

[10] Moreaux, L.; Sandre, O.; Charpak, S.; Blanchard-Desce, M.; Mertz, J. Coherent Scattering in Multi-Harmonic Light Microscopy. *Biophys. J.* **2001**, *80*, 1568–1574.

[11] Champagne, B.; Botek, E.; Nakano, M.; Nitta, T.; Yamaguchi, K. Basis set and electron correlation effects on the polarizability and second hyperpolarizability of model open-shell π-conjugated systems. *J. Chem. Phys.* **2005**, *122*, 114315.

[12] Christiansen, O.; Coriani, S.; Gauss, J.; Hättig, C.; JØrgensen, P.; Pawłowski, F.; Rizzo, A. In *Non-Linear Optical Properties of Matter: From Molecules to Condensed Phases*; Papadopoulos, M. G., Sadlej, A. J., Leszczynski, J., Eds.; Springer Netherlands: Dordrecht, 2006; pp 51–99.

[13] de Wergifosse, M.; Champagne, B. Electron correlation effects on the first hyperpolarizability of push–pull π-conjugated systems. *J. Chem. Phys.* **2011**, *134*, 074113.

[14] Champagne, B.; Perpète, E. A.; Jacquemin, D.; van Gisbergen, S. J. A.; Baerends, E.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. Assessment of Conventional Density Functional Schemes for Computing the Dipole Moment and (Hyper)polarizabilities of Push−Pull π-Conjugated Systems. *J. Phys. Chem. A* **2000**, *104*, 4755–4763.

[15] Zaleśny, R.; Medved', M.; Sitkiewicz, S. P.; Matito, E.; Luis, J. M. Can Density Functional Theory Be Trusted for High-Order Electric Properties? The Case of Hydrogen-Bonded Complexes. *J. Chem. Theory Comput.* **2019**, *15*, 3570–3579.

[16] Besalú-Sala, P.; Sitkiewicz, S. P.; Salvador, P.; Matito, E.; Luis, J. M. A new tuned range-separated density functional for the accurate calculation of second hyperpolarizabilities. *Phys. Chem. Chem. Phys.* **2020**, *22*, 11871–11880.

[17] Lescos, L.; Sitkiewicz, S. P.; Beaujean, P.; Blanchard-Desce, M.; Champagne, B.; Matito, E.; Castet, F. Performance of DFT functionals for calculating the second-order nonlinear optical properties of dipolar merocyanines. *Phys. Chem. Chem. Phys.* **2020**, *22*, 16579–16594.

[18] Naim, C.; Castet, F.; Matito, E. Impact of van der Waals interactions on the structural and nonlinear optical properties of azobenzene switches. *Phys. Chem. Chem. Phys.* **2021**, *23*, 21227–21239.

[19] Norman, P.; Ruud, K.; Saue, T. *Principles and practices of molecular properties: theory, modeling, and simulations*; Wiley, 2018; p 480 P.

[20] Richardson, L. F.; Gaunt, J. A. VIII. The deferred approach to the limit. *Philos. Trans. Royal Soc. A* **1927**, *226*, 299–361.

[21] Medved', M.; Stachová, M.; Jacquemin, D.; André, J.; Perpète, E. A. A generalized Romberg differentiation procedure for calculation of hyperpolarizabilities. *J. Mol. Struct.* **2007**, *847*, 39–46.

[22] Mohammed, A. A. K.; Limacher, P. A.; Champagne, B. Finding optimal finite field strengths allowing for a maximum of precision in the calculation of polarizabilities and hyperpolarizabilities. *J. Comput. Chem.* **2013**, *34*, 1497–1507.

[23] de Wergifosse, M.; Liégeois, V.; Champagne, B. Evaluation of the molecular static and dynamic first hyperpolarizabilities. *Int. J. Quantum Chem.* **2014**, *114*, 900–910.

[24] Christiansen, O.; Jørgensen, P.; Hättig, C. Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quantum Chem.* **1998**, *68*, 1–52.

[25] Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations. *Chem. Rev.* **2012**, *112*, 543–631.

[26] Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[27] Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

[28] Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange. *Chem. Phys.* **2009**, *356*, 98–109, Moving Frontiers in Quantum Chemistry:.

[29] Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladóczki, B.; Kállay, M. An efficient linear-scaling CCSD(T) method based on local natural orbitals. *J. Chem. Phys.* **2013**, 094105.

[30] Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.

[31] Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, 034108.

[32] Nagy, P. R.; Samu, G.; Kállay, M. An integral-direct linear-scaling second-order Møller–Plesset approach. *J. Chem. Theory Comput.* **2016**, 4897–4914.

[33] Nagy, P. R.; Samu, G.; Kállay, M. Optimization of the Linear-Scaling Local Natural Orbital CCSD(T) Method: Improved Algorithm and Benchmark Applications. *J. Chem. Theory Comput.* **2018**, *14*, 4193–4215.

[34] Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.

[35] Bertels, L. W.; Lee, J.; Head-Gordon, M. Third-Order Møller–Plesset Perturbation Theory Made Useful? Choice of Orbitals and Scaling Greatly Improves Accuracy for Thermochemistry, Kinetics, and Intermolecular Interactions. *J. Phys. Chem. Lett.* **2019**, *10*, 4170–4176.

[36] Rettig, A.; Hait, D.; Bertels, L. W.; Head-Gordon, M. Third-Order Møller–Plesset Theory Made More Useful? The Role of Density Functional Theory Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 7473–7489.

[37] Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

[38] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

[39] Schweighauser, L.; Strauss, M. A.; Bellotto, S.; Wegner, H. A. Attraction or Repulsion? London Dispersion Forces Control Azobenzene Switches. *Angew. Chem. Int. Ed.* **2015**, *54*, 13436–13439.

[40] Fabrizio, A.; Corminboeuf, C. How do London Dispersion Interactions Impact the Photochemical Processes of Molecular Switches? *J. Phys. Chem. Lett.* **2018**, *9*, 464–470.

[41] Alain, V.; Blanchard-Desce, M.; Ledoux-Rak, I.; Zyss, J. Amphiphilic polyenic push–pull chromophores for nonlinear optical applications. *Chem. Commun.* **2000**, 353–354.

Chapter 1

Introduction to nonlinear optical properties

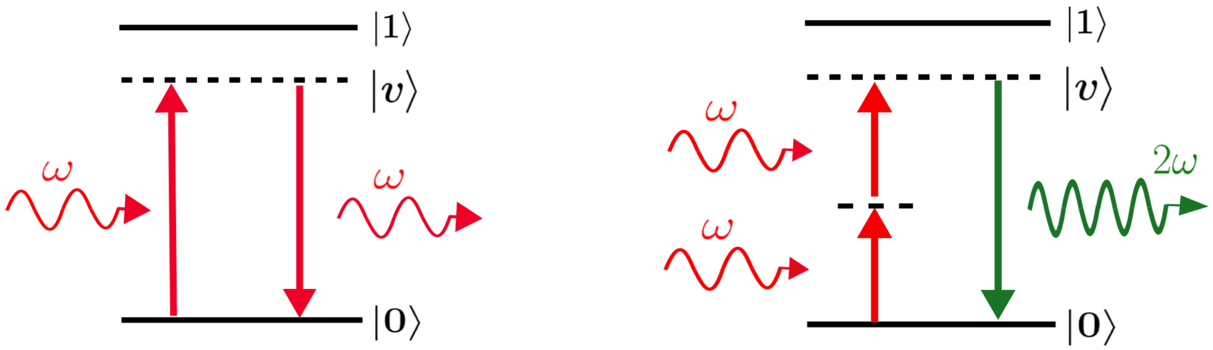
1.1 The source of nonlinear optical properties

Optical properties find their origin in the response of molecules or materials to an external electric field, *i.e.*, to the electric component of the electromagnetic wave in the case of light irradiation. In order to trigger electronic and vibrational excitations of molecules, external oscillatory radiations generally have a wavelength of a few hundred nanometers, thus ranging from the infrared to the visible. Having molecules the size of a few nanometers, the electric fields can be considered uniform over the molecular volume. If the energy carried by the incident photons is not in resonance with the excitation energy of the system, they might interact with the molecule through a scattering process. In this case, each incident photon transfers its momentum to the molecule, which absorbs it and re-emits it without losing any energy, namely, the total energy of the scattered photons is the same as the incident ones.

From the molecular point of view, a scattering process can be described by considering that the molecule absorbs a photon and lays in a so-called virtual state with a short lifetime ($\propto$ fs) and energy far from any of the electronic excited states. From a classical point of view, when the electric field $\mathbf{F}(t)$ interacts with the molecule, the electrons and the nuclei start to oscillate as antennas and propagate the signal. For low field strengths, these oscillating charges generate an induced dipole $\boldsymbol{\mu}^{ind}(t)$ moment that responds linearly to the applied field and depends on the electric polarizability tensor $\boldsymbol{\alpha}$. The i^{th} component of $\boldsymbol{\mu}^{ind}(t)$ is expressed as follows:

$$\mu_i^{ind}(t) = \mu_i^0 + \alpha_{ij} F_j(t), \tag{1.1}$$

where $\boldsymbol{\mu}^0$ is the intrinsic dipole moment of the molecule, and the convention of intrinsic summation over the repeated j is used. This process is referred to as linear scattering and it is illustrated in Fig. 1.1a.



(a) (b)

Figure 1.1: Schematic representation of a linear scattering process (a) and a Second Harmonic Generation process (b). In the figure $|0\rangle$ indicates the ground state, $|1\rangle$ the first excited state and $|v\rangle$ a virtual state in which the scattering process takes place.

If the radiation is intense and focalized on the system (typically produced by a laser field), two or more photons can interact in a single process, and the response of the molecule is no longer linear. The induced dipole moment is then expressed as a Taylor expansion with respect to the external electric fields:

$$\mu_i^{ind}(t) = \mu_i^0 + \alpha_{ij} F_j(t) + \beta_{ijk} F_j(t) F_k(t) + \gamma_{ijkl} F_j(t) F_k(t) F_l(t) + ..., \tag{1.2}$$

where the coefficients α_{ij} are the components of the linear polarizability tensor defined previously, while β_{ijk} and γ_{ijkl} are the components of the tensors describing the nonlinear optical (NLO) behavior of the system, respectively referred to as the first and second hyperpolarizabilities. Equation 1.2 describes the scattering processes in which several incident photons can be combined to produce a resulting photon with a frequency corresponding to the sum of the incident ones. If the incident light can be decomposed in different spectral components, these might combine giving rise to several effects, which can be described by expanding the electric field components as a Fourier series:

$$F_i(t) = \sum_{\omega} F_i^{\omega} e^{-i\omega t} \tag{1.3}$$

where F_i^{ω} represents the amplitude of the field associated to each frequency ω. By reporting Eq. 1.3 into Eq. 1.2, the induced dipole moment can be then expressed as:

$$\begin{aligned} \mu_i^{ind}(t) =& \mu_i^0 + \sum_{\omega} \alpha_{ij}(-\omega;\omega) F_j^{\omega} e^{-i\omega t} + \sum_{\omega_1,\omega_2} \beta_{ijk}(-\omega_\sigma;\omega_1,\omega_2) F_j^{\omega_1} F_k^{\omega_2} e^{-i\omega_\sigma t} + \\ &+ \sum_{\omega_1,\omega_2,\omega_3} \gamma_{ijkl}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) F_j^{\omega_1} F_k^{\omega_2} F_l^{\omega_3} e^{-i\omega_\sigma t} + ... \end{aligned} \tag{1.4}$$

where ω_σ represents the frequency of the scattered photons, which conventionally appear with a negative sign in order to differentiate from the incident ones. The values of ω_σ for the second- and third-order responses are obtained by summing the frequencies of the incident photons, $\omega_\sigma = \omega_1 + \omega_2$ and $\omega_\sigma = \omega_1 + \omega_2 + \omega_3$ respectively. The sum runs also over negative frequencies, indeed both of sum and differences of frequencies are possible. According to the number of photons that are combined and their frequencies, different NLO phenomena can occur. The most common first-, second-, and third-order NLO processes are reported in Table 1.1.

Table 1.1: Most common linear and second- and third-order nonlinear optical processes.

Response	Property
$\alpha(0;0)$	Static polarizability
$\alpha(-\omega;\omega)$	Light Scattering
$\beta(0;0,0)$	Static first hyperpolarizability
$\beta(0;\omega,-\omega)$	Optical Rectification (OR)
$\beta(-\omega;\omega,0)$	Linear change of birefringence (EO-Pockels)
$\beta(-2\omega;\omega,\omega)$	Second Harmonic Generation (SHG)
$\gamma(0;0,0,0)$	Static second hyperpolarizability
$\gamma(0;\omega,-\omega,0)$	dc optical rectification (dc-OR)
$\gamma(-\omega;\omega,0,0)$	Quadratic change of birefringence (EO-Kerr)
$\gamma(-\omega;\omega,-\omega,\omega)$	Intensity dependent refractive index (DFWM)
$\gamma(-2\omega;\omega,\omega,0)$	Electric field induced SHG (EFISHG)
$\gamma(-3\omega;\omega,\omega,\omega)$	Third-harmonic generation (THG)

A large part of this work is dedicated to the study of the Second Harmonic Generation (SHG), in which two incident photons with the same frequency ω interact through the molecule to generate a photon of frequency 2ω (see a schematic representation of this process in Fig. 1.1b).

1.1.1 Impact of symmetry on the (hyper)polarizability tensors

According to Eq. 1.4, α, β, and γ tensors must be fully determined to decipher the linear, second-, and third-order NLO responses of a molecule. However, their expression can be simplified according to the symmetry, which is linked to the experimental conditions and the molecular structure. In practice, the tensors β and γ have, respectively, 27 and 81 independent components. However, when the incident photons have the same frequency, the tensor components stay invariant upon permutation of the last indices:

$$\beta_{ijk} = \beta_{ikj} \tag{1.5}$$

$$\gamma_{ijkl} = \gamma_{ijlk} = \gamma_{ikjl} = \gamma_{iklj} = \gamma_{ilkj} = \gamma_{iljk}. \tag{1.6}$$

Therefore, the number of independent components is reduced to 18 for β and 30 for γ. Moreover, if ω and ω_σ are far from any resonance frequency of the system, the number of independent coefficients of β and γ can be further reduced by applying the Kleinman's permutation rules [1]:

$$\beta_{kij} = \beta_{ijk} \tag{1.7}$$

$$\gamma_{lijk} = \gamma_{ijkl}, \tag{1.8}$$

which is strictly valid only for static properties ($\omega = 0$). With this formulation, the independent components of β and γ tensors are respectively reduced to 10 and 15.

If the system possesses a particular spatial symmetry, the number of independent components in the β and γ tensors can be further reduced. A very particular situation is when the system has an inversion center, *i.e.*, it is centrosymmetric. In this case, each of the components of the β_{ijk} tensor is 0. This can be easily understood by invoking Neumann's principle [2], which states that if a system is invariant with respect to a symmetry operation, each of its properties is also invariant to it. Consequently, for a centrosymmetric system subjected to an inversion of the coordinates (described by the operator $\hat{I}$), the electric field and the dipole moment transform by inversion as $\mathbf{F} \rightarrow -\mathbf{F}$ and $\boldsymbol{\mu}_0 \rightarrow -\boldsymbol{\mu}_0$. The expression of Eq. 1.2 under the inversion operation is thus:

$$\hat{I}\mu_i^{ind} = -\mu_i^{ind} = -\mu_i^0 - \alpha_{ij}F_j + \beta_{ijk}F_jF_k - \gamma_{ijkl}F_jF_kF_l + ... \tag{1.9}$$

By comparing term by term Eq. 1.2 and Eq. 1.9, we deduce that β (and any even order NLO property) must be 0. More generally, if a molecule belongs to a given symmetry group, it is possible to straightforwardly identify the components of the β tensor that are different from 0 (the form of the first hyperpolarizabilty tensor of molecules belonging to the most current point groups can be found in [3]). It must be noticed that in the case of supramolecular aggregates, the entire system must have a non-centrosymmetric spatial organization in order to have a non-zero second-order response.

1.1.2 Macroscopic nonlinear optical properties

We focused so far on the microscopic responses of molecules to external electric fields. However, since experimental measurements are usually performed in liquids or solids, it is also crucial to interpret the optical responses of bulk systems. At the macroscopic level, the components of the polarization vector $\mathbf{P}(t)$ of a medium polarized by a time-dependent electric field $\mathbf{F}(t)$, can be expanded as a sum of terms of different orders:

$$P_i(t) = P_i^{(0)}(t) + P_i^{(1)}(t) + P_i^{(2)}(t) + P_i^{(3)}(t) + ..., \tag{1.10}$$

where $\mathbf{P}^{(0)}(t)$ is the intrinsic polarization, $\mathbf{P}^{(1)}(t)$ the linear polarization, $\mathbf{P}^{(2)}(t)$ the quadratic one, and so on. Similarly to the microscopic properties, if the electric field has different spectral components, by including Eq. 1.3 in Eq. 1.10 we obtain:

$$\begin{aligned} P_i(t) =& P_i^{(0)}(t) + \epsilon_0 \sum_{\omega} \chi_{ij}^{(1)}(-\omega;\omega) F_j^{\omega} e^{-i\omega t} + \epsilon_0 \sum_{\omega_1,\omega_2} \chi_{ijk}^{(2)}(-\omega_\sigma;\omega_1,\omega_2) F_j^{\omega_1} F_k^{\omega_2} e^{-i\omega_\sigma t} \\ &+ \epsilon_0 \sum_{\omega_1,\omega_2} \chi_{ijkl}^{(3)}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) F_j^{\omega_1} F_k^{\omega_2} F_l^{\omega_3} e^{-i\omega_\sigma t} + ..., \end{aligned} \tag{1.11}$$

where ϵ_0 is the dielectric constant of the vacuum, while $\chi^{(1)}$, $\chi^{(2)}$ and $\chi^{(3)}$ are respectively the linear, quadratic and cubic susceptibilities, which are expressed as tensors of rank 2, 3, and 4. By comparing Eq. 1.11 and Eq. 1.4 and considering that the electric field is constant in the system, we observe a direct correspondence between the microscopic (hyper)polarizabilities and the macroscopic susceptibilities:

$$\chi^{(1)} = \frac{\alpha}{\epsilon_0} \tag{1.12}$$

$$\chi^{(2)} = \frac{\beta}{2\epsilon_0} \tag{1.13}$$

$$\chi^{(3)} = \frac{\gamma}{6\epsilon_0} \tag{1.14}$$

However, it is important to stress that the electric fields experienced by a given molecule within a macroscopic medium are different from the external one. In fact, the presence of the medium perturbs the macroscopic electric field and each molecule experiences a so-called local field E_{loc}^{ω}. This effective field depends on the shape, the dimension, and the kind of medium which surrounds the molecules and remains challenging to predict [4].

1.2 Practical applications of Second Harmonic Generation

The first measure of SHG was made by Franken and coworkers in 1961 [5]. Since then, SHG has received an ever-growing attention from the scientific community, owing to the wide range of applications in which this phenomenon can be exploited, ranging from data storage, sensing, to bioimaging [6–9]. Systems displaying large first hyperpolarizability are for instance used as functional components in electro-optical materials or as exogenous probes in microscopy techniques used for investigating biological media such as cell membranes.

Among NLO materials, organic compounds are nowadays popular systems for the fabrication of NLO devices, as they can display faster responses, higher flexibility, lower environmental impact, and higher NLO response than inorganic systems [10]. A particular relevant class of compounds are push-pull D$-\pi-$A systems, which are composed of an electron-donor unit (D) and an electron-acceptor (A) one linked by a π-conjugated bridge.

These molecules usually exhibit intense NLO responses due to the intramolecular charge transfer between the donor and the acceptor substituents. The conjugated π-bridge acts as a reservoir of easily delocalizable electrons and thus contributes to the enhancement of the NLO response, as schematized in Figure 1.2.

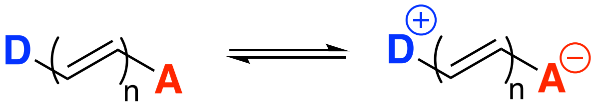


Figure 1.2: Schematic representation of a generic D-π-A push-pull system with the associated Lewis resonance structures.

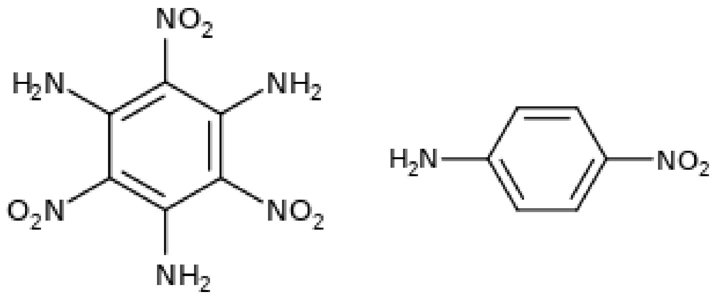
A large variety of D$-\pi-$A systems have been studied both from the experimental and computational sides. In fact, exploring different combinations of D/A components and optimizing the nature and length of the π-bridge is the most effective way to design novel materials with enhanced NLO properties [10–13]. As detailed in Chapter 3, it is possible to rationalize the response of D$-\pi-$A systems by using the two-level approximation first formulated by Oudar [14], which connects the response of push-pull systems to spectroscopic quantities related to the one-photon absorption of the molecule.

The magnitude of the nonlinear optical properties of D$-\pi-$A conjugated systems is directly related to the strength of the electron conjugation within the molecule, which is reflected in the Bond Length Alternation (BLA) [15–17] that measures the average difference in length between double and single bonds along the π-bridge. For a π-conjugated chain composed of N carbon atoms with an interatomic distance $d_{i,j}$ between carbons i and j, the BLA is defined as:

$$BLA = \frac{1}{N-2}\sum_{i=1}^{N-2}(-1)^{i+1}(d_{i+1,i+2} - d_{i,i+1}) \tag{1.15}$$

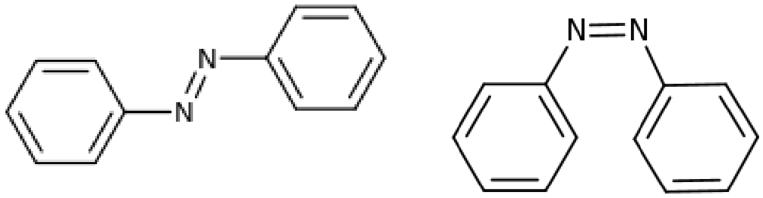
Another class of compounds exhibiting large NLO responses are octupolar molecules, *i.e.*, systems which belong to the D_{3h}, T, or T_d groups of symmetry with a null dipole moment [18]. The simplest example is the 1,3,5-triamino-2,4,6-trinitro-benzene (TATB) [19] which has a response twice larger than the para-nitroaniline molecule, its dipolar counterpart (Fig. 1.3). A large variety of structures belonging to this family have been designed and studied both computationally and experimentally [20–22].

Since they can be used as building blocks for designing responsive materials, molecular systems that exhibit switchable second-order NLO responses, *i.e.*, that can commute between two forms displaying contrast in their first hyperpolarizabilities, have also witnessed a growing interest during the last three decades [23–25]. The commutation between the two forms can be triggered using different techniques, such as variation of pH [26], of temperature [27], application of a redox potential [28] or illumination with an external light [29]. Photochromic molecules often constitute the most appropriate candidates for technological applications since light is a fast and non-invasive way to remotely control



(a) TATB (b) PNA

Figure 1.3: Molecular representation of: (a) 1,3,5-triamino-2,4,6-trinitro-benzene and (b) para-nitroaniline.



(a) *trans (E)* (b) *cis (Z)*

Figure 1.4: Molecular representation of: (a) *trans*-azobenzene, and (b) *cis*-azobenzene.

the state of a system. Nowadays, azobenzene derivatives are the most prominent class of photochromic molecules employed in the design of NLO devices and have been incorporated into Langimur-Blodgett films [30], liquid-crystalline polymers [31] or self-assembled monolayers [32, 33]. They include two phenyl rings linked by an azo bond, and can interconvert between *trans* (E) and *cis* (Z) isomers through light irradiation (Fig. 1.4). The two forms present different first hyperpolarizability values, whose contrast can be enhanced by adding appropriate substituents onto the terminal phenyl groups [34, 35]. Deep knowledge of the relationships linking the chemical structure of azobenzenes to their NLO properties is thus fundamental for a rational design of efficient NLO devices. In chapter 6 of this thesis, we investigate how van der Waals interactions between substituents grafted in *meta* positions of the phenyl groups impact the relative stability of the two isomers of various azobenzene derivatives, as well as the first hyperpolarizability of the *cis* form.

Besides photoresponsive materials, the second-order NLO responses of organic compounds are also largely employed in imaging microscopy. Second harmonic imaging microscopy (SHIM) is complementary and presents advantages compared to two-photon absorption microscopy, because it allows a faster and more sensitive detection while avoiding issues such as dye bleaching, signal saturation or tissue autofluorescence [36]. SHIM allows to investigate biological structures such as cells or entire tissues which have a non-centrosymmetric structure. In this case, the SHG signal originates from their constituent proteins such as collagen [37], myosin [38], and microtubules [39]. Moreover, by labeling centrosymmetric tissues with functional NLO dyes, it is possible to characterize biological media which in principle lack or have a small SHG signal [40]. In this context, organic dyes with a large SHG response such as amphiphilic chromophores are frequently employed as exogenous probes for bi-lipid membranes in high-resolution SHG microscopy techniques [40–43]. These dyes are characterized by a polar head that can be anchored into the membrane and a hydrophobic tail which interacts with the cellular fluid. Their response can be enhanced by their functionalization with D and A moieties. A series of positively amphiphilic chromophores relevant for SHIM applications and incorporating pyridinium or quinolinium acceptor moieties are characterized both computationally and experimentally in Chapter 7 of this thesis.

1.3 Experimental techniques for measuring molecular second harmonic responses

As this thesis focuses on the characterization of molecular NLO responses, we limit this section to the description of experimental techniques allowing the measurement of the SHG signal of molecules in solution, namely the Electric Field-Induced Second Harmonic Generation (EFISHG) and the Hyper-Rayleigh Scattering (HRS) [3]. As mentioned above, in order to have an appreciable SHG signal, the target systems must be non-centrosymmetric. However, molecules in solutions tend to be on average isotropically oriented, and therefore the resulting medium is on average centrosymmetric. Therefore, even molecules with large first hyperpolarizability such as the dipolar or octupolar systems discussed above are challenging to characterize when they are solvated. We discuss in the next sections how EFISHG and HRS techniques overcome this limitation.

1.3.1 Electric Field-Induced Second Harmonic Generation

The EFISHG method is the first technique employed for measuring SHG in solids [44], gas [45, 46] and solutions [47, 48]. Besides the oscillating electric field used for evaluating the SHG signals, an additional static electric field is applied in order to give a preferential orientation to the bulk system (see a schematic representation of this process in Fig. 1.5).

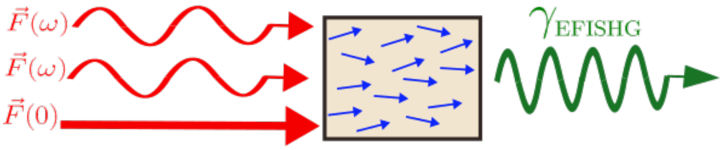


Figure 1.5: Schematic representation of an EFISHG process.

Therefore, EFISHG is a technique based on the third-order response. The EFISHG response is obtained by measuring the intensity $I_{2\omega}$ of the resulting field, which depends quadratically on the γ_{EFISHG} second-hyperpolarizability:

$$I_{2\omega} \propto (\gamma_{EFISHG})^2 I_\omega^2 \tag{1.16}$$

Formally, γ_{EFISHG} can be expressed as a sum of two contributions:

$$\gamma_{EFISHG} = \gamma_{||}(-2\omega;\omega,\omega,0) + \frac{\mu\beta_{||}(-2\omega;\omega,\omega)}{3kT} \tag{1.17}$$

The first term ($\gamma_{||}$) is a pure third-order term that can be expressed as a function of the components of the molecular γ tensor:

$$\gamma_{||} = \frac{1}{15}\sum_{i,j}^{x,y,z}(2\gamma_{iijj} + \gamma_{ijji}) \tag{1.18}$$

The second term involves the norm of the permanent dipole moment $\mu = |\boldsymbol{\mu}|$ and a second order term ($\beta_{||}$), which is the projection of the tensor $\boldsymbol{\beta}$ along the dipole moment direction:

$$\beta_{||} = \frac{3}{5}\frac{\mu_i\beta_i}{\mu} \tag{1.19}$$

where β_i are the components of the β vector defined as :

$$\beta_i = \frac{1}{3}\sum_{j}^{x,y,z}(\beta_{ijj} + 2\beta_{jij}) \tag{1.20}$$

Obtaining separately the two components of Eq. 1.17 is experimentally challenging because it requires a complex experimental setup for measuring γ_{EFISHG} at different temperatures. However, in practice, it has been observed that the term $\gamma_{||}$ is generally negligible with respect to $\mu\beta_{||}/(3kT)$. Therefore, the EFISHG response is usually reported as an *effective* second-order response:

$$\gamma_{EFISHG} = \frac{[\mu\beta_{||}(-2\omega;\omega,\omega)]_{eff}}{3kT}, \tag{1.21}$$

which results convenient from the experimental point of view considering that measurements can be then performed at a single temperature. Although this approximation is generally valid for dipolar molecules with large hyperpolarizabilities, computational studies nevertheless proved that $\gamma_{||}$ can be large and even predominant in particular conditions (see for example [49] and Chapter 7 of this thesis).

An important limitation of the EFISHG technique is that it can not be employed for systems with null dipole moments such as octupolar molecules, as well as for charged molecules, for which the external static electric field would cause a migration of the ions. However, if the solvent employed has a low polarity (such as chloroform), ions and counterions stay in close proximity in the solution, forming neutral entities that allow performing EFISHG measurements [50–52]. This specific case is discussed in detail in Chapter 7.

1.3.2 Hyper-Rayleigh Scattering

The HRS technique, developed in 1965, [53] is a method that allows measuring directly the second-Harmonic generation, without applying a static field to break the isotopic orientational order of the solution, and thus without including third-order contributions. This approach is based on the assumption that, although molecules in solution are on average distributed centrosymmetrically, some noncentrosymmetric deviations can be observed because of density or rotational fluctuations. In fact, the molecular scatterers are timely uncorrelated, *i.e.*, there is no phase relation between scattered fields; therefore, the resulting signal is incoherent and the time-average of $\boldsymbol{\mu}(t)$ is different from 0. Thus, contrary to the EFISHG technique, HRS can be applied straightforwardly to ions and octupolar molecules. The global signal is related to the first hyperpolarizability of each individual molecule, which is averaged in all possible orientations in the term $\langle\beta^2_{HRS}\rangle$. Moreover, the total intensity depends on the molecular concentration C, so that [3]:

$$I_{2\omega} \propto C\,\langle\beta^2_{HRS}\rangle\, I^2_{\omega}. \tag{1.22}$$

The connection between $\langle\beta^2_{HRS}\rangle$ and the β_{ijk} tensor components depends on the polarization of the incident light. Denoting by X, Y, and Z are the axis of the laboratory frame, a typical experimental HRS setup is schematized in Fig. 1.6. The incident light is propagated in the Y-direction and polarized in the XZ-plane, while the HRS signal is measured in the X-direction. The incident beams are usually polarized either horizontally (parallel to X) or vertically (parallel to Z), while the scattered light is measured using vertical polarization.

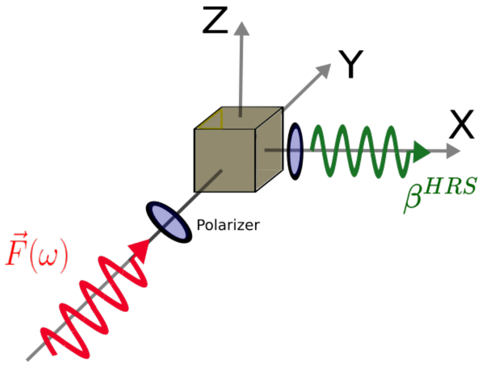


Figure 1.6: Schematic representation of the experimental setup of an HRS measurement.

With this setup, the total HRS hyperpolarizability, β_{HRS}, can be expressed as:

$$\beta_{HRS} = \sqrt{\langle\beta^2_{HRS}\rangle} = \sqrt{\langle\beta^2_{ZZZ}\rangle + \langle\beta^2_{ZXX}\rangle}, \tag{1.23}$$

where $\langle\beta^2_{ZZZ}\rangle$ and $\langle\beta^2_{ZXX}\rangle$ correspond to the scattered signal measured for a vertical polarization using incident beams polarized along Z and X, respectively. The relationships between the experimental HRS components expressed in the laboratory frame (XYZ) and the components of the first hyperpolarizability tensor expressed in the molecular frame (xyz) are obtained by the following equations [54]:

$$\begin{aligned}
\langle\beta^2_{ZZZ}\rangle = {} & \frac{1}{7}\sum_{\zeta}^{x,y,z}\beta^2_{\zeta\zeta\zeta} + \frac{4}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta^2_{\zeta\zeta\eta} + \frac{2}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta_{\zeta\zeta\zeta}\beta_{\zeta\eta\eta} \\
& + \frac{4}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta_{\eta\zeta\zeta}\beta_{\zeta\zeta\eta} + \frac{4}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta_{\zeta\zeta\zeta}\beta_{\eta\eta\zeta} + \frac{1}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta^2_{\eta\zeta\zeta} \\
& + \frac{4}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\zeta\zeta\eta}\beta_{\eta\xi\xi} + \frac{1}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\eta\zeta\zeta}\beta_{\eta\xi\xi} \\
& + \frac{4}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\zeta\zeta\eta}\beta_{\xi\xi\eta} \\
& + \frac{2}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta^2_{\zeta\eta\xi} + \frac{4}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\zeta\eta\xi}\beta_{\eta\zeta\xi}
\end{aligned} \tag{1.24}$$

$$\begin{aligned}
\langle\beta^2_{ZXX}\rangle = {} & \frac{1}{35}\sum_{\zeta}^{x,y,z}\beta^2_{\zeta\zeta\zeta} + \frac{4}{105}\sum_{\zeta\neq\eta}^{x,y,z}\beta_{\zeta\zeta\zeta}\beta_{\zeta\eta\eta} - \frac{2}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta_{\zeta\zeta\zeta}\beta_{\eta\eta\zeta} \\
& + \frac{8}{105}\sum_{\zeta\neq\eta}^{x,y,z}\beta^2_{\zeta\zeta\eta} + \frac{3}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta^2_{\zeta\eta\eta} - \frac{2}{35}\sum_{\zeta\neq\eta}^{x,y,z}\beta_{\zeta\zeta\eta}\beta_{\eta\zeta\zeta} \\
& + \frac{1}{35}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\zeta\eta\eta}\beta_{\zeta\xi\xi} - \frac{2}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\zeta\zeta\xi}\beta_{\eta\eta\xi} \\
& - \frac{2}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\zeta\zeta\eta}\beta_{\eta\xi\xi} \\
& + \frac{2}{35}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta^2_{\zeta\eta\xi} - \frac{2}{105}\sum_{\zeta\neq\eta\neq\xi}^{x,y,z}\beta_{\zeta\eta\xi}\beta_{\eta\zeta\xi}
\end{aligned} \tag{1.25}$$

Interestingly, the HRS signals can be also employed to investigate the molecular symmetry of a molecule. An indicator of the symmetry is the ratio between the two polarization components of the scattered signal, namely, the depolarization ratio (DR):

$$DR = \frac{\langle\beta^2_{ZZZ}\rangle}{\langle\beta^2_{ZXX}\rangle}. \tag{1.26}$$

DR adopts specific values according to the symmetry of the system under study. For instance, for a system with a symmetry C_{2v} with the charge transfer direction oriented along the z axis, the Eqs. 1.24 and 1.25 can be simplified as:

$$\langle\beta^2_{ZZZ}\rangle = \frac{1}{7}\beta^2_{zzz} + \frac{6}{35}\beta_{zzz}\beta_{zxx} + \frac{9}{35}\beta^2_{zxx} \tag{1.27}$$

$$\langle\beta^2_{ZXX}\rangle = \frac{1}{35}\beta^2_{zzz} - \frac{2}{105}\beta_{zzz}\beta_{zxx} - \frac{38}{105}\beta^2_{zxx} \tag{1.28}$$

For a system where the longitudinal β response strongly dominates ($\beta_{zzz} >> \beta_{zxx}$), $DR = 5$, and the system has a 1D response. For an octupolar molecule with T_d symmetry, the only non-zero tensor component is β_{xyz}, so that:

$$\langle\beta^2_{ZZZ}\rangle = \frac{12}{35}\beta^2_{xyz} \qquad \langle\beta^2_{ZXX}\rangle = \frac{8}{35}\beta^2_{xyz} \tag{1.29}$$

and therefore $DR = 1.5$.

It is worth mentioning that, for systems in solution, both the solute and the solvent can contribute to the optical response, and the total response is the sum of these two contributions:

$$I_{2\omega} \propto (C_s\langle\beta^2_{HRS}\rangle_s + C_x\langle\beta^2_{HRS}\rangle_x)I^2_\omega, \tag{1.30}$$

where C_s is the concentration of the solvent, C_x the concentration of the solute, while $\langle\beta^2_{HRS}\rangle_s$ and $\langle\beta^2_{HRS}\rangle_x$ are their respective HRS signals. The response of the solute is then obtained by excluding the contributions of the solvent. This calibration can be made by two different approaches. In the first, defined as the *internal reference* method, $\langle\beta^2_{HRS}\rangle_x$ is obtained by performing the measurements at different concentrations of the solute, while in the second one, the "*external reference*" method, the solvent contribution $\langle\beta^2_{HRS}\rangle_s$ is measured independently.

The HRS method is generally preferred to EFISHG because it can be straightforwardly applied to measure SHG on charged and octupolar molecules. However, since it is an incoherent process with no collective enhancement of the NLO response by means of an artificial orientation of the molecules, the intensity is low and high-sensitivity detection tools are necessary.

Bibliography

[1] Kleinman, D. A. Nonlinear Dielectric Polarization in Optical Media. *Phys. Rev.* **1962**, *126*, 1977–1979.

[2] Neumann, F. E.; Meyer, O. E. *Vorlesungen über die Theorie der Elasticität der festen Körper und des Lichtäthers* ; Teubner, Leipzig, 1885.

[3] Verbiest, T.; Clays, K.; V., R. *Second-order Nonlinear Optical Characterization Techniques* ; CRC Press, 2009.

[4] Medved', M.; Stachová, M.; Jacquemin, D.; André, J.; Perpète, E. A. A generalized Romberg differentiation procedure for calculation of hyperpolarizabilities. *J. Mol. Struct.* **2007**, *847*, 39–46.

[5] Franken, P. A.; Hill, A. E.; Peters, C. W.; Weinreich, G. Generation of Optical Harmonics. *Phys. Rev. Lett.* **1961**, *7*, 118–119.

[6] Freund, I.; Deutsch, M. Second-harmonic microscopy of biological tissue. *Opt. Lett.* **1986**, *11*, 94–96.

[7] Dalton, L. R. et al. Organic electro-optic materials. Optical Materials in Defence Systems Technology. 2004; pp 93 – 104.

[8] Papadopoulos, M. G.; Sadlej, A. J.; Leszczynsk, E., J. Non-Linear Optical Properties of Matter, from Molecules to Condensed Phases. 2006.

[9] Stegeman, G. I.; Stegeman, R. A. Nonlinear Optics: Phenomena, Materials, and Devices; Boreman, G. 2012.

[10] Kanis, D. R.; Ratner, M. A.; Marks, T. J. Design and construction of molecular assemblies with large second-order optical nonlinearities. Quantum chemical aspects. *Chem. Rev.* **1994**, *94*, 195–242.

[11] Marder, S. R.; Beratan, D. N.; Cheng, L.-T. Approaches for Optimizing the First Electronic Hyperpolarizability of Conjugated Organic Molecules. *Science* **1991**, *252*, 103–106.

[12] Würthner, F.; Effenberger, F.; Wortmann, R.; Krämer, P. Second-order polarizability of donor—acceptor substituted oligothiophenes: substituent variation and conjugation length dependence. *Chem. Phys.* **1993**, *173*, 305–314.

[13] Plaquet, A.; Champagne, B.; Kulhánek, J.; Bureš, F.; Bogdan, E.; Castet, F.; Ducasse, L.; Rodriguez, V. Effects of the Nature and Length of the π-Conjugated Bridge on the Second-Order Nonlinear Optical Responses of Push–Pull Molecules Including 4,5-Dicyanoimidazole and Their Protonated Forms. *ChemPhysChem* **2011**, *12*, 3245–3252.

[14] Oudar, J. L.; Chemla, D. S. Hyperpolarizabilities of the nitroanilines and their relations to the excited state dipole moment. *J. Chem. Phys.* **1977**, *66*, 2664–2668.

[15] Bourhill, G.; Bredas, J.; Cheng, L.; Marder, S. R.; Meyers, F.; Perry, J. W.; Tiemann, B. G. Experimental Demonstration of the Dependence of the First Hyperpolarizability of Donor-Acceptor-Substituted Polyenes on the Ground-State Polarization and Bond Length Alternation. *J. Am. Chem. Soc.* **1994**, *116*, 2619–2620.

[16] Marder, S. R.; Cheng, L.-T.; Tiemann, B. G.; Friedli, A. C.; Blanchard-Desce, M.; Perry, J. W.; Skindhøj, J. Large First Hyperpolarizabilities in Push-Pull Polyenes by Tuning of the Bond Length Alternation and Aromaticity. *Science* **1994**, *263*, 511–514.

[17] Gorman, C. B.; Marder, S. R. Effect of Molecular Polarization on Bond-Length Alternation, Linear Polarizability, First and Second Hyperpolarizability in Donor-Acceptor Polyenes as a Function of Chain Length. *Chem. Mater.* **1995**, *7*, 215–220.

[18] Zyss, J. Molecular engineering implications of rotational invariance in quadratic nonlinear optics: From dipolar to octupolar molecules and materials. *J. Chem. Phys.* **1993**, *98*, 6583–6599.

[19] Bredas, J. L.; Meyers, F.; Pierce, B. M.; Zyss, J. On the second-order polarizability of conjugated .pi.-electron molecules with octupolar symmetry: the case of triaminotrinitrobenzene. *J. Am. Chem. Soc.* **1992**, *114*, 4928–4929.

[20] Campo, J.; Painelli, A.; Terenziani, F.; Van Regemorter, T.; Beljonne, D.; Goovaerts, E.; Wenseleers, W. First Hyperpolarizability Dispersion of the Octupolar Molecule Crystal Violet: Multiple Resonances and Vibrational and Solvation Effects. *J. Am. Chem. Soc.* **2010**, *132*, 16467–16478.

[21] Castet, F.; Blanchard-Desce, M.; Adamietz, F.; Poronik, Y. M.; Gryko, D. T.; Rodriguez, V. Experimental and Theoretical Investigation of the First-Order Hyperpolarizability of Octupolar Merocyanine Dyes. *ChemPhysChem* **2014**, *15*, 2575–2581.

[22] Beaujean, P.; Sanguinet, L.; Rodriguez, V.; Castet, F.; Champagne, B. Multi-State Second-Order Nonlinear Optical Switches Incorporating One to Three Benzazolo-Oxazolidine Units: A Quantum Chemistry Investigation. *Molecules* **2022**, *27*.

[23] Coe, B. J. Molecular Materials Possessing Switchable Quadratic Nonlinear Optical Properties. *Chem. Eur. J.* **1999**, *5*, 2464–2471.

[24] Wegner, H. A. Molecular Switches. Second Edition. Edited by Ben L. Feringa and Wesley R. Browne. *Angew. Chem.* **2012**, *51*, 2281–2281.

[25] Castet, F.; Rodriguez, V.; Pozzo, J.; Ducasse, L.; Plaquet, A.; Champagne, B. Design and Characterization of Molecular Nonlinear Optical Switches. *Acc. Chem. Res.* **2013**, *46*, 2656–2665.

[26] Sanguinet, L.; Pozzo, J.; Guillaume, M.; Champagne, B.; Castet, F.; Ducasse, L.; Maury, E.; Soulié, J.; Mançois, F.; Adamietz, F.; Rodriguez, V. Acidoswitchable NLO-phores: Benzimidazolo[2,3-b]oxazolidines. *J. Phys. Chem. B* **2006**, *110*, 10672–10682.

[27] Sliwa, M.; Spangenberg, A.; Malfant, I.; Lacroix, P. G.; Métivier, R.; Pansu, R. B.; Nakatani, K. Structural, Optical, and Theoretical Studies of a Thermochromic Organic Crystal with Reversibly Variable Second Harmonic Generation. *Chem. Mater.* **2008**, *20*, 4062–4068.

[28] Boubekeur-Lecaque, L.; Coe, B. J.; Clays, K.; Foerier, S.; Verbiest, T.; Asselberghs, I. Redox-Switching of Nonlinear Optical Behavior in Langmuir−Blodgett Thin Films Containing a Ruthenium(II) Ammine Complex. *J. Am. Chem. Soc.* **2008**, *130*, 3286–3287.

[29] Berkovic, G.; Krongauz, V.; Weiss, V. Spiropyrans and Spirooxazines for Memories and Switches. *Chem. Rev.* **2000**, *100*, 1741–1754.

[30] Delaire, J. A.; Nakatani, K. Linear and Nonlinear Optical Properties of Photochromic Molecules and Materials. *Chem. Rev.* **2000**, *100*, 1817–1846.

[31] Priimagi, A.; Ogawa, K.; Virkki, M.; Mamiya, J.; Kauranen, M.; Shishido, A. High-Contrast Photoswitching of Nonlinear Optical Response in Crosslinked Ferroelectric Liquid-Crystalline Polymers. *Adv. Mater.* **2012**, *24*, 6410–6415.

[32] Schulze, M.; Utecht, M.; Moldt, T.; Przyrembel, D.; Gahl, C.; Weinelt, M.; Saalfrank, P.; Tegeder, P. Nonlinear optical response of photochromic azobenzene-functionalized self-assembled monolayers. *Phys. Chem. Chem. Phys.* **2015**, *17*, 18079–18086.

[33] Tonnelé, C.; Champagne, B.; Muccioli, L.; Castet, F. Nonlinear Optical Contrast in Azobenzene-Based Self-Assembled Monolayers. *Chem. Mater.* **2019**, *31*, 6759–6769.

[34] Ledoux, I.; Zyss, J.; Barni, E.; Barolo, C.; Diulgheroff, N.; Quagliotto, P.; Viscardi, G. Properties of novel azodyes containing powerful acceptor groups and thiophene moiety. *Synth. Met.* **2000**, *115*, 213–217.

[35] Krawczyk, P.; Kaczmarek, A.; Zaleśny, R.; Matczyszyn, K.; Bartkowiak, W.; Ziółkowski, M.; Cysewski, P. Linear and nonlinear optical properties of azobenzene derivatives. *J. Mol. Model.* **2009**, *15*, 581—590.

[36] Pantazis, P.; Maloney, J.; Wu, D.; Fraser, S. E. Second harmonic generating (SHG) nanoprobes for in vivo imaging. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14535–14540.

[37] Campagnola, P. J. *3D SHG imaging and analysis of fibrillar collagen, Second harmonic generation imaging*; Boca Raton, London, New York: CRC Press, Taylor & Francis Group, 2013; p 125.

[38] Yokota, H.; Kaneshiro, J.; Uesu, Y. Optical second harmonic generation microscopy as a tool of material diagnosis. *Phys. Res. Int.* **2012**, *2012*, 704634.

[39] Kwan, A. C. *Second-harmonic generation imaging of microtubules, Second harmonic generation imaging*; Boca Raton, London, New York: CRC Press, Taylor & Francis Group, 2013; p 151.

[40] Reeve, J. E.; Anderson, H. L.; Clays, K. Dyes for biological second harmonic generation imaging. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13484–13498.

[41] Campagnola, P. J.; Wei, M. D.; Loew, L. M. High-Resolution Nonlinear Optical Imaging of Live Cells by Second Harmonic Generation. *Biophys. J.* **1999**, *77*, 3341–3349.

[42] Moreaux, L.; Sandre, O.; Charpak, S.; Blanchard-Desce, M.; Mertz, J. Coherent Scattering in Multi-Harmonic Light Microscopy. *Biophys. J.* **2001**, *80*, 1568–1574.

[43] Dombeck, D. A.; Blanchard-Desce, M.; Webb, W. W. Optical Recording of Action Potentials with Second-Harmonic Generation Microscopy. *J. Neurosci.* **2004**, *24*, 999–1003.

[44] Terhune, R. W.; Maker, P. D.; Savage, C. M. Optical harmonic generation in calcite. *Phys. Rev. Lett.* **1962**, *8*, 2550.

[45] Finn, R. S.; Ward, J. F.; Ward, F.; Finn, R. S. DC-induced optical second-harmonic generation in the inert gases. *Phys. Rev. Lett.* **1971**, *26*, 285.

[46] Hauchecorne, G.; Kerhervé, F.; Mayer, G. Measurement of light wave interactions in various substances. *J. Phys. (France)* **1971**, *32*, 47.

[47] Levine, B. F.; Bethea, C. G. Molecular hyperpolarizabilities determined from conjugated and nonconjugated organic liquids. *App. Phys. Lett.* **1974**, *24*, 445.

[48] Levine, B. F.; Bethea, C. G. Second and third order hyperpolarizabilities of organic molecules. *J. Chem. Phys.* **1975**, *63*, 2666.

[49] Pielak, K.; Tonnelé, C.; Sanguinet, L.; Cariati, E.; Righetto, S.; Muccioli, L.; Castet, F.; Champagne, B. Dynamical behavior and second harmonic generation responses in acido-triggered molecular switches. *J. Phys. Chem. C* **2018**, *122*, 26160.

[50] Alain, V.; Blanchard-Desce, M.; Ledoux-Rak, I.; Zyss, J. Amphiphilic polyenic push–pull chromophores for nonlinear optical applications. *Chem. Commun.* **2000**, 353–354.

[51] Ramos, T. N.; Canuto, S.; Champagne, B. Unraveling the Electric Field-Induced Second Harmonic Generation Responses of Stilbazolium Ion Pairs Complexes in Solution Using a Multiscale Simulation Method. *J. Chem. Inf. Model.* **2020**, *60*, 4817–4826.

[52] Ramos, T. N.; Castet, F.; Champagne, B. Second Harmonic Generation Responses of Ion Pairs Forming Dimeric Aggregates. *J. Phys. Chem. B* **2021**, *125*, 3386–3397.

[53] Terhune, R. W.; Maker, P. D.; Savage, C. M. Measurements of Nonlinear Light Scattering. *Phys. Rev. Lett.* **1965**, *14*, 681–684.

[54] Bersohn, R.; Pao, Y.; Frisch, H. L. Double-Quantum Light Scattering by Molecules. *J. Chem. Phys.* **1966**, *45*, 3184–3198.

Chapter 2

Theoretical methods

2.1 The many-body problem

The main goal of computational chemistry is to investigate and predict chemical and physical properties of molecules and materials. This objective requires an accurate description of the interactions of electrons and nuclei in atoms and molecules which could be guaranteed only by a proper treatment of quantum mechanical effects. The position and the time evolution of a system are described at the quantum mechanical level by the wavefunction Ψ_t. In order to determine it, it is necessary to solve the so-called time-dependent Schrödinger equation, namely:

$$\hat{H}\Psi_t = i\frac{\partial}{\partial t}\Psi_t \tag{2.1}$$

The quantity $\hat{H}$ is the Hamiltonian of the system which describes the interactions among electrons and nuclei as well as with the external environment. The first step toward the solution of Eq. 2.1 is to find the so-called stationary states which describe the system at a specific time ($t_0 = 0$ for convenience), and then compute its time-evolution. If the Hamiltonian is time independent, the stationary states are found from the solution of the time-independent non-relativistic Schrödinger equation, namely,

$$\hat{H}\Psi_0 = E\Psi_0. \tag{2.2}$$

If the system is isolated, the Hamiltonian is composed of two operators, the kinetic energy ($\hat{T}$) and the Coulomb potential ($\hat{V}$),

$$\hat{H} = \hat{T} + \hat{V}. \tag{2.3}$$

$\hat{T}$ can be separated into electronic ($\hat{T}_e$) and nuclear contributions ($\hat{T}_N$). Therefore, for a system of n electrons and M nuclei it can be expressed as:

$$\hat{T} = \hat{T}_e + \hat{T}_N = -\sum_{i}^{n}\frac{\nabla_i^2}{2} - \sum_{I}^{M}\frac{\nabla_I^2}{2m_I}, \tag{2.4}$$

where the summation on lowercase letters runs over the electrons, the summation on uppercase letters runs over nuclei, and m_I stays for the mass of the nucleus I. The Coulomb potential can be decomposed in three terms: the attraction between electrons and nuclei ($\hat{V}_{eN}$), the Coulomb repulsion between electrons $\hat{V}_{ee}$, and the repulsion between nuclei ($\hat{V}_{NN}$). The expression of these terms is collected in the following Eq.:

$$\hat{V} = \hat{V}_{eN} + \hat{V}_{ee} + \hat{V}_{NN} = -\sum_{i,I}^{n,M} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_{i>j}^{n} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{I>J}^{M} \frac{Z_I Z_J}{|\mathbf{R}_J - \mathbf{R}_J|} \tag{2.5}$$

where $\mathbf{R}_I$ identifies the position of the nucleus I, $\mathbf{r}_i$ of the electron i, while Z_I indicates the charges of the nucleus. Unfortunately, an exact analytic solution of Eq. 2.2 is possible just for systems of one electron and one nucleus; for more complicated systems it is necessary to find appropriate approximations. During the last century, several methods have been developed to find approximated solutions of Eq. 2.2 able to reproduce accurately chemical and physical phenomena with an affordable computational cost. The methods for solving Eq. 2.2 that do not employ empirical parameters are called *ab initio*. One of the most basic approximations is the Born-Oppenheimer (BO) approximation [1] which consists in decoupling the wavefunction in nuclear and electronic contributions. Indeed, this approximation is based on the fact that the nuclei are much heavier than the electrons and therefore slower. Following BO, the wavefunction of Eq. 2.2 could be decoupled as:

$$\Psi(\mathbf{r}, \mathbf{R}) = \psi(\mathbf{r}, \mathbf{R})\phi(\mathbf{R}) \tag{2.6}$$

According to this approximation, we can identify the electronic part of the Hamiltonian, where the positions of the nuclei are parameters. Therefore, in order to determine the electronic contribution in terms of nuclei positions, we solve the following eigenvalues equation:

$$\hat{H}_e \psi(\mathbf{r}, \mathbf{R}) = E_e(\mathbf{R})\psi(\mathbf{r}, \mathbf{R}), \tag{2.7}$$

Where $\hat{H}_e$ is:

$$\hat{H}_e = \hat{T}_e + \hat{V}_{NN} + \hat{V}_{eN} + \hat{V}_{ee}. \tag{2.8}$$

A similar equation is obtained for the nuclei, which lie in the potential energy surface (PES) set up by the electrons. Although the BO approximation is a crude approximation, it is in general very accurate for most of systems. However, in some cases, it is necessary to include explicitly the interactions of electrons and nuclei due to the non-adiabatic effects, for instance in the case of conical intersections [2]. In order to solve Eq. 2.7, it is necessary to find appropriate approximations to the actual wavefunction ψ. The wavefunction of a system of N electrons can be expressed as a product of n molecular spin-orbitals. Each molecular spin-orbital is composed of a molecular orbital (MO) $\phi(\mathbf{r})$, which gives the location of the electrons in the molecule, and a spin function $\chi(\sigma_i)$, which, considering that the electrons are fermions, has two possible values for identifying the spin-up or the spin-down states. Moreover, according to the Pauli principle, the wavefunction of the system must be antisymmetric with respect to the exchange of the coordinates of any pair of electrons.

2.2 Hartree Fock method

For solving the electronic structure problem the easiest way is to employ the independent particle model, in which the instantaneous interactions between electrons are replaced by an average field generated by all the electrons. In particular, in the Hartree Fock (HF) method, each electron is described by a molecular orbital and for a system of N electrons, the wavefunction is made by an antysimmetrized product of N molecular orbitals. This ensures the fulfillment of the Pauli principle for the total wavefunction. The resulting wavefunction is called Slater Determinant and its expression is:

$$\psi^{\mathrm{HF}}(\mathbf{x_1},...\mathbf{x_N}) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \tilde{\phi}_1(\mathbf{x_1}) & \tilde{\phi}_2(\mathbf{x_1}) & ... & \tilde{\phi}_N(\mathbf{x_1}) \\ \tilde{\phi}_1(\mathbf{x_2}) & \tilde{\phi}_2(\mathbf{x_2}) & ... & \tilde{\phi}_N(\mathbf{x_2}) \\ ... & ... & ... & ... \\ \tilde{\phi}_1(\mathbf{x_N}) & \tilde{\phi}_2(\mathbf{x_N}) & ... & \tilde{\phi}_N(\mathbf{x_N}) \end{vmatrix} \tag{2.9}$$

where $\tilde{\phi}(\mathbf{x}_i)$ is the product of a spatial orbital $\phi(\mathbf{r}_i)$ and a spin function $\chi(\sigma_i)$. The factor $1/\sqrt{N!}$ is a normalization coefficient. The energy of this wavefunction can be found by minimizing the expectation value of the Hamiltonian of Eq. 2.10. Indeed, according to the variational principle, the best wavefunction of the system is the one that minimizes the energy (see Eq. 2.10).

$$E^{HF} = \langle \psi^{HF}_{trial} | \hat{H} | \psi^{HF}_{trial} \rangle \geq E_{exact} \tag{2.10}$$

By including Eq. 2.9 in Eq. 2.10, we can obtain the expression of the HF energy, namely:

$$E^{HF} = \sum_i h_i + \frac{1}{2} \sum_{i,j} (J_{ij} - K_{ij}), \tag{2.11}$$

where the quantity h_i is the independent particle energy:

$$h_i = \langle \phi_i | - \frac{\nabla_i^2}{2} + \sum_I \frac{Z_I}{r_{Ii}} | \phi_i \rangle \tag{2.12}$$

J_{ij} accounts for the repulsion between the electrons, and K_{ij} is the so-called exchange term that arises because the wavefunction is antisymmetric. The expression of these integrals are:

$$\begin{aligned} J_{ij} &= \langle \phi_i \phi_j | \frac{1}{r_{ij}} | \phi_i \phi_j \rangle \\ K_{ij} &= \langle \phi_i \phi_j | \frac{1}{r_{ij}} | \phi_j \phi_i \rangle \end{aligned} \tag{2.13}$$

By minimizing the Eq. 2.11 with respect to the spin-orbitals ϕ we obtain the so-called HF eigenvalues equations, namely:

$$\hat{f}_i \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}) \tag{2.14}$$

where $\hat{f}$ is a one-electron operator called Fock operator, while ϵ_i are the single-particle energies associated to each orbital. These quantities have in general no physical interpretations, with the exceptions of the energies corresponding to the HOMO (ϵ_{HOMO}) and LUMO (ϵ_{LUMO}) which according to the Koopman's theorem [3] can be associated respectively to the ionization potential and electron affinity. This approach is defined as restricted HF because in this formulation electrons with different spins have an identical molecular orbital which is multiplied by the corresponding spin function. The Fock operator has the following expression:

$$\hat{f}_i = \hat{h}_i + \sum_j^m (\hat{J}_j - \hat{K}_j) \tag{2.15}$$

where the quantities $\hat{J}_j$ and $\hat{K}_j$ are the Coulomb and exchange operators, which are defined as:

$$\begin{aligned} \hat{J}_j \phi_i &= \left(\langle \phi_j | \frac{1}{r_{ij}} | \phi_j \rangle \right) \phi_i \\ \hat{K}_j \phi_i &= \left(\langle \phi_j | \frac{1}{r_{ij}} | \phi_i \rangle \right) \phi_j \end{aligned} \tag{2.16}$$

The expressions in Eq. 2.16 show that each orbital feels the presence on average of all the orbitals.

2.2.1 Unrestricted approach

For the description of open-shell systems, the restricted Hartree Fock framework in which electrons with different spins occupy the same orbitals is not appropriate. The standard way to overcome this problem is through the so-called unrestricted HF method (UHF) which consists in computing two different sets of equations 2.14 for electrons with different spins (α and β), namely:

$$\hat{f}_i^\alpha \phi_i^\alpha(\mathbf{r}) = \epsilon_i^\alpha \phi_i^\alpha(\mathbf{r}) \qquad \hat{f}_i^\beta \phi_i^\beta(\mathbf{r}) = \epsilon_i^\beta \phi_i^\beta(\mathbf{r}) \tag{2.17}$$

This method increases the accuracy of HF calculations for which the RHF picture is completely wrong (such as for the study of bond dissociations), and allows the treatment of open-shell systems. However, the UHF wavefunction is not anymore an eigenfunction of the spin operator $\hat{S}^2$, bringing to the system the so-called "spin-contamination". Therefore, even though through the unrestricted approach the energy could be improved, the wavefunction is generally unreliable, affecting particularly the computation of molecular properties. The amount of spin contamination can be easily monitored by computing

$< \hat{S}^2 >$, which identifies the average value of the $\hat{S}^2$ operator.

2.2.2 Basis sets approximation

For solving the HF equations it is convenient to use a basis of functions for which the integrals of equations 2.15 are easy to calculate. In principle, any kind of function can be used, however, in Quantum Chemistry the most common approximation is the so-called linear combination of atomic orbitals (LCAO), which consists in expanding the molecular orbitals in a M_b basis of functions centered into the positions of the nuclei ($\{\psi_\mu\}$), namely:

$$\phi_i = \sum_{\mu}^{M_b} c_{\mu i} \psi_\mu. \tag{2.18}$$

Therefore, by including Eq. 2.18 in Eq. 2.14 and by projecting to the set of basis functions, we obtain a set of equations which could be expressed in matrix notation as:

$$\mathbf{FC} = \mathbf{SC}\epsilon, \tag{2.19}$$

Where **F** and **S** are respectively the matrix representation of the Fock operator in the atomic basis set and the overlap matrix, with respectively elements:

$$F_{\mu\nu} = \langle \psi_\mu | \hat{f} | \psi_\nu \rangle \qquad S_{\mu\nu} = \langle \psi_\mu | \psi_\nu \rangle \tag{2.20}$$

Eq. 2.19 are called the Roothan (Poble-Nesbet-Berthier equations for the unrestricted approach) equations and can be solved iteratively by the so-called Self Consistent Procedure (SCF). The selection of the basis set is arbitrary, however, in general, the most common choices are the Slater Types Orbitals (STOs) and the Gaussian Type Orbitals (GTOs). STOs integrals reproduce the correct electron-nuclear cusp conditions at short distance ($\propto e^{-r}$) and therefore, they reproduce the exact behavior of the wavefunction. However, the computation of two-electron integrals of Eq. 2.16 with STOs basis require numerical integration and therefore they are more challenging and time-consuming. On the other hand, GTOs functions are Gaussians ($\propto e^{-r^2}$) and, therefore, the resulting integrals can be computed analytically. These basis functions can be further improved for tackling challenging systems, for example, we have the so-called diffuse functions which are designed for taking into account long-range interactions, and polarization functions which increase the accuracy in bond regions. In this work, we employed just GTOs basis sets, in particular, testing two different basis sets schemes the Pople's [4] basis sets and the Dunning's [5], where the second ones have been designed mostly for post-Hartree Fock methods.

2.3 Post-HF methods

Employing the HF approximation and a large basis set it is possible to reproduce up to ∼99% of the exact energy of most of molecular systems, however, a reproduction of the remaining ∼1% is fundamental to properly understanding the chemistry of the system. This energy difference has been firstly defined by Löwdin [6] as electron correlation energy, namely:

$$E^{corr} = E^{exact} - E^{HF} \tag{2.21}$$

The source of this discrepancy is related to the neglect of the instant interactions between electrons due to the mean-field approach. There are different ways of classifying electron correlation, and this is strictly related to the system and the context that we are investigating. One example of correlation effect is the Fermi repulsion of electrons which is due to their fermionic nature. It is automatically taken into account in the HF method by the Slater determinant wavefunction. In fact, in the HF method two electrons with the same spin can not occupy the same position because of the antisymmetry of the wavefunction. However, we will see that this is not generally true for DFT approximations. Moreover, the so-called Coulomb correlation refers to the fact that the electrons repel each other because of their electrostatic interactions. In the HF model, we do not observe this correlation between electrons with opposite spins because of the nature of the $\hat{f}^{HF}$ operator: indeed, if two electrons have opposite spins they can occupy the same position in space. The most relevant way of classifying correlation effects is by the separation between the *dynamic* and *nondynamic* correlation. The name dynamic correlation is referred to the motion of electrons and it is important in systems for which the ground state is generally well described by the single determinant wavefunction and could be improved just by small corrections to it. On the other hand, nondynamic (or static) correlation arises when two or more configurations are needed to properly describe a system, in this case, the HF picture is completely wrong and more sophisticated methods are needed to be employed. In the last decades, several different methods have been developed to face these problems. Although there is not a clear-cut separation between dynamic and nondynamic correlation, in general, quantum chemistry methods can include one or the other type of correlation. Methods for treating dynamic correlation effects are known as post-HF methods and include: the Configuration Interaction (CI), the Møller Plesset Perturbation theory (MPn) and the Coupled Clusters approximations. We will introduce these methods in the following subsections. On the other hand, for reproducing nondynamic correlation effects, it is necessary to employ specific methods that take into account degeneracies or pseudo-degeneracies of orbital energies such as the Multi Configuration Self Consistent Field (MCSCF) or the Quantum Monte Carlo methods. These methods will not be studied in this thesis because we did not include in our studies systems in which nondynamic correlation is relevant.

2.3.1 Configuration Interaction

The conceptually easiest way to overcome the single determinant HF formulation is to build a wavefunction made of several determinants. The full configuration interaction method consists in minimizing the expansion coefficients following the variational principle. A preliminary HF calculation is performed, and a set of *excited* determinants are constructed by exciting electrons from the HF occupied orbitals to the virtual ones. According to the number of orbitals that we excite with respect to the HF configuration, we can separate the determinants into singly, doubly, triply, etc.. excited determinants. The resulting wavefunction could be expressed as:

$$\Psi = a_0\Phi_{HF} + \sum_{S} a_S\Phi_S + \sum_{D} a_D\Phi_D + \sum_{T} a_T\Phi_T + ... \tag{2.22}$$

If m is the number of orbitals and N is the number of electrons, the resulting number of determinants in the restricted formulation will be $\binom{m}{N/2}$. Unfortunately, computing and storing all these determinants and diagonalizing the resulting CI matrix is prohibitive even for the most modern computational resources, and therefore the full CI (FCI) method can be applied only to systems of few electrons. On the other hand, if the HF calculation is already a good approximation we can decide to reduce the number of determinants by truncating the CI expansion considering just the most important excitations. Therefore, by truncating the CI expansion to singles(S), doubles(D) or triples(T), etc. We will obtain, respectively, the methods CIS, CISD and CISDT. Although these methods allow to save computational time with respect to FCI, they have some limitations that may affect their performances. For example, the CISD method scales as $\binom{N_{occ}}{2}\binom{N_{vir}}{2}$, however, it is neither size consistent nor size extensive, and therefore it might give misleading results for example in calculations of barriers, thermodynamic or dissociation energies.

2.3.2 Rayleigh–Schrödinger perturbation theory

An alternative way to include correlation in a many-body system is through perturbation theory. If the Hamiltonian of a system can be decoupled into a reference Hamiltonian (for which the solution of the Schrödinger equation is known), and a small perturbation, the energy can be found as a perturbative expansion in terms of a coupling parameter. Indeed, the total Hamiltonian of the system can be written as:

$$\hat{H}^{\lambda} = \hat{H}^{(0)} + \lambda\hat{V}, \tag{2.23}$$

$\hat{H}_0$ is the unperturbed Hamiltonian, $\hat{V}$ is the perturbation and λ is the coupling parameter. Therefore, the Schrödinger Equation is expressed as:

$$\hat{H}^{\lambda}\psi_n^{\lambda} = E_n^{\lambda}\psi_n^{\lambda}, \tag{2.24}$$

where the energy and the wavefunction can be expressed as a sum of different perturbative terms:

$$E_n^{\lambda} = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + ... \tag{2.25}$$

$$\psi_n^{\lambda} = \psi_n^{(0)} + \lambda \psi_n^{(1)} + \lambda^2 \psi_n^{(2)} + ..., \tag{2.26}$$

and $\psi_n^{(0)}$ is the wavefunction of the unperturbed Hamiltonian $\hat{H}_n^{(0)}$:

$$\hat{H}_n^{(0)} \psi_n^{(0)} = E_n^{(0)} \psi_n^{(0)} \tag{2.27}$$

To find the solution of Eq. 2.24 it is convenient to impose normalization constraints as:

$$\langle \psi_n^{(0)} | \psi_n^{(0)} \rangle = 1 \tag{2.28}$$

$$\langle \psi_n^{(0)} | \psi_n^{(i)} \rangle = 0 \qquad i = 1, 2, ... \tag{2.29}$$

Substituing Eq. 2.26 in Eq. 2.24, equating the coefficients of the same power of λ, and then multiplying by $\langle \psi_n^{(0)}|$ we obtain the expression of the perturbed energies:

$$E_n^{(0)} = \langle \psi_n^{(0)} | \hat{H}^{(0)} | \psi_n^{(0)} \rangle \tag{2.30}$$

$$E_n^{(1)} = \langle \psi_n^{(0)} | \hat{V} | \psi_n^{(0)} \rangle \tag{2.31}$$

$$E_n^{(2)} = \langle \psi_n^{(0)} | \hat{V} | \psi_n^{(1)} \rangle \tag{2.32}$$

$$E_n^{(3)} = \langle \psi_n^{(0)} | \hat{V} | \psi_n^{(2)} \rangle \tag{2.33}$$

$$...$$

To find the perturbed wavefunctions $\langle \psi_n^{(1)}|$, $\langle \psi_n^{(2)}|$, ... we need to expand the solution in a set of known basis functions, which for convenience is usually $\{\psi_n^{(0)}\}$. The first two perturbations of the wavefunction can be thus written as:

$$\psi_n^{(1)} = \sum_m \langle \psi_m^{(0)} | \psi_n^{(1)} \rangle \, \psi_m^{(0)} \tag{2.34}$$

$$\psi_n^{(2)} = \sum_m \langle \psi_m^{(0)} | \psi_n^{(2)} \rangle \, \psi_m^{(0)} \tag{2.35}$$

In order to find the second-order perturbed energy, we can substitute Eq. 2.34 in Eq. 2.32, obtaining:

$$E_n^{(2)} = \langle \psi_n^{(0)} | \hat{V} | \psi_n^{(1)} \rangle = \sum_m{}' \frac{\langle \psi_n^{(0)} | \hat{V} | \psi_m^{(0)} \rangle \, \langle \psi_m^{(0)} | \hat{V} | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \tag{2.36}$$

where $\sum'$ indicates that the terms $m = n$ are excluded from the summation. Similarly we can obtain the third-order expression:

$$E_n^{(3)} = \sum_{pq}{}' \frac{\langle \psi_n^{(0)} | \hat{V} | \psi_p^{(0)} \rangle \langle \psi_p^{(0)} | \hat{V} | \psi_q^{(0)} \rangle \langle \psi_q^{(0)} | \hat{V} | \psi_n^{(0)} \rangle}{(E_n^{(0)} - E_p^{(0)})(E_n^{(0)} - E_q^{(0)})} - \langle \psi_n^{(0)} | \hat{V} | \psi_n^{(0)} \rangle \sum_{p}{}' \frac{| \langle \psi_n^{(0)} | \hat{V} | \psi_p^{(0)} \rangle |^2}{(E_n^{(0)} - E_p^{(0)})^2} \tag{2.37}$$

2.3.3 Møller-Plesset Perturbation Theory

The Møller-Plesset (MP) theory is a specific application of the RS perturbation theory. In this case, the ψ_0 is the HF wavefunction, therefore, the unperturbed Hamiltonian can be expressed as the sum of the 1-particle Hamiltonian and the HF potential:

$$\hat{H}^{(0)} = \sum_i \hat{f}_i = \sum_i [\hat{h}_i + \hat{v}_i^{HF}] \tag{2.38}$$

where the operator $\hat{v}_i^{HF}$ is the Fock potential and $\hat{f}_i$ is defined by solving Eq. 2.14. The perturbation operator is:

$$\hat{V} = \sum_{i<j} \frac{1}{r_{ij}} - \sum_i \hat{v}_i^{HF} \tag{2.39}$$

The expression of the averaging at the zeroth and first order of perturbation are therefore:

$$E^{\mathrm{MP0}} = \sum_i \epsilon_i \tag{2.40}$$

$$E^{\mathrm{MP1}} = \sum_i \epsilon_i - \frac{1}{2} \sum_{i,j} \langle \phi_i \phi_j || \phi_i \phi_j \rangle = E^{HF}, \tag{2.41}$$

where $\langle \phi_i \phi_j || \phi_i \phi_j \rangle = J_{ij} - K_{ij}$.

In this formulation, the zero-th order excited states wavefunction $\{\psi_n^{(0)}\}$, are created by combination of excited determinants obtained by promoting electrons from occupied to virtual orbitals. According to the number of electrons that we excite we can obtain single $\{\psi_a^r\}$, double $\{\psi_{ab}^{rs}\}$, ... excited determinants. The subscript indicates the occupied orbitals from which the electrons are removed and the superscript the virtual ones to which they are moved. However, considering that that $\hat{V}$ is a two-particle operator, triple and higher excitations do not mix with $\psi_0^{(0)}$. Moreover, because the orbitals are orthonormal we can apply the Brillouin theorem obtaining null contribution also from integrals computing the reference HF configuration to single excitations, namely:

$$\langle \psi_0^{(0)} | \hat{V} | \psi_a^r \rangle = 0. \tag{2.42}$$

Therefore, we can only have contributions from the double excitations $\{\psi_{ab}^{rs}\}$. Considering

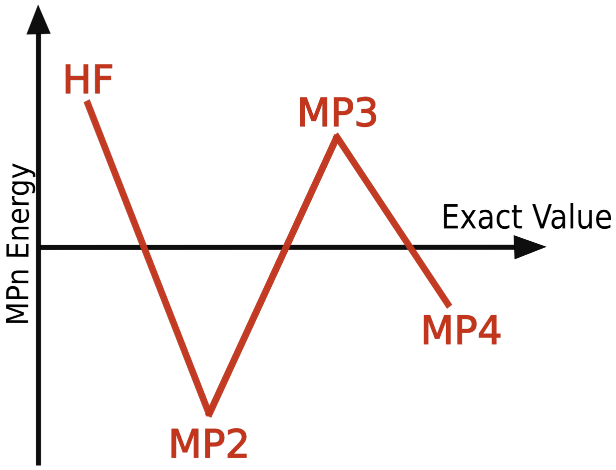


Figure 2.1: Schematic representation of the oscillatory behavior of MPn energy calculations.

that the contribution of the doubly excited energy is:

$$E_{ab}^{rs} = E^0 + \epsilon_r + \epsilon_s - \epsilon_a - \epsilon_b. \tag{2.43}$$

The expression of the second-order correction to the energy is:

$$E^{MP2} = E^{HF} - \sum_{a<b,r<s} \frac{|\langle ab||rs\rangle|^2}{\epsilon_r + \epsilon_s - \epsilon_a - \epsilon_b} \tag{2.44}$$

Where the sum on a, b run over occupied orbitals and the sum on r, s runs over virtual ones. Similarly to Eq. 2.44, we can obtain the results for MP3:

$$\begin{aligned} E^{MP3} = E^{MP2} &+ \frac{1}{8} \sum_{a,b,c,d,r,s} \frac{\langle ab||rs\rangle \langle rs||cd\rangle \langle cd||ab\rangle}{(\epsilon_r + \epsilon_s - \epsilon_a - \epsilon_b)(\epsilon_r + \epsilon_s - \epsilon_c - \epsilon_d)} \\ &+ \frac{1}{8} \sum_{a,b,r,s,t,u} \frac{\langle ab||rs\rangle \langle rs||tu\rangle \langle tu||ab\rangle}{(\epsilon_r + \epsilon_s - \epsilon_a - \epsilon_b)(\epsilon_t + \epsilon_u - \epsilon_a - \epsilon_b)} \\ &+ \frac{1}{8} \sum_{a,b,c,s,t,u} \frac{\langle ab||rs\rangle \langle cs||tb\rangle \langle rt||ac\rangle}{(\epsilon_r + \epsilon_s - \epsilon_a - \epsilon_b)(\epsilon_r + \epsilon_t - \epsilon_a - \epsilon_c)} \end{aligned}$$

In general, going beyond the third-order perturbation, results too computational demanding and generally not convenient with respect to other alternatives such as Coupled Cluster Methods. Moreover, the improvement obtained by expanding to larger order it is not systematic. Indeed, it is widely known that for several systems including larger perturbation terms brings to an oscillating series which slowly converge to the exact value (see Fig. 2.1). However, it has been observed in particularly challenging cases a divergent behavior of MPn series linked mainly to nondynamic correlation effects [7].

2.3.4 The Møller-Plesset Kohn-Sham method

Another limitation of standard MPn approaches is that the employment of HF wavefunction as reference. For example, in UHF calculations with large spin contamination, MPn calculations performs poorly [8] giving results even worse than the uncorrelated ones [9]. Several schemes for improving the reference orbitals used in MP2 calculations have been developed specifically for open-shell systems, such as orbital optimized MP2 (OO-MP2) methods [10] and the regularized OO-MP2 methods (the so-called δ-OO-MP2 [11] and κ-OO-MP2 [12] methods). Recently, a novel scheme for MP2 and MP3 methods which employs as a reference wavefunction k-OO-MP2 [13] or Kohn-Sham (KS) [14] orbitals has been proposed. In this paragraph, we will briefly introduce the construction of the second one and we will refer to it as MPn:KS to distinguish it from the standard MPn:HF. If we apply Møller-Plesset perturbation theory to Kohn-Sham orbitals we need to take into account that the orbitals corresponding to the zeroth-order DFT computation come from the KS operator $\hat{f}^{KS}$, which expression will be defined in the DFT section of this chapter. The matrix representation of this operator is not diagonal in the basis of the HF orbitals, therefore, in order to employ the same perturbation operator $\hat{V}$ as in Eq. 2.39, we need to solve the eigenvalue equations separately for occupied and virtual orbitals:

$$\hat{f}_i^{HF}\phi_i^{KS,p} = \epsilon_i^p \phi_i^{KS,p} \tag{2.45}$$

This passage is called semi-canonicalization of molecular orbitals and the new set of orbitals $\phi^{KS,p}$ are called canonical orbitals. To calculate the excited states of this wavefunction we must take into account that the Fock matrix in the basis of canonical orbitals is not completely diagonal, and couples between occupied and virtual orbitals. For this reason, the Brillouin theorem is not anymore valid and we need to include in the expression of the perturbation energy, besides the standard double excitations contributions (DE) also the so-called non Brillouin single excitations (SE). Therefore, the expressions for MP2:KS and MP3:KS are the following:

$$E^{MP2:KS} = E^{DFT} + E^{MP2-SE} + E^{MP2-DE} \tag{2.46}$$

$$E^{MP3:KS} = E^{DFT} + E^{MP2-SE} + E^{MP2-DE} + E^{MP3-SE} + E^{MP3-DE} \tag{2.47}$$

It has been observed [13, 14] that while E^{MP2-SE} is crucial for correcting the energy, E^{MP3-SE} does not significantly affect the results and, therefore, could be neglected to save computational time. For energy calculations, MP2:KS generally does not improve the performance of standard MP2:HF methods. However, MP3:KS improves significantly MP3:HF calculations overperforming also more costly methods such as CCSD. The performance of this method in the calculation of nonlinear optical properties will be discussed in chapter 5.

2.3.5 Spin component scaling MP2

When evaluating the correlation of an MPn calculation it is possible to split the contributions to the energy into antiparallel- and parallel-spin pairs of electrons. Indeed, it has been observed that MP2 tends to overestimate the contributions of the same-spin (SS) electrons and to underestimate the opposite-spin (OS) ones. The SCS-MPn method compensates this error by introducing two parameters to scale the SS and OS correlation contributions. The SS contribution mainly controls the long-range behavior of the MP2 energy taking into account moderate nondynamic correlation effects. Conversely, the OS contribution refers to short-range effects and therefore it is associated with the dynamic correlation of the system. In general, it has been observed that MP2 tends to overestimate the nondynamic correlation and underestimate short-range ones. Therefore, being c_{SS} and c_{OS} the scaling coefficients, an effective parameterization is achieved by decreasing the SS (c_{SS}<1) and increasing the OS (c_{OS}>1) energy contributions. The expression of the SCS-MP2 method will be then:

$$E^{SCS-MP2} = E^{HF} + c_{SS} \cdot E^{MP2}_{SS} + c_{OS} \cdot E^{MP2}_{OS} \tag{2.48}$$

For standard MP2 calculation we have c_{SS}=c_{OS}=1. The values of the coefficients c_{SS} and c_{OS} have been parameterized by Grimme [15] and they are: $c_{SS} = 0.3$ and $c_{OS} = 1.2$. It is important to remark that this method improves the MP2 results without increasing the computational cost. An analogue scheme has been developed also for MP3 [16]. Another scaling method is the scale opposite spin MP2 (SOS-MP2) [17] in which the SS component is neglected ($c_{SS} = 0$) and $c_{OS} = 1.3$. Interestingly, this method can scale as O(M^4) (being more convenient than MP2 which scales as O(M^5)). The performances of SCS-MP2 and SOS-MP2 in the calculation of nonlinear optical properties will be discussed in Chapter 4.

2.3.6 Coupled Cluster Theory

Coupled Cluster (CC) methods are another popular family of methods used in computational chemistry for evaluating the correlation energy. These methods are based on the hypothesis that the CC wavefunction is the eigenfunction of the many-body Hamiltonian. This wavefunction consists in applying an exponential operator $e^{\hat{T}}$ to the HF wavefunction (Φ_0):

$$\Phi_{CC} = e^{\hat{T}}\Phi_0, \tag{2.49}$$

where the operator $\hat{T}$ contains all the excitations of electrons from occupied to virtual orbitals, namely:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + ... + \hat{T}_N \tag{2.50}$$

If we perform a Taylor expansion on the operator $e^{\hat{T}}$ we obtain:

$$e^{\hat{T}} = 1 + \hat{T} + \frac{1}{2}\hat{T}^2 + \frac{1}{6}\hat{T}^3 + ... = \sum_{k=0}^{\infty} \frac{1}{k!}\hat{T}^k. \tag{2.51}$$

Using Eq. 2.50, the expression in Eq. 2.51 can be separated by the order of excitations:

$$e^{\hat{T}} = 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2}\hat{T}_1^2) + (\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3) + ... \tag{2.52}$$

The operators $\hat{T}_i$ are defined with respect to the HF reference wavefunction as the excited Slater determinants, for example the first two terms in Eq. 2.50 are:

$$\hat{T}_1\Phi_0 = \sum_{i}^{occ}\sum_{a}^{vir} t_i^a \Phi_i^a \tag{2.53}$$

$$\hat{T}_2\Phi_0 = \sum_{i<j}^{occ}\sum_{a<b}^{vir} t_{ij}^{ab} \Phi_{ij}^{ab}, \tag{2.54}$$

where the sum over i and j runs over the occupied orbitals, while the one over a and b spans all the virtuals. The coefficients t_i^a and t_{ij}^{ab} in Eqs. 2.53 and 2.54 correspond to the a_S and a_D coefficients in Eq. 2.22 and are defined as excitation amplitudes. In order to solve the many-body problem we need to solve the following Schrödinger equation:

$$\hat{H}e^{\hat{T}}\Phi_0 = Ee^{\hat{T}}\Phi_0. \tag{2.55}$$

For extracting the energy it is thus necessary to evaluate the excitation amplitudes of the cluster expansion. Unfortunately, obtaining a solution by the variational method is not affordable for systems bigger than few electrons (in fact it is even more expensive than FCI), therefore the common way to proceed is to project $< \Phi_0|$ into Eq. 2.55, obtaining an expression for the CC energy:

$$\langle\Phi_0|\hat{H}e^{\hat{T}}|\Phi_0\rangle = E_{CC}\,\langle\Phi_0|e^{\hat{T}}\Phi_0\rangle \tag{2.56}$$
$$E_{CC} = \langle\Phi_0|\hat{H}e^{\hat{T}}|\Phi_0\rangle\,. \tag{2.57}$$

The explicit expansion of E_{CC} is obtained by expanding the $e^{\hat{T}}$ operator. Considering that the Coulomb interaction in the many-body Hamiltonian allows excitations up to two electrons, the expression of the energy becomes:

$$\begin{aligned} E_{CC} =& \langle \Phi_0|\hat{H}(1+\hat{T}_1+\hat{T}_2+\frac{1}{2}\hat{T}_1^2)|\Phi_0\rangle \\ &= \langle \Phi_0|\hat{H}|\Phi_0\rangle + \langle \Phi_0|\hat{H}|\hat{T}_1\Phi_0\rangle + \langle \Phi_0|\hat{H}|\hat{T}_2\Phi_0\rangle + \frac{1}{2}\langle \Phi_0|\hat{H}|\hat{T}_1^2\Phi_0\rangle \\ &= E_0 + \sum_{i}^{occ}\sum_{a}^{vir} t_i^a \langle \Phi_0|H|\Phi_i^a\rangle + \sum_{i<j}^{occ}\sum_{a<b}^{vir} t_{ij}^{ab}\langle \Phi_0|\hat{H}|\Phi_{ij}^{ab}\rangle + \\ &+ \sum_{i<j}^{occ}\sum_{a<b}^{vir}(t_i^a t_j^b - t_i^b t_j^a)\langle \Phi_0|\hat{H}|\Phi_{ij}^{ab}\rangle \end{aligned} \tag{2.58}$$

As it could be observed E_{CC} depends only on singles (t_i^a) and doubles (t_{ij}^{ab}) amplitudes. In order to determine these amplitudes it is necessary to project the excited determinants in Eq. 2.55. Being the wavefunction orthonormal, we obtain then the following equations:

$$\begin{aligned} \langle \Phi_i^a|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = 0 \\ \langle \Phi_{ij}^{ab}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = 0 \\ \langle \Phi_{ijk}^{abc}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = 0 \\ \ldots \end{aligned} \tag{2.59}$$

In Eqs. 2.59, the terms involving single excitations and the reference state are 0 because of the Brillouin theorem, and therefore the remaining terms generate a set of equations in which the singles, doubles, triples, etc. configurations are coupled. These equations are the so-called CC equations, which are solved iteratively. In order to perform a CC calculation, the Cluster operator must be truncated according to the accuracy needed. One of the most popular CC approximations is the CCSD method which consists in including just the single and double excitations, therefore the Φ_{CCSD} wavefunction will have the following expression:

$$\Phi_{CCSD} = e^{\hat{T}_1+\hat{T}_2}\Phi_0 \tag{2.60}$$

Unlike CISD, this method is both size extensive and size consistent and it scales as $O(nM^6)$, where M is the size of the basis sets and n are the number of iterations necessary for solving the CC equations (usually ∼10). If also the triples are included explicitly we refer to the CCSDT method, which scales as $O(nM^8)$ and, therefore, requires an exceptional computational effort both regarding the time necessary to solve the CC equations and the disk space. A more affordable method is CCSD(T), in which triples excitations are included perturbatively after the calculation of the E_{CCSD} energies. This method is one of the most efficient methods for including dynamic correlation. Indeed, it is usually referred as the "*gold standard*" of computational chemistry.

Table 2.1: Thresholds values of different diagnostic tools for the CCSD wavefunction.

Indicator	Value
T1	0.002
D1	0.05
D2	0.18

2.3.6.1 CC2 and CC3 methods

Other populars CCs approximations are CC2 and CC3 [18]. The first method is an approximation of a CCSD calculation, in which the double contributions are computed in a perturbative way, as in the MP perturbation theory. In this scheme, the t_{ij}^{ab} amplitudes are expressed as a combination of t_i^a, which are the only terms that are evaluated iteratively. This method scales as $O(nM^5)$ and it has a similar accuracy to MP2 for ground-state calculations. However, it admits a linear response formulation that allows well-reproduced excited states. In the same way, the CC3 method consists in to simplify the CCSDT method by evaluating the triples excitations in a perturbative scheme.

2.3.6.2 Multireference diagnostic

Post-HF methods can be used only for systems where dynamic correlation is predominant, i.e., systems which are well described with a single Slater Determinant. In order to assess the multiconfigurational character of a particular wavefunction, several diagnostic tools have been developed. The first example is given by the T_1 diagnostic [19], which has been developed for coupled-cluster type wavefunctions. This indicator is constructed as the Euclidean norm of the single excitations coefficients (t_1) divided by the square root of the number of correlated electrons. This diagnostic takes advantage of the property of single excitations to control the orbital relaxations responsible for the multiconfigurational character of the system. Other diagnostics have been developed for CC-type wavefunctions such as D1 [20] and D2 [21] diagnostics. For these methods, reference thresholds for assessing the quality of the reference calculations have been investigated for different computational methods. For instance, for CCSD the thresholds values of T1, D1 and D2 exceeded which the system is considered multireference are reported in Table 2.1. Other kind of indicators, based on the natural occupation numbers, have been also developed. To this group belong the indicators based on the entanglement entropy [22] or the I_D and I_{ND} [23] which are based on a decomposition of the Coulomb hole. Natural occupation based indicators are not strictly limited to CC methods but they can be employed to all the methods for which the wavefunctions are available.

2.4 Accelerated wavefunction methods

Post-HF methods provide very accurate results, reaching up to the chemical accuracy for small molecules. However, at present, their computational cost scales unfavorably with the size of the system, preventing the application of these methods in medium to large molecules. For this reason, a huge effort has been made in the last years to reduce the computational cost of these methods. These developments are mainly related to more efficient implementations but also to the development of a novel class of methods for which the wavefunctions are optimized and less computational heavy. These methods, that are known as accelerated wavefunction techniques, through numerical fitting, rotation of orbitals and truncation techniques, etc.. are able to simplify or neglect superfluous components of standard wavefunction methods without a significant loss of accuracy. In this section, we present some of these methods.

2.4.1 Resolution of Identity

When performing a post-HF calculation, the most time-consuming part is the transformation of the electronic repulsion integrals (ERI) from the AO basis to the MOs, a process that scales as $O(N^5)$. Indeed, it is necessary to perform a four-index transformation:

$$\langle \phi_i \phi_j | \phi_k \phi_l \rangle = \sum_{\mu,\nu,\lambda,\sigma}^{N} C_{\mu i} C_{\nu j} C_{\lambda k} C_{\sigma l} \langle \chi_\mu \chi_\nu | \chi_\lambda \chi_\sigma \rangle \tag{2.61}$$

One very popular way of reducing the computational effort is by the resolution of identity approximations (RI). Following this scheme, the ERIs are approximated as two- or three-index integrals, thus reducing the scaling with respect to the basis set size. Indeed the orbital products are expanded into auxiliary basis functions centered in the position of the atoms ($\bar{\chi}_P$):

$$\chi_\mu \chi_\nu \approx \sum_{P}^{m} \tilde{C}_{\mu\nu}^{P} \bar{\chi}_P \tag{2.62}$$

where m is the number of auxiliary basis functions. The expansion coefficients $\tilde{C}_{\mu P}$ are obtained by minimizing the error integral:

$$\frac{\partial}{\partial \tilde{C}_{\mu\nu}^{P}} \langle R_{\mu\nu} | R_{\lambda\sigma} \rangle = \frac{\partial}{\partial \tilde{C}_{\mu\nu}^{P}} \int R_{\mu\nu} r_{12}^{-1} R_{\lambda\sigma} = 0 \tag{2.63}$$

where $R_{\mu\nu}$ is

$$R_{\mu\nu} = \chi_\mu \chi_\nu - \sum_{P} \tilde{C}_{\mu\nu}^{P} \bar{\chi}_P. \tag{2.64}$$

R$_{\mu\nu}$ is orthogonal to the $\{\chi_P\}$, and, therefore, the following expression containing three- and two-index integrals can be derived:

$$\langle \chi_\mu \chi_\nu | \chi_\lambda \chi_\sigma \rangle = \sum_{PQ} \langle \chi_\mu \chi_\nu | \bar{\chi}_P \rangle [\langle \bar{\chi}_P | \bar{\chi}_Q \rangle]^{-1} \langle \chi_\lambda \chi_\sigma | \bar{\chi}_Q \rangle. \tag{2.65}$$

Although we do not reduce the formal scaling, with Eq. 2.65 it is possible to increase the efficiency for solving Eq. 2.61 by a factor of ∼10. Moreover, a reduction of the disk space to store the integrals is obtained by this transformation, passing from $O(N^4)$ for MP2 to $O(N^2m)$ for RI-MP2. The accuracy of this approximation depends on the size of the auxiliary basis set chosen, which is usually parameterized according to the standard basis employed. The computational time and disk space needs increase with the size of the AO basis. Hence, employing the RI approximation is particularly appropriate for performing calculations with large basis sets, for example, the Dunning ones [5]. RI algorithms have been implemented for the CC2 method [24] in Turbomole, resulting in one of the methods with the best balance between accuracy and computational cost for evaluating excited states. Generally, when RI approximations are applied to other coupled cluster methods, such as RI-CCSD and RI-CCSD(T), they do not improve the computational efficiency because of the way the coupled cluster equations are implemented. However, they provide a consistent saving of storing resources. As a result of this approximation, generally, HF becomes the bottleneck for computations of small- and medium-size molecules with large basis sets. RI approximations have been implemented also for SCF methods. In fact, RI approximation on the Coulomb integrals reduces the computational time by a factor of 10-100. However, if the exact exchange must be computed, such as for HF case or hybrid density functional calculations, it becomes the most time-consuming part of the SCF procedure scaling as $O(N^4)$. A popular implementation of the resolution of identity of the exchange part is the RI-J-K method [25], which reduces the computational cost of the exchange integrals by one order of magnitude. Another popular approximation is the so-called "Resolution of identity chain of spheres method" (RI-J-COSX)[26], which substitutes the exchange part by a numerical integration which scales linearly with the system size. This method underperforms the RI-J-K method for small and medium-size systems. However, it brings large time savings for systems bigger than 100 atoms [27]. Each of these methods have specific auxiliary basis sets which are different from the RI-MP2 ones.

2.4.2 Localization techniques

Post-HF methods are designed to appropriately reproduce dynamic correlation, which is prevalently of short-range nature [28]. However, standard HF orbitals which are employed for constructing excitation amplitudes are generally delocalized over the whole molecule (especially virtual orbitals) and, therefore, they are not the optimal choice for recovering the dynamic correlation effects. This difficulty is even enhanced for KS orbitals because of the delocalization error (see section 2.5). For this reason, novel approximation techniques of Post-HF methods have been developed. They are based on the idea of simplifying the standard correlated wavefunction to obtain a more compact form in which the orbitals employed are localized. These methods, through a complex combination of screening thresholds, cutoffs, parameters and through a large employment of the resolution of identity approximation, can reach linear scaling with the size of the system.

Nowadays, one of the most popular implementations following this philosophy is the do-

main local pair natural orbital (DLPNO) method by Neese and coworkers implemented in Orca [29]. This technique exploits the concept of localization, in fact localized occupied orbitals are associated with a reduced number of virtual orbitals localized on the atoms (the so-called "*domains*") which are employed to compute the correlation energy.
The occupied and virtual orbitals are localized following different schemes. The occupied ones are obtained by popular localization schemes such as the Foster-Boys [30] or the Pipek-Mezey [31] algorithms. To each orbital i is assigned a domain $\{A\}_i$ of atoms A on which it has a significant amplitude p_i^A. Only the atoms in which the Mulliken population of the orbital i is larger than an arbitrary cutoff (T_{CutMKN}) are included in the domain $\{A\}_i$. The virtual orbitals are then constructed in each of the domains obtained from the occupied ones. In an older implementation [32] (the so-called LPNO method), the virtual orbitals associated with each domain were expanded on the AOs. However, in novel implementations, it was preferred to employ the so-called Projected Atomic Orbitals (PAOs) [33], which are obtained by applying a projecting operator on the AOs to remove the contribution of localized internal orbitals, namely each μ^{PAO} could be expressed as:

$$\mu^{PAO} = (1 - |i >< i|)\mu^{AO}. \tag{2.66}$$

This transformation allows to span all the virtual orbitals keeping the locality. For the evaluation of the electron pairs contributions, the pair domains $\{A\}_{ij}$ are defined as $\{A\}_{ij} = \{A\}_i \cup \{A\}_j$. The number of electron pairs which should be included in the calculation is not defined by a spatial criteria but from the semicanonical correlation of the MP2 pair energy ϵ_{ij}^{SC-MP2}, which is computed at the MP2 level:

$$\epsilon_{ij}^{SC-MP2} = -\sum_{ab} \frac{4\,\langle ia|jb\rangle \langle ia|jb\rangle - 2\,\langle ia|jb\rangle \langle ib|ja\rangle}{\epsilon_a + \epsilon_b - F_{ii} - F_{jj}}. \tag{2.67}$$

The number of correlated orbitals is further reduced by employing a cutoff ($T_{CutPairs}$) on ϵ_{ij}^{SC-MP2}. The virtual orbitals are extracted for each pair by the diagonalization of the pair-specific density ($\mathbf{D}^{ij}$) computed at the MP2 level:

$$\mathbf{D}^{ij}\mathbf{d}^{ij} = n^{ij}\mathbf{d}^{ij}, \tag{2.68}$$

where in Eq. 2.68 $\mathbf{d}^{ij}$ and n^{ij} are the Pair Natural Orbitals (PNOs) and the respective occupancy. The extended expression of $\mathbf{D}^{ij}$ can be found in [34]. Even in this case, the number of PNOs associated with each domain is reduced by a cutoff (T_{CutPNO}), which results crucial for reducing the computational time of a DLPNO calculation. The values of these thresholds: T_{CutMKN}, $T_{CutPairs}$, and T_{CutPNO} are the main parameters that rule the accuracy of these methods. In particular, the developers of Orca, through several benchmark studies, have identified three levels of ascending accuracy for these thresholds, which have been defined as [35]: $LoosePNO$, $NormalPNO$, and $TightPNO$, which give these methods a black-box character.

These acceleration techniques have been efficiently implemented for single-point energy calculations of the DLPNO-MP2 [36], DLPNO-CCSD [37] and DLPNO-CCSD(T) methods. In particular, for the DLPNO-CCSD(T) method, two different ways of computing the perturbative triple corrections have been developed. In the standard DLPNO-CCSD(T) [38] the off-diagonal terms of the Fock matrix elements are neglected, corresponding to the analog CCSD(T_0) method for standard CC. The second way of computing the triples leads to the DLPNO-CCSD(T_1) method [39], in which the perturbative triples are computed entirely, resulting in a considerable increase of the computational cost. Analytical gradients for both DLPNO-MP2 [40] and DLPNO-CCSD [41] methods have been implemented, making possible the geometry optimization and computation of several molecular properties such as analytic dipole moments. Interestingly, analytic second derivatives have been also implemented for DLPNO-MP2, which allows to compute analytic polarizabilities and NMR shieldings [42].

Other linear scaling localization techniques have been developed for MP2 and CC methods, particularly relevant are PNO-LCCSD and PNO-LCCSD(T) [43]. These methods have been implemented in MOLPRO [44] and are conceptually very similar to the DLPNO methods. Their main differences is the implementation: even if they both construct the domains using the PAOs, they differ on how the domains are constructed. Indeed, in this method, PNO based methods use a spatial criteria to identify and separate the domains. As a result, the domains obtained are more compact. Considering the similarity of this method with the DLPNO approach, and considering that we have not employed this method in any part of this thesis, we will not go further in the description of this method.

A different approach comes from the localizations developed by Kallay and coworkers in MRCC [45]. They have developed the LMP2 [46], LNO-CCSD [47] and the LNO-CCSD(T) [48, 49] methods, in which the correlation energy is partitioned into localized molecular orbital (LMOs) contributions. In this technique each domain is adjusted individually, and even in this case the virtual orbitals are expanded on the PAOs basis. The pair contributions are then evaluated through the MP2 correlation energy. Irrelevant contributions are neglected even in the LNO case by a particular cutoff. The main difference between this method and the DLPNO one lies in the evaluation of amplitudes, which consists in dividing the molecule into different fragments and to compute the energies and the amplitudes independently for each fragment. Also for LNO methods a set of thresholds have been parameterized according to the different levels of accuracy and efficiency needed. The latter are ruled by the keyword *lcorthr* in an MRCC input and it can be defined as: $Normal$, $Tight$ and $VTight$. The DLPNO and LNO methods have similar performances for the evaluation of energy and low-order molecular properties. However, it has been shown that the LNO-CCSD(T) reproduce better moderate nondynamic correlation effects as compared to DLPNO-CCSD(T_0/T_1) [50].

2.5 Density functional theory

During the last 40 years, Density Functional Theory (DFT) became the most popular electronic structure method. With the development of the theory and the increase of computational facilities its audience of users has expanded from specialized theoretical chemists to experimentalists, giving this theory an almost black-box character at least for the simplest calculations. This popularity is due to its unbeatable balance between accuracy and computational cost. In fact, while for standard quantum chemical methods the fundamental quantity to be determined is the electronic wavefunction (a function of $3N$ spatial variables), in DFT the main unknown variable is the electronic density (a function of three spatial variables), which is directly connected to the wavefunction by the following relation:

$$\rho(\mathbf{r}) = N \int d\sigma \int d\mathbf{x}_2 ... d\mathbf{x}_n |\psi(\{\mathbf{x}_i\})|^2, \tag{2.69}$$

where the integral over σ indicates the integral over the spin, and N is a normalization coefficient that depends on the number of electrons in the system. The first attempt to describe the energy in function of the electronic density has been the Thomas-Fermi model [51, 52]. This method, which is considered a precursor of DFT, describes the kinetic energy of a uniform electron gas surrounded by a uniform positive charge only as a function of the electronic density. Later, Hohenberg and Kohn proved that the solution of the Schrödinger equation for the ground state depends on the knowledge of the electronic density. For a system of electrons in the BO approximation they formulate the Hohenberg and Kohn's theorems [53]:

- The ground-state energy of a system of electrons is uniquely determined by the electronic density. Therefore the ground state energy can be written as a functional of the electronic density $E_0[\rho]$.
- The electronic density that minimizes $E_0[\rho]$ is the density of the ground state ρ_0, i.e., the variational principle is valid, therefore if $\tilde{\rho}$ is a trial density:

$$E_0[\tilde{\rho}] \geq E_0[\rho_0] \tag{2.70}$$

The expression of the energy can be decomposed in three terms:

$$E[\rho] = T[\rho] + U[\rho] + V[\rho], \tag{2.71}$$

where T is the kinetic energy of the system, U is the electronic potential between the electrons and V is the energy of the electrons due to the external potential of the nuclei. The quantity $T + V$ is defined as the HK functional (F^{HK}) and it has a universal expression, while V is system dependent and can be expressed as:

$$V[\rho] = \int d\mathbf{r} v_{ext} \rho(\mathbf{r}), \tag{2.72}$$

where v_{ext} is the external potential generated by the nuclei. In order to determine the energy in Eq. 2.71 it is necessary to determine the F^{HK}, which unfortunately it is unknown in its explicit form. Therefore, the goal of DFT developers is to find proper approximations to it. The most employed technique in DFT is the Kohn-Sham method(KS), which allows to reintroduce the spin-orbitals in DFT [54]. In this formulation the electron density of the KS System (i.e. a fictitious system of non interacting electrons) have the same electronic density as the real one, namely:

$$\tilde{\rho}(\mathbf{r}) = \rho(\mathbf{r}) \tag{2.73}$$

If $\phi(\mathbf{r})$ are the occupied molecular orbitals also the KS system $\tilde{\rho}$ has the expression:

$$\tilde{\rho}(\mathbf{r}) = \sum_{i}^{occ} \phi_i^*(\mathbf{r}) \phi_i(\mathbf{r}) \tag{2.74}$$

In the formulation of KS $\tilde{\rho}$ is employed in order to obtain a simpler expression of Eq. 2.71:

$$E[\tilde{\rho}] = T_S[\tilde{\rho}] + J[\tilde{\rho}] + E_{XC}[\tilde{\rho}] \tag{2.75}$$

where T_S and J are respectively the kinetic energy of the non-interacting system and the classical Coulomb potential. Their expression is known explicitly:

$$T_S[\tilde{\rho}] = -\frac{1}{2} \sum_i \langle \phi_i | \nabla_i^2 | \phi_i \rangle \tag{2.76}$$

$$J[\tilde{\rho}] = \sum_i \langle \phi_i | \hat{J} | \phi_i \rangle \qquad \text{where: } \hat{J} = \int d\mathbf{r_1} \frac{\tilde{\rho}(\mathbf{r}_1)}{|\mathbf{r} - \mathbf{r}_1|} \tag{2.77}$$

The term E_{XC} is called exchange correlation functional and contains all the terms which can not be described by the non-interacting framework. In principle, Eq. 2.75 is exact. However, the exact expression of E_{XC} is not known and must be approximated. Analogously to the HF calculation, the MOs minimizing the total energy in Eq. 2.75 are obtained through a set of equations called KS equations:

$$\hat{f}^{KS} \phi_i = \epsilon_i^{KS} \phi_i \tag{2.78}$$

where $\hat{f}^{KS}$ is the KS operator and has the following expression:

$$\hat{f}^{KS}(\mathbf{r}) = -\frac{\nabla_i^2}{2} + \hat{V}_{ne}(\mathbf{r}) + \hat{J}(\mathbf{r}) + \hat{V}_{XC}(\mathbf{r}) \tag{2.79}$$

where the exchange-correlation potential $\hat{V}_{XC}$ is defined as:

$$\hat{V}_{XC}(\mathbf{r}) = \frac{\delta E_{XC}}{\delta \tilde{\rho}(\mathbf{r})} \tag{2.80}$$

By expanding Eq. 2.78 in a basis set we obtain a set of equations analogous to the Roothan equations for the HF method.

2.5.1 Classification of Density Functional Approximations

The most critical aspect of DFT is the choice of the appropriate density functional approximation (DFA). Indeed, in the last years, hundreds of DFAs have been constructed according to the different strategies, targets, and levels of accuracy needed. Although DFT is usually considered an ab initio method, several DFAs make use of semi-empirical parameters obtained from extensive benchmarks of experimental data or wavefunction methods. An exhaustive general classification of all the DFAs is not possible, however, a general used ranking of the different DFAs is given by the so-called " *Jacob's ladder*" [55]. Here, the functionals are ranked in different rungs according to the complexity of the density functional construction and therefore the computational cost. Although by climbing Jacob's ladder we expect methods with increased accuracy, it is not always the case. The performance of each DFA is strongly dependent on the system and the target property we want to investigate. Jacob's ladder is composed of 5 rungs: functionals constructed by the local spin-density approximation (LSDA), the Generalized Gradient approximation (GGA), the meta-GGAs, hybrids functionals and finally the double hybrids or RPA-based functionals. The LSDA functionals are the most simple approximations in DFT and they depend only on the electronic density. This formulation works properly for the homogenous electronic gas but it is inappropriate for real systems because it can not reproduce the inhomogeneities of the electronic density. An example is the SVWN [56] functional, for which empirical parameters are extracted from Quantum Monte Carlo computations. The GGA functionals belong to the second rung in which the construction of the functional besides $\rho(\mathbf{r})$, also includes its gradient ($\nabla\rho(\mathbf{r})$) to take into account the variations of the electronic density. Examples of these functionals are PBE [57] and BLYP [58]. A further step is given by meta-GGA functionals which also include second derivatives (or kinetic energies) of $\rho(\mathbf{r})$. The most popular functionals which belong to this family are TPSS [59] and M06L [60].

One of the main problems of the functionals of the first rungs is the so-called self interaction error (SIE) [61]. This error is due to a fictitious interaction of the electrons with themselves and is absent in HF because the exchange term exactly cancels out the contribution of the Coulomb potential. This issue leads to a delocalization of the electronic density which is usually referred to as " *delocalization error*" and it is associated with the inability of semi-local density functionals to describe fractional charges [62]. A possible solution to this problem is obtained by replacing the local GGA exchange with the exact HF exchange. The theoretical justification for these functionals is given by the adiabatic connection [63] which allows to connect the non-interacting KS energy with the exact one through one external parameter. This principle guided the development of hybrid functionals, which belong to the fourth rung of Jacob's ladder. In fact, in these functionals, the exchange part of E_{XC} contains a fraction of the exact HF exchange and a fraction of the GGA's:

$$E_X = aE_X^{HF} + (1-a)E_X^{GGA} \tag{2.81}$$

Where a is a parameter determined empirically. The famous B3LYP [64] belongs to this group of functionals, which, especially in the 00's found huge popularity in quantum chemistry. Usually, the parameter a varies between 0.2 to 0.6, therefore, these methods only partially account for the SIE. Indeed, it has been observed that with a bigger percentage of HF exchange, even if the delocalization error is absent, the performance of these functionals is drastically reduced compared to GGA's. A further step in hybrid functionals consists in including the contribution of the exact exchange depending on the interelectronic distance. In this formulation, the two-electron operator of the exact exchange integrals is splitted in a short range part and a long range part by the erf function:

$$\frac{1}{r_{12}} = \frac{1-erf(\omega r_{12})}{r_{12}} + \frac{erf(\omega r_{12})}{r_{12}}, \tag{2.82}$$

Where ω is the parameter that controls the relevance of the long-range term of the exchange. If ω is equal to 0 the exchange part will be only short-range, while if it is ∞ it will be only long-range. Several schemes of range separation have been developed: the long range correction (LC) [65], which allows to develop the LC-BLYP [65] and LC-ωPBE [66] functionals, while the Coulomb attenuated correction is used for the functionals CAM-B3LYP[67] and for the ωB97 family of functionals [68, 69]. In the DFAs presented so far the expression of the E_{XC} depends on the occupied orbitals which describe the electronic density. In the fifth (and last until now) rung of Jacob's ladder, there are the functionals that included also the dependency on the virtual orbitals. In particular, double hybrids are functionals that include to a standard DFA a correction coming from wavefunction theories, commonly from MP2 calculations. The most famous example is the B2-PLYP functional [70]. These functionals have a large % of the HF exchange (>50%), therefore, they do not suffer from SIE, while excessive localization effects are balanced by the inclusion of virtual orbitals. However, they significantly increase the computational cost (see below).

2.5.2 General problems in DFT

Although standard DFAs are generally able to target most of the chemical systems, there are still some problems that can not be solved in a systematic way. These problems are the already mentioned delocalization error, which could be partially avoided by employing range separated or double hybrid functionals, the London's dispersion interactions and the problem of nondynamic correlation. The problem of a correct reproduction of dispersion interactions, is related to the fact that standard DFAs are not able to reproduce their long-range C_6/R^6 asymptotic behavior. Although there are some functionals, such as the Minnesota functionals [71], for which the parameterization is able to include the effects of dispersion, standard DFAs are generally not able to account for these effects in a systematic way. The most common solution to this issue is obtained by including empirical corrections to the standard DFT energy. This is the case for the Grimme's D [15], D2 [72] and D3 [73] corrections. Other approximations include the explicit dependency on the electronic density through a nonlocal potential. The most common functionals of these kinds are the D4 corrections [74] and the VV10 [75] functional. The problem of nondynamic correlation is related to the single-determinant nature of the KS wavefunction. Standard density functional approximations are generally not able to take into account nondynamic correlation effects or they reproduce them in a partial and non systematic manner [76]. As in the HF method, an alternative solution is to employ the unrestricted formalism. However, this approach besides suffering from spin contamination results effective for some systems but inadequate for others [77, 78]. These issues led to the development to a series of functionals able to reproduce strong correlation effects to some extent, some example are the B05 [79], B13 [80] and KP16/B13 [81] functionals. These functionals are able to recover most nondynamic correlation in systems where it is very strong but they underperform for regular systems.

2.5.2.1 The Computational Cost of DFT approximations

DFT presents nowadays the most convenient balance between accuracy and computational cost in computational chemistry. If N is the dimension of the basis set, GGA and meta-GGA functionals usually scale asymptotically as $O(N^3)$, while for hybrid functionals the computing of the exchange part scales as $O(N^4)$. However, this scaling could be reduced one order of magnitude also for DFAs by applying RI-J approximations for Coulomb terms and RI-JK or RIJCOSX approximations for the exchange. Double hybrids scale as $O(N^5)$ (the same scaling as MP2), however, in this case it is possible to apply the same RI approximation employed in the RI-MP2 method, which can reduce the scaling one order of magnitude.

2.6 Solvent effects

All methods described above consider the molecules as isolated in the gas phase. However, for a proper reproduction of experimental results it is necessary to include the effects of the external environments. The ideal way for treating solvation at the QM level is to include explicitly the solvent molecules in the calculation. However, in order to reproduce long-range effects a large number of molecules should be included, which is prohibitive from the computational point of view. The most popular approaches nowadays to account for solvent effects are the continuum solvation models. In these models, the solute is introduced in a cavity that is surrounded by the solvent substituted by a continuous dielectric medium. Using the approximation that the dielectric polarization is described by the linear response (LR), the total electrostatic potential $V(\mathbf{r})$ of the system solute+continuous solvent can be evaluated by solving the Poisson equation:

$$-\nabla \cdot [\epsilon(\mathbf{r})\nabla V(\mathbf{r})] = 4\pi\rho(\mathbf{r}), \tag{2.83}$$

where $\rho(\mathbf{r})$ is the electronic density of the solute, and $\epsilon(\mathbf{r})$ is the dielectric constant. Using the approximation that the medium is isotropic and homogeneous, $\epsilon(\mathbf{r})$ has the following expression:

$$\epsilon(\mathbf{r}) = \begin{cases} 1, & \mathbf{r} \in \textit{cavity} \\ \epsilon_s, & \mathbf{r} \notin \textit{cavity} \end{cases} \tag{2.84}$$

Therefore, the only external parameters needed to define the solute are the dielectric constant ϵ_s and the shape of the cavity. Usually, the cavity is modeled following the Van der Waals surface of the molecule. The most effective way to solve this electrostatic problem is to decompose the potential $V(\mathbf{r})$ in two contributions: one coming from the solute charge density $V^{\rho}(\mathbf{r})$ and the other generated from a surface charge σ located on the boundary(Γ) of the cavity:

$$V(\mathbf{r}) = V^{\rho}(\mathbf{r}) + V^{\sigma}(\mathbf{r}) = \int d\mathbf{r}'\frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} + \int_{\Gamma} d\mathbf{r}'\frac{\sigma(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \tag{2.85}$$

where the expression of $\sigma(\mathbf{r})$ depends on the model and implementation used. The most common are the polarizable continuum model (PCM) [82, 83] and the conductor-like screening model (COSMO) [27], which are largely employed in computational chemistry because of their great versatility for different molecular solutes and high accuracy/cost ratio. In this thesis we employed the PCM model by using the Integral Equation Formalism (IEP) [84] implemented in Gaussian. For the numerical implementation of the PCM algorithm it is necessary to divide the surface of the cavity in discrete parts in order to solve the problem numerically. To each element of the surface of the cavity a_i we can associate the corresponding superficial charge σ_i. If we assume that the charge is constant in each a_i, we can rewrite it in terms of point charges q_i, which can be expressed as:

$$q_i = \sigma_i a_i. \tag{2.86}$$

This discretization can be made following different schemes such as the so-called GEPOL method [85] or the continuous surface charge (CSC) formalism [86]. The coupling with the QM part is made by defining an effective Hamiltonian ($\hat{H}^{eff}$) obtained by adding a reaction potential $\hat{V}^R$ to the vacuum Hamiltonian $\hat{H}^0$:

$$\hat{H}^{eff} = \hat{H}^0 + \hat{V}^R, \tag{2.87}$$

where $\hat{V}^R$ is:

$$\hat{V}^R = \sum_i q_i \hat{V}_i, \tag{2.88}$$

in which q_i are the point charges defined in Eq. 2.86 and $\hat{V}_i$ is the potential defined in each area a_i. Therefore, for evaluating the electronic solvation energy, the Schrödinger equation corresponding to this Hamiltonian must be solved. Considering that there is a mutual polarization between the solute and the solvent, the most efficient way to compute the solvation energy is solving iteratively the KS equations with the effective Hamiltonian 2.87 and Eq. 2.83. These calculations do not increase significantly the computational cost of standard DFT calculations, considering that the expression for the evaluation of the $\hat{V}^R$ is a one-particle operator.

2.7 Molecular dynamics

Through QM methods, it is generally too computational demanding to explicitly evaluate the effects of the environment and the dynamic evolution of a system. One efficient alternative is to capture these effects by classical mechanics. Among the available modelling techniques at the classical level, atomistic Molecular dynamics (MD) is a widely-used computational method that study the time-evolution of the interactions of atoms and molecules. MD simulation consists in describing the trajectories of the nuclei, which are computed by solving Newton's equations of motion:

$$m_i \ddot{\mathbf{r}}_i = \mathbf{f}_i = -\nabla V_i \tag{2.89}$$

The forces acting on the nuclei $\mathbf{f}_i$ are determined according to a potential energy V_i (see below) which act on the atom i. Therefore, the motion of the electrons is not considered and quantum effects are neglected. The solution of these equations for each time t will define the microstate of the system characterized by the positions and the momenta of all of its particles. Therefore, for a system composed of N particles, each microstate depends on $6N$ variables. Given a thermodynamic state, the set of all possible microstates is defined as an ensemble. Fixing the thermodynamic state of the system (i. e., fixing some state functions such as the pressure or the temperature), knowing the position $\mathbf{r}$ and the momentum $\mathbf{p}$, we can compute any macroscopic observable which can be expressed as ensemble averages. According to the quantity we keep fixed, several possible statistical ensembles can be considered:

- The microcanonical ensemble (NVE) is the standard ensemble for MD; it consists in creating a thermodynamically isolated system with a constant number of particles, volume and energy.
- The canonical ensemble (NVT) consists in a system with a fixed number of particles,

a fixed volume and temperature. It can be modeled by coupling the system with a thermostat in the MD simulation, such as for example the Noose-Hoover thermostat [87, 88].

- The isothermal isobaric ensemble (NPT) for which we keep fixed the number of particles, the pressure and the temperature. To maintain the pressure constant the system is coupled with a barostat such as the one developed by Berendsen [89] and a thermostat.

Once the macrostate is fixed, for evaluating a macroscopic observable A is evaluated from its statistical distribution over the ensemble $< A >_{ens}$:

$$A_{obs} =< A >_{ens}= \sum_i A_i \rho_{ens}(i) \tag{2.90}$$

where in Eq. 2.90 the sum over i runs over all the microstates defined by the selected ensemble and $\rho_{ens}(i)$ represents the probability distribution of each microstate in the ensemble. However, in practical applications, we consider just a finite number of configurations on which we evaluate the observables. In MD each configuration is determined by Eqs. 2.89 by numerical techniques. Therefore, Eq. 2.90 isreduced to:

$$A_{obs} =< A >_{ens}\simeq \frac{1}{N} \sum_{t_i} A_{t_i}, \tag{2.91}$$

Where N is the number of configurations computed and A_{t_i} is the value of A at each timestep t_i. This process is analogous to perform a measurement of a real experiment. First the sample is prepared in a *"box"*, with specific initial conditions, before the system is left to evolve in time. The properties are evaluated from the configurations generated, and the observable is computed as a statistical average of them.

2.7.1 Solution of Newton equations

An analytic solution of equations 2.89 does not exist for a many-body system, therefore, they need to be solved numerically. There are several numerical schemes employed nowadays to integrate the equations of motions [90]. One of the most efficient is the so-called velocity Verlet algorithm [91], which allows to compute positions and momenta as Taylor expansions with respect to the time. If we fix the initial conditions at $t = 0$ as $\mathbf{r}(0) = \mathbf{r}_0$ and $\mathbf{p}(0) = \mathbf{p}_0$ and Δt is the time step chosen for the numerical integration, we can define the evolution of $\mathbf{r}$ as:

$$\mathbf{r}(\Delta t) = \mathbf{r}_0 + \Delta t \frac{\mathbf{p}_0}{m} + \Delta t^2 \frac{\mathbf{f_0}}{2m}, \tag{2.92}$$

where $\mathbf{f_0}$ are the forces evaluated at time $t = 0$. The forces $\mathbf{f}(\Delta t)$ are then evaluated according to the new positions and, therefore, an analogue equation for $\mathbf{p}(\Delta t)$ is computed:

$$\mathbf{p}(\Delta t) = \mathbf{p}_0 + \Delta t \frac{\mathbf{f_0} + \mathbf{f}(\Delta t)}{2m} \tag{2.93}$$

The most time-consuming part of the procedure is the evaluation of the forces. Indeed for N particles, it scales asymptotically as $O(N^2)$. While the choice of the time step is arbitrary, it should be long enough to allow a proper sample of the phase space without performing a large number of integration steps, but small enough to ensure a proper numerical integration (a proper choice of the time step is usually of few fs).

In order to simulate periodic systems, such as crystals or polymers or to remove unphysical edge effects given by the choice of the simulation box, periodic boundary conditions (PBC) are usually employed. Using this approximation the simulation cell is replicated all over the space in order to form an infinite system. If a molecule moves in the central cell its copies in the other cells also will move following the same trajectory. Therefore, if a molecule gets out to one cell it will be replaced by one from another cell. In order to avoid the interaction between each molecule replicated in different cells, particular care should be taken when choosing the size of the box and the cutoff parameters for evaluating the long-range interactions.

2.7.2 Description of the Force Field

In MD simulations, the interactions between atoms is described using a set of potential functions and associated parameters commonly referred as force field (FF). Considering that the electrons are not included in the MD description, the different chemical properties of the nuclei are reproduced by defining different atom types that are associated to different parameters of the FF. For instance, if we want to reproduce three different hybridization states of a carbon atom, we need to include 3 different atom types and for each one a different set of FF parameters. Obviously, the more atom types we include in the model, the more costly will be the computation. Each FF has been parameterized for different types of systems and for the specific property one wants to evaluate; in this work, we employed the generalized amber force field (GAFF) [92] although several other FFs have been produced, such as the DREIDING [93] or CHARMM [94]. For the GAFF model, the atom types can not change during the simulation and therefore the connectivity must be defined before starting the simulation. The FFs can be generally decomposed in *bonded* and $non-bonded$ components:

$$V^{FF} = V_{bonded} + V_{non-bonded} \tag{2.94}$$

The bonded part includes the stretching of the bonds (V_{bond}), the bending of the angles (V_{angles}) and the torsion of dihedrals (V_{dih}). The expression of these functions are:

$$V_{bond} = \sum_{i}^{bonds} \frac{k_i^r}{2}(r_i - r_i^{eq})^2 \qquad V_{angles} = \sum_{l}^{angles} \frac{k_l^\theta}{2}(\theta_l - \theta_l^{eq})^2 \qquad V_{dih} = \sum_{n}^{dih} \frac{k_n}{2}(1 + cos(n\alpha - \beta)) \tag{2.95}$$

where k_i^r, r_i^{eq}, k_l^θ, θ_i^{eq}, k_n, α and β are the empirical parameters of the model. In the $non-bonded$ interactions, all the other contributions which are not related to bonding termsare included. In standard GAFF we have two terms which are described by pairwise interactions: the Van der Waals (V_{VdW}) and Coulomb electrostatic potentials (V_{elec}), namely:

$$V_{VdW} = \sum_{i,j} 4\epsilon_{i,j} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] \qquad V_{elec} = \sum_{i,j} \frac{q_i q_j}{r_{ij}} \tag{2.96}$$

where i and j indicate two different atoms of the system, r_{ij} identifies the distance between them, q_i the atomic charge, σ_{ij} and ϵ_{ij} are the empirical parameters related to London's dispersion forces. The number of pairs included in the computation of these terms is defined by a spatial criterium and limited by a truncation parameter (cutoff). All the parameters of the FF are determined by benchmark studies on more accurate methods (such as DFT or wavefunction methods) or by direct interpolation of experimental data.

2.7.2.1 Reparameterization of a Force Field

As FF are generally derived for a given type of systems their straightforward transferability to all kind of compounds should not be assumed. Therefore, depending what is being investigated, it might be necessary to actually reparameterize it according to the specific systems studied. For example, for the study of the NLO properties, it is necessary to accurately reproduce the response of the molecules. Generally, this can be made by parameterizing the BLA and the torsion angles according to reference QM calculations. In order to reproduce the BLA of a system with a π-conjugated bridge, different atom types for adjacent carbons must be defined in order to obtain different bond lengths for each component of the bridge. The corresponding equilibrium bond lengths and angles are usually taken from optimized DFT geometries. For the reparameterization of the force constants of V_{bond} and V_{angles} and the coefficients k_n, n of V_{dih} more complex methods are needed. One of the most employed techniques is the adaptive biased force (ABF) method. This is an iterative procedure that consists in mimicking the reference potential energy surface (which might be obtained from DFT calculations) of the interaction that we want to parameterize, with the analogue obtained by an MD sampling. More details of this procedure can be found in the literature [95]. Moreover, the atomic charges of Coulomb potentials are also generally reparameterized specifically for the systems studied. Usually, they are extracted from DFT calculations on optimized molecular geometries by evaluating the electrostatic fitted potentials (ESP). In chapter 8, starting from the standard GAFF parameterization [92], we applied this procedure to reproduce the NLO response of a polyenic system.

Bibliography

[1] Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.* **1927**, *389*, 457–484.

[2] Cederbaum, L. S. Born–Oppenheimer approximation and beyond for time-dependent electronic processes. *J. Chem. Phys.* **2008**, *128*, 124101.

[3] T., K. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica* **1934**, *1*, 104–113.

[4] Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728.

[5] Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[6] Löwdin, P.-O. *Advances in Chemical Physics*; John Wiley and Sons, Ltd, 1958; pp 207–322.

[7] Olsen, J.; Jørgensen, P.; Helgaker, T.; Christiansen, O. Divergence in Møller–Plesset theory: A simple explanation based on a two-state model. *J. Chem. Phys.* **2000**, *112*, 9736–9748.

[8] Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *WIREs Computational Molecular Science* **2011**, *1*, 509–530.

[9] Slow convergence of the møller-plesset perturbation series: the dissociation energy of hydrogen cyanide and the electron affinity of the cyano radical. *Chem. Phys. Lett.* **1987**, *138*, 481–485.

[10] Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *J. Chem. Phys.* **2007**, *126*, 164101.

[11] Stück, D.; Head-Gordon, M. Regularized orbital-optimized second-order perturbation theory. *J. Chem. Phys.* **2013**, *139*, 244109.

[12] Lee, J.; Head-Gordon, M. Regularized Orbital-Optimized Second-Order Møller–Plesset Perturbation Theory: A Reliable Fifth-Order-Scaling Electron Correlation Model with Orbital Energy Dependent Regularizers. *J. Chem. Theory Comput.* **2018**, *14*, 5203–5219.

[13] Bertels, L. W.; Lee, J.; Head-Gordon, M. Third-Order Møller–Plesset Perturbation Theory Made Useful? Choice of Orbitals and Scaling Greatly Improves Accuracy for Thermochemistry, Kinetics, and Intermolecular Interactions. *J. Phys. Chem. Lett.* **2019**, *10*, 4170–4176.

[14] Rettig, A.; Hait, D.; Bertels, L. W.; Head-Gordon, M. Third-Order Møller–Plesset Theory Made More Useful? The Role of Density Functional Theory Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 7473–7489.

[15] Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

[16] Grimme, S. Improved third-order Møller–Plesset perturbation theory. *J. Comput. Chem.* **2003**, *24*, 1529–1537.

[17] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

[18] Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

[19] Lee, T. J.; Taylor, P. R. A diagnostic for determining the quality of single-reference electron correlation methods. *Int. J. Quantum Chem.* **1989**, *36*, 199–207.

[20] Janssen, C. L.; Nielsen, I. M. B. New diagnostics for coupled-cluster and Møller–Plesset perturbation theory. *Chem. Phys. Lett.* **1998**, *290*, 423–430.

[21] Nielsen, I. M. B.; Janssen, C. L. Double-substitution-based diagnostics for coupled-cluster and Møller–Plesset perturbation theory. *Chem. Phys. Lett.* **1999**, *310*, 568–576.

[22] Ziesche, P. Correlation strength and information entropy. *International Journal of Quantum Chemistry* **1995**, *56*, 363–369.

[23] Ramos-Cordoba, E.; Salvador, P.; Matito, E. Separation of dynamic and nondynamic correlation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24015–24023.

[24] Hättig, C.; Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **2000**, *113*, 5154–5161.

[25] Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

[26] Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange. *Chem. Phys.* **2009**, *356*, 98–109, Moving Frontiers in Quantum Chemistry:.

[27] Kossmann, S.; Neese, F. Comparison of two efficient approximate Hartee–Fock approaches. *Chem. Phys. Lett.* **2009**, *481*, 240–243.

[28] Pulay, P. Localizability of dynamic electron correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

[29] Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, 224108.

[30] Foster, J. M.; Boys, S. F. Canonical configurational interaction procedure. *Rev. Mod. Phys.* **1960**, *32*, 300–302.

[31] Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for *ab initio* and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

[32] Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.

[33] Boughton, J. W.; Pulay, P. Comparison of the boys and Pipek–Mezey localizations in the local correlation approach and automatic virtual basis selection. *J. Comput. Chem.* **1993**, *14*, 736–740.

[34] Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *J. Chem. Phys.* **2009**, *130*, 114108.

[35] Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the accuracy limits of local pair natural orbital coupled-cluster theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.

[36] Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple

linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, 034108.

[37] Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.

[38] Schütz, M. Low-order scaling local electron correlation methods. III. Linear scaling local perturbative triples correction (T). *J. Chem. Phys.* **2000**, *113*, 9986–10001.

[39] Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.

[40] Pinski, P.; Neese, F. Communication: Exact analytical derivatives for the domain-based local pair natural orbital MP2 method (DLPNO-MP2). *J. Chem. Phys.* **2018**, *148*, 031101.

[41] Datta, D.; Kossmann, S.; Neese, F. Analytic energy derivatives for the calculation of the first-order molecular properties using the domain-based local pair-natural orbital coupled-cluster theory. *J. Chem. Phys.* **2016**, *145*, 114101.

[42] Stoychev, G. L.; Auer, A. A.; Gauss, J.; Neese, F. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *J. Chem. Phys.* **2021**, *154*, 164110.

[43] Ma, Q.; Werner, H.-J. Scalable Electron Correlation Methods. 5. Parallel Perturbative Triples Correction for Explicitly Correlated Local Coupled Cluster with Pair Natural Orbitals. *J. Chem. Theory Comput.* **2018**, *14*, 198–215.

[44] Werner, H. J. MOLPRO, a package of ab initio programs. *http://www. molpro. net* **2010**,

[45] Kállay, M. et al. The MRCC program system: Accurate quantum chemistry from water to proteins. *J. Chem. Phys.* **2020**, *152*, 074107.

[46] Nagy, P. R.; Samu, G.; Kállay, M. An integral-direct linear-scaling second-order Møller–Plesset approach. *J. Chem. Theory Comput.* **2016**, 4897–4914.

[47] Rolik, Z.; Kállay, M. A general-order local coupled-cluster method based on the cluster-in-molecule approach. *J. Chem. Phys.* **2011**, *135*, 104111.

[48] Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladóczki, B.; Kállay, M. An efficient linear-scaling CCSD(T) method based on local natural orbitals. *J. Chem. Phys.* **2013**, 094105.

[49] Nagy, P. R.; Samu, G.; Kállay, M. Optimization of the Linear-Scaling Local Natural Orbital CCSD(T) Method: Improved Algorithm and Benchmark Applications. *J. Chem. Theory Comput.* **2018**, *14*, 4193–4215.

[50] Sylvetsky, N.; Banerjee, A.; Alonso, M.; Martin, J. M. L. Performance of localized coupled cluster methods in a moderately strong correlation regime: Hückel–Möbius interconversions in expanded porphyrins. *J. Chem. Theory Comput.* **2020**, *16*, 3641–3653.

[51] Thomas, L. H. The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society* **1927**, *23*, 542–548.

[52] Fermi, E. Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente. *Zeitschrift für Physik* **1928**, *48*, 73–79.

[53] Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136*, B864–B871.

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

[55] Perdew, J. P.; Schmidt, K. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conference Proceedings* **2001**, *577*, 1–20.

[56] Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

[57] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[58] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[59] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.

[60] Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

[61] Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Many-electron self-interaction error in approximate density functionals. *J. Chem. Phys.* **2006**, *125*, 201102.

[62] Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into Current Limitations of Density Functional Theory. *Science* **2008**, *321*, 792–794.

[63] Yang, W. Generalized adiabatic connection in density functional theory. *J. Chem. Phys.* **1998**, *109*, 10107–10110.

[64] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[65] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

[66] Vydrov, O. A.; Scuseria, G. E. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.* **2006**, *125*, 234109.

[67] 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.

[68] Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

[69] Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[70] Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

[71] Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theoretical Chemistry Accounts* **2008**, *120*, 215–241.

[72] Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[73] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

[74] Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, 034112.

[75] Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.

[76] Gräfenstein, J.; Kraka, E.; Filatov, M.; Cremer, D. Can Unrestricted Density-Functional Theory Describe Open Shell Singlet Biradicals? *Int. J. Mol. Sci.* **2002**, *3*, 360–394.

[77] Zhang, D.; Truhlar, D. G. Unmasking Static Correlation Error in Hybrid Kohn-Sham Density Functional Theory. *J. Chem. Theory Comput.* **2020**, *16*, 5432–5440.

[78] Naim, C.; Amovilli, C. Extraction of a One-Particle Reduced Density Matrix from a Quantum Monte Carlo Electronic Density: A New Tool for Studying Nondynamic Correlation. *Computation* **2021**, *9*.

[79] Becke, A. D. Real-space post-Hartree–Fock correlation models. *J. Chem. Phys.* **2005**, *122*, 064101.

[80] Becke, A. D. Density functionals for static, dynamical, and strong correlation. *J. Chem. Phys.* **2013**, *138*, 074109.

[81] Kong, J.; Proynov, E. Density Functional Model for Nondynamic and Strong Correlation. *J. Chem. Theory Comput.* **2016**, *12*, 133–143.

[82] Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117–129.

[83] Mennucci, B. Polarizable continuum model. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 386–404.

[84] Cancès, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

[85] Silla, E.; Tuñón, I.; Pascual-Ahuir, J. L. GEPOL: An improved description of molecular surfaces II. Computing the molecular area and volume. *J. Comput. Chem.* **1991**, *12*, 1077–1088.

[86] Scalmani, G.; Frisch, M. J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* **2010**, *132*, 114110.

[87] Evans, D. J.; Holian, B. L. The Nose–Hoover thermostat. *J. Chem. Phys.* **1985**, *83*, 4069–4074.

[88] Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635–2643.

[89] Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

[90] Leimkuhler, B. J.; Reich, S.; Skeel, R. D. In *Mathematical Approaches to Biomolecular Structure and Dynamics*; Mesirov, J. P., Schulten, K., Sumners, D. W., Eds.; Springer New York: New York, NY, 1996; pp 161–185.

[91] Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637–649.

[92] Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

[93] Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

[94] Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

[95] Pizzirusso, A.; Di Pietro, M. E.; De Luca, G.; Celebre, G.; Longeri, M.; Muccioli, L.; Zannoni, C. Order and Conformation of Biphenyl in Cyanobiphenyl Liquid Crystals: A Combined Atomistic Molecular Dynamics and 1H NMR Study. *ChemPhysChem* **2014**, *15*, 1356–1367.

Chapter 3

Computation of Nonlinear optical properties

3.1 Sum over states method

Linear and nonlinear optical properties can be modeled as a response of a quantum mechanical system to a classical external electric field. In the general case, the electric fields are time-dependent, therefore the wavefunction of this system is determined by the time-dependent Schrödinger equation (2.1). The Hamiltonian of the system can be expressed as:

$$\hat{H} = \hat{H}_0 + \hat{V}(t), \tag{3.1}$$

where $\hat{H}_0$ is the Hamiltonian of the isolated molecule and $\hat{V}(t)$ is the potential associated with the external electric field $\mathbf{F}(t)$, namely:

$$\hat{V}(t) = -\hat{\mu}_i F_i(t), \tag{3.2}$$

where $\hat{\boldsymbol{\mu}}$ is the permanent dipole moment and the implicit summation notation is used. Generally, the field coming from the external source is much weaker than the fields that hold atoms and molecules, this interaction can be then treated in the framework of perturbation theory. Therefore, the wavefunction can be expressed in terms of eigenstates of the unperturbed system, so if $|n\rangle$ are the exact eigenstates of the Hamiltonian $\hat{H}_0$, we have:

$$\hat{H}_0 |n\rangle = E_n |n\rangle, \tag{3.3}$$

where E_n are the energies of each state. The resulting perturbed time-dependent wavefunction can be expressed as:

$$|\psi(t)\rangle = \sum_n d_n(t) e^{-iE_n t/\hbar} |n\rangle, \tag{3.4}$$

where $d_n(t)$ are the unknown expansion coefficients which satisfy $d_n(-\infty) = \delta_{n0}$. We can expand $d_n(t)$ in a perturbation series, namely:

$$d_n(t) = d_n^{(0)} + d_n^{(1)}(t) + d_n^{(2)}(t) + d_n^{(3)}(t) + ... = \sum_N^{\infty} d_n^{(N)}(t). \tag{3.5}$$

By including Eq. 3.5 in Eq. 3.4 and multiplying Eq. 2.1 by $\langle m| e^{iE_m t/\hbar}$ we can obtain the equations to determine the $d_m^{(N)}(t)$:

$$i\hbar \frac{\partial}{\partial t} d_m^{(N)}(t) = \sum_n \langle m|\hat{\mu}_i|n\rangle F_i(t) e^{i\omega_{mn}t} d_n^{(N-1)}(t), \tag{3.6}$$

where $\omega_{mn} = \frac{E_m - E_n}{\hbar}$. These equations are recursive with $d_n^{(0)} = \delta_{n0}$. We can then express also $|\psi(t)\rangle$ in a perturbative series as:

$$|\psi(t)\rangle = |\psi^{(0)}(t)\rangle + |\psi^{(1)}(t)\rangle + |\psi^{(2)}(t)\rangle + ... = \sum_N^{\infty} |\psi^{(N)}(t)\rangle , \tag{3.7}$$

where:

$$|\psi^{(N)}(t)\rangle = \sum_n d_n^{(N)} e^{\frac{-iE_n t}{\hbar}} |n\rangle . \tag{3.8}$$

The expectation values of the Hamiltonian or any other operator will be therefore time-dependent. In particular, the expectation value of the electric dipole moment $\hat{\boldsymbol{\mu}}$ can be expressed as a perturbation expansion:

$$\langle \psi(t)|\hat{\mu}_i|\psi(t)\rangle = \langle \hat{\mu}_i \rangle^{(0)} + \langle \hat{\mu}_i \rangle^{(1)} + \langle \hat{\mu}_i \rangle^{(2)} + \langle \hat{\mu}_i \rangle^{(3)} + ... \tag{3.9}$$

The first four terms of Eq. 3.9 are:

$$\langle \hat{\mu}_i \rangle^{(0)} = \langle 0|\hat{\mu}_i|0\rangle \tag{3.10}$$

$$\langle \hat{\mu}_i \rangle^{(1)} = \langle 0|\hat{\mu}_i|\psi^{(1)}\rangle + \langle \psi^{(1)}|\hat{\mu}_i|0\rangle \tag{3.11}$$

$$\langle \hat{\mu}_i \rangle^{(2)} = \langle 0|\hat{\mu}_i|\psi^{(2)}\rangle + \langle \psi^{(2)}|\hat{\mu}_i|0\rangle + \langle \psi^{(1)}|\hat{\mu}_i|\psi^{(1)}\rangle \tag{3.12}$$

$$\langle \hat{\mu}_i \rangle^{(3)} = \langle 0|\hat{\mu}_i|\psi^{(3)}\rangle + \langle \psi^{(3)}|\hat{\mu}_i|0\rangle + \langle \psi^{(1)}|\hat{\mu}_i|\psi^{(2)}\rangle + \langle \psi^{(2)}|\hat{\mu}_i|\psi^{(1)}\rangle \tag{3.13}$$

In Eq. 3.10 we have the permanent dipole moment, while from Eqs 3.11-3.13 we can recover the expressions of α, β, and γ. For instance, expressing the electric field in its spectral components as in Eq. 1.3, we can expand Eq. 3.11 as:

$$\langle \hat{\mu}_i \rangle^{(1)} = \sum_\omega \frac{1}{\hbar} \sum_n{}' \left[\frac{\langle 0|\hat{\mu}_i|n\rangle \langle n|\hat{\mu}_j|0\rangle}{\omega_{n0} - \omega} + \frac{\langle 0|\hat{\mu}_j|n\rangle \langle n|\hat{\mu}_i|0\rangle}{\omega_{n0} + \omega} \right] F_j^{\omega} e^{-i\omega t}, \tag{3.14}$$

where $\sum'$ indicates that $n = 0$ is excluded from the summation. By a direct comparison between Eq. 3.14 and Eq. 1.4 we can obtain the expression of the dynamic polarizability:

$$\alpha_{ij}(-\omega;\omega) = \frac{1}{\hbar}\sum_n{}' \Big[\frac{\langle 0|\hat{\mu}_i|n\rangle\,\langle n|\hat{\mu}_j|0\rangle}{\omega_{n0}-\omega} + \frac{\langle 0|\hat{\mu}_j|n\rangle\,\langle n|\hat{\mu}_i|0\rangle}{\omega_{n0}+\omega}\Big], \tag{3.15}$$

where we assumed that the frequency ω of the incident radiation is different from the resonance frequencies ω_{n0} of the system to avoid divergences. In the same way, we can express the first hyperpolarizabilty as:

$$\beta_{ijk}(-\omega_\sigma;\omega_1,\omega_2) = \frac{1}{\hbar^2}\sum \mathcal{P}_{-\sigma,1,2}\sum_{np}{}' \frac{\langle 0|\hat{\mu}_i|n\rangle\,\langle n|\overline{\hat{\mu}_j}|p\rangle\,\langle p|\hat{\mu}_k|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{p0}-\omega_2)}, \tag{3.16}$$

where $\sum \mathcal{P}_{-\sigma,1,2}$ indicates the 6 elements obtained by permuting the pairs of indices (i, $-\omega_\sigma$), (j, ω_1) and (k, ω_2), while $\overline{\hat{\mu}_j}$ indicates the difference between the excited state dipole moment with respect to the permanent one: $\overline{\hat{\mu}_j} = \hat{\mu}_j - \langle 0|\hat{\mu}_j|0\rangle$. An analog expression can be obtained for the second hyperpolarizability:

$$\begin{aligned}\gamma_{ijkl}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) =& \frac{1}{\hbar^3}\sum \mathcal{P}_{-\sigma,1,2,3}\Bigg[\sum_{nmp}{}' \frac{\langle 0|\hat{\mu}_i|n\rangle\,\langle n|\overline{\hat{\mu}_j}|m\rangle\,\langle m|\overline{\hat{\mu}_k}|p\rangle\,\langle p|\hat{\mu}_l|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{p0}-\omega_2-\omega_3)(\omega_{p0}-\omega_3)}+\\ &-\sum_{nm}{}' \frac{\langle 0|\hat{\mu}_i|n\rangle\,\langle n|\overline{\hat{\mu}_j}|0\rangle\,\langle 0|\overline{\hat{\mu}_k}|m\rangle\,\langle m|\hat{\mu}_l|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{m0}-\omega_3)(\omega_{m0}+\omega_2)}\Bigg],\end{aligned} \tag{3.17}$$

where in this case $\sum \mathcal{P}_{-\sigma,1,2,3}$ indicates the 24 elements obtained by permuting the pairs of indices (i, $-\omega_\sigma$), (j, ω_1), (k, ω_2) and (l, ω_3). The Eqs. 3.15-3.17 are the so-called sum-over-states (SOS) expressions of linear and nonlinear optical properties. We made this derivation by assuming that these states are the exact eigenstates of the Hamiltonian and consequently the properties are exact. Therefore, both electronic and nuclear degrees of freedom are included and each excited state is the exact non-adiabatic vibronic state of the system. However, if we employ the Born-Oppenheimer approximation, the wavefunction can be described as a product of the electronic and vibrational parts. In this case, the electronic and nuclear contributions can be separated and for this thesis, we focus only on the electronic part. The SOS scheme is rigorous only if the exact wavefunction is known (impossible in practical cases), however, it can be employed also for standard electronic structure calculations (HF, DFT, CC,...) although it has been observed that in order to reach the convergence a large number of states must be included in the summation. Some schemes based on the truncation of the SOS series are effective for some particular systems, the most popular being the two-level approximation for charge transfer systems (see next sections).

3.1.1 Residues Analysis

If the explicit expressions of the optical properties are known it is possible to determine the transition moments between excited states which can be connected to spectroscopic quantities. This analysis can be made by computing the residues of Eqs. 3.15-3.17 [1]. In fact, being $f(z)$ a function of a complex variable z, divergent in a which can be expressed as :

$$f(z) = \frac{g(z)}{(z-a)^n} \tag{3.18}$$

where a is defined as pole of order n of $f(z)$ and $g(z)$ is a function non singular in a. The residues related to the pole a can be computed as:

$$Res(f,a) = \frac{1}{(n-1)!} \lim_{z \to a} \frac{d^{n-1}}{dz^{n-1}}[(z-a)^n f(z)]. \tag{3.19}$$

For example, from Eq. 3.15 the frequencies ω_{n0} are poles of order 1 of the polarizability (which corresponds to the excitation energies of the system), while the corresponding residues are:

$$\lim_{\omega \to \omega_{n0}} (\omega_{n0} - \omega)\alpha_{ij}(-\omega;\omega) = \langle 0|\hat{\mu}_i|n\rangle \langle n|\hat{\mu}_j|0\rangle \,, \tag{3.20}$$

which are the transition dipole moments between the ground and excited states, associated with the intensity of the one-photon absorption. Other transition matrix elements can be evaluated by computing the residues of β and γ, a list of the main optical properties which can be obtained from residue analysis is collected in Table 3.1.

3.1.2 Two-level approximation

When computing the optical properties of systems that have a single dominant one-photon transition (from the ground state $|0\rangle$ to the excited state $|f\rangle$), it is convenient to include in the SOS expansion only the state $|f\rangle$ and neglect contributions from all other excited states. This model is called two-level approximation (TLA) and allows a qualitative description of linear and nonlinear optical properties of specific systems. For instance, TLA is particularly performing for push-pull systems which are made of a donor and an acceptor tied by a π-bridge, having a strong charge transfer along a particular direction. If the charge transfer is along z the only relevant component of the polarizability tensor is α_{zz}, which, applying the TLA on Eq. 3.15, can be expressed as:

$$\alpha_{zz}^{TLA}(-\omega;\omega) = \alpha_{zz}^{TLA}(0)F(\omega), \tag{3.21}$$

Table 3.1: Scheme of the main properties which can be obtained by residue analysis of the polarizability $\alpha(-\omega;\omega)$, first-hyperpolarizability $\beta(-\omega_\sigma;\omega_1,\omega_2)$ and second-hyperpolarizability $\gamma(-\omega_\sigma;\omega_1,\omega_2,\omega_3)$.

Response	Pole	Residue property
$\alpha(-\omega;\omega)$	$\omega=\omega_{f0}$	One-photon transition matrix elements between the ground state $\lvert 0\rangle$ and the excited state $\lvert f\rangle$
$\beta(-\omega_\sigma;\omega_1,\omega_2)$	$\omega_2=\omega_{f0}$	Two-photon transition matrix elements between the ground state $\lvert 0\rangle$ and the excited state $\lvert f\rangle$
	$\omega_1=-\omega_{f0}$, $\omega_2=\omega_{g0}$	One-photon transition matrix elements between the excited states $\lvert f\rangle$ and $\lvert g\rangle$
	$\omega_1=-\omega_{f0}$, $\omega_2=\omega_{f0}$	Electric dipole moment of the excited state $\lvert f\rangle$
$\gamma(-\omega_\sigma;\omega_1,\omega_2,\omega_3)$	$\omega_3=\omega_{f0}$	Three-photon transition matrix elements between the ground state $\lvert 0\rangle$ and the excited state $\lvert f\rangle$
	$\omega_2=-\omega_{f0}$, $\omega_3=\omega_{g0}$	Two-photon transition matrix elements between the excited states $\lvert f\rangle$ and $\lvert g\rangle$
	$\omega_2=-\omega_{f0}$, $\omega_3=\omega_{f0}$	Linear electric dipole polarizability of the excited state $\lvert f\rangle$

where $\alpha_{zz}^{TLA}(0)$ is the static polarizability and $F(\omega)$ is the dispersion correction to the polarizability. $\alpha_{zz}^{TLA}(0)$ can be expressed as:

$$\alpha_{zz}^{TLA}(0)=2\frac{|\mu_z^{0f}|^2}{\hbar\omega_{f0}}, \tag{3.22}$$

where μ_z^{0f} indicates the transition dipole moment matrix element $\langle 0|\hat{\mu}|f\rangle$. The expression of $F(\omega)$ is:

$$F(\omega)=\frac{1}{1-\omega^2/\omega_{f0}^2}. \tag{3.23}$$

With the TLA we can also approximate β_{zzz}:

$$\beta_{zzz}^{TLA}(-\omega_\sigma;\omega_1,\omega_2)=\beta^{TLA}(0)G(\omega_1,\omega_2), \tag{3.24}$$

where the static component $\beta_{zzz}^{TLA}(0)$ is:

$$\beta_{zzz}^{TLA}(0)=3\frac{|\mu_z^{0f}|^2(\mu_z^{ff}-\mu_z^{00})}{\hbar^2\omega_{f0}^2}, \tag{3.25}$$

and the frequency dispersion factor $G(\omega_1,\omega_2)$ is:

$$G(\omega_1,\omega_2)=\frac{1-\omega_1\omega_2/\omega_{f0}^2}{(1-\omega_1^2/\omega_{f0}^2)(1-\omega_2^2/\omega_{f0}^2)(1-(\omega_1+\omega_2)^2/\omega_{f0}^2)}. \tag{3.26}$$

Eq. 3.24 was used by Oudar and Chemla [2] to study the para-nitroaniline molecule. From Eq. 3.25 we can therefore identify the characteristics that enhance the first hyperpolarizability for push-pull systems: a small excitation energy, a large difference between the ground and excited state dipole moments, and a large transition dipole moments between the ground and the excited state. Push-pull systems have a low-energy charge-transfer state, which generally consists of an excitation of the electron from the HOMO orbital (mainly localized on the donor) to the LUMO orbital (mainly localized on the acceptor). By increasing the relative strength of the D/A units and the length of the π-conjugated bridge the difference in energy between the ground and the excited state decreases and therefore the optical response is enhanced.

3.2 Practical computation of static optical properties

Considering that the SOS approximation is not an efficient way of computing the NLOPs with standard electronic structure methods, other techniques have been developed for this purpose. In this section, we will focus on the evaluation of static NLOPs. If the system is subject to a weak static electric field $\mathbf{F}$, the external potential defined in Eq. 3.2 is time-independent and defined as:

$$\hat{V} = -\mu_i F_i. \tag{3.27}$$

The perturbed ground state energy $E(\mathbf{F})$ can be then expressed as a Taylor expansion with respect to the field components as:

$$\begin{aligned} E(\mathbf{F}) = \langle\psi|\hat{H}_0 + \hat{V}|\psi\rangle = E_0 + \frac{\partial E(\mathbf{F})}{\partial F_i}\bigg|_0 F_i + \frac{\partial^2 E(\mathbf{F})}{\partial F_i \partial F_j}\bigg|_0 F_i F_j + \frac{\partial^3 E(\mathbf{F})}{\partial F_i \partial F_j \partial F_k}\bigg|_0 F_i F_j F_k + \\ + \frac{\partial^4 E(\mathbf{F})}{\partial F_i \partial F_j \partial F_k \partial F_l}\bigg|_0 F_i F_j F_k F_l + ... \end{aligned} \tag{3.28}$$

The expansion coefficients are the static linear and nonlinear optical properties:

$$\mu_i(0) = \frac{\partial E(\mathbf{F})}{\partial F_i}\bigg|_0 \tag{3.29}$$

$$\alpha_{ij}(0) = \frac{\partial^2 E(\mathbf{F})}{\partial F_i \partial F_j}\bigg|_0 \tag{3.30}$$

$$\beta_{ijk}(0) = \frac{\partial^3 E(\mathbf{F})}{\partial F_i \partial F_j \partial F_k}\bigg|_0 \tag{3.31}$$

$$\gamma_{ijkl}(0) = \frac{\partial^4 E(\mathbf{F})}{\partial F_i \partial F_j \partial F_k \partial F_l}\bigg|_0 \tag{3.32}$$

The NLOPs are therefore derivatives of the perturbed energy and can be evaluated by using two main computational approaches, namely numerical or analytic differentiation methods.

3.2.1 Numerical differentiation

The numerical or finite field method (FF) is the easiest and most straightforward way to evaluate static optical properties. The idea of the FF method is that if $f(x)$ is a function of one variable x it is possible to estimate its derivatives by using some specific values of the function. In fact, if h is a small perturbation of the variable x we can expand, according to the Taylor formula, the function $f(x+h)$ in terms of derivatives of $f(x)$, namely:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \frac{h^3}{3!}f'''(x) + .. + \frac{h^n}{n!}f^n(x) + O(h^n), \tag{3.33}$$

where $O(h^n)$ indicates that the polynomials with an exponent larger than n are neglected. The most simple estimation of the first derivative is then:

$$f'(x) \approx \frac{f(x+h) - f(x)}{h} + O(h). \tag{3.34}$$

The difference between the exact value of the derivative and the approximated one determines the accuracy of the numerical derivative. There are two different sources of errors in numerical differentiation: the rounding error and the truncation error. The rounding error arises because computers work with finite-precision numbers. Indeed, each number computed by a machine is not exact but it comes from a rounding operation, which plays a major role in numerical differentiation. When we use finite precision numbers, the digits after a certain precision are assigned randomly by the calculator. This issue is particularly relevant when computing differences between nearly equal numbers. In fact, if x and y are two numbers with p equal digits on a total of k, the difference between the two is a number with $k-p$ digits of significance, and the remaining k-digits are assigned randomly. In the case of numerical derivatives, the magnitude of this contribution is further enhanced by dividing by the small parameter h . One way to reduce this error is by using larger values of the interval h.

The other source of error is the truncation error, which comes from the contributions of the higher order derivatives neglected in Eq. 3.34. The formula used for defining the numerical derivative in Eq. 3.34 is called forward difference, another expression employed for reducing the truncation error is the so-called central difference formula, *i.e.*:

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} + O(h^2), \tag{3.35}$$

which, according to Eq. 3.33, reduces the error from $O(h)$ to $O(h^2)$. Moreover, the truncation error can be further reduced by combining differences from different steps. If for example we define the differences $d_h^{(1)}$ and $d_{2h}^{(1)}$ (where $x = 0$ for convenience) as:

$$d_h^{(1)} = \frac{f(h) - f(-h)}{2h} \tag{3.36}$$

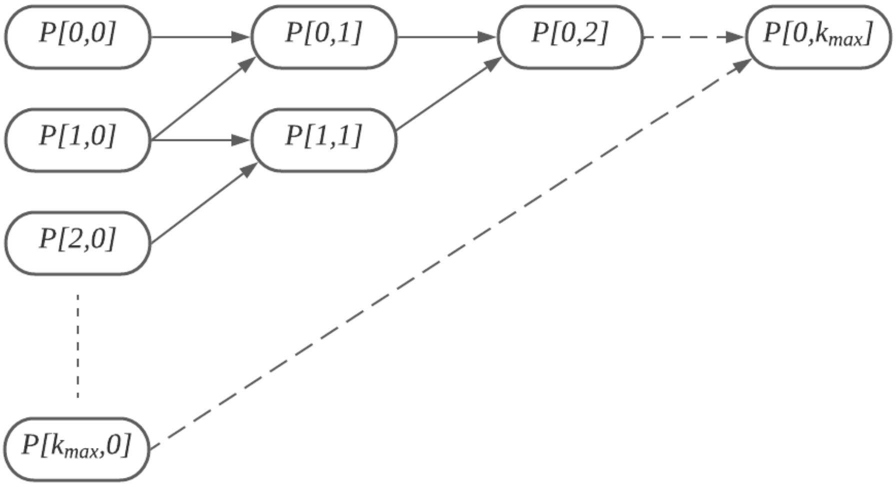


Figure 3.1: Schematic representation of the Romberg triangle. The arrows identify the elements which are combined for each step of the triangle.

$$d_{2h}^{(1)} = \frac{f(2h) - f(-2h)}{2h}. \tag{3.37}$$

We notice that, by simple algebraic operations, we can obtain the expression of $f'(0)$ as:

$$f'(0) \approx \frac{4d_h^{(1)} - d_{2h}^{(1)}}{3} + O(h^4), \tag{3.38}$$

obtaining an approximation of the fourth order with respect to h. This expression can be generalized by an arbitrary number of steps k resulting in the so-called Rutishauser–Romberg (RR) procedure [3–5]. Being a an arbitrary value for the quotient, the expressions of Eqs 3.36 and 3.37 can be then generalized as:

$$d_{kh}^{(1)} = \frac{f(a^k h) - f(-a^k h)}{2a^k h}. \tag{3.39}$$

The RR procedure consists in evaluating f on a geometrical sequence of points with the step a, namely: 0, $\pm h$, $\pm ah$, $\pm a^2 h$, ..., $\pm a^{k_{max}} h$. The expressions of Eq. 3.39 for different k can be combined in order to remove higher order differentiation terms as observed in Eq. 3.38 by the recursive formula:

$$P[k,m] = \frac{a^{2m} P[k-1,m] - P[k-1,m+1]}{a^{2m} - 1} \qquad k = 1, 2, ..., k_{max}, \tag{3.40}$$

where $P[k,m]$ is the required derivative of the m-th iteration of the RR procedure. For first order derivatives, the starting points of the recursive procedure are $P[k,0] = d_{kh}^{(1)}$. The result of this analysis can be represented as a triangle (Fig. 3.1) which can be used to monitor the convergence of the derivative. The values in the RR triangle in principle tend to converge when moving from left to right and from down to up. The value with the lowest numerical error is chosen by evaluating the minimum value of the matrix $\epsilon[k,m]$, that contains the differences between the $P[k,m]$ in consecutive lines, namely:

$$min\{\epsilon[k,m]\} = min\{P[k+1,m] - P[k,m]\} = \epsilon[\tilde{m},\tilde{k}], \tag{3.41}$$

where $\tilde{m}$ and $\tilde{k}$ are the indices of the matrix ϵ associated with the minimum value. Then, according to this scheme, the corresponding number in the RR procedure with the lowest numeric error is $P[\tilde{k}+1,\tilde{m}]$. This formulation can be extended to higher-order derivatives, changing the initial steps of the RR procedure. However, it must be taken into account that the higher the order of the derivative, the higher the numerical error associated. The expressions of the central difference for the second, third, and fourth numerical derivatives are:

$$d^{(2)}_{kh} = \frac{f(-a^k h) + f(a^k h) - 2f(0)}{(a^k h)^2} \tag{3.42}$$

$$d^{(3)}_{kh} = 3\frac{-f(-a^{k+1}h) + af(-a^k h) - af(a^k h) + f(a^{k+1}h)}{a(a^2-1)(a^k h)^3} \tag{3.43}$$

$$d^{(4)}_{hk} = 12\frac{f(-a^{k+1}h) - a^2 f(-a^k h) + 2(a^2-1)f(0) - a^2 f(a^k h) + f(a^{k+1}h)}{a^2(a^2-1)(a^k h)^4} \tag{3.44}$$

3.2.1.1 Computation of static NLOP by numerical differentiation

By using the RR presented in this section, it is possible to compute the static NLOPs. In fact, if for instance the function $f(x)$ is the perturbed energy $E(\mathbf{F})$ and the step h is the amplitude $\mathbf{F}$ of the field, according to Eqs. 3.39, 3.42-3.44, it is possible to obtain the expression of μ, α, β, and γ as first, second, third, and fourth order numerical derivatives of $E(\mathbf{F})$. This method can be applied to any electronic structure calculation for which field-dependent energies can be calculated. For HF or DFT calculations, the static electric field corrects the expression of the Fock operator by including another term in the one-electron Hamiltonian at each step of the SCF procedure:

$$\langle\phi|\hat{f}|\phi\rangle \rightarrow \langle\phi|\hat{f}|\phi\rangle + \langle\phi|\hat{r}_j|\phi\rangle\, F_j, \tag{3.45}$$

where $\hat{r}_j$ is the position vector operator. This change on the Fock operator will perturb the converged orbitals, which afterward depend on the magnitude of the electric field. Usually, this procedure does not increase the computational cost of the energy calculation. However, this calculation generally is more time-consuming because the SCF procedure requires more steps to reach convergence. In post-HF methods, the effects of the electric field are included in the correlation energy by constructing the excitation amplitudes with the field-dependent orbitals. As explained in the previous section, for computing the NLOPs, the energies must be computed with a large number of digits to reduce rounding errors. However, it is worth remembering that in practical QC computations several truncation schemes and cutoffs are employed, and these approximations affect the numerical precision. It is therefore advised to use the tightest criteria affordable for the QC computations. For example, in the SCF all two-electron integrals larger than 10^{-12} should be included, and a threshold at least of 10^{-10} a.u. should be used for the convergence of the SCF procedure, while for CC calculations a threshold at least of 10^{-8} a.u. should be chosen for solving the CC equations [4, 6, 7]. The magnitude of the electric fields must also be selected carefully in the RR procedure. If the fields are too small (smaller than 10^{-4} a.u.) the energy differences will suffer from large rounding errors; while if the fields are too large the SCF might converge to an excited state. The value of the quotient a is also relevant for reducing the truncation error. Generally, it is taken as 2, however, it has been observed that the efficiency of the RR procedure can be increased by reducing the value of a in a range between 1 and 2. In particular, optimal results have been obtained with $a = \sqrt{2}$ [4].

3.2.2 Analytic computation of static optical properties

Analytic computations of NLOPs rely on response theory. According to RS perturbation theory (see Chapter 2), the electronic energy of a system under a small linear perturbation defined by a parameter λ can be expressed as:

$$E(\lambda, \mathbf{p}) = \langle \psi(\mathbf{p}) | \hat{H}^{\lambda} | \psi(\mathbf{p}) \rangle \,, \tag{3.46}$$

where with $\mathbf{p} = (p1, p2, ...)$ we indicate the parameters of the electronic wavefunction, such as the orbital parameters for an HF/DFT calculation, or the coefficients of the determinants for a CI calculation, while $\hat{H}^{\lambda}$ is defined in Eq. 2.23. The optimal parameters obtained from a converged energy calculation are indicated as $\mathbf{p}_0$, and the optimal energy is defined as:

$$E(\lambda) = E(\lambda, \mathbf{p}_0) \tag{3.47}$$

The first order derivative of $E(\mathbf{F})$ with respect to the parameter λ can be written as:

$$\frac{dE(\lambda)}{d\lambda} = \frac{\partial E(\lambda, \mathbf{p}_0)}{\partial \lambda} + \sum_i \left. \frac{\partial E(\lambda, \mathbf{p})}{\partial p_i} \right|_{\mathbf{p}=\mathbf{p}_0} \frac{\partial p_i}{\partial \lambda} . \tag{3.48}$$

For a stationary wavefunction the second term of Eq. 3.48 is equal to 0, in fact we have:

$$\left. \frac{\partial E(\lambda, \mathbf{p})}{\partial p_i} \right|_{\mathbf{p}=\mathbf{p}_0} = 0 \qquad \forall \, p_i . \tag{3.49}$$

As a result of this expression, the coefficients $\partial p_i / \partial \lambda$, which describe how the wavefunction parameters evolve with respect to the perturbation, are not necessary for the determination of first-order derivatives. By writing explicitly Eq. 3.48 we obtain the Hellmann–Feynman theorem [8] for a wavefunction stationary with respect to all wavefunction parameters:

$$\frac{dE(\lambda)}{d\lambda} = \langle \psi(\mathbf{p}_0) | \frac{\partial \hat{H}^{\lambda}}{\partial \lambda} | \psi(\mathbf{p}_0) \rangle \,. \tag{3.50}$$

The second derivative with respect to λ can be obtained by deriving Eq. 3.48:

$$\frac{d^2E}{d^2\lambda} = \sum_i \left. \frac{\partial^2 E(\lambda, \mathbf{p})}{\partial\lambda\partial p_i} \right|_{\mathbf{p}=\mathbf{p}_0} \frac{\partial p_i}{\partial\lambda}. \tag{3.51}$$

For second-order derivatives, the variation of the parameters $\mathbf{p}$ can be determined by deriving Eq. 3.49, and we obtain the following set of equations:

$$\sum_j \left. \frac{\partial^2 E(\lambda, \mathbf{p})}{\partial p_i \partial p_j} \right|_{\mathbf{p}=\mathbf{p}_0} \frac{\partial p_i}{\partial\lambda} = -\frac{\partial^2 E(\lambda, \mathbf{p})}{\partial\lambda\partial p_i} \tag{3.52}$$

These are the so-called linear response equations. In the next section, we will show how these equations are expressed for the HF and DFT methods, and how to use those results to compute the polarizability.

3.2.2.1 Coupled Perturbed Hartree Fock and DFT

We start the derivation from the HF wavefunction Φ_0. The perturbed wavefunction ϕ can be expressed through an exponential parameterization:

$$|\phi(\boldsymbol{\kappa})\rangle = e^{\hat{\kappa}(\boldsymbol{\kappa})} |\Phi_0\rangle, \tag{3.53}$$

where $\hat{\kappa}$ is the orbital excitation operator which includes the one-particle excitation parameters from occupied to virtual orbitals $\boldsymbol{\kappa} = \{\kappa_{ar}\}$, namely:

$$\hat{\kappa}(\boldsymbol{\kappa}) = \sum_a^{occ} \sum_r^{vir} (\kappa_{ar} \hat{c}_r^\dagger \hat{c}_a - \kappa_{ar}^* \hat{c}_a^\dagger \hat{c}_r), \tag{3.54}$$

where the operators $\hat{c}_r^\dagger$ and $\hat{c}_a$ are, respectively, the creation and destruction operators in the second quantization formalism. The expression of the HF energy in the presence of a perturbation field is:

$$E^{HF}(\boldsymbol{F}, \boldsymbol{\kappa}) = \langle \Phi(\boldsymbol{\kappa}) | \hat{H}_0 - \hat{\mu}_i F_i | \Phi(\boldsymbol{\kappa}) \rangle. \tag{3.55}$$

The expression of the polarizability in the CPHF approximation will be then:

$$\alpha_{ij}^{CPHF} = -\frac{\partial^2 E^{HF}(\mathbf{F})}{\partial F_i \partial F_j} = -\sum_a^{occ} \sum_r^{vir} \left. \frac{\partial^2 E^{HF}(\mathbf{F}, \boldsymbol{\kappa})}{\partial F_i \partial\kappa_{ar}} \right|_{\boldsymbol{\kappa}=0} \frac{\partial\kappa_{ar}}{\partial F_j}. \tag{3.56}$$

The perturbed electronic gradient can be evaluated by some algebraic operations as:

$$\left. \frac{\partial^2 E^{HF}}{\partial F_i \partial\kappa_{ar}} \right|_{k=0} = -2 \langle r | \hat{\mu}_i | a \rangle. \tag{3.57}$$

For evaluating the coefficients $\partial\kappa_{ar}/\partial F_j$ we need to compute the linear response equations 3.52. The Hessian of E^{HF} for the parameters k is:

$$\left.\frac{\partial^2 E^{HF}(\boldsymbol{\kappa})}{\partial\kappa_{ar}\partial\kappa_{bs}}\right|_{k=0} = 2(A_{ar,bs} + B_{ar,bs}), \tag{3.58}$$

where:

$$A_{ar,bs} = (\epsilon_r - \epsilon_a)\delta_{ab}\delta_{rs} + \langle ra|bs\rangle - \langle ra|sb\rangle \tag{3.59}$$

$$B_{ar,bs} = \langle rb|as\rangle - \langle rb|sa\rangle\,. \tag{3.60}$$

In which ϵ_a and ϵ_r are the single-particle energies of occupied and virtual orbitals, while $\langle ra|bs\rangle$ and $\langle ra|sb\rangle$ represent, respectively, the Coulomb and the exchange integrals. Inserting Eq. 3.58 and Eq. 3.57 into Eq. 3.52 we obtain the CPHF equations:

$$\sum_b^{occ}\sum_s^{vir}(A_{ar,bs} + B_{ar,bs})\frac{\partial\kappa_{bs}}{\partial F_j} = \langle r|\hat{\mu}_j|a\rangle\,. \tag{3.61}$$

An analogous formulation can be obtained for DFT calculations by slightly modifying the A and B matrices (see below). The solution of the linear Eqs. 3.61 is found by a combination of direct and iterative methods implemented for a solution of large systems of linear equations called direct inversion in the iterative subspace (DIIS) method [9, 10]. The computation of the first hyperpolarizability in principle requires solving the quadratic response equations, however, according to the $2n + 1$ rule [11, 12], it can be obtained as a function of the linear response vectors. This rule states that by the knowledge of a wavefunction of order n it is possible to obtain derivatives on the energy up to order $2n + 1$.

3.2.2.2 Comparison between numerical and analytical derivatives

In principle, for static optical properties, analytic and numerical derivatives are expected to give identical results, both approaches having advantages and limitations. In fact, through FF calculations, we can compute derivatives in a simple way for each computational method which allows the calculation of the energy perturbed by an electric field, while analytic derivatives need complicated implementations which are specific for each method. On the other hand, analytic derivatives are not sensitive to numerical errors, which are significant for FF derivatives. Moreover, analytic derivatives allow to compute efficiently cross derivatives, for which several single-point calculations are necessary in the case of numerical ones (see for example, ref. [7] for the expression of each component of the tensor β). However, solving the CPHF equations might result challenging (or even out of reach) for big systems, because it requires larger memory (RAM) than energy computations. Therefore, we generally advise employing analytic derivatives when available and computationally feasible, and performing FF differentiation on the highest order analytical derivatives available (for example, compute the first-order numerical derivatives of analytic α to obtain β).

It is worth reminding that numerical procedures can be applied only for static properties, however, they can be combined with analytical frequency-dependent quantities to obtain some specific dynamic properties. For example, the numerical derivative of the dynamic polarizability $\alpha_{ij}(\omega;\omega)$ allows to evaluate the Pockel effect $\beta_{ijk}(\omega;\omega,0)$ and the numerical derivative of the analytic first-hyperpolarizability $\beta_{ijk}(2\omega;\omega,\omega)$ is need to compute the EFISHG response $\gamma_{ijkl}(2\omega;\omega,\omega,0)$.

3.2.3 Inclusion of electron correlation in static NLOP calculations

In NLOP calculations the inclusion of electron correlation is fundamental. Nowadays, the reference wavefunction method for computing NLOPs including correlation is CCSD(T). However, its application is limited to small systems considering the big computational effort needed. When CCSD(T) is out of reach, CCSD is generally considered the best method. However, MP2 is also considered an excellent alternative in terms of computational savings. MP2 has been shown to give accurate hyperpolarizabilities for various systems (see ref [6] and Chapter 4) with a more favorable computational scaling with respect to CCSD. Higher expansion series MPn generally do not considerably improve the MP2 description of hyperpolarizabilities [6]. As an example, with respect to CCSD(T), we compare in Table 3.2 the performances of HF, MP2, MP3, and CCSD methods for the calculation of γ_{zzzz} for 5 polydiacetylene molecules (PDAn) with increasing size of the π-bridge ($n = 1 - 5$), where the z axis is orientated along the direction of the principal inertia axis (see Chapter 4 for more information on these geometries).

Table 3.2: Comparison of γ_{zzzz} computed with different wavefunction methods for PDAn. The values in the table are expressed in $[10^5\, a.\,u.]$. The values between parenthesis represent the ratio with the reference CCSD(T) values. The derivatives have been computed numerically, see Chapter 4 for more details on the numerical procedure employed.

Molecule	HF	MP2	MP3	CCSD	CCSD(T)
PDA1	0.8(0.7)	1.3(1.1)	1.1(1.0)	1.1(1.0)	1.1(1.0)
PDA2	5.4(0.7)	9.0(1.2)	7.6(1.0)	6.7(0.9)	7.6(1.0)
PDA3	17.4(0.7)	31.6(1.3)	25.2(1.0)	20.3(0.8)	24.4(1.0)
PDA4	38.0(0.7)	64.0(1.2)	51.5(1.0)	44.9(0.8)	53.0(1.0)
PDA5	65.8(0.6)	121.1(1.1)	90.9(0.8)	83.9(0.8)	111.0(1.0)

From this table, we have that HF generally underestimates the γ values but remains with a constant ratio with respect to the reference CCSD(T). MP3 and CCSD show very similar performances, MP3 being slightly better than CCSD. On the other hand, MP2 despite overestimating the CCSD(T) values, gives excellent results considering the saving of computational efforts. The selection of an appropriate basis set is also fundamental for the accurate evaluation of NLOPs. It has been observed [6, 13–15] that polarization and diffuse functions are generally needed for the accurate computation of first- and second-hyperpolarizabilities. On the other hand, passing from double to triple zeta basis sets gives a marginal improvement of the absolute magnitude of these properties.

3.2.3.1 Computing Static NLOP with DFT

For medium- and large-size systems, evaluating NLOPs through wavefunction methods is too computationally demanding. Nowadays, the best alternative is DFT. However, standard DFAs are generally designed and optimized for reproducing energies and not molecular properties. As a result, the accurate prediction of NLOPs is still very challenging for DFAs. The main issues for computing high order NLOPs are due to the wrong asymptotic behavior of standard DFAs, which is related to the SIE [6, 16–20]. It has been observed that LDA and GGA functionals tend to drastically overestimate first- and second-hyperpolarizabilities because of the delocalization error [7, 20]. For systems in which the exchange is dominant with respect to correlation, this behavior can be explained by an excessive delocalization of the electronic density which leads to an overestimation of the response properties. By increasing the percentage of exact exchange this phenomenon is partially corrected, however, including the exact HF exchange might lead to an overlocalization of the electronic density and, therefore, an underestimation of the property (such as in the HF case). Excellent static NLOPs are usually obtained with hybrid functionals with an intermediate percentage of HF exchange (such as M06-2X [18, 20]) and range-separated functionals (in which the high percentage of HF exchange at a long-range assures accurate results [16, 19]). For the range-separated functionals, it is also relevant the value of the parameter ω in Eq. 2.82, which controls the balance between the short range and long range part.

For instance, in Table 3.3 we report, for the calculation of γ_{zzzz} in the PDAn molecules, the performances of 5 functionals with the same approximation for the correlation and different ways of including the exchange part: BLYP which is a GGA functional, B3LYP and BH&HLYP which are global hybrids with respectively 20% and 50% of exchange, CAM-B3LYP($\omega = 0.33$) and LC-BLYP($\omega = 0.47$) which are two range separated functionals with a percentage of explicit exchange included between 20-65% and 0-100% respectively.

Table 3.3: Comparison of γ_{zzzz} computed with different DFAs for the PDAn molecules. The values in the table are expressed in $[10^5\, a.u.]$. The values between parenthesis represent the ratio with the reference CCSD(T) values. The derivatives have been computed numerically. More details on the numerical procedure employed can be found in Chapter 5.

Molecule	BLYP	B3LYP	BH&HLYP	CAM-B3LYP	LC-BLYP
PDA1	1.7(1.5)	1.5(1.3)	1.2(1.0)	1.2(1.1)	1.0(0.8)
PDA2	17.8(2.3)	13.9(1.8)	15.9(2.1)	9.7(1.3)	6.7(0.9)
PDA3	98.8(4.0)	67.4(2.8)	37.8(1.5)	37.8(1.5)	22.3(0.9)
PDA4	365.6(6.9)	212.5(4.0)	98.5(1.9)	96.4(1.8)	49.6(0.9)
PDA5	1022.5(9.2)	502.0(4.5)	196.0(1.8)	188.1(1.7)	95.4(0.9)

As observed in this table, by increasing the amount of HF exchange γ_{zzzz} decreases, passing from a huge overestimation with BLYP to a slight underestimation with LC-BLYP (and HF in Table 3.2). The best results for these molecules are obtained with the LC-BLYP functionals. Worth mentioning in this context are the novel DFAs, which tend to minimize the delocalization error and therefore improve the accuracy of NLOPs computations by optimizing the parameter ω with different strategies [16–19].

3.3 Time-dependent response theory

If the external potential is time-dependent and periodic, it is necessary to construct the dynamic response functions to compute molecular properties. If $\hat{V}(\lambda, t)$ is the perturbation potential that oscillates with a frequency ω, the time-dependent Hamiltonian is:

$$\hat{H}(\lambda, t) = \hat{H}_0 + \hat{V}(\lambda, t) \tag{3.62}$$

The exact wavefunction $\psi(t)$ will be therefore time-dependent and can be obtained by solving the time-dependent Schrödinger equation, Eq. 2.1. For computing response properties we consider the wavefunction $\psi(\boldsymbol{\kappa}, t)$ dependent of a general set of parameters $\boldsymbol{\kappa}_i$ as for the time-independent case. Being $\mathcal{F}(t)$ a unitary matrix, it is possible to rotate the wavefunction as:

$$|\bar{\psi}(\boldsymbol{\kappa}, t)\rangle = e^{-i\mathcal{F}(t)} |\bar{\psi}(\boldsymbol{\kappa}, t)\rangle . \tag{3.63}$$

Including Eq. 3.63 in Eq. 2.1 we can obtain the following Eq. :

$$\left[\hat{H}(t) - i\frac{\partial}{\partial t}\right] |\bar{\psi}(\boldsymbol{\kappa}, t)\rangle = Q(t)|\bar{\psi}(\boldsymbol{\kappa}, t)\rangle, \tag{3.64}$$

in which $Q(t)$ is defined as:

$$Q(t) = \frac{\delta\mathcal{F}(t)}{\delta t} = \langle\bar{\psi}(\boldsymbol{\kappa}, t)| \left(\hat{H}(t) - i\frac{\partial}{\partial t}\right) |\bar{\psi}(\boldsymbol{\kappa}, t)\rangle , \tag{3.65}$$

which is named quasi-energy and it is usually expressed as the time-average $\mathcal{Q}$:

$$\mathcal{Q} = \int_0^{\frac{2\pi}{\omega}} Q(t)dt = \{Q(t)\}_T. \tag{3.66}$$

If we restrict our analysis to variational methods, by differentiating Eq. 3.66 we can obtain the response functions analogously to the energy in the time-independent case. The time-averaged quasi-energy in the time-independent limit recovers the expression of the ground state energy. The advantage of the quasi-energy is given by the time-averaged variational conditions which determine the time evolution of $\mathcal{Q}$:

$$\frac{\partial\mathcal{Q}}{\partial\kappa_i} = 0. \tag{3.67}$$

Moreover, it admits the TD extension of the Hellmann-Feynmann theorem [21]:

$$\frac{dQ}{d\lambda} = \left\{ \left\langle \bar{\psi}(\{\kappa_i\},t) \middle| \frac{\partial H(\lambda,t)}{\partial \lambda} \middle| \bar{\psi}(\boldsymbol{\kappa},t) \right\rangle \right\}_T . \tag{3.68}$$

For computing explicitly the terms in Eq. 3.68, we can express the external perturbation in terms of its Fourier expansion as:

$$\hat{V}(\lambda^{\{\omega_k\}}t) = \sum_k \lambda^{\omega_k} \hat{V} e^{-i\omega_k t}, \tag{3.69}$$

where $\{\omega_k\}$ are multiple of the perturbation frequency ω, λ^{ω_k} are the perturbation parameters, and $\hat{V}$ is a time-independent operator. The parameters of Eq. 3.69 are chosen in order to leave the Hamiltonian hermitian, therefore: $\hat{V} = \hat{V}^\dagger$, $\omega_{-k} = -\omega_k$ and $\lambda^{-\omega} = (\lambda^{\omega})^*$. From Eq. 3.69 we can obtain the derivative with respect to each perturbation parameter as:

$$\frac{\partial \hat{V}(\lambda^{\{\omega_k\}}t)}{\partial \lambda^{\omega_k}} = \hat{V} e^{-i\omega_k t}. \tag{3.70}$$

Therefore, by substituting Eq. 3.70 in Eq. 3.68 we can obtain:

$$\frac{dQ}{d\lambda^{\omega_0}} = \left\{ \left\langle \bar{\Psi}(\{\kappa_i\},t) \middle| \hat{V} \middle| \bar{\Psi}(\{\kappa_i\},t) \right\rangle e^{-i\omega_0 t} \right\}_T . \tag{3.71}$$

The expression in the bracket represents the expectation value of an operator $\hat{V}$ on a time-dependent wavefunction, it could be expressed in the perturbation theory framework as:

$$\begin{aligned} \langle \hat{V} \rangle (t) = {} & \langle\langle \hat{V} \rangle\rangle_0 + \sum_k \lambda^{\omega_k} \langle\langle \hat{V}; \hat{V} \rangle\rangle_{\omega_k} e^{-i\omega_k t} + \\ & + \frac{1}{2!} \sum_{k,l} \lambda^{\omega_k} \lambda^{\omega_l} \langle\langle \hat{V}; \hat{V}, \hat{V} \rangle\rangle_{\omega_k \omega_l} e^{-i(\omega_k+\omega_l)t} + \\ & + \frac{1}{3!} \sum_{k,l,m} \lambda^{\omega_k} \lambda^{\omega_l} \lambda^{\omega_m} \langle\langle \hat{V}; \hat{V}, \hat{V}, \hat{V} \rangle\rangle_{\omega_k \omega_l \omega_m} e^{-i(\omega_k+\omega_l+\omega_m)t} + \dots, \end{aligned} \tag{3.72}$$

where the terms $\langle\langle \hat{V} \rangle\rangle_0$, $\langle\langle \hat{V}; \hat{V} \rangle\rangle_{\omega_k}$, $\langle\langle \hat{V}; \hat{V}, \hat{V} \rangle\rangle_{\omega_k \omega_l}$ and $\langle\langle \hat{V}; \hat{V}, \hat{V}, \hat{V} \rangle\rangle_{\omega_k \omega_l \omega_m}$ are respectively the zero-order, linear, quadratic and cubic dynamic response functions which respectively rule how a time-dependent operator respond to perturbations of zeroth-, first-, second- and third-order. By including Eq. 3.72 into Eq. 3.71 and by employing the equality $\{e^{i\omega t}\}_T = \delta(\omega)$ we obtain:

$$\begin{aligned} \frac{dQ}{d\lambda^{\omega}} = {} & \langle\langle \hat{V} \rangle\rangle_0 \delta(\omega) + \sum_k \lambda^{\omega_k} \langle\langle \hat{V}; \hat{V} \rangle\rangle_{\omega_k} \delta(\omega_0 + \omega_k) + \\ & + \frac{1}{2!} \sum_{k,l} \lambda^{\omega_k} \lambda^{\omega_l} \langle\langle \hat{V}; \hat{V}, \hat{V} \rangle\rangle_{\omega_k \omega_l} \delta(\omega_0 + \omega_k + \omega_l) + \\ & + \frac{1}{3!} \sum_{k,l,m} \lambda^{\omega_k} \lambda^{\omega_l} \lambda^{\omega_m} \langle\langle \hat{V}; \hat{V}, \hat{V}, \hat{V} \rangle\rangle_{\omega_k \omega_l \omega_m} \delta(\omega_0 + \omega_k + \omega_l + \omega_m) + \dots, \end{aligned} \tag{3.73}$$

where energy conservation is provided by the Dirac δ which allows that the combination of the incident frequencies sums up to the response frequency, *i.e.*, $\omega_0 = \sum_k \omega_k$. Therefore, we can connect the response equations to the perturbation expansion of the time-averaged quasi-energy as:

$$\langle\langle \hat{V}; \hat{V} \rangle\rangle_{\omega} = \frac{d\mathcal{Q}}{d\lambda^{-\omega} d\lambda^{\omega}} \tag{3.74}$$

$$\langle\langle \hat{V}; \hat{V}, \hat{V} \rangle\rangle_{\omega_A \omega_B} = \frac{d\mathcal{Q}}{d\lambda^{-\omega_\sigma} d\lambda^{\omega_A} d\lambda^{\omega_B}} \tag{3.75}$$

$$\langle\langle \hat{V}; \hat{V}, \hat{V}, \hat{V} \rangle\rangle_{\omega_A \omega_B \omega_C} = \frac{d\mathcal{Q}}{d\lambda^{-\omega_\sigma} d\lambda^{\omega_A} d\lambda^{\omega_B} d\lambda^{\omega_C}}, \tag{3.76}$$

where in Eqs. 3.75 and 3.76 the response frequency ω_σ is respectively equal to $\omega_A + \omega_B$ and $\omega_A + \omega_B + \omega_C$. For finding the analytical expression of the response functions it is necessary to perturbatively expand the parameters $\boldsymbol{\kappa}$ which are time and perturbation dependent:

$$\boldsymbol{\kappa}(\lambda, t) = \boldsymbol{\kappa}^{(0)} + \boldsymbol{\kappa}^{(1)}(\lambda, t) + ..., \tag{3.77}$$

where for example the first terms $\boldsymbol{\kappa}^{(1)}$ can be expressed as:

$$\boldsymbol{\kappa}^{(1)}(\lambda, t) = \sum_{\omega_k} \boldsymbol{\kappa}^{\omega_k} \lambda^{\omega_k} e^{-i\omega_k t}. \tag{3.78}$$

The expression of the linear response function can be then obtained in terms of the $\boldsymbol{\kappa}^{\omega}$ parameters as:

$$\frac{d^2\mathcal{Q}}{d\lambda^{-\omega} d\lambda^{\omega}} = \sum_i \frac{\partial^2 \mathcal{Q}}{\partial\lambda^{-\omega} \partial\kappa_i^{\omega}} \frac{\partial\kappa_i^{\omega}}{\partial\lambda^{\omega}}. \tag{3.79}$$

The expression of $\partial\kappa_i^{\omega}/\partial\lambda^{\omega}$ are evaluated by deriving Eq. 3.67, resulting into the linear first-order response equations:

$$\sum_j \frac{\partial^2 \mathcal{Q}}{\partial\kappa_i^{-\omega} \partial\kappa_j^{\omega}} \frac{\partial\kappa_j^{\omega}}{\partial\lambda^{\omega}} = -\frac{\partial^2 \mathcal{Q}}{\partial\kappa_i^{-\omega} \partial\lambda^{\omega}}. \tag{3.80}$$

By solving these equations we can compute the dynamic polarizability $\alpha(-\omega; \omega)$. Equivalent equations can be obtained for the quadratic response functions to compute the first hyperpolarizability $\beta(-2\omega; \omega, \omega)$. However, as already observed for the static case, because of the $2n + 1$ rule [11, 12], it can be expressed in terms of the linear response vectors.

3.3.1 Applications to TD-HF and TD-DFT

Through response theory is possible to compute all the dynamic optical properties. In this thesis, we have employed response functions to compute the dynamic polarizability $\alpha(-\omega; \omega)$ and the dynamic first hyperpolarizability for the second harmonic generation

$\beta(-2\omega;\omega,\omega)$, which corresponds to the following linear and quadratic response properties:

$$\alpha_{ij}(-\omega;\omega) = \langle\langle \hat{\mu}_i; \hat{\mu}_j \rangle\rangle_\omega \tag{3.81}$$

$$\beta_{ijk}(-2\omega;\omega,\omega) = \langle\langle \hat{\mu}_i; \hat{\mu}_j, \hat{\mu}_k \rangle\rangle_{\omega\omega} . \tag{3.82}$$

For an HF wavefunction we can rewrite Eq. 3.80 as:

$$\left[\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} - \omega \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}\right] \begin{pmatrix} \boldsymbol{X}_j^{\boldsymbol{\omega}} \\ \mathbf{Y}_j^{\boldsymbol{\omega}} \end{pmatrix} = \begin{pmatrix} \mu_j \\ \mu_j \end{pmatrix}, \tag{3.83}$$

where the matrices $\mathbf{A}$ and $\mathbf{B}$ have been already defined, respectively, in Eqs. 3.59 and 3.60, and $\boldsymbol{X}_j^\omega$ and $\mathbf{Y}_j^\omega$ are the frequency-dependent linear response vectors which contain the coefficients of excitations and de-excitations. Solving these Eqs. we can then calculate α_{ij}^{TD-HF} and β_{ijk}^{TD-HF}. For example, the expression of $\alpha^{TD-HF}(-\omega,\omega)$ is:

$$\alpha^{TD-HF}(-\omega,\omega) = -\begin{pmatrix} \mu_i \\ \mu_i^* \end{pmatrix}^\dagger \begin{pmatrix} \boldsymbol{X}_j^\omega \\ \mathbf{Y}_j^{\omega *} \end{pmatrix} = -2 \sum_a^{occ.} \sum_r^{vir.} \langle a|\mu_i|r\rangle \, (\boldsymbol{X}_{j,ar}^\omega + \mathbf{Y}_{j,ar}^\omega), \tag{3.84}$$

where $\boldsymbol{X}_{j,ar}^\omega$ and $\mathbf{Y}_{j,ar}^\omega$ are the coefficients of the frequency-dependent linear response vectors which contain the excitations $a \rightarrow r$ and de-excitations $r \rightarrow a$. An extension of this formalism to DFT is given by the Runge and Groos theorem [22], which extends the Hohenberg-Kohn theorem to time-dependent Hamiltonian showing that there is a one-to-one correspondence between the time-dependent electronic density and the external potential. If the adiabatic approximation is employed, the time-dependent exchange-correlation functional can be approximated as:

$$\hat{V}_{XC}(\mathbf{r},t) \approx \frac{\delta E_{XC}[\rho]}{\delta\rho(\mathbf{r},t)} \tag{3.85}$$

The inclusion of the $\hat{V}_{XC}$ potential will slightly modify the expression of Eqs. 3.59 and 3.60, which will become:

$$A_{ar,bs} = (\epsilon_r - \epsilon_a)\delta_{ab}\delta_{rs} + 2\langle ra|bs\rangle - a_x \langle ab|rs\rangle + (1-a_x)\langle ra|f_{XC}|sb\rangle \tag{3.86}$$

$$B_{ar,bs} = 2\langle rb|as\rangle - a_x \langle rb|sa\rangle + (1-a_x)\langle ra|f_{XC}|sb\rangle , \tag{3.87}$$

where a_x is the amount of the exact exchange included in the definition of the DFA and f_{XC} is the exchange-correlation kernel, defined as:

$$f_{XC}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\delta V_{XC}(\mathbf{r}_1)}{\delta\rho(\mathbf{r}_2)} \tag{3.88}$$

3.3.1.1 Photoabsorption spectrum

Through TD-HF/TD-DFT it is possible to compute also the absorption spectrum of the system. In fact, by neglecting the perturbation on Eq. 3.83, we obtain the so-called Casida's Equations [23]:

$$\left[\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} - \omega_n \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}\right] \begin{pmatrix} \boldsymbol{X}_i^{\omega_n} \\ \mathbf{Y}_i^{\omega_n} \end{pmatrix} = 0 \tag{3.89}$$

By solving these pseudo-eigenvalues equations we can obtain the excitation frequencies ω_n from the ground state to the n-th excited state. The explicit computation of the inversion matrix is usually prohibitive for practical applications, therefore, usually, these equations are solved by the so-called Davidson method [24], which allows to obtain low-lying excited state frequencies through an iterative procedure that consists in expanding the matrices and vectors in smaller subspaces. The eigenvectors $\boldsymbol{X}_i^{\omega_n}$ and $\boldsymbol{Y}_i^{\omega_n}$ can be used for determining the oscillator strengths of these transitions:

$$f_n = \frac{2}{3}\omega_n \sum_i [(\boldsymbol{X}_i^{\omega_n} + \mathbf{Y}_i^{\omega_n})^T \mu_i] = \frac{2}{3}\omega_n \sum_i \langle 0|\hat{\mu}_i|n\rangle \langle n|\hat{\mu}_i|0\rangle \tag{3.90}$$

3.3.2 Computation of experimental NLO response

The degree of reliability of computational protocols is subjected to their capability to reproduce experimental results. However, the exact reproduction of experimental conditions through computational methods is a highly challenging task. In fact, beyond the chemical complexity of the system under investigation, the computational method should take into account effects such as frequency dispersion, environmental effects arising from interactions with the molecular surrounding, as well as dynamic fluctuations.

Frequency dispersion occurs when the frequency ω of the laser probe enters in resonance with one of the possible electron excitations of the investigated chromophore, i.e. when the chromophore absorbs light in a spectral region close to ω. Even if the use of tunable lasers in experimental setups allows in principle to select a laser frequency far from any absorption band of the chromophore, it is often not possible to completely remove frequency dispersion effects from experimental data. Accounting accurately for these effects is thus fundamental for obtaining reliable predictions of the NLO responses. As discussed previously, dynamic NLOP of large systems are not computationally affordable using correlated wavefunction-based approximations; therefore, they are generally evaluated at the DFT level by solving the TD-DFT equations. However, it has been demonstrated in several previous works [18, 25, 26] that XC functionals that reproduce well the absorption properties of π-conjugated molecules are not the same as those reliable for predicting the NLO responses, which requires making a compromise. A common alternative allowing includes frequency dispersion effects consists in correcting the static response (computed at any level of approximation including electron correlation) by a multiplicative factor, defined as the ratio between the dynamic and static hyperpolarizabilities (both computed at the DFT level). This approximation assumes that frequency dispersion and electron correlation effects can be treated independently, which is suitable for a variety of compounds [14].

Another key element for reproducing experimental data is the appropriate modeling of the environment. In the second part of this thesis, we focus on the simulation of the NLO responses of molecules in solution, for which a proper treatment of the interaction between the solute and the solvent is necessary. The solute might polarize or induce structural changes in the solute, which can strongly affect its NLO properties. Computational models for describing the solute-solvent interactions can be divided into two groups, the implicit and the explicit ones [27]. The PCM method presented in Chapter 2 belongs to the first group: the solvent is treated as a continuous medium characterized by its macroscopic dielectric constants. The interaction between the solvent and solute is then modeled by an interaction potential that implicitly takes into account the mutual polarization effects. This method often ensures a good approximation of homogenous solutions. However, it might result inadequate when there exist specific and spatially oriented solute-solvent interactions at the atomic level, such as hydrogen bonds or π-stacking. In these cases, models in which the solvent molecules are explicitly included in the calculations are more appropriate. A compromise combining both the implicit and explicit approaches consists in including explicitly a few solvent molecules around the solute, all embedded within a dielectric continuum. Another alternative is given by hybrid QM/QM' or QM/MM calculations. The two levels of theory can be combined in a single run treating the core-shell at a QM level and the outer shell (the solvent molecules) at a lower (QM') level, or with a polarizable force field (MM) accounting for polarization effects in a self-consistent way. The interface between the two regions can be treated at different levels of complexity such as by employing fixed charges models [28–30], induced dipoles [31] or fluctuating charges [32].

However, the theoretical approaches described above rely on the assumption that the investigated system and its environment have a rigid nature. Yet, recent theoretical studies have evidenced the significant role of structural fluctuations on the NLO responses of neutral or ionic chromophores in solution [33–38]. A cost-effective strategy for including structural dynamics effects consists of the sequential application of MD simulations and QM calculations. The accuracy of these approaches depends on the choice of the MM force field, which should be specifically parameterized to reproduce the most relevant structural features, i.e. those having the highest impact on the targeted NLO property. In Chapter 7, we report a study in which such a sequential MD/DFT approach is used to investigate the EFISHG response of an ionic amphiphilic dye in solution. In principle, dynamical fluctuations could also be modeled at a full quantum mechanical level by relying on *ab initio* molecular dynamics methods [39]. However, these approaches are nowadays still too computationally demanding to be employed in real-life systems.

Bibliography

[1] Norman, P.; Ruud, K.; Saue, T. *Principles and practices of molecular properties: theory, modeling, and simulations*; Wiley, 2018; p 480 P.

[2] Oudar, J. L.; Chemla, D. S. Hyperpolarizabilities of the nitroanilines and their relations to the excited state dipole moment. *J. Chem. Phys.* **1977**, *66*, 2664–2668.

[3] Richardson, L. F.; Gaunt, J. A. VIII. The deferred approach to the limit. *Philos. Trans. Royal Soc. A* **1927**, *226*, 299–361.

[4] Medved', M.; Stachová, M.; Jacquemin, D.; André, J.; Perpète, E. A. A generalized Romberg differentiation procedure for calculation of hyperpolarizabilities. *J. Mol. Struct.* **2007**, *847*, 39–46.

[5] Mohammed, A. A. K.; Limacher, P. A.; Champagne, B. Finding optimal finite field strengths allowing for a maximum of precision in the calculation of polarizabilities and hyperpolarizabilities. *J. Comput. Chem.* **2013**, *34*, 1497–1507.

[6] de Wergifosse, M.; Champagne, B. Electron correlation effects on the first hyperpolarizability of push–pull π-conjugated systems. *J. Chem. Phys.* **2011**, *134*, 074113.

[7] de Wergifosse, M.; Liégeois, V.; Champagne, B. Evaluation of the molecular static and dynamic first hyperpolarizabilities. *Int. J. Quantum Chem.* **2014**, *114*, 900–910.

[8] Feynman, R. P. Forces in Molecules. *Phys. Rev.* **1939**, *56*, 340–343.

[9] Pulay, P. Second and third derivatives of variational energy expressions: Application to multiconfigurational self-consistent field wave functions. *J. Chem. Phys.* **1983**, *78*, 5043–5051.

[10] Handy, N. C.; Schaefer, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.

[11] Karna, S. P.; Dupuis, M. Frequency dependent nonlinear optical properties of molecules: Formulation and implementation in the HONDO program. *J. Comput. Chem.* **1991**, *12*, 487–504.

[12] Thorvaldsen, A. J.; Ruud, K.; Kristensen, K.; Jørgensen, P.; Coriani, S. A density matrix-based quasienergy formulation of the Kohn–Sham density functional response theory using perturbation- and time-dependent basis sets. *J. Chem. Phys.* **2008**, *129*, 214108.

[13] Jacquemin, D.; Champagne, B.; Hättig, C. Correlated frequency-dependent electronic first hyperpolarizability of small push–pull conjugated chains. *Chem. Phys. Lett.* **2000**, *319*, 327–334.

[14] Castet, F.; Rodriguez, V.; Pozzo, J.; Ducasse, L.; Plaquet, A.; Champagne, B. Design and Characterization of Molecular Nonlinear Optical Switches. *Acc. Chem. Res.* **2013**, *46*, 2656–2665.

[15] Champagne, B.; Botek, E.; Nakano, M.; Nitta, T.; Yamaguchi, K. Basis set and electron correlation effects on the polarizability and second hyperpolarizability of model open-shell π-conjugated systems. *J. Chem. Phys.* **2005**, *122*, 114315.

[16] Nénon, S.; Champagne, B.; Spassova, M. I. Assessing long-range corrected functionals with physically-adjusted range-separated parameters for calculating the polarizability and the second hyperpolarizability of polydiacetylene and polybutatriene chains. *Phys. Chem. Chem. Phys.* **2014**, *16*, 7083–7088.

[17] Wang, C.; Yuan, Y.; Tian, X. Assessment of range-separated exchange functionals and nonempirical functional tuning for calculating the static second hyperpolarizabilities of streptocyanines. *J. Comput. Chem.* **2017**, *38*, 594–600.

[18] Lescos, L.; Sitkiewicz, S. P.; Beaujean, P.; Blanchard-Desce, M.; Champagne, B.; Matito, E.; Castet, F. Performance of DFT functionals for calculating the second-order nonlinear optical properties of dipolar merocyanines. *Phys. Chem. Chem. Phys.* **2020**, *22*, 16579–16594.

[19] Besalú-Sala, P.; Sitkiewicz, S. P.; Salvador, P.; Matito, E.; Luis, J. M. A new tuned range-separated density functional for the accurate calculation of second hyperpolarizabilities. *Phys. Chem. Chem. Phys.* **2020**, *22*, 11871–11880.

[20] Naim, C.; Castet, F.; Matito, E. Impact of van der Waals interactions on the structural and nonlinear optical properties of azobenzene switches. *Phys. Chem. Chem. Phys.* **2021**, *23*, 21227–21239.

[21] Christiansen, O.; Jørgensen, P.; Hättig, C. Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quantum Chem.* **1998**, *68*, 1–52.

[22] Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

[23] Casida, M. E. In *Recent Developments and Applications of Modern Density Functional Theory*; Seminario, J., Ed.; Theoretical and Computational Chemistry; Elsevier, 1996; Vol. 4; pp 391–439.

[24] Davidson, E. R. The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. *J. Comput. Phys.* **1975**, *17*, 87–94.

[25] Champagne, B.; Perpète, E. A.; Jacquemin, D.; van Gisbergen, S. J. A.; Baerends, E.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. Assessment of Conventional Density Functional Schemes for Computing the Dipole Moment and (Hyper)polarizabilities of Push−Pull π-Conjugated Systems. *J. Phys. Chem. A* **2000**, *104*, 4755–4763.

[26] Chołuj, M.; Alam, M. M.; Beerepoot, M. T. P.; Sitkiewicz, S. P.; Matito, E.; Ruud, K.; Zaleśny, R. Choosing Bad versus Worse: Predictions of Two-Photon-Absorption Strengths Based on Popular Density Functional Approximations. *J. Chem. Theory Comput.* **2022**, *18*, 1046–1060.

[27] Medved', M.; Budzák, Š.; Bartkowiak, W.; Reis, H. In *Handbook of Computational Chemistry*; Leszczynski, J., Kaczmarek-Kedziera, A., Puzyn, T., G. Papadopoulos, M., Reis, H., K. Shukla, M., Eds.; Springer International Publishing: Cham, 2017; pp 741–794.

[28] Nielsen, C. B.; Christiansen, O.; Mikkelsen, K. V.; Kongsted, J. Density functional self-consistent quantum mechanics/molecular mechanics theory for linear and nonlinear molecular properties: Applications to solvated water and formaldehyde. *J. Chem. Phys.* **2007**, *126*, 154112.

[29] Hidalgo Cardenuto, M.; Champagne, B. QM/MM investigation of the concentration effects on the second-order nonlinear optical responses of solutions. *J. Chem. Phys.* **2014**, *141*, 234104.

[30] Le Breton, G.; Bonhomme, O.; Benichou, E.; Loison, C. First hyperpolarizability of water in bulk liquid phase: long-range electrostatic effects included via the second hyperpolarizability. *Phys. Chem. Chem. Phys.* **2022**, *24*, 19463–19472.

[31] Steindal, A. H.; Ruud, K.; Frediani, L.; Aidas, K.; Kongsted, J. Excitation Energies in Solution: The Fully Polarizable QM/MM/PCM Method. *J. Phys. Chem. B* **2011**, *115*, 3027–3037.

[32] Giovannini, T.; Ambrosetti, T.; C., C. A polarizable embedding approach to second harmonic generation (SHG) of molecular systems in aqueous solutions. *Theor. Chem. Acc.* **2018**, *137*.

[33] Pielak, K.; Tonnelé, C.; Sanguinet, L.; Cariati, E.; Righetto, S.; Muccioli, L.; Castet, F.; Champagne, B. Dynamical behavior and second harmonic generation responses in acido-triggered molecular switches. *J. Phys. Chem. C* **2018**, *122*, 26160.

[34] Lescos, L.; Beaujean, P.; Tonnelé, C.; Aurel, P.; Blanchard-Desce, M.; Rodriguez, V.; de Wergifosse, M.; Champagne, B.; Muccioli, L.; Castet, F. Self-assembling, structure and nonlinear optical properties of fluorescent organic nanoparticles in water. *Phys. Chem. Chem. Phys.* **2021**, *23*, 23643–23654.

[35] Ramos, T. N.; Canuto, S.; Champagne, B. Unraveling the Electric Field-Induced Second Harmonic Generation Responses of Stilbazolium Ion Pairs Complexes in Solution Using a Multiscale Simulation Method. *J. Chem. Inf. Model.* **2020**, *60*, 4817–4826.

[36] Ramos, T. N.; Castet, F.; Champagne, B. Second Harmonic Generation Responses of Ion Pairs Forming Dimeric Aggregates. *J. Phys. Chem. B* **2021**, *125*, 3386–3397.

[37] Bouquiaux, C.; Tonnelé, C.; Castet, F.; Champagne, B. Second-Order Nonlinear Optical Properties of an Amphiphilic Dye Embedded in a Lipid Bilayer. A Combined Molecular Dynamics–Quantum Chemistry Study. *J. Phys. Chem. B* **2020**, *124*, 2101–2109.

[38] Bouquiaux, C.; Castet, F.; Champagne, B. Unravelling the Effects of Cholesterol on the Second-Order Nonlinear Optical Responses of Di-8-ANEPPS Dye Embedded in Phosphatidylcholine Lipid Bilayers. *J. Phys. Chem. B* **2021**, *125*, 10195–10212.

[39] Dyer, P. J.; Cummings, P. T. Hydrogen bonding and induced dipole moments in water: Predictions from the Gaussian charge polarizable model and Car-Parrinello molecular dynamics. *J. Chem. Phys.* **2006**, *125*, 144519.

Chapter 4

Acceleration and enhancement wavefunction methods for calculations of static linear and nonlinear optical properties

This chapter has been realized in the framework of a collaboration involving:
Besalú-Sala P. (Univ. of Girona), Zaleśny R. (Univ. of Wrocław), and Luis J. M. (Univ. of Girona)

4.1 Introduction

In several fields as different as molecular biology or material science, the demand for functional materials bearing specific electro-optical features is increasing yearly [1–3] as for instance in the construction of two-photon absorption or non-invasive three-dimensional fluorescence microscopy devices [4, 5]. The key compounds used for building such devices are, however, difficult to design or optimize since most of the newest applications are based on the nonlinear response of these molecular units upon interaction with light, which is a physical process difficult to model or rationalize.

As we have already observed in Section 3.2, the energy of a molecule subjected to an external static electric field **F** can be expressed as a Taylor expansion of its unperturbed energy, E_0, with respect to **F**:

$$E(F) = E_0 - \sum_{i}^{x,y,z} \mu_i F_i - \frac{1}{2!} \sum_{i,j}^{x,y,z} \alpha_{ij} F_i F_j - \frac{1}{3!} \sum_{i,j,k}^{x,y,z} \beta_{ijk} F_i F_j F_k - \frac{1}{4!} \sum_{i,j,k,l}^{x,y,z} \gamma_{ijkl} F_i F_j F_k F_l \quad ... \tag{4.1}$$

The expansion coefficients in Eq. 4.1 are respectively the components of the dipole moment μ_i, polarizability α_{ij}, first hyperpolarizabilty β_{ijk}, and second hyperpolarizability γ_{ijkl} tensors, which can be expressed as consecutive derivatives of the energy with respected to F_i calculated at $F_i = 0$. Considering electric fields applied along the z direction

($F_i = F_z$), the corresponding diagonal components of the tensors are:

$$\mu_z = -\frac{\partial E}{\partial F_z}\Big|_{F_z=0} \tag{4.2}$$

$$\alpha_{zz} = \frac{\partial \mu_z}{\partial F_z}\Big|_{F_z=0} = -\frac{\partial^2 E}{\partial F_z^2}\Big|_{F_z=0} \tag{4.3}$$

$$\beta_{zzz} = \frac{\partial \alpha_{zz}}{\partial F_z}\Big|_{F_z=0} = \frac{\partial^2 \mu_{zz}}{\partial F_z^2}\Big|_{F_z=0} = -\frac{\partial^3 E}{\partial F_z^3}\Big|_{F_z=0} \tag{4.4}$$

$$\gamma_{zzzz} = \frac{\partial \beta_{zzz}}{\partial F_z}\Big|_{F_z=0} = \frac{\partial^2 \alpha_{zz}}{\partial F_z^2}\Big|_{F_z=0} = \frac{\partial^3 \mu_{zz}}{\partial F_z^3}\Big|_{F_z=0} = -\frac{\partial^4 E}{\partial F_z^4}\Big|_{F_z=0} \tag{4.5}$$

These quantities describe the magnitude of the linear and nonlinear responses of the chemical system to an external electric field, hence their accurate computation is crucial for the bottom-up design of optic, electro-optic, and optoelectronic devices. Despite the broad scope of application of density functional approximations (DFAs), these methods often struggle at reproducing linear and nonlinear optical responses of molecular systems. Despite some DFAs (usually implying hybrid exchange-correlation functionals with a large percentage of Hartree-Fock exchange) can reproduce the correct trends in the evolution of properties within series of molecules, they often fail to accurately reproduce the magnitude of the electrical response properties [6–10]. At the heart of this problem is the delocalization error [11], inducing the overdelocalization of electrons, which also leads to the underestimation of reaction barriers and charge-transfer excitation energies and rate-constants [11, 12], the overestimation of the conductance of molecular junctions, the magnetizability of strong antiaromatic molecules [13], electron conjugation [14], and aromaticity [15–20]. A necessary condition to avoid the consequences of the delocalization error on electrical responses is the correct asymptotic decay of the exchange-correlation potentials [21, 22]. The latter is easily imposed using range-separated (RS) DFA. However, even state-of-the-art DFAs using optimally-tuned range-separation parameters sometimes incorrectly reproduce the magnitude of β and γ for relatively simple molecules [7, 23]. Even though the delocalization error is often the main problem in DFAs, electron correlation (beyond the local or semi-local approximations included in most DFAs) is also an essential factor to consider. Indeed, double hybrids often improve the performance of their hybrid or range-separated peers for computing nonlinear optical (NLO) properties [24].

On the other hand, wavefunction methods (WFMs) are exempt from many problems of DFAs, in particular from the delocalization error. The hierarchical structure of WFMs, such as configuration interaction (CI), Møller-Plesset perturbation theory, or coupled-cluster (CC), provides a systematic way toward the exact solution for a given basis set. As a DFT counterpart, Perdew defined the Jacob ladder in the DFT framework, which gives a qualitative indication of the expected accuracy of DFA according to its type; unfortunately, these expectations are not always met for NLO Properties (NLOPs) [6, 25]. High-order WFMs are often considered more accurate than DFAs. In particular, CC including single and double excitations with a perturbative estimation of triples [CCSD(T)][26] is often regarded as the gold standard of WFMs. The computational time of canonical CCSD(T) single-point energy calculation scales as $\mathcal{O}(N^3M^4)$, where N is the number of electrons and M is the number of basis functions of the system. Hence, despite the advantages of WFMs over DFAs, the computational cost of the former usually prevents the calculation of NLOPs beyond cost-effective methods such as the second-order Møller-Plesset perturbation theory (MP2) [24]. Besides, MP2 still presents an unfavorable scaling ($\mathcal{O}(M^5)$) compared to most DFAs and lacks the accuracy to compete with CCSD(T) in a number of situations [27–29].

Many attempts have been made to increase the cost-efficiency of WFMs [30–36]. They can be classified into two groups: methods aiming at increasing the accuracy of the low-cost WFMs (enhanced WFMs) and techniques developed to bring down the computational cost of WFMs (accelerated WFMs). Among the available acceleration techniques, resolution of identity (RI)[37, 38] approximations have become of routine use in many WFMs, the most popular being RI-MP and RI-CC methods [30, 39]. RI techniques have also been introduced for Hartree-Fock (HF) and DFT methods [40–42]. These methods show excellent performance in calculating energies, with considerable time savings [43–46]. Other methods are based on orbital localization, exploiting the local nature of dynamic correlation [47]. They are usually coupled with RI approximations and by localizing natural orbitals they can drastically reduce the computational cost and reach an almost linear scaling with the size of the system, *i.e.*, $\mathcal{O}(M)$ [48–53]. Enhancement techniques exploit some of the systematic deficiencies of WFMs. For instance, MP2 underestimates the opposite-spin (OS) correlation, which is unbalanced with respect to the amount of same-spin correlation (SS) because it is based on Hartree-Fock, which considers the Pauli principle but treats OS pairs as statistically independent pairs. One way to compensate for it is to introduce variable amounts of SS and OS MP2 correlation in what is known as the spin-component scaled MP2 (SCS-MP2) method [54].

Benchmark studies of thermodynamics, kinetics, and molecular properties have been performed on accelerated and enhanced methods [55–58]. However, thus far, a systematic study of the performance of these methods for computing NLO properties is missing in the literature. In this work, we assess the accuracy and computational cost of several enhanced and accelerated techniques applied to CCSD, CCSD(T) and MP2 methods against their canonical counterparts, focusing on the calculation of dipole moments, polarizabilities, and first and second hyperpolarizabilities.

4.2 Methodology

4.2.1 Theoretical Methods

In this section, we briefly review various accelerated and enhanced WFMs. Acceleration techniques aim to reduce the genuine (unaccelerated) method's computational time without sacrificing accuracy. On the other hand, enhancement techniques aim at increasing the accuracy of the canonical method without increasing the computational cost. A summary of the methods considered in this study is provided in Figure 4.1.

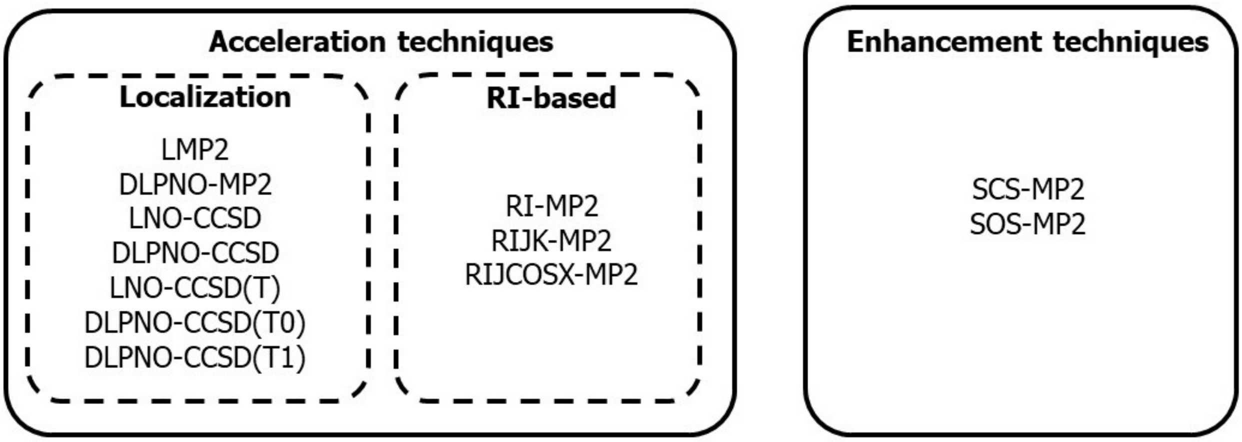


Figure 4.1: Summary of methods studied in this work

4.2.1.1 Accelerated Methods

Among acceleration techniques, one of the most popular methods is the resolution of identity (RI). Within this scheme, the two-electron (four-index) integrals, $\langle ab|cd\rangle$, are approximated as two- or three-index integrals through a density-fitting procedure, thus reducing the scaling with respect to the basis set size. The expression of the fundamental equations regarding RI approximations are reported in Section 2.4.1. The main limitation of the RI method is that the density fitting has to be parameterized for a specific canonical basis set [59, 60], and auxiliary basis sets are not available for all basis sets reported in the literature. Implementations of RI for MP2, CC, or DFT are available. However, the perfomance of most RI-CCSD implementations is not as good as expected. For instance, the implementation of RI-CCSD in ORCA [61] is in practice slower than the canonical CCSD. For this reason, in this chapter, we have limited our assessment to RI-MP2 variants.

The RI approximation can be applied independently to the self-consistent field (SCF) part of the calculation, to the post-HF, or to both. In the present work, we applied the RI either to the MP2 part only [62–64] (RI-MP2, hereafter) or to both the SCF and MP2 parts. Two different versions of RI-SCF calculations have been tested, the RI-JK-SCF method [65], in which both the Coulomb (J) and exchange (K) integrals are treated with the RI method, and the *chain of spheres* method (RI-J-COSX-SCF), in which the Coulomb part is computed with the RI approximation and the exchange part is computed by numerical integration over a predefined grid [66]. Both methods are implemented in ORCA [61] and we refer to them throughout as RIJK-MP2 and RIJCOSX-MP2, respectively. Additionally, we have tested a tighter COSX grid referred as COSX2 in this chapter (GridX=4 in ORCA input).

The second family of acceleration techniques tested in this work is designed to take advantage from the eminently local character of electron correlation [47]. In this regard, the transformation from canonical to localized orbitals can be achieved by a unitary transformation of the wavefunction. Depending on the constraints added to the unitary transformation, several localization schemes may arise. Recently, methods based on the localized pair natural orbitals (LPNOs) are becoming very popular as they introduce a drastic reduction of the computational cost, resulting in an almost-linear scaling with the molecular size. These methods employ the pair natural orbitals (PNOs) formulation, reducing the virtual space of the calculation [67] by localizing its orbitals through the Foster-Boys algorithm [68, 69], which consists in minimizing $\langle \widehat{L} \rangle$ (with $\widehat{L} = |\vec{r}_1 - \vec{r}_2|^2$). Other important localization schemes used in alternative contexts are the Edmiston-Rudenberg [70] or the Pipek-Mezey [71] ones, which impose the minimization of the orbital self-repulsion and of the atomic Mulliken charges, respectively. It is important to distinguish methods that localize the orbitals after the SCF calculations from those like the Extremely Localized Molecular Orbitals (ELMO)[72] scheme, which applies directly the variational principle on the constrained many-body Slater determinant. Among the localization methods available in the literature, we decided to test two different schemes for the calculation of the nonlinear optical properties. First, the domain localized pair natural orbital (DLPNO) method [49, 50, 53] implemented in ORCA [61]. Second, the localized natural orbital method (LNO) developed by Kállay and coworkers [73] implemented in the MRCC package [74]. These two methodologies use different strategies to construct the virtual domain. The machinery behind DLPNO is rather convoluted and can be summarized as follows. A set of pair natural orbitals (PNOs) providing the most compact description of the virtual space is constructed. The latter PNOs are obtained through the diagonalization of the pair density matrix for every pair of localized occupied orbitals. Finally, the DLPNO method expands the PNOs in terms of certain basis functions, more specifically, into the set of Pulay's projected atomic orbitals (PAOs) [47], belonging to a specific electron-pair domain [32]. Alternatively, the LNO method first localizes the MOs using a distance criterion. Subsequently, each localized MO is assigned to a local subspace of occupied and virtual orbitals, which is constructed from approximate Møller-Plesset frozen natural orbitals. Finally, the CC equations are solved for each LNO subspace and the total correlation energy is obtained from the summation over all the sub-spaces. The main difference between DLPNO and LNO schemes is that the former defines the interacting subspaces from electron pairs, while the latter uses individual electrons [51, 74–76].

Such definitions for the localized orbitals (either LNO or DLPNO) can be used to efficiently compute post-HF energies and wavefunctions. More informataions over DLPNO and LNO methods can be found in Section 2.4.2. In this chapter, we have assessed LMP2 [35, 52], which uses the LNO localization method, LNO-CCSD [52, 74], DLPNO-MP2 [49], and DLPNO-CCSD energy derivatives [32]. We also computed analytic DLPNO-MP2 polarizabilities using ORCA 5.0 [57, 77]. We refer to this method as DLPNO-MP2-α. We compared DLPNO-MP2-α analytic polarizabilities with numerical (from energy) DLPNO-MP2 polarizabilities to evaluate the magnitude of the numerical errors in the calculation of polarizabilities using the Rutishauser–Romberg technique (see below). We also considered triple perturbative corrections to DLPNO-CCSD using the two alternative approximations available: DLPNO-CCSD(T0) [48], in which the triples corrections are calculated following a purely perturbational approach, and DLPNO-CCSD(T1) [53], which is more expensive (and considered more accurate) because it is partially self-consistent. The triple perturbative corrections have been also tested for the LNO scheme, referred as LNO-CCSD(T) [73, 78].

4.2.1.2 Enhancement Methods

Regarding the enhancement methodology, we assessed the spin-scaled component MP2 method. As it has been already anticipated in Section 2.3.5, SCS-MP2 does not reduce the computational time explicitly, but it effectively improves the quality of the results of a canonical MP2 calculation by increasing the amount of opposite-spin (OS) correlation and scaling down same-spin (SS) correlation,

$$E^{SCS-MP2} = E^{HF} + c_{SS}E^{MP2}_{c,SS} + c_{OS}E^{MP2}_{c,OS}, \tag{4.6}$$

where $c_{OS} = 6/5$ and $c_{SS} = 1/3$ for Grimme's SCS-MP2 (as opposed to canonical MP2, where $c_{OS} = 1$ and $c_{SS} = 1$) [79]. Head-Gordon and co-workers suggested a scaled opposite-spin MP2 (SOS-MP2), which takes values $c_{OS} = 1.3$ and $c_{SS} = 0$. By excluding the same-spin correlation the computational complexity might be reduced from fifth to fourth order [80].

4.2.2 Computational details

Single-point calculations have been performed using an energy threshold of 10^{-9} a.u. for convergence of both the SCF and CC calculations. A tighter convergence criterion (10^{-14} a.u.) was also tested, however, the results for the (hyper)polarizabilities showed no significant improvement, whereas medium/large systems showed hampered convergence. All calculations have been performed with the aug-cc-pVDZ in conjunction with the corresponding auxiliary basis set when needed. Test on the time performances of the different methods employed in this chapter with our computational facilities have been collected in Section A.1 of the Annex.

Static linear polarizabilities and first and second hyperpolarizabilities have been evaluated numerically through finite-field central derivatives of the total energy. These calculations have been performed on the range of external electric fields $\pm 2^j \cdot 10^{-4}$ a.u. with $j = [0, 9]$ (1 a.u.= 51.422 V·Å^{-1}). By construction, the finite-field central derivatives remove the truncation error due to the higher-order terms of the Taylor expansion of the field-dependent energy with different parity of the derivative evaluated. In order to reduce the truncation error coming from neglecting the latter higher-order terms with the same parity than the derivative evaluated, the Rutishauser–Romberg (RR) formula has been employed [81, 82], which we recall from Section 3.2.1:

$$P^{i,j} = \frac{4^i \cdot P^{(i-1),j} - P^{(i-1),(j+1)}}{4^i - 1}, \tag{4.7}$$

where P is the calculated property, i is the RR iteration number, and j is the exponent entering the expression of the electric field amplitude ($\pm 2^j \cdot 10^{-4}$ a.u.). In order to choose the i and j values minimizing the truncation error, the minimum of the difference between the j^{th} and $(j+1)^{\text{th}}$ rows for the same i column of the matrix $P^{i,j}$ is evaluated and defined as the Romberg Error (RE), namely:

$$RE = \min_{i,j} |P^{i,j} - P^{i,(j+1)}| \tag{4.8}$$

The generalized RR formula was employed for systems presenting convergence problems on the RR iterations (*i.e.*, presenting a relatively large rounding error) [81]. However, for the molecules studied in this chapter, the rounding error coming from the generalized RR expression did not decrease significantly compared to the standard RR. In order to verify which methods are numerically stable, we reported the Mean Absolute Romberg Error (MARoE) of each property and the relative MARoE (%MARoE) calculated by dividing MARoE by the average value.

The localization schemes depend on a wide range of cutoffs, thresholds, and parameters that control the accuracy of the energy calculations and, subsequently, their derivatives [56]. The developers of the DLPNO method identified three different sets of thresholds for the localization schemes, associated with a particular computational cost and accuracy. They employ the keywords LoosePNO, NormalPNO, and TightPNO to refer to these approximations [83]. At the same time, LNO developers identified three sets of parameters controlled by the variable *lcorthr* in the MRCC input [84]. After a few tests, it became obvious that to reduce the numerical error associated to each single-point calculation, TightPNO should be used for DLPNO calculations and *lcorthr=VeryTight* for LNO. Calculations with looser cuttoffs are included in the Annex for comparison. As an illustrative example, the relative errors committed by the numerical differentiation (measured by %MARoE) of α, β and γ applying several thresholds at the DLPNO-CCSD(T) level of theory for the γ-NLO set (see below) are summarized in Table 4.1.

Table 4.1: Relative mean absolute Romberg error (%MARoE) on α, β and γ values obtained from the numerical derivatives of the energy, calculated for molecules of the γ-NLO set (see next section) by using various DLPNO methods employing different thresholds for the SCF/CC equations and orbital localization scheme.

Method	**SCF/CC**	**Localization**	**%MARoE**		
			α	β	γ
DLPNO-CCSD(T0)	VeryTight	NormalPNO	1	50	70
DLPNO-CCSD(T0)	VeryTight	TightPNO	0	40	64
DLPNO-CCSD(T0)	ExtremeSCF	TightPNO	0	35	64
DLPNO-CCSD(T1)	VeryTight	NormalPNO	0	53	98
DLPNO-CCSD(T1)	VeryTight	TightPNO	0	29	51

Improving the SCF/CC convergence from VeryTight to ExtremeSCF criterion provides comparable %MARoE for α and β, and γ values. Therefore, considering the computational cost of using ExtremeSCF, we only employed the former criterion. The numerical derivative errors also show a large dependency on the localization cutoffs. In order to minimize these errors, we employed the highest TightPNO criterion. We also tested user-tailored combinations of the set of parameters, seeking an increase of accuracy for NLOP calculations but we did not find a situation where one particular parameter was singled out as the most relevant or dominant to improve the quality of the NLOs. In practice, the errors increase significantly with the order of the energy derivatives, hence, β and γ require the tightest criteria. Notice that in some cases, the numerical instability of the energies are so large that the error committed can exceed %MARoE=95%. These conclusions about the need of high localization cutoffs are in line with the findings of Alonso, Martin, and co-workers [56], who identified that very tight cuttoffs for DLPNO are needed to reproduce the relative energy of extended porphyrins.

The performance of enhanced and accelerated CCSD(T), CCSD, and MP2 WFMs has been assessed by comparison to reference values obtained using the corresponding canonical methods, by considering four statistical measures: the Mean Absolute Error (MAE), the Root Mean-Square Error (RMSE), the Maximum Error (MAX), and the percentage MAE (%MAE)

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - t_i| \tag{4.9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - t_i)^2} \tag{4.10}$$

$$\mathrm{MAX} = \max_i |x_i - t_i| \tag{4.11}$$

$$\%\mathrm{MAE} = \frac{\mathrm{MAE}}{\frac{1}{n}\sum_{i=1}^{n} t_i} \tag{4.12}$$

where t_i and x_i are the reference and the predicted values for system i, respectively.

4.2.3 Benchmark sets

Two benchmark sets are used in this chapter: the γ-NLO set[7] and the β-NLO set (see Figure 4.2). The γ-NLO set contains 60 molecules, formed by 2 to 36 atoms of the second period and/or hydrogen. The latter set can be split into two subsets: The first one (γ-NLO-A, 37 molecules) contains molecules that, for the adopted orientation, are symmetric along the z-axis; while the second set (γ-NLO-B, 23 molecules) includes polar molecules oriented aligning their inertia axis to z axis, and thus they are not symmetric along z-axis. The γ-NLO-A set includes the first oligomers of two series of well-known NLO compounds —the all-trans polyacetylene (PA), and the polydiacetilene (PDA)— as well as some small organic and inorganic molecules, and weakly-interacting H_2 chains, which are particularly challenging systems for the computation of second hyperpolarizabilities [85]. This set has been only employed for the evaluation of α_{zz} and γ_{zzzz} (even derivatives with respect to the electric field) because these molecules present, by construction, null μ_z and β_{zzz} (odd derivatives). Conversely, the γ-NLO-B has been employed to evaluate μ_z, α_{zz}, β_{zzz}, and γ_{zzzz}. This set includes the first six oligomers of all-trans polymethineimine (PMI), the NLOP calculation of which proves difficult for electronic structure methods.

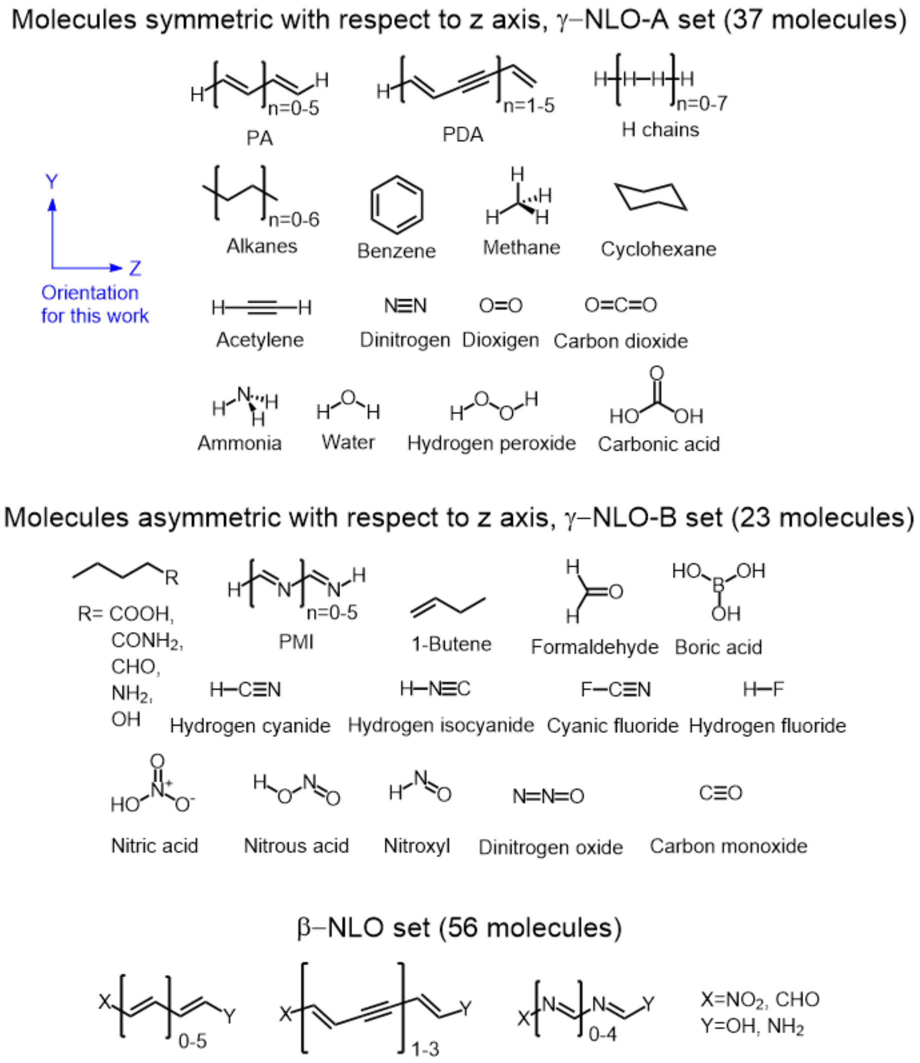


Figure 4.2: Benchmark γ- and β-NLO sets studied in this chapter. The subset γ-NLO-A contains molecules that, on the adopted orientation, are symmetric along the z-axis; while the subset γ-NLO-B contains molecules that are not symmetric along the z-axis.

The β-NLO set contains molecules with expected large β and γ. In particular, it consists of 56 π-conjugated push-pull systems that result from the functionalization of the terminal positions of PA_{1-6}, PDA_{1-3}, and PMI_{1-5} oligomers with two possible electron-withdrawing (-NO_2, -CHO) and two possible electron-donor (-NH_2, -OH) substituents.

Molecular geometries and reference CCSD(T) energy derivatives for γ-NLO set are available and have been published elsewhere [7], whereas the data corresponding to the β-NLO is included in the Annex of this chapter. All the molecules have a singlet ground state, with the exception of O_2, which is a triplet molecule. Unrestricted calculations with localized methods are still not implemented in MRCC, therefore, O_2 was excluded from this study.

The single-reference character of the molecules was assessed through a series of multireferences diagnostic criteria. On one side, we computed D1[86], D2 [87], and T1 [88–90] over the CCSD wavefunctions. We employed various natural-occupancy-based diagnostics for MP2 calculations, namely: NON, V, MRI [91], and I_{ND} [92–95]. According to the latter diagnostics, none of the molecules presents a high multiconfigurational character (see A.7 and A.8) and, therefore, coupled-cluster and MP2 wavefunction are appropriate methodologies to assess the electronic structure of such systems. Only the longitudinal components of the dipole moment vector (μ_z), and of the polarizability (α_{zz}) and hyperpolarizability tensors (β_{zzz} and γ_{zzzz}) were computed for all the molecules (see Figure 4.2). For simplicity, hereafter the indices will be dropped and the diagonal tensor components will be noted μ, α, β and γ. All properties have been reported without sign in order to allow direct comparison between the different sign conventions adopted by the packages utilized, and facilitate the analysis.

4.3 Results

In this section, we will only show statistical errors with respect to some reference values. Absolute magnitudes of the linear and nonlinear optical properties are available in the Annex. The results are organized as follows: the performance of accelerated methods are first checked against their canonical counterparts. Then, we consider their accuracy by comparing the computed optical responses with CCSD(T)/aug-cc-pVDZ reference values. Finally, we assess enhanced wavefunction methods.

4.3.1 Relative performance of accelerated methods

In this section, we consider the performance of accelerated MP2 and CCSD calculations, whereas accelerated CCSD(T) methods will be assessed in Section 4.3.2. As detailed above, different statistical measures were collected to quantify the errors. However, all statistical parameters generally provide a similar assessment of the methods. Hence, unless otherwise indicated, we mainly use the relative mean average error (%MAE) to analyze the data.

4.3.1.1 Accelerated MP2 calculations

Table 4.2 reports the statistical measures assessing the performance of six accelerated MP2 methods with respect to canonical MP2. The errors committed by MP2 accelerated methods for the lowest-order properties (dipole moment and polarizability) are minimal (ca. %1). Therefore, employing any of these methods is advisable to reduce the computational cost of these properties. However, the highest accuracy is achieved by RIJK-MP2, closely followed by RI-MP2 and the analytical calculation of the polarizability at the DLPNO-MP2 level (DLPNO-MP2-α). The remarkable accuracy of the analytical field-free and field-dependent linear polarizabilities obtained from DLPNO-MP2 is reflected by the accuracy of DLPNO-MP2-α first and second hyperpolarizabilities (calculated from

Table 4.2: Performance of accelerated MP2 methods with respect to canonical MP2 for the γ-NLO set. Units are a.u.

	Method	**RI-MP2**	**RIJK-MP2**	**RIJCOSX2-MP2**	**LMP2**	**DLPNO-MP2-α**	**DLPNO-MP2**
μ	MAE	1.9E-04	1.9E-04	2.2E-04	4.3E-04	2.2E-04	5.8E-04
	RMSE	3.3E-04	3.8E-04	1.7E-03	8.3E-04	3.9E-04	1.2E-03
	MAX	1.2E-03	1.4E-03	2.6E-03	3.1E-03	1.5E-03	4.7E-03
	%MAE	0	0	0	0	0	0
α	MAE	4.6E-02	3.9E-02	1.2E+00	5.1E-01	7.6E-02	9.5E-01
	RMSE	1.6E-01	1.2E-01	7.6E+00	1.8E+00	2.8E-01	3.0E+00
	MAX	1.1E+00	8.4E-01	5.8E+01	1.2E+01	1.8E+00	2.0E+01
	%MAE	0	0	1	0	0	1
β	MAE	5.7E-01	3.7E+00	1.2E+01	7.1E+01	2.9E+00	1.2E+02
	RMSE	1.7E+00	1.5E+01	3.8E+01	2.3E+02	8.9E+00	2.5E+02
	MAX	6.6E+00	7.0E+01	1.7E+02	9.9E+02	3.9E+01	8.8E+02
	%MAE	0	2	5	29	1	49
γ	MAE	2.7E+04	5.5E+04	5.4E+04	2.6E+05	3.3E+04	1.9E+05
	RMSE	1.3E+05	2.6E+05	1.4E+05	1.3E+06	1.5E+05	4.7E+05
	MAX	8.7E+05	1.7E+06	6.0E+05	9.5E+06	1.1E+06	1.8E+06
	%MAE	5	11	10	51	6	36

numerical derivatives of the former). Together with RI-MP2, DLPNO-MP2-α provides the most accurate values for the whole range of optical properties. Since RIJK-MP2 and RIJCOSX2-MP2 only display a slight increase of the relative MAE on β and γ while they reduce the computational cost of RI-MP2 by applying the resolution of identity also at the SCF level, one should likewise consider these methods to compute hyperpolarizabilities to obtain values close to the MP2 accuracy.

4.3.1.2 Accelerated CCSD methods

Table 4.3 collects data to assess the performance of DLPNO-CCSD and LNO-CCSD against canonical CCSD. The errors committed by CCSD accelerated methods for the dipole moment and the polarizability are slightly larger than their counterparts at the MP2 level, presenting also very small (%MAE $\leq$ 2%). However, none of these methods provides reasonable values for the first and the second polarizabilities. The excellent results obtained from analytical DLPNO-MP2-α suggest that if analytical values of the polarizability at the DLPNO-CCSD were available, we would also obtain accurate hyperpolarizabilities at this level of theory.

The poor results obtained for the high-order optical properties from DLPNO-CCSD and LNO-CCSD are due to numerical errors (see Tables A.9 and A.10 for the Romberg errors) that cannot be avoided using other numerical differentiation techniques. The poor quality of the single-point DLPNO-CCSD and LNO-CCSD field-dependent energies is responsible for it, and it cannot be solved by using tighter convergence criteria for the localization schemes. Table A.11 shows that these results are even worse if we employ looser criteria.

Table 4.3: Performance of accelerated CCSD methods with respect to canonical CCSD for the γ-NLO set. Units are a.u.

	Method	**DLPNO-CCSD**	**LNO-CCSD**
μ	MAE	2.4E-03	1.6E-03
	RMSE	4.7E-03	3.5E-03
	MAX	1.8E-02	1.3E-02
	%MAE	0	0
α	MAE	1.5E+00	1.9E+00
	RMSE	8.5E+00	1.2E+01
	MAX	3.8E+01	4.4E+01
	%MAE	1	2
β	MAE	2.1E+02	4.7E+01
	RMSE	3.0E+02	1.2E+02
	MAX	6.5E+02	4.6E+02
	%MAE	114	25
γ	MAE	4.7E+05	2.1E+05
	RMSD	1.3E+06	1.2E+06
	MAX	5.9E+06	4.1E+06
	%MAE	122	54

Table 4.4: Performance of acceleration methods with respect to CCSD(T) references for the evaluation of μ and α for the γ-NLO set. Units are a.u.

	Method	**DLPNO-CCSD(T0)**	**DLPNO-CCSD(T1)**	**LNO-CCSD(T)**	**MP2**	**CCSD**
μ	MAE	3.4E-03	3.2E-03	2.3E-03	2.9E-02	2.2E-02
	RMSE	6.6E-03	7.5E-03	6.3E-03	4.3E-02	3.4E-02
	MAX	2.6E-02	2.6E-02	2.6E-02	1.3E-01	9.5E-02
	%MAE	1	1	0	4	3
α	MAE	1.5E+00	8.2E-01	1.2E+00	3.7E+00	3.0E+00
	RMSE	4.2E+00	2.1E+00	3.9E+00	7.2E+00	3.0E+00
	MAX	2.0E+01	1.0E+01	2.1E+01	3.1E+01	6.5E+01
	%MAE	1	1	1	3	3

4.3.2 Absolute performance of accelerated methods

Thus far, we have evaluated the efficiency of accelerated MP2 and CCSD to reproduce their canonical counterparts. In this section, we benchmark accelerated methods against the reference CCSD(T) calculations for the calculation of NLOPs, including accelerated CCSD(T) variants, which are assessed for the first time.

Table 4.4 collects the statistical data for the dipole moment and the polarizability. Accelerated MP2 and CCSD methods are omitted from these analyses, since we demonstrated above that they have the same accuracy as their canonical counterparts. DLPNO-CCSD(T) variants show excellent performance with errors below or equal to 1% with respect to the canonical CCSD(T). MP2 and CCSD methods also provide very good approximations of the two properties.

The data in Tables 4.5 and 4.6 illustrate the performance of various methods to compute the first and second hyperpolarizability, respectively. We have omitted the data for the RI approximations and DLPNO-MP2-α, since we can expect that the accuracy of

Table 4.5: Performance of acceleration methods with respect to CCSD(T) references for the evaluation of β for the γ-NLO set.

		DLPNO			**LNO**			**Canonical**	
	Method	**CCSD**	**CCSD(T0)**	**CCSD(T1)**	**CCSD**	**CCSD(T)**	**LMP2**	**MP2**	**CCSD**
β	MAE	2.3E+02	1.1E+02	2.7E+01	9.2E+01	6.4E+01	5.7E+01	1.1E+02	5.1E+01
	RMSE	3.6E+02	3.1E+02	4.8E+01	2.5E+02	2.0E+02	1.2E+02	2.6E+02	1.3E+02
	MAX	1.1E+03	1.3E+03	1.9E+02	9.2E+02	9.1E+02	3.6E+02	8.3E+02	4.6E+02
	%MAE	165	78	19	67	46	41	78	37

Table 4.6: Performances of acceleration methods with respect to CCSD(T) references for the evaluation of γ for the γ-NLO set. Units are a.u.

		DLPNO			**LNO**			**Canonical**	
	Method	**CCSD**	**CCSD(T0)**	**CCSD(T1)**	**CCSD**	**CCSD(T)**	**LMP2**	**MP2**	**CCSD**
γ	MAE	5.3E+05	7.0E+05	2.3E+05	2.2E+05	4.6E+05	2.6E+05	7.7E+04	8.4E+04
	RMSE	1.9E+06	3.8E+06	1.1E+06	1.5E+06	2.2E+06	1.2E+06	2.4E+05	6.0E+05
	MAX	8.6E+06	2.8E+07	8.3E+06	6.8E+06	1.5E+07	8.9E+06	1.2E+06	2.7E+06
	%MAE	117	155	51	49	103	58	17	19

these methods is similar to that of MP2 (see Table 4.2). None of the methods tested give a relative MAE below 15% for the first hyperpolarizability, the best methods being DLPNO-CCSD(T1) followed by canonical CCSD. The latter results indicate that the accurate evaluation of triples is crucial in reproducing CCSD(T) values. All DLPNO-based methods show significant numerical derivative errors (see MARoE values in Tables A.9 and A.10), which can be partially (but not sufficiently) reduced by employing tighter cutoffs (see Table A.12). Interestingly, the accuracy of DLPNO methods in reproducing triple excitations goes hand in hand with the numerical stability of the energies. Indeed, for the same cutoffs, we find more stable energies (lower MARoE values) for DLPNO-CCSD(T1) than for DLPNO-CCSD(T0) —see Table A.12. The LNO-CCSD (LNO-CCSD(T)) method gives somewhat more accurate first hyperpolarizabilities than DLPNO-CCSD (DLPNO-CCSD(T0)); however, both LNO methods exhibit relative MAE errors above 45%. Interestingly, LMP2 outperforms MP2, mainly because of a better reproduction of the first hyperpolarizabilities of the PMI oligomers by LMP2 —which is probably due to a fortuitous cancellation of errors.

In Table 4.6, we collect the statistics for the second hyperpolarizabilities. In this case, none of the accelerated methods gives a relative MAE below 50%. All the methods also show substantial numerical derivative errors, evidencing that numerical instabilities in the energy values hinder an accurate calculation of their fourth-order derivatives. Interestingly, MP2 performs better than CCSD. In this sense, RI-based accelerated MP2 methods (see Table 4.2) are an economical alternative for computing second hyperpolarizabilities.

4.3.3 Performance of enhanced MP2 methods: SCS-MP2

In this section, we assess the accuracy of several spin-scaled methods designed to improve the performance of MP2 energies by adjusting the amount of same-spin (c_{SS}) and opposite-

Table 4.7: Performance of different spin-component scaled MP2 methods with respect to CCSD(T) for the evaluation of the linear and nonlinear optical properties for the γ-NLO set. Units are a.u.

	Method	**MP2**	**SCS-MP2**	**SOS-MP2**
	c_{SS}	1	0.3	0
	c_{OS}	1	1.2	1.3
μ	MAE	2.9E-02	2.0E-02	1.5E-02
	RMSE	4.3E-02	2.7E-02	1.9E-02
	MAX	1.3E-01	6.6E-02	4.3E-02
	%MAE	4	3	2
α	MAE	3.7E+00	2.9E+00	4.8E+00
	RMSE	7.2E+00	8.4E+00	1.4E+01
	MAX	3.1E+01	5.0E+01	8.3E+01
	%MAE	3	3	4
β	MAE	1.1E+02	1.7E+02	2.0E+02
	RMSE	2.6E+02	4.4E+02	5.4E+02
	MAX	8.3E+02	1.6E+03	2.0E+03
	%MAE	78	125	148
γ	MAE	7.7E+04	6.4E+04	5.6E+04
	RMSE	2.4E+05	1.9E+05	1.7E+05
	MAX	1.2E+06	9.3E+05	8.5E+05
	%MAE	17	14	12

spin (c_{OS}) correlation. As described in above, we tested two different popular schemes: the Grimme's SCS-MP2[79] in which $c_{SS} = 1/3$ and $c_{OS} = 6/5$, and SOS-MP2[96] that takes $c_{SS} = 0$ and $c_{OS} = 1.3$. The computational cost of SCS-MP2 is insignificantly larger than canonical MP2, whereas SOS-MP2 can be formally implemented to reduce the scaling by one order in the number of basis functions. We collected the statistics of their performance in Table 4.7.

The results for the dipole moment and the polarizability show only a marginal improvement over MP2, which already exhibits pretty accurate results. According to all statistical measures, the first hyperpolarizability is better estimated by canonical MP2 than by spin-scaled methods. Reversely, both SCS-MP2 and SOS-MP2 marginally better reproduce the second hyperpolarizability. Although there is not much improvement, it might be worth exploring the possibility of an accelerated SOS-MP2 method as an economical way to compute γ.

4.3.4 Performances of accelerated methods on push-pull systems

From the results of the previous sections, we have identified β and γ as the most challenging properties for accelerated wavefunction methods. We have also pinpointed the best-performing methods for β is DLPNO-CCSD(T1), and for gamma are CCSD, RI-MP2, and MP2. In this section, we put these methods to the test by analyzing further

Table 4.8: Performance of DLPNO-CCSD(T1), MP2, RI-MP2, and CCSD methods with respect to the CCSD(T) references for the evaluation of β and γ for the β-set. Units are a.u.

	Method	**DLPNO-CCSD(T1)**	**RI-MP2**	**MP2**	**CCSD**
β	MAE	1.7E+03	2.8E+03	2.7E+03	7.7E+02
	RMSE	2.6E+03	4.5E+03	4.5E+03	1.2E+03
	MAX	6.8E+03	1.3E+04	1.3E+04	4.5E+03
	%MAE	17	28	27	8
γ	MAE	-	5.6E+05	5.2E+05	3.9E+05
	RMSE	-	9.8E+05	8.7E+05	7.0E+05
	MAX	-	3.1E+06	2.9E+06	3.2E+06
	%MAE	-	21	19	14

their relative accuracy for computing the NLO properties of the push-pull molecules contained in the β-NLO set. These molecules are expected to exhibit a significant β response that is susceptible to be impacted by larger errors. In the case of γ, we have excluded DLPNO-CCSD(T1) because it presented too large energy derivative errors. The statistical results are collected in Table 4.8.

Although the molecules in the β-NLO set present larger absolute errors than those in the γ-NLO set, the relative errors are smaller because of the larger average β and γ values. The performance of DLPNO-CCSD(T1) compared to CCSD is not as good as in γ-NLO set. DLPNO-CCSD(T1) hyperpolarizabilities have average errors about twice as large as those of CCSD, although it outperforms MP2. The results of MP2 and RI-MP2 are comparable for the first and second polarizability, showing that the resolution of identity methods can be safely employed as substitutes for MP2 also for π-conjugated push-pull derivatives. The comparison between MP2 and CCSD shows that for molecules with larger responses, MP2 exhibit larger deviations than CCSD for the first hyperpolarizability, while both methods lead to similar %MAEs for the second hyperpolarizability.

4.4 Conclusions

In this chapter, we have benchmarked various alternatives to wavefunction methods that either reduce the computational cost or improve the performance of the canonical methods. In particular, we have tested RI-MP2, RIJK-MP2, RIJCOSX2-MP2, LMP2, SCS-MP2, SOS-MP2, DLPNO-MP2, LNO-CCSD, LNO-CCSD(T), DLPNO-CCSD, DLPNO-CCSD(T0), and DLPNO-CCSD(T1). Our results indicate that all these methods produce numerically stable energies to compute first and second derivatives of the energy with respect to an external electric field. Since, in general, these derivatives are not highly affected by correlation energy, we can safely employ any of the latter methods to calculate the dipole moment and the polarizability with average relative errors below 5%.

On the other hand, the calculation of higher-order derivatives represents a challenge for accelerated and enhanced wavefunction methods. In particular, the third and fourth derivatives of the energy, required to compute the first and second polarizability, critically depend on the numerical stability of the single-point field-dependent energy calculations.

Our results show that RI-based methods produce reliable energies from which to compute up to fourth-order derivatives of the energy with respect to an external field. Hence, methods like RI-MP2, RIJK-MP2, or RIJCOSX2-MP2 are a cost-effective way to obtain first and second hyperpolarizabilities with a marginal average error with respect to canonical MP2 (up to 5% for β and up to 11% for γ). Conversely, methods based on orbital localizations (LNO and DLPNO techniques) applied to MP2 suffer from large numerical instabilities that result in large errors for β (29-49%) and γ (36-51%). The same techniques applied to CCSD and CCSD(T) result in even larger errors, which exceed 100% in the worse cases. The only exception is DLPNO-CCSD(T1), which produces an acceptable relative error of 19% for the calculation of β.

The precision of single-point energy calculations with LNO and DLPNO critically depends on the cutoffs for the SCF/CC equations and the orbital localization scheme. In addition, the accuracy of DLPNO methods in reproducing triple excitations goes hand in hand with the numerical stability of the energies. Hence, for the same cutoffs, we have more stable energies for DLPNO-CCSD(T1) than for DLPNO-CCSD(T0).

Analytical field-dependent polarizabilities are available at the DLPNO-MP2 level of theory, from which we have numerically computed first and second hyperpolarizabilities that are in excellent agreement with their canonical MP2 counterparts. We can thus anticipate that if analytical DLPNO-CCSD polarizabilities were available, we would have a cost-effective method to compute accurate hyperpolarizabilities.

Finally, we assessed spin-component scaled methods as techniques that improve the performance of MP2 at the same cost. Unfortunately, these techniques produce only produce only a marginal improvement in the case of the second polarizability.

To summarize, we recommend RIJK-MP2, RIJCOSX2-MP2, and DLPNO-MP2 (using analytical polarizabilities to compute γ) to compute the dipole moment, the polarizability, and the second hyperpolarizability, whereas only DLPNO-CCSD(T1) using tight cutoffs can be employed to obtain reasonably accurate first hyperpolarizabilities. We hope that these results will prompt the implementation of analytical low-order properties for accelerated wavefunction methods and/or more precise single-point energies that can be employed to compute numerical derivatives.

Bibliography

[1] Tadepalli, S.; Slocik, J. M.; Gupta, M. K.; Naik, R. R.; Singamaneni, S. Bio-optics and bioinspired optical materials. *Chem. Rev.* **2017**, *117*, 12705–12763.

[2] Dini, D.; Calvete, M. J. F.; Hanack, M. Nonlinear optical materials for the smart filtering of optical radiation. *Chem. Rev.* **2016**, *116*, 13043–13233.

[3] Liu, W.; Liu, M.; Liu, X.; Wang, X.; Deng, H.-X.; Lei, M.; Wei, Z.; Wei, Z. Recent advances of 2D materials in nonlinear photonics and fiber lasers. *Adv. Opt. Mater.* **2020**, *8*, 1901631.

[4] Drobizhev, M.; Makarov, N. S.; Tillo, S. E.; Hughes, T. E.; Rebane, A. Two-photon absorption properties of fluorescent proteins. *Nat. Methods* **2011**, *8*, 393–399.

[5] Denk, W.; Strickler, J.; Webb, W. Two-photon laser scanning fluorescence microscopy. *Science* **1990**, *248*, 73–76.

[6] Zaleśny, R.; Medved', M.; Sitkiewicz, S. P.; Matito, E.; Luis, J. M. Can Density Functional Theory Be Trusted for High-Order Electric Properties? The Case of Hydrogen-Bonded Complexes. *J. Chem. Theory Comput.* **2019**, *15*, 3570–3579.

[7] Besalú-Sala, P.; Sitkiewicz, S. P.; Salvador, P.; Matito, E.; Luis, J. M. A new tuned range-separated density functional for the accurate calculation of second hyperpolarizabilities. *Phys. Chem. Chem. Phys.* **2020**, *22*, 11871–11880.

[8] Naim, C.; Castet, F.; Matito, E. Impact of van der Waals interactions on the structural and nonlinear optical properties of azobenzene switches. *Phys. Chem. Chem. Phys.* **2021**, *23*, 21227–21239.

[9] Johnson, L. E.; Dalton, L. R.; Robinson, B. H. Optimizing calculations of electronic excitations and relative hyperpolarizabilities of electrooptic chromophores. *Acta Chem. Scand.* **2014**, *47*, 3258–3265.

[10] Champagne, B.; Perpete, E. A.; Jacquemin, D.; van Gisbergen, S. J.; Baerends, E.-J.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. Assessment of conventional density

functional schemes for computing the dipole moment and (hyper) polarizabilities of push- pull π-conjugated systems. *J. Phys. Chem. A* **2000**, *104*, 4755–4763.

[11] Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Localization and delocalization errors in density functional theory and implications for band-gap prediction. *Phys. Rev. Lett.* **2008**, *100*, 146401.

[12] Besalú-Sala, P.; Voityuk, A. A.; Luis, J. M.; Solà, M. Evaluation of charge-transfer rates in fullerene-based donor–acceptor dyads with different density functional approximations. *Phys. Chem. Chem. Phys.* **2021**, *23*, 5376–5384.

[13] Valiev, R. R.; Benkyi, I.; Konyshev, Y. V.; Fliegl, H.; Sundholm, D. Computational studies of aromatic and photophysical properties of expanded porphyrins. *J. Phys. Chem. A* **2018**, *122*, 4756–4767.

[14] Sancho-García, J.; Pérez-Jiménez, A. Improved accuracy with medium cost computational methods for the evaluation of bond length alternation of increasingly long oligoacetylenes. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5874–5879.

[15] Torrent-Sucarrat, M.; Navarro, S.; Cossío, F. P.; Anglada, J. M.; Luis, J. M. Relevance of the DFT method to study expanded porphyrins with different topologies. *J. Comput. Chem.* **2017**, *38*, 2819–2828.

[16] Casademont-Reig, I.; Woller, T.; Contreras-García, J.; Alonso, M.; Torrent-Sucarrat, M.; Matito, E. New electron delocalization tools to describe the aromaticity in porphyrinoids. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2787–2796.

[17] Casademont-Reig, I.; Ramos-Cordoba, E.; Torrent-Sucarrat, M.; Matito, E. How do the Hückel and Baird Rules Fade away in Annulenes? *Molecules* **2020**, *25*, 711.

[18] Torrent-Sucarrat, M.; Navarro, S.; Marcos, E.; Anglada, J. M.; Luis, J. M. Design of Hückel–Möbius topological switches with high nonlinear optical properties. *J. Phys. Chem. C* **2017**, *121*, 19348–19357.

[19] Casademont-Reig, I.; Guerrero-Avilés, R.; Ramos-Cordoba, E.; Torrent-Sucarrat, M.; Matito, E. How aromatic ara molecular nanorings? The case of a six-porphyrin nanoring. *Angew. Chem. Int. Ed.* **2021**, *60*, 24080–24088.

[20] Casademont-Reig, I.; Soriano-Agueda, L.; Ramos-Cordoba, E.; Torrent-Sucarrat, M.; Matito, E. Reply to Correspondence on How aromatic ara molecular nanorings? The case of a six-porphyrin nanoring. *Angew. Chem. Int. Ed.* **2021**, *61*, e202206836.

[21] Autschbach, J.; Srebro, M. Delocalization error and “functional tuning” in Kohn–Sham calculations of molecular properties. *Acc. Chem. Res.* **2014**, *47*, 2592–2602.

[22] Gritsenko, O.; van Leeuwen, R.; van Lenthe, E.; Baerends, E. J. Self-consistent approximation to the Kohn-Sham exchange potential. *Phys. Rev. A* **1995**, *51*, 1944–1954.

[23] Lescos, L.; Sitkiewicz, S.; Beaujean, P.; Blanchard-Desce, M.; Champagne, B. R.; Matito, E.; Castet, F. Performance of DFT functionals for calculating the second-order nonlinear optical properties of dipolar merocyanines. *Phys. Chem. Chem. Phys.* **2020**, 16579–16594.

[24] de Wergifosse, M.; Champagne, B. Electron correlation effects on the first hyperpolarizability of push–pull π-conjugated systems. *J. Chem. Phys.* **2011**, *134*, 074113.

[25] Chołuj, M.; Alam, M. M.; Beerepoot, M. T. P.; Sitkiewicz, S. P.; Matito, E.; Ruud, K.; Zaleśny, R. Choosing bad versus worse: predictions of two-photon-absorption strengths based on popular density functional approximations. *J. Chem. Theory Comput.* **2022**, *18*, 1046–1060.

[26] Purvis III, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[27] Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 509–530.

[28] Leininger, M. L.; Allen, W. D.; Schaefer, H. F.; Sherrill, C. D. Is Møller–Plesset perturbation theory a convergent ab initio method? *J. Chem. Phys.* **2000**, *112*, 9213–9222.

[29] Olsen, J.; Christiansen, O.; Koch, H.; Jørgensen, P. Surprising cases of divergent behavior in Møller–Plesset perturbation theory. *J. Chem. Phys.* **1996**, *105*, 5082–5090.

[30] Schurkus, H. F.; Luenser, A.; Ochsenfeld, C. Communication: Almost error-free resolution-of-the-identity correlation methods by null space removal of the particle-hole interactions. *J. Chem. Phys.* **2017**, *146*, 211106.

[31] Duchemin, I.; Li, J.; Blase, X. Hybrid and Constrained Resolution-of-Identity Techniques for Coulomb Integrals. *J. Chem. Theory Comput.* **2017**, *13*, 1199–1208.

[32] Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.

[33] Pulay, P.; Saebø, S. Orbital-invariant formulation and second-order gradient evaluation in Møller-Plesset perturbation theory. *Theor. Chim. Acta* **1986**, 357–368.

[34] Saebø, S.; Pulay, P. Local configuration interaction: An efficient approach for larger molecules. *Chem. Phys. Lett.* **1985**, 13–18.

[35] Nagy, P. R.; Samu, G.; Kállay, M. An integral-direct linear-scaling second-order Møller–Plesset approach. *J. Chem. Theory Comput.* **2016**, 4897–4914.

[36] Ochsenfeld, C.; Kussmann, J.; Lambrecht, D. S. *Rev. Comput. Chem.*; John Wiley & Sons, Ltd, 2007; Chapter 1, pp 1–82.

[37] Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496–4501.

[38] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On some approximations in applications of Xα theory. *The Journal of Chemical Physics* **1979**, *71*, 3396–3402.

[39] Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[40] Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

[41] Kossmann, S.; Neese, F. Comparison of two efficient approximate Hartee–Fock approaches. *Chem. Phys. Lett.* **2009**, *481*, 240–243.

[42] Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *J. Chem. Phys.* **2009**, *130*, 114108.

[43] Van Alsenoy, C. Ab initio calculations on large molecules: The multiplicative integral approximation. *J. Comput. Chem.* **1988**, *9*, 620–626.

[44] Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496–4501.

[45] Florian Weigend, H. P., Marco Häser; Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143 – 152.

[46] Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On first-row diatomic molecules and local density models. *J. Chem. Phys.* **1979**, *71*, 4993–4999.

[47] Pulay, P. Localizability of Dynamic Electron Correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

[48] Schütz, M. Low-order scaling local electron correlation methods. III. Linear scaling local perturbative triples correction (T). *J. Chem. Phys.* **2000**, *113*, 9986–10001.

[49] Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, 034108.

[50] Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, 034106.

[51] Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladóczki, B.; Kállay, M. An efficient linear-scaling CCSD(T) method based on local natural orbitals. *J. Chem. Phys.* **2013**, 094105.

[52] Nagy, P. R.; Samu, G.; Kállay, M. An integral-direct linear-scaling second-order Møller–Plesset approach. *J. Chem. Theory Comput.* **2016**, 4897–4914.

[53] Guo, Y.; Riplinger, C.; Becker, U.; Liakos, D. G.; Minenkov, Y.; Cavallo, L.; Neese, F. Communication: An improved linear scaling perturbative triples correction for the domain based local pair-natural orbital based singles and doubles coupled cluster method [DLPNO-CCSD(T)]. *J. Chem. Phys.* **2018**, *148*, 011101.

[54] Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

[55] Pinski, P.; Neese, F. Communication: Exact analytical derivatives for the domain-based local pair natural orbital MP2 method (DLPNO-MP2). *J. Chem. Phys.* **2018**, *148*, 031101.

[56] Sylvetsky, N.; Banerjee, A.; Alonso, M.; Martin, J. M. L. Performance of localized coupled cluster methods in a moderately strong correlation regime: Hückel–Möbius interconversions in expanded porphyrins. *J. Chem. Theory Comput.* **2020**, *16*, 3641–3653.

[57] Stoychev, G. L.; Auer, A. A.; Gauss, J.; Neese, F. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *J. Chem. Phys.* **2021**, *154*, 164110.

[58] Grimme, S.; Goerigk, L.; Fink, R. F. Spin-component-scaled electron correlation methods. *WIREs, Comput. Mol. Sci.* **2012**, *2*, 886–906.

[59] Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials (Chem. Phys. Letters 240 (1995) 283-290). *Chem. Phys. Lett.* **1995**, *242*, 652–660.

[60] Boström, J.; Aquilante, F.; Pedersen, T. B.; Lindh, R. Ab initio density fitting: Accuracy assessment of auxiliary basis sets from Cholesky decompositions. *J. Chem. Theory Comput.* **2009**, *5*, 1545–1553.

[61] Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, 224108.

[62] Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

[63] Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

[64] Bernholdt, D. E.; Harrison, R. J. Large-scale correlated electronic structure calculations: the RI-MP2 method on parallel computers. *Chem. Phys. Lett.* **1996**, *250*, 477–484.

[65] Kendall, R. A.; Früchtl, H. A. The impact of the resolution of the identity approximate integral method on modern ab initio algorithm development. *Theor. Chem. Acc.* **1997**, *97*, 158–163.

[66] Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree–Fock and hybrid DFT calculations. A 'chain-of-spheres' algorithm for the Hartree–Fock exchange. *Chem. Phys.* **2009**, *356*, 98–109.

[67] Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103.

[68] Sparta, M.; Neese, F. Chemical applications carried out by local pair natural orbital based coupled-cluster methods. *Chem. Soc. Rev.* **2014**, *43*, 5032–5041.

[69] Foster, J. M.; Boys, S. F. Canonical configurational interaction procedure. *Rev. Mod. Phys.* **1960**, *32*, 300–302.

[70] Edmiston, C.; Ruedenberg, K. Localized Atomic and molecular orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457–464.

[71] Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for *ab initio* and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

[72] Sironi, M.; Genoni, A.; Civera, M.; Pieraccini, S.; Ghitti, M. Extremely localized molecular orbitals: theory and applications. *Theor. Chem. Acc.* **2007**, *117*, 685–698.

[73] Nagy, P. R.; Samu, G.; Kállay, M. Optimization of the linear-scaling local natural orbital CCSD(T) method: Improved algorithm and benchmark applications. *J. Chem. Theory Comput.* **2018**, *14*, 4193–4215.

[74] Kállay, M. et al. The MRCC program system: Accurate quantum chemistry from water to proteins. *J. Chem. Phys.* **2020**, *152*, 074107.

[75] Rolik, Z.; Szegedy, L.; Ladjánszki, I.; Ladóczki, B.; Kállay, M. An efficient linear-scaling CCSD(T) method based on local natural orbitals. *J. Chem. Phys.* **2013**, 094105.

[76] Kállay, M. Linear-scaling implementation of the direct random-phase approximation. *J. Chem. Phys.* **2015**, *142*, 204105.

[77] Neese, F. Software update: The ORCA program system—Version 5.0. *WIREs Computational Molecular Science* **2022**, *12*, e1606.

[78] Nagy, P. R.; Kállay, M. Optimization of the linear-scaling local natural orbital CCSD(T) method: Redundancy-free triples correction using Laplace transform. *J. Chem. Phys.* **2017**, *146*, 214106.

[79] Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

[80] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

[81] Medved', M.; Stachova, M.; Jacquemin, D.; André, J.-M.; Perpète, E. A generalized Romberg differentiation procedure for calculation of hyperpolarizabilities. *J. Mol. Struct.: THEOCHEM* **2007**, *847*, 39–46.

[82] Rutishauser, H. Ausdehnung des Rombergschen prinzips. *Numer. Math. (Heidelb.)* **1963**, *5*, 48–54.

[83] Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. Exploring the accuracy limits of local pair natural orbital coupled-cluster theory. *J. Chem. Theory Comput.* **2015**, *11*, 1525–1539.

[84] Nagy, P. R.; Kállay, M. Approaching the basis set limit of CCSD(T) energies for large molecules with local natural orbital coupled-cluster methods. *J. Chem. Theory Comput.* **2019**, *15*, 5275–5298.

[85] Wouters, S.; Limacher, P. A.; Van Neck, D.; Ayers, P. W. Longitudinal static optical properties of hydrogen chains: Finite field extrapolations of matrix product state calculations. *J. Chem. Phys.* **2012**, *136*, 134110.

[86] Janssen, C. L.; Nielsen, I. M. New diagnostics for coupled-cluster and Møller–Plesset perturbation theory. *Chem. Phys. Lett.* **1998**, *290*, 423–430.

[87] Nielsen, I. M.; Janssen, C. L. Double-substitution-based diagnostics for coupled-cluster and Møller–Plesset perturbation theory. *Chem. Phys. Lett.* **1999**, *310*, 568–576.

[88] Lee, T. J.; Head-Gordon, M.; Rendell, A. P. Investigation of a diagnostic for perturbation theory. Comparison to the T1 diagnostic of coupled-cluster theory. *Chem. Phys. Lett.* **1995**, *243*, 402–408.

[89] Lee, T. J.; Taylor, P. R. A diagnostic for determining the quality of single-reference electron correlation methods. *Int. J. Quantum Chem.* **1989**, *36*, 199–207.

[90] Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. A simple DFT-based diagnostic for nondynamical correlation. *Theor. Chem. Acc.* **2013**, *132*, 1–9.

[91] Bartlett, R. J.; Park, Y. C.; Bauman, N. P.; Melnichuk, A.; Ranasinghe, D.; Ravi, M.; Perera, A. Index of multi-determinantal and multi-reference character in coupled-cluster theory. *J. Chem. Phys.* **2020**, *153*, 234103.

[92] Ramos-Cordoba, E.; Salvador, P.; Matito, E. Separation of dynamic and nondynamic correlation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24015–24023.

[93] Ramos-Cordoba, E.; Matito, E. Local Descriptors of dynamic and nondynamic correlation. *J. Chem. Theory Comput.* **2017**, *13*, 2705–2711.

[94] Via-Nadal, M.; Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Matito, E. Singling out Weak and Strong Correlation. *J. Phys. Chem. Lett.* **2019**, *10*, 4032–4037.

[95] Via-Nadal, M.; Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Matito, E. Natural range separation of the Coulomb hole. *J. Chem. Phys.* **2022**, *156*, 184106.

[96] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

Chapter 5

Performance of MP3:KS methods for the computation of static linear and nonlinear optical properties

This chapter has been realized in the framework of a collaboration involving:
Besalú-Sala P. (Univ. of Girona), Ramos-Cordoba E. (Univ. of Basque Country), Zaleśny R. (Univ. of Wrocław), and Luis J. M. (Univ. of Girona)

5.1 Introduction

As discussed in the previous chapters, the inclusion of electron correlation is crucial for an accurate evaluation of molecular nonlinear optical properties. However, high-level CCSD(T) calculations are out of the reach of current computational resources for real-life molecules, as this method scales as $\mathcal{O}(N^7)$ (with N the number of basis functions used for expanding the molecular orbitals). Therefore, it is often necessary to make the compromise of decreasing the level of theory in order to reduce the computational cost.

In this context, the preferred wavefunction-based methods for computing NLO properties of medium-size molecules are CCSD, which needs several iterations scaling as $\mathcal{O}(N^6)$, and MP2, which scales as $\mathcal{O}(N^5)$. MP3 methods also scale $\mathcal{O}(N^6)$ but they are less time consuming because they do not need several iterations to reach the convergence. However, they are rarely used because they do not systematically improve the MP2 results.

Nevertheless, we address in this chapter the performance of a recent implementation of the MP3 method, in which the reference molecular orbitals are obtained from DFT (KS-MOs) instead of the standard HF orbitals. The MP3:KS method, has been shown to increase the accuracy of MP3:HF, and provide results comparable to CCSD for thermochemical barriers, noncovalent interactions and dipole moments [1, 2]. The success of this method is attributed to the deficiencies of the HF method which brings a fictitious spin contamination even in closed-shell systems and tends to give poor reference MOs. In fact, HF tends to overlocalize the MOs giving a poor reference to the subsequent wavefunction calculations [3]. This issue led to the development of the so-called orbital optimized MP2 method (OO-MP2) [4], in which HF orbitals are optimized in the presence of MP2 correlation. A more appropriate reference is given by KS-MOs, which conversely, might suffer from the delocalization error depending on the choice of the functional (see section 2.5). However, it has been observed [2] that MP2:KS and MP3:KS methods are not particularly sensitive to the choice of the reference functional for the computation of energies. Because of the use of canonical orbitals, non-Brillouin single excitations between occupied and virtual orbitals are present at both MP2 and MP3 level. However, MP3:KS method is optimized by including MP2 single excitations, and neglecting the corresponding MP3 ones. Interesting, with this construction MP2:KS overestimates the correlation effects underperforming with respect to MP2:HF (see the expression of MP2:KS and MP3:KS energies in Section 2.3.4). On the other hand, in MP3:KS this contribution is mitigated by the presence of MP3 doubles which reduces it and allows to improve canonical MP3. In this chapter, we address the efficiency of MP3:KS for computing the static linear and nonlinear optical properties (α, β, and γ) of reference molecules.

5.2 Computational details

The main equations of the MP3:KS method are provided in Section 2.3.4. Calculations have been performed by neglecting the non-Brillouin single excitations of the MP3 perturbation while including the corresponding MP2 ones, as suggested by the developers of the method [2]. The systems studied are the molecules belonging to the γ-NLO (both γ-NLO-A and γ-NLO-B) and β-NLO sets employed in Chapter 4. The single-reference character of the molecules of the β-NLO-set has been further assessed through the T1 diagnostic criterion [5–7] over the CCSD wavefunctions. To reduce the computational cost, MP3 corrections were calculated by making use of the resolution of identity (RI) approximation. We demonstrated in Chapter 4 that the use of this acceleration scheme at the MP2 level does not significantly impact the computed NLO properties. The very good agreement between MP3 and RI-MP3 results was further demonstrated here (see Table B.4) by computing the NLO responses of molecules from the γ-set (excluding the largest molecules PDA4 and PDA5 that are not computationally affordable at the MP3 level). Thus, we will not make any distinction between MP3 and RI-MP3.

Reference CCSD(T) values of the NLO properties are taken from Chapter 4. All the calculations have been performed on a developer's version of Qchem [8] using the aug-cc-pVDZ basis set, while the corresponding aug-cc-pVDZ auxiliary basis set was employed whenever RI is applied. Single-point calculations have been performed using an energy threshold of 10^{-9} a.u. Several density functional approximations (DFAs) belonging to different rungs of Jacob's ladder have been used to generate the reference orbitals used in the MP3:KS scheme. In particular, we tested three functionals making use of the Becke88 [9] exchange functional and LYP correlation functional [10], differing in the amount (if any) and treatment of exact HF exchange: BLYP, [9], B3LYP [11], and CAM-B3LYP [12], which belong to the GGA, hybrid GGA and long-range corrected hybrid GGA families, respectively (see section 2.5.1). In addition, we tested three range-separated functionals of the ωB97 series, namely ωB97-XD [13], ωB97-XV [14], and ωB97-MV [15]. NLO properties have been computed as numerical derivatives of the field-dependent electronic energies using the finite-field procedure, as described in the previous chapter. The performance of the various MP3:KS approximations are evaluated by using the same statistical measures as in Chapter 4, namely the Mean Value (MV), Mean Absolute Error (MAE), the Root Mean-Square Error (RMSE), the Maximum Error (MAX), and the percentage MAE (%MAE).

5.3 Results

5.3.1 Performance of unrestricted MP3 calculations

Before testing the MP3:KS methods, we evaluated the error originating from spin contamination when using an unrestricted framework. When using the UHF MOs, we found that 21 molecules over the 60 ones composing the γ-set have a broken symmetry ground state, as reflected by the mean values of the $\hat{S}^2$ operator reported in Table B.2, which are significantly larger than the expected value of 0 for closed-shell systems. (Note that $\langle \hat{S}^2 \rangle = 2.0$ for the O_2 molecule owing to its triplet ground state). Very large spin contamination errors were also obtained for molecules of the β-set, as reported in Table B.3.

Table 5.1: Performance of HF:MP3 and UHF:MP3 for calculating NLOPs with respect to CCSD(T) for the γ-NLO-set.

		MP3	
Ref. orbitals		**HF**	**UHF**
μ	MV	7.0E-01	1.0E+00
	MAE	2.0E-02	3.0E-01
	RMSE	4.0E-02	5.0E-01
	MAX	9.0E-02	1.0E+00
	%MAE	3	41
α	MV	1.1E+02	1.6E+02
	MAE	3.4E+00	4.6E+01
	RMSE	1.2E+01	2.7E+02
	MAX	4.3E+01	1.2E+03
	%MAE	3	40
β	MV	2.2E+02	4.9E+03
	MAE	8.9E+01	4.8E+03
	RMSE	2.1E+02	1.3E+04
	MAX	6.6E+02	5.2E+04
	%MAE	65	3443
γ	MV	4.2E+05	3.5E+06
	MAE	6.4E+04	3.0E+06
	RMSE	2.8E+05	1.9E+07
	MAX	2.0E+06	1.5E+08
	%MAE	14	668

The NLO properties calculated using the MP3 and UMP3 methods are compared to reference CCSD(T) calculations for the γ-set in Table 5.1. When allowing the wavefunction to relax towards the unrestricted solution, the NLO properties get much worse, showing error indices one order of magnitude larger than those associated to restricted calculations. This behavior can be related to a well-known problem of MP2 and MP3 calculations, which lose accuracy when moving to an unrestricted frame [16–18]. Unrestricted HF orbitals should be therefore avoided for a proper reproduction of molecular NLO responses.

On the other hand, when using DFT we observe that the only molecule which presents a broken symmetry solution for some of the selected functionals is HNO, while for the rest of the systems the most stable solution is the closed-shell one (see Table B.2). Unrestricted KS MOs can thus be used as reference orbitals in the MP3:KS scheme, at least from the viewpoint of spin contamination issues.

5.3.2 Performance of MP3:KS methods for α and γ

Table 5.2 reports the performance of MP3:KS methods with respect to CCSD(T) reference for computing even-order NLO properties. Molecules of both the γ-NLO and β-NLO sets have been considered. Results provided by standard CCSD, MP2 and MP3 methods using HF orbitals are also reported for comparison.

Table 5.2: Performance of MP2, MP3, CCSD, and various MP3:KS methods with respect to the reference CCSD(T) for computing odd-order NLO responses (α and γ) of molecules of the γ-NLO and β-NLO sets. CAM-B3LYP, ωB97-XD, ωB97-XV, and ωB97-MV functionals are abbreviated respectively as CAM, ωXD, ωXV, and ωMV.

		MP2	**MP3**	**CCSD**			**MP3:KS**			
Ref.	**orbitals**	**HF**	**HF**	**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
α	MV	2.4E+02	2.2E+02	2.2E+02	2.3E+02	2.3E+02	2.3E+02	2.3E+02	2.3E+02	2.3E+02
	MAE	1.0E+01	5.4E+00	7.6E+00	1.5E+01	8.6E+00	4.5E+00	5.7E+00	3.5E+00	3.4E+00
	RMSE	2.0E+01	9.9E+00	1.4E+01	3.7E+01	1.8E+01	7.6E+00	1.2E+01	5.1E+00	5.1E+00
	MAX	9.4E+01	4.3E+01	5.9E+01	2.7E+02	9.9E+01	3.4E+01	8.5E+01	1.8E+01	1.9E+01
	%MAE	4	2	3	7	4	2	3	2	2
γ	MV	1.7E+06	1.4E+06	1.3E+06	3.6E+06	2.7E+06	1.9E+06	1.7E+06	1.6E+06	1.6E+06
	MAE	2.9E+05	1.9E+05	2.3E+05	2.2E+06	1.3E+06	4.1E+05	3.1E+05	1.5E+05	1.6E+05
	RMSE	6.3E+05	3.9E+05	5.6E+05	7.2E+06	4.4E+06	1.1E+06	7.8E+05	3.0E+05	3.5E+05
	MAX	2.9E+06	2.0E+06	3.2E+06	4.7E+07	3.2E+07	6.5E+06	4.2E+06	1.4E+06	2.0E+06
	%MAE	19	12	15	140	85	27	20	10	11

For calculations of α, MP2, MP3 and CCSD present similar errors and could be employed indistinctly as alternative to CCSD(T). MP3 is the best method, slightly outperforming CCSD for all error indicators. For γ, MP3 still outperforms the other canonical wavefunction methods for all of the indices considered (MAE, RMSE, MAX and %MAE).

Moving to MP3:KS results, we observe that varying the density functional approximations (*i.e.*, the reference orbitals included in the perturbational analysis) has a significant impact on the quality of the results. MP3:BLYP and MP3:B3LYP underperform regular MP3 for both α and γ. On the other hand, employing as reference range-separated functionals CAM-B3LYP, ωB97-XD, ωB97-XV and ωB97-MV, generally increases the accuracy of the calculations.

Interestingly, in the case of α, smaller or equal values for all error indicators are observed for MP3:CAM-B3LYP, MP3:ωB97-XV and MP3:ωB97-MV with respect to regular MP3 calculations, while MP3:ωB97-XD performs slightly worse for all the indicators. The best results are obtained with MP3:ωB97-XV ad MP3:ωB97-MV which show very similar indicators. In fact, with respect to MP3:HF, they reduce the MAE from 5.4 to 3.5 and 3.4 a.u. respectively, the RMSE from 9.9 to 5.1 a.u., the MAX from 43 to 18 and 19 a.u., and have the same %MAE equal to 2.%. Considering γ, MP3:ωB97-XV and MP3:ωB97-MV outperform MP3 for all statistical indices, while MP3:CAM-B3LYP and MP3:ωB97-XD underperform for all the indicators. As we have observed for α, the best results are obtained with MP3:ωB97-XV, which slightly outperforms MP3:ωB97-MV for all indicators.

Interestingly, this method outperform even CCSD for all the indicators. The same considerations are obtained from the separated analysis of NLO properties calculated from the γ-NLO and β-NLO sets (see Tables B.6 and B.7).

5.3.3 Performance of MP3:KS methods for μ and β

Dipole moment values calculated at the MP3:KS level for the various functionals are compared in Table B.5 to values calculated using standard post-HF methods (MP2, MP3 and CCSD). MP3:KS calculations show homogeneous results and very little variations when changing the exchange-correlation functional. However, differently from what observed by Bertels and coworkers [1], none of the MP3:KS method beats MP3:HF for the dipole moments, although they have a comparable accuracy.

Table 5.3: Performance of MP2, MP3, CCSD and of various MP3:KS methods with respect to the reference CCSD(T) for computing β of molecules of the γ-NLO-B and β-NLO sets. CAM-B3LYP, ωB97-XD, ωB97-XV and ωB97-MV functionals are abbreviated respectively as CAM, ωXD, ωXV and ωMV.

		MP2	**MP3**	**CCSD**	**MP3:KS**					
Ref. orbitals		**HF**	**HF**	**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
β	MV	8.5E+03	7.5E+03	6.6E+03	1.6E+04	1.2E+04	8.3E+03	8.0E+03	7.3E+03	7.3E+03
	MAE	2.0E+03	1.3E+03	5.6E+02	9.4E+03	5.0E+03	1.8E+03	1.5E+03	9.1E+02	9.5E+02
	RMSE	4.5E+03	2.7E+03	1.2E+03	2.7E+04	1.3E+04	4.1E+03	3.2E+03	1.8E+03	1.9E+03
	MAX	1.3E+04	7.8E+03	4.5E+03	1.2E+05	5.3E+04	1.3E+04	9.5E+03	4.9E+03	5.3E+03
	%MAE	28	18	8	131	70	26	21	13	13

The results related to first hyperpolarizabilities are collected in Table 5.3. Among the three standard *ab initio* methods, CCSD provides the most accurate β values, while the MP3 scheme is better than MP2 for all indicators.

Considering MP3:KS approximations, MP3:BLYP and MP3:B3LYP perform worse than MP3, as it was already observed for odd-order NLO properties. MP3:CAM-B3LYP and MP3:ωB97-XD provide similar errors as MP3 without outperforming it. Interestingly, the two best methods identified for computing α and γ, namely MP3:ωB97-MV and more specifically MP3:ωB97-XV, are also those giving the best performance for β calculations imrpoving MP3. Overall, these methods give results comparable to CCSD, without outperforming the latter in none of the indicators.

In order to gain better insights into the nature of the systems for which MP3:KS methods show worse performances than CCSD for computing β, we compare in Tables 5.4 and 5.5 the statistical indicators separately for molecules of the γ-NLO-B and β-NLO sets. For the γ-NLO-B set, which contains molecules with small β responses, all the methods which include range separated functionals MP3:CAM-B3LYP, MP3:ωB97-XD, MP3:ωB97-XV and MP3:ωB97-MV outperform CCSD for all indicators. For the β-NLO set, which contains molecules with larger β responses (including donor-acceptor push-pull systems), none of the MP3:KS methods outperfoms CCSD. Nevertheless, MP3:ωB97-XV and MP3:ωB97-MV show performances similar to CCSD, and provide better results than the regular MP3 scheme.

Table 5.4: Performance of MP2, MP3, CCSD and various MP3:KS methods with respect to the reference CCSD(T) for computing β of molecules of the γ-NLO-B set. CAM-B3LYP, ωB97-XD, ωB97-XV, and ωB97-MV functionals are abbreviated respectively as CAM, ωXD, ωXV, and ωMV.

		MP2	**MP3**	**CCSD**			**MP3:KS**			
Ref. orbitals		**HF**	**HF**	**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
	MV	2.4E+02	2.2E+02	1.8E+02	2.3E+02	2.0E+02	1.1E+02	1.1E+02	9.7E+01	1.0E+02
	MAE	1.1E+02	8.9E+01	5.0E+01	1.0E+02	6.7E+01	2.6E+01	2.8E+01	4.3E+01	3.9E+01
β	RMSE	2.7E+02	2.1E+02	1.3E+02	3.0E+02	2.2E+02	7.3E+01	7.6E+01	1.2E+02	1.1E+02
	MAX	9.1E+02	6.6E+02	4.7E+02	1.2E+03	9.4E+02	3.1E+02	3.3E+02	5.2E+02	4.7E+02
	%MAE	80	65	37	72	48	19	20	31	28

Table 5.5: Performance of MP2, MP3, CCSD and various MP3:KS methods with respect to the reference CCSD(T) for computing β of molecules of the β-NLO set. CAM-B3LYP, ωB97-XD, ωB97-XV and ωB97-MV functionals are abbreviated respectively as CAM, ωXD, ωXV and ωMV.

		MP2	**MP3**	**CCSD**			**MP3:KS**			
Ref. orbitals		**HF**	**HF**	**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
	MV	1.2E+04	1.0E+04	9.2E+03	2.2E+04	1.6E+04	1.2E+04	1.1E+04	1.0E+04	1.0E+04
	MAE	2.8E+03	1.8E+03	7.7E+02	1.3E+04	7.0E+03	2.6E+03	2.1E+03	1.3E+03	1.3E+03
β	RMSE	4.5E+03	2.7E+03	1.2E+03	2.7E+04	1.3E+04	4.1E+03	3.2E+03	1.8E+03	1.9E+03
	MAX	1.3E+04	7.8E+03	4.5E+03	1.2E+05	5.3E+04	1.3E+04	9.5E+03	4.9E+03	5.3E+03
	%MAE	28	18	8	131	70	26	21	13	13

5.3.4 Performance of MP3:KS for large molecules

In this section, the accuracy of the most promising functionals according to the results of the previous section, namely ωB97-XV and ωB97-MV, is further analyzed for computing the γ and β responses of the largest molecules belonging to the γ-NLO and β-NLO sets. For the evaluation of β, this subset comprises 16 push-pull derivatives: PA10, PA12, PMI6, and PDA3 combining four donor-acceptor pairs at their extremities. The non-substituted PMI6 molecule was also considered, since it has a non centrosymmetric shape for a total of 17 molecules. For the evaluation of γ, the considered subset includes the 17 molecules used for β, plus PA10, PA12, PDA3, PDA4, and PDA5 without any substituents, for a total of 22 molecules.

The results for γ are collected in Table 5.6. As can be seen, MP2 shows error indices comparable to CCSD, with slightly worse MAE, RMSE and %MAE values but better MAX. On the other hand, the standard MP3 scheme outperforms CCSD and MP2 for all indicators. Employing MP3:ωB97-XV and MP3:ωB97-MV improves also the accuracy of the MP3 γ values, roughly reducing by half all error indicators with respect to CCSD.

The results for β are collected in Table 5.7. In this case, we observe that CCSD outperforms all the other methods with respect to all statistical indicators. Overall, MP3 outperforms MP2 for all error indices, but the difference is too small to justify the computational effort. On the other hand, MP3:KS methods, and more especially MP3:ωB97-XV as already observed for γ, result particularly effective for this specific set of molecules, with a reduction of the relative errors with respect to canonical MP3. In particular, the %MAE value is lowered more than half compared to MP2 calculations.

Table 5.6: Performance of MP2, MP3, CCSD and various MP3:KS methods with respect to the reference CCSD(T) for the calculation of γ, for a subset of 22 molecules selected in the γ-NLO and β-NLO sets (see text). CAM-B3LYP, ωB97-XD, ωB97-XV and ωB97-MV functionals are abbreviated respectively as CAM, ωXD, ωXV and ωMV.

		MP2	**MP3**	**CCSD**	**MP3:KS**	
Ref. orbitals		**HF**	**HF**	**HF**	**ωXV**	**ωMV**
γ	MV	4.5E+06	3.8E+06	3.5E+06	4.2E+06	4.2E+06
	MAE	9.0E+05	5.4E+05	7.0E+05	4.3E+05	4.8E+05
	RMSE	1.2E+06	7.2E+05	1.1E+06	5.6E+05	6.6E+05
	MAX	2.9E+06	2.0E+06	3.2E+06	1.4E+06	2.0E+06
	%MAE	22	13	17	10	12

Table 5.7: Performance of MP2, MP3, CCSD and of various MP3:KS methods with respect to the reference CCSD(T) for the calculation of β, for a subset of 17 molecules selected in the γ-NLO-B and β-NLO sets (see text). CAM-B3LYP, ωB97-XD, ωB97-XV and ωB97-MV functionals are abbreviated respectively as CAM, ωXD, ωXV and ωMV.

		MP2	**MP3**	**CCSD**	**MP3:KS**	
Ref. orbitals		**HF**	**HF**	**HF**	**ωXV**	**ωMV**
β	MV	1.7E+06	1.5E+06	1.3E+06	1.4E+06	1.4E+06
	MAE	4.4E+03	3.0E+03	1.1E+03	2.0E+03	2.0E+03
	RMSE	5.9E+03	3.5E+03	1.6E+03	2.4E+03	2.5E+03
	MAX	1.3E+04	7.8E+03	4.5E+03	4.9E+03	5.3E+03
	%MAE	31	21	8	14	15

5.3.5 Performance of the MP2:KS and DFT methods for computing NLO properties

For the sake of completeness, we also addressed the performance of the MP2:KS method for computing NLO properties. The results are reported in Tables B.8-B.10. As already observed for the calculation of ground-state energies and dipole moments, [2] these methods do not introduce any improvement with respect to standard MP2 methods. Moreover, in Tables B.11-B.13 we report the results obtained with the standard DFAs for the computation of static linear and nonlinear optical properties. By comparing these results with those gathered in Tables 5.2, 5.3, 5.4 and 5.5 we can see that MP3:KS calculations have a better accuracy than regular DFT calculations using the same functional, pointing out the need of wavefunction methods to optimally reproduce NLOPs.

5.4 Conclusions

In this study, we have addressed the performance of MP3:KS schemes for computing static NLO properties of small organic molecules. We observed that for closed-shell molecules NLO properties evaluated with unrestricted HF reference orbitals are not reliable because of spin contamination, therefore, using reference KS orbitals which avoid in most of the cases the break of the spin symmetry allows to improve the results. Moreover, we observed that even for MP3 the employment of the RI approximation reduces substantially the computational cost without affecting accuracy of the properties evaluated. We have demonstrated that, when employed in association with range separated functionals involving a large amount of HF exchange, MP3:KS improves the accuracy of the computed NLO responses with respect to standard MP3. For α and γ, MP3:ωB97-XV provides accurate results and can be used as a cheaper alternative to CCSD. For β we observe two distinct behaviors. For molecules from the γ-NLO set, some of the MP3:KS schemes outperform CCSD, in particular MP3:CAM-B3LYP. For the β-NLO set containing systems with large β, CCSD remains the best method among the tested ones. However, MP3:KS schemes still provides better results than the standard MP3 method, and MP3:ωB97-XV emerges as the best among the latter.

As mentioned in the computational section, all MP3:KS calculations reported in this chapter have been performed by neglecting the non-Brillouin single excitations from the MP3 corrections. Further tests are thus needed to address the reliability of this approximation for computing NLO properties, especially in the case of large π-conjugated molecules, for which the MP3:KS method does not improve the quality of the results compared to CCSD calculations.

Bibliography

[1] Bertels, L. W.; Lee, J.; Head-Gordon, M. Third-Order Møller–Plesset Perturbation Theory Made Useful? Choice of Orbitals and Scaling Greatly Improves Accuracy for Thermochemistry, Kinetics, and Intermolecular Interactions. *J. Phys. Chem. Lett.* **2019**, *10*, 4170–4176.

[2] Rettig, A.; Hait, D.; Bertels, L. W.; Head-Gordon, M. Third-Order Møller–Plesset Theory Made More Useful? The Role of Density Functional Theory Orbitals. *J. Chem. Theory Comput.* **2020**, *16*, 7473–7489.

[3] Li, C.; Yang, W. On the piecewise convex or concave nature of ground state energy as a function of fractional number of electrons for approximate density functionals. *J. Chem Phys.* **2017**, *146*, 074107.

[4] Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *J. Chem. Phys.* **2007**, *126*, 164101.

[5] Lee, T. J.; Head-Gordon, M.; Rendell, A. P. Investigation of a diagnostic for perturbation theory. Comparison to the T1 diagnostic of coupled-cluster theory. *Chem. Phys. Lett.* **1995**, *243*, 402–408.

[6] Lee, T. J.; Taylor, P. R. A diagnostic for determining the quality of single-reference electron correlation methods. *Int. J. Quantum Chem.* **1989**, *36*, 199–207.

[7] Fogueri, U. R.; Kozuch, S.; Karton, A.; Martin, J. M. L. A simple DFT-based diagnostic for nondynamical correlation. *Theor. Chem. Acc.* **2013**, *132*, 1–9.

[8] Shao, Y. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

[9] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

[10] 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*, 785–789.

[11] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[12] A new hybrid exchange–correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chemical Physics Letters* **2004**, *393*, 51–57.

[13] Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

[14] Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

[15] Mardirossian, N.; Head-Gordon, M. ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.

[16] Nobes, R. H.; Pople, J. A.; Radom, L.; Handy, N. C.; Knowles, P. J. Slow convergence of the møller-plesset perturbation series: the dissociation energy of hydrogen cyanide and the electron affinity of the cyano radical. *Chem. Phys. Lett.* **1987**, *138*, 481–485.

[17] Gill, P. M. W.; Pople, J. A.; Radom, L.; Nobes, R. H. Why does unrestricted Møller–Plesset perturbation theory converge so slowly for spin-contaminated wave functions? *J. Chem. Phys.* **1988**, *89*, 7307–7314.

[18] Jensen, F. A remarkable large effect of spin contamination on calculated vibrational frequencies. *Chem. Phys. Lett.* **1990**, *169*, 519–528.

Chapter 6

Impact of van der Waals interactions on the structural and nonlinear optical properties of azobenzene switches

This chapter is based on the publication:
Naim C., Castet F. and Matito E., *Phys. Chem. Chem. Phys.*, 2021,23, 21227-21239

6.1 Introduction

Materials and molecules that exhibit large second-order nonlinear optical properties (NLOPs), *i.e.*, whose dielectric polarization responds quadratically to the electric-field component of light, find applications in laser components, optical communications, data processing, and storage [1, 2] as well as in bioimaging and therapy [3–5]. In addition to the quest for materials able to deliver high second harmonic generation (SHG) responses, an important research topic in the field is the design of new molecular systems with commutable second-order NLOPs. Indeed, such systems that can isomerize between two (or more) stable and reversibly interconvertible forms with a large difference in their first hyperpolarizabilities (β) present high interest for exploitation in optoelectronic and photonic devices, such as high-density optical memories with multiple storages and nondestructive readout capacity [6–8].

Since light is a fast, non-invasive, and low-cost way to trigger the interconversion between the different isomeric forms, photochromic molecules constitute the most prominent class of candidates for eventual exploitation in NLO devices. One of the most studied families of photochromes that exhibit large β contrasts is that of double-bond isomer derivatives, which undergo a reversible isomerization between a stable *trans* (E) and a metastable *cis* (Z) form upon illumination in the UV range. Among them, azobenzene derivatives have been the subject of many theoretical and experimental investigations, mostly due to the high first hyperpolarizabilities obtained when adding donor and acceptor groups to both ends of the azobenzene core [9–17]. Reversible SHG switch-

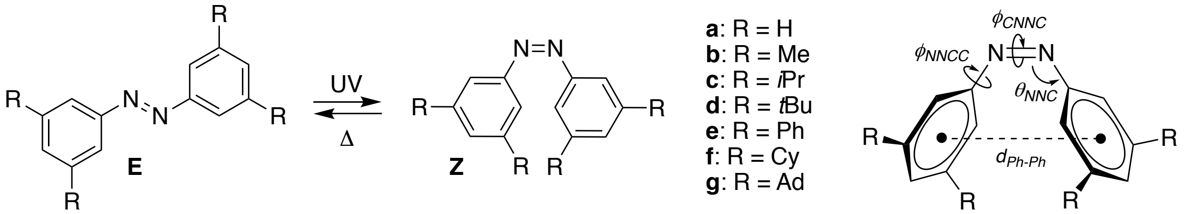


Figure 6.1: Substituted azobenzene derivatives investigated in this study (Cy = cyclohexyl, Ad = adamantyl) and geometrical parameters used to characterize the *cis* conformers.

ing in solid-state structures incorporating azobenzenes as NLO functional units was also demonstrated, whether in poled polymers [18, 19], LB films [7], liquid-crystalline polymers [20, 21], or self-assembled monolayers [22–25]. Recent theoretical works also highlighted that azobenzene NLO switches could be exploited as versatile probes for phase recognition in biological environments [26, 27].

At the molecular level, it is well known that the nature of the chemical substituents determines the relative thermodynamic stability of the *trans* and *cis* forms of azobenzene derivatives and thus plays a critical role in their photo-isomerization process. In particular, chemical functionalization in the *meta*-position with bulky substituents has been shown to lower the *cis* → *trans* thermal reaction rates due to the attractive London dispersion forces that stabilize the *cis* isomer [28]. In addition to their key role in the structure and photochemical processes of molecular switches, Van der Waals interactions are also expected to impact their NLO properties. However, although the impact of London dispersion forces on molecular properties has recently received some attention [29–31], their influence on the NLOPs of molecular switches has been overlooked thus far.

In this work, we report a comprehensive theoretical study of the structural and NLO properties of a collection of azobenzene molecules symmetrically substituted in *meta*-position with functional groups of different bulkiness (Figure 6.1). In the first part, we address the performance of a large set of density functional approximations (DFAs) for calculating the geometries and relative energies of the E and Z isomers by using the second-order approximate coupled-cluster singles and doubles (CC2) results as the theoretical reference. Computational results are also compared to previously reported data [28]. In the second part, we investigate the accuracy of the different DFAs for reproducing the static first hyperpolarizability of the *cis* forms compared to *ab initio* second-order Møller–Plesset (MP2) and CC2 calculations. Finally, we use the partition of NLOPs in terms of orbital contributions (PNOC) [32] to decipher the individual role of the *meta*-substituents on the NLO responses of this series of compounds.

6.2 Computational methodology

Reference geometries were optimized at the CC2 level [33] using the Resolution of Identity (RI) [34] and frozen core approximations. Despite MP2 has been reported to fail for some intermolecular complexes [35], it remains an important approach for evaluating interaction energies, particularly of large molecules for which higher-level calculations are prohibitive. In Table C.1, based on the results of the Diet-GMTKN55 benchmark set [36], we show that CC2 and MP2 offer very similar performance for intramolecular noncovalent interactions (MAE below 2 kcal/mol and RMSE below 1.5 kcal/mol for both methods). The cc-pVDZ basis set [37] was used in all calculations and employed as the auxiliary basis for RI computations [38]. This basis set guarantees a good compromise between accuracy and computational cost because it provides very similar geometrical structures and energy differences between Z and E isomers as those obtained using the more extended cc-pVTZ basis set (see Tables C.2 and C.3). The energy convergence threshold in both single-point SCF and CC calculations has been set to 10^{-6} a.u. RI-MP2, a wide range of DFAs (see Table C.4), and the Hartree-Fock (HF) method were tested to address their ability to reproduce reference RI-CC2 structures. HF completely neglects electron correlation and thus is expected to perform poorly in the systems affected by dispersion. It is included here as a means to quantify the amount of electron correlation (and dispersion) introduced by the other methods. The selected DFAs involve two different dispersion correction schemes: the Grimme's D3 scheme [39–42] and the nonlocal Van Der Waals density functional (VV) scheme developed by Vydrov and Van Voorhis [43]. To size up the impact of these corrections on the molecular structures, calculations were also performed using the uncorrected counterparts of the DFAs. We have also employed Tα-LC-BLYP, recently designed by Besalú-Sala et al [44]. Tα-LC-BLYP employs LC-BLYP with a molecule-dependent value of the range-separation parameter. This parameter is determined from the static polarizability obtained with the original LC-BLYP employing an empirical equation obtained from the calibration of the CCSD(T) static second hyperpolarizabilities [44].

In order to confirm the single-reference nature of the title molecules, we have performed the D1 diagnostics [45] on the CC2 wavefunctions. In all cases, D1 values are close to the threshold value of 0.05, indicating the partial multireference character of these wavefunctions (see C.5). We have also computed the I_{ND} index [46, 47] obtained from the range-separation partition of the Coulomb hole [48, 49]. I_{ND} is proportional to the deviation from idempotency of the first-order reduced density matrix [50, 51]. In all cases, I_{ND} values are rather small compared to other multireference molecules [52] (see C.5). All in all, we can conclude that these molecules present a rather mild multireference character and they can be reasonably well described with single-reference methods such as MP2 and CC2.

The selected set of DFAs was then employed to evaluate the static components of the first hyperpolarizability tensor (*i.e.*, in the zero frequency limit) of the azobenzene derivatives, as well as the total first hyperpolarizability defined as follows:

$$\beta(0;0,0) = \beta = \frac{1}{5}\sqrt{\beta_x^2 + \beta_y^2 + \beta_z^2} \tag{6.1}$$

where

$$\beta_i = \sum_{j=x,y,z} (\beta_{ijj} + \beta_{jij} + \beta_{jji}) \tag{6.2}$$

and i=x,y,z. Since *trans* forms have a centrosymmetric shape with no second-order NLO responses, β values were calculated for *cis* forms only. Reference values were computed using a finite field (FF) differentiation procedure of the CC2 and MP2 electronic energies. The performance of RI-MP2 and MP2 were evaluated for the molecule R=H in Section C.10 of the Annex in order to verify that the RI approximation does not influence the numerical stability of third-order derivatives. For the compound with the adamantyl substituent, NLOP calculations were performed only at the CAM-B3LYP level due to computational limitations. Linear and nonlinear optical properties were calculated using the coupled-pertubed Kohn-Sham (CPKS) equations implemented in Gaussian [53] for all the DFAs, except for the VV ones. In all other cases, the optical response was calculated using numerical differentiation. A Romberg scheme [54, 55] was employed to control and improve the accuracy of the numerical derivatives, employing field amplitudes starting from ±0.001 a.u. with a multiplicative step of 1.4142. All first hyperpolarizability calculations were performed using the aug-cc-pVTZ basis set. For some of the molecules, the combination of aug-cc-pVTZ and M06L converged to the wrong SCF minimum. Preliminary calculations with the cc-pVTZ basis set were performed and used as a guess to avoid this problem. The pertinent integrals resulting from the electronic energy calculation using DFAs were done numerically using the "Ultrafine" integration grid involving 99 radial shells around each atom and 590 angular points per shell. RI-CC2 and RI-MP2 calculations were performed with TURBOMOLE [56], where the RIJK approximation was employed in addition to the RI approximation for some of the largest molecules (**d**-Z, **e**-Z, and **f**-Z) since it has been shown to reduce the computational cost without affecting the accuracy of the energies [57]. Calculations using DFAs involving the VV approximation were performed with Qchem [58], whereas calculations involving all other DFAs were performed with Gaussian 16 [53]. The root mean square deviation (RMSD) of distances between equivalent atoms of the two molecules has been calculated to quantify the structural differences of two optimized geometries. The RMSD of two structures was calculated using the software provided in Ref. 59. This program calculates the RMSD between two Cartesian geometries using the Kabsch algorithm (1976) [60] or the Quaternion algorithm [61] if rotation is needed before calculating the RMSD. Following the authors recommendation, we have excluded the hydrogen atoms in the calculation of the RMSD. For the sake of completeness, the RMSD values including all the atoms in the molecule are given in Tables C.10-C.11.

6.3 Results and discussion

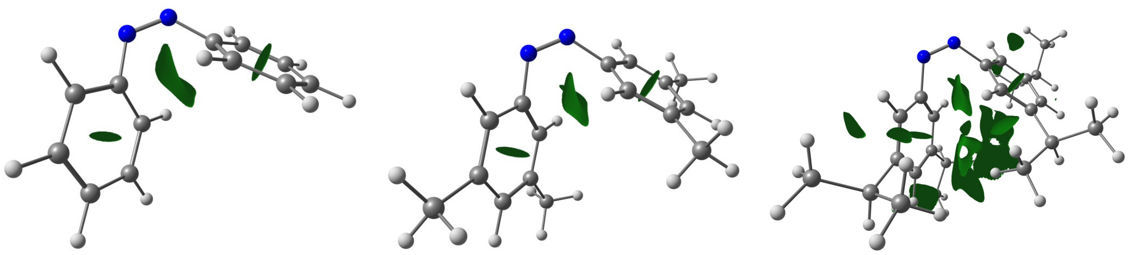
6.3.1 Structure and relative energies of the conformers

6.3.1.1 Reference *ab initio* calculations

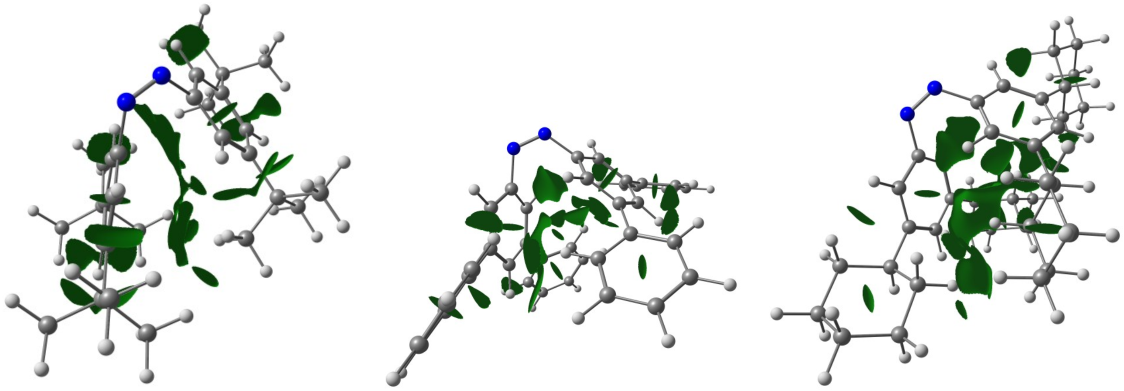
Table 6.1 reports selected structural parameters of the *cis* isomers calculated at the HF, RI-MP2, and RI-CC2 levels with the cc-pVDZ basis set, namely, the N=N and C-N bond lengths (d_{NN} and d_{CN}), the N=N-C bond angles (θ_{NNC}), the torsional angles around the N=N and C-N bonds (ϕ_{CNNC} and ϕ_{NNCC}), as well as the distance d_{PhPh} between the centroids of the two phenyl rings (see Figure 6.1). As expected, geometries calculated using HF, which does not account for dispersion effects, significantly differ from the MP2 and CC2 ones. In particular, HF N=N distances are shorter, while the θ_{NNC} and ϕ_{NNCC} angles are larger, resulting in larger d_{PhPh} distances between the lateral phenyl rings.

The geometries calculated using MP2 and CC2 are quite similar (largest RMSD=0.08 Å, see Table C.2), d_{NN}, d_{CN}, θ_{NNC}, and ϕ_{CNNC} barely changing from one compound to another. In contrast, d_{PhPh} and ϕ_{NNCC}, which are directly linked to the relative orientation of the *meta*-substituents, show slight variations. These two geometrical parameters reflect the interplay between attractive dispersion interactions and repulsive steric hindrance. For MP2 and CC2, the d_{PhPh} distance mostly decreases when increasing the size of the R groups as a consequence of the increase of the dispersion interactions. However, the size of the substituents is not the only factor controlling the strength of the noncovalent interaction. Compound **e** (R = Ph) exhibits the smallest d_{NN} value within the series, owing to the planar shape of the phenyl substituents that allows larger spatial overlap and hence larger attractive London dispersion forces, together with the possibility to adopt a relative spatial orientation minimizing the steric hindrance (**e** also has the smallest ϕ_{NNCC} and the largest ϕ_{CNNC}). Taken as a whole, the geometrical parameters of the *cis* conformer computed at the CC2 and MP2 levels clearly evidence the balance between repulsive steric hindrance and attractive VdW interactions: increasing the bulkiness of the substituents in *meta* position both enhances steric repulsion and dispersion attraction, resulting in slight changes in the geometry of the central part of the compounds. The enhancement of attractive non-covalent interactions (NCI) as the bulkiness of *meta*-substituents increases is further illustrated in Figure 6.2, where the surface area of the NCI isosurface clearly increases with the size of R. In Table C.10, we collect the geometrical data of the *trans* conformers, which are far less affected by electron correlation. Indeed, only d_{NN} is stretched upon the inclusion of electron correlation, the differences between CC2 and MP2 being insignificant.

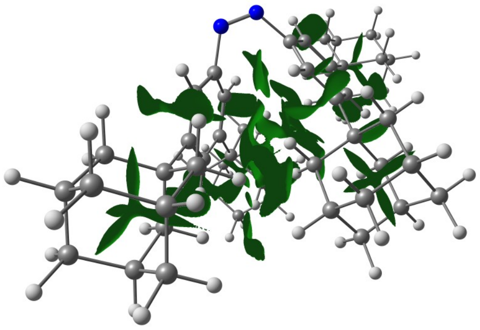
As demonstrated recently [28], the existence of these attractive VdW interactions has a crucial impact on the thermodynamics of the isomerization reaction by stabilizing the *cis* form relative to the *trans* one. As shown in Figure 6.3 and Table C.12, HF calculations do not reproduce these effects correctly and provide very similar energy differences between



(a) R = H (b) R = Me (c) R = iPr



(d) R = tBu (e) R = Ph (f) R = Cy



(g) R = Ad

Figure 6.2: Isosurfaces indicating the regions of weakly attractive intramolecular noncovalent interactions (NCI) in the different molecules, generated using the NCI software [62, 63] with the promolecular density obtained from the CC2 geometries.

Table 6.1: Characteristic distances and angles of the *cis* isomers of compounds **a**-**g** (see Figure 6.1), as calculated at the RI-CC2/cc-pVDZ level (see also Table C.6 for *trans* isomers).

Molecule	d_{NN} [Å]	d_{NC} [Å]	d_{PhPh} [Å]	θ_{NNC} [°]	ϕ_{CNNC} [°]	ϕ_{NNCC} [°]
HF/cc-pVDZ						
a (R = H)	1.215	1.43	4.60	123.9	-5.0	-58.2
b (R = Me)	1.215	1.44	4.59	123.8	-4.9	-60.3
c (R = *i*Pr)	1.215	1.44	4.58	123.8	-5.0	-60.0
d (R = *t*Bu)	1.215	1.44	4.60	124.0	-5.2	-59.7
e (R = Ph)	1.215	1.44	4.58	123.6	-4.7	-58.9
f (R = Cy)	1.215	1.44	4.58	123.8	-5.0	-59.8
g (R = Ad)	1.215	1.44	4.60	124.0	-5.2	-59.0
RI-MP2/cc-pVDZ						
a (R = H)	1.271	1.44	4.42	120.7	-6.8	-53.9
b (R = Me)	1.272	1.44	4.43	121.0	-6.6	-52.5
c (R = *i*Pr)	1.275	1.44	4.15	119.1	-6.2	-56.9
d (R = *t*Bu)	1.276	1.44	4.12	119.1	-6.8	-55.1
e (R = Ph)	1.277	1.45	3.91	117.8	-4.7	-60.2
f (R = Cy)	1.276	1.44	4.14	119.0	-6.7	-56.9
g (R = Ad)	1.277	1.44	4.05	118.6	-6.9	-55.8
RI-CC2/cc-pVDZ						
a (R = H)	1.269	1.43	4.40	121.0	-7.6	-52.7
b (R = Me)	1.279	1.45	4.40	120.9	-5.9	-51.1
c (R = *i*Pr)	1.282	1.45	4.16	119.1	-6.5	-56.2
d (R = *t*Bu)	1.283	1.45	4.12	119.0	-6.8	-55.0
e (R = Ph)	1.284	1.45	3.92	117.8	-5.3	-59.6
f (R = Cy)	1.282	1.45	4.14	119.0	-7.0	-56.4
g (R = Ad)	1.284	1.45	4.07	118.9	-5.7	-54.1

the two isomers (ΔE_{EZ}) for all compounds, as a result of the neglect of dispersion interactions. On the contrary, both CC2 and MP2 calculations confirm the lowering of ΔE_{EZ} when increasing the bulkiness of the R groups. Although CC2 provides slightly smaller ΔE_{EZ} values than MP2, we note the good agreement between these two levels of approximation, with differences in the ΔE_{EZ} values smaller than 1.0 kcal/mol for compounds **a**-**e** and equal to 1.25 kcal/mol for **f**. A larger deviation (1.87 kcal/mol) is obtained for the largest compound **g** (R = Ad).

6.3.1.2 DFT calculations

Using the CC2 results as a reference, we now analyze the performance of different DFAs in reproducing the geometrical structures and relative energies of the *trans* and *cis* forms of the investigated series of azobenzenes. As above, the similarity between DFT and CC2 geometries is measured for both the *trans* and *cis* isomers forms using the RMSD of atomic positions.

In the Annex, we report the differences in the geometries provided by the DFAs including D3 or VV dispersion corrections and their uncorrected counterparts. As expected,

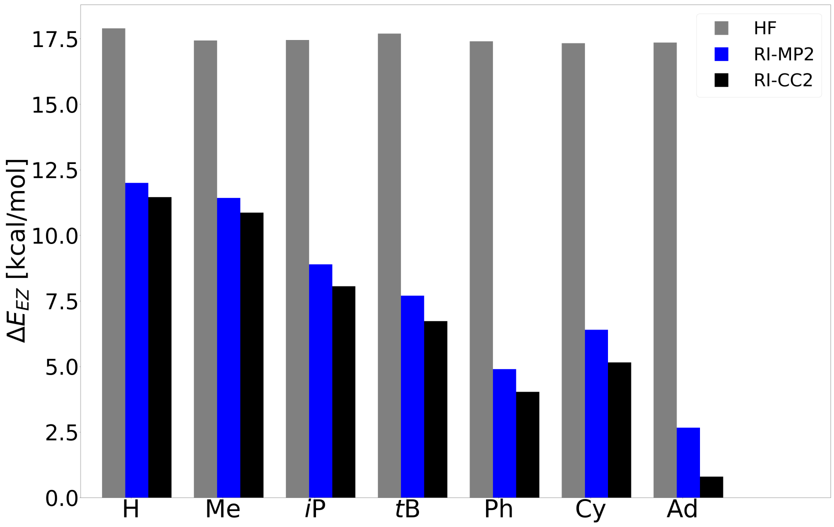


Figure 6.3: Difference in the energy of the *trans* (E) and *cis* (Z) forms, $\Delta E_{EZ} = E_Z - E_E$, as calculated at the HF, RI-MP2, and RI-CC2 levels with the cc-pVDZ basis set.

the RMSD values reported in Table C.7 and Figures C.1 and C.1 show that structural differences are small for *trans* isomers and large for *cis* isomers (in particular for compounds incorporating bulky substituents). The RMSD values calculated for *cis* isomers using ωB97X-D and, especially, M06-2X-D3 are much smaller than those calculated using the other selected DFAs, which confirms that these functionals already include some dispersion corrections in their native form [64, 65].

Geometries of the *cis* isomers optimized using dispersion-corrected DFAs are compared to CC2 geometries in Figure 6.4. The results clearly show that adding dispersion corrections using either the D3 or VV scheme largely improves the matching with the reference geometries. Although acceptable results are obtained with all the tested dispersion-corrected functionals, ωB97-X-D shows the smallest average RMSD (0.10 Å), followed by M06-2X-D3 (0.13 Å), LC-ωPBE08-VV (0.14 Å), B3LYP-D3 (0.16 Å), PBE0-D3 and CAM-B3LYP-D3 (both at 0.17 Å). The values of all the RMSD averaged over the total number of compounds are collected in Table C.9.

We now analyze the performance of the selected DFAs in reproducing the energy difference between the two isomers optimized at the corresponding level of theory (Figure 6.5 and Table C.13). The best agreement with CC2 ΔE_{ZE} values is obtained with the LC-ωPBE-D3 and ωB97-X-D functionals, which display average errors of 1.40 and 1.57 kcal/mol, respectively, for the molecular series. Interestingly, among the studied DFAs, ωB97-X-D was already the best functional to obtain accurate geometries. On the other hand, notice that the molecule with R = Ph gives rise to errors of at least 3 kcal/mol for all the functionals considered. As we can check in Table C.14, it is not due to a wrong estimation of the geometry with these DFAs. Finally, for the sake of completeness, we also calculated the differences in the Gibbs free energies (Table C.13). In general, ΔG_{ZE}

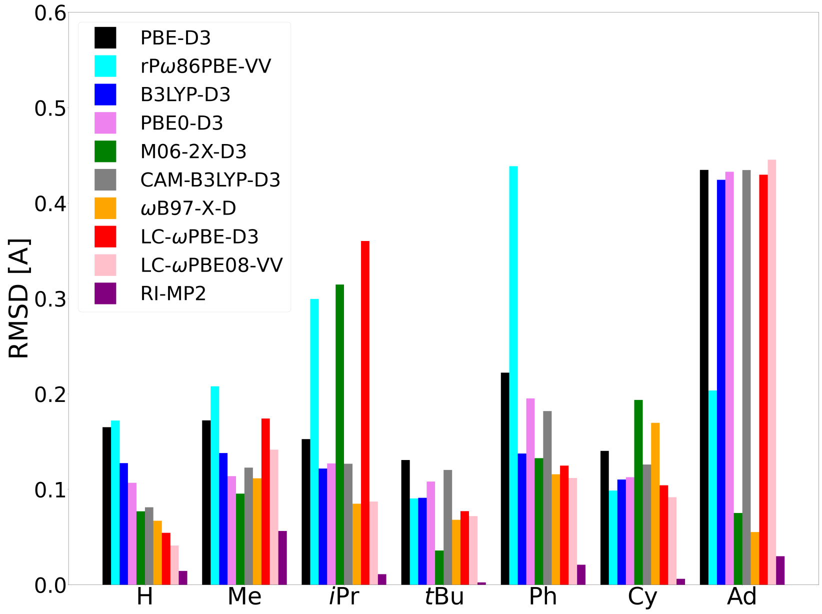


Figure 6.4: RMSD excluding hydrogen atoms between DFAs and CC2 geometries for *cis* isomers. RI-MP2 results are also shown for comparison.

follows the same trend as ΔE_{ZE}, showing a strong reduction upon addition of dispersion corrections to compounds bearing bulky *meta*-substituents.

6.3.2 Second-order NLO properties

6.3.2.1 Reference *ab initio* calculations

Hereafter, all the calculations will be performed using the geometries optimized at RI-CC2/cc-pVDZ level of theory. The total static first hyperpolarizabilities (Eq. 6.1) of the *cis* isomers, calculated at the HF, MP2, and CC2 *ab initio* levels are reported in Table 6.2. Comparison of MP2 and CC2 β values with HF ones provides a direct assessment of the magnitude of electron correlation effects. The first hyperpolarizability increases from HF to MP2, as expected from the localized nature of HF densities. Going from MP2 to CC2 further increases the β values. An excellent linear correlation is obtained between MP2 and CC2 β values (see Figure 6.6), with a very small intercept, suggesting that MP2 values are systematically 64% smaller than CC2 ones. Hence, the trends among different compounds are perfectly reproduced using either MP2 or CC2, and it thus indifferent which method we employ for benchmarking. We will choose MP2 values as the reference ones because their absolute values provide a better agreement with most DFAs (*vide infra*). Notice also that despite the complete neglect of electron correlation, for the first five molecules of the series, HF static polarizabilities also show a very good linear correlation with the MP2 counterparts (see Figure 6.6). Compound **e** (R = Ph) is not shown in the latter figure and it will be discussed in detail in the following section.

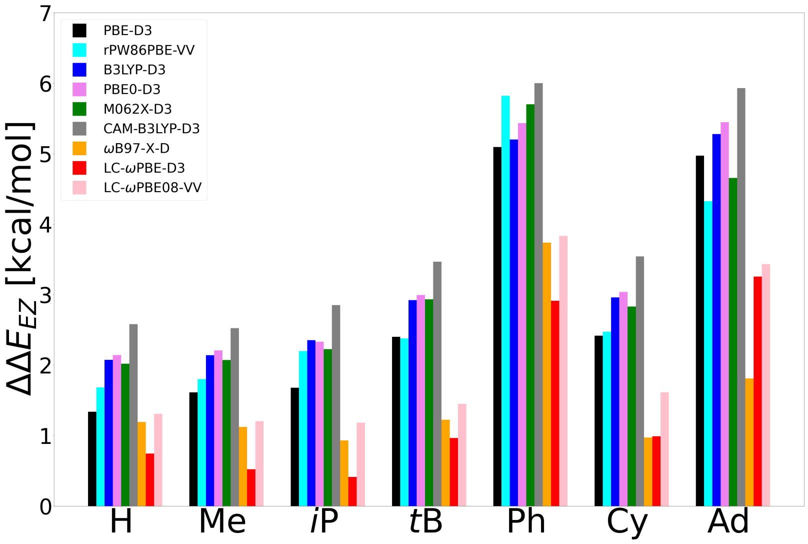


Figure 6.5: Differences in the Z-E isomerization energies calculated at the CC2 level and at the DFT level using DFAs with dispersion corrections ($\Delta\Delta E_{ZE} = \Delta E_{ZE}(DFT) - \Delta E_{ZE}(CC2)$, in kcal/mol).

Table 6.2: Total static first hyperpolarizabilities (β, a.u.) calculated using different *ab initio* levels of approximation together with the aug-cc-pVTZ basis set.

Molecule	HF	MP2	CC2
a (R = H)	7.7	52.5	80.8
b (R = Me)	40.7	119.3	161.7
c (R = *i*Pr)	31.6	99.8	148.5
d (R = *t*Bu)	32.9	84.1	136.1
e (R = Ph)	21.3	152.4	239.8
f (R = Cy)	50.7	127.7	191.0
g (R = Ad)	60.2	-	-

6.3.2.2 DFT calculations

We now address the performance of DFAs in reproducing the static first hyperpolarizability of the azobenzene derivatives **a**-**f**. Since D3 corrections do not affect the electronic density at fixed geometry, D3-corrected functionals are excluded from the benchmark. DFAs incorporating dispersion effects through the VV scheme are considered. The latter will be informative about the effect of dispersion corrections in the calculation of nonlinear optical properties, beyond the indirect effect of dispersion on the geometries.

In the last three columns of Table 6.3, we collect three different measures of the error committed in the hyperpolarizability using a collection of DFAs and RI-CC2 with respect to RI-MP2. These data are reported in Figure 6.6 for all molecules except **g** (R = Ad). The three DFAs that provide the lowest mean absolute error (MAE), mean absolute

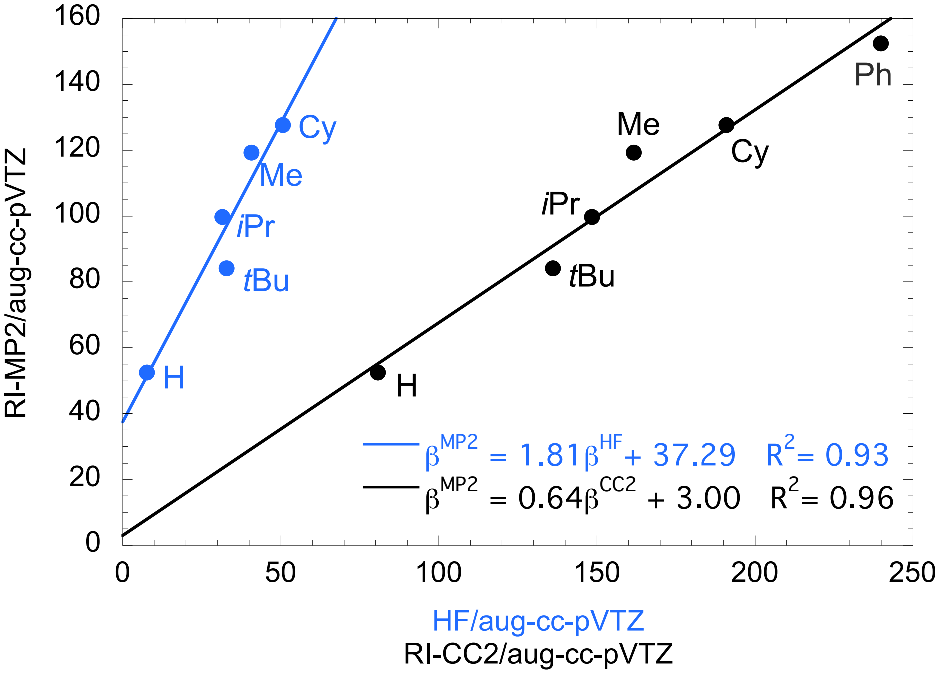


Figure 6.6: Correlation between static hyperpolarizabilities (β, a.u.) of the series of azobenzenes calculated at the MP2 level with respect to the values calculated using CC2 and HF. The β value calculated for compound **e** (R = Ph) is excluded from the fit in HF calculations.

percentage error (MAPE), and root mean square deviation (RMSD) are CAM-B3LYP, M06-2X, and ωB97X-D (unlike other dispersion-corrected DFAs, ωB97-X-D employs a different parameterization than the native functional, ωB97X). Hence, the latter functionals, which outperform RI-CC2, provide the closest agreement to the RI-MP2 hyperpolarizability values. However, these results would hide the fact that if we are only concerned about the relative magnitude of hyperpolarizability among the compounds, many other methods work just fine. The latter is even more relevant if we consider the difficulty in reproducing absolute values of nonlinear optical properties and the fact that we are using RI-MP2 values for comparison, which are approximate too. Table 6.3 provides the results for the linear regression of RI-MP2 vs. other methods using all the molecules except for compounds **e** (R = Ph) and **g** (R = Ad); we will use the former compound's results to measure the ability of the methods to provide a correct trend of the hyperpolarizability for this series of compounds. The Pearson coefficient, R^2, shows that 9 out of 20 methods (including RI-CC2, CAM-B3LYP, and M06-2X) give excellent results (R$^2 \geq 0.95$) for the five smallest compounds of the series. The difference between the value predicted by the linear regression and the MP2 hyperpolarizability, $\Delta\beta^{Ph} = \hat{\beta}^{Ph,\text{meth}} - \beta^{Ph,\text{MP2}}$, provides an estimate of the predictability of the linear regression for molecule **e** including phenyl *meso*-substituents. Despite the excellent correlations found for the first five compounds, only six methods give an estimate of $\beta^{Ph,\text{MP2}}$ with an error below 20 a.u. (RI-CC2, PBE, rPW86-PBE-VV, M06, M06L, and M06-2X). For many DFAs and RI-CC2, the value of the intercept (b) is rather small (below 20 a.u.) and the Pearson coefficient is rather high. In these cases, the prediction of β is dominated by the slope, a, which determines whether the quantity is overestimated ($a < 1$) or underestimated ($a > 1$) with respect to the RI-MP2 values. Our conjecture is that, in the latter DFAs, the error on the hyperpolarizability is dominated by the delocalization error [66]. In Figure 6.7, we represent the % of HF exchange at long range (large r_{12}) that is included in the XCF against the value of a. The plot reflects clearly how DFAs with a small percentage of HF exchange tend to overestimate the hyperpolarizability with respect to RI-MP2, whereas the opposite occurs for DFAs with a large percentage of HF exchange (see also Tables C.16-C.17). If we consider all the measures included in Table 6.3, the best performing DFA is M06-2X. We also obtain quite good results using CAM-B3LYP, ωB97-X-D, and Tα-LC-BLYP. All these functionals include at least 54% of HF exchange at long range. Hence, a percentage of nearly 50% between DFT and HF exchange is the adequate balance for computing the NLO properties of these systems, as it was also concluded from previous works on smaller molecules [67] and conjugated push-pull chromophores [68, 69]. These results thus contribute to the recent findings about the importance of the delocalization error in the calculation of linear and nonlinear optical properties [44, 68, 70–72] as well as other properties such as aromaticity [73–76] and conjugation [77]. It is worth noticing that the recent strategy [44] of finding the optimal range-separation parameter to calculate γ from the values of α works quite well to calculate the value of β for these systems. Indeed, Tα-LC-BLYP and LC-BLYP give a similar Pearson coefficient, whereas for the former a and b are closer to 1 and 0, respectively. As a result, Tα-LC-BLYP significantly improves MAE, MAPE, and RMSD.

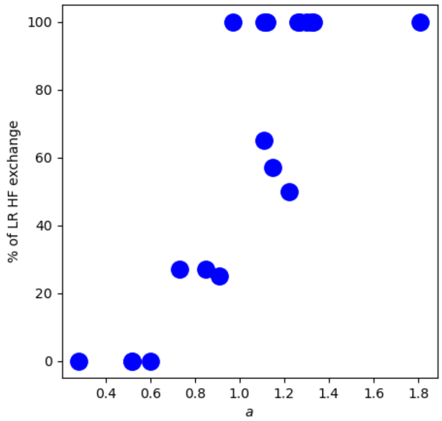


Figure 6.7: Results of the fit of the a parameter correlated with the % of Long Range HF exchange for each functional included in the set.

Regarding the impact of dispersion corrections, β values computed using rPW86PBE-VV show an improvement compared to those computed with its uncorrected parent. Although some error measurements like RSMD, MAE, or MAPE are only marginally improved upon including of VV dispersion corrections, $\Delta\beta^{Ph}$, a and the Pearson coefficient are much better for rPW86PBE-VV. Conversely, LC-ωPBE08-VV and LC-ωPBE08 show

Table 6.3: Results of the linear correlation between the MP2 hyperpolarizabilities and several other methods, together with several error measures (see text). The linear regression function is: $\hat{\beta}^{\mathrm{MP2}} = a\beta^{\mathrm{meth}} + b$, where β^{meth} is value of the hyperpolarizability for one the methods listed in the table. $\Delta\beta^{Ph} = \hat{\beta}^{Ph,\mathrm{meth}} - \beta^{Ph,\mathrm{MP2}}$.

Method	a	b [a.u.]	$\Delta\hat{\beta}^{Ph}$ [a.u.]	R^2	RMSD [a.u.]	MAPE [%]	MAE [a.u.]
RI-CC2	0.72	-6.98	-13.72	0.96	56.75	34	53.68
HF	1.81	37.29	76.48	0.93	80.18	318	75.15
PBE	0.52	4.57	10.23	0.84	90.69	45	86.87
rPW86-PBE	0.28	38.85	41.73	0.68	119.55	49	107.03
rPW86-PBE-VV	0.52	-0.87	-0.77	0.88	102.73	48	98.33
PBE0	0.91	-14.05	22.00	0.97	23.18	18	21.55
B3LYP	0.85	-18.98	22.37	0.96	37.21	27	36.14
BH&H	1.22	4.26	25.71	0.96	29.44	32	26.27
CAM-B3LYP	1.11	-7.42	49.61	0.96	22.22	13	13.24
ωB97-X	1.12	3.78	54.46	0.91	31.32	27	22.52
ωB97X-D	0.97	-5.19	53.57	0.88	21.62	16	16.19
LC-BLYP	1.32	20.43	60.79	0.93	54.40	94	48.89
Tα-LC-BLYP	1.11	0.72	57.32	0.94	29.75	24	20.02
LC-ωPBE	1.26	10.86	57.70	0.95	44.17	59	37.93
LC-ωPBE08	1.33	15.59	56.13	0.95	50.45	79	45.18
LC-ωPBE08-VV	1.30	17.54	48.07	0.95	48.37	75	43.99
M06	0.73	-16.27	17.24	0.86	59.20	36	57.85
M06L	0.59	-25.23	14.96	0.84	115.56	53	113.92
M06-2X	1.15	-0.06	19.53	0.97	19.45	18	16.36
M06-HF	1.27	47.23	59.12	0.86	71.55	693	67.37

Table 6.4: Vector (β_i) and tensor (β_{ijk}) components of the total first hyperpolarizability, as calculated at the CAM-B3LYP/aug-cc-pVTZ level. The Cartesian frame used for the calculations is shown in Figure 6.8. Units are a.u.

Molecule	β	β_x	β_y	β_z	β_{zxx}	β_{zyy}	β_{zzz}
a (R = H)	52	0	-17	-261	-135	21	27
b (R = Me)	115	0	-30	-576	-159	-14	-19
c (R = *i*Pr)	93	-4	-7	-463	-149	16	-22
d (R = *t*Bu)	91	-82	50	-447	-167	43	-25
e (R = Ph)	100	-1	22	-498	-122	-62	19
f (R = Cy)	119	60	28	-592	-144	-33	-20
g (R = Ad)	148	-9	-201	-714	-170	2	-69

a similar performance. Finally, going from ωB97X to ωB97X-D significantly improves the RSMD, MAE, and MAPE, while only slightly modifies the linear regression plot; *i.e.*, both functionals give similar trends among the series of compounds but ωB97X-D provide better estimates of the MP2 hyperpolarizabilities. It is worth stressing that the differences in the β values computed using the two latter functionals cannot be only ascribed to dispersion corrections since the whole set of their parameters (including the range separation, ω) is different.

6.3.2.3 Partition of the hyperpolarizability into orbital contributions

To gain a deeper insight into the contribution of the *meta*-substituents to the total first hyperpolarizability, we now examine in more detail the vector (β_i) and tensor (β_{ijk}) components of the total first hyperpolarizability (Eqs. 6.1 and 6.2) of the investigated compounds. A common Cartesian frame, in which the origin is placed at the center of the N=N bond, the x axis is oriented along this bond, and the xz plane contains the two nitrogen atoms and one of the adjacent carbon atom (see Figure 6.8) is used for all the molecules. CAM-B3LYP/aug-cc-pVTZ hyperpolarizabilities are considered in the analyses since, as discussed in the previous section, they are in close agreement with the absolute MP2 reference values. As shown in Table 6.4, the β_z component is one order of magnitude larger than β_x and β_y in all compounds, except for compound **g** (R = Ad) for which β_z/β_y = 3.55, the larger contribution from the y being due to bulkiness of the adamantyl substituent. In turn, decomposition of β_z in terms of tensor elements evidence that the dominant contribution to the first hyperpolarizability arises from the transverse β_{zxx} component. Note that, as shown in the Annex (Table C.19), a very similar picture of the relative magnitude of the β components can be drawn from CAM-B3LYP calculations carried out using the smaller 6-311++G** basis set.

To rationalize the role of molecular substituents in the value of the hyperpolarizability, we use the PNOC scheme developed by Sitkiewicz and coworkers [32] to decompose the NLO responses into contributions of the molecular orbitals (MOs) of the unperturbed system:

$$\beta_{ijk} = \sum_p \beta_{ijk,p} = -\sum_p \Delta_{pp}^{(jk)} M_{pp}^{(i)} - \sum_{p \neq q} \Delta_{pq}^{(jk)} M_{pq}^{(i)}, \tag{6.3}$$

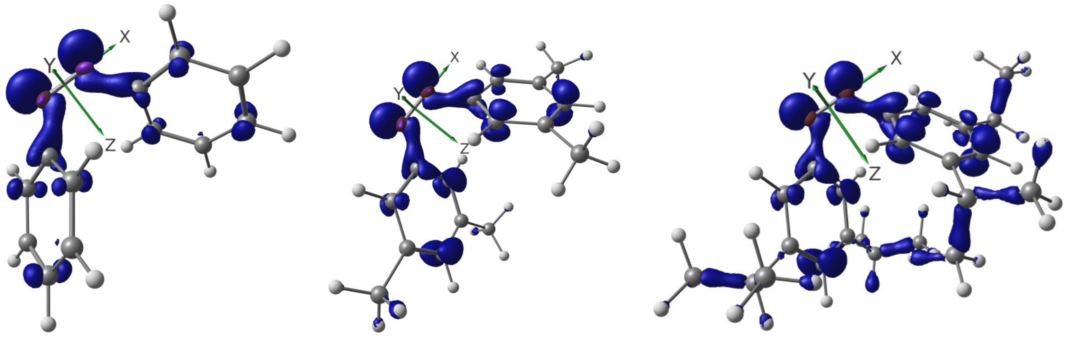
where $\mathbf{M}^i$ is the transition dipole matrix along the i direction, and $\Delta^{(jk)}$ is the second-order derivative of the one-particle reduced density matrix with respect to external fields oriented along j and k axes. This expression is an exact decomposition of β based on the assumption that the terms that depend on two different MOs can be equally distributed between them. In this framework, for an orthonormal set of orbitals, a real-space representation of the largest β_{zxx} component can be obtained according to:

$$\beta_{zxx}(\mathbf{r}) = \sum_p \beta_{zxx,p} |\phi_p(\mathbf{r})|^2, \tag{6.4}$$

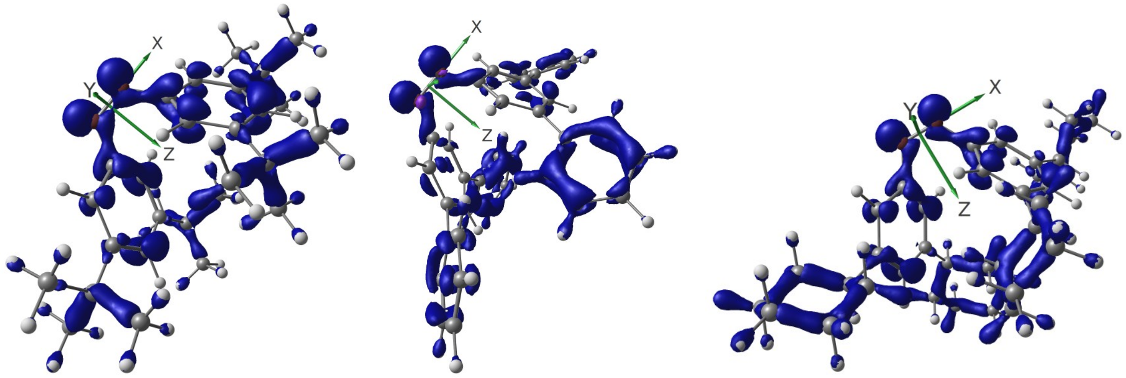
where $|\phi_p(\mathbf{r})|^2$ is the square of the amplitude of molecular orbital p. The PNOC representations of $\beta_{zxx}(\mathbf{r})$ in the Cartesian space are displayed in Figure 6.8 for the azobenzene series at the CAM-B3LYP/6-311++G** level. The main contribution to the first hyperpolarizability arises from the N=N bond for all the molecules, although for bulkier substituents there are also important contributions. Some contributions are due to *ortho* and *para* carbon atoms of the adjacent phenyl rings, although, in general, the adjacent phenyl ring shows rather small contributions to β. Interestingly, the phenyl substituents in compound **e** (R = Ph) show an important contribution to β from π orbitals. In compounds **c**, **d**, **f**, and **g**, $\beta_{zxx}(\mathbf{r})$ also shows contributions due to C-C σ-bonds of the R alkyl groups, although larger contributions from the substituents are obtained for bulkier substituents such as cyclohexyl and adamantyl.

Individual orbital contribution analysis reveals that the hyperpolarizibility of these compounds cannot be qualitatively explained from the contribution of a few molecular orbitals. Many MOs must be included in the summation of the Eq. 6.3 to obtain a good estimate of the total β_{zxx} component (see Figure C.12), hampering a simple deconvolution of the total NLO response in terms of individual MOs. However, for all compounds except **e** (R = Ph), the highest occupied MO, which is mainly associated with the p orbitals of the nitrogen atoms (the π orbitals and the lone pairs) and to the C-N σ bonds, provides the largest contribution to β_{zxx} (Table C.20).

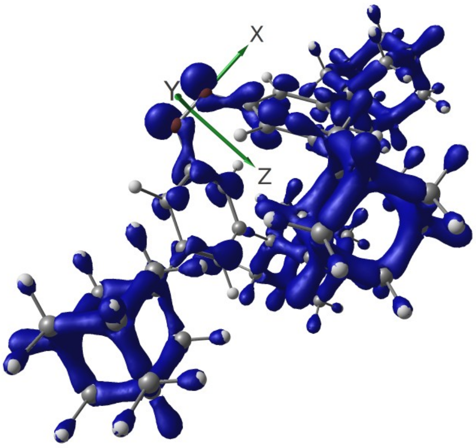
Finally, we have used the PNOC as a tool to analyze the delocalization errors committed by some DFAs. As a measure of the extent of electron delocalization, we have computed the Laplacian of the electron density [78], $\nabla^2\rho(\mathbf{r})$, which attains large negative values in delocalized regions. Table 6.5 gives PNOC and $\nabla^2\rho(\mathbf{r})$ values for **e** (R = Ph) and **c** (R = iPr) at the PBE and CAM-B3LYP levels of theory. Since PBE suffers from larger delocalization errors than CAM-B3LYP, it is expected that both the PNOC and -$\nabla^2\rho(\mathbf{r})$ display regions with larger values for PBE. In order to clearly visualize these differences, in Table 6.5, we have also plotted the difference between these two DFAs. For both molecules, the main difference in the distribution of the hyperpolarizability across



(a) R = H (b) R = Me (c) R = iPr



(d) R = tBu (e) R = Ph (f) R = Cy

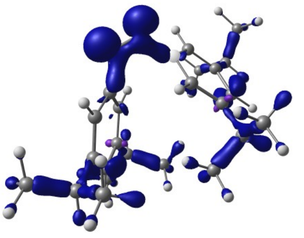
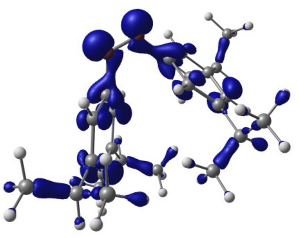
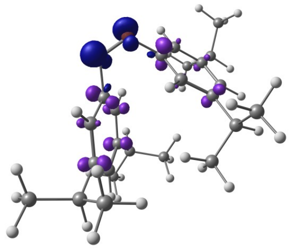
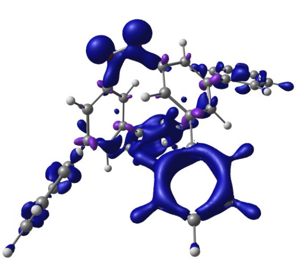
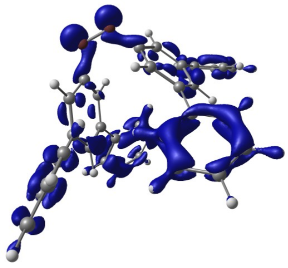
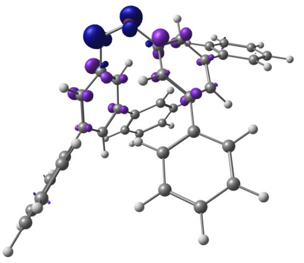
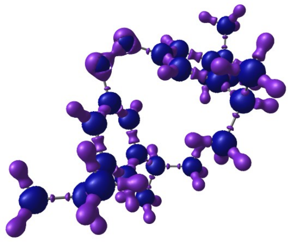
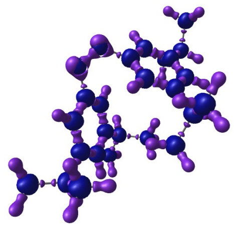
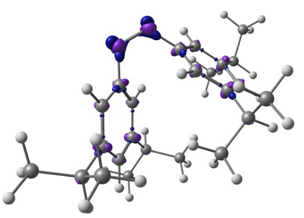
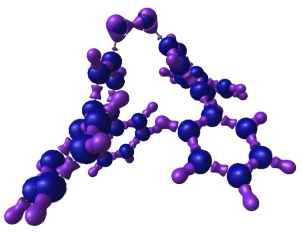
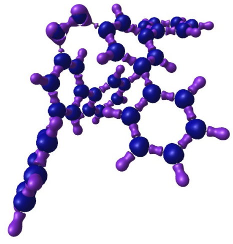
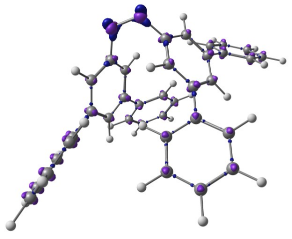


(g) R = Ad

Figure 6.8: Isosurfaces of $\beta_{zxx}(\mathbf{r})$ for the azobenzene molecules obtained using the PNOC partition at the CAM-B3LYP/6-311++G** level. Isocontour values of ±5 a.u. were used for all compounds.

the molecule comes from the lone pair of the nitrogen atoms. $\Delta\nabla^2\rho(\mathbf{r})$ values are also larger around the nitrogen atoms. Hence, one is deemed to conclude that the delocalization error is the main source of error in the calculation of the hyperpolarizability and it comes mostly from a wrong description of the electron density in the vicinity of the N=N bond (in particular, the p orbitals involved in the π bond and the lone pairs).

Table 6.5: Comparison of $\beta_{zxx}(\mathbf{r})$ calculated through PNOC for the functionals PBE and CAM-B3LYP evaluated for the molecules **c** (R = *i*Pr) and **e** (R = Ph). The isosurface $\Delta\beta_{zxx}(\mathbf{r})$ is calculated as: $\Delta\beta_{zxx}(\mathbf{r}) = \beta_{zxx}^{(\mathrm{PBE})} - \beta_{zxx}^{(\mathrm{CAM\text{-}B3LYP})}$. Isocontour values of ±5 a.u. were used for all the β_{zxx} . For the Laplacian of the electronic density isocontour ±0.7 was used for the Laplacians, while ±0.09 was used for the difference between the laplacian ($\Delta\nabla^2$) of the PBE and the CAM-B3LYP densities. Positive numbers are displayed in blue, negative ones in purple.

Molecule	$\beta_{zxx}(\mathbf{r})^{(PBE)}$	$\beta_{zxx}(\mathbf{r})^{(CAM-B3LYP)}$	$\Delta\beta_{zxx}(\mathbf{r})$
R = *i*Pr			
R = Ph			
Molecule	$\nabla^2\rho(\mathbf{r})^{(PBE)}$	$\nabla^2\rho(\mathbf{r})^{(CAM-B3LYP)}$	$\Delta\nabla^2\rho(\mathbf{r})$
R = Ph			
R = Ph			

6.4 Conclusions

The reliability of a large selection of exchange-correlation density functional approximations to predict the geometrical structure, relative Z-E energies (ΔE_{ZE}), and second-order nonlinear responses of a series of azobenzene molecules has been assessed with respect to reference correlated *ab initio* calculations carried out at the RI-MP2 and RI-CC2 levels. Our calculations show that RI-MP2 and RI-CC2 approximations provide very similar performance for intramolecular noncovalent interactions, based on the results of Diet-GMTKN55 and the series of azobenzene molecules studied in this paper. Comparisons with Hartree-Fock results show that electron correlation and, in particular, dispersion interactions are the driving forces behind the stabilization of the *cis* conformer with respect to the *trans* one. Indeed, the *cis* isomer stabilizes upon functionalization of the *meta*-position with bulky substituents that develop mutual attractive London interactions.

Among the selected dispersion-corrected exchange-correlation functionals, ωB97-X-D provides the closest agreement with reference RI-CC2 geometries and ΔE_{ZE} values. CAM-B3LYP-D3 and LC-ωPBE-D3 also provide reasonably good geometries, the latter giving excellent ΔE_{ZE} values too.

This study shows an excellent linear correlation between static first hyperpolarizabilities computed at the RI-MP2 and RI-CC2 levels using RI-CC2 geometries. RI-MP2 provides β values systematically underestimated compared to RI-CC2. Despite density functional approximations do not give accurate first hyperpolarizabilities, many of the selected functionals qualitatively reproduce the evolution of β along the series of azobenzene derivatives. The most accurate functionals are CAM-B3LYP, ωB97-X-D, and Tα-LC-BLYP. In general, the best performing functionals include at least 50% of the exact HF exchange. ωB97X-D seems to be the best suited functional to describe all the relevant features of these compounds: the geometry, the energy difference between the isomers, and the first hyperpolarizabilities.

Subsequently, we analyzed the property-structure relationship for the first hyperpolarizability of azobenzene derivatives to identify the effect of the *meta*-substituents on the NLO responses. For all molecules, the dominant contribution to β comes from the β_{zxx} tensor component. The PNOC decomposition analysis [32] was employed to analyze the molecular orbitals and the real-space contributions to the optical response. A simple inspection of the orbital contributions shows that it is impossible to obtain a simple description of the total first hyperpolarizability in terms of a few orbital contributions. However, the PNOC real-space analysis reveals that the main contribution to β comes mostly from N=N, followed by smaller contributions from the adjacent phenyl rings. Interestingly, phenyl substituents attached in *meso* of the adjacent phenyl rings also show important contributions from the π system to the first hyperpolarizability.

Finally, we studied the relationship between the real-space PNOC analysis and the delocalization error. We found a qualitative correlation between the electron delocalization differences (measured through the Laplacian of the electron density) and the real-space

PNOC representations. Hence, it seems that the regions of the molecule most affected by the delocalization error are, in turn, the ones responsible for the inaccuracy of the first hyperpolarizability.

Bibliography

[1] Dalton, L. R.; Sullivan, P. A.; Bale, D. H. Electric Field Poled Organic Electro-optic Materials: State of the Art and Future Prospects. *Chem. Rev.* **2010**, *110*, 25–55.

[2] Stegeman, G. I.; Stegeman, R. A. In *Nonlinear Optics: Phenomena, Materials and Devices*; Boreman, G., Ed.; John Wiley and Sons: Hoboken, NJ, 2012.

[3] Campagnola, P. J.; Wei, M. D.; Loew, L. M. High-Resolution Nonlinear Optical Imaging of Live Cells by Second Harmonic Generation. *Biophys. J.* **1999**, *77*, 3341–3349.

[4] Moreaux, L.; Sandre, O.; Charpak, S.; Blanchard-Desce, M.; Mertz, J. Coherent Scattering in Multi-Harmonic Light Microscopy. *Biophys. J.* **2001**, *80*, 1568–1574.

[5] Reeve, J. E.; Anderson, H. L.; Clays, K. Dyes for biological second harmonic generation imaging. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13484–13498.

[6] Coe, B. J. Molecular Materials Possessing Switchable Quadratic Nonlinear Optical Properties. *Chem. Eur. J.* **1999**, *5*, 2464–2471.

[7] Delaire, J. A.; Nakatani, K. Linear and Nonlinear Optical Properties of Photochromic Molecules and Materials. *Chem. Rev.* **2000**, *100*, 1817–1846.

[8] Castet, F.; Rodriguez, V.; Pozzo, J.-L.; Ducasse, L.; Plaquet, A.; Champagne, B. Design and Characterization of Molecular Nonlinear Optical Switches. *Acc. Chem. Res.* **2013**, *46*, 2656–2665.

[9] Ledoux, I.; Zyss, J.; Barni, E.; Barolo, C.; Diulgheroff, N.; Quagliotto, P.; Viscardi, G. Properties of novel azodyes containing powerful acceptor groups and thiophene moiety. *Synth. Met.* **2000**, *115*, 213 – 217.

[10] Yesodha, S. K.; Sadashiva Pillai, C. K.; Tsutsumi, N. Stable polymeric materials for nonlinear optics: a review based on azobenzene systems. *Prog. in Polym. Sci.* **2004**, *29*, 45 – 74.

[11] Krawczyk, P.; Kaczmarek, A.; Zaleśny, R.; Matczyszyn, K.; Bartkowiak, W.; Ziółkowski, M.; Cysewski, P. Linear and nonlinear optical properties of azobenzene derivatives. *J. Mol. Model.* **2009**, *15*, 581–590.

[12] Pérez-Moreno, J.; Zhao, Y.; Clays, K.; Kuzyk, M. G.; Shen, Y.; Qiu, L.; Hao, J.; Guo, K. Modulated Conjugation as a Means of Improving the Intrinsic Hyperpolarizability. *J. Am. Chem. Soc.* **2009**, *131*, 5084–5093.

[13] Kleinpeter, E.; Bölke, U.; Kreicberga, J. Quantification of the push-pull character of azo dyes and a basis for their evaluation as potential nonlinear optical materials. *Tetrahedron* **2010**, *66*, 4503–4509.

[14] Cardoso, C.; Abreu, P. E.; Milne, B. F.; Nogueira, F. Computational Study of Molecules with High Intrinsic Hyperpolarizabilities. *J. Phys. Chem. A* **2010**, *114*, 10676–10683.

[15] Jaunet-Lahary, T.; Chantzis, A.; Chen, K. J.; Laurent, A. D.; Jacquemin, D. Designing Efficient Azobenzene and Azothiophene Nonlinear Optical Photochromes. *J. Phys. Chem. C* **2014**, *118*, 28831–28841.

[16] Ghanavatkar, C. W.; Mishra, V. R.; Sekar, N. Review of NLOphoric azo dyes–Developments in hyperpolarizabilities in last two decades. *Dyes Pigm.* **2021**, *191*, 109367.

[17] Zaleśny, R.; Bulik, I. W.; Bartkowiak, W.; Luis, J. M.; Avramopoulos, A.; Papadopoulos, M. G.; Krawczyk, P. Electronic and vibrational contributions to first hyperpolarizability of donor–acceptor-substituted azobenzene. *J. Chem. Phys.* **2010**, *133*, 244308.

[18] Loucif-Saibi, R.; Nakatani, K.; Delaire, J. A.; Dumont, M.; Sekkat, Z. Photoisomerization and second harmonic generation in disperse red one-doped and -functionalized poly(methyl methacrylate) films. *Chem. Mater.* **1993**, *5*, 229–236.

[19] Sekkat, Z.; Prêtre, P.; Knoesen, A.; Volksen, W.; Lee, V. Y.; Miller, R. D.; Wood, J.; Knoll, W. Correlation between polymer architecture and sub-glass-transition-temperature light-induced molecular movement in azo-polyimide polymers: influence on linear and second- and third-order nonlinear optical processes. *J. Opt. Soc. Am. B* **1998**, *15*, 401–413.

[20] Vijay Srinivasan, M.; Kannan, P. Photo-switching and nonlinear optical behaviors of center linked bent-core azobenzene liquid crystalline polymers. *J. Mat. Sci.* **2011**, *46*, 5029–5043.

[21] Priimagi, A.; Ogawa, K.; Virkki, M.; Mamiya, J.-i.; Kauranen, M.; Shishido, A. High-Contrast Photoswitching of Nonlinear Optical Response in Crosslinked Ferroelectric Liquid-Crystalline Polymers. *Adv. Mat.* **2012**, *24*, 6410–6415.

[22] Yamada, K.; Otsubo, H.; Yonemura, H.; Yamada, S.; Matsuo, T. Nonlinear Optical Responses of Dialkoxyazobenzene Isomers Adsorbed to Interface between Heptane and Viologen-Modified Quartz Plates. *Chem. Lett.* **1997**, *26*, 451–452.

[23] Schulze, M.; Utecht, M.; Moldt, T.; Przyrembel, D.; Gahl, C.; Weinelt, M.; Saalfrank, P.; Tegeder, P. Nonlinear optical response of photochromic azobenzene-functionalized self-assembled monolayers. *Phys. Chem. Chem. Phys.* **2015**, *17*, 18079–18086.

[24] Tonnelé, C.; Champagne, B.; Muccioli, L.; Castet, F. Nonlinear Optical Contrast in Azobenzene-Based Self-Assembled Monolayers. *Chem. Mater.* **2019**, *31*, 6759–6769.

[25] Riaz, S.; Friedrichs, G. Vibrational sum-frequency generation study of molecular structure, sterical constraints and nonlinear optical switching contrast of mixed alkyl-azobenzene self-assembled monolayers. *Zeitschrift für Physikalische Chemie* **2020**, *234*, 1427 – 1452.

[26] Osella, S.; Knippenberg, S. Triggering On/Off States of Photoswitchable Probes in Biological Environments. *J. Am. Chem. Soc.* **2017**, *139*, 4418–4428.

[27] Knippenberg, S.; Osella, S. Push/Pull Effect as Driving Force for Different Optical Responses of Azobenzene in a Biological Environment. *J. Phys. Chem. C* **2020**, *124*, 8310–8322.

[28] Schweighauser, L.; Strauss, M. A.; Bellotto, S.; Wegner, H. A. Attraction or Repulsion? London Dispersion Forces Control Azobenzene Switches. *Angew. Chem. Int. Ed.* **2015**, *54*, 13436–13439.

[29] Aradhya, S. V.; Frei, M.; Hybertsen, M. S.; Venkataraman, L. Van der Waals interactions at metal/organic interfaces at the single-molecule level. *Nature Mat.* **2012**, *11*, 872–876.

[30] Wagner, J. P.; Schreiner, P. R. London Dispersion in Molecular Chemistry—Reconsidering Steric Effects. *Angew. Chem. Int. Ed.* **2015**, *54*, 12274–12296.

[31] Fabrizio, A.; Corminboeuf, C. How do London Dispersion Interactions Impact the Photochemical Processes of Molecular Switches? *J. Phys. Chem. Lett.* **2018**, *9*, 464–470.

[32] Sitkiewicz, S. P.; Rodríguez-Mayorga, M.; Luis, J. M.; Matito, E. Partition of optical properties into orbital contributions. *Phys. Chem. Chem. Phys.* **2019**, *21*, 15380–15391.

[33] Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

[34] Hättig, C.; Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **2000**, *113*, 5154–5161.

[35] Cybulski, S. M.; Lytle, M. L. The origin of deficiency of the supermolecule second-order Møller-Plesset approach for evaluating interaction energies. *J. Chem. Phys.* **2007**, *127*, 141102.

[36] Gould, T. 'Diet GMTKN55' offers accelerated benchmarking through a representative subset approach. *Phys. Chem. Chem. Phys.* **2018**, *20*, 27735–27739.

[37] Dunning Jr., T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[38] Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core–valence and quintuple-ζ basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

[39] Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[40] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

[41] Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.

[42] Grimme, S. Density functional theory with London dispersion corrections. *WIREs Computational Molecular Science* **2011**, *1*, 211–228.

[43] Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.

[44] Besalú-Sala, P.; Sitkiewicz, S. P.; Salvador, P.; Matito, E.; Luis, J. M. A new tuned range-separated density functional for the accurate calculation of second hyperpolarizabilities. *Phys. Chem. Chem. Phys.* **2020**, *22*, 11871–11880.

[45] Lee, T. J. Comparison of the T1 and D1 diagnostics for electronic structure theory: a new definition for the open-shell D1 diagnostic. *Chem. Phys. Lett.* **2003**, *372*, 362–367.

[46] Ramos-Cordoba, E.; Salvador, P.; Matito, E. Separation of dynamic and nondynamic correlation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24015–24023.

[47] Ramos-Cordoba, E.; Matito, E. Local Descriptors of dynamic and nondynamic correlation. *J. Chem. Theory Comput.* **2017**, *13*, 2705–2711.

[48] Via-Nadal, M.; Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Matito, E. Singling out Weak and Strong Correlation. *J. Phys. Chem. Lett.* **2019**, *10*, 4032–4037.

[49] Via-Nadal, M.; Rodríguez-Mayorga, M.; Ramos-Cordoba, E.; Matito, E. Natural range separation of the Coulomb hole. *J. Chem. Phys.* **2022**, *156*, 184106.

[50] Löwdin, P.-O. Quantum theory of many-particle systems. I. Physical interpretations by means of density matrices, natural spin-orbitals, and convergence problems in the method of configurational interaction. *Phys. Rev.* **1955**, *97*, 1474–1489.

[51] Smith Jr, V. H. Approximate natural orbitals for carbon 1 S. *Theor. Chim. Acta (Berlin)* **1967**, *7*, 245.

[52] Sitkiewicz, S. P.; Ramos-Cordoba, E.; Luis, J. M.; Matito, E. How Many Electrons Does a Molecular Electride Hold? *J. Phys. Chem. A* **2021**, *125*, 4819–4835.

[53] Frisch, M. J. et al. Gaussian˜16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

[54] de Wergifosse, M.; Liégeois, V.; Champagne, B. Evaluation of the molecular static and dynamic first hyperpolarizabilities. *Int. J. Quant. Chem.* **2014**, *114*, 900–910.

[55] Champagne, B.; Beaujean, P.; de Wergifosse, M.; Cardenuto, M. H.; Liégeois, V.; Castet, F. In *Frontiers of Quantum Chemistry*; Wojcik, M., Nakatsuji, H., Kirtman, B., Ozaki, Y., Eds.; Springer, Singapore, 2018; Chapter Quantum Chemical Methods for Predicting and Interpreting Second-Order Nonlinear Optical Properties: From Small to Extended π-Conjugated Molecules, pp 117–138.

[56] Balasubramani, S. G. et al. TURBOMOLE: Modular program suite for ab initio quantum-chemical and condensed-matter simulations. *J. Chem. Phys.* **2020**, *152*, 184107.

[57] Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

[58] Shao, Y. et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184–215.

[59] Charnley, J.; Bratholm, L. Calculate Root-mean-square deviation (RMSD) of Two Molecules Using Rotation. GitHub, v.1.3.2.

[60] Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crys. A* **1976**, *32*, 922–923.

[61] Walker, M. W.; Shao, L.; Volz, R. A. Estimating 3-D location parameters using dual number quaternions. *CVGIP: image understanding* **1991**, *54*, 358–367.

[62] Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing noncovalent interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.

[63] Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. NCIPLOT: a program for plotting noncovalent interaction regions. *J. Chem. Theory Comput.* **2011**, *7*, 625–632.

[64] Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[65] Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

[66] Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into Current Limitations of Density Functional Theory. *Science* **2008**, *321*, 792–794.

[67] Castet, F.; Champagne, B. Assessment of DFT Exchange-Correlation Functionals for Evaluating the Multipolar Contributions to the Quadratic Nonlinear Optical Responses of Small Reference Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 2044–2052.

[68] Johnson, L. E.; Dalton, L. R.; Robinson, B. H. Optimizing Calculations of Electronic Excitations and Relative Hyperpolarizabilities of Electrooptic Chromophores. *Acc. Chem. Res.* **2014**, *47*, 3258–3265.

[69] Lescos, L.; Sitkiewicz, S.; Beaujean, P.; Blanchard-Desce, M.; Champagne, B. R.; Matito, E.; Castet, F. Performance of DFT Functionals for Calculating the Second-Order Nonlinear Optical Properties of Dipolar Merocyanines. *Phys. Chem. Chem. Phys.* **2020**, 16579–16594.

[70] Champagne, B.; Perpète, E. A.; Jacquemin, D.; van Gisbergen, S. J. A.; Baerends, E.-J.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. Assessment of Conventional Density Functional Schemes for Computing the Dipole Moment and (Hyper)polarizabilities of Push−Pull π-Conjugated Systems. *J. Phys. Chem. A* **2000**, *104*, 4755–4763.

[71] de Wergifosse, M.; Champagne, B. Electron correlation effects on the first hyperpolarizability of push–pull π-conjugated systems. *J. Chem. Phys.* **2011**, *134*, 074113.

[72] Zaleśny, R.; Medved', M.; Sitkiewicz, S. P.; Matito, E.; Luis, J. M. Can Density Functional Theory Be Trusted for High-Order Electric Properties? The Case of Hydrogen-Bonded Complexes. *J. Chem. Theory Comput.* **2019**, *15*, 3570–3579.

[73] Casademont-Reig, I.; Woller, T.; Contreras-García, J.; Alonso, M.; Torrent-Sucarrat, M.; Matito, E. New electron delocalization tools to describe the aromaticity in porphyrinoids. *Phys. Chem. Chem. Phys.* **2018**, *20*, 2787–2796.

[74] Casademont-Reig, I.; Ramos-Cordoba, E.; Torrent-Sucarrat, M.; Matito, E. How do the Hückel and Baird Rules Fade away in Annulenes? *Molecules* **2020**, *25*, 711.

[75] Casademont-Reig, I.; Guerrero-Avilés, R.; Ramos-Cordoba, E.; Torrent-Sucarrat, M.; Matito, E. How Aromatic Are Molecular Nanorings? The Case of a Six-Porphyrin Nanoring**. *Angew. Chem. Int. Ed.* **2021**, *60*, 24080–24088.

[76] Casademont-Reig, I.; Ramos-Cordoba, E.; Torrent-Sucarrat, M.; Matito, E. In *Aromaticity: Modern Computational Methods and Applications*; Fernández, I., Ed.; Elsevier: The Netherlands, 2021; Chapter Aromaticity descriptors based on electron delocalization, pp 235–258.

[77] Sancho-García, J.; Pérez-Jiménez, A. Improved accuracy with medium cost computational methods for the evaluation of bond length alternation of increasingly long oligoacetylenes. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5874–5879.

[78] Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, 1990.

Chapter 7

Electric-field induced second harmonic generation responses of push–pull polyenic dyes: Experimental and theoretical characterizations

This chapter has been realized in the framework of a collaboration involving: Vangheluwe R. (Univ. of Bordeaux), Ledoux-Rak I. (Univ. of Paris-Saclay), Champagne B. (Univ. of Namur), Tonnelé C. (Donostia International Physics Center), and Blanchard-Desce M. (Univ. of Bordeaux)

7.1 Introduction

The design of organic dyes delivering high second-order nonlinear optical (NLO) properties is an important issue in many (bio)technological fields for probing asymmetric media such as artificial interfaces or cell membranes [1]. In particular, the exogenous labeling of lipid bilayers by amphiphilic potential-sensitive dyes displaying large second harmonic generation (SHG) responses is at the heart of high resolution imaging microscopy techniques [2–5]. SHG probes are usually designed by functionalizing the two extremities of a π-conjugated linker by electron-donating and electron-withdrawing substituents, which provide the asymmetry required for quadratic NLO phenomena. The elaboration of SHG chromophores based on this dipolar architecture [6–9], including responsive systems such as NLO switches [10–12], has been the object of intense research in the last 30 years. In this context, the synthesis and characterization of the SHG responses of amphiphilic chromophores incorporating a pyridinium acceptor and a dibutylaminophenyl donor as terminal groups, connected through increasingly large polyenic bridges (series **A**, Figure 7.1) it has been reported [13]. The presence of the hydrophobic butyl chains on the donor group, together with the positively-charged hydrophilic acceptor moiety, confers an amphiphilic character to these push–pull dyes, which facilitates their interaction with a lipidic membrane. In addition to their use as SHG probes, pyridinium-based derivatives have been widely used in the last decades for producing materials with large quadratic NLO responses, owing to their structural diversity and ability to form different non-centrosymmetric crystal packings when associated with different anions [14].

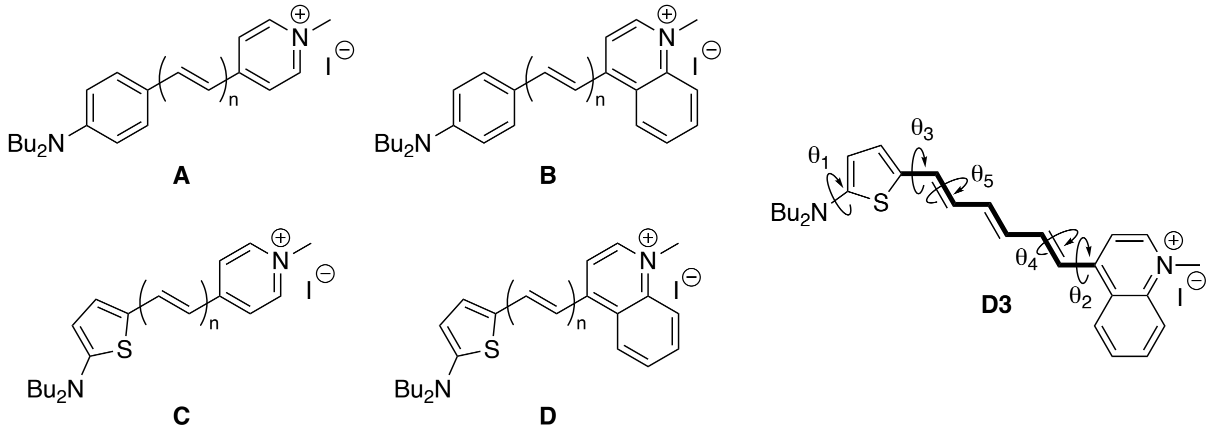


Figure 7.1: Structure of the push–pull polyenic dyes **A**-**D** investigated in this study ($n = 1 - 5$). On the right: structure of **D3** with the conjugated segment (in bold) used to calculate the bond length alternation and the torsional angles θ_i discussed in the text. See the Annex for a more detailed definition of the geometrical parameters.

In continuation to the efforts in designing new SHG chromophores, we report in this contribution three new families of derivatives (Figure 7.1), for which enhanced optical nonlinearities are expected compared to the original series. In series **B** and **D** the pyridinium acceptor is replaced by a quinolinium, while **C** and **D** involve a thienyl linker in place of the original phenyl. The SHG responses of these dyes are probed by means of electric-field induced second harmonic generation (EFISHG). The use of the EFISHG technique was made possible for these cationic chromophores by using a solvent of low relative permittivity, which prevents the ion pairs from dissociating into the postively-charged dye and its iodide counterion. The role of such ion pairs interactions in the second-order NLO responses of pyridinium-based salts was investigated in several experimental [15] and theoretical works [16, 17].

In this study, experimental characterizations are complemented by theoretical chemistry calculations, which provide a fundamental understanding on the origin of the NLO responses in these complexes. The computational approach is based on a methodology developed in previous works, combining classical molecular dynamics (MD) simulations and quantum chemical (QM) calculations based on time-dependent density functional theory (TD-DFT). This MD/QM scheme was formerly employed to investigate ion pairs in solution [16–18], but also more complex systems such as organic nanoparticles [19], self-assembled monolayers [20, 21], and stained lipid bilayers [22, 23]. Here, MD simulations carried out on the reference **D3**/iodine complex bring key information on the average position of the iodide anion with respect to the organic chromophore, as well as on the effect of dynamical structural fluctuations on the NLO properties. Then, systematic TD-DFT calculations performed on the whole set of systems provide a rationale to experimental data and allow to establish precise relationships between the structure of the dye and the magnitude of the second- and third-order contributions to the EFISHG intensity.

7.2 Experimental and computational details

7.2.1 EFISHG measurements

In the EFISHG setup [24], the probed SHG response results from a third-order process described by $\gamma(-2\omega;\omega,\omega,0)$. The expression of the NLO quantities connected to the EFISHG response have been presented in Section 1.3.1 and reported here for convenience. γ_{EFISHG} can be decomposed as :

$$\gamma(-2\omega;\omega,\omega,0) = \gamma_{EFISHG} = \gamma_{||}(-2\omega;\omega,\omega,0) + \frac{\mu\beta_{||}(-2\omega;\omega,\omega)}{3kT} \tag{7.1}$$

where μ is the norm of the ground state dipole moment, k is the Boltzmann constant and $3kT = 2.833 \times 10^{-3}$ a.u. at room temperature (298.15 K). The $\gamma_{||}(-2\omega;\omega,\omega,0)$ contribution corresponds to the isotropic invariant of the γ tensor:

$$\gamma_{||}(-2\omega;\omega,\omega,0) = \gamma_{||} = \frac{1}{15}\sum_{i,j}^{x,y,z}(2\gamma_{iijj} + \gamma_{ijji}) \tag{7.2}$$

The second term in equation 7.1 implies the projection of the vectorial representation of the β tensor on the dipole moment:

$$\beta_{||}(-2\omega;\omega,\omega) = \beta_{||} = \frac{3}{5\mu}\vec{\mu}\cdot\vec{\beta} \tag{7.3}$$

where the β vector components read:

$$\beta_i = \frac{1}{3}\sum_{j}^{x,y,z}(\beta_{ijj} + 2\beta_{jij}) \tag{7.4}$$

The relative amplitudes of the second- and third-order contributions to the global EFISHG response can be analyzed using the $R_{3/2}$ ratio:

$$R_{3/2} = 3kT\frac{\gamma_{||}}{\mu\beta_{||}} \tag{7.5}$$

As already anticpated in Section 1.3.1, in principle, the use of the EFISHG technique is precluded in the case of ionic species, because the dc electric field necessary to break the centrosymmetry induces the migration of ions. However, operating in a solvent of low polarity substantiates the formation of electrically neutral dye/iodine ion pairs, which enables the characterization of their NLO responses by means of EFISHG. In this work, all experiments were carried out in chloroform, with dielectric constants $\epsilon_0 = 4.711$ in the static limit and $\epsilon_\infty = 2.091$ at infinite frequency. An incident laser wavelength of 1907 nm was used, in order to minimize the frequency dispersion effects due to absorption of the second harmonic light. Moreover, as evidenced from equation 7.1, measurements should be performed at different temperatures to evaluate the relative amplitude of the second- and third-order contributions in the total EFISHG response. Here, experimental results were analyzed assuming that the $\gamma_{||}$ contribution is negligible with respect to the $\mu\beta_{||}/3kT$ term, so that $R_{3/2} \sim 0$. Therefore, EFISHG responses are analyzed as *effective* second-order responses:

$$\gamma_{EFISHG} \equiv \frac{\left[\mu\beta_{||}\right]_{eff}}{3kT} \tag{7.6}$$

The validity of this assumption is discussed hereafter at the light of computational results.

7.2.2 Computational methodology

The computational methodology consisted of two steps. By using **D3** as reference system, we first addressed the dynamical behavior of the complex composed by the dye and its counterion, as well as the impact of structural fluctuations on the EFISHG response. This first step was achieved by adopting the same framework as used in previous studies [16–18]: i) the structures of chloroform-solvated ion pairs were first generated using classical molecular dynamics (MD) simulations employing a system-specific force field (see next section), and ii) non-correlated structural snapshots were eventually extracted at regular time intervals of the trajectories to calculate their EFISHG responses by means of time-dependent density functional theory (TD-DFT). This MD/TD-DFT approach allowed us to assess the relative magnitude of the $\gamma_{//}$ and $\mu\beta_{||}/3kT$ contributions ($R_{3/2}$, equation 7.5), as well as to identify the average position of the iodine with respect to the cationic chromophore. In a second step, we extended the calculations to the whole set of molecules by optimizing the structure of the dye/iodine complexes at the DFT level. The average position of the iodine with respect to the chromophore derived from MD simulations on **D3** was used to prepare the initial dye/iodine structures for geometry optimization. On the basis of the relaxed structures, the EFISHG responses of all complexes were calculated using TD-DFT and compared to experimental results.

7.2.2.1 Force field parameterization and MD simulations

The general AMBER force field (FF) was partially modified to finely reproduce the equilibrium geometry and torsional degrees of freedom of **D3**, which is prerequisite to obtain reliable description of the NLO responses. In particular, we have modified the FF bond lengths to accurately reproduce (with a mean absolute error of 0.002 Å) those calculated at the M06-2X/6-311G(d) level, in which solvent effects (chloroform) were accounted for by using the integral equation formalism of the polarizable continuum model (IEF-PCM)

[25]. Electrostatic potential-fitted (ESP) atomic charges of the chromophore, iodide anion and chloroform have been obtained at the same level of theory. In addition, relaxed potential energy scans were performed at the M06-2X/6-311G(d) level in gas phase for five relevant dihedral angles (θ_1-θ_5, see Figure 7.1) and fitted by employing the procedure described in Ref. [26]. These torsional potentials were calculated using a simplified structure of the **D3** molecule, in which butyl chains were removed from the chromophore since they are irrelevant for the FF parameterization. The FF potentials are free-energy profiles obtained using the adapting biasing force (ABF) method considering the chromophore inside a box with 40 Ar atoms at the temperature of 298.15 K. Reparameterization of the bonds and torsion potentials was done iteratively until convergence. All the details of force field parameterization are reported in the Annex.

7.2.2.2 Molecular dynamics simulations

The MD simulations have been performed using the NAMD software [27], starting from a low density cubic box of size 120 Å and equilibrated for 10 ns in the NpT ensemble (p = 1 atm and T = 298.15 K). The production run was performed in the NVT ensemble. In order to better span the conformational degrees of freedom of the dye, two different initial conformations corresponding to different values of the θ_3 dihedral were considered. Each MD run was of 20 ns with timestep of 1 fs and performed by rescaling the temperature to 298.15 K every 100 steps. 200 structural snapshots were extracted from each MD trajectory, providing a total set of 400 structures for calculating the NLO properties. The probability distributions of the values of relevant geometrical parameters (bond length alternation, torsional angles, anion-cation distances, see Annex B) for the 400 geometrical snapshots were found to coincide with the distributions obtained by using the 40000 structures extracted at every timestep of the simulation, confirming that the set of selected geometries is representative of the dynamics of the system.

7.2.2.3 Calculation of the EFISHG responses

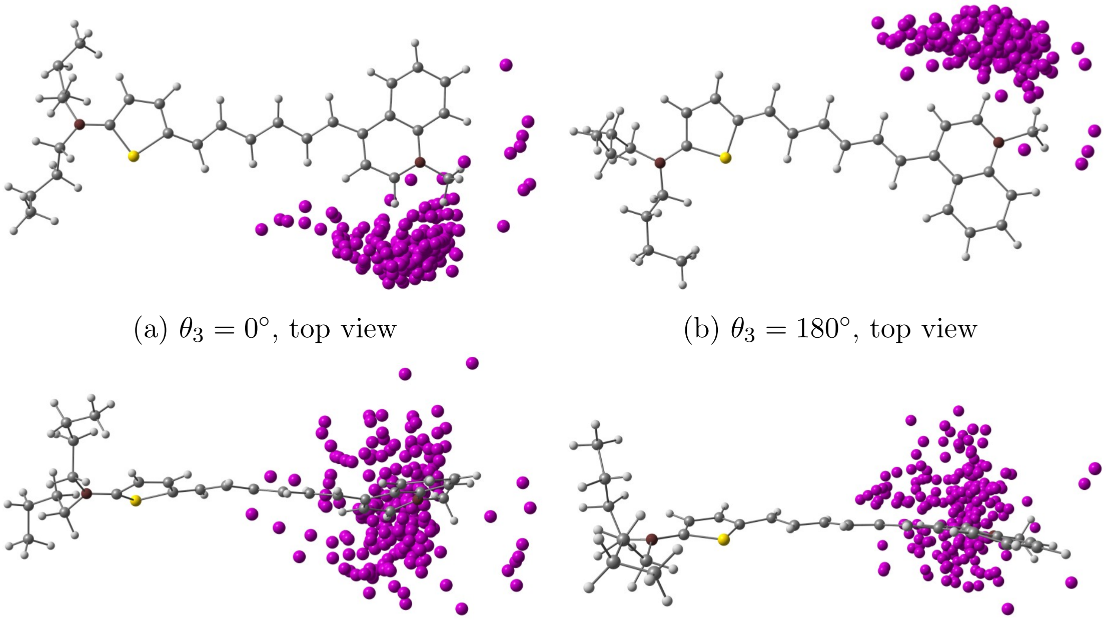
For all compounds, the molecular geometries of the dye/iodide complexes were optimized at IEF-PCM:M06-2X/6-311G(d) level in chloroform. The optimized structures were confirmed to be real minima of the potential energy surface on the basis of their harmonic vibrational frequencies, which were found real for all normal modes. Calculations of hyperpolarizabilities have been evaluated at the TD-DFT level employing the M06-2X exchange-correlation (XC) functional [28]. First hyperpolarizabilities were computed analytically using the standard TD-DFT method [29], while calculations of the second hyperpolarizabilities were performed by using the first-order numerical derivatives of the analytical first hyperpolarizabilities. The suitability of M06-2X for computing the second-order NLO properties of push-pull π-conjugated dyes was demonstrated in previous computational works [30, 31]. To further assess its adequacy for the compounds investigated here, static first hyperpolarizabilities were also calculated using the second-order

Møller–Plesset perturbation theory (MP2), in which the energy derivatives are calculated using a numerical finite field (FF) procedure and refined by using a Romberg scheme. As reported in Annex (Figure D.15), a very good correlation was found between the two sets of data. Preliminary calculations on series **D** were also performed for choosing the most appropriate basis set. Three basis sets of decreasing complexity were tested for computing the static first hyperpolarizability of the chromophores, namely aug-cc-pVTZ, aug-cc-pVDZ and 6-311+G(d). As reported in Figure D.16, both the 6-311+G(d) and aug-cc-pVDZ basis sets provide static β values very similar to those obtained using the larger aug-cc-pVTZ basis. Therefore, the computationally-cheaper 6-311+G(d) basis was chosen to describe the chromophore. To model the iodide anion, we employed the aug-cc-pVDZ basis set and the associated pseudopotential, although the basis set used for the counterion does not have a significant impact on the first hyperpolarizability of the complexes (Table D.5). All DFT calculations were performed using Gaussian 16 [32]. Graphical representations of the molecules were realized with the Chemcraft package [33].

7.3 Results and discussion

7.3.1 Dynamics of the EFISHG response of the D3/iodine complex

The probability distributions issued from MD samplings reveal large fluctuations in the BLA values along the conjugated linker, with BLA = (-0.045 ± 0.023) Å when considering the 400 geometries used for NLO calculations. MD simulations also show that the position of the iodine with respect to the chromophore is highly dynamical, the counterion nevertheless remaining in close proximity to the dye with an average distance from the quinolinium nitrogen $d_{NI} = (4.57 \pm 0.47)$ Å. Moreover, the iodide anion oscillates around an average position located within the mean plane of the quinolinium moiety, with values of the dihedral angle between the terminal phenyl and the anion $\theta_{PhI} = 175.1° \pm 45.9°$. (Figure D.9). Figure 7.2 illustrates how the iodide anion is fluctuating around the dye for the 400 selected snapshots extracted from the MD simulations.



(a) $\theta_3 = 0°$, top view
(b) $\theta_3 = 180°$, top view
(c) $\theta_3 = 0°$, side view
(d) $\theta_3 = 180°$, side view

Figure 7.2: Distribution of the iodide counterion positions with respect to the **D3** chromophore (fixed in its initial position) for the two different initial conformations of the dihedral θ_3.

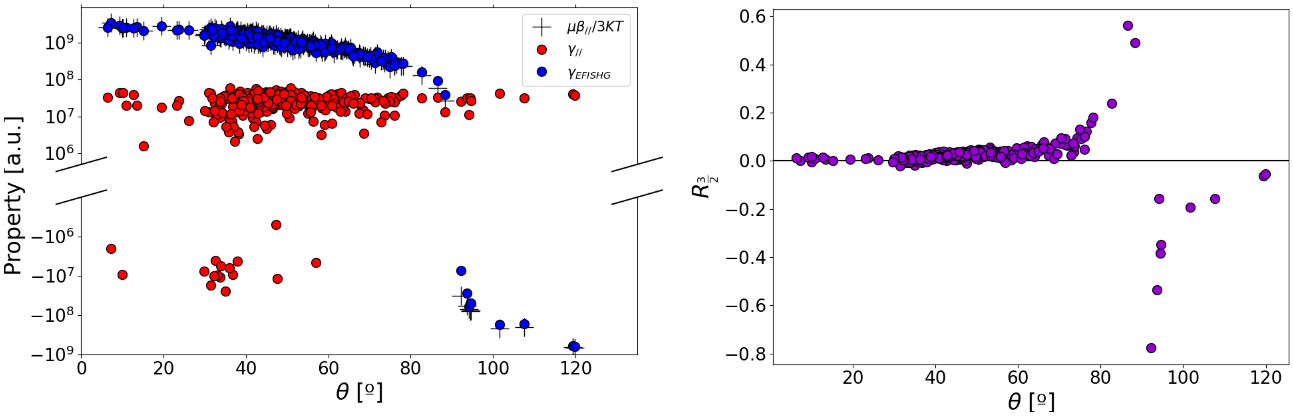
The time evolution of the various terms involved in the EFISHG signal (namely γ_{EFISHG}, $\gamma_{||}$, μ, $\beta_{||}$ and $\mu\beta$, see equations 7.1-7.4), as well as of the angle θ between the $\vec{\mu}$ and $\vec{\beta}$ vectors, are collected in Figure D.10. Their average values and standard deviations are reported in Table 7.1. The results show that both the second- and third-order contributions of the EFISHG response display high sensitivity to structural fluctuations, with standard deviations reaching about 50% of their average values. Moreover, the EFISHG response is strongly dominated by the second-order term $\mu\beta_{||}/3kT$, which is two orders of magnitude larger than $\gamma_{||}$ and confirms the experimental hypothesis assuming that $R_{3/2} \sim 0$. Therefore, the effective EFISHG response defined in equation 7.6, regrouping both second- and third- order NLO contributions, can be adopted for direct comparisons with experiment. As reported in Table 7.1, the $[\mu\beta_{||}]_{eff}$ value predicted by MD/TD-DFT calculations (3.5 10^{-6} a.u.) is overestimated compared to experimental measurements (1.5 10^{-6} a.u.), but the order of magnitude of the NLO response is well reproduced.

Table 7.1: Average values and standard deviations (σ) of the EFISHG properties computed at the IEF-PCM:M06-2X/6-311+G(d) level in chloroform for the **D3** molecule, using the 400 snapshots extracted from the MD trajectories. The NLO properties and the norm of the dipole moment are given in atomic units, the angle θ is given in degrees.

Property	Average	σ
$\gamma_{EFISHG} \times 10^{-7}$	124.1	55.4
$\mu\beta_{\|\|}/3kT \times 10^{-7}$	121.6	55.5
$\gamma_{\|\|} \times 10^{-7}$	2.6	1.3
$R_{3/2}$	0.02	0.07
$\mu\beta_{\|\|} \times 10^{-6}$	3.4	1.6
$\beta_{\|\|} \times 10^{-4}$	16.8	6.1
μ	12	2
θ	48	15
$[\mu\beta_{\|\|}]_{eff} \times 10^{-6}$	3.5	1.6
$[\mu\beta_{\|\|}]_{eff} \times 10^{-6}$ (exp.)	1.5	-

To gain further insight into the relationship between the EFISHG response and the dynamical structure of the dye/iodine complex, Figure 7.3a reports the distribution of the second- and third-order contributions to the EFISHG response with respect to the angle θ between the $\vec{\mu}$ and $\vec{\beta}$ vectors. These plots show that $\gamma_{||}$ displays weak variations with respect to θ, despite a small set of structures give rise to negative $\gamma_{||}$ values, as discussed later on. Figure D.13e further shows that the variations of the $\gamma_{||}$ values are not correlated to those of the θ angle. On the contrary, the second-order contribution $\mu\beta_{||}/3kT$ strongly depends on θ, since $\beta_{||}$ involves the scalar product between the $\vec{\mu}$ and $\vec{\beta}$ vectors (equation 7.3). Therefore, $\mu\beta_{||}/3kT$ progressively decreases as the θ value increases, and cancels out for $\theta = 90°$ before changing sign. In the $\theta \sim 90°$ region, the EFISHG response is thus dominated by the third-order term $\gamma_{||}$, which translates into a divergence of the $R_{3/2}$ ratio (Figure 7.3b). As illustrated in Figure 7.4, the value of the θ angle is entirely driven by the position of the iodine anion with respect to the chromophore. When I^- is located close to the polyenic bridge, the $\vec{\mu}$ and $\vec{\beta}$ vectors are quasi perpendicular ($\theta \sim 90°$), giving rise to a vanishing $\mu\beta_{||}/3kT$ contribution, while when I^- is found close to the electron-withdrawing extremity of the dye ($\theta \sim 0°$), the $\mu\beta_{||}/3kT$ contribution is maximized. The average position of I^- deduced from the MD trajectories corresponds to an intermediate value of $\theta = 48°$, situation in which $\gamma_{||}$ is negligible with respect to $\mu\beta_{||}/3kT$ ($R_{3/2} = 0.02$). Note that, in addition to $\beta_{||}$, the norm of the total dipole moment is also strongly correlated with the θ angle, and thus with the position of the iodide anion. The evolution of $\beta_{||}$ and μ with respect to θ is illustrated in Figure D.13. Consistent with the dipole variations, we can also notice in Figure 7.4 that the position of the counterion induces significant variation in the electrostatic potential within the molecule. For $\theta \sim 0°$, the charge transfer along the long molecular axis is enhanced, reinforcing the asymmetry of the electron density, while for $\theta \sim 90°$ the potential displays a more symmetrical shape. However, as illustrated in

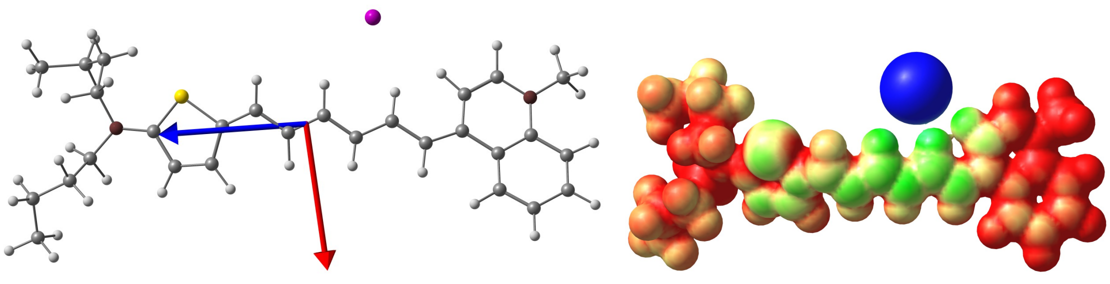
Figure D.13, the total first hyperpolarizability (β_{tot}) does not exhibit any clear correlation with θ, indicating that the changes in the electronic density of the chromophore induced by the iodide position has no significant impact on the second-order NLO response itself, but only on its projection onto the dipole moment direction. This is further illustrated in Figure D.12, which shows that β_{tot} values calculated in presence of the iodide and those calculated for the same geometries of the chromophore after removing the anion evolve similarly.



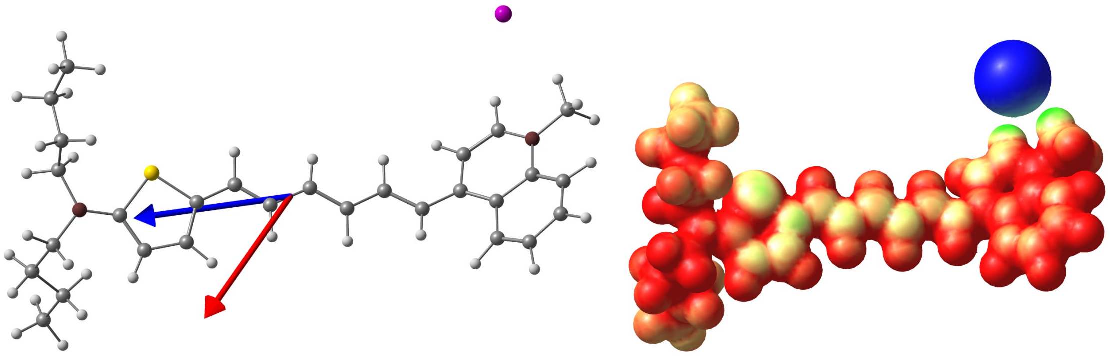
(a)

(b)

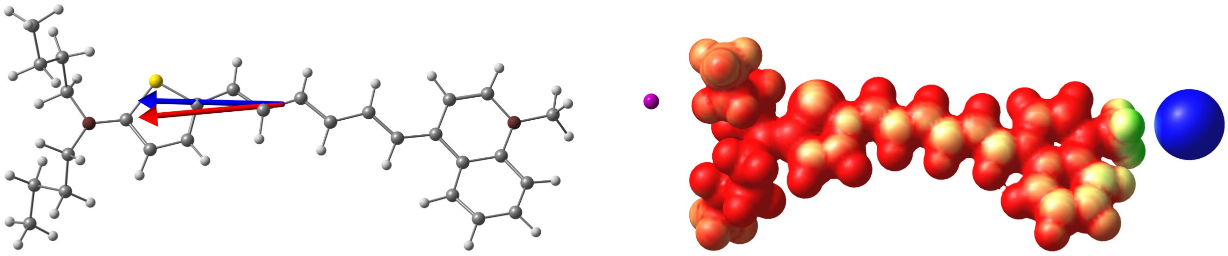
Figure 7.3: Distribution of the total EFISHG response and of its second- and third-order contributions ($\mu\beta_{||}/3kT$ and $\gamma_{||}$) (a) and of the $R_{3/2}$ ratio (b) as a function of the angle θ for the **D3**. The y-axis in plot (a) is in logarithmic scale.



(a) $\theta = 94°$ (b) MEP



(c) $\theta = 48°$ (d) MEP



(e) $\theta = 7°$ (f) MEP

Figure 7.4: Left: Structure of the **D3**/iodine complex for three different positions of the anion, with the value of the θ angle between the (normalized) $\vec{\mu}$ (red) and $\vec{\beta}$ (blue) vectors; Right: Molecular Electrostatic Potential (MEP) for the three configurations.

It is also instructive to address the variations of the NLO properties with respect to the BLA along the conjugated segment of the chromophore. Figure D.11 first reveals that the BLA values calculated along the MD trajectories are spread over a broad range, from negative to positive values. In addition, the BLA globally increases as the θ angle decreases, which indicates that the position of the iodide anion somehow influences the conjugation along the polyenic linker, although the two quantities do not show a clear direct correlation. Recent calculations also demonstrated a similar influence of the position of the anion on the BLA of the organic cation in cyanine crystals [34].

As illustrated in Figure D.14, $\gamma_{||}$ evidences a global reverse correlation with BLA, negative $\gamma_{||}$ values being associated to highly conjugated structures with BLA values close to zero. A similar relationship between $\gamma_{||}$ and BLA values was recently observed for the phenol blue, another typical cyanine dye [35]. Interestingly, the variations of the isotropic linear polarizability (α_{iso}) of **D3** are also clearly correlated with BLA, while μ and β_{tot} are not. This evidences that odd-order optical quantities are mainly impacted by the fluctuations in the degree of conjugation within the chromophore, while even-order ones mainly depends on the fluctuations in the iodide position.

7.3.2 EFISHG response of the dye/iodide complexes

In a second step, the structure of all the dye/iodide complexes represented in Figure 7.1 were optimized at the IEF-PCM:M06-2X/6-311G(d) level, by setting the initial position of the anion to its average position according to the MD samplings. The BLA along the polyenic linker of the dye, the value of the θ_2 and θ_3 torsional angles, as well as the d_{NI} and θ_{PhI} values associated to the position of the iodide anion are reported in Table 7.2 for all complexes. As indicated by the latter parameters, the optimization process did not change significantly the position of the iodine (see Figures D.17, D.18, D.19, D.20 for a scheme of the optimized geometries). In series **A** and **B**, the absolute BLA values slightly decrease with increasing n, while they are significantly smaller and regularly increase in series **C** and **D**, indicating that the conjugation along the polyenic bridge is mostly driven by the nature of the donor moiety. Note that, for very extended linkers ($n >> 5$), the BLA is expected to converge towards similar values independently on the nature of the donor and acceptor moieties, as the conjugated segment resembles in this case an unsubstituted polyene. Furthermore, the values of θ_2 show that molecules incorporating a quinolinium acceptor (series **B** and **D**) display larger deviation from planarity than their pyridinium analogs.

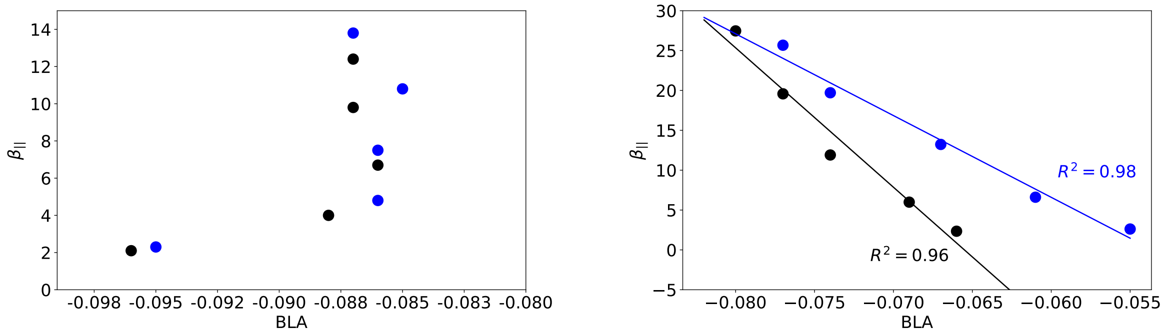
Table 7.2: Bond length alternation (BLA, Å), torsional angles θ_2 and θ_3 (degrees), as well as d_{NI} (Å) and θ_{PhI} (degrees) values associated to the position of the iodide anion in all dye/iodide complexes.

Molecule	BLA	θ_2	θ_3	d_{NI}	θ_{PhI}
A1	-0.096	-177.6	-177.6	4.49	-177.6
A2	-0.088	-179.5	-178.0	4.49	174.2
A3	-0.086	179.2	177.3	4.49	-177.2
A4	-0.087	-178.4	179.3	4.49	-179.3
A5	-0.087	178.9	178.1	4.47	172.8
B1	-0.095	156.9	174.0	4.50	-175.2
B2	-0.086	-160.2	-178.9	4.51	-179.3
B3	-0.086	-159.7	-177.6	4.50	177.0
B4	-0.085	-161.6	178.5	4.51	-177.9
B5	-0.087	-159.4	179.4	4.50	-177.0
C1	-0.066	-179.8	-179.3	4.50	-178.7
C2	-0.069	-179.6	-179.6	4.49	-178.1
C3	-0.074	177.6	179.8	4.48	-173.0
C4	-0.077	179.8	-179.5	4.49	177.9
C5	-0.080	-179.3	-179.3	4.49	177.5
D1	-0.055	-172.1	-178.3	4.53	-179.1
D2	-0.061	-167.7	-179.2	4.52	-179.9
D3	-0.067	167.2	-179.1	4.52	178.3
D4	-0.074	163.0	-180.0	4.51	-177.9
D5	-0.077	163.8	-179.1	4.51	176.8

The computed NLO responses collected in Table 7.3 show that, whatever the nature of the dye, the third-order contribution to the total EFISHG response remains small with respect to the second-order contribution, with $R_{3/2}$ ratios not exceeding 8%. This result is consistent with an older report on stilbazolium-anion complexes [16]. Nevertheless, the $R_{3/2}$ ratio monotically increases in the four series when elongating the polyenic linker, indicating that the $\gamma_{||}$ contribution should be considered in the case of highly extended systems. Series **C** and **D** incorporating a dibutyl-aminothienyl donor display larger second-order $\beta_{||}$ responses, which inversely correlate with the BLA values (Figure 7.5). Consistently, series **A** and **B**, which display small variations of BLA with n, show a lower enhancement of $\beta_{||}$ when increasing the size of the bridge, although without showing any BLA-$\beta_{||}$ correlation. Interestingly, the values of the dipole moment do not vary significantly with n in any of the molecular series, so that the magnitude of the second-order contribution of the EFISHG response, $\mu\beta_{||}/3kT$, is essentially driven by the value of $\beta_{||}$ and that of the θ angle between the $\vec{\mu}$ and $\vec{\beta}$ vectors. While $\beta_{||}$ smoothly increases with n in the four series, θ increases in the **A**, **B** and **D** families, which has for effect to damp the increase of the $\mu\beta_{||}/3kT$ term. On the contrary, the θ angle decreases in the **C** series, making the $\vec{\mu}$ and $\vec{\beta}$ vectors more and more parallel as elongating the linker, and thus contributing to the enhancement of the second-order response.

Table 7.3: Components of the dynamic (λ = 1907 nm) EFISHG responses of the dye/iodine complexes calculated at the IEF-PCM:M06-2X/6-311+G(d) level: dipole moment norm (μ, a.u.), parallel first hyperpolarizability ($\beta_{||}, 10^4$ a.u.), angle between $\vec{\mu}$ and $\vec{\beta}$ (θ, degrees), $\mu\beta_{||}$ (10^5 a.u.), $\mu\beta_{||}/3kT$ (10^8 a.u.), parallel second hyperpolarizability ($\gamma_{||}$, 10^6 a.u.), total EFISHG signal (γ_{EFISHG}, 10^8 a.u.), effective EFISHG response ($[\mu\beta_{||}]_{eff}$, 10^5 a.u.), and $R_{3/2}(\times 10^2)$ ratio.

Molecule	μ	$\beta_{\|\|}$	θ	$\mu\beta_{\|\|}$	$\mu\beta_{\|\|}/3kT$	$\gamma_{\|\|}$	γ_{EFISHG}	$[\mu\beta_{\|\|}]_{eff}$	$R_{3/2}$
A1	10.5	2.1	45	3.7	1.3	2.4	1.3	3.8	1.8
A2	10.5	4.0	47	6.9	2.4	6.5	2.5	7.1	2.7
A3	10.6	6.7	47	11.9	4.2	15.2	4.3	12.3	3.6
A4	10.7	9.8	48	17.4	6.2	28.5	6.4	18.2	4.6
A5	10.5	12.4	51	21.7	7.7	46.2	8.1	23.0	6.0
B1	10.8	2.3	53	4.2	1.5	3.7	1.5	4.3	2.5
B2	10.7	4.8	53	8.5	3.0	10.3	3.1	8.8	3.4
B3	10.7	7.5	54	13.4	4.7	22.5	5.0	14.1	4.7
B4	10.6	10.8	56	19.1	6.7	42.2	7.1	20.3	6.3
B5	10.6	13.8	56	24.3	8.6	66.0	9.2	26.1	7.7
C1	11.5	2.3	37	4.5	1.6	1.3	1.6	4.5	0.8
C2	12.2	6.0	30	12.2	4.3	5.5	4.4	12.4	1.3
C3	12.6	11.9	27	25.0	8.8	17.4	9.0	25.5	2.0
C4	12.8	19.6	24	41.9	14.8	40.6	15.2	43.0	2.7
C5	13.0	27.5	22	59.3	20.9	74.2	21.7	61.4	3.5
D1	12.1	2.6	44	5.3	1.9	1.3	1.9	5.3	0.7
D2	12.0	6.6	47	13.2	4.6	7.6	4.7	13.4	1.6
D3	12.0	13.2	48	26.4	9.3	28.1	9.6	27.2	3.0
D4	11.8	19.7	51	38.7	13.7	68.4	14.3	40.6	5.0
D5	11.5	25.7	53	49.4	17.4	124.6	18.7	52.9	7.1



(a) (b)

Figure 7.5: Correlation between $\beta_{||}$ (10^4 a.u.) and BLA (Å) values in (a) series **A** (black) and **B** (blue) and (b) in series **C** (black) and **D** (blue).

7.3.3 Comparison to experiment

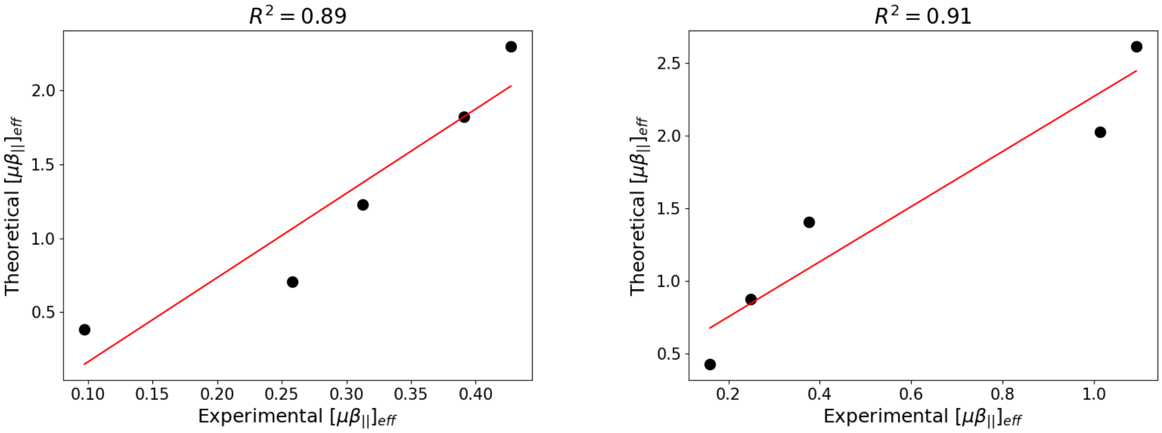
The experimental EFISHG data are collected in Table 7.4, together with the maximum absorption wavelength. The static EFISHG responses extrapolated using the two-state approximation are also reported. Consistently with the red shift of the main absorption band, the EFISHG intensity increases in each series with the size of the polyenic bridge. As expected, removing frequency dispersion effects by extrapolating responses to the infinite-wavelength limit attenuates the increase of $\mu\beta_{||}$ with chain length. As shown in Figure 7.6, theoretical results are in good agreement with measured data. All series display Pearson correlation coefficients larger than 0.89 between computed and experimental $[\mu\beta_{||}]_{eff}$ values, with R^2 values very close to 1.0 for series **C** and **D**. However, the relative ordering of the $[\mu\beta_{||}]_{eff}$ values in the different series is only partially reproduced: while DFT calculations predict that series **A** and **B** display the smallest NLO responses in agreement with experimental results, the relative magnitude of the EFISHG signal is inverted in series **C** and **D**. The enhancement of the $[\mu\beta_{||}]_{eff}$ values with the elongation of the polyenic chain is also underestimated in the two latter series.

Table 7.4: Experimental EFISHG responses ($\mu\beta_{||}(2\omega) \equiv \left[\mu\beta_{||}(-2\omega;\omega,\omega)\right]_{eff}$) measured at $\lambda = 1907$ nm, maximum absorption wavelength (λ_{max}, nm), two-state frequency dispersion factors $F(\omega)$,[a] and static EFISHG responses extrapolated as $\mu\beta_{||}(0) = \mu\beta_{||}(2\omega)/F(\omega)$. All NLO data are given in 10^{-48} esu (1 a.u. of $\beta = 8.6392\ 10^{-33}$ esu).

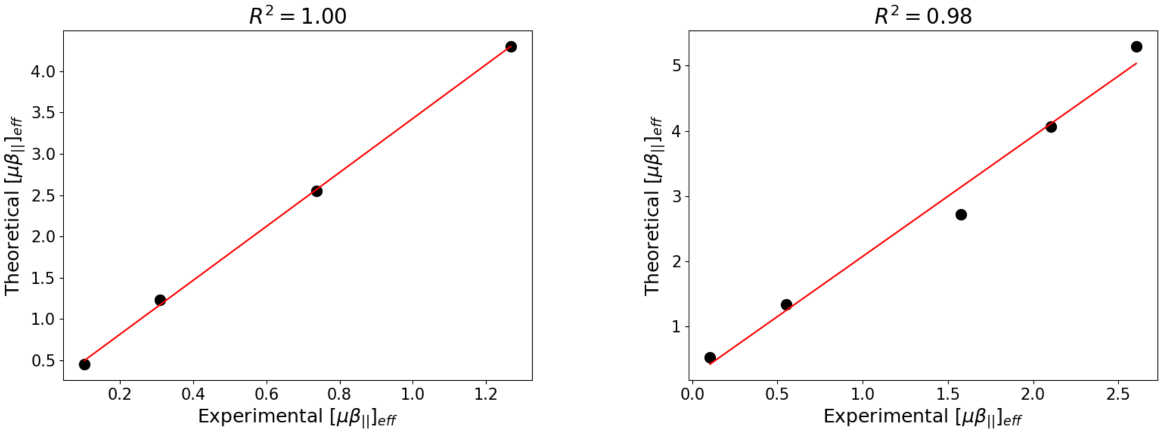
Molecule	$\mu\beta_{\|\|}(0)$	$\mu\beta_{\|\|}(2\omega)$	λ_{max}	$F(\omega)$
A1[b]	549	841	517	1.53
A2[b]	1338	2230	558	1.67
A3[b]	1537	2700	581	1.76
A4[b]	1887	3380	589	1.79
A5[b]	2014	3690	598	1.83
B1	783	1380	582	1.76
B2	1098	2156	582	1.96
B3	1566	3250	624	2.08
B4	4050	8750	624	2.16
B5	4284	9435	643	2.20
C1	523	891	568	1.70
C2	1332	2669	631	2.00
C3	2946	6366	656	2.16
C4	4954	10946	663	2.21
D1	445	900	634	2.02
D2	1793	4770	714	2.66
D3	4083	13600	763	3.33
D4	5710	18170	754	3.18
D5	7281	22500	748	3.09

[a] $1/F(\omega) = (1 - \lambda_{max}^2/\lambda^2)\,(1 - 4\lambda_{max}^2/\lambda^2)$

[b] From Ref. 13



(a) series **A** (b) series **B**



(c) series **C** (d) series **D**

Figure 7.6: Correlation between experimental and computational $\mu\beta_{||}(2\omega)$ values (in 10^7 a.u.) for the four series of molecules.

7.4 Conclusions

In this work, the NLO responses of four series of amphiphilic cationic chromophores have been investigated by means of EFISHG experiments and quantum chemical calculations. EFISHG measurements were made possible owing to the electric neutrality of the pairs formed, in chloroform, by the cationic dye and the iodine anion. The computational approach combining MD simulations and time-dependent DFT calculations performed on a representative dye/iodine pair allowed us to describe the geometrical fluctuations of the complexes, and to highlight their impact on the NLO responses. These calculations confirmed that the iodine anion remains in the proximity of the dye all along the simulations, with an average position close to the charged heterocyclic acceptor. The position of the iodine relative to the dye was also shown to be at the origin of the relative magnitude of the second- and third-order contributions of the EFISHG signal. In the investigated system, the third-order component was found negligible compared to the second-order one, supporting the experimental assumptions. Overall, the good agreement between experimental and theoretical characterizations demonstrates the reliability of the computational protocol, which provides a tractable strategy towards systematic and accurate *in silico* characterization of novel molecular probes for SHG imaging.

Bibliography

[1] Dalton, L. R.; Sullivan, P. A.; Bale, D. H. Electric Field Poled Organic Electro-optic Materials: State of the Art and Future Prospects. *Chem. Rev.* **2010**, *110*, 25–55.

[2] Campagnola, P. J.; Wei, M. D.; Loew, L. M. High-Resolution Nonlinear Optical Imaging of Live Cells by Second Harmonic Generation. *Biophys. J.* **1999**, *77*, 3341–3349.

[3] Moreaux, L.; Sandre, O.; Charpak, S.; Blanchard-Desce, M.; Mertz, J. Coherent Scattering in Multi-Harmonic Light Microscopy. *Biophys. J.* **2001**, *80*, 1568–1574.

[4] Dombeck, D. A.; Blanchard-Desce, M.; Webb, W. W. Optical Recording of Action Potentials with Second-Harmonic Generation Microscopy. *J. Neurosci.* **2004**, *24*, 999–1003.

[5] Reeve, J. E.; Anderson, H. L.; Clays, K. Dyes for biological second harmonic generation imaging. *Phys. Chem. Chem. Phys.* **2010**, *12*, 13484–13498.

[6] Marder, S. R.; Cheng, L.-T.; Tiemann, B. G.; Friedli, A. C.; Blanchard-Desce, M.; Perry, J. W.; Skindhøj, J. Large First Hyperpolarizabilities in Push-Pull Polyenes by Tuning of the Bond Length Alternation and Aromaticity. *Science* **1994**, *263*, 511–514.

[7] Bourhill, G.; Bredas, J.-L.; Cheng, L.-T.; Marder, S. R.; Meyers, F.; Perry, J. W.; Tiemann, B. G. Experimental Demonstration of the Dependence of the First Hyperpolarizability of Donor-Acceptor-Substituted Polyenes on the Ground-State Polarization and Bond Length Alternation. *J. Am. Chem. Soc.* **1994**, *116*, 2619–2620.

[8] Parthasarathy, V.; Pandey, R.; Stolte, M.; Ghosh, S.; Castet, F.; Würthner, F.; Das, P. K.; Blanchard-Desce, M. Combination of Cyanine Behaviour and Giant Hyperpolarisability in Novel Merocyanine Dyes: Beyond the Bond Length Alternation (BLA) Paradigm. *Chem. Eur. J.* **2015**, *21*, 14211–14217.

[9] Parthasarathy, V.; Pandey, R.; Das, P. K.; Castet, F.; Blanchard-Desce, M. Linear and Nonlinear Optical Properties of Tricyanopropylidene-Based Merocyanine Dyes:

Synergistic Experimental and Theoretical Investigations. *ChemPhysChem* **2018**, *19*, 187–197.

[10] Coe, B. J. Molecular Materials Possessing Switchable Quadratic Nonlinear Optical Properties. *Chem. Eur. J.* **1999**, *5*, 2464–2471.

[11] Castet, F.; Rodriguez, V.; Pozzo, J.-L.; Ducasse, L.; Plaquet, A.; Champagne, B. Design and Characterization of Molecular Nonlinear Optical Switches. *Acc. Chem. Res.* **2013**, *46*, 2656–2665.

[12] Beaujean, P.; Bondu, F.; Plaquet, A.; Garcia-Amorós, J.; Cusido, J.; Raymo, F. M.; Castet, F.; Rodriguez, V.; Champagne, B. Oxazines: A New Class of Second-Order Nonlinear Optical Switches. *J. Am. Chem. Soc.* **2016**, *138*, 5052–5062.

[13] Alain, V.; Blanchard-Desce, M.; Ledoux-Rak, I.; Zyss, J. Amphiphilic polyenic push–pull chromophores for nonlinear optical applications. *Chem. Commun.* **2000**, 353–354.

[14] Kannan, S.; Sekar, A.; Sivaperuman, K. Effects of the molecular structure on the second-order nonlinear optical properties of stilbazolium derivative single crystals: a review. *J. Mater. Chem. C* **2020**, *8*, 16668–16690.

[15] Tessore, F.; Cariati, E.; Cariati, F.; Roberto, D.; Ugo, R.; Mussini, P.; Zuccaccia, C.; Macchioni, A. The Role of Ion Pairs in the Second-Order NLO Response of 4-X-1-Methylpiridinium Salts. *ChemPhysChem* **2010**, *11*, 495–507.

[16] Ramos, T. N.; Canuto, S.; Champagne, B. Unraveling the Electric Field-Induced Second Harmonic Generation Responses of Stilbazolium Ion Pairs Complexes in Solution Using a Multiscale Simulation Method. *J. Chem. Inf. Model.* **2020**, *60*, 4817–4826.

[17] Ramos, T. N.; Castet, F.; Champagne, B. Second Harmonic Generation Responses of Ion Pairs Forming Dimeric Aggregates. *J. Phys. Chem. B* **2021**, *125*, 3386–3397.

[18] Pielak, K.; Tonnelé, C.; Sanguinet, L.; Cariati, E.; Righetto, S.; Muccioli, L.; Castet, F.; Champagne, B. Dynamical behavior and second harmonic generation responses in acido-triggered molecular switches. *J. Phys. Chem. C* **2018**, *122*, 26160−26168.

[19] Lescos, L.; Beaujean, P.; Tonnelé, C.; Aurel, P.; Blanchard-Desce, M.; Rodriguez, V.; de Wergifosse, M.; Champagne, B.; Muccioli, L.; Castet, F. Self-assembling, structure and nonlinear optical properties of fluorescent organic nanoparticles in water. *Phys. Chem. Chem. Phys.* **2021**, *23*, 23643–23654.

[20] Tonnelé, C.; Pielak, K.; Deviers, J.; Muccioli, L.; Champagne, B.; Castet, F. Nonlinear optical responses of self-assembled monolayers functionalized with indolino–oxazolidine photoswitches. *Phys. Chem. Chem. Phys.* **2018**, *20*, 21590–21597.

[21] Tonnelé, C.; Champagne, B.; Muccioli, L.; Castet, F. Nonlinear Optical Contrast in Azobenzene-Based Self-Assembled Monolayers. *Chem. Mater.* **2019**, *31*, 6759–6769.

[22] Bouquiaux, C.; Tonnelé, C.; Castet, F.; Champagne, B. Second-Order Nonlinear Optical Properties of an Amphiphilic Dye Embedded in a Lipid Bilayer. A Combined Molecular Dynamics - Quantum Chemistry Study. *J. Phys. Chem. B* **2020**, *124*, 2101–2109.

[23] Bouquiaux, C.; Castet, F.; Champagne, B. Unravelling the Effects of Cholesterol on the Second-Order Nonlinear Optical Responses of Di-8-ANEPPS Dye Embedded in Phosphatidylcholine Lipid Bilayers. *J. Phys. Chem. B* **2021**, *125*, 10195–10212.

[24] Ledoux, I.; Zyss, J. Influence of the molecular environment in solution measurements of the Second-order optical susceptibility for urea and derivatives. *Chemical Physics* **1982**, *73*, 203–213.

[25] Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999–3093.

[26] Pizzirusso, A.; Savini, M.; Muccioli, L.; Zannoni, C. An atomistic simulation of the liquid-crystalline phases of sexithiophene. *J. Mater. Chem.* **2011**, *21*, 125–133.

[27] Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

[28] Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[29] van Gisbergen, S. J. A.; Snijders, J. G.; Baerends, E. J. Calculating frequency-dependent hyperpolarizabilities using time-dependent density functional theory. *J. Chem. Phys.* **1998**, *109*, 10644–10656.

[30] Johnson, L. E.; Dalton, L. R.; Robinson, B. H. Optimizing Calculations of Electronic Excitations and Relative Hyperpolarizabilities of Electrooptic Chromophores. *Acc. Chem. Res.* **2014**, *47*, 3258–3265.

[31] Lescos, L.; Sitkiewicz, S.; Beaujean, P.; Blanchard-Desce, M.; Champagne, B. R.; Matito, E.; Castet, F. Performance of DFT Functionals for Calculating the Second-Order Nonlinear Optical Properties of Dipolar Merocyanines. *Phys. Chem. Chem. Phys.* **2020**, *22*, 16579–16594.

[32] Frisch, M. J. et al. Gaussian 16 Revision C.01. 2016; Gaussian Inc. Wallingford CT.

[33] Chemcraft - graphical software for visualization of quantum chemistry computations. `https://www.chemcraftprog.com`.

[34] Eskandari, M.; Roldao, J. C.; Cerezo, J.; Milián-Medina, B.; Gierschner, J. Counterion-Mediated Crossing of the Cyanine Limit in Crystals and Fluid Solution: Bond Length Alternation and Spectral Broadening Unveiled by Quantum Chemistry. *J. Am. Chem. Soc.* **2020**, *142*, 2835–2843.

[35] Brandão, I.; Fonseca, T. L.; Franco, L. R.; Georg, H. C.; Castro, M. A. Density functional theory investigation of the second hyperpolarizability of the phenol blue in solution. *Chem. Phys. Lett.* **2022**, *796*, 139549.

Conclusions and perspectives

This thesis work contributes to the field of computational chemistry applied to the study, interpretation, and rationalization of electronic nonlinear optical (NLO) properties of molecular systems. We especially focused on assessing the most appropriate computational protocols depending on the nature of the investigated systems. Important challenges have been addressed: i) the accurate description of NLO properties of small molecules in gas phase through state-of-the-art wavefunction methods, ii) the investigation by means of Density Functional Theory (DFT) of the impact of dispersion effects on the second-order NLO properties of molecular switches, and iii) the rationalization of experimental second-order responses of ion pairs in solution, including dynamic effects through a multiscale Molecular Dynamics (MD)/DFT computational protocol.

In Chapters 4 and 5, the performances of some state-of-the-art wavefunction methods have been evaluated for the computation of linear and nonlinear optical properties of small molecules. This study focused on static electronic properties, namely: the dipole moment (μ), the polarizability (α), the first-hyperpolarizability (β), and the second-hyperpolarizability (γ), which are computed numerically through the Finite Field derivatives of field dependent energies. The methods tested either reduce the computational cost or improve the performance of canonical methods for the computation of energies by reformulation of the wavefunction employing judicious mathematical transformations, empirical parameters, thresholds, and cutoffs.

In particular, Resolution of Identity (RI) approximation techniques (RI-MP2, RIJK-MP2, and RIJCOSX-MP2), methods based on the localization of molecular orbitals (DLPNO-MP2, DLPNO-CCSD, DLPNO-CCSD(T0), DLPNO-CCSD(T1), LMP2, LNO-CCSD, and LNO-CCSD(T)) and Spin Component Scaling methods (SCS-MP2 and SOS-MP2) were assessed in Chapter 4. Our results evidenced that all these methods accurately reproduce the μ and α values provided by reference CCSD(T) calculations. On the other hand, we observed that the numerical errors arising from numerical instabilities of the field-dependent energies highly impact the higher-order derivatives. Methods based on the RI approximation give stable high-order derivatives, and therefore they are a cost-effective way to obtain first and second hyperpolarizabilities with a negligible error with respect to canonical MP2 for β and γ. Reversely, methods based on orbital localization (namely, LNO and DLPNO schemes) applied to MP2, CCSD, and CCSD(T) wavefunctions, suffer from numerical instabilities that result in large numerical errors for β and γ. The only exception is DLPNO-CCSD(T1), which produces acceptable errors for the calculation of β. Interestingly, numerically β and γ values computed from analytical polarizabilities with the DLPNO-MP2 method, are in excellent agreement with the canonical MP2 counterparts, suggesting that the implementation of analytical DLPNO-CCSD polarizabilities would offer a cost-effective method to compute accurate hyperpolarizabilities. Finally, spin-component scaled methods do not produce any relevant improvement compared to canonical MP2, except in the case of γ.

In Chapter 5, we have addressed the performance of MP3:KS schemes for computing static NLO properties. This method, by a formulation that combines wavefunction theory and density functional approximations (DFAs), aims at outperforming standard MP3 and challenges CCSD. We have demonstrated that, differently from what has been observed in previous works for the computation of energies, selecting a reference DFA that does not suffer from the delocalization error is fundamental for calculating the NLO properties. In particular, the MP3:ωB97-XV method can be used as a cheaper alternative to CCSD for computations of α and γ. For β, MP3:KS schemes outperform CCSD only for molecules with a small response, the best performance being obtained with MP3:CAM-B3LYP. For systems with large β, CCSD remains the best method among all the approximations considered. However, MP3:KS schemes still improve the standard MP3, and MP3:ωB97-XV emerges as the best among the latter.

In Chapter 6, we investigated the reliability of a series of DFAs and wavefunction methods as tools for modeling the structure and interpreting the second-order NLO response of a series of azobenzene switches bearing different *meta*-substituents. We observed that electron correlation and, in particular, dispersion interactions are the driving forces behind the stabilization of the *cis* conformer with respect to the *trans* one. Indeed, the *cis* isomer stabilizes upon functionalization of the *meta*-position with bulky substituents because of van der Waals interactions. Despite most of DFAs do not accurately reproduce the magnitude of β, DFAs including at least 50% of the exact HF exchange reproduce the evolution of β along the series of azobenzene derivatives. By comparing to RI-CC2 and RI-MP2 methods, we found that ωB97X-D is the best functional to describe the relevant features of these compounds: the geometry, the energy difference between the isomers, and the magnitude of β. Through the PNOC decomposition analysis, we further demonstrated that the main contribution to β mostly comes from the double bond between the nitrogen atoms, followed by smaller contributions arising from the adjacent phenyl rings. Moreover, by a comparison between the PNOC decomposition and the laplacian of the electronic density, we showed that the regions of the molecule that are the most affected by the delocalization error are the ones responsible for the inaccuracy of β.

In Chapter 7, the Second Harmonic Generation (SHG) responses of a series of amphiphilic cationic dyes and their corresponding iodine counterions in chloroform solution have been investigated using a sequential MD/DFT approach. These calculations confirmed that the iodine anion remains in the proximity of the dye, with an average position close to the charged acceptor. The position of the iodine relative to the dye was also shown to be at the origin of the relative magnitude of the second- and third-order contributions of the EFISHG signal. In the amphiphilic chromophores studied, the third-order component was found negligible compared to the second-order one, supporting the experimental assumptions. The excellent agreement between experimental and theoretical results proves that this MD/QM scheme constitutes a useful tool for the design of SHG dyes driven by computational studies.

Overall, this thesis gives an overview of some of the capabilities and limitations of modern computational chemistry methods for computing NLO properties of small and medium-size organic molecules. For small molecules in the gas phase, accurate static NLO properties can be obtained by means of correlated wavefunction methods, the unfavorable size/cost scaling of which can be improved by means of acceleration techniques. For medium size molecules, which are still out of the reach of highly accurate wavefunction methods, DFT remains the workhorse for a qualitative reproduction and interpretation of NLO processes. We have demonstrated in Chapter 6 that recent DFAs such as range separated hybrids have the potential to reach the accuracy of wavefunction methods. However, these methods are designed through an extensive parameterization of the ground-state energies, and are generally not transferable to systems or molecular properties different from those they have been designed for.

Research progresses in the field of computational chemistry applied to NLO could thus be divided in two complementary strategies, namely i) the development of *ab initio* schemes enabling a systematic improvement of the accuracy and/or a lowering of the computational cost, and ii) the development of methods specifically tailored in order to reproduce target systems and properties. Among the latter, recent schemes that mix DFT response theory and semi-empirical parameterizations, referred to as *simplified* (s)DFT schemes, have been proven successful for computing the NLO responses of systems with thousands of electrons [1, 2], these methods are promising for the study of complex molecular and supramolecular systems.

In the aim of reproducing experimental NLO measurements, other important effects such as structural fluctuations and interactions of the investigated system with its environment should be also taken into account. As illustrated in Chapter 7 of this thesis and recently reviewed [3], the former effects can be included in the simulations by using multiscale methods combining sequentially classical molecular dynamics and QM calculations. The association of MD simulations with low-cost methods will certainly broaden the scope of systems affordable for computational chemistry in the near future.

Bibliography

[1] de Wergifosse, M.; Grimme, S. Perspective on Simplified Quantum Chemistry Methods for Excited States and Response Properties. *J. Phys. Chem. A* **2021**, *125*, 3841–3851.

[2] Lescos, L.; Beaujean, P.; Tonnelé, C.; Aurel, P.; Blanchard-Desce, M.; Rodriguez, V.; de Wergifosse, M.; Champagne, B.; Muccioli, L.; Castet, F. Self-assembling, structure and nonlinear optical properties of fluorescent organic nanoparticles in water. *Phys. Chem. Chem. Phys.* **2021**, *23*, 23643–23654.

[3] Castet, F.; Tonnelé, C.; Muccioli, L.; B., C. Predicting the second-order nonlinear optical Responses of Organic Materials: the role of dynamics. *In preparation*

Appendix A

Supporting information of Chapter 4

A.1 Computational time

Before computing the optical properties of the molecules belonging to the γ-NLO set, we performed a test to have a rough idea about the relative computational time required for computing the electrical (hyper)polarizabilities with all the methods considered in this study. Regarding CCSD(T) and CCSD calculations, we performed free-field single point (SP) calculations on the five PA oligomers of the γ-NLO set, as they cover a reasonable range of molecular sizes representative of the whole set. For MP2, the SP calculations were much faster. For this reason, we decided to switch from PA oligomers to PDA oligomers, which are the biggest molecules of the γ-NLO set. All the calculations have been performed on a single node with 16 processors Intel Xeon Gold 6140, with a disk memory of 384 GB and 64 GB of RAM. For CCSD and CCSD(T) calculations all the processors were used in the calculations, while MP2 calculations were run serially. The all times in minutes for each calculation are collected in Tables A.2, A.1 and A.3. The asymptotic behavior of these methods has been evaluated by a polynomial fit correlating the time consumed to the number of basis functions associated with each studied system. The fitting has been performed by adjusting a and b in the function $y = ax^b$, and including the point (0,0) in the fit. We also included CAM-B3LYP SP calculations for the same molecules in order to verify how these methods perform time-wise with respect to a range separated density functional approach time scaling. The results are collected in Figures A.1, A.2, A.3 and Table A.4. From the resulting fit of CC approximations, in general CCSD and CCSD(T) are faster than their corresponding RI counterparts. This behavior is associated to the density fitting implementations, which might not be optimal for these methods. On the contrary, when employing RI-MP2 and RIJK-MP2 schemes, the computational cost drops substantially with respect to MP2. Regarding the RIJCOSX method we observed that moving from the default COSX grid to the COSX2 one (more dense grid) does not substantially affect the cost, and therefore only the second one is presented in the manuscript. Surprisingly, RI approximations result computationally cheaper even with respect to CAM-B3LYP computations, confirming

Table A.1: Wall time for approximate CCSD calculations tested on PA chains of different lengths. These calculations have been performed employing 16 cpus. The values in the table are in minutes / 1 cpu.

			DLPNO-CCSD		LNO-CCSD		
Molecule	CCSD	RI-CCSD	Normal	Tight	Normal	Tight	VeryTight
PA2	3	25	12	5	5	5	5
PA4	14	53	7	14	19	19	24
PA6	52	260	15	31	46	59	87
PA8	185	1041	28	67	96	140	234
PA10	594	3349	44	95	175	259	469
PA12	1538	8803	71	139	281	423	773

that the bottleneck of the calculation for systems of this size is the SCF and not anymore the (accelerated) MP2 part. Considering localized methods, moving to tighter convergence criteria does not drastically increase the computational cost, and therefore owing to the huge accuracy gained in the NLOPs, we only present the calculations employing the Tight criteria on the DLPNO and the veryTight criteria on the LNO construction. For CCSD and CCSD(T) calculations, LNO based methods generally need more time than the corresponding DLPNO counterparts, showing a larger exponent of the fit. On the other hand, T1 corrections for DLPNO-CCSD present a smaller exponent with respect to standard T corrections, but the parameter a is one order of magnitude larger, indicating that time savings start to become relevant only for big molecules (more than 400 basis functions). To finalize, DLPNO-MP2 schemes show comparable timings with respect to MP2. Instead, LMP2 is more efficient than DLPNO and show a scaling comparable to CAM-B3LYP.

Table A.2: Wall time for approximate CCSD(T) calculations tested on PA chains of different lengths. These calculations have been performed employing 16 cpus. The values in the table are in minutes / 1 cpu.

			DLPNO-CCSD(T0)		DLPNO-CCSD(T1)		LNO-CCSD(T)		
Molecule	CCSD(T)	RI-CCSD(T)	Normal	Tight	Normal	Tight	Normal	Tight	vTight
PA2	18	6	4	6	12	6	5	7	6
PA4	12	64	9	15	23	31	20	27	36
PA6	78	253	18	34	38	39	63	86	135
PA8	301	1083	34	78	73	83	131	203	350
PA10	1068	3572	50	113	111	121	221	374	737
PA12	3212	9519	73	172	146	171	339	576	1171

Table A.3: Wall time for approximate MP2 calculations tested on PDA chains of different lengths. These calculations have been performed employing 1 cpu. The values in the table are in minutes / 1 cpu.

			RI-MP2			DLPNO-MP2		LMP2		
Molecule	MP2	SCF	RI-JK-SCF	RI-JCOSX	RI-JCOSX2	Normal	Tight	Normal	Tight	vTight
PDA1	3	2	1	1	2	5	3	1	1	1
PDA2	9	7	2	2	4	11	12	5	5	5
PDA3	22	16	4	5	8	28	27	12	12	12
PDA4	46	31	8	8	13	48	50	23	23	27
PDA5	72	48	14	13	21	72	84	39	37	45

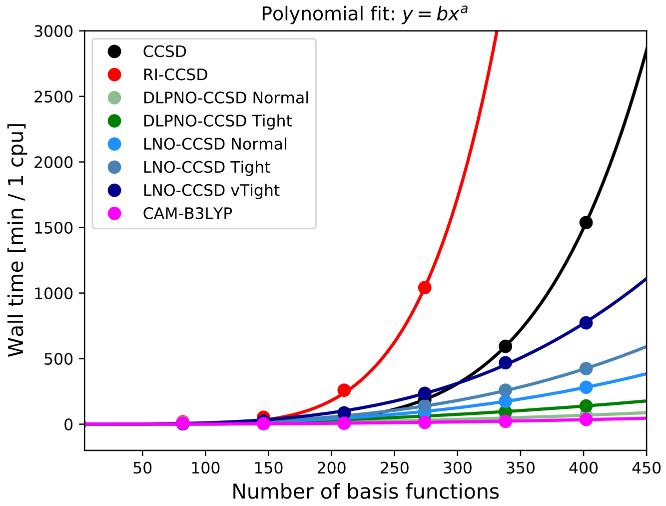


Figure A.1: Wall time for approximate CCSD calculations tested on PA chains of different lengths with respect to the number of electrons. These calculations have been performed employing 16 cpus. The time values are in minutes.

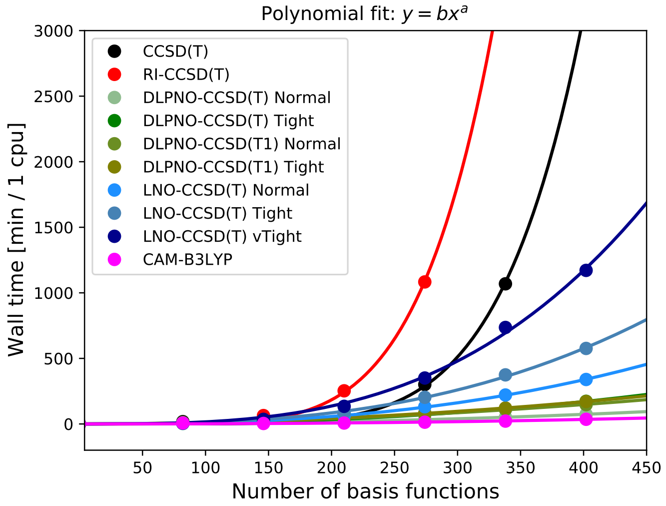


Figure A.2: Wall time in minutes for approximate CCSD(T) calculations tested on PA chains of different lengths with respect to the number of electrons and the corresponding best fits. Data computed using 16 CPUs in a single node.

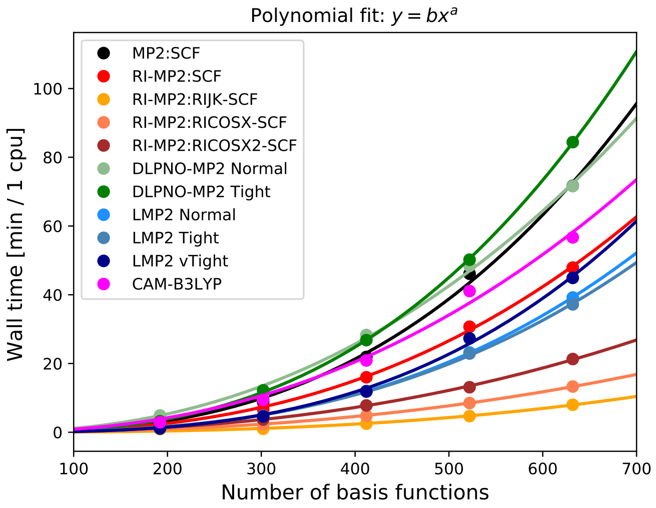


Figure A.3: MP2 timings in minutes for the PDA oligomers with respect to the number of electrons. Data computed using 1 CPU.

Table A.4: Parameters obtained from the best fits of the wall time with respect to the number of electrons of approximated CCSD(T), CCSD, CAM-B3LYP calculations performed using 16 cpus for the PA chains, and from those of MP2 and CAM-B3LYP calculations performed with 1 cpu for the PDA chains.

Method	Approx	a	b [min / 1 cpu]	r^2
CCSD(T)	-	6.28	1.44E-13	1.00
RI-CCSD(T)	-	5.65	1.80E-11	1.00
DLPNO-CCSD(T0)	Normal	2.13	2.11E-04	1.00
DLPNO-CCSD(T0)	Tight	2.34	1.39E-04	1.00
DLPNO-CCSD(T1)	Normal	1.90	1.68E-03	1.00
DLPNO-CCSD(T1)	Tight	1.97	1.27E-03	1.00
LNO-CCSD(T)	Normal	2.56	7.50E-05	1.00
LNO-CCSD(T)	Tight	2.76	3.67E-05	1.00
LNO-CCSD(T)	vTight	3.10	1.01E-05	1.00
CCSD	-	5.48	8.24E-12	1.00
RI-CCSD	-	5.56	2.87E-11	1.00
DLPNO-CCSD	Normal	2.17	1.57E-04	0.97
DLPNO-CCSD	Tight	2.13	4.05E-04	1.00
LNO-CCSD	Normal	2.77	1.72E-05	1.00
LNO-CCSD	Tight	2.93	1.02E-05	1.00
LNO-CCSD	vTight	3.13	5.36E-06	1.00
MP2:SCF	-	2.68	1.33E-04	1.00
RI-MP2:SCF	-	2.55	2.13E-04	1.00
RI-MP2:RIJK-SCF	-	2.98	3.74E-06	1.00
RI-MP2:RICOSX-SCF	-	2.33	2.43E-04	1.00
RI-MP2:RICOSX2-SCF	-	2.35	3.33E-04	1.00
DLPNO-MP2	Normal	2.27	1.92E-03	1.00
DLPNO-MP2	Tight	2.68	1.62E-04	1.00
LMP2	Normal	2.78	3.88E-05	1.00
LMP2	Tight	2.70	5.99E-05	1.00
LMP2	vTight	2.96	1.43E-05	1.00
CAM-B3LYP (from PAi)	-	2.35	2.63E-05	0.95
CAM-B3LYP (from PDAi)	-	2.28	1.40E-03	0.99

A.1.0.1 Comparison between DLPNO-MP2 and DLPNO-MP2-α

We compare here the performance of analytic DLPNO-MP2 calculations for computing polarizabilities (DLPNO-MP2-α) with numerical (from 9 energy computations) DLPNO-MP2 calculations in Orca 5.

Table A.5: Wall time in minutes for approximate DLPNO-MP2 calculations tested on the PDA chains of different lengths. The values of DLPNO-MP2 have been multiplied for 9 in order to properly estimate the time necessary to compute NLOPs. These calculations have been performed employing 16 cpus. The values in the table are in minutes / 1 cpu.

	DLPNO-MP2		DLPNO-MP2-α	
Molecule	Normal	Tight	Normal	Tight
PDA1	90	72	191	55
PDA2	153	180	824	142
PDA3	324	378	267	311
PDA4	576	648	452	532
PDA5	873	1026	705	844

Figure A.4: Wall time for approximate DLPNO-MP2 calculations tested on the PDA chains of different lengths with respect to the number of electrons and the corresponding best fit. The values of DLPNO-MP2 have been multiplied for 9 in order to properly estimate the time necessary to compute NLOPs. Data computed using 16 CPUs in a single node.

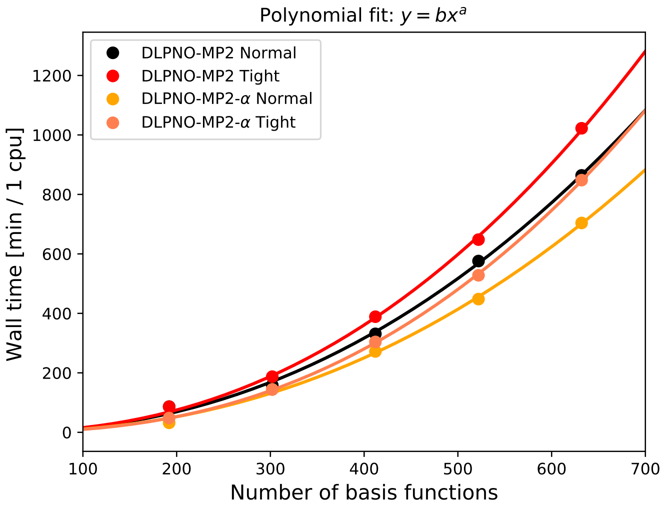


Table A.6: Parameters obtained from the best fits of the wall time with respect to the number of electrons of approximated MP2 calculations computed with 16 cpus for PDA chains.

Method	Approx	a	b [min /1 cpu]	r^2
DLPNO-MP2	Normal	2.20	2.28E-03	0.98
DLPNO-MP2	Tight	2.26	1.73E-03	0.99
DLPNO-MP2-α	Normal	2.25	1.29E-03	1.00
DLPNO-MP2-α	Tight	2.41	5.54E-04	1.00

A.2 Indicators of correlation

Table A.7: D1, D2 and T1 diagnostics calculated on the CCSD(T) wavefunction for each molecule of the set. In red are highlighted the molecules which are multireference according to the indicator proposed, namely: $T_1 \geq 0.002$, $D_1 \geq 0.05$ and $D_2 \geq 0.18$.

Molecule	**D1**	**D2**	**T1**	**Molecule**	**D1**	**D2**	**T1**
Carbon Monoxide	0.042	0.164	0.020	Pentane	0.017	0.146	0.008
Hydrogen fluoride	0.022	0.120	0.013	Cyclohexane	0.018	0.146	0.009
Dinitrogen	0.024	0.175	0.013	Hexane	0.017	0.147	0.008
Dioxygen	0.038	0.245	0.016	Heptane	0.017	0.147	0.008
Water	0.027	0.127	0.012	Octane	0.018	0.148	0.008
Carbon dioxide	0.051	0.156	0.020	H1	0.017	0.184	0.012
Cyanogen fluoride	0.031	0.180	0.016	H2	0.017	0.191	0.011
Hydrogen cyanide	0.028	0.185	0.015	H3	0.017	0.193	0.011
Hydrogen isocyanide	0.028	0.185	0.015	H4	0.017	0.195	0.011
Nitroxyl	0.045	0.213	0.018	H5	0.017	0.195	0.011
Dinitrogen oxide	0.049	0.192	0.021	H6	0.017	0.196	0.011
Ammonia	0.023	0.134	0.010	H7	0.017	0.196	0.011
Acetylene	0.027	0.188	0.013	H8	0.017	0.197	0.011
Formaldehyde	0.048	0.186	0.017	PMI1	0.034	0.197	0.013
Hydrogen Peroxide	0.030	0.183	0.015	PMI2	0.046	0.208	0.015
Nitrous acid	0.067	0.199	0.024	PMI3	0.058	0.213	0.016
Nitric acid	0.065	0.200	0.020	PMI4	0.070	0.215	0.017
Metane	0.011	0.135	0.007	PMI5	0.081	0.216	0.018
Carbonic acid	0.058	0.158	0.016	PMI6	0.090	0.217	0.019
Boric acid	0.048	0.129	0.014	PA2	0.028	0.197	0.010
Ethane	0.014	0.140	0.007	PA4	0.031	0.211	0.011
Propane	0.015	0.143	0.008	PA6	0.033	0.218	0.011
1-Butane	0.029	0.197	0.010	PA8	0.035	0.222	0.011
Benzene	0.024	0.191	0.010	PA10	0.036	0.225	0.011
Butane	0.016	0.144	0.008	PA12	0.037	0.227	0.011
1-Butanol	0.029	0.145	0.010	PDA1	0.033	0.208	0.011
Pentanal	0.051	0.179	0.013	PDA2	0.033	0.213	0.012
n-Buthylamine	0.027	0.145	0.009	PDA3	0.033	0.215	0.012
1-Penthanoic acid	0.058	0.163	0.014	PDA4	0.034	0.220	0.012
1-Pentanamide	0.059	0.163	0.013	PDA5	0.036	0.232	0.012

Table A.8: Indicators based on natural orbital occupancies on the MP2 wavefunction for each molecule of the set. In red are highlighted the molecules which are multireference according to the indicator proposed. The threshold values have been published for a CCSD wavefunction: $NON \geq 0.058$, $\bar{V} \geq 0.03$, $MRI \leq 0$, and $I_{ND} \geq 0.03$.

Molecule	NON	$\bar{V}$	MRI	I_{ND}
1-Butane	0.024	0.022	0.971	0.013
Water	0.013	0.020	0.997	0.012
Ammonia	0.011	0.021	0.997	0.012
Benzene	0.035	0.026	0.786	0.015
Butane	0.013	0.020	0.990	0.012
Pentanal	0.028	0.021	0.971	0.012
1-Pentanamide	0.027	0.021	0.974	0.012
1-Penthanoic acid	0.027	0.021	0.974	0.012
n-Buthylamine	0.015	0.021	0.987	0.012
1-Butanol	0.017	0.021	0.987	0.012
Acetylene	0.027	0.028	0.943	0.016
Ciclohexane	0.013	0.021	0.983	0.012
Carbon Monoxide	0.026	0.025	0.981	0.014
Carbon dioxide	0.031	0.024	0.967	0.013
Ethane	0.011	0.019	0.995	0.012
Cyanogen fluoride	0.033	0.025	0.926	0.014
H1	0.009	0.023	0.999	0.014
H2	0.012	0.025	0.999	0.015
Formaldehyde	0.029	0.023	0.977	0.013
Carbonic acid	0.027	0.021	0.983	0.012
Hydrogen Peroxide	0.030	0.023	0.971	0.013
H3	0.013	0.025	0.998	0.015
Boric acid	0.018	0.020	0.991	0.011
H4	0.014	0.026	0.997	0.016
H5	0.014	0.026	0.996	0.016
H6	0.014	0.026	0.996	0.016
H7	0.015	0.026	0.995	0.016
H8	0.015	0.026	0.994	0.016
Hydrogen cyanide	0.032	0.030	0.919	0.016
Heptane	0.014	0.021	0.982	0.012

Molecule	NON	$\bar{V}$	MRI	I_{ND}
Hexane	0.014	0.020	0.985	0.012
Hydrogen fluoride	0.012	0.017	0.998	0.009
Hydrogen isocyanide	0.032	0.030	0.919	0.016
Nitroxyl	0.038	0.026	0.916	0.014
Nitrous acid	0.034	0.024	0.941	0.013
Nitric acid	0.046	0.025	0.854	0.014
Metane	0.008	0.018	0.998	0.011
Dinitrogen	0.032	0.029	0.935	0.016
Dinitrogen oxide	0.048	0.031	0.603	0.017
Dioxygen	0.036	0.024	0.935	0.013
Octane	0.015	0.021	0.980	0.012
PA10	0.037	0.024	0.806	0.014
PA12	0.039	0.024	0.763	0.014
PA2	0.022	0.022	0.982	0.013
PA4	0.029	0.023	0.947	0.014
PA6	0.033	0.024	0.902	0.014
PA8	0.036	0.024	0.853	0.014
PDA1	0.035	0.025	0.887	0.014
PDA2	0.039	0.026	0.784	0.015
PDA3	0.041	0.026	0.695	0.015
PDA4	0.043	0.026	0.618	0.015
PDA5	0.043	0.027	0.550	0.015
Pentane	0.013	0.020	0.987	0.012
PMI1	0.028	0.024	0.972	0.014
PMI2	0.032	0.025	0.931	0.014
PMI3	0.034	0.025	0.892	0.014
PMI4	0.034	0.025	0.865	0.014
PMI5	0.034	0.025	0.847	0.014
PMI6	0.034	0.025	0.835	0.014
Propane	0.012	0.020	0.992	0.012

A.3 Numerical errors

Before verifying the accuracy of the NLOPs, the associated numerical errors were addressed. In Table A.9, we collect the MARoE and the ratio %MARoE of the Romberg error associated to each property for for the canonical and accelerated MP2 methods. We shall stress here that the Romberg error is not equal to the numerical error, but constitutes an estimation of its order of magnitude.

Table A.10: Mean Absolute Romberg Error (MARoE) and the ratio between MARoE and the mean value (%MARoE) obtained using the canonical and accelerated CCSD and CCSD(T) methods for molecules of the of γ-NLO set.

		DLPNO-CCSD	**DLPNO-CCSD(T0)**	**DLPNO-CCSD(T1)**	**LNO-CCSD**	**LNO-CCSD(T)**	**CCSD**	**CCSD(T)**
μ	MARoE	3.5E-04	2.2E-04	1.7E-04	1.2E-04	1.5E-04	1.1E-04	8.4E-05
	%MARoE	0	0	0	0	0	0	0
α	MARoE	3.5E-01	6.0E-01	3.0E-01	2.1E-01	3.8E-01	2.4E-01	2.5E-01
	%MARoE	0	1	0	0	0	0	0
β	MARoE	2.9E+02	7.0E+01	4.3E+01	3.8E+01	5.0E+01	3.3E+01	1.2E+01
	%MARoE	59	40	29	17	26	15	15
γ	MARoE	4.2E+05	6.6E+05	3.4E+05	1.7E+05	3.4E+05	3.1E+04	3.1E+04
	%MARoE	73	175	98	40	37	5	6

Table A.9: Mean Absolute Romberg Error (MARoE) and ratio between MARoE and the mean value (%MARoE) of γ-NLO set for the acceleration techniques of MP2 method. For the DLPNO-MP2-α method, the values of μ and α are computed analytically and therefore there is no associated RE.

		RI-MP2	**RIJK-MP2**	**RIJCOSX2-MP2**	**LMP2**	**DLPNO-MP2**	**DLPNO-MP2(α)**	**MP2**
μ	MARoE	2.3E-07	1.1E-07	6.6E-07	1.9E-04	1.5E-04	-	2.3E-07
	%MARoE	0	0	0	0	0	-	0
α	MARoE	7.4E-03	1.2E-03	2.7E-02	1.2E-01	4.0E-01	-	1.5E-03
	%MARoE	0	0	0	0	0	-	0
β	MARoE	5.9E-01	2.1E-01	1.3E+00	2.9E+01	4.8E+01	2.4E-01	6.2E-01
	%MARoE	0	0	1	16	18	0	0
γ	MARoE	2.2E+03	4.4E+03	1.6E+04	3.2E+05	4.0E+05	5.8E+03	2.3E+03
	%MARoE	0	1	3	87	80	1	0

In Table A.10, we collect the MARoE and %MARoE of CCSD and CCSD(T) methods and the corresponding approximations.

A.4 Impact of the different convergence criteria on the accuracy of localization schemes

Table A.11: Performance of localized CCSD methods with respect to CCSD(T) references for molecules of the γ-NLO set. MV, MAE , RMSE and MAX are in [a.u.]

		DLPNO-CCSD		**LNO-CCSD**			
		Normal	Tight	Normal	Tight	vTight	**CCSD**
μ	MV	7.2E-01	7.2E-01	7.2E-01	7.0E-01	7.2E-01	7.2E-01
	MARoE	4.8E-04	3.5E-04	5.0E-04	1.3E-04	1.2E-04	1.1E-04
	MAE	1.8E-02	2.1E-02	2.1E-02	2.2E-02	2.2E-02	2.2E-02
	RMSE	2.7E-02	3.1E-02	3.1E-02	3.1E-02	3.2E-02	3.4E-02
	MAX	7.1E-02	7.7E-02	8.9E-02	7.2E-02	8.2E-02	9.5E-02
	%MAE	3	3	3	3	3	3
α	MV	1.2E+02	1.1E+02	1.1E+02	1.1E+02	1.1E+02	1.1E+02
	MARoE	1.8E+00	3.5E-01	2.0E+00	5.4E-01	2.1E-01	2.5E-01
	MAE	1.7E+01	3.8E+00	6.9E+00	4.5E+00	3.6E+00	3.0E+00
	RMSE	7.1E+01	1.0E+01	2.2E+01	1.1E+01	8.5E+00	8.8E+00
	MAX	4.3E+02	6.1E+01	1.4E+02	5.5E+01	4.4E+01	5.9E+01
	%MAE	15	3	6	4	3	3
β	MV	7.5E+02	3.5E+02	3.7E+02	2.5E+02	2.2E+02	1.9E+02
	MARoE	3.7E+02	2.1E+02	9.9E+01	4.5E+01	3.8E+01	3.3E+01
	MAE	6.2E+02	2.3E+02	2.4E+02	1.2E+02	9.2E+01	5.1E+01
	RMSE	2.0E+03	3.6E+02	6.2E+02	3.1E+02	2.5E+02	1.3E+02
	MAX	7.9E+03	1.1E+03	2.3E+03	1.2E+03	9.2E+02	4.6E+02
	%MAE	445	165	172	90	67	37
γ	MV	4.9E+06	5.7E+05	1.5E+07	1.2E+06	4.3E+05	3.8E+05
	MARoE	3.1E+06	4.2E+05	6.7E+06	6.5E+05	1.7E+05	3.1E+04
	MAE	4.5E+06	5.3E+05	1.5E+07	7.2E+05	2.2E+05	8.4E+04
	RMSE	4.5E+07	1.9E+06	6.6E+07	5.3E+06	1.5E+06	6.0E+05
	MAX	2.1E+08	8.6E+06	3.9E+08	2.5E+07	6.8E+06	2.7E+06
	%MAE	992	117	3202	159	49	19

Table A.12: Performance of localized CCSD(T) methods with respect to CCSD(T) references for molecules of the γ-NLO set. MV, MAE , RMSE and MAX are in [a.u.]

		DLPNO-CCSD(T0)		**DLPNO-CCSD(T1)**		**LNO-CCSD(T)**		
		Normal	Tight	Normal	Tight	Normal	Tight	vTight
μ	MV	7.0E-01	7.0E-01	7.0E-01	7.0E-01	7.0E-01	7.0E-01	7.0E-01
	MARoE	3.5E-04	2.1E-04	4.1E-04	1.7E-04	3.2E-04	3.6E-04	1.5E-04
	MAE	7.3E-03	3.4E-03	9.6E-03	3.2E-03	6.7E-03	4.1E-03	2.3E-03
	RMSE	1.7E-02	6.6E-03	2.5E-02	7.5E-03	1.4E-02	8.9E-03	6.3E-03
	MAX	7.6E-02	2.6E-02	8.2E-02	2.6E-02	5.2E-02	2.9E-02	2.6E-02
	%MAE	1	1	1	1	1	1	0
α	MV	1.2E+02	1.1E+02	1.2E+02	1.2E+02	1.1E+02	1.2E+02	1.2E+02
	MARoE	1.3E+00	6.0E-01	4.9E-01	3.0E-01	3.1E+00	6.4E-01	3.9E-01
	MAE	1.2E+01	1.5E+00	2.9E+01	8.2E-01	7.3E+00	2.7E+00	1.2E+00
	RMSE	5.7E+01	4.2E+00	1.1E+02	2.1E+00	2.2E+01	7.7E+00	3.9E+00
	MAX	3.3E+02	2.0E+01	5.3E+02	1.0E+01	1.2E+02	4.0E+01	2.1E+01
	%MAE	10	1	25	1	6	2	1
β	MV	7.3E+02	1.7E+02	6.9E+02	1.5E+02	4.2E+02	2.6E+02	1.9E+02
	MARoE	3.7E+02	7.0E+01	3.7E+02	4.3E+01	1.6E+02	1.0E+02	5.0E+01
	MAE	5.9E+02	1.1E+02	5.7E+02	2.7E+01	2.9E+02	1.3E+02	6.4E+01
	RMSE	1.9E+03	3.1E+02	1.8E+03	4.8E+01	7.7E+02	3.2E+02	2.0E+02
	MAX	7.7E+03	1.3E+03	7.5E+03	1.9E+02	3.0E+03	1.1E+03	9.1E+02
	%MAE	433	78	414	19	209	92	46
γ	MV	1.1E+07	1.0E+06	3.7E+05	3.5E+05	6.9E+06	2.1E+06	9.1E+05
	MARoE	7.9E+06	6.6E+05	1.9E+05	3.4E+05	7.9E+06	1.7E+08	3.4E+05
	MAE	1.1E+07	7.0E+05	5.2E+05	2.3E+05	6.5E+06	1.8E+06	4.6E+05
	RMSE	6.2E+07	3.8E+06	1.8E+06	1.1E+06	3.5E+07	7.4E+06	2.2E+06
	MAX	4.4E+08	2.8E+07	1.0E+07	8.3E+06	2.7E+08	5.0E+07	1.5E+07
	%MAE	2416	155	115	51	1430	403	103

Table A.13: Performance of RI-MP2 methods with respect to CCSD(T) references for molecules of the γ-NLO set. MV, MAE , RMSE and MAX are in [a. u.]

		RI-MP2	**RI-MP2:RIJK**	**RIJCOSX-MP2**	**RIJCOSX2-MP2**	**MP2**
μ	MV	7.2E-01	7.2E-01	7.2E-01	7.2E-01	7.2E-01
	MARoE	2.3E-07	1.1E-07	1.0E-05	6.1E-07	2.3E-07
	MAE	2.9E-02	2.9E-02	3.0E-02	3.0E-02	2.9E-02
	RMSE	4.4E-02	4.4E-02	4.5E-02	4.4E-02	4.3E-02
	MAX	1.3E-01	1.3E-01	1.3E-01	1.3E-01	1.3E-01
	%MAE	4	4	4	4	4
α	MV	1.2E+02	1.2E+02	1.2E+02	1.2E+02	1.2E+02
	MARoE	7.4E-03	1.2E-03	2.1E-02	2.7E-02	1.5E-03
	MAE	3.7E+00	3.7E+00	5.0E+00	4.1E+00	3.7E+00
	RMSE	7.2E+00	7.3E+00	1.4E+01	9.1E+00	7.3E+00
	MAX	3.1E+01	3.1E+01	9.1E+01	4.6E+01	3.1E+01
	%MAE	3	3	4	4	3
β	MV	2.4E+02	2.4E+02	2.2E+02	2.4E+02	2.4E+02
	MARoE	5.9E-01	2.1E-01	1.7E+01	3.1E+00	6.2E-01
	MAE	1.1E+02	1.1E+02	9.2E+01	1.1E+02	1.1E+02
	RMSE	2.6E+02	2.7E+02	2.2E+02	2.8E+02	2.6E+02
	MAX	8.3E+02	9.0E+02	8.4E+02	1.0E+03	8.3E+02
	%MAE	77	80	67	81	78
γ	MV	6.0E+05	5.7E+05	5.6E+05	5.6E+05	5.2E+05
	MARoE	2.2E+03	4.4E+03	1.3E+04	1.6E+04	2.3E+03
	MAE	1.5E+05	1.3E+05	1.2E+05	1.1E+05	7.7E+04
	RMSE	6.0E+05	4.4E+05	4.2E+05	3.4E+05	2.4E+05
	MAX	3.4E+06	2.3E+06	2.3E+06	1.6E+06	1.2E+06
	%MAE	34	28	27	25	17

Table A.14: Performance of localized MP2 methods with respect to CCSD(T) references for molecules of the γ-NLO set. MV, MAE , RMSE and MAX are in [a. u.].

		LMP2			**DLPNO-MP2**		**DLPNO-MP2-α**	
		Normal	Tight	vTight	Normal	Tight	Normal	Tight
μ	MV	7.2E-01	7.2E-01	7.2E-01	7.2E-01	7.2E-01	7.2E-01	7.2E-01
	MARoE	1.9E-04	1.4E-05	1.9E-04	2.1E-04	1.5E-04	x	x
	MAE	3.0E-02	3.0E-02	2.9E-02	2.9E-02	2.9E-02	2.9E-02	2.9E-02
	RMSE	4.6E-02	4.5E-02	4.3E-02	4.3E-02	4.3E-02	4.4E-02	4.4E-02
	MAX	1.4E-01	1.3E-01	1.2E-01	1.3E-01	1.3E-01	1.3E-01	1.3E-01
	%MAE	4	4	4	4	4	4	4
α	MV	1.2E+02	1.2E+02	1.2E+02	1.2E+02	1.2E+02	1.2E+02	1.2E+02
	MARoE	2.0E+00	3.0E-01	1.2E-01	6.0E-01	4.0E-01	x	x
	MAE	4.4E+00	3.8E+00	3.4E+00	4.1E+00	4.3E+00	3.5E+00	3.6E+00
	RMSE	9.9E+00	7.8E+00	6.7E+00	9.2E+00	9.3E+00	6.8E+00	7.1E+00
	MAX	4.5E+01	3.4E+01	2.8E+01	4.8E+01	5.0E+01	2.9E+01	3.0E+01
	%MAE	4	3	3	4	4	3	3
β	MV	3.1E+02	1.5E+02	1.8E+02	2.2E+02	2.7E+02	2.3E+02	2.4E+02
	MARoE	4.2E+02	1.3E+01	2.9E+01	1.3E+02	4.8E+01	4.7E-01	2.4E-01
	MAE	1.8E+02	6.1E+01	5.7E+01	9.5E+01	1.5E+02	1.0E+02	1.1E+02
	RMSE	4.3E+02	1.3E+02	1.2E+02	2.4E+02	4.0E+02	2.5E+02	2.6E+02
	MAX	1.7E+03	4.4E+02	3.6E+02	1.1E+03	1.7E+03	8.2E+02	8.7E+02
	%MAE	133	44	41	69	105	74	78
γ	MV	1.2E+06	1.2E+06	3.7E+05	4.4E+05	5.0E+05	5.7E+05	5.5E+05
	MARoE	6.0E+06	8.6E+05	3.2E+05	6.2E+05	4.0E+05	5.6E+03	5.8E+03
	MAE	7.9E+05	7.6E+05	2.6E+05	1.9E+05	1.5E+05	1.3E+05	1.0E+05
	RMSE	3.4E+06	4.2E+06	1.2E+06	6.6E+05	3.8E+05	4.5E+05	3.7E+05
	MAX	2.0E+07	3.2E+07	8.9E+06	4.5E+06	1.9E+06	2.2E+06	2.1E+06
	%MAE	174	166	58	42	32	28	23

A.5 Performances of accelerated methods on the β-Set

Table A.15: Performance of DLPNO-CCSD(T1), MP2, RI-MP2 and CCSD methods for the β-NLO set with respect to the reference CCSD(T). MV, MAE , RMSE and MAX are in [a. u.].

		DLPNO-CCSD(T1)	RI-MP2	MP2	CCSD
μ	MV	3.3E+00	3.4E+00	3.4E+00	3.3E+00
	MARoE	1.9E-03	2.8E-05	1.7E-08	5.0E-06
	MAE	2.3E-02	1.0E-01	1.0E-01	5.6E-02
	RMSE	3.1E-02	1.3E-01	1.2E-01	6.8E-02
	MAX	9.8E-02	2.9E-01	2.9E-01	1.4E-01
	%MAE	1	3	3	2
α	MV	3.5E+02	3.6E+02	3.6E+02	3.3E+02
	MARoE	6.0E+00	6.0E-02	9.9E-05	7.3E-03
	MAE	1.0E+01	1.7E+01	1.7E+01	1.3E+01
	RMSE	2.3E+01	2.7E+01	2.7E+01	1.7E+01
	MAX	1.2E+02	9.5E+01	9.4E+01	2.8E+01
	%MAE	3	5	5	4
β	MV	1.0E+04	1.2E+04	1.2E+04	9.2E+03
	MARoE	6.6E+03	1.0E+02	4.7E-01	1.8E+01
	MAE	1.7E+03	2.8E+03	2.7E+03	7.7E+02
	RMSE	2.6E+03	4.5E+03	4.5E+03	1.2E+03
	MAX	6.8E+03	1.3E+04	1.3E+04	4.5E+03
	%MAE	17	28	27	8
γ	MV	1.2E+07	2.9E+06	2.9E+06	2.3E+06
	MARoE	2.3E+04	1.1E+05	1.6E+03	5.6E+04
	MAE	9.6E+06	5.6E+05	5.2E+05	3.9E+05
	RMSE	2.9E+07	9.8E+05	8.7E+05	7.0E+05
	MAX	1.3E+08	3.1E+06	2.9E+06	3.2E+06
	%MAE	355	20	21	14

A.6 Performances of the accelerated methods merging the β-NLO and γ-NLO-B sets.

Table A.16: Performance of DLPNO-CCSD(T1), MP2, RI-MP2 and CCSD methods for β merging the β-NLO and γ-NLO-B set with respect to the reference CCSD(T). MV, MAE , RMSE and MAX are in [a. u.].

		DLPNO-CCSD(T1)	**RI-MP2**	**MP2**	**CCSD**
	MV	7.1E+03	8.5E+03	8.5E+03	6.6E+03
	MAE	1.2E+03	2.0E+03	2.0E+03	5.6E+02
β	RMSE	2.2E+03	3.8E+03	3.8E+03	1.0E+03
	MAX	6.8E+03	1.3E+04	1.3E+04	4.5E+03
	%MAE	17	28	27	8

Appendix B

Supporting information of Chapter 5

B.1 T1 diagnostics for molecules of the β-set

Table B.1: Values of T1 diagnostics for molecules of the β-set computed on the CCSD(T) wavefunctions. The system is considered multireference if the value of the T1 diagnostic is larger than 0.02.

Molecule	T1	Molecule	T1
PA2-NH$_2$-CHO	0.017	PMI4-OH-NO$_2$	0.021
PA2-NH$_2$-NO$_2$	0.020	PMI5-NH$_2$-CHO	0.020
PA2-OH-CHO	0.017	PMI5-OH-CHO	0.020
PA2-OH-NO$_2$	0.020	PMI5-OH-NO$_2$	0.022
PA4-NH$_2$-CHO	0.016	PA10-NH$_2$-CHO	0.015
PA4-NH$_2$-NO2	0.020	PA10-NH$_2$-NO$_2$	0.018
PA4-OH-CHO	0.016	PA10-OH-CHO	0.015
PA4-OH-NO$_2$	0.019	PA10-OH-NO$_2$	0.018
PA6-NH$_2$-CHO	0.016	PA12-NH$_2$-CHO	0.015
PA6-OH-CHO	0.016	PA12-NH$_2$-NO$_2$	0.017
PDA1-CHO-OH	0.017	PA12-OH-CHO	0.015
PDA1-NH$_2$-CHO	0.016	PA12-OH-NO$_2$	0.017
PDA1-OH-NO$_2$	0.019	PA6-OH-NO$_2$	0.019
PMI2-NH$_2$-CHO	0.019	PA8-NH$_2$-CHO	0.016
PMI2-NH$_2$-NO$_2$	0.021	PA8-NH$_2$-NO$_2$	0.018
PMI2-OH-CHO	0.019	PDA2-NH$_2$-CHO	0.015
PMI2-OH-NO$_2$	0.021	PDA2-NH$_2$-NO$_2$	0.018
PMI3-NH$_2$-CHO	0.019	PDA2-OH-CHO	0.015
PMI3-NH$_2$-NO$_2$	0.021	PDA2-OH-NO$_2$	0.018
PMI3-OH-CHO	0.019	PDA3-NH$_2$-CHO	0.015
PMI3-OH-NO$_2$	0.021	PDA3-NH$_2$-NO$_2$	0.017
PMI4-OH-CHO	0.019	PDA3-OH-CHO	0.015
PA6-NH2-NO$_2$	0.019	PDA3-OH-NO$_2$	0.017
PA8-OH-CHO	0.015	PMI5-NH$_2$-NO$_2$	0.022
PA8-OH-NO$_2$	0.018	PMI6-NH$_2$-CHO	0.020
PDA1-NH$_2$-NO$_2$	0.019	PMI6-NH$_2$-NO$_2$	0.022
PMI4-NH$_2$-CHO	0.020	PMI6-OH-CHO	0.020
PMI4-NH$_2$-NO$_2$	0.021	PMI6-OH-NO$_2$	0.022

B.2 Evaluation of spin contamination in the unrestricted framework

Table B.2: $\langle \hat{S}^2 \rangle$ values calculated for molecules of the γ-NLO-set using unrestricted HF and DFT methods with different exchange-correlation functionals.

Molecule	Functional	$\langle \hat{S}^2 \rangle$	Molecule	Functional	$\langle \hat{S}^2 \rangle$
O_2	UHF	2.0	PA10	UHF	1.3
HNO	UHF	0.5	PDA3	UHF	1.6
N_2O	UHF	0.3	PA12	UHF	1.6
HNO_2	UHF	0.2	PDA4	UHF	2.2
HNO3	UHF	0.2	PDA5	UHF	2.7
PMI1	UHF	0.1	O_2	UCAM-B3LYP	2.0
PMI2	UHF	0.4	HNO	UCAM-B3LYP	0.2
PMI3	UHF	0.7	O_2	UB3LYP	2.0
Benzene	UHF	0.5	HNO	UB3LYP	0.2
PDA1	UHF	0.5	O_2	UBLYP	2.0
PA6	UHF	0.7	HNO	UBLYP	0.2
PMI4	UHF	1.0	O_2	UωB97X-D	2.0
PMI5	UHF	1.4	HNO	UωB97X-D	0.1
PA8	UHF	1.0	O_2	UωB97M-D	2.0
PDA2	UHF	1.1	O_2	UωB97X-D	2.0
PMI6	UHF	1.7	HNO	UωB97X-V	0.0

Table B.3: $\langle \hat{S}^2 \rangle$ values calculated for molecules of the β-NLO-set using the unrestricted HF method.

Molecule	Functional	$\langle \hat{S}^2 \rangle$	Molecule	Functional	$\langle \hat{S}^2 \rangle$
PA2-OH-CHO	UHF	0.1	PA8-OH-CHO	UHF	1.4
PA2-OH-NO_2	UHF	0.5	PA8-OH-NO_2	UHF	1.7
PA2-NH_2-NO_2	UHF	0.4	PDA2-OH-CHO	UHF	1.8
PMI2-OH-CHO	UHF	0.0	PDA2-OH-NO_2	UHF	2.1
PMI2-OH-NO_2	UHF	0.5	PMI5-NH_2-CHO	UHF	0.9
PMI2-NH_2-NO_2	UHF	0.4	PMI5-NH_2-NO_2	UHF	1.2
PA4-OH-CHO	UHF	0.6	PA8-NH_2-CHO	UHF	1.4
PA4-OH-NO2	UHF	0.9	PA8-NH_2-NO_2	UHF	1.6
PA4-NH_2-CHO	UHF	0.6	PDA2-NH_2-CHO	UHF	1.8
PA4-NH_2-NO_2	UHF	0.9	PDA2-NH_2-NO_2	UHF	2.1
PMI3-OH-CHO	UHF	0.5	PMI6-OH-CHO	UHF	1.6
PMI3-OH-NO_2	UHF	0.8	PMI6-OH-NO_2	UHF	1.8
PDA1-CHO-OH	UHF	1.1	PMI6-NH_2-CHO	UHF	1.3
PDA1-OH-NO_2	UHF	1.3	PMI6-NH_2-NO_2	UHF	1.6
PMI3-NH_2-NO_2	UHF	0.4	PA10-OH-CHO	UHF	1.7
PDA1-NH_2-CHO	UHF	1.0	PA10-OH-NO_2	UHF	2.0
PDA1-NH_2-NO_2	UHF	1.3	PA10-NH_2-CHO	UHF	1.7
PA6-OH-CHO	UHF	1.0	PA10-NH_2-NO_2	UHF	2.0
PA6-OH-NO_2	UHF	1.3	PDA3-OH-CHO	UHF	2.5
PMI4-OH-CHO	UHF	0.9	PDA3-OH-NO_2	UHF	2.8
PMI4-OH-NO_2	UHF	1.2	PDA3-NH_2-CHO	UHF	2.5
PA6-NH_2-CHO	UHF	1.0	PDA3-NH_2-NO_2	UHF	2.8
PA6-NH_2-NO_2	UHF	1.3	PA12-OH-CHO	UHF	2.1
PMI4-NH_2-CHO	UHF	0.5	PA12-OH-NO_2	UHF	2.4
PMI4-NH_2-NO_2	UHF	0.8	PA12-NH_2-CHO	UHF	2.0
PMI5-OH-CHO	UHF	1.2	PA12-NH_2-NO_2	UHF	2.3
PMI5-OH-NO_2	UHF	1.5			

B.3 Comparison between MP3:HF and RI-MP3:HF

Table B.4: Performance of MP3:HF and RI-MP3:HF methods for calculating NLO properties with respect to the reference CCSD(T) values, for molecules of the γ-NLO-set. Units are a. u.

		MP3	**RI-MP3**
Ref. orbitals		**HF**	**HF**
	MV	7.3E-01	7.3E-01
	MAE	2.4E-02	2.4E-02
μ	RMSD	3.6E-02	3.6E-02
	MAX	9.4E-02	9.4E-02
	%MAE	3	3
	MV	9.3E+01	9.3E+01
	MAE	2.3E+00	2.3E+00
α	RMSD	4.2E+00	4.1E+00
	MAX	1.8E+01	1.8E+01
	%MAE	2	2
	MV	2.2E+02	2.2E+02
	MAE	8.9E+01	8.6E+01
β	RMSD	2.1E+02	2.0E+02
	MAX	6.6E+02	6.1E+02
	%MAE	65	62
	MV	1.9E+05	1.9E+05
	MAE	2.9E+04	2.9E+04
γ	RMSD	9.5E+04	1.0E+05
	MAX	5.0E+05	5.6E+05
	%MAE	6	6

B.4 Performance of MP3:KS for computing dipole moments

Table B.5: Performances of MP3:KS methods for the calculation of the dipole moment with respect to reference CCSD(T) values, for molecules from the joint γ-NLO-B and β-NLO sets. Units are a.u.

		MP2	**MP3**	**CCSD**			**MP3:KS**			
Ref.	**orbitals**	**HF**	**HF**	**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
μ	MV	2.6E+00	2.6E+00	2.6E+00	2.7E+00	2.7E+00	2.7E+00	2.7E+00	2.7E+00	2.7E+00
	MAE	8.0E-02	7.2E-02	8.0E-02	1.4E-01	1.3E-01	1.3E-01	1.3E-01	1.3E-01	1.3E-01
	RMSD	1.1E-01	1.0E-01	1.1E-01	1.7E-01	1.7E-01	1.6E-01	1.6E-01	1.6E-01	1.6E-01
	MAX	2.9E-01	3.1E-01	2.9E-01	4.3E-01	4.2E-01	3.9E-01	3.9E-01	3.9E-01	4.0E-01
	%MAE	3	3	3	5	5	5	5	5	5

B.5 Performance of MP3:KS for computing NLO properties

Table B.6: Performance of MP3:KS methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules from the γ-NLO set. Units are a.u.

		MP2	**MP3**	**CCSD**	**MP3:KS**					
Ref. orbitals		**HF**	**HF**	**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
μ	MV	7.2E-01	7.3E-01	7.2E-01	7.5E-01	7.4E-01	7.4E-01	7.4E-01	7.4E-01	7.5E-01
	MAE	3.0E-02	2.0E-02	2.0E-02	5.0E-02	4.0E-02	4.0E-02	4.0E-02	4.0E-02	4.0E-02
	RMSD	4.0E-02	4.0E-02	3.0E-02	8.0E-02	7.0E-02	7.0E-02	6.0E-02	7.0E-02	7.0E-02
	MAX	1.0E-01	9.0E-02	9.0E-02	2.0E-01	2.0E-01	2.0E-01	2.0E-01	2.0E-01	2.0E-01
	%MAE	4	3	3	7	6	6	6	6	6
α	MV	1.2E+02	1.1E+02	1.1E+02	1.1E+02	1.1E+02	1.2E+02	1.2E+02	1.2E+02	1.2E+02
	MAE	3.7E+00	3.4E+00	3.0E+00	8.4E+00	3.5E+00	1.5E+00	1.5E+00	1.5E+00	1.3E+00
	RMSD	1.1E+01	1.2E+01	1.4E+01	6.2E+01	2.3E+01	2.9E+00	2.9E+00	2.4E+00	2.3E+00
	MAX	3.1E+01	4.3E+01	5.9E+01	2.7E+02	9.9E+01	7.3E+00	7.5E+00	8.3E+00	5.1E+00
	%MAE	3	3	3	7	3	1	1	1	1
β	MV	2.4E+02	2.2E+02	1.8E+02	2.3E+02	2.0E+02	1.1E+02	1.1E+02	9.7E+01	1.0E+02
	MAE	1.1E+02	8.9E+01	5.0E+01	1.0E+02	6.7E+01	2.6E+01	2.8E+01	4.3E+01	3.9E+01
	RMSD	2.7E+02	2.1E+02	1.3E+02	3.0E+02	2.2E+02	7.3E+01	7.6E+01	1.2E+02	1.1E+02
	MAX	9.1E+02	6.6E+02	4.7E+02	1.2E+03	9.4E+02	3.1E+02	3.3E+02	5.2E+02	4.7E+02
	%MAE	80	65	36	72	48	19	20	31	28
γ	MV	5.2E+05	4.2E+05	3.8E+05	8.8E+05	3.8E+05	8.8E+04	4.6E+05	4.5E+05	4.5E+05
	MAE	8.2E+04	6.4E+04	8.4E+04	4.4E+05	1.6E+05	5.0E+04	4.9E+04	3.8E+04	4.1E+04
	RMSD	2.6E+05	2.8E+05	3.7E+05	1.8E+06	7.9E+05	1.6E+05	1.6E+05	1.3E+05	1.5E+05
	MAX	1.1E+06	2.0E+06	2.7E+06	1.2E+07	5.9E+06	7.5E+05	9.6E+05	8.7E+05	9.9E+05
	%MAE	18	14	19	98	36	11	11	8	9

Table B.7: Performance of MP3:KS methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules from the β-NLO set. Units are a.u.

		MP2	**MP3**	**CCSD**	**MP3:KS**					
Ref. orbitals		**HF**	**HF**	**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
μ	MV	3.4E+00	3.4E+00	3.3E+00	3.5E+00	3.5E+00	3.5E+00	3.5E+00	3.5E+00	3.5E+00
	MAE	1.0E-01	9.2E-02	5.6E-02	1.8E-01	1.7E-01	1.6E-01	1.6E-01	1.6E-01	1.7E-01
	RMSD	1.2E-01	1.2E-01	6.7E-02	2.0E-01	2.0E-01	1.8E-01	1.8E-01	1.8E-01	1.9E-01
	MAX	2.9E-01	3.1E-01	1.4E-01	4.3E-01	4.2E-01	3.9E-01	3.9E-01	3.9E-01	4.0E-01
	%MAE	3	3	2	5	5	5	5	5	5
α	MV	3.6E+02	3.4E+02	3.3E+02	3.7E+02	3.6E+02	3.5E+02	3.5E+02	3.5E+02	3.5E+02
	MAE	1.7E+01	7.5E+00	1.3E+01	2.2E+01	1.4E+01	7.7E+00	1.0E+01	5.6E+00	5.7E+00
	RMSD	2.7E+01	1.2E+01	1.7E+01	3.5E+01	2.2E+01	1.1E+01	1.7E+01	7.0E+00	7.1E+00
	MAX	9.4E+01	3.8E+01	5.1E+01	1.2E+02	7.8E+01	3.4E+01	8.5E+01	1.8E+01	1.9E+01
	%MAE	5	2	4	6	4	2	3	2	2
β	MV	1.2E+04	1.0E+04	9.2E+03	2.2E+04	1.6E+04	1.2E+04	1.1E+04	1.0E+04	1.0E+04
	MAE	2.8E+03	1.8E+03	7.7E+02	1.3E+04	7.0E+03	2.6E+03	2.1E+03	1.3E+03	1.3E+03
	RMSD	4.5E+03	2.7E+03	1.2E+03	2.7E+04	1.3E+04	4.1E+03	3.2E+03	1.8E+03	1.9E+03
	MAX	1.3E+04	7.8E+03	4.5E+03	1.2E+05	5.3E+04	1.3E+04	9.5E+03	4.9E+03	5.3E+03
	%MAE	28	18	8	131	70	26	21	13	13
γ	MV	2.9E+06	2.5E+06	2.3E+06	6.6E+06	5.1E+06	3.3E+06	3.1E+06	2.7E+06	2.8E+06
	MAE	5.2E+05	3.2E+05	3.9E+05	4.0E+06	2.5E+06	8.0E+05	5.9E+05	2.7E+05	3.0E+05
	RMSD	8.7E+05	4.8E+05	7.0E+05	1.0E+07	6.3E+06	1.6E+06	1.1E+06	4.1E+05	4.9E+05
	MAX	2.9E+06	1.6E+06	3.2E+06	4.7E+07	3.2E+07	6.5E+06	4.2E+06	1.4E+06	2.0E+06
	%MAE	19	12	14	147	94	30	22	10	11

B.6 Performance of MP2:KS for computing NLO properties

Table B.8: Performance of MP2:KS methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules from both the γ-NLO and β-NLO sets. Units are a.u.

		MP2	**MP2:KS**					
Ref. orbitals		**HF**	**BLYP**	**B3LYP**	**CAM**	ω**XD**	ω**XV**	ω**MV**
μ	MV	2.6E+00	9.7E-01	7.3E-01	7.4E-01	7.4E-01	7.4E-01	7.5E-01
	MAE	8.0E-02	6.6E-01	5.4E-01	4.7E-01	4.5E-01	4.3E-01	4.4E-01
	RMSD	1.1E-01	9.6E-01	7.1E-01	6.5E-01	6.4E-01	6.2E-01	6.3E-01
	MAX	2.9E-01	4.4E+00	2.8E+00	2.9E+00	2.9E+00	2.9E+00	2.8E+00
	%MAE	3	68	73	63	61	58	58
α	MV	2.4E+02	1.4E+02	1.3E+02	1.3E+02	1.3E+02	1.9E+02	1.1E+02
	MAE	1.0E+01	8.2E+01	6.2E+01	4.2E+01	3.8E+01	6.7E+01	3.8E+01
	RMSD	2.0E+01	1.5E+02	1.1E+02	7.1E+01	6.4E+01	2.7E+02	6.8E+01
	MAX	9.4E+01	6.9E+00	5.1E+00	4.1E+00	4.0E+00	3.6E+00	3.8E+00
	%MAE	4	71	53	36	33	58	33
β	MV	1.7E+06	1.9E+04	1.4E+04	1.1E+04	1.0E+04	9.4E+03	9.5E+03
	MAE	2.9E+05	1.3E+04	8.2E+03	4.7E+03	3.7E+03	2.8E+03	2.9E+03
	RMSD	6.3E+05	2.8E+04	2.0E+04	1.0E+04	7.9E+03	5.7E+03	6.0E+03
	MAX	2.9E+06	1.4E+05	1.1E+05	4.7E+04	3.4E+04	2.3E+04	2.4E+04
	%MAE	19	176	114	65	52	39	41
γ	MV	8.5E+03	5.1E+06	3.8E+06	2.6E+06	3.0E+06	2.2E+06	1.8E+06
	MAE	2.0E+03	4.0E+06	2.3E+06	1.1E+06	1.6E+06	7.1E+05	6.9E+05
	RMSD	4.5E+03	1.9E+07	8.3E+06	3.5E+06	4.8E+06	2.3E+06	2.0E+06
	MAX	1.3E+04	1.1E+04	1.0E+03	2.7E+03	3.3E+03	2.8E+03	3.1E+03
	%MAE	28	262	150	73	104	46	45

Table B.9: Performance of MP2:KS methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules of the γ-NLO set. Units are a.u.

		MP2	**MP2:KS**					
Ref. orbitals		**HF**	**BLYP**	**B3LYP**	**CAM**	ω**XD**	ω**XV**	ω**MV**
μ	MV	7.0E-01	9.7E-01	7.3E-01	7.4E-01	7.4E-01	7.4E-01	7.5E-01
	MAE	3.0E-02	2.7E-01	4.5E-02	4.9E-02	4.9E-02	5.0E-02	4.4E-02
	RMSD	4.0E-02	5.0E-01	6.7E-02	7.9E-02	7.8E-02	8.2E-02	7.3E-02
	MAX	1.0E-01	1.8E+00	1.8E-01	2.5E-01	2.5E-01	2.6E-01	2.3E-01
	%MAE	4	38	6	7	7	7	6
α	MV	1.2E+02	1.5E+02	1.4E+02	1.3E+02	1.3E+02	1.8E+02	1.1E+02
	MAE	3.7E+00	2.9E+01	2.2E+01	1.5E+01	1.4E+01	6.5E+01	1.8E+01
	RMSD	1.1E+01	1.5E+02	1.0E+02	6.2E+01	5.7E+01	3.5E+02	9.8E+01
	MAX	3.1E+01	5.5E+02	3.7E+02	2.1E+02	1.9E+02	1.7E+03	3.7E+02
	%MAE	3	25	19	13	12	56	15
β	MV	2.4E+02	7.7E+02	3.4E+02	6.5E+01	3.7E+01	5.8E+01	2.4E+02
	MAE	1.1E+02	6.3E+02	2.0E+02	8.9E+01	1.2E+02	9.3E+01	1.1E+02
	RMSD	2.7E+02	2.2E+03	7.1E+02	2.5E+02	3.4E+02	2.7E+02	2.6E+02
	MAX	9.1E+02	1.0E+04	3.2E+03	1.0E+03	1.5E+03	1.1E+03	8.3E+02
	%MAE	80	456	147	65	84	67	78
γ	MV	5.2E+05	4.0E+06	2.0E+06	1.0E+04	9.0E+05	7.0E+05	1.5E+05
	MAE	8.2E+04	3.5E+06	1.6E+06	5.8E+05	4.4E+05	2.8E+05	3.6E+05
	RMSD	2.6E+05	2.1E+07	8.5E+06	2.7E+06	2.0E+06	1.2E+06	1.6E+06
	MAX	1.1E+06	1.5E+08	6.1E+07	1.8E+07	1.3E+07	7.3E+06	1.1E+07
	%MAE	18	776	363	129	98	62	80

Table B.10: Performance of MP2:KS methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules of the β-NLO set. Units are in a.u.

		MP2	**MP2:KS**					
Ref. orbitals		**HF**	**BLYP**	**B3LYP**	**CAM**	ω**XD**	ω**XV**	ω**MV**
μ	MV	3.4E+00	3.8E+00	3.8E+00	3.7E+00	3.7E+00	3.6E+00	3.6E+00
	MAE	1.0E-01	5.3E-01	4.6E-01	3.6E-01	3.3E-01	3.0E-01	3.1E-01
	RMSD	1.2E-01	6.7E-01	5.5E-01	4.2E-01	3.9E-01	3.5E-01	3.6E-01
	MAX	2.9E-01	1.4E+00	1.2E+00	9.1E-01	8.4E-01	7.4E-01	7.5E-01
	%MAE	3	16	14	11	10	9	9
α	MV	3.6E+02	4.9E+02	4.5E+02	4.2E+02	4.1E+02	4.0E+02	4.0E+02
	MAE	1.7E+01	1.4E+02	1.0E+02	6.9E+01	6.0E+01	5.1E+01	5.3E+01
	RMSD	2.7E+01	2.0E+02	1.5E+02	9.6E+01	8.5E+01	7.1E+01	7.3E+01
	MAX	9.4E+01	5.9E+02	4.1E+02	2.5E+02	2.2E+02	1.9E+02	1.9E+02
	%MAE	5	42	30	20	17	15	15
β	MV	1.2E+04	2.4E+04	2.0E+04	1.6E+04	1.4E+04	1.3E+04	1.3E+04
	MAE	2.8E+03	1.5E+04	1.1E+04	6.4E+03	5.2E+03	3.9E+03	4.1E+03
	RMSD	4.5E+03	3.1E+04	2.3E+04	1.2E+04	9.3E+03	6.8E+03	7.1E+03
	MAX	1.3E+04	1.4E+05	1.1E+05	4.7E+04	3.4E+04	2.3E+04	2.4E+04
	%MAE	28	149	112	64	52	39	41
γ	MV	2.9E+06	6.1E+06	5.6E+06	4.4E+06	5.3E+06	3.8E+06	3.7E+06
	MAE	5.2E+05	4.0E+06	3.1E+06	1.8E+06	3.0E+06	1.1E+06	1.0E+06
	RMSD	8.7E+05	8.9E+06	8.3E+06	4.3E+06	6.8E+06	3.0E+06	2.3E+06
	MAX	2.9E+06	3.2E+07	4.0E+07	2.0E+07	2.7E+07	1.8E+07	9.4E+06
	%MAE	19	148	113	65	109	42	38

B.7 Performance of DFT

Table B.11: Performance of HF and DFT methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules from the γ-NLO and β-NLO sets. Units are a.u.

Ref. orbitals		**HF**	**BLYP**	**B3LYP**	**CAM**	**ωXD**	**ωXV**	**ωMV**
μ	MV	3.2E+00	3.2E+00	3.2E+00	3.0E+00	3.0E+00	3.0E+00	3.9E+00
	MAE	6.5E-01	7.9E-01	6.8E-01	6.7E-01	6.1E-01	6.1E-01	6.1E-01
	RMSD	8.0E-01	9.6E-01	8.4E-01	8.3E-01	7.8E-01	7.9E-01	7.9E-01
	MAX	1.8E+00	3.0E+00	3.2E+00	3.2E+00	3.3E+00	3.2E+00	3.2E+00
	%MAE	25	31	27	26	24	24	21
α	MV	2.3E+02	2.1E+02	2.2E+02	2.3E+02	2.3E+02	2.3E+02	2.3E+02
	MAE	2.2E+01	7.2E+01	4.5E+01	2.3E+01	2.2E+01	2.0E+01	2.7E+01
	RMSD	3.5E+01	1.4E+02	8.9E+01	4.3E+01	4.0E+01	3.3E+01	4.8E+01
	MAX	4.9E+00	9.8E+00	7.5E+00	6.7E+00	6.7E+00	6.4E+00	6.5E+00
	%MAE	10	32	20	10	10	9	12
β	MV	3.5E+03	9.0E+03	8.3E+03	6.9E+03	6.4E+03	5.8E+03	5.9E+03
	MAE	3.7E+03	4.5E+03	4.1E+03	2.3E+03	2.0E+03	1.7E+03	1.7E+03
	RMSD	5.5E+03	1.1E+04	9.2E+03	4.2E+03	3.4E+03	2.7E+03	2.7E+03
	MAX	1.9E+04	5.8E+04	5.1E+04	2.0E+04	1.3E+04	9.5E+03	9.6E+03
	%MAE	52	63	57	32	28	24	24
γ	MV	7.4E+05	2.8E+06	2.2E+06	1.6E+06	1.4E+06	1.3E+06	1.4E+06
	MAE	8.0E+05	2.1E+06	1.2E+06	5.1E+05	4.2E+05	3.5E+05	4.7E+05
	RMSD	1.7E+06	9.2E+06	4.4E+06	1.3E+06	9.2E+05	7.4E+05	1.3E+06
	MAX	1.4E+04	3.3E+03	8.4E+03	9.3E+03	9.8E+03	9.5E+03	9.3E+03
	%MAE	52	134	80	33	27	23	30

Table B.12: Performance of HF and DFT methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules of the γ-NLO set. Units are a.u.

Ref. orbitals		**HF**	**BLYP**	**B3LYP**	**CAM**	ω**XD**	ω**XV**	ω**MV**
μ	MV	9.0E-01	7.0E-01	8.0E-01	8.0E-01	8.0E-01	8.0E-01	8.0E-01
	MAE	2.4E-01	3.1E-02	6.3E-02	9.2E-02	9.6E-02	9.9E-02	1.2E-01
	RMSD	4.2E-01	4.3E-02	1.2E-01	1.7E-01	1.8E-01	2.0E-01	2.4E-01
	MAX	1.4E+00	1.2E-01	4.2E-01	5.9E-01	6.3E-01	6.7E-01	7.4E-01
	%MAE	34	4	9	13	14	14	17
α	MV	1.2E+02	1.7E+02	1.5E+02	1.3E+02	1.3E+02	1.3E+02	1.4E+02
	MAE	9.4E+00	5.2E+01	3.4E+01	1.9E+01	1.7E+01	1.2E+01	2.9E+01
	RMSD	3.1E+01	2.3E+02	1.5E+02	7.4E+01	6.8E+01	4.4E+01	1.4E+02
	MAX	8.0E+01	8.7E+02	5.4E+02	2.5E+02	2.3E+02	1.4E+02	5.4E+02
	%MAE	8	44	30	16	15	10	25
β	MV	2.4E+02	9.0E+02	5.5E+02	2.2E+02	1.9E+02	1.2E+02	1.4E+02
	MAE	1.1E+02	7.6E+02	4.1E+02	8.4E+01	5.8E+01	4.9E+01	5.1E+01
	RMSD	2.4E+02	2.4E+03	1.2E+03	2.0E+02	1.3E+02	1.6E+02	1.4E+02
	MAX	7.5E+02	1.0E+04	5.2E+03	6.3E+02	4.7E+02	7.4E+02	6.0E+02
	%MAE	76	552	298	61	42	35	37
γ	MV	2.9E+05	2.8E+06	1.6E+06	7.1E+05	6.2E+05	5.2E+05	6.3E+05
	MAE	1.6E+05	2.4E+06	1.1E+06	2.6E+05	1.7E+05	7.1E+04	2.1E+05
	RMSD	1.6E+05	1.2E+07	5.5E+06	1.2E+06	7.3E+05	2.3E+05	1.0E+06
	MAX	4.5E+06	9.1E+07	3.9E+07	7.7E+06	4.7E+06	1.3E+06	7.9E+06
	%MAE	35	527	244	57	37	16	46

Table B.13: Performance of HF and DFT methods for the calculation of the NLO properties with respect to reference CCSD(T) values, for molecules of the β-NLO set. Units are a.u.

Ref. orbitals		**HF**	**BLYP**	**B3LYP**	**CAM**	ω**XD**	ω**XV**	ω**MV**
μ	MV	4.1E+00	4.2E+00	4.1E+00	4.0E+00	3.9E+00	3.9E+00	3.9E+00
	MAE	8.2E-01	8.8E-01	8.0E-01	6.3E-01	6.2E-01	5.3E-01	5.3E-01
	RMSD	9.1E-01	1.1E+00	9.1E-01	6.9E-01	6.7E-01	5.7E-01	5.7E-01
	MAX	1.8E+00	2.4E+00	1.9E+00	1.2E+00	1.2E+00	1.0E+00	9.7E-01
	%MAE	25	26	24	19	19	16	16
α	MV	3.5E+02	2.5E+02	2.9E+02	3.3E+02	3.3E+02	3.4E+02	3.4E+02
	MAE	3.5E+01	9.4E+01	5.6E+01	2.8E+01	2.7E+01	2.9E+01	2.9E+01
	RMSD	4.6E+01	1.4E+02	8.2E+01	3.8E+01	3.7E+01	3.8E+01	3.8E+01
	MAX	1.1E+02	5.9E+02	4.1E+02	2.5E+02	2.2E+02	1.9E+02	1.9E+02
	%MAE	10	27	16	8	8	8	8
β	MV	4.9E+03	1.2E+04	1.2E+04	9.7E+03	9.0E+03	8.2E+03	8.3E+03
	MAE	5.2E+03	5.9E+03	5.5E+03	3.1E+03	2.8E+03	2.4E+03	2.4E+03
	RMSD	6.6E+03	1.3E+04	1.1E+04	5.0E+03	4.0E+03	3.2E+03	3.2E+03
	MAX	1.9E+04	5.8E+04	5.1E+04	2.0E+04	1.3E+04	9.5E+03	9.6E+03
	%MAE	52	59	55	31	28	24	24
γ	MV	1.2E+06	2.8E+06	2.9E+06	2.5E+06	2.3E+06	2.1E+06	2.1E+06
	MAE	1.5E+06	1.7E+06	1.4E+06	7.8E+05	6.9E+05	6.4E+05	6.1E+05
	RMSD	2.3E+06	2.9E+06	2.9E+06	1.4E+06	1.1E+06	1.0E+06	9.8E+05
	MAX	7.5E+06	8.7E+06	1.2E+07	5.0E+06	3.5E+06	3.9E+06	3.7E+06
	%MAE	55	64	50	29	26	24	23

Appendix C

Supporting information of Chapter 6

C.1 Performance of RI-CC2 and RI-MP2 methods for dispersion interactions

In order to verify that the choice of RI-CC2 as reference geometry is valid for the systems under investigation, we evaluated the performances of both RI-CC2 and RI-MP2 methods on the benchmark dataset Diet-GMTKN55[1]. Among all the molecules of the set we reproduced 20 energy differences which belong to the subset of "Intramolecular noncovalent interactions", namely the molecules of Diet-GMTKN55 which also belong to the sets: ACONF, BUT14DIOL, IDISP, MCONF, PCONF21, SCONF and UPU23. The calculations have been performed using the cc-pVDZ basis set. The results are reported in Table C.1. Both RI-CC2 and RI-MP2 methods reproduce the reference data to a good accuracy (the MAE is 1.93 for RI-CC2 and 1.76 for RI-MP2).

Table C.1: Collection of the energies of the Intramolecular non-covalent interactions on the Diet-GMTKN55. On the first column is indicated the name of the corresponding subset, while on the second and third columns the molecules considered as reactant and product for the calculation of the energy difference. All the energies are in kcal/mol. The geometries and reference methods can be found in the reference paper[1].

Dataset	Reactants	Products	ΔE_{Ref}	ΔE_{RI-CC2}	ΔE_{RI-MP2}
ACONF	H_ttt	H_g+x-t+	2.63	2.78	2.79
	H_ttt	H_x+g-g-	3.08	3.26	3.29
BUt14DIOL	B1	B3	0.30	0.50	0.44
	B1	B30	2.85	4.24	4.15
	B1	B33	2.63	5.57	5.25
	B1	B39	3.06	4.71	4.58
	B1	B40	3.10	5.40	5.05
	B1	B43	3.29	4.97	4.75
	B1	B44	3.59	6.04	5.67
	B1	B54	3.15	6.04	5.67
IDISP	octane1	octane2	-1.21	4.27	3.53
MCONF	1	12	4.45	4.31	3.83
	1	21	3.11	5.60	5.38
	1	31	4.86	6.62	6.42
PCONF21	99	366	0.70	3.96	3.62
SCONF	G1	G3	6.16	-0.63	0.29
UPU23	2p	1b	2.97	1.20	1.80
	2p	1g	2.20	1.58	1.81
	2p	7a	7.26	7.43	5.80
	2p	4b	5.48	5.74	5.80
			RMSE	1.41	1.28
			MAE	1.93	1.76

C.2 Impact of the basis set

Table C.2: *Z*-*E* energy differences (ΔE_{EZ}, kcal/mol) calculated at the RI-CC2 level using the cc-pVDZ and cc-pVTZ basis sets.

	a (R = H)	**b** (R = Me)
cc-pVDZ	11.5	10.9
cc-pVTZ	10.8	10.0
Difference	0.7	0.9

Table C.3: RMSD (Å) between geometries optimized at the RI-CC2/cc-pVDZ and RI-CC2/cc-pVTZ levels for *E* and *Z* isomers.

Molecule	RMSD (*E*)	RMSD (*Z*)
a (R = H)	0.000	0.000
b (R = Me)	0.000	0.026

C.3 Density functional approximations

Table C.4: DFT exchange-correlation functionals (DFAs) considered in this study, with the type of approximation, amount of exact HF exchange (HFX) and applied dispersion correction (Disp.).

DFA acronym	Type	% of HFX	Ref.	Disp.
PBE	pure GGA	0	2	PBE-D3
M06L	pure GGA	0	3	-
rPW86PBE	pure GGA	0	4	rPW86PBE-VV
B3LYP	hybrid GGA	20	5	B3LYP-D3
PBE0	hybrid GGA	25	6	PBE0-D3
M06	hybrid GGA	27	3	-
BH&H	hybrid GGA	50	7 (G16 version)	-
M06-2X	hybrid GGA	57	3	M06-2X-D3
M06-HF	hybrid GGA	100	3	-
CAM-B3LYP	hybrid RS (0.33)	19 (SR), 65 (LR)	8	CAM-B3LYP-D3
LC-ωPBE	hybrid RS (0.40)	0 (SR), 100 (LR)	9	LC-ωPBE-D3
LC-ωPBE08	hybrid RS (0.45)	0 (SR), 100 (LR)	4	LC-ωPBE08-VV
LC-BLYP	hybrid RS (0.47)	0 (SR), 100 (LR)	10	-
Tα-BLYP	hybrid RS (variable)	0 (SR), 100 (LR)	11	-
ωB97X	hybrid RS (0.30)	15.8 (SR), 100 (LR)	12	ωB97X-D
ωB97X-D	hybrid RS (0.20)	22.2 (SR), 100 (LR)	13	-

GGA: Generalized Gradient Approximation; RS: Range Separated (with standard range parameter ω given in Bohr^{-1}); SR: Short Range; LR: Long Range.

C.4 Correlation indicators for CC2 calculations

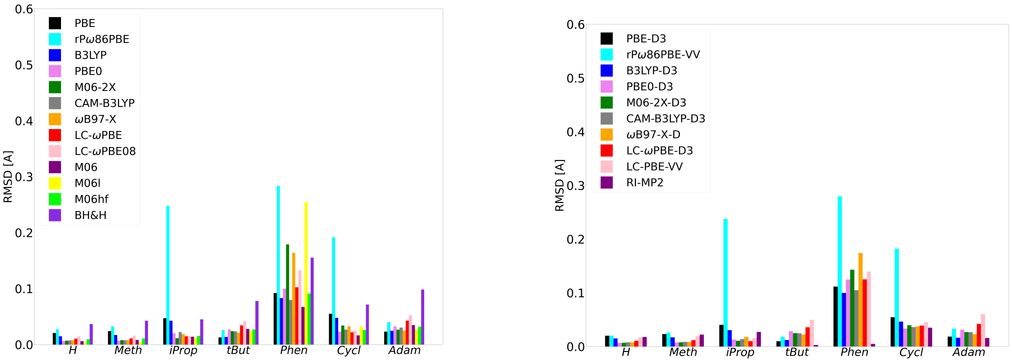
Table C.5: Evaluation of D1 and I_{ND} diagnostic calculated on the optimized geometry at the CC2 level with the cc-pVDZ basis set. The D1 threshold for molecules with a potential multireference character is 0.05. In the case of I_{ND}, we can take the dissociation of the H_2 molecule as a reference. The dissociated molecule gives $I_{ND} = 0.5$, whereas $I_{ND} = 0.025$ corresponds to the equilibrium geometry of H_2 molecule.[14] In the present case, all the molecules display $I_{ND} \leq 0.025$, and, hence, they are not expected to present large multireference character.

Isomer	*E*		*Z*	
Molecule	D1	I_{ND}	D1	I_{ND}
a (R = H)	0.046	0.023	0.053	0.024
b (R = Me)	0.049	0.022	0.053	0.022
c (R = *i*Pr)	0.049	0.021	0.051	0.021
d (R = *t*Bu)	0.050	0.021	0.052	0.021
e (R = Ph)	0.048	0.024	0.050	0.025
f (R = Cy)	0.050	0.020	0.051	0.020
g (R = Ad)	0.051	0.020	0.052	0.021

C.5 Geometrical parameters of *E* isomers

Table C.6: Characteristic distances and angles in the *E* isomers of the investigated molecules (Figure 1), as calculated at the RI-CC2, RI-MP2 and HF methods levels.

Molecule	d_{NN}[Å]	d_{NC}[Å]	θ_{NNC}[°]	ϕ_{CNNC}[°]
HF/cc-pVDZ				
a (R = H)	1.22	1.43	116	180
b (R = Me)	1.22	1.43	116	180
c (R = *i*Pr)	1.22	1.43	116	180
d (R = *t*Bu)	1.22	1.43	116	180
e (R = Ph)	1.22	1.43	116	180
f (R = Cy)	1.22	1.43	116	180
g (R = Ad)	1.22	1.43	116	180
RI-MP2/cc-pVDZ				
a (R = H)	1.28	1.43	113	180
b (R = Me)	1.28	1.42	113	180
c (R = *i*Pr)	1.28	1.42	113	180
d (R = *t*Bu)	1.28	1.42	113	180
e (R = Ph)	1.28	1.42	113	180
f (R = Cy)	1.28	1.42	113	180
g (R = Ad)	1.28	1.42	113	180
RI-CC2/cc-pVDZ				
a (R = H)	1.28	1.43	113	180
b (R = Me)	1.29	1.43	113	180
c (R = *i*Pr)	1.29	1.43	113	180
d (R = *t*Bu)	1.29	1.42	113	180
e (R = Ph)	1.29	1.43	113	180
f (R = Cy)	1.29	1.42	114	180
g (R = Ad)	1.29	1.42	113	180



(a) (b)

Figure C.1: RMSD without hydrogens between DFT and CC2 geometries for *trans* isomers, calculated using DFAs without (left) and with (right) dispersion corrections. RI-MP2 results are also shown for comparison.

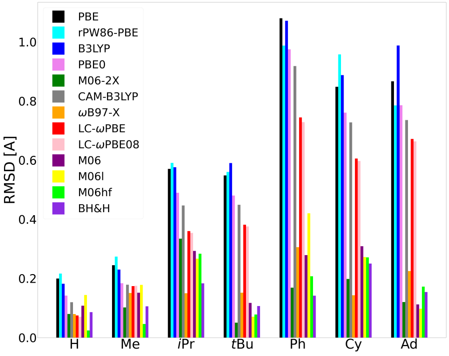


Figure C.2: RMSD without hydrogens between DFAs and CC2 geometries for *cis* isomers, calculated using DFAs without dispersion corrections. RI-MP2 results are also shown for comparison.

C.6 RMSD analysis without hydrogens

Table C.7: RMSD without hydrogens (Å) between geometries optimized using DFAs with and without dispersion corrections (in parenthesis) for *E* and *Z* isomers.

Molecule	Approx.	RMSD (*E*)	RMSD (*Z*)
a (R = **H**)	PBE(D3)	0.00	0.04
	rPW86PBE(VV)	0.00	0.00
	PBE0(D3)	0.00	0.04
	B3LYP(D3)	0.00	0.00
	M06-2X(D3)	0.00	0.00
	CAM-B3LYP(D3)	0.00	0.00
	LC-ωPBE(D3)	0.00	0.05
	LC-ωPBE(VV)	0.00	0.00
	ωB97X(D)	0.00	0.01
b (R = **Meth**)	PBE(D3)	0.00	0.07
	rPW86PBE(VV)	0.00	0.00
	PBE0(D3)	0.00	0.07
	B3LYP(D3)	0.00	0.00
	M06-2X(D3)	0.00	0.01
	CAM-B3LYP(D3)	0.00	0.00
	LC-ωPBE(D3)	0.00	0.01
	LC-ωPBE08(VV)	0.00	0.00
	ωB97-X(D)	0.00	0.05
c (R = **iProp**)	PBE(D3)	0.01	0.52
	rPW86PBE(VV)	0.00	0.00
	PBE0(D3)	0.01	0.45
	B3LYP(D3)	0.01	0.00
	M06-2X(D3)	0.00	0.04
	CAM-B3LYP(D3)	0.01	0.44
	LC-ωPBE(D3)	0.01	0.00
	LC-ωPBE08(VV)	0.00	0.00
	ωB97-X(D)	0.00	0.12

Molecule	Approx.	RMSD (E)	RMSD (Z)
d (R = **tBut**)	PBE(D3)	0.00	0.42
	rPW86PBE(VV)	0.00	0.00
	PBE0(D3)	0.00	0.63
	B3LYP(D3)	0.01	0.00
	M06-2X(D3)	0.00	0.04
	CAM-B3LYP(D3)	0.00	0.35
	LC-ωPBE(D3)	0.00	0.38
	LC-ωPBE08(VV)	0.00	0.00
	ωB97-X(D)	0.00	0.18
e (R = **Phen**)	PBE(D3)	0.03	0.87
	rPW86PBE(VV)	0.00	0.00
	PBE0(D3)	0.03	0.79
	B3LYP(D3)	0.02	0.00
	M062X(D3)	0.04	0.04
	CAM-B3LYP(D3)	0.03	0.80
	LC-ωPBE(D3)	0.03	0.65
	LC-ωPBE08(VV)	0.00	0.00
	ωB97-X(D)	0.02	0.20
f (R = **Cycl**)	PBE(D3)	0.02	0.82
	rPW86PBE(VV)	0.00	0.00
	PBE0(D3)	0.02	0.75
	B3LYP(D3)	0.03	0.00
	M06-2X(D3)	0.01	0.03
	CAM-B3LYP(D3)	0.02	0.75
	LC-ωPBE(D3)	0.02	0.63
	LC-PBE08(VV)	0.00	0.00
	ωB97-X(D)	0.01	0.13
g (R = **Adam**)	PBE(D3)	0.01	0.79
	rPW86PBE(VV)	0.00	0.00
	PBE0(D3)	0.01	0.70
	B3LYP(D3)	0.02	0.00
	M06-2X(D3)	0.00	0.06
	CAM-B3LYP(D3)	0.01	0.61
	LC-ωPBE(D3)	0.01	0.54
	LC-ωPBE08(VV)	0.00	0.00
	ωB97-X(D)	0.01	0.22

Table C.8: RMSD without hydrogens (Å) between DFAs and CC2 geometries for *E* and *Z* isomers, calculated using DFAs without and with dispersion corrections. RMSD between MP2 and CC2 geometries are also reported.

Compound	Approx.	RMSD (*E*)	RMSD (*Z*)
a (R = **H**)	RI-MP2	0.02	0.01
	PBE	0.02	0.20
	PBE-D3	0.02	0.17
	rPW86PBE	0.03	0.22
	rPW86PBE-VV	0.02	0.17
	PBE0	0.01	0.14
	PBE0-D3	0.01	0.11
	B3LYP	0.01	0.18
	B3LYP-D3	0.01	0.13
	CAM-B3LYP	0.01	0.12
	CAM-B3LYP-D3	0.01	0.08
	LC-ωPBE	0.01	0.07
	LC-ωPBE-D3	0.01	0.05
	LC-ωPBE08	0.01	0.07
	LC-ωPBE08-VV	0.02	0.04
	ωB97-X	0.01	0.08
	ωB97-X-D	0.01	0.07
	BH&H	0.04	0.09
	M06	0.01	0.11
	M06L	0.00	0.14
	M06-2X	0.01	0.08
	M06-2X-D3	0.01	0.08
	M06-HF	0.01	0.02

Compound	Approx.	RMSD (E)	RMSD (Z)
b (R = **Me**)	RI-MP2	0.02	0.06
	PBE	0.02	0.24
	PBE-D3	0.02	0.17
	rPW86PBE	0.03	0.27
	rPW86PBE-VV	0.03	0.21
	PBE0	0.01	0.18
	PBE0-D3	0.01	0.11
	B3LYP	0.02	0.23
	B3LYP-D3	0.02	0.14
	CAM-B3LYP	0.01	0.18
	CAM-B3LYP-D3	0.01	0.12
	LC-ωPBE	0.01	0.17
	LC-ωPBE-D3	0.01	0.17
	LC-ωPBE08	0.02	0.17
	LC-ωPBE08-VV	0.02	0.14
	ωB97-X	0.01	0.15
	ωB97-X-D	0.01	0.11
	BH&H	0.04	0.11
	M06	0.01	0.15
	M06L	0.01	0.18
	M06-2X	0.01	0.10
	M06-2X-D3	0.01	0.10
	M06-HF	0.01	0.05
c (R = **iPr**)	RI-MP2	0.03	0.01
	PBE	0.05	0.57
	PBE-D3	0.04	0.15
	rPW86PBE	0.25	0.59
	rPW86PBE-VV	0.24	0.30
	PBE0	0.02	0.49
	PBE0-D3	0.01	0.13
	B3LYP	0.04	0.58
	B3LYP-D3	0.03	0.12
	CAM-B3LYP	0.02	0.45
	CAM-B3LYP-D3	0.01	0.13
	LC-ωPBE	0.01	0.36
	LC-ωPBE-D3	0.01	0.36
	LC-ωPBE08	0.01	0.35
	LC-ωPBE08-VV	0.02	0.09
	ωB97-X	0.02	0.15
	ωB97-X-D	0.02	0.08
	BH&H	0.05	0.18
	M06	0.01	0.29
	M06L	0.01	0.27
	M06-2X	0.01	0.33
	M06-2X-D3	0.01	0.31
	M06-HF	0.02	0.28

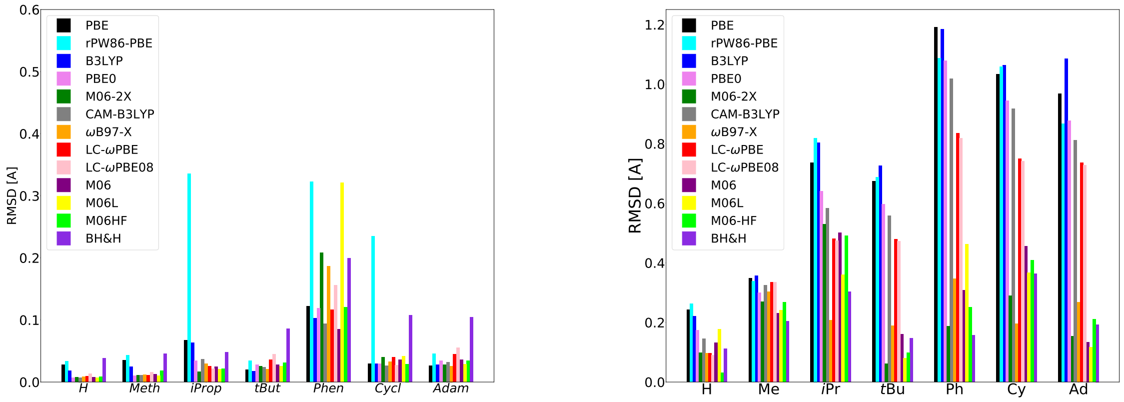
Compound	Approx.	RMSD (E)	RMSD (Z)
d (R = **tBu**)	RI-MP2	0.00	0.00
	PBE	0.01	0.55
	PBE-D3	0.01	0.13
	rPW86PBE	0.03	0.56
	rPW86PBE-VV	0.02	0.09
	PBE0	0.03	0.48
	PBE0-D3	0.03	0.11
	B3LYP	0.01	0.59
	B3LYP-D3	0.01	0.09
	CAM-B3LYP	0.02	0.45
	CAM-B3LYP-D3	0.02	0.12
	LC-ωPBE	0.03	0.38
	LC-ωPBE-D3	0.04	0.08
	LC-ωPBE08	0.04	0.38
	LC-ωPBE08-VV	0.05	0.07
	ωB97-X	0.02	0.15
	ωB97-X-D	0.02	0.07
	BH&H	0.08	0.11
	M06	0.03	0.12
	M06L	0.02	0.07
	M06-2X	0.02	0.05
	M06-2X-D3	0.02	0.04
	M06-HF	0.03	0.08
e (R = **Ph**)	RI-MP2	0.01	0.02
	PBE	0.09	1.08
	PBE-D3	0.11	0.22
	rPW86PBE	0.28	0.99
	rPW86PBE-VV	0.28	0.44
	PBE0	0.10	0.97
	PBE0-D3	0.12	0.20
	B3LYP	0.08	1.07
	B3LYP-D3	0.10	0.14
	CAM-B3LYP	0.08	0.92
	CAM-B3LYP-D3	0.10	0.18
	LC-ωPBE	0.10	0.74
	LC-ωPBE-D3	0.12	0.12
	LC-ωPBE08	0.13	0.73
	LC-ωPBE08-VV	0.14	0.11
	ωB97-X	0.16	0.31
	ωB97-X-D	0.17	0.12
	BH&H	0.16	0.14
	M06	0.07	0.28
	M06L	0.26	0.42
	M06-2X	0.18	0.17
	M06-2X-D3	0.14	0.13
	M06-HF	0.09	0.21

Compound	Approx.	RMSD (E)	RMSD (Z)
f (R = **Cy**)	RI-MP2	0.03	0.01
	PBE	0.05	0.85
	PBE-D3	0.05	0.14
	rPW86PBE	0.19	0.96
	rPW86PBE-VV	0.18	0.10
	PBE0	0.02	0.76
	PBE0-D3	0.03	0.11
	B3LYP	0.05	0.89
	B3LYP-D3	0.05	0.11
	CAM-B3LYP	0.03	0.73
	CAM-B3LYP-D3	0.04	0.13
	LC-ωPBE	0.02	0.61
	LC-ωPBE-D3	0.04	0.10
	LC-ωPBE08	0.02	0.60
	LC-ωPBE08-VV	0.05	0.09
	ωB97-X	0.03	0.14
	ωB97-X-D	0.04	0.17
	BH&H	0.07	0.25
	M06	0.02	0.31
	M06L	0.03	0.27
	M06-2X	0.03	0.20
	M06-2X-D3	0.04	0.19
	M06-HF	0.03	0.27
g (R = **Ad**)	RI-MP2	0.02	0.03
	PBE	0.02	0.87
	PBE-D3	0.02	0.43
	rPW86PBE	0.04	0.79
	rPW86PBE-VV	0.03	0.20
	PBE0	0.03	0.79
	PBE0-D3	0.03	0.43
	B3LYP	0.02	0.99
	B3LYP-D3	0.02	0.42
	CAM-B3LYP	0.03	0.74
	CAM-B3LYP-D3	0.03	0.43
	LC-ωPBE	0.04	0.67
	LC-ωPBE-D3	0.04	0.43
	LC-ωPBE08	0.05	0.66
	LC-ωPBE08-VV	0.06	0.45
	ωB97-X	0.02	0.23
	ωB97-X-D	0.02	0.06
	BH&H	0.10	0.15
	M06	0.03	0.11
	M06L	0.03	0.10
	M06-2X	0.03	0.12
	M06-2X-D3	0.03	0.08
	M06-HF	0.03	0.17

Table C.9: Average values of the RMSD calculations (Å) without the hydrogens for each set of molecules in E and Z isomers.

Functional	Average RMSD (E)	Average RMSD (Z)
RI-MP2	0.02	0.02
PBE	0.04	0.62
PBE-D3	0.04	0.20
rPW86PBE	0.12	0.62
rPW86PBE-VV	0.11	0.22
B3LYP	0.03	0.65
B3LYP-D3	0.03	0.16
PBE0	0.03	0.55
PBE0-D3	0.03	0.17
M06-2X	0.04	0.15
M06-2X-D3	0.04	0.13
CAM-B3LYP	0.03	0.51
CAM-B3LYP-D3	0.03	0.17
ωB97X	0.04	0.17
ωB97X-D	0.04	0.10
LC-ωPBE	0.03	0.43
LC-ωPBE-D3	0.04	0.19
LC-ωPBE08	0.04	0.42
LC-ωPBE08-VV	0.05	0.14
M06	0.02	0.20
M06L	0.05	0.21
M06-HF	0.03	0.15
BH&H	0.08	0.15

C.7 RMSD analysis with hydrogens



(a)

(b)

Figure C.3: RMSD with hydrogens (Å) between DFAs and CC2 geometries for *trans* (left) and *cis* isomers (right), calculated using DFAs without dispersion corrections. RI-MP2 results are also shown for comparison.

Table C.10: RMSD (Å) between geometries optimized using DFAs with and without dispersion corrections (in parenthesis) for *E* and *Z* isomers.

Molecule	Approx.	RMSD (*E*)	RMSD (*Z*)
a (R = H)	PBE(D3)	0.00	0.05
	rPW86PBE (VV)	0.01	0.06
	PBE0(D3)	0.00	0.05
	B3LYP(D3)	0.00	0.07
	M06-2X(D3)	0.00	0.00
	CAM-B3LYP(D3)	0.00	0.05
	LC-ωPBE(D3)	0.00	0.06
	LC-ωPBE08(VV)	0.01	0.01
	ωB97X(D)	0.00	0.02
b (R = Me)	PBE(D3)	0.00	0.01
	rPW86PBE(VV)	0.01	0.08
	PBE0(D3)	0.00	0.10
	B3LYP(D3)	0.01	0.12
	M06-2X(D3)	0.00	0.03
	CAM-B3LYP(D3)	0.00	0.10
	LC-ωPBE(D3)	0.00	0.12
	LC-ωPBE08(VV)	0.01	0.16
	ωB97X(D)	0.00	0.06
c (R = *i*Pr)	PBE(D3)	0.01	0.72
	rPW86PBE(VV)	0.03	0.73
	PBE0(D3)	0.01	0.64
	B3LYP(D3)	0.02	0.77
	M06-2X(D3)	0.00	0.05
	CAM-B3LYP(D3)	0.02	0.62
	LC-ωPBE(D3)	0.02	0.65
	LC-ωPBE08(VV)	0.01	0.53
	ωB97X(D)	0.00	0.15

Molecule	Approx.	RMSD (E)	RMSD (Z)
d (R = tBu)	PBE(D3)	0.01	0.53
	rPW86PBE(VV)	0.02	0.59
	PBE0(D3)	0.01	0.47
	B3LYP(D3)	0.01	0.64
	M06-2X(D3)	0.00	0.05
	CAM-B3LYP(D3)	0.01	0.44
	LC-ωPBE(D3)	0.01	0.49
	LC-ωPBE08(VV)	0.01	0.49
	ωB97X(D)	0.01	0.23
e (R = Ph)	PBE(D3)	0.03	0.97
	rPW86PBE(VV)	0.02	0.74
	PBE0(D3)	0.03	0.88
	B3LYP(D3)	0.03	1.07
	M06-2X(D3)	0.04	0.05
	CAM-B3LYP(D3)	0.03	0.89
	LC-ωPBE(D3)	0.03	0.72
	LC-ωPBE08(VV)	0.03	0.71
	ωB97X(D)	0.03	0.22
f (R = Cy)	PBE(D3)	0.03	1.00
	rPW86PBE(VV)	0.03	1.08
	PBE0(D3)	0.03	0.94
	B3LYP(D3)	0.04	1.11
	M06-2X(D3)	0.01	0.04
	CAM-B3LYP(D3)	0.03	0.94
	LC-ωPBE(D3)	0.03	0.81
	LC-ωPBE08(VV)	0.03	0.76
	ωB97X(D)	0.01	0.18
g (R = Ad)	PBE(D3)	0.02	0.90
	rPW86PBE(VV)	0.02	1.15
	PBE0(D3)	0.02	0.79
	B3LYP(D3)	0.02	1.02
	M06-2X(D3)	0.00	0.07
	CAM-B3LYP(D3)	0.02	0.69
	LC-ωPBE(D3)	0.02	0.61
	LC-ωPBE08(VV)	0.02	0.61
	ωB97X(D)	0.01	0.27

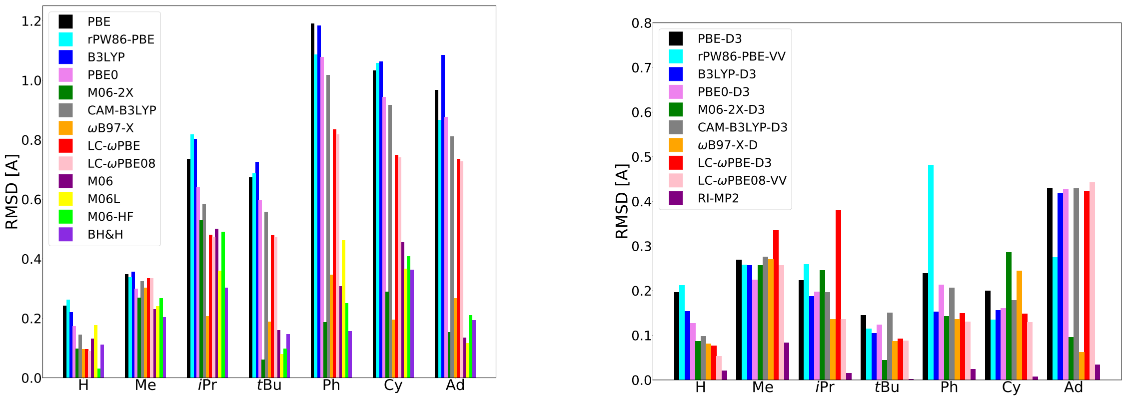
Table C.11: RMSD (Å) between DFAs and CC2 geometries for E and Z isomers, calculated using DFAs without and with dispersion corrections. RMSD between MP2 and CC2 geometries are also reported.

Compound	Approx.	RMSD (E)	RMSD (Z)
a (R = H)	RI-MP2	0.02	0.02
	PBE	0.03	0.24
	PBE-D3	0.03	0.20
	rPW86PBE-VV	0.03	0.21
	rPW86PBE	0.03	0.26
	PBE0	0.01	0.17
	PBE0-D3	0.01	0.13
	B3LYP	0.02	0.22
	B3LYP-D3	0.02	0.15
	CAM-B3LYP	0.01	0.14
	CAM-B3LYP-D3	0.01	0.10
	LC-ωPBE	0.01	0.10
	LC-ωPBE-D3	0.01	0.08
	LC-ωPBE08	0.01	0.09
	LC-ωPBE08-VV	0.02	0.26
	ωB97-X	0.01	0.10
	ωB97-X-D	0.01	0.08
	BH&H	0.04	0.11
	M06	0.01	0.13
	M06L	0.01	0.18
	M06-2X	0.01	0.10
	M06-2X-D3	0.01	0.09
	M06-HF	0.01	0.03

Compound	Approx.	RMSD (E)	RMSD (Z)
b (R = Me)	RI-MP2	0.03	0.08
	PBE	0.04	0.35
	PBE-D3	0.03	0.27
	rPW86PBE-VV	0.04	0.26
	rPW86PBE	0.04	0.34
	PBE0	0.01	0.30
	PBE0-D3	0.01	0.22
	B3LYP	0.02	0.36
	B3LYP-D3	0.02	0.26
	CAM-B3LYP	0.01	0.32
	CAM-B3LYP-D3	0.01	0.28
	LC-ωPBE	0.01	0.34
	LC-ωPBE-D3	0.01	0.34
	LC-ωPBE08	0.01	0.34
	LC-ωPBE08-VV	0.02	0.26
	ωB97-X	0.01	0.30
	ωB97-X-D	0.01	0.27
	BH&H	0.05	0.20
	M06	0.01	0.23
	M06L	0.01	0.24
	M06-2X	0.01	0.27
	M06-2X-D3	0.01	0.26
	M06-HF	0.02	0.27
c (R = iPr)	RI-MP2	0.04	0.01
	PBE	0.07	0.74
	PBE-D3	0.06	0.22
	rPW86PBE-VV	0.32	0.26
	rPW86PBE	0.34	0.68
	PBE0	0.03	0.64
	PBE0-D3	0.02	0.20
	B3LYP	0.06	0.80
	B3LYP-D3	0.04	0.19
	CAM-B3LYP	0.04	0.58
	CAM-B3LYP-D3	0.02	0.20
	LC-ωPBE	0.03	0.48
	LC-ωPBE-D3	0.01	0.38
	LC-ωPBE08	0.02	0.47
	LC-ωPBE08-VV	0.02	0.14
	ωB97-X	0.03	0.21
	ωB97-X-D	0.03	0.14
	BH&H	0.05	0.30
	M06	0.02	0.50
	M06L	0.02	0.36
	M06-2X	0.02	0.53
	M06-2X-D3	0.02	0.25
	M06-HF	0.02	0.49

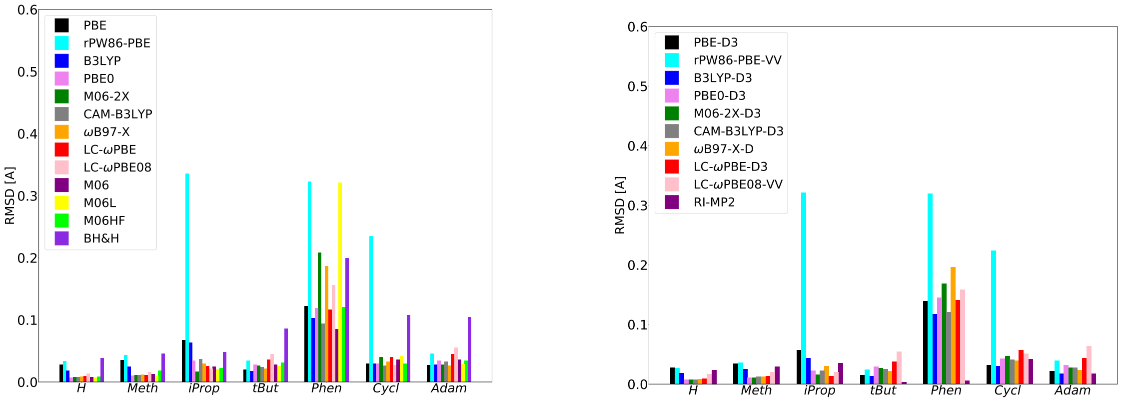
Compound	Approx.	RMSD (E)	RMSD (Z)
d (R = tBu)	RI-MP2	0.003	0.002
	PBE	0.02	0.67
	PBE-D3	0.02	0.14
	rPW86PBE-VV	0.02	0.11
	rPW86PBE	0.03	0.69
	PBE0	0.03	0.60
	PBE0-D3	0.03	0.12
	B3LYP	0.02	0.73
	B3LYP-D3	0.01	0.10
	CAM-B3LYP	0.02	0.56
	CAM-B3LYP-D3	0.02	0.15
	LC-ωPBE	0.04	0.48
	LC-ωPBE-D3	0.04	0.09
	LC-ωPBE08	0.04	0.47
	LC-ωPBE08-VV	0.05	0.09
	ωB97-X	0.02	0.19
	ωB97-X-D	0.02	0.09
	BH&H	0.09	0.15
	M06	0.03	0.16
	M06L	0.03	0.08
	M06-2X	0.03	0.06
	M06-2X-D3	0.03	0.04
	M06-HF	0.03	0.10
e (R = Ph)	RI-MP2	0.01	0.03
	PBE	0.12	1.19
	PBE-D3	0.14	0.24
	rPW86PBE-VV	0.32	0.48
	rPW86PBE	0.32	1.09
	PBE0	0.12	1.08
	PBE0-D3	0.14	0.21
	B3LYP	0.10	1.18
	B3LYP-D3	0.12	0.15
	CAM-B3LYP	0.09	1.02
	CAM-B3LYP-D3	0.12	0.21
	LC-ωPBE	0.12	0.84
	LC-ωPBE-D3	0.14	0.15
	LC-ωPBE08	0.15	0.81
	LC-ωPBE08-VV	0.16	0.13
	ωB97-X	0.19	0.35
	ωB97-X-D	0.20	0.14
	BH&H	0.20	0.16
	M06	0.09	0.31
	M06L	0.32	0.46
	M06-2X	0.21	0.19
	M06-2X-D3	0.17	0.14
	M06-HF	0.12	0.25

Compound	Approx.	RMSD (E)	RMSD (Z)
f (R = Cy)	RI-MP2	0.04	0.01
	PBE	0.03	1.03
	PBE-D3	0.03	0.20
	rPW86PBE-VV	0.22	0.13
	rPW86PBE	0.23	1.06
	PBE0	0.03	0.94
	PBE0-D3	0.04	0.16
	B3LYP	0.03	1.06
	B3LYP-D3	0.03	0.16
	CAM-B3LYP	0.03	0.92
	CAM-B3LYP-D3	0.04	0.18
	LC-ωPBE	0.04	0.75
	LC-ωPBE-D3	0.06	0.15
	LC-ωPBE08	0.03	0.74
	LC-ωPBE08-VV	0.05	0.13
	ωB97-X	0.03	0.20
	ωB97-X-D	0.04	0.24
	BH&H	0.11	0.36
	M06	0.04	0.46
	M06L	0.04	0.37
	M06-2X	0.04	0.29
	M06-2X-D3	0.05	0.29
	M06-HF	0.03	0.41
g (R = Ad)	RI-MP2	0.02	0.03
	PBE	0.03	0.97
	PBE-D3	0.02	0.43
	rPw86PBE-VV	0.04	0.27
	rPw86PBE	0.05	0.87
	PBE0	0.03	0.88
	PBE0-D3	0.03	0.43
	B3LYP	0.03	1.09
	B3LYP-D3	0.02	0.42
	CAM-B3LYP	0.03	0.81
	CAM-B3LYP-D3	0.03	0.43
	LC-ωPBE	0.04	0.74
	LC-ωPBE-D3	0.04	0.42
	LC-ωPBE08	0.06	0.73
	LC-ωPBE08-VV	0.06	0.44
	wB97-X	0.03	0.27
	wB97-X-D	0.02	0.06
	BH&H	0.10	0.19
	M06	0.04	0.13
	M06L	0.03	0.12
	M06-2X	0.03	0.15
	M06-2X-D3	0.03	0.10
	M06-HF	0.03	0.21



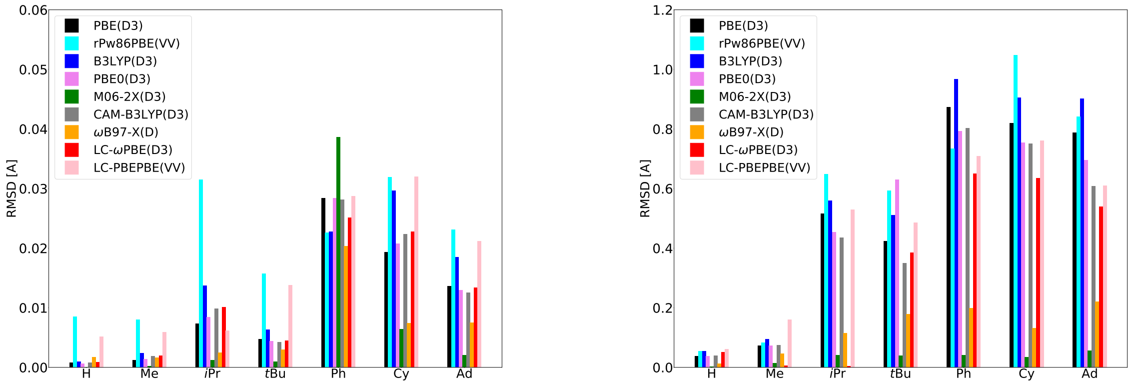
(a) (b)

Figure C.4: RMSD between DFAs and CC2 geometries for *cis* isomers, calculated using DFAs without (left) and with (right) dispersion corrections. RI-MP2 results are also shown for comparison.



(a) (b)

Figure C.5: RMSD between DFAs and CC2 geometries for *trans* isomers, calculated using DFAs without (left) and with (right) dispersion corrections. RI-MP2 results are also shown for comparison.



(a) (b)

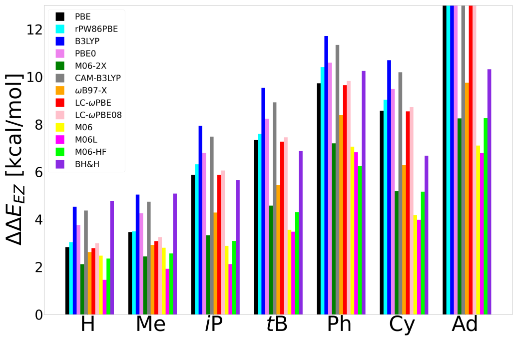
Figure C.6: RMSD (Å) excluding hydrogen atoms between geometries optimized using DFAs with and without dispersion corrections for *trans* (left) and *cis* (right) isomers.

C.8 Relative *Z*-*E* energies

C.8.1 CC2 *vs* MP2

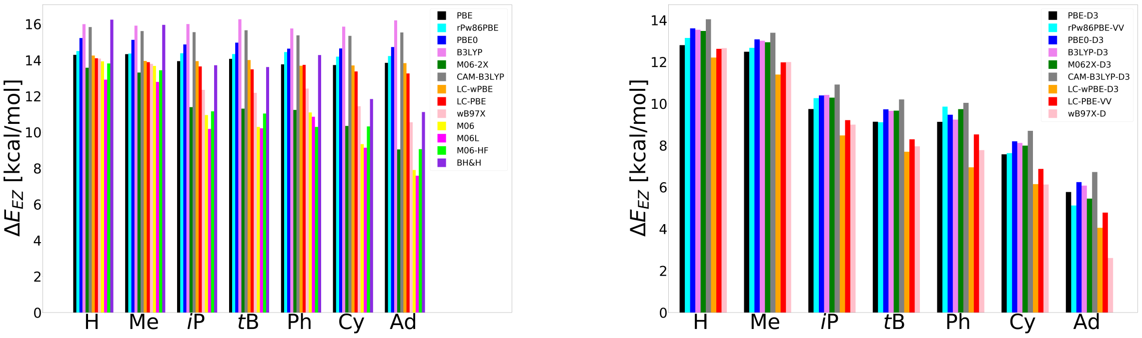
Table C.12: Difference in the energy of the *trans* (E) and *cis* (Z) forms, $\Delta E_{EZ} = E_Z - E_E$, as calculated at the HF, RI-CC2 and RI-MP2 levels with the cc-pVDZ basis set. All energy values are in kcal.mol^{-1}.

Molecule	Approx.	ΔE_{EZ}
a (R = H)	RI-CC2	11.46
	RI-MP2	12.01
	HF	17.9
b (R = Me)	RI-CC2	10.87
	RI-MP2	11.43
	HF	17.4
c (R = *i*Pr)	RI-CC2	8.06
	RI-MP2	8.90
	HF	17.5
d (R = *t*Bu)	RI-CC2	6.73
	RI-MP2	7.70
	HF	17.7
e (R = Ph)	RI-CC2	4.04
	RI-MP2	4.9
	HF	17.4
f (R = Cy)	RI-CC2	5.16
	RI-MP2	6.41
	HF	17.3
g (R = Ad)	RI-CC2	0.80
	RI-MP2	2.67
	HF	17.4



(a)

Figure C.7: *Z*-*E* energy differences calculated at the CC2 level and DFAs without dispersion corrections ($\Delta\Delta E_{ZE} = \Delta E_{ZE}(DFT) - \Delta E_{ZE}(CC2)$, in kcal/mol).



(a) (b)

Figure C.8: Difference in the energy of the *trans* (E) and *cis* (Z) forms, $\Delta E_{EZ} = E_Z - E_E$, as calculated using DFAs without (left) and with (right) dispersion corrections and the cc-pVDZ basis set.

C.8.2 DFT relative *Z*-*E* energies

Table C.13: Relative electronic energies of the E and Z conformers, $\Delta E_{ZE} = E_E - E_Z$, and relative Gibbs free energies ($\Delta G_{ZE} = G_E - G_Z$), calculated at T = 300 K and P = 1 atm. All calculations were done using the cc-pVDZ basis set. All energy values are in kcal.mol^{-1}. The Gibbs free energy have been calculated for the functionals implemented in Gaussian using the quasi-harmonic approximation (qa-G) employing the software goodvibes.py [15]. The vibrational entropy was tested using the Grimme scheme [16], while the enthalpy contributions were evaluated with the Head-Gordon scheme [17]. The threshold for which the qa-G was employed is 100 cm^{-1}. For the functionals calculated with Q-Chem (namely, rPW86-PBE, rPW86-PBE-VV, LC-ωPBE08 and LC-ωPBE08-VV) frequency calculations do not have been performed because analytical calculations are not implemented, and therefore free energies are not available. For RI-CC2 and RI-MP2 the Gibbs free energies correspond to normal computations.

Compound	Approx.	ΔE_{ZE}	ΔG_{ZE}
a (R = H)	RI-CC2	-11.5	-11.6
	RI-MP2	-12.0	-12.3
	PBE	-14.3	-13.6
	PBE-D3	-12.8	-12.2
	PBE0	-15.2	-14.7
	PBE0-D3	-13.6	-13.1
	rPW86-PBE	-14.5	
	rPW86-PBE-VV	-13.1	
	B3LYP	-16.0	-15.4
	B3LYP-D3	-13.5	-13.0
	BH&H	-16.2	-15.7
	CAM-B3LYP	-15.8	-15.3
	CAM-B3LYP-D3	-14.0	-13.6
	ωB97-X	-14.1	-13.7
	ωB97-X-D	-12.7	-12.4
	LC-ωPBE	-14.2	-14.0
	LC-ωPBE-D3	-12.2	-12.1
	LC-ωPBE08	-14.5	
	LC-ωPBE08-VV	-12.8	
	M06	-13.9	-13.5
	M06L	-12.9	-12.4
	M06-2X	-13.6	-13.1
	M06-2X-D3	-13.5	-12.8
	M06-HF	-13.8	-13.4
	HF	-17.9	-17.5

Compound	Approx.	ΔE_{ZE}	ΔG_{ZE}
b (R = Me)	RI-CC2	-10.9	-11.4
	RI-MP2	-11.4	-12.1
	PBE	-14.3	-13.5
	PBE-D3	-12.5	-11.9
	PBE0	-15.1	-14.7
	PBE0-D3	-13.1	-12.7
	rPW86-PBE	-14.4	
	rPW86-PBE-VV	-12.7	
	B3LYP	-15.9	-15.4
	B3LYP-D3	-13.0	-12.7
	BH&H	-16.0	-15.2
	CAM-B3LYP	-15.6	-15.3
	CAM-B3LYP-D3	-13.4	-13.2
	ωB97-X	-13.8	-13.7
	ωB97-X-D	-12.0	-11.4
	LC-ωPBE	-14.0	-14.0
	LC-ωPBE-D3	-11.4	-11.5
	LC-ωPBE08	-14.1	
	LC-ωPBE08-VV	-12.1	
	M06	-13.6	-13.8
	M06L	-12.8	-12.3
	M06-2X	-13.3	-13.7
	M06-2X-D3	-12.9	-13.2
	M06-HF	-13.4	-13.9
	HF	-17.4	-17.2

Compound	Approx.	ΔE_{ZE}	ΔG_{ZE}
c (R = iPr)	RI-CC2	-8.1	-8.4
	RI-MP2	-8.9	-9.5
	PBE	-13.9	-14.1
	PBE-D3	-9.7	-10.5
	PBE0	-14.9	-15.1
	PBE0-D3	-10.4	-11.1
	rPW86-PBE	-14.4	
	rPW86-PBE-VV	-10.3	
	B3LYP	-16.0	-16.1
	B3LYP-D3	-10.4	-11.5
	BH&H	-13.7	-14.5
	CAM-B3LYP	-15.5	-15.7
	CAM-B3LYP-D3	-10.9	-11.8
	ωB97-X	-12.3	-13.6
	ωB97-X-D	-9.0	-10.1
	LC-ωPBE	-13.9	-14.6
	LC-ωPBE-D3	-8.5	-10.2
	LC-ωPBE08	-14.1	
	LC-ωPBE08-VV	-9.2	
	M06	-11.0	-13.0
	M06L	-10.2	-11.3
	M06-2X	-11.4	-12.6
	M06-2X-D3	-10.3	-11.4
	M06-HF	-11.2	-14.0
	HF	-17.5	-17.7

Compound	Approx.	ΔE_{ZE}	ΔG_{ZE}
d (R = tBu)	RI-CC2	-6.7	-7.5
	RI-MP2	-7.7	-8.6
	PBE	-14.1	-13.9
	PBE-D3	-9.1	-9.5
	PBE0	-15.0	-15.0
	PBE0-D3	-9.7	-10.3
	rPW86-PBE	-14.3	
	rPW86-PBE-VV	-9.1	
	B3LYP	-16.3	-16.2
	B3LYP-D3	-9.7	-10.5
	BH&H	-13.6	-14.2
	CAM-B3LYP	-15.7	-15.8
	CAM-B3LYP-D3	-10.2	-11.0
	ωB97-X	-12.2	-13.0
	ωB97-X-D	-8.0	-9.9
	LC-ωPBE	-14.0	-14.4
	LC-ωPBE-D3	-7.7	-8.8
	LC-ωPBE08	-14.2	
	LC-ωPBE08-VV	-8.2	
	M06	-10.3	-11.5
	M06L	-10.2	-11.6
	M06-2X	-11.3	-11.9
	M06-2X-D3	-9.7	-10.2
	M06-HF	-11.0	-11.6
	HF	-17.7	-18.0

Compound	Approx.	ΔE_{ZE}	ΔG_{ZE}
e (R = Ph)	RI-CC2	-4.0	-5.9
	RI-MP2	-4.9	-7.0
	PBE	-13.8	-13.3
	PBE-D3	-9.1	-9.3
	PBE0	-14.6	-14.3
	PBE0-D3	-9.5	-9.7
	rPW86-PBE	-14.4	
	rPW86-PBE-VV	-9.9	
	B3LYP	-15.8	-15.3
	B3LYP-D3	-9.2	-10.0
	BH&H	-14.3	-14.4
	CAM-B3LYP	-15.4	-15.0
	CAM-B3LYP-D3	-10.0	-10.7
	ωB97-X	-12.4	-12.4
	ωB97-X-D	-7.8	-8.5
	LC-ωPBE	-13.7	-13.5
	LC-ωPBE-D3	-7.0	-8.0
	LC-ωPBE08	-13.9	
	LC-ωPBE08-VV	-7.9	
	M06	-11.1	-11.6
	M06L	-10.9	-10.9
	M06-2X	-11.2	-11.0
	M06-2X-D3	-9.7	-9.9
	M06-HF	-10.3	-10.5
	HF	-17.4	-17.2

Compound	Approx.	ΔE_{ZE}	ΔG_{ZE}
f (R = Cy)	RI-CC2	-5.2	-5.5
	RI-MP2	-6.4	-7.0
	PBE	-13.7	-13.5
	PBE-D3	-7.6	-8.6
	PBE0	-14.6	-14.4
	PBE0-D3	-8.2	-9.4
	rPW86-PBE	-14.2	
	rPW86-PBE-VV	-7.6	
	B3LYP	-15.9	-15.8
	B3LYP-D3	-8.1	-9.7
	BH&H	-11.8	-13.2
	CAM-B3LYP	-15.3	-15.3
	CAM-B3LYP-D3	-8.7	-10.3
	ωB97-X	-11.4	-12.5
	ωB97-X-D	-6.1	-8.2
	LC-ωPBE	-13.7	-13.8
	LC-ωPBE-D3	-6.1	-7.9
	LC-ωPBE08	-13.9	
	LC-ωPBE08-VV	-6.8	
	M06	-9.3	-10.1
	M06L	-9.1	9.0
	M06-2X	-10.3	-11.5
	M06-2X-D3	-8.0	-9.5
	M06-HF	-10.3	-12.2
	HF	-17.3	-17.5

Compound	Approx.	ΔE_{ZE}	ΔG_{ZE}
g (R = Ad)	RI-CC2	-0.8	
	RI-MP2	-2.7	-4.9
	PBE	-13.8	-13.7
	PBE-D3	-5.8	-6.9
	PBE0	-14.7	-14.5
	PBE0-D3	-6.2	-7.5
	rPW86-PBE	-14.2	
	rPW86-PBE-VV	-5.1	
	B3LYP	-16.2	-15.7
	B3LYP-D3	-6.1	-7.7
	BH&H	-11.1	-12.4
	CAM-B3LYP	-15.5	-14.7
	CAM-B3LYP-D3	-6.7	-8.3
	ωB97-X	-10.6	-12.1
	ωB97-X-D	-2.6	-3.7
	LC-ωPBE	-13.8	-13.9
	LC-ωPBE-D3	-4.1	-6.2
	LC-ωPBE08	-14.0	
	LC-ωPBE08-VV	-4.2	
	M06	-7.9	8.2
	M06L	-7.6	-9.1
	M06-2X	-9.0	-10.4
	M06-2X-D3	-5.5	-6.8
	M06-HF	-9.1	-10.4
	HF	-17.4	-17.8

C.9 Calculations for the Ph molecule

DFAs with dispersion corrections are able to reproduce qualitatively well the energy gap of all the systems considered. However, we observed that are much less accurate for the compound **f**, which includes phenyl substituents in the *meso* position. In order to investigate these discrepancies, we performed single-point calculations on the RI-CC2 geometries for this molecule. The results are collected in Table C.14 and Fig. C.9.

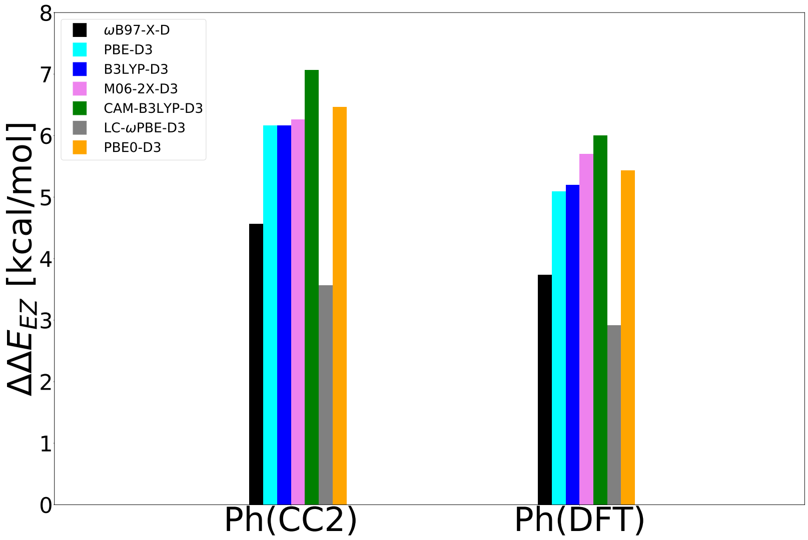


Figure C.9: $Z-E$ energy differences with respect to RI-CC2 issued from DFAs for the molecule with R = Ph, using reference RI-CC2 geometries (left) and optimized DFAs geometries (right) with dispersion corrections ($\Delta\Delta E_{ZE} = \Delta E_{ZE}^{DFT}(DFT/CC2)$ $-\Delta E_{ZE}^{CC2}(CC2)$, in kcal/mol).

Table C.14: $Z-E$ energy differences issued from RI-CC2 and DFAs for the molecule with R = Ph, using reference RI-CC2 geometries (left) and optimized DFAs' geometries (right) with dispersion corrections (in kcal/mol).

Compound	Approx.	$\Delta E_{EZ}(RI-CC2)$	$\Delta E_{EZ}(DFT)$
Ph	RI-CC2	-4.0	-
	ωB97-X-D	-8.6	-7.8
	PBE-D3	-10.2	-9.1
	PBE0-D3	-10.5	-9.5
	B3LYP-D3	-10.2	-9.2
	M06-2X-D3	-10.3	-9.7
	CAM-B3LYP-D3	-11.1	-10.0
	LC-ωPBE-D3	-7.6	-6.9

C.10 Nonlinear Optical properties

C.10.1 Reference RI-MP2 results

The accuracy of the RI approximation in the evaluation of β has been tested for the molecule R=H, with respect to the standard MP2 calculation.

Molecule	β_{MP2}	β_{RI-MP2}
R= H	52.55	52.52

C.10.2 *ab initio* results

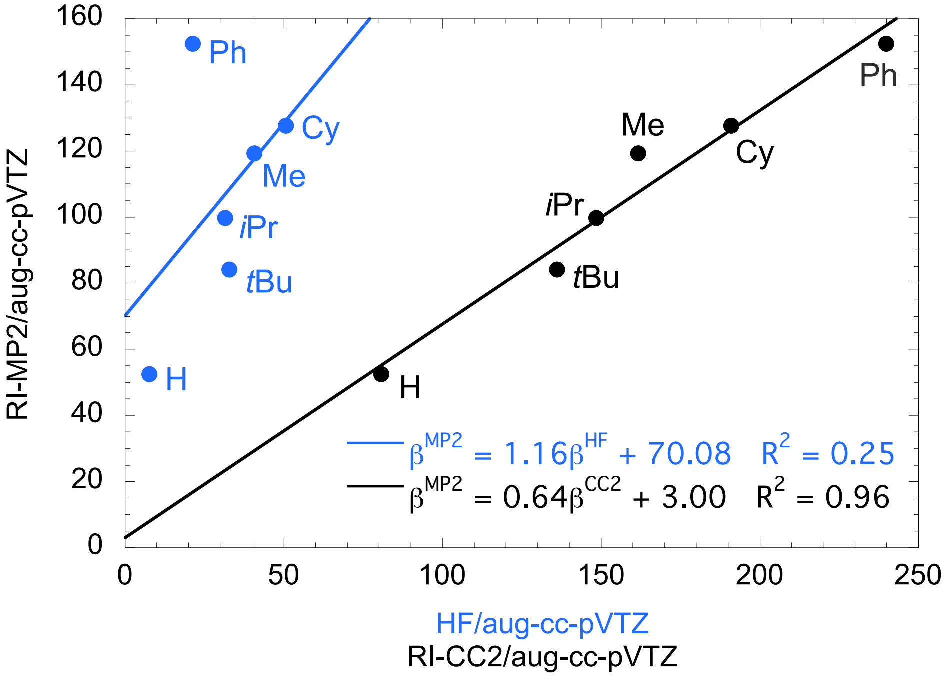


Figure C.10: Correlation between static hyperpolarizabilities (a.u.) of the series of azobenzenes calculated at the MP2 level with respect to the values calculated using CC2 and HF.

C.10.3 DFT results

Table C.15: Total first hyperpolarizabilities (β, a.u.) of the *cis* isomers, calculated using various levels of approximation with the aug-cc-pVTZ basis set, using RI-CC2/cc-pVDZ geometries.

Compound	Approx.	β
a (R = H)	RI-MP2	52.5
	RI-CC2	80.8
	PBE	99.6
	rPW86-PBE	107.2
	rPW86-PBE-VV	104.0
	PBE0	73.1
	B3LYP	83.3
	BH&H	41.1
	CAM-B3LYP	52.4
	ωB97X	44.0
	ωB97X-D	61.5
	LC-BLYP	22.2
	Tα-LC-BLYP	44.3
	LC-PBE	23.6
	LC-ωPBE	31.2
	LC-PBE-VV	42.5
	M06	96.8
	M06L	136.3
	M06-2X	43.8
	M06-HF	1.5
	HF	7.7

Compound	Approx.	β
b (R = Me)	RI-MP2	119.3
	RI-CC2	161.7
	PBE	186.6
	rPW86-PBE	196.4
	rPW86-PBE-VV	200.4
	PBE0	142.8
	B3LYP	155.8
	BH&H	99.3
	CAM-B3LYP	115.4
	ωB97X	107.2
	ωB97X-D	132.0
	LC-BLYP	75.6
	Tα-LC-BLYP	109.6
	LC-PBE	76.6
	LC-ωPBE	87.5
	LC-PBE-VV	85.5
	M06	166.0
	M06L	227.1
	M06-2X	100.0
	M06-HF	46.7
	HF	40.7
c (R = *i*Pr)	RI-MP2	99.8
	RI-CC2	148.5
	PBE	176.1
	rPW86-PBE	183.9
	rPW86-PBE-VV	186.0
	PBE0	117.3
	B3LYP	132.7
	BH&H	71.6
	CAM-B3LYP	92.7
	ωB97X	76.8
	ωB97X-D	95.4
	LC-BLYP	56.7
	Tα-LC-BLYP	86.1
	LC-PBE	55.3
	LC-ωPBE	66.9
	LC-PBE-VV	58.9
	M06	150.4
	M06L	196.5
	M06-2X	84.8
	M06-HF	46.3
	HF	31.6

Compound	Approx.	β
d (R = tBu)	RI-MP2	84.1
	RI-CC2	136.1
	PBE	179.2
	rPW86-PBE	188.2
	rPW86-PBE-VV	187.5
	PBE0	115.5
	B3LYP	130.8
	BH&H	68.3
	CAM-B3LYP	91.3
	ωB97X	83.2
	ωB97X-D	107.1
	LC-BLYP	58.4
	Tα-LC-BLYP	84.9
	LC-PBE	56.4
	LC-ωPBE	67.5
	LC-PBE-VV	70.4
	M06	159.5
	M06L	218.0
	M06-2X	81.6
	M06-HF	41.6
	HF	32.9
e (R = Ph)	RI-MP2	152.4
	RI-CC2	239.8
	PBE	266.2
	rPW86-PBE	253.3
	rPW86-PBE-VV	294.2
	PBE0	158.3
	B3LYP	174.7
	BH&H	100.2
	CAM-B3LYP	99.7
	ωB97X	84.4
	ωB97X-D	107.6
	LC-BLYP	53.9
	Tα-LC-BLYP	85.0
	LC-PBE	59.5
	LC-ωPBE	66.7
	LC-PBE-VV	72.8
	M06	207.8
	M06L	277.9
	M06-2X	115.9
	M06-HF	36.4
	HF	21.3

Compound	Approx.	β
f (R = Cy)	RI-MP2	127.7
	RI-CC2	191.0
	PBE	294.3
	rPW86-PBE	348
	rPW86-PBE-VV	253.7
	PBE0	158.0
	B3LYP	175.3
	BH&H	97.8
	CAM-B3LYP	119.2
	ωB97X	105.1
	ωB97X-D	131.0
	LC-BLYP	75.7
	Tα-LC-BLYP	107.2
	LC-PBE	75.7
	LC-ωPBE	88.4
	LC-PBE-VV	79.3
	M06	202.3
	M06L	263.4
	M06-2X	111.6
	M06-HF	59.1
	HF	50.1

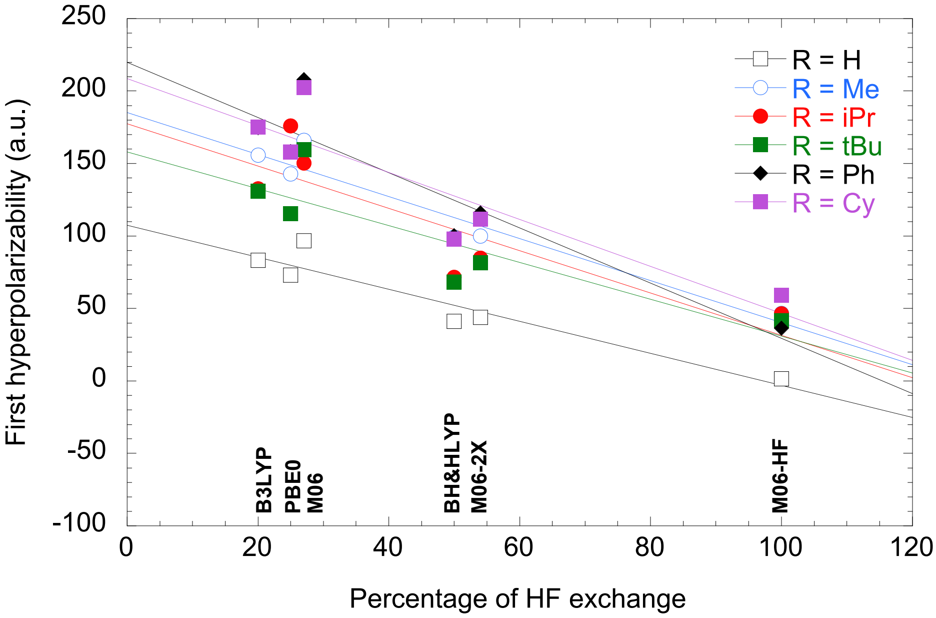


Figure C.11: Evolution of the static hyperpolarizabilities (a.u.) of the series of azobenzenes calculated using various DFAs in combination with the aug-cc-pVTZ basis set, as a function of the percentage of exact Hartree-Fock exchange.

Table C.18: Vector (β_i) and tensor (β_{ijk}) components of the total first hyperpolarizability, as calculated summing all the orbital contributions of the hyperpolarizzability tensors obtained by the PNOC analysis at CAM-B3LYP/6-311++G** level.

Molecule	β_{xxx}	β_{yyx}	β_{zzx}	β_x	β_{yyy}	β_{xxy}	β_{zzy}	β_y	β_{zzz}	β_{xxz}	β_{yyz}	β_z	β_{tot}
a (R = **H**)	-0.0	-0.1	0.0	-0.4	2.9	-10.4	0.1	-22.3	31.4	-155.9	14.5	-330.2	66.2
b (R = **Me**)	-0.4	-2.2	1.9	-2.1	4.2	8.3	-1.1	34.3	-30.7	-160.7	-26.7	-654.4	131.1
c (R = ***i*Pr**)	-2.3	-0.9	0.6	-7.9	-16.0	8.6	8.2	2.0	-15.6	-145.6	17.7	-430.6	86.1
d (R = ***t*Bu**)	-32.2	10.1	-9.5	-94.9	-3.0	0.5	18.4	47.8	-21.7	-170.2	52.4	-418.5	86.3
e (R = **Ph**)	-0.3	-6.3	5.9	-1.9	13.7	6.6	-10.2	30.3	22.7	-148.1	-98.6	-672.3	134.6
f (R = **Cy**)	5.4	6.0	-12.1	-2.2	7.7	5.4	-1.7	34.2	-13.5	-142.5	-39.9	-587.8	117.8
g (R = **Ad**)	2.7	-22.4	14.3	-15.9	-54.3	-18.5	4.1	-206.0	-70.2	-170.0	4.2	-708.0	147.5

Table C.16: Absolute errors relative to MP2 (AE(%) = $|\beta_{DFT} - \beta_{MP2}|/\beta_{MP2} \times 100$) on the total first hyperpolarizabilities of *cis* isomers, as calculated for selected hybrid functionals.

DFA	R = H	R = Me	R = *i*Pr	R = *t*Bu	R = Ph	R = Cy	Mean AE
PBE0	52.4	38.7	43.1	109.4	24.2	23.7	48.6
B3LYP	61.9	30.2	56.1	57.8	14.2	44.0	37.3
BH&H	4.8	65.5	0.5	14.9	55.2	23.4	27.4
M06	48.6	18.9	66.3	78.8	4.7	58.4	46.0
M06-2X	2.9	63.3	0.2	0.8	46.5	12.6	21.0
M06-HF	90.5	98.7	53.2	44.9	72.7	53.7	69.0
best DFA	M06-2X	M06	M06-2X	M06-2X	M06	M06-2X	M06-2X

Table C.17: Absolute errors relative to MP2 (AE(%) = $|\beta_{DFT} - \beta_{MP2}|/\beta_{MP2} \times 100$) on the total first hyperpolarizabilities of *cis* isomers, as calculated for selected hybrid range-separated functionals.

DFA	R = H	R = Me	R = *i*Pr	R = *t*Bu	R = Ph	R = Cy	Mean AE
CAM-B3LYP	0.2	3.3	7.1	8.6	34.6	6.7	10.1
LC-BLYP	57.7	36.6	43.2	30.6	64.6	40.7	45.6
LC-PBE	55.0	35.8	44.6	32.9	61.0	40.7	45.0
LC-ωPBE	40.6	26.7	33.0	19.7	56.2	30.8	34.5
ωB97X	16.2	10.1	23.0	1.1	44.6	17.7	18.8
ωB97X-D	17.1	10.6	4.4	27.3	29.4	2.6	15.3
best DFA	CAM	CAM	ωB97X-D	ωB97X	ωB97X-D	ωB97X-D	CAM

C.11 PNOC analysis

Table C.19: Vector (β_i) and tensor (β_{ijk}) components of the total first hyperpolarizability, as calculated through the CPKS method at CAM-B3LYP/6-311++G** level.

Molecule	β_{xxx}	β_{yyx}	β_{zzx}	β_x	β_{yyy}	β_{xxy}	β_{zzy}	β_y	β_{zzz}	β_{xxz}	β_{yyz}	β_z	β_{tot}
a (R = **H**)	0.0	0.0	0.1	0.2	2.8	-10.5	0.2	-22.4	31.3	-155.9	14.4	-330.3	66.2
b (R = **Me**)	0.1	-2.0	2.1	0.8	4.1	8.3	-1.1	33.6	-30.2	-160.5	-26.5	-651.5	130.5
c (R = ***i*P**)	-1.8	-0.7	0.7	-5.4	-16.4	8.5	8.2	0.8	-15.3	-145.7	18.1	-428.8	85.8
d (R = ***t*Bu**)	-31.7	10.1	-9.4	-92.7	-3.2	0.7	18.6	48.4	-21.8	-170.3	52.3	-419.7	86.5
e (R = **Ph**)	0.1	-6.1	6.0	0.1	13.7	6.7	-10.3	30.3	23.1	-148.3	-98.5	-671.2	134.4
f (R = **Cy**)	6.0	6.3	12.1	73.4	8.3	5.4	-1.5	36.6	-12.9	-142.2	-39.2	-582.9	117.7
g (R = **Ad**)	0.8	-22.6	16.6	-16.6	-55.7	-8.3	8.7	-166.1	-68.0	-170.4	0.6	-713.3	146.5

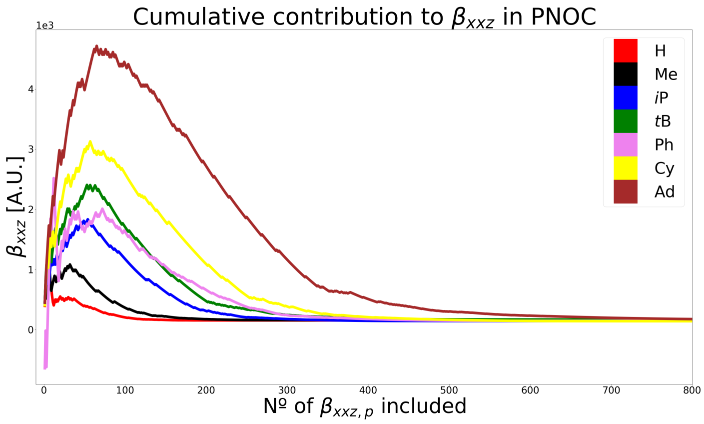


Figure C.12: Accumulation of the first most important 800 orbital contributions of $\beta_{xyz,p}$ for the azobenzene derivatives. The index p indicates the corresponding orbital that is included.

Table C.20: Predominant orbitals and corresponding hyperpolarizability contribution according to PNOC decomposition, ordered by importance. Units are a.u. Calculations done with CAM-B3LYP functional for the azobenzene series. H and L indicate the HOMO and LUMO for each compound. N_H is the total number of occupied MOs.

R=H N_H=48

N_{orb}	β_{zxx}	N_{orb}	β_{zyy}	N_{orb}	β_{zzz}
H	397.0	H-1	137.0	H-3	182.5
H-3	268.5	L+2	-127.4	H-2	181.2
H-2	261.6	H-3	118.7	L+3	-169.9
H-4	240.2	H-2	109.9	L+7	-114.3
L+3	-200.2	L+3	-94.7	L+1	-103.7
L+8	-193.5	H	79.6	H-1	100.4

R=Me N_H=64

N_{orb}	β_{zxx}	N_{orb}	β_{zyy}	N_{orb}	β_{zzz}
H	464.7	H-1	204.5	H-3	221.8
H-3	356.5	H-2	181.5	H-2	212.8
L+4	-300.3	L+2	-149.0	L+4	-156.9
H-2	268.7	L+7	-146.2	H-1	152.6
L+2	-244.7	H-3	142.2	L+2	-145.4
H-1	224.3	H	122.5	L+7	-132.5

R=iPr N_H=96

N_{orb}	β_{zxx}	N_{orb}	β_{zyy}	N_{orb}	β_{zzz}
H	410.9	H-1	263.8	H-3	216.7
H-3	337.4	H-2	215.0	H-2	187.9
H-2	246.0	L+4	-199.6	H-1	143.0
L+5	-231.1	H-3	182.8	L+5	-113.2
H-1	184.5	H-5	157.4	L+8	-111.0
L+6	-181.5	H	148.4	L+4	-107.4

R=tBu N_H=112

N_{orb}	β_{zxx}	N_{orb}	β_{zyy}	N_{orb}	β_{zzz}
H	459.0	H-1	253.9	H-3	222.5
H-3	381.5	H-2	203.2	H-2	209.7
H-2	239.9	H-5	163.5	L+6	-151.8
L+5	-219.2	H-4	162.9	H-1	141.6
H-1	190.5	H	143.0	L+5	-134.6
105	186.9	L+4	-138.1	L+1	-100.8

R=Ph N_H=128

N_{orb}	β_{zxx}	N_{orb}	β_{zyy}	N_{orb}	β_{zzz}
L+2	-638.7	H-1	1030.9	H-1	631.0
H-3	624.6	H-2	997.7	H-2	533.1
L+1	-605.1	L+2	-953.8	L+2	-451.5
H-1	528.4	H-3	942.6	L+3	-434.8
H-2	505.2	L+3	-862.9	H+2	-285.8
H-5	494.8	H+2	-860.3	H-6	272.5

R=Cy N_H=140

N_{orb}	β_{zxx}	N_{orb}	β_{zyy}	N_{orb}	β_{zzz}
H	408.4	H-1	327.0	H-3	214.1
H-3	332.7	H-2	286.2	H-2	193.2
H-2	265.1	H-3	224.2	H-1	158.6
H-1	207.8	H-5	213.8	L+5	-151.1
L+5	-185.3	H-16	198.8	H-16	149.2
H-7	169.9	L+5	-197.4	H-13	148.8

R=Ad N_H=196

N_{orb}	β_{zxx}	N_{orb}	β_{zyy}	N_{orb}	β_{zzz}
H	448.86	H-1	330.88	H-3	229.41
H-3	352.31	H-2	287.14	H-2	205.59
H-2	265.25	H-3	208.47	H-1	175.41
H-1	235.28	H-5	207.32	H-9	158.35
H-6	224.77	H	189.34	H-12	157.32
H-7	205.63	H-7	184.2	H-15	134.75

Table C.21: Comparison between the predominant orbital contribution of the hyperpolarizability (in a. u.) tested for the functionals CAM-B3LYP and PBE performed on the substituents R=*i*Pr and R=Ph. *H* indicates the number of the HOMO orbital for each compound while *L* indicates the corresponding LUMO.

R=iP N_H=96

CAM-B3LYP		PBE	
N_{orb}	β_{zxx}	N_{orb}	β_{zxx}
H	410.9	H-3	561.7
H-3	337.4	H-4	514.8
H-2	246.0	L+3	-405.5
L+4	-231.1	L+2	-357.6
H-1	184.5	L+1	-335.4
L+5	-181.5	H	268.4

R=Ph N_H=128

CAM-B3LYP		PBE	
N_{orb}	β_{zxx}	N_{orb}	β_{zxx}
L+2	-638.7	L+2	-918.82
H-3	624.6	H-3	862.83
L+1	-605.1	L+1	-795.17
H-1	528.4	H-2	685.38
H-2	505.2	H-4	608.30
H-5	494.8	H-1	593.97

Bibliography

[1] Gould, T. 'Diet GMTKN55' offers accelerated benchmarking through a representative subset approach. *Phys. Chem. Chem. Phys.* **2018**, *20*, 27735–27739.

[2] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[3] Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[4] Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better.

[5] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields.

[6] Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

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

[8] 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.

[9] Vydrov, O. A.; Scuseria, G. E. Assessment of a long-range corrected hybrid functional. *J. Chem. Phys.* **2006**, *125*, 234109.

[10] Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. A long-range correction scheme for generalized-gradient-approximation exchange functionals. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

[11] Besalú-Sala, P.; Sitkiewicz, S. P.; Salvador, P.; Matito, E.; Luis, J. M. A new tuned range-separated density functional for the accurate calculation of second hyperpolarizabilities. *Phys. Chem. Chem. Phys.* **2020**, *22*, 11871–11880.

[12] Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.

[13] Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[14] Sitkiewicz, S. P.; Ramos-Cordoba, E.; Luis, J. M.; Matito, E. How Many Electrons Does a Molecular Electride Hold?

[15] Luchini, G.; Alegre-Requena, J.; Funes-Ardoiz, I.; Paton, R. GoodVibes: automated thermochemistry for heterogeneous computational chemistry data [version 1; peer review: 2 approved with reservations]. *F1000Research* **2020**, *9*.

[16] Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chemistry – A European Journal* **2012**, *18*, 9955–9964.

[17] Li, Y.-P.; Gomes, J.; Mallikarjun Sharada, S.; Bell, A. T.; Head-Gordon, M. Improved Force-Field Parameters for QM/MM Simulations of the Energies of Adsorption for Molecules in Zeolites and a Free Rotor Correction to the Rigid Rotor Harmonic Oscillator Model for Adsorption Enthalpies. *J. of Phys. Chem. C* **2015**, *119*, 1840–1850.

Appendix D

Supporting information of Chapter 7

D.1 Definition of the geometrical parameters

D.1.1 Bond length alternation

The bond length alternation (BLA) along the conjugated polyenic bridge connecting the donor and acceptor units is calculated as:

$$BLA = \frac{1}{N-2} \sum_{i=1}^{N-2} (-1)^{i+1} (d_{i+1,i+2} - d_{i,i+1}) \tag{D.1}$$

where N is the total number of carbon atoms in the π-conjugated chain and $d_{i,j}$ is the interatomic distance between carbons i and j. According to the 2-state resonance picture schematized in Figure D.1 for **D3**, a negative (positive) value of the BLA calculated along the C_1-C_8 chain indicates that the ground-state electronic structure is dominated by structure 1 (2). A value equal to zero indicates that the two resonance forms equivalently contribute to the ground-state.

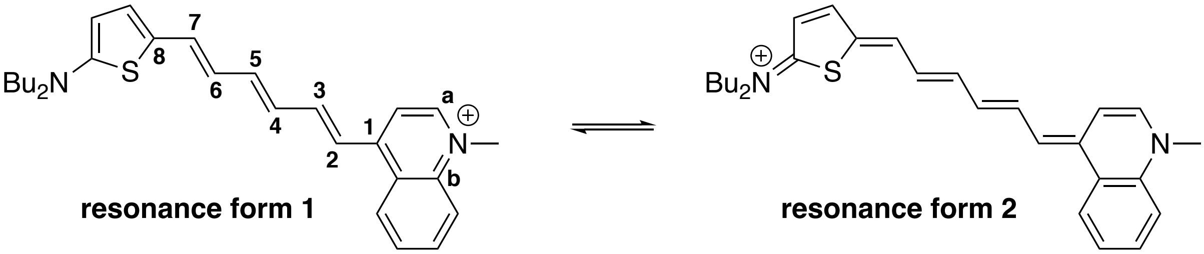


Figure D.1: Resonance structures of **D3** and labels of atoms used to calculate the bond length alternation.

D.1.2 Torsional angles

Table D.1: Definition of the dihedral angles θ_1-θ_8 schematized in Figure D.2b.

Angle	atoms
θ_1	S-C-N-C
θ_2	C-1-2-3
θ_3	6-7-8-S
θ_4	1-2-3-4
θ_5	5-6-7-8
θ_6	2-3-4-5
θ_7	4-5-6-7
θ_8	3-4-5-6

D.2 Derivation of an optimal force field for the D3 dye

D.2.1 Iterative parameterization of bond lengths and torsional potentials

At the DFT level, the BLA along the polyenic bridge amounts to -0.043 Å while the original GAFF (Generalized Amber Force Field) relaxed geometry yields a value of -0.054 Å due to the limited number of atom types used (ca and c2). Besides, the planarity of the DFT structure is not well reproduced, with out-of-plane deviations of up to 14°. These discrepancies in the molecular structure with respect to DFT expectedly translate into poor estimations of the first hyperpolarizability, namely $\beta_{||}(\text{GAFF}) = 2.269\ 10^5$ a.u. against $1.924\ 10^5$ a.u. with the reference DFT geometry.

To achieve a more accurate description of the molecular geometry, we derived consistently new parameters for both bonds and torsional potentials following an iterative approach using a simplified structure of the **D3** molecule, in which butyl chains were removed from the chromophore (Figure D.2). Practically, several new atom types were added to the original GAFF (Figure S2a) to properly reproduce the bond lengths (and thus, the BLA) and the torsional potentials around the dihedral angles θ_1-θ_5, fitted on relaxed potential energy scans performed at the M06-2X/6-311G(d) level in the gas phase (Figure D.3). Note that the torsional potentials around the θ_6-θ_8 angles were not reparameterized and used as defined in the standard version of GAFF in all MD simulations. Converged parameters were used as starting point to describe the whole **D3** molecule, including butyl chains. A final refinement of the bonds was done so as to reproduce accurately the DFT geometry of the whole system and proper description of the dihedrals was checked.

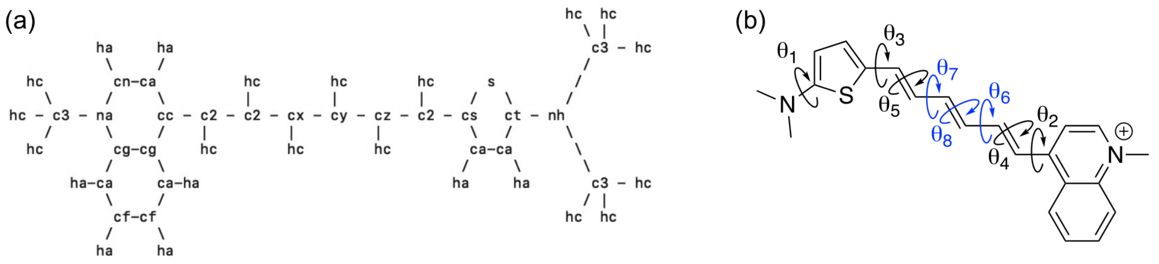
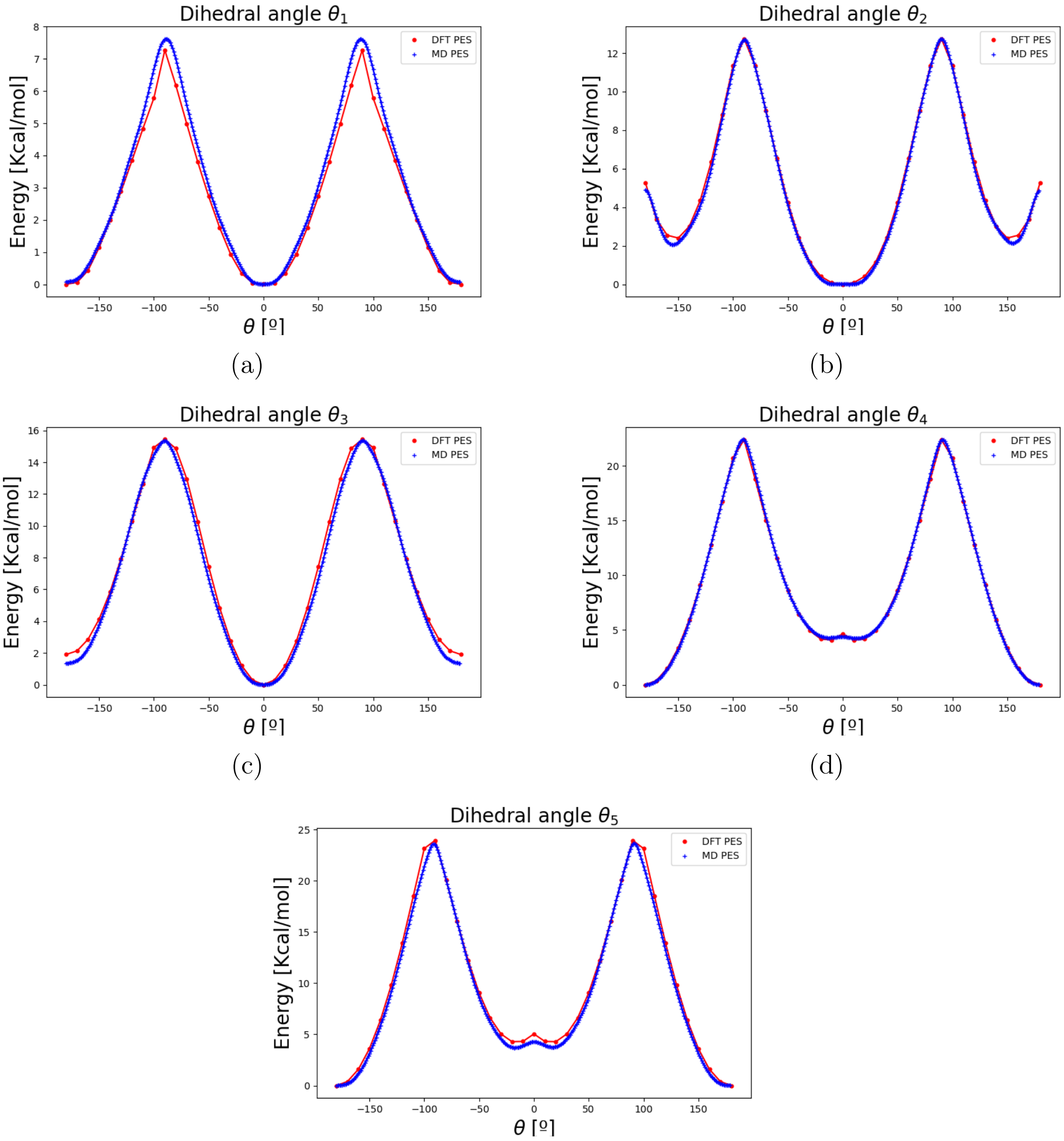


Figure D.2: Geometry of the fragment employed in the optimization of the AMBER force field. (a) Atom types defined in the modified GAFF, (b) The reparametrized dihedrals are θ_{1-5}. The torsional potentials around θ_6-θ_8 (in blue) were not reparameterized.



(a) (b)

(c) (d)

(e)

Figure D.3: Relaxed potential energy scans calculated at the M06-2X/6-311G(d) level (red) and force field fitted potentials (blue).

D.2.2 Assessment of the quality of the modified force field

In order to assess the quality of the reparameterized force field, the geometry of the **D3** molecule was optimized using molecular mechanics (MM) with the original and system-specific derived force fields and compared to the geometry optimized using the reference DFT level. Results gathered in Table D.3 show that MM calculations using the reparameterized GAFF reproduce well the BLA along the π-conjugated bridge, while they give rise to a slightly less planar structure compared to DFT, with differences in the θ_2 and θ_5 dihedrals of the order of 10°. Note that the deviation from planarity resulting from the MM optimization is canceled out by the subsequent MD simulations, which predict a strictly planar average structure (Table D.4).

Table D.2: Values of the bond distances d_{ij} and BLA (Å), and of dihedrals (degrees) optimized at the DFT and MM levels with the original and reparameterized GAFF. Absolute differences with DFT values are also reported.

Parameters	DFT	MM (GAFF rep.)	Diff.	MM (GAFF orig.)	Diff.
d_{12}	1.424	1.423	-0.001	1.506	0.082
d_{23}	1.372	1.373	0.001	1.334	-0.038
d_{34}	1.411	1.411	0.000	1.332	-0.079
d_{45}	1.372	1.374	0.002	1.330	-0.042
d_{56}	1.412	1.411	-0.001	1.331	-0.081
d_{67}	1.370	1.371	0.001	1.333	-0.036
d_{78}	1.415	1.416	0.001	1.490	0.075
BLA	-0.043	-0.041	0.002	-0.054	-0.097
θ_1	1.4	4.6	3.3	-4.6	3.2
θ_2	2.8	14.0	11.2	16.9	14.1
θ_3	-178.9	-176.8	2.0	-179.9	1.1
θ_4	-179.3	179.7	0.4	179.8	0.6
θ_5	179.9	171.8	8.0	180.0	0.1
θ_6	-179.3	-179.3	0.1	-179.3	0.0
θ_7	179.8	179.0	0.8	179.9	0.1
θ_8	179.9	179.4	0.5	179.9	0.0

To assess the impact of the geometry mismatches on the NLO responses, $\beta_{||}$ has been computed at the IEF-PCM/M06-2X/6-311+G(d) level by using both the DFT and MM geometries (Table D.3). Using the MM structure optimized with the original GAFF provides a $\beta_{||}$ value overestimated by more than 15%, while the overestimation is reduced to about 7% with the reparameterized force field. Note that fixing all dihedral angles to their DFT-optimized values only slightly reduces the overestimation of $\beta_{||}$, indicating that the difference can only be partly ascribed to the deviation from planarity of the MM geometry.

Table D.3: Static $\beta_{||}$ values (a.u.) calculated at the IEF-PCM/M06-2X/6-311+G(d) level in chloroform using the DFT and MM geometries, and using the MM geometry while fixing all dihedral angles listed in Table D.2 to their DFT-optimized values. Deviations (in %) with respect to values computed using the DFT geometry are listed in the last column.

Geometry	$\beta_{\|\|}(0;0,0)$	% wrt DFT
DFT	1.92E+05	-
MM (GAFF original)	2.27E+05	15.2
MM (GAFF reparameterized)	2.08E+05	7.3
MM (GAFF rep. + DFT dihedrals)	2.05E+05	6.0

D.2.3 Assessment of the quality of the force field used for chloroform

The quality of the solvent description using standard GAFF parameters and DFT-derived charges has been verified by comparing the calculated mass density with the experimental one (1.468 g/cm^{-3}). For this purpose we performed an NPT simulation on 1600 chloroform molecules at 1 atm and 298.15 K obtaining a density of 1.415 g/cm^{-3}.

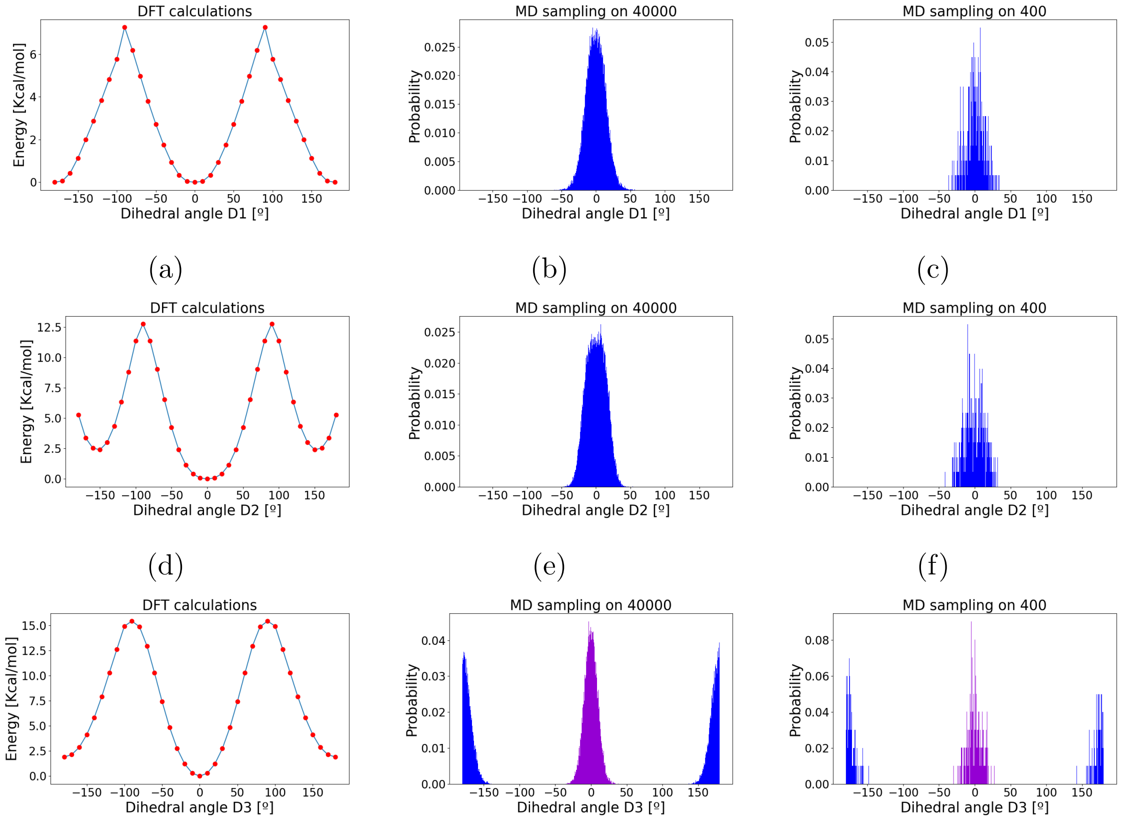
D.3 MD simulations

D.3.1 Probability distributions of geometrical parameters

The distributions of the dihedral angles along the π-conjugated linker, of the average bond length alternation (BLA) and of the average distance (d_{NI}) between the counterion I$^-$ and the nitrogen of the quinolinium group have been calculated for the 20000 frames extracted from each NVT simulation starting at two different initial values of the dihedral θ_3, *i.e.* for 40000 geometrical structures. These distributions are compared in Figure D.4 with those obtained using the 400 geometries used for computing the NLO properties. The average values of the geometrical parameters, extracted from the full and reduced sets of structures, are reported in Table D.4. The correlation of the results show that the 400 snapshots selected for calculation of the NLO properties are representative of the dynamics of the system.

Table D.4: Values of geometrical parameters (BLA and d_{NI} in Å, and dihedral angles θ_i in degrees) averaged over 40000 structures and over the 400 snapshots selected for calculation of the NLO properties.

Property	MD (40000 snapshots)	MD (400 snapshots)
BLA	-0.042 ± 0.022	-0.045 ± 0.023
d_{NI}	4.56 ± 0.48	4.57 ± 0.47
θ_1	-0.4 ± 14.1	0.4 ± 12.9
θ_2	0.0 ± 14.1	-0.7 ± 14.6
θ_3 (sim 1)	180.1 ± 11.1	180.1 ± 10.1
θ_3 (sim 2)	0.0 ± 9.4	-0.3 ± 10.5
θ_4	180.0 ± 8.9	180.4 ± 8.8
θ_5	180.0 ± 10.2	179.4 ± 10.5
θ_6	180.0 ± 6.3	180.3 ± 6.5
θ_7	180.0 ± 6.1	179.9 ± 6.1
θ_8	180.0 ± 6.4	180.2 ± 6.3

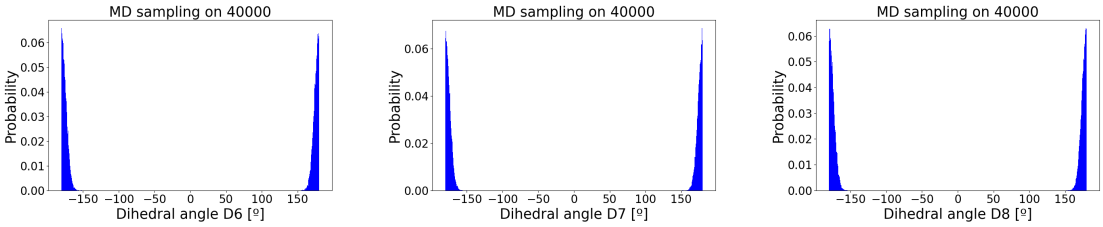


(a) (b) (c)

(d) (e) (f)

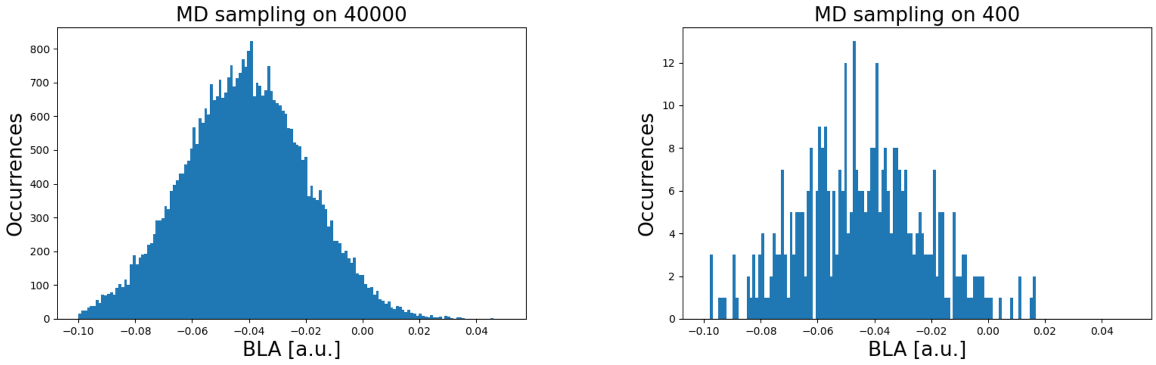
(g) (h) (i)

Figure D.4: Potential energy curves (left) and probability distributions associated to the rotation of the dihedrals θ_1, θ_2 and θ_3 of **D3** obtained using 40000 structures (middle) and the selected 400 structural snapshots (right) extracted from the MD trajectories. The two colors in the distributions of θ_3 (plots (h) and (i)) correspond to two different trajectories starting from $\theta_3 = 0°$ (purple) and $\theta_3 = 180°$ (blue).



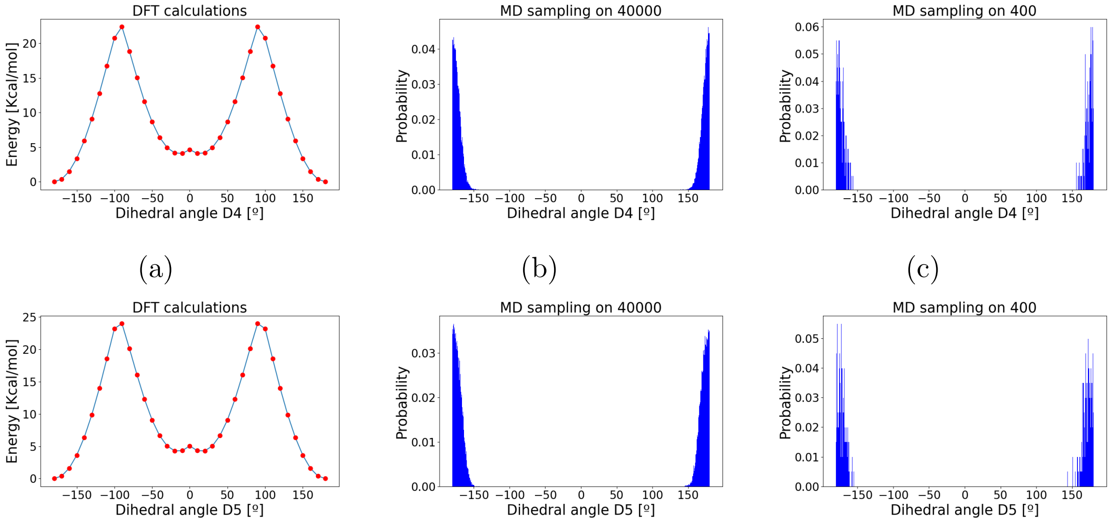
(a) θ_6 (b) θ_7 (c) θ_8

Figure D.6: Probability distributions associated to the rotation around the internal dihedrals θ_6, θ_7 and θ_8 of **D3**.



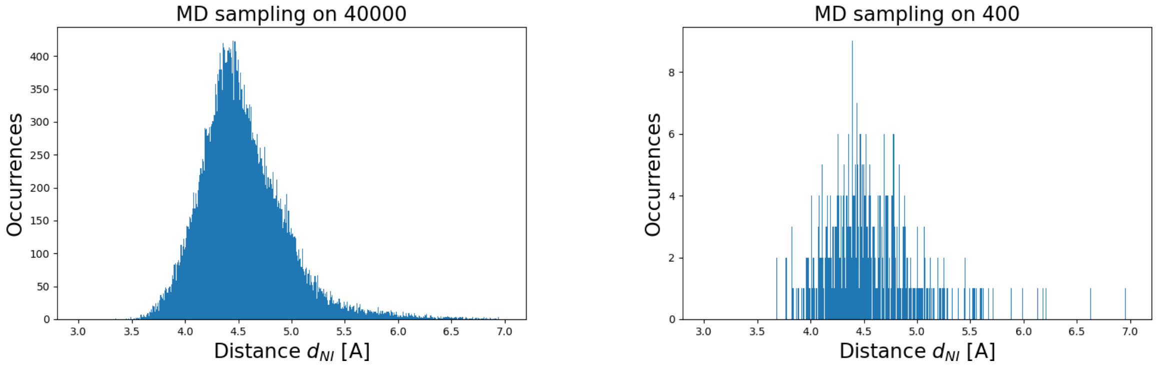
(a) (b)

Figure D.7: Probability distributions of the BLA along the conjugated bridge of **D3**, obtained using 40000 structures (left) and the selected 400 structural snapshots (right) extracted from the MD trajectories.



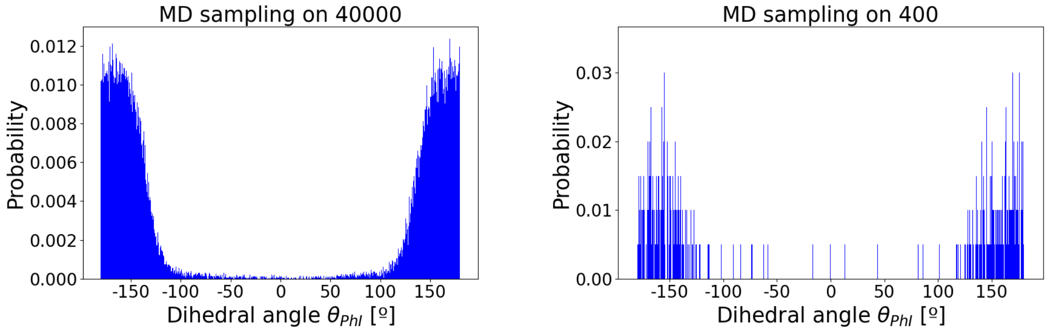
(d) (e) (f)

Figure D.5: Potential energy curves (left) and probability distributions associated to the rotation of the dihedrals θ_4 and θ_5 of **D3** obtained using 40000 structures (middle) and the selected 400 structural snapshots (right) extracted from the MD trajectories.



(a) (b)

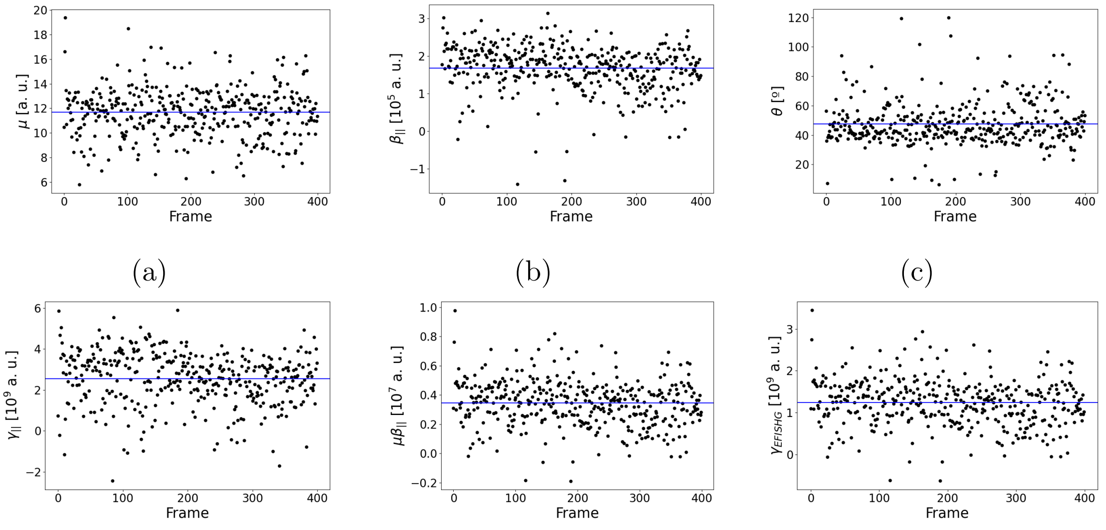
Figure D.8: Probability distributions of the distances between the nitrogen of the quinolinium group and the iodine atom in **D3**, obtained using 40000 structures (left) and the selected 400 structural snapshots (right) extracted from the MD trajectories.



(a) (b)

Figure D.9: Probability distributions of the dihedral angle θ_{PhI} = (I-a-N-b) (see Figure D.1 for atom labels) in **D3**, obtained using 40000 structures (left) and the selected 400 structural snapshots (right) extracted from the MD trajectories.

D.3.2 Time evolution of the EFISHG properties over the MD trajectories



(a) (b) (c)

(d) (e) (f)

Figure D.10: Time evolution of the EFISHG properties along the MD trajectories.

D.3.3 Structure - NLO properties relationships

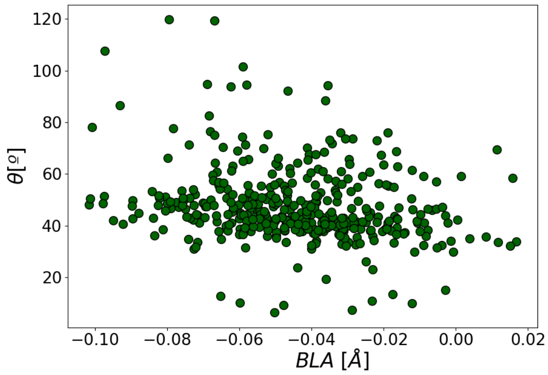


Figure D.11: Evolution of the θ angle (degrees) between the $\vec{\mu}$ and $\vec{\beta}$ vectors with respect to the bond length alternation (Å).

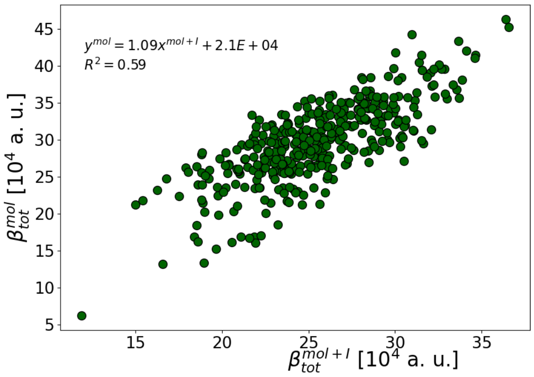
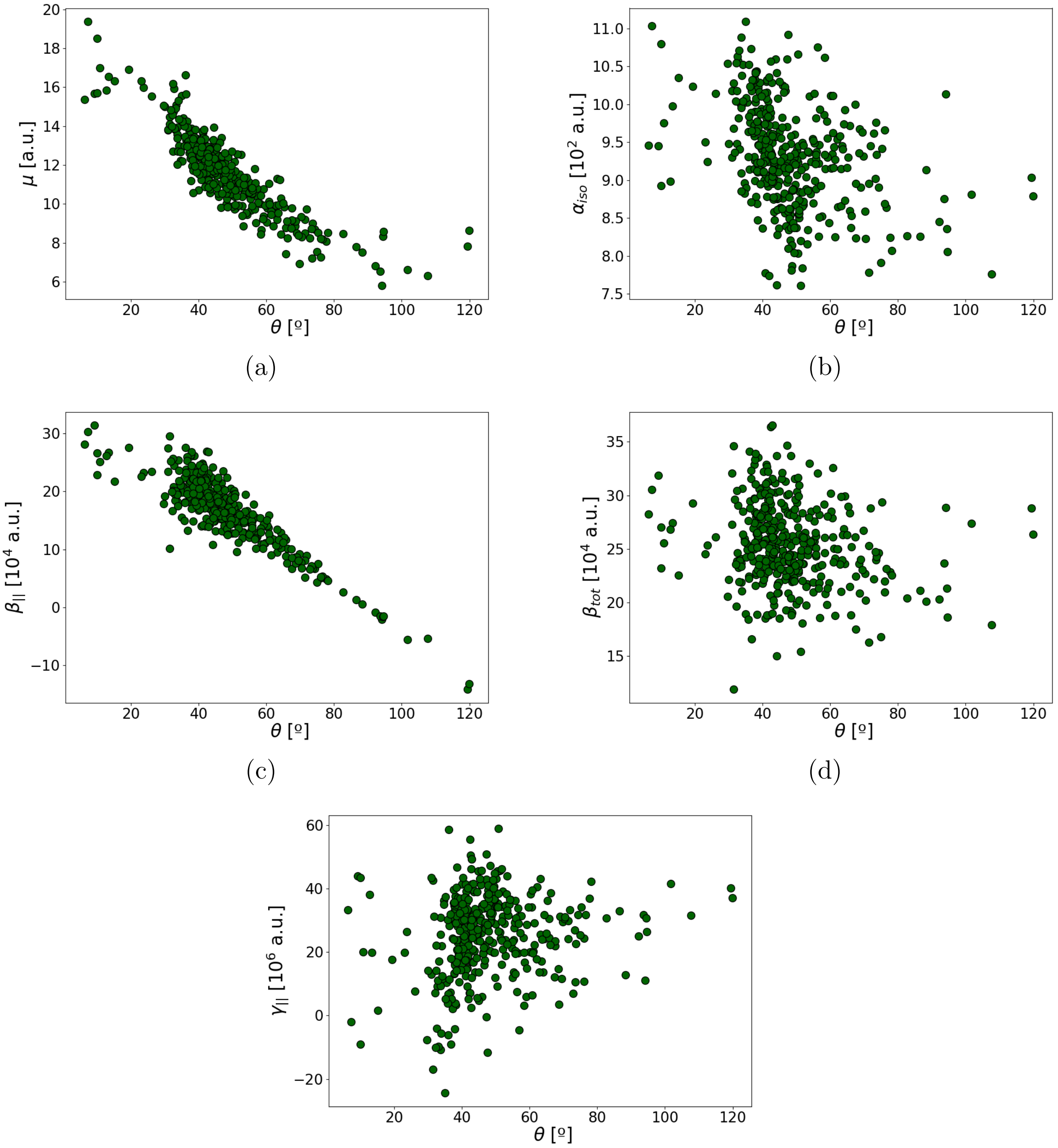


Figure D.12: Evolution of β_{tot} values (in 10^4 a.u.) calculated in presence of the iodide with respect to those calculated for the same geometries of the chromophore after removing the anion.

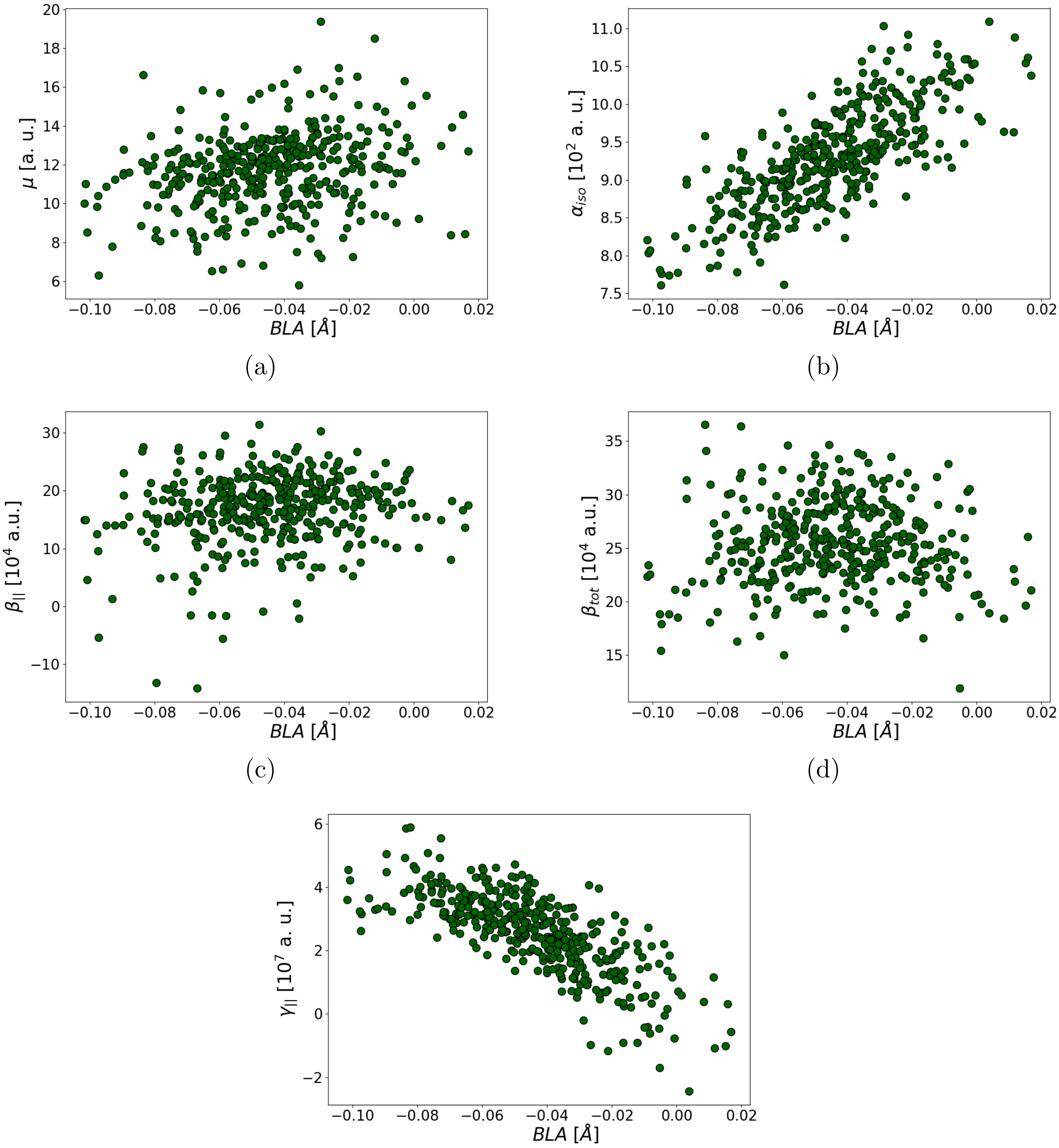


(a) (b)

(c) (d)

(e)

Figure D.13: Evolution of the optical quantities with respect to the θ angle (degrees) between the $\vec{\mu}$ and $\vec{\beta}$ vectors: (a) dipole moment (a.u.), (b) isotropic polarizability (α_{iso}, a.u.), $\beta_{||}$ (10^4 a.u.), β_{tot} (10^4 a.u.), and $\gamma_{||}$ (10^7 a.u.).



(a) (b)

(c) (d)

(e)

Figure D.14: Evolution of the optical quantities with respect to the bond length alternation (Å): (a) dipole moment (a.u.), (b) isotropic polarizability (α_{iso}, a.u.), $\beta_{||}$ (10^4 a.u.), β_{tot} (10^4 a.u.), and $\gamma_{||}$ (10^7 a.u.).

D.4 Assessment of the DFT approximation for computing NLO properties

D.4.1 Performance of M06-2X with respect to MP2

To assess the reliability of the selected exchange-correlation functional, the static first hyperpolarizabilities (in vacuum) of compounds of series **D** were calculated at the M06-2X/6-311+G(d) level and compared to MP2 calculations carried out with the same basis

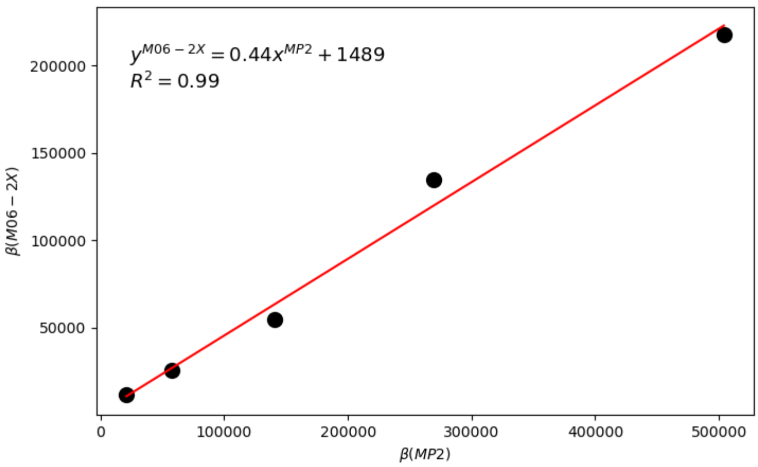


Figure D.15: Comparison of static first hyperpolarizabilities (equation D.2, in a.u.) of compounds of series **D**, computed at the M06-2X and MP2 levels using the 6-311+G(d) basis set.

set. Comparisons were made considering the norm of the static beta vector, defined as:

$$\beta = \beta_{tot} = \frac{1}{5}\sqrt{\beta_x^2 + \beta_y^2 + \beta_z^2} \tag{D.2}$$

As shown in Figure D.15, the very good correlation observed for the two sets of values validates the suitability of the M06-2X exchange-correlation functional for the series of investigated systems.

D.4.2 Impact of the basis set

Three basis sets of increasing size (6-311+G(d), aug-cc-pVDZ and aug-cc-pVTZ) have been tested for molecules of series **D**, by calculating their static first hyperpolarizabilities at the M06-2X level. The results displayed in Figure D.16 show that the two smaller basis sets (6-311+G(d) and aug-cc-pVDZ) provide similar results to the ones obtained using aug-cc-pVTZ, validating the choice of 6-311+G(d) for calculation of the NLO properties of the dyes.

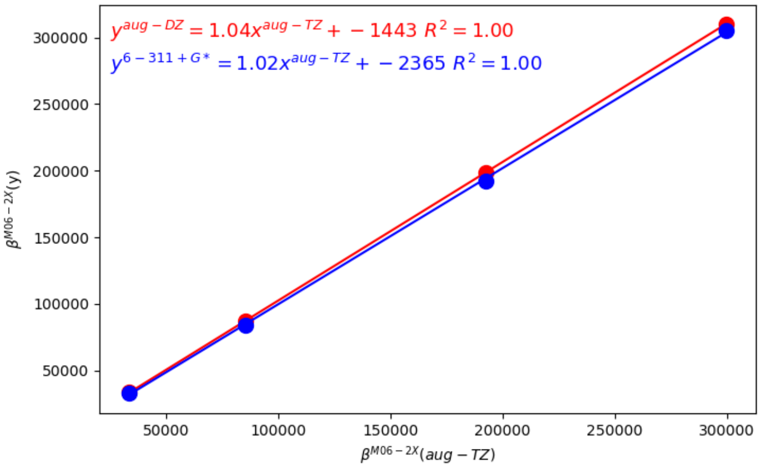


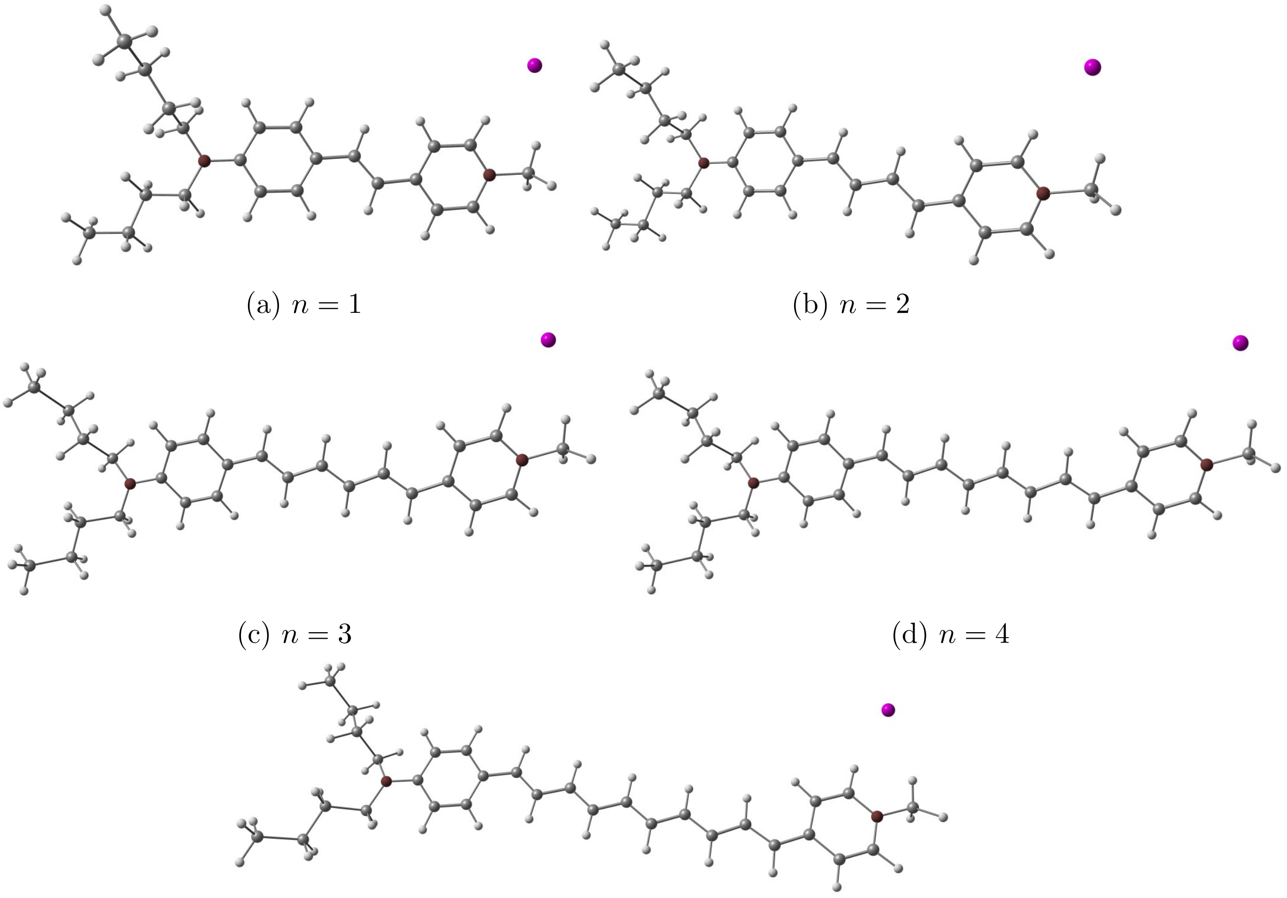
Figure D.16: Comparison of static first hyperpolarizabilities (equation D.2, in a.u.) of compounds **D1-D4**, calculated at the TD-DFT/M06-2X level using different basis sets.

Three basis set used for describing the iodine counterion were also evaluated, by calculating the static first hyperpolarizability of the **D3**/iodine complex. In these tests the 6-311+G(d) basis set was used for the dye. The results collected in Table D.5 show that varying the basis set for iodine does not impact significantly the values of β.

Table D.5: Static first hyperpolarizabilities (equation D.2, in 10^4 a.u.) of the **D3**/iodine complex as a function of the basis set used for describing the iodine anion. These computations have been performed on 24 cpus on a Xeon node with 64 Gb of RAM. The total number N_{basis} of basis functions and the computation time (min) are also provided.

Dye basis	Iodine basis	β	N_{basis}	Time
6-311+G(d)	6-311G(d)	16.8	857	19
6-311+G(d)	aug-cc-pvdz(PP)	17.0	827	18
6-311+G(d)	aug-cc-pvtz(PP)	17.0	850	35

D.5 Optimized structures of the dye/iodine complexes



(a) $n = 1$

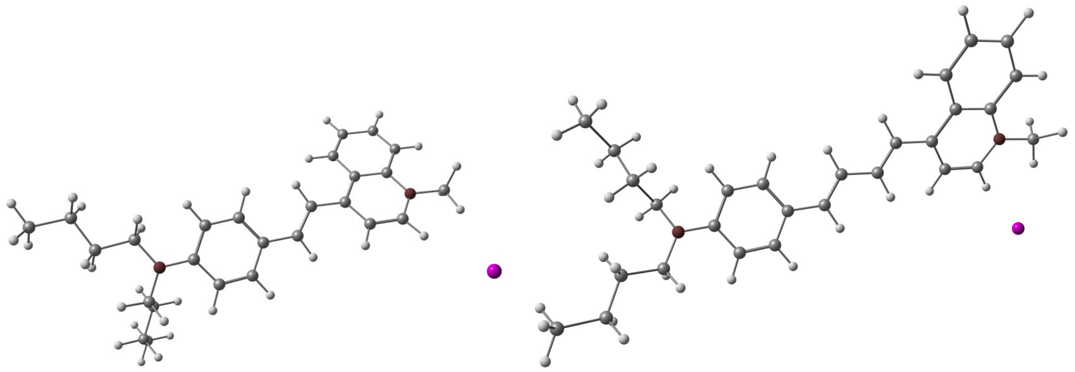
(b) $n = 2$

(c) $n = 3$

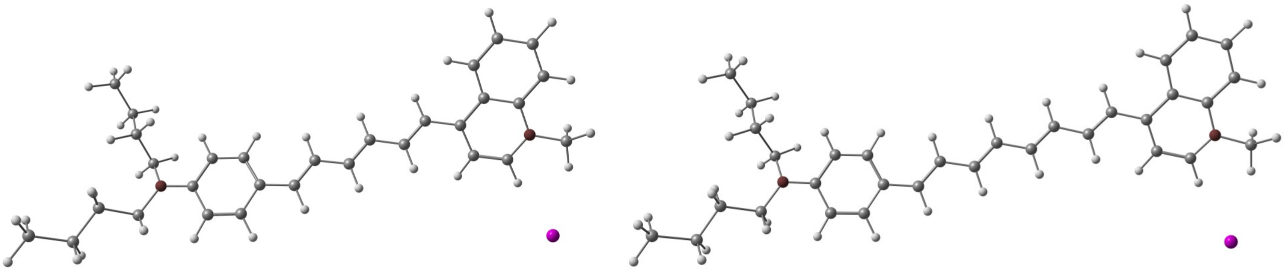
(d) $n = 4$

(e) $n = 5$

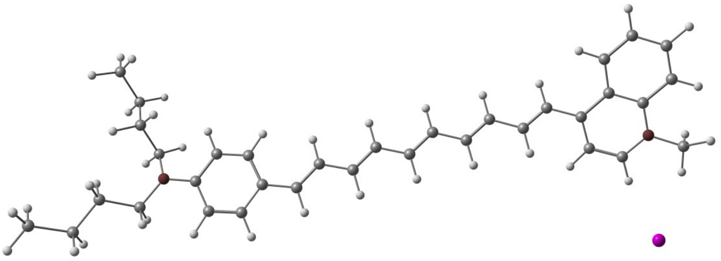
Figure D.17: Optimized structures of the dye/iodine complexes for dyes of series **A**.



(a) $n = 1$ (b) $n = 2$

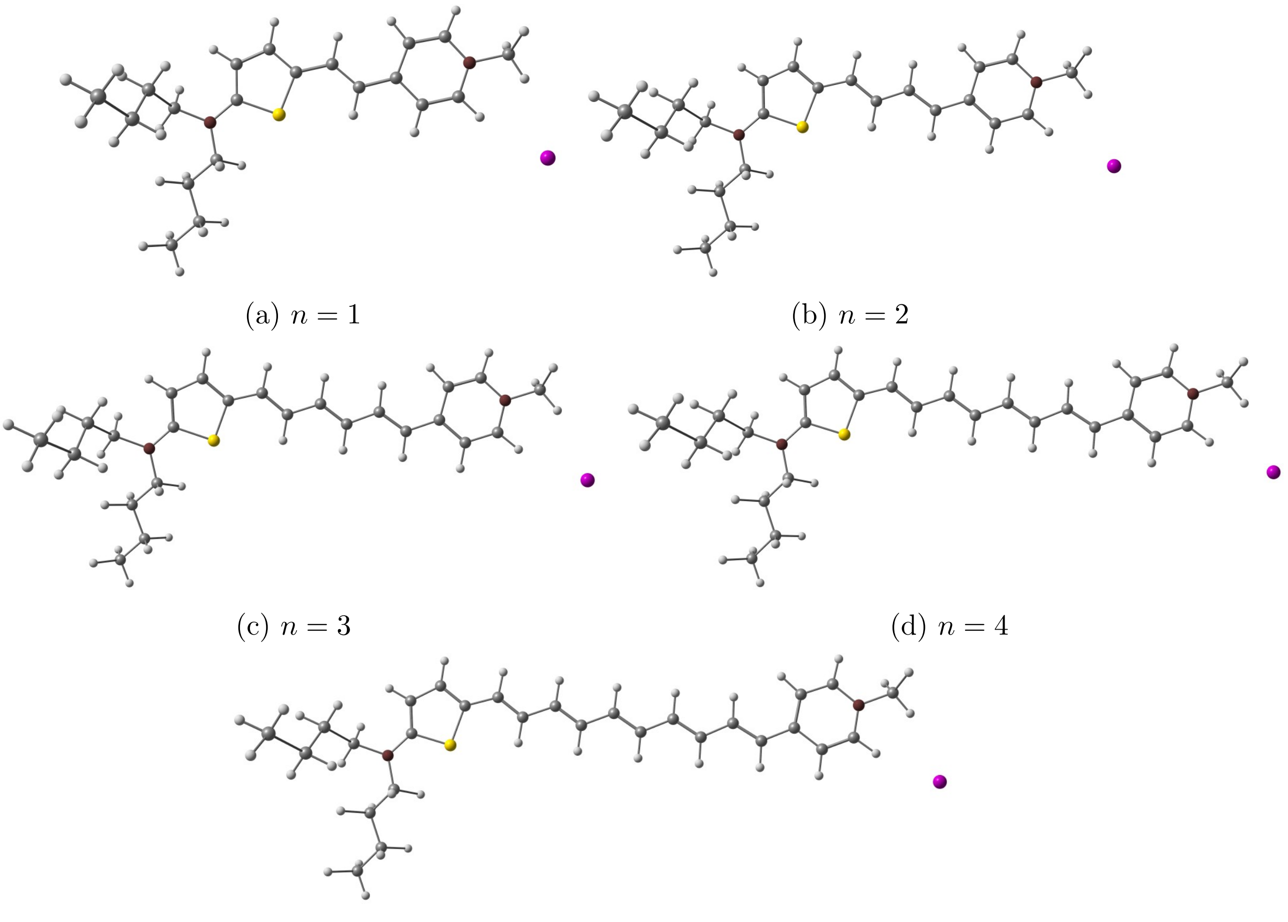


(c) $n = 3$ (d) $n = 4$



(e) $n = 5$

Figure D.18: Optimized structures of the dye/iodine complexes for dyes of series **B**.

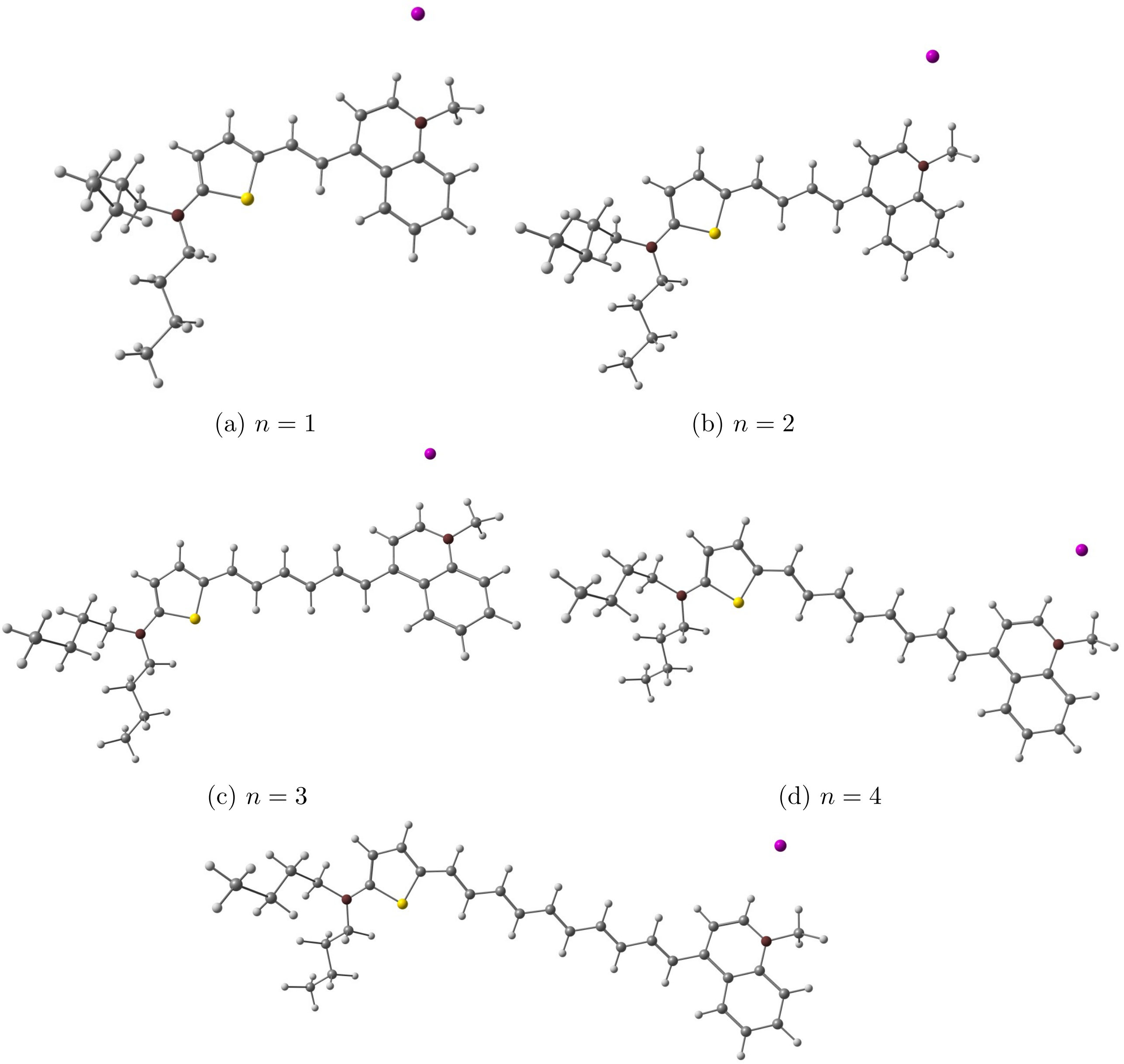


(a) $n = 1$ (b) $n = 2$

(c) $n = 3$ (d) $n = 4$

(e) $n = 5$

Figure D.19: Optimized structures of the dye/iodine complexes for dyes of series **C**.



(a) $n = 1$

(b) $n = 2$

(c) $n = 3$

(d) $n = 4$

(e) $n = 5$

Figure D.20: Optimized structures of the dye/iodine complexes for dyes of series **D**.