eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA

CAMPUS DE EXCELENCIA INTERNACIONAL

Doctoral Thesis

Insights into the Efficient Quantum Chemical Simulations of Nonlinear Optical Properties

Sebastian Paweł Sitkiewicz

2022

Donostia - San Sebastián

Doctoral Thesis
Completed within the Doctoral Program
Theoretical Chemistry and Computational Modelling
at the University of the Basque Country

Insights into the Efficient Quantum Chemical Simulations of Nonlinear Optical Properties

by
Sebastian Paweł Sitkiewicz

Supervised by
Dr. Eduard Matito and Dr. Josep Maria Luis

Tutored by
Prof. Xabier López Pestaña

Donostia-San Sebastián
June 2022

Acknowledgments

The first and the most important person whom I want to thank is my wife, Inna. You have always been by my side as a dearest friend, great support, and a dedicated and caring mother of the two most precious beings in the world — Mikołaj and Bartek. I am grateful to you for all of the strength and efforts that you have put into supporting me and my ambitions. It was a long journey, maybe sometimes rough, but for sure unforgettable, adventurous, and joyful and I am so happy that we took this journey together.

And you little imps... When I am writing these words, you are 3 and 1 years old and you have just gone to sleep at 11 PM, after helping to repot a plant and "accidentally" spilling the soil all over the bedroom. Every day with you feels like a roller-coaster, with no chances to get bored for sure. You fill my life with pure joy and you are my greatest motivation to keep on going and not give up. I hope that one day you get to read these words and invoke some of our best memories from this period, as we have plenty of those... And hey, check out these awesome paintings below that you made with me the other day!

I want to also thank my parents, grandparents, and of course my little sisters, Karolina and Weronika. I know that during these years of my absence I was not the best son, grandson, nor the best brother, but finally, I have come back home and now I can share with you the joy of my little family too

Now, I would like to thank my great supervisors, Dr. Eduard Matito and Dr. Josep Maria Luis. You have been the exceptional mentors and I am grateful that you passed your knowledge and your broad expertise to me. Life as the *extranjeros* was not easy for me and my family but both of you had always made sure that we were always in the safest and the most comfortable place possible.

I want to especially thank Dr. Robert Zaleśny, my "older colleague", for being such a great coworker and even better friend thorough the years, and to Dr. Mauricio Antonio Rodríguez Mayorga, the most romantic and ambitious scientist that I had the chance to meet, who also has stayed a true friend of mine. Many thanks to Dr. Irene Casademont Reig for always being so warm and helpful to us, especially to kiddos. Thanks to the other colleagues from the group with whom I could always share a laugh on a daily basis: Dr. Mireia Via Nadal, Carmelo Naim, Xiang Xu, Aarón Rodriguez, Dr. Rubén Ferradás, and Dr. Eloy Ramos Cordoba. I am also grateful to the Kimika Teorika group and the rest of the DIPC and UPV community, for always being kind to me and hosting me in Donostia during all of these years.

I would like to thank Lizaveta Petrusevich for designing and painting such a creative cover for my Thesis.

And at the very, I would like to acknowledge the funding from the Gobierno Vasco, which I received through the predoctoral grants during all my years in the Doctorate School: PRE_2017_1_0403, PRE_2018_2_0200, PRE_2019_2_0168, and PRE_2020_-2_0015.

„Żyrotrony, Kocurro i rozgwiazda Korona Cierniowa”
("Gyrotrons, Kwazii, and a Crown-of-thorns starfish")
by Mikołaj Sitkiewicz, Wrocław 2022

"Dinozaur, Grrrr!"
("Dinosaur, Grrrr!")
by Bartłomiej Sitkiewicz, Wrocław 2022

Abstract

Nonlinear Optics (NLO) has hugely contributed to many fields of modern science and engineering. NLO processes occur due to the interaction of molecules and materials with intense (optical) electromagnetic fields. NLO stimulated the development of more precise methods for spectroscopic and structural characterization of materials, emerging photon technologies, optical signal processing, optical computers, ultra-fast switchers, laser amplifiers, and many others. Two of the most notable NLO processes are multi-photon absorption, which is the simultaneous absorption of several photons by a molecule, and the sum-frequency generation, the conversion of two optical beams of lower frequency into a one beam of higher frequency. The probability of the NLO processes is usually quantified with the (molecular) NLO properties (NLOPs), such as the first and second hyperpolarizabilities.

Quantum chemistry provides an additional framework for the interpretation of the NLO phenomena and corresponding NLOPs, and computational tools for an efficient design of the new NLO molecules and materials. However, the simulation of NLOPs requires a precise description of many quantum phenomena, with the electron correlation being the most fundamental one. The search for a theoretical method that is accurate and cost-effective for the computation of NLOPs started several decades ago but despite the progress, most methods that are adequate to calculate electronic energies struggle to reproduce NLOPs to a good accuracy. The most prominent NLO responses are often observed for large molecular structures, thus one is often restricted to cost-efficient methods. The most notable examples of such type of methods are Density Functional Approximations (DFAs), which are the most widely used in computational physics and chemistry, due to their the overall good ratios between the cost and efficiency. Unfortunately, for the computation of NLOPs, it has been repeatedly shown that the performance of DFAs may vary dramatically from one system to another. Therefore, there is still a lot of room for improvement in the field of computational chemistry.

This Thesis provides new insights into the computations of NLOPs of molecular systems using modern quantum chemistry methods. It covers the topics related to i) the application and assessment of a broad range of theoretical methods to compute response properties; ii) an analysis of the nature of the nonlinear response and its dependence on the electronic correlation; and iii) the assessment of some DFAs, by uncovering hidden sources of errors affecting the computation of fundamental molecular properties.

In the Chapter 7, we present a novel decomposition scheme for nonresonant NLOPs, namely the Partition of NLOPs Into Orbital Contributions (PNOC). It expresses the electronic response in the terms of the Natural Orbitals (NOs) of the unperturbed molecular system, and for each NO assigns a contribution of the NLOP. PNOC provides an insight into the character of the nonlinear response. For example, using PNOC one can track the changes in the electric response due to the alteration of the chemical structure, such as the formation of the radical structures, or quantifying the contribution to the (hyper)polarizabilities of π-type NOs in conjugated systems. PNOC does not pose a large computational cost and can be applied to most of the electronic structure methods. With

its aid, one can also study the importance of electronic correlation for static NLOPs and, for example, prove how important is the proper description of the multireference character in open-shell systems.

The Chapters 8, 9, and 10 present a coherent analysis on the hidden problem of spurious oscillations affecting computation of vibrational NLOPs using Kohn-Sham DFT (KS-DFT). First, in Chapter 8, we present a benchmark of the DFA performance to simulate nuclear relaxation contributions to static (hyper)polarizabilities of hydrogen-bonded systems. In this test, the best performing DFA yielded average relative errors below 20% for all tested properties. However, more interesting observations were made for the worst performing DFAs, namely ωB97X, M06, and M06-2X. They yielded huge errors, exceeding hundreds of percent, especially for the anharmonic vibrational terms. We show that this is due to the inherited numerical instabilities of these DFAs caused by the emergence of spurious oscillations which affect the derivatives with respect to some vibrational modes. In Chapter 9, we introduce a new computational procedure to detect and quantify the errors arising from the spurious oscillations. It is based on the Fourier spectral analysis and employs several techniques from the digital signal processing. We test dozens of DFAs and show the true extent of their internal grid instabilities. For more than half of the studied functionals we report huge relative errors caused by the spurious oscillations in the derivatives of the energy, dipole moment, and static polarizability, which exceed hundreds and thousands of percent. We also report the highest order of the derivatives that one can safely compute for each DFA, which is greatly limited when the spurious oscillations are present. In Chapter 10, we analyze the dependence of the spurious oscillations on the basis set or the quality of the electronic density. We also propose a substantial improvement of our original algorithm, which greatly reduces the cost while preserving the original accuracy in the determination of errors. Moreover, we discuss the effect of spurious oscillations on other molecular properties through the indirect errors arising from the Self-Consistent Field procedure adopted in KS-DFT. Such indirect errors affect Kohn-Sham molecular orbitals and, therefore, propagate to all energy components and molecular properties. Lastly, we split the errors coming from the spurious oscillations among the exchange and correlation components and unveil which of those are responsible for the spurious oscillations in popular DFAs.

In this Thesis, besides studying NLOPs, we also investigate the topological features of some molecular systems with enhanced NLOPs, namely the molecular electrides. Electrides are novel ionic substances where electrons act as anions. In this work, using first-principle methods, we try to answer a fundamental question related to the electrides, namely if the localized electrons that define the anion of the electride formally constitute as an isolated integer number of electrons. For that purpose, we evaluate several indicators of electronic localization and quantify the number of isolated electrons, as well as directly compute electron distribution functions (EDFs) which determine the probabilities of having a specific number of isolated electrons. Moreover, using the effective oxidation state (EOS) approach, we assign formal oxidation states of such isolated 'centers'. The EOS analysis and the EDFs allow us to formally classify the localized electrons in the studied molecular electrides.

Findings presented in this Thesis expand the current knowledge on some of the crucial aspects of simulations of nonlinear optical properties, and may be of great interest for the development of more efficient and accurate quantum chemical methods.

Resumen

La óptica no lineal (NLO) ha contribuido enormemente a muchos campos de la ciencia y la ingeniería modernas. Los procesos de la NLO ocurren debido a la interacción de moléculas y materiales con campos eléctricos (ópticos) intensos. La NLO dio origen a métodos más precisos para la caracterización espectroscópica y estructural de materiales, así como nuevas técnicas ópticas para la conversión de energía. Los procesos de la NLO más notables son la absorción multifotónica, que es una absorción simultánea de varios fotones por molécula, y la generación de suma de frecuencias, una conversión de dos haces ópticos de menor energía en un solo haz de mayor energía. La probabilidad de los procesos NLO generalmente se cuantifica con las propiedades (moleculares) ópticas no lineales (NLOP), como la primera y la segunda hiperpolarizabilidad.

La química cuántica proporciona un marco adicional para la interpretación de los fenómenos NLO y las NLOP correspondientes y, además, herramientas computacionales para un diseño eficiente de los nuevos materiales y moléculas con NLOP. Por precisa que sea, la simulación de NLOP requiere una descripción precisa de muchos fenómenos cuánticos, siendo la correlación de electrones el más fundamental. En las últimas décadas se ha buscado un método teórico que sea preciso y rentable para el cálculo de NLOP y, a pesar del progreso, la mayoría de los métodos que son adecuados para calcular energías electrónicas no permiten reproducir las NLOP con una precisión equivalente. Las respuestas más prominentes se observan con frecuencia en estructuras moleculares grandes, lo cual limita el uso de métodos computacional a métodos de bajo coste computacional, como la Teoría del Funcional de la Densidad (DFT), que es el método más utilizado en física y química computacionales. Dado que se desconoce el funcional de densidad exacto, se utilizan aproximaciones al funcional de la densidad (DFA). Desafortunadamente, se ha demostrado repetidamente que la precisión de las DFA en el cálculo de las NLOP puede variar drásticamente de un sistema a otro. Por lo tanto, todavía hay mucho margen de mejora en este campo de la química computacional.

Esta tesis proporciona nuevos conocimientos sobre cálculos eficientes de NLOP de sistemas moleculares utilizando métodos modernos de la química cuántica. Abarca los temas relacionados con i) la aplicación y evaluación de una amplia gama de métodos teóricos para calcular propiedades respuesta; ii) un análisis de la naturaleza de la respuesta no lineal y su dependencia de la correlación electrónica; y iii) el desarrollo de métodos computacionales más precisos, en particular DFA, al revelar el origen de los errores que afectan al cálculo de propiedades moleculares fundamentales.

En el capítulo 7, presentamos un esquema de descomposición novedoso para NLOP no resonantes, al que llamaremos partición de NLOP en contribuciones orbitales (PNOC). PNOC expresa la respuesta electrónica en los términos de los Orbitales Naturales (NO) del sistema molecular no perturbado, y para cada NO asigna una contribución a la NLOP. PNOC también proporciona una idea del carácter de la respuesta no lineal. Por ejemplo, utilizando PNOC se pueden monitorizar los cambios en la respuesta eléctrica debido a la alteración de la estructura química, como la formación de estructuras radicales, o cuantificar la contribución a las (hiper)polarizabilidades de los NO de tipo π en sistemas con-

jugados. PNOC no conlleva un gran coste computacional y se puede aplicar a cualquier método de estructura electrónica. Con la ayuda de PNOC también se puede estudiar la importancia de la correlación electrónica para NLOP estáticas y, por ejemplo, demostrar qué tan adecuada es la descripción del carácter multireferencial en sistemas de capa abierta.

Los capítulos 8, 9 y 10 presentan el desconocido problema de las oscilaciones espurias que afectan el cálculo de las NLOP vibracionales usando Kohn-Sham DFT (KS- DFT). Primero, en el Capítulo 8, usamos como punto de referencia el hecho de que las DFA no pueden reproducir correctamente las contribuciones de la relajación nuclear a las (hiper) polarizabilidades estáticas de los sistemas con enlaces por puente de hidrógeno. En esta prueba, el DFA que da mejorer resultados da errores relativos promedio por debajo del 20% en todas las propiedades probadas. Sin embargo, se realizaron observaciones más interesantes para los DFA con peor rendimiento, a saber, ωB97X, M06 y M06-2X. En concreto, estos funcionales dieron grandes errores, superando el 100%, especialmente para los términos vibracionales anarmónicos, y como mostramos, esto se debe a las inestabilidades numéricas heredadas de estos DFA que dan lugar a oscilaciones espurias que afectan a las derivadas de alto orden con respecto a algunos modos vibracionales. En el Capítulo 9 presentamos un nuevo procedimiento computacional para detectar y cuantificar los errores que surgen de las oscilaciones espurias. Se basa en el análisis espectral de Fourier y emplea varias técnicas del procesamiento de señales digitales. Probamos docenas de DFA y mostramos hasta que punto los funcionales dependen del tamaño de la malla de puntos de integración. Para más de la mitad de los funcionales estudiados reportamos enormes errores relativos de las oscilaciones espurias en las derivadas de alto orden de la energía, del momento dipolar y de la polarizabilidad estática, que en algunos casos exceden el 1000%. También calculamos el orden más alto de las derivadas que se pueden estudiar de manera eficiente para cada DFA, que está muy limitado cuando las oscilaciones espurias están presentes. En el Capítulo 10, revelamos más información sobre las oscilaciones espurias, por ejemplo, su dependencia con el conjunto de bases de cálculo o con la calidad de la densidad electrónica. También proponemos una gran mejora de nuestro algoritmo original, que reduce en gran medida el tiempo de costo conservando la precisión original en la determinación de errores. Además, discutimos el efecto de las oscilaciones espurias en otras propiedades moleculares, a través de los errores indirectos que surgen del procedimiento Self-Consistent Field adoptado en KS-DFT.

En esta Tesis, además de estudiar los NLOP, también investigamos las características topológicas de algunos sistemas moleculares con NLOP grandes tales como los electruros moleculares. Los electruros son unos nuevos compuestos iónicos donde los electrones actúan como aniones. En este trabajo, utilizando métodos de primeros principios, tratamos de responder una pregunta fundamental relacionada con los electruros: podemos asignar un número entero de electrones a las regiones del espacio que identificamos como responsables del carácter electrúrico? Para ello, evaluamos varios indicadores de localización electrónica y cuantificamos el número de electrones aislados y calculamos directamente las funciones de distribución de electrones (EDF) que están relacionadas con la probabilidad de tener un número específico de electrones aislados. Además, con el uso de herramientas de la química cuántica, asignamos estados de oxidación efectivos formales (EOS) a tales electrones aislados. El análisis de EOS y los EDFs nos permiten clasificar formalmente los electrones aislados de los electruros moleculares estudiados.

Los hallazgos presentados en esta Tesis amplían el conocimiento actual sobre varios aspectos cruciales de las simulaciones de propiedades ópticas no lineales, que pueden ser de gran interés para el desarrollo de DFA más eficientes y precisos.

List of Abbreviations

Acronym	Description	Acronym	Description
1-RDM	One-Particle Reduced Density Matrix	LOL	Localized-Orbital Locator
AO	Atomic Orbital	LSDA	Local Spin Density Approximation
BCP	Bond Critical Point	MCSCF	Multi-Configurational Self-Consistent Field
BKPT	Bishop-Kirtman Perturbation Theory	MO	Molecular Orbital
CC	Coupled Cluster	NACP	Nuclear Attractor Critical Point
CCP	Cage Critical Point	NCI	Non-covalent Interactions
CI	Configuration Interaction	NLO	Nonlinear Optical / Nonlinear Optics
CP	Critical Point	NLOP	Nonlinear Optical Properties
CPHF	Coupled-Perturbed Hartree-Fock	NNA	Non-Nuclear Attractor
CPKS	Coupled-Perturbed Kohn-Sham	NNACP	Non-Nuclear Attractor Critical Point
DFA	Density Functional Approximation	NO	Natural Orbital
DFT	Density Functional Theory	NR	Nuclear Relaxation
DH	Double Hybrid	OS	Oxidation State
EDF	Electron Distribution Function	PEC	Potential Energy Curve
eff-AOs	effective Atomic Orbitals	PNOC	Partition of NLOPs Into Orbital Contributions
ELF	Electron Localization Function	PT	Perturbation Theory
EOS	Effective Oxidation State	QTAIM	Quantum Theory of Atoms in Molecules
FF	Finite-Field	RCP	Ring Critical Point
FF-NR	Finite-Field Nuclear Relaxation	RMS	Root Mean Square
FIC	Field-Induced Coordinate	RMSE	Root Mean Square Error
FIC1	First-Order Field-Induced Coordinate	RR	Romberg-Rutishauser
FIC2	Second-Order Field-Induced Coordinate	RRMSE	Relative Root Mean Square Error
FMO	Frozen Molecular Orbitals	RSH	Range-Separated Hybrid
G	Grid	SAD	Superposition of Atomic Densities
GGA	Generalized Gradient Approximation	SADMO	Purified Superposition of Atomic Densities
GH	Global Hybrid	SCF	Self-Consistent Field
GTO	Gaussian Type Orbital	SOS	Sum-Over-States
HF	Hartree-Fock	TD-HF	Time-Dependent Hartree-Fock
KS-DFT	Kohn-Sham Density Functional Theory	TD-PT	Time-Dependent Perturbation Theory
LCAO	Linear Combination of Atomic Orbitals	VSCF	Vibrational Self-Consistent Field
LG	Large Grid	ZPVA	Zero Point Vibrational Average

Table of Contents

Part I

Prologue

Chapter 1

Introduction

The emergence of Nonlinear Optics (NLO) has revolutionized many fields of science and engineering and evolved the way we utilize various properties of matter. NLO describes the interaction of matter with very intense optical electromagnetic fields. Effects of such perturbations on the molecular structure are vast and cover a plethora of changes — the electronic densities are polarized, the positions of nuclei and their vibrational motion are altered, and excitations to higher quantum states occur, to name just a few. What distinguishes nonlinear from the linear optics is that the probability of NLO processes depend nonlinearly on the intensity of the applied optical field and very intense optical fields are required. Two of the fundamental types of NLO processes are multi-photon absorption and sum-frequency generation. In the former, several photons are simultaneously absorbed by the molecule, whereas in the latter, several incident photons are combined into one photon of larger energy. The first prediction of the existence of the NLO phenomena, a two-photon absorption, was done in 1931 by Maria Göppert-Mayer [1]. It took nearly 30 years to finally realize it experimentally by Kaiser and Garret [2], when lasers had been already invented [3]. Moreover, at the same time, another landmark discovery was done by Franken, Hill, and Peters, namely, the experimental demonstration of the second-harmonic generation [4].

There are two main types of optical processes, resonant and nonresonant. The resonant processes occur when the frequency of the incident beam matches a particular transition frequency between the ground state and some targeted (electronic or vibrational) state. The resonant NLO phenomena possess high spatial selectivity, which can be greatly utilized in (bio)imaging [5–9] and structural studies of interfaces [10–14]. The nonresonant optical processes take place when the perturbing frequency is different from any of the transition frequencies. Whereas most NLO phenomena can occur both in the resonant and nonresonant frequency regions, the ones related to the static electric fields, that is time-independent ones, are strictly nonresonant. Moreover, in comparison to the equivalent processes in the resonant region, the nonresonant ones have usually lower enhancement factors and worse spatial resolution.

The evolution of computational quantum chemistry has hugely contributed to NLO with the possibility to accurate design new NLO molecules and materials. In addition, it provided the theoretical framework to effectively analyze the relationship between the electronic structure of molecules and their optical response. However, from the point of view of computational chemistry, the molecular NLO properties (NLOPs) also provide an almost unlimited ”playground” for testing the boundaries of modern electronic structure theory methods. The latter is due to the theoretical complexity of these properties, which requires a very accurate description of most quantum effects.

Perhaps, conceptually, the most simple (and universal) method to compute the static

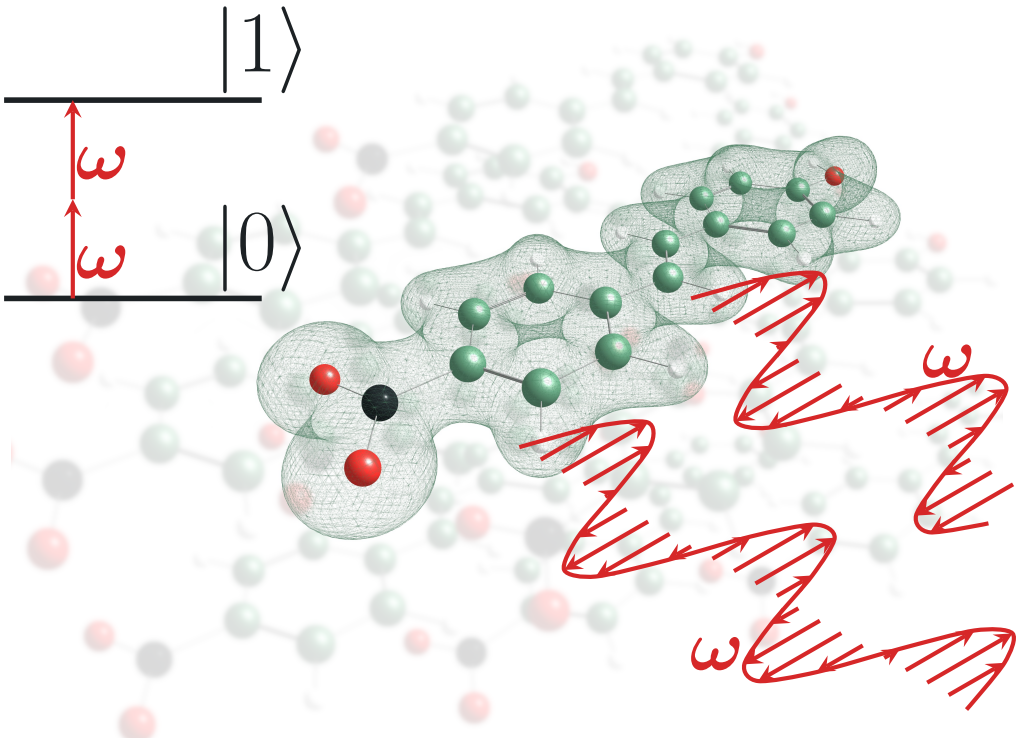

Figure 1.1: Schematic representation of two-photon absorption process where two photons of frequency ω are simultaneously absorbed by the molecule.

electronic NLOPs is the Finite-Field (FF) technique [15, 16]. Static NLOPs are defined as the consecutive derivatives that appear in the Taylor expansion of the total electronic energy with respect to the perturbing electric field $\mathbf{F} = (F_x, F_y, F_z)$ [15, 16],

$$E(\mathbf{F}) = E_0 - \sum_{i}^{x,y,z} \mu_i^0 F_i - \frac{1}{2}\sum_{i,j}^{x,y,z} \alpha_{ij} F_i F_j - \frac{1}{6}\sum_{i,j,k}^{x,y,z} \beta_{ijk} F_i F_j F_k - \frac{1}{24}\sum_{i,j,k,l}^{x,y,z} \gamma_{ijkl} F_i F_j F_k F_l + \cdots, \tag{1.1}$$

where μ_i^0, α_{ij}, β_{ijk}, γ_{ijkl} are the components of the permanent dipole moment, static polarizability, and the first and second hyperpolarizabilities, respectively. In FF, one computes the total energies of a system perturbed with a set of electric fields and calculates the above derivatives numerically via central differences, polynomial fitting, or any other preferred method of numerical differentiation. However, since most of the NLO phenomena originate from the interaction of quantum systems with time-dependent periodic optical fields, it was natural that the theoretical foundations of NLO rooted in the Time-Dependent Perturbation Theory (TD-PT) [17, 18]. TD-PT is one of the most powerful tools of quantum mechanics and allows to interpret a plethora of physical processes, including NLO. However, TD-PT is also a demanding theory, as it has a set of quite strict requirements and assumptions, which reflect the complexity of the studied processes. To accurately simulate the NLOPs and the related NLO phenomena, one must obtain the exact electronic and nuclear parts of the wavefunction (usually treated separately under the Born-Oppenheimer approximation). For the former, one has to fully account for the electronic correlation, whereas for the latter, a proper representation of a vibrational motion is needed. Moreover, to study most of the NLO phenomena with TD-PT, one has to determine the exact electronic and vibrational wavefunctions not only for the ground state but also for (all) excited states. Except for some theoretical models, these conditions are not

fulfilled in the computational studies of many-body systems such as molecules. Therefore, the TD-PT formulas derived for the NLO phenomena, also denoted as Sum-Over-States (SOS), are rigorously valid for Full Configuration Interaction (FCI) wavefunctions, and then they cannot be directly applied to the approximate wavefunctions obtained with Hartree-Fock (HF), Configuration Interaction (CI), truncated Multi-Configurational Self-Consistent Field (MCSCF), or even Coupled-Cluster (CC) theories. Moreover, one usually cannot compute all possible excited states. Nevertheless, the SOS expressions utilizing the approximated wavefunction can still provide some valuable information. Further simplifications of the SOS formula were proposed for analysis purposes. The most notable one is the two-state approach of Chemla and Oudar [19], in which the expressions for the (non)resonant electronic hyperpolarizabilities are reduced by including only the ground state and one excited state. Despite its simplicity, this approach can unveil the most important molecular features responsible for the enhancement of the nonlinear response [20].

A powerful (and more modern) alternative formulation of electronic NLOPs is provided by the response theory, which has been developed both for the exact and approximated wavefunctions [15, 16, 21–25]. Response theory relies on the time-evolution of the observables and does not impose any conditions on the exactness of the Hamiltonian and the wavefunction. One of the most important features of the response theory is that one can obtain properties of the excited states from the ground state response functions, that is, without the explicit computation of the excited state wavefunction [15, 16, 23]. For example, from the linear response functions one can obtain information on the excited state energies and the one-photon transition matrix elements between the ground state and the excited states. The quadratic response functions give access to the properties related to the two-photon absorption from the ground state, as well as the dipole moments of the excited states. The cubic response functions provide information on the three-photon absorption from the ground state, two-photon absorption from the lower excited states to the higher excited states, and the polarizabilities of the excited states.

There are two main formalisms within the response theory, namely the polarization propagator [21, 26] and the quasi-energy derivative approach [27–30]. The former one is based on the expansion of the observables with respect to the perturbation and nowadays is usually combined with the Ehrenfest theorem [16, 21, 24]. The quasi-energy approach is considered to be more attractive and flexible because it provides a uniform formulation for both variational and nonvariational methods [16, 24, 25, 27–29, 31]. The quasi-energy approach can be viewed as an extension of the time-independent response theory to the time-dependent regime. That is, for the static electric fields (the time-independent perturbations), the corresponding NLOPs can be defined as the consecutive derivatives of the total energy system with respect to the applied electric field (see Eq. 1.1). The explicit analytic formulas are obtained by a tedious differentiation of all parameters that define the total energy [15, 16, 23]. In particular, one needs to determine the response of the wavefunction parameters, such as the molecular orbital (MO) coefficients or configuration interaction (CI) expansion coefficients, which are related to the Wigner's $2n + 1$ rule. For variational wavefunctions, one can further simplify the formulas by the means of the Hellmann-Feynman theorem. In the case of the oscillating electric fields (the periodic time-dependent perturbations), one can express frequency-dependent NLOPs as the time-averaged derivatives of the quasi-energy with respect to the amplitude of the field [29]. In such picture, one can also define the time-dependent version of the Hellmann-Feynman theorem.

Over the decades, numerous works have been dedicated to the development of analytical formulas for the response properties obtained from various levels of quantum theories

[16, 21, 25, 29, 31–47]. Undoubtedly, they have been mostly focused on the ab-initio correlated wavefunction methods, such as MCSCF and CC [21, 29, 31, 34–39, 48–63], as these methods define the upper standards of computational methods due to their high accuracy. CC singles, doubles and perturbatively accounted triple excitations, CCSD(T), is usually used as the golden-standard for most of the ground-state molecular properties, including the static NLOPs [23], although, for the excited state properties and the dynamic NLOPs, CCSD(T) is outperformed, for example, by the CC3 method [23]. However, the CC methods scale very quickly with the size of the basis set, which limits their applicability. Even with the cost-reducing approximations, such as resolution of identity (RI), CCSD(T) or CC3 are usually too expensive to study nonlinear response for systems with more than a few dozens of atoms. Still, thanks to the regular development of effective electronic structure methods, the parallelization techniques, and the continuously increasing capabilities of modern computer units, nowadays, we can tackle much larger chemical systems than a decade ago. This is important because the molecular systems that show good NLO capabilities are usually large and, moreover, one should also account for their experimental environment such as the solvent or bulk [64–71]. Therefore, as long term goals one should consider using electronic structure methods with a modest computational cost, but still of reasonable accuracy. For this reason, Density Functional Theory (DFT) along with the Kohn-Sham approach (KS-DFT) have been the preferred method for the simulations of various molecular and bulk properties [72–74].

Given its effectiveness, low cost, and conceptual simplicity, KS-DFT has become the true workhorse of modern computational chemistry [72–74]. Since the exact density functional is unknown, a race for constructing the most accurate and versatile density functional approximation (DFA) is still on. In the recent years, several DFAs that provided accurate electronic energies and geometrical structures have been proposed [75–78]. They partially overcome the main challenges of DFAs, such as the delocalization error [79], dispersion forces [80], and strong correlation [81]. Still, the DFAs' performance in the area of molecular linear and nonlinear optics has been under active assessment for decades [66–68, 82–108]. Most of such works attempt to find the best performing and well-rounded DFA for the evaluation of NLOPs but, unfortunately, there are as many different conclusions as studies. Most of these works rely on the implementation of the response theory in time-dependent DFT (TD-DFT), which allowed to study one- and two-photon phenomena in a cost-effective way. A key difference between the time-dependent HF (TD-HF) and TD-DFT, both approximating the wavefunction with a single Slater determinant, is the inclusion of time-propagation of the approximate exchange-correlation kernel in the latter. Development of the response theory for TD-DFT began with the implementation of linear response functions [48–57], and later was extended to the nonlinear response functions [31, 58–63]. More recently, the computationally flexible procedures to construct arbitrary-order response functions in KS-DFT were proposed [41, 42, 109]. They are based on a powerful computational technique of automatic differentiation, which allows to exactly compute arbitrary-order derivatives of any molecular property, without the need to derive the explicit formulas for a given DFA.

The vibrational contribution to NLO phenomena, in general, requires a more elaborate computational treatment because one has to accurately describe both the electronic correlation and the nuclear motion. One can distinguish three main types of methodologies for the computation of vibrational NLOPs. The first one, a perturbative approach, is Bishop-Kirtman Perturbation Theory [110–112] which originates from TD-PT for exact quantum states. In BKPT, the electronic and vibrational motions are decoupled by means of the Born-Oppenheimer approximation and several other assumptions. This leads to partitioning of the total NLOP into the electronic and vibrational contributions. The latter

ones involve summations in the basis of normal modes and are separately derived for each particular NLO phenomenon. Furthermore, in BKPT, one separates the vibrational contributions into the harmonic and anharmonic components. The second type of methods for vibrational NLOPs are derivative-based and mostly developed for nonresonant processes [112–119]. In these methods, the dependency of the electronic energy and dipole moment with the strength of an external electric field is expressed through a double expansion in electric field and the field-dependent normal coordinates. Some derivative-based methods, such as Bishop-Hasan-Kirtman Finite-Field Nuclear Relaxation (FF-NR) or the Field-Induced Coordinates (FICs), may be used to cost-efficiently simulate the leading (an)harmonic corrections to NLOPs of large molecular systems. The third type of methods for the vibrational NLOPs involves the vibrational self-consistent field (VSCF) procedure [120–122]. In VSCF, one solves a set of single-mode vibrational Hartree equations and as a result obtains a fully anharmonic wavefunction [123, 124]. The corresponding single-mode effective potential is defined taking into account the coupling between the modes. The VSCF results can be further corrected by the vibrational equivalents of electronic post-HF methods, such as the vibrational PT [125–127], CI [128, 129], or CC [130, 131] methods. Then, with such (post-)VSCF wavefunctions one can directly compute the vibrational contributions to NLOPs that already account for the anharmonic effects [123, 124, 131–133].

One of the landmark papers on the analysis of NLOPs is due to Chopra and co-workers [134], who studied orbital contributions to the linear and nonlinear response to static electric fields. Although the decomposed contributions sum up properly, they are dependent on the choice of the coordinate origin, which is popularly named as origin dependency. To alleviate the serious problem of the origin dependency one can fix the coordinate axis at the center of mass or nuclear charge. Another important work was done by Nakano *et al.*, who established the concept of NLOP density derivatives, which are n-order derivatives of the electron density with respect to the external electric field (see Section 4.6) [135–137]. The integrals of these quantities do not add up to the total values of the NLOP, but they are not origin dependent. Later, this analysis was extended to open-shell systems through the usage of the density of effectively unpaired electrons [138–140]. Geskin and Brédas adapted the Mulliken partitioning and defined the atomic contributions to NLOPs using the derivatives of the Mulliken charges with respect to the electric field, and this approach was later extended to dynamic properties within the time-dependent frameworks [141, 142]. Hieringer and Baerends decomposed the first hyperpolarizability into pairwise contributions [143]. Mandado and co-workers attempted to decompose the linear response in terms of field-induced orbitals [144]. Recently, Montilla, Luis and Salvador proposed a genuine origin-independent methodology for the decomposition of the static (hyper)polarizabilities [145]. Unfortunately, the above mentioned approaches are not suitable to compare the performance of quantum chemical methods and analyze the role of the electronic correlation. Most of them cannot be adopted for the more advanced methods or would require a very high computational cost. Decomposing NLOPs into orbital contributions would help identifying pitfalls in electronic structure methods such as the wrong selection of active orbital spaces or the adequate size of orbital basis sets.

This Thesis is a collection of three projects, each devoted to a different aspect of computer simulations of molecular nonlinear optics. Although the nonlinear optical phenomena most often involve at least one time-dependent electromagnetic field, in this Thesis, for the practical reasons, we consider the response to the static electric fields only. The new computational procedures introduced in Chapters 7, 9, and 10 were initially applied for the static NLOPs, nevertheless, they could be extended to the frequency-dependent NLOPs as well.

Chapter 7 discusses the role of electron correlation in the NLOPs using the partitioning scheme developed by the Candidate in the first year of the doctorate studies, the Partition of Optical Properties Into Orbital Contributions (PNOC) [146].

The next three chapters, namely Chapters 8, 9, and 10, form a coherent trilogy devoted to the spurious oscillations in DFAs and how they affect the computation of the vibrational properties. Chapter 8 serves as an introduction explaining the origin and motivation for the main project. The starting point of this trilogy was a collaborative work on the DFAs computations of vibrational contributions to NLOPs of weakly-bonded molecular complexes [108]. From the Candidate's perspective, the most interesting result from this work was the pinpointing the origin of a dramatic failure of several density functional approximations. In a series of subsequent tests, the very source of these errors was finally unveiled — spurious oscillations which were directly related to the accuracy of the numerical integration in KS-DFT. After the identification of the source of the problem, a diagnosis of its full extent had to be done. This is the topic of Chapter 9, the main part of this trilogy, where a new computational protocol to measure spurious oscillations based on the Fourier spectral analysis and techniques from the digital signal processing, is presented. This method allowed us to efficiently analyze the errors arising solely from the spurious oscillations [147]. With its aid, a broad series of tests were carried out for dozens of popular density functional approximations which directly provided the information on their numerical stability. Afterwards, in Chapter 10, a follow-up study exploring the nature of the spurious oscillations was carried out. We analyzed the influence of the spurious oscillations in various parts of the KS-DFT energy terms. Some observations motivated us to adopt further approximations in the electronic structure computations to reduce the cost of the analysis performed with our new method without affecting its accuracy.

Lastly, Chapter 11 has more application-related character and, while not directly involving any computations of NLOPs, addresses an important topological question about a new and promising NLO materials called *molecular electrides*. In these very peculiar chemical systems, highly localized electrons, formally not belonging to any of the atoms, act as the anions of ionic species. Such an unusual electronic structure leads to a set of unique properties, among which are huge NLO responses. However, despite their large potential in many fields of science, their topological classification of their electronic structure is not yet completed. The work presented in Chapter 11 tries to fill this gap [148]. In particular, it discusses the suitability of several ab-initio and DFT methods to describe topological quantities related to electron localization and provides an answer on the most likely number of localized electrons in electride anions. Furthermore, in this work, the first attempt to assign the formal oxidation state of the electride isolated electrons was made.

Chapter 2

Objectives

The aim of this Thesis is to provide new tools and insights for developing more efficient computational protocols for the accurate simulation of the NLOPs of molecular systems, focusing on the assessment of Density Functional Approximations (DFAs). This is achieved by combining different aspects of modern computational chemistry. This Thesis has the following three general goals:

Assessment of the performance of various QC methods in the simulation of the NLOPs. The focus is put on the modern DFAs, which show great cost-effectiveness and are routinely applied in most NLO studies, and the noncovalently bonded complexes, which recently have gained a lot of attention due to their indisputable stabilizing role in all surrounding matter. While in recent years a few DFAs that properly predict the geometrical and energetical features of the molecular complexes have been proposed, their suitability for the description of the vibrational and nonlinear optical response properties of such systems still needs further assessment. Besides simulations of NLOPs, our goal is to put additional emphasis on the topological characterization of a promising new kind of NLO compounds — molecular electrides — and the highly localized (excess) electrons which lead to the enhanced response to electric fields.

Explanation and analysis of the nature of the molecular response to static electric fields and its dependence on the electron correlation. Such a problem is of special interest for the systems with open-shell electronic structures, which are also characterized by large nonresonant hyperpolarizabilities and carry a large computational cost to properly describe their multireference character. Our goal is to aid the analysis proposing a decomposition method operating in the basis of the natural orbitals. Developing a computational tool to provide additional insight into the role of different types of correlation in such systems is of great interest.

Development of a method to unveil a hidden problem of DFAs, namely, the grid-related spurious oscillations that drastically hinder the simulations of vibrational optical properties. These oscillations are directly related to one of the most fundamental approximations adopted in DFT, namely, the numerical integration of exchange-correlation components of the functionals. Since it is almost impossible to bypass the numerical integration in DFT, this problem becomes a big issue for the development of new DFAs. Besides being interested in the physical accuracy of DFAs, one should also be aware of the limitations which originate from the constantly increasing structural complexity of the newly developed functionals. Therefore, our main goal is split into more concrete objectives: The development of a novel computational program for the cost-efficient quantification of the errors in molecular properties arising from the spurious oscillations, and to raise awareness of the limitations of various DFAs, which include the accurate computation of vibrational responses.

Chapter 3

List of Scientific Publications

1. Sebastian P. Sitkiewicz, Mauricio Rodríguez-Mayorga, Josep M. Luis, and Eduard Matito *"Partition of Optical Properties into Orbital Contributions"*, Phys. Chem. Chem. Phys., 2019, **21**, 15380

2. Robert Zaleśny, Miroslav Medved', Sebastian P. Sitkiewicz, Eduard Matito, and Josep M. Luis *"Can Density Functional Theory Be Trusted for High-Order Electric Properties? The Case of Hydrogen-Bonded Complexes"*, J. Chem. Theory Comput., 2019, **15(6)**, 3570

3. Sebastian P. Sitkiewicz, Eloy Ramos-Cordoba, Josep M. Luis, and Eduard Matito *"How Many Electrons Does a Molecular Electride Hold?"*, J. Phys. Chem. A, 2021, **125(22)**, 4819

4. Sebastian P. Sitkiewicz, Robert Zaleśny, Eloy Ramos-Cordoba, Josep M. Luis, and Eduard Matito *"How Reliable are Modern Density Functional Approximations to Simulate Vibrational Spectroscopies"*, J. Phys. Chem. Lett., doi:10.1021/acs.jpclett.2c01278

5. Sebastian P. Sitkiewicz, Robert Zaleśny, Eloy Ramos-Cordoba, Josep M. Luis, and Eduard Matito *"The Conundrum of Numerical Integration Grids in Density Functional Theory"* (in preparation)

Part II

Methodology and Theoretical Foundations

Chapter 4

Quantum Chemical Evaluation of the Nonlinear Optical Properties

4.1 Introduction

This Chapter introduces the methodology of the nonlinear optical response required to follow most of the Results section of this Thesis. These theoretical foundations are shown from the perspective of molecular quantum chemistry, i.e., the microscopic representation of the response to external electric fields. Moreover, most of the weight is put on the nonresonant response, and in particular, response to the static electric fields, as it was the sole regime of fields studied in this Thesis.

Most of the sections of this Chapter are based on Refs [15, 112, 149–151], hence the corresponding notation is also adapted (and unified across this Chapter where it was needed).

4.2 Multipole Expansion

Unless stated otherwise, in this Chapter, the α, β, γ, and δ subscripts denote the x, y, and z Cartesian components, $\mathbf{F}$ is the electric field vector with three Cartesian components, $\mathbf{F} = (F_x, F_y, F_z)$.

In general, the external electric fields $\mathbf{F}$ applied to a molecule may be fully inhomogeneous through the space. In such cases, a set of particles with charges e_i at the positions $\vec{r}_i = (r_{i,\alpha}, r_{i,\beta}, r_{i,\gamma})$ interacts with the associated scalar potential $\phi^{\mathbf{F}}$ according to the multipole expansion [152]:

$$W = q\phi^{\mathbf{F}}(\vec{r}'_O) - \sum_{\alpha} \hat{\mu}_\alpha F_\alpha - \frac{1}{3}\sum_{\alpha,\beta} \hat{\Theta}_{\alpha\beta} F'_{\alpha\beta} - \frac{1}{15}\sum_{\alpha,\beta,\gamma} \hat{\Omega}_{\alpha\beta\gamma} F''_{\alpha\beta\gamma} - \cdots, \tag{4.1}$$

with W being the interaction energy between the molecule and the electric potential, q being the total charge of the system and $\hat{\mu}_\alpha$, $\hat{\Theta}_{\alpha\beta}$, and $\hat{\Omega}_{\alpha\beta\gamma}$ being the dipole, quadrupole, and octupole moment operators, respectively. In the above expansion, F_α is the α-th component of the electric field, $F'_{\alpha\beta}$ is the $\alpha\beta$-th component of the field gradient, and $F''_{\alpha\beta\gamma}$ is the $\alpha\beta\gamma$-th component of the field Hessian (all of them measured at the origin of the coordinate system, $\vec{r}'_O$):

$$F_\alpha = -\left(\frac{\partial \phi^{\mathbf{F}}(\vec{r}')}{\partial \vec{r}'_\alpha}\right)_{\vec{r}'=\vec{r}'_O}, \tag{4.2}$$

$$F'_{\alpha\beta} = -\left(\frac{\partial^2 \phi^{\mathbf{F}}(\vec{r}')}{\partial \vec{r}'_\alpha \partial \vec{r}'_\beta}\right)_{\vec{r}'=\vec{r}'_O}, \tag{4.3}$$

$$F''_{\alpha\beta\gamma} = -\left(\frac{\partial^3 \phi^{\mathbf{F}}(\vec{r}')}{\partial \vec{r}'_\alpha \partial \vec{r}'_\beta \partial \vec{r}'_\gamma}\right)_{\vec{r}'=\vec{r}'_O}. \tag{4.4}$$

It has to be stressed that Eq. 4.1 is valid only for static, i.e., time-independent, electric fields.

4.3 Nonlinear Optical Properties

Molecular nonlinear optical properties (NLOPs) describe the response of molecules in the presence of (intense) radiation electric fields. Such electric fields are time-dependent and

described as the sum of electromagnetic waves with the ω angular frequencies,

$$F_\alpha(t) = \sum_\omega F_\alpha^\omega e^{-i\omega t}\,, \tag{4.5}$$

with F_α^ω being the Fourier amplitudes of the oscillating fields (the summation includes both positive and negative frequencies).

For most of the optical phenomena and related experiments, it is assumed that the molecules are subdued with an homogeneous electric field. This is because the size of the molecules is much smaller than the wavelength of the applied electric field. In such case, the interaction of nuclei and electrons with such fields (see Eq. 4.1) is governed only through the dipole moment, whereas the higher moments do not contribute.

In the quantum mechanical nomenclature, the total Hamiltonian, $\hat{H}$, of the molecule interacting with the classical electric field $\mathbf{F}(t)$, is

$$\hat{H} = \hat{H}_0 - \sum_\alpha \mu_\alpha(t, \mathbf{F}(t)) F_\alpha(t)\,, \tag{4.6}$$

where $\hat{H}_0$ is the (time-independent) Hamiltonian without the presence of the field, $\mu_\alpha(t)$ and $F_\alpha(t)$ are the α-th components of time-dependent dipole moment vector and electric field vector. Time-dependent polarization described with $\boldsymbol{\mu}(t)$ and $\mathbf{F}(t)$ is most generally expressed as the Taylor series in the electric field strength,

$$\begin{aligned}\mu_\alpha(t, \mathbf{F}(t)) = \mu_\alpha^0 &+ \sum_\omega \sum_\beta \alpha_{\alpha\beta}(-\omega;\omega) F_\beta^\omega e^{-i\omega t} \\ &+ \frac{1}{2} \sum_{\omega_1,\omega_2} \sum_{\beta,\gamma} \beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) F_\beta^{\omega_1} F_\gamma^{\omega_2} e^{-i\omega_\sigma t} \\ &+ \frac{1}{6} \sum_{\omega_1,\omega_2,\omega_3} \sum_{\beta,\gamma,\delta} \gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) F_\beta^{\omega_1} F_\gamma^{\omega_2} F_\delta^{\omega_3} e^{-i\omega_\sigma t} + \cdots\,,\end{aligned} \tag{4.7}$$

where μ_α^0 is the α-th component of the permanent dipole moment (vector) and the consecutive terms in the expansion stand for the linear and nonlinear optical properties. The linear response is governed by the polarizability (matrix), $\boldsymbol{\alpha}(-\omega;\omega)$. Nonlinear response is mainly governed by the quadratic response of the first hyperpolarizability (3^{rd}-rank tensor), $\boldsymbol{\beta}(-\omega_\sigma;\omega_1,\omega_2)$, and the cubic response of second hyperpolarizability (4^{th}-rank tensor), $\boldsymbol{\gamma}(-\omega_\sigma;\omega_1,\omega_2,\omega_3)$. Higher order nonlinearities are not commonly studied. In the above equation, ω_σ is the frequency of the observed induced polarization and is given by $\omega_\sigma = \omega_1 + \omega_2$ for terms involving β and $\omega_\sigma = \omega_1 + \omega_2 + \omega_3$ for terms involving γ.

Linear and nonlinear optical properties are universal quantities, which for different combinations of involved light waves, may describe different optical phenomena. Examples of those are listed in Table 4.1. It is important to stress that all optical phenomena occur at the same time, but with different probabilities. The intensity of the different nonlinear optical phenomena may be controlled through a particular experimental realization. For example, if the applied electric field, $\mathbf{F}_{\text{A}}(t)$, contains both static and frequency dependent parts,

$$F_{\text{A},\alpha}(t) = F_{\text{A},\alpha}^0 + F_{\text{A},\alpha}^\omega \cos(\omega t)\,, \tag{4.8}$$

the dipole polarization takes the form [15]:

$$\mu_\alpha(t) = \tilde{\mu}_\alpha^0 + \tilde{\mu}_\alpha^\omega \cos(\omega t) + \tilde{\mu}_\alpha^{2\omega} \cos(2\omega t) + \tilde{\mu}_\alpha^{3\omega} \cos(3\omega t)\,. \tag{4.9}$$

Table 4.1: Selected linear and nonlinear properties and the optical phenomena they may describe [15, 151, 153].

Property	Process
$Re[\alpha(-\omega;\omega)]$	Refractive Index
$Im[\alpha(-\omega;\omega)]$	One-Photon Absorption
$\beta(-(\omega_1+\omega_2);\omega_1,\omega_2)$	Sum Frequency Generation
$\beta(-(\omega_1-\omega_2);\omega_1,-\omega_2)$	Difference Frequency Generation
$\beta(-2\omega;\omega,\omega)$	Second Harmonic Generation
$\beta(-\omega;\omega,0)$	Electro-Optical Pockels Effect
$\beta(0;\omega,-\omega)$	Optical Rectification
$\gamma(-3\omega;\omega,\omega,\omega)$	Third Harmonic Generation
$\gamma(-2\omega;\omega,\omega,0)$	Electric Field-Induced Second Harmonic Generation
$Re[\gamma(-\omega;\omega,-\omega,\omega)]$	Intensity-Dependent Refractive Index
$Im[\gamma(-\omega;\omega,-\omega,\omega)]$	Two-Photon Absorption
$\gamma(-\omega_1;\omega_1,\omega_2,-\omega_2)$	Optical Kerr Effect
$\gamma(-\omega;\omega,0,0)$	Electro-Optical Kerr Effect
$\gamma(0;\omega,-\omega,0)$	Electric Field-Induced Optical Rectification

In above equation, terms up to the third order in the electric field were included, and the factors in front of the cosine packets are the Fourier amplitudes:

$$\tilde{\mu}_\alpha^0 = \mu_\alpha^0 + \sum_\beta \alpha_{\alpha\beta}(0;0)F_\beta^0 + \frac{1}{2}\sum_{\beta,\gamma}\beta_{\alpha\beta\gamma}(0;0,0)F_\beta^0F_\gamma^0 + \frac{1}{6}\sum_{\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}(0;0,0,0)F_\beta^0F_\gamma^0F_\delta^0$$
$$+\frac{1}{4}\sum_{\beta,\gamma}\beta_{\alpha\beta\gamma}(0;-\omega,\omega)F_\beta^\omega F_\gamma^\omega + \frac{1}{4}\sum_{\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}(0;-\omega,\omega,0)F_\beta^0F_\gamma^\omega F_\delta^\omega\,, \tag{4.10}$$

$$\tilde{\mu}_\alpha^\omega = \sum_\beta \alpha_{\alpha\beta}(-\omega;\omega)F_\beta^\omega + \sum_{\beta,\gamma}\beta_{\alpha\beta\gamma}(-\omega;\omega,0)F_\beta^\omega F_\gamma^0 + \frac{1}{2}\sum_{\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}(-\omega;\omega,0,0)$$
$$\times F_\beta^\omega F_\gamma^0F_\delta^0 + \frac{1}{8}\sum_{\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}(-\omega;\omega,-\omega,\omega)F_\beta^\omega F_\gamma^\omega F_\delta^\omega\,, \tag{4.11}$$

$$\tilde{\mu}_\alpha^{2\omega} = \frac{1}{4}\sum_{\beta,\gamma}\beta_{\alpha\beta\gamma}(-2\omega;\omega,\omega)F_\beta^\omega F_\gamma^\omega + \frac{1}{4}\sum_{\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}(-2\omega;\omega,\omega,0)F_\beta^\omega F_\gamma^\omega F_\delta^0\,, \tag{4.12}$$

$$\tilde{\mu}_\alpha^{3\omega} = \frac{1}{24}\sum_{\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}(-3\omega;\omega,\omega,\omega)F_\beta^\omega F_\gamma^\omega F_\delta^\omega\,. \tag{4.13}$$

In the $\tilde{\mu}_\alpha^0$ term, one can find the polarization of the system due to the static field and the optical rectification; $\tilde{\mu}_\alpha^\omega$ is responsible for the (intensity-dependent) refractive index, the electro-optical Pockels and Kerr effects, and the one- and two-photon absorption probabilities; $\tilde{\mu}_\alpha^{2\omega}$ determines the (electric field-induced) second harmonic generation; and lastly, $\tilde{\mu}_\alpha^{3\omega}$ quantifies the third-harmonic generation.

Although not useful for the description of the light-matter interaction, some chemical phenomena involve only static electric fields, i.e. time and frequency independent $\mathbf{F}(t) = \mathbf{F}$. In such a scenario, for a system in a stationary state, the dipole moment no longer changes in time, $\mu_\alpha(t,\mathbf{F}(t)) = \mu_\alpha(\mathbf{F})$, and Eq. 4.7 takes the form

$$\mu_\alpha(\mathbf{F}) = \mu_\alpha^0 + \sum_\beta \alpha_{\alpha\beta}F_\beta + \frac{1}{2}\sum_{\beta,\gamma}\beta_{\alpha\beta\gamma}F_\beta F_\gamma + \frac{1}{6}\sum_{\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}F_\beta F_\gamma F_\delta\,, \tag{4.14}$$

where, for clarity, the frequency dependence of α, β, γ has been dropped, i.e.,

$$\alpha_{\alpha\beta} = \alpha_{\alpha\beta}(0;0)\,,$$
$$\beta_{\alpha\beta\gamma} = \beta_{\alpha\beta\gamma}(0;0,0)\,,$$
$$\gamma_{\alpha\beta\gamma\delta} = \gamma_{\alpha\beta\gamma\delta}(0;0,0,0)\,.$$

Because Eq. 4.14 is a Taylor expansion of $\mu_\alpha(\mathbf{F})$ around $\mathbf{F} = \mathbf{0}$, the static (hyper)polarizabilities may be defined as the consecutive derivatives with respect to the magnitude of the electric field:

$$\alpha_{\alpha\beta} = \left.\frac{\mathrm{d}\mu_\alpha(\mathbf{F})}{\mathrm{d}F_\beta}\right|_{\mathbf{F}=\mathbf{0}}, \tag{4.15}$$

$$\beta_{\alpha\beta\gamma} = \left.\frac{\mathrm{d}^2\mu_\alpha(\mathbf{F})}{\mathrm{d}F_\gamma \mathrm{d}F_\beta}\right|_{\mathbf{F}=\mathbf{0}}, \tag{4.16}$$

$$\gamma_{\alpha\beta\gamma\delta} = \left.\frac{\mathrm{d}^3\mu_\alpha(\mathbf{F})}{\mathrm{d}F_\delta \mathrm{d}F_\gamma \mathrm{d}F_\beta}\right|_{\mathbf{F}=\mathbf{0}}. \tag{4.17}$$

4.4 Sum-Over-States

Microscopic representation of the optical response is traditionally done in the framework of time-dependent perturbation theory (TD-PT) [15, 17, 18, 150]. Response to external electric fields of any kind is expressed in terms of the excited states, leading to the Sum-Over-States (SOS) representation. The derivation presented here follows the one given in Ref [15]. The main assumption of TD-PT is that the interaction with the perturbation constitutes only a fraction of the total energy of the system. Then, the total time-dependent Hamiltonian of the interacting system $\hat{H}(t)$ is partitioned as [18]

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t)\,, \tag{4.18}$$

where $\hat{H}_0$ is the time-independent Hamiltonian of the unperturbed system and $\hat{V}(t)$ is time-dependent perturbation operator. $\hat{V}(t)$ involves the dipole moment operator (treated quantum mechanically) and the time-dependent electric field $F_\alpha(t)$ (treated classically and defined by Eq. 4.5),

$$\hat{V}(t) = -\sum_\alpha \hat{\mu}_\alpha F_\alpha(t) e^{\varepsilon t}\,. \tag{4.19}$$

To ensure that the perturbation is turned on adiabatically at $t = -\infty$, an additional exponential factor with a small positive infinitesimal ε is used.

As in a regular perturbation theory, it is assumed that one knows a complete set of exact eigensolutions $|n\rangle$ of the unperturbed Hamiltonian $\hat{H}_0$,

$$\hat{H}_0|n\rangle = E_n|n\rangle\,, \tag{4.20}$$

with E_n being the corresponding eigenvalues. Time propagation of a (ground-state) wavefunction can be expressed using the whole set of exact time-independent eigenfunctions,

$$|\Psi(t)\rangle = \sum_n c_n(t)|n\rangle = \sum_n d_n(t) e^{-iE_n t/\hbar}|n\rangle\,, \tag{4.21}$$

where the time evolution is included in the admixture coefficients $c_n(t) = d_n(t)e^{-iE_n t/\hbar}$. Following PT approach, the $d_n(t)$ coefficients are expanded in a perturbation series,

$$d_n(t) = d_n^{(0)}(t) + d_n^{(1)}(t) + d_n^{(2)}(t) + d_n^{(3)}(t) + \cdots\,, \tag{4.22}$$

with $d_n^{(0)}(t)$ denoting the time evolution of the unperturbed system (only $d_n^{(0)}(t) = 1$ prevails) and $d_n^{(1)}(t)$, $d_n^{(2)}(t)$, $d_n^{(3)}(t)$, ..., being perturbative corrections of consecutive order. For that reason, the time evolution of the perturbed wavefunction is also expanded in a similar fashion:

$$|\Psi(t)\rangle = |\Psi^{(0)}(t)\rangle + |\Psi^{(1)}(t)\rangle + |\Psi^{(2)}(t)\rangle + |\Psi^{(3)}(t)\rangle + \cdots , \tag{4.23}$$

where

$$|\Psi^{(N)}(t)\rangle = \sum_n d_n^{(N)} e^{-iE_n t/\hbar} |n\rangle . \tag{4.24}$$

The expectation value of the time-dependent polarization is expanded in the perturbative series as well,

$$\langle \Psi(t)|\hat{\mu}_\alpha|\Psi(t)\rangle = \langle \hat{\mu}_\alpha\rangle^{(0)} + \langle \hat{\mu}_\alpha\rangle^{(1)} + \langle \hat{\mu}_\alpha\rangle^{(2)} + \langle \hat{\mu}_\alpha\rangle^{(3)} + \cdots , \tag{4.25}$$

and the consecutive terms are described with mixed terms involving various orders of correction to the eigenfunction:

$$\langle \hat{\mu}_\alpha\rangle^{(0)} = \langle 0|\hat{\mu}_\alpha|0\rangle , \tag{4.26}$$

$$\langle \hat{\mu}_\alpha\rangle^{(1)} = \langle \Psi^{(0)}(t)|\hat{\mu}_\alpha|\Psi^{(1)}(t)\rangle + \langle \Psi^{(1)}(t)|\hat{\mu}_\alpha|\Psi^{(0)}(t)\rangle , \tag{4.27}$$

$$\langle \hat{\mu}_\alpha\rangle^{(2)} = \langle \Psi^{(0)}(t)|\hat{\mu}_\alpha|\Psi^{(2)}(t)\rangle + \langle \Psi^{(1)}(t)|\hat{\mu}_\alpha|\Psi^{(1)}(t)\rangle + \langle \Psi^{(2)}(t)|\hat{\mu}_\alpha|\Psi^{(0)}(t)\rangle , \tag{4.28}$$

$$\begin{aligned}\langle \hat{\mu}_\alpha\rangle^{(3)} = {} & \langle \Psi^{(0)}(t)|\hat{\mu}_\alpha|\Psi^{(3)}(t)\rangle + \langle \Psi^{(1)}(t)|\hat{\mu}_\alpha|\Psi^{(2)}(t)\rangle + \langle \Psi^{(2)}(t)|\hat{\mu}_\alpha|\Psi^{(1)}(t)\rangle \\ & + \langle \Psi^{(3)}(t)|\hat{\mu}_\alpha|\Psi^{(0)}(t)\rangle ,\end{aligned} \tag{4.29}$$

Each correction is responsible for a particular optical response, namely $\langle \hat{\mu}_\alpha\rangle^{(1)}$ will be responsible for the linear polarization, whereas $\langle \hat{\mu}_\alpha\rangle^{(2)}$ and $\langle \hat{\mu}_\alpha\rangle^{(3)}$ (and higher) for the nonlinear polarization. Detailed derivation of these terms follow general rules of TD-PT and can be found in Ref. [15].

In the following part of this section, final formulae for linear and nonlinear optical properties are presented and commented on. The linear dipole polarization can be concisely written as [15, 18]

$$\langle \hat{\mu}_\alpha\rangle^{(1)} = \sum_\omega \alpha_{\alpha\beta}(-\omega;\omega) F_\beta^\omega e^{-i\omega t} e^{\varepsilon t} , \tag{4.30}$$

with a direct dependence on the microscopic polarizability $\alpha_{\alpha\beta}(-\omega;\omega)$. As a result of the TD-PT derivation, it is expressed in the sum-over-state manner as [15, 18]

$$\alpha_{\alpha\beta}(-\omega;\omega) = \frac{1}{\hbar}\sum_n \left[\frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\mu}_\beta|0\rangle}{\omega_{n0} - \omega - i\varepsilon} - \frac{\langle 0|\hat{\mu}_\beta|n\rangle\langle n|\hat{\mu}_\alpha|0\rangle}{\omega_{n0} + \omega + i\varepsilon} \right] . \tag{4.31}$$

All the SOS formulae for NLOPs utilize two main ingredients, namely the transition dipole moment between two energetic states, $\langle n|\hat{\mu}_\alpha|m\rangle$, and the transition frequency $\omega_{nm} = (E_n - E_m)/\hbar$. In Eq. 4.31, only transitions from the ground-state $|0\rangle$ to a particular excited state $|n\rangle$ are included. First order nonlinearity of the dipole polarization [15, 18],

$$\langle \hat{\mu}_\alpha\rangle^{(2)} = \frac{1}{2}\sum_{\omega_1,\omega_2} \beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) F_\beta^{\omega_1} F_\gamma^{\omega_2} e^{-i\omega_\sigma t} e^{2\varepsilon t} , \tag{4.32}$$

depends on the hyperpolarizability $\beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2)$ which is expressed in the SOS form as [15, 18]:

$$\begin{aligned}\beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = \frac{1}{\hbar^2}\sum\mathscr{P}_{1,2}\sum_{n,p}\Bigg[&\frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\mu}_\beta|p\rangle\langle p|\hat{\mu}_\gamma|0\rangle}{(\omega_{n0}-\omega_\sigma-i2\varepsilon)(\omega_{p0}-\omega_2-i\varepsilon)}\\ &+\frac{\langle 0|\hat{\mu}_\beta|n\rangle\langle n|\hat{\mu}_\alpha|p\rangle\langle p|\hat{\mu}_\gamma|0\rangle}{(\omega_{n0}+\omega_1+i\varepsilon)(\omega_{p0}-\omega_2-i\varepsilon)}\\ &+\frac{\langle 0|\hat{\mu}_\gamma|n\rangle\langle n|\hat{\mu}_\beta|p\rangle\langle p|\hat{\mu}_\alpha|0\rangle}{(\omega_{p0}+\omega_\sigma+i2\varepsilon)(\omega_{n0}+\omega_2+i\varepsilon)}\Bigg]. \end{aligned} \tag{4.33}$$

In above equation, $\sum\mathscr{P}_{1,2}$ permutes the pairs of indices (β,ω_1) and (γ,ω_2). In contrast to Eq. 4.31, the SOS expression for hyperpolarizability involves transition dipole moment integrals between the n–th and p–th excited states. Lastly, the second nonlinear part of the dipole polarization [15, 18],

$$\langle\hat{\mu}_\alpha\rangle^{(3)} = \frac{1}{6}\sum_{\omega_1,\omega_2,\omega_3}\gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3)F_\beta^{\omega_1}F_\gamma^{\omega_2}F_\delta^{\omega_3}e^{-i\omega_\sigma t}e^{3\varepsilon t}, \tag{4.34}$$

directly depends on the second hyperpolarizability $\gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3)$. The SOS formula for second hyperpolarizability is finally derived as [15, 18]

$$\begin{aligned}\gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) = \frac{1}{\hbar^3}&\sum\mathscr{P}_{1,2,3}\\ &\times\sum_{n,m,p}\Bigg[\frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\mu}_\beta|m\rangle\langle m|\hat{\mu}_\gamma|p\rangle\langle p|\hat{\mu}_\delta|0\rangle}{(\omega_{n0}-\omega_\sigma-i3\varepsilon)(\omega_{m0}-\omega_2-\omega_3-i2\varepsilon)(\omega_{p0}-\omega_3-i\varepsilon)}\\ &+\frac{\langle 0|\hat{\mu}_\beta|n\rangle\langle n|\hat{\mu}_\alpha|m\rangle\langle m|\hat{\mu}_\gamma|p\rangle\langle p|\hat{\mu}_\delta|0\rangle}{(\omega_{n0}+\omega_1+i\varepsilon)(\omega_{m0}-\omega_2-\omega_3-i2\varepsilon)(\omega_{p0}-\omega_3-i\varepsilon)}\\ &+\frac{\langle 0|\hat{\mu}_\gamma|n\rangle\langle n|\hat{\mu}_\beta|m\rangle\langle m|\hat{\mu}_\alpha|p\rangle\langle p|\hat{\mu}_\delta|0\rangle}{(\omega_{m0}+\omega_1+\omega_2+i2\varepsilon)(\omega_{n0}+\omega_2+i\varepsilon)(\omega_{p0}-\omega_3-i\varepsilon)}\\ &+\frac{\langle 0|\hat{\mu}_\delta|n\rangle\langle n|\hat{\mu}_\gamma|m\rangle\langle m|\hat{\mu}_\beta|p\rangle\langle p|\hat{\mu}_\alpha|0\rangle}{(\omega_{p0}+\omega_\sigma+i3\varepsilon)(\omega_{m0}+\omega_2+\omega_3+i2\varepsilon)(\omega_{n0}+\omega_3+i\varepsilon)}\Bigg], \end{aligned} \tag{4.35}$$

where $\sum\mathscr{P}_{1,2,3}$ permutes the pairs of indices (β,ω_1), (γ,ω_2) and (δ,ω_3). This expression involves transitions to three intermediate states n, m, and p. Due to their generality, all above SOS formulae in Eqs. 4.31 - 4.35 describe both resonant and nonresonant responses. The former one occurs when the frequency of incident light ω matches a particular transition frequency between the ground state and the targeted excited state. Such processes possess high spatial selectivity and probability and are greatly enhanced. Moreover, the dipole polarization of resonant phenomena contains both real and imaginary terms, the latter describing absorption/emission processes. In the nonresonant region of the response, when the perturbing frequency is significantly different from any of the transition frequencies, nonlinear polarization also occurs, but usually with a lower enhancement factor and poorer spatial resolution. In contrast to the resonant polarization, nonresonant polarization has only the real part, since the imaginary contribution is negligible.

To facilitate the usage of the SOS formula, improve their convergence, one can develop reduced formulas including only the key terms. This can be done separately for each type of optical processes, both resonant and nonresonant ones. Since this Thesis focuses mostly on the response to static electric fields, the reduced SOS formulae for nonresonant processes are presented.

The molecular polarizability in the nonresonant region is reduced to

$$\alpha_{\alpha\beta}(-\omega;\omega) = \frac{1}{\hbar}{\sum_{n}}' \left[\frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\mu}_\beta|0\rangle}{\omega_{n0}-\omega} - \frac{\langle 0|\hat{\mu}_\beta|n\rangle\langle n|\hat{\mu}_\alpha|0\rangle}{\omega_{n0}+\omega} \right], \tag{4.36}$$

where prime in the summation means skipping the ground state $|0\rangle$. In the above formula, the only difference lies in the denominator where imaginary terms have been excluded. Including a canonical transformation to the interaction representation suggested by Bates [18, 154], the formula for the nonresonant molecular hyperpolarizability is simplified to

$$\beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = \frac{1}{\hbar^2}\sum \mathscr{P}_{-\sigma,1,2}{\sum_{n,p}}' \frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\overline{\hat{\mu}_\beta}|p\rangle\langle p|\hat{\mu}_\gamma|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{p0}-\omega_2)}, \tag{4.37}$$

where $\overline{\hat{\mu}_\alpha} = \hat{\mu}_\alpha - \langle 0|\hat{\mu}_\alpha|0\rangle$ is the fluctuation dipole moment operator, $\sum \mathscr{P}_{-\sigma,1,2}$ permutes the pairs $(\alpha,-\omega_\sigma)$, (β,ω_1), (γ,ω_2). Lastly, the nonresonant molecular second hyperpolarizability can be also concisely written as

$$\begin{aligned}\gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) = \frac{1}{\hbar^3}\sum \mathscr{P}_{-\sigma,1,2,3} \\ \times \left[{\sum_{n,m,p}}' \frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\overline{\hat{\mu}_\beta}|m\rangle\langle m|\overline{\hat{\mu}_\gamma}|p\rangle\langle p|\hat{\mu}_\delta|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{m0}-\omega_2-\omega_3)(\omega_{p0}-\omega_3)} \right. \\ \left. - {\sum_{n,m}}' \frac{\langle 0|\hat{\mu}_\alpha|n\rangle\langle n|\hat{\mu}_\beta|0\rangle\langle 0|\hat{\mu}_\gamma|m\rangle\langle m|\hat{\mu}_\delta|0\rangle}{(\omega_{n0}-\omega_\sigma)(\omega_{m0}-\omega_3)(\omega_{m0}+\omega_2)} \right],\end{aligned} \tag{4.38}$$

where $\sum \mathscr{P}_{-\sigma,1,2,3}$ permutes the pairs $(\alpha,-\omega_\sigma)$, (β,ω_1), (γ,ω_2), and (δ,ω_3).

4.5 Response to Static Electric Fields

Description of the perturbation with static electric fields through static NLOPs, while not essential for most of the real-life experiments, is of key importance in the field of molecular quantum chemistry. They allow to quantify the effects of polarization of electronic density and changes in the molecular structure in the present of electric perturbation. Permanent dipole moment and static polarizability are used to describe the noncovalent interactions between molecular systems and they are also used as a first approach to design molecules with high NLO properties, which may be candidates to be used in the new opto-electronic devices. Static polarizability is crucial for the description of the mechanisms of chemical reactions in the presence of oriented external electric fields (OEEFs). Recently, controlling chemical reactions using OEEFs has gained much attention [155–160].

Accurate computations of static NLOPs are very difficult. They require large basis sets, including polarization and diffuse functions, sometimes even specially tailored solely for the computations of NLOPs such as Sadlej's POL basis set [161–165]. Static NLOPs are also very dependent on the electronic correlation. For that reason, the static response properties are usually used to further test the accuracy of quantum chemical methods.

From a methodological point of view, Eqs. 4.36 to 4.38 are also valid for static properties. In real applications, however, the response to static electric fields is described using alternative approaches to TD-PT, which usually fulfill the Hellmann-Feynman theorem. When the static electric field perturbs a molecular system that initially was in a stationary

state for a given Hamiltonian $\hat{H}_0$, both electrons and nuclei relax their position and the whole system is described with the new stationary state $|\Psi(\mathbf{F})\rangle$,

$$\hat{H}(\mathbf{F})|\Psi(\mathbf{F})\rangle = E(\mathbf{F})|\Psi(\mathbf{F})\rangle\,. \tag{4.39}$$

In such case, the field-dependent perturbed Hamiltonian $\hat{H}$,

$$\hat{H}(\mathbf{F}) = \hat{H}_0 - \sum_\alpha \hat{\mu}_\alpha F_\alpha\,, \tag{4.40}$$

includes the interaction with the external field through the dipole moment operator (defined by the positions of electrons $r_{i,\alpha}$ and nuclei $R_{I,\alpha}$),

$$\hat{\mu}_\alpha = \sum_I R_{I,\alpha} - \sum_i r_{i,\alpha}\,. \tag{4.41}$$

To quantify the linear dependence of the electronic energy on the electric field, one has to compute its total first derivative with respect to the external field,

$$\begin{aligned}
\frac{\mathrm{d}E(\mathbf{F})}{\mathrm{d}F_\alpha} &= \frac{\mathrm{d}}{\mathrm{d}F_\alpha}\langle\Psi(\mathbf{F})|\hat{H}(\mathbf{F})|\Psi(\mathbf{F})\rangle\,,\\
&= \langle\frac{\mathrm{d}\Psi(\mathbf{F})}{\mathrm{d}F_\alpha}|\hat{H}(\mathbf{F})|\Psi(\mathbf{F})\rangle + \langle\Psi(\mathbf{F})|\hat{H}(\mathbf{F})|\frac{\mathrm{d}\Psi(\mathbf{F})}{\mathrm{d}F_\alpha}\rangle + \langle\Psi(\mathbf{F})|\frac{\mathrm{d}\hat{H}(\mathbf{F})}{\mathrm{d}F_\alpha}|\Psi(\mathbf{F})\rangle\,,\\
&= E(\mathbf{F})\langle\frac{\mathrm{d}\Psi(\mathbf{F})}{\mathrm{d}F_\alpha}|\Psi(\mathbf{F})\rangle + E(\mathbf{F})\langle\Psi(\mathbf{F})|\frac{\mathrm{d}\Psi(\mathbf{F})}{\mathrm{d}F_\alpha}\rangle + \langle\Psi(\mathbf{F})|\frac{\mathrm{d}\hat{H}(\mathbf{F})}{\mathrm{d}F_\alpha}|\Psi(\mathbf{F})\rangle\,,\\
&= E(\mathbf{F})\frac{\mathrm{d}}{\mathrm{d}F_\alpha}\langle\Psi(\mathbf{F})|\Psi(\mathbf{F})\rangle + \langle\Psi(\mathbf{F})|\frac{\mathrm{d}\hat{H}(\mathbf{F})}{\mathrm{d}F_\alpha}|\Psi(\mathbf{F})\rangle\,,\\
&= \langle\Psi(\mathbf{F})|\frac{\mathrm{d}\hat{H}(\mathbf{F})}{\mathrm{d}F_\alpha}|\Psi(\mathbf{F})\rangle = -\langle\Psi(\mathbf{F})|\hat{\mu}_\alpha|\Psi(\mathbf{F})\rangle\,,\\
&= -\mu_\alpha(\mathbf{F})\,,
\end{aligned} \tag{4.42}$$

which equals to the negative field-dependent dipole moment. The derivation in Eq. 4.42 directly follows the Hellmann-Feynman theorem. To obtain the final result, Eq. 4.39, the hermiticity of $\hat{H}(\mathbf{F})$ and the normalization condition of $|\Psi(\mathbf{F})\rangle$ were adopted.

Combining Eqs. 4.42 and 4.15-4.17, one can express the linear and nonlinear static electrical properties as derivatives of total energy:

$$\mu^0_\alpha = \mu_\alpha(\mathbf{0}) = -\left.\frac{\mathrm{d}E(\mathbf{F})}{\mathrm{d}F_\alpha}\right|_{\mathbf{F}=\mathbf{0}}\,, \tag{4.43}$$

$$\alpha_{\alpha\beta} = \left.\frac{\mathrm{d}\mu_\alpha(\mathbf{F})}{\mathrm{d}F_\beta}\right|_{\mathbf{F}=\mathbf{0}} = -\left.\frac{\mathrm{d}^2E(\mathbf{F})}{\mathrm{d}F_\beta\mathrm{d}F_\alpha}\right|_{\mathbf{F}=\mathbf{0}}\,, \tag{4.44}$$

$$\beta_{\alpha\beta\gamma} = \left.\frac{\mathrm{d}^2\mu_\alpha(\mathbf{F})}{\mathrm{d}F_\gamma\mathrm{d}F_\beta}\right|_{\mathbf{F}=\mathbf{0}} = -\left.\frac{\mathrm{d}^3E(\mathbf{F})}{\mathrm{d}F_\gamma\mathrm{d}F_\beta\mathrm{d}F_\alpha}\right|_{\mathbf{F}=\mathbf{0}}\,, \tag{4.45}$$

$$\gamma_{\alpha\beta\gamma\delta} = \left.\frac{\mathrm{d}^3\mu_\alpha(\mathbf{F})}{\mathrm{d}F_\delta\mathrm{d}F_\gamma\mathrm{d}F_\beta}\right|_{\mathbf{F}=\mathbf{0}} = -\left.\frac{\mathrm{d}^4E(\mathbf{F})}{\mathrm{d}F_\delta\mathrm{d}F_\gamma\mathrm{d}F_\beta\mathrm{d}F_\alpha}\right|_{\mathbf{F}=\mathbf{0}}\,. \tag{4.46}$$

With such a definition of static NLOPs, one can define an equivalent Taylor expansion of the field-dependent energy,

$$\begin{aligned}
E(\mathbf{F}) = E_0 &- \sum_\alpha \mu^0_\alpha F_\alpha - \frac{1}{2}\sum_{\alpha,\beta}\alpha_{\alpha\beta}F_\alpha F_\beta - \frac{1}{6}\sum_{\alpha,\beta,\gamma}\beta_{\alpha\beta\gamma}F_\alpha F_\beta F_\gamma\\
&- \frac{1}{24}\sum_{\alpha,\beta,\gamma,\delta}\gamma_{\alpha\beta\gamma\delta}F_\alpha F_\beta F_\gamma F_\delta + \cdots\,.
\end{aligned} \tag{4.47}$$

This representation is of fundamental value and is adopted both in the analytical and numerical computations of static NLOPs.

4.6 Density-Based Representation of Static NLOPs

If one considers only the electronic part of the Hamiltonian, an equivalent description of the static NLOPs can be obtained in terms of the electronic densities. In a similar fashion to Eq. 4.14, the charge density function $\rho(\mathbf{r},\mathbf{F})$ can be also expressed by the means of a Taylor expansion with respect to the electric field strength [134–136]:

$$\rho(\mathbf{r},\mathbf{F}) = \rho(\mathbf{r},\mathbf{0}) + \sum_{\beta}\rho^{(\beta)}(\mathbf{r})F_\beta + \frac{1}{2}\sum_{\beta,\gamma}\rho^{(\beta\gamma)}(\mathbf{r})F_\beta F_\gamma + \frac{1}{6}\sum_{\beta,\gamma,\delta}\rho^{(\beta\gamma\delta)}(\mathbf{r})F_\beta F_\gamma F_\delta + \cdots . \tag{4.48}$$

In the above equation, the following shorthand notation for the density derivatives has been adopted [135, 136]:

$$\rho^{(\beta)}(\mathbf{r}) = \left.\frac{\partial\rho(\mathbf{r},\mathbf{F})}{\partial F_\beta}\right|_{\mathbf{F}=\mathbf{0}}, \tag{4.49}$$

$$\rho^{(\beta\gamma)}(\mathbf{r}) = \left.\frac{\partial^2\rho(\mathbf{r},\mathbf{F})}{\partial F_\gamma\partial F_\beta}\right|_{\mathbf{F}=\mathbf{0}}, \tag{4.50}$$

$$\rho^{(\beta\gamma\delta)}(\mathbf{r}) = \left.\frac{\partial^3\rho(\mathbf{r},\mathbf{F})}{\partial F_\delta\partial F_\gamma\partial F_\beta}\right|_{\mathbf{F}=\mathbf{0}}. \tag{4.51}$$

Combining Eq. 4.48 with the definition of the dipole moment in terms of the charge density integral,

$$\mu_\alpha^{el}(\mathbf{F}) = -\int \mathrm{d}\mathbf{r}\; r_\alpha\rho(\mathbf{r},\mathbf{F}), \tag{4.52}$$

$$= -\int \mathrm{d}\mathbf{r}\; r_\alpha\rho(\mathbf{r},\mathbf{0}) - \sum_{\beta}F_\beta\int \mathrm{d}\mathbf{r}\; r_\alpha\rho^{(\beta)}(\mathbf{r}) - \frac{1}{2}\sum_{\beta,\gamma}F_\beta F_\gamma\int \mathrm{d}\mathbf{r}\; r_\alpha\rho^{(\beta\gamma)}(\mathbf{r}) - \frac{1}{6}\sum_{\beta,\gamma,\delta}F_\beta F_\gamma F_\delta\int \mathrm{d}\mathbf{r}\; r_\alpha\rho^{(\beta\gamma\delta)}(\mathbf{r}) + \cdots , \tag{4.53}$$

one can compare it with Eq. 4.14 and isolate the relations for higher nonlinear response functions in terms of the electronic density derivatives [135, 136]:

$$\mu_\alpha^0 = -\int \mathrm{d}\mathbf{r}\; r_\alpha\rho(\mathbf{r},\mathbf{0}), \tag{4.54}$$

$$\alpha_{\alpha\beta}^{el} = -\int \mathrm{d}\mathbf{r}\; r_\alpha\rho^{(\beta)}(\mathbf{r}), \tag{4.55}$$

$$\beta_{\alpha\beta\gamma}^{el} = -\int \mathrm{d}\mathbf{r}\; r_\alpha\rho^{(\beta\gamma)}(\mathbf{r}), \tag{4.56}$$

$$\gamma_{\alpha\beta\gamma\delta}^{el} = -\int \mathrm{d}\mathbf{r}\; r_\alpha\rho^{(\beta\gamma\delta)}(\mathbf{r}). \tag{4.57}$$

For that reason, the above–defined derivatives $\rho^{(\beta)}(\mathbf{r})$, $\rho^{(\beta\gamma)}(\mathbf{r})$, and $\rho^{(\beta\gamma\delta)}(\mathbf{r})$ have been named after their NLOPs analogs, respectively, α–, β–, γ–densities [135, 136]. However, these are not rigorous property densities, because their sole integration, i.e. $\int \mathrm{d}\mathbf{r}\rho^{(\beta\gamma\delta\ldots)}(\mathbf{r})$ does not yield the corresponding NLOP (actually, it yields zero) — one needs to include the position operator r_α in the integration.

One can also represent electronic NLOPs using the reduced density matrices expressed in terms of the atomic or molecular orbital basis. Namely, the expectation value

of an operator $\hat{O}$, expressed as the sum of n one–electron operators $\hat{O}_1 = \sum_n \hat{o}_n$, can be written as

$$\begin{aligned}\langle\Psi|\hat{O}_1|\Psi\rangle &= \sum_\sigma \int \mathrm{d}\mathbf{r}\, \hat{o}_1(\mathbf{r}')\rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma)\big|_{\mathbf{r}'=\mathbf{r}}\,, \\ &= \sum_\sigma \sum_{\mu,\nu} D^\sigma_{\mu\nu}\langle\chi_\mu|\hat{o}_1|\chi_\nu\rangle\,,\end{aligned} \tag{4.58}$$

where $D^\sigma_{\mu\nu}$ is the orbital representation of the σ part of 1-RDM, $\rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma)$:

$$D^\sigma_{\mu\nu} = \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \chi^*_\mu(\mathbf{r}')\rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma)\chi_\nu(\mathbf{r})\,. \tag{4.59}$$

The total $D_{\mu\nu}$ is given by the sum of $D^\alpha_{\mu\nu}$ and $D^\beta_{\mu\nu}$:

$$D_{\mu\nu} = D^\alpha_{\mu\nu} + D^\beta_{\mu\nu}\,. \tag{4.60}$$

Therefore, one can further express the field-dependent dipole moment $\mu_\alpha(\mathbf{F})$ of Eq. 4.52 in terms of the field-dependent one-particle reduced density matrix, $\mathbf{D}(\mathbf{F})$, and the one-electron transition dipole moment matrix $\mathbf{h}^{(\alpha)}$ defined as $h^{(\alpha)}_{\mu\nu} = \langle\chi_\mu|r_\beta|\chi_\nu\rangle$:

$$\mu^{el}_\alpha(\mathbf{F}) = -\sum_{\mu,\nu} D_{\mu\nu}(\mathbf{F}) h^{(\alpha)}_{\mu\nu}\,. \tag{4.61}$$

For such purposes, it is natural to utilize an atomic orbital (AO) basis, that is, a set of the contracted GTOs centered at the atomic positions. The AOs are invariant respect to the strength of the external electric field. Therefore the static polarizability, the first and second hyperpolarizabilities can be written as:

$$\alpha^{el}_{\alpha\beta} = -\sum_{\mu\nu} D^{(\beta)}_{\mu\nu} h^{(\alpha)}_{\mu\nu}\,, \tag{4.62}$$

$$\beta^{el}_{\alpha\beta\gamma} = -\sum_{\mu\nu} D^{(\beta\gamma)}_{\mu\nu} h^{(\alpha)}_{\mu\nu}\,, \tag{4.63}$$

$$\gamma^{el}_{\alpha\beta\gamma\delta} = -\sum_{\mu\nu} D^{(\beta\gamma\delta)}_{\mu\nu} h^{(\alpha)}_{\mu\nu}\,. \tag{4.64}$$

In these equations, $D^{(\beta)}_{\mu\nu}$, $D^{(\beta\gamma)}_{\mu\nu}$, and $D^{(\beta\gamma\delta)}_{\mu\nu}$ are the derivatives of the field-dependent 1-RDM with respect to the external field denoted with the equivalent notation of Eqs. 4.49 - 4.51.

4.7 Separation of Electronic and Vibrational Contributions to NLOPs

The SOS expressions in Eqs. 4.31 to 4.35 are derived for the most general case of any eigenstate. The same holds for the formulae for the nonresonant response in Eqs. 4.36 to 4.38. However, the description of the molecular NLOPs is only feasible when one adapts the most fundamental approximation adapted in quantum chemistry, namely the Born-Oppenheimer (BO) approximation, which decouples the motion of the electrons and nuclei.

The total Hamiltonian, without the presence of the electric field, is split into electronic and nuclear parts,

$$\hat{H}_0(\mathbf{R},\mathbf{r}) = \hat{H}^{el}_0(\mathbf{R},\mathbf{r}) + \hat{H}^{nuc}_0(\mathbf{R})\,, \tag{4.65}$$

and the total wavefunction is expressed as a product of the electronic and nuclear parts,

$$|\Psi(\mathbf{R},\mathbf{r})\rangle = |\Psi^{el}(\mathbf{R},\mathbf{r})\rangle \times |\Psi^{nuc}(\mathbf{R})\rangle\,. \tag{4.66}$$

Firstly, the eigenproblem of the electronic part is being solved,

$$\hat{H}_0^{el}(\mathbf{R},\mathbf{r})|\Psi_K^{el}(\mathbf{R},\mathbf{r})\rangle = V_K^{el}(\mathbf{R})|\Psi_K^{el}(\mathbf{R},\mathbf{r})\rangle\,, \tag{4.67}$$

and the resulting electronic energy $V_K^{el}(\mathbf{R})$, which parametrically depends on the position of the nuclei, is used as a part of the $\hat{H}_{0,K}^{nuc}(\mathbf{R})$ operator, which is then used to solve the eigenproblem of the nuclear part of the wavefunction,

$$\hat{H}_{0,K}^{nuc}(\mathbf{R})|\Psi_{kK}^{nuc}(\mathbf{R})\rangle = E_{kK}|\Psi_{kK}^{nuc}(\mathbf{R})\rangle\,. \tag{4.68}$$

Then, one can express any NLOP, generally denoted here as P, as a sum of the electronic, P^{el}, and vibrational, P^{vib}, parts:

$$P = P^{el} + P^{vib}\,. \tag{4.69}$$

If the BO approximation is assumed in the SOS expression one has to simply utilize the vibronic product states. These are denoted as $|N,n\rangle$, where the first index denotes N-th electronic state, and the second n-th vibrational state. In particular, the ground state is denoted as $|0,0\rangle$. The formulae for NLOPs in the nonresonant region (Eqs. 4.36-4.38) in terms of the vibronic states are rewritten as:

$$\begin{aligned}
\alpha_{\alpha\beta}(-\omega_\sigma;\omega) &= \frac{1}{\hbar}\sum \mathscr{P}_{-\sigma,1}{\sum_{k,K}}'\langle 0,0|\hat{\mu}_\alpha|K,k\rangle\langle k,K|\hat{\mu}_\beta|0,0\rangle \\
&\quad\times\left(\frac{1}{\omega_{kK}-\omega_\sigma}+\frac{1}{\omega_{kK}+\omega_\sigma}\right), \qquad (4.70)\\
\beta_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) &= \frac{1}{\hbar^2}\sum \mathscr{P}_{-\sigma,1,2}{\sum_{k,K,l,L}}'\langle 0,0|\hat{\mu}_\alpha|K,k\rangle\langle k,K|\overline{\hat{\mu}_\beta}|L,l\rangle\langle l,L|\hat{\mu}_\gamma|0,0\rangle \\
&\quad\times\frac{1}{\omega_{kK}-\omega_\sigma}\cdot\frac{1}{\omega_{lL}-\omega_2}\,, \qquad (4.71)\\
\gamma_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) &= \frac{1}{\hbar^3}\sum \mathscr{P}_{-\sigma,1,2,3} \\
&\quad\Bigg[{\sum_{k,K,l,L,m,M}}'\langle 0,0|\hat{\mu}_\alpha|K,k\rangle\langle k,K|\overline{\hat{\mu}_\beta}|L,l\rangle\langle l,L|\overline{\hat{\mu}_\gamma}|M,m\rangle \\
&\quad\times\langle m,M|\hat{\mu}_\delta|0,0\rangle\frac{1}{\omega_{kK}-\omega_\sigma}\cdot\frac{1}{\omega_{lL}-\omega_2-\omega_3}\cdot\frac{1}{\omega_{mM}-\omega_3} \\
&\quad-{\sum_{k,K,l,L}}'\langle 0,0|\hat{\mu}_\alpha|K,k\rangle\langle k,K|\hat{\mu}_\beta|0,0\rangle\langle 0,0|\hat{\mu}_\gamma|L,l\rangle\langle l,L|\hat{\mu}_\delta|0,0\rangle \\
&\quad\times\frac{1}{\omega_{kK}-\omega_\sigma}\cdot\frac{1}{\omega_{lL}-\omega_3}\cdot\frac{1}{\omega_{lL}+\omega_2}\Bigg]\,. \qquad (4.72)
\end{aligned}$$

In the above expressions, $\omega_{kK} = (E_{kK}-E_{00})/\hbar$ is the transition frequency to the $|K,k\rangle$ vibronic state measured from the ground $|0,0\rangle$ vibronic state.

4.8 Bishop-Kirtman Perturbation Theory

In their seminal work, Bishop and Kirtman proposed an efficient way for the calculation of vibrational contributions to NLOPs in a perturbative manner, by adopting few additional approximations related to the denominators of the SOS expressions [110–112]. This method has been named as the Bishop-Kirtman Perturbation Theory (BKPT). It expresses the overall vibrational part of the NLOP as a sum of a pure vibrational term, P^v, and the well-known Zero Point Vibrational Average (ZPVA) correction, ΔP^{ZPVA}:

$$P^{vibr} = P^v + \Delta P^{\mathrm{ZPVA}}. \tag{4.73}$$

Firstly, in BKPT is assumed that ω_{kK} transition frequencies (i.e. transition to vibronic states of excited electronic states) are much larger than frequencies of the incident light beams, $\omega_{kK} \gg \omega_i$. Secondly, it is assumed that the energy differences between vibrational levels in higher excited states are negligible in comparison to the differences in energies between the electronic states, which can be summarized as $\omega_{kK} \cong \omega_{0K}$. To facilitate the formulae, the BKPT expressions utilize the following shorthand notation: $\omega_{0K} \equiv \omega_K$ is the transition frequency to higher electronic state K, $\omega_{k0} \equiv \omega_k$ is the transition frequency to higher vibrational state of the ground electronic state.

When those two approximations are adapted, one can obtain the following formula for the nonresonant $\alpha_{\alpha\beta}(-\omega_\sigma;\omega)$,

$$\begin{aligned}\alpha_{\alpha\beta}(-\omega_\sigma;\omega) = \frac{1}{\hbar}\sum_{K\neq 0}\langle 0|\left[\langle 0|\hat{\mu}_\alpha|K\rangle\langle K|\hat{\mu}_\beta|0\rangle\cdot\left(\frac{1}{\omega_K-\omega}+\frac{1}{\omega_K+\omega}\right)\right]|0\rangle \\ + \frac{1}{\hbar}\sum_{k\neq 0}\left[\langle 0|\mu_\alpha^{00}|k\rangle\langle k|\mu_\beta^{00}|0\rangle\cdot\left(\frac{1}{\omega_k-\omega}+\frac{1}{\omega_k+\omega}\right)\right],\end{aligned} \tag{4.74}$$

where $\mu_\beta^{00} = \langle 0|\hat{\mu}_\beta|0\rangle$. In the above expression, the first term including the summation over the electronic excited states, constitutes for the ZPVA part of the response. The second term constitutes entirely for the pure vibrational part, P^v. It includes a summation over the vibrational states of the ground electronic state. An equivalent splitting is done for the nonlinear response of $\beta_{\alpha\beta\gamma}$ and $\gamma_{\alpha\beta\gamma\delta}$.

Bishop and Kirtman proposed a new concise "square bracket" notation to represent vibrational contributions to NLOPs:

$$\alpha^v_{\alpha\beta}(-\omega_\sigma;\omega) = [\mu^2], \tag{4.75}$$

$$\beta^v_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = [\mu\alpha]+[\mu^3], \tag{4.76}$$

$$\gamma^v_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) = [\alpha^2]+[\mu\beta]+[\mu^2\alpha]+[\mu^4]. \tag{4.77}$$

Each square involves the following transition dipole moments between different vibrational states of the ground electronic state,

$$(\mu_\alpha)_{ij} = \langle i|\mu_\alpha^{00}|j\rangle, \tag{4.78}$$

$$(\alpha_{\alpha\beta})_{ij} = \langle i|\alpha_{\alpha\beta}^{00}|j\rangle, \tag{4.79}$$

$$(\beta_{\alpha\beta\gamma})_{ij} = \langle i|\beta_{\alpha\beta\gamma}^{00}|j\rangle, \tag{4.80}$$

as well as transition frequencies between vibrational states ω_i. The required square

bracket terms are calculated as follows:

$$[\mu^2] = \frac{1}{\hbar}\sum \mathscr{P}_{-\sigma,1}\sum_k{}'(\mu_\alpha)_{0k}(\mu_\beta)_{k0}\cdot\frac{1}{\omega_k-\omega_\sigma}\,, \tag{4.81}$$

$$[\mu\alpha] = \frac{1}{2\hbar}\sum \mathscr{P}_{-\sigma,1,2}\sum_k{}'(\mu_\alpha)_{0k}(\alpha_{\beta\gamma})_{k0}\cdot\left(\frac{1}{\omega_k+\omega_\sigma}+\frac{1}{\omega_k-\omega_\sigma}\right), \tag{4.82}$$

$$[\mu^3] = \frac{1}{\hbar^2}\sum \mathscr{P}_{-\sigma,1,2}\sum_{k,l}{}'(\mu_\alpha)_{0k}(\overline{\mu_\beta})_{kl}(\mu_\gamma)_{l0}\cdot\frac{1}{\omega_k-\omega_\sigma}\cdot\frac{1}{\omega_1-\omega_2}\,, \tag{4.83}$$

$$[\alpha^2] = \frac{1}{4\hbar}\sum \mathscr{P}_{-\sigma,1,2,3}\sum_k{}'(\alpha_{\alpha\beta})_{0k}(\alpha_{\gamma\delta})_{k0}\cdot\frac{1}{\omega_k-\omega_2-\omega_3}\,, \tag{4.84}$$

$$[\mu\beta] = \frac{1}{6\hbar}\sum \mathscr{P}_{-\sigma,1,2,3}\sum_k{}'(\mu_\alpha)_{0k}(\beta_{\beta\gamma\delta})_{k0}\cdot\left(\frac{1}{\omega_k+\omega_\sigma}+\frac{1}{\omega_k-\omega_\sigma}\right), \tag{4.85}$$

$$\begin{aligned}[\mu^2\alpha] = \frac{1}{2\hbar^2}\sum \mathscr{P}_{-\sigma,1,2,3}\sum_{k,l}{}'\Bigg[&(\mu_\alpha)_{0k}(\overline{\mu_\beta})_{kl}(\alpha_{\gamma\delta})_{l0}\cdot\left(\frac{1}{\omega_k+\omega_\sigma}+\frac{1}{\omega_k-\omega_\sigma}\right)\\ &\times\left(\frac{1}{\omega_l+\omega_2+\omega_3}+\frac{1}{\omega_l-\omega_2-\omega_3}\right)+(\mu_\alpha)_{0k}(\overline{\alpha_{\beta\gamma}})_{kl}(\mu_\delta)_{l0}\\ &\times\frac{1}{\omega_k-\omega_\sigma}\cdot\frac{1}{\omega_l-\omega_3}\Bigg]\,,\end{aligned} \tag{4.86}$$

$$\begin{aligned}[\mu^4] = \frac{1}{\hbar^3}\sum \mathscr{P}_{-\sigma,1,2,3}\Bigg[&\sum_{k,l,m}{}'(\mu_\alpha)_{0k}(\overline{\mu_\beta})_{kl}(\overline{\mu_\gamma})_{lm}(\mu_\delta)_{m0}\\ &\times\frac{1}{\omega_k-\omega_\sigma}\cdot\frac{1}{\omega_l-\omega_2-\omega_3}\cdot\frac{1}{\omega_m-\omega_3}\\ &-\sum_{k,l}{}'(\mu_\alpha)_{0k}(\mu_\beta)_{k0}(\mu_\gamma)_{0l}(\mu_\delta)_{l0}\cdot\frac{1}{\omega_k-\omega_\sigma}\cdot\frac{1}{\omega_l-\omega_3}\cdot\frac{1}{\omega_l+\omega_2}\Bigg]\,.\end{aligned} \tag{4.87}$$

As earlier in this Chapter, prime in the summation excludes ground vibrational states and $(\overline{\mu_\alpha})_{ij}$ is the dipole fluctuation coupling the ith and jth vibrational states:

$$(\overline{\mu_\alpha})_{ij} = \langle i|\overline{\mu_\alpha^{00}}|j\rangle = \langle i|\mu_\alpha^{00}|j\rangle - \delta_{ij}\langle 0,0|\hat{\mu}_\alpha|0,0\rangle\,. \tag{4.88}$$

However, a direct evaluation of Eqs. 4.81-4.87 is possible only for diatomics, when one generates the property profiles along the internuclear distance and utilizes the numerical technique of Numerov-Cooley to evaluate the necessary integrals [15, 112].

For the polyatomic molecules, one solves this problem by expanding static electric properties into the Taylor series in the normal modes $\mathbf{Q}$, which in the case of μ_α is:

$$\begin{aligned}\mu_\alpha(\mathbf{Q}) = \mu_\alpha^0 &+ \sum_a \left.\frac{\partial\mu_\alpha}{\partial Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0}}\cdot Q_a + \frac{1}{2}\sum_{a,b}\left.\frac{\partial^2\mu_\alpha}{\partial Q_a\partial Q_b}\right|_{\mathbf{Q}=\mathbf{Q_0}}\cdot Q_aQ_b\\ &+\frac{1}{6}\sum_{a,b,c}\left.\frac{\partial^3\mu_\alpha}{\partial Q_a\partial Q_b\partial Q_c}\right|_{\mathbf{Q}=\mathbf{Q_0}}\cdot Q_aQ_bQ_c+\cdots\,.\end{aligned} \tag{4.89}$$

In the above equation, μ_α^0 is the permanent dipole moment for the optimized geometry, denoted as $\mathbf{Q_0}$. Consecutive terms involve the derivatives of the dipole moment with respect to the normal coordinates (all evaluated for the optimized geometry). Equivalent expansions are adopted for other electric properties such as polarizability and hyperpolarizabilities. The first derivative of the dipole moment with respect to the normal coordinate corresponds to the purely harmonic approximation, whereas higher derivatives

define *electrical* anharmonicity. Namely, the second derivative of the dipole moment with respect to the normal coordinates is considered as the first-order electrical anharmonicity, and the third derivative is considered as the second-order electrical anharmonicity. For the (hyper)polarizabilities one utilizes the equivalent expansions and terminology.

Moreover, one expresses the vibrational wavefunction as a linear combination of harmonic oscillators, obtained by solving the eigenproblem for the following perturbed anharmonic potential:

$$V(\mathbf{Q}) = V^0 + \frac{1}{2}\sum_a \omega_a^2 Q_a^2 + \frac{1}{6}\sum_{a,b,c} F_{abc} Q_a Q_b Q_c + \frac{1}{24}\sum_{a,b,c,d} F_{abcd} Q_a Q_b Q_c Q_d + \cdots . \quad (4.90)$$

In the above definition, ω_a is the harmonic vibrational frequency of mode Q_a (please note the change of the notation). The first order *mechanical* anharmonicity of the potential energy is accounted with the cubic force constant, F_{abc}. The second order mechanical anharmonicity is described by the quartic force constants, F_{abcd}. These are defined as the high-order derivatives of the electronic energy $V(\mathbf{Q})$:

$$F_{abc} = \left.\frac{\mathrm{d}^3 V(\mathbf{Q})}{\mathrm{d}Q_a \mathrm{d}Q_b \mathrm{d}Q_c}\right|_{\mathbf{Q}=\mathbf{Q_0}}, \quad (4.91)$$

$$F_{abcd} = \left.\frac{\mathrm{d}^4 V(\mathbf{Q})}{\mathrm{d}Q_a \mathrm{d}Q_b \mathrm{d}Q_c \mathrm{d}Q_d}\right|_{\mathbf{Q}=\mathbf{Q_0}}. \quad (4.92)$$

.

In their original work [110, 111], Bishop and Kirtman accounted for the second-order electrical anharmonicity, i.e. Eq. 4.89, up to the cubic terms and the first-order mechanical anharmonicity, i.e. Eq. 4.90, up to the cubic terms. Few missing terms in the original formulation were added later in the work of Bishop, Luis and Kirtman [166]. This led to the separation of the square bracket terms in Eqs. 4.81–4.87 into contributions $[A]^{n,m}$ with n-th and m-th orders of electrical and mechanical anharmonicity, respectively. New square bracket formulas for vibrational NLOPs are [110–112]:

$$\alpha^v_{\alpha\beta}(-\omega_\sigma;\omega) = [\mu^2]^{0,0} + [\mu^2]^{2,0} + [\mu^2]^{1,1} + [\mu^2]^{0,2}, \quad (4.93)$$

$$\begin{aligned}\beta^v_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = {} & [\mu\alpha]^{0,0} + [\mu\alpha]^{2,0} + [\mu\alpha]^{1,1} + [\mu\alpha]^{0,2} \\ & + [\mu^3]^{1,0} + [\mu^3]^{0,1},\end{aligned} \quad (4.94)$$

$$\begin{aligned}\gamma^v_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) = {} & [\alpha^2]^{0,0} + [\alpha^2]^{2,0} + [\alpha^2]^{1,1} + [\alpha^2]^{0,2} \\ & + [\mu\beta]^{0,0} + [\mu\beta]^{2,0} + [\mu\beta]^{1,1} + [\mu\beta]^{0,2} \\ & + [\mu^2\alpha]^{1,0} + [\mu^2\alpha]^{0,1} + [\mu^4]^{2,0} + [\mu^4]^{1,1} + [\mu^4]^{0,2}.\end{aligned} \quad (4.95)$$

Formulas for the $[A]^{n,m}$ terms are expressed using the following shorthand notation for the property derivatives,

$$(\alpha/a) = \left.\frac{\mathrm{d}\mu_\alpha}{\mathrm{d}Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0}} \quad \text{and} \quad (\alpha/ab) = \left.\frac{\mathrm{d}^2\mu_\alpha}{\mathrm{d}Q_a\mathrm{d}Q_b}\right|_{\mathbf{Q}=\mathbf{Q_0}},$$

$$(\alpha\beta/a) = \left.\frac{\mathrm{d}\alpha_{\alpha\beta}}{\mathrm{d}Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0}} \quad \text{and} \quad (\alpha\beta/ab) = \left.\frac{\mathrm{d}^2\alpha_{\alpha\beta}}{\mathrm{d}Q_a\mathrm{d}Q_b}\right|_{\mathbf{Q}=\mathbf{Q_0}},$$

$$(\alpha\beta\gamma/a) = \left.\frac{\mathrm{d}\beta_{\alpha\beta\gamma}}{\mathrm{d}Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0}} \quad \text{and} \quad (\alpha\beta\gamma/ab) = \left.\frac{\mathrm{d}^2\beta_{\alpha\beta\gamma}}{\mathrm{d}Q_a\mathrm{d}Q_b}\right|_{\mathbf{Q}=\mathbf{Q_0}},$$

and the denominators involving harmonic vibrational frequencies of normal modes and incident frequencies,

$$\lambda_a^{\pm\sigma} = \lambda_a^{+\sigma}\lambda_a^{-\sigma},$$
$$\lambda_a^{\pm(\sigma-1)} = \lambda_a^{+(\sigma-1)}\lambda_a^{-(\sigma-1)},$$
$$\lambda_{ab}^{\pm\sigma} = \lambda_{ab}^{+\sigma}\lambda_{ab}^{-\sigma},$$

and the signed lambda terms are defined as:

$$\lambda_a^{+\sigma} = (\omega_a + \omega_\sigma)^{-1} \quad \text{and} \quad \lambda_a^{-\sigma} = (\omega_a - \omega_\sigma)^{-1},$$
$$\lambda_a^{+(\sigma-1)} = (\omega_a + \omega_\sigma - \omega_1)^{-1} \quad \text{and} \quad \lambda_a^{-(\sigma-1)} = (\omega_a - \omega_\sigma + \omega_1)^{-1},$$
$$\lambda_{ab}^{+\sigma} = (\omega_a + \omega_b + \omega_\sigma)^{-1} \quad \text{and} \quad \lambda_{ab}^{+-\sigma} = (\omega_a + \omega_b - \omega_\sigma)^{-1}.$$

Within the adopted anharmonic expansions, $[\mu^2]^{n,m}$ involve the first and second derivatives of μ with respect to normal modes and cubic force constants:

$$[\mu^2]^{0,0} = \frac{1}{2}\sum \mathscr{P}_{-\sigma,1}\sum_a (\alpha/a)(\beta/a)\lambda_a^{\pm\sigma}, \tag{4.96}$$

$$[\mu^2]^{2,0} = -\frac{\hbar}{8}\sum \mathscr{P}_{-\sigma,1}\sum_{a,b} (\alpha/ab)(\beta/ab)\lambda_b^{\pm\sigma}(\omega_a^{-1} + \omega_b^{-1}), \tag{4.97}$$

$$[\mu^2]^{1,1} = -\frac{\hbar}{4}\sum \mathscr{P}_{-\sigma,1}\sum_{a,b,c} \big[F_{abc}(\alpha/ab)(\beta/c)(\omega_a^{-1} + \omega_b^{-1})\lambda_{ab}^{\pm\sigma}\lambda_c^{\pm\sigma} + F_{bcc}(\alpha/ab)(\beta/a)\omega_b^{-2}\omega_c^{-1}\lambda_a^{\pm\sigma}\big], \tag{4.98}$$

$$[\mu^2]^{0,2} = -\frac{\hbar}{4}\sum \mathscr{P}_{-\sigma,1}\sum_{a,b,c} \omega_a^{-1}\Big[F_{aabc}(\alpha/b)(\beta/c)\lambda_b^{\pm\sigma}\lambda_c^{\pm\sigma} - \sum_d \big\{F_{aab}F_{bcd}(\alpha/c)(\beta/d)\omega_b^{-2}\lambda_c^{\pm\sigma}\lambda_d^{\pm\sigma} + 2F_{abc}F_{abd}(\alpha/c) \times(\beta/d)\lambda_{ab}^{\pm\sigma}\lambda_c^{\pm\sigma}\lambda_d^{\pm\sigma}\big\}\Big]. \tag{4.99}$$

Although not given in this manuscript, $[\alpha^2]^{n,m}$, $[\mu\alpha]^{n,m}$, and $[\mu\beta]^{n,m}$ terms are derived from the corresponding $[\mu^2]^{n,m}$ equation by performing few simple substitutions listed in Table V of Ref. [110].

Square bracket terms $[\mu^3]^{n,m}$ involve only the terms corresponding to the first-order electrical and mechanical anharmonicity:

$$[\mu^3]^{1,0} = \frac{1}{2}\sum \mathscr{P}_{-\sigma,1,2}\sum_{a,b} (\alpha/a)(\beta/ab)(\delta/b)\lambda_b^{\pm\sigma}\lambda_b^{\pm\sigma-1}, \tag{4.100}$$

$$[\mu^3]^{0,1} = -\frac{1}{6}\sum \mathscr{P}_{-\sigma,1,2}\sum_{a,b,c} F_{abc}(\alpha/a)(\beta/b)(\delta/c)\lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-(\sigma-1)}\lambda_c^{\pm\sigma-1}. \tag{4.101}$$

Similarly, $[\mu^2\alpha]^{n,m}$ involves not only first-order anharmonic terms, but also polarizability derivatives:

$$[\mu^2\alpha]^{1,0} = \frac{1}{4}\sum \mathscr{P}_{-\sigma,1,2,3}\sum_{a,b} \big[(\alpha/a)(\beta\gamma/ab)(\delta/b)\lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-1-2} + 2(\alpha/a)(\beta/ab)(\gamma\delta/b)\lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-1}\big], \tag{4.102}$$

$$[\mu^2\alpha]^{0,1} = -\frac{1}{4}\sum \mathscr{P}_{-\sigma,1,2,3}\sum_{a,b,c} F_{abc}(\alpha/a)(\beta/b)(\gamma\delta/c)\lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-(\sigma-1)}\lambda_c^{\pm\sigma-1}. \tag{4.103}$$

On the other hand, the $[\mu^4]^{n,m}$ terms involve only the terms second-order anharmonicity:

$$[\mu^4]^{2,0} = \frac{1}{2}\sum \mathscr{P}_{-\sigma,1,2,3} \sum_{a,b,c} (\alpha/a)(\beta/ab)(\gamma/bc)(\delta/c)\lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-1}\lambda_c^{\pm\sigma-1-2}, \quad (4.104)$$

$$[\mu^4]^{1,1} = -\frac{1}{6}\sum \mathscr{P}_{-\sigma,1,2,3} \sum_{a,b,c,d} F_{abc}(\alpha/a)(\beta/b)(\delta/cd)(\delta/d) \times \lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-(\sigma-1)}\lambda_c^{\pm\sigma-1}\lambda_d^{\pm\sigma-1-2}, \quad (4.105)$$

$$[\mu^4]^{2,0} = -\frac{1}{24}\sum \mathscr{P}_{-\sigma,1,2,3} \sum_{a,b,c,d} \Big[F_{abcd}(\alpha/a)(\beta/b)(\gamma/c)(\delta/d) \times \lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-(\sigma-1)}\lambda_c^{\pm\sigma-(\sigma-2)}\lambda_d^{\pm\sigma-1-2} - 3\sum_e F_{abc}F_{cde}(\alpha/a)(\beta/b) \times (\gamma/d)(\delta/e)\lambda_a^{\pm\sigma}\lambda_b^{\pm\sigma-(\sigma-1)}\lambda_c^{\pm\sigma-1}\lambda_d^{\pm\sigma-(\sigma-2)}\lambda_e^{\pm\sigma-1-2}\Big]. \quad (4.106)$$

By omitting all of the anharmonic terms in Eqs. 4.93-4.95, one is left with the double harmonic approximation in BKPT:

$$\alpha^{v,\mathrm{har}}_{\alpha\beta}(-\omega_\sigma;\omega) = [\mu^2]^{0,0}, \quad (4.107)$$

$$\beta^{v,\mathrm{har}}_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2) = [\mu\alpha]^{0,0}, \quad (4.108)$$

$$\gamma^{v,\mathrm{har}}_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3) = [\alpha^2]^{0,0} + [\mu\beta]^{0,0}. \quad (4.109)$$

Lastly, within the BKPT one can express the ZPVA correction as the sum of square bracket terms corresponding to different anharmonicity. For example, if one only considers the first-order electrical and mechanical anharmonicity, one can express the ZPVA correction for a property P (equivalent for α, β, γ) as:

$$\Delta P^{\mathrm{ZPVA}} = [P]^{0,1} + [P]^{1,0}. \quad (4.110)$$

with the square bracket terms defined as:

$$[P]^{0,1} = -\frac{\hbar}{4}\sum_{a,b} F_{abb}(\partial P/\partial Q_a)\omega_a^{-2}\omega_b^{-1}, = -\sum_a (\partial E^{\mathrm{ZP}}/\partial Q_a)(\partial P/\partial Q_a)\omega_a^{-2}, \quad (4.111)$$

$$[P]^{1,0} = \frac{\hbar}{4}\sum_a (\partial^2 P/\partial Q_a^2)\omega_a^{-2}. \quad (4.112)$$

In Eq. 4.111, E^{ZP} is the zero point energy obtained with the harmonic approximation. Even with such reduced level of anharmonicity, ZPVA corrections require more costly high-order derivatives of the property P. For example, $\Delta\gamma^{\mathrm{ZPVA}}_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3)$ involves already second order derivatives of γ^{el} with respect to normal modes. On the contrary, $\gamma^{v}_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2,\omega_3)$ defined by Eq. 4.95, does not involve derivatives of γ^{el}, but the derivatives of lower order NLOPs, such as α^{el} or β^{el}.

4.9 Alternative Representation of Nuclear Relaxation and Curvature Contributions

BKPT allows to describe the nonresonant vibrational response to both time-dependent and static electric fields. However, for the static fields there exists an alternative formulation

directly based on the derivatives of electronic properties [112–115, 149]. This approach is based on the double-series expansion in the normal modes and electric field of the electronic energy $V(\mathbf{Q},\mathbf{F})$:

$$V(\mathbf{Q},\mathbf{F}) = \sum_{n=0}^{} \sum_{a_1=1}^{3N-6} \cdots \sum_{a_n=1}^{3N-6} \sum_{m=0}^{} \sum_{\alpha_1=1}^{x,y,z} \cdots \sum_{\alpha_m=1}^{x,y,z} a_{nm}^{a_1\cdots a_n,\alpha_1\cdots\alpha_m} Q_{a_1}\cdots Q_{a_n} F_{\alpha_1}\cdots F_{\alpha_m}\,, \quad (4.113)$$

where Q_a is the a-th normal mode, and F_x, F_y, and F_z are the components of the static electric field vector $\mathbf{F}$. The consecutive factors are defined as the (mixed) derivatives:

$$a_{nm}^{a_1\cdots a_n,\alpha_1\cdots\alpha_m} = \frac{1}{n!m!}\cdot \left.\frac{\partial^{n+m}V(\mathbf{Q},\mathbf{F})}{\partial Q_{a_1}\ldots\partial Q_{a_n}\partial F_{\alpha_1}\ldots\partial F_{\alpha_m}}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}}. \quad (4.114)$$

These factors include the electronic contributions to the basic molecular electrical properties and their derivatives with respect to the normal coordinates, such as

$$\begin{aligned}
a_{01}^{\alpha} &= \left.\frac{\partial V}{\partial F_\alpha}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = -\mu_\alpha^0\,,\\
a_{02}^{\alpha\beta} &= \frac{1}{2}\left.\frac{\partial^2 V}{\partial F_\alpha \partial F_\beta}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = -\frac{1}{2}\alpha_{\alpha\beta}^{el}\,,\\
a_{03}^{\alpha\beta\gamma} &= \frac{1}{6}\left.\frac{\partial^3 V}{\partial F_\alpha \partial F_\beta \partial F_\gamma}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = -\frac{1}{6}\beta_{\alpha\beta\gamma}^{el},\\
a_{20}^{ab} &= \frac{1}{2}\left.\frac{\partial^2 V}{\partial Q_b \partial Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = \frac{1}{2}k_{ab}\,,\\
a_{30}^{abc} &= \frac{1}{6}\left.\frac{\partial^3 V}{\partial Q_c \partial Q_b \partial Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = \frac{1}{6}F_{abc}\,,\\
a_{11}^{a,\alpha} &= \frac{1}{2}\left.\frac{\partial^2 V}{\partial Q_a \partial F_\alpha}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = -\frac{1}{2}\left.\frac{\partial \mu_\alpha^0}{\partial Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}},\\
a_{21}^{ab,\alpha} &= \frac{1}{6}\left.\frac{\partial^3 V}{\partial Q_b \partial Q_a \partial F_\alpha}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = -\frac{1}{2}\left.\frac{\partial^2 \mu_\alpha^0}{\partial Q_b \partial Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}},\\
a_{12}^{a,\alpha\beta} &= \frac{1}{6}\left.\frac{\partial^3 V}{\partial Q_a \partial F_\beta \partial F_\alpha}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}} = -\frac{1}{6}\left.\frac{\partial \alpha_{\alpha\beta}^{el}}{\partial Q_a}\right|_{\mathbf{Q}=\mathbf{Q_0},\mathbf{F}=\mathbf{0}}.
\end{aligned}$$

Moreover, it is convenient to define the following q_1 and q_2 terms, which are derivatives with respect to a given normal mode of the dipole moment and the static polarizability, scaled with the harmonic frequency of this mode:

$$q_1^{a,\alpha} = \frac{a_{11}^{a,\alpha}}{2a_{20}^{aa}}\,, \quad (4.115)$$

$$q_2^{a,\alpha\beta} = \frac{a_{12}^{a,\alpha\beta}}{2a_{20}^{aa}}\,. \quad (4.116)$$

Imposing the condition that the field-dependent molecular structure is always optimized ($\partial V(\mathbf{Q},\mathbf{F})/\partial Q_i = 0, \forall i$), one can obtain the following expression for the field-

dependent normal coordinates displacements:

$$Q_a(\mathbf{F}) = -\sum_{\alpha}^{x,y,z} q_1^{a,\alpha} F_\alpha - \sum_{\alpha,\beta}^{x,y,z} \left[q_2^{a,\alpha\beta} - \sum_{b}^{3N-6} \frac{a_{21}^{ab,\alpha}}{a_{20}^{aa}} q_1^{a,\beta} + \sum_{b,c}^{3N-6} \frac{3a_{30}^{abc}}{2a_{20}^{aa}} q_1^{b,\alpha} q_1^{c,\beta} \right] F_\alpha F_\beta + \cdots . \tag{4.117}$$

Inserting the above equation back into Eq. 4.113 and grouping the terms by power in electric field strength, one can reach new definitions of vibrational NLOPs.

Using this approach, in contrast to the Eq. 4.73, the overall response of the vibrational structure of the system is partitioned into nuclear relaxation (NR) and curvature (curv) terms,

$$P^{vibr} = P^{NR} + P^{curv}, \tag{4.118}$$

Interpretation of the NR terms is very simple - they quantify changes in the polarization due to the reorganization of the molecular structure due to the electric field perturbation. On the other hand, the curvature contributions quantify the change in the polarization due field-induced perturbation of the PES curvature.

The final expressions to the static linear and nonlinear optical properties are listed below. The NR contribution to polarizability is simply described as

$$\alpha_{\alpha\beta}^{NR}(0;0) = \frac{1}{2} \sum \mathscr{P}_{\alpha\beta} \sum_{a}^{3N-6} a_{11}^{a,\alpha} q_1^{a,\beta}, \tag{4.119}$$

with $\sum \mathscr{P}_{\alpha\beta}$ denoting permutation operator exchanging solely the α and β upper indices. $\alpha_{\alpha\beta}^{NR}(0;0)$ depends solely on the quantities obtained at the harmonic approximation, i.e., no electrical and mechanical anharmonic terms are involved. Next, the NR contribution to the first hyperpolarizability is defined as

$$\beta_{\alpha\beta\gamma}^{NR}(0;0,0) = \sum \mathscr{P}_{\alpha\beta\gamma} \left[\sum_{a}^{3N-6} a_{12}^{a,\alpha\beta} q_1^{a,\gamma} - \sum_{a,b}^{3N-6} a_{21}^{ab,\alpha} q_1^{a,\beta} q_1^{b,\gamma} + \sum_{a,b,c}^{3N-6} a_{30}^{abc} q_1^{a,\alpha} q_1^{b,\beta} q_1^{c,\gamma} \right], \tag{4.120}$$

with $\sum \mathscr{P}_{\alpha\beta\gamma}$ denoting permutation operator exchanging the α, β, and γ upper indices. The first summation constitutes for the purely harmonic term, while the second and third involve electrical ($a_{21}^{ab,\alpha}$) and mechanical (a_{30}^{abc}) anharmonicity. Lastly, NR of the second

hyperpolarizability is computed as,

$$\begin{aligned}\gamma^{NR}_{\alpha\beta\gamma\delta}(0;0,0,0) = \sum \mathscr{P}_{\alpha\beta\gamma\delta} \Bigg[& \sum_{a}^{3N-6} \frac{1}{2} a_{12}^{a,\alpha\beta} q_2^{a,\gamma\delta} + \sum_{a}^{3N-6} a_{13}^{a,\alpha\beta\gamma} q_1^{a,\delta} \\ & - \sum_{a,b}^{3N-6} \left(a_{22}^{ab,\alpha\beta} q_1^{a,\gamma} q_1^{b,\delta} + 2 a_{21}^{ab,\alpha} q_2^{b,\gamma\delta} \right) \\ & + \sum_{a,b,c}^{3N-6} 3 a_{30}^{abc} q_1^{a,\alpha} q_1^{b,\beta} q_2^{c,\gamma\delta} \\ & + \sum_{a,b,c}^{3N-6} \left(a_{31}^{abc,\alpha} q_1^{a,\beta} q_1^{b,\gamma} q_1^{c,\delta} + \frac{a_{21}^{ab,\alpha} a_{21}^{bc,\beta}}{a_{20}^{bb}} q_1^{a,\gamma} q_1^{c,\delta} \right) \\ & + \sum_{a,b,c,d}^{3N-6} \left(-a_{40}^{abcd} q_1^{a,\alpha} q_1^{b,\beta} q_1^{c,\gamma} q_1^{d,\delta} + \sum_{e}^{3N-6} \frac{9 a_{30}^{abc} a_{30}^{cde}}{4 a_{20}^{cc}} q_1^{a,\alpha} q_1^{b,\beta} q_1^{d,\gamma} q_1^{e,\delta} \right) \\ & - \sum_{a,b,c,d}^{3N-6} \frac{3 a_{30}^{abc} a_{21}^{cd,\alpha}}{a_{20}^{cc}} q_1^{a,\gamma} q_1^{b,\delta} q_1^{d,\beta} \Bigg] , \end{aligned} \tag{4.121}$$

where, again, $\sum \mathscr{P}_{\alpha\beta\gamma\delta}$ denotes the permutation operator exchanging the α, β, γ, and δ upper indices. Except the first summation, all summations involve anharmonic derivatives. Normally, the analytical expressions for those high-order derivatives are not available (or are not implemented), and one has to obtain them using numerical differentiation. This fact combined with the extremely large number of mixed derivatives to be computed, makes a direct usage of Eq. 4.121 very costly.

There exist a direct correspondence of the NR terms with the BKPT square bracket terms present in Eqs. 4.93 - 4.95 (when one sets all incident frequencies $\omega = 0$):

$$\alpha^{NR}_{\alpha\beta}(0;0) = [\mu^2]^{0,0}_{\omega=0} , \tag{4.122}$$

$$\beta^{NR}_{\alpha\beta\gamma}(0;0,0) = [\mu\alpha]^{0,0}_{\omega=0} + [\mu^3]^{1,0}_{\omega=0} + [\mu^3]^{0,1}_{\omega=0} , \tag{4.123}$$

$$\begin{aligned}\gamma^{NR}_{\alpha\beta\gamma\delta}(0;0,0,0) = & [\alpha^2]^{0,0}_{\omega=0} + [\mu\beta]^{0,0}_{\omega=0} + [\mu^2\alpha]^{1,0}_{\omega=0} + [\mu^2\alpha]^{0,1}_{\omega=0} \\ & + [\mu^4]^{2,0}_{\omega=0} + [\mu^4]^{0,2}_{\omega=0} + [\mu^4]^{1,1}_{\omega=0} . \end{aligned} \tag{4.124}$$

In the above relations, each of the square bracket terms corresponds to the different summation in Eqs. 4.119-4.121 (the order of the terms in each set of equations is preserved). Nuclear relaxation terms involve the lowest order terms of the different square bracket BK contributions, defined as the sum $n+m$ in $[A]^{n+m}$. Square bracket terms of the higher anharmonicity, along with the ZPVA correction, constitute to the curvature terms. For example, the curvature contribution to the polarizability, $\alpha^{curv}_{\alpha\beta}(0,0)$, can be expressed with the BKPT square brackets as:

$$\alpha^{curv}_{\alpha\beta}(0,0) = \Delta\alpha^{\mathrm{ZPVA}} + [\mu^2]^{2,0}_{\omega=0} + [\mu^2]^{1,1}_{\omega=0} + [\mu^2]^{0,2}_{\omega=0} + \dots . \tag{4.125}$$

Concluding, the response to the static electric fields of molecular systems can be expressed both with BKPT and the derivative NR+curvature scheme. Both approaches are simply different partitionings of the same overall vibrational response. However, the calculation of the nuclear relaxation terms can be more feasible when one adapts one of the two methods described in the following sections, namely, the finite-field nuclear relaxation approach or evaluation of NR expressions using the field-induced coordinates.

4.10 Bishop-Hasan-Kirtman Finite-Field Nuclear Relaxation Approach

Bishop, Hasan and Kirtman presented a very efficient and concise approach to calculate nuclear relaxation and curvature terms, the Finite-Field Nuclear Relaxation (FF-NR) approach [116]. In FF-NR, one does not need to explicitly calculate the high-order derivatives, which is especially useful for larger chemical systems.

It requires computations of the static electronic NLOPs under the perturbation with static field $\mathbf{F}$ for both the equilibrium geometry relaxed in the electric field, $\mathbf{R_F}$, and "frozen" equilibrium geometry of the system, $\mathbf{R_0}$, i.e. the one optimized without the presence of the electric field. During the geometry optimizations, the field-free Eckart conditions must be enforced to prohibit reorientation [117].

One defines two main quantities, namely a difference in electronic properties due to the field perturbation, $(\Delta P^{el})_{\mathbf{R_0}}$, calculated for the same field-free optimized structure,

$$(\Delta P^{el})_{\mathbf{R_0}} = P^{el}(\mathbf{F},\mathbf{R_0}) - P^{el}(\mathbf{0},\mathbf{R_0})\,, \tag{4.126}$$

and a similar difference, $(\Delta P^{el})_{\mathbf{R_F}}$ involving the corresponding field-free and field-relaxed equilibrium geometries,

$$(\Delta P^{el})_{\mathbf{R_F}} = P^{el}(\mathbf{F},\mathbf{R_F}) - P^{el}(\mathbf{0},\mathbf{R_0})\,. \tag{4.127}$$

For different dipolar properties μ_α, $\alpha_{\alpha\beta}$, and $\beta_{\alpha\beta\gamma}$ one can expand $(\Delta P^{el})_{\mathbf{R_0}}$ as the following Taylor series in the field strength:

$$\begin{aligned}(\Delta\mu^{el}_\alpha)_{\mathbf{R_0}} = &\sum_\beta \alpha^{el}_{\alpha\beta}(0;0)F_\beta + \frac{1}{2}\sum_{\beta,\gamma}\beta^{el}_{\alpha\beta\gamma}(0;0,0)F_\beta F_\gamma \\ &+ \frac{1}{6}\sum_{\beta,\gamma,\delta}\gamma^{el}_{\alpha\beta\gamma\delta}(0;0,0,0)F_\beta F_\gamma F_\delta + \cdots\,,\end{aligned} \tag{4.128}$$

$$(\Delta\alpha^{el}_{\alpha\beta})_{\mathbf{R_0}} = \sum_\gamma \beta^{el}_{\alpha\beta\gamma}(0;0,0)F_\gamma + \frac{1}{2}\sum_{\gamma,\delta}\gamma^{el}_{\alpha\beta\gamma\delta}(0;0,0,0)F_\gamma F_\delta + \cdots\,, \tag{4.129}$$

$$(\Delta\beta^{el}_{\alpha\beta\gamma})_{\mathbf{R_0}} = \sum_\delta \gamma^{el}_{\alpha\beta\gamma\delta}(0;0,0,0)F_\delta + \cdots\,. \tag{4.130}$$

The above equations allow to obtain the consecutive electronic parts of NLOPs through numerical differentiation or polynomial fitting:

$$\alpha^{el}_{\alpha\beta}(0;0) = \left.\frac{\partial(\Delta\mu^{el}_\alpha)_{\mathbf{R_0}}}{\partial F_\beta}\right|_{\mathbf{F}=\mathbf{0}}\,, \tag{4.131}$$

$$\beta^{el}_{\alpha\beta\gamma}(0;0,0) = \left.\frac{\partial^2(\Delta\mu^{el}_\alpha)_{\mathbf{R_0}}}{\partial F_\beta\partial F_\gamma}\right|_{\mathbf{F}=\mathbf{0}} = \left.\frac{\partial(\Delta\alpha^{el}_{\alpha\beta})_{\mathbf{R_0}}}{\partial F_\gamma}\right|_{\mathbf{F}=\mathbf{0}}\,, \tag{4.132}$$

$$\gamma^{el}_{\alpha\beta\gamma\delta}(0;0,0,0) = \left.\frac{\partial^3(\Delta\mu^{el}_\alpha)_{\mathbf{R_0}}}{\partial F_\beta\partial F_\gamma\partial F_\delta}\right|_{\mathbf{F}=\mathbf{0}} = \left.\frac{\partial^2(\Delta\alpha^{el}_{\alpha\beta})_{\mathbf{R_0}}}{\partial F_\gamma\partial F_\delta}\right|_{\mathbf{F}=\mathbf{0}} = \left.\frac{\partial(\Delta\beta^{el}_{\alpha\beta\gamma})_{\mathbf{R_0}}}{\partial F_\delta}\right|_{\mathbf{F}=\mathbf{0}}\,. \tag{4.133}$$

The authors showed that $(\Delta P^{el})_{\mathbf{R_F}}$ can be expanded in a similar manner,

$$(\Delta\mu_\alpha^{el})_{\mathbf{R_F}} = \sum_\beta a^I_{\alpha\beta}F_\beta + \frac{1}{2}\sum_{\beta,\gamma} b^I_{\alpha\beta\gamma}F_\beta F_\gamma + \frac{1}{6}\sum_{\beta,\gamma,\delta} g^I_{\alpha\beta\gamma\delta}F_\beta F_\gamma F_\delta + \cdots, \tag{4.134}$$

$$(\Delta\alpha_{\alpha\beta}^{el})_{\mathbf{R_F}} = \sum_\gamma b^{II}_{\alpha\beta\gamma}F_\gamma + \frac{1}{2}\sum_{\gamma,\delta} g^{II}_{\alpha\beta\gamma\delta}F_\gamma F_\delta + \cdots, \tag{4.135}$$

$$(\Delta\beta_{\alpha\beta\gamma}^{el})_{\mathbf{R_F}} = \sum_\delta g^{III}_{\alpha\beta\gamma\delta}F_\delta + \cdots, \tag{4.136}$$

but the factors, defined as the consecutive derivatives, contribute both to the electronic and NR terms. Therefore, the differentiation of $(\Delta\mu_\alpha^{el})_{\mathbf{R_F}}$ with respect to the applied field yields,

$$a^I_{\alpha\beta} = \left.\frac{\partial(\Delta\mu_\alpha^{el})_{\mathbf{R_F}}}{\partial F_\beta}\right|_{\mathbf{F}=\mathbf{0}} = \alpha^{el}_{\alpha\beta}(0;0) + \alpha^{NR}_{\alpha\beta}(0;0), \tag{4.137}$$

$$b^I_{\alpha\beta\gamma} = \left.\frac{\partial^2(\Delta\mu_\alpha^{el})_{\mathbf{R_F}}}{\partial F_\beta \partial F_\gamma}\right|_{\mathbf{F}=\mathbf{0}} = \beta^{el}_{\alpha\beta\gamma}(0;0,0) + \beta^{NR}_{\alpha\beta\gamma}(0;0,0), \tag{4.138}$$

$$g^I_{\alpha\beta\gamma\delta} = \left.\frac{\partial^3(\Delta\mu_\alpha^{el})_{\mathbf{R_F}}}{\partial F_\beta \partial F_\gamma \partial F_\delta}\right|_{\mathbf{F}=\mathbf{0}} = \gamma^{el}_{\alpha\beta\gamma\delta}(0;0,0,0) + \gamma^{NR}_{\alpha\beta\gamma\delta}(0;0,0,0). \tag{4.139}$$

On the contrary, the differentiation of $(\Delta\alpha_{\alpha\beta}^{el})_{\mathbf{R_F}}$ and $(\Delta\beta_{\alpha\beta\gamma}^{el})_{\mathbf{R_F}}$ leads to NR terms of the first and second hyperpolarizabilities in the infinity frequency regime:

$$b^{II}_{\alpha\beta\gamma} = \left.\frac{\partial(\Delta\alpha_{\alpha\beta}^{el})_{\mathbf{R_F}}}{\partial F_\gamma}\right|_{\mathbf{F}=\mathbf{0}} = \beta^{el}_{\alpha\beta\gamma}(0;0,0) + \beta^{NR}_{\alpha\beta\gamma}(-\omega;\omega,0)_{\omega\to\infty}, \tag{4.140}$$

$$g^{II}_{\alpha\beta\gamma\delta} = \left.\frac{\partial^2(\Delta\alpha_{\alpha\beta}^{el})_{\mathbf{R_F}}}{\partial F_\gamma \partial F_\delta}\right|_{\mathbf{F}=\mathbf{0}} = \gamma^{el}_{\alpha\beta\gamma\delta}(0;0,0,0) + \gamma^{NR}_{\alpha\beta\gamma\delta}(-\omega;\omega,0,0)_{\omega\to\infty}, \tag{4.141}$$

$$g^{III}_{\alpha\beta\gamma\delta} = \left.\frac{\partial(\Delta\beta_{\alpha\beta\gamma}^{el})_{\mathbf{R_F}}}{\partial F_\delta}\right|_{\mathbf{F}=\mathbf{0}} = \gamma^{el}_{\alpha\beta\gamma\delta}(0;0,0,0) + \gamma^{NR}_{\alpha\beta\gamma\delta}(-2\omega;\omega,\omega,0)_{\omega\to\infty}. \tag{4.142}$$

By subtracting purely electronic terms, one can obtain NR contributions to NLOPs.

4.11 Field-Induced Coordinates

In their seminal work, Luis *et al.* proposed another efficient way of computing the NR terms with the usage of the Field-Induced Coordinates (FICs) [118, 119]. FICs are defined as linear combinations of normal modes and describe the displacements of atoms induced by the external electric field. Based on the Taylor expansions of the potential and normal coordinates (see Eqs. 4.113 and 4.117) authors define FICs of different orders. The first-order FIC (FIC1), denoted as χ_1^α and having three Cartesian components, is defined as:

$$\chi_1^\alpha = \sum_a^{3N-6} \frac{\partial Q_a^F}{\partial F_\alpha} Q_a = -\sum_a^{3N-6} q_1^{a,\alpha} Q_a, \tag{4.143}$$

FIC1 is defined within the double harmonic approximation (only the linear derivatives in normal modes are involved), however, as it will be shown later, it is used to describe the

anharmonic response as well. The second-order FIC (FIC2), denoted as $\chi_2^{\alpha\beta}$ and having nine Cartesian components, is defined as:

$$\begin{aligned}\chi_2^{\alpha\beta} &= \frac{1}{2}\sum_{a,b}^{3N-6}\frac{\partial^2 Q_a^F}{\partial F_\alpha \partial F_\beta}Q_a\,,\\ &= -\sum_{a}^{3N-6}\left[q_2^{a,\alpha\beta} - \sum_{b}^{3N-6}\frac{a_{21}^{ab,\alpha}}{a_{20}^{aa}}q_1^{b,\beta} + \sum_{b,c}^{3N-6}\frac{3a_{30}^{abc}}{2a_{20}^{aa}}q_1^{b,\alpha}q_1^{c,\beta}\right]Q_a\,,\\ &= \chi_{2,har}^{\alpha\beta} + \Delta\chi_{2,anh}^{\alpha\beta}\,. \end{aligned} \tag{4.144}$$

FIC2 constitutes terms involving both the harmonic and anharmonic parts:

$$\chi_{2,har}^{\alpha\beta} = -\sum_{a}^{3N-6} q_2^{a,\alpha\beta}Q_a\,, \tag{4.145}$$

$$\Delta\chi_{2,anh}^{\alpha\beta} = -\sum_{a}^{3N-6}\left[-\sum_{b}^{3N-6}\frac{a_{21}^{ab,\alpha}}{a_{20}^{aa}}q_1^{b,\beta} + \sum_{b,c}^{3N-6}\frac{3a_{30}^{abc}}{2a_{20}^{aa}}q_1^{b,\alpha}q_1^{c,\beta}\right]. \tag{4.146}$$

The main feature of FICs is that they drastically reduce the cost of computation of the NR contributions to NLOPs. The formulae for the properties are equivalent to Eqs. 4.119 - 4.121 but with a_{nm}, q_1, and q_2 being evaluated as the derivatives with respect to the three FIC1 and the six FIC2, and not the $3N-6$ normal coordinates. This exceptional feature is shown below for the example of NR contribution to static polarizability, which requires usage of only two FIC1 as shown by the authors. The derivation involves a formal representation of FICs as a new set of coordinates denoted as ϕ,

$$\phi_i = \sum_b M_{ia}Q_a\,. \tag{4.147}$$

and from the definition of FIC1, one gets $M_{1a} = -q_1^{a,\alpha}$. With few basic operations such as the derivative chain rule, one can change the basis of $\alpha_{\alpha\beta}^{NR}(0;0)$ as follows:

$$\begin{aligned}\alpha_{\alpha\beta}^{NR}(0;0) &= \sum\mathscr{P}_{\alpha\beta}\sum_{a}^{3N-6}\frac{1}{2}a_{11}^{a,\alpha}q_1^{a,\beta}\,,\\ &= \sum\mathscr{P}_{\alpha\beta}\sum_{a}^{3N-6}\frac{\partial\mu_\alpha}{\partial Q_a}\frac{\partial Q_a^F}{\partial F_\beta} = \sum\mathscr{P}_{\alpha\beta}\sum_{a,b}^{3N-6}\frac{\partial\mu_\alpha}{\partial\phi_b}\frac{\partial\phi_b}{\partial Q_a}\frac{\partial Q_a^F}{\partial F_\beta}\,,\\ &= \sum\mathscr{P}_{\alpha\beta}\sum_{a,b}^{3N-6}\frac{\partial\mu_\alpha}{\partial\phi_b}M_{ba}M_{1a} = \sum\mathscr{P}_{\alpha\beta}\frac{\partial\mu_\alpha}{\partial\phi_1}\sum_{a}^{3N-6}M_{1a}M_{1a}\,,\\ &= \sum\mathscr{P}_{\alpha\beta}\frac{\partial\mu_\alpha}{\partial\phi_1}\sum_{a}^{3N-6}M_{1a}\frac{\partial Q_a^F}{\partial F_\alpha} = \sum\mathscr{P}_{\alpha\beta}\frac{\partial\mu_\alpha}{\partial\phi_1}\frac{\partial\phi_1^F}{\partial F_\alpha}\,,\\ &= \sum\mathscr{P}_{\alpha\beta}\frac{1}{2}a_{11}^{\phi_1,\alpha}q_1^{\phi_1,\beta}\,. \end{aligned} \tag{4.148}$$

As a result, one expresses $\alpha_{\alpha\beta}^{NR}(0;0)$ with the exactly the same formulae as before, but in the FIC1 basis. The above evaluation requires the usage of only two Cartesian components of FIC1, namely χ_1^α and χ_1^β. Moreover, one has to always use just these two FIC1 coordinates, regardless of number of atoms in the system. An equivalent demonstration could be done for any NR component of static NLOPs, as well as those for the infinite frequency limit. To calculate $\beta_{\alpha\beta\gamma}^{NR}(0;0,0)$ using FICs, one only has to utilize all

three components of FIC1, χ_1^i (i=α, β, γ). Computation of $\gamma^{NR}_{\alpha\beta\gamma\delta}(0;0,0,0)$ is much more costly, because one has to utilize also all of the FIC2 components, χ_2^{ij} (i, j =α, β, γ). This still requires performing some numerical differentiation to obtain the high-order derivatives, such as a_{30} and a_{40} (expressed in the FIC1 and FIC2 basis), however, it involves many orders of magnitude less than if one calculated everything in the normal modes basis. For example, to compute $\gamma^{NR}_{\alpha\beta\gamma\delta}(0;0,0,0)$ using the normal modes basis, one requires roughly $(3N-6)^4$ mixed derivatives of the a_{40} type, among others. Moreover, as shown by the authors, some terms of $\gamma^{NR}_{\alpha\beta\gamma\delta}(0;0,0,0)$ can be calculated using FIC1 only (that is without FIC2), and, furthermore, one can use the harmonic part of FIC2 (see Eq. 4.145) to calculate $\gamma^{NR}_{\alpha\beta\gamma\delta}(-\omega;\omega,0,0)$ and $\gamma^{NR}_{\alpha\beta\gamma\delta}(-\omega;\omega,-\omega,\omega)$ in the infinite frequency limit $\omega\to\infty$, as well as the $[\alpha^2]^{0,0}_{\omega=0}$ part of $\gamma^{NR}_{\alpha\beta\gamma\delta}(0;0,0,0)$.

By comparison to the arbitrary frequency BKPT terms, FICs have been expanded also to frequency-dependent NLOPs in the nonresonant region. It simply requires changing the denominators of several terms, such as q_1, to include frequency dispersion [119]. For example, FD-FIC of the first order is defined as [119],

$$\chi^{\alpha}_{1,|\omega|} = -\sum_{a}^{3N-6} q^{a,\alpha}_{1,|\omega|} Q_a\,, \tag{4.149}$$

which is an equivalent definition of Eq. 4.143, but with the modified q_1 term,

$$q^{a,\alpha}_{1,|\omega|} = \frac{a^{a,\alpha}_{11}}{2a^{aa}_{20}-\omega^2}\,. \tag{4.150}$$

Then, $\alpha^{NR}_{\alpha\beta}(-\omega;\omega)$ and $\beta^{NR}_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2)$ are defined again with the equivalent formulae 4.119 and 4.120, but with frequency dependent a_{nm} and $q^{a,\alpha}_{1,|\omega|}$ quantities. As expected, formulae for $\alpha^{NR}_{\alpha\beta}(-\omega;\omega)$ and $\beta^{NR}_{\alpha\beta\gamma}(-\omega_\sigma;\omega_1,\omega_2)$ reduce properly to their static and infinite frequency limit counterparts, when one sets $\omega=0$ or $\omega\to\infty$. As in the case of static FICs, the number of necessary FD-FIC coordinates is still far smaller than if one utilized all normal modes in the evaluation. However, the usage of FD-FICs in the computation of $\gamma^{NR}_{\alpha\beta\gamma\delta}(-\omega_\sigma;\omega_1,\omega_2;\omega_3)$ requires much more computational effort, but still far less than if one utilized the normal mode representation of Eq. 4.121.

Chapter 5

Density Functional Theory

5.1 Introduction

This chapter is devoted to the introduction of fundamental aspects of Density Functional Theory (DFT), shifted towards computational chemistry. The discussion is more focused on the practical aspects of DFT, not on the strict theoretical properties and constraints, on which DFT is built. A part of this Chapter belongs to a fundamental approximation made in DFT, namely the computation of integrals through the numerical integration, which is directly related to the work presented in Chapters 8, 9, and 10.

5.2 Wavefunction Representation of a Quantum System

The most fundamental differential equation of quantum mechanics, the Schrödinger equation,

$$\hat{H}|\Psi\rangle = E|\Psi\rangle\,, \tag{5.1}$$

defines a set of pure quantum states of any quantum-mechanical system. Each of the states is described with its unique wavefunction $|\Psi\rangle$ and is characterized with a total energy E. These so-called eigensolutions are specific for a given energy operator, the Hamiltonian $\hat{H}$, which uniquely defines the quantum system.

The wavefunction vectors describing the states of N electrons are very complicated quantities, depending on the 4N coordinates ($3N$ spatial and N spin coordinates),

$$|\Psi\rangle = \Psi(\mathbf{r}\sigma, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_N\sigma_N)\,, \tag{5.2}$$

where $\mathbf{r}_i = (x_i, y_i, z_i)$ is the vector of position of the i-th electron, and the spin of the electron is represented with σ_i=α or β (the first electron denoted without the lower index). $|\Psi\rangle$ contains all possible information on the particular quantum state. Among others, one of the most important features of $|\Psi\rangle$ is its probabilistic interpretation and connection to an experimentally measurable quantity, the electronic density $\rho(\mathbf{r})$ and its spin component $\rho_\sigma(\mathbf{r})$. The electronic spin density, $\rho_\sigma(\mathbf{r})$, is defined through the integration of $|\Psi\rangle$ over all electronic coordinates but one, namely

$$\rho_\sigma(\mathbf{r}) = N \sum_{\sigma_2 \ldots \sigma_N} \int \mathrm{d}\mathbf{r}_2 \cdots \int \mathrm{d}\mathbf{r}_N |\Psi(\mathbf{r}\sigma, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_N\sigma_N)|^2\,, \tag{5.3}$$

where $\mathrm{d}\mathbf{r} = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ denotes the infinitesimal volume element. Then, the total electronic density, $\rho(\mathbf{r})$, is defined as a sum of its two spin components:

$$\rho(\mathbf{r}) = \rho_\alpha(\mathbf{r}) + \rho_\beta(\mathbf{r})\,. \tag{5.4}$$

With the use of $\rho(\mathbf{r})$, one can quantify the probability of finding any electron within the volume element $\mathrm{d}\mathbf{r}$ centered at position r, which is simply evaluated as $\rho(\mathbf{r})\mathrm{d}\mathbf{r}$.

On the other hand, the Hamiltonian contains the information on the number and kind of particles, how they interact with each other, and how they interact with an external perturbation, if present. This Thesis focuses solely on the electronic structure of the molecular systems. The separation of the nuclear motion is done by the means of the Born-Oppenheimer approximation, and furthermore, the nuclei are treated as classical point charges. Hence in this work, only the electronic part of the Hamiltonian is considered. In such case, the Hamiltonian of the system containing N_{el} electrons and N_{nuc} nuclei is defined as a sum of several components:

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V}_{ext}\,, \tag{5.5}$$

where each of them has its own physical meaning. The two first operators in Eq. 5.5 are universal for any kind of system containing electrons, being the one-electron kinetic energy operator $\hat{T}$,

$$\hat{T} = -\frac{1}{2}\sum_{i=1}^{N_{el}} \nabla_i^2 , \tag{5.6}$$

and the potential energy of the electrostatic repulsion between electron pairs, $\hat{V}_{ee}$ (which is a two-electron operator),

$$\hat{V}_{ee} = \frac{1}{2}\sum_{i}^{N_{el}}\sum_{j\neq i}^{N_{el}} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} . \tag{5.7}$$

The third term in Eq. 5.5, $\hat{V}_{ext}$, is actually system dependent and describes the interaction of electrons with an unique external potential. In general, this potential may be one-electron (i.e. affects each electron separately) or many-electron (i.e. affects several of them simultaneously). In this work, only one-electron external potentials are considered, which makes $\hat{V}_{ext}$ to be expressed as a sum of the N_{el} one-electron components,

$$\hat{V}_{ext} = \sum_{i=1}^{N_{el}} \hat{v}_{\mathrm{ext},\sigma}(\mathbf{r}_i) , \tag{5.8}$$

In general, $\hat{v}_{\mathrm{ext},\sigma}(\mathbf{r}_i)$ may be any type of potential, such as the electrostatic attraction of electrons and nuclei, the interaction with external electric and magnetic fields (the latter depends on the spin of electrons as well), mechanic pressure spatially confining electrons, or any model interaction as well, such as the model of Hooke's atom. For a molecular system without any external perturbation, $\hat{v}_{\mathrm{ext},\sigma}(\mathbf{r}_i)$ simply constitutes the potential energy between the i-th electron and the set of N_{nuc} nuclei,

$$\hat{v}_{\mathrm{ext},\sigma}(\mathbf{r}_i) = \hat{v}_{\mathrm{ext}}(\mathbf{r}_i) = \sum_{I}^{N_{nuc}} \frac{-Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} . \tag{5.9}$$

Accordingly, the total energy of the system can also be partitioned into several components,

$$\begin{aligned} E &= \langle \hat{H} \rangle , \\ &= \langle \hat{T} \rangle + \langle \hat{V}_{ee} \rangle + \langle \hat{V}_{ext} \rangle , \end{aligned} \tag{5.10}$$

where $\langle \hat{O} \rangle$ is the expectation value of the operator $\hat{O}$ and is formally calculated as

$$\begin{aligned} \langle \hat{O} \rangle &= \langle \Psi | \hat{O} | \Psi \rangle , \\ &= \sum_{\sigma_1 \ldots \sigma_N} \int \mathrm{d}\mathbf{r} \cdots \int \mathrm{d}\mathbf{r}_N \Psi^*(\mathbf{r}\sigma, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_N\sigma_N) \hat{O} \Psi(\mathbf{r}\sigma, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_N\sigma_N) . \end{aligned} \tag{5.11}$$

However, in addition to the tremendous amount of work required to solve the Schrödinger equation (which has been solved exactly for a limited number of model systems), working with such complex structure as $|\Psi\rangle$ is expensive from the computational point of view. Because the electronic Hamiltonian of Eq. 5.5 consists of only one- and two-electron operators, one can formally work with simpler quantities called reduced density matrices [167], to express total energy and other (molecular) properties.

The one-particle reduced density matrix (1-RDM), $\rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma)$, is defined as [167]

$$\begin{aligned} \rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma) = N \sum_{\sigma_2 \ldots \sigma_N} \int \mathrm{d}\mathbf{r}_2 \cdots \int \mathrm{d}\mathbf{r}_N \Psi^*(\mathbf{r}'\sigma, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_N\sigma_N) \\ \times \Psi(\mathbf{r}\sigma, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_N\sigma_N) , \end{aligned} \tag{5.12}$$

and is a more generalized version of Eq. 5.3 Actually, its diagonal component, $\mathbf{r}' = \mathbf{r}$, is equal to the electronic spin density, namely $\rho_1(\mathbf{r}\sigma;\mathbf{r}\sigma) = \rho_\sigma(\mathbf{r})$. The one-particle reduced density matrix can be used to efficiently calculate the expectation values of $\langle\hat{T}\rangle$ and $\langle\hat{V}_{ext}\rangle$,

$$\langle\hat{T}\rangle = -\frac{1}{2}\sum_\sigma \int \mathrm{d}\mathbf{r}\nabla^2_{\mathbf{r}}\rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma)|_{\mathbf{r}'=\mathbf{r}}\,, \tag{5.13}$$

$$\langle\hat{V}_{ext}\rangle = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\hat{v}_{\mathrm{ext},\sigma}(\mathbf{r})\,. \tag{5.14}$$

The pair density, $\rho_2(\mathbf{r}',\mathbf{r})$, which is the diagonal part of two-particle reduced density matrix, is calculated by the integration of $|\Psi\rangle$ over all but two position coordinates $\mathbf{r}$ and $\mathbf{r}'$,

$$\rho_2(\mathbf{r}',\mathbf{r}) = N(N-1)\sum_{\sigma_1\ldots\sigma_N}\int \mathrm{d}\mathbf{r}_3\cdots\int \mathrm{d}\mathbf{r}_N|\Psi(\mathbf{r}'\sigma_1,\mathbf{r}\sigma_2,\ldots,\mathbf{r}_N\sigma_N)|^2\,. \tag{5.15}$$

Similarly to density, it also has a probability interpretation, that is, $\rho_2(\mathbf{r}',\mathbf{r})\mathrm{d}\mathbf{r}'\mathrm{d}\mathbf{r}$ is the joint probability of finding one electron in a volume element $\mathrm{d}\mathbf{r}'$ at position $\mathbf{r}'$, and a second electron in a volume element $\mathrm{d}\mathbf{r}$ at position $\mathbf{r}$. More importantly, $\rho_2(\mathbf{r}',\mathbf{r})$ can be used to evaluate the interelectronic repulsion energy $\langle\hat{V}_{ee}\rangle$,

$$\langle\hat{V}_{ee}\rangle = \frac{1}{2}\int \mathrm{d}\mathbf{r}\int \mathrm{d}\mathbf{r}'\frac{\rho_2(\mathbf{r}',\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|}\,. \tag{5.16}$$

Actually, Eqs. 5.13, 5.14, and 5.16 represent these energy components as *the functionals of the reduced density matrices*. A functional is a mathematical rule which maps a function (defined over a particular argument space) into a single real or complex number [72], with a definite integral being the most notable example of a functional.

5.3 Density Functional Theory

In Density Functional Theory (DFT), one makes one step further in the "simplification" of the energy expression. In DFT, the total electronic energy of the system is represented solely in terms of the density [72, 168, 169]:

$$E^{\mathrm{DFT}}[\rho] = T[\rho] + V_{ee}[\rho] + V_{ext}[\rho]\,. \tag{5.17}$$

Evaluation of such energy expression should be formally very efficient, since the density is a function of only 3 spatial variables (in contrast to the wavefunction which depends on $3N$ spatial coordinates or $\rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma)$ and $\rho_2(\mathbf{r}',\mathbf{r})$, which depend on 6 spatial coordinates).

Almost a century ago, Thomas and Fermi proposed the first approximation to the energy, which was evaluated solely using the electronic density, and it did not depend on the wavefunction or the reduced density matrix [72, 170, 171]. A key component of their model was the fully local kinetic energy functional, $T^{\mathrm{TF}}[\rho]$, based on the infinite uniform electron gas,

$$\langle\hat{T}\rangle \approx T^{\mathrm{TF}}[\rho] = \frac{3}{10}(3\pi^2)^{2/3}\int \mathrm{d}\mathbf{r}\rho^{5/3}(\mathbf{r})\,. \tag{5.18}$$

The interelectronic repulsion was treated fully classically, i.e., as an interaction between two independent electronic densities.

$$\begin{aligned}\langle\hat{V}_{ee}\rangle &= U[\rho] + U_{ncl} \approx U[\rho]\,,\\ &= \frac{1}{2}\int \mathrm{d}\mathbf{r}\int \mathrm{d}\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,,\end{aligned} \tag{5.19}$$

where U_{ncl} is the non-classical part of the interelectronic repulsion, which was omitted in the Thomas-Fermi model. Nuclear attraction energy could already be exactly calculated using only the density,

$$\langle \hat{V}_{ext} \rangle = V^{ext}[\rho] = -Z \int \mathrm{d}\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}|} \,. \tag{5.20}$$

The overall expression for the total energy of their atomic model was:

$$\begin{aligned} E \approx E^{\mathrm{TF}}[\rho] &= T^{\mathrm{TF}}[\rho] + U[\rho] + V^{ext}[\rho] \,, \\ &= \frac{3}{10}(3\pi^2)^{2/3} \int \mathrm{d}\mathbf{r} \rho^{5/3}(\mathbf{r}) - Z \int \mathrm{d}\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}|} \\ &+ \frac{1}{2} \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \,. \end{aligned} \tag{5.21}$$

Despite its performance was very limited, it was the first realization of a true *functional of density*. Another historical functional was proposed by Slater in 1951, who approximated the nonclassical part of the electron interaction, namely the exchange interaction:

$$U_{ncl} \approx E_x^S[\rho] = \frac{9}{8} \left(\frac{3}{\pi} \right)^3 \int \mathrm{d}\mathbf{r} \rho^{4/3}(\mathbf{r}) \,. \tag{5.22}$$

While originally Slater used different assumptions, an equivalent result was obtained when developing an exact exchange functional for the model of infinite uniform electron gas [72, 172]. Later, the Slater functional became a starting point in the construction of all modern exchange functionals.

In 1964, Hohenberg and Kohn built the theoretical foundations of DFT as we know them today [168]. In their seminal works, they proved two fundamental theorems. The first Hohenberg-Kohn theorem is a proof of Eq. 5.17, i.e. the existence of the functional of the density yielding the exact total energy of the ground state. The theorem states that the external potential $\hat{V}_{ext}$ is uniquely determined by the density of the system, $\rho(\mathbf{r})$, therefore it is a density functional. Hence, the total electronic Hamiltonian, $\hat{H}$, and its corresponding expectation value, the energy, are also functionals of the density.

The second Hohenberg-Kohn theorem proves the variational character of the exact density functional of the ground-state energy, that is, for any input density different from the ground-state density, the exact functional will yield an energy higher than the true ground-state energy E_0:

$$E_0 = E[\rho_0] < E[\rho] \quad \text{if} \quad \rho(\mathbf{r}) \neq \rho_0(\mathbf{r}) \,. \tag{5.23}$$

Within the proof, they define the total energy with the use of a universal functional $F_{\mathrm{HK}}[\rho]$ (named as the Hohenberg-Kohn functional), which is system independent:

$$E[\rho_0] = F_{\mathrm{HK}}[\rho] + V^{ext}[\rho] \,, \tag{5.24}$$

where

$$F_{\mathrm{HK}}[\rho] = T[\rho] + V_{ee}[\rho] = \min_{\Psi \to \rho_0} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle \,, \tag{5.25}$$

In the above definition, $|\Psi\rangle$ must be the wavefunction that gives the exact density $\rho_0(\mathbf{r})$ (Eqs. 5.3 and 5.4) and yields the lowest expectation value.

Although being an exact theory, the biggest obstacle of DFT is the description of highly non-local quantities, namely the kinetic energy and interelectronic repulsion, in terms of a very local variable, the electronic density. An accurate approximation of the kinetic energy functional turned out to be the most challenging task, giving large errors even in the simplest atomic systems.

5.4 Kohn-Sham DFT

Within the field of quantum chemistry, modern applications of DFT almost always rely on the Kohn-Sham version of DFT (KS-DFT) [173]. Kohn and Sham proposed to circumvent the problem of kinetic energy by evaluating it for the set of one-particle spinorbitals obtained from a fictitious noninteracting model. This model was very special - it has some resemblance with a real (molecular) system, i.e., the same number of electrons and nuclei at the same positions, but with the interelectronic interaction completely turned off, while still preserving the exact density of the true (interacting) system. While electrons were not explicitly feeling the interaction of each other, they are were placed into a new unique one-particle external potential $\hat{v}_{S,\sigma}(\mathbf{r})$, which exactly reproduces the same (interacting) density.

Such noninteracting quantum system is described with the Hamiltonian $\hat{H}_{KS}$, expressed as a sum of single-particle operators $\hat{f}_{KS,i,\sigma}$:

$$\begin{aligned}\hat{H}_{KS} &= -\frac{1}{2}\sum_{i}^{N_{el}}\nabla_i^2 + \sum_{i}^{N_{el}}\hat{v}_{S,\sigma}(\mathbf{r}_i)\,,\\ &= \sum_{i}^{N_{el}}\hat{f}_{i,\sigma}^{KS}\,.\end{aligned} \tag{5.26}$$

The latter does not contain any operator explicitly accounting for the interelectron interaction. The wavefunction $|\Phi\rangle$ is represented by a single Slater determinant formed from a set of N Kohn-Sham spinorbitals $\{\phi_{i,\sigma}\}$ [174]:

$$\begin{aligned}|\Phi\rangle &= |\phi_{1,\alpha}(\mathbf{r}_1)\phi_{1,\beta}(\mathbf{r}_2)\cdots\phi_{N/2,\alpha}(\mathbf{r}_{N-1})\phi_{N/2,\beta}(\mathbf{r}_N)\rangle\,,\\ &= \frac{1}{\sqrt{N!}}\sum_{n}^{N!}(-1)^{p_n}\mathscr{P}_n\{\phi_{1,\alpha}(\mathbf{r}_1)\phi_{1,\beta}(\mathbf{r}_2)\cdots\phi_{N/2,\alpha}(\mathbf{r}_{N-1})\phi_{N/2,\beta}(\mathbf{r}_N)\}\,,\end{aligned} \tag{5.27}$$

where $\mathscr{P}_n$ is an operator that generates the n-th permutation of the electron labels 1, 2, ..., N and p_n is the number of transpositions required to obtain such permutation [174]. Following the general rules of quantum mechanics, the density of a system of N independent particles can be expressed as a sum of the squares of spinorbitals' amplitudes. Therefore, the density of the Kohn-Sham system, $\rho_S(\mathbf{r})$, is evaluated as

$$\rho_S(\mathbf{r}) = \sum_{\sigma}\sum_{i}^{N/2}|\phi_{i,\sigma}(\mathbf{r})|^2 = \rho(\mathbf{r})\,, \tag{5.28}$$

and, as required by the model, $\rho_S(\mathbf{r})$ is equal to the exact density $\rho(\mathbf{r})$ of the interacting system.

The total electronic energy in KS-DFT is defined as the sum of the following components,

$$E^{\mathrm{KS-DFT}} = T_s[\{\phi_s\}] + V_{ext}[\rho] + U[\rho] + E_{xc}[\rho]\,. \tag{5.29}$$

The first term, the kinetic energy evaluated within the KS-DFT scheme, $T_s[\{\phi_s\}]$, is obtained from the expectation value of $\hat{T}$ for the KS determinant Φ:

$$\begin{aligned}T_s[\{\phi_s\}] &= \langle\Phi|\hat{T}|\Phi\rangle\,,\\ &= -\frac{1}{2}\sum_{\sigma}\sum_{i}^{N/2}\int \mathrm{d}\mathbf{r}\phi_{i,\sigma}^{*}(\mathbf{r})\nabla^2\phi_{i,\sigma}(\mathbf{r})\,,\end{aligned} \tag{5.30}$$

and is evaluated for a given set of Kohn-Sham orbitals. The second term in Eq. 5.29 is the potential energy, $V_{ext}[\rho]$, corresponding to the interaction of electrons with the external potential. As mentioned in Eq. 5.14, $V_{ext}[\rho]$ is an explicit functional of the density and, in the case of molecular systems (without the presence of any external perturbations), is defined as the integral:

$$V_{ext}[\rho] = \int \mathrm{d}\mathbf{r} \sum_{I}^{N_{nucl}} \frac{-Z_I}{|\mathbf{r}-\mathbf{R}_J|}\rho(\mathbf{r})\,. \tag{5.31}$$

The third term in Eq. 5.29 is the classical part of the interelectronic repulsion, namely the Coulomb repulsion energy $U[\rho]$:

$$U[\rho] = \frac{1}{2}\int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,. \tag{5.32}$$

Lastly, $E_{xc}[\rho]$ is the exchange-correlation energy, the main unknown in KS-DFT, which incorporates the effects of all approximations made by adapting the Kohn-Sham non-interacting model. $E_{xc}[\rho]$ is defined through two separate components, the exchange energy, $E_x[\rho]$, and the correlation energy, $E_c[\rho]$,

$$E_{xc}[\rho] = E_x[\rho] + E_c[\rho]\,. \tag{5.33}$$

The exchange energy incorporates the Pauli exclusion effect that affects the electrons of the same spin, and is simply defined as the difference between the two-electron energy calculated for the Kohn-Sham wavefunction $|\Phi\rangle$ and the classical Coulomb repulsion:

$$E_x[\rho] = \langle\Phi|\hat{V}_{ee}|\Phi\rangle - U[\rho]\,. \tag{5.34}$$

Correlation energy constitutes the remaining energetic effects arising from the approximating the real interacting system with the Kohn-Sham model, including the incompleteness of $|\Phi\rangle$ for the description of the true ground-state wavefunction $|\Psi\rangle$:

$$\begin{aligned} E_c[\rho] &= \langle\Psi|\hat{T}+\hat{V}_{ee}|\Psi\rangle - \langle\Phi|\hat{T}+\hat{V}_{ee}|\Phi\rangle\,, \\ &= (T[\rho]-T_s[\rho]) + (V_{ee}[\rho]-U[\rho]-E_x[\rho])\,, \\ &= T_c[\rho]+U_c[\rho]\,. \end{aligned} \tag{5.35}$$

As shown above, one can further separate E_c into the kinetic part, $T_c[\rho]$, accounting for the lack of correlation in kinetic energy, and potential part, $U_c[\rho]$, accounting for the lack of correlation in interelectronic repulsion. Both $E_x[\rho]$ and $E_c[\rho]$ are always non-positive numbers. The Kohn-Sham approach alleviates the problem of the nonlocal character of kinetic energy by using an exact quantum mechanical representation in T_s (although for an approximate wavefunction). T_s approximates most of the exact $T[\rho]$ and correctly mimics the atomic shell structure. Then, the small fraction of the energy which is put into E_{xc} is somewhat better approximated by local and semi-local approximations. However, as a side effect, in KS-DFT one actually works with orbitals and not only the density, which increases the cost of the actual calculation.

Kohn-Sham spinorbitals are obtained by solving the set of N independent one-particle equations, which are defined for the one-particle Kohn-Sham operator $\hat{f}_\sigma^{KS}$,

$$\begin{aligned} \hat{f}_\sigma^{KS}\phi_{i,\sigma}(\mathbf{r}) &= \varepsilon_{i,\sigma}\phi_{i,\sigma}(\mathbf{r})\,, \\ \left\{-\frac{1}{2}\nabla^2 + \hat{v}_{S,\sigma}(\mathbf{r})\right\}\phi_{i,\sigma}(\mathbf{r}) &= \varepsilon_{i,\sigma}\phi_{i,\sigma}(\mathbf{r})\,. \end{aligned} \tag{5.36}$$

The above set of equations are equivalent to the ones obtained from the Hartree-Fock method, the difference lying in the one-electron potential $\hat{v}_{S,\sigma}(\mathbf{r})$. In the Kohn-Sham method, $\hat{v}_{S,\sigma}(\mathbf{r})$ is defined as the sum of the external potential, $\hat{v}_{\text{ext},\sigma}(\mathbf{r})$ (nuclear attraction), the classical Hartree (Coulomb) potential $\hat{v}_{\text{H}}(\mathbf{r})$, and, the exchange-correlation potential $\hat{v}_{\text{xc},\sigma}(\mathbf{r})$:

$$\hat{v}_{S,\sigma}(\mathbf{r}) = \hat{v}_{\text{ext},\sigma}(\mathbf{r}) + \hat{v}_{\text{H}}(\mathbf{r}) + \hat{v}_{\text{xc},\sigma}(\mathbf{r})\,. \tag{5.37}$$

Those potentials are defined as the functional derivatives of corresponding potential energies:

$$\hat{v}_{\text{H}}(\mathbf{r}) = \frac{\delta U[\rho]}{\delta \rho} = \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}\,, \tag{5.38}$$

$$\hat{v}_{\text{xc},\sigma}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho_\sigma}\,, \tag{5.39}$$

$$= \frac{\delta E_x[\rho]}{\delta \rho_\sigma} + \frac{\delta E_c[\rho]}{\delta \rho_\sigma}\,, \tag{5.40}$$

$$= \hat{v}_{\text{x},\sigma}(\mathbf{r}) + \hat{v}_{\text{c},\sigma}(\mathbf{r})\,, \tag{5.41}$$

and $\hat{v}_{\text{ext},\sigma}(\mathbf{r})$ is defined according to Eq. 5.9.

5.4.1 KS-DFT Equations in a Finite Basis

In quantum chemistry, the Kohn-Sham equations are solved iteratively using a finite atomic-centered basis. This was formally established in the work of Pople, Gill, and Johnson and was the DFT equivalent of the Roothaan-Hall (for closed-shell systems) and Pople-Nesbet (for open-shell systems) equations [175]. This section summarizes the most important theoretical aspects of this implementation. Unless stated otherwise, the Latin letters (i, j, k, l) label the molecular spinorbitals, the Greek letters (μ, ν, λ, δ) label the atomic orbitals, whereas α and β denote the spin component.

In most quantum chemical computations, molecular spinorbitals are expressed as a linear combination of atomic orbitals (LCAO),

$$\phi_{i,\sigma}(\mathbf{r}) = \sum_{\mu}^{M} c_{\mu i}^{\sigma} \chi_\mu(\mathbf{r})\,, \tag{5.42}$$

with $\chi_\mu(\mathbf{r})$ being an atomic-centered basis function, named also Atomic Orbitals (AOs), and $c_{\mu i}^{\sigma}$ being the corresponding orbital coefficient. The most common AOs are the Gaussian type orbitals (GTOs). The spin density and its gradient are expressed in terms of AOs as

$$\rho_\sigma(\mathbf{r}) = \sum_{\mu,\nu}^{M} P_{\mu\nu}^{\sigma} \chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r})\,, \tag{5.43}$$

$$\nabla\rho_\sigma(\mathbf{r}) = \sum_{\mu,\nu}^{M} P_{\mu\nu}^{\sigma} \nabla(\chi_\mu(\mathbf{r}) \chi_\nu(\mathbf{r}))\,. \tag{5.44}$$

$P_{\mu\nu}^{\sigma}$ is the spin one-particle density matrix and when the wavefunction is described with a single Slater determinant, it is expressed in the atomic basis as:

$$P_{\mu\nu}^{\sigma} = \sum_{i}^{N/2} (c_{\mu i}^{\sigma})^* c_{i\nu}^{\sigma}\,. \tag{5.45}$$

Like in the case of the total electronic density, one defines the total density matrix as the sum of two spin components:

$$P_{\mu\nu} = P^{\alpha}_{\mu\nu} + P^{\beta}_{\mu\nu}\,. \tag{5.46}$$

Most of the functional approximations, as will be showed later in the following sections, have semi-local form, depending solely on the spin densities and its gradient. One can express Eq. 5.33 as the integral of a general integrand of density and its gradient, $f(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta})$,

$$E_{xc}[\rho] = \int \mathrm{d}\mathbf{r} f(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta})\,, \tag{5.47}$$

where

$$\gamma_{\alpha\alpha}(\mathbf{r}) = |\boldsymbol{\nabla}\rho_\alpha(\mathbf{r})|^2\,, \tag{5.48}$$

$$\gamma_{\beta\beta}(\mathbf{r}) = |\boldsymbol{\nabla}\rho_\beta(\mathbf{r})|^2\,, \tag{5.49}$$

$$\gamma_{\alpha\beta}(\mathbf{r}) = \boldsymbol{\nabla}\rho_\alpha(\mathbf{r}) \cdot \boldsymbol{\nabla}\rho_\beta(\mathbf{r})\,. \tag{5.50}$$

For $E_{xc}[\rho]$ in Eq. 5.47, one can derive the expressions for the exchange-correlation potentials:

$$\hat{v}_{\mathrm{xc},\alpha}(\mathbf{r}) = \frac{\partial f}{\partial \rho_\alpha} - 2\boldsymbol{\nabla}\cdot\left(\frac{\partial f}{\partial \gamma_{\alpha\alpha}}\boldsymbol{\nabla}\rho_\alpha(\mathbf{r})\right) - \boldsymbol{\nabla}\cdot\left(\frac{\partial f}{\partial \gamma_{\alpha\beta}}\boldsymbol{\nabla}\rho_\beta(\mathbf{r})\right)\,, \tag{5.51}$$

$$\hat{v}_{\mathrm{xc},\beta}(\mathbf{r}) = \frac{\partial f}{\partial \rho_\beta} - 2\boldsymbol{\nabla}\cdot\left(\frac{\partial f}{\partial \gamma_{\beta\beta}}\boldsymbol{\nabla}\rho_\beta(\mathbf{r})\right) - \boldsymbol{\nabla}\cdot\left(\frac{\partial f}{\partial \gamma_{\alpha\beta}}\boldsymbol{\nabla}\rho_\alpha(\mathbf{r})\right)\,. \tag{5.52}$$

One can naturally expand the above formula for $E_{xc}[\rho]$ in terms other density-related quantities, such as its Laplacian, or terms based on the kinetic-energy density.

In the finite basis, the one-particle Kohn-Sham eigenequations (Eq. 5.36) are transformed into a set of secular equations that must be solved to obtain the orbital coefficients $c^{\sigma}_{\nu i}$,

$$\sum_{\nu}^{M}\left(F^{\sigma}_{\mu\nu} - \varepsilon^{\sigma}_i S_{\mu\nu}\right)c^{\sigma}_{\nu i} = 0\,, \tag{5.53}$$

which involve the matrices expressed in the AO basis. Namely, these are the overlap $S_{\mu\nu}$ matrix,

$$S_{\mu\nu} = \int \mathrm{d}\mathbf{r}\chi^*_\mu(\mathbf{r})\chi_\nu(\mathbf{r})\,. \tag{5.54}$$

and the Fock matrix, $F^{\sigma}_{\mu\nu}$,

$$\begin{aligned}
F^{\sigma}_{\mu\nu} &= \int \mathrm{d}\mathbf{r}\chi^*_\mu(\mathbf{r})\left\{-\frac{1}{2}\nabla^2 + \hat{v}_{S,\sigma}(\mathbf{r})\right\}\chi_\nu(\mathbf{r})\,,\\
&= \int \mathrm{d}\mathbf{r}\chi^*_\mu(\mathbf{r})\left\{-\frac{1}{2}\nabla^2 + \hat{v}_{\mathrm{ext},\sigma}(\mathbf{r}) + \hat{v}_{\mathrm{H}}(\mathbf{r}) + \hat{v}_{\mathrm{xc},\sigma}(\mathbf{r})\right\}\chi_\nu(\mathbf{r})\,,\\
&= \int \mathrm{d}\mathbf{r}\chi^*_\mu(\mathbf{r})\left\{-\frac{1}{2}\nabla^2 + \hat{v}_{\mathrm{ext},\sigma}(\mathbf{r})\right\}\chi_\nu(\mathbf{r})\\
&+ \int \mathrm{d}\mathbf{r}\chi^*_\mu(\mathbf{r})\hat{v}_{\mathrm{H}}(\mathbf{r})\chi_\nu(\mathbf{r})\\
&+ \int \mathrm{d}\mathbf{r}\chi^*_\mu(\mathbf{r})\hat{v}_{\mathrm{xc},\sigma}(\mathbf{r})\chi_\nu(\mathbf{r})\,, \qquad (5.55)\\
&= H^{\mathrm{core}}_{\mu\nu} + J_{\mu\nu} + F^{xc,\sigma}_{\mu\nu}\,. \qquad (5.56)
\end{aligned}$$

In KS-DFT, the Fock matrix can be split into three components. The first two are also common for the Hartree-Fock method: the bare nucleus Hamiltonian matrix, $H_{\mu\nu}^{\text{core}}$, and the Coulomb matrix, $J_{\mu\nu}$,

$$\begin{aligned} H_{\mu\nu}^{\text{core}} &= \int \mathrm{d}\mathbf{r}\chi_\mu^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2+\hat{v}_{ext}(\mathbf{r})\right)\chi_\nu(\mathbf{r}), \\ &= T_{\mu\nu}+V_{\mu\nu}^{ext}, \end{aligned} \tag{5.57}$$

$$J_{\mu\nu} = \sum_{\lambda,\delta}^{M} P_{\lambda\delta}(\mu\nu|\lambda\delta), \tag{5.58}$$

where $(\mu\nu|\lambda\delta)$ is a two-electron integral expressed in a quantum mechanical notation

$$(\mu\nu|\lambda\delta) = \int \mathrm{d}\mathbf{r}\int \mathrm{d}\mathbf{r}'\chi_\mu^*(\mathbf{r})\chi_\nu(\mathbf{r})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\chi_\lambda^*(\mathbf{r}')\chi_\delta(\mathbf{r}'). \tag{5.59}$$

The third term in Eq. 5.56 is the exchange-correlation part of the Fock matrix (whereas in Hartree-Fock, there is solely the exact exchange instead). By inserting Eqs. 5.51 and 5.52 into the integral of Eq. 5.55, one obtains the following expressions for the α and β parts of $F_{\mu\nu}^{xc}$:

$$\begin{aligned} F_{\mu\nu}^{xc,\alpha} = \int \mathrm{d}\mathbf{r}\Bigg[&\frac{\partial f}{\partial\rho_\alpha}\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}) \\ &+\left(2\frac{\partial f}{\partial\gamma_{\alpha\alpha}}\nabla\rho_\alpha(\mathbf{r})+\frac{\partial f}{\partial\gamma_{\alpha\beta}}\nabla\rho_\beta(\mathbf{r})\right)\cdot\nabla(\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}))\Bigg], \end{aligned} \tag{5.60}$$

$$\begin{aligned} F_{\mu\nu}^{xc,\beta} = \int \mathrm{d}\mathbf{r}\Bigg[&\frac{\partial f}{\partial\rho_\beta}\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}) \\ &+\left(2\frac{\partial f}{\partial\gamma_{\beta\beta}}\nabla\rho_\beta(\mathbf{r})+\frac{\partial f}{\partial\gamma_{\alpha\beta}}\nabla\rho_\alpha(\mathbf{r})\right)\cdot\nabla(\chi_\mu(\mathbf{r})\chi_\nu(\mathbf{r}))\Bigg]. \end{aligned} \tag{5.61}$$

After solving Eq. 5.53 in an iterative manner, one obtains the set of LCAO coefficients $c_{\nu i}^{\sigma}$ defining the Kohn-Sham spinorbitals. As a result, one expresses the total energy of Eq. 5.29 in a finite AO basis:

$$E^{\text{KS}-\text{DFT}} = \sum_{\mu,\nu}^{M} P_{\mu\nu}H_{\mu\nu}^{\text{core}} + \frac{1}{2}\sum_{\mu,\nu,\lambda,\delta}^{M} P_{\mu\nu}P_{\lambda\delta}(\mu\nu|\lambda\delta) + E_{xc}[\rho], \tag{5.62}$$

where $E_{xc}[\rho]$ is defined by Eq. 5.47 and in almost all cases is calculated using numerical integration.

5.5 Jacob's Ladder of Density Functional Approximations

In principle, DFT is as an exact theory, as the Schrödinger equation is. However, its exactness requires the exact density functional, which is still unknown (and may remain as such forever). Besides hints on the properties of exact functional, in a form of strict theoretical constraints, no general prescription of how to approximate it exists. However, over the decades, some relatively successful Density Functional Approximations (DFAs) have been proposed.

Nowadays, it is common to discuss the increasing complexity and performance of DFAs using the so-called Jacob's ladder of DFT. This simple hierarchy, introduced by

John Perdew and Karla Schmidt, classifies DFAs onto different rungs (i.e., tiers) based on the type of density-based components they utilize [176, 177]. Going up in the ladder, each rung incrementally adds more density components in the functional, while also increasing their computational cost. In general, DFAs from higher rungs tend to show better performance in reproducing experimental or high-level ab-initio data. Below, a very brief description of each rung of the DFT Jacob's ladder is given, along with some examples of the most popular approximations of each type.

The 1st Rung - Local Spin Density Approximations (LSDAs)

The lowest rung of the ladder belongs to the LSDAs, includes local functionals, depending solely on the values of the electronic spin densities, $\rho_\alpha(\mathbf{r})$ and $\rho_\beta(\mathbf{r})$, and are represented as

$$E_{xc}^{\mathrm{LSDA}}[\rho_\alpha,\rho_\beta] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\varepsilon_{xc}^{\mathrm{unif}}(\rho_\alpha,\rho_\beta)\,, \tag{5.63}$$

where $\varepsilon_{xc}^{\mathrm{unif}}(\rho_\alpha,\rho_\beta)$ is the exchange-correlation energy per particle of the electron gas of density ρ, and in the case of LSDAs, it is derived from the infinite uniform electron gas model. The LSDA exchange part can be derived exactly from first principles and is simply the Slater-Dirac functional:

$$E_x^{\mathrm{LSDA}}[\rho_\alpha,\rho_\beta] = \sum_\sigma \int \mathrm{d}\mathbf{r}\rho_\sigma(\mathbf{r})\varepsilon_{x,\sigma}^{\mathrm{Slater}}\,, \tag{5.64}$$

$$\varepsilon_{x,\sigma}^{\mathrm{Slater}} = -\frac{3}{2}\left(\frac{3}{4\pi}\right)^{1/3}\rho_\sigma^{1/3}(\mathbf{r})\,. \tag{5.65}$$

On the contrary, no explicit exact expression for the LSDA correlation has been derived. The most widely used LSDA approximations are the ones of Vosko, Wilk, and Nuisar (VWN, several versions available) and the one of Perdew and Wang (PW92). Those have very complicated formula and were derived as interpolations of highly accurate quantum Monte-Carlo calculations. Despite its very local character, LSDA works very well for metallic solid-state systems because of the very delocalized and slowly varying electronic density of such systems. In contrast, in molecular systems where the density changes are very rapid and of high amplitudes, LSDA performs badly and is not reliable even for a qualitative analysis.

The 2nd Rung - Generalized Gradient Approximations (GGAs)

The second rung consists of functionals that also include the gradients of the spin density, $\nabla\rho_\alpha(\mathbf{r})$ and $\nabla\rho_\beta(\mathbf{r})$, namely

$$E_{xc}^{\mathrm{GGA}}[\rho_\alpha,\rho_\beta] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\varepsilon_{xc}^{\mathrm{GGA}}(\rho_\alpha,\rho_\beta,\nabla\rho_\alpha,\nabla\rho_\beta)\,. \tag{5.66}$$

In many cases, dependency on $\nabla\rho(\mathbf{r})$ is introduced through the reduced spin density gradient, $s_\sigma(\mathbf{r})$, which is defined as

$$s_\sigma(\mathbf{r}) = \frac{|\nabla\rho_\sigma(\mathbf{r})|}{\rho_\sigma^{4/3}(\mathbf{r})}\,, \tag{5.67}$$

and is a convenient parameter of density inhomogenicity. Large values of $s_\sigma(\mathbf{r})$ correspond not only to regions of large gradients but also to regions of very small densities, such as exponential tails far from the nuclei. Conversely, small values of $s_\sigma(\mathbf{r})$ correspond

to regions with slowly varying densities or regions with very large densities, such as in the closest vicinity of nuclei. In principle, there are many separate exchange and correlation GGAs that one can mix together; therefore, it is better to introduce them separately. In quantum chemistry, the two most popular exchange functionals are B88 [178] and PBE [179]. The former one, developed by Becke in 1988 and fitted to reproduce the exact exchange of rare gas atoms, is defined as the extension of the Slater exchange [178]:

$$E_x^{\mathrm{B88}}[\rho_\alpha,\rho_\beta] = \sum_\sigma \int \mathrm{d}\mathbf{r}\rho_\sigma(\mathbf{r})\varepsilon_{x,\sigma}^{\mathrm{B88}}\,, \tag{5.68}$$

$$\varepsilon_{x,\sigma}^{\mathrm{B88}} = \varepsilon_{x,\sigma}^{\mathrm{Slater}} - \rho_\sigma^{1/3}(\mathbf{r})F_{x,\sigma}^{\mathrm{B88}}\,. \tag{5.69}$$

In the above equation, $F_{x,\sigma}^{\mathrm{B88}}$ is the enhancement factor,

$$F_{x,\sigma}^{\mathrm{B88}} = \beta\frac{s_\sigma(\mathbf{r})^2}{1+6\beta s_\sigma(\mathbf{r})\sinh^{-1}s_\sigma(\mathbf{r})}\,, \tag{5.70}$$

with the only fitted parameter of $\beta = 0.0042$. $F_{x,\sigma}^{\mathrm{B88}}$ vanishes for ideally homogeneous spin densities, as expected.

Regarding the GGA correlation functionals, two of the most widely used are LYP [180] and PBE [181].

The 3rd Rung - meta-Generalized Gradient Approximations (meta-GGAs)

The third rung belongs to the functionals depending on the kinetic energy density, τ_σ, and less frequently on the Laplacian of the density $\nabla^2\rho(\mathbf{r})$:

$$E_{xc}^{\mathrm{meta-GGA}}[\rho_\alpha,\rho_\beta] = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\varepsilon_{xc}^{\mathrm{meta-GGA}}(\rho_\alpha,\rho_\beta,\nabla\rho_\alpha,\nabla\rho_\beta,\tau_\alpha,\tau_\beta)\,, \tag{5.71}$$

and the kinetic energy density calculated within KS-DFT as

$$\tau_\sigma = \sum_i |\nabla\phi_{i,\sigma}|^2\,, \tag{5.72}$$

with $\phi_{i,\sigma}$ being the i-th Kohn-Sham spinorbital. Due to the dependence on τ_σ in meta-GGA, one can distinguish the regions of space in which one orbital dominates the electron spin density [177]. Almost all meta-GGAs have very complicated, layered expressions. The most popular meta-GGAs used in chemistry are TPSS [182], SCAN [183], M06-L [184] and MN15-L [76].

The 4th Rung - Global and Range-Separated Hybrids (GHs and RSHs)

The fourth rung belongs to hybrid functionals, in which one introduces a fully non-local exact exchange, E_x^{exact}, i.e., the Hartree-Fock exchange evaluated for the occupied Kohn-Sham orbitals:

$$E_x^{\mathrm{exact}} = -\frac{1}{2}\sum_\sigma\sum_{i,j}^{\mathrm{occ}}\int \mathrm{d}\mathbf{r}\int \mathrm{d}\mathbf{r}'\frac{\phi_{i,\sigma}^*(\mathbf{r})\phi_{j,\sigma}^*(\mathbf{r}')\phi_{i,\sigma}(\mathbf{r})\phi_{j,\sigma}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,. \tag{5.73}$$

Such types of hybrid functionals are used only within the KS-DFT scheme, which explicitly uses one-particle orbitals. While in general one has to construct functionals in which exchange and correlation parts have similar level of locality, it is formally possible to introduce the exact exchange through the adiabatic connection [169, 185–188].

Two of the most common types of hybrid DFAs are Global Hybrids (GHs) and Range-Separated Hybrids (RSHs). In the former ones, the mixing of the exact and (meta-)GGA exchange is the same for all points in space, i.e., it is set globally:

$$E_{xc}^{\mathrm{GH}} = E_{xc}^{(\mathrm{meta-})\mathrm{GGA}} + a_x(E_x^{\mathrm{exact}} - E_x^{(\mathrm{meta-})\mathrm{GGA}}), \tag{5.74}$$

where a_x is one of the possible admixture coefficients, which increases the exact exchange, while decreasing the (meta-)GGA exchange.

One of the other possibilities to incorporate the exact exchange, which is adapted in RSHs, is to vary the amount of exact exchange depending on the interelectronic distance. Namely, exact exchange is introduced only at long interelectronic distances (LR), whereas at short interelectronic ranges (SR) mostly the (meta-)GGA exchange prevails. In general, such partition can be expressed as

$$E_{xc,\omega}^{\mathrm{RSH}} = E_{x,\omega}^{(\mathrm{meta-})\mathrm{GGA,SR}} + E_{x,\omega}^{\mathrm{exact,LR}} + E_c^{(\mathrm{meta-})\mathrm{GGA}}. \tag{5.75}$$

Such inclusion of the exact exchange at long distances corrects the exact asymptotic behavior of the exchange interaction, which should asymptotically decay as $\lim_{r_{12}\to\infty} E_x = 1/r_{12}$ [169, 189]. The asymptotic limit can be achieved through the Ewald splitting of the two-electron repulsion operator, which in a more general way is defined as [189]

$$\frac{1}{r_{12}} = \frac{1-[\alpha+\beta\mathrm{erf}(\omega r_{12})]}{r_{12}} + \frac{\alpha+\beta\mathrm{erf}(\omega r_{12})}{r_{12}}. \tag{5.76}$$

The first term of the r.h.s. of Eq. 5.76 corresponds to the SR part, whereas the second corresponds to the LR part; erf is the standard error function, α governs the minimal amount of exact exchange at both SR and LR, β controls the additional amount of exact exchange at different ranges. The most important parameter defining the RSH performance is ω, the range separation parameter, which determines the transition from SR to LR. Larger values of ω result in a quicker transition from the DFA to the exact exchange, that is, the transition already occurs at the shorter interelectronic distances.

From Eq. 5.76 one can obtain the following equations for the SR and LR exchange terms:

$$E_{x,\omega}^{(\mathrm{meta-})\mathrm{GGA,SR}} = (1-\alpha)E_x^{(\mathrm{meta-})\mathrm{GGA}} - \beta E_{x,\omega}^{(\mathrm{meta-})\mathrm{GGA}}, \tag{5.77}$$

$$E_{x,\omega}^{\mathrm{exact,LR}} = \alpha E_x^{\mathrm{exact}} + \beta E_{x,\omega}^{\mathrm{exact}}. \tag{5.78}$$

While the expression for $E_{x,\omega}^{\mathrm{exact}}$ is obtained by changing $1/r_{12}$ in Eq. 5.73 by the second term of Eq. 5.76, the short range version of the DFA exchange $E_{x,\omega}^{(\mathrm{meta-})\mathrm{GGA}}$ must be derived from scratch.

Some RSHs utilize only the short-range part of the DFA exchange at SR, whereas at LR only the exact exchange prevails. This corresponds to setting $\alpha = 0$ and $\beta = 1$ in Eqs. 5.77 and 5.78. Most notable examples of such RSHs are LC-BLYP [190], LC-ωPBE [191], and ωB97X [192], with the parameter ω set to 0.47, 0.30, and 0.30 Bohr^{-1}, respectively. The CAM-B3LYP functional includes the GGA exchange at all distances, by using α = 0.19, β = 0.46 and ω = 0.33 Bohr^{-1} [193]. RSHs improve the accuracy of the excitation energies to Rydberg and charge transfer states, when compared to GHs and other types of DFAs [194–199].

Chemical performance of RSHs can be improved by a system-specific tuning of the range-separation parameter ω, giving rise to RSHs commonly known as Optimally Tuned RSHs (OT-RSHs) [92, 100, 102, 107, 200–206]. Several ω-tuning techniques have been

proposed, adopting different constraints and targeting different molecular properties. Most often, the range separation parameter is adjusted to impose Janak's theorem, which states that for the exact Kohn-Sham exchange correlation potential, the energy of the highest occupied molecular orbital (HOMO) should be equal to the negative value of the ionization potential [207]. A popular procedure for the optimization of range separation parameter requires to minimize the following function,

$$J(\omega) = \sqrt{\left[\varepsilon^{N}_{\mathrm{HOMO}}(\omega) + \mathrm{IP}^{N}(\omega)\right]^2 + \left[\varepsilon^{N+1}_{\mathrm{HOMO}}(\omega) + \mathrm{IP}^{N+1}(\omega)\right]^2}, \tag{5.79}$$

where for the system with N electrons, $\varepsilon^{N}_{\mathrm{HOMO}}$ is the energy of HOMO orbital and IP^{N} is the ionization potential.

The 5$^{\mathrm{th}}$ Rung - Double Hybrids (DHs)

The last rung of Jacob's ladder, the fifth one, belongs to the functionals which additionally utilize the information on the virtual Kohn-Sham orbitals. This is done to more precisely account for nonlocal correlation effects, which are omitted in most DFAs. Inclusion of the nonlocal correlation can be done in many ways, but one of the most popular and conceptually simplest methods is the so-called Double Hybrids (DHs). In DHs, in addition to mixing of the (meta-)GGA and exact exchanges (as in GHs), one combines the higher-level correlation with the correlation part of DFAs:

$$E^{\mathrm{DH}}_{xc} = E^{(\mathrm{meta-})\mathrm{GGA}}_{xc} + a_x(E^{\mathrm{exact}}_{x} - E^{(\mathrm{meta-})\mathrm{GGA}}_{x}) + a_c(E^{\mathrm{high-level}}_{c} - E^{(\mathrm{meta-})\mathrm{GGA}}_{c}). \tag{5.80}$$

In the most popular DHs, $E^{\mathrm{high-level}}_{c}$ is evaluated by the second-order perturbation theory in the framework of DFA, E^{PT2}_{c}, by the means of the Görling-Levy theory [208, 209]. Shortly, E^{PT2}_{c} is evaluated in similar manner to the second-order Moller-Plesset (MP2) corrections for the Hartree-Fock energy [174, 210], but employing the Kohn-Sham orbitals. The most important feature of DHs is their improved description of long-range effects in non-covalent interactions. Unfortunately, the usage of DFAs from this rung is very costly (formal cost equivalent to MP2), which prohibits using them for large chemical systems. Most notable examples of modern DHs are B2PLYP [211], mPW2PLYP [212], PBE-QIDH [213], and PBE0-DH [214].

5.6 Numerical Integration in KS-DFT

This section introduces the key elements of the numerical integration performed when solving the KS-DFT equations. Although by many considered as just a technical issue, numerical integration is a fundamental approximation that determines the accuracy of modern DFAs. Besides constructing a physically accurate exchange-correlation functional, one has to solve two key types of integrals in KS-DFT. These are the integrals for the exchange-correlation energy $E_{xc}[\rho(\mathbf{r})]$ (Eq. 5.47) and the elements of the Fock matrix corresponding to the exchange-correlation potential (Eqs.5.60 and 5.61). Exact, i.e., fully analytical evaluation of those integrals is possible only for the simplest DFAs, such as Slater exchange, and mostly for model systems or plain atoms.

5.6.1 Multicenter Integration. Becke Partitioning Scheme

The key concept of solving such complicated integrals defined over whole molecular space (i.e. solving molecular integrals) is to first partition them into many integrals over

smaller subspaces. Multicenter partition is an exact decomposition and does not introduce any approximation. If the I_A integrals were solved analytically, the total I integral would be exact. By referring to the nature of atoms in molecules, in his fundamental work, Becke proposed to partition the molecular integrals simply into integrals defined over the atomic domains [215]. The goal was to express the integral I of a general integrand $F(\mathbf{r})$ as a sum of atomic contributions,

$$I = \int \mathrm{d}\mathbf{r} F(\mathbf{r}) = \sum_A I_A \,, \tag{5.81}$$

where A denotes the A-th atom and the integral I_A is simply defined as

$$I_A = \int \mathrm{d}\mathbf{r} F_A(\mathbf{r}) \,. \tag{5.82}$$

While still defined over the whole Cartesian space, the origin of integration in I_A is chosen to be exactly at the position of the A-th nucleus. The partitioning itself is achieved through the usage of atomic weights w_A, which for each point in space, $\mathbf{r}$, sum to one:

$$\sum_A \omega_A(\mathbf{r}) = 1 \,. \tag{5.83}$$

Then, one finally partitions the integrand $F(\mathbf{r})$ as

$$F(\mathbf{r}) = \sum_A \omega_A(\mathbf{r}) F(\mathbf{r}) = \sum_A F_A(\mathbf{r}) \,, \tag{5.84}$$

which leads to Eqs. 5.81 and 5.82. In Becke's partitioning scheme, the weights are defined as

$$\omega_A(\mathbf{r}) = \frac{P_A(\mathbf{r})}{\sum_B P_B(\mathbf{r})} \,, \quad \text{where} \quad P_A = \prod_{B \neq A} s_B^k(\mu_{AB}) \,, \tag{5.85}$$

with s_B^k designed to be (smooth) step-like functions. For a given atom A, its weight should decay smoothly from the value $\omega_A(R_A) \approx 1$ at the position of the A-th nucleus, R_A, to zero. Moreover, at R_A all the weights of the other atoms are (nearly) zero. Becke proposed to define atomic cells in a strictly geometrical manner using fuzzy Voronoi polyhedra (also called fuzzy Voronoi cells) [215]. For that purpose, Becke utilized the elliptical coordinate μ_{AB},

$$\mu_{AB} = \frac{r_A - r_B}{R_{AB}} \,, \tag{5.86}$$

where for a given point in space $\mathbf{r}$, r_A and r_B denote distances to nuclei A and B from that point, and R_{AB} is the internuclear separation between this pair of centers:

$$\begin{aligned} r_A &= |\mathbf{r} - \mathbf{r}_A| \,, \\ r_B &= |\mathbf{r} - \mathbf{r}_B| \,, \\ R_{AB} &= |\mathbf{r_A} - \mathbf{r}_B| \,. \end{aligned}$$

If one used non-overlapping cells (where weights ω are ideal step functions), all points for which $-1 \leq \mu_{AB} \leq 0$, $\forall B \neq A$ would belong exclusively to the polyhedron of the atom A. This simplified scenario would correspond to the following ideal step function $s_B^k(\mu_{AB})$:

$$s_B^k(\mu_{AB}) = \begin{cases} 1, & \text{if} \quad -1 \leq \mu_{AB} \leq 0 \\ 0, & \text{if} \quad 0 < \mu_{AB} \leq 1 \end{cases} . \tag{5.87}$$

Instead, Becke proposed a following smoothing function,

$$s_A^k(\mu_{AB}) = \frac{1}{2}\left[1 - f_k(\mu_{AB})\right], \tag{5.88}$$

which was constructed iteratively as

$$f_1(\mu_{AB}) = \frac{3}{2}\mu_{AB} - \frac{1}{2}\mu_{AB}^3, \tag{5.89}$$

$$f_k(\mu_{AB}) = f[f_{k-1}(\mu_{AB})]. \tag{5.90}$$

After thorough exploration, Becke finally adopted $k = 3$. Formally, the above formulation was given for homoatomic pairs. In the case of heteroatomic pairs, Becke proposed to use $s_A^k(\nu_{AB})$ instead, with ν_{AB} calculated as

$$\nu_{AB} = \mu_{AB} + a_{AB}(1 - \mu_{AB}^2). \tag{5.91}$$

In the above formula, the role of the parameter a_{AB} was to account for the different sizes of the atoms A and B,

$$a_{AB} = \frac{u_{AB}}{u_{AB}^2 - 1} \quad \text{and} \quad u_{AB} = \frac{\chi_{AB} - 1}{\chi_{AB} + 1}, \tag{5.92}$$

where $\chi_{AB} = R_{0,A}/R_{0,B}$ is the ratio of the Slater-Bragg radii of atoms A, $R_{0,A}$, and B, $R_{0,B}$.

Becke's scheme adapting the fuzzy Voronoi cells is the most widely applied multicenter partitioning of molecular integrals in quantum chemistry. However, there have been proposed other partitioning schemes, the most notable of those are Hirshfeld partitioning and the one based on the Quantum Theory of Atoms in Molecules (QTAIM). Salvador and Ramos-Cordoba extended Becke's scheme to include the density topological descriptors to define ν_{AB}, rather than the fixed ratio of the Slater-Bragg radii of atoms [216]. This allowed to account for the differences in sizes of atomic basins in different chemical environments, especially those due to the partial ionic character of atoms. In the Hirshfeld partitioning, the atomic weight in Eq. 5.84 is defined as the ratio [217]

$$\omega_A(\mathbf{r}) = \frac{\rho_A^0(\mathbf{r})}{\sum_B \rho_B^0(\mathbf{r})}, \tag{5.93}$$

where $\rho_X^0(\mathbf{r})$ correspond to the spherically averaged density of isolated atom X placed at the equivalent position in space as in the studied molecule.

5.6.2 Atomic Integration in the Spherical Polar Coordinate System

After defining atomic contributions to the F integrand, one has to solve separately the I_A integrals. Because I_A are atomic centered it is naturally to express them in the spherical polar coordinate system, (r, θ, and ϕ), and treating the origin of the integration at the position of the nucleus:

$$\begin{aligned} I_A &= \int \mathrm{d}\mathbf{r} F_A(\mathbf{r}), \\ &= \int_0^\infty \mathrm{d}r\, r^2 \int_0^\pi \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi F_A(r,\theta,\phi). \end{aligned} \tag{5.94}$$

Then, one calculates atomic integral in a stepwise manner, namely, firstly integrating over the angular part,

$$\overline{F_A}(r) = \int_0^\pi \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi F_A(r,\theta,\phi), \tag{5.95}$$

and secondly, integrating the radial part,

$$I_A = \int_0^\infty \mathrm{d}r\, r^2 \overline{F_A}(r)\,. \tag{5.96}$$

At this point, one introduces the actual numerical integration to approximate the Eqs. 5.95 and 5.96. To do so, for each radial shell r_i (defining the sphere of radius r_i) the angular integral is evaluated using the angular quadrature,

$$\overline{F_A}(r_i) \approx \sum_j^{N_\Omega} w_j^{\theta,\phi} F_A(r_i, \theta_j, \phi_j)\,, \tag{5.97}$$

where $w_j^{\theta,\phi}$ is the angular grid-weight and θ_j, ϕ_j are the angular coordinates of the j-th angular grid-point. Subsequently, the radial quadrature is utilized to finally evaluate the I_A integral,

$$I_A \approx \sum_i^{N_r} w_i^r \overline{F_A}(r_i)\,, \tag{5.98}$$

where, equivalently, w_i^r is the radial grid-weight and r_i is the radial distance from the origin of integration (i.e. the radial shell) of the i-th radial grid point.

Therefore, the overall numerical approximation of the I_A integral can be concisely written as:

$$I_A \approx \sum_i^{N_r} w_i^r \sum_j^{N_\Omega} w_j^{\theta,\phi} F_A(r_i, \theta_j, \phi_j)\,. \tag{5.99}$$

Then, the problem of numerical integration boils down to determining grid points and weights. Traditionally, the size of the integration grid in the KS-DFT calculations is denoted as (N_r, N_Ω), where N_r and N_Ω, respectively, indicate the number of radial shells and angular points per atom, hence, the total number of grid points per atom is $N_r N_\Omega$.

5.6.3 Angular Grids

In most of the applications in quantum chemistry, the angular part is integrated via the Lebedev quadrature [218–221]. Each integral over the unit sphere $S^2 = \{\mathbf{r} \in R^3 : |\mathbf{r}| = 1\}$ is expressed with the Lebedev quadrature as [218–221]

$$\int_{S^2} \mathrm{d}\Omega f(\mathbf{r}) = \int_0^\pi \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi\, f(\theta,\phi) = 4\pi \sum_i^{N_\Omega} w_i^\Omega f(\mathbf{r}_i)\,, \tag{5.100}$$

where w_i^Ω are the grid weights and r_i are the points located in the unit sphere. Actually, one can always expand any function $f(\theta,\phi)$ defined over a unit sphere as a sum of the spherical harmonics,

$$f(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_{lm} Y_l^m(\theta,\phi)\,. \tag{5.101}$$

Spherical harmonic Y_l^m of the l-th degree and the m-th order is defined using the normalized associated Legendre function of corresponding degree and order:

$$Y_l^m(\theta,\phi) = \frac{1}{\sqrt{2}} P_l^m(\cos\theta) e^{im\phi}\,. \tag{5.102}$$

Grid weights and sampling point positions are obtained by solving a set of non-linear equations that were derived by enforcing symmetry and exactness conditions. Lebedev

solved and tabulated angular grids up to $l_{\max} = 131$, which corresponds to a grid of 5810 points.

Lebedev's angular grid owes its popularity to its very unique property the grid of the l-th order integrates exactly all spherical harmonics up to the l-th degree [218–223]. This is of special importance because the spherical harmonics actually define the angular part of the atom-centered basis functions utilized in quantum chemistry. This corresponds to a finite sum equivalent to Eq. . However, more complex integrands involving kernels of DFAs, are much more complicated and combine linear harmonic in a nonlinear manner. It has also been shown that Lebedev grids are very cost-efficient when compared to other schemes for integration of the angular part, such as angular product grids involving Gauss-Legendre quadratures. Additionally, the Lebedev grid is invariant under symmetry operations of the octahedral point group. Unfortunately, this particular feature does not provide any gain in the real calculations of molecules because of the adoption of multi-center partitioning of molecular integrals.

5.6.4 Radial Grids

Radial integration of the atomic integral is treated as a typical 1D problem,

$$I = \int_0^\infty \mathrm{d}r\, r^2 f(r)\,, \tag{5.103}$$

where for brevity, following simplification of symbols is used from now on, namely $I = I_A$, $f(r) = \overline{F}_A(r)$, and $w_i = w_i^r$. However, the radial integral defined in the unbounded range $[0;\infty)$ requires either the application of a special quadrature or the mapping to a finite integral before applying the quadrature. In almost all radial grids used in quantum chemistry, the latter approach is adapted, which allows for more tunable integration. This is important because most molecular integrands are concentrated in the core regions of atoms.

The two most widely used quadratures for the radial grid of DFAs are Euler-Maclaurin quadrature of second order and Chebyshev quadrature of the second kind. Euler-Maclaurin quadrature, in general, can approximate any type of function, usually defined over the range $[0,1]$,

$$\int_0^1 \mathrm{d}x f(x) \approx \sum_i^n w_i f(x_i)\,, \tag{5.104}$$

where n is the number of radial points. Euler-Maclaurin uses equally spaced points x_i and equal weights w_i [224],

$$x_i = \frac{i}{n+1},, \tag{5.105}$$

$$w_i = 1\,. \tag{5.106}$$

In the Chebyshev quadrature of the second kind, defined over the range $[-1,1]$, one approximates a particular type of integral [225],

$$\int_{-1}^1 \mathrm{d}x \sqrt{1-x^2} f(x) \approx \sum_i^n w_i f(x_i)\,, \tag{5.107}$$

with the sampling points x_i being the nodes of the Chebyshev polynomials of the second

kind with the corresponding weights w_i,

$$x_i = \cos\left(\frac{i\pi}{n+1}\right), \tag{5.108}$$

$$w_i = \frac{\pi}{n+1}\sin^2\left(\frac{i\pi}{n+1}\right). \tag{5.109}$$

However, prior to the application of such quadratures, one has to map the semi-infinite domain $[0;\infty)$ into the finite domain for which a given quadrature has been developed. For that reason, one has to apply the change of variables $r = R(t)$, with t being the new integration coordinate for the finite domain. In general, after applying the $R(t)$ mapping, one converts the integral in Eq. 5.103 to

$$I = \int_a^b \mathrm{d}t R'(t) R^2(t) f(R(t)), \tag{5.110}$$

where $R'(t) = \mathrm{d}R(t)/\mathrm{d}t$ is the derivative coming from the Jacobian transformation of the integral. The choice of the mapping function is crucial, it has to provide the highest resolution in the core region of the atoms and a bit sparser resolution in the valence region of the atoms (to minimize the number of redundant points r which do not contribute to the property). The resolution of the mapping is controlled with $R'(t)$, small value of $R'(t)$ tells that a given interval for r is mapped onto a large interval of t (implying high resolution) [226]. For an efficient mapping, $R'(t)$ should be small where $R(t)$ is small, and large where $R(t)$ is large [226].

In the rest of this section a few of commonly applied radial grids are briefly introduced. This short compilation corresponds to the one given in the work of Gill and Chien [223]. For each of the presented radial grids, the description includes its adopted mapping $R(t)$, the form of the transformed integral I on the finite ranges, sampling points r_i and their corresponding weights w_i. Moreover, the type of function integrated exactly by such given radial grid is also shown.

In his work on multicenter partitioning, Becke also proposed the radial grid to solve the atomic integrals [215]. It adapted the following mapping $R(t)$ onto the $[-1,1]$ interval:

$$r = R(t) = R_0\frac{1+t}{1-t}, \tag{5.111}$$

$$I = 2R^3\int_{-1}^{1}\mathrm{d}t\frac{(1+t)^2}{(1-t)^4}f(R(t)). \tag{5.112}$$

Then, to integrate Eq. 5.112, Becke used the Chebyshev quadrature of the second kind (see Eqs. 5.107 to 5.109). This leads to the following equations for radial points r_i and weights w_i:

$$r_i = R_0\frac{1+t_i}{1-t_i}, \tag{5.113}$$

$$w_i = \frac{2\pi}{n+1}R_0^3\frac{(1+t_i)^{5/2}}{(1-t_i)^{7/2}}, \tag{5.114}$$

$$t_i = \cos\left(\frac{i\pi}{n+1}\right). \tag{5.115}$$

As shown by Gill and Chien, Becke's radial grid with n points can integrate exactly the generic function $f(r)$,

$$f(r) = \frac{1}{r^{3/2}(r+R_0)^3}P_{2n-1}\left(\frac{r-R_0}{r+R_0}\right), \tag{5.116}$$

where P_{2n-1} is a polynomial of the $2n$-1-th order.

Murray, Handy and Laming proposed a radial grid, named Handy's grid, which was based on the following mapping to the $[0,1]$ interval [224],

$$r = R(t) = R_0 \frac{t^2}{(1-t)^2}\,, \tag{5.117}$$

$$I = 2R_0^3 \int_0^1 \mathrm{d}t \frac{t^5}{(1-t)^7} f\left(R(t)\right)\,. \tag{5.118}$$

Handy's grid utilizes the second-order Euler-Maclaurin summation formula (see Eqs. 5.104 to 5.106) to calculate the above integral, which yields the following radial points and weights:

$$r_i = R_0 \frac{t_i^2}{(1-t_i)^2}\,, \tag{5.119}$$

$$w_i = R_0^3 \frac{2t_i^5}{(n+1)(1-t_i)^7}\,, \tag{5.120}$$

$$t_i = \frac{i}{n+1}\,. \tag{5.121}$$

The n-point Handy grid integrates exactly the generic function $f(r)$:

$$f(r) = \frac{1}{r^{5/2}(\sqrt{r}+\sqrt{R_0})^2} L\left(\frac{\sqrt{r}}{\sqrt{r}+\sqrt{R_0}}\right)\,, \tag{5.122}$$

where $L(x)$ is a continuous function linear between the roots x_i and vanishing at the boundaries. Treutler and Ahlrichs proposed a more complicated transformation of the coordinates to the $[-1,1]$ interval [227],

$$r = R(t) = -R_0 \frac{(1+t)^\alpha}{\ln 2} \ln\left(\frac{1-t}{2}\right)\,, \tag{5.123}$$

$$I = \left(\frac{R_0}{\ln 2}\right)^3 \int_{-1}^1 \mathrm{d}t (1+t)^{3\alpha} \left[\frac{1}{1-t}\ln^2\left(\frac{1-t}{2}\right) - \frac{\alpha}{1+t}\ln\left(\frac{1-t}{2}\right)\right] f\left(R(t)\right)\,, \tag{5.124}$$

and combined it with the Chebyshev quadrature, yielded the following set of points and weights,

$$r_i = -R_0 \frac{(1+t_i)^\alpha}{\ln 2} \ln\left(\frac{1-t_i}{2}\right)\,, \tag{5.125}$$

$$w_i = R_0^3 \frac{\pi}{n+1} \frac{(1+t_i)^{3\alpha}}{\ln^3 2} \left[\sqrt{\frac{1+t_i}{1-t_i}}\ln^2\left(\frac{1-t_i}{2}\right) - \alpha\sqrt{\frac{1-t_i}{1+t_i}}\ln^3\left(\frac{1-t_i}{2}\right)\right]\,, \tag{5.126}$$

$$t_i = \cos\left(\frac{i\pi}{n+1}\right)\,, \tag{5.127}$$

where, basing on their extensive tests, α was set to 0.6 [227]. Gill and Chien could not determine the type of functions which are exactly integrated by Ahlrichs' grid, due to the complicated mapping in Eq. 5.123. Mura and Knowles suggested the following transformation $R(t)$ [228],

$$r = R(t) = -R_0 \ln(1-t^3)\,, \tag{5.128}$$

$$I = 3R_0^3 \int_0^1 \mathrm{d}t \frac{t^2 \ln^2(1-t^3)}{1-t^3} f\left(R(t)\right) , \tag{5.129}$$

which maps the integral over the range $[0,1]$. They combined it with the Euler-Maclaurin integration, yielding the following set of radial points and weights,

$$r_i = -R_0 \ln(1-t_i^3) , \tag{5.130}$$

$$w_i = R_0^3 \frac{3t_i^2 \ln^2(1-t_i^3)}{(n+1)(1-t_i^3)} , \tag{5.131}$$

$$t_i = \frac{i}{n+1} , \tag{5.132}$$

$$f(r) = \frac{e^{-r/R_0}}{r^2(1-e^{-r/R_0})^{2/3}} L[(1-e^{-r/R_0})^{1/3}] . \tag{5.133}$$

In the same work, Gill and Chien proposed their own radial grid, namely the MultiExp grid [223]. It was constructed to properly integrate linear combinations of exponential functions, namely,

$$f(r) = \sum_{k=1}^{2m} c_k e^{-kr/R_0} . \tag{5.134}$$

The authors were motivated by the fact that the basis sets applied in most of the KS-DFT computations are of Slater (STO) or Gaussian (GTO) type. In the latter case, GTO basis functions are still contracted to mimic the sharpness of STOs. In the MultiExp grid, the following mapping is applied to the $[0,1]$ interval:

$$r = R(t) = -R_0 \ln t , \tag{5.135}$$

$$I = R_0^3 \int_0^1 \mathrm{d}t t^{-1} (\ln t)^2 f\left(R(t)\right) . \tag{5.136}$$

In contrast to other radial grids, the MultiExp grid has its own unique set of grid points and weights,

$$r_i = -R_0 \ln t_i , \tag{5.137}$$

$$w_i = R_0^3 \frac{a_i}{x_i} , \tag{5.138}$$

where a_i and x_i are found numerically by inverting the Cholesky triangle of the associated Gram matrix of polynomials $Q_n(t)$, which are orthogonal to the weight function $\ln^2 t$.

None of the radial quadratures ideally approaches the problem of integration in KS-DFT. Many works tried to assess the efficiency of each radial quadrature, along with new improved schemes based on local augmentation or adaptive grid schemes [222, 226–230].

Chapter 6

Chemical Bonding Analysis

6.1 Introduction

In this Chapter, we briefly introduce the methodology to perform a topological analysis of the electronic density within the Quantum Theory of Atoms in Molecules (QTAIM) and to compute formal oxidation states of the atoms in a molecule using the effective oxidation states approach. In Chapter 11 of this Thesis, these two theories are used to characterize very unique and rare types of systems called molecular electrides. In molecular electrides, one can identify isolated localized electrons that formally do not belong to any of the atoms and may act as the smallest anions possible.

6.2 Quantum Theory of Atoms in Molecules

There exist plethora of ways of defining the atomic moieties in a molecule from first principles. In general, they can be categorized into those operating in the Hilbert space (i.e., orbital space) or in the Cartesian space (i.e., the real space of positions of electrons and nuclei). The most notable partitionings belonging to the former group are the ones developed by Mulliken[231] and Löwdin[232]. Shortly, they define an atom in terms of the AOs localized in the corresponding nucleus, and all expectation values of local atomic properties are evaluated only using such set of atomic orbitals. Bonds, on the other hand, are mostly characterized through the overlap of the functions belonging to a pair of atoms. The problem of such types of partitionings is their heavy dependence on the type and size of the basis set and the level of computational theory. The second group of partitionings operates purely in the real space and atoms are defined based on the geometry of the system, or on the local character of the electronic density. The most commonly used in quantum chemistry are the ones based on the geometrical partitioning of the electron density using the Voronoi polyhedra, [215, 216] Hirshfeld-based partitionings, [217, 233–236], and the QTAIM proposed by Richard Bader [237]. These approaches can be applied to most of the computational methods, and are based on the information stored in the electronic density.

In QTAIM, a definition of an atom within a molecule is rigorously based on the topology of the electronic density. The boundaries of the atoms are limited by the zero-flux surface of the gradient vector field of the charge density, $\Omega(\mathbf{r})$. This surface is defined as the set of points in space fulfilling the following zero-flux condition:

$$\nabla\rho(\mathbf{r})\cdot\mathbf{n}(\mathbf{r}) = 0 \quad \forall\mathbf{r}\in\Omega(\mathbf{r})\,, \tag{6.1}$$

where $\mathbf{n}(\mathbf{r})$ is the unit vector perpendicular to $\Omega(\mathbf{r})$. This boundary condition partitions the Cartesian space into a set of disjoint regions, where each region is defined by the basin of a single nucleus (which formally is a nuclear attractor, see below). One of the main concepts of QTAIM, in addition to the formal topological definition of atoms, is the atomic additivity of energies and average properties. As proved by Bader, the average value of the observable $\langle\hat{O}\rangle$ for the total system (O can be both a one- or two-particle operator), is expressed as the sum of the atomic contributions $O(\Omega)$ [237, 238]:

$$\langle\hat{O}\rangle = \sum_{\Omega} O(\Omega)\,. \tag{6.2}$$

Within QTAIM, one can partition most of the properties into atomic or functional group contributions and use those to predict properties in new molecules. Another key feature of QTAIM is definition of the topological descriptors, which mark particular type of interactions between atoms and may also recover the Lewis model of chemical bonds [239]. The

former is directly related to the topology of the density and the presence of the critical points, whereas the latter is related to the description of the localization of electrons using the Laplacian of the density.

Critical points of the density, r_c, are the points in the (Cartesian) space, where the gradient of the density $\nabla\rho(\mathbf{r}_c)$ vanishes,

$$\nabla\rho(\mathbf{r}_c) = 0\,. \tag{6.3}$$

The character of these critical points is determined from the Hessian matrix of the electronic density, $\mathbf{H}(\mathbf{r}_c)$, which contains all partial second derivatives with respect to the x, y, and z Cartesian components:

$$\mathbf{H}(\mathbf{r}_c) = \nabla^T\nabla\rho(\mathbf{r})|_{\mathbf{r}=\mathbf{r}_c} = \begin{pmatrix} \frac{\partial^2\rho(\mathbf{r})}{\partial x^2} & \frac{\partial^2\rho(\mathbf{r})}{\partial x\partial y} & \frac{\partial^2\rho(\mathbf{r})}{\partial x\partial z} \\ \frac{\partial^2\rho(\mathbf{r})}{\partial y\partial x} & \frac{\partial^2\rho(\mathbf{r})}{\partial y^2} & \frac{\partial^2\rho(\mathbf{r})}{\partial y\partial z} \\ \frac{\partial^2\rho(\mathbf{r})}{\partial z\partial x} & \frac{\partial^2\rho(\mathbf{r})}{\partial z\partial y} & \frac{\partial^2\rho(\mathbf{r})}{\partial z^2} \end{pmatrix}_{\mathbf{r}=\mathbf{r}_c}, \tag{6.4}$$

which is diagonalized to obtain the eigenvalue matrix $\mathbf{\Lambda}(\mathbf{r}_c)$:

$$\mathbf{H}(\mathbf{r}_c)\mathbf{L} = \mathbf{L}\mathbf{\Lambda}(\mathbf{r}_c)\,, \tag{6.5}$$

$$\mathbf{\Lambda}(\mathbf{r}_c) = \begin{pmatrix} \frac{\partial^2\rho(\mathbf{r})}{\partial x_i^2} & 0 & 0 \\ 0 & \frac{\partial^2\rho(\mathbf{r})}{\partial y_i^2} & 0 \\ 0 & 0 & \frac{\partial^2\rho(\mathbf{r})}{\partial z_i^2} \end{pmatrix}_{\mathbf{r_1}=\mathbf{r}_c} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \tag{6.6}$$

These eigenvalues determine the (principal) curvatures of $\rho(\mathbf{r})$ at $\mathbf{r}_c$, namely the positive eigenvalue corresponds to a minimum of $\rho(\mathbf{r})$ along the corresponding direction, whereas the negative eigenvalue corresponds to the maximum of $\rho(\mathbf{r})$ at $\mathbf{r}_c$. The type of each critical point, which is usually denoted as (σ,λ), is determined by the rank σ of the Hessian, i.e., the number of nonzero eigenvalues, and its signature λ, defined as the sum of the signs of the eigenvalues λ_1, λ_2, and λ_3. There are four types of CPs (corresponding to four possible signatures), each pointing to a different element of the chemical structure:

- Nuclear attractor critical points (NACPs) with CP = (3,-3) have all negative λ_i eigenvalues and are local maxima of $\rho(\mathbf{r})$. Each atomic basin in QTAIM contains one nuclear attractor, at which all trajectories of $\nabla\rho(\mathbf{r})$ terminate. Nuclear attractors are usually found at the position of the nuclei. One can also find attractors that do not coincide with any nucleus — such unique points in the density topology are named non-nuclear attractor critical points (NNACPs). This very rare feature is a footprint of an exceptional electron localization. However, the values of $\rho(\mathbf{r})$ at NNACPs are orders of magnitude smaller than at NACPs.

- Bond critical points (BCPs) with CP = (3,-1) have two negative λ_i eigenvalues and one positive. The latter corresponds to the minimum along the axis connecting two NACPs. BCP is an universal indicator of bonding between two highly interacting atoms.

- Ring critical points (RCPs) with CP = (3,+1) have two positive eigenvalues and one negative. It indicates a ring-like pattern in $\rho(\mathbf{r})$, where at $\mathbf{r}_c$ there are minima in two directions forming a plane, and a maximum along the direction perpendicular to ring plane. The most notable example of RCP is the center of benzene ring.

- Cage critical points (CCPs) with CP = (3,+3) have all positive eigenvalues making this point a local minimum in three dimensions. Its presence indicates a multi-atom cage-like structure and is usually located near the center of such a cage.

The number of critical points satisfy a relationship given by the topological Poincaré-Hopf theorem:

$$N_{NACP} + N_{NNACP} - N_{BCP} + N_{RCP} - N_{CCP} = 1\,, \tag{6.7}$$

If this relation is violated, most likely not all critical points were found in the search or some critical points are due to discontinuities of the electronic density.

The Laplacian of the electronic density, $\nabla^2\rho(\mathbf{r})$, is a topological descriptor of the electron localization [238, 240]:

$$\nabla^2\rho(\mathbf{r}) = \frac{\partial^2\rho(\mathbf{r})}{\partial x^2} + \frac{\partial^2\rho(\mathbf{r})}{\partial y^2} + \frac{\partial^2\rho(\mathbf{r})}{\partial z^2}\,. \tag{6.8}$$

The negative values of $\nabla^2\rho(\mathbf{r})$ are natural signs of the electron localization - they correspond to regions where the potential energy is dominant over the kinetic energy [238, 240]. Regions with large negative $\nabla^2\rho(\mathbf{r})$ are usually found very close to the nuclei, at chemical bonds, and the positions of lone pairs of electrons. In contrast, the positive values of $\nabla^2\rho(\mathbf{r})$ denote the regions where the kinetic energy dominates, and hence there is electron depletion [238, 240].

6.3 Effective Oxidation State Analysis

The oxidation state (OS) is a fundamental concept in chemistry used to rationalize observations and make further predictions. It is used in the analysis of the chemical structures and redox properties, mechanisms and bonding patters in organometallic chemistry, and the rationalization of the spectroscopic properties related to electro–magnetic response of matter. While being the bread and butter of all chemists, its formal definition still has few ambiguities. The latest formal definition of OS established by IUPAC in 2016 says *"the oxidation state of an atom is the charge of this atom after ionic approximation of its heteronuclear bonds"* [241]. Hence, it is justified to say that OS is a property of atoms embedded within a molecule and in many cases OS values may be transferable within the same functional group. However, despite its simple definition, the realization of the OS concept from first principles is not trivial. Several studies have been devoted to establish a connection between particular quantum chemistry magnitudes and the values of OS obtained following the IUPAC recommendations; moreover, several authors have proposed their own computational routines to predict the most likely OS of a given atom [241–248].

In 2015, Ramos-Cordoba, Postils, and Salvador proposed a universal procedure for the determination of oxidation state of atomic-moieties from first principles, namely the Effective Oxidation State (EOS) analysis [248, 249]. It utilizes the concept of effective atomic orbitals (eff-AOs), which are distorted atomic orbitals that come from the real space partitioning (or localization) of MOs into atomic regions [250–253].

For all the electrons of the same spin in each atom A, the procedure for the construction of eff-AOs is as follows. First, one defines the intra-atomic contribution of the i-th MO to atom A, $\phi^A_{i,\sigma}(\mathbf{r})$, as [250–253]

$$\phi^A_{i,\sigma}(\mathbf{r}) = \omega_A(\mathbf{r})\phi_{i,\sigma}(\mathbf{r})\,, \tag{6.9}$$

where $\omega_A(\mathbf{r})$ is the atomic-partitioning weight (which considering all the atoms of the molecule sum up to one for all electron coordinates, $\sum_A \omega_A(\mathbf{r}) = 1$). These atomic weights

are the same as those applied to electronic densities to define the atomic basins, and can come from any type of atomic partitioning, both in the Cartesian space, such as the Becke, Hirshfeld-based, or QTAIM partitionings, as well as in the Hilbert space, such as the Mulliken or Löwdin population analyses [248, 249]. Second, the overlap matrix $\mathbf{S}^A_\sigma$ between the intra-atomic parts of MOs is defined as:

$$S^A_{ij,\sigma} = \int \mathrm{d}\mathbf{r}\phi^{A*}_{i,\sigma}(\mathbf{r})\phi^A_{j,\sigma}(\mathbf{r})\,. \tag{6.10}$$

Third, $\mathbf{S}^A_\sigma$ is diagonalized to yield the eff-AOs occupancies, $\boldsymbol{\lambda}^A_\sigma$ (the eigenvalues) and a unitary matrix of the transformation coefficients, $\mathbf{U}^A_\sigma$ (the eigenvectors)

$$\boldsymbol{\lambda}^A_\sigma = \mathbf{U}^{A\dagger}_\sigma \mathbf{S}^A_\sigma \mathbf{U}^A_\sigma\,. \tag{6.11}$$

The $\mathbf{U}^A_\sigma$ are used to finally define eff-AO, $\eta^A_{i,\sigma}(\mathbf{r})$, as:

$$\eta^A_{i,\sigma}(\mathbf{r}) = \frac{1}{\sqrt{\lambda^A_{i,\sigma}}}\sum_j^{n_\sigma} U^A_{ji,\sigma}\phi^A_{j,\sigma}(\mathbf{r})\,. \tag{6.12}$$

Each $\eta^A_{i,\sigma}(\mathbf{r})$ has its unique effective occupation number $\boldsymbol{\lambda}^A_\sigma$, which has a similar interpretation as the NO occupancy. In the above procedure, only eff-AOs with nonnegative eigenvalues are constructed and for most of the partitioning schemes, $0 < \lambda^A_{i,\sigma} \leq 1$.

Afterwards, one finally calculates the formal oxidation state with the EOS scheme. The eff-AO occupancies of all atoms (or fragments), $\lambda^A_{i,\sigma}$, are gathered and sorted descendingly. The first $N_{el,\sigma}$ eff-AOs, with $N_{el,\sigma}$ being the number of σ electrons, are set as fully occupied, that is, $\lambda^A_{i,\sigma} \to \lambda^{A,id}_{i,\sigma} = 1$, while the rest are set as fully empty, that is, $\lambda^A_{i,\sigma} \to \lambda^{A,id}_{i,\sigma} = 0$. Such a formal integer assignment of occupancies is a representation of the discrete ionic model, which is used to define the oxidation state by IUPAC. Then, for each atom (or fragment), one calculates its formal effective oxidation state, EOS$[A]$, as the difference between the total number of electrons in the isolated atom/fragment, N^A_{el}, and the number of fully occupied eff-AOs of this particular fragment, which equals to $\sum_i^{n_\sigma} \lambda^{A,id}_{i,\sigma}$:

$$\mathrm{EOS}[A] = N^A_{el} - \sum_i^{n_\sigma} \lambda^{A,id}_{i,\sigma}\,. \tag{6.13}$$

To provide some insight on the reliability of the assigned oxidation state, the authors have proposed the following index, R, as the minimal value between its two spin components, namely $R = \min(R_\alpha, R_\beta)$, where each spin component is defined as

$$R_\sigma(\%) = 100 \cdot \min\left(1, \max\left(0, \lambda_{\mathrm{LO},\sigma} - \lambda_{\mathrm{FU},\sigma} + \frac{1}{2}\right)\right)\,, \tag{6.14}$$

where $\lambda_{\mathrm{LO},\sigma}$ is the occupation of the lowest occupied eff-AO (LO), that is, the smallest value of $\lambda^A_{i,\sigma}$ within the set of occupied eff-AOs, and $\lambda_{\mathrm{FU},\sigma}$ is the occupation of the first unoccupied eff-AO (FU), i.e., the largest value of $\lambda^A_{i,\sigma}$ within the set of unoccupied eff-AOs. The reliability index R takes values in the range between 50% and 100%, where $R \approx 50\%$ indicates a fully uncertain assignment of the oxidation state. In contrast, $R \approx 100\%$ indicates a very reliable prediction of the oxidation state which comes from a large difference between $\lambda_{\mathrm{LO},\sigma}$ and $\lambda_{\mathrm{FU},\sigma}$.

Two most important features of the EOS analysis are that: 1) it provides almost equivalent results for all real-space definitions of atoms (i.e., it is robust among different weight definition, $\omega_A(\mathbf{r})$) and 2) it can be applied to any electronic structure method for which the first reduced density matrix, $\rho_1(\mathbf{r}'\sigma;\mathbf{r}\sigma)$, is available.

Part III

Results

Chapter 7

Partition of Optical Properties Into Orbital Contributions

7.1 Introduction

Quantum chemistry provides a mathematical framework that allows one to compute and simulate NLOPs of promising materials and also analyze their nature and origin. The former has especially benefited from the development of computational methods, which currently allow to tackle the energy and NLOPs of large molecular structures and can guide the design of new nonlinear materials. The most used method to assess the role of the different electronic states in NLOPs is the SOS approach using the few-states approximation. Such kind of analysis is of relevance mostly for specific situations, for example, when one is interested in the resonant response related to the transition to one precisely defined excited state. In the nonresonant response region, in particular, for static hyperpolarizabilities, it is sometimes hard to distinguish one dominating electronic state describing almost all the response. Moreover, the rationalization of NLOPs using excited states is not as straightforward as if it is done in terms of the (molecular) orbitals.

This Chapter of Thesis is devoted to the decomposition procedure developed by the Candidate, namely Partition of NLOPs Into Orbital Contributions (PNOC) [146]. It expresses the electronic response in the terms of the Natural Orbitals (NOs) of the unperturbed system and serves as a tool for two main purposes: i) to study the structure-property relations for NLOPs and to provide more insight into the character of the response, and ii) to provide a basis for assessing the performance of the computational methods for the evaluation of the NLOPs, especially to analyze the origin of the errors arising from different types of electronic correlation, which can be understood in terms of the NOs. The first section is devoted to the description of the PNOC procedure, the second section describes the details of the computational protocol of the implementation of PNOC, and the remaining sections present the PNOC analysis of the NLOPs a few selected molecules to illustrate the potential of this tool. Since the PNOC is currently only developed to analyze the electronic part of the response, the corresponding *el* upper index is removed from the presented formulae. Moreover, the discussion focuses on the diagonal components of the NLOP tensors, that is, $\alpha_{\alpha\alpha}$, $\beta_{\alpha\alpha\alpha}$, and $\gamma_{\alpha\alpha\alpha\alpha}$.

7.2 The Decomposition Procedure

PNOC operates in the Hilbert space and is based on the 1-RDM representation of the NLOPs (see Section 4.6). Flowchart of the PNOC procedure is shown in Fig. 7.1 for the example of the partitioning of second hyperpolarizability $\gamma_{\alpha\beta\gamma\delta}$. One proceeds equivalently for $\alpha_{\alpha\beta}$ and $\beta_{\alpha\beta\gamma}$, using the corresponding derivative of 1-RDM.

First, one needs to obtain the derivatives of 1-RDM in the AO basis. In the current implementation, these are calculated with the finite field procedure using the following central differences formula,

$$D_{\mu\nu}^{(\alpha)} = \frac{1}{12F_\alpha}\left[D_{\mu\nu}(-2F_\alpha) - 8D_{\mu\nu}(-F_\alpha) + 8D_{\mu\nu}(F_\alpha) - D_{\mu\nu}(2F_\alpha)\right], \tag{7.1}$$

$$D_{\mu\nu}^{(\alpha\beta)} = \frac{1}{12F_\alpha^2}\left[-D_{\mu\nu}(-2F_\alpha) + 16D_{\mu\nu}(-F_\alpha) - 30D_{\mu\nu}(\mathbf{0}) + 16D_{\mu\nu}(F_\alpha) - D_{\mu\nu}(2F_\alpha)\right], \tag{7.2}$$

$$D_{\mu\nu}^{(\alpha\beta\gamma)} = \frac{1}{48F_\alpha^3}\left[D_{\mu\nu}(-4F_\alpha) - 34D_{\mu\nu}(-2F_\alpha) + 64D_{\mu\nu}(F_\alpha) - 64D_{\mu\nu}(F_\alpha) + 34D_{\mu\nu}(2F_\alpha) - D_{\mu\nu}(4F_\alpha)\right], \tag{7.3}$$

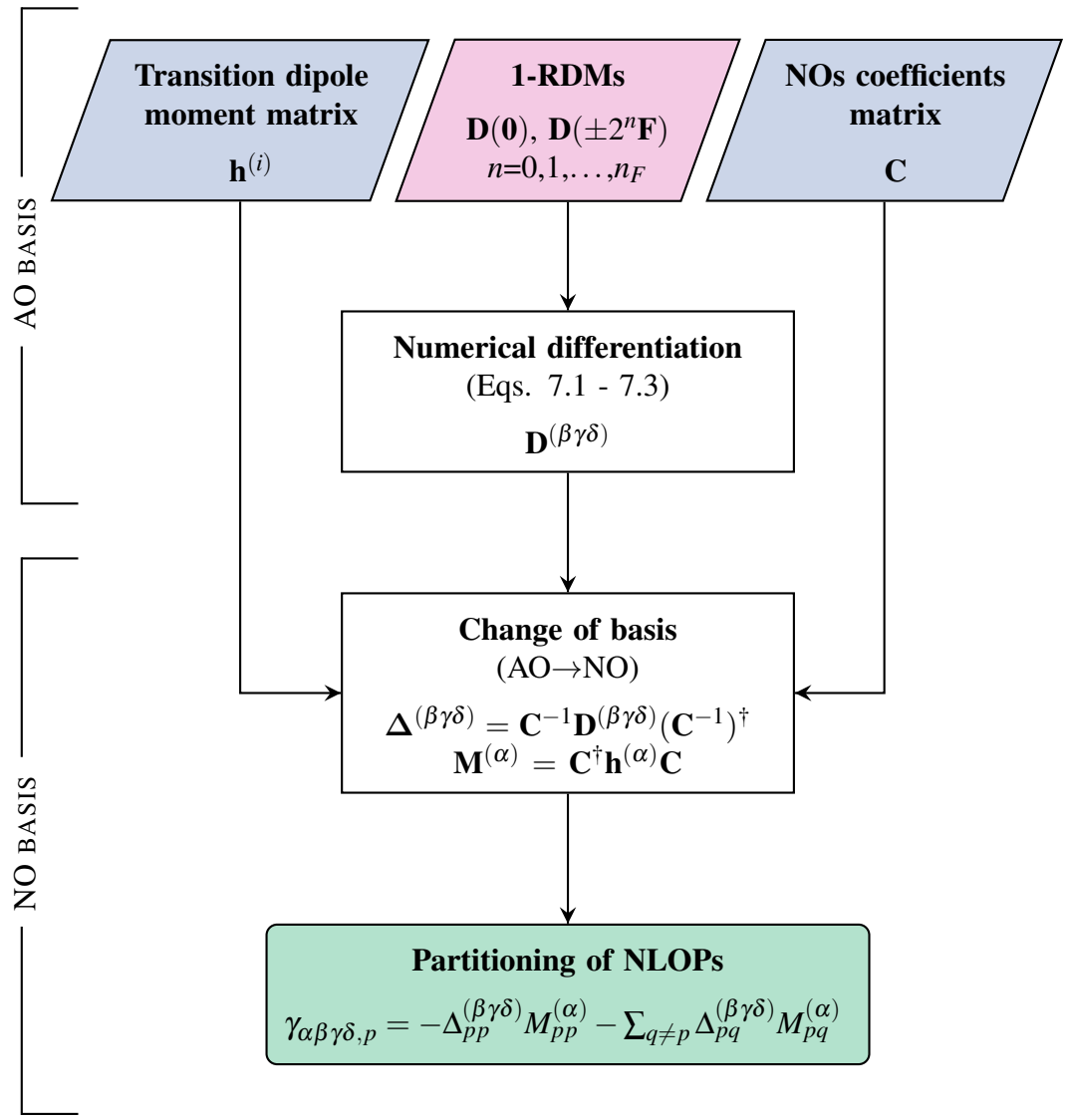

Figure 7.1: Flowchart of the PNOC procedure to obtain the NOs contributions to the second hyperpolarizability, $\gamma_{\alpha\beta\gamma\delta}$. An equivalent procedure is done for the partitioning of $\alpha_{\alpha\beta}$ and $\beta_{\alpha\beta\gamma}$ by the usage of the corresponding derivative of the 1-RDM, namely $\mathbf{D}^{(\beta)}$ and $\mathbf{D}^{(\beta\gamma)}$.

where $D_{\mu\nu}(F_\alpha)$ denotes the 1-RDM in the AO basis of the system perturbed with electric field, F_α applied along the α-th direction. These matrices are obtained from the electronic structure calculations and are available for most computational methods. Such numerical computation of the 1-RDM derivatives is straightforward because the AO basis is invariant to the applied electric field (the AO basis depends only on the position of the atoms).

Second, one projects the $\mathbf{h}^{(\alpha)}$ and $\mathbf{D}^{(\alpha\ldots)}$ from the AO to NO basis of the unperturbed, i.e., field-free, molecule. The transformation is achieved using the matrix $\mathbf{C}$, corresponding to the change between AOs, χ^{AO}, and NOs, ϕ^{NO}:

$$\phi^{\mathrm{NO}} = \chi^{\mathrm{AO}}\mathbf{C}\,, \tag{7.4}$$

where ϕ^{NO} and χ^{AO} are row vectors, and the linear combination of atomic orbitals (LCAO) coefficients are organized column wise in $\mathbf{C}$. With such matrix $\mathbf{C}$, one performs the projection and obtains an equivalent representation of NLOPs in a differently chosen

basis:

$$\begin{aligned}
\alpha_{\alpha\beta} &= -\sum_{\mu,\nu} D^{(\beta)}_{\mu\nu} h^{(\alpha)}_{\nu\mu} = -\mathrm{Tr}\left(\mathbf{D}^{(\beta)}\mathbf{h}^{(\alpha)}\right), \\
&= -\mathrm{Tr}\left(\mathbf{C}^{-1}\mathbf{D}^{(\beta)}(\mathbf{C}^{-1})^{\dagger}\mathbf{C}^{\dagger}\mathbf{h}^{(\alpha)}\mathbf{C}\right), \\
&= -\mathrm{Tr}\left(\boldsymbol{\Delta}^{(\beta)}\mathbf{M}^{(\alpha)}\right) = -\sum_{p,q}\Delta^{(\beta)}_{pq} M^{(\alpha)}_{qp}, \qquad (7.5) \\
\beta_{\alpha\beta\gamma} &= -\sum_{\mu,\nu} D^{(\beta\gamma)}_{\mu\nu} h^{(\alpha)}_{\nu\mu} = -\mathrm{Tr}\left(\mathbf{D}^{(\beta\gamma)}\mathbf{h}^{(\alpha)}\right), \\
&= -\mathrm{Tr}\left(\mathbf{C}^{-1}\mathbf{D}^{(\beta\gamma)}(\mathbf{C}^{-1})^{\dagger}\mathbf{C}^{\dagger}\mathbf{h}^{(\alpha)}\mathbf{C}\right), \\
&= -\mathrm{Tr}\left(\boldsymbol{\Delta}^{(\beta\gamma)}\mathbf{M}^{(\alpha)}\right) = -\sum_{p,q}\Delta^{(\beta\gamma)}_{pq} M^{(\alpha)}_{qp}, \qquad (7.6) \\
\gamma_{\alpha\beta\gamma\delta} &= -\sum_{\mu,\nu} D^{(\beta\gamma\delta)}_{\mu\nu} h^{(\alpha)}_{\nu\mu} = -\mathrm{Tr}\left(\mathbf{D}^{(\beta\gamma\delta)}\mathbf{h}^{(\alpha)}\right), \\
&= -\mathrm{Tr}\left(\mathbf{C}^{-1}\mathbf{D}^{(\beta\gamma\delta)}(\mathbf{C}^{-1})^{\dagger}\mathbf{C}^{\dagger}\mathbf{h}^{(\alpha)}\mathbf{C}\right), \\
&= -\mathrm{Tr}\left(\boldsymbol{\Delta}^{(\beta\gamma\delta)}\mathbf{M}^{(\alpha)}\right) = -\sum_{p,q}\Delta^{(\beta\gamma\delta)}_{pq} M^{(\alpha)}_{qp}. \qquad (7.7)
\end{aligned}$$

In above equations, $\mathbf{M}^{(\alpha)}$ is the transition dipole matrix and $\Delta^{(\alpha\ldots)}$ are the n-th order derivative of the 1-RDM calculated in the NO basis:

$$\mathbf{M}^{(\alpha)} = \mathbf{C}^{\dagger}\mathbf{h}^{(\alpha)}\mathbf{C}, \qquad (7.8)$$

$$\boldsymbol{\Delta}^{(\beta)} = \mathbf{C}^{-1}\mathbf{D}^{(\beta)}(\mathbf{C}^{-1})^{\dagger}, \qquad (7.9)$$

$$\boldsymbol{\Delta}^{(\beta\gamma)} = \mathbf{C}^{-1}\mathbf{D}^{(\beta\gamma)}(\mathbf{C}^{-1})^{\dagger}, \qquad (7.10)$$

$$\boldsymbol{\Delta}^{(\beta\gamma\delta)} = \mathbf{C}^{-1}\mathbf{D}^{(\beta\gamma\delta)}(\mathbf{C}^{-1})^{\dagger}. \qquad (7.11)$$

Finally, one can proceed with the actual partitioning. A selected (non)linear property is expressed as a sum of p NO components,

$$\alpha_{\alpha\beta} = \sum_{p}\alpha_{\alpha\beta,p}, \qquad (7.12)$$

$$\beta_{\alpha\beta\gamma} = \sum_{p}\beta_{\alpha\beta\gamma,p}, \qquad (7.13)$$

$$\gamma_{\alpha\beta\gamma\delta} = \sum_{p}\gamma_{\alpha\beta\gamma\delta,p}. \qquad (7.14)$$

Each component is obtained under the assumption that the pair contribution of p and q NOs is equally distributed among them, which leads to

$$\alpha_{\alpha\beta,p} = -\Delta_{pp}^{(\beta)} M_{pp}^{(\alpha)} - \frac{1}{2}\sum_{q\neq p}\left(\Delta_{pq}^{(\beta)} M_{qp}^{(\alpha)} + \Delta_{qp}^{(\beta)} M_{pq}^{(\alpha)}\right),$$
$$= -\Delta_{pp}^{(\beta)} M_{pp}^{(\alpha)} - \sum_{q\neq p}\Delta_{pq}^{(\beta)} M_{qp}^{(\alpha)}, \tag{7.15}$$

$$\beta_{\alpha\beta\gamma,p} = -\Delta_{pp}^{(\beta\gamma)} M_{pp}^{(\alpha)} - \frac{1}{2}\sum_{q\neq p}\left(\Delta_{pq}^{(\beta\gamma)} M_{qp}^{(\alpha)} + \Delta_{qp}^{(\beta\gamma)} M_{pq}^{(\alpha)}\right),$$
$$= -\Delta_{pp}^{(\beta\gamma)} M_{pp}^{(\alpha)} - \sum_{q\neq p}\Delta_{pq}^{(\beta\gamma)} M_{qp}^{(\alpha)}, \tag{7.16}$$

$$\gamma_{\alpha\beta\gamma\delta,p} = -\Delta_{pp}^{(\beta\gamma\delta)} M_{pp}^{(\alpha)} - \frac{1}{2}\sum_{q\neq p}\left(\Delta_{pq}^{(\beta\gamma\delta)} M_{qp}^{(\alpha)} + \Delta_{qp}^{(\beta\gamma\delta)} M_{pq}^{(\alpha)}\right),$$
$$= -\Delta_{pp}^{(\beta\gamma\delta)} M_{pp}^{(\alpha)} - \sum_{q\neq p}\Delta_{pq}^{(\beta\gamma\delta)} M_{qp}^{(\alpha)}. \tag{7.17}$$

The analysis employs NOs of the unperturbed system as the target basis, as this orbital basis is often used in the description of 1-RDM for the many-body wavefunction. If one performs the spin unrestricted computations, such as Unrestricted HF (UHF) or broken spin-symmetry KS-DFT, unrestricted NOs defined by Pulay and Hamilton are adopted instead [254]. PNOC is a universal method which can be applied to any quantum chemistry method for which 1-RDMs are available. In contrast to most of the NLOP decomposition schemes found in the literature [134, 141, 142, 255–258], PNOC is free of the origin-dependency problem in some situations. Origin–dependency in PNOC comes directly from $\mathbf{M}^{(\alpha)}$. Upon a translation of the system into a vector $\mathbf{T}$, the partitioned values of $\alpha_{\alpha\alpha,p}$ and $\gamma_{\alpha\alpha\alpha\alpha,p}$ change according to the following equations:

$$\alpha_{\alpha\alpha,p}(T_\alpha) = -\Delta_{pp}^{(\alpha)}\langle\phi_p|r_\alpha - T_\alpha|\phi_p\rangle - \sum_{q\neq p}\Delta_{pq}^{(\alpha)}\langle\phi_p|r_\alpha - T_\alpha|\phi_q\rangle,$$
$$= \alpha_{\alpha\alpha,p}(\mathbf{0}) + T_\alpha\Delta_{pp}^{(\alpha)}, \tag{7.18}$$

$$\gamma_{\alpha\alpha\alpha\alpha,p}(T_\alpha) = -\Delta_{pp}^{(\alpha\alpha\alpha)}\langle\phi_p|r_\alpha - T_\alpha|\phi_p\rangle - \sum_{q\neq p}\Delta_{pq}^{(\alpha\alpha\alpha)}\langle\phi_p|r_\alpha - T_\alpha|\phi_q\rangle,$$
$$= \gamma_{\alpha\alpha\alpha,p}(\mathbf{0}) + T_\alpha\Delta_{pp}^{(\alpha\alpha\alpha)}. \tag{7.19}$$

In the above equations, the orthonormality condition between NOs is utilized. For the systems which are centrosymmetric along the studied axis α, all diagonal elements of $\mathbf{\Delta}^{(\alpha)}$ and $\mathbf{\Delta}^{(\alpha\alpha\alpha)}$ are equal to zero, thus making the decomposed values of $\alpha_{\alpha\alpha,p}$ and $\gamma_{\alpha\alpha\alpha\alpha,p}$ origin independent. Noteworthy to mention is that $\beta_{\alpha\alpha\alpha}$ vanishes in the centrosymmetric systems.

7.3 Tested Systems and Computational Details

In this work, we study NLOPs of the three different chemical models shown in Fig. 7.2. In the first part of the Results, devoted to the role of the static and dynamic electron correlation in NLOPs, we compared the NLOP contributions in benzene, C_6H_6 and *para* - benzyne, C_6H_4. The first molecule is a typical example of a closed-shell system for which a single HF Slater determinant is a good reference for the correlated methods. Therefore, for benzene the static correlation is non-relevant and most of the properties are governed by the dynamic correlation. On the contrary, *para*-benzyne is a diradical molecule in its

singlet state, with two unpaired electrons in the *para* position. *Para*-benzyne is an open-shell system that requires to include several determinants to the reference wavefunction, and therefore it is a typical case for which the static correlation is crucial and determines most of the chemical properties. For both systems, we mostly focused on the NO contributions to the polarizability and second hyperpolarizabilities of the orbitals shown in Fig. 7.3.

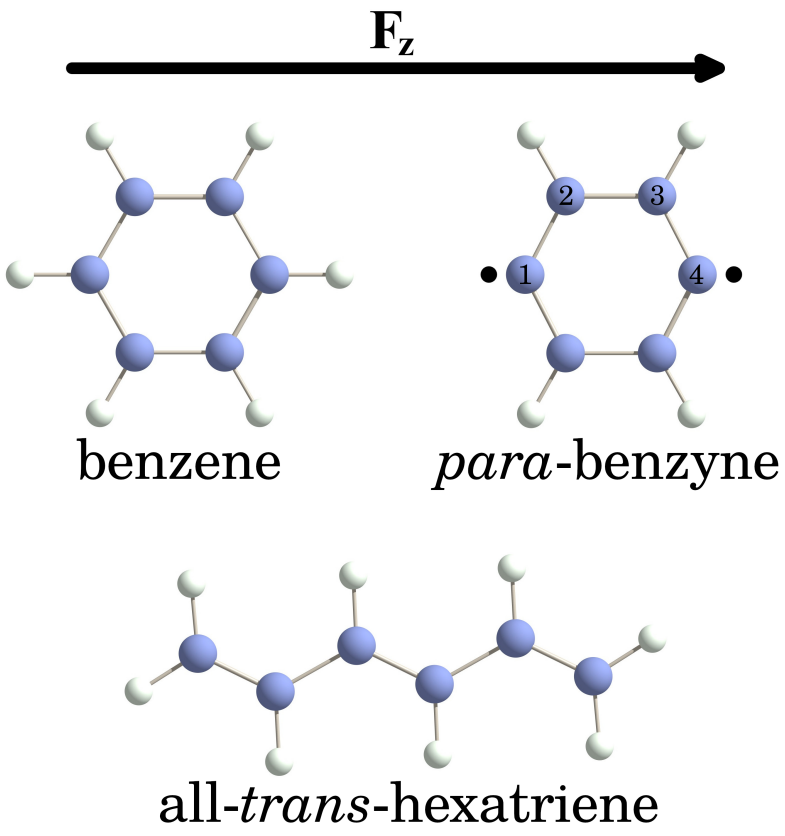

Figure 7.2: Graphical representation of the studied molecular systems, benzene and *para*-benzyne, and all-*trans*-hexatriene.

In C_6H_4, which belongs to the D_{2h} symmetry point group, we focus on the decomposition of NLOPs in the following valence NO contributions: the singly occupied σ-type NOs $5b_{1u}$ and $6a_g$ (occupied by the unpaired electrons), the occupied π-type orbitals, denoted as π(occ): $1b_{3u}$, $1b_{1g}$, $1b_{2g}$, the virtual π-type orbitals, denoted as π(vir): $2b_{3u}$, $1a_u$, $1b_{2g}$. This set of NOs was correlated in the corresponding CASSCF(8,8) calculation. In C_6H_6, which belongs to the D_{6h} symmetry point group, the valence π orbitals are the relevant NOs, namely the occupied π(occ) NOs: $1a_{2u}$, $1e_{1g}$, and $2e_{1g}$, and the virtual π(vir) NOs: $1e_{2u}$, $2e_{2u}$, and $1b_{2g}$. For the completeness, we also investigated contributions of the σ-type NOs $6e_{1u}$ and $6e_{2g}$, which are the doubly occupied counterparts of the singly occupied $5b_{1u}$ and $6a_g$ in C_6H_4. In the CASSCF computations for benzene, only the π(occ) and π(vir) were included in the active space, leading to CASSCF(6,6). Additionally, for both systems we analyzed the overall contributions of the following sets of NOs: $1s_C$ — carbon core orbitals; σ(occ) — all occupied σ-type NOs (including the above-mentioned $5b_{1u}$ and $6a_g$); σ(vir) — all virtual σ-type NOs, as well as total contributions from the π (occ) and π (vir) groups.

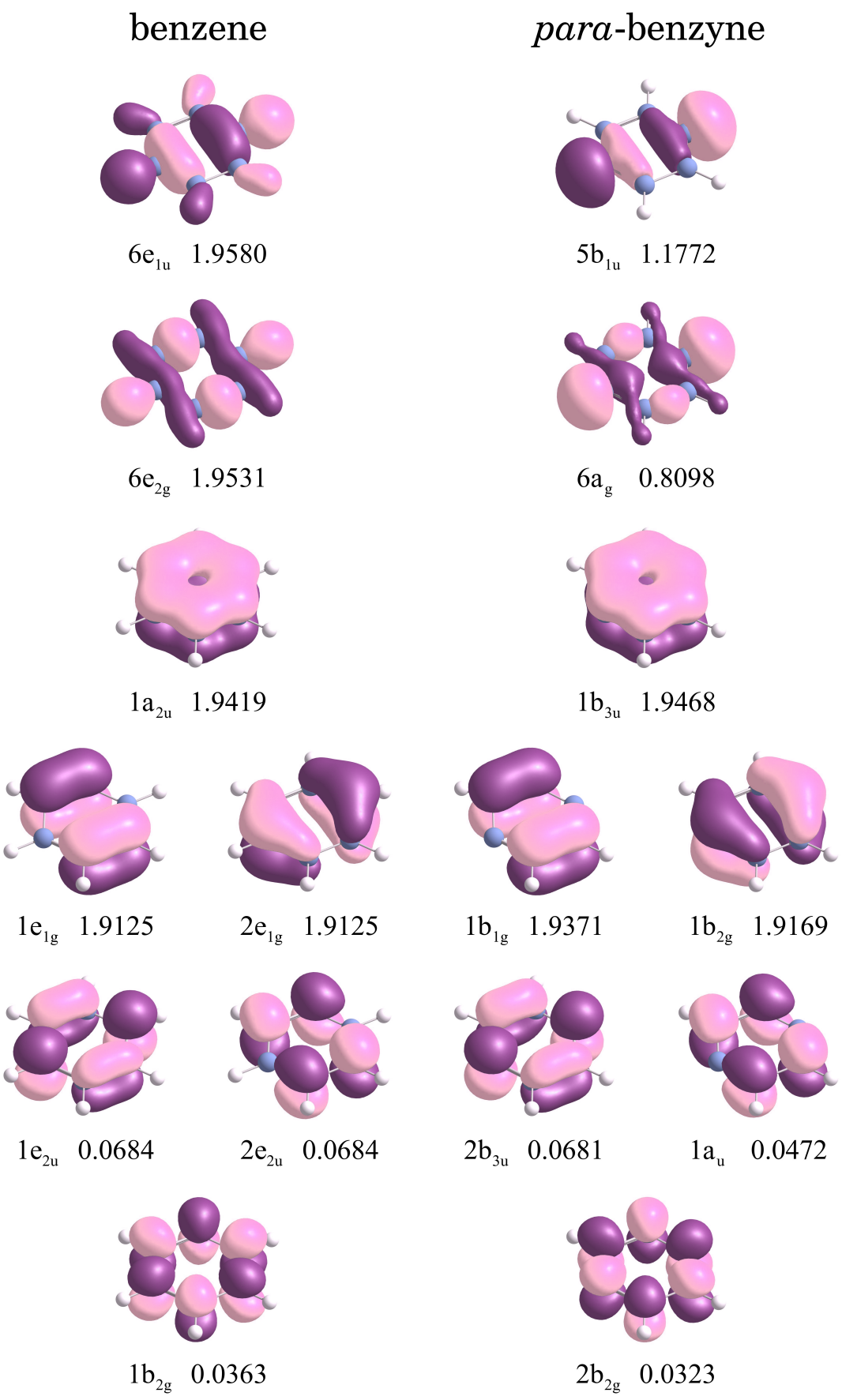

Figure 7.3: Graphical representation of the studied NOs of benzene (left panel) and *para*-benzyne (right panel). Numbers below the orbital pictures denote their natural occupancies. Obtained at the (U)CCSD/aug-cc-pVDZ level of theory.

In the second part of the Results section, using PNOC we analyze the effect of type and size of the atomic basis set on the static response of all-*trans*-hexatriene, C_6H_8. Hexatriene is a conjugated closed-shell system for which proper description of the dynamical correlation is necessary. In this system, we focus on the overall contributions of NOs belonging to two defined orbital subspaces, namely the full valence (FV) and higher virtual (HV) orbitals. The former one consists of the first 19 σ-type occupied NOs (including core 1*s* carbon orbitals), 3 occupied π-type NOs, the first 3 virtual π-type NOs, and the first 16 virtual σ-type NOs. All remaining virtual σ and π NOs are classified as HV orbitals (and their number changes with the size of the basis set).

Geometries of benzene and *para*-benzyne were optimized using (U)B3LYP/aug-cc-pVDZ, whereas the hexatriene was optimized with CAM-B3LYP/cc-PVDZ. In all calculations of static NLOPs, the systems were placed in such a way that their centers of mass coincided with the origin of the Cartesian coordinate system. Moreover, they were rotated until the main component of the diagonalized inertia tensor coincided with the z-axis of the Cartesian system. In such geometrical arrangement, the longitudinal components of polarizability and second hyperpolarizability, α_{zz} and γ_{zzzz} were studied.

The electronic structure calculations were performed with the Gaussian09 [259] computational package and PNOC was implemented in an in-house FORTRAN90 code. In all computations, the highest possible symmetry constraints were applied (accounting for electric field perturbation). Tight convergence criteria in all procedures were used, namely, in (U)HF the DIIS convergence threshold was set to 10^{-11}, in CASSCF energy convergence criteria was set to 10^{-11}, and in CCSD convergence criteria for the total energy and the norm of amplitude vector in CCSD were set to 10^{-10} and 10^{-8}, respectively. Such convergence criteria allowed to achieve the best stability in the numerical differentiation of the elements of 1-RDMs (according to Eqs. 7.1 - 7.3). For those methods that do not satisfy the Hellmann–Feynman theorem, namely (U)MP2 and (U)CCSD, relaxed and unrelaxed 1-RDMs can be often obtained. The former are more adequate for the calculation of electronic properties and, hence, we adopt them in this work.

7.4 Nature of the Static Response in Open-Shell Systems

In this section, we focus on the changes in the nature of the response when one forms the diradical structure. This analysis is based on the results obtained with (U)CCSD, which is also considered a reference method for assessing the performance of other levels of theory in the next section. In Table 7.1, the PNOC contributions of the selected NOs (and orbital subgroups) for benzene and *para*-benzyne are compiled. As expected, the core orbitals are not involved in the static response as the core electrons are highly confined, therefore these are not going to be further discussed. For each system, we first discuss the key observations for α_{zz} and then for γ_{zzzz}.

In benzene, the overall contributions of the σ- and π-type NOs to α_{zz} are very similar. The largest contributions correspond to the four frontier π-type NOs, namely $1e_{1g}$ (10.5%), $2e_{1g}$ (13.7%), $1e_{2u}$ (9.1%), and $2e_{2u}$ (8.8%). In the case of γ_{zzzz}, the main role of the response shifts to the π NOs. In general, these contribute 60% to γ_{zzzz}, with $2e_{1g}$ (11.4%) and $1e_{2u}$ (8.0%) having the highest individual contributions. Moreover, one can find the opposite (negative) contributions coming from other NOs.

Table 7.1: PNOC contributions of selected NOs to longitudinal α_{zz} and γ_{zzzz} of benzene, C_6H_6, and *para*-benzyne, C_6H_4, obtained at the reference (U)CCSD/aug–cc–pVDZ level of theory. Values are given in the atomic units (relative contributions to the total property are given in parenthesis).

	C_6H_6		C_6H_4	
Orbitals	$\alpha_{zz,p}$	$\gamma_{zzzz,p}$	$\alpha_{zz,p}$	$\gamma_{zzzz,p}$
$6e_{1u}$ ($5b_{1u}$)	4.36 (5.5%)	-373 (-3.0%)	7.83(10.5%)	8536 (29.0%)
$6e_{2g}$ ($6a_g$)	4.57 (5.7%)	-77 (-0.6%)	4.48 (6.0%)	5112 (17.4%)
$1a_{2u}$ ($1b_{3u}$)	0.66 (0.8%)	186 (1.5%)	0.64 (0.9%)	-80 (-0.3%)
$1e_{1g}$ ($1b_{1g}$)	8.39(10.5%)	-411 (-3.4%)	8.78(11.8%)	-1387 (-4.7%)
$2e_{1g}$ ($1b_{2g}$)	10.94(13.7%)	1402(11.4%)	7.40 (9.9%)	-540 (-1.8%)
$1e_{2u}$ ($2b_{3u}$)	7.25 (9.1%)	978 (8.0%)	4.58 (6.1%)	-489 (-1.7%)
$2e_{2u}$ ($1a_u$)	7.06 (8.8%)	-343 (-2.8%)	6.88 (9.2%)	-1120 (-3.8%)
$1b_{2g}$ ($2b_{2g}$)	0.26 (0.3%)	451 (3.7%)	0.20 (0.3%)	103 (0.4%)
$1s_c$	0.01 (0.0%)	0 (0.0%)	0.01 (0.0%)	-2 (0.0%)
σ(occ)	18.04(22.6%)	-1153 (-9.4%)	24.01(32.2%)	12602 (42.9%)
σ(vir)	20.68(25.9%)	6018(49.1%)	16.18(21.7%)	17879 (60.8%)
σ(all)	38.72(48.5%)	4865(39.7%)	40.19(53.9%)	30481(103.7%)
π(occ)	19.99(25.0%)	1177 (9.6%)	16.81(22.5%)	-2007 (-6.8%)
π(vir)	21.12(26.5%)	6213(50.7%)	17.57(23.6%)	932 (3.2%)
π(all)	41.11(51.5%)	7390(60.3%)	34.39(46.1%)	-1076 (-3.7%)
Total NLOP				
CCSD	79.84	12255	74.58	29403
CCSD(T)	80.39	13180	76.70	37152

In *para*-benzyne, formation of the diradical open-shell electronic structure leads to huge changes to the static electronic response. Longitudinal α_{zz} decreases by 5.3 au for UCCSD, but the PNOC analysis reveals a shift of the major contribution to the σ-type of NOs, which constitute now 53% of the property. In particular, one observes a large increase in the absolute contributions for the singly occupied $5b_{1u}$ and $6a_g$ NOs. On the contrary, contributions of the π-type NOs are decreased being the largest decrease reported for the frontier π-type NOs, from 10.9 to 7.4 for $1e_{1g} \rightarrow 1b_{2g}$, and from 7.2 to 4.6 for $1e_{2u} \rightarrow 2b_{3u}$. Equivalent but much more pronounced changes are observed for γ_{zzzz}.

7.5 Role of Electron Correlation in Static NLOPs

We discuss the role of static and dynamic electron correlation in the static response by comparison of the PNOC contributions obtained with (U)HF, (U)MP2, CASSCF, and (U)CCSD. Each of these methods is considered the main representative of a group of quantum mechanical methods that introduce electron correlation in a different manner: HF – no correlation, CASSCF – static correlation, MP2 and CCSD – dynamic correlation. Moreover, breaking of the spin symmetry in HF (and later in MP2) allows to partially deal with the problems arising from the absolute near-degeneracy of the HOMO and LUMO orbitals. This phenomenon has been labeled as the type A static correlation [260]. Results of the decomposition are compiled in Tables 7.2 and 7.3 and the ones obtained with (U)CCSD (given in Table 7.1) are considered as the reference.

Table 7.2: PNOC contributions to longitudinal α_{zz} (in au) of benzene, C_6H_6, and *para*-benzyne, C_6H_4, obtained at different levels of theory. For the computations of benzene, HF, MP2, and CASSCF adopting the (6,6) active space, CAS(6,6), were used. For *para*-benzyne unrestricted variants UHF, UMP2 and CASSCF adopting the (8,8) active space, CAS(8,8), were used.

	C_6H_6			C_6H_4		
Orbitals	HF	MP2	CAS(6,6)	UHF	UMP2	CAS(8,8)
$6e_{1u}$ ($5b_{1u}$)	4.43	4.31	4.58	4.67	6.56	6.70
$6e_{2g}$ ($6a_g$)	4.67	4.21	4.93	2.70	3.38	3.76
$1a_{2u}$ ($1b_{3u}$)	0.38	0.74	0.81	1.06	0.29	0.84
$1e_{1g}$ ($1b_{1g}$)	9.50	8.60	7.44	6.93	12.96	7.30
$2e_{1g}$ ($1b_{2g}$)	12.07	11.24	10.01	7.05	9.13	6.91
$1e_{2u}$ ($2b_{3u}$)	8.42	7.59	6.72	4.17	6.02	4.18
$2e_{2u}$ ($1a_u$)	8.42	7.29	6.63	6.60	11.58	6.51
$1b_{2g}$ ($2b_{2g}$)	0.00	0.38	0.25	0.12	0.01	0.19
$1s_c$	0.01	0.01	0.07	0.01	0.01	0.01
σ(occ)	18.08	18.41	19.00	19.97	21.12	22.65
σ(vir)	18.09	20.84	19.01	14.55	14.72	14.59
σ(all)	36.18	39.25	38.01	34.52	35.84	37.24
π(occ)	21.96	20.58	18.26	15.03	22.39	15.04
π(vir)	21.96	22.04	18.43	14.90	22.22	14.98
π(all)	43.92	42.62	36.68	29.93	44.60	30.02
Total α_{zz}	80.10	81.89	74.70	64.47	80.45	67.27

Looking simply at the total values of α_{zz}, one could reach a conclusion about the performance of a given method. In benzene, α_{zz} obtained with HF is very similar to the CCSD one. MP2 yields a value overestimated by 2.0 au, and CASSCF underestimated it by 5.1 au. This suggests that HF describes the linear response better than CASSCF, which in principle is a much more superior method and yields energies variationally closer to the exact ones. With a tool like PNOC, one can compare not only the total values but also the NO contributions to NLOP, which provide a more detailed analysis of the performance of the different methods. PNOC reveals that the small overall error in HF α_{zz} comes from a fortunate compensation of errors — an underestimation of the contributions of σ-type NOs (by 2.5 au) and an overestimation of the contribution of the π-type NOs (by 2.8 au). In CASSCF, one can observe a systematic underestimation in the contributions of the π-type NOs. In MP2, the orbital-wise contributions to α_{zz} closely follow the reference trends. In *para*-benzyne, for which the static correlation comes into play, the performance of these methods changes. UHF underestimates the total α_{zz} by 10.1 au, UMP2 overestimates it by 5.9 au, and CASSCF underestimates it by 7.3 au, which could lead to the conclusion that UMP2 gives the best description of this linear response property. However, when one examines the NO contributions one can observe that it is CASSCF which captures the changes due to the formation of the diradical. Namely, CASSCF correctly predicts a dominance of contributions of the σ-type orbitals - σ(all) and π(all) are 55% and 45% of total α_{zz}, respectively. In contrast, in UMP2 the response is incorrectly dominated by the π-type NOs - σ(all) and π(all) are 45% and 55% of total α_{zz}. Moreover, UMP2 significantly overestimates the relative contributions of $1b_{1g}$ and $1a_u$ NOs.

Table 7.3: PNOC contributions to the longitudinal γ_{zzzz} (in au) of benzene, C_6H_6, and *para*-benzyne, C_6H_4, obtained at different levels of theory. For the computations of benzene, HF and MP2 and CASSCF adopting the (6,6) active space, CAS(6,6), were used. For *para*-benzyne unrestricted variants UHF, UMP2 and CASSCF adopting the (8,8) active space, CAS(8,8), were used.

	C_6H_6			C_6H_4		
Orbitals	HF	MP2	CAS(6,6)	UHF	UMP2	CAS(8,8)
$6e_{1u}$ ($5b_{1u}$)	-377	-338	-377	733	2796	6544
$6e_{2g}$ ($6a_g$)	-157	75	-229	-288	364	3683
$1a_{2u}$ ($1b_{3u}$)	311	119	269	45	-558	103
$1e_{1g}$ ($1b_{1g}$)	-1300	-493	486	1983	2109	43
$2e_{1g}$ ($1b_{2g}$)	1555	686	1706	931	1523	-41
$1e_{2u}$ ($2b_{3u}$)	3475	296	1363	843	1694	-65
$2e_{2u}$ ($1a_u$)	-396	-375	411	1688	1866	22
$1b_{2g}$ ($2b_{2g}$)	411	423	391	154	211	137
$1s_c$	0	0	0	-1	-1	-2
σ(occ)	-1018	-987	-1246	35	2457	9113
σ(vir)	4305	6739	4327	7559	13542	10140
σ(all)	3287	5752	3080	7594	15999	19253
π(occ)	567	311	2461	2959	3074	105
π(vir)	5352	6089	5488	4610	7612	1634
π(all)	5919	6400	7948	7568	10686	1739
Total γ_{zzzz}	9206	12152	11028	15162	26683	20990

The differences are even more pronounced for the description of γ_{zzzz}. In C_6H_6, compared to the reference value obtained with CCSD, γ_{zzzz} = 13180 au, HF, MP2 and CASSCF underestimate this value by 3974, 1028 and 2152 au, respectively. The decomposed values of γ_{zzzz} support the superior performance of MP2, which for benzene provides a well-balanced description of response coming from the σ- and π-type NOs. On the contrary, CASSCF(6,6) strongly underestimates the σ(occ) contribution and yields overestimated individual contributions of the frontier π NOs, especially $1e_{1g}$ and $2e_{2u}$ which are positive, but in fact, should be negative. Upon the formation of the diradical C_6H_4, the errors in the second hyperpolarizability are enhanced. Compared to UCCSD, for which γ_{zzzz} is 29403 au, UMP2 seems to provide the best total value of 26683 au, while UHF and CASSCF (8,8) produce underestimated values of 15162 and 20990 au, respectively. Interestingly, PNOC's detailed analysis provides the opposite conclusions. First, the PNOC analysis unveils that a small error of UMP2 in total γ_{zzzz} of C_6H_4 is accidental and comes from the cancellation of large errors. Namely, UMP2 provides a very underestimated contribution of the σ-type NOs, σ(all) = 15999 au against 30481 au obtained with UCCSD. Furthermore, UMP2 strongly overestimates the contribution of π-type NOs, π (all) = 10686 au versus the reference value -1076 au. Orbital-wise, UMP2 hugely underestimates the contribution of $5b_{1u}$ and $6a_g$ NOs (the only ones occupied by radical electrons) and greatly overestimates the contributions of all valence π NOs. Even worse errors are found in the orbital contributions described with UHF. In contrast, for CASSCF(8,8), one observes a strong but systematic underestimation of orbital contributions, especially of σ-type NOs. However, CASSCF(8,8) correctly describes the key role of $5b_{1u}$ and $6a_g$ in static γ_{zzzz}, and the negligible role of valence π -NOs.

Concluding this section, the inefficiencies due to a lack of electronic correlation can be examined with the PNOC contributions. Moreover, with PNOC one can detect cases where good performance is only accidental due to the cancellation of errors. These two scenarios are especially reflected in the contributions to higher-order nonlinear response of γ_{zzzz}. When the static correlation is not relevant, MP2 provides a reasonable description of the orbital nature of the response. In such cases, the CASSCF calculations including only a few valence orbitals are insufficient because of the unbalanced treatment of the correlation. Namely, one can observe the correct trends of the orbital contributions to γ_{zzzz} but with the underestimated values. Situation changes when the static correlation plays the main role (like in C_6H_4). For the accurate calculation γ_{zzzz}, UMP2 provides an incorrect orbital-based picture of γ_{zzzz} and one has to account for the static correlation via multiple determinants, for example, by using the CASSCF method.

7.6 Assessing the Role of Basis Functions

PNOC can be also used to assess the role of the particular type of atomic basis functions in the NLOP computations. In this section, we study the influence of diffuse functions of different angular momentum in simulations of static NLOPs. In particular, we focus on the longitudinal component of the second hyperpolarizability, γ_{zzzz}, of π-conjugated system, all-*trans*-hexatriene, C_6H_8. Starting from the cc-pVDZ/cc-pVDZ for carbon/hydrogen basis set (also abbreviated here as VDZ/VDZ), diffuse functions of selected angular momentum, *s*, *p* or *d* type are separately added to the Carbon (C) and/or Hydrogen (H) centers (original primitive exponents are utilized). The addition of diffuse functions for selected elements is denoted with a + sign. We tested the following basis sets, where the first basis corresponds to C and the second to H (the total number of basis functions is given in parenthesis): VDZ/VDZ (124), VDZ+*s*/VDZ (130), VDZ+*p*/VDZ (142), VDZ+*d*/VDZ (154), VDZ+*sp*/VDZ (148), VDZ+*p*/VDZ+*p* (166), AVDZ/VDZ (178), VDZ/AVDZ (156) and AVDZ/AVDZ (210). In such notation, standard AVDZ is equivalent to VDZ+*spd* for C and to VDZ+*sp* for H.

The results of total and decomposed NLOPs are compiled in Table 7.4. It is well known that diffuse functions are essential for the computation of high-order nonlinear response, whereas non-augmented basis sets tend to greatly underestimate these properties. Full augmentation from VDZ/VDZ to AVDZ/AVDZ leads to a 48% increase of γ_{zzzz}.

PNOC reveals that these large changes in γ_{zzzz} are mainly due to the increased response of the orbitals from the π(HV) subspace. In particular, an increase in the π(HV) contribution accounts for the 64% of the increase of the total γ_{zzzz}. PNOC also reveals that the addition of diffuse functions to carbon atoms is much more important than to the hydrogen centers. Using the AVDZ/VDZ basis, one can already reach very good correspondence to the ones of AVDZ/AVDZ for all orbital components, namely σ(FV), σ(HF), π(FV), and π(HV). On the other hand, if one added the diffuse functions solely to the hydrogen centers it would lead to the incorrect response of the π-type NOs. Among the set of diffuse functions of C, the *p*-type diffuse functions seem to be the most important ones, while the *s*-type is the least important. For γ_{zzzz}, the addition of *d*-type diffuse functions also marginally affects the response, and it is through a slight increase in the contribution of π(HV). Since diffuse functions of *p* angular momentum in carbon are crucial for γ_{zzzz}, the effect of further addition of *p* diffuse functions to hydrogen centers is also examined. The VDZ+p/VDZ+p basis set has substantially better performance than VDZ+p/VDZ and yields results comparable to those of AVDZ/VDZ. PNOC reveals that it correctly describes both the FV and the HV components of the response.

Table 7.4: Basis set dependency of the PNOC contributions to longitudinal γ_{zzzz} of hexatriene. Decomposed values were obtained at the CCSD level of theory. Additionally, CCSD(T) total values are given.

C: H:	VDZ VDZ	VDZ+s VDZ	VDZ+p VDZ	VDZ+d VDZ	VDZ+sp VDZ	VDZ+p VDZ+p	AVDZ VDZ	VDZ AVDZ	AVDZ AVDZ
σ(FV)	-11175	-11591	-14565	-12884	-14905	-15929	-15725	-14256	-16120
σ(HV)	4006	5105	6368	5196	7349	6941	7352	7887	8057
σ(all)	-7169	-6486	-8198	-7688	-7557	-8988	-8373	-6369	-8063
π(FV)	109754	110375	130891	111891	131418	131879	128768	119306	129463
π(HV)	5756	5779	34215	11305	34681	39289	37697	16528	39135
π(all)	115511	116154	165106	123195	166099	171169	166465	135834	168598
Sum(FV)	98579	98784	116326	99006	116513	115950	113043	105050	113343
Sum(HV)	9762	10883	40583	16501	42029	46231	45049	24415	47192
Total γ_{zzzz}									
CCSD	108342	109668	156909	115507	158542	162181	158092	129465	160535
CCSD(T)	104299	105860	155439	112562	157482	161993	158102	126938	161292

7.7 Local Representation of Static NLOPs

PNOC scheme can also provide a real-space representation of the NLOPs using the partitioned values of the properties. One can define a local version of a $\alpha_{zz}(\mathbf{r})$ and $\gamma_{zzzz}(\mathbf{r})$ through multiplication of each NO contribution to the property by its NO real-space amplitude:

$$\alpha_{zz}(\mathbf{r}) = \sum_p \alpha_{zz,p} \left|\phi_p^{\mathrm{NO}}(\mathbf{r})\right|^2 \tag{7.20}$$

$$\gamma_{zzzz}(\mathbf{r}) = \sum_p \gamma_{zzzz,p} \left|\phi_p^{\mathrm{NO}}(\mathbf{r})\right|^2 \tag{7.21}$$

This definition assumes a uniform contribution of NO to the property. Because NOs are orthonormal, integration of $\alpha_{zz}(\mathbf{r})$ and $\gamma_{zzzz}(\mathbf{r})$ retrieves the total values of α_{zz} and γ_{zzzz}:

$$\alpha_{zz} = \int \mathrm{d}\mathbf{r}\alpha_{zz}(\mathbf{r}) \tag{7.22}$$

$$\gamma_{zzzz} = \int \mathrm{d}\mathbf{r}\gamma_{zzzz}(\mathbf{r}) \tag{7.23}$$

On the contrary, the total values of the properties are not recovered by the integration of density derivatives $\rho^{(1)}(\mathbf{r})$ and $\rho^{(3)}(\mathbf{r})$ which always integrate to zero.

For the two systems studied, benzene and *para*-benzyne, local representations of $\alpha_{zz}(\mathbf{r})$ and $\gamma_{zzzz}(\mathbf{r})$, as well as $\rho^{(z)}(\mathbf{r})$ and $\rho^{(zzz)}(\mathbf{r})$, are shown in Figure 7.4. These were obtained at the reference (U)CCSD/aug-cc-pVDZ level of theory. Both PNOC and density derivatives provide a similar picture, with the former being easier to analyze. To interpret the density derivatives, one has to establish the patterns of regions with positive and negative values, which is much more troublesome to be done for larger chemical systems. In C_6H_6, $\alpha_{zz}(\mathbf{r})$ is almost equally distributed among all carbon atoms. Except for those in the *para* position, hydrogen atoms do not seem to contribute to the response. Picture for $\gamma_{zzzz}(\mathbf{r})$ is different, as the larger local contribution is found for the C1 and C4 atoms (see labels in Figure 7.2). Moreover, the shape of the contribution of each carbon atom resembles the shape of the p-type AO. In contrast, in C_6H_4, both $\alpha_{zz}(\mathbf{r})$ and $\gamma_{zzzz}(\mathbf{r})$ have enhanced values at the position of unpaired electrons. This is especially clear in $\gamma_{zzzz}(\mathbf{r})$, which in shape resembles frontier σ-type NOs $5b_{1u}$ and $6a_g$. Interestingly, one can also observe negative contributions in the shape of p-type AOs at the carbon atoms without unpaired electrons. This originates from the negative contributions of the frontier π-type NOs (see Table 7.1).

Additionally, through the PNOC local representation of NLOPs one can also establish a connection with electron correlation. We compared $\alpha_{zz}(\mathbf{r})$ and $\gamma_{zzzz}(\mathbf{r})$ with the local representation of the nondynamic (static) and dynamic correlation, $I_{nd}(\mathbf{r})$ and $I_d(\mathbf{r})$ [261]. These two indices and their local representations are defined using the occupancies of the NOs, n_i^σ, namely

$$I_{nd}(\mathbf{r}) = \frac{1}{2}\sum_\sigma \sum_p n_i^\sigma (1-n_i^\sigma) \left|\phi_p^{\mathrm{NO}}(\mathbf{r})\right|^2, \tag{7.24}$$

$$I_d(\mathbf{r}) = \frac{1}{4}\sum_\sigma \sum_p \left([n_i^\sigma (1-n_i^\sigma)]^{1/2} - 2n_i^\sigma (1-n_i^\sigma) \right) \left|\phi_p^{\mathrm{NO}}(\mathbf{r})\right|^2. \tag{7.25}$$

It has been shown that $I_{nd}(\mathbf{r})$ and $I_d(\mathbf{r})$ depict regions in space in which a particular type of electronic correlation is important. In comparison with the local representations of NLOPs, in benzene, $\gamma_{zzzz}(\mathbf{r})$ follows the profile of the dynamic correlation index $I_d(\mathbf{r})$.

In contrast, in *para*-benzyne, $\gamma_{zzzz}(\mathbf{r})$ follows the profile of nondynamic correlation index $I_{nd}(\mathbf{r})$. This simple local representation provides a quick depiction of the connection between the static NLOPs and the electronic correlation of different types.

7.8 Conclusions

Partition of NLOPs Into Orbital Contributions decomposes the total values of the linear and nonlinear optical properties into the natural orbitals of the unperturbed system. It can be applied regardless the electronic structure method. It does not pose a large computational cost, since PNOC requires only the information on the field-perturbed first-order reduced density matrix (from which the numerical derivatives are computed). The most important feature of PNOC is its ability to detect flaws in quantum chemical methods to compute NLOPs. It can unveil the compensation of errors of electronic structure methods through the orbital-wise contributions. As shown for *para*-benzyne, while the total value of γ_{zzzz} computed using UMP2 is much closer to the reference values than the one obtained using CASSCF(8,8), the orbital picture of the response is wrong. UMP2 contributions to γ_{zzzz} of the NOs occupied by the radical electrons are incorrect, whereas the ones obtained at CASSCF(8,8) are only systematically underestimated. This directly concludes that UMP2, despite adopting unrestricted formalism, cannot capture the effect of static correlation on the static response. One can also utilize PNOC to assess the role of particular basis functions in the computation of response properties. For example, we have proven that for the second hyperpolarizabilities of π-conjugated systems, the addition of the *p*-type diffuse functions to the basis set is the most beneficial and one does not need to employ *s*- or *d*-type diffuse functions. Moreover, since PNOC shows that the biggest contributions belong to the π-type NOs, one can solely apply these diffuse functions at the carbon atoms. Lastly, without any extra cost, PNOC can be used to obtain a real-space representation of the NLOPs, which can work as a visualization aid to connect the magnitude of the optical property with some parts of the molecule and design of molecules with the tailored NLOPs.

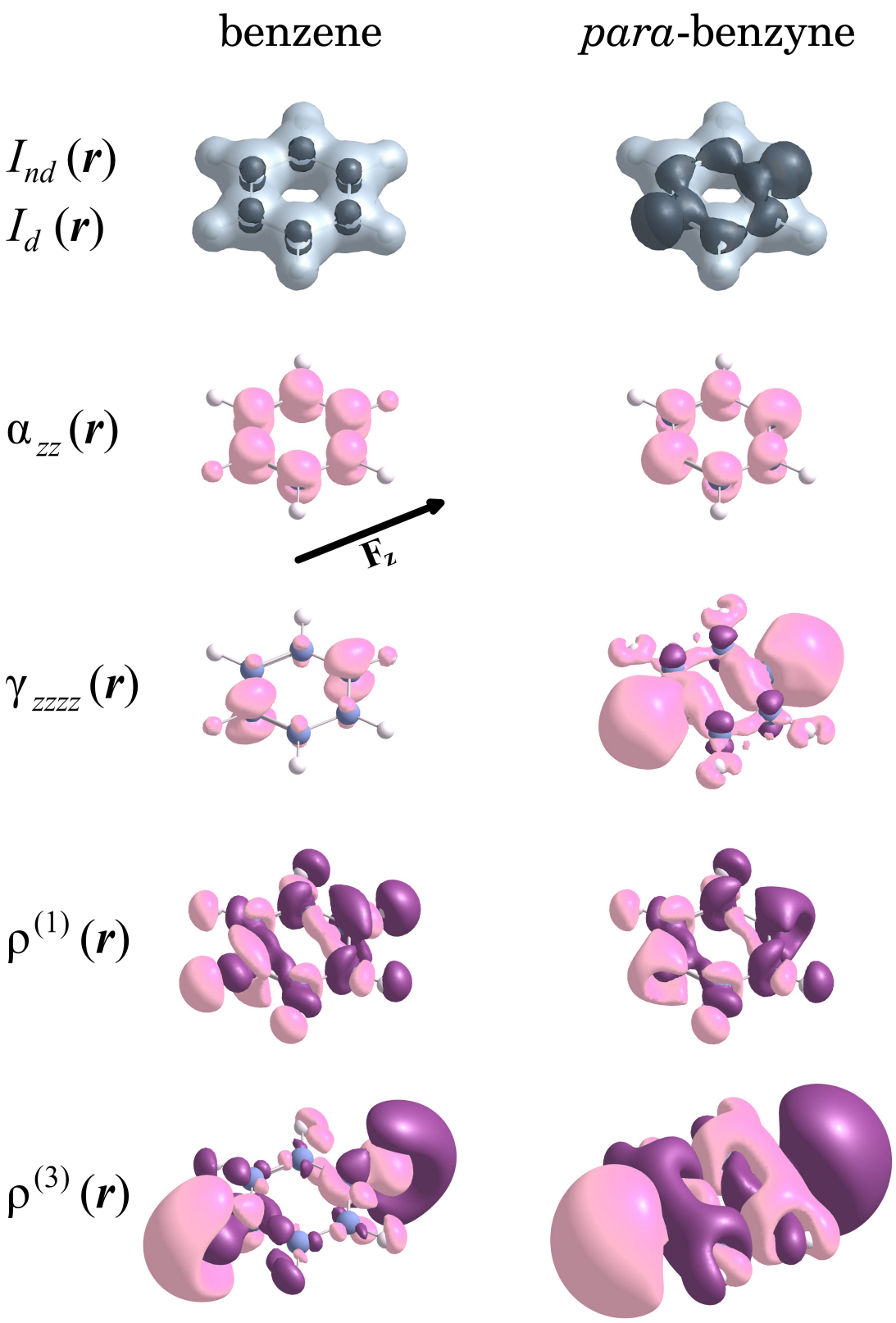

Figure 7.4: Comparison of functions representing the NLOPs in the Cartesian space: $\alpha_{zz}(\mathbf{r})$, $\gamma_{zzzz}(\mathbf{r})$, $\rho^{(z)}(\mathbf{r})$, and $\rho^{(zzz)}(\mathbf{r})$. The last row presents local nondynamic and dynamic correlation indices $I_{nd}(\mathbf{r})$ (light gray) and $I_d(\mathbf{r})$ (dark gray). Left panel depicts functions obtained for benzene, and right panel for *para*-benzyne. Light and dark violet colors represent positive and negative values of some isosurfaces: $\alpha_{zz}(\mathbf{r})$ = 0.40, $\gamma_{zzzz}(\mathbf{r})$ = 40.0, $\rho^{(z)}(\mathbf{r})$ = 0.085, $\rho^{(zzz)}(\mathbf{r})$ = 5.00, and $I_{nd}(\mathbf{r})$ = $I_d(\mathbf{r})$ = 0.006. Obtained at the (U)CCSD/aug-cc-pVDZ level of theory.

Chapter 8

Performance of DFAs in the Simulations of Nuclear Relaxation Contributions to NLOPs

8.1 Introduction

The design of new molecular materials with a large nonlinear optical properties remains a challenging task for computational chemistry. This is due to the necessity of the accurate description of both electronic and vibrational (hyper)polarizabilities of individual molecules as well as an appropriate inclusion of the intermolecular interactions effects in the bulk [262–266]. To accurately simulate the molecular response properties one usually requires high-level quantum-chemical methods and large basis sets [267–271]. However, very often the molecular size becomes a limiting factor for the application of such methods. On the other hand, the interactions with surrounding molecules are by definition long-range and have dynamic and many-body nature [272, 273]. This leads to even stricter requirements on the accuracy of the computational protocols [265, 274]. DFT is a cost-effective method for the simulations of the electronic and vibrational structures of molecules, however its suitability for molecular NLO is still being validated.

As a natural step from the isolated molecules towards more complex systems, the *electronic* NLOPs of various types of weakly-bound systems were intensively studied [64, 65, 95, 96, 272, 273, 275–286]. In particular, it was found that the electronic properties of hydrogen bonded systems can be well described by RSHs and also by some GHs employing a large amount of the exact exchange (e.g. M06-2X) [95]. In the case of the dispersion-bonded systems, these methods also performed well provided that dispersion interaction corrections were applied [285, 286]. The good performance of RSHs could be traced back not only to an improved description of the NLO properties of the individual subsystems, but also to a better description of the *interaction-induced* (*excess*) component of NLO properties involving the effects of intermolecular interactions in the complex [287].

Thus far, only few studies addressed the *vibrational* part of the NLOPs of noncovalently-bonded complexes [280, 287]. One could expect that the low-frequency vibrational modes in such systems corresponding to the intermolecular nuclear motion will have a higher impact on the vibrational contributions to NLOPs than in the isolated molecules. The accurate computation of vibrational NLOPs requires approaches computationally more expensive than the ones for the calculation of their electronic counterparts. In particular, one has to evaluate the equilibrium geometry, vibrational frequencies and electrical and mechanical anharmonic corrections at the same level of theory. For some molecular systems, the nonresonant static vibrational contributions to hyperpolarizability were of the same magnitude (or even larger) than their electronic counterparts [288–290]. Moreover, vibrational hyperpolarizabilities are important for electro-optical phenomena like the Pockels and DC Kerr effects, and for optical processes like the intensity-dependent refractive index [119]. On the contrary, for the second and third harmonic generation, when the optical frequencies are much higher than the vibrational frequencies, the vibrational hyperpolarizabilities become far smaller than their electronic counterparts [291].

The main goal of this Chapter is to benchmark modern DFAs for the simulation of the anharmonic vibrational and electronic (hyper)polarizabilities of the noncovalently-bonded molecular complexes [108]. We provide the answer to which functionals are the most and the least reliable for simulations of these properties. Focusing on the latter ones, by a consecutive decomposition of the NR contributions, we narrowed down which anharmonic BKPT terms and which property derivatives are described incorrectly. Additionally, in a subsequent study, we unveil that the large errors in the high-order property derivatives are due to spurious oscillations of the potential energy along the vibrational coordinates. These spurious oscillations have origin in the numerical integration adopted in the KS-DFT computations. This Chapter, and especially its last two sections, can be considered as a prequel to the main work of this Thesis described in Chapters 9 and 10.

8.2 Systems and Computational Details

In this study, the focus is put on the electric properties of molecular complexes. We study the longitudinal electronic and NR contributions to the (hyper)polarizabilities of the following set of hydrogen-bonded systems: HCN·HCl, HCN·HCN, N_2·HF, OC·HF, HCN·HF, HCN·HNC, HNC·HCN, FCCH·NCF, and HCCH·NCF. All systems are co-linear with the z Cartesian axis, and therefore, we study the diagonal components of NLOPs along the z direction. Unless otherwise indicated, the electronic and vibrational contributions to properties and equilibrium geometries were determined at the very same level of theory.

The NR contributions are calculated using the FF-NR approach [116] (see Section 4.10). The necessary field-dependent optimizations maintain the Eckart conditions and are carried out with the procedure developed by Luis *et al*. [117]. For the numerical differentiation in the FF-NR approach, we employ the Romberg–Rutishauser (RR) scheme [292]. In the RR scheme, the correction of order p to the derivative P is determined iteratively according to the formula:

$$P^{p,k} = \frac{4^p \cdot P^{p-1,k} - P^{p-1,k-1}}{4^p - 1}, \tag{8.1}$$

where k is related to the strength of the field used in the FF-NR calculations. The value of $P^{0,k}$ is determined from the finite–difference expression for the derivative of the energy (E) or dipole moment with respect to the electric field F. The smallest electric field amplitude F_z used in FF-NR was equal to 0.0002 a.u. In all calculations we employ the Romberg–Rutishauser scheme using the following sequence of electric fields $\pm F_z$, $\pm 2F_z$, …, $\pm 2^k F_z$, where k=7. Moreover, all values of NLOPs obtained with the FF-NR procedure were re-checked using the field-induced coordinates [118, 119] and BKPT [110–115] methods.

The DFT calculations were performed employing 10 functionals: BLYP [178, 180], CAM-B3LYP [193], LC-BLYP [190], ωB97X [192], M06 [293], M06-2X [293], MN15 [75], PBE0 [294], and HSE06 [295]. In the KS-DFT computations, the pruned SuperFine-Grid, implemented in Gaussian package, was applied [296]. Moreover, the reference values were determined at the CCSD(T) level of theory [297]. As a reference to benchmark, various anharmonic terms of Eqs. 4.93 to 4.95, we also employed the MP2 method [210]. All the calculations of the properties were performed using the aug-cc-pVTZ basis set[298–305]. Moreover, to achieve the best stability in the numerical differentiation, tight convergence criteria for energy and density were imposed (energy convergence criterion was set to 10^{-12} Hartree). FF-NR, FIC and BKPT calculations were performed using custom computer routines based on the energies (FF-NR method) and property derivatives (FIC and BKPT methods). Quantum chemical computations were done using the Gaussian package (ver. 09 and 16) [259, 296].

8.3 Errors in the Harmonic and Anharmonic Nuclear Relaxation Contributions

We start the discussion from the overall performance of DFAs in the estimations of both electronic and vibrational contributions to NLOPs. Table 8.3 compiles the average absolute relative errors in these properties computed with respect to the reference values obtained with CCSD(T). All of the tested functionals accurately describe electronic polarizability, α_{zz}^{el}, yielding average errors from 0.8% (CAM-B3LYP) to 8% (BLYP). In the case of electronic nonlinear response, the average errors are much larger — for β_{zzz}^{el}, they range from 13.4% (M06-2X) to 118.1% (BLYP), and for γ_{zzzz}^{el} they range from 4.0% (MN15) to 128% (BLYP). Such unsatisfactory performance of the BLYP functional is not unexpected, since it has been shown that GGA functionals struggle to predict accurate values of electronic (hyper)polarizabilities [87, 88, 306]. On the other hand, the smallest errors are achieved with RSHs, such as CAM-B3LYP, and GHs with a high percentage of the exact exchange, such as M06-2X and MN15.

In comparison to the electronic part, one observes larger average errors for $\alpha_{zz}^{\mathrm{NR}}$. The smallest average error in $\alpha_{zz}^{\mathrm{NR}}$ is 3.6% (ωB97X), and the largest one 31.4% (BLYP). In the description of $\beta_{zzz}^{\mathrm{NR}}$, all DFAs, except CAM-B3LYP, LC-BLYP, and MN15, yield average errors larger than 20%. Most notably, M06-2X, M06, and ωB97X yield large average errors of 50.5%, 56.0%, and 85.6%, respectively. The performance of these three DFAs is even worse for $\gamma_{zzzz}^{\mathrm{NR}}$. While other DFAs yield average errors for $\gamma_{zzzz}^{\mathrm{NR}}$ in the range 13.6% (CAM-B3LYP) - 30.3% (BLYP), M06-2X, M06, and ωB97X functionals yield average errors exceeding hundreds and thousands of percent.

Table 8.1: The average absolute relative errors (in percentages) in the longitudinal electronic and nuclear-relaxation (hyper)polarizabilities of molecular complexes. Relative errors were calculated with respect to the reference values calculated at the CCSD(T)/aug-cc-pVTZ level of theory.

	$\alpha_{zz}^{\mathrm{el}}$	$\beta_{zzz}^{\mathrm{el}}$	$\gamma_{zzzz}^{\mathrm{el}}$	$\alpha_{zz}^{\mathrm{NR}}$	$\beta_{zzz}^{\mathrm{NR}}$	$\gamma_{zzzz}^{\mathrm{NR}}$
MP2	0.1	27.9	5.4	2.4	8.7	10.0
BLYP	8.0	118.1	128.2	31.4	36.0	30.3
B3LYP	2.0	33.1	37.1	19.8	19.5	16.5
CAM-B3LYP	0.8	19.7	6.9	14.2	13.7	13.6
LC-BLYP	2.1	24.0	18.4	10.8	15.7	15.4
ωB97X	0.8	24.1	5.8	3.6	85.6	1767.6
M06	1.8	48.1	32.5	13.4	56.0	1072.1
M06-2X	1.2	13.4	7.6	7.4	50.5	891.0
MN15	0.9	27.4	4.0	7.6	15.2	14.3
PBE0	0.9	32.1	24.2	25.6	30.9	24.7
HSE06	1.0	35.7	27.2	24.3	29.0	27.5

To investigate the possible source of these unusually high errors in $\beta_{zzz}^{\mathrm{NR}}$ and $\gamma_{zzzz}^{\mathrm{NR}}$, we further partition the NR contributions to the first and second hyperpolarizabilities into the harmonic and anharmonic terms. The harmonic contributions correspond to the 0,0 square bracket terms of Eqs. 4.123 and 4.124, and the anharmonic corrections constitute for the remaining terms in these equations.

For this analysis, we use MP2 as the reference method, because the computations of explicit harmonic and anharmonic terms to $\gamma_{zzzz}^{\mathrm{NR}}$ at the CCSD(T) level using the BKPT method require massive computational resources. The accuracy of MP2 method in predicting the NR properties is very good. Namely, the MP2 average relative errors for $\alpha_{zz}^{\mathrm{NR}}$, $\beta_{zzz}^{\mathrm{NR}}$ and $\gamma_{zzzz}^{\mathrm{NR}}$ are 2.37%, 8.72% and 9.95%, respectively. It can thus be concluded that

MP2 can be used as a reference method to determine which harmonic or anharmonic terms are the origin of the poor performance of some of the DFAs.

Before the discussion on the DFAs errors, let us first comment on the actual character of the NR response of these chemical systems. Table 8.2 presents the percentage contribution of harmonic and anharmonic terms obtained at the MP2/aug-cc-pVTZ. As observed, the anharmonicity plays a crucial role and its effects cannot be neglected when studying the vibrational part of the response property. When described at the MP2/aug-cc-pVTZ level of theory, the anharmonicity in $\beta^{\mathrm{NR}}_{zzz}$ can account for 7% ($N_2 \cdot$HF) of a total value of $\beta^{\mathrm{NR}}_{zzz}$, and may reach up to 50% (HCN$\cdot$HF). In the case of a second hyperpolarizability, the anharmonic corrections become more important and prevail over the harmonic terms. Except HCCH$\cdot$NCF (34%) and FCCH$\cdot$NCF (49.63%) complexes, the anharmonicity contributions exceed 50% of the total $\gamma^{\mathrm{NR}}_{zzzz}$ value and reach up to 87% (HCN$\cdot$HCl).

Table 8.2: Percentage weight of net harmonic and anharmonic contributions to nuclear-relaxation first and second hyperpolarizabilities computed at the MP2/aug-cc-pVTZ level of theory.

System	$\beta^{\mathrm{NR}}_{zzz}$		$\gamma^{\mathrm{NR}}_{zzzz}$	
	[%] harmonic	[%] anharmonic	[%] harmonic	[%] anharmonic
HCN$\cdot$HF	50.1	49.9	26.3	73.8
HCN$\cdot$HCl	51.1	49.0	13.0	87.0
OC$\cdot$HF	64.1	36.0	38.9	61.1
$N_2\cdot$HF	93.0	7.0	37.0	63.0
HCN$\cdot$HNC	55.7	44.3	21.9	78.1
HNC$\cdot$HCN	71.4	28.6	47.0	53.0
HCN$\cdot$HCN	71.8	28.2	42.5	57.5
HCCH$\cdot$NCF	81.6	18.5	65.5	34.5
FCCH$\cdot$NCF	76.3	23.7	50.4	49.6

Except for the ωB97X, M06, and M06-2X, all functionals show consistent performance in the computations of both the harmonic and anharmonic contributions to $\beta^{\mathrm{NR}}_{zzz}$ and $\gamma^{\mathrm{NR}}_{zzzz}$. For example, the CAM-B3LYP functional yields average errors of 11.1% and 5.6% for $\beta^{\mathrm{NR,har}}_{zzz}$ and $\beta^{\mathrm{NR,anh}}_{zzz}$, and 3.1% and 7.0% for $\gamma^{\mathrm{NR,har}}_{zzzz}$ and $\gamma^{\mathrm{NR,anh}}_{zzzz}$. In contrast, M06-2X, M06, and ωB97X provide a rather good description of the harmonic contributions to $\beta^{\mathrm{NR}}_{zzz}$ (errors in the range 7.3% - 21.1%) and $\gamma^{\mathrm{NR}}_{zzzz}$ (errors in the range 3.1% - 7.9%), but fail dramatically in the estimation of the anharmonic contributions. Most notably, one can observe huge average errors in the $\gamma^{\mathrm{NR,anh}}_{zzzz}$, again reaching thousands of percent.

In order to further pinpoint the source of these huge errors in $\gamma^{\mathrm{NR,anh}}_{zzzz}$, we investigate each of the anharmonic square bracket term from Eq. 4.124. That is, we analyze the average absolute relative errors for $[\mu^2\alpha]^{1,0}_{\omega=0}$, $[\mu^2\alpha]^{0,1}_{\omega=0}$, $[\mu^4]^{2,0}_{\omega=0}$, $[\mu^4]^{0,2}_{\omega=0}$, and $[\mu^4]^{1,1}_{\omega=0}$ (again, relative errors measured with respect to the values obtained with MP2/aug-cc-pVTZ). Average errors in these components obtained by the DFAs in question (and for CAM-B3LYP) are compiled in Table 8.4. Huge discrepancies are observed for the $[\mu^4]^{2,0}_{\omega=0}$ and $[\mu^4]^{0,2}_{\omega=0}$ terms. Most notably, the largest average errors are observed for ωB97X and the $[\mu^4]^{2,0}_{\omega=0}$ (635.6%) and $[\mu^4]^{0,2}_{\omega=0}$ (871.4%). In contrast, for CAM-B3LYP the average errors in these properties are about ten times smaller.

Within the BKPT formalism, the anharmonic $[\mu^4]_{\omega=0}$ terms involve the high-order derivatives of total energy and dipole moment with respect to the normal modes (or equivalently, FICs). As described in Section 4.9 (see Eqs. 4.121 and 4.122-4.124), $[\mu^4]^{2,0}_{\omega=0}$,

the second-order electrical anharmonicity, is defined as

$$[\mu^4]^{2,0}_{\omega=0} = \sum_{a,b,c}^{3N-6} \left(a_{31}^{abc,\alpha} q_1^{a,\beta} q_1^{b,\gamma} q_1^{c,\delta} + \frac{a_{21}^{ab,\alpha} a_{21}^{bc,\beta}}{a_{20}^{bb}} q_1^{a,\gamma} q_1^{c,\delta} \right), \tag{8.2}$$

and, in particular, involves the second ($a_{21}^{ab,\alpha}$) and third ($a_{31}^{abc,\alpha}$) derivatives of dipole moment (see notation defined in Eq. 4.114). $[\mu^4]^{0,2}_{\omega=0}$, which corresponds to the second-order mechanical anharmonicity, is defined as

$$[\mu^4]^{0,2}_{\omega=0} = \sum_{a,b,c,d}^{3N-6} \left(-a_{40}^{abcd} q_1^{a,\alpha} q_1^{b,\beta} q_1^{c,\gamma} q_1^{d,\delta} + \sum_{e}^{3N-6} \frac{9 a_{30}^{abc} a_{30}^{cde}}{4 a_{20}^{cc}} q_1^{a,\alpha} q_1^{b,\beta} q_1^{d,\gamma} q_1^{e,\delta} \right), \tag{8.3}$$

and, in particular, involves the third (a_{30}^{abc}) and fourth (a_{40}^{abcd}) derivatives of energy. A simple test allowed to preliminarily establish that the problem lies almost exclusively in the high-order derivatives, such as a_{30}^{abc} and a_{40}^{abcd}. Namely, we computed $[\mu^4]^{0,2}_{\omega=0}$ by combining the values of a_{30}^{abc} and a_{40}^{abcd} obtained with MP2 and the values of $q_1^{a,\beta}$ obtained with ωB97X. Such values of $[\mu^4]^{0,2}_{\omega=0}$ are characterized by much smaller errors.

Table 8.3: The average absolute relative errors in harmonic, $\beta_{zzz}^{\mathrm{NR,har}}$ and $\gamma_{zzzz}^{\mathrm{NR,har}}$, and anharmonic, $\beta_{zzz}^{\mathrm{NR,anh}}$ and $\gamma_{zzzz}^{\mathrm{NR,anh}}$, NR contributions for a set of 9 molecular complexes. Relative errors were calculated with respect to the values calculated at the MP2/aug-cc-pVTZ level of theory.

	$\beta_{zzz}^{\mathrm{NR,har}}$	$\beta_{zzz}^{\mathrm{NR,anh}}$	$\gamma_{zzzz}^{\mathrm{NR,har}}$	$\gamma_{zzzz}^{\mathrm{NR,anh}}$
BLYP	49.5	24.1	25.2	14.8
B3LYP	22.5	14.1	8.3	13.1
CAM-B3LYP	11.1	5.6	3.1	7.0
LC-BLYP	5.7	7.1	4.9	8.1
ωB97X	7.3	76.9	3.1	1630.3
M06	21.1	61.5	7.9	1059.0
M06-2X	12.3	48.8	3.2	836.5
MN15	9.2	9.6	3.1	18.5
PBE0	25.1	12.6	7.9	12.3
HSE06	25.2	12.9	8.4	19.6

Table 8.4: Average absolute relative error in predicting anharmonic BKPT contributions to $\gamma_{zzzz}^{\mathrm{NR}}$ for a set of 9 molecular complexes. Relative errors were calculated with respect to the values of the same contributions calculated at the MP2/aug-cc-pVTZ level of theory.

	CAM-B3LYP	ωB97X	M06	M06-2X
$[\mu^2\alpha]^{1,0}_{\omega=0}$	19.0	154.9	87.7	72.1
$[\mu^2\alpha]^{0,1}_{\omega=0}$	24.8	40.7	84.8	52.1
$[\mu^4]^{1,1}_{\omega=0}$	123.6	135.5	341.0	165.4
$[\mu^4]^{2,0}_{\omega=0}$	58.5	635.6	459.7	250.4
$[\mu^4]^{0,2}_{\omega=0}$	79.0	871.4	429.5	477.2

8.4 Grid-Related Spurious Oscillations in the Energy Derivatives Curves

This poor performance in computation of high-order derivatives of molecular properties motivated us to further investigate this problem. One can get more insight into the observed (erroneous) values of the molecular properties by establishing their dependence along a particular (geometrical) coordinate.

In this test, we are interested in the errors in $[\mu^4]^{0,2}_{\omega=0}$ arising from the high-order derivatives a^{abc}_{30} and a^{abcd}_{40}. To reduce the dimensionality of the problem, we switched to the FIC representation of these properties. In fact, to evaluate the first summation in Eq. 8.3, one can utilize solely the z-component of FIC1, $\chi_{1,z}$, which drastically reduces the dimensionality of the problem since only the diagonal derivatives with respect to $\chi_{1,z}$ are needed. In contrast, the second summation in 8.3 requires both z-component of FIC1, $\chi_{1,z}$ and *harmonic* part of FIC2, $\chi_{2,har,zz}$. In this preliminary analysis, we solely focus on the $\mathrm{d}^m E/\mathrm{d}{\chi_{1,z}}^m$ derivatives. For that reason, we compute the curves of $\mathrm{d}^m E/\mathrm{d}{\chi_{1,z}}^m$ (m = 0 - 4) for different geometrical displacements along the $\chi_{1,z}$ coordinate of the HCN$\cdot$ HF complex. We obtain those for three DFAs, namely B3LYP, ωB97X, and M06-2X, combined with the unpruned version of SuperFineGrid, i.e., (250, 974) grid, as well for the CCSD(T) method.

Results for the $\mathrm{d}^m E/\mathrm{d}{\chi_{1,z}}^m$ derivatives are depicted in Figure 8.1 and in these plots $\Delta\chi_{1,z}$ = 0 corresponds to the optimized geometry. While one does not observe any dramatic differences in the curves of $E(R)$, $\mathrm{d}E/\mathrm{d}\chi_{1,z}$, and $\mathrm{d}^2E/\mathrm{d}{\chi_{1,z}}^2$, interesting spurious oscillations emerged in the curves of $\mathrm{d}^3E/\mathrm{d}{\chi_{1,z}}^3$ and $\mathrm{d}^4E/\mathrm{d}{\chi_{1,z}}^4$ obtained with M06-2X and ωB97X. These oscillations introduce large inconsistencies in the values of the properties. In the previous works, similar spurious oscillations were reported in PECs along the interatomic separation R in Ar_2 (the $E(R)$ curve) and were attributed to an insufficient size of the grid used in the numerical integration [307, 308].

For that reason, we evaluate $\mathrm{d}^4E/\mathrm{d}{\chi_{1,z}}^4$ with several integration grids, namely (99, 590), SuperFineGrid and its unpruned version (250, 974), and (750, 974) (see nomenclature of grid sizes in Section 5.6.2). These results are depicted in the last row in Figure 8.1. No significant differences are observed for B3LYP when combined with different integration grids. In contrast, one notices big changes in the curves for M06-2X and ωB97X obtained with different grids. Most notably, for these DFAs, the oscillations have much larger amplitudes when the smaller grid, (99, 590) is used. In the case of the (250, 974) grid, oscillations are much smaller but still considerably large. It seems that the oscillations can be alleviated by using a huge integration grid, such as (750, 974).

The values of $\mathrm{d}^4E/\mathrm{d}{\chi_{1,z}}^4$ at the optimized geometries, for the (99, 590), SuperFineGrid, (250, 974), and (750, 974) grids are respectively: $8.10\cdot10^{-4}$, $8.07\cdot10^{-4}$, $8.01\cdot10^{-4}$, and $8.00\cdot10^{-4}$ for B3LYP; $1.34\cdot10^{-3}$, $-9.20\cdot10^{-5}$, $2.52\cdot10^{-4}$, and $4.08\cdot10^{-4}$ for M06-2X; $6.72\cdot10^{-3}$, $7.76\cdot10^{-3}$, $1.92\cdot10^{-4}$, and $7.96\cdot10^{-4}$ for ωB97X. Therefore, for M06-2X and ωB97X combined with the SuperFineGrid, one observes the relative errors in $\mathrm{d}^4E/\mathrm{d}{\chi_{1,z}}^4$ due to the insufficient integration grid equal to -122% and 875%.

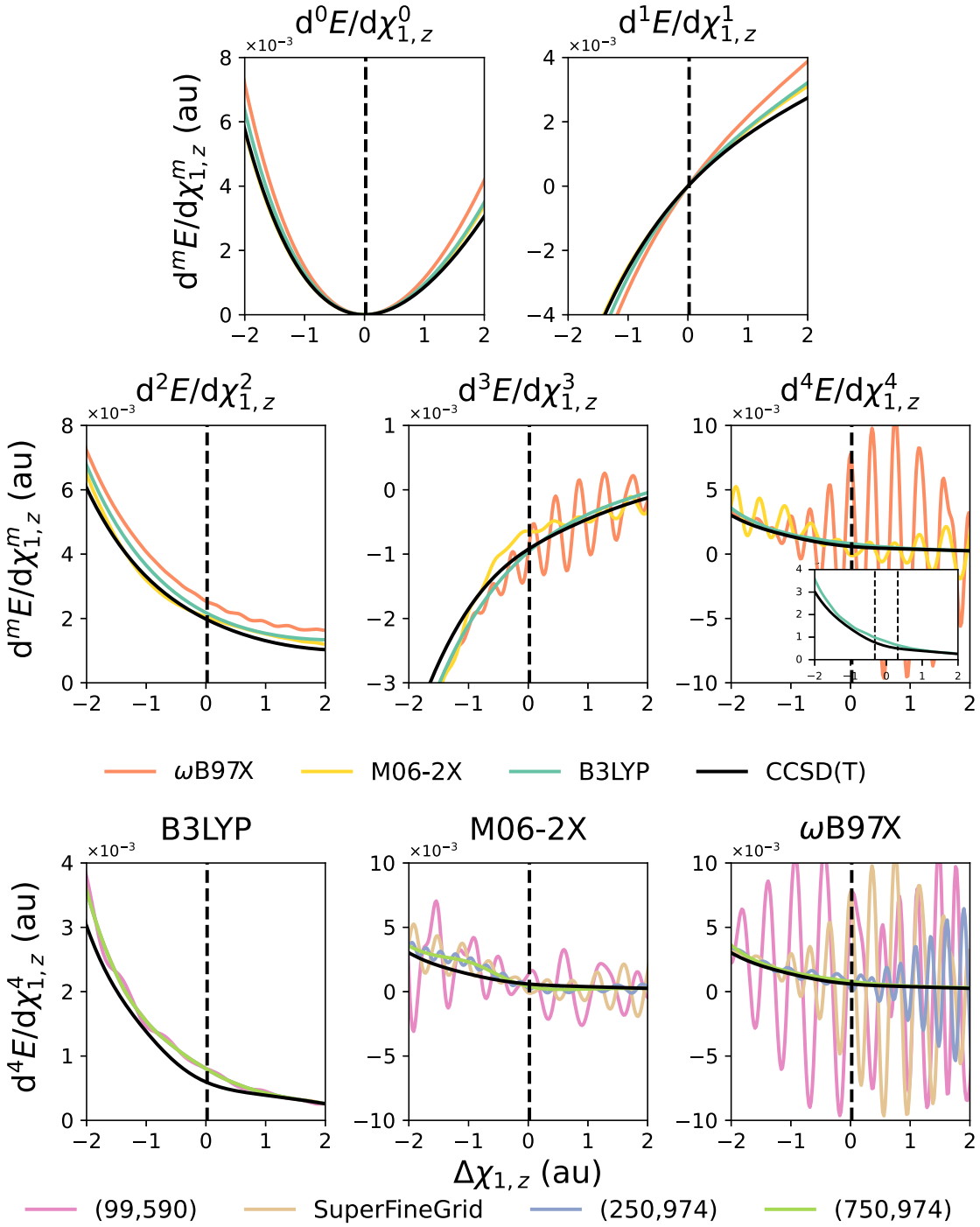

Figure 8.1: Grid-related spurious oscillations affecting the derivatives of the total energy with respect to $\chi_{1,z}$ of HCN · HF, $d^mE/d{\chi_{1,z}}^m$ (m = 0 - 4). Top and middle row panels show the results of DFAs combined with the (99, 590) grid. In the bottom row panels, curves of $d^4E/d{\chi_{1,z}}^4$ obtained with various integration grids are shown. In all plots, the black solid curve represents the CCSD(T) results and the vertical lines mark the equilibrium distance. All properties are given in atomic units.

8.5 Grid-Size Dependence of NR Contributions

Lastly, to fully confirm that the huge errors in the NR contributions to the hyperpolarizabilities are in fact due to the insufficient size of the grid used in the numerical integration, for all molecular systems we reevaluated electronic and NR NLOPs using the huge (750, 974) integration grid. Table 8.5 compiles the average absolute relative errors obtained with two integration grids, (SuperFineGrid) and (750, 974). The average errors of the electronic part of the response remain unchanged when computed with a larger grid. In contrast, there is a dramatic increase in the performance of all three problematic DFAs for the NR contributions. For $\beta^{\mathrm{NR}}_{zzz}$, the average errors decrease from 56% to 14% for M06, from 51% to 18% for M06-2X, and from 86% to 4% for ωB97X. More strikingly, one observes even larger reduction of errors for $\gamma^{\mathrm{NR}}_{zzzz}$. With the (750, 974) they are reduced from hundreds and thousands of percent to 9, 34, and 53% for ωB97X, M06, and M06-2X, respectively. Surprisingly, ωB97X in combination with the (750, 974) grid, outperforms other functionals in the computations of NR contributions, and shows almost equivalent performance as MP2.

Table 8.5: Effect of the adopted numerical integration grids, (SuperFineGrid) and (750, 974), on the performance of some DFAs in computations of electronic (el) and nuclear relaxation (NR) contributions to NLOPs. Numbers correspond to the average absolute relative errors. The relative errors were calculated with respect to the reference values calculated at the CCSD(T)/aug-cc-pVTZ level of theory. The averaging included all nine hydrogen-bonded complexes. Data for the SuperFineGrid corresponds to the one from Table 8.1.

	$\alpha^{\mathrm{el}}_{zz}$	$\beta^{\mathrm{el}}_{zzz}$	$\gamma^{\mathrm{el}}_{zzzz}$	$\alpha^{\mathrm{NR}}_{zz}$	$\beta^{\mathrm{NR}}_{zzz}$	$\gamma^{\mathrm{NR}}_{zzzz}$
M06(SuperFineGrid)	2	48	35	13	56	1072
M06(750, 974)	2	47	32	11	14	34
M062X(SuperFineGrid)	1	13	8	7	51	869
M062X(750, 974)	1	14	7	7	18	53
ωB97X(SuperFineGrid)	1	24	6	4	86	1768
ωB97X(750, 974)	1	26	6	3	4	9

8.6 Conclusions

In this work, we assessed the performance of DFAs in predicting the molecular (hyper)polarizabilities, including both the electronic and vibrational contributions, of nine hydrogen-bonded complexes. There are two main outcomes of this particular study.

The first one has a more general and pragmatic character and is directly related to the main results of the benchmark. Namely, the best accuracy in simulations of the electronic and NR (hyper)polarizabilities of the hydrogen-bonded complexes was observed for the CAM-B3LYP functional. For this DFA, the average absolute errors for all studied properties were below 20%. Next in performance were the LC-BLYP and MN15 functionals, with errors not exceeding 30%. In contrast, the worst performance was observed for the ωB97X, M06, and M06-2X, for which the NR contributions to the second hyperpolarizability yielded average relative errors exceeding hundreds of percent.

The second conclusion is more case-specific, nevertheless being as much (or even more) important as the first conclusion. By performing several subsequent analyses, we have managed to pinpoint the source of huge errors in $\gamma^{\mathrm{NR}}_{zzzz}$ observed for ωB97X, M06, and M06-2X. Namely, when combined with the integration grids of insufficient sizes, these

functionals provided very inaccurate high-order derivatives of the energy with respect to the nuclear coordinates, which are an essential factor of the expression for the anharmonic terms of vibrational hyperpolarizabilities. Such high-order derivatives turned out to be affected by the grid-related spurious oscillations of the corresponding curves along the selected displacement coordinate. This result was unexpected because in the main part of this work we applied SuperFineGrid — the largest predefined grid implemented in Gaussian package [259, 296]. For almost all applications of KS-DFT, SuperFineGrid is believed to be of an army-grade quality. Still, integration grid of such size turned out to be too small for the ωB97X, M06, and M06-2X functionals. Surprisingly, when the huge (750, 974) grid was adopted in FF-NR, the ωB97X functional unlocked its true potential and provided the best predictions of NLOPs in the whole set of tested methods, on par with the MP2 method. These latter observations and, especially, the phenomenon of the spurious oscillations, became the motivation for the two of main projects of this Thesis, to which Chapters 9 and 10 are devoted to.

Chapter 9

Spurious Oscillations Affecting Molecular Properties in KS-DFT

9.1 Introduction

Unlike most wavefunction methods, the computation of electronic energies from DFAs requires the numerical integration of the exchange and correlation functionals, which depend on the size and type of the numerical integration grid (see Sections 5.6.2 - 5.6.4). In most *in-silico* studies, the inaccuracy of the numerical integration is usually considered to yield small random numerical errors. Standard numerical integration techniques to calculate the energy density functionals have been mostly developed to integrate fairly smooth varying integrands with high atomic character, such as the electronic density.

In their seminal work, Johnson, Wolkow, and DiLabio showed that grid instability manifests itself through regular spurious oscillations of the Potential Energy Curve (PEC) [307]. Spurious oscillations were present in the PECs of the dispersion bound complexes, such as Ar_2 or benzene dimer, computed using the meta-GGA functionals combined with small integration grids. These oscillations were alleviated when large (radial) grids were used. Following this observation, Johnson *et al.* performed a detailed study of the most grid–susceptible components of some meta-GGAs. They proved that in the case of the dispersion bound complexes, the problem lies in the insufficiently sampled midpoint regions between the interacting fragments [308]. At those points, most of the kinetic density-based components of the meta-GGA energy functional integrands display enhanced values which are not accurately enough integrated. This is opposite to the behavior of the reduced density gradient, the main component of GGA functionals, which goes to zero at such points. Such midpoints are not properly sampled, mostly because of the nature of atomic–centered quadratures, which are designed to integrate efficiently the core and valence regions of the atoms. As shown and explained by the authors, one of the solutions is to simply increase the number of radial shells of the integration grid.

There are two main disadvantages of such a strategy. The first one is that the radial sampling does not change uniformly with the size of the radial grid. For example, unless the algorithm is changed, a two-fold increase in the number of points in the radial grid will not guarantee a twice better sampling in the region between two atoms — since most of the new radial shells will be added to the core region of atoms. The second one is that the computational cost of the integration increases linearly with the size of the applied numerical integration grid. A cost-effective solution was proposed by Grafenstein and Cremer through the local augmentation of the radial grid [309]. It provides a smooth increase in radial sampling in a selected part of the radial space (for example, in the far valence region), without the need to add redundant radial shells in other regions of space (for example, in the core regions). Despite its efficiency and simplicity, this approach has not been implemented in standard computational packages thus far.

Only a few works have addressed the issue of grid-related spurious oscillations [307–318]. Peverati, Macrina, and Baldridge commented on the grid-related oscillations in PECs of triangular structures of rare gases when described with the M06-L functional [311]. Contreras-Garcia and Yang observed similar spurious oscillations in the curves of the noncovalent interaction energy [312]. Mardirossian and Head-Gordon, when designing the B97M-V, ωB97M-V and ωB97X-V functionals, included information on the grid stability and spurious oscillations found in PECs of dissociating dispersion–bounded systems [77, 78, 319]. Grafenstein, Izotov, and Cremer studied the grid-dependent singularities occurring at the critical points of the electronic density for several meta-GGAs [310]. This problem is partially related to the grid-dependent oscillations, but leads to much more dramatic errors. Such singularities may produce great blow-ups in the total energy and may even prevent obtaining a stable SCF solution of the Kohn-Sham equations. As a possible cure to this problem, they remove the singularity–prone terms composing some

meta-GGA functionals. Yu, He and Truhlar did similar tests for the newly developed meta-GGA MN15-L functional [76]. Dasgupta and Herbert, in tests of their newly proposed SG2 and SG3 integration grids, check for the oscillations in PECs obtained with several grid-demanding DFAs [316]. Gould, Johnson, and Tawfik observed oscillations in PECs obtained with M06-L and SCAN when benchmarking popular dispersion corrections for non-equilibrium geometries. More recently, Bootsma and Wheeler discussed the practical issues arising from the lack of rotational invariance of the integration grids in DFT, which can lead to great uncertainties in the computed thermodynamic properties of organic reactions [320]. In all the examples mentioned above, except the last one, the grid-related spurious oscillations were analyzed only in PECs, i.e., the dependence of the derivatives of the energy or any other molecular properties along the vibrational coordinates were not explored. Under such "mild" conditions, spurious oscillations were observed only for the meta-GGA functionals.

First of all, let us show that the problem affects various energy derivatives and it cannot be always solved by increasing the numerical grid size. Figure 9.1 illustrates the spurious oscillations of $\mathrm{d}^m E/\mathrm{d}R^m$ in dissociating Ar_2, when calculated with three popular DFAs, namely B3LYP, M06-2X, and ωB97X. Although these oscillations are not evident from the potential energy curve of Ar_2 [308, 317, 321] (see the left graph in the first row in Figure 9.1), they compromise the calculation of accurate force constants and higher-order derivatives. In some cases, at the expense of a large computational cost, the problem can be partially alleviated if grids much larger than those predefined in standard packages are employed. However, for some DFAs, such as M06-2X and ωB97X, the oscillations persist after a large increase of the grid size (see the graphs in the bottom row in Figure 9.1). In such situations obtaining an accurate response property seems to be out of reach for atom-centered grids.

Therefore, there was an urge to perform a broad study and to establish the full extent of the spurious oscillation problem [147]. First of all, for such benchmark study one must acquire suitable reference results not affected by the spurious oscillations. However, there are two main difficulties to accomplish that: i) One cannot directly compare the DFA and another computational method because the differences could be either due to the spurious oscillations or to the methods' performance, hence, the reference must be based on the same DFA; ii) As shown in Figure 9.1, one should not blindly trust even a very large grid, because oscillations may still be present. To circumvent these two problems, we benefited from the fact that these grid-related errors have regular oscillatory patterns. We designed an algorithm that employs Fourier spectral analysis and filtering techniques to identify and quantify oscillations in property profiles along some nuclear displacement coordinates [147]. With the aid of this algorithm, we show that the spurious oscillations that arise from the sizable numerical integration errors of DFAs are neither a problem limited to meta-GGAs and the PES of dispersion compounds nor can be solved using the largest predefined grids available in most computational packages [147]. Moreover, these spurious oscillations easily occur in molecular complexes (and even a single molecule) with low-frequency vibrational modes, affect various molecular properties, and vary with the grid size, making most DFA predictions strongly grid-size dependent.

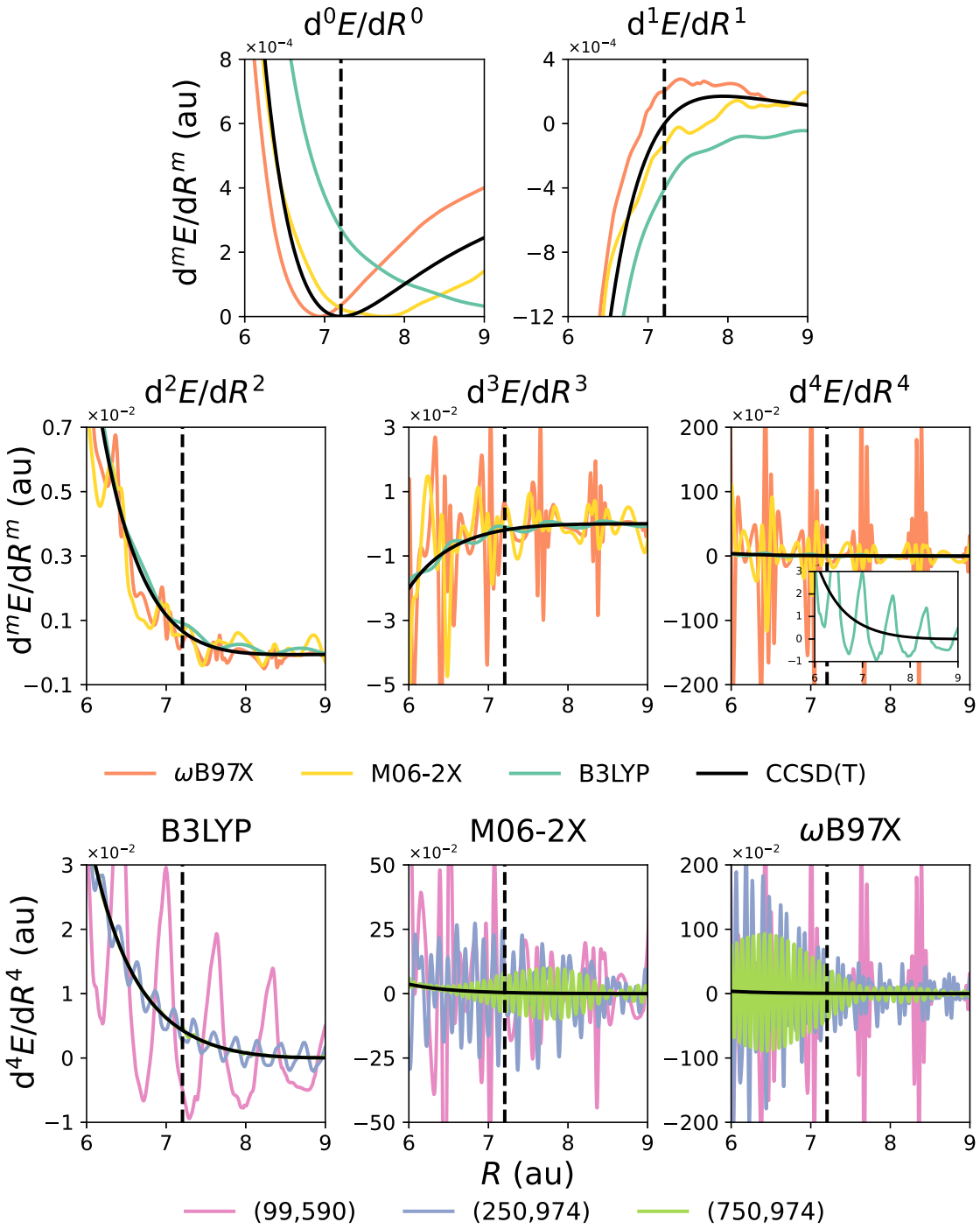

Figure 9.1: Grid-related spurious oscillations affecting the derivatives of the total energy with respect to interatomic separation R of Ar_2, d^mE/dR^m (m = 0 - 4). Top and middle row panels show the results of DFAs combined with the (99, 590) grid. On the bottom row panels, curves of d^4E/dR^4 obtained with various integration grids are shown. In all plots, the black solid curve represents the CCSD(T) results and the vertical lines mark the equilibrium distance. All units are a.u.

This Chapter, has the following structure. The second section describes the studied systems and the computational details. The third section describes a novel algorithm to efficiently detect and quantify the errors in molecular properties arising from the spurious oscillations. The fourth and fifth sections discuss the results of a broad benchmark on the grid-stability of DFAs, done with the use of the newly proposed algorithm. The last section shows additional examples of spurious oscillations affecting low-frequency modes of the single molecules.

9.2 Computational Details

9.2.1 Systems and Properties

In this work, we focused on the systems which characterize with a very loose and anharmonic motion (also referred as a floppy motion), dimolecular complexes stabilized by different type of intermolecular interaction. Namely, we included hydrogen-bonded ($HCN \cdot HF$, $HCN \cdot HCl$, $CO \cdot HF$, and $N_2 \cdot HF$), halogen-bonded ($HCN \cdot BrH$ and $HCN \cdot BrF$), and dispersion-bonded (Ar_2 and He_2) complexes (see Table 9.1). Selected hydrogen-bonded systems were also used in the previous study on the NR contributions (see Chapter 8), whereas dispersion-bonded Ar_2 and He_2 have been frequently used in other studies to test the grid-dependency of DFAs [76–78, 307–310, 316, 319]. All of the systems are co-linear with the Cartesian Z-axis and are characterized by a low-frequency mode corresponding to the intermolecular stretching (their harmonic frequencies are compiled in Table 9.1).

Table 9.1: Molecular complexes included in the benchmark and the type of intermolecular interaction responsible for their binding. The third column contains the harmonic vibrational frequency of the the intermolecular stretching mode (belonging to the $\Sigma^+_{(g)}$ irreducible representation), obtained at the CCSD(T)/aug-cc-pVTZ level of theory.

Complex	Intermolecular interaction	ω_{har} [cm^{-1}]
$HCN \cdot HF$	hydrogen-bond	187.3
$HCN \cdot HCl$	hydrogen-bond	117.8
$N_2 \cdot HF$	hydrogen-bond	121.5
$OC \cdot HF$	hydrogen-bond	137.3
$HCN \cdot BrH$	halogen-bond	63.3
$HCN \cdot BrF$	halogen-bond	124.7
He_2	dispersion	27.8
Ar_2	dispersion	29.5

In this work, we refer to the property curve as the dependence of the property (for instance, an energy derivative) along a nuclear displacement coordinate. We analyzed oscillations in the property curves along the nuclear displacement coordinate ξ of derivatives of the total electronic energy, $\mathrm{d}^m E/\mathrm{d}\xi^m$ (m = 0 - 4), the dipole moment, $\mathrm{d}^m \mu_z/\mathrm{d}\xi^m$ (m = 0 - 3), and the static polarizability, $\mathrm{d}^m \alpha_{zz}/\mathrm{d}\xi^m$ (m = 0 - 2) (all derivatives are with respect to the same nuclear displacement coordinate, i.e., ξ). The choice of displacement coordinates is discussed in Section 9.2.2. These derivatives are of key importance to describe basic molecular properties related to the vibrational structure of the molecules, such as IR and Raman spectroscopies. In Table 9.2, the connection between these derivatives and various molecular properties is given.

Table 9.2: Property derivatives studied in this work and the experimental molecular properties they are describing, including vibrational frequencies, IR and Raman intensities, and NR contributions to static non-linear optical properties. In the case of higher anharmonic corrections, lower-order mechanical derivatives are also utilized. Moreover, in the case of IR and Raman intensities, as well as NR contributions, purely mechanical derivatives $\mathrm{d}^m E/\mathrm{d}\xi^m$ are also used.

Derivative	Molecular Property
$\mathrm{d}E/\mathrm{d}\xi$	Forces acting on nuclei
$\mathrm{d}^2E/\mathrm{d}\xi^2$	Quadratic force constants; Harmonic vibrational frequencies; NR contribution to vibrational polarizability
$\mathrm{d}^3E/\mathrm{d}\xi^3$	Cubic force constants; Cubic anharmonic correction to vibrational frequencies; Quadratic and Cubic anharmonic corrections to the IR and Raman intensities; NR contribution to vibrational hyperpolarizability
$\mathrm{d}^4E/\mathrm{d}\xi^4$	Quartic force constants; Quartic anharmonic correction to vibrational frequencies; Cubic anharmonic correction to the IR and Raman intensities; NR contribution to vibrational second hyperpolarizability
μ_i	Permanent dipole moment
$\mathrm{d}\mu_i/\mathrm{d}\xi$	Harmonic IR intensity; NR contribution to vibrational polarizability
$\mathrm{d}^2\mu_i/\mathrm{d}\xi^2$	Quadratic anharmonic correction to IR intensity; NR contribution to vibrational hyperpolarizability
$\mathrm{d}^3\mu_i/\mathrm{d}\xi^3$	Cubic anharmonic correction to IR intensity; NR contribution to vibrational second hyperpolarizability
α_{ij}	Static electronic polarizability
$\mathrm{d}\alpha_{ij}/\mathrm{d}\xi$	Harmonic Raman intensity; NR contribution to vibrational hyperpolarizability
$\mathrm{d}^2\alpha_{ij}/\mathrm{d}\xi^2$	Quadratic anharmonic correction to Raman intensity; NR contribution to vibrational second hyperpolarizability
[a] $\mathrm{d}^3\alpha_{ij}/\mathrm{d}\xi^3$	Cubic anharmonic correction to Raman intensity

[a] Derivative not studied in this work

9.2.2 Nuclear Displacement Coordinates

To effectively describe the most anharmonic nuclear displacements in polyatomic systems, we adopted the first–order Field-Induced Coordinate (FIC), $\chi_{1,i}$ [118] (see Section 4.11). In the case of hydrogen- and halogen-bonded systems, the property curves were obtained for linear displacements along the first-order FIC vector, $\chi_{1,z}$. For these systems, $\chi_{1,z}$ is parallel to the main intermolecular axis and involves only vibrational modes of the σ-type. In fact, $\chi_{1,z}$ is almost exclusively given by the low-frequency intermolecular stretching mode between the two molecules held by a noncovalent interaction (see example in Figure 9.2).

Figure 9.2: Graphical representation of $\chi_{1,z}$ (in normalized Cartesian coordinates) of HCN $\cdot$ BrH obtained with B3LYP/aug-cc-pVTZ and the (250, 974) integration grid. For this system, according to Eq. 4.143, $\chi_{1,z} = -219.48Q_1 + 0.21Q_2 + 0.19Q_3 + 0.62Q_4$, where Q_1-Q_4 are Σ-type normal modes, and Q_1 is the intermolecular stretching.

Therefore, one can try to express $\chi_{1,z}$ in terms of the separation between two molecules, where $\pm\Delta\chi_{1,z}$ corresponds approximately to ± 0.15Å in the intermolecular separation. However, for the sake of generality, in this paper, we have preferred to use FICs which would be the most efficient way of studying the spurious oscillations of more complex polyatomic systems. In this work, for polyatomic systems, curves were sampled with 400 points and $\Delta(\Delta\chi_{1,z})$ = 0.02 au, and $\Delta\chi_{1,z}$ = 0 corresponds to the optimized equilibrium geometry.

In the case of diatomic dispersion-bonded systems, scan along the interatomic distance R has been done. The curves were sampled with $\Delta R = 0.005$ Bohr and consisted of 1000 points. In all cases, irrespective if the methods give a bounded system or not, starting points (i.e. ΔR = 0.0 Bohr) $R = 7.2$ Bohr and $R = 6.0$ Bohr were taken for Ar_2 and He_2, respectively. Afterwards, the derivatives of the properties were calculated numerically (see Section 9.2.4).

9.2.3 Electronic Structure Calculations

We investigated a broad range of forty-five DFAs including generalized gradient approximations (GGAs), meta-GGAs, global hybrids, range-separated DFAs, and double hybrids (see Table 9.3). For all DFAs (see Table 9.3), KS-DFT computations were performed using Gaussian16 [296], except for B97, B97-D, SCAN, SCAN0, ωB97X-D3, B97M-V, ωB97X-V, ωB97M-V, which were performed using QChem 5.1 [322]. CCSD(T) computations were done using CFOUR [323]. In this work, all of the computations employed the aug-cc-pVTZ [298–305] basis set, except in Section 9.6 where the aug-cc-pVDZ [298–305] basis set was employed.

Table 9.3: List of DFAs used thorough this work; GGA — Generalized Gradient Approximation, GH — global hybrids, RSH — range-separated hybrids, DH — double hybrids.

DFA	Type	Ref.	DFA	Type	Ref.
BLYP	GGA	[178, 180]	PBE	GGA	[179]
B1LYP	GH GGA	[324]	PBE0	GH GGA	[294]
B3LYP	GH GGA	[325, 326]	LR-wPBE	RSH GGA	[191]
BH&H	GH GGA[a]	[296]	LR-wHPBE	RSH GGA	[296]
BH&HLYP	GH GGA	[296]	B97	GH GGA	[327]
LC-BLYP	RSH GGA	[190]	B97-D	GGA	[328]
CAM-B3LYP	RSH GGA	[193]	ωB97	RSH GGA	[192]
M06	GH meta-GGA	[293]	ωB97X	RSH GGA	[192]
M06-L	meta-GGA	[184]	ωB97X-D	RSH GGA	[329]
M06-HF	GH meta-GGA	[330]	ωB97X-D3	RSH GGA	[331]
M06-2X	GH meta-GGA	[293]	B97M-V	meta-GGA	[319]
SOGGA11	GGA	[332]	ωB97M-V	RSH meta-GGA	[77]
SOGGA11-X	GH GGA	[333]	ωB97X-V	RSH GGA	[78]
M11	RSH meta-GGA	[334]	TPSS	meta-GGA	[182]
M11-L	meta-GGA	[335]	RevTPSS	meta-GGA	[336]
MN12-L	meta-GGA	[337]	TPSSh	GH meta-GGA	[338]
N12-SX	RSH GGA	[339]	PBE-0DH	DH GGA	[214]
N12	GGA	[334]	PBE-QIDH	DH GGA	[213]
MN12-SX	RSH meta-GGA	[339]	B2PLYP	DH GGA	[211]
MN15	GH meta-GGA	[75]	mPW2PLYP	DH GGA	[212]
MN15-L	meta-GGA	[76]	HSE03	RSH GGA	[340, 341]
SCAN	meta-GGA	[183]	HSE06	RSH GGA	[295]
SCAN0	GH meta-GGA	[342]			

[a] No GGA exchange is used, i.e., $E_{xc}^{BHH} = 0.5\, E_x^{Slater} + 0.5\, E_x^{HF} + E_c^{LYP}$

The SCF DIIS converge criteria was set to 10^{-12} in the HF and KS–DFT computations, whereas in CCSD(T) the convergence threshold for maximum amplitude was set to 10^{-10}. In the coupled–perturbed equations, convergence threshold was set to 10^{-11}. In all methods, the screening threshold for the two–electron integrals was set to 10^{-14}. Except Ar_2 and He_2, all molecular systems were optimized at the corresponding level of theory with the convergence criteria of Root Mean Square (RMS) of the gradient set below 10^{-6} (and 10^{-8} in case of the CCSD(T)). In the optimization procedure with DFAs, only the (250, 974) grid has been used.

In the KS-DFT computations, we use Handy-Lebedev grids [218–221, 224] (see Sections 5.6.3 and 5.6.4) , (N_r, N_Ω), where N_r is number of radial shells and N_Ω is number of points in the Lebedev angular grid. In this study, the following unpruned integration grids were tested: (99, 590), (250, 974), (500, 974) and (750, 974), along with some of their pruned versions: UltraFine, SG3, and SuperFineGrid. For the integration of nonlocal VV10 [343] correlation in B97M-V, ωB97X-V and ωB97M-V, the SG3 grid has been used. The smallest unpruned grid tested in our study, (99, 590), is an unpruned version of UltraFine [296] (a default predefined grid in Gaussian16) and SG3 [316] (the largest predefined grid in QChem 5.1 [322]). The (250, 974) grid is the unpruned version of a popular (and the largest) predefined grid in Gaussian16, known as the SuperFineGrid. The unpruned (250, 974) grid is already larger than most of the largest predefined grids in standard computational packages. Table 9.4 contains the list of largest predefined grids utilized in the chosen computational packages. Besides the names of the grids, their un-

pruned parent grid, as well as the radial grid are also given. The cost of the calculations multiply by four going from (99, 590) to (250, 974) and triples from (250, 974) to (750, 974). Hence, the calculation with the latter grid increases by more than one order of magnitude the cost of most calculations performed nowadays.

Table 9.4: The largest predefined grids in some quantum chemical computational packages. The fourth column lists the parent unpruned grids (shown for *C* atom).

Program	Grid	Radial Quadrature	Parent Grid Size
Gaussian16	SuperFineGrid	Handy [224]	(250, 974)
QChem 5.1	SG3	DE [229, 344–346]	(99, 590)
ORCA 5.0.1	DEFGRID3	Ahlrichs [227]	(45, 590)
GAMESS US	JANS	Handy [224]	(150, 974)
NWChem 7.0.2	huge	Knowles [228]	(300, 1454)
DALTON	UltraF	LMG [230]	(111, 1454)

9.2.4 Numerical Differentiation

Molecular properties E, μ_z, and α_{zz} have been obtained analytically. In the case of double hybrid functionals and CCSD(T), the relaxed densities have been used to compute μ_z and α_{zz}. On the other hand, the derivatives $\mathrm{d}^m E/\mathrm{d}\xi^m$ (m = 1 - 4), $\mathrm{d}^m \mu_z/\mathrm{d}\xi^m$ (m = 1 - 3), and $\mathrm{d}^m \alpha_{zz}/\mathrm{d}\xi^m$ (m = 1 - 2) were obtained through numerical differentiation. For that purpose, we have constructed an algorithm to automatically choose the Savitzky-Golay [347] differentiation filter (i.e., differentiation via polynomial fitting). We will refer to this method as AutoSG. The order of the filter (i.e., the order of the interpolating polynomial) has been chosen automatically, with the criterion to quantitatively differentiate all relevant (i.e., non-noisy) bands in the Fourier Transform spectrum of the property. In this procedure, a maximum order of 30 has been used as the limit. The length of the filter (i.e., the number of points used in the fitting) has been set to 51, 75, and 101 for hydrogen-, halogen- and dispersion-bonded systems, respectively. Such number of points provided a good trade-off between the cost and steepness of the cutoff region.

The most important feature of this algorithm is that it differentiates hidden spurious oscillatory bands and keeps the noise contribution to the minimum. Such numerical derivatives have proven to be of very good quality and superior to ones obtained with a finite (central) differentiation or even the Romberg-Rutishauser procedure [292]. If one would use the finite (central) differentiation with a very large step size, one could miss the differentiation of high-frequency spurious oscillations. However, the latter does not represent a solution to the spurious oscillations because there is no guarantee that finite-difference numerical derivatives are not affected by the spurious oscillations. Besides, since spurious oscillations are an inherent problem of the DFA, analytical derivatives — which are employed in many computational packages for low-order derivatives — would also present spurious oscillations on the property curve.

In Figures 9.3-9.5, we show a performance comparison of numerical differentiation techniques, including our newly implemented AutoSG. In the test, the formula for central differentiation (utilizing Richardson extrapolation once, which we label here as $CD(n,1)$) has been used:

$$\begin{aligned} CD(n,1) = \frac{1}{48\Delta R^4}\,[&-E(R-4\Delta R)+68E(R-2\Delta R)-256E(R-\Delta R)+378E(R) \\ &-256E(R+\Delta R)+68E(R+2\Delta R)-E(R+4\Delta R)]\,, \end{aligned} \tag{9.1}$$

where $\Delta R = 2^n \cdot 0.005$ au.

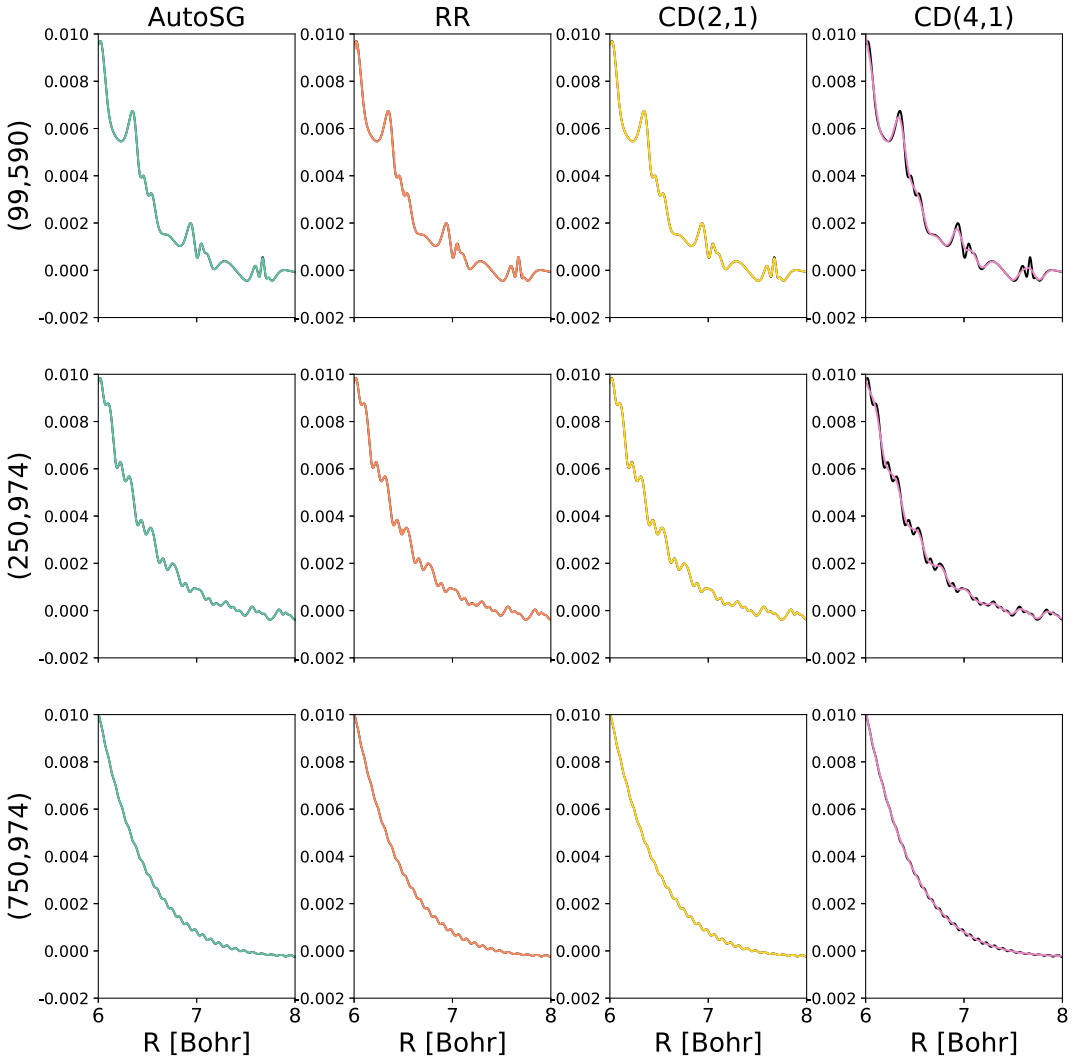

Figure 9.3: Accuracy of different numerical differentiation procedures to compute $\mathrm{d}^2E/\mathrm{d}R^2$ in Ar_2. Black curve represents the reference curve obtained analytically. Other curves were obtained as numerical second derivatives of the total energy. AutoSG is our automatic differentiation using Savitzky-Golay filtering; RR corresponds to Romberg-Rutishauser differentiation involving up to 4 different fields and the starting perturbation of $\Delta R = 0.005$ a.u.; CD(2,1) corresponds to the central differentiation and uses $\Delta R = 0.02$ a.u. starting step; CD(4,1) corresponds to the central differentiation and uses $\Delta R = 0.08$ a.u. starting step; all curves were obtained using the ωB97X functional combined with different integration grids, namely (99, 590) — top row, (250, 974) — middle row, and (750, 974) — bottom row.

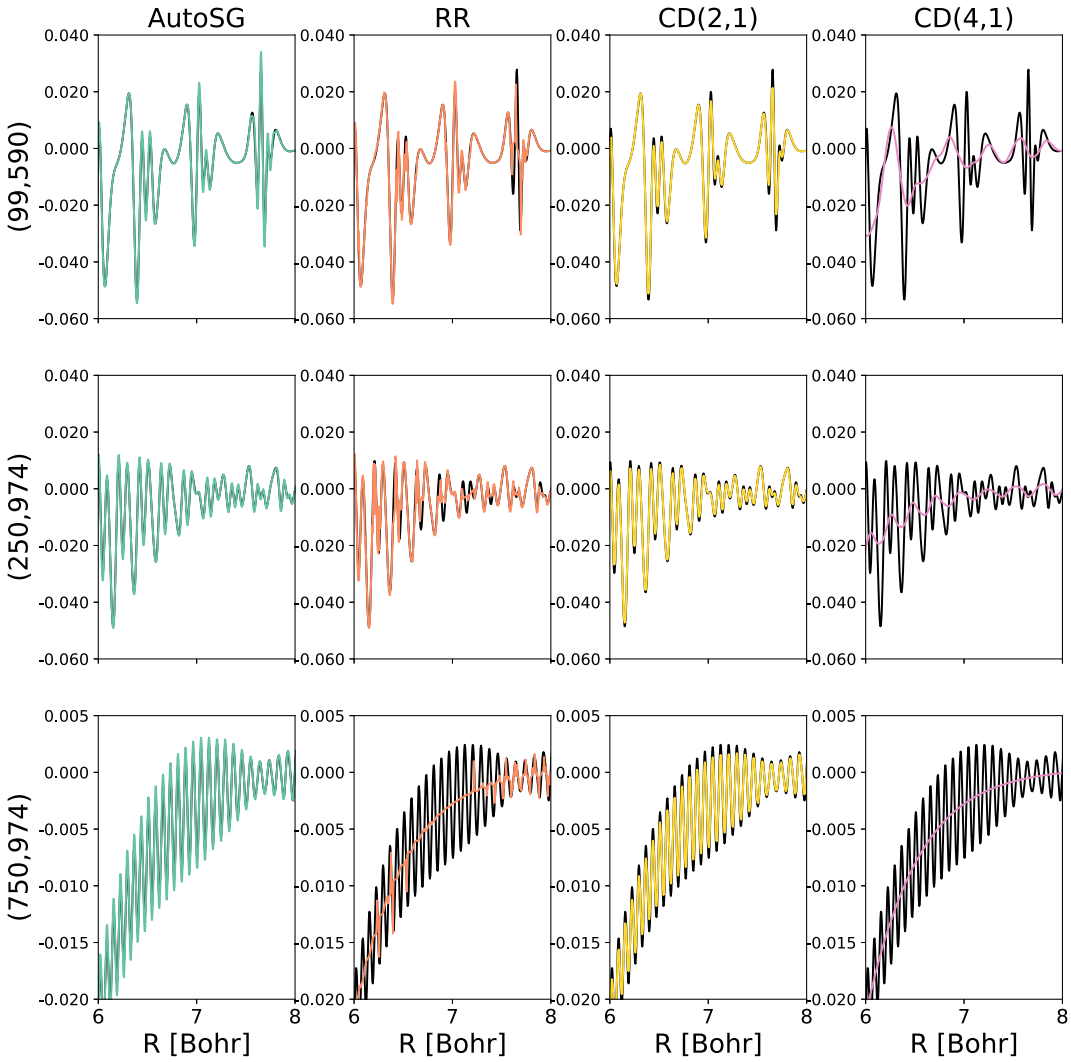

Figure 9.4: Accuracy of different numerical differentiation procedures to compute $\mathrm{d}^3E/\mathrm{d}R^3$ in Ar_2. Black curve represents the reference curve obtained as numerical first derivative (utilizing CD(2,1)) of the analytically obtained $\mathrm{d}^2E/\mathrm{d}R^2$. Other curves were obtained as numerical third derivatives from the total energies. AutoSG is our automatic differentiation using Savitzky-Golay filtering; RR corresponds to Romberg-Rutishauser differentiation involving up to 4 different fields and the starting perturbation of $\Delta R = 0.005$ a.u.; CD(2,1) corresponds to the central differentiation and uses $\Delta R = 0.02$ a.u. starting step; CD(4,1) corresponds to the central differentiation and uses $\Delta R = 0.08$ a.u. starting step; all curves were obtained using the ωB97X functional combined with different integration grids, namely (99, 590) — top row, (250, 974) — middle row, and (750, 974) — bottom row.

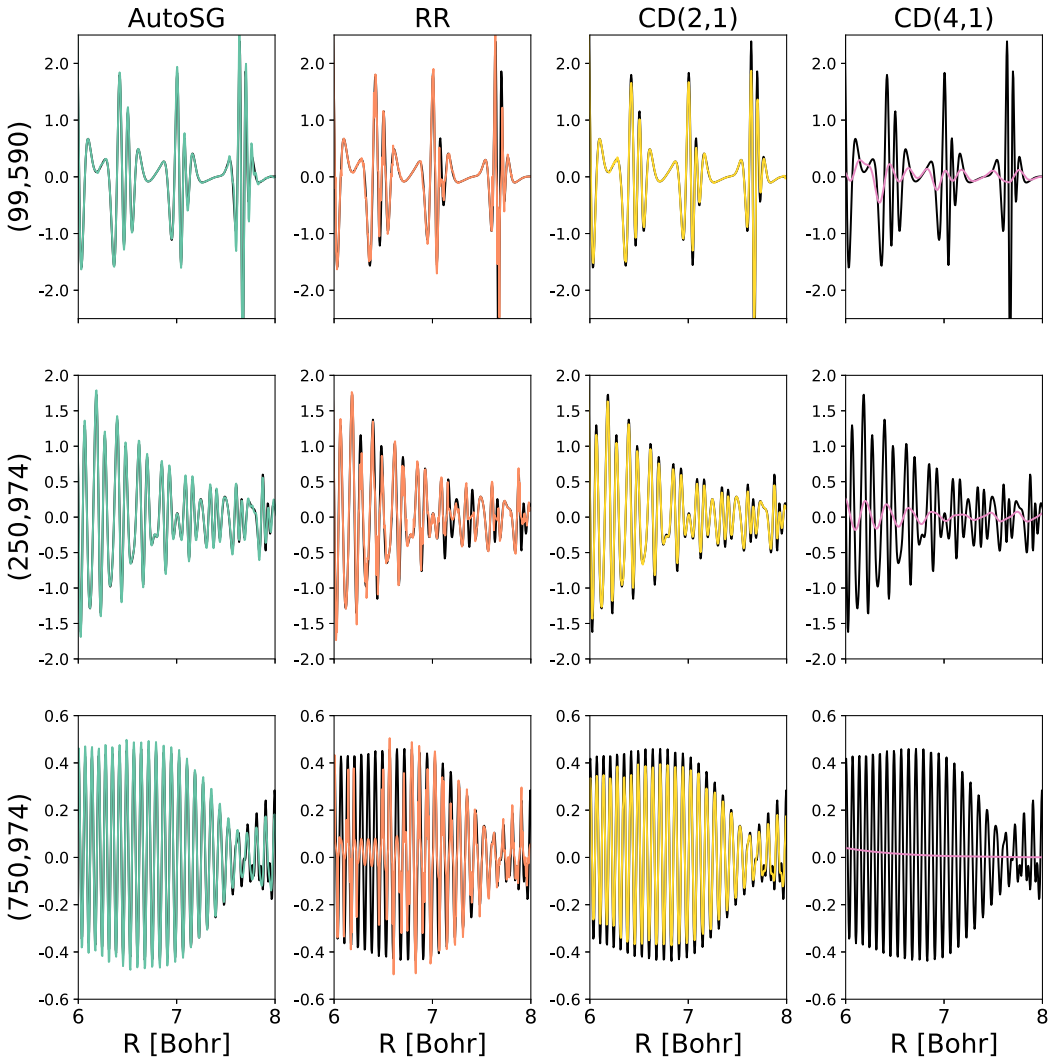

Figure 9.5: Accuracy of different numerical differentiation procedures to compute $\mathrm{d}^4E/\mathrm{d}R^4$ in Ar_2. Black curve represents the reference curve obtained as numerical second derivative (utilizing CD(2,1)) of the analytically obtained $\mathrm{d}^2E/\mathrm{d}R^2$. Other curves were obtained as numerical fourth derivatives from the total energies. AutoSG is our automatic differentiation using Savitzky-Golay filtering; RR corresponds to Romberg-Rutishauser differentiation involving up to 4 different fields and the starting perturbation of $\Delta R = 0.005$ a.u.; CD(2,1) corresponds to the central differentiation and uses $\Delta R = 0.02$ a.u. starting step; CD(4,1) corresponds to the central differentiation and uses $\Delta R = 0.08$ a.u. starting step; all curves were obtained using the ωB97X functional combined with different integration grids, namely (99, 590) — top row, (250, 974) — middle row, and (750, 974) — bottom row.

9.3 Procedure for Quantification of the Grid-Related Spurious Oscillations

9.3.1 The Algorithm

In this section, the algorithm for the detection and quantification of spurious oscillations along a nuclear displacement is discussed in detail. The algorithm will need two inputs: the property curve we want to analyze, calculated with a certain DFA combined with the grid G, $P^{\mathrm{DFA(G)}}$, and some spurious-oscillation free reference, P^{ref}. In order to guarantee that the latter does not have spurious oscillations it is convenient to choose a computational method that does not rely on numerical integration, such as HF, MP2 or coupled-cluster. The number of points that will be analyzed on the property curve defines the set of geometries indicated as $\{\boldsymbol{\xi_0}\}$, and we will refer to them as *displacement range*. The algorithm will provide a quantitative measure of the degree of oscillations that are present in the property curve within the displacement range.

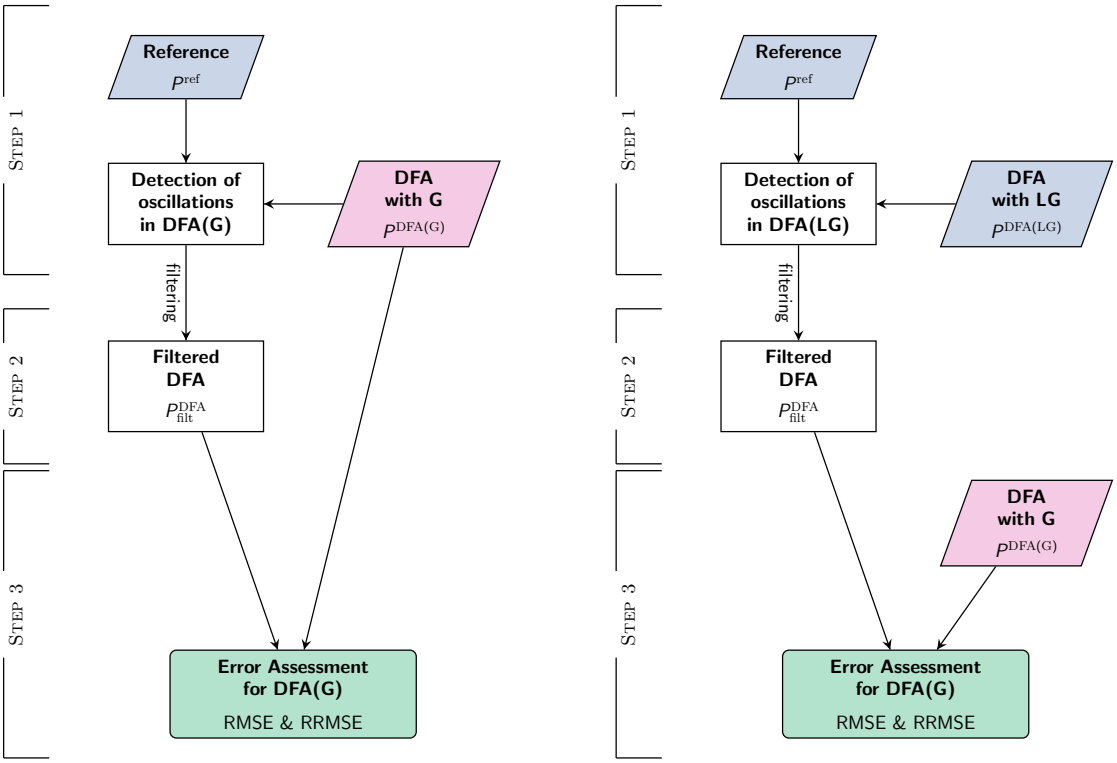

Figure 9.6: *Left: Algorithm 1*, which tests the property curve obtained with grid G, and also uses this property curve to obtain the filtered curve. *Right: Algorithm 2*, which uses a larger-grid (LG) to obtain the filtered curve that will be used to quantify the spurious oscillations on a smaller grid (G). In all cases, the same DFA and basis are employed.

The flowchart of the algorithm is shown in Figure 9.6. The algorithm consists of three stages: 1) detection of the oscillations in $P^{\mathrm{DFA(G)}}$, 2) obtaining an oscillation-free version of P^{DFA}, $P^{\mathrm{DFA}}_{\mathrm{filt}}$, and 3) quantifying the spurious oscillations by comparing $P^{\mathrm{DFA}}_{\mathrm{filt}}$ and $P^{\mathrm{DFA(G)}}$. There are two variants of the algorithm, which will be commented below. The algorithm has been implemented in an in-house Python3 program, which uses NumPy [348], SciPy [349], SymPy [350], and Matplotlib [351] libraries. All graphs shown in this Chapter were prepared using the Matplotlib package [351].

The main goal of the first and second stages is to obtain $P^{\mathrm{DFA}}_{\mathrm{filt}}$. To do so it is convenient to use the same DFA one wants to test but with a larger integration grid (LG), which will

have much less spurious oscillations than a lower-quality grid (G) that one wants to test (*Algorithm 2* in Figure 9.6). However, as we will comment later, the algorithm can be easily adapted to confidently filter the oscillations of any grid size (presuming these will be easily separable). *Algorithm 1* in Figure 9.6 uses the same grid G of the tested DFA to generate $P_{\text{filt}}^{\text{DFA}}$. In this first work, to reduce the uncertainty about the presence of spurious oscillations, we have decided to employ the Algorithm 2 and the (750, 974) grid as LG ($P^{\text{DFA(LG)}} = P^{\text{DFA(750,974)}}$).

In this work, we have used reference property curves obtained with CCSD(T), $P^{\text{ref}} = P^{\text{CCSD(T)}}$. However, the choice of the *ab-initio* method used to distinguish the (eventual) oscillations in $P^{\text{DFA(LG)}}$ does not impact the final results in detection and measurement of the spurious oscillations. Despite the evident differences in performance between HF, MP2, and CCSD(T) all these methods work equally well as spurious-oscillation-free references for this algorithm.

At the first stage, the discrete Fourier transforms of $P^{\text{DFA(LG)}}$ and P^{ref} are compared. In these spectra, only bands which are ten times larger than the baseline of the noise level (determined as the average of last 30% of the spectrum) are considered. The comparison of their frequency spectra will reveal bands that are present on $P^{\text{DFA(LG)}}$ but not in P^{ref}, which will be marked as potentially spurious oscillatory bands. The bands of both curves are evenly matched using an overlap criterion larger than 0.90. Any extra bands present in the spectrum of P^{DFA}, but not in P^{ref}, are labeled as spurious oscillatory bands (spurious bands, hereafter). For a better distinction of the spurious bands, which may have much lower amplitude than the main low-frequency band, the Kaiser windowing technique [352] is adapted to P^{DFA} and P^{ref} (with a different relative side-lobe attenuation for each type of derivative). Figure 9.7 illustrates the property curves of $P^{\text{DFA(750,974)}}$ and $P^{\text{CCSD(T)}}$ and their Fourier transform, for HCN · BrH along $\Delta\chi_{1,z}$.

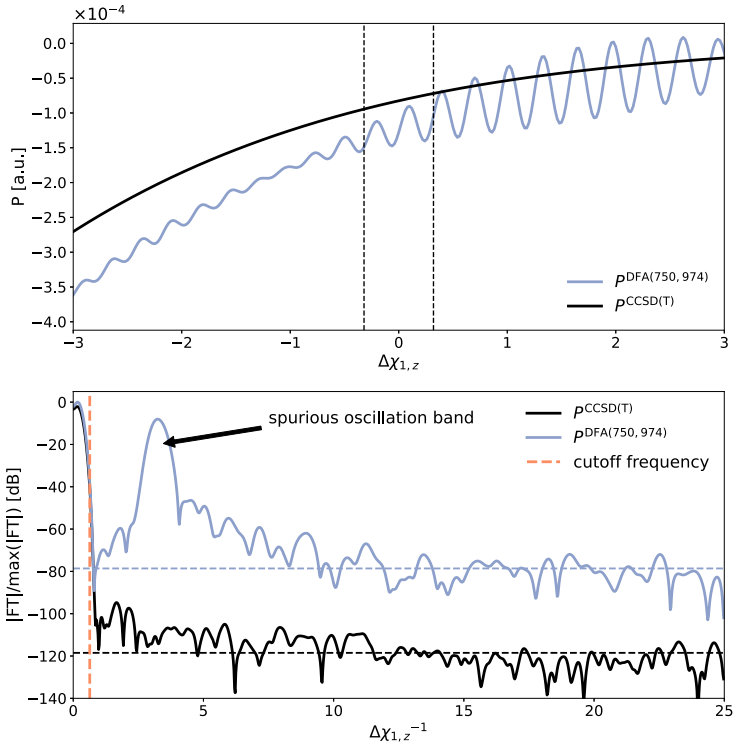

Figure 9.7: *Top:* $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ curves for HCN · BrH along the $\Delta\chi_{1,z}$ displacement, obtained with the ωB97X/aug-cc-pVTZ combined with the (750, 974) grid, $P^{\mathrm{DFA(LG)}} = P^{\mathrm{DFA(750,974)}}$ (blue). Property curve obtained with CCSD(T), $P^{\mathrm{ref}} = P^{\mathrm{CCSD(T)}}$, is shown in black. *Bottom:* Frequency spectra of $P^{\mathrm{DFA(750,974)}}$ and $P^{\mathrm{CCSD(T)}}$. The vertical dashed line marks the cutoff frequency for the low-pass filtering of $P^{\mathrm{DFA(750,974)}}$. The horizontal dashed line marks the noise level in the $P^{\mathrm{DFA(750,974)}}$ and $P^{\mathrm{CCSD(T)}}$ spectra. The band corresponding to the spurious oscillation present in $P^{\mathrm{DFA(750,974)}}$, is indicated with a black arrow.

At the second stage, the spurious bands detected in $P^{\mathrm{DFA(LG)}}$ are filtered out (along with high-frequency numerical noise), yielding a filtered reference curve, $P^{\mathrm{DFA}}_{\mathrm{filt}}$. This is done using a low-pass finite impulse response (FIR) filter of type I, designed with the Remez exchange algorithm [353]. Filter specifications (frequency cutoff, stop– and pass–band attenuation) are automatically chosen by the program. To this end, the position and height of the oscillation bands are used:

- The cutoff frequency corresponds to the right-most side of the main non-spurious band (which determines the general shape of the curve), which has the same height as the top of the detected oscillatory band.
- The stop-band attenuation is chosen to be 60dB smaller than the maximum of the highest oscillatory band.
- The pass-band attenuation is chosen to be 40dB.
- The number of filter coefficients (i.e., the number of filter taps) is set to the first odd number smaller than the total number of points minus the number of points in the displacement range, $\{\boldsymbol{\xi_0}\}$.
- The width of the transition is estimated using Bellanger's estimate [354], and if needed, is iteratively changed to yield a stable filter.

The latter parameters permit to filter the oscillations within $\{\boldsymbol{\xi_0}\}$ with a very good precision and in an automatized manner. It is important to stress out that such designed FIR filter will provide a valid part of the filtered curve only within $\{\boldsymbol{\xi_0}\}$. The resulting curve

outside the displacement range is corrupted by the initial conditions of the FIR filter, which is an indispensable feature of the FIR filters.

The results of filtering of $P^{\mathrm{DFA(LG)}}$, along with the Fourier transform of an automatically designed FIR filter (also called the response of the filter), are shown in Figure 9.8. The separation of the spurious band from the main low-frequency band is the key factor to a successful filtering. In principle, if the filtering is done accurately enough, any grid will provide the same true reference $P^{\mathrm{DFA}}_{\mathrm{filt}}$. We have checked that the main low-frequency band is well separated from the spurious bands when the (750, 974) grid is used as LG. Using (250, 974) as LG, in most cases, also yields the correct filtered curve $P^{\mathrm{DFA}}_{\mathrm{filt}}$. Difficulties arise for (99, 590) or smaller grids because, in these cases, the spurious oscillations are very strong, span already at low frequencies, and cannot be easily filtered out. For these difficult cases, the filtering using $P^{\mathrm{DFA(99,590)}}$ requires a much a higher resolution of the $P^{\mathrm{DFA(99,590)}}$ spectra and, hence, more points in $P^{\mathrm{DFA(99,590)}}$, which is not always feasible. Therefore, with the current version of the filtering technique we recommend the use of large grids as LG.

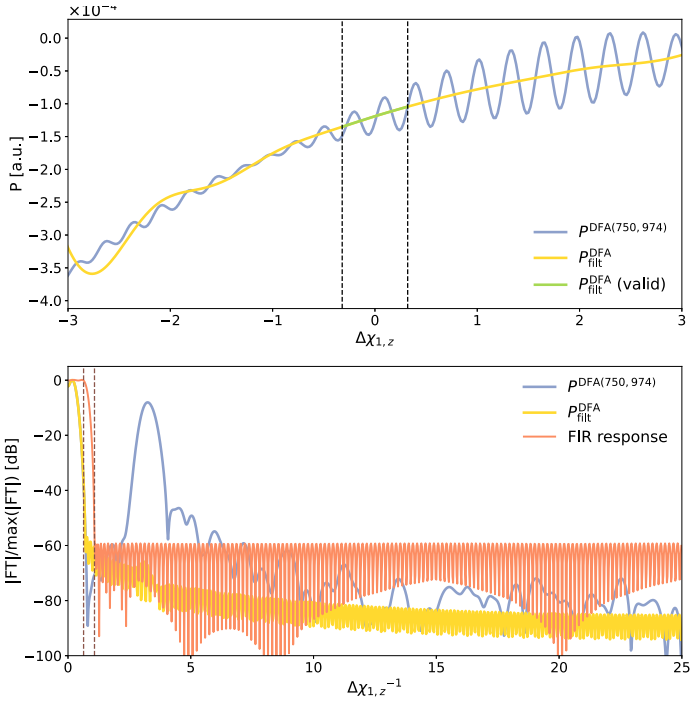

Figure 9.8: Results of the filtering of $P^{\mathrm{DFA(LG)}} = P^{\mathrm{DFA(750,974)}}$ with an automatically designed FIR filter, shown for $P = \mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ of HCN · BrH using ωB97X/aug-cc-pVTZ. *Top:* Original (i.e., unfiltered) and filtered curves in the $\chi_{1,z}$ domain. One can observe the effect of the initial conditions in the left part of the yellow curve. The valid part of the filtered curve within the $\{\boldsymbol{\xi_0}\}$ studied range, is shown in green color. *Bottom:* Frequency spectra of the original and filtered curve (in reciprocal domain), along with the response of the designed FIR filter (orange, shown in the normalized decibel scale). In the top plot, dashed black vertical lines mark the displacement range of the property curve $\{\boldsymbol{\xi_0}\}$, for which RMS, RMSE, and RRMSE are calculated. In the bottom plot, dashed orange vertical lines mark the transition band of the designed low-pass filter.

At the third (and last) stage of the algorithm, we use $P^{\mathrm{DFA}}_{\mathrm{filt}}$ and $P^{\mathrm{DFA(G)}}$ to quantify the spurious grid oscillations in the property curve obtained from some DFA and integration grid G, $P^{\mathrm{DFA(G)}}$. We employ the RMS and the root mean square error (RMSE) to define the relative root mean square error (RRMSE), which quantifies the distortion of the

property curve in the displacement range due to spurious oscillations.

$$RMSE = RMS\left[P^{\mathrm{DFA(G)}} - P^{\mathrm{DFA}}_{\mathrm{filt}}\right], \tag{9.2}$$

and

$$RRMSE = \frac{RMSE}{RMS\left[P^{\mathrm{DFA}}_{\mathrm{filt}}\right]} \cdot 100\,. \tag{9.3}$$

In the current work, the RMS involves the geometries within the displacement range, which we have selected following these criteria:

- For the hydrogen– and halogen– bonded systems: 33 points in the range 0 ± 0.32 $\Delta\chi_{1,z}$ (i.e., displacement along the normalized FIC coordinate, starting from the optimized geometry).
- For the dispersion-bonded systems: 129 points in the range 7.2 ± 0.32 a.u. for Ar_2, and 6.0 ± 0.32 a.u. for He_2 (i.e., displacements along the interatomic separation close to the overall minimum in the potential energy of Ar_2 and He_2).

Because the grid errors change in a (rapid) oscillatory manner along the nuclear coordinates, it is more wisely to quantify those with RMSE rather than investigate the error of a single geometry. If the observed grid-related oscillation would consist only of one particular frequency (i.e., be a single sinusoid) and if within the displacement range there would be an integer number of cycles, the Maximum Absolute Error (MaxAbsErr) would be $\sqrt{2}$ RMSE (which is a well-known property for the root mean square of a single sinusoid). In reality, the grid-related spurious oscillations consist of a broad range of dispersed frequencies and they have irregular shapes. Figure 9.9 presents the distribution of the MaxAbsErr/RMSE ratio among the studied systems. The ratio MaxAbsErr/RMSE is on average 1.87 with a standard deviation 0.47 (i.e., MaxAbsErr is about twice larger than RMSE) within $\{\boldsymbol{\xi_0}\}$. For example, if for a property one obtains RMSE = $1{\cdot}10^{-3}$ au, the maximum error coming from the spurious oscillations within the $\{\boldsymbol{\xi_0}\}$ range is expected to be about $2{\cdot}10^{-3}$ au.

RMS and RMSE (with opposite signs) are plotted in Figure 9.10. In this example, the curve obtained with ωB97X and the (750, 974) grid for HCN$\cdot$BrH, $P^{\mathrm{DFA(LG)}} = P^{\mathrm{DFA(750,974)}}$ (purple color in Figure 9.10) still suffers from spurious oscillations. The spurious oscillations are qualitatively filtered out in $P^{\mathrm{DFA}}_{\mathrm{filt}}$ (green color in Figure 9.10). Curves obtained with other grids, $P^{\mathrm{DFA(G)}} = P^{\mathrm{DFA(99,590)}}$, and $P^{\mathrm{DFA(250,974)}}$, are also shown for comparison. The difference between the latter curves and $P^{\mathrm{DFA}}_{\mathrm{filt}}$ yields the red curves in the plots of Figure 9.10, which represent the spurious oscillations in $P^{\mathrm{DFA(G)}}$ within the $\{\boldsymbol{\xi_0}\}$ range.

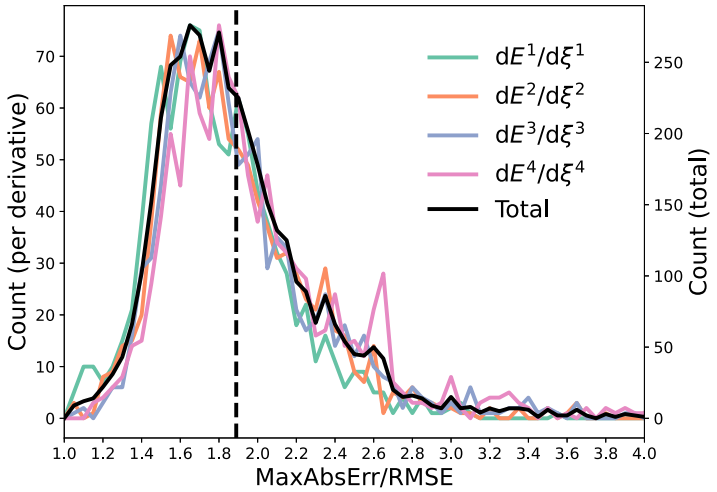

Figure 9.9: Distribution of values of the ratio between the maximum absolute error (MaxAbsErr) and the Root Mean Square Error (RMSE) measured over the same range of interatomic displacements for each property, with the data bin interval of 0.05, for different order of the $d^m E/d\xi^m$ derivatives (m = 1 - 4), as well as their overall sum (labeled as Total). The values included all tested DFAs, three integration grids, and all molecular complexes. The vertical dash line marks the average value of MaxAbsErr/RMSE (1.87).

9.3.2 Spurious Oscillations and the Errors in Vibrational Properties

It is difficult to anticipate the final effect of the spurious oscillations on the vibrational contributions to the molecular properties because they depend on various derivatives in quite different ways. For example, for the i-th normal mode, the corresponding harmonic vibrational frequency, ω_{har} is defined as the second order diagonal derivative of total energy. The corresponding harmonic IR and Raman intensities, IR_{har} and RA_{har}, are proportional to the first derivative of the dipole moment and polarizability with respect to the selected normal mode, respectively. On the contrary, the anharmonic corrections to vibrational frequencies, IR and Raman intensities, which are commonly evaluated at the second-order vibrational perturbation theory (VPT2) developed by Barone and coworkers [355, 356], involve many crossed terms between derivatives of different types. For example, the anharmonic corrections to vibrational frequencies, ω_{anh} involve mixed third and fourth order partial derivatives of the energy, and usually one does not further track the values coming solely from the diagonal derivatives. The anharmonic contributions to IR and Raman intensities, IR_{anh} and RA_{anh}, have very a complicated form. In the corresponding formulae [355, 356], one can find terms depending solely on the mixed third-order derivatives of dipole moment (or polarizability), several cross terms between mixed third-order derivatives of the energy and the second-order derivatives of the dipole moment (or polarizability), and the cross term involving the mixed fourth-order derivatives of the energy and the first-order derivatives of the dipole moment (or polarizability). Moreover, almost all these terms are multiplied by scale factors depending on various combinations of harmonic frequencies (usually in the denominators).

RRMSE is a convenient measure of the spurious oscillations because it only requires the computation of the derivative profile at the DFT level and a cost-efficient *ab initio* reference such as HF. However, one should still establish the possible connection between the RRMSE values and the potential errors in the vibrational properties. For that means, in Table 9.5, we collect the RRMSE values and relative errors for the selected vibrational spectroscopic properties of $N_2 \cdot$HF. It is important to stress that while errors in vibrational properties correspond to a single optimized geometry of the system, RRMSE measures the

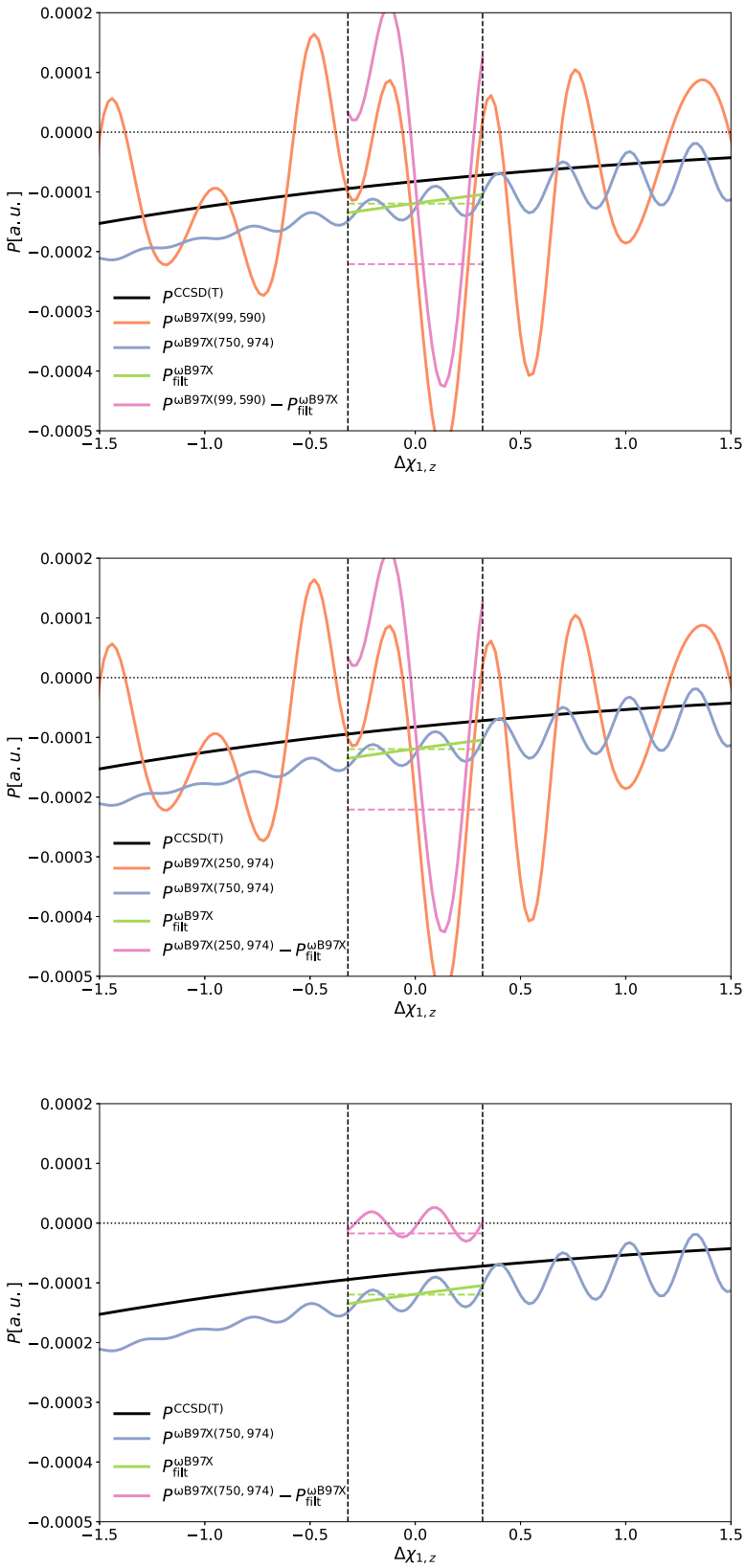

Figure 9.10: The $P^{\mathrm{DFA(G)}}$, $P^{\mathrm{DFA(LG)}}$ and $P^{\mathrm{DFA}}_{\mathrm{filt}}$ curves of $\mathrm{d}^3E/\mathrm{d}{\chi_{1,z}}^3$ of the HCN · BrH system. Calculated with ωB97X/aug-cc-pVTZ and three tested grids G= (99, 590) *(top)*, with RMSE = $2.21 \cdot 10^{-4}$ and RRMSE = 184.7%; (250, 974) *(middle)*, with RMSE = $1.51 \cdot 10^{-4}$ and RRMSE = 126.3%; (750, 974) *(bottom)*, with RMSE = $1.72 \cdot 10^{-5}$ and RRMSE = 14.4%. In all computations, LG = (750, 974) was utilized. Dashed colored horizontal limes mark the negative value of RMSE (pink) and RMS of the filtered profile (green). Dashed black vertical lines mark the studied range of the property curve $\{\boldsymbol{\xi_0}\}$, for which RMS, RMSE, and RRMSE are calculated.

errors for a set of slightly distorted geometries (along a chosen displacement coordinate). The numbers show that the relative errors in the property might be as large as twice the RRMSE value. For instance, 17% RRMSE in $\mathrm{d}\alpha_{zz}/\mathrm{d}\xi$ gives a relative error on the harmonic Raman intensity of 38%; 64% RRMSE in $\mathrm{d}^2\alpha_{zz}/\mathrm{d}\xi^2$ gives a relative error on the anharmonic correction to the Raman intensity of 163%. Certain properties depend on more than one derivative and the RRMSE of all the derivatives need to be considered. For instance, RRMSEs of 16% and 665% for $\mathrm{d}^3E/\mathrm{d}\xi^3$ and $\mathrm{d}^4E/\mathrm{d}\xi^4$, respectively, must be considered to explain the 234% error on the calculation of IR anharmonic corrections to the intermolecular stretching mode of $N_2 \cdot HF$ using ωB97X with the (250, 974) grid. Therefore, the large RRMSE values for $\mathrm{d}^mE/\mathrm{d}\xi^m$, $\mathrm{d}^m\mu_z/\mathrm{d}\xi^m$, and $\mathrm{d}^m\alpha_{zz}/\mathrm{d}\xi^m$ indicate significant spurious oscillations and may warn about the presence of potentially large errors in vibrational properties.

Table 9.5: RRMSE of various derivatives of E, μ_z, and α_{zz} with respect to the floppy intermolecular stretching mode Q_3 of $N_2 \cdot HF$ using three integration grids: G1 - (99, 590), G2 - (250, 974) and G3 - (500, 974). The last 6 columns contain the relative errors (in %) of the selected (an)harmonic vibrational properties: ω_{har} and $\Delta\omega_{\text{anh}}$ — harmonic vibrational frequency and its anharmonic correction; IR_{har} and $\Delta\text{IR}_{\text{anh}}$ — harmonic IR intensities and its anharmonic correction; RA_{har} and RA_{anh} — harmonic Raman intensities and its anharmonic correction. The error estimations were obtained with the (750, 974) grid results used as the reference.

	RRMSE [%]									Prop. Rel. Err. [%]					
	$d^2E/d\chi_{1,z}{}^2$			$d^3E/d\chi_{1,z}{}^3$			$d^4E/d\chi_{1,z}{}^4$			$\omega_{\text{har}}(Q_3)$			$\Delta\omega_{\text{anh}}(Q_3)$		
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
B3LYP	0	0	0	0	0	0	4	0	0	0	0	0	2	0	0
CAM-B3LYP	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0
N12	7	1	0	119	31	2	1973	874	70	2	0	0	625	175	0
M062X	2	0	0	21	3	0	388	70	1	1	0	0	62	14	0
ωB97X	3	1	0	52	16	1	1447	665	48	2	0	0	195	234	9

	RRMSE [%]									Prop. Rel. Err. [%]					
	$d^1\mu_z/d\chi_{1,z}{}^1$			$d^2\mu_z/d\chi_{1,z}{}^2$			$d^3\mu_z/d\chi_{1,z}{}^3$			$\text{IR}_{\text{har}}(Q_3)$			$\Delta\text{IR}_{\text{anh}}(Q_3)$		
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
B3LYP	0	0	0	0	0	0	7	0	0	0	0	0	1	0	0
CAM-B3LYP	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0
N12	1	0	0	56	14	1	4548	1983	148	2	0	0	303	62	4
M062X	0	0	0	12	1	0	373	42	1	0	0	0	9	1	0
ωB97X	0	0	0	26	9	1	1650	794	50	2	0	0	31	126	4

	RRMSE [%]						Prop. Rel. Err. [%]					
	$d^1\alpha_{zz}/d\chi_{1,z}{}^1$			$d^2\alpha_{zz}/d\chi_{1,z}{}^2$			$\text{RA}_{\text{har}}(Q_3)$			$\Delta\text{RA}_{\text{anh}}(Q_3)$		
	G1	G2	G3	G1	G2	G3	G1	G2	G3	G1	G2	G3
B3LYP	0	0	0	5	0	0	0	0	0	6	0	0
CAM-B3LYP	0	0	0	2	0	0	0	0	0	4	0	0
N12	17	4	0	8588	3726	257	38	11	1	32656	15921	192
M062X	5	0	0	907	114	29	11	0	0	171	110	9
ωB97X	12	4	0	2103	973	64	15	0	0	60435	915	163

9.4 Analysis of Relative Errors from the Spurious Oscillations

After showing that RRMSE is a meaningful indicator of the errors in the vibrational properties due to the spurious oscillations, we can use it to quantify the errors for different property derivatives, systems, and DFAs. An example of such analysis is given in Table 9.6, which compiles the RRMSE values for various derivatives of the HCN·BrH system and a small group of selected functionals. In general, the errors are the largest for the smallest grid and higher-order derivatives. Errors arising from the spurious oscillations are usually not observable in the curves of $E(R)$, $\mu_z(R)$ and $\alpha_{zz}(R)$. However, for the smallest grid tested, namely (99, 590), RRMSE may reach already substantial values for the gradients of the energy (this is mostly due to the fact that the values of gradients are very close to zero, because we are studying slightly displaced optimized geometries). Nevertheless, for the second derivatives of energy, ωB97X, SCAN, and M06-2X combined with the (99, 590) grid yield RRMSE already larger than 10%. For the same functionals, one observes the largest RRMSE values for higher-order derivatives, $\mathrm{d}^4E/\mathrm{d}\xi^4$, reaching thousands of percent. Similarly, one observes propagation of errors for $\mathrm{d}^m\mu_z/\mathrm{d}\xi^m$ and $\mathrm{d}^m\alpha_{zz}/\mathrm{d}\xi^m$, although this increase seems to be much quicker than in the case of $\mathrm{d}^mE/\mathrm{d}\xi^m$ (roughly, one derivative "faster"). Some of the functionals, such as BH&H, B3LYP, and CAM-B3LYP may yield small RRMSE when the (250, 974) grid is used. However, in the case of the latter two, when those are combined with the (99, 590) grid one can observe RRMSE in the range 17- 74% for the highest derivatives, namely $\mathrm{d}^4E/\mathrm{d}\xi^4$, $\mathrm{d}^3\mu_z/\mathrm{d}\xi^3$, and $\mathrm{d}^2\alpha_{zz}/\mathrm{d}\xi^2$

To depict the true extent of the problem, for each DFA and integration grid, we collect the maximum RRMSE values within the set of all systems. Figure 9.11 graphically depict those errors found for $\mathrm{d}^mE/\mathrm{d}\xi^m$, Figure 9.12 for $\mathrm{d}^m\mu_z/\mathrm{d}\xi^m$, and Figure 9.13 for $\mathrm{d}^m\alpha_{zz}/\mathrm{d}\xi^m$. In most of the cases, the maximum values of errors are found for HCN·BrH, Ar_2, He_2, as these systems turned out to be the most grid demanding.

The (99, 590) grid — which is equal or larger than the largest predefined grid in QChem 5.1 [322] or ORCA 5.0.1 [357–359] — gives maximum RRMSE values that exceed 10%, 40%, and 500%, respectively, for $\mathrm{d}^2E/\mathrm{d}\xi^2$, $\mathrm{d}^3E/\mathrm{d}\xi^3$, and $\mathrm{d}^4E/\mathrm{d}\xi^4$. The latter is a clear indication that (99, 590) is definitely an insufficient grid to compute vibrational spectroscopic data for molecules with at least a low-frequency vibrational mode. For the (250, 974) grid, the maximum values of RRMSE for $\mathrm{d}^4E/\mathrm{d}\xi^4$ are greater than 160% for all DFAs, except for BH&H (18%) and LC-BLYP (11%). The largest grid, (750, 974), reduces the maximum RRMSE significantly for the first and second energy derivatives but some popular DFAs —such as ωB97X, SCAN, or M06-2X— give extremely large errors for the third and fourth derivatives; in fact, even with the largest grid, only a handful of functionals gives a maximum RRMSE below 50% for $\mathrm{d}^4E/\mathrm{d}\xi^4$. Grid-dependence may look slightly less problematic for the computations of $\mathrm{d}^m\mu_z/\mathrm{d}\xi^m$ (see Figure 9.12). For example, for the (250, 974) grid, most of the functionals yield maximum RRMSE below 10% for $\mathrm{d}^2\mu_z/\mathrm{d}\xi^2$ and 50% for $\mathrm{d}^3\mu_z/\mathrm{d}\xi^3$. However, one has to bear in mind that Ar_2 and He_2 could not be included in the statistical analysis of $\mathrm{d}^m\mu_z/\mathrm{d}\xi^m$ (Ar_2 and He_2 are centrosymmetric, hence $\mathrm{d}^m\mu_z/\mathrm{d}\xi^m$ is always zero). In the case of $\mathrm{d}^m\alpha_{zz}/\mathrm{d}\xi^m$, grid errors enhance very quickly. For half of the DFAs combined with the (250, 974) grid, maximum RRMSEs exceed 100% already for $\mathrm{d}\alpha_{zz}/\mathrm{d}\xi$ (see Figure 9.13).

Table 9.6: The RRMSE values (in percentages) for various derivatives of energy, dipole moment, and static polarizability, shown for the selected DFAs and three integration grids. Obtained for the HCN · BrH system and the displacements along the $\chi_{1,z}$ coordinate.

		$d^m E/d\chi_{1,z}^m$					$d^m \mu_z/d\chi_{1,z}^m$				$d^m \alpha_{zz}/d\chi_{1,z}^m$		
Method	Grid	0	1	2	3	4	0	1	2	3	0	1	2
BH&H	(99, 590)	0.0	0.4	0.2	1	7	0.0	0.0	0	0	0.0	0.0	0
	(250, 974)	0.0	0.0	0.0	0	0	0.0	0.0	0	0	0.0	0.0	0
	(750, 974)	0.0	0.0	0.0	0	1	0.0	0.0	0	0	0.0	0.0	0
CAM-B3LYP	(99, 590)	0.0	0.4	0.3	2	17	0.0	0.2	3	43	0.0	0.3	11
	(250, 974)	0.0	0.0	0.0	0	5	0.0	0.0	0	4	0.0	0.0	2
	(750, 974)	0.0	0.0	0.0	0	3	0.0	0.0	0	0	0.0	0.0	0
LC-ωHPBE	(99, 590)	0.0	10.8	3.2	25	104	0.0	1.0	11	215	0.0	2.9	36
	(250, 974)	0.0	0.0	0.0	0	6	0.0	0.0	0	3	0.0	0.0	0
	(750, 974)	0.0	0.0	0.0	0	5	0.0	0.0	0	2	0.0	0.0	0
B3LYP	(99, 590)	0.0	1.7	0.6	4	20	0.0	0.3	9	74	0.0	0.6	13
	(250, 974)	0.0	0.1	0.1	1	13	0.0	0.0	0	9	0.0	0.1	4
	(750, 974)	0.0	0.0	0.0	0	0	0.0	0.0	0	0	0.0	0.0	0
MN15	(99, 590)	0.0	3.5	2.5	11	180	0.0	1.6	38	328	0.0	2.4	69
	(250, 974)	0.0	0.0	0.0	0	6	0.0	0.0	0	1	0.0	0.0	0
	(750, 974)	0.0	0.0	0.0	0	6	0.0	0.0	0	0	0.0	0.0	0
TPSSh	(99, 590)	0.0	20.8	10.4	53	455	0.0	2.8	23	482	0.0	4.3	74
	(250, 974)	0.0	0.3	0.5	5	100	0.0	0.0	1	21	0.0	0.3	14
	(750, 974)	0.0	0.0	0.0	0	7	0.0	0.0	0	1	0.0	0.0	0
MN15-L	(99, 590)	0.0	8.5	6.6	29	374	0.0	8.3	103	1542	0.1	7.7	171
	(250, 974)	0.0	0.0	0.0	0	5	0.0	0.0	0	4	0.0	0.5	33
	(750, 974)	0.0	0.0	0.0	0	8	0.0	0.0	0	3	0.0	0.3	25
ωB97X	(99, 590)	0.0	20.9	18.0	185	5147	0.0	5.2	163	7145	0.1	28.9	2509
	(250, 974)	0.0	4.8	7.9	126	3393	0.0	1.5	112	6359	0.0	18.4	1517
	(750, 974)	0.0	0.1	0.4	14	765	0.0	0.2	22	1972	0.0	1.8	302
SCAN	(99, 590)	0.0	19.7	14.9	129	3908	0.0	2.1	99	7689	0.1	33.2	1456
	(250, 974)	0.0	1.6	2.7	48	1993	0.0	0.9	54	3241	0.0	6.6	573
	(750, 974)	0.0	0.1	0.5	21	1002	0.0	0.1	10	916	0.0	4.6	551
M06-2X	(99, 590)	0.0	31.4	11.9	72	1068	0.0	8.6	117	4339	0.0	10.4	622
	(250, 974)	0.0	4.6	4.9	58	780	0.0	0.5	23	881	0.0	7.3	370
	(750, 974)	0.0	0.0	0.0	0	0	0.0	0.0	0	11	0.0	0.0	1

9.5 Limits in the Computation of Property Derivatives. Ranking of DFA's Grid-Stability

Since the errors arising from the spurious oscillations may increase dramatically fast and reach thousands of percent, it seems convenient to ask a pragmatic question: up to which order of derivative one can obtain small errors caused by the spurious oscillations for a given DFA combined with the selected grid? In other words, to which derivative order a given DFA yields reliable values? To this end, we defined a *safe* calculation of the property derivative as a calculation with RRMSE value below 10%. This value was chosen heuristically through numerical tests, which indicated that the maximum relative errors in the derivatives (measured with respect to RMS of the filtered profile) are usually below 20% (see Figure 9.9). For each tested DFA, integration grid, and molecular system, we established the maximum order of derivatives which can be safely obtained. Derivatives up to the fourth, third, and second order along the nuclear displacement were considered for the energy and the main symmetry axis component of the dipole moment and static polarizability, respectively. The results of this analysis for DFAs combined with the (250, 974) grid are compiled in Tables 9.7 to 9.9 (each table contains the results for the derivatives of different property) An equivalent analysis was done for the (99, 590) and (750, 974) grids, however, those results are shown only in the summarized version in Figures A1.1 and A1.2.

One can observe a similar behavior of DFAs for the systems characterized with the same type of intermolecular interaction, namely hydrogen-bonded (group A), halogen-bonded (group B), or dispersion-bonded (group C) molecules. In general, the highest derivatives are safely reached for the hydrogen-bonded complexes. This means that these systems, although still grid-wise challenging for a lot of DFAs, they suffer from the spurious oscillations the least. The halogen-bonded molecules exhibit much more significant oscillations hence one can safely reach only lower derivatives. The dispersion-bonded systems are the species affected by this problem the most. Moreover, within the same group of molecules, the magnitude of the spurious oscillations increases in the following order: $\mathrm{d}^m E/\mathrm{d}\xi^m$, $\mathrm{d}^m \mu_z/\mathrm{d}\xi^m$, $\mathrm{d}^m \alpha_{zz}/\mathrm{d}\xi^m$. These observations allowed to reduce the dimensionality of the analysis and discuss only the worst cases among the whole group of molecules. These results are compiled in Figures 9.14 for the (250, 974) grid and in Figures A1.1, and A1.2, for the (99, 590) and (750, 974) grids, respectively. For example, for the energy derivatives calculated with MN12-SX(250, 974), the smallest number within groups A and B is 3, and within group C is 1 (see Table 9.7).

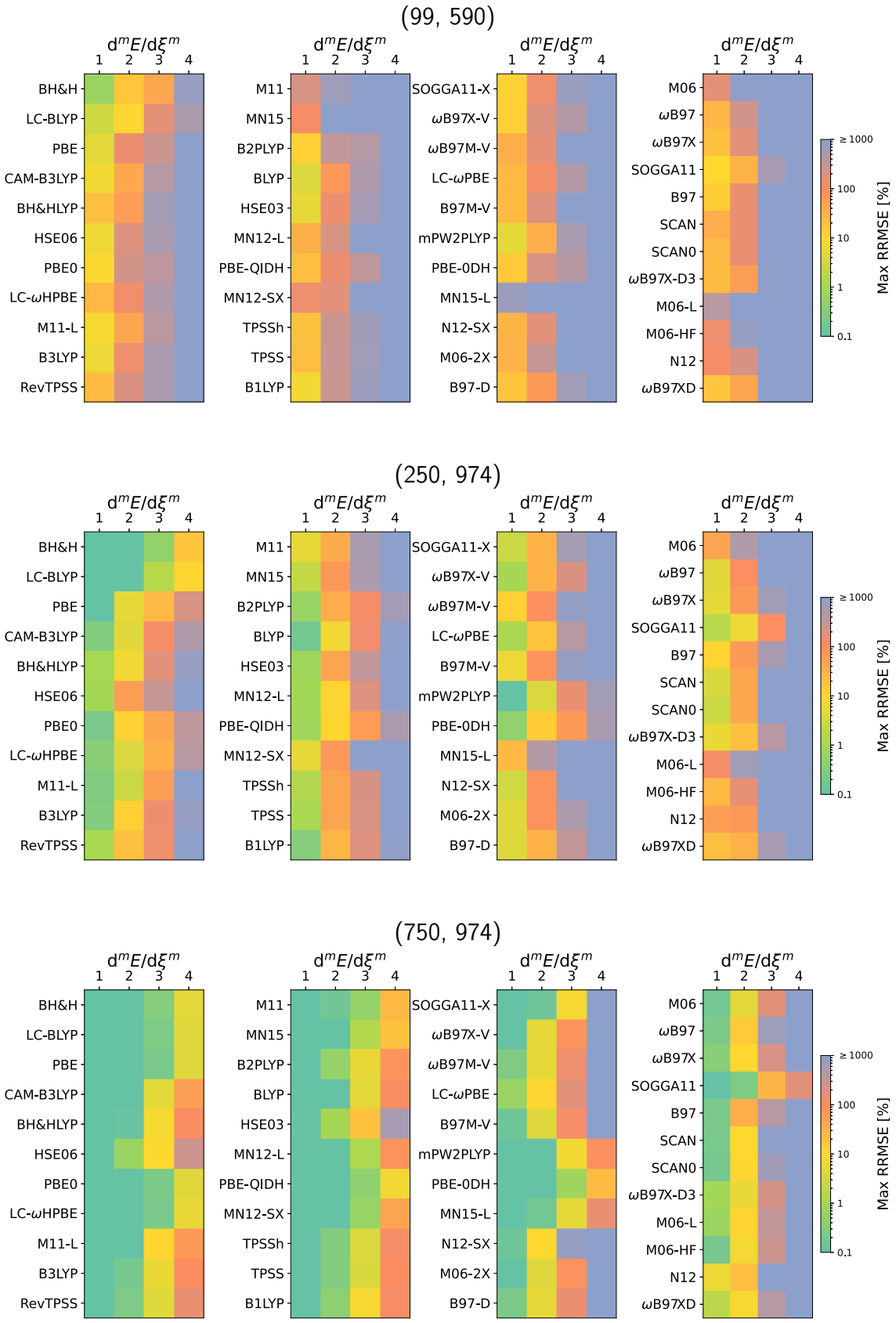

Figure 9.11: Maximum values of RRMSE for the derivatives of the total energy found within the set of all molecular complexes, obtained with the (99, 590) grid (top panel), the (250, 974) grid (middle panel), and the (750, 974) grid (bottom panel).

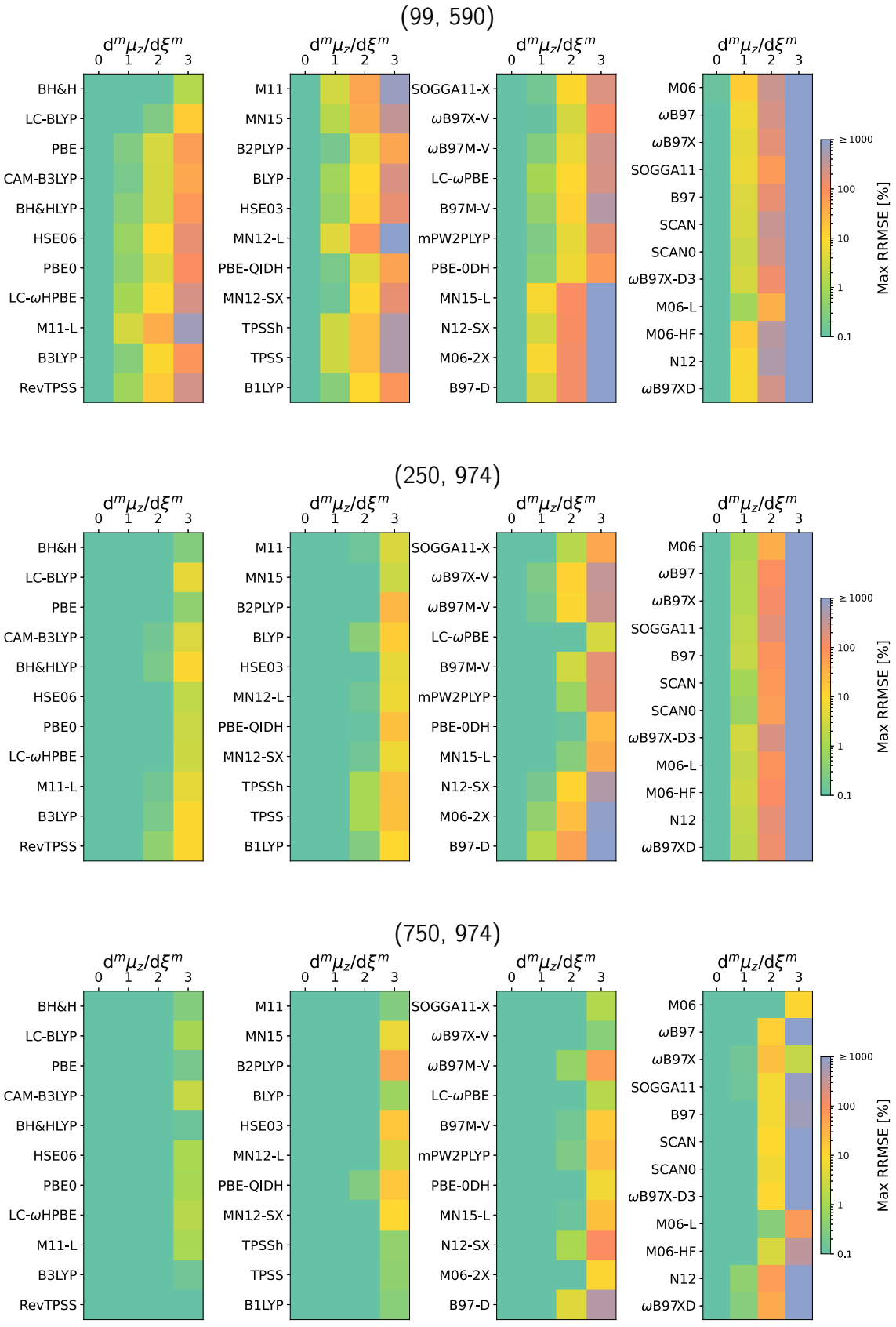

Figure 9.12: Maximum values of RRMSE for the derivatives of μ_z found within the set of hydrogen- and halogen-bonded molecular complexes, obtained with the (99, 590) grid (top panel), the (250, 974) grid (middle panel), and the (750, 974) grid (bottom panel).

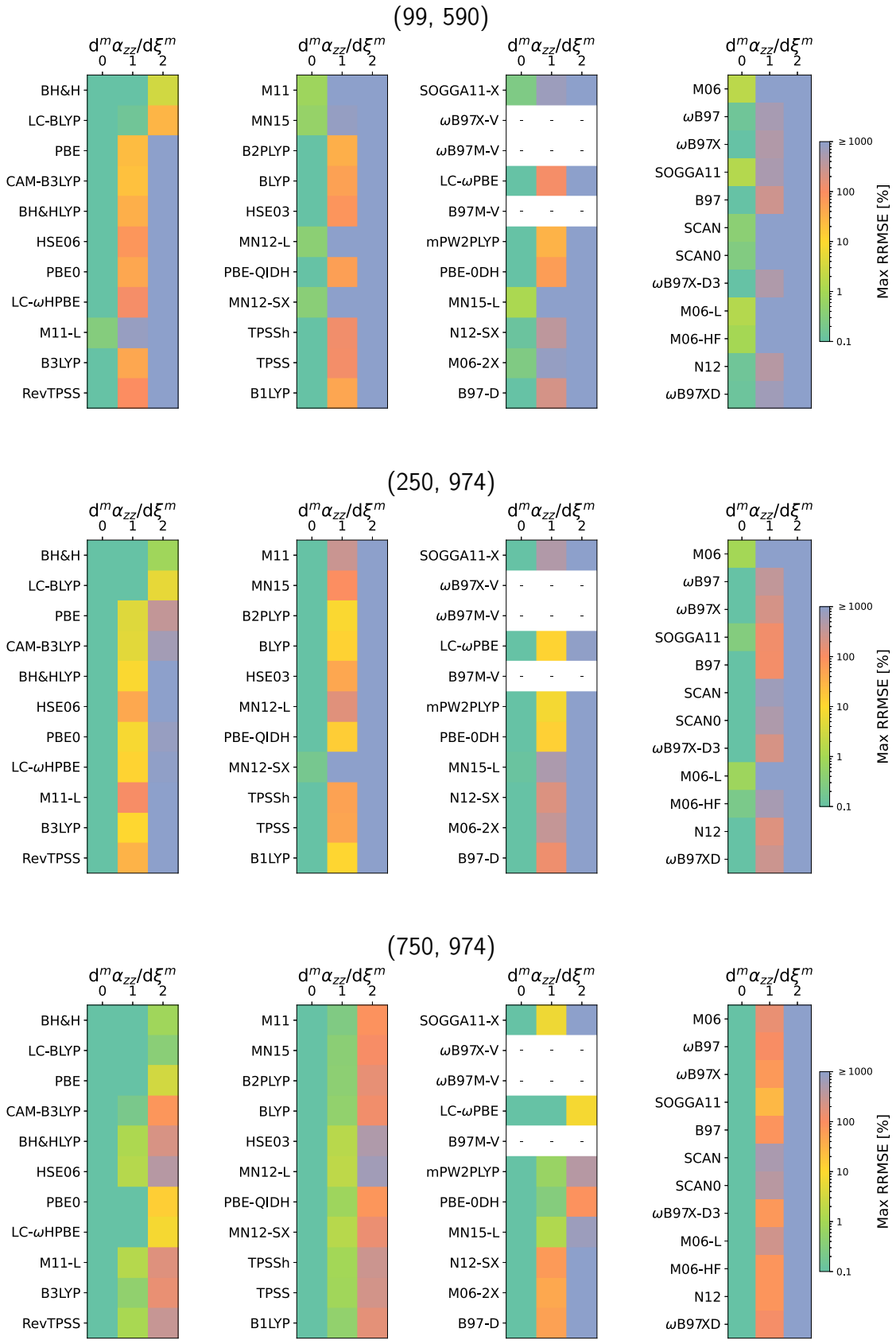

Figure 9.13: Maximum values of RRMSE for the derivatives of α_{zz} found within the set of all molecular complexes, obtained with the (99, 590) grid (top panel), the (250, 974) grid (middle panel), and the (750, 974) grid (bottom panel). Data for $d^m\alpha_{zz}/d\xi^m$ is not available for ωB97X-V, ωB97M-V and B97M-V, for which static polarizabilities are not implemented in QChem 5.1.

Table 9.7: Maximum order m of $\mathrm{d}^m E/\mathrm{d}\xi^m$ derivatives (m = 0 - 4), for which RRMSE was lower than 10%. All calculations use the (250, 974) grid.

	HCN · HF	HCN · HCl	OC · HF	N_2 · HF	HCN · BrH	HCN · BrF	Ar_2	He_2
BH&H	4	4	4	4	4	4	4	3
LC-BLYP	4	4	4	4	4	4	4	3
PBE	4	4	4	4	4	4	4	2
CAM-B3LYP	4	4	4	4	4	4	3	2
BH&HLYP	4	4	4	4	3	4	3	2
HSE06	4	4	4	4	4	4	3	1
PBE0	4	4	4	4	4	4	4	1
LC-ωHPBE	4	4	4	4	4	4	4	2
M11-L	3	3	3	3	3	4	2	2
B3LYP	4	4	4	4	3	4	3	1
RevTPSS	4	4	4	4	3	4	3	1
M11	3	3	3	3	4	3	2	1
MN15	4	4	4	4	4	4	3	1
B2PLYP	4	4	4	4	4	4	3	1
BLYP	4	4	4	4	3	4	3	2
HSE03	3	3	3	3	2	3	3	1
MN12-L	4	4	4	4	4	3	2	1
PBE-QIDH	4	4	4	4	4	4	4	1
MN12-SX	4	4	3	4	4	3	2	1
TPSSh	4	4	4	4	3	3	3	1
TPSS	4	4	4	4	3	4	3	1
B1LYP	4	4	4	4	3	4	3	1
SOGGA11-X	4	4	4	4	3	4	2	1
ωB97X-V	4	4	4	4	3	3	2	1
ωB97M-V	4	3	4	4	3	3	2	0
LC-ωPBE	2	3	3	2	2	2	2	1
B97M-V	4	3	3	4	3	3	3	1
mPW2PLYP	4	3	4	3	3	3	3	2
MN15-L	4	4	4	4	4	4	3	0
PBE-0DH	4	4	4	4	4	3	4	1
N12-SX	4	3	3	3	2	3	1	1
M06-2X	3	3	3	3	2	2	1	1
B97-D	3	3	3	2	2	3	2	1
M06	4	-	3	3	2	3	1	0
ωB97	3	2	3	2	2	2	1	1
ωB97X	3	2	3	2	2	2	1	1
SOGGA11	3	2	3	2	2	4	2	3
B97	3	3	3	3	0	2	1	1
SCAN	2	2	2	2	2	2	1	1
SCAN0	2	2	2	2	2	2	1	1
ωB97X-D3	3	2	3	2	2	2	1	1
M06-L	3	3	3	3	0	3	1	0
M06-HF	3	3	3	2	0	2	1	0
N12	3	2	3	2	0	2	1	1
ωB97XD	3	2	3	2	0	2	1	1

Table 9.8: Maximum order m of $d^m\mu_z/d\xi^m$ derivatives (m = 0 - 3), for which RRMSE was lower than 10%. All calculations use the (250, 974) grid.

	$HCN \cdot HF$	$HCN \cdot HCl$	$OC \cdot HF$	$N_2 \cdot HF$	$HCN \cdot BrH$	$HCN \cdot BrF$
BH&H	3	3	3	3	3	3
LC-BLYP	3	3	3	3	3	3
PBE	3	3	3	3	3	3
CAM-B3LYP	3	3	3	3	3	3
BH&HLYP	3	3	3	3	3	3
HSE06	3	3	3	3	3	3
PBE0	3	3	3	3	3	3
LC-ωHPBE	3	3	3	3	3	3
M11-L	3	3	3	3	3	3
B3LYP	3	3	3	3	3	3
RevTPSS	3	3	3	3	3	3
M11	3	3	3	3	3	3
MN15	3	3	3	3	3	3
B2PLYP	2	3	3	3	3	2
BLYP	3	3	3	3	2	3
HSE03	3	3	3	3	3	3
MN12-L	3	3	3	3	3	3
PBE-QIDH	2	2	3	3	3	3
MN12-SX	3	3	3	3	3	3
TPSSh	3	3	3	3	2	3
TPSS	3	3	3	3	2	3
B1LYP	3	3	3	3	2	3
SOGGA11-X	3	3	3	3	2	3
ωB97X-V	3	3	3	3	1	3
ωB97M-V	3	3	3	3	2	2
LC-ωPBE	3	3	3	3	3	3
B97M-V	3	3	2	3	2	2
mPW2PLYP	3	2	2	2	2	2
MN15-L	2	3	3	3	3	2
PBE-0DH	2	3	3	3	3	2
N12-SX	2	2	2	2	1	2
M06-2X	3	2	2	2	1	2
B97-D	2	2	2	2	1	2
M06	3	-	2	2	1	2
ωB97	2	2	2	2	1	1
ωB97X	2	2	2	2	1	1
SOGGA11	1	2	3	2	1	2
B97	2	2	2	2	1	2
SCAN	1	1	1	1	1	1
SCAN0	1	1	1	1	1	1
ωB97X-D3	2	2	2	1	1	1
M06-L	2	2	2	2	1	3
M06-HF	2	2	2	2	1	2
N12	2	2	2	1	1	1
ωB97XD	2	2	2	1	1	1

Table 9.9: Maximum order m of the $d^m\alpha_{zz}/d\xi^m$ derivatives (m = 0 - 2), for which RRMSE was lower than 10%. All calculations use the (250, 974) grid.

	HCN · HF	HCN · HCl	OC · HF	N_2 · HF	HCN · BrH	HCN · BrF	Ar_2	He_2
BH&H	2	2	2	2	2	2	2	2
LC-BLYP	2	2	2	2	2	2	2	2
PBE	2	2	2	2	2	2	1	1
CAM-B3LYP	2	2	2	2	2	2	1	1
BH&HLYP	2	2	2	2	2	2	1	1
HSE06	2	2	2	2	2	2	1	0
PBE0	1	2	2	2	2	2	1	1
LC-ωHPBE	1	2	2	2	2	2	1	0
M11-L	2	2	2	2	2	2	1	0
B3LYP	2	2	2	2	2	2	1	0
RevTPSS	2	2	2	2	2	2	1	0
M11	2	2	2	2	2	2	1	0
MN15	1	2	2	2	2	2	1	0
B2PLYP	2	2	2	2	2	2	1	1
BLYP	2	2	2	2	2	2	1	0
HSE03	2	2	2	2	2	2	1	0
MN12-L	1	2	2	2	2	2	1	0
PBE-QIDH	1	1	2	2	2	2	1	0
MN12-SX	1	2	2	2	2	2	0	0
TPSSh	2	2	2	2	1	2	1	0
TPSS	2	2	2	2	1	2	1	0
B1LYP	1	2	2	2	2	2	1	0
SOGGA11-X	2	2	2	2	1	2	1	0
ωB97X-V	-	-	-	-	-	-	-	-
ωB97M-V	-	-	-	-	-	-	-	-
LC-ωPBE	1	2	2	2	2	2	1	0
B97M-V	-	-	-	-	-	-	-	-
mPW2PLYP	2	2	1	1	1	1	1	1
MN15-L	1	2	2	2	1	1	1	0
PBE-0DH	1	2	2	2	1	2	1	0
N12-SX	2	1	1	1	1	2	0	0
M06-2X	2	1	1	1	1	2	0	0
B97-D	1	1	1	1	1	2	0	0
M06	1	-	2	1	1	2	0	0
ωB97	1	1	1	1	0	1	0	0
ωB97X	1	1	1	1	0	1	0	0
SOGGA11	0	1	1	1	1	1	0	0
B97	2	1	1	1	0	1	0	0
SCAN	0	0	0	0	1	1	0	0
SCAN0	0	1	0	0	1	1	0	0
ωB97X-D3	1	1	1	1	0	1	0	0
M06-L	1	1	1	1	0	1	0	0
M06-HF	2	1	1	1	0	1	0	0
N12	1	1	1	1	0	1	0	0
ωB97XD	2	1	1	1	0	1	0	0

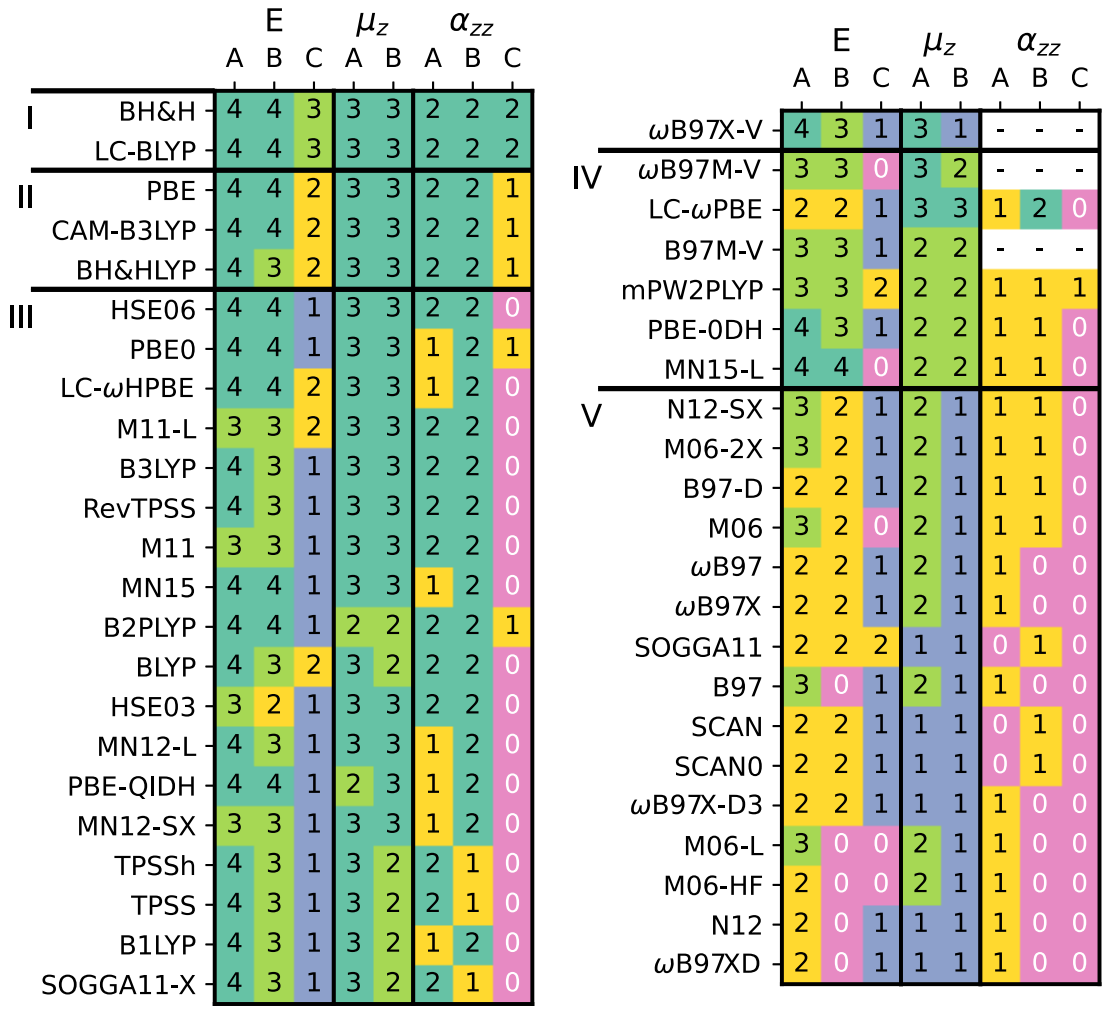

Figure 9.14: Maximum order of derivatives that can be safely obtained (i.e., RRMSE$\leq$ 10%) with the (250, 974) grid. The maximum order of the derivatives computed is fourth for the energy, third for the dipole moment, and second for the polarizability. Groups A, B, and C, correspond to hydrogen-, halogen-, and dispersion-bonded complexes, respectively. Data not available for ωB97X-V, ωB97M-V, and B97M-V, for which static polarizabilities are not implemented in QChem 5.1.

In Figure A1.1, one can observe that modern DFAs with the (99, 590) integration grid cannot be safely used to calculate high-order derivatives needed to describe the vibrational structure of the highly anharmonic systems. This is very surprising because the integration grid of the corresponding size is the default choice in the most popular computational packages (UltraFine in Gaussian [296] and SG3 in QChem [322]). We ranked the grid-errors of DFAs when combined with (250, 974) grid, accordingly to the following weight:

$$w^{\mathrm{DFA}} = \frac{1}{3}\left(\frac{1}{12}\sum_{G}^{\mathrm{A,B,C}}\max_{i\in G}(n_{E_i}) + \frac{1}{9}\sum_{G}^{\mathrm{A,B}}\max_{i\in G}(n_{\mu_{z,i}}) + \frac{1}{6}\sum_{G}^{\mathrm{A,B,C}}\max_{i\in G}(n_{\alpha_{zz,i}})\right), \tag{9.4}$$

where n_{E_i}, $n_{\mu_{z,i}}$ and $n_{\alpha_{zz,i}}$ are, respectively, the lowest maximum order of the derivative that can be safely calculated with that DFA (RRMSE$\leq$ 10%) for the energy, the dipole moment, and the polarizability of molecule i. If the weight is one for a given DFA, it means that such DFA can be used to safely compute all the high-order derivatives of E, μ_z, and α_{zz}.

We classified the DFAs into five rungs according to their overall stability across various molecular properties.

Rung I includes grid-robust DFAs, namely, BH&H and LC-BLYP. These DFAs provide safe results for almost all derivatives using the (250, 974) grid. BH&H is the only functional in our set which does not employ a GGA or meta-GGA exchange. In practice, BH&H is seldom employed because it is not accurate for most chemical purposes (this DFA is not even considered in two of the most extensive recent benchmark studies [360, 361]). Conversely, LC-BLYP has been extensively used to study molecular properties and, hence, it is considered the safest choice among the DFAs covered in this study.

DFAs in *Rung II* —PBE, CAM-B3LYP, and BH&HLYP— suffer from some grid dependency. They can provide stable anharmonic vibrational corrections for the hydrogen- and halogen-bonded complexes, and harmonic contributions for the dispersion-bonded systems. On the contrary, they should not be employed to calculate anharmonic contributions of dispersion-bonded systems.

Rung III includes DFAs suffering from grid-dependent oscillations for all the studied systems. Widely used DFAs, such as B3LYP, PBE0, MN15, and ωB97X-V belong to this group. The DFAs in this rung cannot provide stable force constants or the Raman intensities (see $\mathrm{d}\alpha_{zz}/\mathrm{d}\xi$) of dispersion-bonded systems.

Rung IV encloses DFAs which are highly grid dependent. Most of these DFAs encounter serious difficulties to study molecules with low-frequency vibrations and, within this group of molecules, they are limited to the basic anharmonic corrections to frequencies and IR intensities (see $\mathrm{d}^3E/\mathrm{d}\xi^3$ and $\mathrm{d}^2\mu_z/\mathrm{d}\xi^2$) of hydrogen- and halogen-bonded complexes. We advise to use these DFAs with a great caution, as they may require giant integration grids to avoid spurious oscillations (see Figure A1.2).

DFAs with the highest grid dependency belong to *Rung V*. This group consists mostly of DFAs from the B97, M06, SCAN, and N12 families, including very popular DFAs such as M06-2X and ωB97-X. Within the set of molecules exhibiting low-frequency vibrations, most of *Rung V* DFAs are limited to the calculation of harmonic vibrational properties of hydrogen- and halogen-bonded complexes. Some DFAs, namely, B97, M06-L, M06-HF, N12, and ωB97XD, cannot provide stable energy or polarizability gradients for the halogen-bonded systems. Grids smaller than (250, 974) seriously compromise the calculation of energy gradients even for hydrogen-bonded complexes (see Figure S18). A much larger grid, such as (750, 974), increases the stability of some DFAs in *Rung V*, but even with this enormous grid, these DFAs cannot be trusted to calculate harmonic frequencies of dispersion-bonded complexes (see Figure S19).

9.6 Spurious Oscillations in Intramolecular Normal Modes

In the previous sections, it has been extensively shown that the grid-related spurious oscillation may dramatically affect noncovalently bounded systems. One can further ask if this problem can also affect the atomic motion within a single molecule.

This Section presents the first evidences of spurious oscillations affecting intramolecular normal modes. This is shown for the several systems: the allyl anion, phenanthrene, and H_2S_2 (in Appendix, additional results are shown for the butadiene, cyclobutadiene, benzene, naphthalene, and H_2O_2). For each of the system, the $\mathrm{d}^m E/\mathrm{d}Q^m$ and $\mathrm{d}^m \mu_\alpha/\mathrm{d}Q^m$ derivatives with respect to one selected normal mode of low-frequency (or two in the case of butadiene) were studied. We have tested only modes for which one could observe a noticeable change of distances between pair(s) of nuclei that are not covalently bonded, but are sufficiently close enough to interact with each other (and therefore they mimic the displacements studied in the noncovalently bonded complexes). All these modes, despite having low vibrational frequencies, are well defined with the harmonic approximation. The VPT2 anharmonic corrections to the vibrational frequencies, IR and Raman intensities are very small (as shown in Tables 9.10 to 9.14 for the VPT2 results obtained with HF). Since the purpose of this new set of calculations was to demonstrate the broad scope of the spurious oscillations problem, the analysis is limited to a few integration grids, a handful of DFAs and only the energy and dipole moment derivatives. In most of the studied examples, we tested only the ωB97X and M06-2X functionals combined with the (99, 590), (250, 974), and (750, 974) grids. Moreover, as the oscillation-free (ab-initio) reference, we chose the HF method. The property curves contain 300-400 points and are sampled with system-specifically chosen steps.

9.6.1 Allyl anion (Q_2 mode)

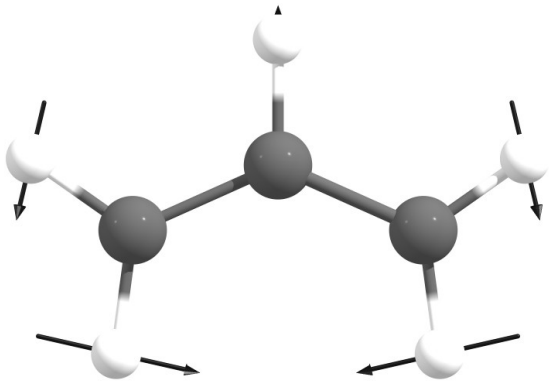

Figure 9.15: Graphical representation of the Q_2 (A_1) normal mode of allyl anion. Obtained at the M06-2X/aug-cc-pVDZ level of theory.

It is a smallest organic system in which we already noticed the effect of the spurious oscillations. We studied the effect of spurious oscillations in the computation of derivatives with respect to the Q_2 normal mode (see graphical representation in Figure 9.15). This mode belongs to the A_1 irreducible representation and corresponds to the second-lowest vibrational frequency. In this mode, in particular, two hydrogen atoms belonging to C1 and C3 atoms are getting closer. The Q_2 mode is very harmonic. The anharmonic corrections to vibrational frequency, and IR intensities account only for 1% of their corresponding total values (see Table 9.10).

Table 9.10: Vibrational properties of the Q_2 normal mode of the allyl anion obtained with HF, ωB97X, and M06-2X: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, IR_{har} and IR_{anh} — harmonic and anharmonic IR intensities, RA_{har} and RA_{anh} — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

	Q_2 (A_1) mode					
Method	$\omega_{\text{har}}\left[\frac{1}{\text{cm}}\right]$	$\omega_{\text{anh}}\left[\frac{1}{\text{cm}}\right]$	$\text{IR}_{\text{har}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{IR}_{\text{anh}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{RA}_{\text{har}}\left[\text{Å}^6\right]$	$\text{RA}_{\text{anh}}\left[\text{Å}^6\right]$
HF	457	459	2.983	3.019	0.705	0.264
ωB97X	427		2.561			
M06-2X	422		2.109			

For the selected ωB97X and M06-2X functionals, the spurious oscillations affect the high order derivatives $\mathrm{d}^3E/\mathrm{d}Q_2^3$, $\mathrm{d}^4E/\mathrm{d}Q_2^4$, and $\mathrm{d}^3\mu_z/\mathrm{d}Q_2^3$. Examples of these oscillations are depicted in Figure 9.16 and the RRMSE values are compiled in Table 9.11. When the (99, 590) grid is utilized, RRMSE are above 10% and reach 38% for $\mathrm{d}^3\mu_z/\mathrm{d}Q_2^3$ obtained using M06-2X. On the contrary, when the larger grids are applied, RRMSE errors are below 2% for ωB97X and 8% for M06-2X. While these errors are not as large as the ones observed in the previous sections, one can expect more dramatic errors in larger and more complicated systems.

Table 9.11: Values of RRMSE (in percentages) for the selected property derivatives with respect to the Q_2 normal mode of allyl anion, obtained for some selected DFAs and integration grids.

Method	Grid	$\mathrm{d}^3E/\mathrm{d}Q_2^3$	$\mathrm{d}^4E/\mathrm{d}Q_2^4$	$\mathrm{d}^2\mu_z/\mathrm{d}Q_2^2$	$\mathrm{d}^3\mu_z/\mathrm{d}Q_2^3$
	(99, 590)	7.3	12.9	0.2	12.9
ωB97X	(250, 974)	0.7	1.9	0.0	1.0
	(750, 974)	0.7	1.9	0.0	1.0
	(99, 590)	17.7	15.0	0.7	38.3
M06-2X	(250, 974)	3.1	7.5	0.0	1.2
	(750, 974)	3.1	7.6	0.0	1.2

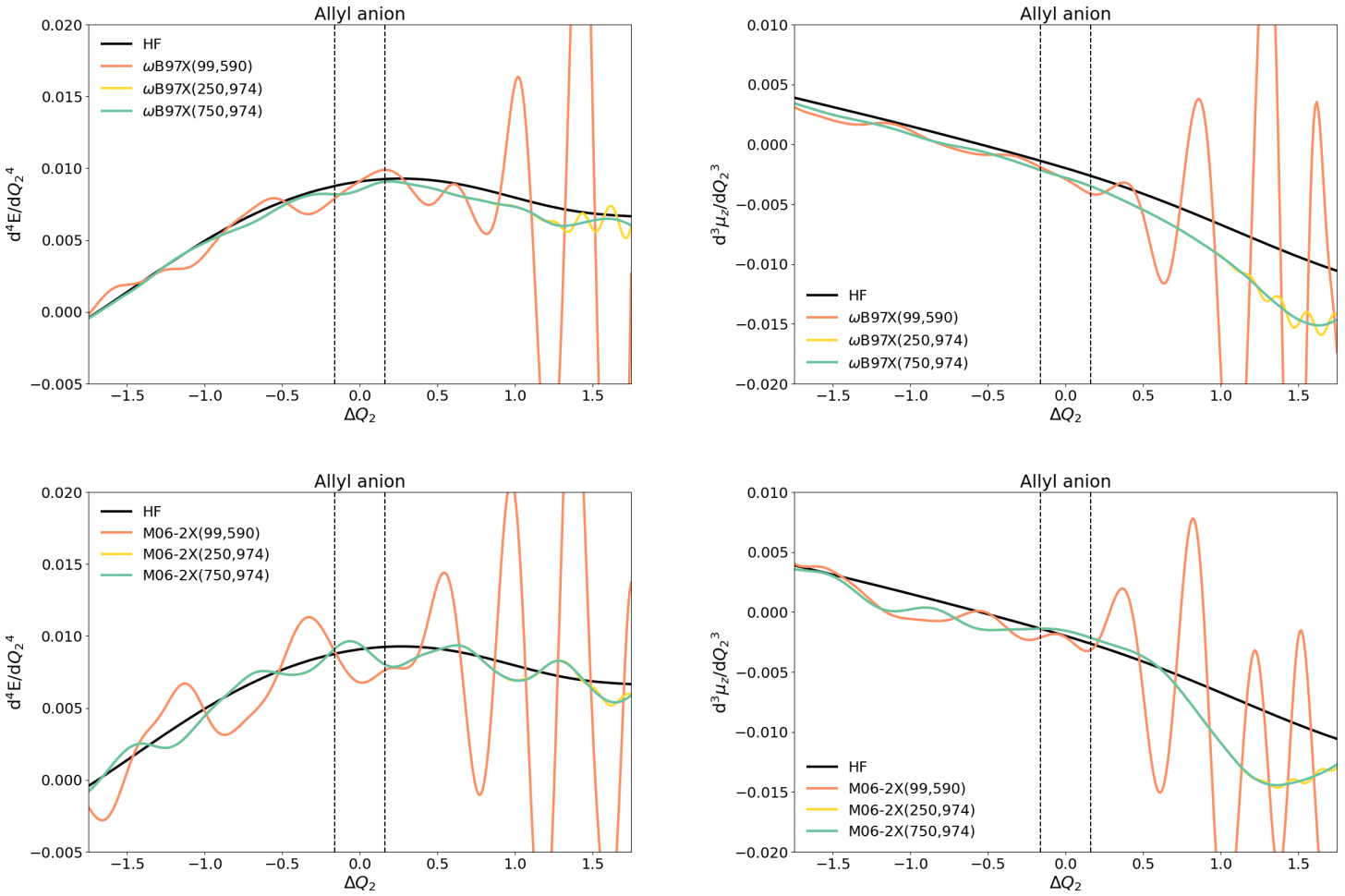

Figure 9.16: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_2^4$ (left column) and $\mathrm{d}^3\mu_z/\mathrm{d}Q_2^3$ (right column) for the displacements along the Q_2 normal coordinate of allyl anion. Data obtained with ωB97X (top row) and M06-2X (bottom row).

9.6.2 Phenanthrene (Q_4 mode)

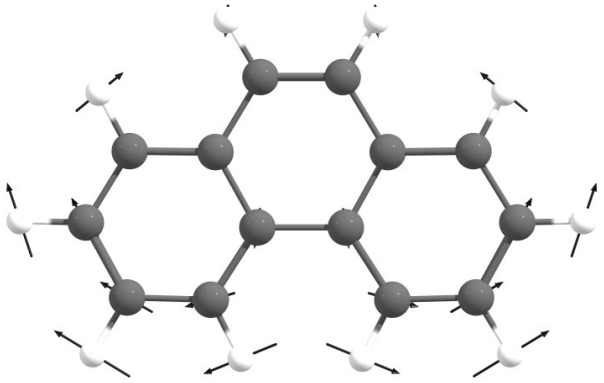

Figure 9.17: Graphical representation of the Q_4 (A_1) normal mode of phenanthrene. Obtained at M06-2X/aug-cc-pVDZ level of theory.

The largest molecule we investigated was phenanthrene, an example of a polyaromatic hydrocarbon. We investigated the Q_4 normal mode, which shares similarities with the Q_2 mode of allyl anion (see Figure 9.17). However, in this mode the key motion is the one of the two hydrogen atoms belonging to the C1 and C4 carbons and now are slightly farther apart than the ones bounded to C1 and C3 atoms of allyl anion. The vibrational frequency is rather small (around 255 cm$^-$1) but still very harmonic.

Table 9.12: Vibrational properties of the Q_4 normal mode of phenanthrene obtained with HF, ωB97X, and M06-2X: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, IR_{har} and IR_{anh} — harmonic and anharmonic IR intensities, RA_{har} and RA_{anh} — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

	Q_4 (A_1) mode					
Method	$\omega_{\text{har}}\left[\frac{1}{\text{cm}}\right]$	$\omega_{\text{anh}}\left[\frac{1}{\text{cm}}\right]$	$\text{IR}_{\text{har}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{IR}_{\text{anh}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{RA}_{\text{har}}\left[\text{Å}^6\right]$	$\text{RA}_{\text{anh}}\left[\text{Å}^6\right]$
HF	263	261	0.495	0.486	0.177	0.165
ωB97X	253		0.4278			
M06-2X	252		0.4133			

Surprisingly, the effect of spurious oscillations is dramatic. For the high-order energy derivatives, $\text{d}^4E/\text{d}Q_4^4$ and $\text{d}^3\mu_z/\text{d}Q_4^3$, the oscillations have huge amplitudes and very high frequencies (see Figure 9.18). When the (99, 590) grid is utilized, RRMSE exceeds 146% and 3030% for $\text{d}^3E/\text{d}Q_4^3$ and $\text{d}^4E/\text{d}Q_4^4$ using the ωB97X functional. For M06-2X, the corresponding errors are much lower but still unacceptably large; namely RRMSE equals to 15% and 116% for $\text{d}^3E/\text{d}Q_4^3$ and $\text{d}^4E/\text{d}Q_4^4$, respectively. Dramatic errors of the same magnitudes are also observed for the derivatives of the dipole moment, for both ωB97X and M06-2X. Moreover, the errors are only partially alleviated with the huge (750, 974) grid, which for a molecule of such size would already be computationally very costly. Therefore, if one performs the full VPT2 computations with these functionals and integration grids, one should expect huge errors in the resulting the anharmonic vibrational frequencies of this mode (as well as other modes which were vibrationally coupled).

Table 9.13: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_4 normal mode of phenanthrene, obtained for some DFAs and integration grids.

Method	Grid	d^3E/dQ_4^3	d^4E/dQ_4^4	$d^2\mu_z/dQ_4^2$	$d^3\mu_z/dQ_4^3$
	(99, 590)	146.8	3030.2	116.6	1136.2
ωB97X	(250, 974)	100.6	1716.4	69.5	1086.2
	(750, 974)	7.4	41.2	11.0	84.9
	(99, 590)	15.0	116.4	37.3	516.8
M06-6X	(250, 974)	10.2	188.7	8.8	218.2
	(750, 974)	3.6	22.6	1.9	27.3

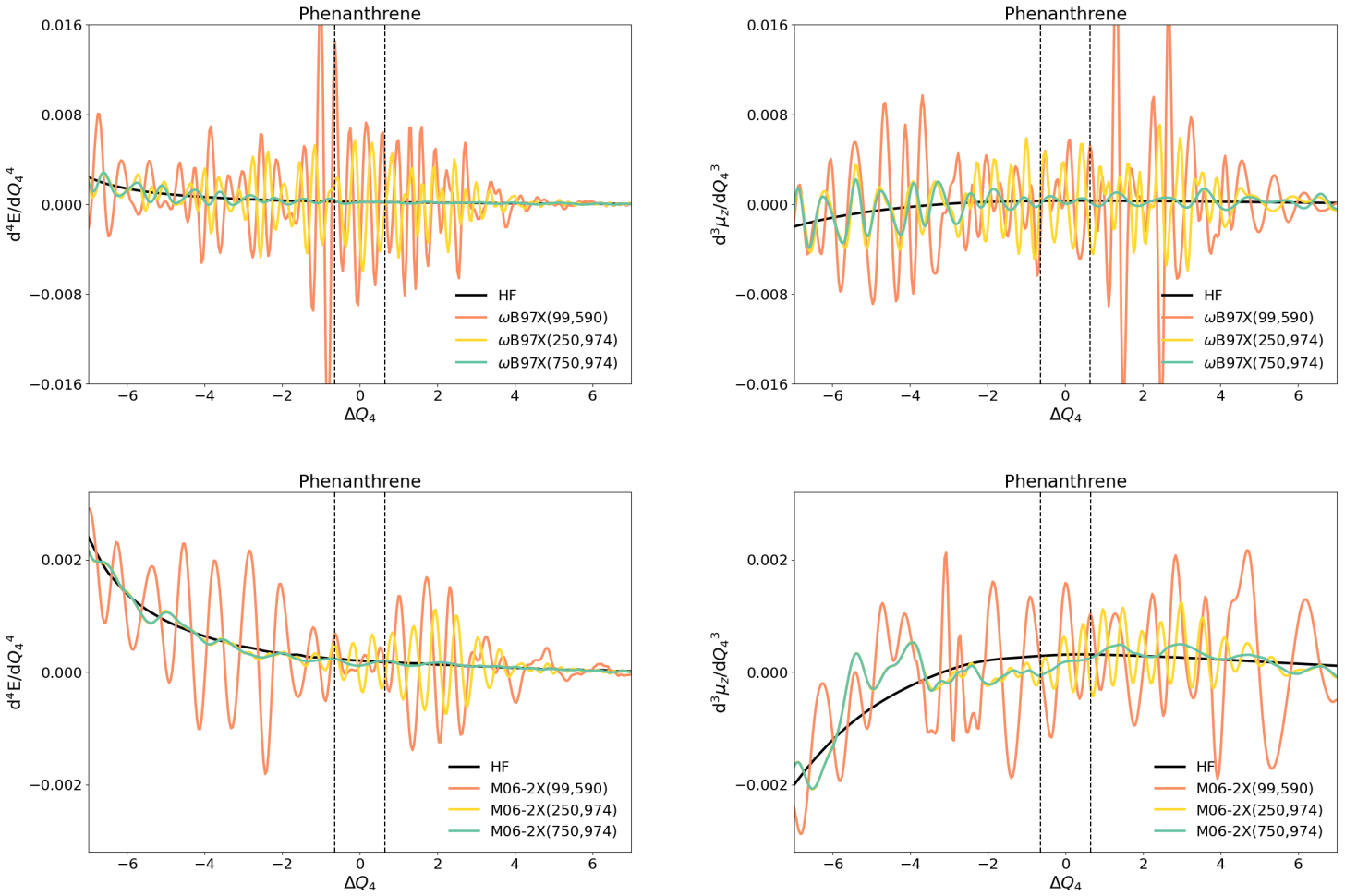

Figure 9.18: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_4^4$ (left column) and $\mathrm{d}^3\mu_z/\mathrm{d}Q_4^3$ (right column) for the displacements along the Q_4 normal coordinate of phenanthrene. Data obtained with ωB97X (top row) and M06-2X (bottom row).

9.6.3 H_2S_2 (Q_1 mode)

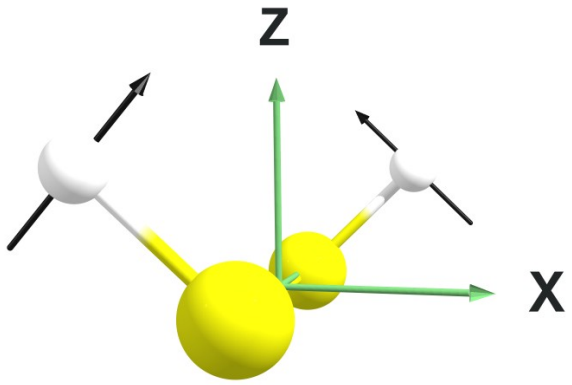

Figure 9.19: Graphical representation of the Q_1 (A) normal mode of H_2S_2. Obtained at M06-2X/aug-cc-pVDZ level of theory.

In this last section, we present the spurious oscillations affecting the description of a different type of nuclear motion, that is, a displacement mimicking changes in dihedral angle, and in much simpler chemical system, namely H_2S_2. We studied the displacements along the Q_1 normal mode of H_2S_2 (see Figure 9.19). To be precise, such of nuclear displacements along Q_1 are of linear type, i.e., atoms are proportionally displaced along the normalized Q_1 vector. In contrast, if one decided to study the displacement along the dihedral angle, these would be described by a curvilinear mode. We decided to choose the former option to stay consistent with the normal coordinate system. The Q_1 normal mode in H_2S_2 is characterized with low vibrational frequency and slightly larger anharmonic contributions than the previously studied modes (see Tables 9.14). At HF level, the anharmonic correction reduce the harmonic frequency from 447 cm^{-1} to 421 cm^{-1}.

Table 9.14: Vibrational properties of the Q_1 normal mode of H_2S_2 obtained with some methods: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, IR_{har} and IR_{anh} — harmonic and anharmonic IR intensities, RA_{har} and RA_{anh} — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

	Q_1 (A) mode					
Method	$\omega_{har}\left[\frac{1}{cm}\right]$	$\omega_{anh}\left[\frac{1}{cm}\right]$	$IR_{har}\left[\frac{km}{mol}\right]$	$IR_{anh}\left[\frac{km}{mol}\right]$	$RA_{har}\left[Å^6\right]$	$RA_{anh}\left[Å^6\right]$
HF	447	421	15.918	15.937	0.0427	0.0583
B3LYP	428		12.270		0.0370	
CAM-B3LYP	432		13.689		0.0402	
BH&H	450		15.281		0.0392	
LC-BLYP	438		15.865		0.0480	
ωB97X	433		14.684		0.0346	
ωB97XD	429		13.927		0.0310	
PBE	435		11.857		0.0346	
PBE0	445		13.423		0.0378	
TPSSh	449		12.539		0.0385	
M06-2X	415		13.132		0.0398	
N12	433		13.520		0.0305	
MN15	452		14.369		0.0440	
MN15-L	468		13.535		0.0403	

For these molecular systems, we included 13 different DFAs, constituting a few representatives of each rung from Section 9.5, and tested the grids with larger angular grid

as well. Namely, besides including the "default" (99, 590) grid, we tested the importance of the radial part of the integration by comparison with the results for the (250, 974) and (750, 974) grids, and the importance of the angular part by comparison of the (750, 974), (750, 3074), and (750, 5294) grids. The 5294-point grid is the largest Lebedev grid and we consider it as the best possible angular grid to be used in such computations. The (750, 5294) grid is 69 times larger than (99, 590), 16 times larger than (250, 974), and 5 times larger than the (750, 974) grid. This implies a great increase in the cost of the calculation, which disqualifies the choice of this grid in any real life computations, especially VPT2 calculations which involves the numerical computation of high-order derivatives and solving dozens of CPKS equations (for the Hessian and dipole and polarizability derivatives). All electronic structure calculations employed the aug-cc-pVTZ basis set and were done using the QCHEM 5.1 computational package.

Results for H_2S_2 are compiled in Tables 9.15 and 9.16 and the oscillations are shown in Figures 9.20 and 9.21. One finds substantial errors for the third derivatives of energy, $\mathrm{d}^3E/\mathrm{d}Q_1^3$, and in the case of M06-2X, already for the second derivative, $\mathrm{d}^2E/\mathrm{d}Q_1^2$. In the latter case, RRMSE for $\mathrm{d}^2E/\mathrm{d}Q_1^2$ are above 20% for the (99, 590), (250, 974) and (750, 974) grids, implying a 20% of error in the vibrational frequencies of Q_1 modes obtained at the harmonic level. Secondly, the 974-point Lebedev grid is not large enough to permit the safe calculations of the high-order nuclear derivatives, which in practical terms means that one might not reach a correct answer with a program such as Gaussian. Thirdly, the rather grid-robust DFAs from rung 3, such as PBE, PBE0, and MN15 (see Table 9.15) also suffer from considerable grid-errors. For the latter one, the RRMSE for $\mathrm{d}^3E/\mathrm{d}Q_1^3$ and $\mathrm{d}^4E/\mathrm{d}Q_1^4$ exceed 330% for the (99, 590) grid, and 55% for the (250, 974) and (750, 974) grids. Surprisingly, for PBE and PBE0, the RRMSE values are much larger for the (250, 974) and (750, 974) grids than for (99, 590). This is due to a fortunate lowering of amplitude of oscillations for very small displacements $\Delta Q_1 \approx 0$ in the case of the (99, 590) grid (see Figure 9.20). Lastly, for the most grid-demanding DFAs from rung 5, namely ωB97X, ωB97XD, N12 and M06-2X, errors can exceed dozens and hundreds of percent. In this test, the largest RRMSE values are observed for M06-2X combined with the (99, 590) grid. Moreover, for those DFAs, the spurious oscillations of substantial amplitudes persist even if the huge (750, 3074) grid is used.

Table 9.15: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_1 normal mode of H_2S_2, obtained for some DFAs and integration grids.

Method	Grid	d^2E/dQ_1^2	d^3E/dQ_1^3	d^4E/dQ_1^4
B3LYP	(99, 590)	0.6	5.8	6.7
	(250, 974)	0.4	2.4	3.5
	(750, 974)	0.4	2.4	3.5
	(750, 3074)	0.0	0.0	0.1
	(750, 5294)	0.0	0.0	0.1
CAM-B3LYP	(99, 590)	0.2	1.0	1.4
	(250, 974)	0.3	0.4	1.3
	(750, 974)	0.3	0.4	1.3
	(750, 3074)	0.0	0.0	0.1
	(750, 5294)	0.0	0.0	0.1
BH&H	(99, 590)	0.4	2.3	4.9
	(250, 974)	0.2	0.8	1.2
	(750, 974)	0.2	0.8	1.3
	(750, 3074)	0.0	0.0	0.1
	(750, 5294)	0.0	0.0	0.1
LC-BLYP	(99, 590)	0.4	3.6	6.3
	(250, 974)	0.3	1.3	1.7
	(750, 974)	0.3	1.2	1.5
	(750, 3074)	0.0	0.0	0.0
	(750, 5294)	0.0	0.0	0.1
PBE	(99, 590)	0.3	6.1	3.7
	(250, 974)	0.6	35.1	35.4
	(750, 974)	0.6	35.2	35.4
	(750, 3074)	0.0	0.8	0.8
	(750, 5294)	0.0	0.0	0.1
PBE0	(99, 590)	0.3	3.9	3.8
	(250, 974)	0.5	23.3	28.7
	(750, 974)	0.5	23.3	28.8
	(750, 3074)	0.0	0.3	0.7
	(750, 5294)	0.0	0.0	0.1
TPSSh	(99, 590)	2.8	99.4	97.9
	(250, 974)	0.6	77.1	76.8
	(750, 974)	0.6	77.3	76.8
	(750, 3074)	0.1	5.6	5.2
	(750, 5294)	0.0	0.1	0.3

Table 9.16: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_1 normal mode of H_2S_2, obtained for some DFAs and integration grids.

Method	Grid	d^2E/dQ_1^2	d^3E/dQ_1^3	d^4E/dQ_1^4
	(99, 590)	2.4	40.9	45.3
	(250, 974)	0.4	48.2	56.2
ωB97X	(750, 974)	0.4	48.3	56.3
	(750, 3074)	0.1	14.9	20.4
	(750, 5294)	0.0	0.2	0.3
	(99, 590)	2.2	64.2	61.9
	(250, 974)	0.9	33.0	38.4
ωB97XD	(750, 974)	0.9	33.1	38.7
	(750, 3074)	0.2	21.8	28.6
	(750, 5294)	0.0	1.4	2.1
	(99, 590)	21.9	711.9	817.2
	(250, 974)	20.1	497.3	676.4
M06-2X	(750, 974)	20.1	499.1	678.4
	(750, 3074)	0.8	49.5	59.1
	(750, 5294)	0.0	2.1	3.6
	(99, 590)	1.3	199.3	203.9
	(250, 974)	1.0	129.3	140.2
N12	(750, 974)	1.0	129.1	140.1
	(750, 3074)	0.2	31.3	44.9
	(750, 5294)	0.0	2.2	3.6
	(99, 590)	12.0	334.1	353.5
	(250, 974)	2.2	55.4	58.6
MN15	(750, 974)	2.2	54.9	58.3
	(750, 3074)	0.1	1.9	2.2
	(750, 5294)	0.0	0.1	0.1
	(99, 590)	2.0	39.4	38.6
	(250, 974)	0.5	37.8	41.8
MN15-L	(750, 974)	0.5	37.7	41.9
	(750, 3074)	0.3	11.1	10.5
	(750, 5294)	0.0	0.9	1.3

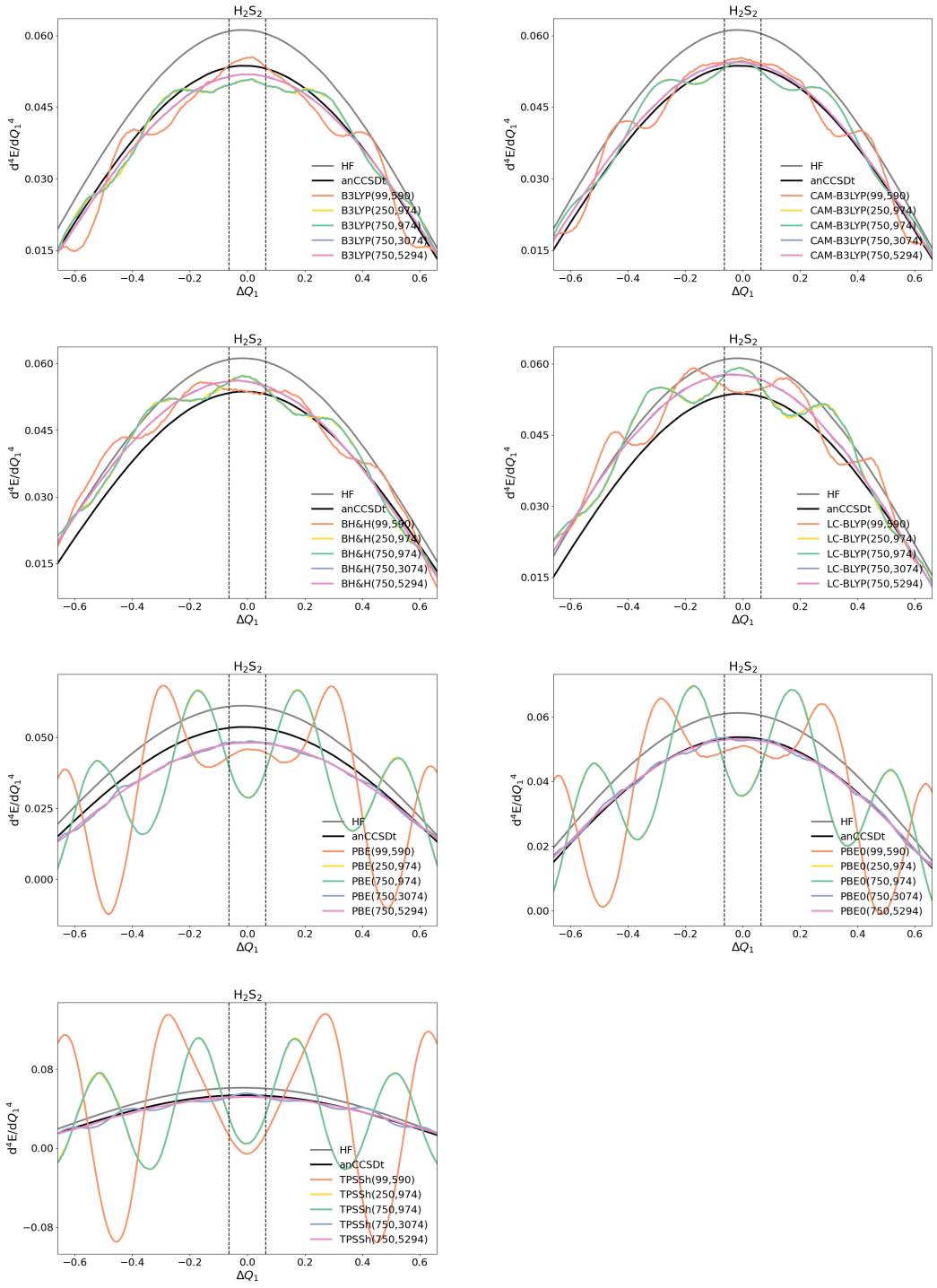

Figure 9.20: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_1^4$ for the displacements along the Q_1 normal coordinate of H_2S_2. Left column shows the results for B3LYP, BH&H, PBE, and TPSSh, and right column for CAM-B3LYP, LC-BLYP, and PBE0.

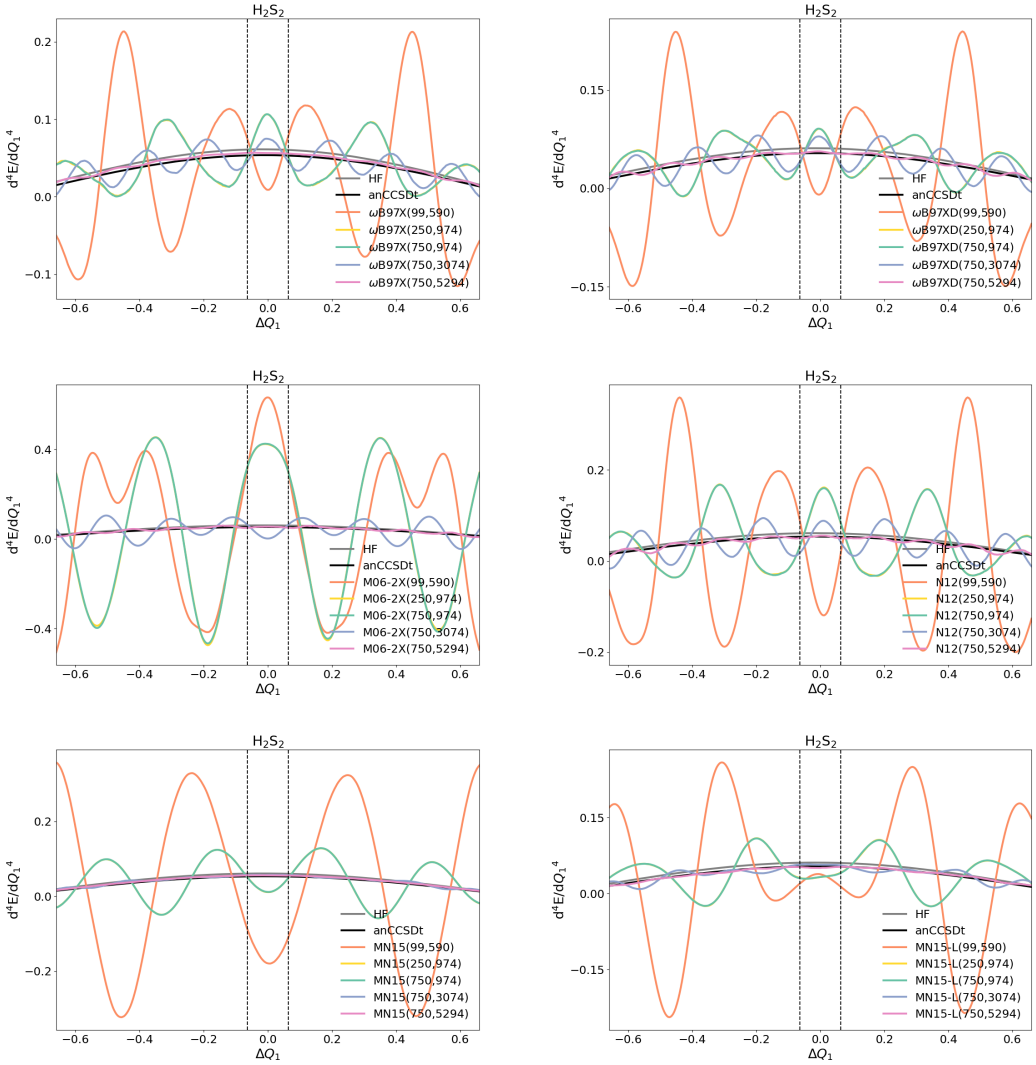

Figure 9.21: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_1^4$ for the displacements along the Q_1 normal coordinate of H_2S_2. Left column shows the results for ωB97X, M06-2X, MN15, and right column for ωB97XD, N12, and MN15-L.

9.7 Conclusions

In summary, we have uncovered an important limitation of most modern density functional approximations, which might suffer from spurious oscillations of molecular properties along the nuclear displacement coordinate. The oscillations are due to numerical integration errors, which sometimes can be alleviated using large integration grids that compromise the favorable cost-accuracy ratio of DFAs. However, in some cases, the spurious oscillations cannot be avoided even using enormous grids such as (750, 974).

With the newly developed algorithm, based on the Fourier spectral analysis and digital signal processing techniques, one can easily identify and quantify these oscillations. With its use we performed a wide benchmarking of the grid-stability of DFAs and classified forty-five popular DFAs into five groups. Among the widely employed functionals, only LC-BLYP shows robustness against spurious oscillations of various molecular properties examined along a nuclear displacement coordinate. Hardly a handful of DFAs are a safe choice to compute the high-order derivatives, such as $\mathrm{d}^3E/\mathrm{d}\xi^3$, $\mathrm{d}^4E/\mathrm{d}\xi^4$, $\mathrm{d}^3\mu_z/\mathrm{d}\xi^3$, and $\mathrm{d}^3\alpha_{zz}/\mathrm{d}\xi^3$, which are key quantities to define the basic anharmonic corrections to the vibrational properties. This criterion is especially harder to fulfill for the loosely noncovalently bonded complexes, especially halogen– and dispersion–bonded complexes.

While the spurious oscillations are the most prominent and evident for the noncovalently-bonded complexes, they are not limited to this kind of chemical systems. As shown in the last section of this Chapter, oscillations could be identified in different sorts of single-molecules normal modes. In particular, we show that molecules like H_2O_2, H_2S_2, the allyl anion, butadiene, cyclobutadiene, benzene, naphthalene, and phenanthrene also exhibit spurious oscillations for various energy and dipole derivatives over certain normal modes. All the evidence indicates that the oscillations are a universal feature of modern DFAs that can affect molecules with low-frequency vibrational modes. Given the ubiquitous nature of the latter, the grid-dependency problem of modern DFAs constitutes a challenge to the development of new approximations that calls for the inclusion of molecular properties and the study of spurious oscillations in the construction of new DFAs.

Chapter 10

The Origin of Spurious Oscillations in Density Functional Approximations

10.1 Introduction

In the previous Chapter, it was demonstrated that the grid dependence of DFAs and the resulting spurious oscillations are much broader problems than was originally anticipated. Although they are not initially observable in the PECs, grid-related spurious oscillations emerge in the energy derivatives with respect to the nuclear displacement, which are key quantities to simulate the vibrational IR and Raman spectra. We showed that DFAs from all rungs of Jacob's ladder suffer from spurious oscillations and the corresponding errors may reach hundreds of percent for high-order derivatives. DFAs from the BLYP and PBE families were the least affected by the spurious oscillations, whereas the functionals from the B97, M06, and SCAN families were the most affected by them. Due to the large errors in the high-order derivatives of E, μ, and α with respect to the normal coordinates, the latter DFAs turned out to be unsuitable for the computation of VPT2 anharmonic corrections. As a natural follow-up question, one could further ask which parts of these DFAs are responsible of the spurious oscillations.

The newly developed algorithm has proven to be an indispensable tool for studying the phenomenon of spurious oscillations, giving a measure of the grid-related errors even for the high-order derivatives of molecular properties. However, in its current version, the procedure needs some computing effort to compute property curves. For example, to evaluate energy derivatives errors, one has to evaluate hundreds of energy calculations for different molecular geometries. Accomplishing that for all normal modes would require a substantial computational cost. In this Chapter, we will explore some of the aspects of electronic structure methods that could lead to a reduction of the computational complexity of the procedure, without affecting the errors measures.

We start the discussion with a rather technical aspect: the choice of the size of the basis set. Next, we test, if one can skip the SCF procedure and utilize the arbitrary input densities for the computation of electronic energies and their derivatives with respect to the nuclear displacements. We also discuss direct and indirect effects of the spurious oscillations in energy components. Last but not least, we unveil the main source of spurious oscillations in the most popular DFAs. We show which exchange and correlation functionals are the most grid robust, and which are the ones that are most grid demanding.

10.2 Systems and Computational Details

In this Chapter, we employ the computational procedure established in Chapter 9. Namely, the detection of grid-related spurious oscillations and the quantification of the corresponding errors are done with the algorithm presented in Section 9.3.1. We test forty-five DFAs, however, the set slightly differs from the one in Chapter 9 (see Table 10.1). Now, we include LSDA functionals (SVWN5 and SPW92), the Perdew-Wang functionals (PW91, mPW91, and B1PW91), the VSXC meta-GGA functional, but do not consider double hybrids (mPW2PLYP, B2PLYP, PBE-0DH, and PBE-QIDH) and the HSE functionals (HSE03 and HSE06). We investigate the same type of molecular displacements as in the previous Chapter for HCN$\cdot$HF, HCN$\cdot$HCl, $N_2\cdot$HF, OC$\cdot$HF, HCN$\cdot$BrF, Ar_2. We optimize the geometries of polyatomic systems at the corresponding level of theory, using the aug-cc-pVTZ basis set and the (250, 974) grid. Then, for each optimized geometry, a set of displacements along the $\chi_{1,z}$ coordinate is generated. For SVWN5 and SPW92, the optimization procedure is skipped and the B3LYP geometries are adopted instead. In this work, we mostly focus on the errors in the derivatives of various energy components with respect to nuclear displacements ξ. Additionally, in Section 10.3, the errors in the derivatives of dipole moment and static polarizability are also reported. Derivatives are

obtained using the numerical technique described in Section 9.2.4. In this Chapter, we quantify the errors coming from spurious oscillations using RMSE (see Eq. 9.2).

Through this Chapter, we test the (99, 590), (250, 974), (750, 974), and (1500, 974) integration grids, and we adopt the last one as the large grid (LG) grid that is required to obtain the filtered property profiles (see Section 9.3.1). For the integration of the nonlocal VV10 correlation in ωB97M-V, ωB97X-V, and B97M-V, a pruned SG3 grid is used [316]. In this Chapter, except for the optimization procedures and tests from Section 10.3, we utilize the 6-31+G* basis set in all electronic structure calculations. In the KS-DFT SCF calculations, the DIIS convergence criterion is set to 10^{-11} and almost all the two-electron integrals are included (cutoff set to 10^{-30}). Static polarizabilities, reported in Section 10.3, are obtained analytically using the CPKS equations. All HF and KS-DFT calculations are performed with QCHEM 5.1 [322] and all graphs are prepared with the Matplotlib package [351].

Table 10.1: List of DFAs used in this Chapter; GGA — Generalized Gradient Approximation, GH — global hybrids, RSH — range-separated hybrids.

DFA	Type	Ref.	DFA	Type	Ref.
BLYP	GGA	[178, 180]	PBE	GGA	[179]
B1LYP	GH GGA	[324]	PBE0	GH GGA	[294]
B3LYP	GH GGA	[325, 326]	PBE50	GH GGA	[322]
BH&H	GH GGA[a]	[296]	LC-ωPBE	RSH GGA	[191]
BH&HLYP	GH GGA	[296]	VSXC	meta-GGA	[362]
LC-BLYP	RSH GGA	[190]	B97	GH GGA	[327]
CAM-B3LYP	RSH GGA	[193]	B97-D	GGA	[328]
M06	GH meta-GGA	[293]	ωB97	RSH GGA	[192]
M06-L	meta-GGA	[184]	ωB97X	RSH GGA	[192]
M06-HF	GH meta-GGA	[330]	ωB97X-D	RSH GGA	[329]
M06-2X	GH meta-GGA	[293]	ωB97X-D3	RSH GGA	[331]
SOGGA11	GGA	[332]	B97M-V	meta-GGA	[319]
SOGGA11-X	GH GGA	[333]	ωB97M-V	RSH meta-GGA	[77]
M11	RSH meta-GGA	[363]	ωB97X-V	RSH GGA	[78]
M11-L	meta-GGA	[335]	TPSS	meta-GGA	[182]
MN12-L	meta-GGA	[337]	RevTPSS	meta-GGA	[336]
N12-SX	RSH GGA	[339]	TPSSh	GH meta-GGA	[338]
N12	GGA	[334]	SPW92	LDA	[364, 365]
MN12-SX	RSH meta-GGA	[339]	SVWN5	LDA	[364, 366]
MN15	GH meta-GGA	[75]	PW91	GGA	[367]
MN15-L	meta-GGA	[76]	mPW91	GGA	[368]
SCAN	meta-GGA	[183]	B1PW91	GH GGA	[324]
SCAN0	GH meta-GGA	[342]			

[a] No GGA exchange is used, *i.e.*, $E_{xc}^{BHH} = 0.5\, E_x^{Slater} + 0.5\, E_x^{HF} + E_c^{LYP}$

10.3 The Effect of the Basis Set

We start the discussion by analyzing the dependency of the errors coming from spurious oscillations on the atom-centered basis set. We tested 26 basis sets of different sizes and families: STO-3G, Pople's basis sets with different levels of contraction, 3-21G, 6-31G and 6-311G (with and without diffuse and polarization functions), Karlsruhe basis set (ma-)Def2SVP and (ma-)Def2TZVP, and Dunning's basis sets (aug-)cc-pVDZ and (aug)-cc-pVTZ. In the case of Pople's basis sets, the importance of the diffuse and polarization functions in hydrogen atoms was also tested (by analyzing the differences between 6-31G, 6-31G*, 6-31G**, as well as between 6-31G, 6-31+G, 6-31++G). For three representative DFAs, ωB97X, M06-2X, and B3LYP, we calculated the RMSE of the derivatives of the energy, dipole moment, and polarizability of the HCN$\cdot$HF system. These molecular properties, when listed in that order, show an increasing dependence on the quality of the basis set, especially on the inclusion of diffuse functions.

Results of this test are collected in Figures 10.1 and A2.3, and the raw data is given in Tables A3.1 to A3.6 (see Appendix). The errors coming from grid-related spurious oscillations in the $\mathrm{d}^m E/\mathrm{d}\xi^m$, $\mathrm{d}^m \mu_z/\mathrm{d}\xi^m$, and $\mathrm{d}^m \alpha_{zz}/\mathrm{d}\xi^m$ derivatives depend slightly on the basis set. This observation goes in line with the fact that most exchange-correlation functionals have a semi-local character and do not depend strongly on the size of the basis set. Except for a few cases, for all properties and DFAs, the magnitude of the error obtained with aug-cc-pVTZ can be predicted using a much smaller basis set. With some exceptions, small basis sets, such as Def2SVP or 6-31G, can be used to quantify the errors coming from spurious oscillations in a qualitative manner. Remarkably, one can predict the magnitude of the grid errors even for high-order derivatives such as $\mathrm{d}^4 E/\mathrm{d}\xi^4$ using a very small basis set. Hence, the actual cost of our procedure to identify grid-related oscillations is dramatically lowered using a small basis set.

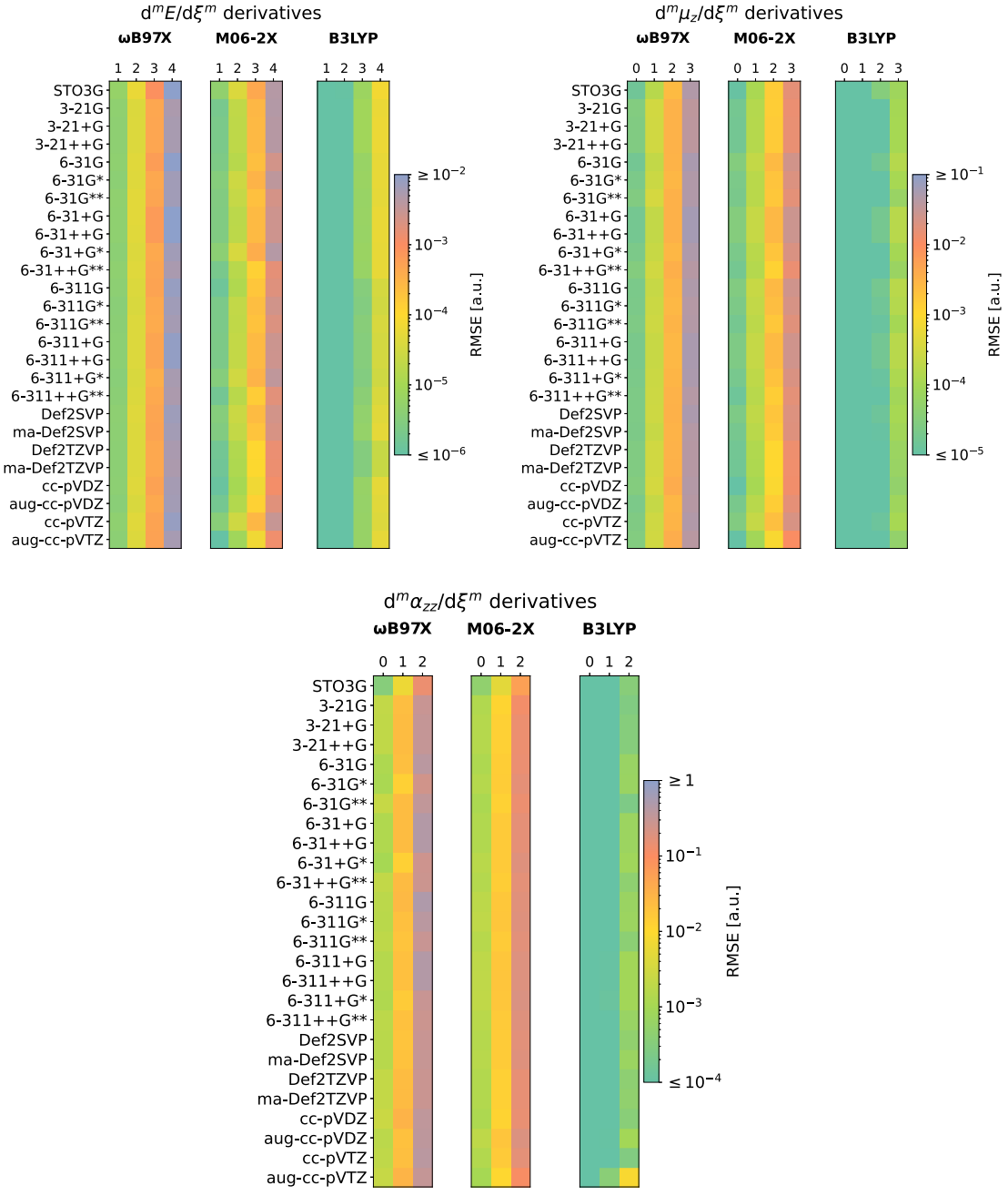

Figure 10.1: Study of the basis set dependence of the grid-related errors in the d$^m E/$d$\chi_{1,z}{}^m$ derivatives (*left panel*, m= 1 - 4, stored column-wise), the d$^m \mu_z/$dξ^m derivatives (*middle panel*, m = 0 - 3, stored column-wise), and the d$^m \alpha_{zz}/$dξ^m derivatives (*right panel*, m = 0 - 2, stored column-wise) of HCN · HF using ωB97X, M06-2X, and B3LYP combined with the (99, 590) integration grid. Colors reflect the values of the errors in the property measured using RMSE (note the logarithmic scale). Raw data compiled in Tables A3.1 to A3.6.

10.4 The Usage of Approximate Densities

Since spurious oscillation errors can be measured using the small basis sets (hence, densities of lower quality), we will investigate the importance of the quality of the density to estimate errors coming from spurious oscillations. To this end, we analyze the grid-related oscillations in the curves of $\mathrm{d}^m E/\mathrm{d}\xi^m$ computed using different types of input densities:

- The SCF density, which was the density used thus far;
- The trial densities from the superposition of atomic densities (SAD), and its purified version (SADMO), both of which are popular guess densities used to initialize the SCF procedure [322, 369, 370];
- The trial density constructed from *frozen* molecular orbitals (which we will refer as FMO), i.e., the SCF procedure is only performed for the equilibrium geometry and then the same set of MO coefficients are used for all points of the PES.

The energy computation coming from a DFA and a trial density X will be indicated as DFA@X. It is important to stress that, except the SCF density, all of the other densities do not correspond to a stationary wavefunction. Moreover, the SAD and FMO densities are not idempotent and should never be used to calculate molecular property. Conversely, the SADMO density is obtained from an initial diagonalization of the SAD density matrix, followed by the aufbau occupation of the corresponding NOs that are employed to construct the density.

There are two main advantages in favor of using trial densities to check for spurious oscillations. First, it dramatically reduces the cost of the procedure. Instead of performing hundreds of SCF calculations for the displaced geometries to construct the whole PEC (and its derivatives), one only has to perform one SCF calculation (of the optimized geometry) and the rest of points of the PEC can be calculated at almost no cost (since they do not require SCF or to construct a Fock matrix, or calculate the 1-and 2-electron integrals). Formally, employing the FMO densities (and orbitals) to search for potential oscillations of N nuclear displacements by studying M points of the PEC provides the $N \times M$-fold reduction of computer time. For example, the search for oscillations in only one coordinate with 400 points in PEC requires 400 times less of computer time. If all $3N-6$ normal coordinates are studied, $(3N-6)400$ times less CPU time is required. Using FMO densities we can study only the *direct* grid errors in $\mathrm{d}^m E_x/\mathrm{d}\xi^m$ and $\mathrm{d}^m E_c/\mathrm{d}\xi^m$, as it will be explained in Section 10.5.

For this analysis, we tested 20 DFAs and computed the $\mathrm{d}^m E/\mathrm{d}\xi^m$ curves for two noncovalently bonded systems: HCN$\cdot$HF and Ar_2. Figures 10.2 and 10.3 include only the spurious oscillations obtained using the SCF, SAD, SADMO, and FMO densities (see also Figures A2.4 to A2.10 in the Appendix). These spurious oscillations are calculated from the difference between the results obtained with two different grids: the (99, 590) grid and the reference (1500, 974) grid, $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$. For these cases, no filtering procedure was utilized as the (1500, 974) grid yields oscillation-free $\mathrm{d}^3 E/\mathrm{d}\xi^3$ and $\mathrm{d}^4 E/\mathrm{d}\xi^4$ curves for HCN$\cdot$HF. Figures 10.4 to 10.6 present correlation plots of the RMSE values obtained using the SCF and trial densities.

As expected, the FMO, SAD, and SADMO trial densities generate property curves of $\mathrm{d}^3 E/\mathrm{d}\xi^3$ and $\mathrm{d}^4 E/\mathrm{d}\xi^4$ that are incorrect and unphysical. However, the description of the grid-related spurious oscillations using the trial densities is remarkably good. In particular, the spurious oscillations perfectly match the SCF ones when the FMO trial densities are used. For all the examples, ΔP curves obtained from FMOs and SCF densities match very well even in the case of highly displaced geometries. The latter is true regardless

of the DFA type. On the contrary, SAD and SADMO provide limited predictions of the shape of spurious oscillations. With these input densities, one observes a good characterization of spurious oscillations for SVWN5, B3LYP, MN15 functionals, whereas for the other tested DFAs the oscillations are rather poorly described.

These conclusions are further confirmed by the linear regression analysis, which reveals that all trial densities can predict the existence of spurious oscillations, FMO densities being clearly the best ones for this purpose. The Pearson correlation between the errors predicted by the FMO and SCF densities is very close to one for all $\mathrm{d}^m E/\mathrm{d}\xi^m$ derivatives. In the case of SAD and SADMO, the correlation with the SCF errors is slightly lower (R^2 larger than 0.95). The largest grids, (750, 974) and (1500, 974), usually do not produce large spurious oscillations and, therefore, it is difficult to reproduce their small errors using approximate densities (see Figures 10.4 to 10.6). A similar effect occurs when the RMSE errors are very small (see the bottom left parts of the correlation plots). The latter occurs mostly for the 1$^{\mathrm{st}}$ and 2$^{\mathrm{nd}}$ derivatives, whereas for the 3$^{\mathrm{rd}}$ and 4$^{\mathrm{th}}$ derivatives situation is a bit better (because the spurious oscillations have much larger amplitudes). Nevertheless, even when the RMSE is very small the trial densities provide the same qualitative answer than the SCF densities. Therefore, the trial densities may be used to detect and quantify the spurious oscillations of DFA, for a fraction of the cost of a single SCF cycle.

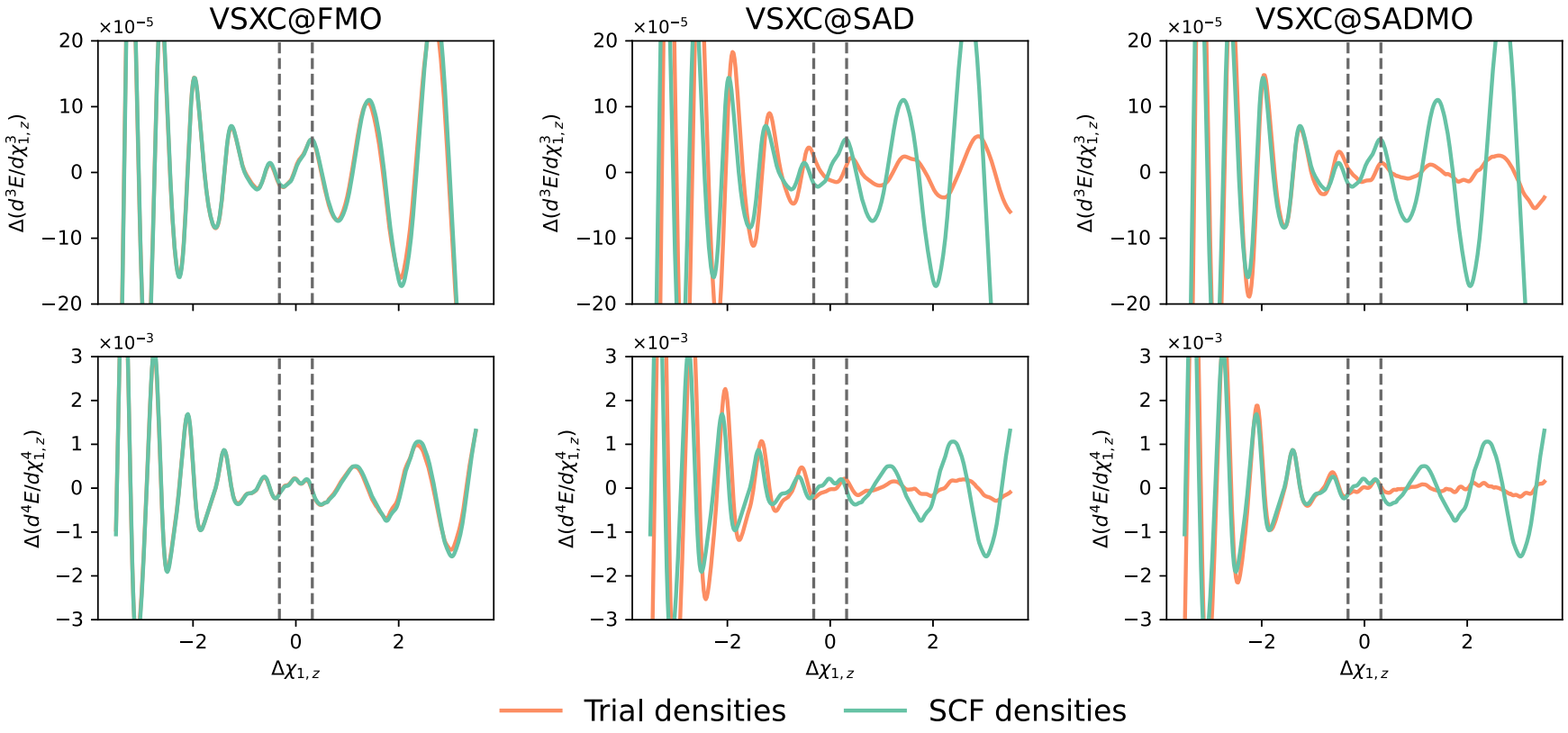

Figure 10.2: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with VSXC/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

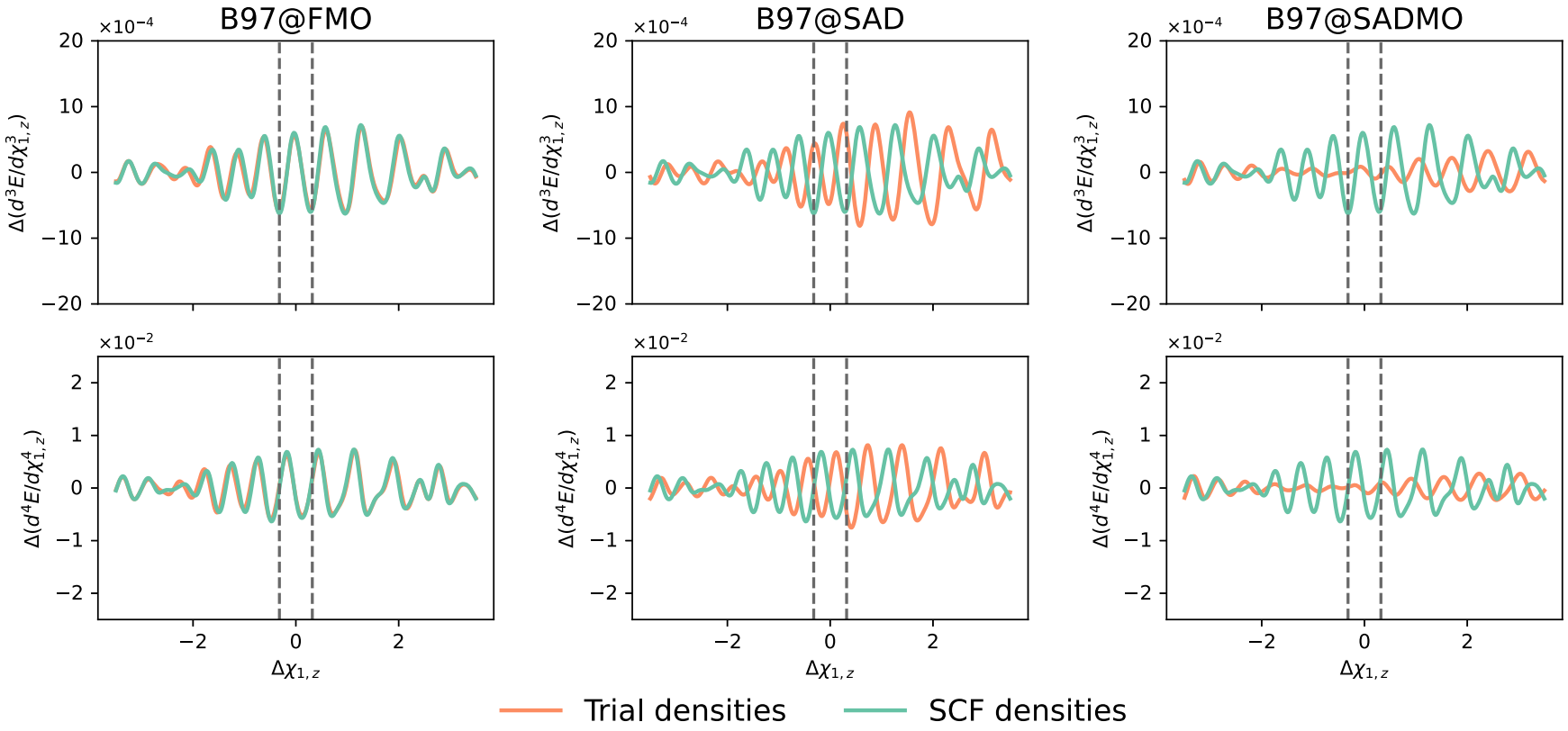

Figure 10.3: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with B97/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

10.5 Direct and Indirect Errors in Energy Components Derivatives

The quality of the integration grid directly affects the computation of E_{xc}, since it is the only part of the electronic energy that is evaluated using numerical quadratures. We will call this type of error as *direct* error. However, since the KS-DFT equations are solved until self consistency, one can expect the numerical integration errors to also affect the Kohn-Sham MOs. Then, those MOs are used to evaluate all energy components (see Eq. 5.29): the Kohn-Sham kinetic energy T_s, the nuclei-electron potential energy V_{en}, the Coulomb electron repulsion J, the exact HF-like exchange $E_{x,\mathrm{HF}}$, the exchange energy E_x, and the correlation energy E_c, as well as any other molecular property. It is important to stress that for a given set of MOs, the T_s, V_{en}, J, and $E_{x,\mathrm{HF}}$ energy components are evaluated analytically using closed-form expressions (when Gaussian type of basis functions are used). Therefore, the computation of expectation values does not rely on the numerical integration (and are free of *direct* errors). We will call *indirect* grid error to the error propagated by affected MOs to other energy components and molecular properties. In fact, spurious oscillations affecting the derivatives of dipole moment and static polarizabilities, which were already discussed in Chapter 9, partially suffer from the indirect error (both $\boldsymbol{\mu}$ and $\boldsymbol{\alpha}$ are density-based electronic properties, see Eqs. 4.54 and 4.55 in Section 4.6).

In order to analyze the magnitude of these error types, we analyzed the curves of derivatives of all components of the electronic energy: $\mathrm{d}^m T_s/\mathrm{d}\xi^m$, $\mathrm{d}^m V_{en}/\mathrm{d}\xi^m$, $\mathrm{d}^m J/\mathrm{d}\xi^m$, $\mathrm{d}^m E_{x,\mathrm{HF}}/\mathrm{d}\xi^m$, $\mathrm{d}^m E_x/\mathrm{d}\xi^m$, and $\mathrm{d}^m E_c/\mathrm{d}\xi^m$. We performed the KS-DFT SCF computations using several DFAs (B3LYP, VSXC, B97, ωB97M-V, M06-2X, and MN15) and two integration grids, (99, 590) and (1500, 974). Like in Section 10.4, we considered the results obtained with the (1500, 974) grid as the oscillation-free references.

The differences in the properties obtained with the (99, 590) and (1500, 974) grids are shown in Figure 10.7 (see additional Figures A2.11 to A2.14 in the Appendix). Equivalent observations and conclusions can be made by analyzing the derivatives of any order. In this Section, we present the results of the third derivatives with respect to the nuclear displacements. For all DFAs tested, one observes spurious oscillations affecting derivatives of all energy components. The magnitude of the spurious oscillations in the derivatives of T, V_{en}, and J is much larger than those in $E_{x,\mathrm{HF}}$, E_x, and E_c. Errors have opposite sign for different energy components, which in general, follows the sign of the value of the original energy component. For T_s and J, which are positive energy components, the oscillations in $\mathrm{d}^3 T_s/\mathrm{d}\xi^3$ and $\mathrm{d}^3 J/\mathrm{d}\xi^3$ are in-phase and out-of-phase with the ones of $\mathrm{d}^3 V_{en}/\mathrm{d}\xi^3$, $\mathrm{d}^3 E_{x,\mathrm{HF}}/\mathrm{d}\xi^3$, $\mathrm{d}^m E_x/\mathrm{d}\xi^m$, and $\mathrm{d}^m E_c/\mathrm{d}\xi^m$ (which correspond to the energy components that have a negative contribution to the total energy).

Interestingly, for the components evaluated fully analytically, i.e., T_s, V_{en}, J, and $E_{x,\mathrm{HF}}$, there is a cancellation of errors. Namely, for the sum of these energy components, $\mathrm{d}(T + V_{en} + J + V_{nn} + E_{x,\mathrm{HF}})^3/\mathrm{d}\xi^3 = \mathrm{d}(E - E_{xc})^3/\mathrm{d}\xi^3$, the spurious oscillations have magnitudes similar to those of $\mathrm{d}^3 E_{xc}/\mathrm{d}\xi^3$ but are of the opposite phase. Consequently, the total spurious oscillations in the $\mathrm{d}^3 E/\mathrm{d}\xi^3$ curves have smaller amplitudes.

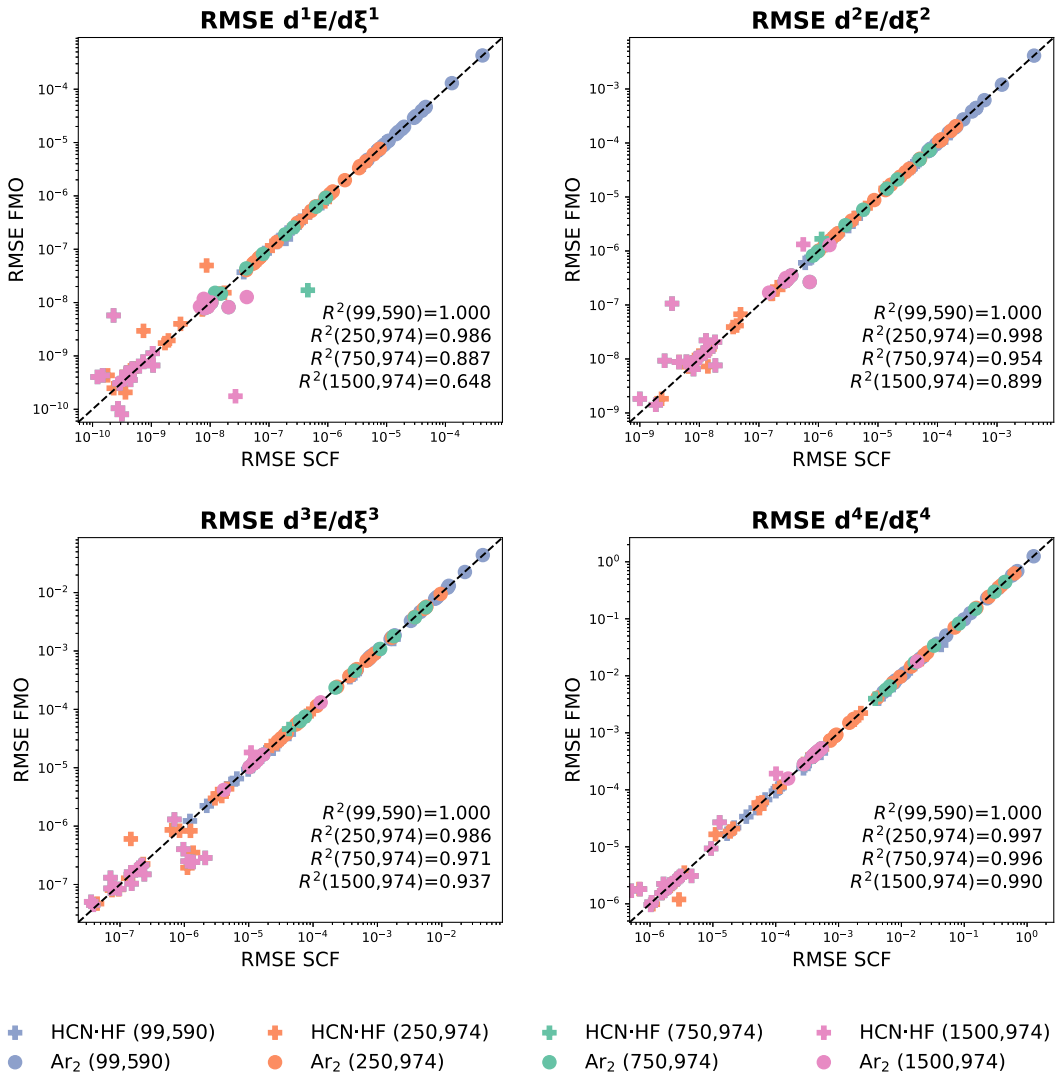

Figure 10.4: Correlation between RMSE in $\mathrm{d}^m E/\mathrm{d}\xi^m$ obtained using the SCF and FMO densities for 20 DFAs combined with the 6-31+G* basis set. Colors represent the integration grid used to calculate RMSE: (99, 590) — blue, (250, 974) — orange, (500, 974) — green, and (1500, 974) — pink. Point shapes represent different chemical systems: HCN · HF — plus, Ar_2 — circle. The line representing ideal correlation, i.e. $y = x$, is shown with a dashed line.

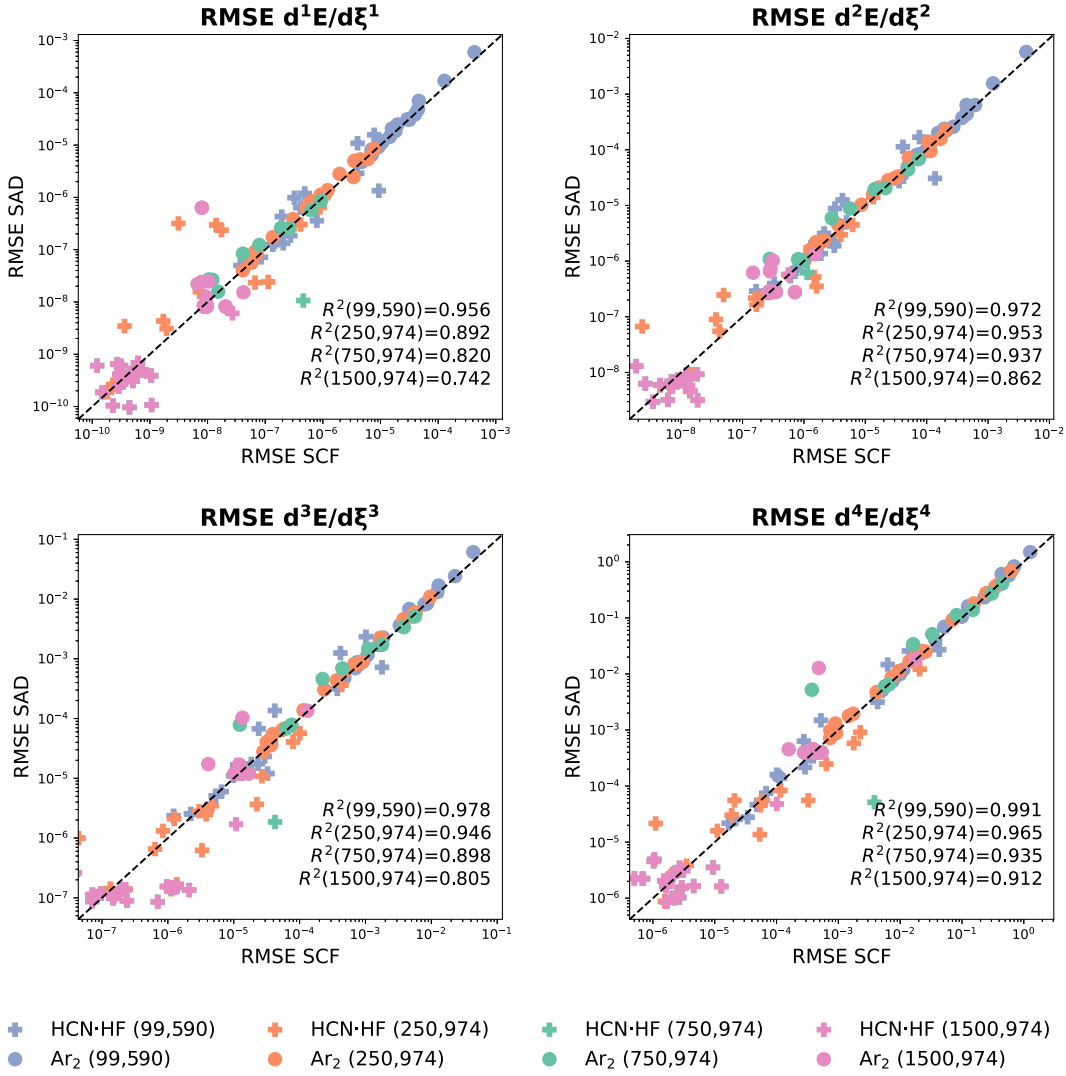

Figure 10.5: Correlation between RMSE in $\mathrm{d}^m E/\mathrm{d}\xi^m$ obtained using the SCF and SAD densities for 20 DFAs combined with the 6-31+G* basis set. Colors represent the integration grid used to calculate RMSE: (99, 590) — blue, (250, 974) — orange, (500, 974) — green, and (1500, 974) — pink. Point shapes represent different chemical systems: HCN · HF — plus, Ar_2 — circle. The line representing ideal correlation, i.e. $y = x$, is shown with a dashed line.

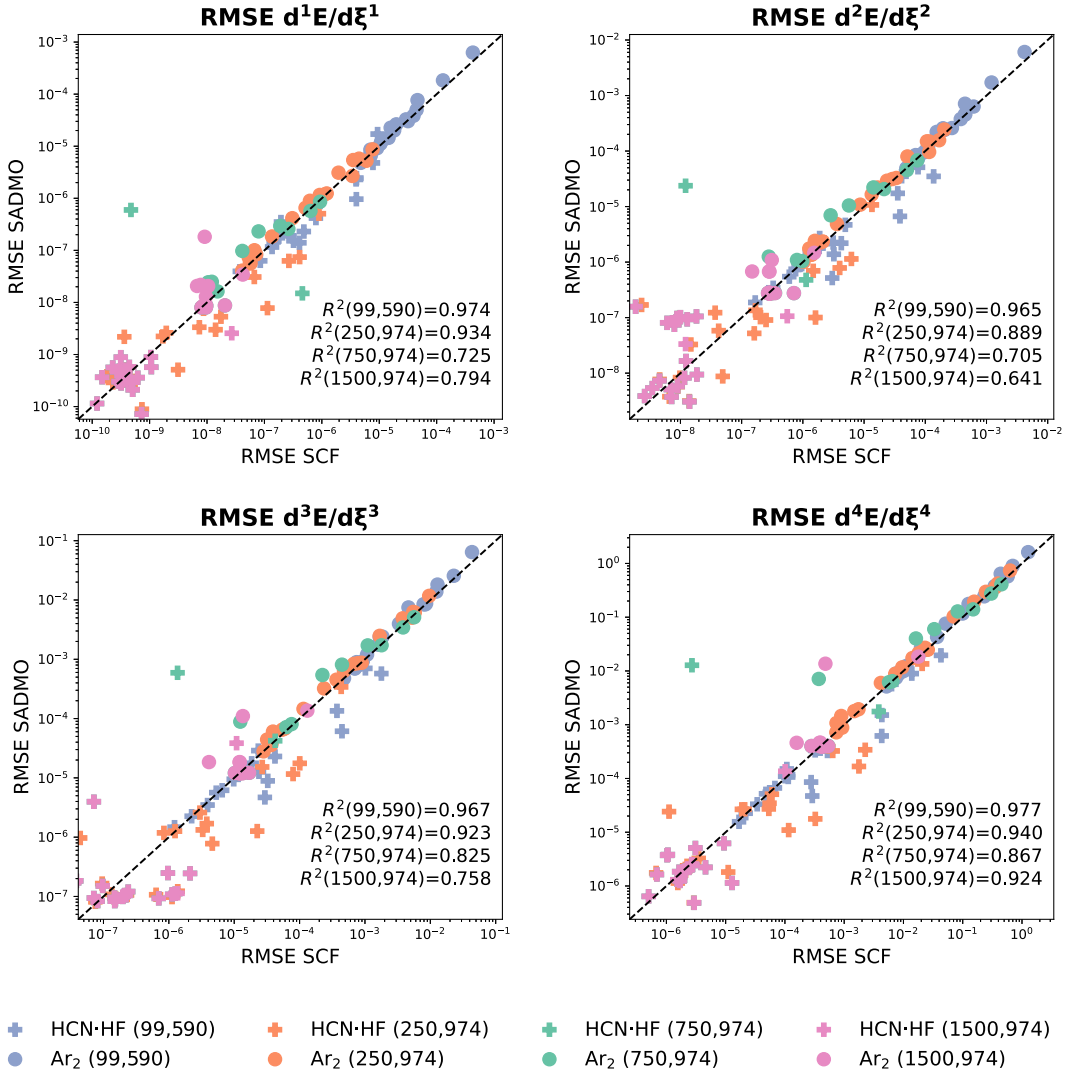

Figure 10.6: Correlation between RMSE in $\mathrm{d}^m E/\mathrm{d}\xi^m$ obtained using the SCF and SADMO densities for 20 DFAs combined with the 6-31+G* basis set. Colors represent the integration grid used to calculate RMSE: (99, 590) — blue, (250, 974) — orange, (500, 974) — green, and (1500, 974) — pink. Point shapes represent different chemical systems: HCN · HF — plus, Ar_2 — circle. The line representing ideal correlation, i.e. $y = x$, is shown with a dashed line.

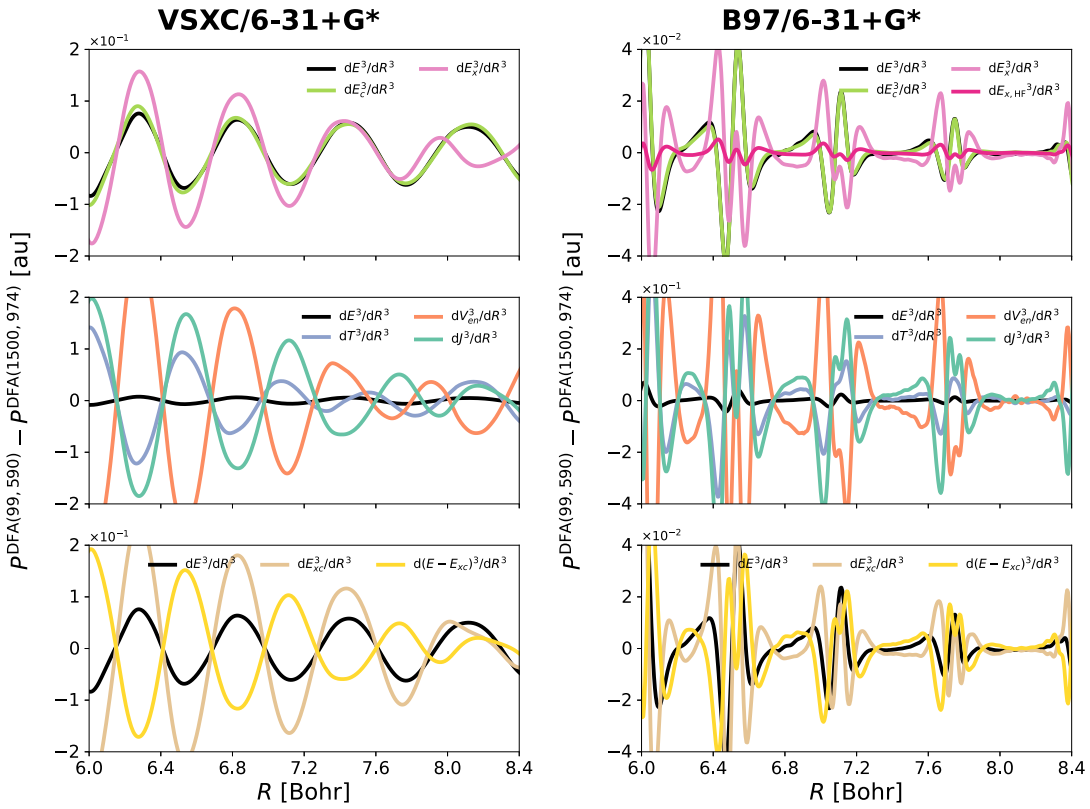

Figure 10.7: Grid-dependent spurious oscillations in the third derivatives of energy components of Ar_2: total electronic (E), kinetic (T), nuclei attraction (V_{en}), Coulomb repulsion (J), exact HF-like exchange $E_{x,\mathrm{HF}}$, exchange (E_x), and correlation (E_c), as well as their sums. Obtained with VSXC/6-31+G* (left panel) and B97/6-31+G* (right panel) as the difference between (99, 590) and (1500, 974) integration grids.

Due to the indirect errors, the spurious oscillations cannot be easily separated between the E_x and E_c using the SCF MOs. However, there are two possibilities to overcome this issue and study only the direct errors affecting the E_x and E_c derivatives.

The first solution is to use MOs obtained from a SCF calculation using a huge grid for the evaluation of the E_x and E_c integrals. However, this solution is time demanding. The second solution is to use the FMO densities, which reproduce the spurious oscillations at almost no computational cost. In fact, using the same set of MOs, the values of T_s, V_{en}, J, and E_x^{HF} are independent of the integration grid.

Proof of this concept is shown in Figure 10.8 (see additional Figures A2.15 to A2.18 in the Appendix), where spurious oscillations for the (99, 590) grid in $\mathrm{d}^3E_x/\mathrm{d}R^3$, and $\mathrm{d}^3E_c/\mathrm{d}R^3$ are calculated using different input densities. In these figures, the black curves were obtained by using MOs from the SCF performed with (1500, 974) grid and only contain the direct errors. The orange curves were obtained with the SCF densities using the (99, 590) grid, and include both the direct and indirect errors. And finally, the green curves were obtained using the FMO densities with a small grid, and reproduce the direct errors calculated using the trial densities. In Figure 10.8, we observe huge discrepancies between the red and black (or green) curves of $\mathrm{d}^3E_x/\mathrm{d}R^3$ due to the effects of the indirect errors. In contrast, the FMO densities reproduce the direct spurious oscillations for $\mathrm{d}^3E_x/\mathrm{d}R^3$, and $\mathrm{d}^3E_c/\mathrm{d}R^3$. The latter is another advantage of using FMO trial densities to study the effects of spurious oscillations.

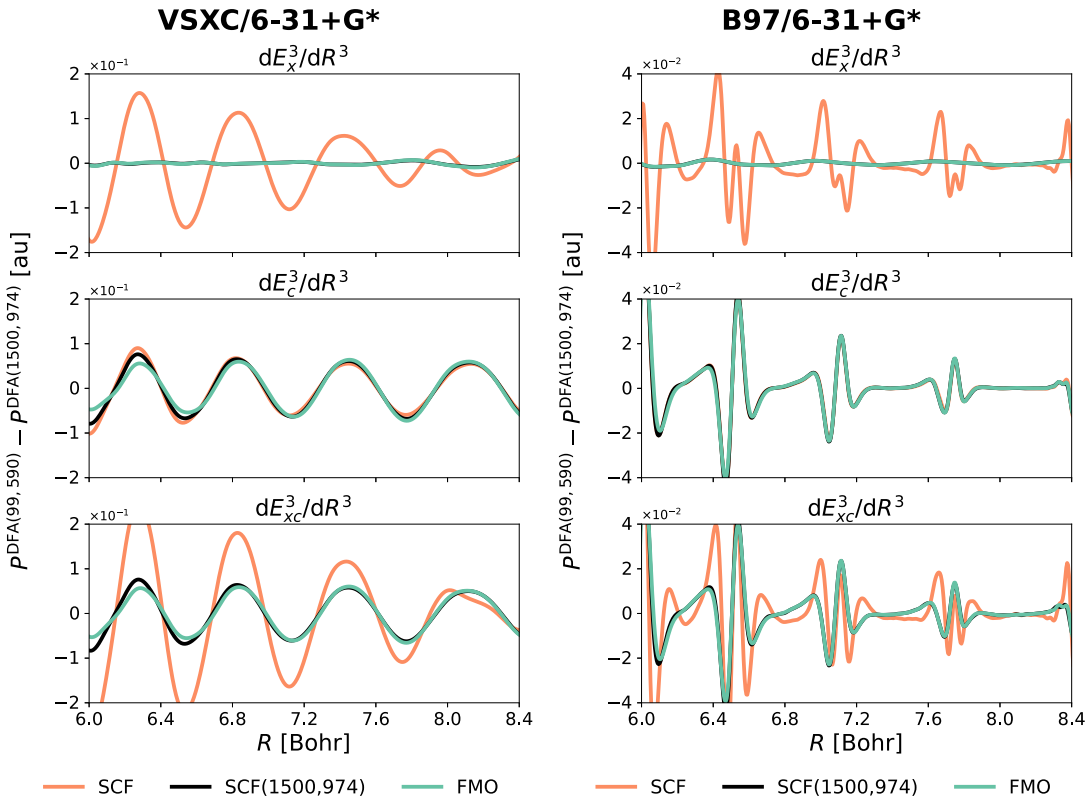

Figure 10.8: Spurious oscillations in the third derivatives of exchange (E_x), correlation (E_c) energy components of Ar_2, as well as their sum, obtained using VSXC/6-31+G* (left panel) and B97/6-31+G* (right panel), and the (99, 590) and (1500, 974) integration grids. Three different types of the input densities are employed: SCF — density obtained in SCF using corresponding grids, SCF(1500, 974) — density obtained in SCF using the reference (1500, 974) grid, FMO — trial density constructed using FMOs obtained for one geometry (R=7.2 Bohr) and the (250, 974) integration grid.

10.6 The Origin of Grid Dependency in DFAs: Exchange and Correlation

With our algorithm, we can analyze the origin of the direct errors due to spurious oscillations. In particular, we can separately inspect the errors coming from spurious oscillations in the derivatives of exchange and correlation energies, $\mathrm{d}^m E_x/\mathrm{d}\xi^m$ and $\mathrm{d}^m E_c/\mathrm{d}\xi^m$, respectively. However, to that end one has to first circumvent the problem of indirect errors – which, as we demonstrated in the previous section, can be accomplished using the FMO trial densities.

For forty-five DFAs, we computed the average ratio between the RMSE of $\mathrm{d}^m E_x/\mathrm{d}\xi^m$ and $\mathrm{d}^m E_c/\mathrm{d}\xi^m$ using the FMO trial densities. We included the ratios calculated for the third and fourth derivatives, using two integration grids (99, 590) and (250, 974), for $HCN \cdot HF$, $HCN \cdot HCl$, $OC \cdot HF$, $N_2 \cdot HF$, $HCN \cdot BrF$, and Ar_2. The final results are summarized in Table 10.2, whereas the system-specific data is compiled in Tables 10.3 and 10.4, for the (250, 974) grid, and Tables A3.7 to A3.8, for the (99, 590) grid.

The results may be summarized with the following three statements. First, for half of the studied DFAs, the total error in $\mathrm{d}^m E_{xc}/\mathrm{d}\xi^m$ is dominated by the error in just one component. For the second half, which mostly consists of standalone functionals, the errors in the exchange and correlation components have comparable magnitude. Second, the relative grid stability of DFAs is very similar for all energy derivatives. Third, the main source of error (exchange or correlation) may change with the grid and the chemical

system.

To facilitate the analysis, we first sorted the DFAs according to their relative stability in the derivatives of the exchange-correlation energy, $d^m E_{xc}/d\xi^m$. Namely, for a given property and molecular system, we established the relative error of a given DFA with respect to the lowest RMSE value among all 45 DFAs. Examples depicting the relative grid stability of $d^m E_{xc}/d\xi^m$ (as well as of $d^m E_x/d\xi^m$ and $d^m E_c/d\xi^m$) are shown in Figures 10.9 and 10.10. For example, for the $d^3 E_{xc}/d\xi^3$ derivatives, one finds the lowest RMSE values of $HCN \cdot HF$, $HCN \cdot HCl$, and Ar_2 using BH&H; $OC \cdot HF$, $N_2 \cdot HF$ using SPW92; and $HCN \cdot BRF$ using BH&HLYP. Second, we construct a vector $\overline{K}^{\mathrm{DFA}}$ with the lowest RMSE value for a selection of grids and properties. Namely, we choose the (99, 590) and (250, 974) grids, and $d^3 E_{xc}/d\xi^3$ and $d^4 E_{xc}/d\xi^4$ properties, giving rise to fourth-dimensional $\overline{K}^{\mathrm{DFA}}$ vector for each DFA, characterized by the lowest and largest components, $\overline{K}^{\mathrm{DFA}}_{\mathrm{min}}$ and $\overline{K}^{\mathrm{DFA}}_{\mathrm{max}}$. Third, we identify the lowest $\overline{K}_{\mathrm{min}}$ among all the DFAs, $\overline{K}_{\mathrm{min}}$, and use it to define $\overline{K} = \overline{K}_{\mathrm{max}}/\overline{K}_{\mathrm{min}}$

For example, the BH&H functional yields RMSE values which can be three times larger than the smallest RMSE found (for a given property, grid, and molecular system). On the contrary, N12 yields RMSE values of $d^3 E_{xc}/d\xi^3$ and $d^4 E_{xc}/d\xi^4$ which can be 70000 times larger than the smallest RMSE found.

Now we will discuss the main source of grid instability and spurious oscillations in the selected DFAs. This is briefly summarized in Figure A3.9. Starting from the three most grid-robust DFAs, namely BH&H, SVWN5, and SPW92 the grid errors coming from the exchange part are much larger than the ones from the correlation part. All these DFAs incorporate the LSDA exchange and correlation (except BH&H which includes the LYP correlation [296]).

Within the DFAs from the BLYP family (BLYP, B1LYP, B3LYP, BH&HLYP, LC-BLYP, and CAM-B3LYP), for which $\overline{K}$ ranges between 5 and 10, the B88 exchange part is responsible for the spurious oscillations. On average, it yields 10-20 larger RMSE than the LYP correlation part. In fact, the LYP correlation yields a RMSE as small as the one of LSDA correlation functionals (see Figures 10.9 and 10.10).

In contrast, within the DFAs from the PBE family (PBE, PBE0, PBE50, and LC-ωPBE), which are also grid-robust functionals with $\overline{K}$ between 5 and 30, errors coming from the exchange and correlation parts have a very similar magnitude. On average, it is the exchange part that yields larger RMSE. A closer inspection reveals that when the (99, 590) grid is used, the PBE exchange introduces much more grid-related errors than the PBE correlation (see Table A3.7), whereas, for the (250, 974) grid, it is the correlation which yields slightly larger RMSE errors (see Table 10.3). The PBE correlation yields few times larger RMSE than the LSDA and LYP correlation.

Next in stability are the functionals from the TPSS family (TPSS, RevTPSS, and TPSSh), the MN15 family (MN15 and MN15-L) and the B1PW91 functional. It is clear that the TPSS exchange is much more grid dependent than the TPSS correlation, with average ratios in the range of 10 - 24 (see Table 10.2). For the MN15, MN15-L, and B1PW91 functionals, on average, the exchange and correlation parts yield quite similar RMSE. However, a closer inspection of the data reveals that, in fact, for almost all chemical systems but $HCN \cdot BrF$, it is the exchange part which is more stable (ratios are below 1, see Tables A3.8 and 10.4).

Among the B97-V functionals (ωB97X-V, ωB97M-V, and B97M-V), which have values of $\overline{K}$ in the range 400 - 700, and the B97-based functionals (B97, B97-D, ωB97, ωB97X, ωB97X-D, and ωB97X-D3), which have values of $\overline{K}$ in the range 5000 - 70000, the correlation part is the main (and most likely the only) source of grid instabilities and spurious oscillations. Ratios between the errors from oscillations in exchange and correla-

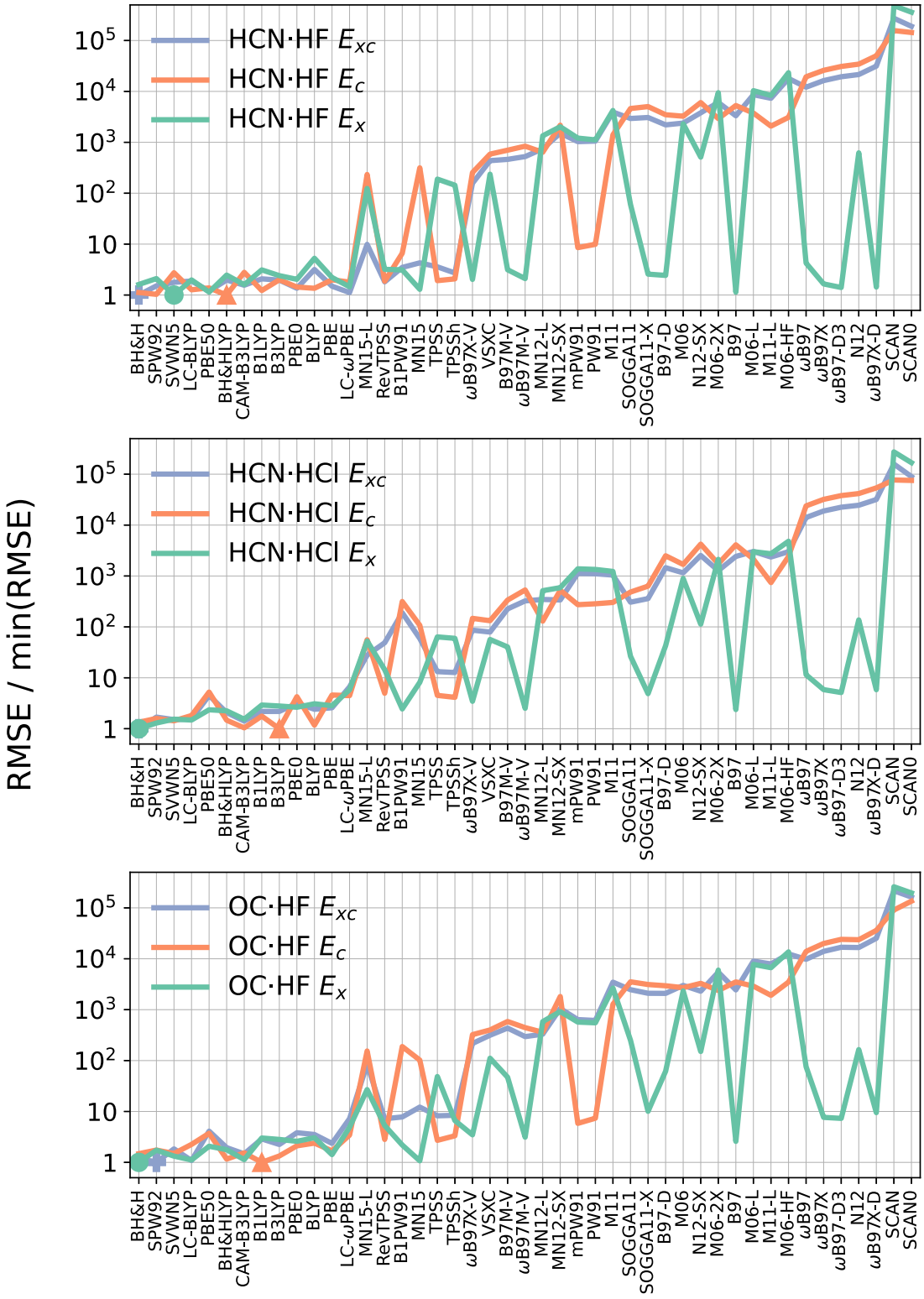

Figure 10.9: The relative values of RMSE in $\mathrm{d}^4E_{xc}/\mathrm{d}\xi^4$, $\mathrm{d}^4E_x/\mathrm{d}\xi^4$, and $\mathrm{d}^4E_c/\mathrm{d}\xi^4$ for HCN · HF, HCN · HCl, and OC · HF, with respect to the minimal RMSE value found within the whole set. The minimum values are marked with a + sign for $\mathrm{d}^4E_{xc}/\mathrm{d}\xi^4$, ▲ for $\mathrm{d}^4E_c/\mathrm{d}\xi^4$, and • for $\mathrm{d}^4E_x/\mathrm{d}\xi^4$. The ordering of DFAs corresponds to the one in Table 10.2. Results obtained using the (250, 974) grid.

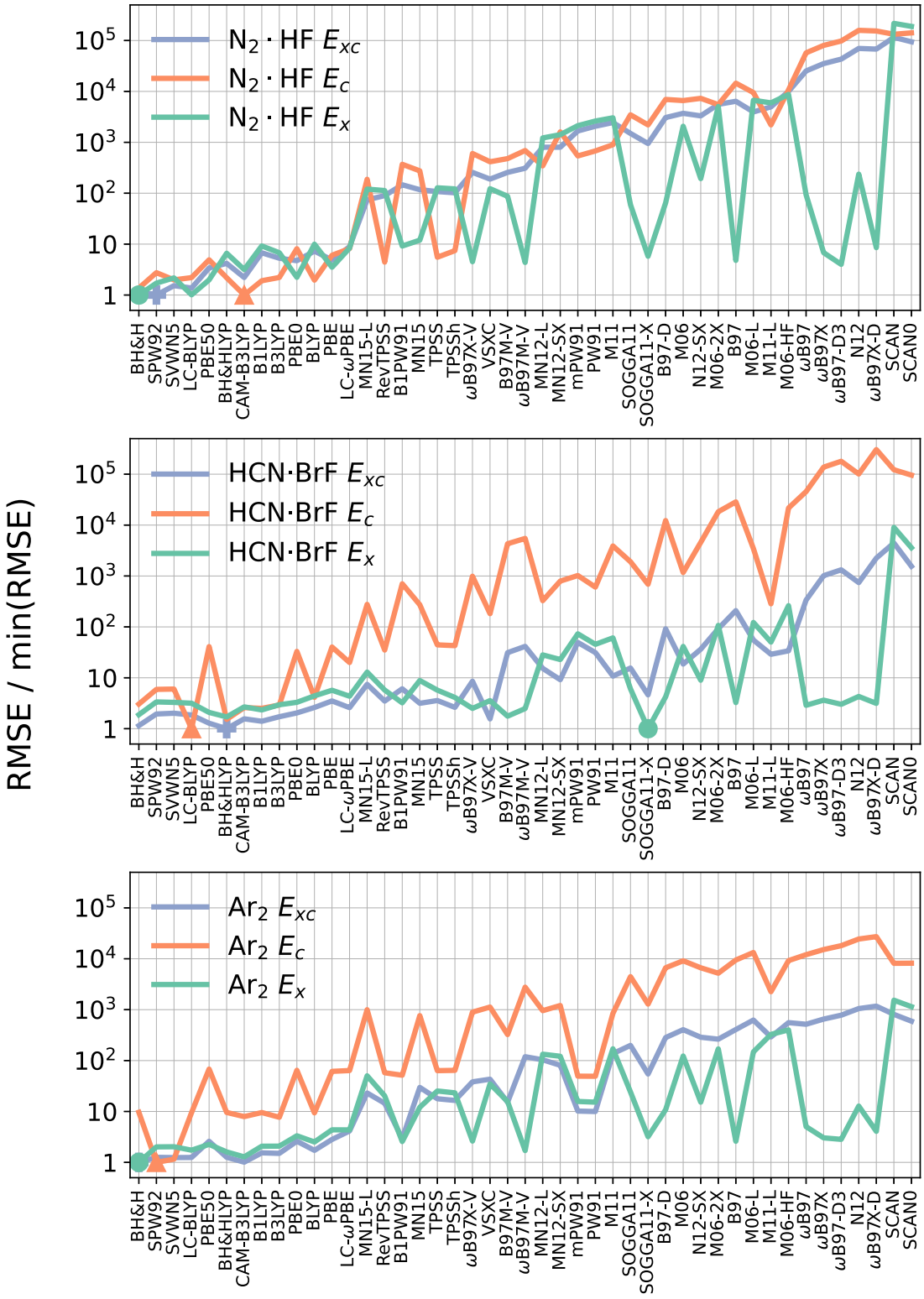

Figure 10.10: The relative values of RMSE in $\mathrm{d}^4E_{xc}/\mathrm{d}\xi^4$, $\mathrm{d}^4E_x/\mathrm{d}\xi^4$, and $\mathrm{d}^4E_c/\mathrm{d}\xi^4$ for $N_2\cdot$HF, HCN$\cdot$BrF, and Ar_2, with respect to the minimal RMSE value found within the whole set. The minimum values are marked with a + sign for $\mathrm{d}^4E_{xc}/\mathrm{d}\xi^4$, ▲ for $\mathrm{d}^4E_c/\mathrm{d}\xi^4$, and • for $\mathrm{d}^4E_x/\mathrm{d}\xi^4$. The ordering of DFAs corresponds to the one in Table 10.2. Results obtained using the (250, 974) grid.

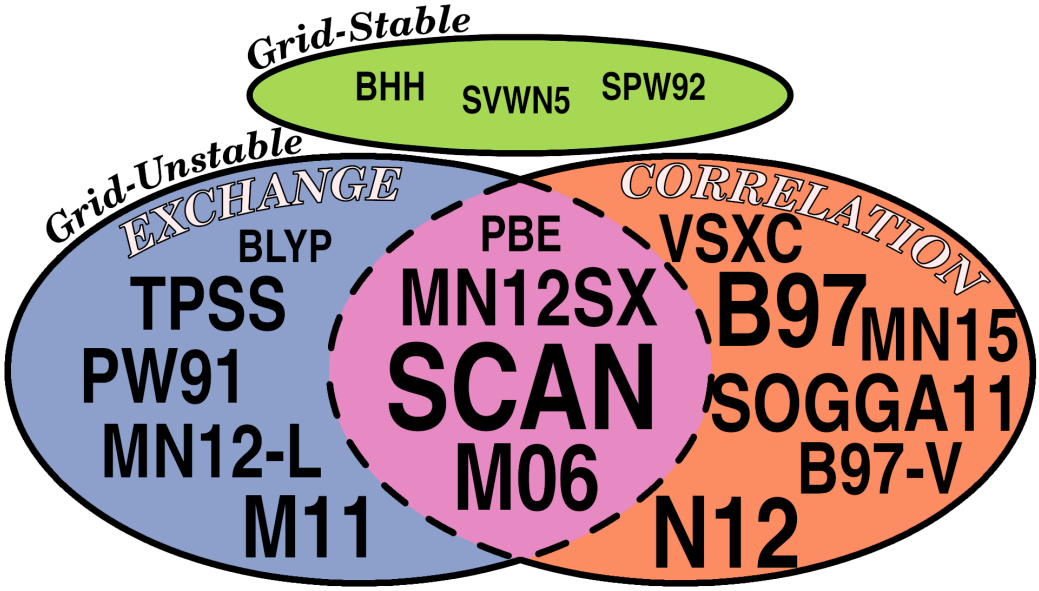

Figure 10.11: Families of DFAs grouped by the main origin of their grid instability, namely, exchange (blue area), correlation (orange area), and both (pink area). Functionals in which both exchange and correlation are exceptionally stable are shown in green area. The font size indicates the DFAs overall grid instability, namely, the larger the names the more grid unstable a given DFA is. The description of the functional families is given in Table A3.9.

tion are much smaller than one, namely in the range of 10^{-5} — 10^{-3}, especially when the (250, 974) grid is used. Interestingly, the exchange part of the B97 and B97-V functionals is as stable as the B88 exchange (see Figures 10.9 and 10.10).

The functionals from the PW91 family (PW91 and mPW91), suffer from large errors due to spurious oscillations (i.e., $\overline{K}$ = 2000 - 3000). In the case of these functionals, both exchange and correlation are highly grid sensitive to spurious oscillations (see Figures 10.9 and 10.10), although, the exchange yields errors which, on average, are 24 - 28 times larger than those due to correlation.

Lastly, we discuss the remaining DFAs, which suffer from huge spurious oscillations and almost all of them belong to rungs 4 and 5 from Section 9.5 (and correspond to $\overline{K}$ = 400 - 10^5). All of them were developed as the standalone functionals, that is, the exchange and correlation parts were developed simultaneously and are not separable. For the VSXC functional, which suffers from the smallest errors in this group of DFAs, the correlation yields on average two times larger errors than the exchange. On the contrary, for the SOGGA11, SOGGA11-X, N12, and N12-SX functionals, the correlation part is responsible of the spurious oscillations. More strikingly, in the case of SOGGA11-X and N12, the exchange part is almost as stable as the B88, B97 and PBE exchange and it is the correlation part that most suffers from spurious oscillations. In M11, M11-L, MN12-L, MN12-SX, SCAN, and SCAN0, both exchange and correlation cause large errors. Similarly, in the DFAs from the M06 family, both components are greatly grid sensitive to spurious oscillations and there is no clear dominance of one component (trends are system and grid specific).

Table 10.2: The average ratios between the RMSE of $d^m E_x/d\xi^m$ and $d^m E_c/d\xi^m$ for some DFAs (the values in parentheses). DFAs and they their relative grid stability (see description in the text): numbers in first column denote the maximum desired relative errors in $d^m E_{xc}/d\xi^m$ with respect to the most stable DFA, $\overline{K}$ (see description in text). For each DFA, Numbers in parentheses are the average values of ratios between the RMSE for $d^m E_x/d\xi^m$ and $d^m E_c/d\xi^m$, including $m = 3, 4$, all molecules and the (99, 590) and (250, 974) grids. (the presented numbers are obtained by averaging the values from the corresponding rows in Tables A3.7 to 10.4).

$\overline{K}$	DFAs (average RMSE($d^m E_x/d\xi^m$) / RMSE($d^m E_c/d\xi^m$))			
1				
2				
3	BH&H (5.4)			
4	SPW92 (11.0)	SVWN5 (10.8)		
5	LC-BLYP (21.8)	PBE50 (1.8)		
6	BH&HLYP (11.2)	CAM-B3LYP (11.6)		
7				
8	B1LYP (12.6)	B3LYP (13.0)	PBE0 (2.6)	
9				
10	BLYP (18.8)	PBE (3.1)		
20				
30	LC-ωPBE (4.2)			
40				
50				
60				
70				
80				
90	MN15-L (1.3)			
100	RevTPSS (10.8)			
200	B1PW91 (1.1)	MN15 (0.9)	TPSS (23.7)	TPSSh (14.7)
300				
400	ωB97X-V (0.6)			
500				
600	VSXC (0.5)			
700	B97M-V (0.6)	ωB97M-V (0.1)		
800				
900	MN12-L (3.1)			
1000				
2000	MN12-SX (1.6)	mPW91 (28.2)		
3000	PW91 (24.2)			
4000	M11 (3.1)	SOGGA11 (0.2)	SOGGA11-X (0.2)	
5000	B97-D ($6.6{\cdot}10^{-2}$)	M06 (0.6)	N12-SX (0.2)	
6000				
7000	M06-2X (1.1)			
8000	B97 ($2.8{\cdot}10^{-2}$)			
9000	M06-L (1.2)	M11-L (4.0)		
10000				
20000	M06-HF (1.8)			
30000	ωB97 ($1.6{\cdot}10^{-2}$)			
40000	ωB97X ($1.0{\cdot}10^{-2}$)			
50000	ωB97X-D3 ($8.5{\cdot}10^{-3}$)			
60000				
70000	N12 ($2.8{\cdot}10^{-2}$)	ωB97X-D ($6.1{\cdot}10^{-3}$)		
80000				
90000				
$> 10^5$	SCAN (3.6)	SCAN0 (2.7)		

Table 10.3: *Part 1)* Ratios between the RMSE values of E_x and E_c for some DFAs combined with the (250, 974) grid. The numbers correspond to the averages of ratios for the $\mathrm{d}^3E/\mathrm{d}\xi^3$ and $\mathrm{d}^4E/\mathrm{d}\xi^4$ derivatives, computed separately for each chemical system. The second and third columns show the names of the (meta-)GGAs exchange and correlation functionals included in a given DFA (information not shown for standalone functionals, for which exchange and correlation parts are not meant to be used separately).

DFA	E_x^{DFA}	E_c^{DFA}	HCN · HF	HCN · HCl	OC · HF	N_2 · HF	HCN · BrF	Ar_2	He_2
SVWN5	S	VWN5	0.5	1.7	1.4	1.9	35.0	25.5	5.2
SPW92	S	PW92	2.1	1.3	1.5	0.8	36.0	28.5	6.7
BH&H	S	LYP	1.2	0.9	1.1	1.1	37.3	1.6	2.9
BLYP	B88	LYP	4.8	4.2	2.0	9.6	101.1	4.0	$1.1\cdot10^{4}$
B1LYP	B88	LYP	2.4	2.5	4.8	8.9	58.4	3.3	$5.0\cdot10^{3}$
B3LYP	B88	LYP	1.4	4.1	3.3	4.3	61.8	4.1	$8.1\cdot10^{3}$
BH&HLYP	B88	LYP	3.6	2.5	2.1	5.0	67.0	2.5	$4.4\cdot10^{3}$
LC-BLYP	B88	LYP	1.6	0.9	0.7	0.8	208.1	2.7	5.1
CAM-B3LYP	B88	LYP	0.9	2.2	1.2	3.9	62.7	2.5	$3.3\cdot10^{3}$
PBE	PBE	PBE	1.1	0.6	1.4	0.8	11.2	0.9	0.6
PBE0	PBE	PBE	1.5	0.9	2.0	0.5	8.9	0.7	0.5
PBE50	PBE	PBE	0.8	0.6	0.7	0.6	5.2	0.4	0.3
LC-ωPBE	PBE	PBE	0.9	1.7	7.9	1.5	16.1	0.9	0.1
TPSS	TPSS	TPSS	103.7	20.4	30.6	73.3	11.8	5.5	0.9
RevTPSS	RevTPSS	RevTPSS	2.7	3.3	2.7	61.8	13.0	4.9	0.9
TPSSh	TPSS	TPSS	74.5	13.6	3.3	36.5	8.2	5.0	0.8
B97			$3.0\cdot10^{-4}$	$5.3\cdot10^{-4}$	$7.3\cdot10^{-4}$	$4.1\cdot10^{-4}$	$8.8\cdot10^{-3}$	$4.7\cdot10^{-3}$	$1.7\cdot10^{-1}$
B97-D			$1.2\cdot10^{-3}$	$1.7\cdot10^{-2}$	$2.6\cdot10^{-2}$	$1.3\cdot10^{-2}$	$2.8\cdot10^{-2}$	$3.1\cdot10^{-2}$	$4.3\cdot10^{-1}$
ωB97			$4.8\cdot10^{-4}$	$5.2\cdot10^{-4}$	$8.5\cdot10^{-3}$	$2.8\cdot10^{-3}$	$4.5\cdot10^{-3}$	$7.1\cdot10^{-3}$	$1.4\cdot10^{-2}$
ωB97X			$1.5\cdot10^{-4}$	$1.9\cdot10^{-4}$	$3.7\cdot10^{-4}$	$9.3\cdot10^{-5}$	$2.1\cdot10^{-3}$	$4.5\cdot10^{-3}$	$3.1\cdot10^{-2}$
ωB97X-D			$8.2\cdot10^{-5}$	$1.1\cdot10^{-4}$	$2.6\cdot10^{-4}$	$6.5\cdot10^{-5}$	$8.4\cdot10^{-4}$	$3.6\cdot10^{-3}$	$6.2\cdot10^{-2}$
ωB97X-D3			$9.6\cdot10^{-5}$	$1.5\cdot10^{-4}$	$3.0\cdot10^{-4}$	$5.2\cdot10^{-5}$	$1.4\cdot10^{-3}$	$3.7\cdot10^{-3}$	$4.1\cdot10^{-2}$
ωB97M-V			$2.9\cdot10^{-3}$	$3.9\cdot10^{-3}$	$7.2\cdot10^{-3}$	$7.5\cdot10^{-3}$	$4.2\cdot10^{-2}$	$1.8\cdot10^{-2}$	$8.7\cdot10^{-3}$
ωB97X-V			$1.0\cdot10^{-2}$	$2.1\cdot10^{-2}$	$1.2\cdot10^{-2}$	$9.0\cdot10^{-3}$	$2.3\cdot10^{-1}$	$5.0\cdot10^{-2}$	$2.4\cdot10^{-2}$
B97M-V			$5.7\cdot10^{-3}$	$1.3\cdot10^{-1}$	$1.4\cdot10^{-1}$	$3.1\cdot10^{-1}$	$3.5\cdot10^{-2}$	$7.9\cdot10^{-1}$	$7.1\cdot10^{-1}$

Table 10.4: *Part 2)* Ratios between the RMSE values of E_x and E_c for some DFAs combined with the (250, 974) grid. The numbers correspond to the averages of ratios for the $d^3E/d\xi^3$ and $d^4E/d\xi^4$ derivatives, computed separately for each chemical system. The second and third columns show the names of the (meta-)GGAs exchange and correlation functionals included in a given DFA (information not shown for standalone functionals, for which exchange and correlation parts are not meant to be used separately).

DFA	E_x^{DFA}	E_c^{DFA}	HCN · HF	HCN · HCl	OC · HF	N_2 · HF	HCN · BrF	Ar_2	He_2
B1PW91	B88	PW91	$4.8{\cdot}10^{-1}$	$1.2{\cdot}10^{-2}$	$2.0{\cdot}10^{-2}$	$4.3{\cdot}10^{-2}$	$4.0{\cdot}10^{-1}$	$7.2{\cdot}10^{-1}$	$4.4{\cdot}10^{-1}$
PW91	PW91	PW91	116.2	5.7	118.5	6.0	5.9	3.3	0.6
mPW91	mPW91	mPW91	151.1	6.2	129.1	6.3	5.6	3.3	0.6
VSXC			0.4	0.5	0.4	0.5	1.9	0.5	0.5
SCAN			3.5	4.4	4.5	2.6	6.5	3.0	6.1
SCAN0			3.0	3.0	2.6	1.9	3.5	2.3	4.5
M06			0.7	0.5	1.1	0.4	2.1	0.2	$2.6{\cdot}10^{-2}$
M06-L			2.7	1.3	3.2	0.9	2.5	0.2	$5.5{\cdot}10^{-2}$
M06-HF			7.1	1.9	5.3	1.2	0.9	0.7	$1.7{\cdot}10^{-2}$
M06-2X			3.1	1.2	3.2	1.3	0.5	0.6	$9.4{\cdot}10^{-2}$
M11			3.5	4.9	3.1	5.1	1.2	2.9	$2.5{\cdot}10^{-2}$
M11-L			4.6	4.4	5.1	4.3	12.5	2.0	1.1
SOGGA11			$1.0{\cdot}10^{-2}$	$5.6{\cdot}10^{-2}$	$9.8{\cdot}10^{-2}$	$2.6{\cdot}10^{-2}$	$3.1{\cdot}10^{-1}$	$7.9{\cdot}10^{-2}$	1.9
SOGGA11-X			$8.5{\cdot}10^{-4}$	$7.6{\cdot}10^{-3}$	$3.3{\cdot}10^{-3}$	$3.0{\cdot}10^{-3}$	$1.3{\cdot}10^{-1}$	$4.8{\cdot}10^{-2}$	$5.0{\cdot}10^{-1}$
N12			$2.1{\cdot}10^{-2}$	$4.5{\cdot}10^{-3}$	$9.6{\cdot}10^{-3}$	$2.7{\cdot}10^{-3}$	$3.9{\cdot}10^{-3}$	$1.2{\cdot}10^{-2}$	$3.7{\cdot}10^{-1}$
N12-SX			$9.0{\cdot}10^{-2}$	$3.0{\cdot}10^{-2}$	$5.9{\cdot}10^{-2}$	$3.8{\cdot}10^{-2}$	$1.3{\cdot}10^{-1}$	$4.5{\cdot}10^{-2}$	$3.6{\cdot}10^{-1}$
MN12-L			2.4	4.6	2.3	5.1	6.7	1.9	0.6
MN12-SX			1.1	1.4	0.8	1.4	2.2	1.4	5.9
MN15			$7.3{\cdot}10^{-3}$	$8.0{\cdot}10^{-2}$	$1.7{\cdot}10^{-2}$	$4.9{\cdot}10^{-2}$	2.6	$2.1{\cdot}10^{-1}$	$3.0{\cdot}10^{-1}$
MN15-L			0.6	1.1	0.3	1.0	3.7	0.7	0.1

10.7 Conclusions

In this Chapter, we have performed a detailed analysis on the origin of spurious oscillations. First, we showed that the oscillations are not strongly basis set dependent; hence, they may be estimated using compact basis sets, such as 6-31G or Def2SVP. For instance, 6-31+G* predicts the magnitude of the error due to the spurious oscillation in the energy and property derivatives computed with aug-cc-pVTZ. This conclusion holds also for the derivatives of dipole moment and static polarizability, which, in general, are more susceptible to the quality of the basis set than the electronic energies.

Second, we unveiled that to identify errors coming from the spurious oscillations, one does not require to utilize the Kohn-Sham MOs and the corresponding densities. One can actually utilize trial densities that were not obtained from an SCF procedure, such as the SAD, SADMO, and FMO densities. The SAD and SADMO densities are already available in most computational packages and are a opular choice for the initial guess in the SCF procedures. The FMO densities can be trivially constructed from LCAO coefficients of the equilibrium geometry and used for all the other points of the property curve. The best accuracy in the description of oscillations was observed for the FMO densities. This approach, while being limited to the derivatives of electronic energy, can provide very accurate estimates of the errors coming from spurious oscillations at a very low computational cost. The cost of our diagnostic reduces to just one KS-DFT computation when the FMO densities are employed.

In this Chapter, we also discussed the consequences of direct and indirect errors due to spurious oscillations. Small integration grids affect not only the evaluation of the exchange-correlation energies, which leads to *direct* errors, but also affect the Kohn-Sham MOs through the SCF procedure; these errors are called *indirect*. Indirect errors affect molecular properties computed with such orbitals. Moreover, the indirect errors hinder a separation of spurious oscillations coming from the exchange and correlation components. However, the separation can be achieved using MOs free of the spurious oscillation error, or the same FMOs for all the points of the PEC.

Lastly, in this work, we study wether the main source of spurious oscillations is the exchange or the correlation energy term. In LSDAs and the functionals from the BLYP family, the spurious oscillations are small and emerge almost exclusively from the Slater and B88 exchange functionals. The same origin of the oscillations is observed for the moderately grid-sensitive functionals from the TPSS, PW91, M11, and MN12 families. Regarding the correlation functionals, only PW92, VWN5 (LSDAs) and LYP (GGA) are exceptionally grid-robust. More interestingly, our study reveals that the huge direct spurious oscillations from the B97 and N12 families are only due to the the correlation part. The exchange part of these functionals is almost as robust as the B88 or PBE exchanges. Lastly, in the case of most standalone functionals, such as the ones of the M06, M11 and MN15 families, which also yield huge spurious oscillation errors, both exchange and correlation suffer from considerable grid dependencies.

Chapter 11

How Many Electrons Does a Molecular Electride Hold?

11.1 Introduction

Electrides are novel ionic substances where the electrons act as the anions in crystal structures and molecules [371–375]. These anionic electrons are localized in particular parts of the crystals and are considered as the individual entities that formally do not belong to any atom. Electrides share some similarities with the F-centers in crystals, which are vacancies in the crystal structure arising from the (random) misalignment of atoms (or their complete lack). For instance, both the anionic electrons of the electrides and the electrons localized in the F-centers are easily photo-excited, which leads to the intense absorption in the UV-Vis range and acute colors of such materials. Actually, the electrides can be treated as the *stoichiometric* F-centers. Although the notion of free electrons may also be found in metals or alkaline metal solutions of ammonia, the character of these electrons is drastically different to the ones that belong to the anions of the electrides. In the former case, electrons are greatly delocalized over the whole lattice and still can be formally assigned to the metallic centers, whereas in electrides they are hugely localized and do not belong to any atom. In the case of alkaline metal solutions of ammonia, free electrons are solvated by the ammonia molecules and form a disordered liquid state, unlike the case of crystal electrides, where the anionic vacancies form ordered geometrical structures.

The first hypothesis on the existence of electrides was done by James L. Dye, who had previously worked with alkalides [376]. In 1983, Dye managed to synthesize the first electride [377], namely $Cs^+(18C6)_2 \cdot e^-$, and a few years later successfully characterized the anionic vacancies in its crystal structure [378]. However, $Cs^+(18C6)_2 \cdot e^-$ was inherently unstable under normal laboratory conditions. It took Dye two decades more, up to 2005, to finally synthesize a first stable organic electride, Na^+[tripip-aza(222)]$\cdot e^-$. However, the actual first realization of a stable *inorganic* electride, $[Ca_{24}Al_{28}O_{64}]^{4+}(4e^-)$, synthesis was achieved in 2003 by the group of Hideo Hosono [379]. $[Ca_{24}Al_{28}O_{64}]^{4+}(4e^-)$ has found many applications, to name the most prominent ones, it can act as an electron emitter [380], an electron injection electrode [381], a synthetic organic reagent in coupling reactions [382], a cathode in the top-emission organic light-emitting diodes [383] and fluorescent lamps [384], a reversible H_2 storage device [385], a catalyst in the synthesis of ammonia [386], an electrode in electrochemical reactions [387], and an electric conductor [388]. Today, there is a large collection of (in)organic electrides, which are classified according to the spatial pattern of the lattice voids with the electride electrons, namely zero- [389], one- [390–393], two- [394–401], and three-dimensional electrides [402–407], as well as electride nanoparticles [408, 409]. In the recent years, the concept of the single-molecule electrides (known as the molecular electrides) emerged [375], and it will be the topic of this Chapter.

Since the electride electrons can be found in the single molecules, it was of key importance to establish a unique set of features which could be used to produce an unbiased classification of molecules as molecular electrides [375]. Electrides present several special properties which may be helpful in their identification. Solid-state electrides, show huge magnetic susceptibilities due to the unpaired electrons in the vacancies; most electrides are expected to be Mott insulators, and some electrides show superconducting behavior at low temperatures [381, 410–412]. Moreover, electrides are characterized by huge nonlinear optical response properties, and indeed some electrides have the largest first hyperpolarizabilities ever reported [413, 414]. However, the electrons localized in electrides can not be experimentally detected because the density of the free electrons is not large enough to be located in the X-ray of the crystal structure. Therefore, most experimental confirmations rely on indirect structural measurements, such as comparisons with other crystal structures of electrides and alkalides [415–417], chemical shifts of the cor-

responding cations, EPR studies, and atomic resolution scanning tunneling microscopy [418]. In contrast, theoretical chemistry provides reliable tools for the topological characterization of the electrides because it has an easy access to direct information on the electronic density of the system. The most common tools to identify localized electrons are the topology of the electronic density (done by the means of Quantum Theory of Atoms in Molecules, QTAIM) and the presence of NNACPs [238], the Electron Localization Function (ELF) [419–421], the non-covalent interactions (NCI) [422, 423], the localized-orbital locator (LOL) [424], and the Laplacian of the electron density [238]. Although the increased values of the localization indicators may be observed in many chemical systems that cannot be considered electrides, molecular electrides are characterized using several of them simultaneously. In 2015, Postils and coworkers concluded that the rarest feature among these indicators is the presence of NNACP in the topology of the electronic density. They considered it as a unique footprint of the electride anionic electrons.

It was important to find a suitable electronic structure method to detect the NNACP and characterize the corresponding basin, which character influences all the properties of the electride. One of the basic questions, which had not been systematically answered, was which is the most likely number of electrons held by the molecular electrides. This question was also directly related to the formal oxidation state of the electride anion, or equivalently, to what extent these localized electrons can be considered as the anions (with a formal charge of -1). This Chapter is a compilation of the main results of the Candidate's study [148], which aims to provide answers to the above questions.

11.2 Studied Systems and Computational Details

In this work, the previously characterized molecular electrides were studied [375], namely, Li@calix[4]pyrrole [375, 413, 425], Na@calix[4]-pyrrole [375, 413, 425], $TCNQLi_2$ [375, 426], $TCNQNa_2$ [375, 426], $TCNENa_3$ [427], $TCNENa_4$(II) [427], Mg_2EP [428], Mg_2C_{60} [429], and $e@C_{60}F_{60}$ [375, 430]. Their structures are shown in Figure 11.1. We study the topology of the electronic density of those systems using the QTAIM analysis (see Section 6.2). All these studied systems contain at least one NNACP, a significant concentration of density in the region of NNACP observable with ELF [419, 420], negative values of the Laplacian of the electron density [238], and huge values of nonlinear optical properties [375, 414, 428, 429]. Thus, we compute the values of electron density, $\rho(\mathbf{r}_{NNACP})$, and the Laplacian of the electron density, $\nabla^2\rho(\mathbf{r}_{NNACP})$ (both measured at the position of NNACP), the average number of electrons in the NNA basin N_{NNA}, and the localization index λ_{NNA}, which quantifies the average number of localized electrons in the NNA basin. With the increased amount of localized electride electrons, one can expect $\rho(\mathbf{r}_{NNACP})$ to rise, $\nabla^2\rho(\mathbf{r}_{NNACP})$ to decrease and be more negative, and N_{NNA} and λ_{NNA} to increase. Additionally, for each NNA basin, we compute its effective oxidation state (EOS) using the method described in Section 6.3.

To compute the EOS, we must define real-space fragments among which the electrons are distributed. We test two partitions: EOS[1], which considers only two fragments: the NNA basin and the rest of the molecule, and EOS[2], which divides the space into several fragments: the NNA basins, one fragment for each metallic atom, and the rest of the molecule. The only exception is the $e@C_{60}F_{60}$ system, in which only type 1 fragmentation can be applied because there are no metallic center.

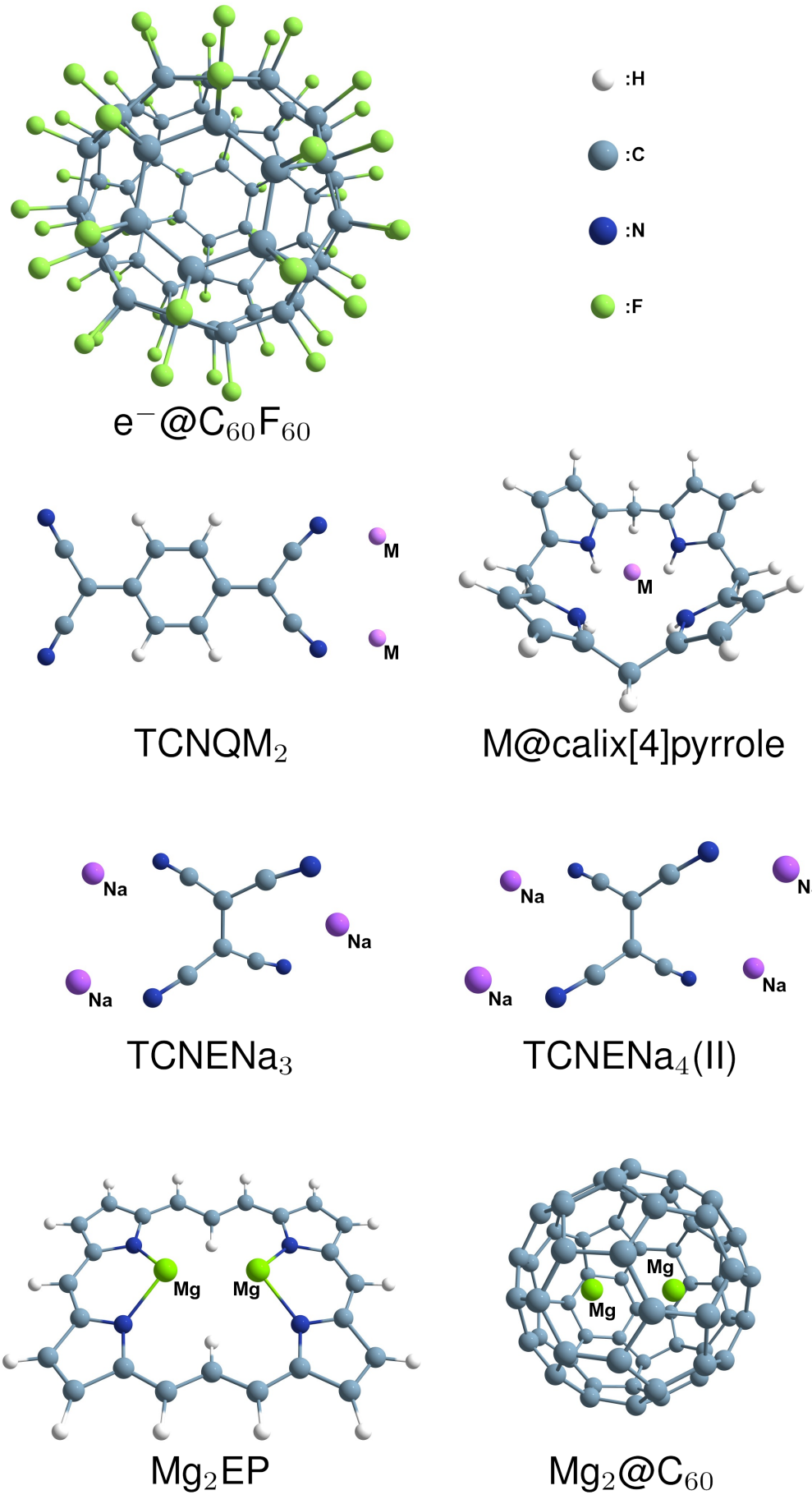

Figure 11.1: From top-left to bottom-right, the molecular structures of e@$C_{60}F_{60}$, $TCNQM_2$, M@calix[4]pyrrole, $TCNENa_3$, $TCNENa_4$(II), Mg_2EP, and Mg_2@C_{60}, where M atom denotes either M = Li or M = Na.

All structures were optimized using CAM-B3LYP/ma-Def2TZVP [193, 431]. For such optimized geometries, single point calculations using different methods, including DFAs such as CAM-B3LYP [193], B3LYP [325, 326], M06-2X [330], and MN15 [75], as well as wavefunction-based ab-initio HF, MP2 [210], CCSD [432] and CCSD(T) [297] methods. For all systems, except $e@C_{60}F_{60}$, several basis sets were tested, namely (aug-)cc-pVDZ and (aug-)cc-pVTZ [298–305], ma-Def2TZVP [431] and ma-Def2QZVP [431]. In the case of $e@C_{60}F_{60}$, the medium-sized basis sets, 6-31G [433], 6-31G(d) [433, 434], 6-31+G [433, 435], and ma-TZVP [431] with the additional diffuse functions placed at the center of the cage, were utilized. The set of additional functions, labeled as DF, consisted of four uncontracted diffuse functions of the *s*- and *p*-types with exponents $1.68714478 \cdot 10^{-n}$ (n = 1-4) [375, 430]. It was previously established that additional DF provide a good description of the NNA localized in the center of the cage [375, 430]. All HF, MP2, and KS-DFT calculations were performed with Gaussian16 (Rev. B01) [436], whereas CCSD and CCSD(T) were done using CFOUR 2.1. [323].

Besides performing the EOS analysis for the NNA basins, the electron distribution functions (EDFs), from the formulation of Cancés *et al*. [437], were also computed. EDFs define the probability of finding exactly zero, one, and two electrons in the particular region of space or atomic basin. The probability of finding zero electrons in the region A is given by:

$$P_0^A = \prod_{i=1}^{N} \left(1 - \lambda_i^A\right), \tag{11.1}$$

and, if $P_0^A \neq 0$, the probability of finding exactly n electrons in the region A, P_n^A, is given by

$$P_n^A = {\sum_{i_1,\ldots,i_n}^{N}}' \prod_{j=1}^{n} \left(\frac{\lambda_{i_j}^A}{1-\lambda_{i_j}^A}\right) P_0^A \quad \forall n > 0, \tag{11.2}$$

where the prime in the summation indicates that we are excluding terms with repeated indices, and λ_i^A are the N non-zero eigenvalues of the atomic overlap matrix for the basin A. In this work, we only consider the basin of the NNA and thus we drop the notation indicating the region ($P_i \equiv P_i^{\mathrm{NNA}}$ and $\lambda_i \equiv \lambda_i^{\mathrm{NNA}}$). EDFs were introduced in chemistry by Daudel [438, 439] and have been extensively used by Martín Pendás and co-workers [440–442] as an elegant tool to characterize chemical bonds [443]. Using the EDFs, one can easily retrieve the average number of electrons in the NNA basin:

$$N_{\mathrm{NNA}} = \sum_i i P_i. \tag{11.3}$$

In the present case, the probability of finding more than two electrons in the NNA basin is expected to be negligible, hence, $N_{\mathrm{NNA}} \approx P_1 + 2P_2$. Despite the MP2 probabilities cannot be computed from Eqs. 11.1 and 11.2, neglecting $P_i \forall i > 2$, and employing the approximate localization index, we can compute approximate MP2 P_0, P_1, and P_2 from the following set of equations:

$$\begin{aligned}
N_{\mathrm{NNA}} &\approx P_1 + 2P_2, && (11.4)\\
\lambda_{\mathrm{NNA}} &\approx N_{\mathrm{NNA}} - P_0 (N_{\mathrm{NNA}})^2 - P_1 (1 - N_{\mathrm{NNA}})^2 - P_2 (2 - N_{\mathrm{NNA}})^2, && (11.5)\\
1 &\approx P_0 + P_1 + P_2, && (11.6)
\end{aligned}$$

which give rise to these approximate probabilities:

$$\begin{aligned}
P_0 &\approx 1 - N_{\mathrm{NNA}} + \frac{1}{2}\left(N_{\mathrm{NNA}}^2 - \lambda_{\mathrm{NNA}}\right), && (11.7)\\
P_1 &\approx \lambda_{\mathrm{NNA}} + N_{\mathrm{NNA}} (1 - N_{\mathrm{NNA}}), && (11.8)\\
P_2 &\approx \frac{1}{2}\left(N_{\mathrm{NNA}}^2 - \lambda_{\mathrm{NNA}}\right). && (11.9)
\end{aligned}$$

Obviously, if $P_i \approx 0 \forall i > 2$ and accurate N_{NNA} and λ_{NNA} are used, the latter expressions provide excellent estimates for P_0, P_1, and P_2, regardless of the wavefunction approximation. We will consider that the molecule is at least a one-electron molecular electride if the probability of having one or more electrons, $1 - P_0$, is higher than P_0.

In the case of post-HF methods, relaxed electronic densities [444] were used in subsequent QTAIM and EOS analyzes. Furthermore, the Muller approximation has been adapted [445–451] to approximate the necessary two-electron densities for computing the (de)localization indices. The topology of the electron density within the QTAIM formalism was done using the AIMAll package (ver. 19.10.12) [452] and the EOS analysis was carried out with APOST3d (dev. ver. 2.0) [453].

11.3 Results and Discussion

In this section, results of the QTAIM, EOS and EDF analyses are presented. The discussion is done separately for each pair of molecular electrides that share a structural motif, that is, $TCNQLi_2$ and $TCNQNa_2$, Li@calix[4]pyrrole and Na@calix[4]pyrrole, $TCNENa_3$ and $TCNENa_4$(II), Mg_2EP and $Mg_2@C_{60}$, and lastly, $e@C_{60}F_{60}$.

11.3.1 $TCNQLi_2$ and $TCNQNa_2$

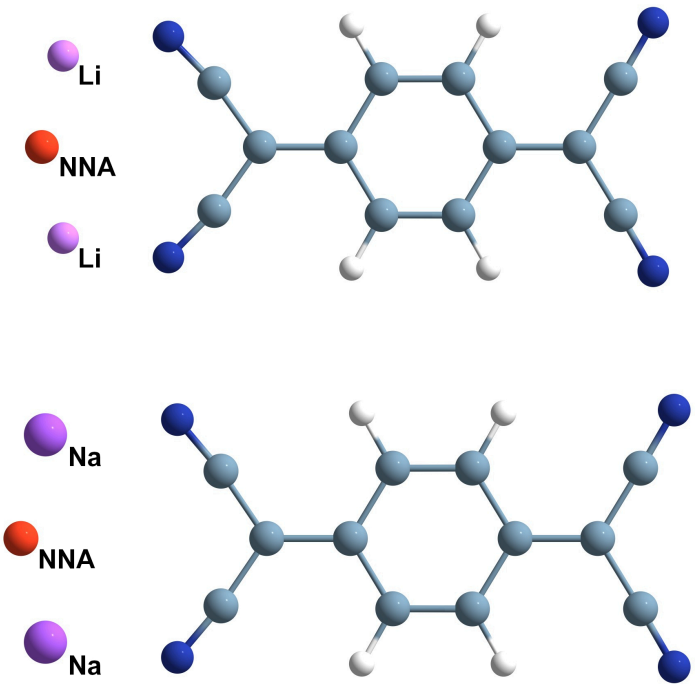

Figure 11.2: Geometries of $TCNQLi_2$ (top) and $TCNQNa_2$ (bottom) optimized at the CAM-B3LYP/ma-Def2TZVP level of theory. The NNA position is indicated with a red ball.

Results are compiled in Tables 11.1 and 11.2. Both $TCNQLi_2$ and $TCNQNa_2$ are molecules with a triplet ground state $^3B_1$. These systems contain one NNACP in between the alkaline atoms (see Figures 11.2), which is detectable with most of the tested DFT and wavefunction-based methods. The exceptions are found for $TCNQLi_2$ and the densities obtained with B3LYP/ma-Def2TZVP, CAM-B3LYP/ma-Def2TZVP, M06-2X/ma-Def2TZVP, MN15/ma-Def2TZVP and MN15/ma-Def2QZVP, and for $TCNQNa_2$ and the

densities obtained with MN15/ma-Def2TZVP and MN15/ma-Def2QZVP. In all of these cases, no NNACP was found in the topology of the density.

Nevertheless, for all methods that correctly report only one NNACP, the values of $\rho(\mathbf{r}_{NNACP})$ and $\nabla^2\rho(\mathbf{r}_{NNACP})$ are very similar. The population of the NNA basin is between 0.5 and 0.7 electrons for $TCNQLi_2$, whereas for $TCNQNa_2$ it is much lower, between 0.2 and 0.5. In general, it can be observed that the values of N_{NNA} and λ_{NNA} for the NNA basin increase with the percentage of HF exchange used in the DFT or wavefunction-based method (20% of the exact exchange), CAM-B3LYP (19 - 65% of the exact exchange), M06-2X (54% of the exact exchange) and finally HF and MP2, which are fully based on the exact exchange. HF provides values of N_{NNA} and λ_{NNA} very close to those obtained with MP2, especially in $TCNQNa_2$, which may suggest that for this system, the average NNA basin populations are underestimated by DFAs. Moreover, in the case of $TCNQNa_2$, both N_{NNA} and λ_{NNA} seem to be more basis set dependent than for $TCNQLi_2$. As compared to $TCNQLi_2$, electrons in the NNA basin of $TCNQNa_2$ are less localized. For the first molecule, one observes that more than 50% of the electrons are localized in the NNA basin and, in the latter case, this number is about 25% for the studied DFAs and below 50% for HF and MP2.

The outcome of the EOS analysis for $TCNQLi_2$ and $TCNQNa_2$ is entirely different. In $TCNQLi_2$, the formal oxidation state of the NNA basin is -1, in spite of the types of fragments used. The values of RI of these predictions increase with the average number of electrons in the NNA basin. Moreover, they are larger for the second fragmentation scheme, namely, EOS[2], which distributes the electrons in the electrides among the NNA, the two Li atoms, and the rest of the molecule. The picture of a negative oxidation state of the NNA basin is further confirmed by the EDFs, for which the highest probability is found for P_1, that is, finding exactly one electron in the basin. One can also find a good correspondence between the values of P_1 and the RI value predicted by the first fragmentation (EOS[1]). In $TCNQNa_2$, the assigned oxidation state is strongly shifted towards zero and depends on the type of fragmentation used. For the fragmentation scheme 1, the oxidation state of the NNA basin is zero, which means that formally one cannot classify this NNA basin as an anion. For the fragmentation scheme 2, for the methods but HF, the oxidation state of NNA is also zero. This observation goes in line with the EDF analysis, where the probabilities of finding exactly zero electrons in the NNA basin are the largest ones.

In summary, according to the EOS and EDF criteria, one can classify $TCNQLi_2$ as a true one-electron molecular electride, while $TCNQNa_2$ represents a borderline situation between a zero-electron and a one-electron molecular electride.

Table 11.1: Results of the QTAIM partitioning for $TCNQLi_2$. The third and fourth columns report values of the density, $\rho(\mathbf{r}_{\mathrm{NNACP}})$, and its Laplacian, $\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$, at the positions of the NNA. The fifth and sixth columns give the population, N_{NNA}, and localization index, λ_{NNA}, for the NNA basin. Starting on the seventh column, results of the EOS and EDF analyses of the NNA basin are given. EOS[1] and EOS[2] correspond to the oxidation state obtained with type 1 and 2 fragmentations, respectively (see text). The last three columns report the probabilities (in percentage) of finding 0, 1, and 2 electrons in the NNA basin.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	no NNAs										
	ma-Def2QZVP	$9.6{\cdot}10^{-3}$	$-8.2{\cdot}10^{-3}$	0.540	0.277	-1	55	-1	83	46.9	52.6	0.6
	cc-pVTZ	$9.6{\cdot}10^{-3}$	$-8.3{\cdot}10^{-3}$	0.563	0.300	-1	60	-1	87	44.7	54.7	0.6
	aug-cc-pVTZ	$9.6{\cdot}10^{-3}$	$-8.3{\cdot}10^{-3}$	0.557	0.295	-1	59	-1	86	45.2	54.2	0.6
B3LYP	ma-Def2TZVP	no NNAs										
	ma-Def2QZVP	$9.4{\cdot}10^{-3}$	$-7.9{\cdot}10^{-3}$	0.505	0.242	0	52	-1	79	50.3	49.2	0.5
	cc-pVTZ	$9.4{\cdot}10^{-3}$	$-8.1{\cdot}10^{-3}$	0.527	0.262	-1	52	-1	82	48.3	51.2	0.6
	aug-cc-pVTZ	$9.4{\cdot}10^{-3}$	$-8.1{\cdot}10^{-3}$	0.524	0.259	-1	52	-1	81	48.5	50.9	0.6
M06-2X	ma-Def2TZVP[a]	$7.8{\cdot}10^{-3}$	$-4.5{\cdot}10^{-3}$	0.424	0.084	0	68	-1	67	58.3	41.2	0.5
	ma-Def2QZVP	$9.3{\cdot}10^{-3}$	$-7.9{\cdot}10^{-3}$	0.607	0.344	-1	67	-1	94	40.6	58.4	1.0
	cc-pVTZ	$9.4{\cdot}10^{-3}$	$-8.3{\cdot}10^{-3}$	0.609	0.345	-1	67	-1	94	40.6	58.4	1.1
	aug-cc-pVTZ	$9.5{\cdot}10^{-3}$	$-8.3{\cdot}10^{-3}$	0.609	0.346	-1	68	-1	94	40.5	58.5	1.0
MN15	ma-Def2TZVP	no NNAs										
	ma-Def2QZVP	found 8 NNAs and Poincare-Hopf not satisfied										
HF	ma-Def2TZVP	$7.8{\cdot}10^{-3}$	$-4.7{\cdot}10^{-3}$	0.593	0.337	-1	66	-1	90	41.5	57.9	0.6
	ma-Def2QZVP	$9.6{\cdot}10^{-3}$	$-8.5{\cdot}10^{-3}$	0.726	0.504	-1	92	-1	100	28.7	70.5	0.9
	cc-pVTZ	$9.7{\cdot}10^{-3}$	$-8.7{\cdot}10^{-3}$	0.719	0.494	-1	91	-1	100	29.4	69.8	0.8
	aug-cc-pVTZ	$9.7{\cdot}10^{-3}$	$-8.7{\cdot}10^{-3}$	0.724	0.502	-1	92	-1	100	28.8	70.3	0.9
MP2	ma-Def2TZVP	$8.0{\cdot}10^{-3}$	$-4.8{\cdot}10^{-3}$	0.550	0.291					45.6	53.9	0.6
	aug-cc-pVTZ	$1.0{\cdot}10^{-2}$	$-9.1{\cdot}10^{-3}$	0.700	0.470					31.0	68.0	1.0

[a] Two closely separated NNAs are found for this basis set (N_{NNA} and λ_{NNA} correspond to the sum of both NNAs).

Table 11.2: Results of the QTAIM partitioning, EOS, and EDF analyses for $TCNQNa_2$. For a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	$6.9{\cdot}10^{-3}$	$-4.4{\cdot}10^{-3}$	0.213	0.044	0	100	0	66	78.9	21.0	0.0
	aug-cc-pVTZ	$7.1{\cdot}10^{-3}$	$-4.9{\cdot}10^{-3}$	0.265	0.068	0	98	0	58	73.7	26.2	0.1
B3LYP	ma-Def2TZVP	$6.6{\cdot}10^{-3}$	$-4.1{\cdot}10^{-3}$	0.170	0.028	0	100	0	72	83.2	16.8	0.0
	aug-cc-pVTZ	$6.8{\cdot}10^{-3}$	$-4.6{\cdot}10^{-3}$	0.217	0.045	0	100	0	65	78.6	21.3	0.0
M06-2X	ma-Def2TZVP	$6.7{\cdot}10^{-3}$	$-4.0{\cdot}10^{-3}$	0.236	0.053	0	100	0	62	76.7	23.2	0.1
	aug-cc-pVTZ	$6.9{\cdot}10^{-3}$	$-4.5{\cdot}10^{-3}$	0.241	0.054	0	100	0	61	76.3	23.6	0.1
MN15	ma-Def2TZVP	no NNAs										
	aug-cc-pVTZ	found 11 NNAs and Poincare-Hopf not satisfied										
HF	ma-Def2TZVP	$7.1{\cdot}10^{-3}$	$-5.2{\cdot}10^{-3}$	0.443	0.191	0	63	-1	68	56.2	43.7	0.2
	aug-cc-pVTZ	$7.1{\cdot}10^{-3}$	$-5.4{\cdot}10^{-3}$	0.495	0.238	0	52	-1	75	51.0	48.8	0.2
MP2	ma-Def2TZVP	$7.3{\cdot}10^{-3}$	$-5.3{\cdot}10^{-3}$	0.418	0.171					58.4	41.4	0.2
	aug-cc-pVTZ	$7.4{\cdot}10^{-3}$	$-5.6{\cdot}10^{-3}$	0.465	0.212					53.7	46.0	0.2

11.3.2 Li@calix[4]pyrrole and Na@calix[4]pyrrole

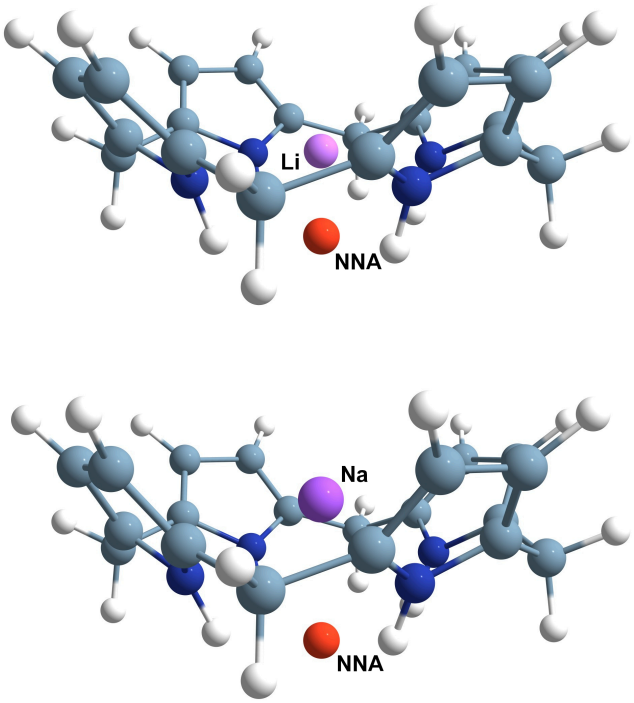

Figure 11.3: Geometries of Li@calix[4]pyrrole (top) and Na@calix[4]pyrrole (bottom) optimized at the CAM-B3LYP/ma-Def2TZVP level of theory. The NNA position is indicated with a red ball.

Results for Li@calix[4]pyrrole and Na@calix[4]pyrrole are compiled in Tables 11.3 and 11.4. These chemical systems are open-shell doublet systems with the $^2A_1$ electronic ground state. A NNACP is found in the center of the pyrrole ring, on the same vertical axis as Li and Na atoms (see Figure 11.3). The presence of NNACP is correctly described by all the methods tested except the MN15 functional. Although the values of density and the Laplacian density are quite similar across all methods and basis sets, larger differences are observed for the averaged quantities. Again, DFAs yield much lower values of N_{NNA} and λ_{NNA} than MP2, while HF drastically overestimates these properties. The best correspondence with the reference MP2 results is found for the CAM-B3LYP functional. Regarding the most likely oxidation state of the electride anion, for Li@calix[4]pyrrole, a situation similar to $TCNQNa_2$ can be observed. That is, the non-zero oxidation state is predicted only by HF and mostly for the second fragmentation state. Although the corresponding RI values of these predictions are much larger for $TCNQNa_2$ (especially for the second fragmentation scheme), the probabilities of founding exactly zero electrons, P_0, are only slightly larger. In the case of Na@calix[4]pyrrole, all methods and fragmentation schemes predict a zero oxidation state and these predictions are fully supported by RI = 100% (except for HF, for which RI = 55 - 64%). The zero oxidation state is supported by the large values of P_0, which are larger than 80% for all methods except HF. Moreover, the probabilities of finding exactly zero electrons in the NNA basin are the highest for Na@calix[4]pyrrole among all systems studied. Therefore, based on the EOS and EDF results, one could assign both Li@calix[4]pyrrole and Na@calix[4]pyrrole as low-electron-number electrides, with the former one being more a borderline case (similar to $TCNQNa_2$).

Table 11.3: Results of the QTAIM partitioning, EOS, and EDF analyses for Li@calix[4]pyrrole. For a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	$7.6{\cdot}10^{-3}$	$-4.0{\cdot}10^{-3}$	0.291	0.079	0	94	0	91	71.5	28.3	0.2
	aug-cc-pVTZ	$8.0{\cdot}10^{-3}$	$-5.2{\cdot}10^{-3}$	0.269	0.067	0	98	0	95	73.7	26.2	0.1
B3LYP	ma-Def2TZVP	$6.9{\cdot}10^{-3}$	$-3.2{\cdot}10^{-3}$	0.215	0.043	0	100	0	100	78.9	21.0	0.1
	aug-cc-pVTZ	$7.2{\cdot}10^{-3}$	$-4.3{\cdot}10^{-3}$	0.193	0.035	0	100	0	100	81.0	18.9	0.1
M06-2X	ma-Def2TZVP	$7.0{\cdot}10^{-3}$	$-2.9{\cdot}10^{-3}$	0.195	0.035	0	100	0	100	81.0	19.0	0.1
	aug-cc-pVTZ	$7.3{\cdot}10^{-3}$	$-4.1{\cdot}10^{-3}$	0.177	0.028	0	100	0	100	82.8	17.1	0.1
MN15	ma-Def2TZVP					no NNAs						
	aug-cc-pVTZ					no NNAs						
HF	ma-Def2TZVP	$7.6{\cdot}10^{-3}$	$-5.2{\cdot}10^{-3}$	0.518	0.250	-0.5	100	-1	52	49.3	50.0	0.7
	aug-cc-pVTZ	$7.8{\cdot}10^{-3}$	$-5.8{\cdot}10^{-3}$	0.511	0.245	0	51	-1	51	49.8	49.5	0.6
MP2	ma-Def2TZVP	$7.7{\cdot}10^{-3}$	$-4.5{\cdot}10^{-3}$	0.321	0.100					68.1	31.8	0.2
	aug-cc-pVTZ	$8.2{\cdot}10^{-3}$	$-5.6{\cdot}10^{-3}$	0.29	0.08					71.2	28.7	0.1

Table 11.4: Results of the QTAIM partitioning, EOS, and EDF analyses for Na@calix[4]pyrrole. For a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	$5.9{\cdot}10^{-3}$	$-2.6{\cdot}10^{-3}$	0.169	0.027	0	100	0	100	83.3	16.7	0.1
	aug-cc-pVTZ	$6.0{\cdot}10^{-3}$	$-2.6{\cdot}10^{-3}$	0.167	0.026	0	100	0	100	83.5	16.4	0.1
B3LYP	ma-Def2TZVP	$5.3{\cdot}10^{-3}$	$-2.0{\cdot}10^{-3}$	0.094	0.008	0	100	0	100	90.8	9.2	0.0
	aug-cc-pVTZ	$5.3{\cdot}10^{-3}$	$-2.0{\cdot}10^{-3}$	0.086	0.007	0	100	0	100	91.5	8.5	0.0
M06-2X	ma-Def2TZVP	$5.7{\cdot}10^{-3}$	$-2.2{\cdot}10^{-3}$	0.100	0.009	0	100	0	100	90.3	9.7	0.0
	aug-cc-pVTZ	$5.5{\cdot}10^{-3}$	$-1.8{\cdot}10^{-3}$	0.062	0.004	0	100	0	100	94.0	6.0	0.0
MN15	ma-Def2TZVP					no NNAs						
	aug-cc-pVTZ					no NNAs						
HF	ma-Def2TZVP	$5.8{\cdot}10^{-3}$	$-3.1{\cdot}10^{-3}$	0.448	0.191	0	63	0	62	56.0	43.7	0.2
	aug-cc-pVTZ	$6.0{\cdot}10^{-3}$	$-3.4{\cdot}10^{-3}$	0.484	0.224	0	55	0	55	52.4	47.3	0.2
MP2	ma-Def2TZVP	$6.2{\cdot}10^{-3}$	$-3.0{\cdot}10^{-3}$	0.202	0.041					79.8	20.2	0.0
	aug-cc-pVTZ	$6.4{\cdot}10^{-3}$	$-3.0{\cdot}10^{-3}$	0.195	0.039					80.5	19.6	0.0

11.3.3 $TCNENa_3$ and $TCNENa_4$(II)

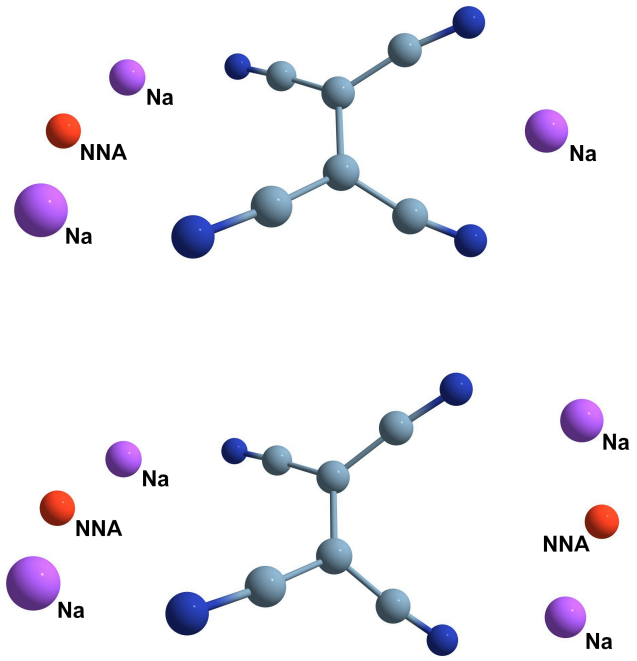

Figure 11.4: Geometries of $TCNENa_3$ (top) and $TCNENa_4$(II) (bottom) optimized at the CAM-B3LYP/ma-Def2TZVP level of theory. The NNA position is indicated with a red ball.

Results for $TCNENa_3$ and $TCNENa_4$(II) are compiled in Tables 11.5 and 11.6. These chemical species are open-shell systems with the ^2A and $^3B_3$ electronic ground states, respectively. Both $TCNENa_3$ and $TCNENa_4$(II) were recently classified as the molecular electrides based on the topological descriptions and reportedly large values of NLOPs. However, the latter ones were presented only for $TCNENa_3$ and only for the static first hyperpolarizability [427]. The first hyperpolarizability of $TCNENa_4$(II) was not included in the discussion, because it was zero due to the centrosymmetry of the system. For the sake of completeness, using CAM-B3LYP and various basis sets, in our work we additionally computed all static NLOPs of these two molecules (up to the second hyperpolarizability). Both systems show remarkably large values of nonlinear $\beta(0;0,0)$ and $\gamma(0;0,0,0)$ and of the same magnitude as in other molecular electrides [414].

Similarly to $TCNQNa_2$, $TCNENa_3$ and $TCNENa_4$(II) posses one and two Na-Na pairs, respectively, and between each of the pairs, the NNACP is located (see Figure 11.4). Due to these structural and topological similarities, the results of the QTAIM, EOS, and EDF analyses on the occupation of the NNA basin are equivalent to those of $TCNQNa_2$. Therefore these two systems can also be considered borderline situations between zero- and one-electron molecular electrides.

Table 11.5: Results of the QTAIM partitioning, EOS, and EDF analyses for $TCNENa_3$. For a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	$7.0{\cdot}10^{-3}$	$-4.4{\cdot}10^{-3}$	0.204	0.040	0	100	0	68	79.8	20.2	0.0
	aug-cc-pVTZ	$7.2{\cdot}10^{-3}$	$-4.9{\cdot}10^{-3}$	0.268	0.069	0	97	0	58	73.5	26.5	0.1
B3LYP	ma-Def2TZVP	$6.7{\cdot}10^{-3}$	$-4.1{\cdot}10^{-3}$	0.162	0.025	0	100	0	74	84.0	16.0	0.0
	aug-cc-pVTZ	$6.9{\cdot}10^{-3}$	$-4.5{\cdot}10^{-3}$	0.218	0.046	0	100	0	66	78.5	21.5	0.0
M06-2X	ma-Def2TZVP	$6.8{\cdot}10^{-3}$	$-4.0{\cdot}10^{-3}$	0.211	0.040	0	100	0	67	79.3	20.6	0.2
	aug-cc-pVTZ	$7.1{\cdot}10^{-3}$	$-4.6{\cdot}10^{-3}$	0.228	0.047	0	100	0	64	77.6	22.2	0.2
MN15	ma-Def2TZVP	no NNAs										
	aug-cc-pVTZ	found 5 NNAs and Poincare-Hopf not satisfied										
HF	ma-Def2TZVP	$7.1{\cdot}10^{-3}$	$-5.2{\cdot}10^{-3}$	0.438	0.187	0	64	-1	66	56.6	43.3	0.1
	aug-cc-pVTZ	$7.2{\cdot}10^{-3}$	$-5.4{\cdot}10^{-3}$	0.495	0.238	0	52	-1	75	51.0	48.8	0.2
MP2	ma-Def2TZVP	$7.0{\cdot}10^{-3}$	$-5.0{\cdot}10^{-3}$	0.410	0.165					59.1	40.7	0.2
	aug-cc-pVTZ	$7.1{\cdot}10^{-3}$	$-5.3{\cdot}10^{-3}$	0.464	0.211					53.8	46.0	0.2

Table 11.6: Results of the QTAIM partitioning, EOS, and EDF analyses for $TCNENa_4$(II). For the a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	$7.0{\cdot}10^{-3}$	$-4.3{\cdot}10^{-3}$	0.202	0.040	0	100	0	68	80.0	20.0	0.0
	aug-cc-pVTZ	$7.2{\cdot}10^{-3}$	$-4.9{\cdot}10^{-3}$	0.266	0.068	0	98	0	59	73.6	26.3	0.1
B3LYP	ma-Def2TZVP	$6.7{\cdot}10^{-3}$	$-4.0{\cdot}10^{-3}$	0.159	0.024	0	100	0	74	84.3	15.7	0.0
	aug-cc-pVTZ	$6.9{\cdot}10^{-3}$	$-4.5{\cdot}10^{-3}$	0.215	0.045	0	100	0	66	78.7	21.2	0.0
M06-2X	ma-Def2TZVP	$6.8{\cdot}10^{-3}$	$-4.0{\cdot}10^{-3}$	0.211	0.041	0	100	0	67	79.2	20.6	0.2
	aug-cc-pVTZ	$7.1{\cdot}10^{-3}$	$-4.6{\cdot}10^{-3}$	0.226	0.046	0	100	0	65	77.8	22.0	0.2
MN15	ma-Def2TZVP					no NNAs						
	aug-cc-pVTZ					no NNAs						
HF	ma-Def2TZVP	$7.1{\cdot}10^{-3}$	$-5.2{\cdot}10^{-3}$	0.439	0.188	0	63	-1	67	56.5	43.4	0.1
	aug-cc-pVTZ	$7.2{\cdot}10^{-3}$	$-5.4{\cdot}10^{-3}$	0.495	0.239	0	52	-1	75	51.0	48.9	0.2
MP2	ma-Def2TZVP	$7.0{\cdot}10^{-3}$	$-5.0{\cdot}10^{-3}$	0.411	0.166					59.1	40.8	0.2
	aug-cc-pVTZ	$7.1{\cdot}10^{-3}$	$-5.2{\cdot}10^{-3}$	0.465	0.211					53.8	46.0	0.2

11.3.4 Mg_2EP and Mg_2C_{60}

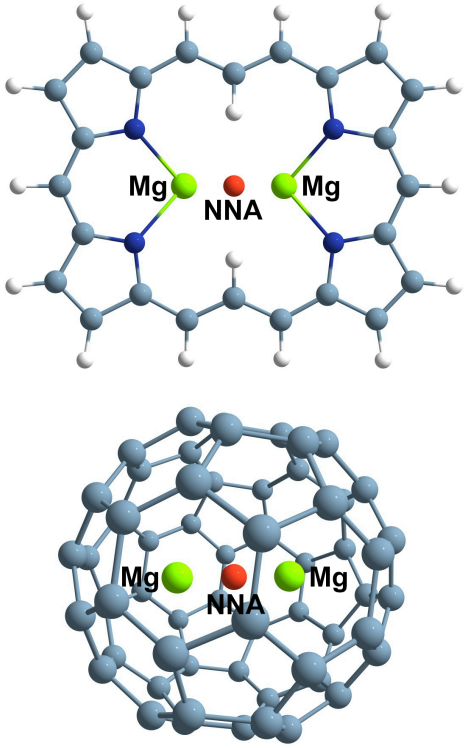

Figure 11.5: Geometries of Mg_2EP (top) and Mg_2C_{60} (bottom) optimized at the CAM-B3LYP/ma-Def2TZVP level of theory. The NNA position is indicated with a red ball.

In contrast to previously discussed systems, Mg_2EP and Mg_2C_{60} do not contain alkaline metals. Instead, they hold a Mg_2 motif which is held by an external structure. Moreover, the electronic ground state of both systems is $^1A_1$. These molecules were classified as electrides by Chattaraj and co-workers [428, 429], and present a NNACP in between Mg_2 (see Figure 11.5). For both systems, all tested DFAs but MN15 yield similar values of the topological descriptors. The average number of electrons in the NNA basin is in the range 0.68 - 0.75, of which 29 - 32% are localized there. On the contrary, for the densities obtained with MN15, N_{NNA} is lower and in the range 0.50 - 0.55, of which only 21 - 23% are localized. Again, HF yields NNA basins holding the largest number of electrons (N_{NNA} = 0.97 - 1.03) and are the most localized (42 - 45%). MP2 yields results which are in between those of HF and DFAs. For all methods but MN15, the EOS analysis assigns two different oxidation states, 0 for EOS[1] and -2 for EOS[2]. Due to the singlet spin symmetry of the molecule, the EOS cannot assign an oxidation state of -1 to the NNA basin. For DFAs, the reliability indices of the assigned oxidation states lie in the range 80 - 87% for EOS[1] and 56 - 61% for EOS[2]. In contrast, for HF the values of RI are lower for EOS[1] (RI = 53 - 60%) but larger for EOS[2] (RI = 75 - 79%). In the case of MN15, despite the NNACP being detected, the assigned oxidation state is zero. The results of the EDF analysis show important differences with the previously discussed molecular electrides. The probability of having exactly two electrons in the NNA basin is not negligible: P_2 is between 11 - 14% for DFAs, 21 - 24% for HF, and 13 - 20% for MP2. On top of that, P_0 and P_1 have very similar values, except for HF, for which P_1 is almost twice as large as P_0. The situation in which we have only one electron is the most likely scenario if one has to choose among the three possibilities, although, no possibility can be ruled out. Therefore, based on the EOS and complementary EDF results one can conclude that Mg_2EP and $Mg_2@C_{60}$ are at least one-electron molecular electrides.

Table 11.7: Results of the QTAIM partitioning, EOS, and EDF analyses for Mg_2EP. For a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	$3.2{\cdot}10^{-2}$	$-3.8{\cdot}10^{-2}$	0.722	0.229	0	82	-2	59	42.1	44.7	12.8
	aug-cc-pVTZ	$3.3{\cdot}10^{-2}$	$-3.9{\cdot}10^{-2}$	0.746	0.244	0	80	-2	60	40.6	45.3	13.6
B3LYP	ma-Def2TZVP	$3.2{\cdot}10^{-2}$	$-3.6{\cdot}10^{-2}$	0.677	0.201	0	87	-2	56	44.9	43.4	11.3
	aug-cc-pVTZ	$3.2{\cdot}10^{-2}$	$-3.8{\cdot}10^{-2}$	0.712	0.222	0	83	-2	58	42.7	44.4	12.4
M06-2X	ma-Def2TZVP	$3.1{\cdot}10^{-2}$	$-3.6{\cdot}10^{-2}$	0.705	0.219	0	84	-2	58	43.1	44.3	12.2
	aug-cc-pVTZ	$3.1{\cdot}10^{-2}$	$-3.7{\cdot}10^{-2}$	0.701	0.216	0	84	-2	58	43.4	44.1	12.1
MN15	ma-Def2TZVP	$2.9{\cdot}10^{-2}$	$-3.3{\cdot}10^{-2}$	0.517	0.116	0	100	0	54	55.9	37.1	6.7
	aug-cc-pVTZ	$3.0{\cdot}10^{-2}$	$-3.7{\cdot}10^{-2}$	0.539	0.125	0	100	0	52	54.4	38.1	7.3
HF	ma-Def2TZVP	$3.4{\cdot}10^{-2}$	$-4.3{\cdot}10^{-2}$	1.018	0.460	0	54	-2	79	25.8	48.8	24.3
	aug-cc-pVTZ	$3.4{\cdot}10^{-2}$	$-4.3{\cdot}10^{-2}$	1.030	0.470	0	53	-2	79	25.3	48.7	24.8
MP2	ma-Def2TZVP	$3.0{\cdot}10^{-2}$	$-3.5{\cdot}10^{-2}$	0.835	0.332					34.8	46.9	18.3
	aug-cc-pVTZ	$3.1{\cdot}10^{-2}$	$-3.6{\cdot}10^{-2}$	0.876	0.366					32.5	47.4	20.1

Table 11.8: Results of the QTAIM partitioning, EOS, and EDF analyses for $Mg_2@C_{60}$. For a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	EOS[2]	RI	P_0	P_1	P_2
CAM-B3LYP	ma-Def2TZVP	$4.0{\cdot}10^{-2}$	$-5.0{\cdot}10^{-2}$	0.699	0.211	0	85	-2	58	43.8	43.9	11.9
	aug-cc-pVTZ	$4.0{\cdot}10^{-2}$	$-5.5{\cdot}10^{-2}$	0.738	0.234	0	82	-2	61	41.4	44.9	13.1
B3LYP	ma-Def2TZVP	$3.9{\cdot}10^{-2}$	$-5.0{\cdot}10^{-2}$	0.685	0.202	0	87	-2	57	44.8	43.4	11.4
	aug-cc-pVTZ	$4.0{\cdot}10^{-2}$	$-5.4{\cdot}10^{-2}$	0.732	0.229	0	82	-2	60	41.9	44.7	12.9
M06-2X	ma-Def2TZVP	$3.7{\cdot}10^{-2}$	$-4.6{\cdot}10^{-2}$	0.687	0.203	0	86	-2	58	44.7	43.5	11.4
	aug-cc-pVTZ	$3.8{\cdot}10^{-2}$	$-5.2{\cdot}10^{-2}$	0.704	0.212	0	85	-2	59	43.7	43.9	12.0
MN15	ma-Def2TZVP	$3.6{\cdot}10^{-2}$	$-4.3{\cdot}10^{-2}$	0.491	0.104	0	100	0	51	58.1	35.8	6.0
	aug-cc-pVTZ	$3.7{\cdot}10^{-2}$	$-5.1{\cdot}10^{-2}$	0.550	0.129	0	99	0	50	53.8	38.5	7.5
HF	ma-Def2TZVP	$4.2{\cdot}10^{-2}$	$-5.9{\cdot}10^{-2}$	0.971	0.404	0	60	-2	75	28.6	48.3	21.9
	aug-cc-pVTZ	$4.2{\cdot}10^{-2}$	$-6.2{\cdot}10^{-2}$	0.984	0.412	0	59	-2	76	28.0	48.3	22.3
MP2	ma-Def2TZVP	$3.3{\cdot}10^{-2}$	$-3.7{\cdot}10^{-2}$	0.691	0.215					44.0	42.9	13.1

11.3.5 $e@C_{60}F_{60}$

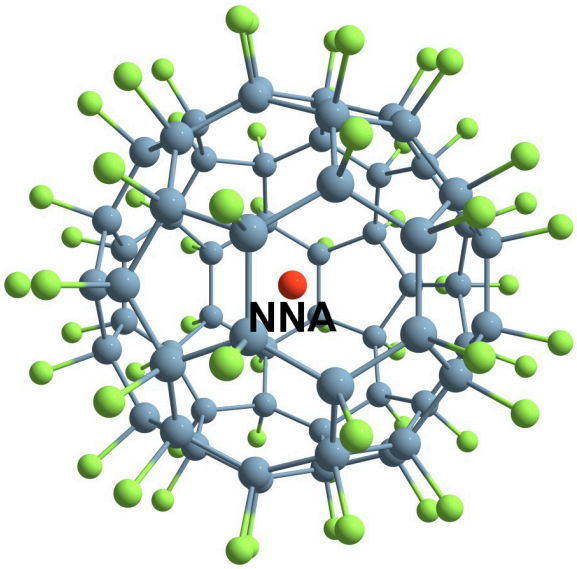

Figure 11.6: Geometries of $e@C_{60}F_{60}$ optimized at the CAM-B3LYP/ma-Def2TZVP+DF level of theory. The NNA position is indicated with a red ball.

The last molecular electride studied in this work, $e@C_{60}F_{60}$, has different character compared to the previous systems. The total charge of this system is -1 (doublet electronic ground-state ^2A_g) and the excess anionic electron is captured and stabilized in the middle of the fullerene cage (see Figure 11.6). This is the largest molecule included in this work, and its size prevents the use of large basis sets. For that reason, to properly capture this effect, one has to augment the basis set with a few DF functions placed in the middle of the cage (see Section 11.2). The results are compiled in Table 11.9. Description of the NNA basin is much more sensitive to the level of theory (and the amount of electron delocalization) for this system than the others. Treating MP2 as the reference, B3LYP underestimates the number of average electrons in the NNA basin and only a small fraction of those is explicitly localized there. In contrast, HF significantly overestimates N_{NNA} and λ_{NNA}. The results obtained with CAM-B3LYP agree well with the MP2 values, and predict N_{NNA} of about 0.3, with about 30% of localization. Nevertheless, the EOS analysis shows a zero oxidation state for the NNA basin with large RI values. This is further supported with the rather large values of the P_0 probabilities of finding exactly zero electrons in the NNA basin, which for MP2 and CAM-B3LYP span in the range 64 - 70%. However, a probability of finding exactly one electron is non-negligible (P_1 = 30 - 34%), which suggests that this system may be considered as a low-electron-number molecular electride.

Table 11.9: Results of the QTAIM partitioning, EOS, and EDF analyses for $e@C_{60}F_{60}$. For a full description, see Table 11.1.

Method	Basis	$\rho(\mathbf{r}_{\mathrm{NNACP}})$	$\nabla^2\rho(\mathbf{r}_{\mathrm{NNACP}})$	N_{NNA}	λ_{NNA}	EOS[1]	RI	P_0	P_1	P_2
CAM-B3LYP	6-31G+DF	$3.99{\cdot}10^{-3}$	$-3.77{\cdot}10^{-3}$	0.314	0.091	0	89	69.7	30.3	0.1
	6-31G(d)+DF	$3.94{\cdot}10^{-3}$	$-3.57{\cdot}10^{-3}$	0.318	0.093	0	88	69.4	30.6	0.0
	6-31+G+DF	$2.85{\cdot}10^{-3}$	$-9.90{\cdot}10^{-4}$	0.343	0.108	0	84	67.1	32.9	0.0
	ma-Def2TZVP+DF	$2.64{\cdot}10^{-3}$	$-6.52{\cdot}10^{-4}$	0.354	0.114	0	83	66.0	33.8	0.2
B3LYP	6-31G+DF	$3.26{\cdot}10^{-3}$	$-3.07{\cdot}10^{-3}$	0.241	0.054	0	100	76.7	23.3	0.0
	6-31G(d)+DF	$3.24{\cdot}10^{-3}$	$-2.92{\cdot}10^{-3}$	0.247	0.057	0	100	76.1	23.8	0.0
	6-31+G+DF	$2.35{\cdot}10^{-3}$	$-8.38{\cdot}10^{-4}$	0.253	0.059	0	100	75.6	24.4	0.0
	ma-Def2TZVP+DF	$2.18{\cdot}10^{-3}$	$-5.76{\cdot}10^{-4}$	0.268	0.065	0	99	74.4	25.6	0.0
HF	6-31G+DF	$6.73{\cdot}10^{-3}$	$-7.52{\cdot}10^{-3}$	0.506	0.236	0	52	51.2	48.6	0.2
	6-31G(d)+DF	$6.55{\cdot}10^{-3}$	$-7.12{\cdot}10^{-3}$	0.504	0.234	0	53	51.6	48.4	0.1
	6-31+G+DF	$4.70{\cdot}10^{-3}$	$-2.54{\cdot}10^{-3}$	0.569	0.296	-1	59	45.2	54.4	0.5
	ma-Def2TZVP+DF	$4.35{\cdot}10^{-3}$	$-1.70{\cdot}10^{-3}$	0.594	0.319	-1	59	43.2	56.3	0.5
	cc-pVDZ+DF	$6.24{\cdot}10^{-3}$	$-6.58{\cdot}10^{-3}$	0.483	0.213	0	57	53.8	46.2	0.1
MP2	6-31G+DF	$4.05{\cdot}10^{-3}$	$-3.40{\cdot}10^{-3}$	0.337	0.111			68.6	31.4	0.0
	6-31G(d)+DF	$3.57{\cdot}10^{-3}$	$-2.61{\cdot}10^{-3}$	0.314	0.098			64.5	35.2	0.2
	6-31+G+DF	$2.51{\cdot}10^{-3}$	$-7.78{\cdot}10^{-5}$	0.357	0.123			66.4	33.4	0.2

11.4 Conclusions

A correct description of a molecular electride hinges on the electronic structure method employed for the analysis. Electron (de)localization has a key importance in the description of the electrons in the NNA basins. In general, the number and the localization of electrons in the NNA basins increase with the amount of the HF exchange included in the method. B3LYP usually gives the lowest average electron numbers in the NNA basins, the largest zero-electron probabilities, and the lowest oxidation states. This result is in line with the large delocalization errors documented in conjugated and aromatic systems [454–457]. M06-2X, a global hybrid functional with 54% of the exact exchange, and CAM-B3LYP, a range-separated hybrid functional, do not suffer from so large delocalization errors and often provide similar NNA electron numbers to MP2. CAM-B3LYP seems to provide the closest agreement with the reference MP2 numbers. On the other hand, HF, which in general overestimates the electron localization, provides the largest electron numbers, localization indices, and most negative oxidation states in the NNA basins. More strikingly, it seems that there exist DFAs which are completely inadequate to study the topology of the molecular electrides. This is the case of MN15, which failed to provide results in almost all studied systems, despite this functional yielding high-quality results in benchmarks on the molecular properties. The basis set employed should include sufficiently flexible functions in the atoms that are close to the NNA. Otherwise, those centers may not be detectable.

Even with a properly suited electronic structure method, the question of how many electrons are held in a given molecular electride cannot be straight forwardly answered. By means of EOS analysis, for most of the electrides, one could not attribute the anionic character to the NNA basin. Moreover, the obtained oxidation state (and the reliability of this assignment) may depend on the choice of the fragmentation. With the use of EDFs and corresponding probabilities, one can support or deny the assignment made by EOS. On the other hand, if one solely took as a reference criterion the probability given by the EDF of finding at least one electron in the NNA basin, one could not make a clear-cut decision. Taking the analysis of all the results into account, in general one can quantitatively indicate that one-electron electrides usually correspond to 1 - P_0 > 0.5, $N_{NNA} \leq 0.5$ and $\lambda_{NNA} \leq 0.2$, and the examples of those are $TCNQLi_2$, Mg_2EP, and $Mg_2@C_{60}$. On the contrary, Li@calix[4]pyrrole, Na@calix[4]pyrrole, and $e@C_{60}F_{60}$ cannot be classified as systems having anionic electrons because the number electrons in the NNA basins is too low. Lastly, the molecular electrides containing the Na_2 motifs, namely $TCNQNa_2$, $TCNENa_3$, and $TCNENa_4$(II), seem to be the borderline cases having between zero and one electron in the NNA basin.

Part IV

Epilogue

Chapter 12

Final Conclusions

In this Thesis, several significant aspects to achieve efficient computations of NLOPs and vibrational properties were addressed. They are also of relevance for the development of quantum methods that are accurate for all purposes, because the proper description of the response properties, such as NLOPs, requires a very rigorous treatment of electron correlation. The present work focused both on the wavefunction-based methods, as well as the DFT methods.

In Chapter 7, we have described a novel decomposition of nonresonant electronic responses, the Partition of Optical Properties Into Orbital Contributions. It expresses the response to electric fields in terms of the NOs of the unperturbed molecular systems. **PNOC is a convenient tool for studying the effects of the static and dynamic correlation on NLOPs and assessing the performance of QC methods more deeply.** For example, using PNOC one can track the changes in the electric response due to the alteration of the chemical structure, such as the formation of radical structures. Moreover, with the aid of PNOC we showed that conclusions drawn solely on the total values of NLOPs may hide a compensation of errors in the description of the response. The latter is especially noticeable for *para*-benzyne, an open-shell diradical system characterized by a strong multireference character for which static correlation drives its molecular response. We unveiled that MP2 provides a good description of the nature of the response for closed-shell systems but it does not characterize well the response of diradical systems, even if one adopts the unrestricted formalism. On the contrary, although the CASSCF method always provides systematically underestimated values of NLOPs in both types of systems, it correctly describes the orbital partition of the response. The smaller values of NLOPs for CASSCF are due to an underestimation of dynamic correlation.

In Chapter 8, we assessed the performance of popular DFAs in the simulations of the nuclear relaxation response of weakly-bonded complexes. As we showed, due to the high anharmonic character of the intermolecular motion, the description of the vibrational response of these systems is demanding for some DFAs. Among the tested functionals, CAM-B3LYP yielded the smallest average errors in the properties, which were smaller than 20%. On the contrary, for ωB97X, M06, and M06-2X, we observed huge relative errors in the NR contributions to the second hyperpolarizability, exceeding hundreds of percent. Through a subsequent analysis, we unveiled that for these DFAs the largest errors corresponded to the the anharmonic NR terms, which depend on high-order derivatives of energy and electronic properties. **We demonstrated that such high-order derivatives are highly susceptible to spurious oscillations which introduce huge errors in the overall properties.** These errors are a consequence of regular instabilities due to insufficient size of the grid used in the numerical integration of exchange-correlation components of DFAs such as ωB97X, M06, and M06-2X. Lastly, we proved that with a large

enough integration grid, (750, 974), which is a few times larger than the SuperFineGrid (the largest predefined grid provided by Gaussian16), ωB97X can provide very accurate values of the NR contributions to NLOPs, even more accurate than those of CAM-B3LYP.

In Chapter 9, we focused on the phenomenon of the grid-related spurious oscillations in DFAs. **We developed a unique tool to efficiently detect and quantify the errors in the property derivatives coming from the spurious oscillations.** With its aid, we broadly inspected dozens of DFAs and assessed their overall grid stability by analyzing the magnitude of errors in the derivatives with respect to the nuclear displacements of the energy, dipole moment, and polarizability. First, we demonstrated that the (99, 590) integration grid, which is an unpruned version of many predefined default grids in quantum chemical packages, is too small to study high-order derivatives that appear in the expressions of anharmonic vibrational properties. With this grid, one can already observe errors exceeding 100% for $\mathrm{d}^2E/\mathrm{d}\xi^2$ and $\mathrm{d}\alpha_{zz}/\mathrm{d}\xi$. For most DFAs, at least the (250, 974) grid must be used to ensure some numerical stability. Second, we grouped the DFAs according to the maximum order of derivatives with errors coming from the spurious oscillations smaller than 10%. Only a handful of DFAs, namely BH&H, LC-BLYP, PBE, CAM-B3LYP, could safely reproduce high-order derivatives, such as $\mathrm{d}^4E/\mathrm{d}\xi^4$ or $\mathrm{d}^3\mu_z/\mathrm{d}\xi^3$. Interestingly, dozens of DFAs do not correctly reproduce derivatives higher than $\mathrm{d}^2E/\mathrm{d}\xi^2$, $\mathrm{d}\mu_z/\mathrm{d}\xi$, and $\mathrm{d}\alpha_{zz}/\mathrm{d}\xi$ due to the grid-related spurious oscillations. In the most demanding systems, such as Ar_2 and He_2, some DFAs could not even yield numerically stable first-order derivatives. Last, in this work, we proved that **the spurious oscillations may affect not only the intermolecular motion of the noncovalently-bonded complexes, but many kinds of vibrations of both organic and inorganic chemical systems.**

In Chapter 10, we tested different approaches to decrease the computational cost of the algorithm to detect the spurious oscillations by reducing the cost of the electronic structure computations of the property profiles. In particular, one can utilize a small basis set for the analysis, or if one is solely interested in the errors arising from the exchange-correlation functional, use trial densities. **The latter approach allows to quantitatively measure the errors coming from the grid-related spurious oscillations almost at no cost.** Moreover, in this Chapter, we analyzed the direct and indirect errors caused by the spurious oscillations. The direct errors not only affect the electron correlation and exchange terms of the electronic energy but also the MOs evaluated in the SCF procedure. Erroneous MOs then induce an indirect error to all the other terms of the electronic energy and other molecular properties. Lastly, we accurately decomposed spurious oscillations into exchange and correlation components. This allowed us to identify the origin of the grid instability and spurious oscillations in particular types of the exchange and correlation functionals. The exchange functionals based on B88, PBE, and B97 were found to be the most grid robust. Among the correlation functionals, the LYP correlation was the most stable one. Most notably, we unveiled that in the B97-based functionals (such as ωB97X), which are greatly affected by spurious oscillations, only the B97 correlation part is responsible for huge errors. In the case of other grid-dependent standalone functionals, such as M06 or M06-2X, both exchange and correlation components suffered from considerable grid errors.

In Chapter 11, we evaluated the performance of several quantum chemical methods to describe the topological features of the electronic structure of molecular electrides. For nine molecular electrides, we aimed to establish the most likely number of electrons localized in the non-nuclear attractor center. For each NNA basin, this was done by computing its average electronic population, effective oxidation state (EOS), and corresponding values of electron distribution functions (EDFs). We observed that the final outcome may

vary depending on the electronic structure method. For example, the electron localization in the NNA basin increases with the amount of HF exchange included in the DFA, which results in larger negative oxidation states. In contrast, upon introduction of dynamic correlation, one observes the opposite effect. For example, MP2 yields lower probabilities of finding electrons in the NNA basins and smaller negative oxidation states than in HF. **For most examples, the EOS approach did not attribute a negative oxidation state to the NNA basin, which does not agree with the formal classification of these molecules as electrides.** The negative EOS of the NNA basins were found only for $TCNQLi_2$, Mg_2EP, and Mg_2C_{60}, whereas for the rest of molecules EOS is zero. However, a deeper analysis using EDFs revealed that the probabilities of finding at least one electron in the NNA basin are large enough to consider most of the studied molecules at least as low-electron number electrides. Beside $TCNQLi_2$, Mg_2EP, and Mg_2C_{60}, which clearly contained at least one localized electron in the NNA basin, in $TCNQNa_2$, $TCNENa_3$, and $TCNENa_4$(II) the MP2 probabilities of having at least one electron is larger than 30%. Therefore, it is justified to use the concept of high- and low-electron number electrides, rather than a binary classification based on the oxidation state.

This Thesis proves that both wavefunction-based methods and DFAs may yield very large errors in the calculation of nonlinear optical and vibrational properties. We believe that the presented findings on the source of errors and instabilities in the electronic structure methods, in particular, DFAs, will have an impact on the development of more efficient and well-rounded methods.

Bibliography

[1] M. Göppert-Mayer, *Ann. Phys.* **1931**, *401*, 273–294.

[2] W. Kaiser, C. Garrett, *Phys. Rev. Lett.* **1961**, *7*, 229.

[3] T. H. Maiman, *Nature* **1960**, *187*, 493–494.

[4] P. A. Franken, A. E. Hill, C. W. Peters, G Weinreich, *Phys. Rev. Lett.* **1961**, *7*, 118.

[5] W. Denk, J. H. Strickler, W. W. Webb, *Science* **1990**, *248*, 73–76.

[6] P. J. Campagnola, A. Lewis, L. M. Loew, *Biophys. J.* **1999**, *77*, 3341–3349.

[7] F. Helmchen, W. Denk, *Nat. Methods* **2005**, *2*, 932–940.

[8] S. Yao, K. D. Belfield, *Eur. J. Org. Chem.* **2012**, *2012*, 3199–3217.

[9] L. Lin, Z. He, T. Zhang, Y. Zuo, X. Chen, Z. Abdelrahman, F. Chen, Z. Wei, K. Si, W. Gong, et al., *J. Mater. Chem. B* **2022**.

[10] X Zhuang, P. Miranda, D Kim, Y. Shen, *Phys. Rev. B* **1999**, *59*, 12632.

[11] Y. Shen, *Nature* **1989**, *337*, 519–525.

[12] M. Raschke, M Hayashi, S. Lin, Y. Shen, *Chem. Phys. Lett.* **2002**, *359*, 367–372.

[13] J. Hunt, P. Guyot-Sionnest, Y. Shen, *Chem. Phys. Lett.* **1987**, *133*, 189–192.

[14] R Superfine, J. Y. Huang, Y. Shen, *Phys. Rev. Lett.* **1991**, *66*, 1066.

[15] P. Norman, K. Ruud in *Non-Linear Optical Properties of Matter*, (Eds.: M. G. Papadopoulos, A. J. Sadlej, J. Leszczynski), Challenges and Advances in Computational Chemistry and Physics, Springer Netherlands, **2006**, pp. 1–49.

[16] P. Norman, K. Ruud, T. Saue, *Principles and Practices of Molecular Properties: Theory, Modeling, and Simulations*, John Wiley & Sons, **2018**.

[17] J. Ward, *Rev. Mod. Phys.* **1965**, *37*, 1–18.

[18] B. Orr, J. Ward, *Mol. Phys.* **1971**, *20*, 513–526.

[19] J. L. Oudar, D. S. Chemla, *J. Chem. Phys.* **1977**, *66*, 2664–2668.

[20] M. M. Alam, M. Chattopadhyaya, S. Chakrabarti, K. Ruud, *Acc. Chem. Res.* **2014**, *47*, 1604–1612.

[21] J. Olsen, P. Jørgensen, *J. Chem. Phys.* **1985**, *82*, 3235–3264.

[22] J. Olsen, P. Jørgensen in *Modern Electronic Structure Theory*, *Vol. 2*, World Scientific, Singapore, **1995**.

[23] O. Christiansen, S. Coriani, J. Gauss, C. Hättig, P. Jørgensen, F. Pawłowski, A. Rizzo in *Non-Linear Optical Properties of Matter*, (Eds.: M. G. Papadopoulos, A. J. Sadlej, J. Leszczynski), Challenges and Advances in Computational Chemistry and Physics, Springer Netherlands, **2006**, pp. 51–99.

[24] P. Norman, *Phys. Chem. Chem. Phys.* **2011**, *13*, 20519–20535.

[25] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, K. Ruud, *Chem. Rev.* **2012**, *112*, 543–631.

[26] J. Linderberg, Y. Öhrn, *Propagators in Quantum Chemistry*, John Wiley & Sons, **2004**.

[27] P. Langhoff, S. Epstein, M Karplus, *Rev. Mod. Phys.* **1972**, *44*, 602.

[28] K. Sasagane, F. Aiga, R. Itoh, *J. Chem. Phys.* **1993**, *99*, 3738–3778.

[29] O. Christiansen, P. Jørgensen, C. Hättig, *Int. J. Quantum Chem.* **1998**, *68*, 1–52.

[30] K. Kristensen, J. Kauczor, T. Kjærgaard, P. Jørgensen, *J. Chem. Phys.* **2009**, *131*, 044112.

[31] P. Sałek, O. Vahtras, T. Helgaker, H. Ågren, *J. Chem. Phys.* **2002**, *117*, 9630–9645.

[32] O. Quinet, B. Kirtman, B. Champagne, *J. Chem. Phys.* **2003**, *118*, 505–513.

[33] H. Koch, P. Jørgensen, *J. Chem. Phys.* **1990**, *93*, 3333–3344.

[34] C. Hättig, O. Christiansen, H. Koch, P. Jørgensen, *Chem. Phys. Lett.* **1997**, *269*, 428–434.

[35] C. Hättig, O. Christiansen, P. Jørgensen, *Chem. Phys. Lett.* **1998**, *282*, 139–146.

[36] O. Christiansen, H. Koch, P. Jørgensen, T. Helgaker, *J. Chem. Phys.* **1998**, *108*, 2801–2816.

[37] C. Hättig, O. Christiansen, P. Jørgensen, *J. Chem. Phys.* **1998**, *108*, 8331–8354.

[38] J. R. Hammond, K. Kowalski, *J. Chem. Phys.* **2009**, *130*, 194108.

[39] D. H. Friese, C. Hättig, K. Ruud, *Phys. Chem. Chem. Phys.* **2012**, *14*, 1175–1184.

[40] M. Pecul, F. Pawłowski, P. Jørgensen, A. Köhn, C. Hättig, *J. Chem. Phys.* **2006**, *124*, 114101.

[41] U. Ekström, L. Visscher, R. Bast, A. J. Thorvaldsen, K. Ruud, *J. Chem. Theory Comput.* **2010**, *6*, 1971–1980.

[42] M. Ringholm, R. Bast, L. Oggioni, U. Ekström, K. Ruud, *J. Chem. Phys.* **2014**, *141*, 134107.

[43] E. Rudberg, P. Sałek, T. Helgaker, H. Ågren, *J. Chem. Phys.* **2005**, *123*, 184108.

[44] A. Masunov, S. Tretiak, *J. Phys. Chem. B* **2004**, *108*, 899–907.

[45] P. Sałek, O. Vahtras, J. Guo, Y. Luo, T. Helgaker, H. Ågren, *Chem. Phys. Lett.* **2003**, *374*, 446–452.

[46] F. Zahariev, M. S. Gordon, *J. Chem. Phys.* **2014**, *140*, 18A523.

[47] C. B. Nielsen, O. Christiansen, K. V. Mikkelsen, J. Kongsted, *J. Chem. Phys.* **2007**, *126*, 154112.

[48] S. Van Gisbergen, J. Snijders, E. Baerends, *J. Chem. Phys.* **1995**, *103*, 9347–9354.

[49] R. Bauernschmitt, R. Ahlrichs, *Chem. Phys. Lett.* **1996**, *256*, 454–464.

[50] C. Jamorski, M. E. Casida, D. R. Salahub, *J. Chem. Phys.* **1996**, *104*, 5134–5147.

[51] M. Petersilka, U. Gossmann, E. Gross, *Phys. Rev. Lett.* **1996**, *76*, 1212.

[52] A. Görling, *Int. J. Quantum Chem.* **1998**, *69*, 265–277.

[53] R. E. Stratmann, G. E. Scuseria, M. J. Frisch, *J. Chem. Phys.* **1998**, *109*, 8218–8224.

[54] D. J. Tozer, N. C. Handy, *J. Chem. Phys.* **1998**, *109*, 10180–10189.

[55] S. Hirata, M. Head-Gordon, *Chem. Phys. Lett.* **1999**, *302*, 375–382.

[56] A. Görling, H. H. Heinze, S. P. Ruzankin, M. Staufer, N. Rösch, *J. Chem. Phys.* **1999**, *110*, 2785–2799.

[57] Z. Rinkevicius, I. Tunell, P. Sałek, O. Vahtras, H. Ågren, *J. Chem. Phys.* **2003**, *119*, 34–46.

[58] S. Van Gisbergen, J. Snijders, E. Baerends, *J. Chem. Phys.* **1998**, *109*, 10644–10656.

[59] S. Van Gisbergen, J. Snijders, E. Baerends, *Comput. Phys. Commun.* **1999**, *118*, 119–138.

[60] S. Tretiak, V. Chernyak, *J. Chem. Phys.* **2003**, *119*, 8809–8823.

[61] I. Tunell, Z. Rinkevicius, O. Vahtras, P. Sałek, T. Helgaker, H. Ågren, *J. Chem. Phys.* **2003**, *119*, 11024–11034.

[62] B. Jansik, D. Jonsson, P. Sałek, H. Ågren, *J. Chem. Phys.* **2004**, *121*, 7595–7600.

[63] B. Jansik, P. Sałek, D. Jonsson, O. Vahtras, H. Ågren, *J. Chem. Phys.* **2005**, *122*, 054107.

[64] A. Zawada, R. W. Góra, M. M. Mikołajczyk, W. Bartkowiak, *J. Phys. Chem. A* **2012**, *116*, 4409–4416.

[65] K. Y. Suponitsky, A. E. Masunov, *J. Chem. Phys.* **2013**, *139*, 094310.

[66] L. Lescos, S. P. Sitkiewicz, P. Beaujean, M. Blanchard-Desce, B. Champagne, E. Matito, F. Castet, *Phys. Chem. Chem. Phys.* **2020**, *22*, 16579–16594.

[67] L. Lescos, P. Beaujean, C. Tonnelé, P. Aurel, M. Blanchard-Desce, V. Rodriguez, M. de Wergifosse, B. Champagne, L. Muccioli, F. Castet, *Phys. Chem. Chem. Phys.* **2021**, *23*, 23643–23654.

[68] M. Chołuj, M. M. Alam, M. T. Beerepoot, S. P. Sitkiewicz, E. Matito, K. Ruud, R. Zaleśny, *J. Chem. Theory Comput.* **2022**.

[69] M. Rosendale, J. Daniel, F. Castet, P. Pagano, J.-B. Verlhac, M. Blanchard-Desce, *Molecules* **2022**, *27*, 2230.

[70] P. Beaujean, L. Sanguinet, V. Rodriguez, F. Castet, B. Champagne, *Molecules* **2022**, *27*, 2770.

[71] T. N. Ramos, F. Castet, B. Champagne, *J. Phys. Chem. B* **2021**, *125*, 3386–3397.

[72] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, **1989**.

[73] A. D. Becke, *J. Chem. Phys.* **2014**, *140*, 18A301.

[74] K. Burke, *J. Chem. Phys.* **2012**, *136*, 150901.

[75] H. S. Yu, X. He, S. L. Li, D. G. Truhlar, *Chem. Sci.* **2016**, *7*, 5032–5051.

[76] H. S. Yu, X. He, D. G. Truhlar, *J. Chem. Theory Comput.* **2016**, *12*, 1280–1293.

[77] N. Mardirossian, M. Head-Gordon, *J. Chem. Phys.* **2016**, *144*, 214110.

[78] N. Mardirossian, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

[79] A. J. Cohen, P Mori-Sánchez, W. Yang, *Science* **2008**, *321*, 792–794.

[80] S. Grimme, *WIREs Comput. Mol. Sci.* **2011**, *1*, 211–228.

[81] A. D. Becke, *J. Chem. Phys.* **2014**, *140*, 18A301.

[82] S. J. A. van Gisbergen, P. R. T. Schipper, O. V. Gritsenko, E. J. Baerends, J. G. Snijders, B. Champagne, B. Kirtman, *Phys. Rev. Lett.* **1999**, *83*, 694–697.

[83] O. V. Gritsenko, E. J. Baerends, *Phys. Rev. A* **2001**, *64*, 042506.

[84] S. Kümmel, L. Kronik, J. P. Perdew, *Phys. Rev. Lett.* **2004**, *93*, 213002.

[85] F. A. Bulat, A. Toro-Labbé, B. Champagne, B. Kirtman, W. Yang, *J. Chem. Phys.* **2005**, *123*, 014319.

[86] A. Kaiser, S. Kümmel, *Phys. Rev. A* **2018**, *98*, 052505.

[87] B. Champagne, E. A. Perpète, S. J. Van Gisbergen, E. J. Baerends, J. G. Snijders, C. Soubra-Ghaoui, K. A. Robins, B. Kirtman, *J. Chem. Phys.* **1998**, *109*, 10489–10498.

[88] B. Champagne, E. A. Perpète, D. Jacquemin, S. J. Van Gisbergen, E. J. Baerends, C. Soubra-Ghaoui, K. A. Robins, B. Kirtman, *J. Phys. Chem. A* **2000**, *104*, 4755–4763.

[89] M. Nakano, R. Kishi, T. Nitta, T. Kubo, K. Nakasuji, K. Kamada, K. Ohta, B. Champagne, E. Botek, K. Yamaguchi, *J. Phys. Chem. A* **2005**, *109*, 885–891.

[90] B. Kirtman, S. Bonness, A. Ramirez-Solis, B. Champagne, H. Matsumoto, H. Sekino, *J. Chem. Phys.* **2008**, *128*, 114108.

[91] A. Ruzsinszky, J. P. Perdew, G. I. Csonka, G. E. Scuseria, O. A. Vydrov, *Phys. Rev. A* **2008**, *77*, 060502.

[92] T. Stein, L. Kronik, R. Baer, *J. Am. Chem. Soc.* **2009**, *131*, 2818–2820.

[93] P. A. Limacher, K. V. Mikkelsen, H. P. Lüthi, *J. Chem. Phys.* **2009**, *130*, 194114.

[94] M. Srebro, J. Autschbach, *J. Chem. Theory Comput.* **2012**, *8*, 245–256.

[95] A. Zawada, A. Kaczmarek-Kędziera, W. Bartkowiak, *J. Mol. Model.* **2012**, *18*, 3073–3086.

[96] G. Maroulis, *Int. J. Quantum Chem.* **2012**, *112*, 2231–2241.

[97] M. de Wergifosse, F. Wautelet, B. Champagne, R. Kishi, K. Fukuda, H. Matsui, M. Nakano, *J. Phys. Chem. A* **2013**, *117*, 4709–4715.

[98] H. Sun, J. Autschbach, *ChemPhysChem* **2013**, *14*, 2450–2461.

[99] J. Autschbach, M. Srebro, *Acc. Chem. Res.* **2014**, *47*, 2592–2602.

[100] K. Garrett, X. Sosa Vazquez, S. B. Egri, J. Wilmer, L. E. Johnson, B. H. Robinson, C. M. Isborn, *J. Chem. Theory Comput.* **2014**, *10*, 3821–3831.

[101] A. J. Garza, O. I. Osman, A. M. Asiri, G. E. Scuseria, *J. Phys. Chem. B* **2015**, *119*, 1202–1212.

[102] M. B. Oviedo, N. V. Ilawe, B. M. Wong, *J. Chem. Theory Comput.* **2016**, *12*, 3593–3602.

[103] T. J. Duignan, J. Autschbach, E. Batista, P. Yang, *J. Chem. Theory Comput.* **2017**, *13*, 3614–3625.

[104] D. Hait, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2018**, *20*, 19800–19810.

[105] A. K. Manna, S. Refaely-Abramson, A. M. Reilly, A. Tkatchenko, J. B. Neaton, L. Kronik, *J. Chem. Theory Comput.* **2018**, *14*, 2919–2929.

[106] M. Chołuj, J. Kozłowska, W. Bartkowiak, *Int. J. Quantum Chem.* **2018**, *118*, e25666.

[107] P. Besalú-Sala, S. P. Sitkiewicz, P. Salvador, E. Matito, J. M. Luis, *Phys. Chem. Chem. Phys.* **2020**, *22*, 11871–11880.

[108] R. Zaleśny, M. Medved', S. P. Sitkiewicz, E. Matito, J. M. Luis, *J. Chem. Theory Comput.* **2019**, *15*, 3570–3579.

[109] R. Bast, U. Ekstrom, B. Gao, T. Helgaker, K. Ruud, A. J. Thorvaldsen, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2627–2651.

[110] D. M. Bishop, B. Kirtman, *J. Chem. Phys.* **1991**, *95*, 2646–2658.

[111] D. M. Bishop, B. Kirtman, *J. Chem. Phys.* **1992**, *97*, 5255–5256.

[112] D. M. Bishop in *Advances in Chemical Physics*, John Wiley & Sons, Ltd, pp. 1–40.

[113] J. Martí, D. M. Bishop, *J. Chem. Phys.* **1993**, *99*, 3860–3864.

[114] J. Martí, J. L. Andrés, J. Bertrán, M. Duran, *Mol. Phys.* **1993**, *80*, 625–633.

[115] J. M. Luis, J. Martí, M. Duran, J. L. Andrés, B. Kirtman, *J. Chem. Phys.* **1998**, *108*, 4123–4130.

[116] D. M. Bishop, M. Hasan, B. Kirtman, *J. Chem. Phys.* **1995**, *103*, 4157–4159.

[117] J. M. Luis, M. Duran, J. L. Andrés, B. Champagne, B. Kirtman, *J. Chem. Phys.* **1999**, *111*, 875–884.

[118] J. M. Luis, M. Duran, B. Champagne, B. Kirtman, *J. Chem. Phys.* **2000**, *113*, 5203–5213.

[119] J. M. Luis, M. Duran, B. Kirtman, *J. Chem. Phys.* **2001**, *115*, 4473–4483.

[120] J. M. Bowman, *J. Chem. Phys.* **1978**, *68*, 608–610.

[121] R. Gerber, M. Ratner, *Chem. Phys. Lett.* **1979**, *68*, 195–198.

[122] J. M. Bowman, *Acc. Chem. Res.* **1986**, *19*, 202–208.

[123] M. Torrent-Sucarrat, J. M. Luis, B. Kirtman, *J. Chem. Phys.* **2005**, *122*, 204108.

[124] J. M. Luis, M. Torrent-Sucarrat, O. Christiansen, B. Kirtman, *J. Chem. Phys.* **2007**, *127*, 084118.

[125] L. S. Norris, M. A. Ratner, A. E. Roitberg, R. Gerber, *J. Chem. Phys.* **1996**, *105*, 11261–11267.

[126] G. M. Chaban, J. O. Jung, R. B. Gerber, *J. Chem. Phys.* **1999**, *111*, 1823–1829.

[127] O. Christiansen, *J. Chem. Phys.* **2003**, *119*, 5773–5781.

[128] J. M. Bowman, K. Christoffel, F. Tobin, *J. Phys. Chem.* **1979**, *83*, 905–912.

[129] J. Bowman, K. Christoffel, *Chem. Phys. Lett.* **1982**, *85*, 220–24.

[130] O. Christiansen, *J. Chem. Phys.* **2004**, *120*, 2149–2159.

[131] P. Seidler, M. Sparta, O. Christiansen, *J. Chem. Phys.* **2011**, *134*, 054119.

[132] A. Zoccante, P. Seidler, O. Christiansen, *J. Chem. Phys.* **2011**, *134*, 154101.

[133] J. Kongsted, O. Christiansen, *J. Chem. Phys.* **2007**, *127*, 154315.

[134] P. Chopra, L. Carlacci, H. F. King, P. N. Prasad, *J. Phys. Chem.* **1989**, *93*, 7120–7130.

[135] M. Nakano, I. Shigemoto, S. Yamada, K. Yamaguchi, *J. Chem. Phys.* **1995**, *103*, 4175–4191.

[136] S. Yamada, M. Nakano, I. Shigemoto, K. Yamaguchi, *Chem. Phys. Lett.* **1996**, *254*, 158–164.

[137] M. Nakano, S. Yamada, K. Yamaguchi, *J. Comput. Methods Sci. Eng.* **2004**, *4*, 677–701.

[138] M. Head-Gordon, *Chem. Phys. Lett.* **2003**, *372*, 508–511.

[139] M. Nakano, H. Fukui, T. Minami, K. Yoneda, Y. Shigeta, R. Kishi, B. Champagne, E. Botek, T. Kubo, K. Ohta, et al., *Theor. Chem. Acc.* **2011**, *130*, 711–724.

[140] T. Yamada, S. Takamuku, H. Matsui, B. Champagne, M. Nakano, *Chem. Phys. Lett.* **2014**, *608*, 68–73.

[141] V. M. Geskin, J. L. Brédas, *J. Chem. Phys.* **1998**, *109*, 6163–6168.

[142] V. M. Geskin, C. Lambert, J. L. Brédas, *J. Am. Chem. Soc.* **2003**, *125*, 15651–15658.

[143] W. Hieringer, E. J. Baerends, *J. Phys. Chem. A* **2006**, *110*, 1014–1021.

[144] N. Otero, P. Karamanis, M. Mandado, *Phys. Chem. Chem. Phys.* **2019**, *21*, 6274–6286.

[145] M. Montilla, J. M. Luis, P. Salvador, *J. Chem. Theory Comput.* **2021**, *17*, 1098–1105.

[146] S. P. Sitkiewicz, M. Rodríguez-Mayorga, J. M. Luis, E. Matito, *Phys. Chem. Chem. Phys.* **2019**, *21*, 15380–15391.

[147] S. P. Sitkiewicz, R. Zaleśny, E. Ramos-Cordoba, J. Luis, E. Matito, *J. Phys. Chem. Lett. (accepted)* **2022**, doi:10.1021/acs.jpclett.2c01278.

[148] S. P. Sitkiewicz, E. Ramos-Cordoba, J. M. Luis, E. Matito, *J. Phys. Chem. A* **2021**, *125*, 4819–4835.

[149] B. Kirtman, J. M. Luis in *Non-Linear Optical Properties of Matter*, (Eds.: M. G. Papadopoulos, A. J. Sadlej, J. Leszczynski), Challenges and Advances in Computational Chemistry and Physics, Springer Netherlands, **2006**, pp. 101–128.

[150] W. Bartkowiak, R. Zaleśny in *Non-Linear Optical Properties of Matter*, (Eds.: M. G. Papadopoulos, A. J. Sadlej, J. Leszczynski), Challenges and Advances in Computational Chemistry and Physics, Springer Netherlands, **2006**, pp. 129–150.

[151] D. M. Bishop, P. Norman in *Handbook of Advanced Electronic and Photonic Materials and Devices, Ten-Volume Set*, *Vol. 9*, (Ed.: H. S. Nalwa), Academic Press, **2000**, Chapter 1.

[152] A. Buckingham, *Q. Rev. Chem. Soc.* **1959**, *13*, 183–214.

[153] R. W. Boyd, *Nonlinear optics. Third Edition*, Academic press, **2008**.

[154] D. R. Bates, *Quantum Theory: Elements*, *Vol. 1*, Academic Press, **1961**, Chapter 8.

[155] S. Shaik, D. Mandal, R. Ramanan, *Nat. Chem.* **2016**, *8*, 1091–1098.

[156] F. Che, J. T. Gray, S. Ha, N. Kruse, S. L. Scott, J.-S. McEwen, *ACS Catal.* **2018**, *8*, 5153–5174.

[157] X. Huang, C. Tang, J. Li, L.-C. Chen, J. Zheng, P. Zhang, J. Le, R. Li, X. Li, J. Liu, et al., *Sci. Adv.* **2019**, *5*, eaaw3072.

[158] S. Shaik, D. Danovich, J. Joy, Z. Wang, T. Stuyver, *J. Am. Chem. Soc.* **2020**, *142*, 12551–12562.

[159] J. Joy, T. Stuyver, S. Shaik, *J. Am. Chem. Soc.* **2020**, *142*, 3836–3850.

[160] P. Besalú-Sala, M. Solá, J. M. Luis, M. Torrent-Sucarrat, *ACS Catal.* **2021**, *11*, 14467–14479.

[161] A. J. Sadlej, *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995–2016.

[162] A. J. Sadlej, M. Urban, *J. Mol. Struct. THEOCHEM* **1991**, *234*, 147–171.

[163] A. J. Sadlej, *Theor. Chim. Acta* **1991**, *81*, 45–63.

[164] A. J. Sadlej, *Theor. Chim. Acta* **1991**, *81*, 45–63.

[165] A. J. Sadlej, *Theor. Chim. Acta* **1992**, *81*, 339–354.

[166] D. M. Bishop, J. M. Luis, B. Kirtman, *J. Chem. Phys.* **1998**, *108*, 10013–10017.

[167] R. McWeeny, *Rev. Mod. Phys.* **1960**, *32*, 335.

[168] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864.

[169] J. P. Perdew, S. Kurth in *A Primer in Density Functional Theory*, *Vol. 1*, (Eds.: C. Fiolhais, F. Nogueira, M. A. L. Marques), Springer-Verlag Berlin Heidelberg New York, **2003**, pp. 1–55.

[170] L. H. Thomas in Mathematical Proceedings of the Cambridge Philosophical Society, *Vol. 23*, Cambridge University Press, **1927**, pp. 542–548.

[171] E. Fermi, *Rend. Accad. Naz. Lincei* **1927**, *6*, 32.

[172] J. C. Slater, *Phys. Rev.* **1951**, *81*, 385.

[173] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133.

[174] A. Szabo, N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover Publications, **1996**.

[175] J. A. Pople, P. M. Gill, B. G. Johnson, *Chem. Phys. Lett.* **1992**, *199*, 557–560.

[176] J. P. Perdew, K. Schmidt in AIP Conference Proceedings, *Vol. 577*, American Institute of Physics, **2001**, pp. 1–20.

[177] J. P. Perdew, *MRS Bull.* **2013**, *38*, 743–750.

[178] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[179] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[180] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.

[181] J. P. Perdew, K. Burke, Y. Wang, *Phys. Rev. B* **1996**, *54*, 16533.

[182] J. Tao, J. P. Perdew, V. N. Staroverov, G. E. Scuseria, *Phys. Rev. Lett.* **2003**, *91*, 146401.

[183] J. Sun, A. Ruzsinszky, J. P. Perdew, *Phys. Rev. Lett.* **2015**, *115*, 036402.

[184] Y. Zhao, D. G. Truhlar, *J. Chem. Phys.* **2006**, *125*, 194101.

[185] J Harris, R. Jones, *J. Phys. F: Met. Phys.* **1974**, *4*, 1170.

[186] J Harris, *Phys. Rev. A* **1984**, *29*, 1648.

[187] O. Gunnarsson, B. I. Lundqvist, *Phys. Rev. B* **1976**, *13*, 4274.

[188] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 1372–1377.

[189] T. Yanai, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, *393*, 51–57.

[190] Y. Tawada, T. Tsuneda, S. Yanagisawa, T. Yanai, K. Hirao, *J. Chem. Phys.* **2004**, *120*, 8425–8433.

[191] M. A. Rohrdanz, J. M. Herbert, *J. Chem. Phys.* **2008**, *129*, 034107.

[192] J.-D. Chai, M. Head-Gordon, *J. Chem. Phys.* **2008**, *128*, 084106.

[193] T. Yanai, D. P. Tew, N. C. Handy, *Chem. Phys. Lett.* **2004**, *393*, 51–57.

[194] A. Dreuw, M. Head-Gordon, *Chem. Rev.* **2005**, *105*, 4009–4037.

[195] D. Jacquemin, V. Wathelet, E. A. Perpete, C. Adamo, *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

[196] A. D. Laurent, D. Jacquemin, *Int. J. Quantum Chem.* **2013**, *113*, 2019–2039.

[197] B. Le Guennic, D. Jacquemin, *Acc. Chem. Res.* **2015**, *48*, 530–537.

[198] A. Charaf-Eddin, A. Planchat, B. Mennucci, C. Adamo, D. Jacquemin, *J. Chem. Theory Comput.* **2013**, *9*, 2749–2760.

[199] M. J. Peach, P. Benfield, T. Helgaker, D. J. Tozer, *J. Chem. Phys.* **2008**, *128*, 044118.

[200] E. Livshits, R. Baer, *Phys. Chem. Chem. Phys.* **2007**, *9*, 2932–2941.

[201] T. Stein, L. Kronik, R. Baer, *J. Chem. Phys.* **2009**, *131*, 244119.

[202] R. Baer, E. Livshits, U. Salzner, *Annu. Rev. Phys. Chem.* **2010**, *61*, 85–109.

[203] A. Karolewski, L. Kronik, S. Kümmel, *J. Chem. Phys.* **2013**, *138*, 204115.

[204] T. Stein, H. Eisenberg, L. Kronik, R. Baer, *Phys. Rev. Lett.* **2010**, *105*, 266802.

[205] Z. Lin, T. Van Voorhis, *J. Chem. Theory Comput.* **2019**, *15*, 1226–1241.

[206] J. Autschbach, M. Srebro, *Acc. Chem. Res.* **2014**, *47*, 2592–2602.

[207] J. F. Janak, *Phys. Rev. B* **1978**, *18*, 7165.

[208] A. Gorling, M. Levy, *Phys. Rev. B* **1993**, *47*, 13105.

[209] A. Gorling, M. Levy, *Phys. Rev. A* **1994**, *50*, 196.

[210] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618.

[211] S. Grimme, *J. Chem. Phys.* **2006**, *124*, 034108.

[212] T. Schwabe, S. Grimme, *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.

[213] E. Brémond, J. C. Sancho-Garcia, A. J. Pérez-Jiménez, C. Adamo, *J. Chem. Phys.* **2014**, *141*, 031101.

[214] E. Brémond, C. Adamo, *J. Chem. Phys.* **2011**, *135*, 024106.

[215] A. D. Becke, *J. Chem. Phys.* **1988**, *88*, 2547–2553.

[216] P. Salvador, E. Ramos-Cordoba, *J. Chem. Phys.* **2013**, *139*, 071103.

[217] F. L. Hirshfeld, *Theor. Chim. Acta* **1977**, *44*, 129–138.

[218] V. I. Lebedev, *USSR Comput. Math. & Math. Phys.* **1975**, *15*, 44–51.

[219] V. I. Lebedev, *USSR Comput. Math. & Math. Phys.* **1976**, *16*, 10–24.

[220] V. I. Lebedev, *Sib. Math. J.* **1977**, *18*, 99–107.

[221] V. I. Lebedev, D. Laikov in Doklady Mathematics, *Vol. 59*, Pleiades Publishing, Ltd., **1999**, pp. 477–481.

[222] P. M. Gill, B. G. Johnson, J. A. Pople, *Chem. Phys. Lett.* **1993**, *209*, 506–512.

[223] P. M. Gill, S.-H. Chien, *J. Comput. Chem.* **2003**, *24*, 732–740.

[224] C. W. Murray, N. C. Handy, G. J. Laming, *Mol. Phys.* **1993**, *78*, 997–1014.

[225] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, *Vol. 10*, **1972**.

[226] J. Gräfenstein, D. Cremer, *J. Chem. Phys.* **2007**, *127*, 164113.

[227] O. Treutler, R. Ahlrichs, *J. Chem. Phys.* **1995**, *102*, 346–354.

[228] M. E. Mura, P. J. Knowles, *J. Chem. Phys.* **1996**, *104*, 9848–9858.

[229] M. Mitani, Y. Yoshioka, *Theor. Chem. Acc.* **2012**, *131*, 1–15.

[230] R. Lindh, P.-Å. Malmqvist, L. Gagliardi, *Theor. Chem. Acc.* **2001**, *106*, 178–187.

[231] R. S. Mulliken, *J. Chem. Phys.* **1955**, *23*, 1833–1840.

[232] P.-O. Löwdin, *J. Chem. Phys.* **1950**, *18*, 365–375.

[233] R. F. Nalewajski, R. G. Parr, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8879–8882.

[234] P. W. Ayers, *J. Chem. Phys.* **2000**, *113*, 10886–10898.

[235] T. Verstraelen, P. W. Ayers, V. Van Speybroeck, M. Waroquier, *J. Chem. Theory Comput.* **2013**, *9*, 2221–2225.

[236] P. Bultinck, C. Van Alsenoy, P. W. Ayers, R. Carbó-Dorca, *J. Chem. Phys.* **2007**, *126*, 144111.

[237] R. F. Bader, *Chem. Rev.* **1991**, *91*, 893–928.

[238] R Bader, *Atoms in Molecules: A Quantum Theory: Oxford Univ*, Press.: Oxford, **1990**.

[239] X. Fradera, M. A. Austen, R. F. Bader, *J. Phys. Chem. A* **1999**, *103*, 304–314.

[240] P. Popelier, *Coord. Chem. Rev.* **2000**, *197*, 169–189.

[241] P. Karen, P. McArdle, J. Takats, *Pure Appl. Chem.* **2016**, *88*, 831–839.

[242] A. J. Thom, E. J. Sundstrom, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2009**, *11*, 11297–11304.

[243] I. D. Brown, *The Chemical Bond in Inorganic Chemistry: The Bond Valence Model*, *Vol. 27*, Oxford University Press, **2016**.

[244] P. H.-L. Sit, F. Zipoli, J. Chen, R. Car, M. H. Cohen, A. Selloni, *Chem.–Eur. J.* **2011**, *17*, 12136–12143.

[245] P. H.-L. Sit, R. Car, M. H. Cohen, A. Selloni, *Inorg. Chem.* **2011**, *50*, 10259–10267.

[246] G. Knizia, *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.

[247] M. Gimferrer, A. Aldossary, P. Salvador, M. Head-Gordon, *J. Chem. Theory Comput.* **2021**.

[248] E. Ramos-Cordoba, V. Postils, P. Salvador, *J. Chem. Theory Comput.* **2015**, *11*, 1501–1508.

[249] V. Postils, C. Delgado-Alonso, J. M. Luis, P. Salvador, *Angew. Chem.* **2018**, *130*, 10685–10689.

[250] I Mayer, *Chem. Phys. Lett.* **1995**, *242*, 499–506.

[251] I Mayer, *Can. J. Chem.* **1996**, *74*, 939–942.

[252] I. Mayer, P Salvador, *J. Chem. Phys.* **2009**, *130*, 234106.

[253] E. Ramos-Cordoba, P. Salvador, I. Mayer, *J. Chem. Phys.* **2013**, *138*, 214107.

[254] P. Pulay, T. P. Hamilton, *J. Chem. Phys.* **1988**, *88*, 4926–4933.

[255] A. Ye, J. Autschbach, *J. Chem. Phys.* **2006**, *125*, 234101.

[256] L. Gagliardi, R. Lindh, G. Karlström, *J. Chem. Phys.* **2004**, *121*, 4494–4500.

[257] K. E. Laidig, R. F. Bader, *J. Chem. Phys.* **1990**, *93*, 7213–7224.

[258] R. Bader, T. Keith, K. Gough, K. Laidig, *Mol. Phys.* **1992**, *75*, 1167–1189.

[259] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al., Gaussian 09 Revision D.01, Gaussian Inc. Wallingford CT 2009.

[260] J. W. Hollett, P. M. Gill, *J. Chem. Phys.* **2011**, *134*, 114111.

[261] E. Ramos-Cordoba, E. Matito, *J. Chem. Theory Comput.* **2017**, *13*, 2705–2711.

[262] S. Di Bella, M. A. Ratner, T. J. Marks, *J. Am. Chem. Soc.* **1992**, *114*, 5842–5849.

[263] A. Facchetti, E. Annoni, L. Beverina, M. Morone, P. Zhu, T. J. Marks, G. A. Pagani, *Nat. Mater.* **2004**, *3*, 910–917.

[264] A. Datta, S. K. Pati, *Chem. Soc. Rev.* **2006**, *35*, 1305–1323.

[265] T. Seidler, K. Stadnicka, B. Champagne, *J. Chem. Theory Comput.* **2014**, *10*, 2114–2124.

[266] M. Medved', S. Budzak, W. Bartkowiak, H. Reis in *Handbook of Computational Chemistry*, *Vol. 17*, (Eds.: J. Leszczynski, A. Kaczmarek-Kędziera, T. Puzyn, M. G. Papadopoulos, H. Reis, M. K. Shukla), Springer International Publishing, **2017**, Chapter 17.

[267] D. P. Shelton, J. E. Rice, *Chem. Rev.* **1994**, *94*, 3–29.

[268] J. López Cacheiro, B. Fernández, D. Marchesan, S. Coriani, C. Hättig, A. Rizzo, *Mol. Phys.* **2004**, *102*, 101–110.

[269] M. Pecul, F. Pawłowski, P. Jørgensen, A. Köhn, C. Hättig, *J. Chem. Phys.* **2006**, *124*, 114101.

[270] J. R. Hammond, K. Kowalski, *J. Chem. Phys.* **2009**, *130*, 194108.

[271] M. de Wergifosse, B. Champagne, *J. Chem. Phys.* **2011**, *134*, 074113.

[272] B. Skwara, W. Bartkowiak, A. Zawada, R. W. Góra, J. Leszczynski, *Chem. Phys. Lett.* **2007**, *436*, 116–123.

[273] R. W. Góra, R. Zaleśny, A. Zawada, W. Bartkowiak, B. Skwara, M. G. Papadopoulos, D. L. Silva, *J. Phys. Chem. A* **2011**, *115*, 4691–4700.

[274] D. R. Kanis, M. A. Ratner, T. J. Marks, *Chem. Rev.* **1994**, *94*, 195–242.

[275] B.-Q. Wang, Z.-R. Li, D. Wu, X.-Y. Hao, R.-J. Li, C.-C. Sun, *J. Phys. Chem. A* **2004**, *108*, 2464–2468.

[276] B. Skwara, A. Kaczmarek, R. W. Góra, W. Bartkowiak, *Chem. Phys. Lett.* **2008**, *461*, 203–206.

[277] A. Baranowska, B. Fernández, A. Rizzo, B. Jansík, *Phys. Chem. Chem. Phys.* **2009**, *11*, 9871–9883.

[278] A. Baranowska, A. Zawada, B. Fernández, W. Bartkowiak, D. Kędziera, A. Kaczmarek-Kędziera, *Phys. Chem. Chem. Phys.* **2010**, *12*, 852–862.

[279] A. Baranowska, B. Fernández, A. J. Sadlej, *Theor. Chem. Acc.* **2011**, *128*, 555–561.

[280] A. Avramopoulos, M. G. Papadopoulos, A. J. Sadlej, *J. Chem. Phys.* **2002**, *117*, 10026–10038.

[281] A. Baranowska-Łączkowska, B. Fernandez, A. Rizzo, B. Jansik, *Mol. Phys.* **2012**, *110*, 2503–2512.

[282] Z. Czyżnikowska, R. W. Góra, R. Zaleśny, W. Bartkowiak, A. Baranowska-Łączkowska, J. Leszczynski, *Chem. Phys. Lett.* **2013**, *555*, 230–234.

[283] R. W. Góra, B. Błasiak, *J. Phys. Chem. A* **2013**, *117*, 6859–6866.

[284] A. Baranowska-Łączkowska, B. Fernández, R. Zaleśny, *J. Comput. Chem.* **2013**, *34*, 275–283.

[285] M. Medved', S. Budzak, A. D. Laurent, D. Jacquemin, *J. Phys. Chem. A* **2015**, *119*, 3112–3124.

[286] O. D. Fominykh, A. V. Sharipova, M. Yu. Balakina, *Int. J. Quantum Chem.* **2016**, *116*, 103–112.

[287] R. Zaleśny, M. Garcia-Borras, R. W. Góra, J. M. Luis, et al., *Phys. Chem. Chem. Phys.* **2016**, *18*, 22467–22477.

[288] B. Kirtman, B. Champagne, J. M. Luis, *J. Comput. Chem.* **2000**, *21*, 1572–1588.

[289] M. Torrent-Sucarrat, M. Sola, M. Duran, J. M. Luis, B. Kirtman, *J. Chem. Phys.* **2004**, *120*, 6346–6355.

[290] J. M. Luis, M. Torrent-Sucarrat, O. Christiansen, B. Kirtman, *J. Chem. Phys.* **2007**, *127*, 084118.

[291] D. M. Bishop, B. Kirtman, H. A. Kurtz, J. E. Rice, *J. Chem. Phys.* **1993**, *98*, 8024–8030.

[292] M. Medved', M. Stachová, D. Jacquemin, J. M. André, E. A. Perpète, *J. Mol. Struc.-THEOCHEM* **2007**, *847*, 39–46.

[293] Y. Zhao, D. Truhlar, *Theor. Chem. Account* **2008**, *120*, 215–241.

[294] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6170.

[295] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, G. E. Scuseria, *J. Chem. Phys.* **2006**, *125*, 224106.

[296] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., Gaussian 16 Revision C.01, Gaussian Inc. Wallingford CT, **2016**.

[297] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, *Chem. Phys. Lett.* **1989**, *157*, 479–483.

[298] T. H. Dunning, *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[299] R. A. Kendall, T. H. Dunning, R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[300] A. K. Wilson, D. E. Woon, K. A. Peterson, T. H. Dunning, *J. Chem. Phys.* **1999**, *110*, 7667–7676.

[301] D. E. Woon, T. H. Dunning, *J. Chem. Phys.* **1993**, *98*, 1358–1371.

[302] D. E. Woon, T. H. Dunning, *J. Chem. Phys.* **1994**, *100*, 2975–2988.

[303] D. Feller, *J. Comput. Chem.* **1996**, *17*, 1571–1586.

[304] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, T. L. Windus, *J. Chem. Inf. Model.* **2007**, *47*, 1045–1052.

[305] B. P. Pritchard, D. Altarawy, B. Didier, T. D. Gibson, T. L. Windus, *J. Chem. Inf. Model.* **2019**, *59*, 4814–4820.

[306] D. Jacquemin, E. A. Perpète, G. Scalmani, M. J. Frisch, R. Kobayashi, C. Adamo, *J. Chem. Phys.* **2007**, *126*, 144105.

[307] E. R. Johnson, R. A. Wolkow, G. A. DiLabio, *Chem. Phys. Lett.* **2004**, *394*, 334–338.

[308] E. R. Johnson, A. D. Becke, C. D. Sherill, G. A. DiLabio, *J. Chem. Phys.* **2009**, *131*, 034111.

[309] J. Grafenstein, D. Cremer, *J. Chem. Phys.* **2007**, *127*, 164113.

[310] J. Grafenstein, D. Izotov, D. Cremer, *J. Chem. Phys.* **2007**, *127*, 214103.

[311] R. Peverati, M. Macrina, K. K. Baldridge, *J. Chem. Theory Comput.* **2010**, *6*, 1951–1965.

[312] J. Contreras-García, W. Yang, E. R. Johnson, *J. Phys. Chem. A* **2011**, *115*, 12983–12990.

[313] N. Mardirossian, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904–9924.

[314] N. Mardirossian, M. Head-Gordon, *J. Chem. Phys.* **2016**, *144*, 214110.

[315] N. Mardirossian, M. Head-Gordon, *Mol. Phys.* **2017**, *115*, 2315–2372.

[316] S. Dasgupta, J. S. Herbert, *J. Comp. Chem.* **2017**, *38*, 869–882.

[317] T. Gould, E. R. Johnson, S. A. Tawfik, *Beilstein J. Org. Chem.* **2018**, *14*, 1181–1191.

[318] P. Morgante, R. Peverati, *Int. J. Quantum Chem.* **2020**, *120*, e26332.

[319] N. Mardirossian, M. Head-Gordon, *J. Chem. Phys.* **2015**, *142*, 074111.

[320] A. N. Bootsma, S. Wheeler, *ChemRxiv* **2019**, doi:10.26434/chemrxiv.8864204.v5.

[321] P. Morgante, R. Peverati, *Int. J. Quantum Chem.* **2020**, *120*, e26332.

[322] E. Epifanovsky, A. T. B. Gilbert, X. Feng, J. Lee, Y. Mao, N. Mardirossian, P. Pokhilko, A. F. White, M. P. Coons, A. L. Dempwolff, et al., *J. Chem. Phys.* **2021**, *155*, 084801.

[323] J. F. Stanton, J. Gauss, L. Cheng, M. E. Harding, D. A. Matthews, P. G. Szalay, *CFOUR*, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package, With contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, S. Blaschke, Y. J. Bomble, S. Burger, O. Christiansen, D. Datta, F. Engel, R. Faber, J. Greiner, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, G.M. KopperW.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, D.P. O'Neill, T. Nottoli, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

[324] C. Adamo, V. Barone, *Chem. Phys. Lett.* **1997**, *274*, 242–250.

[325] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[326] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[327] A. D. Becke, *J. Chem. Phys.* **1997**, *107*, 8554–8560.

[328] S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[329] J.-D. Chai, M. Head-Gordon, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[330] Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2006**, *110*, 13126–13130.

[331] Y.-S. Lin, G.-D. Li, S.-P. Mao, J.-D. Chai, *J. Chem. Theory Comput.* **2013**, *9*, 263–272.

[332] R. Peverati, Y. Zhao, D. G. Truhlar, *J. Phys. Chem. Lett.* **2011**, *2*, 1991–1997.

[333] R. Peverati, D. G. Truhlar, *J. Chem. Phys.* **2011**, *135*, 191102.

[334] R. Peverati, D. G. Truhlar, *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.

[335] R. Peverati, D. G. Truhlar, *J. Phys. Chem. Lett.* **2012**, *3*, 117–124.

[336] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, L. A. Constantin, J. Sun, *Phys. Rev. Lett.* **2009**, *103*, 026403.

[337] R. Peverati, D. G. Truhlar, *Phys. Chem. Chem. Phys.* **2012**, *14*, 13171–13174.

[338] V. N. Staroverov, G. E. Scuseria, J. Tao, J. P. Perdew, *J. Chem. Phys.* **2003**, *119*, 12129–12137.

[339] R. Peverati, D. G. Truhlar, *Phys. Chem. Chem. Phys.* **2012**, *14*, 16187–16191.

[340] J. Heyd, G. E. Scuseria, M. Ernzerhof, *J. Chem. Phys.* **2003**, *118*, 8207–8215.

[341] J. Heyd, G. E. Scuseria, M. Ernzerhof, *J. Chem. Phys.* **2006**, *124*, 219906.

[342] K. Hui, J.-D. Chai, *J. Chem. Phys.* **2016**, *144*, 044114.

[343] O. A. Vydrov, T. Van Voorhis, *J. Chem. Phys.* **2010**, *133*, 244103.

[344] H. Takahasi, M. Mori, *Publ. Res. Inst. Math. Sci.* **1974**, *9*, 721–741.

[345] M. Mori, *Publ. Res. Inst. Math. Sci.* **2005**, *41*, 897–935.

[346] M. Mitani, *Theor. Chem. Acc.* **2011**, *130*, 645–669.

[347] A. Savitzky, M. J. Golay, *Anal. Chem.* **1964**, *36*, 1627–1639.

[348] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, et al., *Nature* **2020**, *585*, 357–362.

[349] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al., *Nat. Methods* **2020**, *17*, 261–272.

[350] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, et al., *PeerJ Comput. Sci.* **2017**, *3*, e103.

[351] J. D. Hunter, *Comput. Sci. Eng.* **2007**, *9*, 90–95.

[352] F. F. Kuo, J. F. Kaiser in *System Analysis by Digital Computer*, (Eds.: F. F. Kuo, J. F. Kaiser).

[353] J. McClellan, T Parks, *IEEE Trans. Circuit Theory* **1973**, *20*, 697–701.

[354] M. Bellanger, *Digital Processing of Signals: Theory and Practice*, New York: Wiley, **1984**.

[355] V. Barone, *J. Chem. Phys.* **2005**, *122*, 014108.

[356] J. Bloino, *J. Phys. Chem. A* **2015**, *119*, 5269–5287.

[357] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73–78.

[358] F. Neese, *WIREs Comput. Mol. Sci.* **2018**, *8*, e1327.

[359] F. Neese, F. Wennmohs, U. Becker, C. Riplinger, *J. Chem. Phys.* **2020**, *152*, 224108.

[360] N. Mardirossian, M. Head-Gordon, *Mol. Phys.* **2017**, *115*, 2315–2372.

[361] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi, S. Grimme, *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.

[362] T. Van Voorhis, G. E. Scuseria, *J. Chem. Phys.* **1998**, *109*, 400–410.

[363] R. Peverati, D. G. Truhlar, *J. Phys. Chem. Lett.* **2011**, *2*, 2810–2817.

[364] P. A. M. Dirac, *Math. Proc. Cambridge Philos. Soc.* **1930**, *26*, 376–385.

[365] J. P. Perdew, Y. Wang, *Phys. Rev. B* **1992**, *45*, 13244–13249.

[366] S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200–1211.

[367] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, C. Fiolhais, *Phys. Rev. B* **1992**, *46*, 6671–6687.

[368] C. Adamo, V. Barone, *J. Chem. Phys.* **1998**, *108*, 664–675.

[369] J. H. Van Lenthe, R. Zwaans, H. J. J. Van Dam, M. F. Guest, *J. Comput. Chem.* **2006**, *27*, 926–932.

[370] S. Lehtola, *J. Chem. Theory Comput.* **2019**, *15*, PMID: 30653322, 1593–1604.

[371] J. L. Dye, *Science* **2003**, *301*, 607–608.

[372] J. L. Dye, *Science* **1990**, *247*, 663–668.

[373] J. L. Dye, M. G. DeBacker, *Annu. Rev. Phys. Chem.* **1987**, *38*, 271–299.

[374] J. L. Dye, *Acc. Chem. Res.* **2009**, *42*, 1564–1572.

[375] V. Postils, M. Garcia-Borrás, M. Solá, J. M. Luis, E. Matito, *Chem. Commun.* **2015**, *51*, 4865–4868.

[376] J. L. Dye, J. M. Ceraso, M. Lok, B. Barnett, F. J. Tehan, *J. Am. Chem. Soc.* **1974**, *96*, 608–609.

[377] A. Ellaboudy, J. L. Dye, P. B. Smith, *J. Am. Chem. Soc.* **1983**, *105*, 6490–6491.

[378] S. B. Dawes, D. L. Ward, R. H. Huang, J. L. Dye, *J. Am. Chem. Soc.* **1986**, *108*, 3534–3535.

[379] S. Matsuishi, Y. Toda, M. Miyakawa, K. Hayashi, T. Kamiya, M. Hirano, I. Tanaka, H. Hosono, *Science* **2003**, *301*, 626–629.

[380] Y. Toda, H. Yanagi, E. Ikenaga, J. J. Kim, M. Kobata, S. Ueda, T. Kamiya, M. Hirano, K. Kobayashi, H. Hosono, *Adv. Mater.* **2007**, *19*, 3564–3569.

[381] H. Yanagi, T. Kuroda, K.-B. Kim, Y. Toda, T. Kamiya, H. Hosono, *J. Mater. Chem.* **2012**, *22*, 4278–4281.

[382] H. Buchammagari, Y. Toda, M. Hirano, H. Hosono, D. Takeuchi, K. Osakada, *Org. Lett.* **2007**, *9*, 4287–4289.

[383] H. Yanagi, K.-B. Kim, I. Koizumi, M. Kikuchi, H. Hiramatsu, M. Miyakawa, T. Kamiya, M. Hirano, H. Hosono, *J. Phys. Chem. C* **2009**, *113*, 18379–18384.

[384] S. Watanabe, T. Watanabe, K. Ito, N. Miyakawa, S. Ito, H. Hosono, S. Mikoshiba, *Sci. Technol. Adv. Mater.* **2011**.

[385] K. Hayashi, *J. Solid State Chem.* **2011**, *184*, 1428–1432.

[386] M. Kitano, Y. Inoue, Y. Yamazaki, F. Hayashi, S. Kanbara, S. Matsuishi, T. Yokoyama, S.-W. Kim, M. Hara, H. Hosono, *Nat. Chem.* **2012**, *4*, 934–940.

[387] J. Li, B. Yin, T. Fuchigami, S. Inagi, H. Hosono, S. Ito, *Electrochem. Commun.* **2012**, *17*, 52–55.

[388] S. W. Kim, T. Shimoyama, H. Hosono, *Science* **2011**, *333*, 71–74.

[389] H. Mizoguchi, M. Okunaka, M. Kitano, S. Matsuishi, T. Yokoyama, H. Hosono, *Inorg. Chem.* **2016**, *55*, 8833–8838.

[390] P. Chanhom, K. E. Fritz, L. A. Burton, J. Kloppenburg, Y. Filinchuk, A. Senyshyn, M. Wang, Z. Feng, N. Insin, J. Suntivich, et al., *J. Am. Chem. Soc.* **2019**, *141*, 10595–10598.

[391] Y. Zhang, B. Wang, Z. Xiao, Y. Lu, T. Kamiya, Y. Uwatoko, H. Kageyama, H. Hosono, *npj Quantum Mater.* **2017**, *2*, 1–7.

[392] J. Wang, K. Hanzawa, H. Hiramatsu, J. Kim, N. Umezawa, K. Iwanaka, T. Tada, H. Hosono, *J. Am. Chem. Soc.* **2017**, *139*, 15668–15680.

[393] Y. Zhang, Z. Xiao, T. Kamiya, H. Hosono, *J. Phys. Chem. Lett.* **2015**, *6*, 4966–4971.

[394] K. Lee, S. W. Kim, Y. Toda, S. Matsuishi, H. Hosono, *Nature* **2013**, *494*, 336–340.

[395] S. Y. Lee, J.-Y. Hwang, J. Park, C. N. Nandadasa, Y. Kim, J. Bang, K. Lee, K. H. Lee, Y. Zhang, Y. Ma, et al., *Nat. Commun.* **2020**, *11*, 1–8.

[396] Y. Lu, J. Wang, J. Li, J. Wu, S. Kanno, T. Tada, H. Hosono, *Phys. Rev. B* **2018**, *98*, 125128.

[397] Y. Lu, T. Tada, Y. Toda, S. Ueda, J. Wu, J. Li, K. Horiba, H. Kumigashira, Y. Zhang, H. Hosono, *Phys. Rev. B* **2017**, *95*, 125117.

[398] J. Yan, M. Ochi, H. B. Cao, B. Saparov, J. Cheng, Y Uwatoko, R. Arita, B. C. Sales, D. G. Mandrus, *J. Phys.: Condens. Matter* **2018**, *30*, 135801.

[399] S. H. Kang, J. Bang, K. Chung, C. N. Nandadasa, G. Han, S. Lee, K. H. Lee, K. Lee, Y. Ma, S. H. Oh, et al., *Sci. Adv.* **2020**, *6*, eaba7416.

[400] D. L. Druffel, K. L. Kuntz, A. H. Woomer, F. M. Alcorn, J. Hu, C. L. Donley, S. C. Warren, *J. Am. Chem. Soc.* **2016**, *138*, 16089–16094.

[401] D. L. Druffel, A. H. Woomer, K. L. Kuntz, J. T. Pawlik, S. C. Warren, *J. Mater. Chem. C* **2017**, *5*, 11196–11213.

[402] J. Wu, Y. Gong, T. Inoshita, D. C. Fredrickson, J. Wang, Y. Lu, M. Kitano, H. Hosono, *Adv. Mater.* **2017**, *29*, 1700924.

[403] T.-N. Ye, Y. Lu, J. Li, T. Nakao, H. Yang, T. Tada, M. Kitano, H. Hosono, *J. Am. Chem. Soc.* **2017**, *139*, 17089–17097.

[404] J. Wu, J. Li, Y. Gong, M. Kitano, T. Inoshita, H. Hosono, *Angew. Chem. Int. Ed.* **2019**, *58*, 825–829.

[405] J. Li, J. Wu, H. Wang, Y. Lu, T. Ye, M. Sasase, X. Wu, M. Kitano, T. Inoshita, H. Hosono, *Chem. Sci.* **2019**, *10*, 5712–5718.

[406] H. Mizoguchi, S.-W. Park, K. Kishida, M. Kitano, J. Kim, M. Sasase, T. Honda, K. Ikeda, T. Otomo, H. Hosono, *J. Am. Chem. Soc.* **2019**, *141*, 3376–3379.

[407] T.-N. Ye, Y. Lu, Z. Xiao, J. Li, T. Nakao, H. Abe, Y. Niwa, M. Kitano, T. Tada, H. Hosono, *Nat. Commun.* **2019**, *10*, 1–10.

[408] Q. Hu, R. Tan, W. Yao, Y. Cui, J. Li, W. Song, *Appl. Surf. Sci.* **2020**, *508*, 145244.

[409] Y. Lu, J. Li, T.-N. Ye, Y. Kobayashi, M. Sasase, M. Kitano, H. Hosono, *ACS Catal.* **2018**, *8*, 11054–11058.

[410] M. Miyakawa, S. W. Kim, M. Hirano, Y. Kohama, H. Kawaji, T. Atake, H. Ikegami, K. Kono, H. Hosono, *J. Am. Chem. Soc.* **2007**, *129*, 7270–7271.

[411] Y. He, *J. Alloys Compd.* **2016**, *654*, 180–184.

[412] Y. Ge, S. Guan, Y. Liu, *New J. Phys.* **2017**, *19*, 123020.

[413] W. Chen, Z.-R. Li, D. Wu, Y. Li, C.-C. Sun, F. L. Gu, *J. Am. Chem. Soc.* **2005**, *127*, 10977–10981.

[414] M. Garcia-Borrás, M. Solá, J. M. Luis, B. Kirtman, *J. Chem. Theory Comput.* **2012**, *8*, 2688–2697.

[415] J. L. Dye, *Nature* **1993**, *365*, 10–11.

[416] J. L. Dye, *Inorg. Chem.* **1997**, *36*, 3816–3826.

[417] D. J. Singh, H. Krakauer, C. Haas, W. E. Pickett, *Nature* **1993**, *365*, 39–42.

[418] Y. Toda, Y. Kubota, M. Hirano, H. Hirayama, H. Hosono, *ACS Nano* **2011**, *5*, 1907–1914.

[419] A. D. Becke, K. E. Edgecombe, *J. Chem. Phys.* **1990**, *92*, 5397–5403.

[420] B. Silvi, A. Savin, *Nature* **1994**, *371*, 683–686.

[421] E. Matito, B. Silvi, M. Duran, M. Solá, *J. Chem. Phys.* **2006**, *125*, 024301.

[422] E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen, W. Yang, *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.

[423] J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J.-P. Piquemal, D. N. Beratan, W. Yang, *J. Chem. Theory Comput.* **2011**, *7*, 625–632.

[424] H. Schmider, A. Becke, *J. Mol. Struct. THEOCHEM* **2000**, *527*, 51–61.

[425] W. Chen, Z.-R. Li, D. Wu, Y. Li, C.-C. Sun, F. L. Gu, Y. Aoki, *J. Am. Chem. Soc.* **2006**, *128*, 1072–1073.

[426] Z.-J. Li, F.-F. Wang, Z.-R. Li, H.-L. Xu, X.-R. Huang, D. Wu, W. Chen, G.-T. Yu, F. L. Gu, Y. Aoki, *Phys. Chem. Chem. Phys.* **2009**, *11*, 402–408.

[427] E. Vessally, S. Majedi, A. Hosseinian, A. Bekhradnia, *J. Mol. Model.* **2020**, *26*, 1–8.

[428] R. Saha, P. K. Chattaraj, *ACS Omega* **2018**, *3*, 17199–17211.

[429] P. Das, R. Saha, P. K. Chattaraj, *J. Comput. Chem.* **2020**, *41*, 1645–1653.

[430] Y.-F. Wang, Z.-R. Li, D. Wu, C.-C. Sun, F.-L. Gu, *J. Comput. Chem.* **2010**, *31*, 195–203.

[431] J. Zheng, X. Xu, D. G. Truhlar, *Theor. Chem. Acc.* **2011**, *128*, 295–305.

[432] G. D. Purvis III, R. J. Bartlett, *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[433] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257–2261.

[434] P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, *28*, 213–222.

[435] T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. V. R. Schleyer, *J. Comput. Chem.* **1983**, *4*, 294–301.

[436] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., Gaussian 16 Revision B.01, Gaussian Inc. Wallingford CT, **2016**.

[437] E. Cancés, R. Keriven, F. Lodier, A. Savin, *Theor. Chem. Acc.* **2004**, *111*, 373–380.

[438] O. Chalvet, R. Daudel, S. Diner, J. P. Malrieu, *Localization and Delocalization in Quantum Chemistry*, *Vol. I*, Reidel, Dordrecht, The Netherlands, **1975**.

[439] O. Chalvet, R. Daudel, S. Diner, J. P. Malrieu, *Localization and Delocalization in Quantum Chemistry*, *Vol. II*, Reidel, Dordrecht, The Netherlands, **1976**.

[440] A. Martín Pendás, E Francisco, M. A. Blanco, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1087–1092.

[441] E. Francisco, A. Martín Pendás, M. A. Blanco, *J. Chem. Phys.* **2007**, *126*, 094102.

[442] E Francisco, A. Martín Pendás, M. A. Blanco, *Chem. Phys. Chem.* **2008**, *178*, 621–634.

[443] A. Martín Pendás, E. Francisco, *Chem. Phys. Chem.* **2019**, *20*, 2722–2741.

[444] N. C. Handy, H. F. Schaefer III, *J. Chem. Phys.* **1984**, *81*, 5031–5033.

[445] A. Müller, *Phys. Lett. A* **1984**, *105*, 446–452.

[446] M. Rodríguez-Mayorga, E. Ramos-Cordoba, M. Via-Nadal, M. Piris, E. Matito, *Phys. Chem. Chem. Phys.* **2017**, *19*, 24029–24041.

[447] M. Rodríguez-Mayorga, E. Ramos-Cordoba, F. Feixas, E. Matito, *Phys. Chem. Chem. Phys.* **2017**, *19*, 4522–4529.

[448] F. Feixas, E. Matito, M. Duran, M. Solá, B. Silvi, *J. Chem. Theory Comput.* **2010**, *6*, 2736–2742.

[449] I. Ruiz, E. Matito, F. J. Holguín-Gallego, E. Francisco, Á. Martín Pendás, T. Rocha-Rinza, *Theor. Chem. Acc.* **2016**, *135*, 1–11.

[450] E. Matito, M. Solà, P. Salvador, M. Duran, *Faraday Discuss.* **2007**, *135*, 325–345.

[451] F. Feixas, M. Solá, J. M. Barroso, J. M. Ugalde, E. Matito, *J. Chem. Theory Comput.* **2014**, *10*, 3055–3065.

[452] T. A. Keith, *AIMAll*, Version 14.11.23, TK Gristmill Software, Overland Park KS, USA (aim.tkgristmill.com), **2014**.

[453] P. Salvador, E. Ramos-Cordoba, *APOST-3D* program, Universitat de Girona (Spain), **2012**.

[454] J. Sancho-García, A. Pérez-Jiménez, *Phys. Chem. Chem. Phys.* **2007**, *9*, 5874–5879.

[455] I. Casademont-Reig, T. Woller, J. Contreras-García, M. Alonso, M. Torrent-Sucarrat, E. Matito, *Phys. Chem. Chem. Phys.* **2018**, *20*, 2787–2796.

[456] I. Casademont-Reig, E. Ramos-Cordoba, M. Torrent-Sucarrat, E. Matito in *Aromaticity*, Elsevier, **2021**, pp. 235–259.

[457] M. Torrent-Sucarrat, S. Navarro, F. P. Cossío, J. M. Anglada, J. M. Luis, *J. Comput. Chem.* **2017**, *38*, 2819–2828.

Appendix

A1 Supplementary Figures: Chapter 9

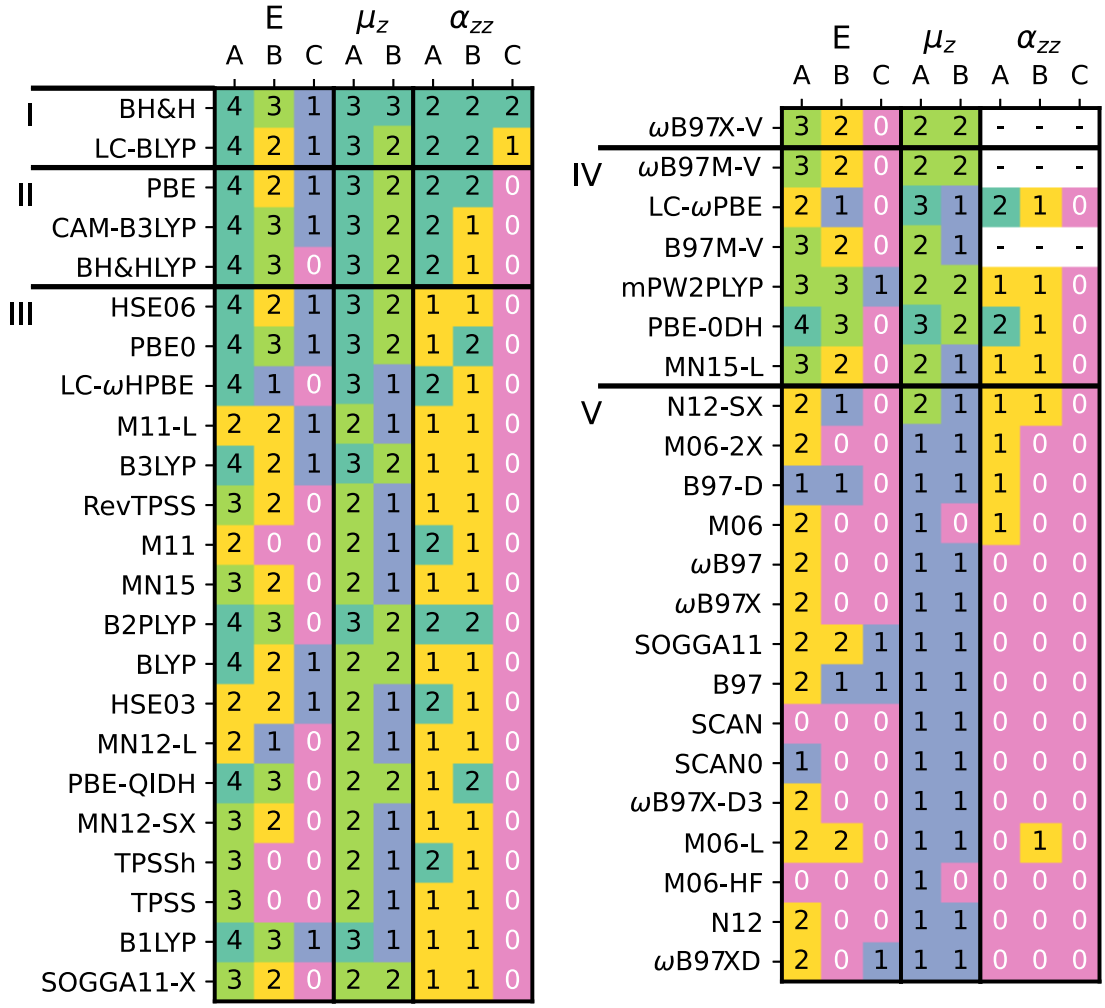

Figure A1.1: Maximum order of derivatives that can be safely obtained (i.e., RRMSE$\leq$ 10%) with the (99, 590) grid. The maximum order of the derivatives computed is fourth for the energy, third for the dipole moment, and second for the polarizability. Groups A, B, and C, correspond to hydrogen-, halogen-, and dispersion-bonded complexes, respectively. Data for $\mathrm{d}^m\alpha_{zz}/\mathrm{d}\xi^m$ is not available for ωB97X-V, ωB97M-V, and B97M-V, for which static polarizabilities are not implemented in QChem 5.1. The rungs maintain the order obtained from the (250, 974) grid.

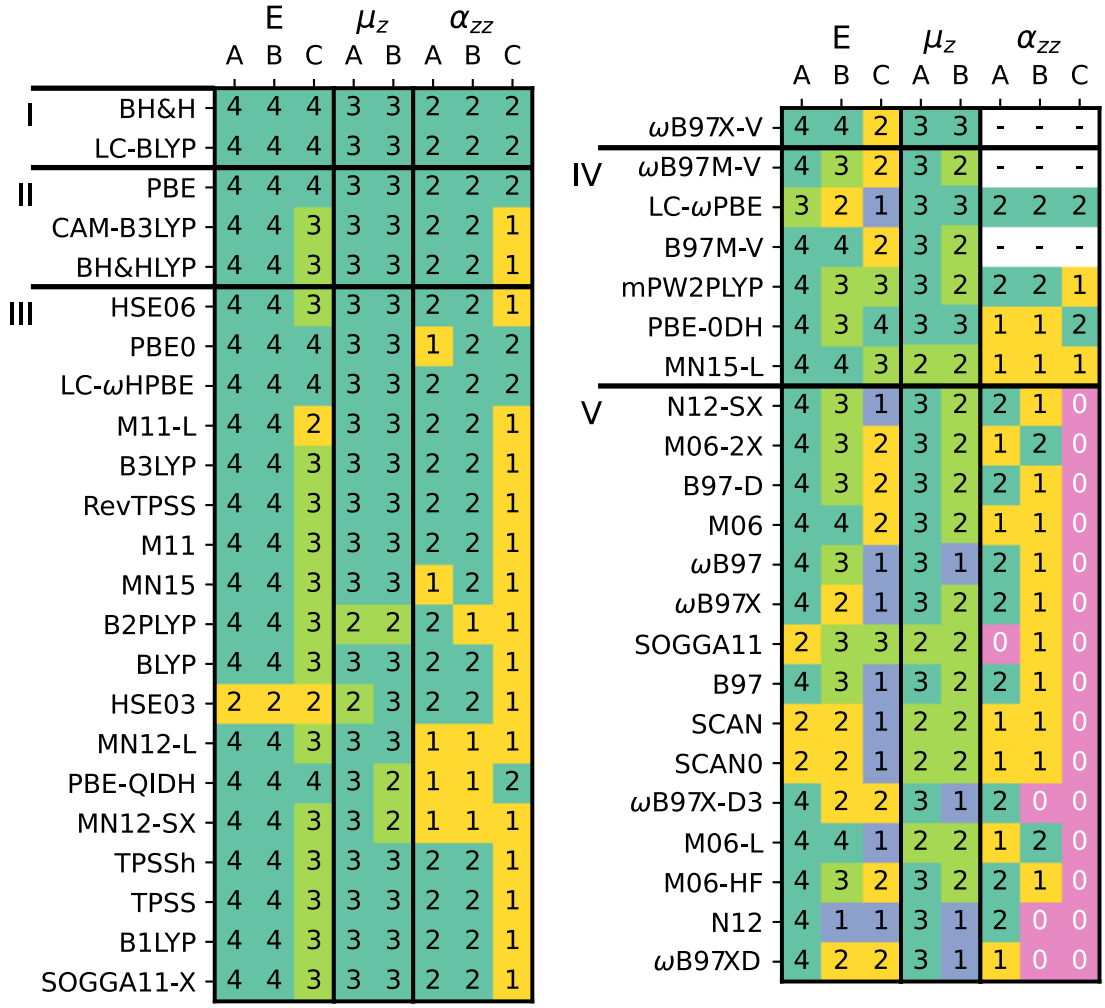

Figure A1.2: Maximum order of derivatives that can be safely obtained (i.e., RRMSE$\leq$ 10%) with the (750, 974) grid. The maximum order of the derivatives computed is fourth for the energy, third for the dipole moment, and second for the polarizability. Groups A, B, and C, correspond to hydrogen-, halogen-, and dispersion-bonded complexes, respectively. Data for $\mathrm{d}^m\alpha_{zz}/\mathrm{d}\xi^m$ is not available for ωB97X-V, ωB97M-V, and B97M-V, for which static polarizabilities are not implemented in QChem 5.1. The rungs maintain the order obtained from the (250, 974) grid.

A2 Supplementary Figures: Chapter 10

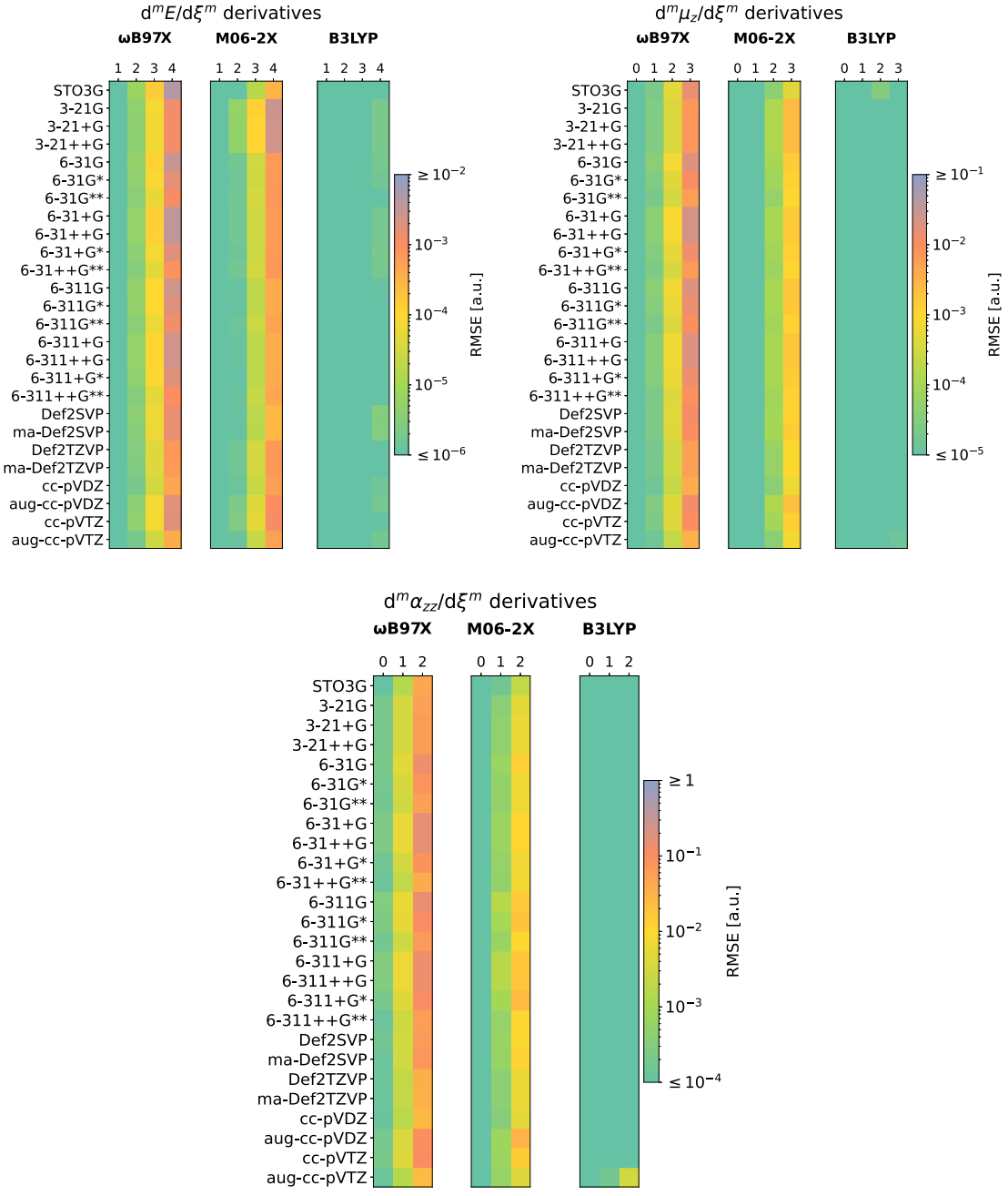

Figure A2.3: Study of the basis set dependence of the grid-related errors in the $d^m E/d\chi_{1,z}{}^m$ derivatives (*left panel*, m= 1 - 4, stored column-wise), the $d^m \mu_z/d\xi^m$ derivatives (*middle panel*, m = 0 - 3, stored column-wise), and the $d^m \alpha_{zz}/d\xi^m$ derivatives (*right panel*, m = 0 - 2, stored column-wise) of HCN · HF using ωB97X, M06-2X, and B3LYP combined with the (250, 974) integration grid. Colors reflect the values of the errors in the property measured using RMSE (note the logarithmic scale). Raw data compiled in Tables A3.1 to A3.6.

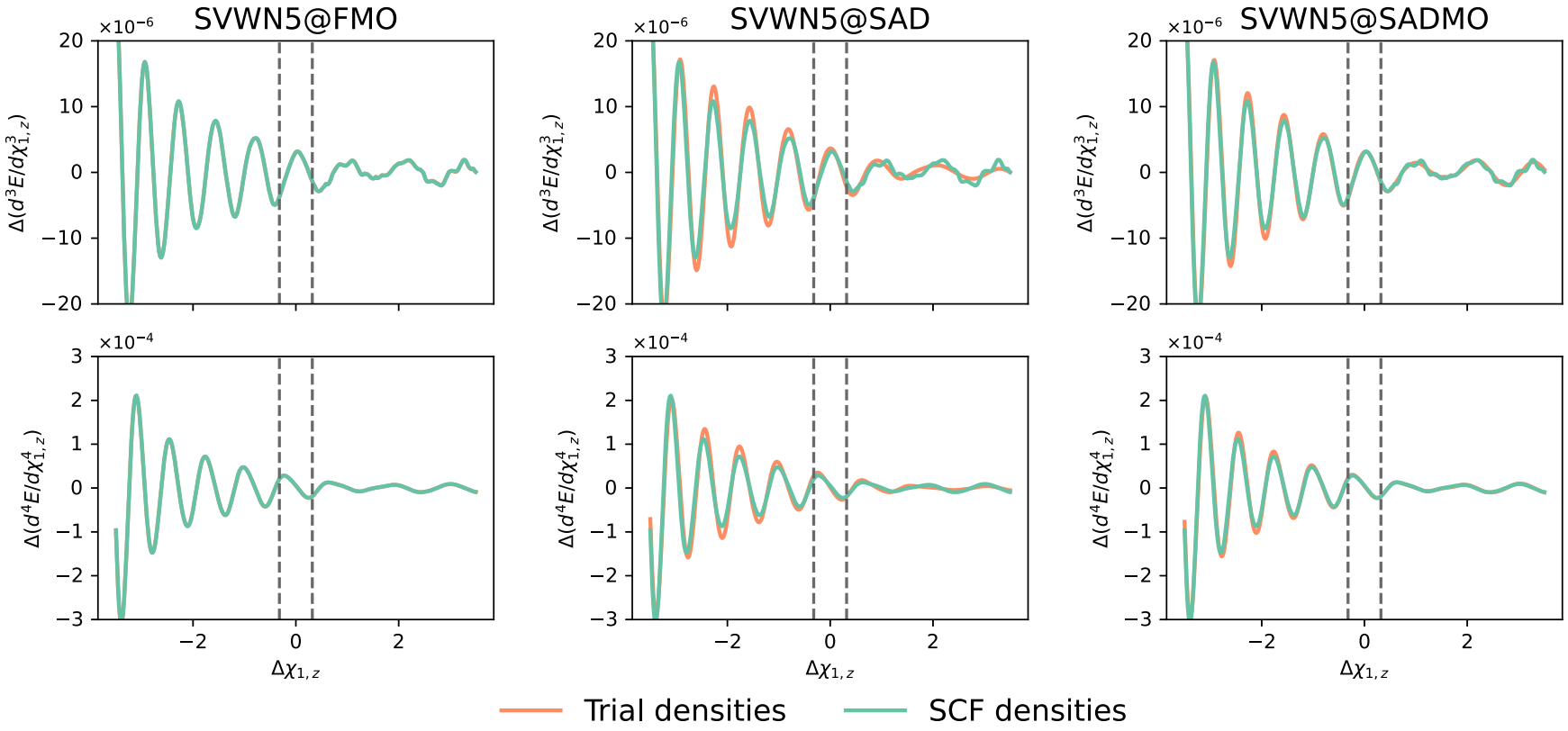

Figure A2.4: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}{\chi_{1,z}}^3$ and $\mathrm{d}^4E/\mathrm{d}{\chi_{1,z}}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with SVWN5/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

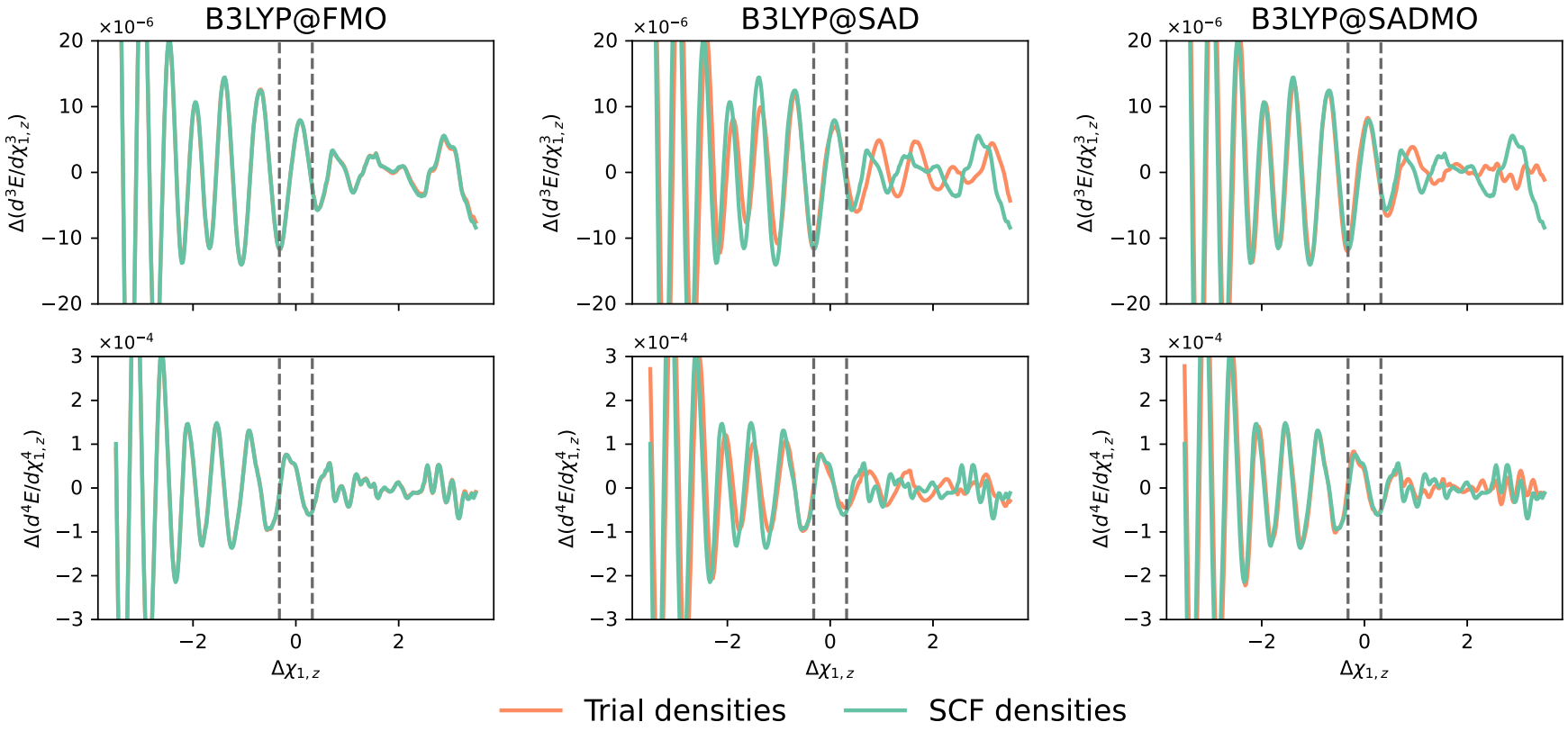

Figure A2.5: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with B3LYP/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

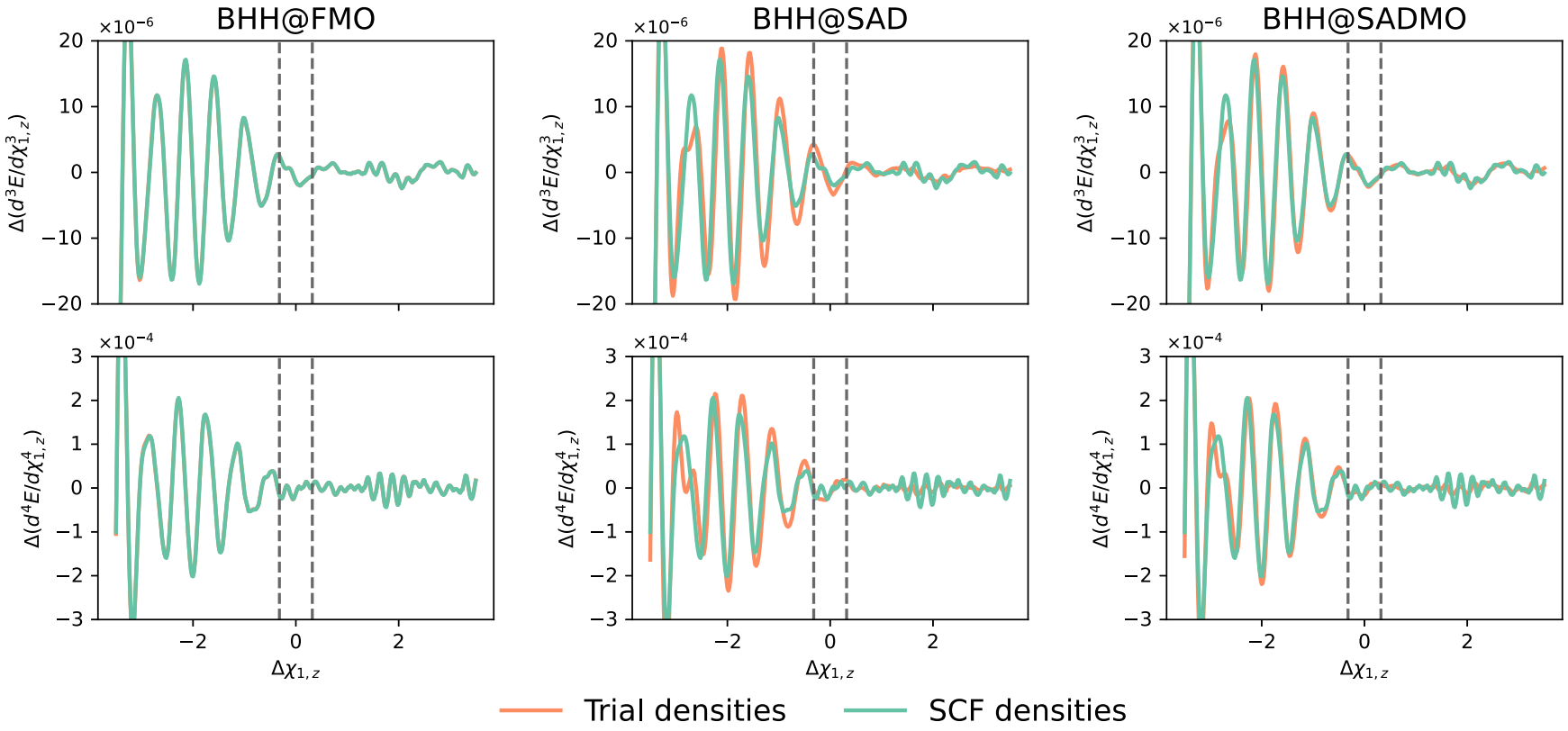

Figure A2.6: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with BH&H/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

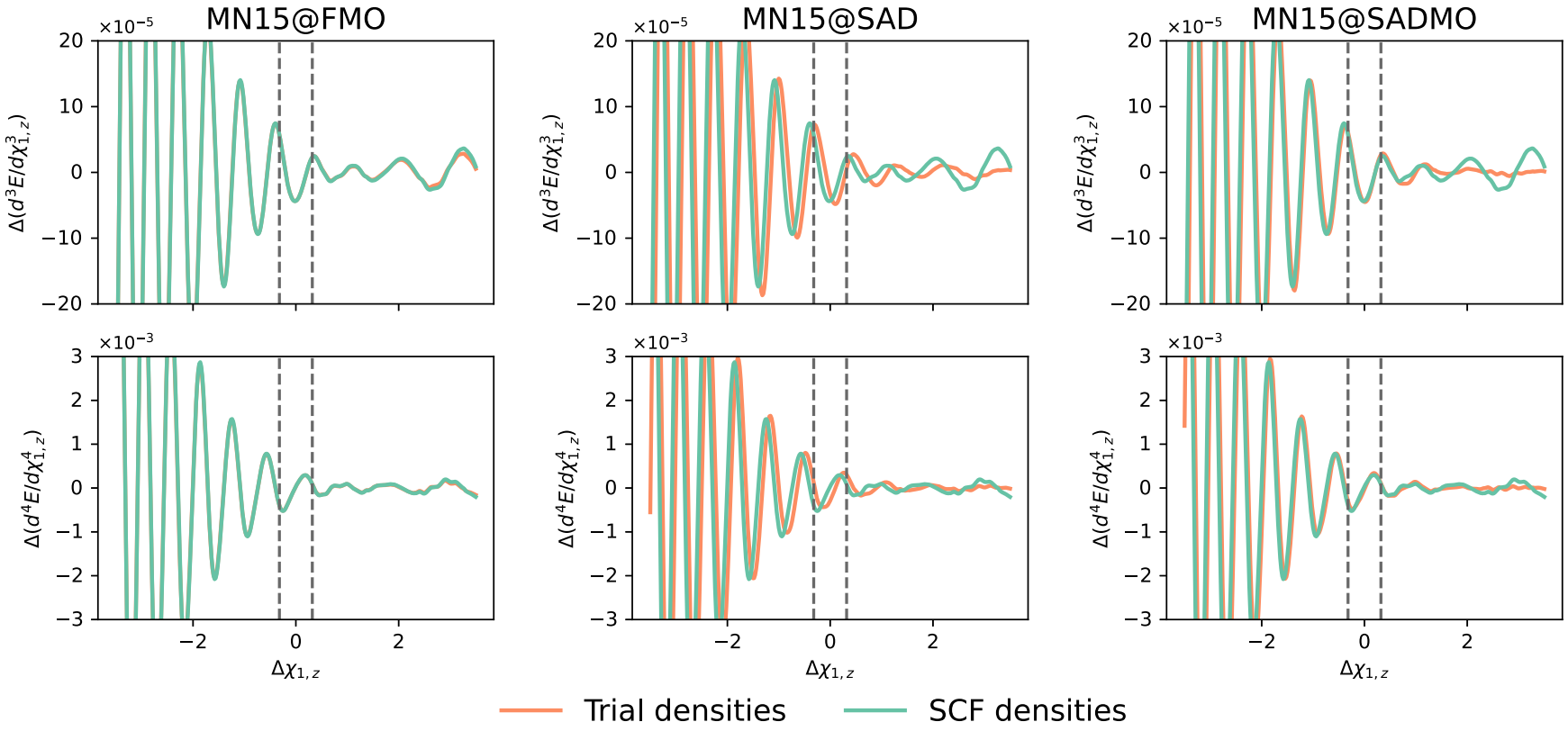

Figure A2.7: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with MN15/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

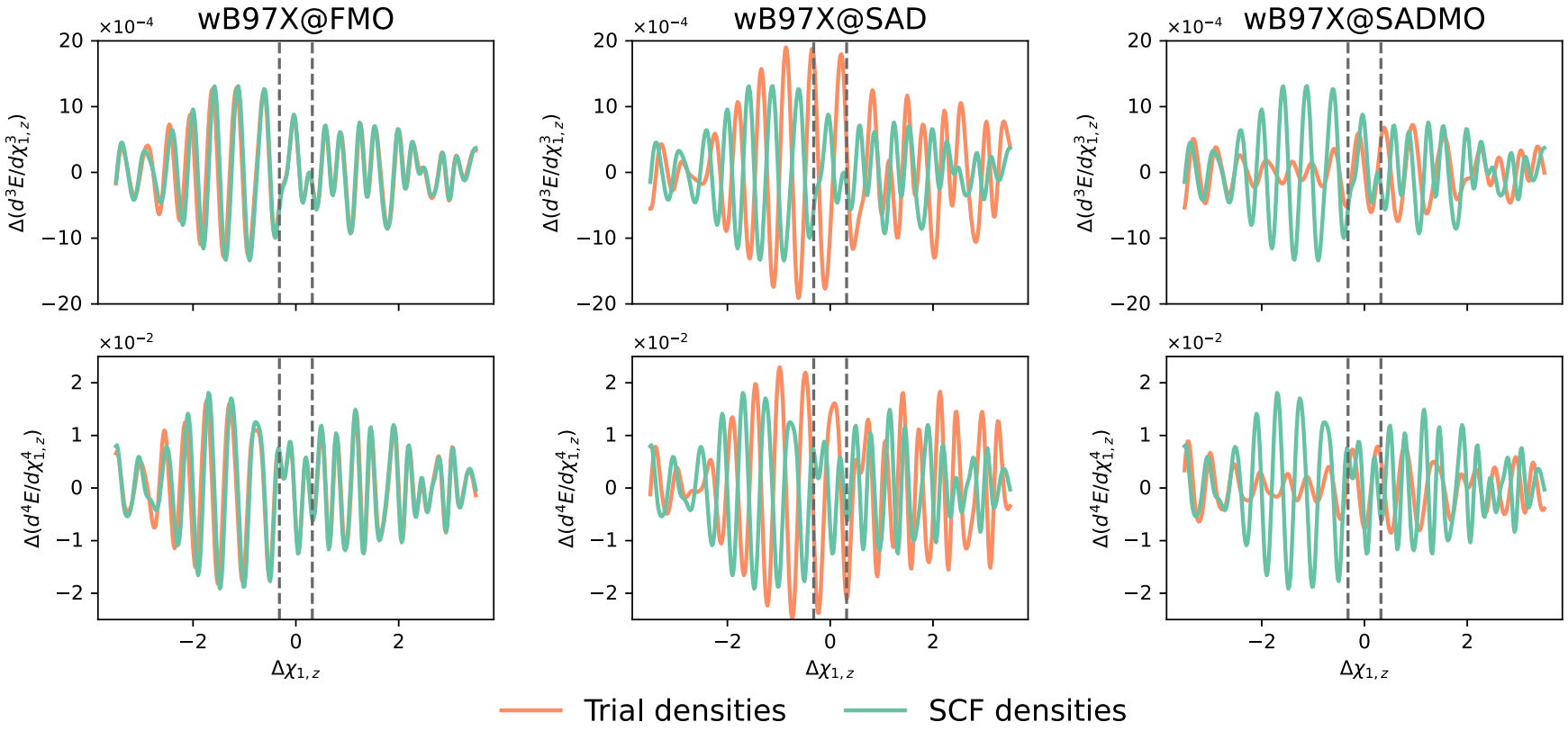

Figure A2.8: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with ωB97X/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

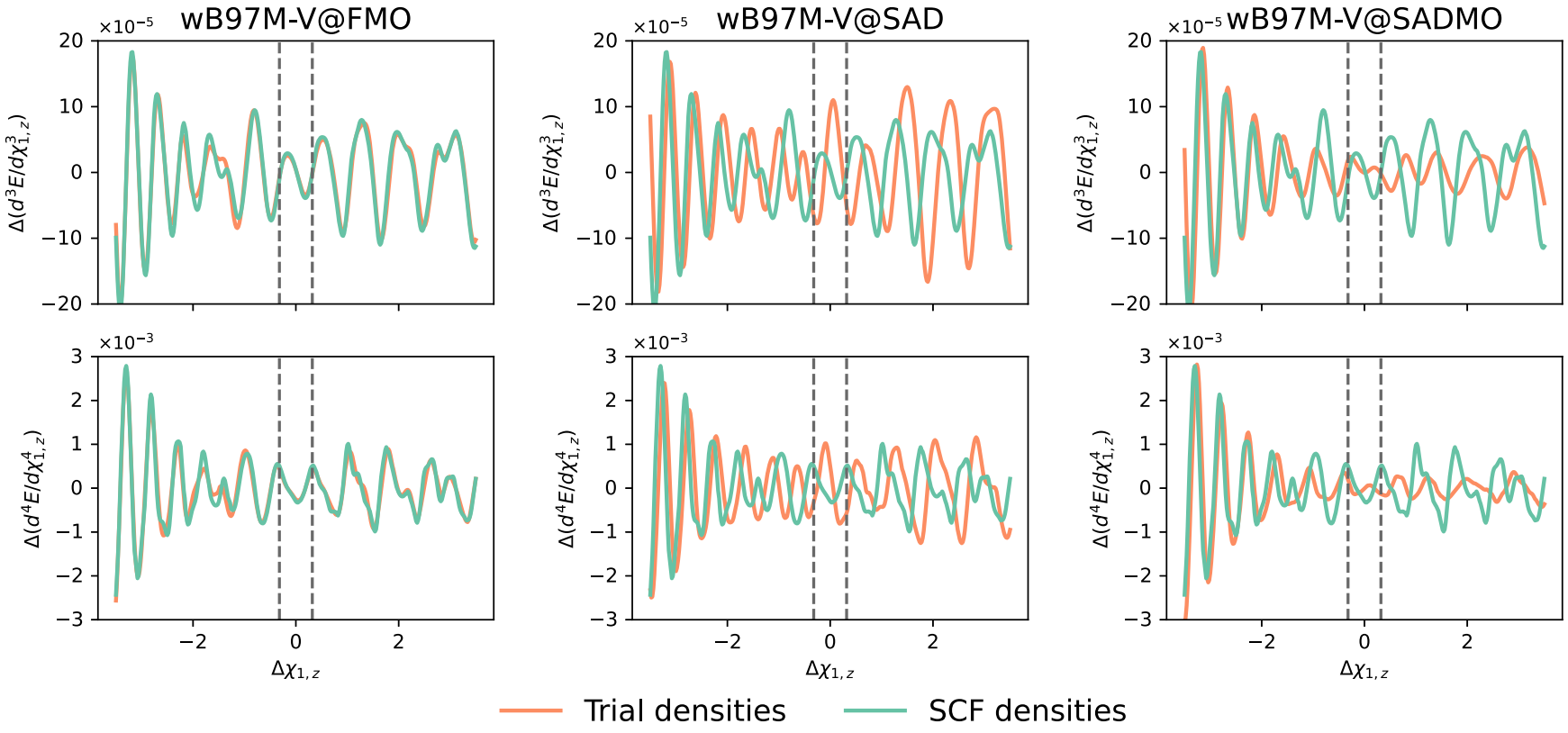

Figure A2.9: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with ωB97M-V/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

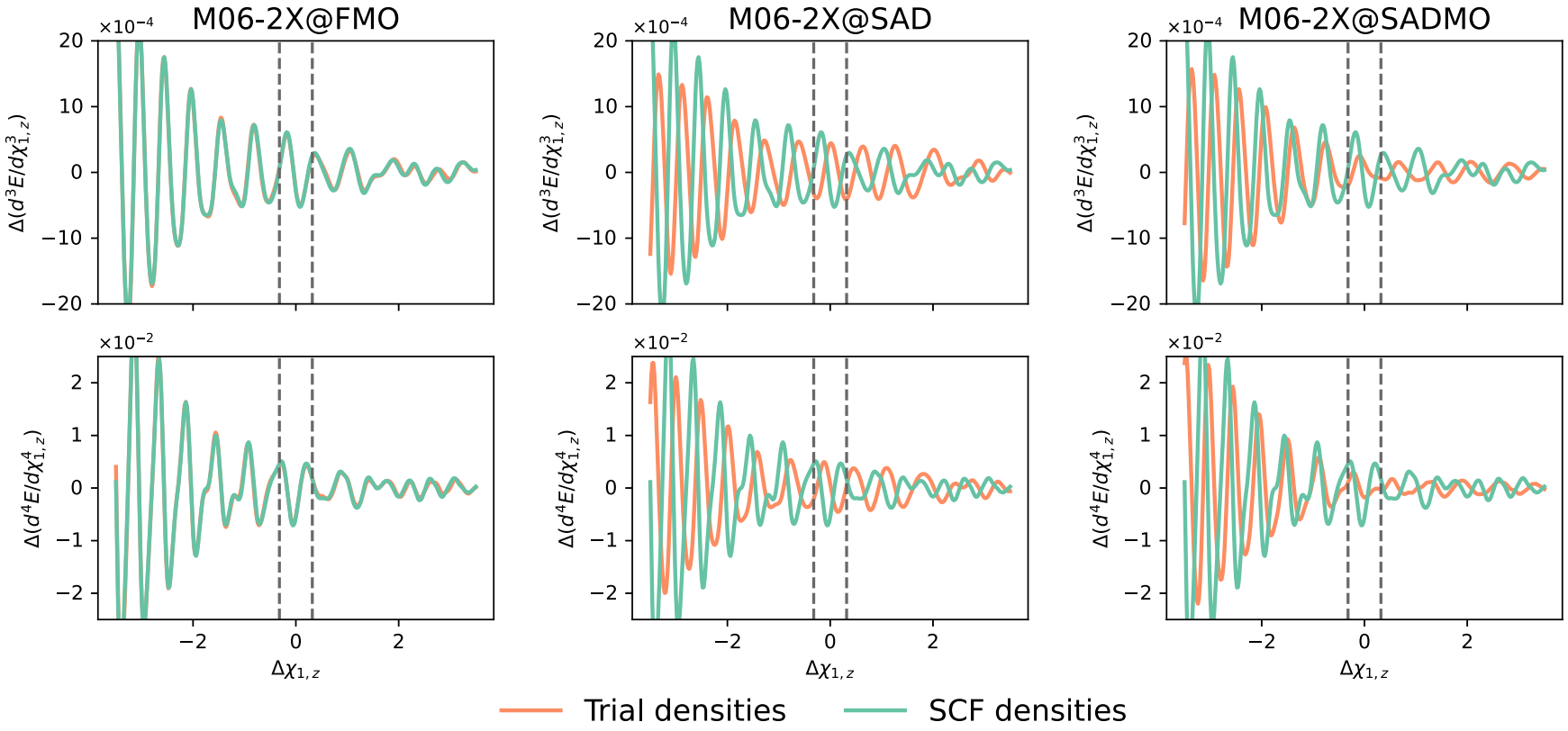

Figure A2.10: Spurious oscillations in $\mathrm{d}^3E/\mathrm{d}\chi_{1,z}{}^3$ and $\mathrm{d}^4E/\mathrm{d}\chi_{1,z}{}^4$ of the HCN · HF system, defined as $\Delta P = P^{\mathrm{DFA}(99,590)} - P^{\mathrm{DFA}(1500,974)}$, obtained with M06-2X/6-31+G* and the trial densities (orange curves) and the SCF density (green curves, shown on each plot).

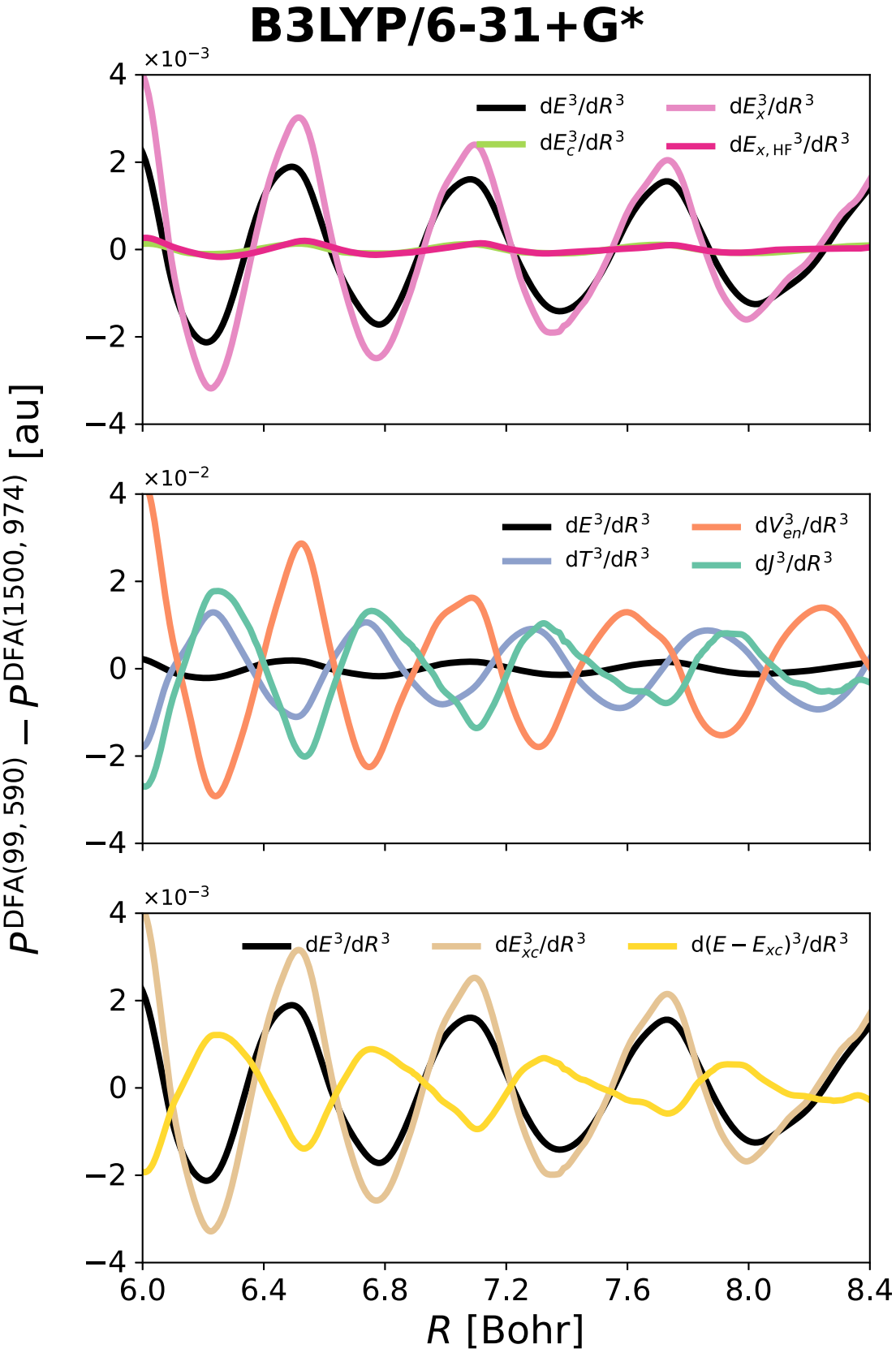

Figure A2.11: Grid-dependent spurious oscillations in the third derivatives of energy components of Ar_2: total electronic (E), kinetic (T), nuclei attraction (V_{en}), Coulomb repulsion (J), exact HF-like exchange $E_{x,\mathrm{HF}}$, exchange (E_x), and correlation (E_c), as well as their sums. Obtained with B3LYP/6-31+G* as the difference between (99, 590) and (1500, 974) integration grids.

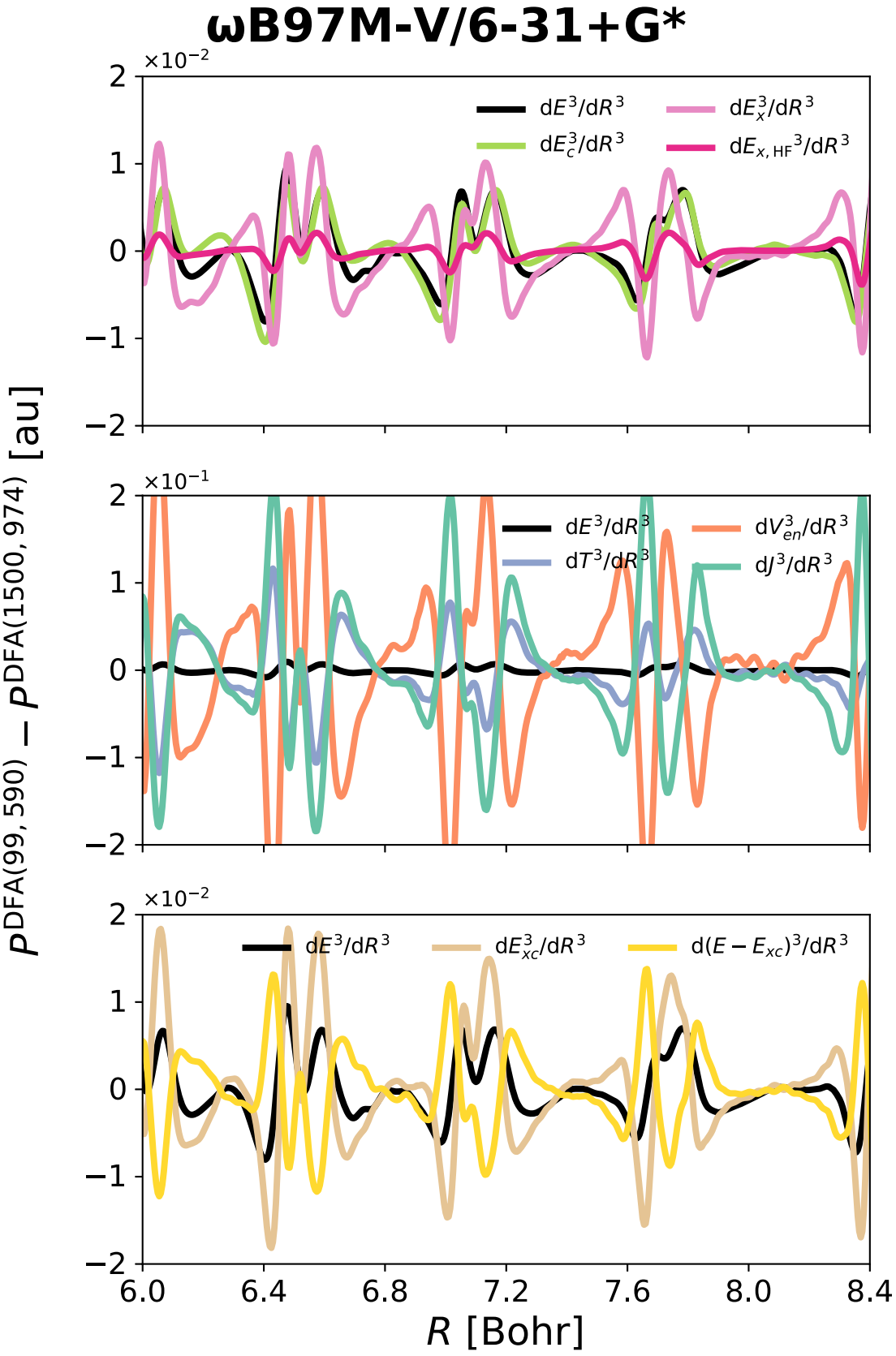

Figure A2.12: Grid-dependent spurious oscillations in the third derivatives of energy components of Ar_2: total electronic (E), kinetic (T), nuclei attraction (V_{en}), Coulomb repulsion (J), exact HF-like exchange $E_{x,\mathrm{HF}}$, exchange (E_x), and correlation (E_c), as well as their sums. Obtained with ωB97M-V/6-31+G* as the difference between (99, 590) and (1500, 974) integration grids.

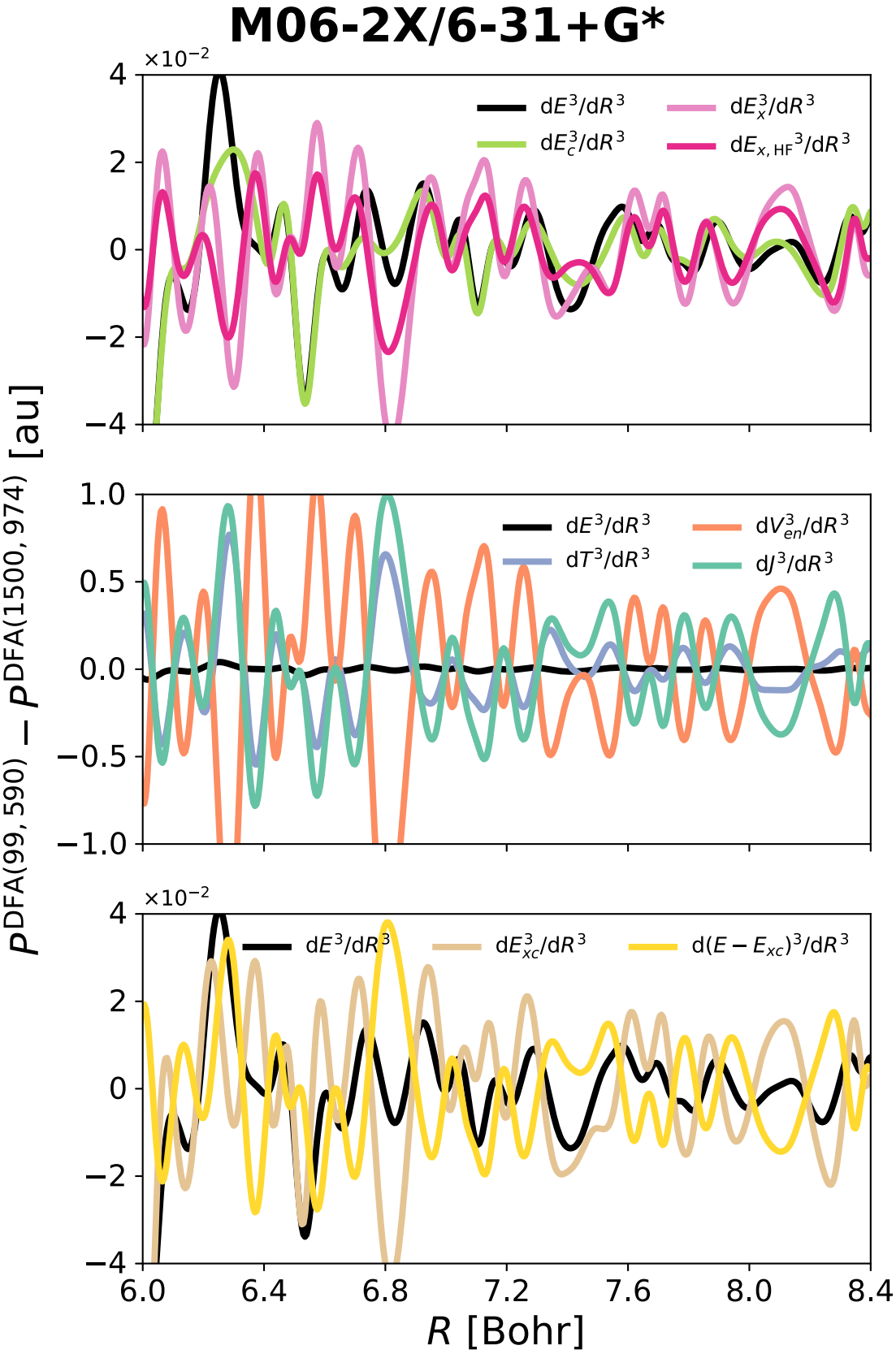

Figure A2.13: Grid-dependent spurious oscillations in the third derivatives of energy components of Ar_2: total electronic (E), kinetic (T), nuclei attraction (V_{en}), Coulomb repulsion (J), exact HF-like exchange $E_{x,\mathrm{HF}}$, exchange (E_x), and correlation (E_c), as well as their sums. Obtained with M06-2X/6-31+G* as the difference between (99, 590) and (1500, 974) integration grids.

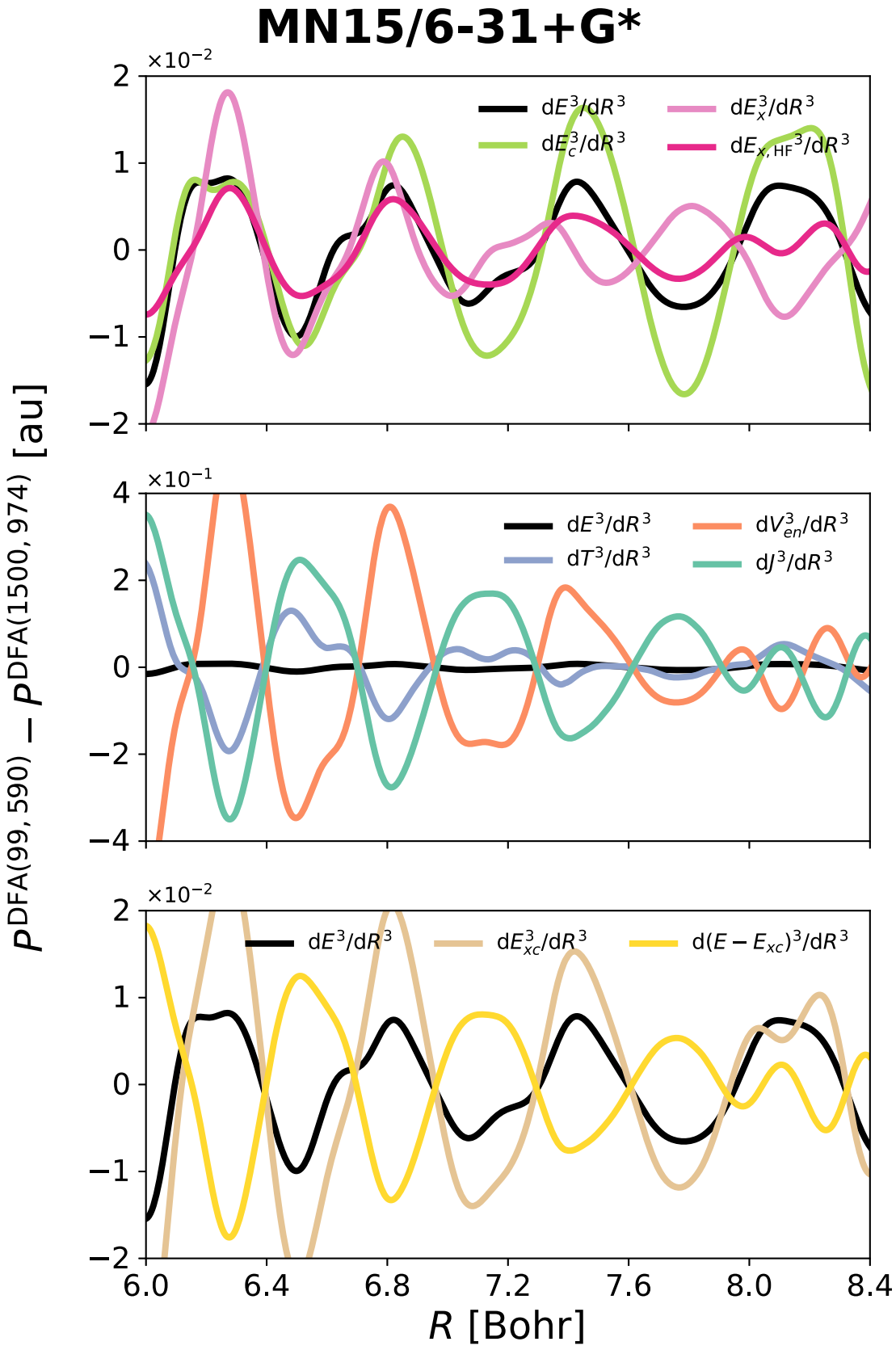

Figure A2.14: Grid-dependent spurious oscillations in the third derivatives of energy components of Ar_2: total electronic (E), kinetic (T), nuclei attraction (V_{en}), Coulomb repulsion (J), exact HF-like exchange $E_{x,\mathrm{HF}}$, exchange (E_x), and correlation (E_c), as well as their sums. Obtained with MN15/6-31+G* as the difference between (99, 590) and (1500, 974) integration grids.

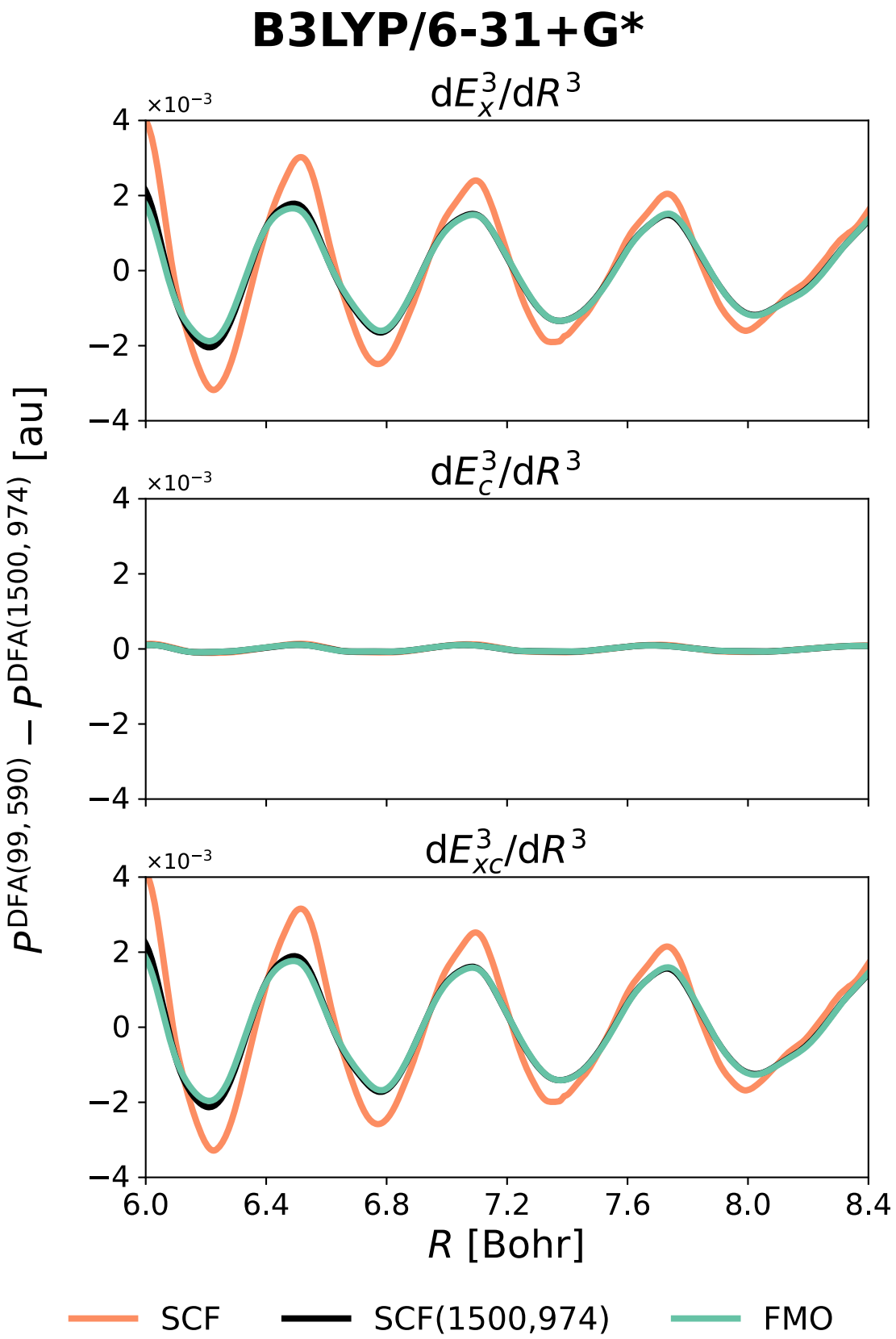

Figure A2.15: Spurious oscillations in the third derivatives of exchange (E_x), correlation (E_c) energy components of Ar_2, as well as their sum, obtained using B3LYP/6-31+G* and the (99, 590) and (1500, 974) integration grids. Three different types of the input densities are employed: SCF — density obtained in SCF using corresponding grids, SCF(1500, 974) — density obtained in SCF using the reference (1500, 974) grid, FMO — trial density constructed using FMOs obtained for one geometry (R=7.2 Bohr) and the (250, 974) integration grid.

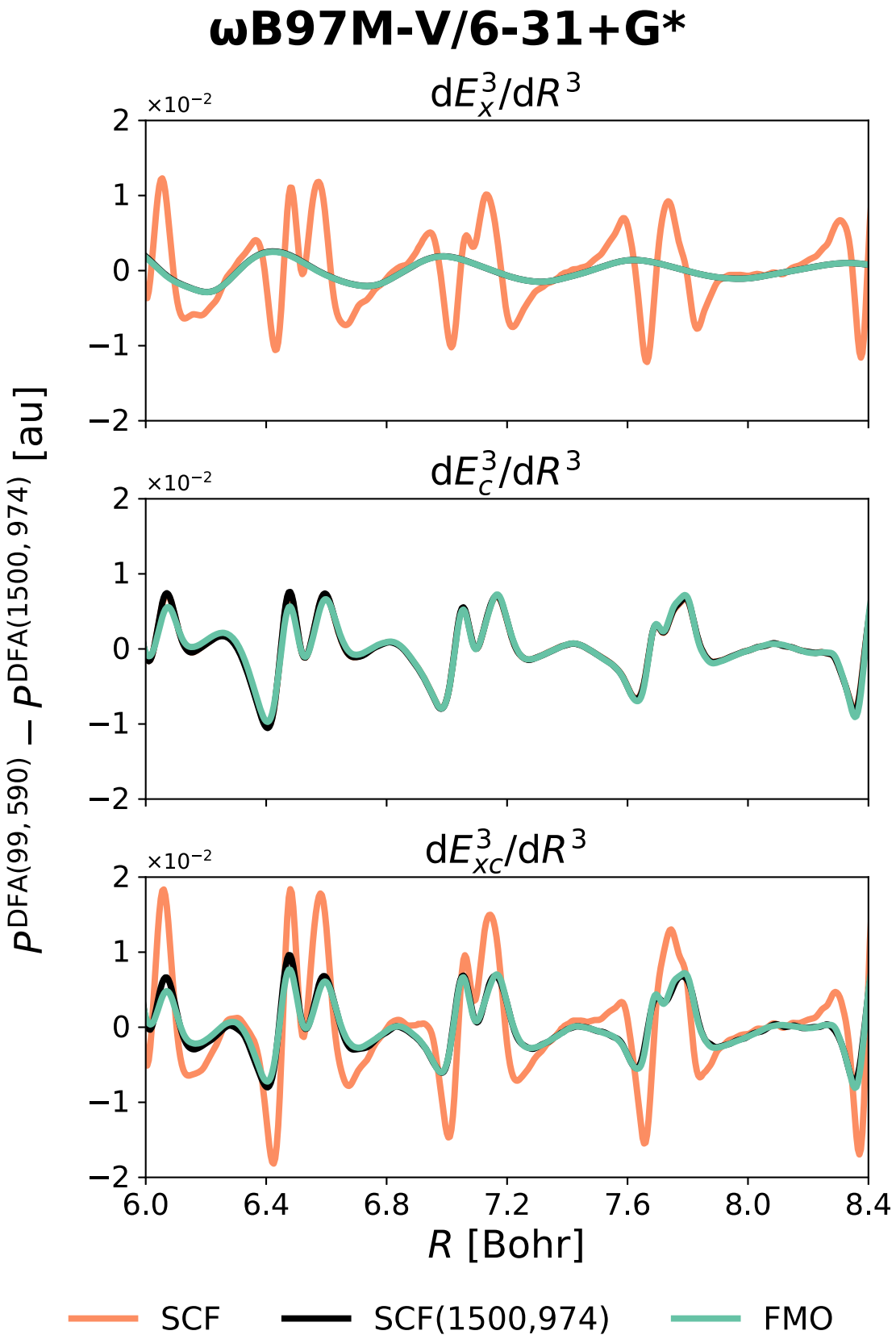

Figure A2.16: Spurious oscillations in the third derivatives of exchange (E_x), correlation (E_c) energy components of Ar_2, as well as their sum, obtained using ωB97M-V/6-31+G* and the (99, 590) and (1500, 974) integration grids. Three different types of the input densities are employed: SCF — density obtained in SCF using corresponding grids, SCF(1500, 974) — density obtained in SCF using the reference (1500, 974) grid, FMO — trial density constructed using FMOs obtained for one geometry (R=7.2 Bohr) and the (250, 974) integration grid.

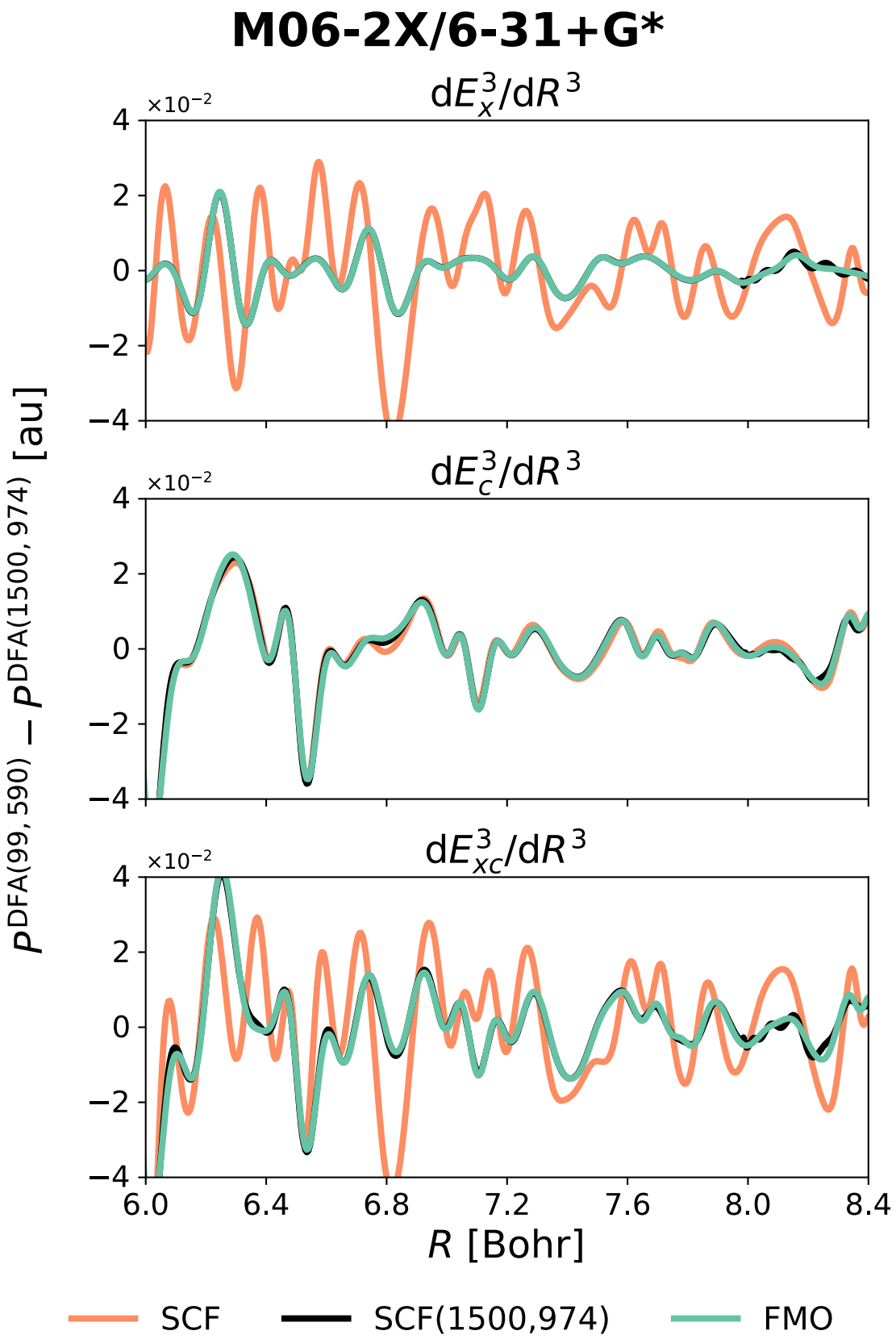

Figure A2.17: Spurious oscillations in the third derivatives of exchange (E_x), correlation (E_c) energy components of Ar_2, as well as their sum, obtained using M06-2X/6-31+G* and the (99, 590) and (1500, 974) integration grids. Three different types of the input densities are employed: SCF — density obtained in SCF using corresponding grids, SCF(1500, 974) — density obtained in SCF using the reference (1500, 974) grid, FMO — trial density constructed using FMOs obtained for one geometry (R=7.2 Bohr) and the (250, 974) integration grid.

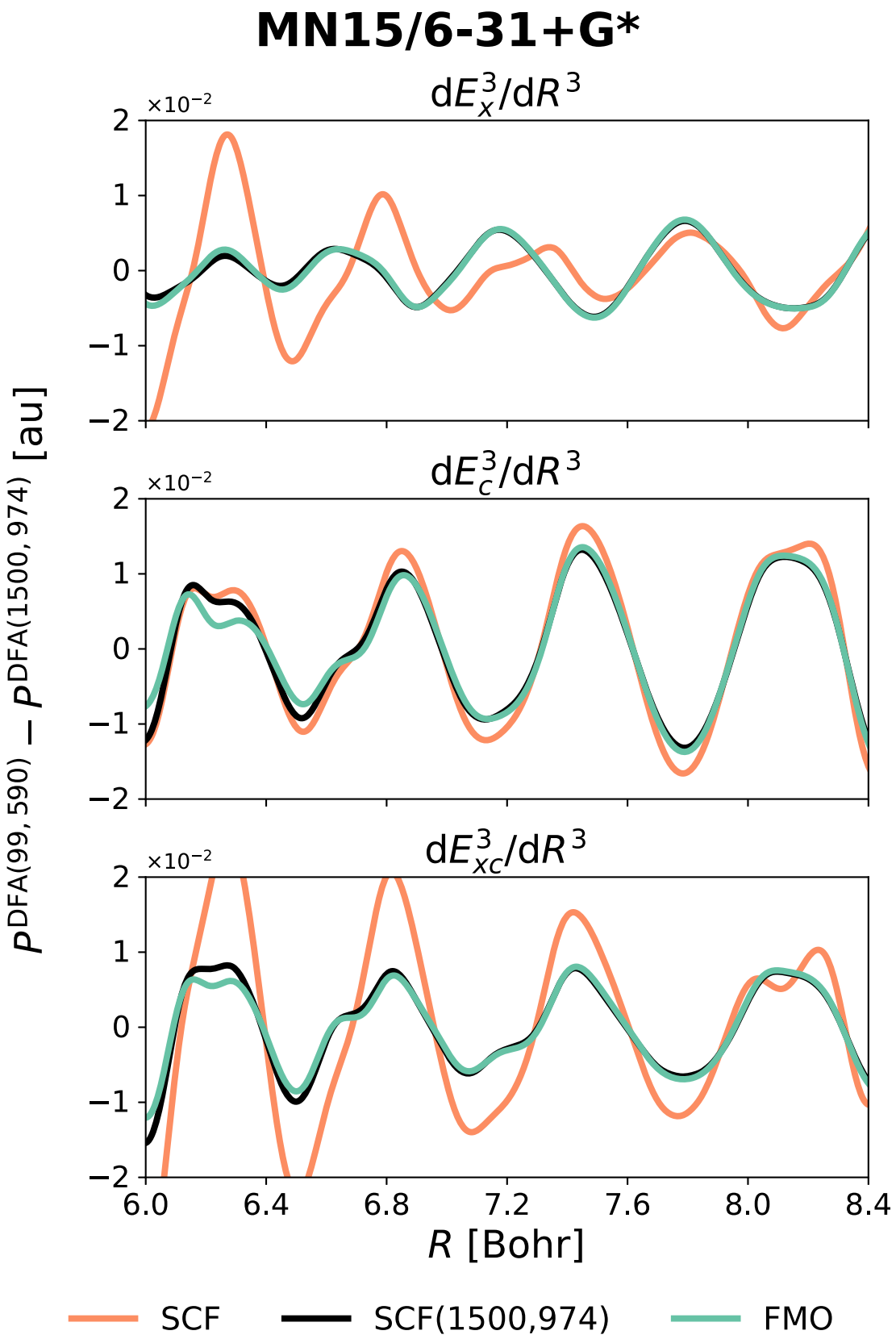

Figure A2.18: Spurious oscillations in the third derivatives of exchange (E_x), correlation (E_c) energy components of Ar_2, as well as their sum, obtained using MN15/6-31+G* and the (99, 590) and (1500, 974) integration grids. Three different types of the input densities are employed: SCF — density obtained in SCF using corresponding grids, SCF(1500, 974) — density obtained in SCF using the reference (1500, 974) grid, FMO — trial density constructed using FMOs obtained for one geometry (R=7.2 Bohr) and the (250, 974) integration grid.

A3 Supplementary Tables: Chapter 10

Table A3.1: Grid errors for various property derivatives obtained with different basis sets and ωB97X combined with the (99, 590) integration grid for the HCN · HF system. For each column, the values in cells have to be multiplied with the exponent given in top.

	$d^m E/d\chi_{1,z}{}^m$				$d^m \mu_z/d\chi_{1,z}{}^m$				$d^m \alpha_{zz}/d\chi_{1,z}{}^m$		
	1	2	3	4	0	1	2	3	0	1	2
Basis set	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]	[$\times 10^{-2}$]	[$\times 10^{-3}$]	[$\times 10^{-2}$]	[$\times 10^{-1}$]
STO-3G	5.7	5.9	8.8	15.8	1.5	1.6	2.4	4.6	0.3	0.6	1.3
3-21G	4.5	4.0	4.5	5.2	2.9	3.1	3.2	3.6	2.1	2.3	2.9
3-21+G	4.3	4.1	4.6	5.6	2.7	3.0	3.2	3.7	2.0	2.3	3.1
3-21++G	4.4	4.1	4.7	5.6	2.7	3.0	3.2	3.7	2.0	2.3	3.1
6-31G	4.4	4.6	5.9	9.5	1.7	2.1	3.1	5.4	1.3	2.2	3.9
6-31G*	4.0	4.3	4.3	6.3	2.2	2.6	2.8	4.6	1.1	1.3	2.4
6-31G**	4.6	4.8	4.7	6.5	2.8	3.4	3.3	4.6	2.3	2.3	3.4
6-31+G	4.4	4.7	6.3	10.5	1.7	2.1	3.4	6.0	1.3	2.3	4.4
6-31++G	4.4	4.7	6.3	10.5	1.7	2.1	3.3	5.9	1.3	2.3	4.5
6-31+G*	4.0	4.1	4.2	6.3	2.2	2.4	2.8	4.9	1.0	1.3	2.6
6-31++G**	4.4	4.4	4.8	5.4	3.0	3.0	3.5	4.2	2.1	2.4	2.7
6-311G	4.2	3.9	4.9	7.9	2.2	2.3	2.9	4.8	1.8	2.6	4.7
6-311G*	3.7	3.8	4.1	6.7	2.3	2.6	2.7	4.4	1.7	2.1	3.9
6-311G**	3.9	4.1	4.0	6.0	2.5	2.8	3.0	4.5	1.9	1.9	2.9
6-311+G	3.9	4.1	4.7	8.5	2.1	2.4	2.7	5.2	1.5	2.1	4.5
6-311++G	3.9	4.1	4.7	8.5	2.1	2.4	2.7	5.2	1.5	2.1	4.5
6-311+G*	3.5	3.8	3.5	5.6	2.3	2.6	2.9	5.1	1.3	1.5	2.9
6-311++G**	3.9	4.0	4.2	5.2	2.4	2.7	3.2	4.4	1.8	2.0	2.6
Def2SVP	4.3	4.5	4.7	6.8	2.4	2.8	3.3	5.1	1.7	1.9	2.8
Def2TZVP	4.0	4.2	4.3	5.3	2.6	3.0	3.3	4.2	2.2	2.3	2.7
ma-Def2SVP	4.2	4.1	5.0	6.2	2.3	2.4	3.2	4.3	1.6	2.0	3.0
ma-Def2TZVP	4.0	4.1	4.3	5.1	2.7	3.0	3.3	4.1	2.2	2.4	2.8
cc-pVTZ	4.4	4.6	4.9	7.9	2.6	2.9	3.0	4.6	1.9	2.1	3.9
cc-pVDZ	4.2	4.4	5.2	6.0	3.0	3.3	3.9	4.5	2.6	3.1	3.5
aug-cc-pVDZ	3.8	3.8	4.2	6.1	2.2	2.2	2.8	4.1	1.7	2.0	3.6
aug-cc-pVTZ	4.0	4.1	5.2	5.6	2.9	3.0	3.8	3.9	2.3	3.0	3.1

Table A3.2: Grid errors for various property derivatives obtained with different basis sets and M06-2X combined with the (99, 590) integration grid for the HCN · HF system. For each column, the values in cells have to be multiplied with the exponent given in top.

	$d^m E/d\chi_{1,z}{}^m$				$d^m \mu_z/d\chi_{1,z}{}^m$				$d^m \alpha_{zz}/d\chi_{1,z}{}^m$		
	1	2	3	4	0	1	2	3	0	1	2
Basis set	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]	[$\times 10^{-2}$]	[$\times 10^{-3}$]	[$\times 10^{-2}$]	[$\times 10^{-1}$]
STO-3G	4.7	4.0	4.3	4.3	1.1	1.2	1.5	1.6	0.5	0.5	0.6
3-21G	1.9	2.0	2.8	4.4	1.7	1.5	1.4	1.4	1.4	1.1	1.2
3-21+G	2.2	1.9	2.6	4.1	1.7	1.5	1.5	1.5	1.5	1.3	1.3
3-21++G	2.1	1.9	2.6	4.1	1.7	1.5	1.5	1.5	1.5	1.3	1.3
6-31G	2.4	1.6	2.2	2.3	3.0	2.2	2.5	2.4	1.3	1.4	1.4
6-31G*	3.3	2.8	3.1	3.6	2.2	1.7	1.9	2.0	1.5	1.5	1.6
6-31G**	2.8	2.2	2.1	2.2	2.1	1.7	1.7	1.6	1.2	1.3	1.4
6-31+G	2.7	1.8	2.5	2.6	3.1	2.3	2.7	2.6	1.4	1.6	1.5
6-31++G	2.7	1.8	2.5	2.6	3.1	2.3	2.7	2.6	1.4	1.6	1.5
6-31+G*	3.9	3.6	3.7	4.3	2.2	1.7	1.8	2.0	1.7	1.6	1.8
6-31++G**	2.0	1.6	1.3	1.6	1.7	1.4	1.2	1.4	1.5	1.4	1.6
6-311G	1.4	1.3	1.8	1.8	2.8	2.4	2.6	2.5	1.8	1.7	1.7
6-311G*	2.3	2.1	2.3	2.4	2.4	2.1	2.2	2.2	1.9	1.7	1.8
6-311G**	2.2	2.0	1.9	2.0	1.8	1.8	1.7	1.8	1.5	1.5	1.7
6-311+G	2.3	2.2	2.5	2.6	2.5	2.3	2.5	2.5	1.9	1.8	1.9
6-311++G	2.3	2.2	2.5	2.6	2.5	2.3	2.5	2.5	1.9	1.8	1.9
6-311+G*	3.2	3.0	3.1	3.5	1.9	1.9	1.9	2.0	2.0	1.8	2.0
6-311++G**	1.7	1.7	1.5	1.7	1.4	1.6	1.4	1.6	1.7	1.5	1.8
Def2SVP	3.3	2.6	2.4	2.3	2.4	2.0	1.8	1.9	1.5	1.5	1.7
Def2TZVP	1.6	1.4	1.1	1.4	1.6	1.3	1.1	1.3	1.4	1.3	1.5
ma-Def2SVP	2.8	2.2	1.9	2.1	1.9	1.7	1.5	1.8	1.6	1.5	1.8
ma-Def2TZVP	1.5	1.3	0.9	1.4	1.5	1.3	1.0	1.3	1.4	1.3	1.5
cc-pVTZ	3.6	2.8	2.8	3.0	2.8	2.2	2.2	2.1	1.4	1.5	1.6
cc-pVDZ	1.2	1.0	0.7	1.2	0.9	1.1	0.9	1.2	1.2	1.2	1.4
aug-cc-pVDZ	1.7	1.4	1.3	1.8	2.1	1.6	1.4	1.6	1.9	1.8	1.9
aug-cc-pVTZ	0.9	0.8	0.7	1.3	0.7	0.9	0.8	1.0	1.0	1.1	1.1

Table A3.3: Grid errors for various property derivatives obtained with different basis sets and B3LYP combined with the (99, 590) integration grid for the HCN · HF system. For each column, the values in cells have to be multiplied with the exponent given in top.

	$d^m E/d\chi_{1,z}{}^m$				$d^m \mu_z/d\chi_{1,z}{}^m$				$d^m \alpha_{zz}/d\chi_{1,z}{}^m$		
	1	2	3	4	0	1	2	3	0	1	2
Basis set	[$\times 10^{-7}$]	[$\times 10^{-7}$]	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-7}$]	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]
STO-3G	1.4	10.5	7.2	7.4	1.5	1.1	3.2	0.8	5.9	3.8	3.4
3-21G	1.6	9.4	8.5	6.0	3.2	1.6	1.0	1.0	6.5	5.6	2.7
3-21+G	1.6	9.4	8.5	6.1	3.0	1.6	1.0	1.0	8.1	6.1	3.2
3-21++G	1.6	9.4	8.5	6.1	3.0	1.6	1.0	1.0	7.5	6.1	3.3
6-31G	1.3	6.2	6.0	5.6	3.0	2.6	1.7	1.5	7.5	6.2	6.3
6-31G*	1.3	7.2	6.6	5.5	2.2	1.6	1.0	1.0	8.8	8.5	6.6
6-31G**	1.2	8.1	6.1	5.2	1.6	1.1	0.9	0.6	6.6	4.4	2.5
6-31+G	1.4	6.3	6.3	5.7	3.3	2.6	1.8	1.6	9.3	7.6	7.0
6-31++G	1.4	6.3	6.3	5.6	3.3	2.6	1.8	1.6	9.3	7.8	7.1
6-31+G*	1.4	7.6	6.5	5.5	2.2	1.4	0.9	1.0	10.9	10.3	8.1
6-31++G**	1.2	8.3	6.2	5.5	1.5	1.2	0.9	0.7	9.4	5.0	4.5
6-311G	0.8	5.7	3.3	3.4	3.7	2.8	1.8	1.7	8.6	9.7	6.5
6-311G*	3.5	5.8	2.9	3.0	3.4	2.4	1.5	1.5	10.8	10.1	7.2
6-311G**	0.8	6.6	3.5	3.5	2.3	1.7	1.2	1.0	10.0	5.1	4.0
6-311+G	0.8	5.9	3.2	3.2	3.5	2.6	1.7	1.6	11.2	11.0	8.8
6-311++G	0.8	5.9	3.2	3.2	3.5	2.6	1.7	1.6	11.2	11.0	8.8
6-311+G*	3.4	6.2	2.9	2.9	2.9	1.9	1.2	1.2	12.8	12.4	8.9
6-311++G**	0.8	6.9	3.8	3.7	1.9	1.6	1.0	0.9	10.6	6.7	6.3
Def2SVP	1.0	9.7	5.0	4.6	2.5	1.8	1.4	1.1	10.2	5.6	4.8
Def2TZVP	2.1	3.7	3.0	2.3	1.3	1.2	0.9	0.7	8.6	4.5	4.2
ma-Def2SVP	1.0	9.0	5.4	5.2	2.0	1.7	1.1	0.9	12.1	7.7	6.8
ma-Def2TZVP	2.1	3.8	3.0	2.4	1.2	1.2	0.8	0.6	8.8	4.9	4.7
cc-pVTZ	1.5	8.0	6.0	4.1	2.4	1.8	1.3	1.0	9.1	5.2	2.8
cc-pVDZ	1.2	8.2	5.8	5.2	1.2	1.0	0.8	0.5	9.6	3.9	3.6
aug-cc-pVDZ	1.4	7.9	5.7	4.5	2.4	1.8	1.0	0.8	15.7	10.7	9.2
aug-cc-pVTZ	3.1	9.1	6.0	5.3	0.6	0.7	0.7	0.5	43.2	37.1	84.8

Table A3.4: Grid errors for various property derivatives obtained with different basis sets and ωB97X combined with the (250, 974) integration grid for the HCN · HF system. For each column, the values in cells have to be multiplied with the exponent given in top.

	$d^m E/d\chi_{1,z}{}^m$				$d^m \mu_z/d\chi_{1,z}{}^m$				$d^m \alpha_{zz}/d\chi_{1,z}{}^m$		
	1	2	3	4	0	1	2	3	0	1	2
Basis set	[$\times 10^{-7}$]	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-3}$]	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]	[$\times 10^{-2}$]
STO-3G	3.9	6.8	15.3	4.6	1.3	2.2	5.2	14.7	0.8	1.5	4.5
3-21G	3.9	4.6	7.8	1.2	1.8	2.6	3.9	6.7	2.0	3.5	5.5
3-21+G	2.8	4.3	7.2	1.2	1.8	2.7	3.9	7.3	2.1	3.5	5.9
3-21++G	2.8	4.3	7.1	1.2	1.8	2.7	3.9	7.3	2.1	3.5	5.9
6-31G	3.2	5.3	12.0	3.0	2.2	3.9	7.6	19.1	2.2	5.0	12.9
6-31G*	3.0	4.6	8.6	1.6	1.7	2.7	4.6	10.3	1.8	3.7	7.6
6-31G**	2.3	3.7	5.6	1.1	1.5	2.2	3.4	5.7	1.8	2.7	5.6
6-31+G	3.2	5.6	13.3	3.5	2.2	4.1	8.4	22.6	2.4	5.7	15.2
6-31++G	3.2	5.6	13.2	3.5	2.1	4.1	8.3	22.6	2.4	5.6	15.1
6-31+G*	2.7	4.0	8.0	1.8	1.8	3.0	5.4	12.7	1.6	3.4	8.2
6-31++G**	2.1	3.2	4.9	0.8	10.7	2.5	3.6	6.6	1.3	2.1	4.1
6-311G	3.3	5.4	10.7	2.4	2.2	4.0	6.9	18.0	2.8	5.8	13.9
6-311G*	4.7	5.1	9.7	1.7	1.9	3.2	5.8	13.5	2.6	5.5	10.3
6-311G**	2.3	3.9	6.2	1.2	1.4	2.2	4.0	9.1	1.7	2.9	6.3
6-311+G	3.3	5.5	11.3	2.4	1.8	3.5	7.1	18.7	2.8	6.1	13.3
6-311++G	3.3	5.5	11.3	2.4	1.9	3.5	7.1	18.8	2.8	6.1	13.3
6-311+G*	4.2	4.5	8.9	1.8	1.9	3.3	5.9	13.2	2.1	4.6	10.2
6-311++G**	2.1	3.2	5.3	1.0	1.4	2.4	4.2	8.7	1.3	2.6	5.7
Def2SVP	2.4	4.0	6.6	1.4	1.5	2.5	4.8	11.2	1.6	3.0	6.6
Def2TZVP	1.9	2.8	4.3	0.7	1.2	1.9	3.0	5.6	1.3	2.2	3.5
ma-Def2SVP	2.2	3.3	6.5	1.4	1.4	2.4	4.3	9.5	1.3	3.0	7.3
ma-Def2TZVP	1.8	2.7	4.1	0.6	1.3	1.9	3.0	5.6	1.3	2.1	3.3
cc-pVTZ	2.7	4.6	8.0	1.8	1.7	2.5	4.4	9.4	2.2	4.1	9.8
cc-pVDZ	1.5	2.1	3.0	0.4	1.1	1.6	2.3	3.7	1.2	1.7	2.5
aug-cc-pVDZ	2.4	4.0	7.8	1.7	1.4	2.8	4.7	11.6	1.8	3.9	9.2
aug-cc-pVTZ	3.5	2.0	2.7	0.4	1.0	1.5	2.2	3.4	1.3	1.6	2.3

Table A3.5: Grid errors for various property derivatives obtained with different basis sets and M06-2X combined with the (250, 974) integration grid for the HCN · HF system. For each column, the values in cells have to be multiplied with the exponent given in top.

	$d^m E/d\chi_{1,z}{}^m$				$d^m \mu_z/d\chi_{1,z}{}^m$				$d^m \alpha_{zz}/d\chi_{1,z}{}^m$		
	1	2	3	4	0	1	2	3	0	1	2
Basis set	[$\times 10^{-8}$]	[$\times 10^{-6}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-7}$]	[$\times 10^{-6}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]	[$\times 10^{-5}$]	[$\times 10^{-4}$]	[$\times 10^{-3}$]
STO-3G	7.3	1.0	1.6	3.0	2.6	3.1	0.4	0.5	1.4	1.7	2.1
3-21G	27.3	5.6	13.0	26.5	3.6	6.3	1.3	2.4	2.3	3.9	4.9
3-21+G	24.9	5.1	11.8	23.9	3.6	6.3	1.3	2.4	2.9	4.6	5.8
3-21++G	24.9	5.1	11.8	23.9	3.6	6.3	1.3	2.4	2.9	4.6	5.9
6-31G	9.5	1.7	3.0	6.4	5.2	7.2	1.0	1.5	5.5	6.7	13.0
6-31G*	6.8	1.5	2.8	6.6	3.8	5.6	0.8	1.3	3.4	5.1	6.6
6-31G**	9.1	1.8	3.4	7.2	3.0	4.0	0.6	0.9	3.3	5.3	6.8
6-31+G	9.4	1.7	2.9	6.1	5.8	7.8	1.1	1.5	5.8	7.0	11.0
6-31++G	9.3	1.7	2.9	6.1	6.0	7.8	1.1	1.5	5.8	7.1	11.1
6-31+G*	6.7	1.5	2.7	6.4	4.3	5.8	0.9	1.3	3.7	5.8	7.0
6-31++G**	8.5	1.7	3.2	7.0	3.3	4.7	0.7	1.1	3.8	5.1	7.4
6-311G	6.9	1.2	1.9	3.8	5.0	8.2	1.1	1.8	5.9	16.4	15.9
6-311G*	5.5	1.1	1.8	4.0	4.6	7.0	1.0	1.6	4.9	10.7	18.8
6-311G**	8.2	1.3	2.6	4.8	13.3	5.7	0.9	1.2	4.4	6.4	9.3
6-311+G	6.5	1.1	1.8	3.6	6.4	8.3	1.1	1.7	6.1	15.8	17.5
6-311++G	6.5	1.1	1.8	3.6	6.4	8.3	1.1	1.7	6.1	16.0	17.6
6-311+G*	5.5	1.0	1.7	3.9	6.2	6.5	1.0	1.5	4.5	9.7	24.5
6-311++G**	7.2	1.2	2.3	4.5	4.0	5.8	0.9	1.3	4.4	5.8	11.1
Def2SVP	16.9	1.1	1.8	2.8	4.5	6.0	0.9	1.2	4.3	6.5	9.3
Def2TZVP	8.1	1.6	3.2	6.8	2.6	3.9	0.6	0.9	3.1	4.1	6.0
ma-Def2SVP	18.7	1.1	1.6	2.7	4.3	6.3	1.0	1.5	4.7	6.2	12.1
ma-Def2TZVP	7.8	1.6	3.2	6.8	2.6	4.0	0.6	0.9	3.1	4.0	6.0
cc-pVTZ	14.2	2.7	5.3	11.0	4.4	6.1	0.9	1.4	4.7	7.4	14.0
cc-pVDZ	6.1	1.3	2.4	5.3	2.2	2.8	0.4	0.6	2.6	3.4	5.1
aug-cc-pVDZ	10.7	2.3	4.4	9.8	5.2	8.6	1.3	2.1	5.5	7.7	30.8
aug-cc-pVTZ	5.7	1.2	2.3	5.2	1.8	2.7	0.4	0.7	10.2	5.8	4.1

Table A3.6: Grid errors for various property derivatives obtained with different basis sets and B3LYP combined with the (250, 974) integration grid for the HCN · HF system. For each column, the values in cells have to be multiplied with the exponent given in top.

	$d^m E/d\chi_{1,z}{}^m$				$d^m \mu_z/d\chi_{1,z}{}^m$				$d^m \alpha_{zz}/d\chi_{1,z}{}^m$		
	1	2	3	4	0	1	2	3	0	1	2
Basis set	[$\times 10^{-10}$]	[$\times 10^{-9}$]	[$\times 10^{-8}$]	[$\times 10^{-7}$]	[$\times 10^{-10}$]	[$\times 10^{-8}$]	[$\times 10^{-7}$]	[$\times 10^{-6}$]	[$\times 10^{-7}$]	[$\times 10^{-6}$]	[$\times 10^{-5}$]
STO-3G	2.1	6.9	7.0	7.9	8.7	1.0	275	.3 2.2	0.4	0.7	5.4
3-21G	6.0	11.9	23.2	22.0	9.1	1.1	2.4	2.8	0.7	1.1	1.7
3-21+G	5.2	15.9	21.8	21.1	10.5	1.2	2.1	2.9	0.7	1.2	1.7
3-21++G	5.3	15.7	21.7	21.1	8.5	1.2	2.1	2.9	0.7	1.2	1.7
6-31G	2.5	6.9	18.8	17.9	14.9	2.6	18.	5 5.9	1.3	1.8	2.9
6-31G*	2.7	6.8	104.0	16.2	9.6	1.4	17.	7 2.8	1.0	1.4	2.3
6-31G**	1.3	2.5	1.3	4.8	8.5	1.2	1.9	2.7	0.8	1.2	1.7
6-31+G	2.5	8.2	19.0	18.1	15.8	32.4	6.6	6.2	1.8	2.0	3.3
6-31++G	2.5	8.2	19.0	18.1	15.8	32.0	6.7	6.8	1.8	2.0	3.3
6-31+G*	4.4	9.3	13.4	20.9	22.5	1.5	2.2	3.0	1.0	1.5	2.2
6-31++G**	2.0	7.4	19.5	18.6	7.4	1.1	1.5	2.3	0.8	1.1	1.7
6-311G	5.7	1.6	3.4	10.6	14.5	28.9	14.	3 5.4	1.4	72.3	8.7
6-311G*	1.4	2.4	2.2	6.6	12.1	1.9	2.8	4.6	1.3	2.2	3.5
6-311G**	0.9	66.8	1.1	5.3	8.4	1.4	1.9	3.1	1.2	1.4	2.7
6-311+G	1.0	1.6	4.0	8.2	14.2	1.8	6.7	4.5	1.5	2.6	3.9
6-311++G	1.0	1.6	4.1	8.2	14.6	1.8	6.6	4.5	1.5	2.5	3.9
6-311+G*	1.3	2.3	2.6	6.7	10.9	1.7	2.6	4.0	1.2	19.3	4.9
6-311++G**	0.9	51.4	1.5	5.9	8.5	1.5	1.8	3.5	1.1	1.3	5.1
Def2SVP	12.7	237.6	28.6	29.5	9.7	1.6	15.	1 3.3	1.0	1.5	2.7
Def2TZVP	0.9	1.9	3.4	9.6	6.1	1.2	4.9	2.1	1.0	0.9	1.8
ma-Def2SVP	12.0	8.5	27.4	28.7	10.4	1.6	2.4	4.1	0.9	1.3	2.2
ma-Def2TZVP	0.9	2.1	3.2	9.2	6.0	1.0	1.3	1.9	1.0	0.9	1.6
cc-pVTZ	3.2	3.3	5.6	8.2	11.2	1.8	3.7	3.7	1.1	1.7	2.9
cc-pVDZ	1.1	2.0	3.1	14.9	4.2	23.6	7.6	1.8	0.5	0.8	1.3
aug-cc-pVDZ	4.4	6.7	10.2	18.3	29.1	1.8	3.0	4.8	6.9	2.7	3.1
aug-cc-pVTZ	0.7	0.6	1.8	16.9	185.4	7.5	7.9	14.6	173.0	176.7	314.5

Table A3.7: *Part 1)* Ratios between the RMSE values of E_x and E_c for some DFAs combined with the (99, 590) grid. The numbers correspond to the averages of ratios for the $d^3E/d\xi^3$ and $d^4E/d\xi^4$ derivatives, computed separately for each chemical system. The second and third columns show the names of the (meta-)GGAs exchange and correlation functionals included in a given DFA (information not shown for the standalone functionals, for which the exchange and correlation parts are not meant to be used separately).

DFA	E_x^{DFA}	E_c^{DFA}	HCN · HF	HCN · HCl	OC · HF	N_2 · HF	HCN · BrF	Ar_2	He_2
SVWN5	S	VWN5	8.1	7.0	8.4	9.6	19.3	11.7	7.3
SPW92	S	PW92	8.3	5.7	7.9	9.3	19.4	11.8	7.4
BH&H	S	LYP	1.4	2.0	0.9	2.2	7.8	7.1	2.2
BLYP	B88	LYP	14.0	27.5	6.9	19.2	11.8	20.9	72.1
B1LYP	B88	LYP	11.1	21.8	6.1	8.2	8.1	16.1	53.8
B3LYP	B88	LYP	14.7	21.3	8.0	8.6	9.5	15.1	54.7
BH&HLYP	B88	LYP	7.7	16.4	5.0	6.5	5.8	11.0	35.9
LC-BLYP	B88	LYP	6.6	8.3	3.2	7.2	10.4	10.8	2.8
CAM-B3LYP	B88	LYP	12.1	18.8	6.7	7.0	9.1	12.0	27.5
PBE	PBE	PBE	2.2	3.3	1.6	3.6	8.9	1.1	0.9
PBE0	PBE	PBE	1.6	2.0	1.3	1.7	9.1	0.8	0.7
PBE50	PBE	PBE	1.1	1.2	0.8	0.8	8.7	0.5	0.4
LC-ωPBE	PBE	PBE	2.1	1.3	1.7	1.1	14.3	0.6	0.1
TPSS	TPSS	TPSS	2.7	9.1	2.2	5.5	19.1	0.8	1.3
RevTPSS	RevTPSS	RevTPSS	3.8	9.5	3.8	5.9	16.7	1.0	1.1
TPSSh	TPSS	TPSS	2.7	8.0	2.2	4.4	16.9	0.8	1.1
B97			$1.4\cdot10^{-2}$	$2.1\cdot10^{-3}$	$2.1\cdot10^{-2}$	$4.6\cdot10^{-3}$	$2.3\cdot10^{-1}$	$4.7\cdot10^{-2}$	$2.2\cdot10^{-1}$
B97-D			$3.3\cdot10^{-2}$	$4.8\cdot10^{-3}$	$4.7\cdot10^{-2}$	$3.2\cdot10^{-2}$	$3.8\cdot10^{-1}$	$1.8\cdot10^{-1}$	$3.9\cdot10^{-1}$
ωB97			$1.5\cdot10^{-2}$	$3.2\cdot10^{-3}$	$1.2\cdot10^{-2}$	$4.7\cdot10^{-3}$	$9.9\cdot10^{-2}$	$3.3\cdot10^{-2}$	$3.3\cdot10^{-2}$
ωB97X			$1.4\cdot10^{-2}$	$2.4\cdot10^{-3}$	$1.1\cdot10^{-2}$	$4.5\cdot10^{-3}$	$5.0\cdot10^{-2}$	$3.5\cdot10^{-2}$	$8.7\cdot10^{-2}$
ωB97X-D			$6.7\cdot10^{-3}$	$1.2\cdot10^{-3}$	$5.8\cdot10^{-3}$	$1.9\cdot10^{-3}$	$2.8\cdot10^{-2}$	$2.4\cdot10^{-2}$	$1.4\cdot10^{-1}$
ωB97X-D3			$1.0\cdot10^{-2}$	$1.8\cdot10^{-3}$	$8.7\cdot10^{-3}$	$3.1\cdot10^{-3}$	$4.2\cdot10^{-2}$	$3.1\cdot10^{-2}$	$1.1\cdot10^{-1}$
ωB97M-V			$2.3\cdot10^{-1}$	$3.0\cdot10^{-2}$	$1.0\cdot10^{-1}$	$6.2\cdot10^{-2}$	$7.0\cdot10^{-1}$	$2.2\cdot10^{-1}$	$1.8\cdot10^{-2}$
ωB97X-V			$2.3\cdot10^{-1}$	$2.9\cdot10^{-2}$	$4.2\cdot10^{-1}$	$5.2\cdot10^{-2}$	5.9	$2.8\cdot10^{-1}$	$7.8\cdot10^{-2}$
B97M-V			$2.2\cdot10^{-1}$	$1.6\cdot10^{-1}$	$3.6\cdot10^{-1}$	$4.0\cdot10^{-1}$	2.7	2.0	$6.8\cdot10^{-1}$

Table A3.8: *Part 2)* Ratios between the RMSE values of E_x and E_c for some DFAs combined with the (99, 590) grid. The numbers correspond to the averages of ratios for the $\mathrm{d}^3E/\mathrm{d}\xi^3$ and $\mathrm{d}^4E/\mathrm{d}\xi^4$ derivatives, computed separately for each chemical system. The second and third columns show the names of the (meta-)GGAs exchange and correlation functionals included in a given DFA (information not shown for the standalone functionals, for which the exchange and correlation parts are not meant to be used separately).

DFA	E_x^{DFA}	E_c^{DFA}	$\mathrm{HCN}\cdot\mathrm{HF}$	$\mathrm{HCN}\cdot\mathrm{HCl}$	$\mathrm{OC}\cdot\mathrm{HF}$	$\mathrm{N_2}\cdot\mathrm{HF}$	$\mathrm{HCN}\cdot\mathrm{BrF}$	$\mathrm{Ar_2}$	$\mathrm{He_2}$
B1PW91	B88	PW91	1.5	1.2	0.7	0.4	6.9	0.9	0.6
PW91	PW91	PW91	4.1	5.7	3.1	6.4	13.8	1.9	1.0
mPW91	mPW91	mPW91	4.5	6.4	3.0	4.9	16.2	1.9	0.9
VSXC			0.4	0.1	0.3	$4.2\cdot10^{-2}$	0.8	0.1	0.6
SCAN			2.6	4.2	3.0	3.3	2.5	2.9	5.4
SCAN0			2.0	3.4	2.2	2.5	3.5	2.2	4.1
M06			0.5	0.1	0.7	0.6	0.6	0.1	0.0
M06-L			0.3	0.1	0.6	0.3	1.5	0.3	0.1
M06-HF			1.1	0.2	1.5	0.4	0.5	0.5	0.0
M06-2X			0.6	0.2	1.2	0.5	0.5	0.5	0.2
M11			3.5	2.2	4.0	2.7	2.4	1.2	0.0
M11-L			3.6	3.5	2.8	2.2	1.6	1.1	1.0
SOGGA11			$9.4\cdot10^{-2}$	$1.2\cdot10^{-1}$	$1.5\cdot10^{-1}$	$4.3\cdot10^{-2}$	$3.9\cdot10^{-1}$	$7.0\cdot10^{-1}$	2.0
SOGGA11-X			$1.1\cdot10^{-1}$	$1.1\cdot10^{-1}$	$7.4\cdot10^{-2}$	$5.2\cdot10^{-2}$	1.8	$1.3\cdot10^{-1}$	$4.3\cdot10^{-1}$
N12			$1.4\cdot10^{-2}$	$3.1\cdot10^{-3}$	$1.6\cdot10^{-2}$	$6.6\cdot10^{-3}$	$1.4\cdot10^{-1}$	$9.8\cdot10^{-2}$	$4.6\cdot10^{-1}$
N12-SX			$5.8\cdot10^{-2}$	$1.9\cdot10^{-2}$	$5.8\cdot10^{-2}$	$7.0\cdot10^{-2}$	1.2	$1.0\cdot10^{-1}$	$5.2\cdot10^{-1}$
MN12-L			1.9	1.4	2.9	1.0	6.2	0.6	0.6
MN12-SX			1.5	0.9	1.5	0.7	5.0	1.7	4.8
MN15			0.6	0.7	0.5	0.7	4.9	0.5	0.3
MN15-L			0.9	1.1	0.6	0.3	5.5	0.2	0.2

Table A3.9: Considered families of DFAs, which are used for summary presented in Figure .

Family	DFAs
BHH	BHH
LSDA	SVWN5, SPW92
BLYP	BLYP, B1LYP, B3LYP, BH&HLYP, LC-BLYP, CAM-B3LYP
PBE	PBE, PBE0, PBE50, LC-ωPBE
TPSS	TPSS, RevTPSS, TPSSh
PW91	PW91, mPW91
M11	M11, M11-L
MN12-L	MN12-L
MN12-SX	MN12-SX
B97	B97, B97-D, ωB97, ωB97X, ωB97X-D, ωB97X-D3
B97-V	ωB97X-V, ωB97M-V, B97M-V
SCAN	SCAN, SCAN0
M06	M06, M06-L, M06-2X, M06-HF
VSXC	VSXC
MN15	MN15, MN15-L
SOGGA11	SOGGA11, SOGGA11-X
N12	N12, N12-SX

A4 Supplementary Examples: Spurious Oscillations in Intramolecular Normal Modes

A4.1 Butadiene (Q_2 and Q_3 modes)

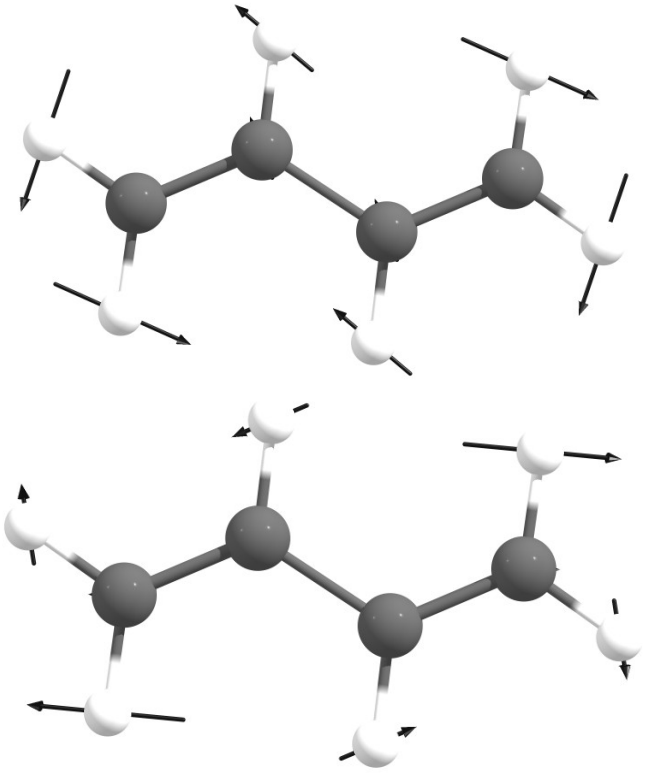

Figure A4.19: Graphical representation of the Q_2 (B_u, left) and Q_3 (A_g, right) normal modes of butadiene. Obtained at the M06-2X/aug-cc-pVDZ level of theory.

One can expect similar problems related to the spurious oscillations as in the allyl anion. For this molecular example, we actually studied two different modes, Q_2 (belonging to the B_u irreducible representation) and Q_3 (belonging to the A_g irreducible representation), which are depicted in Figure A4.19. Both of them have a similar pattern of motion of the H-H pairs in the Q_2 mode of the allyl anion, and they are also very well represented by the harmonic approximation (see Table A4.10).

Again, spurious oscillations are noticeable mostly for the (99, 590) grid. Starting from the derivatives with respect to the Q_2 mode, the oscillations can be observed both for $\mathrm{d}^4E/\mathrm{d}Q_2^4$ and $\mathrm{d}^3\mu_y/\mathrm{d}Q_2^3$ (see Figure A4.20). However, the actual values of RRMSE are quite low for $\mathrm{d}^4E/\mathrm{d}Q_2^4$ (below 10%), but already substantial for $\mathrm{d}^3\mu_y/\mathrm{d}Q_2^3$ — 18.8% for ωB97X and 45.1% for M06-2X. In the case of the derivatives with respect to the Q_3 mode, errors in $\mathrm{d}^4E/\mathrm{d}Q_3^4$ are much larger than in $\mathrm{d}^4E/\mathrm{d}Q_2^4$, namely 14.4% for ωB97X and 48.3% for M06-2X. For the Q_3 mode, dipole derivatives were not analyzed since all their components are zero by the symmetry constraint (Q_3 belongs to the A_g irreducible representation).

Table A4.10: Vibrational properties of the Q_2 (top panel) and Q_3 (bottom panel) normal modes of butadiene obtained with some selected methods: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, IR_{har} and IR_{anh} — harmonic and anharmonic IR intensities, RA_{har} and RA_{anh} — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

Q_2 (B_u) mode

Method	$\omega_{\text{har}}\left[\frac{1}{\text{cm}}\right]$	$\omega_{\text{anh}}\left[\frac{1}{\text{cm}}\right]$	$\text{IR}_{\text{har}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{IR}_{\text{anh}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{RA}_{\text{har}}\left[\text{Å}^6\right]$	$\text{RA}_{\text{anh}}\left[\text{Å}^6\right]$
HF	319	323	4.028	4.041	0.000	0.000
ωB97X	300		2.979			
M06-2X	299		2.878			

Q_3 (A_g) mode

Method	$\omega_{\text{har}}\left[\frac{1}{\text{cm}}\right]$	$\omega_{\text{anh}}\left[\frac{1}{\text{cm}}\right]$	$\text{IR}_{\text{har}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{IR}_{\text{anh}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{RA}_{\text{har}}\left[\text{Å}^6\right]$	$\text{RA}_{\text{anh}}\left[\text{Å}^6\right]$
HF	548	547	0.000	0.000	0.368	0.379
ωB97X	520		0.000			
M06-2X	517		0.000			

Table A4.11: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_2 and Q_3 normal modes of allyl anion, obtained for some DFAs and integration grids.

Method	Grid	$\text{d}^3E/\text{d}Q_2^3$	$\text{d}^4E/\text{d}Q_2^4$	$\text{d}^3\mu_y/\text{d}Q_2^3$	$\text{d}^3E/\text{d}Q_3^3$	$\text{d}^4E/\text{d}Q_3^4$
ωB97X	(99, 590)	1.4	3.8	18.8	2.4	14.4
	(250, 974)	0.1	0.3	2.1	0.0	0.7
	(750, 974)	0.1	0.3	1.2	0.0	0.5
M06-6X	(99, 590)	6.5	16.0	45.1	6.5	48.3
	(250, 974)	0.1	0.4	13.5	0.1	8.2
	(750, 974)	0.1	0.4	10.9	0.1	8.1

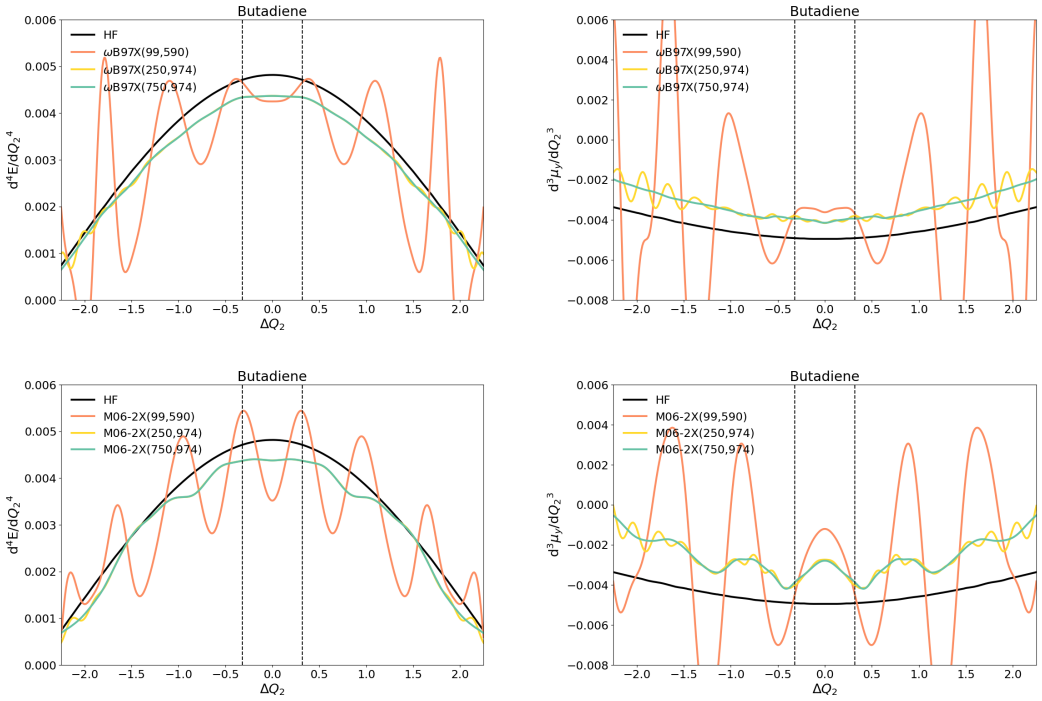

Figure A4.20: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_2^4$ (left column) and $\mathrm{d}^3\mu_y/\mathrm{d}Q_2^3$ (right column) for the displacements along the Q_2 normal coordinate of butadiene. Obtained with ωB97X (top row) and M06-2X (bottom row).

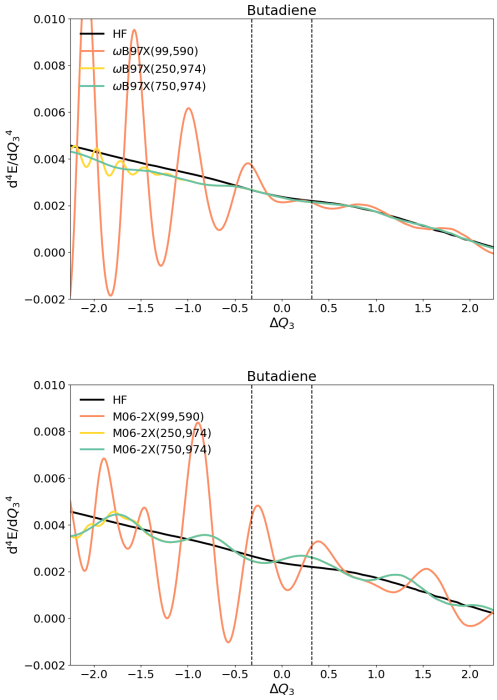

Figure A4.21: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_3^4$ for the displacements along the Q_3 normal coordinate of butadiene. Obtained with ωB97X (top row) and M06-2X (bottom row).

A4.2 Cyclobutadiene (Q_4 mode)

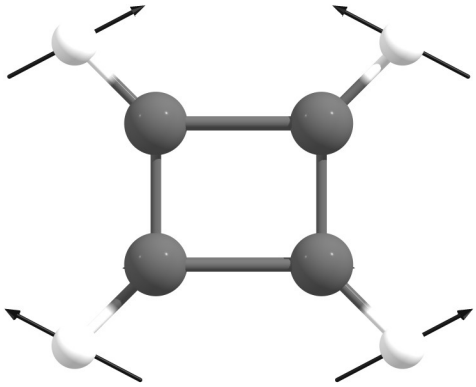

Figure A4.22: Graphical representation of the Q_4 (B_{2u}) normal mode of cyclobutadiene. Obtained at the M06-2X/aug-cc-pVDZ level of theory.

We also investigated the cyclic version of the butadiene, namely cyclobutadiene. It is also more a rigid molecule and the normal mode of our interest, Q_4, is characterized with twice larger vibrational frequencies as the Q_2 and Q_3 modes of butadiene. The Q_4 mode is depicted in Figure A4.22 and the basic quantities related to the IR and Raman spectrum are compiled in Table A4.12.

Despite Q_4 being quite a rigid motion, it is also susceptible to give spurious oscillations in the computations of the high-order derivatives (see Figure A4.23). For M06-2X(99, 590), one observes oscillations across the whole curve of $\mathrm{d}^4E/\mathrm{d}Q_4^4$. However, the corresponding RRMSE value of 6.3% suggests a low and acceptable error. This is due to the fact that it is measured within the selected displacement ranges $\{\boldsymbol{\xi_0}\}$, the oscillations themselves are not that big, but dramatically enhance for larger displacements from the optimized geometry. In contrast, in the curve of $\mathrm{d}^3\mu_y/\mathrm{d}Q_4^3$, the oscillations are much more prominent even for the optimized geometry, with RRMSE exceeding 20%. In the case of ωB97X, one notices the oscillations for all grids tested. In particular, for $\mathrm{d}^4E/\mathrm{d}Q_4^4$ obtained with the (99, 590), (250, 974), and (750, 974) grids the RRMSE are 18.1%, 11.8%, and 11.8%, respectively. For $\mathrm{d}^3\mu_y/\mathrm{d}Q_4^3$, the corresponding errors are approximately twice as large. Interestingly, the grid-errors for the (250, 974) and (750, 974) grids are almost identical, despite the latter grid being three times larger. One can expect that for this studied motion along the Q_4 normal mode of cyclobutadiene, the size of the angular Lebedev grid is also important. Unfortunately, the 974-point Lebedev grid is already the largest Lebedev grid available in Gaussian, and one cannot further test this hypothesis. However, thanks to our algorithm for the detection and quantification of spurious oscillations, this problem could be identified and quantified, without the need of calculations employing a larger angular grid. Effect of the importance of the angular grid is further discussed for the H_2O_2 and H_2S_2 examples in the following sections.

Table A4.12: Vibrational properties of the Q_4 normal mode of cyclobutadiene obtained with some selected methods: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, IR_{har} and IR_{anh} — harmonic and anharmonic IR intensities, RA_{har} and RA_{anh} — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

	Q_4 (B_{2u}) mode					
Method	$\omega_{\text{har}}\left[\frac{1}{\text{cm}}\right]$	$\omega_{\text{anh}}\left[\frac{1}{\text{cm}}\right]$	$\text{IR}_{\text{har}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{IR}_{\text{anh}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{RA}_{\text{har}}\left[\text{Å}^6\right]$	$\text{RA}_{\text{anh}}\left[\text{Å}^6\right]$
HF	804	790	15.557	15.281	0.000	0.000
ωB97X	764		12.255			
M06-2X	749		10.299			

Table A4.13: Values of RRMSE (in percentages) for the selected property derivatives with respect to the Q_4 normal mode of cyclobutadiene, obtained for some DFAs and integration grids.

Method	Grid	$\text{d}^3E/\text{d}Q_4^3$	$\text{d}^4E/\text{d}Q_4^4$	$\text{d}^2\mu_y/\text{d}Q_4^2$	$\text{d}^3\mu_y/\text{d}Q_4^3$
	(99, 590)	26.9	18.1	46.1	41.2
ωB97X	(250, 974)	6.2	11.8	19.4	20.3
	(750, 974)	6.1	11.8	19.4	20.3
	(99, 590)	3.6	6.3	21.8	20.6
M06-6X	(250, 974)	1.5	3.9	1.5	1.5
	(750, 974)	1.5	3.6	1.5	1.6

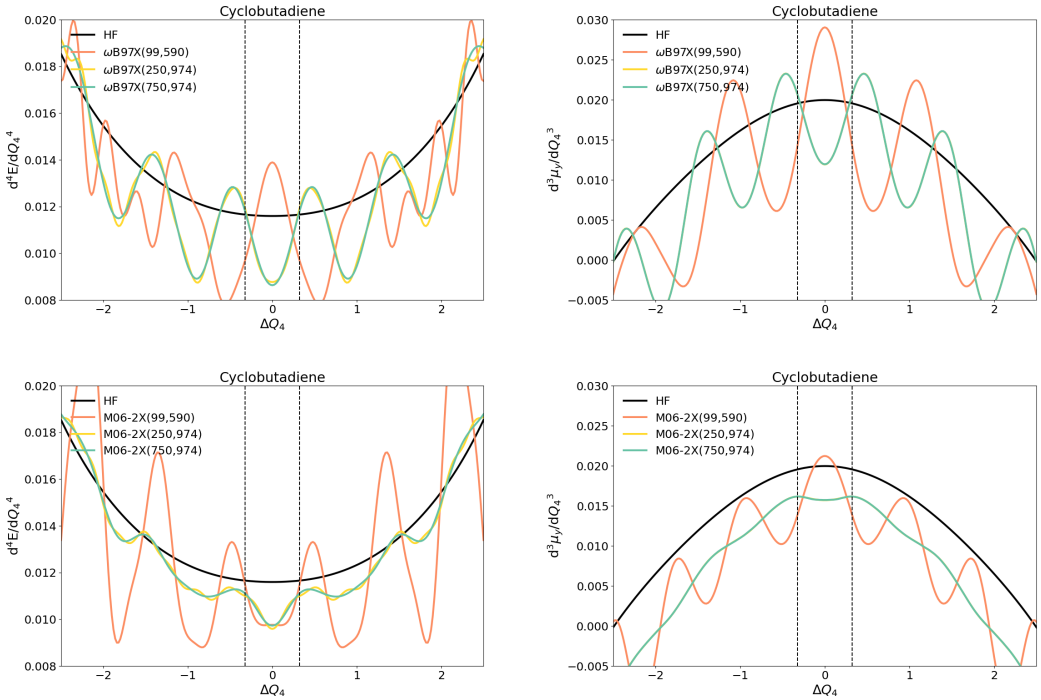

Figure A4.23: Spurious oscillations in $\text{d}^4E/\text{d}Q_4^4$ (left column) and $\text{d}^3\mu_y/\text{d}Q_4^3$ (right column) for the displacements along the Q_4 normal coordinate of cyclobutadiene. Obtained with ωB97X (top row) and M06-2X (bottom row).

A4.3 Benzene (Q_4 mode)

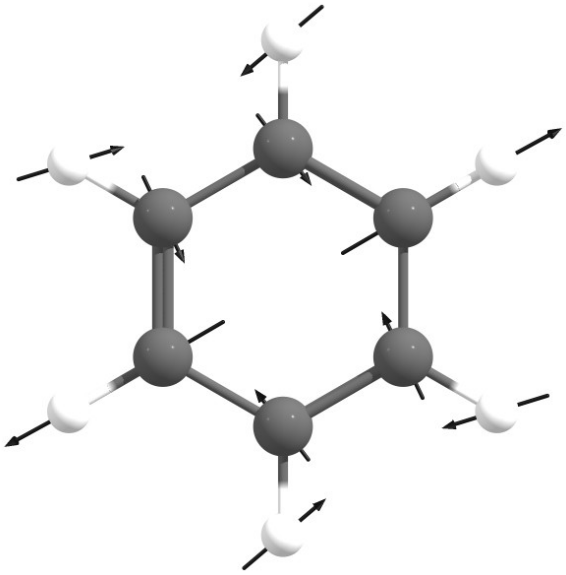

Figure A4.24: Graphical representation of the Q_4 (E_{2g}) normal mode of benzene. Obtained at the M06-2X/aug-cc-pVDZ level of theory.

We investigated the Q_4 mode, one of the low-frequency degenerated E_{2g} modes (see depiction in Figure A4.24), which resembles the motion found in the Q_4 mode of cyclobutadiene. This Q_4 is also quite rigid and well described within the harmonic approximation (see Table A4.14).

The spurious oscillations are especially observable for d^4E/dQ_4^4 obtained with the ωB97X (see Figure A4.25). For the smallest grid tested, namely (99, 590), RRMSE exceeds 70%, whereas for (250, 974) is above 14%. Moreover, one can observe that the oscillations have an amplitude few times larger outside the studied $\{\boldsymbol{\xi_0}\}$ displacement range, and, therefore, for such geometrically distorted system one would observe even larger errors. In this test, the M06-2X functional showed a bit larger grid-stability than ωB97X.

Table A4.14: Vibrational properties of the Q_4 normal mode of benzene obtained with some methods: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, IR_{har} and IR_{anh} — harmonic and anharmonic IR intensities, RA_{har} and RA_{anh} — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

	Q_4 (E_{2g}) mode					
Method	$\omega_{\text{har}}\left[\frac{1}{\text{cm}}\right]$	$\omega_{\text{anh}}\left[\frac{1}{\text{cm}}\right]$	$\text{IR}_{\text{har}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{IR}_{\text{anh}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{RA}_{\text{har}}\left[\text{Å}^6\right]$	$\text{RA}_{\text{anh}}\left[\text{Å}^6\right]$
HF	659	655	0.000	0.000	0.083	0.075
ωB97X	620		0.000			
M06-2X	613		0.000			

Table A4.15: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_4 normal mode of benzene, obtained for some DFAs and integration grids.

Method	Grid	$\mathrm{d}^3E/\mathrm{d}Q_4^3$	$\mathrm{d}^4E/\mathrm{d}Q_4^4$
	(99, 590)	43.1	70.8
ωB97X	(250, 974)	5.2	14.4
	(750, 974)	0.0	6.4
	(99, 590)	16.1	28.1
M06-6X	(250, 974)	2.1	2.9
	(750, 974)	2.2	2.4

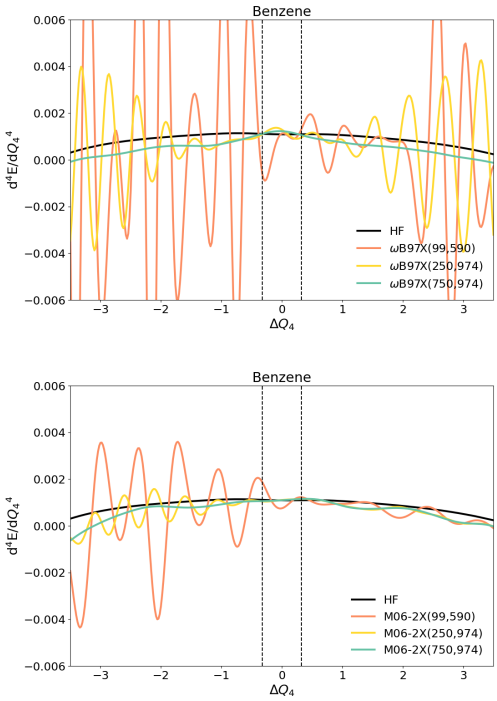

Figure A4.25: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_4^4$ for the displacements along the Q_4 normal coordinate of benzene. Obtained with ωB97X (top row) and M06-2X (bottom row).

A4.4 Naphthalene (Q_3 mode)

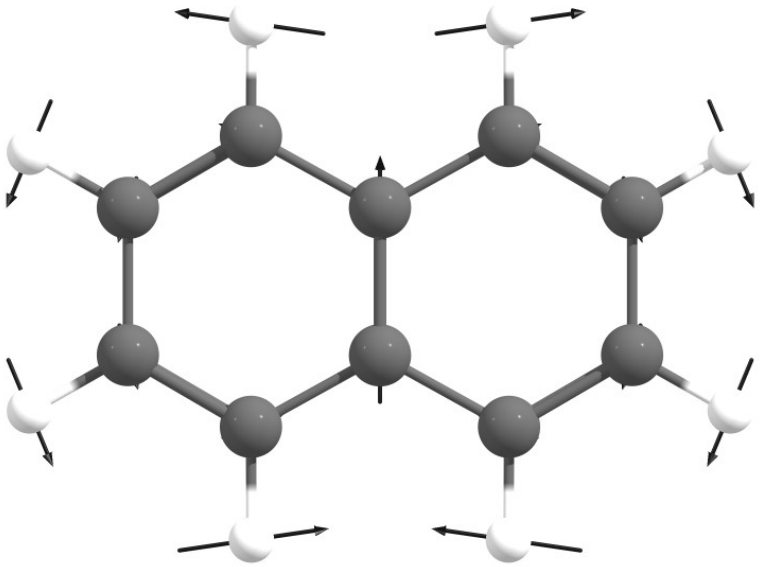

Figure A4.26: Graphical representation of the Q_3 (B_{1u}) normal mode of naphthalene. Obtained at M06-2X/aug-cc-pVDZ level of theory.

Naphthalene is another well-known example of polyaromatic hydrocarbon. For this molecule, we studied the Q_3 normal mode (see Figure A4.26), which shares more similarity with the studied mode of allyl anion and butadiene, rather than the ones of cyclobutadiene and benzene. Namely, it has the characteristic motion of the hydrogen atoms connected the C1 and C8 atoms and characterizes with a smaller vibrational frequencies (of the similar magnitude to the ones of allyl anion and butadiene). For this molecule, we studied few more DFAs, namely B3LYP, BH&H, CAM-B3LYP, ωB97X, M06-2X, and N12. The first three DFAs proven to be quite grid robust in the extensive benchmark from Section 9.5, whereas the latter three turned out to be highly susceptible to the grid-related spurious oscillations.

Indeed, in this test, B3LYP, BH&H, CAM-B3LYP turned out to be also very grid stable even in the computations of $\mathrm{d}^4E/\mathrm{d}Q_3^4$ (see Figure A4.27 and Table A4.17). In contrast, ωB97X, M06-2X, and N12 suffer from huge grid-oscillations (see Figure A4.27) and the resulting values of $\mathrm{d}^4E/\mathrm{d}Q_3^4$ are greatly affected by the errors. All three functionals suffer for large errors already for $\mathrm{d}^3E/\mathrm{d}Q_3^3$. Moreover, this is the first example in this Section for which the RRMSE is larger than 100% (for $\mathrm{d}^4E/\mathrm{d}Q_3^4$ computed with N12(99, 590)). Very similar grid instability is found for ωB97X combined with any of the tested grids. Surprisingly, M06-2X seems to be the least affected from these three DFAs because for larger grids the RRMSE values are drastically reduced.

Table A4.16: Vibrational properties of the Q_3 normal mode of naphthalene obtained with some methods: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, IR_{har} and IR_{anh} — harmonic and anharmonic IR intensities, RA_{har} and RA_{anh} — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

	Q_3 (B_{1u}) mode					
Method	$\omega_{\text{har}}\left[\frac{1}{\text{cm}}\right]$	$\omega_{\text{anh}}\left[\frac{1}{\text{cm}}\right]$	$\text{IR}_{\text{har}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{IR}_{\text{anh}}\left[\frac{\text{km}}{\text{mol}}\right]$	$\text{RA}_{\text{har}}\left[\text{Å}^6\right]$	$\text{RA}_{\text{anh}}\left[\text{Å}^6\right]$
HF	391	373	1.979	2.086	0.000	0.000
B3LYP	372		1.559			
BH&H	380		1.584			
CAM-B3LYP	376		1.601			
ωB97X	371		1.532			
M06-2X	370		1.424			
N12	368		1.435			

Table A4.17: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_3 normal mode of naphthalene, obtained for some DFAs and integration grids.

Method	Grid	$\text{d}^3E/\text{d}Q_3^3$	$\text{d}^4E/\text{d}Q_3^4$
B3LYP	(99, 590)	0.7	1.3
	(250, 974)	0.0	0.1
	(750, 974)	0.0	0.1
BH&H	(99, 590)	6.1	2.2
	(250, 974)	0.1	0.2
	(750, 974)	0.1	0.2
CAM-B3LYP	(99, 590)	1.1	1.0
	(250, 974)	0.1	0.2
	(750, 974)	0.1	0.2
ωB97X	(99, 590)	65.1	97.9
	(250, 974)	15.5	42.4
	(750, 974)	8.5	35.0
M06-6X	(99, 590)	72.9	32.9
	(250, 974)	0.5	0.2
	(750, 974)	0.1	0.1
N12	(99, 590)	54.7	102.0
	(250, 974)	26.6	93.2
	(750, 974)	21.5	74.3

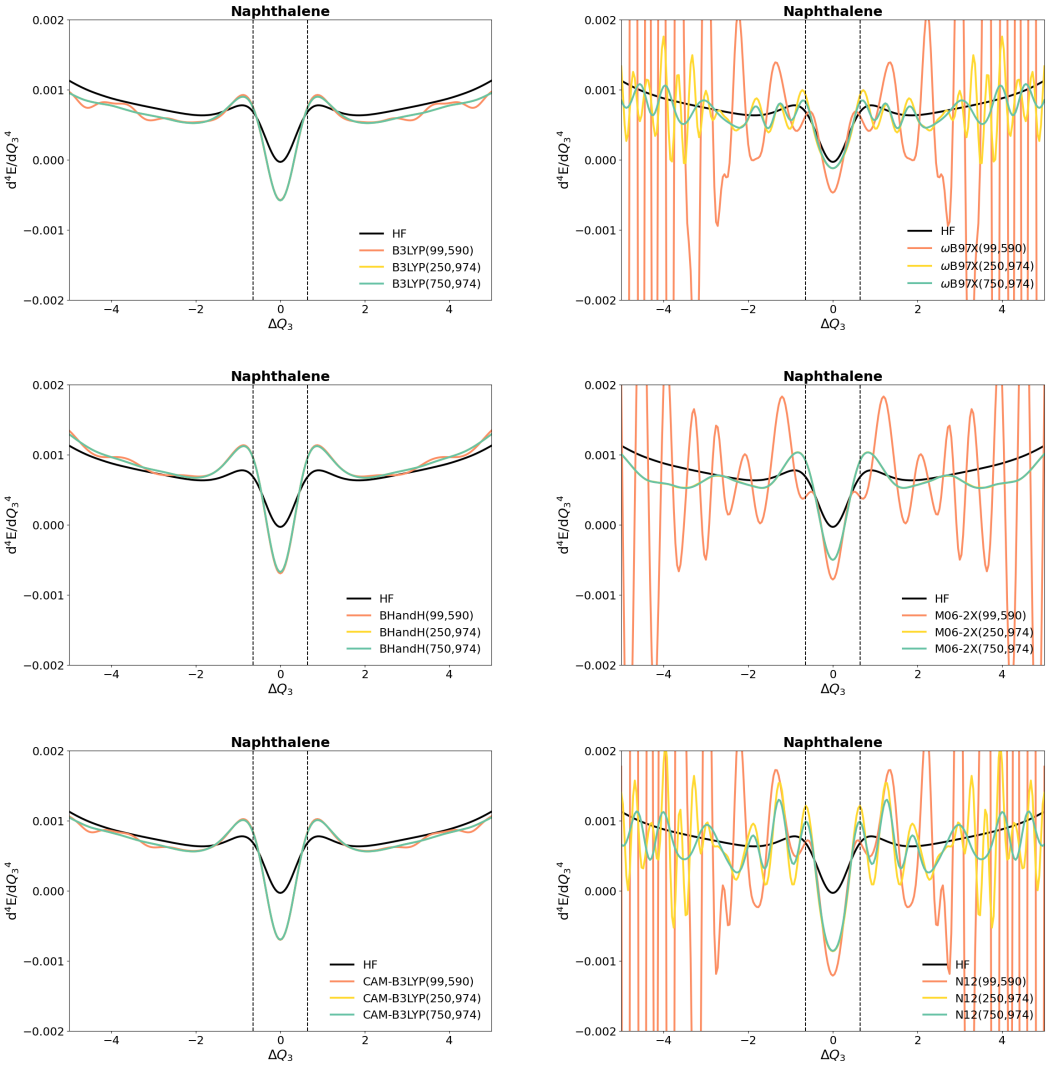

Figure A4.27: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_3^4$ for the displacements along the Q_3 normal coordinate of naphthalene. Left column shows the results for B3LYP, BH&H, and CAM-B3LYP, and right column for ωB97X, M06-2X, and N12.

A4.5 H_2O_2 (Q_1 mode)

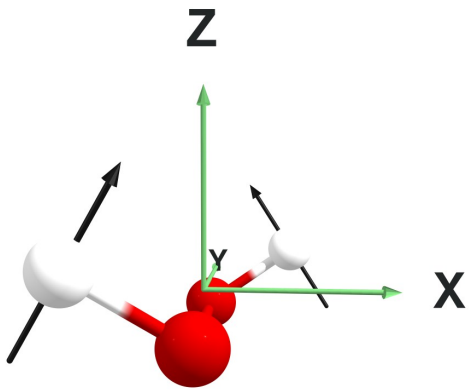

Figure A4.28: Graphical representation of the Q_1 (A) normal mode of H_2O_2. Obtained at M06-2X/aug-cc-pVDZ level of theory.

The results for H_2O_2 are compiled in Tables A4.19 and A4.20 and the oscillations are depicted in Figures A4.29 and A4.30, In H_2O_2 one can observe spurious oscillations in the curves of the highest derivative, namely $\mathrm{d}^4E/\mathrm{d}Q_1^4$, but the final RRMSE values are not that large. The values of RRMSE are below 15% even for the (99, 590) grid, except for M06-2X and MN15-L, for which they exceed 51% and 19%, respectively. In the case of the latter functionals, relative errors of same magnitude are also found for $\mathrm{d}^3E/\mathrm{d}Q_1^3$. One can also notice that the errors depend almost solely on the size of the angular grid, which was expected due to the character of displacements. For almost all properties and DFAs, the RRMSE values for the (250, 974) and (750, 974) are almost equal, despite the latter grid being three times more dense in the radial part than the former.

Table A4.18: Vibrational properties of the Q_1 normal mode of H_2O_2 obtained with some methods: ω_{har} and ω_{anh} — harmonic and anharmonic vibrational frequencies, $\mathrm{IR}_{\mathrm{har}}$ and $\mathrm{IR}_{\mathrm{anh}}$ — harmonic and anharmonic IR intensities, $\mathrm{RA}_{\mathrm{har}}$ and $\mathrm{RA}_{\mathrm{anh}}$ — harmonic and anharmonic Raman intensities. DFA computations utilized the (250, 974) integration grid.

	Q_1 (A) mode					
Method	$\omega_{\mathrm{har}}\left[\frac{1}{\mathrm{cm}}\right]$	$\omega_{\mathrm{anh}}\left[\frac{1}{\mathrm{cm}}\right]$	$\mathrm{IR}_{\mathrm{har}}\left[\frac{\mathrm{km}}{\mathrm{mol}}\right]$	$\mathrm{IR}_{\mathrm{anh}}\left[\frac{\mathrm{km}}{\mathrm{mol}}\right]$	$\mathrm{RA}_{\mathrm{har}}\left[\mathrm{Å}^6\right]$	$\mathrm{RA}_{\mathrm{anh}}\left[\mathrm{Å}^6\right]$
HF	412	344	191.151	193.166	0.053	0.063
B3LYP	374		165.073		0.055	
CAM-B3LYP	383		172.425		0.051	
BH&H	416		180.792		0.047	
LC-BLYP	394		186.539		0.047	
PBE	377		148.887		0.049	
PBE0	398		161.896		0.050	
TPSSh	379		157.529		0.056	
ωB97X	395		166.621		0.043	
ωB97XD	398		164.932		0.042	
M06-2X	392		171.213		0.062	
N12	372		172.413		0.047	
MN15	380		173.555		0.062	
MN15-L	392		151.094		0.058	

Table A4.19: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_1 normal mode of H_2O_2, obtained for some DFAs and integration grids.

Method	Grid	d^2E/dQ_1^2	d^3E/dQ_1^3	d^4E/dQ_1^4
B3LYP	(99, 590)	0.1	0.5	0.4
	(250, 974)	0.1	1.0	1.1
	(750, 974)	0.1	1.0	1.1
	(750, 3074)	0.0	0.0	0.0
	(750, 5294)	0.0	0.0	0.0
CAM-B3LYP	(99, 590)	0.1	0.5	0.3
	(250, 974)	0.1	0.9	0.9
	(750, 974)	0.1	0.9	0.9
	(750, 3074)	0.0	0.0	0.0
	(750, 5294)	0.0	0.0	0.0
BH&H	(99, 590)	0.1	0.8	0.9
	(250, 974)	0.0	0.1	0.0
	(750, 974)	0.0	0.1	0.1
	(750, 3074)	0.0	0.0	0.0
	(750, 5294)	0.0	0.0	0.0
LC-BLYP	(99, 590)	0.1	0.2	0.4
	(250, 974)	0.1	0.7	0.6
	(750, 974)	0.1	0.7	0.6
	(750, 3074)	0.0	0.0	0.0
	(750, 5294)	0.0	0.0	0.0
PBE	(99, 590)	0.2	3.1	3.2
	(250, 974)	0.2	2.2	2.8
	(750, 974)	0.2	2.2	2.8
	(750, 3074)	0.0	0.1	0.1
	(750, 5294)	0.0	0.0	0.1
PBE0	(99, 590)	0.3	2.6	2.5
	(250, 974)	0.2	1.7	2.2
	(750, 974)	0.2	1.7	2.2
	(750, 3074)	0.0	0.1	0.2
	(750, 5294)	0.0	0.0	0.1
TPSSh	(99, 590)	0.2	4.4	6.0
	(250, 974)	0.2	1.5	1.6
	(750, 974)	0.2	1.5	1.6
	(750, 3074)	0.0	0.3	0.8
	(750, 5294)	0.0	0.0	0.1

Table A4.20: Values of RRMSE (in percentages) for some property derivatives with respect to the Q_1 normal mode of H_2O_2, obtained for some DFAs and integration grids.

Method	Grid	d^2E/dQ_1^2	d^3E/dQ_1^3	d^4E/dQ_1^4
	(99, 590)	1.3	9.4	10.1
	(250, 974)	0.2	2.7	3.5
ωB97X	(750, 974)	0.2	2.7	3.5
	(750, 3074)	0.0	0.6	1.6
	(750, 5294)	0.0	0.0	0.1
	(99, 590)	1.7	12.2	13.9
	(250, 974)	0.3	3.5	4.8
ωB97XD	(750, 974)	0.3	3.5	4.8
	(750, 3074)	0.1	0.8	2.1
	(750, 5294)	0.0	0.0	0.2
	(99, 590)	5.6	65.4	51.0
	(250, 974)	3.5	13.9	34.4
M06-2X	(750, 974)	3.5	13.8	34.1
	(750, 3074)	0.2	2.9	6.7
	(750, 5294)	0.0	0.1	0.6
	(99, 590)	0.8	6.0	13.6
	(250, 974)	0.4	3.4	5.3
N12	(750, 974)	0.4	3.4	5.3
	(750, 3074)	0.1	1.7	4.1
	(750, 5294)	0.0	0.1	0.3
	(99, 590)	0.4	3.1	6.8
	(250, 974)	0.1	0.9	1.5
MN15	(750, 974)	0.1	0.9	1.5
	(750, 3074)	0.0	0.6	1.1
	(750, 5294)	0.0	0.0	0.0
	(99, 590)	1.7	16.1	16.3
	(250, 974)	1.4	14.6	19.3
MN15-L	(750, 974)	1.4	14.6	19.3
	(750, 3074)	0.0	0.5	0.8
	(750, 5294)	0.0	0.0	0.2

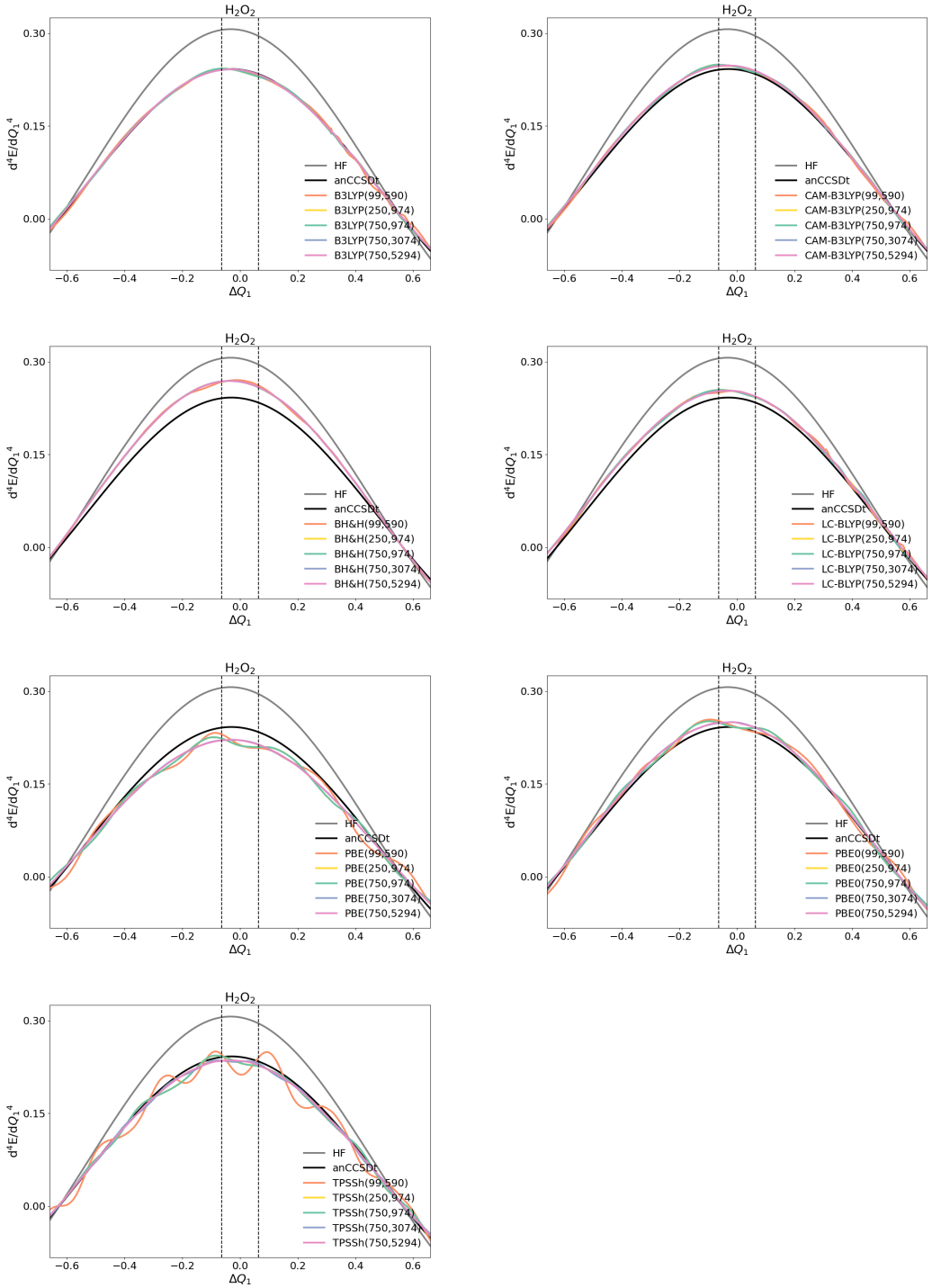

Figure A4.29: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_1^4$ for the displacements along the Q_1 normal coordinate of H_2O_2. Left column shows the results for B3LYP, BH&H, PBE, and TPSSh, and right column for CAM-B3LYP, LC-BLYP, and PBE0.

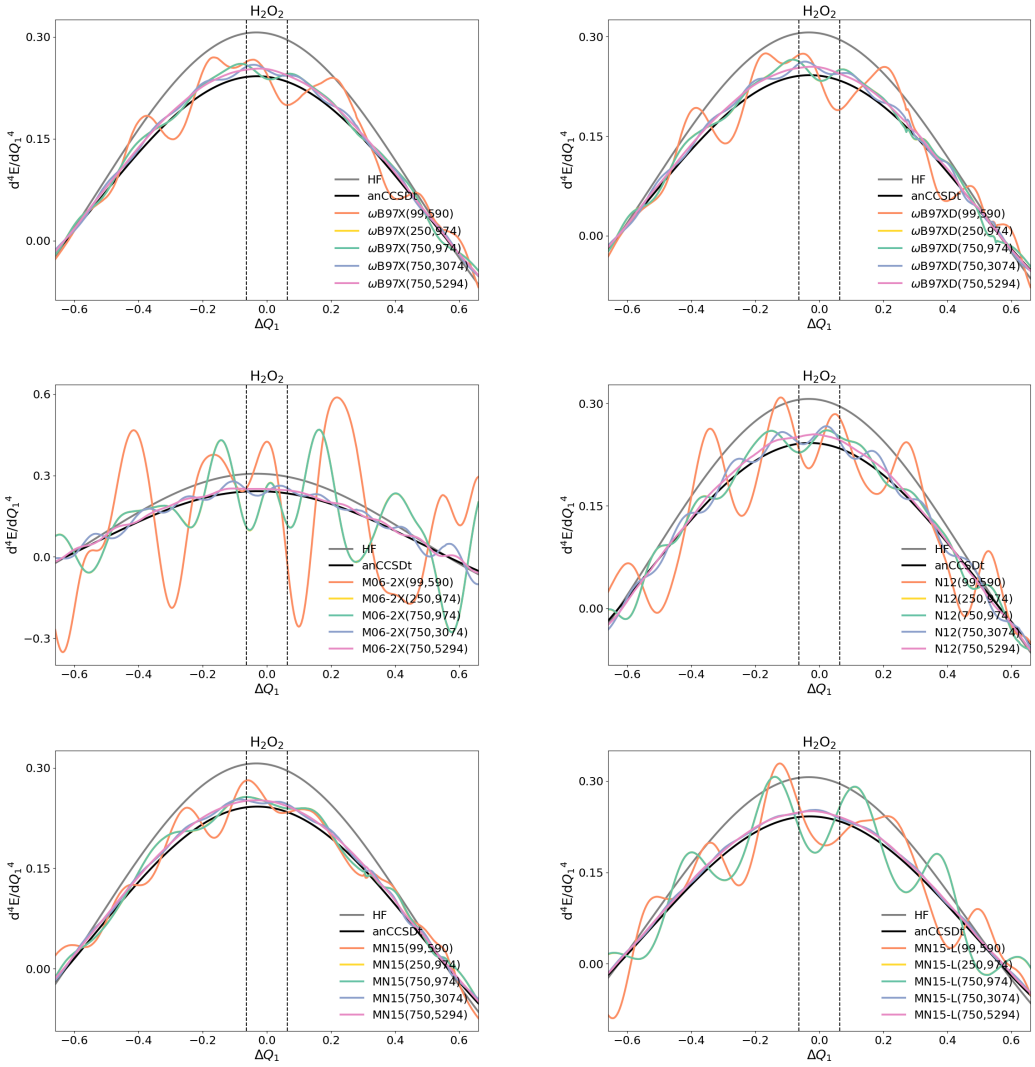

Figure A4.30: Spurious oscillations in $\mathrm{d}^4E/\mathrm{d}Q_1^4$ for the displacements along the Q_1 normal coordinate of H_2O_2. Left column shows the results for ωB97X, M06-2X, MN15, and right column for ωB97XD, N12, and MN15-L.