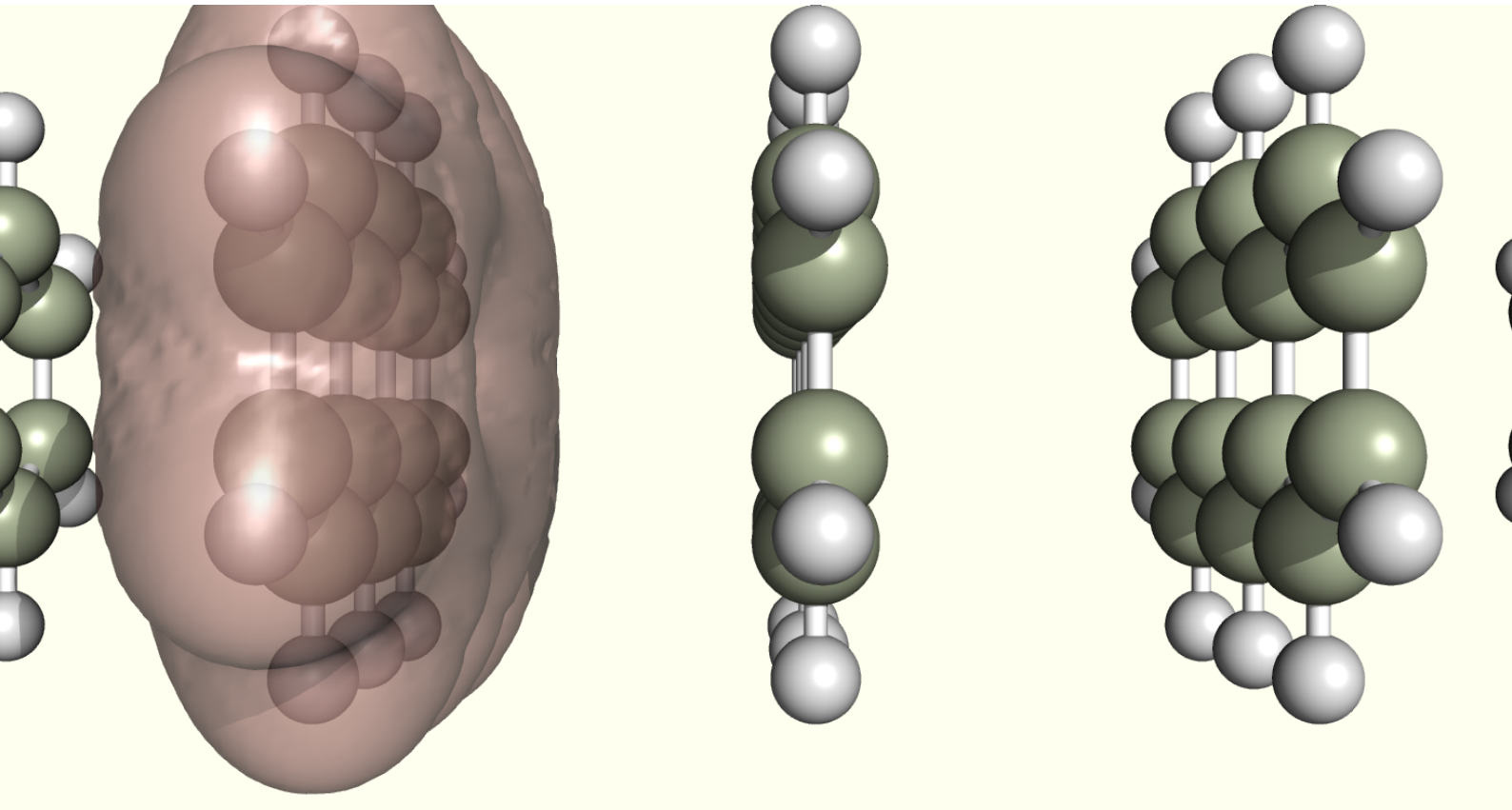
DISSERTATION

ACKS2 electronic polarization: a general dielectric response model

Patrick Gütlein



TECHNISCHE UNIVERSITÄT MÜNCHEN
Fakultät für Chemie
Lehrstuhl für theoretische Chemie

ACKS2 electronic polarization: a general dielectric response model

Patrick Gütlein

Vollständiger Abdruck der von der Fakultät für Chemie der Technischen Universität München zur Erlangung des akademischen Grades eines **Doktors der Naturwissenschaften (Dr. rer. nat.)** genehmigten Dissertation.

Vorsitzender: *Prof. Dr. Karsten Reuter*

Prüfer der Dissertation:

1. *Priv.-Doz. Dr. Harald Oberhofer*
2. *Prof. Dr. David Egger*

Die Dissertation wurde am 05.08.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Chemie am 18.09.2020 angenommen.

do not go gentle into that good night,
old age should burn and rave at close of day,
rage, rage against the dying of the light.

though wise men at their end know dark is right,
because their words had forkened no lightning, they
do not do gentle into that good night.

good men, they last wave by, crying how bright,
their frail deeds might have danced in a green bay,
rage, rage against the dying of the light.

wild men who caught and sang the sun in flight,
and learn, too late, they grieved it on its way,
do not go gentle into that good night.

grave men, near death, who see with blinding sight,
blind eyes could blaze like meteors and be gay,
rage, rage against the dying of the light.

and you, my father, there on the sad height,
curse, bless, me now with your fierce tears, I pray.
do not go gentle into that good night.
rage, rage against the dying of the light.

Dylan Thomas

Preface

This thesis is a publication-based dissertation implying that original research was published in international scientific journals, references [1] and [2], associated articles are appended to this work. Brief summaries of these studies are accompanied by details of the theoretical foundation, further methodological aspects and additional relevant literature.

All work presented in this doctoral thesis was performed between August 2016 and July 2020 at the Chair of Theoretical Chemistry of the Technical University of Munich (TUM), under the supervision of PD Dr. Harald Oberhofer and Prof. Dr. Karsten Reuter. Research stays hosted by Prof. Dr. Jochen Blumberger at the Department of Physics & Astronomy at the University College London (UCL) funded by the TUM Institute for Advanced Study (IAS) complemented the work performed locally at TUM.

München, July 2020

Abstract

Quantum mechanic methods in computational chemistry provide invaluable insights into atom-scale reaction processes and material properties, a fundamental contribution to research and development of current and future technologies. Unfortunately, accurate wave function techniques are limited to simulations of a comparably small number of particles and short time-scales, which is preventing the theoretical investigations of many chemically relevant environments and (dynamical) physical effects in condensed matter. Alternatively, approximate interaction potentials in force field techniques provide a very efficient sampling of the atomistic phase space and significantly reduce the simulation cost and computational time in atom-level material calculations, molecular dynamics and Monte Carlo simulations. Therein, a set of inexpensive mathematical functional forms—historically often motivated by truncated expansion series of quantum mechanical expectation values—represent the total energy, atomic forces and other material properties. A careful choice and validation of the specific description and parametrization implementation are mandatory to avoid an incomplete or even false foundation of the physical effect at scrutiny. One major contribution to the total energy in force fields is the long-range electrostatic interaction, typically considered between atoms that are not connected or separated by a few chemical bonds. An estimate of the electronic charge density distribution and its total Coulomb energy is computationally unfeasible in molecular mechanics, as it requires an (approximate) charge distribution which fundamentally rely on quantum mechanical wave function methods. Instead, they are replaced by simplified atom-related electrostatic interaction potentials like multicenter multipole expansion series of the average atomic charge density, often derived from molecules in vacuum. Deviations from this permanent charge representation due to changes of the local electric field induced by nearby molecules or charges carriers in dense matter can be treated explicitly by additional polarization functions. In this thesis, a recently developed first-principles-based electronic polarization model, the atom-condensed Kohn-Sham density functional theory approximated to second order (ACKS2), has been investigated and was advanced towards practical application as dielectric response extension to force fields. A new minimalistic Cartesian Gaussian function basis set representation and condensation scheme was introduced and optimized in an initial proof-of-principle study. The ACKS2 molecular dielectric response, polarization energy and induced dipole moments, are validated by the underlying density functional reference. A test set of small hydrocarbons and aromats perturbed by a single homogeneous electric field or point charge potentials shows great accuracy, while maintaining beneficial model transferability and numerical efficiency. Next, two flavors of a new fragmentation approach in ACKS2 allow the evaluation of explicit polarization contributions to intermolecular interactions, which is ideally suited for force field techniques. Both techniques again exhibit good agreement for charge-constrained (neutral or charged) dimers of small aromats with their density functional theory reference, concluding with the study of electron-hole pair dissociation and its polarization stabilization in a simple one-dimensional toy model.

Zusammenfassung

In der Computer-gestützten Chemie bieten quantenmechanische Methoden unschätzbare Erkenntnisse zu den Materialeigenschaften und Prozessen chemischer Reaktionen auf atomarer Ebene. Dieses Verständnis ist ein grundlegender Bestandteil in der Erforschung und Entwicklung derzeitiger und zukünftiger Technologien. Bedauerlicherweise sind exakte Methoden, die die Schrödingergleichung mit Hilfe von Wellenfunktionen lösen, rechnerisch sehr aufwendig und dadurch begrenzt auf Simulationen mit einer vergleichweise kleinen Anzhal an Teilchen (Elektronen) und kurzen Zeitskalen. Eine Computer-gestützte Untersuchung vieler chemisch-relevanter Milieus und (dynamischer) physiko-chemischer Effekte in kondensierter Materie ist so leider nicht durchführbar. Dagegen ermöglichen angenäherte Wechselwirkungspotentiale, die in Kraftfeld-Method verwendet werden, eine sehr effiziente Beschreibung des quantenmechanischen Phasenraums. So kann der mathematische Aufwand und die Rechenzeit am Computer für die atomistische Modelierung erheblich reduziert werden, wie beispielsweise in Molekulardynamik oder Monte-Carlo Simulationen. Molekulare Kraftfelder verwenden numerisch effiziente mathematische Funktionen (historisch häufig mit unvollständigen Entwicklungssätzen quantenmechanischer Erwartungswerte begründet) um die totale Energie, atomare Kräfte und andere Materialeigenschaften zu beschreiben. Eine sorgfältige Auswahl und Validierung der jeweiligen mathematischen Darstellung und Parameterisierung sind dabei unerlässlich um eine unvollständige oder gar falsche theoretische Grundlage der physikalischen Effekte zu vermeiden. Ein großer Anteil der Gesamtenergie in diesen Kraftfeldern ist die lang-reichweitige elektrostatische Wechselwirkung. Sie wird typischerweise nur für Atome berücksichtigt, die nicht direkt verbunden beziehungsweise durch mehrere chemische Bindungen getrennt sind. Eine Schätzung der Ladungsverteilung zur Berechnung der Coulomb-Energie ist rechnerisch in Kraftfeldern nicht durchführbar, da die (exakte oder genäherte) Elektronendichte nur mit Hilfe von aufwendigen Wellenfunktionsmethoden berechnet werden kann. Stattdessen werden sie durch vereinfachte atom-zentrierte elektrostatische Potentiale ersetzt, wie beispielsweise eine (unvollständige) Entwicklung in einer Multizentrum-Multipol Darstellung. Als Referenz hierfür dient eine mittlere atomare Ladungsdichte, welche häufig auf Molekülen im Vakuum oder verschiedenen chmeischen Umgebungen basieren. Abweichungen der Ladungsverteilung von dieser permanenten gemittelten Darstellung, aufgrund lokaler Änderungen im elektrischen Feld durch andere Moleküle oder Ladungsträger in kondensierter Materie, können explizit mit Hilfe sogenannter Polarisationsmodelle verarbeitet werden. Der Fokus der vorliegenden Dissertationsschrift liegt auf einem kürzlich entwickeltem expliziten Polarisationsmodel, der atom-kondensierten Kohn-Sham Dichtefunktionaltheorie angenähert zur zweiten Ordnung (Original: *atom-condensed Kohn-Sham density functional theory approximated to second order (ACKS2)*). Diese Methode wurde untersucht und weiterentwickelt, um eine allgemeine praktische Anwendung zur Berechnung expliziter elektronischer Polarisation in Kraftfeldern zu ermöglichen. Hierfür wurde in einer ersten Grundlagenstudie eine neue minimalistische Darstellung in karthesischen

Gaußfunktionen eingeführt und optimiert. Die ACKS2 molekulare dielektrische Antwort, Polarisationsenergie und induzierte Dipolmomente, wurden mit Hilfe der zugrundeliegenden Dichtefunktionaltheorie validiert. Dazu wurde ein die Elektronendichte eines Satzes kleiner Kohlenwasserstoffe und Aromaten mit Hilfe eines simplen elektrostatischen Potentials, einem einzelnen homogenen elektrischen Feld oder einer einzelnen Punktladung, gestört. Dabei zeigte das Gaußsche ACKS2 Model gute Genauigkeit der Polarisationseigenschaften unter Beibehaltung der vorteilhaften Übertragbarkeit und numerischer Effizienz. Im Anschluss wurden zwei Unterarten eines neuen Fragmentierungsschemas in ACKS2 implementiert, um die Beiträge der expliziten Polarisation zur intermolekularen Wechselwirkung berechnen zu können, was sehr gut für die Anwendung in molekularen Kraftfeldern geeignet ist. Beide Techniken weisen sehr gute Übereinstimmung der elektronischen Polarisation mit der Dichtefunktional Referenz für Ladungsbeschrṅkte elektronische Zustände (neutraler oder geladener) molekularer Dimere, bestehend aus kleinen Aromaten. Zum Abschluss zeigt eine Studie der Polarisation während der Ladungstrennung in einer eindimensionalen Ketten von Anthracen-Molekülen illustrativ die praktische Anwendung auf.

Contents

1 *Introduction.* 1

2 *Explicit charge polarization in force field methods.* 5
- 2.1 *The Drude oscillator.* 5
- 2.2 *The induced dipole.* 6
- 2.3 *Charge equilibration.* 6
- 2.4 *Non-additivity of explicit polarizability.* 9

3 *First-principles foundation of the ACKS2 general electronic polarization model.* 13
- 3.1 *Wave-function augmented formulation of density functional theory.* 13
- 3.2 *Perturbative density functional theory.* 14
- 3.3 *Atom-condensed parametrization scheme.* 15
- 3.4 *Semi-local exchange-correlation functionals in ACKS2.* 16
- 3.5 *Dielectric response in intermolecular interactions.* 17
- 3.6 *Electronic charge rearrangements of atoms in a molecule.* 18

4 *Density functional tight-binding.* 21

5 *Publications.* 25
- 5.1 *Toward First-Principles-Level Polarization Energies in Force Fields: A Gaussian Basis for the Atom-Condensed Kohn-Sham Method.* 25
- 5.2 *An iterative fragment-scheme for the ACKS2 electronic polarization model: Application to molecular dimers and chains.* 27

6 *Conclusion & Outlook.* 29

7 *Acknowledgments* 31

A *Electrostatic interactions in multipole expansion.* 33
- A.1 *Series expansion of the classical electrostatic potential.* 33
- A.2 *The multipole electrostatic interaction energy.* 35

B *Density functional theory foundation of charge equilibration.* 37
- B.1 *Derivation of the chemical potential equalization method.* 37
- B.2 *Comparison of CPE to EEM.* 38

C *Formula collection for ACKS2 derivation.* 41
- C.1 *Derivatives related to density functional theory.* 41
- C.2 *Second order energy expansion of electronic response.* 42
- C.3 *Atomic forces in ACKS2.* 43

D *Bibliography* 45

E *Full excerpts of the peer-reviewed scientific publications.* 53
- E.1 *Publication #1* 53
- E.2 *Publication #2* 64

1. Introduction.

THEORETICAL AND COMPUTATIONAL CHEMISTRY focuses on the mathematical and computer-supported modeling of physical and chemical effects of matter on an atomic scale. These methods provide information about atoms in molecules and condensed matter, which can help to understand current (or design new) experimental studies and materials—a crucial feat in a technological world. The application of computational simulations in chemistry and related fields (like biophysics or engineering) is very diverse. For example, computational studies aid in the understanding of protein and DNA folding to predict their structure and functionality [3, 4]. Atom-scale investigations of (chemical) reactivity at surfaces and interfaces are fundamental to heterogeneous catalysis [5] and electrochemistry [6].

The foundation of atom-scale physics rests in quantum mechanics. Typically, the stationary (time-independent) Schrödinger equation is employed and atoms are represented by wave-functions. In many systems, the electrons move orders of magnitudes faster than nuclei, and hence the nuclei appear to be static to the electrons. This justifies a separation of electronic and nuclear degrees of freedom, reducing the latter to a structure-dependent parameter in the electronic Schrödinger equation. At the same time, the nuclei of atoms are typically treated classically as non-massless point charges, and only electrons are treated in a wave-function framework. Additionally, electrons are assumed to remain in the same electronic state (ground state) upon molecular motion, implying weak coupling between the different degrees of freedom. This collection of assumptions is called the adiabatic Born-Oppenheimer (BO) approximation [7]. The solution to the electronic Schrödinger equation yields the energy of the electronic state as function of the nuclear positions, which represents a potential energy (hyper)surface (PES) for the nuclear motion. In theoretical investigations of atom-scale systems, the BO-PES provides forces acting on the different atomic nuclei and determines the (time-dependent) structure of atomic arrangements. In turn, the latter determines other static observables (e.g. bulk modulus) or dynamic (transient) observables (e.g. chemical species on a surface). It is not an overstatement to locate the potential energy surface at the very heart of modern computational chemistry. Unfortunately, usage of wave-function methods and calculation of the PES from the electronic Schrödinger equation is numerically expensive due to the computational cost scaling with the number of electrons in the simulation cell. The wave-function based interaction operator in the Schrödinger equation contains two-electron four-center integrals and hits the limits of modern computers already for a small number of atoms. In consequence, computational studies based on wave-function techniques allow only a small number of atoms (electrons) and short simulation times. This precludes investigations of extended system sizes, which is necessary e.g. for representations of extended and complex surfaces including bulk materials. It also stops any studies of large time scales to capture slow processes (e.g. catalyst degradation) or rare events (transition between meta-stable states). Hence, theoretical and computational chemistry need alternatives to bridge the gap and utilize computationally less expensive energetics.

Here, force field (FF) techniques provide an invaluable set of tools to the modern materials' modeling community. They are designed for efficient sampling of the phase space of condensed matter, thereby enabling study of systems too large and simulation times too long to be tackled by first-principles. Their numerical speedup compared to *ab-initio* calculations is due to simpli-

fied electronic interactions between different atoms and ions, which are typically either trained on extensive high-level first-principles references or empirical properties from experimental studies. The various effective inter-atomic potential contributions to the total energy in a force field are treated on different footings, often transformed to simple mathematical functional forms by physical and chemical heuristics. Historically, the force field total energy comprises chemical bond related energy terms as well as contributions between atoms which are not covalently bonded (non-bonding), as illustrated for the carbon backbone of naphthalene molecules in fig. 1,

$$E^{\mathrm{FF}} = E^{\Delta d} + E^{\Delta \alpha} + E^{\Delta \beta} + E^{\mathrm{el.stat.}} + E^{\mathrm{disp.}} + E^{\mathrm{cross.}}. \tag{1.1}$$

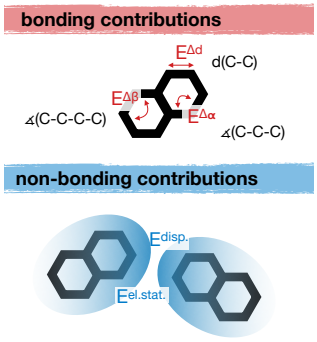


Figure 1 Illustration of the energy contributions in force field techniques for the carbon backbone of a naphthalene molecule/dimer.

In conventional force fields [8–14], the different energy contributions often employ truncated series expansion representations of simple functional forms. In practice, the utility of force fields is facilitated by application of a chemical reference structure during the parametrization process, where e.g. the inter-atomic distance distance d is expressed with respect to a zero-order reference distance, $\Delta d{=}d{-}d_0$. Hence, the bonding energy terms often employ a quadratic functional form, partly supplemented by small higher order corrections, for bond lengths Δd and bond-bond angles $\Delta\alpha$: $E^{\Delta d} \propto \Delta d^2$ and $E^{\Delta \alpha} \propto \Delta \alpha^2$. The bond-dihedral rotation angle $\Delta\beta$ utilizes a fourier series expansion as simple rotation functional form: $E^{\Delta \beta} \propto \cos\left(n_\beta \Delta \beta\right)$, where n_β represents the multiplicity of the rotational symmetry. The total electrostatic interactions are typically projected onto simplified atom- or bond-related charge distributions, most prominently represented by Coulomb interactions between point charges q_i and q_j with distance $r_{i,j}$, $E^{\mathrm{el.stat.}} \propto {}^{q_i q_j}/_{\Delta r_{i,j}}$. These charge representations are fixed and pre-defined for specific atom types to maintain computational efficiency. Additional degrees of freedom to account for explicit electronic charge rearrangments in simulations are provided in a sub-class, so-called polarizable force fields, by e.g. atom-centered inducible dipoles [15–18]. In close relation to the electrostatic interactions, long-range attractive London dispersion effects are expressed as an inverse power expansion series, $E^{\mathrm{disp.}} \propto \sum_n {}^{C_n}/_{\Delta r^n}$, typically corrected by a short-range repulsive term [19, 20]. More sophisticated force field implementations provide additional cross-coupling terms between different energy contributions collected in $E^{\mathrm{cross.}}$ [21]. The truncated expansion and simple functional forms of force fields are typically sufficient to capture most of the total energy during atomic motions, that are similar to the chemical reference morphologies, at great numerical efficiency. Yet, it leads to fixed atomic types and pre-defined chemical connectivity. In contrast, the modeling of chemical bond formation and breaking in terms of transition state theory and complex reaction mechanisms is an elaborate task and requires more sophisticated force field formulations, accompanied by increased computational cost [22]. While the approximate form of the total energy description varies between different effective interaction and energy approaches, in particular with the advent of more flexible functional forms given by machine-learned potentials [23–26], the separation of short-range (bonding) and long-range (electrostatic and dispersion) interactions typically still prevails. The success of force field techniques is confirmed in various applications and research compartments within the materials' modeling community [27–29], here

illustrated by a few recent examples. In a study of battery materials and energy storage, Heenen *et al.* developed a polarizable force field from density functional theory (DFT) reference energies, forces and stresses to study the ion conductivity in antiperovskite glasses [30]. Giannini *et al.* scrutinized polaron mobility and charge transport in a (one-dimensional) model of organic semiconductor materials, where a classical non-polarizable force field provides the energies of the molecular sites [31, 32]. A work by Futera *et al.* [33] focuses on an extension and benchmark of a reactive force field implementation versus DFT for the individual adsorption of all natural amino acids on a gold surface, showcasing the validity of approximate interaction potentials in biomolecular systems.

In atomistic simulations, the (long-range) classical electrostatic interaction energy between two charge densities ρ and ρ' is one of the key ingredients (due to long-range) in an accurate material property calculation, $E^{\text{el.stat.}} = \iint {}^{\rho\rho'}/_{|r-r'|} dr dr'$. Yet, developing a physically reliable and numerically efficient model is a difficult task, as the exact charge density of the atomic structure is unknown in the absence of the (computationally expensive) electronic structure and spatial integration of complex distributions is computationally prohibitive in force fields. A multicenter multipole expansion—different flavors deploy cartesian, polar or internal coordinates—of the classical electrostatic interaction simplifies the interaction integral and reduces the numerical cost while maintaining reasonable accuracy. The true charge density is typically estimated by approximate permanent atomic multipole moments, which in its simplest form constitutes an ideal point charge model, $E = {}^{q_i q_j}/_{\Delta r_{i,j}}$ These point charges are typically corrected for short-range interactions due to the non-ideal nature of atomic charge distributions by Gaussian- or Slater-type functions [34, 35]. A detailed derivation of an ideal Cartesian multicenter multipole expansion of the electrostatic interaction energy is given in appendix A.

The concept of atom-projected properties, like approximate electrostatic interaction potentials, derived from first-principles calculations is well established by various atom-in-molecules and (atom-projected) charge partitioning schemes. Prominent examples of this include e.g. the distributed multipole analysis [36], Bader charge concentrations [37], Hirshfeld atom-in-molecules charges [38], Mulliken population analysis [39] and (restrained) electrostatic potential derivation [40, 41]. Alternatively, more empirical approaches treat the atom-centered multipole moments as independent parameters optimized to match observables, either . Throughout the past decades, approximate atom-centered first order mutipole representations (ideal point- , Gaussian- or Slater-type charges) of the electrostatic energy were the dominant choice for theoretical investigations of condensed phase matter, in particular motivated by their simplicity and numerical efficiency [8–13]. Due to the ambiguous nature of the truncated multipole expansion combined with various partitioning schemes, the (atom-projected) charges in different force field parametrizations lack considerably in transferability [42–44]. Furthermore, recent discussions of both the static, isotropic nature of the point charge and lack of directionality in their respectivey potentials [14, 45–48] prompted the development of more sophisticated models involving point charges distributed outside the nuclei or atom-centered higher angular momentum terms like dipole or quadrupole (or rarely even octopole) moments [49, 50].

Going even further, incorporation of the electronic polarization and charge transfer of atomic structures induced by local electrostatic potential changes, e.g. due to the motion of close molecules,

excess charge carriers or external biases, is not a simple task [51–53]. Their contributions are typically rather small compared to the total interaction energy, yet, they can strongly influence the local processes in condensed phase matter, in particular for weakly-bound systems, for example in protein folding [54, 55] or excitation and charge transport in organic semiconductors [56, 57]. The necessity of explicit electronic polarisation can most easily be illustrated on e.g. the binding energy curves in pi-cation systems [58]. Large portions of the attractive electrostatic interactions between benzene—a non-polar, but highly polarizable molecule—and a potassium ion are mediated by the electronic polarization, which oversimplified representations by fixed partial charges cannot capture. In another example, a recent study predicts that the energetic stabilization of electron-hole pairs by (explicit) electronic polarization is crucial during charge separation at donor-acceptor junctions in organic solar cells [57]. Unfortunately, explicit polarization introduces more model intricacies (choice of functional form, parameters, etc.) into force fields, initially displaying rather mixed success compared to well-tuned non-polarizable electrostatic models [45, 59–61]. The latter frequently employ *ad hoc* corrections to fixed-charge electrostatic interaction representations derived from gas-phase molecules, typically at the cost of seriously limited transferability of parameters [62]. For example, one study suggests the overestimation of the individual charge magnitudes by about 10-20% to account for increased polarization in condensed phase matter [43], whereas another introduces reduction by about 30% for ions to account for electronic screening effects [63].

In this thesis, various polarization extensions to force field techniques are briefly discussed, followed by an in-depth discussion of the methodological details and development of a general dieletric response model based on the atom-condensed Kohn-Sham DFT (KS-DFT) approximated to second order.

2. Explicit charge polarization in force field methods.

MODELS for an explicit, yet efficient charge polarization in force fields can roughly be grouped in three main classes: i) classical Drude oscillators, ii) induced point dipoles, iii) charge equilibration. This chapter is dedicated to a brief summary of the fundamental ideas of the first two dielectric response classes and a more elaborate discussion of the various charge equilibration methods, as they are conceptual predecessors to the general ACKS2 electronic response model. For the interested reader, a good introduction into the approximate treatment of explicit charge polarization in force fields, dicussing model origins, code implementations, parameter evaluation and practical applications is provided by these publications [15–18].

2.1. The Drude oscillator.

The Drude oscillator comprises a virtual charged particle attached to the core of a (typically non-hydrogen) atom by a harmonic spring to account for intra-atomic electronic polarization [64, 65], sometimes also referred to as charge-on-spring, shell or core-shell model. The shell particle of atom i carries a charge of $q_{i,D}$, and its core typcially carries a charge of $q_{i,0}{=}q_i{-}q_{i,D}$ to maintain the total atomic charge q_i. The force constant, $k_{i,D}$, of the harmonic spring between both charges is connected to the atomic polarizability α_i, $k_{i,D}{=}\frac{q_{i,D}^2}{\alpha_i}$. Any electronic polarization effects due to electrostatic interaction potentials are covered by the displacement of the shell charge relative to its respective core. Note, that the Drude oscillator does not incorporate any atom-atom charge transfer. The computational cost of the Drude oscillator polarization representation in FF scales with simple charge-charge interactions between the core and shell of two particles. The energy of a Drude oscillator follows simple charge-charge interactions between core and shell of two different atoms, i and j, and the harmonic spring energy,

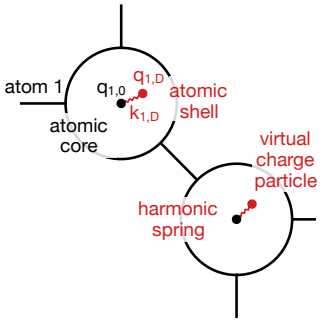


Figure 2 Illustration of the Drude oscillator model for efficient polarization extentions in force fields.

$$E_{\text{Drude}} = \sum_{i<j} \frac{q_{i,0}q_{j,0}}{|\boldsymbol{r}_{i,0} - \boldsymbol{r}_{j,0}|} + \sum_{i<j} \frac{q_{i,D}q_{j,D}}{|\boldsymbol{r}_{i,D} - \boldsymbol{r}_{j,D}|} + \sum_{i,j} \frac{q_{i,0}q_{j,D}}{|\boldsymbol{r}_{i,0} - \boldsymbol{r}_{j,D}|} + \sum_{i} \frac{1}{2} k_{i,D} |\boldsymbol{r}_{i,0} - \boldsymbol{r}_{i,D}|^2. \tag{2.1}$$

Hence, the computational cost of the Drude oscillator model is about four times the computational cost of the fixed point charge non-polarizable electrostatic interaction calculations. The Drude model gained significant popularity in the biological and organic molecular simulation community in the last two decades [66–70]. For example, proper explicit classical water models are indispensable to produce reliable solvation of molecules in a biological environment. The dielectric response of water is accomplished either by a single charge-on-spring at the oxygen [71, 72], several distributed Drude oscillators at each nucleus and lone pair [73] or a charge-transfer extended Drude oscillator [74].

2.2. The induced dipole.

This model comprises an (ideal) inducible dipole located at each nuclear position to capture electronic polarization effects [75], sometimes also referred to as distributed (dipole) polarizability model. The local electric field acting on atom i due to its environment prompts a dielectric response in form of an induced point dipole $\boldsymbol{\mu}_i$ mediated by its effective (isotropic) polarizability α_i,

$$\boldsymbol{\mu}_i = \alpha_i \left[\boldsymbol{E}_i + \sum_{j \neq i} \boldsymbol{T}_{i,j} \boldsymbol{\mu}_j \right] \tag{2.2}$$

$$E_{\text{ind.dip.}} = \frac{1}{2} \sum_i \boldsymbol{\mu}_i \boldsymbol{E}_i. \tag{2.3}$$

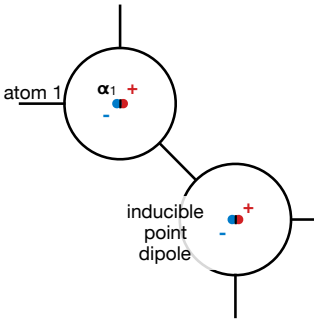


Figure 3 Illustration of the inducible point dipole model for efficient polarization extentions in force fields.

The electric field contribution $\boldsymbol{E}_i$ accounts for the *permanent* charge or multipole representation of the fixed electrostatic interactions, and $\boldsymbol{T}_{i,j}$ provides mutual dipole-dipole interactions between different polarizable sites. Typically, these contributions are corrected for short-range effects, like the non-ideal, continuous nature of the electron distribution including charge accumulation (variable local distribution shape) and charge penetration (charge density overlap), to improve the energetics by short-range dampening functions or non-ideal dipole representations [76–79]. Otherwise this can lead to numerical instabilities of ever-increasing dipole moments called polarization catastrophe. The computational cost scales with the charge-charge, charge-dipole and dipole-dipole interaction evaluation. The model of induced atomic dipoles (or distributed atomic polarisabilities) bears many conceptual similarities to the previously introduced Drude oscillator, which has been studied and compared to non-polarizable force fields [70, 80, 81]. In order to improve the overall (long-range) electrostatic interaction energy, induced dipole models have been extended by higher angular momentum terms of the permanent [52] or inducible [82, 83] distributed multipole representation. Note, the Drude oscillator and inducible dipole models still only represent simplified and efficient intra-atomic polarization contributions. They are able to account for overall molecular dipole (and higher angular) moments, but essentially lack an explicit inter-atomic charge transfer description. Combining induced dipole models with fluctuating charge models, see next section, establishes a computational framework to account for explicit intra-atomic polarization as well as inter-atomic charge transfer [84].

2.3. Charge equilibration.

Charge equilibration (QEq) techniques provide a set of tools to estimate atomic partial charges in molecules as function of their topology and molecular conformation. Conceptually, they slightly differ from both previous sets of explicit polarization models, which only allow an intra-atomic response. QEq models retain the same functional structure of the electrostatic energy of non-polarizable simple fix-charged models—an atom-centerd point charge representation. However, in QEq the atomic partial charges are not fixed *a priori*, i.e. pre-determined parameters fitted to first-principles or empirical data. Instead they are flexible and adapt on-the-fly during molecular

dynamics or Monte-Carlo simulations in terms of atom-atom charge transfer prompted by the molecular topology or conformational changes.

Following the ideas of Sanderson on chemical bond formation, charge density (and atomic electron population) flows between atoms until the electronegativity is equal in the entire molecule, i.e. their electronegativity converges to a common intermediate value [85]. A theoretical foundation is given by Iczkowksi and Margrave [86], who relate the free-atom (or ion) electronegativity χ to the change of total energy E as a function the electronic population N of an atom, $\chi = -{}^{dE}/_{dN}$. A similar description, yet closer to an accurate first-principles formalism and the finite nature of electrons, by Parr *et al.* provides a quantum mechanical footing and clear mathematical foundation for the electronegativity (and chemical potential μ) concept, $\mu = {}^{dE}/_{d\rho} = -\chi$ [87]. Mortier *et al.* develop the foundation for the (atomic) electronegativity equalization method (EEM) in two seminal papers [88–90] (a less conceptual, more fundamental derivation based on a density functional theory formalism by York and Yang [91] is given in appendix B). The authors of EEM expand the electronic energy of a free-atom A as a function of its atom-projected electronic population, $N_A = \int \rho_A dr$, up to second order in a simple point-charge representation. A slightly more convenient form of atomic partial charges q_A instead of electronic population, following the transformation $q_A = (Z_A - N_A)e$, where Z_A is the atomic nuclear charge and explicit mention of the elementary charge e will be disregarded for the sake of brevity, yields

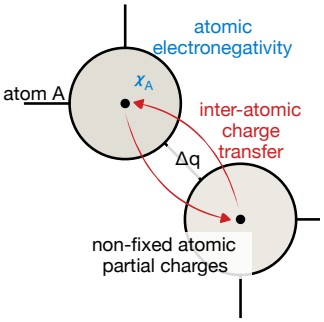


Figure 4 Illustration of equilibration methods by atom-atom charge transfer.

$$E_A(q_A) = E_A^0 + \frac{\partial E}{\partial q_A} q_A + \frac{1}{2} \frac{\partial^2 E}{\partial q_A^2} q_A^2 \tag{2.4}$$

$$= E_A^0 + \chi_A^0 q_A + \frac{1}{2} \eta_{A,A}^0 q_A^2. \tag{2.5}$$

The first term comprises the total energy of a free-atom in an unperturbed reference state—typically a neutral atom, but can in principle be extended to ions. The first and second order energy derivatives are identified with the atomic electronegativity, $\chi_A^0 = {}^{\partial E_A}/_{\partial q_A}$, and atomic hardness [89] (or idempotential [92]), $\eta_A^0 = {}^{\partial \chi_A}/_{\partial q_A} = {}^{\partial^2 E}/_{\partial q_A^2}$. The superscript index 0 indicates a free-atom property (opposed to an atom in a molecule). Similarly, the total energy (not only the classical electrostatic contributions) of a molecule can be expressed as function of its total charge. Projection of the molecular charge distribution onto the atom centers, i.e. an atom-in-molecule partial charge representation, results in

$$E^{\text{mol.}}(q_A, q_B, \cdots, q_K) = \sum_A (E_A^* + \chi_A^* q_A + \frac{1}{2} \eta_{A,A}^* q_A^2) + \sum_{B<A} q_A \eta_{A,B}^* q_B. \tag{2.6}$$

The first three terms account for the (one-body) single free-atom-like contributions to the total energy as function of its charge, describing the energy gain or cost for charge accumulation and depletion at a site. The last contribution represents interactions between different free-atom-like sites in the same molecule. They are typically reduced to simple Coulomb interactions of the atom-centered partial charges corrected for short-range interactions. Inclusion of additional terms—not shown here—could account for further perturbations, like e.g. the presence of other molecules introduced by $q_A \eta_{A,A'}^* q_{A'}$, where A' denotes an atomic site in another molecule. The ground state

charge distribution of a molecule is determined by the energy minimum with respect to the individual atomic partial charges, $\partial E^{\text{mol.}}/\partial q_A$. The latter is equivalent to impose an electronegativity constraint onto the different sites, which inspired the authors to call this method electronegativity equalization method. A final constraint enforces charge neutrality, $\sum_A q_A$-q=0, where q is the net charge of the molecule. Note, the constraint of charge conservation can be adjusted to allow or prohibit charge transfer between different molecules.

During chemical bond formation, the atomic density distribution changes quite vividly. The pure free-atom character of the unperturbed atomic charge sites decreases and the atom-in-molecule (AIM) character increases, which influences all atomic properties including electronegativity and hardness, $\chi_A^0 \rightarrow \chi_A^{\text{AIM}}$ and $\eta_{A,B}^0 \rightarrow \eta_{A,B}^{\text{AIM}}$. Despite ambiguities, theoretical atom-in-molecule representation is a mathematically sound concept and basically projects the electronic structure information of a molecule (or other chemically bonded systems) onto atom centers. Practically, the accurate calculation of a faithful charge partitioning function from first-principles on-the-fly produces large numerical overhead and can generally not be transferred between different atomic structures. Hence, the usage of exact AIM parameters is computationally prohibitive in any practical simulation and force fields resort to approximate schemes of parameter scales to resolve the transition between free-atoms and atoms-in-molecules. In the original electronegativity equilization method (EEM), Mortier *et al.* considered three different approximate (first-principles derived) free-atom electronegativities modifed by a screening correction, combined with simple point charge interactions in $\eta_{A,B}^*$ [88]. Later, they recalibrated their parameters semi-empirically to match Mulliken charge distributions in small molecules, and applied a Slater-type correction to short-range charge-charge interactions [89]. Rappe and Goddard utilized the experimental atomic ionization potential and electron affinity to estimate the intra-atomic charge transfer parameters, $\chi_A^0 = {}^1\!/_2(IP+EA)$ and $\eta_{A,A}^0 = IP$-EA, corrected for exchange interactions [92]. They applied a Slater-function representation of the short-range inter-atomic interactions and proceeded to call this approach charge equilibration (QE, QEq, CHEQ) method. Similarly, the atomic hardness $\eta_{A,B}$ has been expressed in Slater- [92–94], Gaussian- [51, 91, 95] or other function [96–98] representations to produce (long-range) classical Coulomb interactions corrected for short-range effects. In 1994, Rick *et al.* developed a Lagrange-technique to efficiently propagate the molecular charges in MD simulations [99], generally known as fluctuating charge model (fluc-q, FC, FQ). More recently, York and Yang provided a charge equilibration method derived from density functional theory that moves beyond the simple atom-condensed charge populations and partly recovers the nature of a finite electron density distribution [91]. A summary of the derivation of their approach—the chemical potential equalization (CPE, μEq) method—and a comparison to EEM is provided in appendix B. The authors proceeded to represent the total molecular energy as a function of atom-centered (partial) electron density changes—opposed to simple atomic partial charges—along the lines of a rigorous perturbative DFT formulation, which enables a more fundamental investigation and improvement of the dielectric response basis functions and its parameters in charge equilibration schemes. For example, it provides a theoretical foundation to include higher-order fluctuating multipole moments and enables the describtion of intra-atomic polarization. Throughout the past decades, a plethora of different charge equilibration schemes and parametrization efforts have been published [96, 100–104].

Unfortunately, molecular charge equilibration schemes exhibit two major limitations. First, the procedure to estimate transferable, yet accurate response parameter representations is very tedious and daunting [51, 97]. Finding and parametrizing simple functional forms to account for the various local environments even in an average way is not a simple task and often leads to limited transferability. Secondly, the representation of electronic polarization is oversimplified and fails to capture the continuous nature of the electron density (opposed to simple point charges or point dipoles), e.g. present in charge accumulation (local change of charge distribution shape) and penetration effects (charge density overlap) [18]. As a result, an obvious failure of simple point charge QEq is its inability to allow polarization perpendicular to planar molecules or molecular structures like benzene, which is crucial in many materials like organic semiconductors or biomolecules. A simple remedy is the addition of higher order multipole moments, e.g. inducible dipole moments [91, 105]. Another shortcoming has been reported by Chelli *et al.*, where a QEq model parametrized for small hydrocarbons predicts a superlinear scaling of the dipole polarizability for higher linear n-alkanes (instead of linear scaling for dielectric systems in the macroscopic limit) and breaks size-extensivity [94, 106]. However, in a later study, the addition of atom-centered inducible moments combined with an improved parametrization procedure lifted the superficial molecular polarizability scaling [95]. Furthermore, charge equilibration schemes lead to partially charged molecules in the dissociation limit, which is unphysical, and limits its applicability in reactive force fields.

One methodological pathway to overcome the model limitations of QEq, in particular the incorrect asymptotic behaviour for separated atoms, is the transformation from a (one-body) atomic partial charge representation to a (two-body) charge-transfer along bond representation, where a bond is not necessarily a literal chemical bond, but instead represents a pair of individual atoms. Each bond allows the transfer of charge between two atomic sites to contribute to the overall polarization, where the construction of different boundary conditions leads to a variety of flavors, like the methods of atom-atom charge transfer [94], bond-charge increments [84, 97], charge-transfer variable [107] and split-charges [108]. While these techniques solve some of the previous limitations [51, 109, 110], they are merely *ad hoc* corrections and lack any theoretical foundation. Instead, they basically confine the super-linear polarizability effect to short inter-atomic distances and hence reduce erroneous contributions [106].

An alternative pathway beyond simple QEq models—founded in first principles as extension of CPE—is provided by the ACKS2 model in the next chapter.

2.4. Non-additivity of explicit polarizability.

Traditional non-polarizable (and non-reactive) force field techniques are typically additive, meaning the total energy is computed as sum of several independent energy contributions of a set of variables (e.g. atomic coordinates) and several specific pre-determined input properties (e.g. fixed atomic charges) during molecular dynamics simulations. These energy terms are formulated as two-, three-, four-body (or higher) functions of the (pre-determined) atom types and parameters together with variable nuclear positions. Specifically, the (long-range) classical electrostatic energy is a two-body interaction energy determined by the relative position and the (atom-centered) distributed multipole moment of the individual sites.

Introduction of explicit polarization in force field invalidates the independence of the input properties of the atomic structure and thereby breaks the additivity assumption. The multipole moments of a specific site are not fixed and pre-determined anymore, instead they are calculated on-the-fly in molecular simulations subjected to the presence of other atoms (and interaction potentials). The mutual dependence of the different site multipole moments is apparent in the Drude model due to coupling of the relative shell charge position, $\boldsymbol{r}_{i,D}$ in eq. (2.1). Similarly, the individual sites are coupled by classical dipole-dipole interactions, $\boldsymbol{T}\boldsymbol{\mu}$, in the inducible dipole model in eq. (2.2). In the charge equilibration schemes, charge transfer and interaction of the fluctuating charges between (atomic) sites are evident in the approximate first and second order energy derivatives, χ and η in eq. (2.6). The interdependence of (atomic) multipole moments leads to a large (partially sparse) system of linear equations containing all polarizable sites, which introduces a considerable computational bottleneck in efficient force field techniques. Exact solutions like direct matrix inversion and Cholesky decomposition are computationally prohibitve in force field simulations of extended systems. Instead, more efficient (arbitrarily exact) approximate solutions are utilized in force field implementations following either an iterative-convergence or direct time-propagation technique.

In iterative techniques, the simulation system is divided into smaller subunits for which the explicit polarization linear equations are solved, with the smallest set consisting only a single polarizable site. An initial (sophisticated) estimate of the individual multipole moments yields a first approximate local electrostatic potential and interaction energy. The latter are utilized to determine the dielectric response of the polarizable sites, which prompts a new electrostatic potential and change of the response, and so forth. This procedure is continued until the individual inducible multipoles and electrostatic interaction energy are self-consistent, i.e. they do not change more than a given accuracy threshold in consecutive iteration steps. Unfortunately, the computational cost of the self-consistent polarization is still expensive due to the iterative cycling performed for every atomic structure at every molecular dynamics simulatiton time step individually. Furthermore, the iterative solution can lead to numerical instabilites. Mathematical details and development of methods to determine good initial parameter guesses, pre-conditioning, accelerated self-consistency convergence and efficient numerical algorithms are a matter of ongoing research [111, 112], but details of these approaches go beyond the scope of this thesis.

Alternatively, the individual multipole moments are calculated for an initial atomic structure and consecutively treated as fictitious particles propagated in time by an extended Lagrangian scheme [99, 113]. With that comes a fictitious mass of the explicit polarization fictitious particles, which should be chosen small enough so that any charge response adapts immediately to the motion of the atomic nuclei and artificial thermal coupling between them is avoided. This is equivalent to a Born-Oppenheimer like separation of the nuclear and explicit electronic degrees-of-freedom. Unfortunately, too small masses and thereby temperatures (ideally 0K) lead to very small time propagation steps necessary to avoid energy drifts and hence yield prohibitive computational cost in any practical applications. In consequence, a compromise for the temperature of the fictitious particles is met between ideally very small and practically much smaller than the temperature of the atomic nuclei by application of an additional polarization thermostat at low temperature ($\approx$ 5K).

In summary, current charge equilibration methods provide a computationally efficient set of tools to account for (localized) explicit electronic polarization in force field techniques. However, they crucially suffer from two major limitations. First, the entire DFT energy functional is projected onto a small and local atom-centered basis set, see eq. (2.5). Thereby, the (non-classical) kinetic energy and coulomb interaction contributions are condensed together in a single set of polarization parameters, η and χ, and the complex quantum mechanical information of the electronic response is oversimplified. One particular feature is the erroneous contribution of the non-local (long-range) kinetic energy to the electronic response [114]. Second, determination of the fluctuating charge parameters is non-trivial, while at the same time transferability is low. This seriously limits the (wide-spread) adoption of explicit polarization in force fields for complex chemical system and reactivity studies.

3. First-principles foundation of the ACKS2 general electronic polarization model.

THE ATOM-CONDENSED KOHN-SHAM DFT APPROXIMATED TO SECOND ORDER (ACKS2) is a linear electronic response model derived from first-principles by Verstraelen *et al.* [114, 115]. It is based on an Euler-Lagrange formulation of perturbative KS-DFT and applies two major simplifications, a linearization of the energy functional derivative—equivalent to an energy expansion up to second order—and an atom-projected (small) linear expansion series of the polarization properties. This chapter is intended as brief summary of the derivation of the general ACKS2 formalism following the work of T. Verstraelen and co-workers, more details of mathematical formulations are to be found in appendix C. Particular focus is set on the ACKS2 model in the context of local or semi-local exchange-correlation functionals. The last part of this chapter is dedicated to a brief discussion of the implications of different unperturbed reference states—free-atom, free-molecule or extended simulation systems. The choice of the latter is in principle arbitrary and depends on the intended usage. Starting from a free-atom reference allows an estimate of the (molecular) charge distribution in simulation systems (equivalent to EEM and QEq). A reference system of free molecules models electronic polarization contributions to the intermolecular interaction energy (as originally proposed by CPE). Finally, considering the entire simulation system enables the dielectric response to external fields like a time-dependent electro-magnetic field.

3.1. Wave-function augmented formulation of density functional theory.

In DFT, the total energy as functional of the electronic density is

$$E[\rho] = F[\rho] + \int \rho(\boldsymbol{r}) v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} \tag{3.1}$$

The universal functional F contains the pure electronic interactions, including kinetic energy, classical electron-electron interactions and exchange-correlation corrections. The external potential v_{ext} accounts for interactions with the nuclei and other external potentials. Unfortunately, a mathematical expression for many contributions to the total density functional energy are unknown. Instead, the universal functional is decomposed into two functionals,

$$E[\rho] = E^{\text{exp}}[\rho] + E^{\text{imp}}[\rho] + \int \rho(\boldsymbol{r}) v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r}. \tag{3.2}$$

One functional, E^{exp}, contains the energy contributions that explicitly depend on the electron density, e.g. the classical Ccoulomb interaction between electron distributions. The other functional, E^{imp}, contains all energy contributions that cannot be expressed directly as function of the charge density. Following the ideas of KS-DFT, the electron density is transformed to an auxiliary wave-function, $\rho \rightarrow \Psi$, to implicitly evaluate the unknown energy contributions, $E^{\text{imp}}[\rho] = E^{\text{WF}}[\Psi]$. The application of two consecutive Legendre transforms provides a formal recipe to express the implicit density energy functional in terms of wave function energy contributions [116],

$$E^{\text{o}}[u] = \min_{\Psi} \left(\int \langle \Psi | \Psi \rangle \, u(\boldsymbol{r}) d\boldsymbol{r} + W^{\text{WF}}[\Psi] \right) \tag{3.3a}$$

$$-E^{\text{imp}}[\rho] = \inf \left(\int \rho(\boldsymbol{r}) \, u(\boldsymbol{r}) d\boldsymbol{r} - E^{\text{o}}[u] \right). \tag{3.3b}$$

Here, eq. (3.3a) represents the Legendre transformation of the (yet unspecified) energy contribution W^{WF} from an arbitraty wave function formalism to an energy functional of the auxiliary potential $u(\boldsymbol{r})$. In a second step, eq. (3.3b), the potential formulation of the implicit functional is

converted again by method of Legendre transforms to an energy functional of the electron density. Together, these equations pose a double constrained search formulated in terms of Lagrange multipliers. The total DFT energy, separated in explicit density functional dependent and implicit auxiliary potential dependent functionals, is

$$E[\rho, u] = E^{\text{exp}}[\rho] + E^{\text{o}}[u] + \int \rho(\boldsymbol{r})\,(v_{\text{ext}}(\boldsymbol{r}) - u(\boldsymbol{r}))d\boldsymbol{r} \tag{3.4}$$

The electronic ground state in DFT (energy minimum) is formulated in terms of the Lagrange-multiplier method, where μ is the Lagrange multiplier,

$$L[\rho, u, \mu] = E^{\text{exp}}[\rho] + E^{\text{o}}[u] + \int \rho(\boldsymbol{r})\,(v_{\text{ext}}(\boldsymbol{r}) - u(\boldsymbol{r}))d\boldsymbol{r} - \mu\left(\int \rho(\boldsymbol{r})d\boldsymbol{r} - N\right). \tag{3.5}$$

The stationary point, which determines the electronic ground state, is determined by the functional derivative of L with respect to each variable $\boldsymbol{\nabla} L$, i.e. $^{\partial L}/_{\partial \rho}$=0, $^{\partial L}/_{\partial u}$=0 and $^{\partial L}/_{\partial \mu}$=0 (which provides a simple charge constrained by construction),

$$\frac{E^{\text{exp}}[\rho]}{\partial \rho(\boldsymbol{r})} + v_{\text{ext}}(\boldsymbol{r}) - u(\boldsymbol{r}) - \mu = 0 \tag{3.6a}$$

$$\frac{E^{\text{o}}[u]}{\partial u(\boldsymbol{r})} - \rho(\boldsymbol{r}) = 0 \tag{3.6b}$$

$$\int \rho(\boldsymbol{r})d\boldsymbol{r} - N = 0. \tag{3.6c}$$

A complete list of the first and second order derivatives of the Lagrange-formulation of DFT, see eq. (3.5), is given in appendix C.1.

3.2. Perturbative density functional theory.

A change of the external potential with respect to a (unperturbed) reference state, v_{ext}=$v_{\text{ext},0}$ + Δv_{ext}, prompts a response of the electron density and auxiliary potential, where the initial unperturbed state is labeled by index 0,

$$\rho(\boldsymbol{r}) = \rho_0(\boldsymbol{r}) + \Delta\rho(\boldsymbol{r}) \tag{3.7a}$$

$$u(\boldsymbol{r}) = u_0(\boldsymbol{r}) + \Delta u_0(\boldsymbol{r}) \tag{3.7b}$$

$$\mu = \mu_0 + \Delta\mu. \tag{3.7c}$$

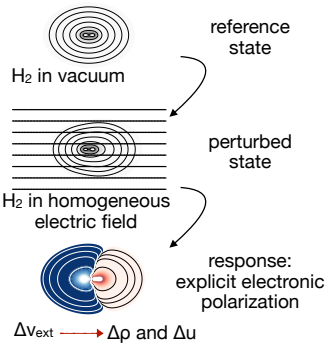


Figure 5 Illustration of electronic polarization in regular and perturbative density functional theory.

The electron density for the perturbed and unperturbed state, ρ and ρ_0, and with them the total energy follows the regular DFT formulation in eqs. (3.6a) to (3.6c), given the external potentials v_{ext} and $v_{\text{ext},0}$. The difference in electronic structure of these two states then gives the polarization response to a change in the external potential as illustrated in fig. 5, e.g. the electronic charge rearrangments $\Delta\rho$ and induced dipole moment $\int \Delta\rho\,\boldsymbol{r}\,d\boldsymbol{r}$. However, this approach is computationally very expensive, as it employs two self-consistent calculations, in particular with a view of explicit polarization corrections in force fields, where the (efficient) calculation of the dielectric response properties is the top priority. A more direct approach to the electronic polarization in response to changes in the external potential is a perturbative approach to density functional theory. The energy functional formulation of the response, including the Lagrange multiplier, is

$$\Delta L^{\text{resp}} = L[\rho_0 + \Delta\rho, u_0 + \Delta u, \mu_0 + \Delta\mu] - L[\rho_0, u_0, \mu_0]. \tag{3.8}$$

The stationary point in a perturbative DFT ansatz is defined by the functional derivatives of the response energy, with respect to $\Delta\rho$, Δu and $\Delta\mu$, and determines the ground state electronic response of the system. A linearized energy functional derivative—equivalent to a Taylor expansion of the response energy with respect to the unperturbed reference state truncated at second order, see appendix C.2 for more details—yields an approximate Euler-Lagrange formulation of the electronic polarization,

$$\int \frac{\partial^2 E^{\text{exp}}[\rho_0]}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')}\Delta\rho(\boldsymbol{r}')d\boldsymbol{r}' + \Delta v_{\text{ext}}(\boldsymbol{r}) - \Delta u(\boldsymbol{r}) - \Delta\mu = 0 \tag{3.9a}$$

$$\int \frac{\partial^2 E^{\text{o}}[u_0]}{\partial u(\boldsymbol{r})\partial u(\boldsymbol{r}')}\Delta u(\boldsymbol{r}')d\boldsymbol{r}' - \Delta\rho(\boldsymbol{r}) = 0 \tag{3.9b}$$

$$\int \Delta\rho(\boldsymbol{r})d\boldsymbol{r} = 0 \tag{3.9c}$$

This triplet of equations is the first-principles foundation of the ACKS2 general explicit polarization model. They define the electronic polarization in the limit of a linear response derived from perturbative DFT—including the transformation to an implicit energy functional of the density—and constitute the theoretical foundation of the central working equations of the general ACKS2 method.

3.3. Atom-condensed parametrization scheme.

In any practical application, the change of electron density $\Delta\rho$ and auxiliary potential Δu are expanded in a linear expansion series,

$$\Delta\rho(\boldsymbol{r}) = \sum_i c_i g_i(\boldsymbol{r}) \tag{3.10a}$$

$$\Delta u(\boldsymbol{r}) = \sum_j d_j h_j(\boldsymbol{r}). \tag{3.10b}$$

This representation is in principle exact in the limit of a complete basis set. For any practically relevant simulation setup, efficient basis functions are necessary, which are capable to condense the electronic structure response information with reasonable accuracy while at the same time they exhibit a sufficiently small computational cost. The choice of an atom-centered (minimalistic) basis set representation is a key component of the ACKS2 approach. Applying a linear expansion of density and auxiliary potential in the Euler-Lagrange response condition above yields the ACKS2 matrix working equations for practical applications,

$$\boldsymbol{\eta c} - \boldsymbol{O d} - \Delta\mu \boldsymbol{D} = -\boldsymbol{V} \tag{3.11a}$$

$$\boldsymbol{\chi d} - \boldsymbol{O}^T \boldsymbol{c} = \boldsymbol{0} \tag{3.11b}$$

$$\boldsymbol{D c} = 0, \tag{3.11c}$$

where the ACKS2 parameters are given by

$$\eta_{k,l} = \iint g_k(\boldsymbol{r}) \frac{\partial^2 E^{\text{exp}}[\rho_0, u_0]}{\partial g_k(\boldsymbol{r}) g_l(\boldsymbol{r}')} g_l(\boldsymbol{r}')\, d\boldsymbol{r} d\boldsymbol{r}' \tag{3.12a}$$

$$\chi_{m,n} = \iint h_m(\boldsymbol{r}) \frac{\partial^2 E^{\text{imp}}[\rho_0, u_0]}{\partial h_m(\boldsymbol{r}) h_n(\boldsymbol{r}')} h_n(\boldsymbol{r}')\, d\boldsymbol{r} d\boldsymbol{r}' \tag{3.12b}$$

$$O_{k,m} = \int g_k(\boldsymbol{r}) h_m(\boldsymbol{r}) d\boldsymbol{r} \tag{3.12c}$$

$$V_k = \int g_k(\boldsymbol{r}) \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} \tag{3.12d}$$

$$D_k = \int g_k(\boldsymbol{r}) d\boldsymbol{r} \tag{3.12e}$$

The hardness kernel $\eta_{k,l}$ and non-interacting linear response kernel $\chi_{m,n}$ include electronic structure information following the choice of separation into explicit and implicit electron density functionals. Details for the parametrization with (local and) semi-local functionals are presented in the next section. The overlap parameters $O_{k,m}$, density integral parameters D_k and perturbation vector V_k are entirely electronic-structure independent and their performance rests on the basis set representation. A useful polarization property is the induced dipole moment as function of the change in electron density,

$$\Delta\mu^{\text{ind.}} = \sum_i c_i \int g_i(\boldsymbol{r})\, \boldsymbol{r} d\boldsymbol{r}. \tag{3.13a}$$

Another important property is the induced polarization energy $E^{\text{pol.}}$, which describes the energetic gain due to rearrangement of charge density, excluding the interaction of the unperturbed reference density with the perturbation potential $\int \rho_0 \Delta v_{\text{ext}} d\boldsymbol{r}$. In the limit of linear response, which should be valid for small perturbations and enforced by the ACKS2 approach, the cost of electronic rearrangements is half the size of the gain by the interaction with the external perturbation.

$$E^{\text{resp.}}_{\text{ACKS2}} = \int \rho_0(\boldsymbol{r}) \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} + \boldsymbol{cV} - \boldsymbol{cOd} + \frac{1}{2}\boldsymbol{c\eta c} + \frac{1}{2}\boldsymbol{d\chi d} \tag{3.14a}$$

$$E^{\text{pol.}}_{\text{ACKS2}} = \boldsymbol{cV} - \boldsymbol{cOd} + \frac{1}{2}\boldsymbol{c\eta c} + \frac{1}{2}\boldsymbol{d\chi d} \tag{3.14b}$$

$$\approx \frac{1}{2}\boldsymbol{cV} \tag{3.14c}$$

Direct computation of atomic forces due to polarization poses a desirable feature for force field simulations. Investigation of atomic forces has not been part of this thesis, nevertheless a derivation of ACKS2 polarization contributions to atomic forces is illustrated in appendix C.3.

3.4. Semi-local exchange-correlation functionals in ACKS2.

The choice of basis set and exchange-correlation (xc) functional influence the accuracy of (static) polarization calculations in DFT [117–119]. Generalized gradient approximations slightly lack in accuracy, PBE-xc [120] exhibits an relative error of 6% for a range of small organic molecules [118] and 10% for small molecules composed of main group elements lighter than Argon [119] compared to coupled cluster calculations. We focus on the application of GGA-level exchange-correlation functionals (in particular PBE), which are explicit with respect to the electron density, i.e. E_{xc}=$E_{\text{xc}}[\rho]$. The explicit functional terms applied in the KS-DFT formulation and the respec-

tive ACKS2 hardness are

$$E^{\text{exp}}[\rho] = E_H[\rho] + E_{\text{xc}}[\rho] \tag{3.15a}$$

$$\eta_{k,l} = \iint g_k(\boldsymbol{r}) \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} g_l(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' + \iint g_k(\boldsymbol{r}) \frac{\partial E_{\text{xc}}[\rho_0]}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')} g_l(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \tag{3.15b}$$

In practice, the Coulomb contributions clearly dominate the hardness elements within a local or semi-local xc-functional framework, particularly for atom centers far apart. However, the same-center contributions (diagonal elements $[\eta_{k,k}]_{\text{xc}}$ and block-diagonals for same atom center) are not negligible and add an intra-atomic exchange-correlation correction. Figure 6 illustratively depicts the distribution of the electronic structure dependent parameters, $[\eta_{k,l}]_{\text{xc}}$ and $\chi_{m,n}$ (see below), for small aromats given the sp-type Cartesian Gaussian function basis set representation developed in publication #1 [1]. The implicit functional, resorting to a wave function technique by Legendre transformation, now only contains the kinetic energy of the Kohn-Sham electronic structure,

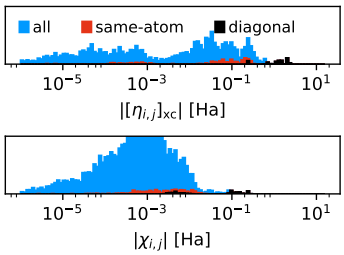


Figure 6 Illustration of the electronic structure dependent ACKS2 parameters, $\eta^{\text{xc}}_{k,l}$ and $\chi_{m,n}$ for a collection of small aromats (benzene, naphthalene, anthracene, tetracene).

$$W^{\text{WF}}[\Psi] = \langle\Psi| - \frac{1}{2}\nabla^2|\Psi\rangle \tag{3.16a}$$

$$E^{\text{o}}[u] = -\inf\left(\int \rho(\boldsymbol{r})u(\boldsymbol{r})d\boldsymbol{r} + W^{\text{WF}}[\Psi]\right) \tag{3.16b}$$

$$= -\inf\left(\int \rho(\boldsymbol{r})u(\boldsymbol{r})d\boldsymbol{r} + \left(E_{\text{KS}}[\Psi] - \int \langle\Psi|\Psi\rangle\, v_{\text{KS}}(\boldsymbol{r})d\boldsymbol{r}\right)\right). \tag{3.16c}$$

This formulation of the implicit density functional in the framework of KS-DFT and (local or) semi-local exchange-correlation functionals helps us identify the auxilary potential with the effective potential of the Kohn-Sham system of non-interacting electrons, $u{=}v_{\text{KS}}$. The non-interacting linear response kernel finally is

$$\chi_{m,n} = \iint h_m(\boldsymbol{r}) \frac{\partial E_{\text{KS}}[u]}{\partial u(\boldsymbol{r})\partial u(\boldsymbol{r}')} h_n(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \tag{3.17a}$$

$$= \sum_{i,j} \frac{1}{\epsilon_i - \epsilon_j}\left(\int \psi_i^*(\boldsymbol{r})h_m(\boldsymbol{r})\psi_j(\boldsymbol{r})d\boldsymbol{r}\right)\left(\int \psi_j^*(\boldsymbol{r})h_m(\boldsymbol{r})\psi_i(\boldsymbol{r})d\boldsymbol{r}\right) + c.c., \tag{3.17b}$$

where ψ_i (ψ_j) are the occupied (virtual) KS orbitals with eigenenergy ϵ_i (ϵ_j) and c.c. is an abbreveation of the complex conjugated term.

3.5. Dielectric response in intermolecular interactions.

The research in this study is dedicated to a free molecule as reference, where the initial unperturbed state ρ_0 (see eq. (3.7a)) is the self-consistent charge density of a molecule in vaccuum. The electronic response $\Delta\rho$ contains the change of the molecular charge density, specifically scrutinized in this thesis for ideal electrostatic potentials—a single point charge or homogeneous electric field—or a simplified Coulomb representation of other molecules. Intended as proof-of-principle study in a bottom-up direction, all ACKS2 parameters are calculated exactly from the self-consistent KS-DFT electronic structure and the ACKS2 response properties—induced dipole moment and polarization energy—are validated against the KS-DFT parent method. Thereby, we assess the ability of our basis set representation to project the response information onto atom-centered ba-

sis functions and get a measure of the *ab-initio* accuracy in ACKS2 (excluding error cancelation due to empirical parameter fitting). This allows us to include efficient, yet accurate first-principles derived response contributions in the intermolecular interaction energy in force fields. Details of the Cartesian Gaussian basis set development for molecules are collected in publication #1, further information about the dielectric response due to interaction of molecules are summarized in publication #2.

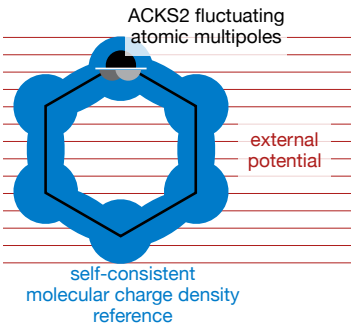


Figure 7 Illustration of molecular ACKS2 model for carbon backbone of benzene.

One critical drawback of this bottom-up approach is that currently we need to perform one self-consistent DFT calculation for each molecular structure or atomic arrangement, which can be quite challenging depending on the system and is practically unfeasible for phase space sampling with force fields. We envision a high-accuracy ACKS2 model for an efficient molecular response, where a basis set representation is developed from KS-DFT (as performed in publication #1) and the first-principles derived parameters are projected onto an efficient descriptor or machine-learned to reduce their computational overhead. Furthermore, the current state of ACKS2 allows—after a single DFT calculation—to calculate the dielectric response of a fixed atomic structure to a varying external perturbation, like e.g. the response of a frozen molecule or cluster to a time-dependent potential.

3.6. Electronic charge rearrangements of atoms in a molecule.

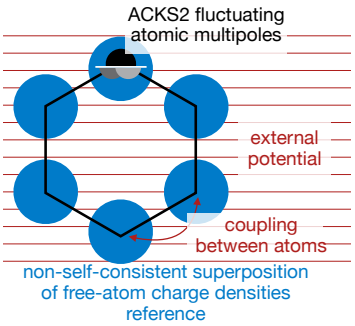


Figure 8 Illustration of free-atom reference ACKS2 model for carbon backbone of benzene.

A free-atom unperturbed reference state is very useful to estimate fluctuating charges and higher multipole moments of the individual atoms in a molecule during molecular motions (atomic multipole changes triggered by conformational and topological changes). It follows the idea of system fragmentation of the ACKS2 reference state introduced in publication #2 [2]. Now, a molecule is decomposed into single-atom fragments, and the ACKS2 model yields an electronic response of each atomic site to the presence of all other sites in the molecule and an external potential if present. Crucially, the free-atom reference ACKS2 needs to approximate the interaction potential between different atomic sites in a molecule (a simplified classical Coulomb interaction approximation suffices for molecular fragments in publication #2). It must contain (approximate) information about the electronic structure at short ranges typical for molecular bond lengths to account for chemical bond interactions between atomic sites. They key ideas for a free-atom ACKS2 are briefly summarized here.

The self-consistent molecular charge density ρ^{mol} is decomposed into atomic contributions (index A for different atom centers), where ρ^{fa} is the unperturbed charge density of an atom in vacuum and $\Delta\rho^{\mathrm{AIM}}$ is its change in a molecule (atom-in-molecule correction),

$$\rho^{\mathrm{mol}}(\boldsymbol{r}) = \sum_A \rho_A^{\mathrm{fa}}(\boldsymbol{r}) + \Delta\rho_A^{\mathrm{AIM}}(\boldsymbol{r}) \tag{3.18}$$

In the free-atom reference ACKS2 model, the unperturbed reference state ρ_0 (see eq. (3.7a)) is defined as (non-self-consistent) superposition of free-atom like charge densities. The molecular (free-atom reference) ACKS2 electronic response $\Delta\rho$ contains the charge rearrangments due to any external potential acting on the entire molecule (calculated individually for each atom center $\Delta\rho_A$). Additionally, it includes an atom-in-molecule correction $\Delta\rho_A^{\mathrm{AIM}}$ to account for the molecular character of the atomic sites. To achieve that, the free-atom ACKS2 perturbation parameter $\Delta v_{\mathrm{ext},A}$ contains the external potential Δv_{ext} (acting on the entire molecule) and an atomic interaction term $\Delta v_{\{A,B\}}^{\mathrm{fa-fa}}$. The later introduces coupling of the atomic site A to the free-atom ground state charge density of another atomic site B. Note, the free-atom ACKS2 interaction potential parameter is atomic site specific, indicated by the atomic label subscript A.

$$\rho_0 = \sum_A \rho_A^{\mathrm{fa}} \tag{3.19a}$$

$$\Delta\rho = \sum_A \Delta\rho_A^{\mathrm{AIM}} + \Delta\rho_A \tag{3.19b}$$

$$\Delta v_{\mathrm{ext},A}(\boldsymbol{r}) = \Delta v_{\mathrm{ext}}(\boldsymbol{r}) + \sum_B \Delta v_{\{A,B\}}^{\mathrm{fa-fa}} \tag{3.19c}$$

The basis set dependent (electronic structure independent) free-atom reference ACKS2 parameters are calculated analytically for a Cartesian Gaussian representation. The electronic-structure dependent ACKS2 parameters, exchange-correlation contributions to η (see eq. (3.15b)) and χ (see eq. (3.17a)), for basis functions located at the same atom are calculated exactly from the KS-DFT electronic structure of an isolated atom in vacuum. For numerical speedup, they can be saved to a file or look-up tables for future use as polarization model. The ACKS2 parameters for basis functions located at different atom centers need to be approximated for any practical implementation, e.g. from a diatomic reference. Alternatively, the coupling of the response at different atomic sites can be treated in the effective inter-atomic interaction potential.

A free-atom reference ACKS2 exhibits two major challenges. First, condensing the *ab-initio* interaction of two atoms into an effective (two-body) potential $\Delta v_{\mathrm{ext}}^{\mathrm{fa-fa}}$ is a non-trivial task, in particular for the short-range interactions and charge penetration (orbital overlap) effects of close atoms. A well-chosen charge distribution function like Slater-type or Gaussian-type may prove a good remedy, as has been applied in previous studies of EEM and CPE [121]. Second, accounting for the changes of the electronic structure and hence DFT-dependent response parameters, $\eta_{i,j}$ and $\chi_{i,j}$, at the atomic site during bond formation is computationally challenging. In principle, the free-atom like character decreases and the atom-in-molecule character increases, $\eta_{i,j}^{\mathrm{fa}} \to \eta_{i,j}^{\mathrm{AIM}}$ and $\chi_{i,j}^{\mathrm{fa}} \to \chi_{i,j}^{\mathrm{AIM}}$. Practically, the bonding effect onto the atoms in a molecule need efficient estimates, for example statistical sampling of atom-in-molecules ACKS2 parameters, $\langle\eta\rangle_{i,j}^{\mathrm{AIM}}$ and $\langle\chi\rangle_{i,j}^{\mathrm{AIM}}$, for a reasonable variety of chemical environments. An effective AIM (linear) scaling term, which specifies the AIM character of a polarizable site versus its free-atom character, can help to account for the transition,

$$\eta_{i,j} = (1-\alpha)\eta_{i,j}^{\mathrm{fa}} + \alpha\,\langle\eta\rangle_{i,j}^{\mathrm{AIM}}\ \chi_{i,j} = (1-\alpha)\chi_{i,j}^{\mathrm{fa}} + \alpha\,\langle\chi\rangle_{i,j}^{\mathrm{AIM}}\,. \tag{3.20}$$

This approach sounds promising as road map for future work of ACKS2 as extention to simple fluctuating charge models, but has not been investigated within this thesis.

4. Density functional tight-binding.

The density functional tight-binding technique , specifically the second order energy expansion version DFTB2, bears many similarities with CPE (and by extension with ACKS2) in their theoretical background. The modeling targets are quite different, as the aim of DFTB methods is to calculate the total energy of a system (and other related properties)—similar to force fields, but on a different functional footing—whereas charge equilibration techniques only cover explicit polarization. In fact, the method of fluctuating atomic charges (FQ) is applied within approximate DFTB models to efficiently evaluate parts of the electrostatic interactions and charge transfer between approximate free-atom-like particles. A brief summary of the origins of DFTB2 is given, following these references [122–124]. The density functional energy in a Kohn-Sham formalism—here conceptually formulated as functional of the electron density similar to the derivation of the CPE model, see appendix B—contains the kinetic energy contributions T_{KS}, classical Coulomb interactions E_{H}, non-classical exchange-correlation interactions E_{xc} (and corrections due to the incorrect kinetic energy functional), interactions with external fields E_{ext}, e.g. due to the atomic nuclei, (and explicit ion-ion interactions E_{ion} between the atomic nuclei)

$$E[\rho] = T_{\mathrm{KS}}[\rho] + E_{\mathrm{H}}[\rho] + E_{\mathrm{ext}}[\rho] + E_{\mathrm{xc}}[\rho] + E_{\mathrm{ion}}[\rho] \tag{4.1}$$

The true kinetic energy is approximated by a system of non-interacting electrons represented by one-particle wave-functions ψ_a^{KS} (occupation $f_a \in [0, 2]$) in an effective potential v_{eff}, which contains classical Coulomb interactions v_{H}, non-classical exchange-correlation effects v_{xc} and external potentials v_{ext}. Additionally, it satisfies the constraint to reproduce the electron density distribution $\sum_a |\psi_a^{\mathrm{KS}}|=\rho$,

$$-\frac{1}{2}\nabla^2 + v_{\mathrm{eff}}^{\mathrm{KS}}(\boldsymbol{r})\,|\psi_a^{\mathrm{KS}}\rangle = \epsilon_a\,|\psi_a^{\mathrm{KS}}\rangle \tag{4.2a}$$

$$v_{\mathrm{eff}}(\boldsymbol{r}) = v_{\mathrm{H}}(\boldsymbol{r}) + v_{\mathrm{xc}}(\boldsymbol{r}) + v_{\mathrm{ext}}(\boldsymbol{r}) \tag{4.2b}$$

$$T_{\mathrm{KS}}[\rho] = \sum_a f_a \,\langle\psi_a^{\mathrm{KS}}| -\frac{1}{2}\nabla^2\,|\psi_a^{\mathrm{KS}}\rangle \tag{4.2c}$$

$$= \sum_a f_a\epsilon_a - \int v_{\mathrm{eff.}}^{\mathrm{KS}}(\boldsymbol{r})\rho(\boldsymbol{r})d\boldsymbol{r} \tag{4.2d}$$

In density functional tight binding DFTB2 [122, 124] the total energy is expanded up to second order with respect to a reference state ρ_0—recent work also includes third order expansion terms [125]—similarly to CPE and ACKS2,

$$E^{\mathrm{DFTB2}} \approx E[\rho_0 + \Delta\rho] + O(\Delta\rho^2) \tag{4.3}$$

$$= \sum_a f_a\epsilon_a - \frac{1}{2}\int v_{\mathrm{H}}(\boldsymbol{r})\rho_0(\boldsymbol{r})d\boldsymbol{r} - \int v_{\mathrm{xc}}(\boldsymbol{r})\rho_0(\boldsymbol{r})d\boldsymbol{r} + E_{\mathrm{xc}}[\rho_0] + E_{\mathrm{ion}} \tag{4.4}$$

$$+ \frac{1}{2}\iint\left(\frac{\partial^2 E_{\mathrm{xc}}[\rho_0]}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')} + \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right)\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')d\boldsymbol{r}d\boldsymbol{r}'$$

The first five terms constitute the energy of the (artificial) reference state energy, $E_0[\rho_0]$, which is not present in CPE and ACKS2, as they only describe the electronic response, i.e. the change of this reference with changes in the external potential. For practical implementations, DFTB contains many approximations to estimate this zero-order reference state to gain numerical speedup compared to conventional DFT. The electronic charge density of the reference state is constructed—somewhat artificially—from optimised pro-atomic (or free-atom-like) distributions, $\rho_0 = \sum_A \rho_A^0$,

in order to utilize an effective and inexpensive inter-atomic interaction framework. The core electrons are typically assumed to be bound tightly (hence the name), so they approximately forgo any changes during molecular or solid state bond formation and are simply added to the effective ion-ion interaction contribution E_{ion}, also referred to as frozen-core approximation. The valence orbital electrons are accounted for explicitly in ψ^{DFTB}, but reduced to a minimal basis representation of a single (Gaussian) radial function per angular momentum, i.e. one primitive function for s-type orbitals, three primtive functions for p-type orbitals, etc.. In the formulation of DFTB, the explicit treatment of the approximate (one-body) wave-function is executed on purpose (opposed to DFT, where the Kohn-Sham electronic structure is a necessary tool) to calculate approximate (molecular) orbital shapes and levels,

$$\begin{aligned} E^{\text{DFTB2}} = \sum_a f_a \Big\{ \langle \psi_a^{\text{DFTB}} | -\frac{1}{2}\nabla^2 + \frac{1}{2}\int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' \\ + v_{\text{H}}[\rho_0](\boldsymbol{r}) + v_{\text{xc}}[\rho_0](\boldsymbol{r}) + v_{\text{ext}}(\boldsymbol{r}) \, |\psi_a^{\text{DFTB}}\rangle \Big\} \\ - \frac{1}{2}\int v_{\text{H}}(\boldsymbol{r})\rho_0(\boldsymbol{r}) d\boldsymbol{r} - \int v_{\text{xc}}(\boldsymbol{r})\rho_0(\boldsymbol{r}) d\boldsymbol{r} + E_{\text{xc}}[\rho_0] + E_{\text{ion}} \\ + \frac{1}{2}\iint \left(\frac{\partial^2 E_{\text{xc}}[\rho_0]}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')} + \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} \right) \partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \end{aligned} \tag{4.5}$$

Changes of the orbitals and the overall charge density upon bond formation, $\rho = \sum_A \rho_A^0 + \Delta\rho_A$, are calculated from the DFTB approximate free-atom-like particle wave functions and density functional in a non-self-consistent [122] or self-consistent manner [123, 126]. The energetic contributions of the reference state $E_0[\rho_0]$ are grouped in band-structure (BS) terms—approximate single-particle orbitals and energies—repulsive (rep) terms—two-body terms dominated by the effective repulsion potential between the ionic cores. Efficient implementation and parametrization details for the band structure and repulsive contribution go beyond the scope of this work.

$$\begin{aligned} E_{\text{BS}}^{\text{DFTB2}} = \sum_a f_a \langle \psi_a^{\text{DFTB}} | -\frac{1}{2}\nabla^2 + \frac{1}{2}\int \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' \\ + v_{\text{H}}[\rho_0](\boldsymbol{r}) + v_{\text{xc}}[\rho_0](\boldsymbol{r}) + v_{\text{ext}}(\boldsymbol{r}) \, |\psi_a^{\text{DFTB}}\rangle \end{aligned} \tag{4.6}$$

$$E_{\text{rep}}^{\text{DFTB2}} = -\frac{1}{2}\int v_{\text{H}}(\boldsymbol{r})\rho_0(\boldsymbol{r}) d\boldsymbol{r} - \int v_{\text{xc}}(\boldsymbol{r})\rho_0(\boldsymbol{r}) d\boldsymbol{r} + E_{\text{xc}}[\rho_0] + E_{\text{ion}} \tag{4.7}$$

The last term in eq. (4.5), provides a second order correction to the interaction potential of the valence electrons for any charge rearrangements of the pro-atomic upon bond formation and other external potentials. They are reduced to simple atomic charge populations—similarly to the initial formulation of the electronegativity equalization method in eq. (2.6)—and hence provide a simple model of charges flowing between different atomic sites. In DFTB the individual partial charges are calculated by simple spatial partitioning of the valence orbitals (core electrons are assigned to nuclei already), typically following a simple radial atom-centered distribution like a s-type Cartesian Gaussian function volume V_A instead of complicated and computationally expensive atom-in-molecule charge partitioning schemes. Note, any first order expansion terms with respect to the reference density are omitted in DFTB2 under the assumption that the approximate reference is close to the true ground state, $\partial E/\partial\rho|_{\rho_0} = \chi = 0$, which is conceptually different to EEM and ACKS2. The second order atom-centered partial charge energy correction in DFTB is simply expressed by an effective two-body interaction potential γ as function of the nuclear positions $\boldsymbol{R}_A$, translating to the inter-atomic and intra-atomic second order terms, $\eta_{A,A}$ and $\eta_{A,B}$ in EEM,

$$q_A^{\mathrm{DFTB2}} = \int_{V_A} \rho(\boldsymbol{r}) d\boldsymbol{r} \tag{4.8}$$

$$\begin{aligned} E_{\mathrm{FQ}}^{\mathrm{DFTB2}} &= \frac{1}{2} \iint \left(\frac{\partial^2 E_{\mathrm{xc}}[\rho_0]}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')} + \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} \right) \partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\ &= \frac{1}{2} \sum_{A,B} q_A \, q_B \, \gamma_{A,B}(|\boldsymbol{R}_A - \boldsymbol{R}_B|) \end{aligned} \tag{4.9}$$

$$\Delta E^{\mathrm{EEM}}(q_A) = \chi_A q_A + \frac{1}{2}\eta_{A,A} q_A^2 + \frac{1}{2}\eta_{A,B} q_A q_B \tag{4.10}$$

The total DFTB2 energy finally reads

$$E^{\mathrm{DFTB2}} = E_{\mathrm{BS}}^{\mathrm{DFTB2}} + E_{\mathrm{rep}}^{\mathrm{DFTB2}} + \frac{1}{2} \sum_{A,B} q_A q_B \gamma_{A,B}(|\boldsymbol{R}_A - \boldsymbol{R}_B|). \tag{4.11}$$

In summary, the theoretical foundation of DFTB2 bears many similarities to CPE and ACKS2. In density functional tight-binding, the total charge density is constructed (self-consistently or non-self-consistently) from parametrized pro-atomic frozen-core charge distributions and minimal valence orbital representations. Charge equilibration techniques provide a second order correction to the total energy based on the change of the atomic charges (with respect to the pro-atomic reference).

5. Publications.

5.1. Toward First-Principles-Level Polarization Energies in Force Fields: A Gaussian Basis for the Atom-Condensed Kohn-Sham Method.

P. Gütlein, L. Lang, K. Reuter, J. Blumberger and H. Oberhofer
J. Chem. Theory Comput. **15**, 8, 4516-4525 (2019)
DOI 10.1021/acs.jctc.9b00415

Content

In this study, we introduce and carefully validate a new basis set representation of the electronic respose in the ACKS2 method as a first step towards a general polarization technique. In their two seminal publications [114, 115], Verstraelen *et al.* initially apply a Hirshfeld- and Fukui-function based description of the electronic polarization derived from the KS-DFT ground state to formulate the theoretical foundation of ACKS2 method in his two seminal publications. While both representations exhibit good accuracy for illustrative test cases, they crucially lack in transferability between different structures and produce a large numerical overhead, which precludes their usage as efficient polarization model. We present a new primitive Cartesian Gaussian function basis set to account for the changes of electron density and Kohn-Sham potential upon perturbation to improve the numerical efficiency and representation transferability of the ACKS2 technique while maintaining accuracy. To this end, we assess the mathematical foundation of Cartesian Gaussian functions and their integrals present in the ACKS2 formulation to calculate all ACKS2 parameters in full detail. Intended as proof-of-concept, we deploy a small basis set of one s- and p-type Gaussian function per atomic site to allow atom-atom charge transfer and intra-atomic polarization. The width of each basis function is optimized independently for different atomic species, here illustrated for carbon and hydrogen, to match the response properties—induced dipole moment and polarization energy—from a full KS-DFT calculation. Thereby, we find a good projection of the full electronic response information onto simple atom-centered Gaussian functions. We constructed a small training set of homonuclear atomic dimers, H_2 and C_2, and added various idealized electrostatic potentials, either a single homogeneous electric field or a point charge, to sample different relative orientations and magnitudes of interaction. To validate our simplistic setup, we built a test set of different molecules, containing small linear hydrocarbons to check for size-extensivity effects, different C_2H_x compounds to include different chemical bonds and atomic connectivity as well as small aromats, all again perturbed by a variety of individual idealized electrostatic perturbation potentials. We demonstrate good agreement of our Gaussian-based ACKS2 induced dipole moments with the parent DFT method. Furthermore, we report and successfully validate (molecular) polarization energies crucial to atom-level simulations for ACKS2, which has been omitted in the initial studies of Verstraelen and co-workers [114, 115]. In a final test, we illustrate the application of ACKS2 for molecular polarizabilities and raman activities of a high-symmetry ring-breathing mode of benzene.

Individual contributions

I produced a python package, which implements all Gaussian function overlap and interaction integrals to evaluate all electronic-structure independent ACKS2 parameters, in order to provide a simple-to-use tool to execute and further develop ACKS2 electronic polarization simulations. All KS-DFT calculations with the FHI-aims code and further post-processing to obtain the various polarization energy contributions for the training and test set data were executed by myself. Finally, I performed the optimization of the primitive Cartesian Gaussian function basis set. Lucas Lang initially researched the mathematical foundation to calculate all ACKS2 parameters for primitive Cartesian Gaussian functions and implemented the evaluation of the electronic-structure dependent contributions in the FHI-aims full-electron DFT code [127]. Dr. Harald Oberhofer edited the manuscript, Prof. Jochen Blumberger and Prof. Karsten Reuter proofread the final draft.

5.2. An iterative fragment-scheme for the ACKS2 electronic polarization model: Application to molecular dimers and chains.

P. Gütlein, J. Blumberger and H. Oberhofer
J. Chem. Theory Comput. (accepted)
DOI 10.1021/acs.jctc.0c00151

Content

In this study, we investigate the electronic polarization contributions to the intermolecular interaction energy of the ACKS2 formalism and Cartesian Gaussian basis set published before. We present two new approaches to construct a polarizable (molecular) fragment representation of a simulation cell in an ACKS2 framework and calculate the mutual polarization in terms of induced dipole moment and polarization energy. Conceptually, both models allow the evaluation of polarization contributions to the intermolecular energy of constrained charge states while at the same time simplifying the ACKS2 parametrization effort. The first approach, fragment-ACKS2 or f-ACKS2, applies a fixed-charge representation of the molecular ground state charge density as approximate interaction potential between different fragments. It is a simple non-additive extension to molecular mechanics simulations and retains the general ACKS2 form from the previous publication. The second approach incorporates an additional coupling between the ACKS2 dielectric response of different fragments in a self-consistent manner, scf-ACKS2. Thereby, it becomes an additive response model, where each fragment is polarized by the *permanent* and the mutually *induced* charge representation. We investigated the efficacy of both approaches for a set of charged hydrocarbon dimers and validated it against its full constrained-DFT parent method. Simultaneously, we estimated the effective deviations in the dielectric response due to a simplified point charge permanent interaction potential versus a full electrostatic potential in DFT. Overall, the trends of induced dipole moment and polarization energy between the fragment-ACKS2 approach and its DFT reference match well, but fragment-ACKS2 systematically miscalculates the properties by a small margin. We find, the self-consistent correction to f-ACKS2 is noticeable and nicely agrees with its DFT parent results, but it is rather small compared to the overall electronic response and can be regarded as high accuracy option for future applications. Finally, we investigate the energetics of an electron-hole pair separation in a model system comprising a linear chain of anthracene molecules, which is computationally unfeasible in DFT.

Individual contributions

I extended the ACKS2 python package from a previous work to allow coupling of different fragment in a non-self-consistent (f-ACKS2) and self-consistent (scf-ACKS2) way. All calculations, ACKS2 and KS-DFT calculations with the FHI-aims code, and further post-processing to obtain the various polarization energy contributions for the training and test set data were executed by myself. Dr. Harald Oberhofer edited the manuscript, Prof. Jochen Blumberger proofread the final draft.

6. Conclusion & Outlook.

Incorporation of the (long-range) classical electrostatic interactions is crucial to study condensed matter with efficient approximate force field techniques and has been an active field of research ever since the first molecular mechanics simulations. Initial attempts, specifically aiming at low computational cost, employed pre-determined (atom-centered) distributed multipole moment representations of the latter. In order to maintain the numerical efficiency of fixed-charge model implementations in additive force fields, implicit *ad hoc* corrections of increasing or decreasing gas phase atomic charges accounting for an average effect of dense phase environments have been applied to mixed success, as they clearly lack the distinct adaptability of the charge density in various media and consequent versatility inherent to the dielectric response. They often, yet not always struggle to account for local effects like a varying response behavior of the same nuclear species in different atomic environments or the directionality of the electronic charge rearrangments. The expected shift of paradigm towards explicit rendition of the electronic polarization observed over the last fifteen years lead to a plethora of flavors and parameter schemes of a small set of simple response models. They mostly employ multicenter (intra-atomic) polarizabilities, Drude oscillator and inducible dipole moments, or fluctuating charges, charge equilibration, and combinations thereof. While they often perform well after a thorough parametrization complying with the distinct model intricacies, which can lead to a very limited transferability, a unified framework on a rigorous theoretical footing in first-principles is ideal for the future development of a general polarization technique.

Here, the recently published ACKS2 technique poses a promising candidate, but so far crucially lacked in a foundation for application in force fields. Methodologically, it hinges on a two-fold simplification of the perturbative KS-DFT polarization, which is exact within the limits of DFT and its exchange-correlation functionals, including a truncated second order energy expansion and a reduced atom-projected representation of the full electronic structure response information. This doctoral thesis was dedicated to the study of the ACKS2 method and progress towards its applicability as general polarization model in a proof-of-principle work. We targeted the transferability of the ACKS2 basis set representation and parameter evaluation while maintaining reasonable accuracy compared to the first-principles reference. Therefore, this study was designed in a bottom-up fashion starting at the quantum mechanical level of molecules, where all ACKS2 parameters are calculated exactly from and the ACKS2 response properties are carefully validated by the full KS-DFT response information in order to gauge the influence of the approximate representation and overall ACKS2 quality. In a first publication, ref [1], a minimalistic Cartesian Gaussian function basis set—including monopole and dipole angular momenta for interatomic charge transfer and intra-atomic polarization—was introduced and optimized for carbon and hydrogen. This simplistic approach exhibited good agreement of polarization properties, namely induced dipole moments and polarization energies, compared to the KS-DFT reference for a variety of small hydrocarbons, which showcased the accuracy and transferability inherent to the ACKS2 model. It provides a clear first-principles based parametrization recipe for future development and applications. With a view on molecular modeling and computational cost, the choice of Gaussian functions allows analytical calculation of many parameters and the number of basis functions per atom center is rather small (i.e. four in a sp-orbital-type representation). Furthermore, the same study points to-

wards the applicability of an average or efficient interpolation estimate of all ACKS2 parameters currently evaluated from an explicit electronic-structure calculation (which is probitive to efficient force field techniques). In a second publication, ref [ADD PAPER], a fragmentation approach to ACKS2 was developed to assess the mutual polarization of (gas-phase reference) molecules in dense matter and its affiliated contribution to the intermolecular energy ideally suited for molecular mechanics simulations. The fragment ACKS2 technique comes in two flavors, a non-self-consistent and a self-consistent version, where the electronic polarization depends only on the permanent charge representation or additionally includes mutual response interactions. The comparison to an exact reference constrained DFT calculations for a set of small aromats exhibits good agreement of the polarization properties and illustrates the validity of the f-ACKS2 approach. A first showcase application of the f-ACKS2 approach was given in the study of charge separation and their distance-dependent stabilization due to explicit electronic polarization in a toy model of anthracene moldecules stacked perpendicularly to a one-dimensional chain. Both works together conclude the intended proof-of-principle work on the validity of the first-principles derived ACKS2 technique as general polarization model focusing on the parametrization, accuracy and transferability in a bottom-up direction.

The next step towards application comprises the development of a more complete ACKS2 (Cartesian Gaussian) basis set that includes more nuclear species (other than just hydrogen and carbon) and flavors of angular momentum, where e.g. a simple s-type (spd-type) representation can be applied as low (high) cost options in force fields. At the same time, numerically efficient estimates of the set of parameters currently still calculated from a KS-DFT electronic structure are indispensable for any extended molecular dynamics and Monte Carlo studies. Finally, a properly optimized implementation of the ACKS2 parameters and (Cartesian Gaussian) basis functions in a production molecular modeling software and consecutive benchmarks of the accuracy and computational cost are valuable final steps to general out-of-the-box applicability of ACKS2.

Conceptually, the first-principles foundation of the ACKS2 model developed within this thesis is arguably incomplete compared to high-level multi-reference wave-function methods like CCSD(T) due to the choice of a semi-local exchange-correlation functional. Extension and cross-comparison to other exchange-correlation functionals including hybrid formulations is an option to improve the underlying theory towards a more exact physical foundation of ACKS2.

Finally, the ACKS2 and f-ACKS2 methods provide valuable post-processing tools. On the foundation of an exactly parametrized ACKS2 formulation off a single KS-DFT calculation, the explicit electronic response of an atomistic structure can be screened for a variety of external perturbation potentials. Furthermore, they allow a decomposition of the polarization properties, like the change of a molecular dipole moment in condensed media. In particular, a comparison between an ACKS2 formulation and a f-ACKS2 (or scf-ACKS2) formulation helps to identify the differences in the dielectric response of a molecule in vacuum and a molecule in dense media and assess their intra- and intermolecular contributions to the total charge rearrangements.

7. Acknowledgments

This PhD thesis would not have been possible without the support and input of many people, but foremost I want to thank my supervisors Prof. Dr. Karsten Reuter, Prof. Dr. Jochen Blumberger and PD Dr. Harald Oberhofer. They initiated this project and trustfully laid it in my hands, giving me a lot of freedom throughout the past years.

Karsten Reuter provided strategic guiding during this project, especially at the start, and anecdotally taught me many lessons of academic work beyond science. He also made it possible to visit many international conferences and learn about the importance of good communication.

Jochen Blumberger invited me to London several times and generally welcomed me into his group with open arms. I really enjoyed the scientific discussions and general atmosphere in this group, and I fondly remember the group dinner outings.

Harald Oberhofer always stood by his open door policy and endured many discussions on all levels in our scientific research and any publication production.

I'm also very grateful to Simon Rittmeyer, my former Master thesis supervisor and colleague. He sparked my interest in a PhD thesis in the computational chemistry community and set me up with many useful tools and tricks from his own experience.

Further, I want to thank David 'Eggere' Egger for the many discussions (on and off-topic) and walks, and the countless bad and good times we shared throughout the last four years.

I also want to thank Christian Kunkel and Simiam Ghan for many great discussions and the numerous attempts to achieve work-life-balance.

Moreover, I am thankful for the relaxed, welcoming and open-minded spirit present in the group at the chair of the Theoretical Chemistry department and the many discussions with my colleagues (A. Bruix, M. Deimel, H. Heenen, H. Margraf, C. Scheurer). Specifically, I want to honor the lost generation of PhD fellows: Christian, Christoph M., David, Jakob T., Matthias, Simiam.

From the bottom of my heart, I appreciate the support and patience of my friends and family. I have not made life easy for you, thank you for enduring my shenanigans.

Finally, I would like to acknowledge financial support from the Technische Universität München-Institute for Advanced Study, funded by the German Excellence Initiative (and the European Union Seventh Framework Programme under Grant Agreement No. 291763) as well as the Solar Technologies Go Hybrid Initiative of the State of Bavaria. On this occasion, the TUM Graduate School's financial travel support is gratefully acknowledged.

A. Electrostatic interactions in multipole expansion.

THE LONG-RANGE CLASSICAL ELECTROSTATIC INTERACTIONS between two (well-separated) molecules, with charge density ρ_1 and ρ_2, constitutes a major contributions to the total intermoleculer interaction energy with

$$E^{\text{el.stat.}} = \iint \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r} d\boldsymbol{r}'. \tag{A.1}$$

Evaluation and implementation of the classical electrostatic interactions in effective force field techniques suffers from two major concerns. First, the exact charge distributions, ρ_1 and ρ_2 are not known, as the calculation of the latter from first-principles techniques like DFT produces large numerical overhead unfeasible in force field calculations. Instead, transferable approximate representations are highly desirable. Secondly, the numerical cost of the real-space double integral above is quite exhaustive, prohibitive to efficient force field simulation. Hence, a choice of approximate representations including efficient potentials for the electrostatic interactions are ideal for effective atomistic simulations. In practice, truncated permanent multicenter multipole expansion series are employed to estimate the electronic dsitribution and (long-range) classical electrostatic interactions, condensing the charge potential onto atom centers, partly extended by off-nuclear sites. In this chapter, the basic concepts of the multicenter multipole expansion (MME) and multipole expansion potential (MEP) are briefly reviewed.

A.1. Series expansion of the classical electrostatic potential.

The classical electrostatic interaction potential ϕ at an arbitrary coordinate $\boldsymbol{r}$ due to a charge distribution ρ_1, where we assume the origin close to (or within) the charge distribution, is

$$\phi_1(\boldsymbol{r}) = \int \frac{\rho_1(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}' \tag{A.2}$$

Note, one can always shift the coordinate system by a constant vector, $\boldsymbol{r}_0$, so the assumption above is valid. Next, the Coulomb interaction operator is transformed to a series expansion in Cartesian coordinates, which for completeness is shown following both the power series and Taylor series expansion approaches.

Power series expansion.

In a first step, the Coulomb operator is transformed following the binomial theorem, where $r{=}|\boldsymbol{r}|$ and $r'{=}|\boldsymbol{r}'|$,

$$\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} = \left[(\boldsymbol{r}-\boldsymbol{r}')^2\right]^{-\frac{1}{2}} \tag{A.3a}$$

$$= \left[r^2 + r'^2 - 2\boldsymbol{r}\cdot\boldsymbol{r}'\right]^{-\frac{1}{2}} \tag{A.3b}$$

$$= \left[r^2\left(1 + \frac{r'^2}{r^2} - \frac{2}{r^2}\boldsymbol{r}\cdot\boldsymbol{r}'\right)\right]^{-\frac{1}{2}} \tag{A.3c}$$

$$= \frac{1}{r}\cdot\frac{1}{\sqrt{1+\frac{1}{r^2}(r'^2 - 2\boldsymbol{r}\cdot\boldsymbol{r}')}}. \tag{A.3d}$$

Next, the linear expansion of a power series is utilized,

$$\frac{1}{\sqrt{1+s}} = 1 - \frac{1}{2}s + \frac{3}{8}s^2 - \frac{15}{48}s^3 + \dots. \tag{A.4}$$

Together, eq. (A.3d) and eq. (A.4) yield (with terms sorted by orders of $1/r$ and truncated at

second order),

$$\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} = \frac{1}{r}\Bigg[1 - \frac{1}{2}\Big(\frac{1}{r^2}(r'^2 - 2\boldsymbol{r}\cdot\boldsymbol{r}')\Big)^1 + \frac{3}{8}\Big(\frac{1}{r^2}(r'^2 - 2\boldsymbol{r}\cdot\boldsymbol{r}')\Big)^2 \tag{A.5a}$$

$$- \frac{15}{48}\Big(\frac{1}{r^2}(r'^2 - 2\boldsymbol{r}\cdot\boldsymbol{r}')\Big)^3 + \dots\Bigg] \tag{A.5b}$$

$$= \frac{1}{r} + \frac{1}{r^3}(\boldsymbol{r}\cdot\boldsymbol{r}') + \frac{1}{r^5}\Big(\frac{3}{2}(\boldsymbol{r}\cdot\boldsymbol{r}')^2 - \frac{1}{2}r'^2\Big) + \dots \tag{A.5c}$$

Finally, the potential at coordinate $\boldsymbol{r}$ due a charge density ρ_1, displayed explicitly up to second order, is

$$\phi_1(\boldsymbol{r}) = \frac{1}{r}\int \rho_1(\boldsymbol{r}')d\boldsymbol{r}' + \frac{1}{r^3}\int \rho_1(\boldsymbol{r}')(\boldsymbol{r}\cdot\boldsymbol{r}')d\boldsymbol{r}' + \frac{1}{r^5}\int \rho_1(\boldsymbol{r}')\Big(\frac{3}{2}(\boldsymbol{r}\cdot\boldsymbol{r}')^2 - r'^2\Big)d\boldsymbol{r}' \tag{A.6a}$$

$$= \frac{1}{r}q_1 + \frac{1}{r^3}\boldsymbol{d}_1\cdot\boldsymbol{r} + \frac{1}{r^5}\boldsymbol{Q}_1 + \dots \tag{A.6b}$$

In here, the expansion coefficients are indentified with the overall net charge q_1, overall net dipole moment $\boldsymbol{d}_1$ and overall net quadrupole moment $\boldsymbol{Q}_1$ of the charge distribution ρ_1.

Taylor series expansion.

The general Taylor expansion series of a one-dimensional function $f(x)$ located at x_0 is defined by

$$f^{\text{Taylor}}(x, x_0) = \sum_n^{\infty}\frac{1}{n!}(x - x_0)\nabla_x f(x) \tag{A.7a}$$

$$= f(x_0) + (x - x_0)\Big[\frac{d}{dx}f(x)\Big]_{x_0} + \frac{1}{2}(x - x_0)^2\Big[\frac{d^2}{dx^2}f(x)\Big]_{x_0} \tag{A.7b}$$

$$+ \frac{1}{6}(x - x_0)^3\Big[\frac{d^3}{dx^3}f(x)\Big]_{x_0} + \dots$$

The expansion of the (three-dimensional) Coulomb operator with respect to the charge density coordinate variable $\boldsymbol{r}'$ around a center $\boldsymbol{r}'_0$, here by choice of simplicity the origin $\boldsymbol{0}$, yields

$$f^{\text{Taylor}}(\boldsymbol{r}', \boldsymbol{r}'_0{=}\boldsymbol{0}) = \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|} \tag{A.8}$$

$$= \frac{1}{r} + x'\frac{\partial}{\partial x'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big|_{\boldsymbol{r}'_0=\boldsymbol{0}} + y'\frac{\partial}{\partial y'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big|_{\boldsymbol{r}'_0=\boldsymbol{0}} \tag{A.9}$$

$$+ z'\frac{\partial}{\partial z'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big|_{\boldsymbol{r}'_0=\boldsymbol{0}} \tag{A.10}$$

$$+ \frac{1}{2!}x'^2\frac{\partial^2}{\partial x'^2}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big|_{\boldsymbol{r}'_0=\boldsymbol{0}} + \frac{1}{2!}x'y'\frac{\partial^2}{\partial x'\partial y'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big|_{\boldsymbol{r}'_0=\boldsymbol{0}} \tag{A.11}$$

$$+ \frac{1}{2!}x'z'\frac{\partial^2}{\partial x'\partial z'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big|_{\boldsymbol{r}'_0=\boldsymbol{0}}\dots \tag{A.12}$$

$$= \frac{1}{r} + \boldsymbol{r}'\cdot\Big[\nabla_{\boldsymbol{r}'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big]_{\boldsymbol{r}'=\boldsymbol{0}} + \frac{1}{2!}\boldsymbol{r}'^{\mathsf{T}}_0\cdot\Big[\nabla^{\mathsf{T}}_{\boldsymbol{r}'}\nabla_{\boldsymbol{r}'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big]_{\boldsymbol{r}'=\boldsymbol{0}}\cdot\boldsymbol{r}' + \dots \tag{A.13}$$

The Taylor expansion of the classical electrostatic potential of a charge density ρ_1 is

$$\phi_1(\boldsymbol{r}) = \frac{1}{r}\int \rho_1(\boldsymbol{r}')d\boldsymbol{r}' + \cdot\Big[\nabla_{\boldsymbol{r}'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big]_{\boldsymbol{r}'=\boldsymbol{0}}\int \boldsymbol{r}'\rho_1(\boldsymbol{r}')d\boldsymbol{r}' \tag{A.14a}$$

$$+ \frac{1}{2!}\int \boldsymbol{r}'^{\mathsf{T}}\cdot\Big[\nabla^{\mathsf{T}}_{\boldsymbol{r}'}\nabla_{\boldsymbol{r}'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\Big]_{\boldsymbol{r}'=\boldsymbol{0}}\boldsymbol{r}'d\boldsymbol{r}'$$

$$= \frac{1}{r}q_1 + \frac{1}{r^3}\boldsymbol{d}_1\cdot\boldsymbol{r} + \frac{1}{r^5}\boldsymbol{Q}_1 + \dots \tag{A.14b}$$

Again, the expansion coefficients are indentified with the overall net charge q_1, overall net dipole moment $\boldsymbol{d}_1$ and overall net quadrupole moment $\boldsymbol{Q}_1$ of the charge distribution ρ_1.

The multipole moments.

The first expansion term is identified with the overall net charge of the charge distribution ρ_1—$q_1 = \int \rho_1$—located at an arbitrary expansion center, chosen here to be the coordinate origin $\mathbf{0}$,

$$\phi_1^{(1)}(\boldsymbol{r}) = \frac{1}{r} q_1 \tag{A.15}$$

The second order expansion contribution, with total dipole moment $\boldsymbol{d}_1 = \int \rho_1 \boldsymbol{r}'$, is

$$\phi_1^{(2)}(\boldsymbol{r}) = \left[\nabla_{\boldsymbol{r}'} \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|}\right]_{\boldsymbol{r}'_0 = \mathbf{0}} \cdot \int \boldsymbol{r}' \rho_1(\boldsymbol{r}') d\boldsymbol{r}' \tag{A.16a}$$

$$= \boldsymbol{d}_1 \cdot \left[(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})\big((x-x')^2 + (y-y')^2 + (z-z')^2\big)^{-\frac{1}{2}}\right]_{\boldsymbol{r}'_0 = \mathbf{0}} \tag{A.16b}$$

$$= \boldsymbol{d}_1 \cdot \left[\big((x-x')^2 + (y-y')^2 + (z-z')^2\big)^{-\frac{3}{2}} (x-x', y-y', z-z')\right]_{\boldsymbol{r}'_0 = \mathbf{0}} \tag{A.16c}$$

$$= \frac{1}{r^3} \boldsymbol{d}_1 \cdot \boldsymbol{r} \tag{A.16d}$$

Note, eq. (A.16c) clearly reveals a dependence of the dipole interaction potential on the multipole expansion center $\boldsymbol{r}'_0$. Any shift of the expansion center ($\boldsymbol{r}' \rightarrow \boldsymbol{r}' + \boldsymbol{r}_c$) yields a correction term to the potential. On the contrary, any changes of the coordinate system ($\boldsymbol{r} \rightarrow \boldsymbol{r} + \boldsymbol{r}_c$ and $\boldsymbol{r}' \rightarrow \boldsymbol{r}' + \boldsymbol{r}_c$) does not affect the dipole interaction potential.

A discussion of the quadrupole and higher multipole moments is omitted for the sake of brevity.

A.2. The multipole electrostatic interaction energy.

Inserting the first multipole expansion of ρ_1 into the total electrostatic energy yields

$$E^{\text{el.stat.}} = \int \rho_0(\boldsymbol{r}) \phi_1(\boldsymbol{r}) d\boldsymbol{r} \tag{A.17a}$$

$$= \int \rho_0(\boldsymbol{r}) \Big(\frac{q_1}{r} + \frac{1}{r^3} \boldsymbol{r} \cdot \boldsymbol{d}_1 + \dots \Big) d\boldsymbol{r} \tag{A.17b}$$

Next, the coordinate system is shifted by an arbitrary offset Δr ($\boldsymbol{r} \rightarrow \boldsymbol{r} + \Delta \boldsymbol{r}$) which length represents the distance between the multipole expansion centers of the charge density distributions ρ_0 and ρ_1. In a second step, the multipole expansion series is repeated for the potential ϕ_0 present at the center of the first expansion series (which is now located at $\Delta \boldsymbol{r}$) due to the charge density ρ_0

$$E^{\text{el.stat.}} = q_1 \Big(\frac{q_0}{\Delta r} + \frac{1}{\Delta r^3} \boldsymbol{d}_0 \cdot \Delta \boldsymbol{r} + \dots \Big) + \boldsymbol{d}_1 \cdot \Big(\frac{q_0}{\Delta r^3} \Delta \boldsymbol{r} + \dots \Big) + \dots \tag{A.18a}$$

$$= \frac{q_0 q_1}{\Delta r} + q_0 \frac{1}{\Delta r^3} \boldsymbol{d}_1 \cdot \Delta \boldsymbol{r} + q_1 \frac{1}{\Delta r^3} \boldsymbol{d}_2 \cdot \Delta \boldsymbol{r} + \Big(\frac{\boldsymbol{d}_0 \cdot \boldsymbol{d}_1}{\Delta r^3} - \frac{(\boldsymbol{d}_0 \cdot \Delta \boldsymbol{r})(\boldsymbol{d}_1 \cdot \Delta \boldsymbol{r})}{\Delta r^5} \Big) + \dots \tag{A.18b}$$

In here, the first term constitutes a simple point charge interaction, the second and third term constitue point charge point dipole interactions, the fourth term constitutes the point dipole point dipole interaction, etc.. Note, the present derivation focused on an idealized Cartesian multipole representation, other techniques are dedicated to capture more of the non-ideal continuous charge distribution nature of matter. Therefore, the overall electron density is expanded in a set of suitable basis functions f_a, e.g. Gaussian functions [34, 36], which allow an efficient calculation of the

interaction integral,

$$\rho(\boldsymbol{r}) = \sum_a c_a f_a(\boldsymbol{r}) \tag{A.19a}$$

$$E^{\text{el.stat.}} = \sum_a \sum_{a'} c_a c_{a'} \iint f_a(\boldsymbol{r}) \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} f_{a'}(\boldsymbol{r}')\, d\boldsymbol{r} d\boldsymbol{r}' \tag{A.19b}$$

$$f_a^{\text{Cart.Gaussian}}(\boldsymbol{r}) = (x - x_0)^i (y - y_0)^j (z - z_0)^k \exp(-\alpha r^2) \tag{A.19c}$$

A major drawback of the expansion of the classical electrostatic interaction potential and energy with respect to a single center, i.e. all multipole moments or basis functions located at the same position, is the number of expansion terms necessary to convergence to a reasonable accuracy, which becomes computationally prohibitive in force fields. Instead multiple centers are utilized to expand only the electrostatic potential of parts of the charge distributions following an (in principle arbitrary) charge partitionig scheme, which leads to distributed multipoles optimzed to represent the electrostatic potential of a specific contribution. In practice, the individual density contributions $\rho_{1,A}$ are typically targeted towards an approximate atomic charge distribution in a molecule to provide transferability of the force field model.

$$\rho = \sum_A \rho_A(\boldsymbol{r}) \tag{A.20}$$

B. Density functional theory foundation of charge equilibration.

In density functional theory, electron density flows upon bond formation until the total energy reaches a (global) minimum, and hence incorporates the same guiding principal as electronegativity equalization methods. The derivation of charge equilibration schemes from DFT follows a more theoretically sound path, which provies a way to identify the (atomic) charge equilibration parameters with its first-principles equivalents and help find the weaknesses of current QEq models. This chapter is dedicated to a summarized derivation of the chemical potential equalization method as presented by York and Yang [91] and a short comparison to the EEM method.

B.1. Derivation of the chemical potential equalization method.

The density functional theory formulation of the energy E of electron density ρ is

$$E[\rho] = F[\rho] + \int \rho(\boldsymbol{r}) v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r}. \tag{B.1}$$

The universal functional F contains the purely electronic contributions to the energy, including Coulomb, exchange and correlation effects as well as the kinetic energy. Interactions with the nuclei and other fields are collected in the external potential v_{ext}. A perturbation of the DFT ground state by a change in the external potential ∂v_{ext} prompts a charge density response $\partial\rho$. The total energy up to second order is

$$\begin{aligned}
&E[\rho_0 + \partial\rho, v_{\text{ext},0} + \partial v_{\text{ext}}] \\
&\quad = E[\rho_0, v_{\text{ext},0}] + \int \frac{\partial E_0}{\partial \rho(\boldsymbol{r})} \partial\rho(\boldsymbol{r}) d\boldsymbol{r} + \int \frac{\partial E_0}{\partial v_{\text{ext}}(\boldsymbol{r})} \partial v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} \qquad \text{(B.2)} \\
&\qquad + \frac{1}{2} \int \frac{\partial^2 E_0}{\partial \rho(\boldsymbol{r}) \partial \rho(\boldsymbol{r}')} \partial\rho(\boldsymbol{r}) \partial\rho(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' + \int \frac{\partial^2 E_0}{\partial \rho(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}')} \partial\rho(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\
&\qquad + \frac{1}{2} \int \frac{\partial^2 E_0}{\partial v_{\text{ext}}(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}')} \partial v_{\text{ext}}(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} d\boldsymbol{r}' \\
&\quad = E_0 + \int \mu_0 \partial\rho(\boldsymbol{r}) d\boldsymbol{r} + \int (\rho_0(\boldsymbol{r}) + \partial\rho(\boldsymbol{r})) \partial v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} \qquad \text{(B.3)} \\
&\qquad + \frac{1}{2} \iint \eta_0(\boldsymbol{r}, \boldsymbol{r}') \partial\rho(\boldsymbol{r})(\boldsymbol{r}) d\boldsymbol{r} d\boldsymbol{r}'.
\end{aligned}$$

The energy of the (molecular) reference ground state is E_0 ($=E[\rho_0, v_{\text{ext},0}]$). The first and second order derivatives are identified with the molecular chemical potential $\mu_0 = {}^{\partial E}/_{\partial\rho}$ and molecular hardness $\eta_0 = {}^{\partial^2 E}/_{\partial\rho^2} = {}^{\partial^2 F}/_{\partial\rho^2} = {}^{\partial\mu}/_{\partial\rho}$. The energy response to the change in the external potential in Euler-Lagrange form is

$$\begin{aligned}
\frac{\partial(E[\rho_0 + \partial\rho, v_{\text{ext},0} + \partial v_{\text{ext}}] - E[\rho_0, v_{\text{ext},0}])}{\partial\rho(\boldsymbol{r})} &= \mu \qquad \text{(B.4)} \\
&= \int \eta_0(\boldsymbol{r}, \boldsymbol{r}') \partial\rho(\boldsymbol{r}') d\boldsymbol{r}' + \partial v_{\text{ext}}(\boldsymbol{r}) + \mu_0. \qquad \text{(B.5)}
\end{aligned}$$

This is the DFT foundation of the chemical potential equalization method. It formally resembles an expansion of the true chemical potential of the perturbed system, μ, in terms of the non-perturbed system chemical potential, μ_0, and first order corrections, $\int \eta(\boldsymbol{r}, \boldsymbol{r}') \partial\rho(\boldsymbol{r}') d\boldsymbol{r}' + \partial v_{\text{ext}}(\boldsymbol{r})$. The electron density changes induced by the external potential change is expanded in a linear series of normalized, well-behaved basis functions ϕ_i ($\lim_{r\to\infty} \phi_i = 0$, $\int \phi_i(\boldsymbol{r}) d\boldsymbol{r} < \infty$),

$$\partial\rho = \sum_i c_i \phi_i(\boldsymbol{r}). \tag{B.6}$$

Hereby, the Euler-Lagrange equation is cast into a matrix. The parameters are the basis function density $d_i = \int \phi_i(\boldsymbol{r}) d\boldsymbol{r}$, perturbation interaction integral $v_i = \int \phi_i(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r}$ and hardness $\eta_{i,j} = \int \phi_i(\boldsymbol{r}) \eta(\boldsymbol{r}, \boldsymbol{r}') \phi_j(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}'$. The foundation of any practical implementation of CPE in force fields is the matrix working equation

$$\sum_j \eta_{i,j} c_j + v_i - (\mu - \mu_0) d_i = 0. \tag{B.7}$$

The polarization energy (excluding the permanent charge distribution energy with the perturbation potential $\int \rho_0(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r}$) is

$$E_{\text{pol.}} = \sum_{i,j} c_i \eta_{i,j} c_j + \sum c_i v_i - (\mu - \mu_0) \sum_i d_i. \tag{B.8}$$

So far, the hardness parameter η has not been defined in details. It contains the charge density second derivatives of Hartree energy E_H, kinetic energy T_{KS} (practical implementations of DFT apply a Kohn-Sham scheme) and any exchange-correlation effects which have not been captured by the previous two E_{xc},

$$\eta = \frac{\partial^2 F}{\partial\rho \partial\rho'} \tag{B.9}$$

$$= \frac{\partial^2 E_H}{\partial\rho \partial\rho'} + \frac{\partial^2 T_{\text{KS}}}{\partial\rho \partial\rho'} + \frac{\partial^2 E_{\text{xc}}}{\partial\rho \partial\rho'} \tag{B.10}$$

$$= \frac{1}{\Delta \boldsymbol{r}} + \frac{\partial^2 T_{\text{KS}}}{\partial\rho \partial\rho'} + \frac{\partial^2 E_{\text{xc}}}{\partial\rho \partial\rho'}. \tag{B.11}$$

Evaluation of classical Coulomb interactions between two basis functions is not a difficult task. Unfortunately, density functional forms of the kinetic and exchange-correlation energies and, in consequence, their derivatives are not known. In any practical implementations, these have to be estimated. York and Yang originally propose a Hückel-Ansatz-like approach, where these contributions are condensed into a set of (local) density independent parameters $\{f_i\}$ (k typically chosen to be 1),

$$\eta_{i,i} = f_i + \int \phi_i(\boldsymbol{r}) \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \phi_i(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \tag{B.12a}$$

$$\eta_{i,j} = k(f_i + f_j) \int \phi_i(\boldsymbol{r}) \phi_j(\boldsymbol{r}) d\boldsymbol{r} + \int \phi_i(\boldsymbol{r}) \frac{1}{|\boldsymbol{r} - \boldsymbol{r}'|} \phi_j(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}'. \tag{B.12b}$$

A constant kinetic-exchange-correlation contribution in f_i is a rather crude approximation. A well parametrized CPE model proves useful if a system stays close to an *equilibrium* state, but lacks considerably in transferability to other systems or fails to capture drastic effects like bond-formation and -breaking in molecular dynamics simulations.

B.2. Comparison of CPE to EEM.

The CPE method has originally been developed for the electronic response of a molecule in its electronic groundstate. The electronegativity equilibration method by Mortier *et al.* [88, 89] and charge equilibration scheme by Rappe and Goddard [92] are special cases of the CPE method, where the unperturbed reference is a free-atom. A direct comparison of the energetics per individual atom reveals their similarities,

$$\Delta E^{\mathrm{CPE}} = \mu_0 \int \partial\rho(\boldsymbol{r}) d\boldsymbol{r} + \frac{1}{2} \int \eta_0(\boldsymbol{r}, \boldsymbol{r}') \partial\rho(\boldsymbol{r}) \partial\rho(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \tag{B.13}$$
$$+ \int (\rho_0(\boldsymbol{r}) + \partial\rho(\boldsymbol{r})) \partial v_{\mathrm{ext}}(\boldsymbol{r}) d\boldsymbol{r}$$

$$\Delta E^{\mathrm{EEM}} = \chi_A^* q_A + \frac{1}{2} \eta_{A,A}^* q_A^2 + \sum_{B \neq A} q_A \eta_{A,B} q_B. \tag{B.14}$$

The energy of the unperturbed free-atom reference has been omitted, yielding only the polarization correction contributions. The first terms provide an first-order correction to the atomic energy with respect to changes of the charge distribution, where the electrnegativity χ is the negative of the chemical potential μ, χ= -μ [87]. The second terms represent the atomic hardness, i.e. the gradient of the change of energy as function of charge rearrangements. The last terms introduce an external perturbation potential due to changes in the environment. While the latter term is in a general form in the CPE model, it comprises the approximate interactions with other atom-centers in the EEM model, typically reduced to screened Coulomb interactions contained in $\eta_{A,B} q_B$. The unperturbed charge contribution, $\int \rho_0(\boldsymbol{r}) \partial v_{\mathrm{ext}}(\boldsymbol{r}) d\boldsymbol{r}$=0, is zero for a neutral atom reference represented by atomic partial charges. Note, the density change basis functions of CPE translate to simple delta-functions for atom-centered point charges of EEM, $\phi_i(\boldsymbol{r}) = \delta(\boldsymbol{r} - \boldsymbol{r}_i)$.

C. Formula collection for ACKS2 derivation.

This chapter provides an addendum to the derivation of the general ACKS2 method from perturbative KS-DFT and contains a summary of extensive formulae.

C.1. Derivatives related to density functional theory.

This section summarizes the functional derivatives for an explicit-implicit separated DFT formulation,

$$L[\rho, u, v_{\text{ext}}, \mu] = E^{\text{exp}}[\rho] + E^{\text{o}}[u] + \int \rho(\boldsymbol{r}) v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} - \int \rho(\boldsymbol{r}) u(\boldsymbol{r}) d\boldsymbol{r} - \mu \left(\int \rho(\boldsymbol{r}) d\boldsymbol{r} - N \right). \tag{C.1}$$

The first order functional derivatives—typically assumed to be small or zero close to the true ground state—of the Legendre-formulation total energy expression, abbreviated $E{=}L[\rho, u, v_{\text{ext}}, \mu]$, are

$$\frac{\partial}{\partial \rho(\boldsymbol{r})} E = \frac{\partial}{\partial \rho(\boldsymbol{r})} E^{\text{exp}}[\rho] + v_{\text{ext}}(\boldsymbol{r}) - u(\boldsymbol{r}) - \mu \tag{C.2a}$$

$$\frac{\partial}{\partial u(\boldsymbol{r})} E = \frac{\partial}{\partial u(\boldsymbol{r})} E^{\text{o}}[u] - \rho(\boldsymbol{r}) \tag{C.2b}$$

$$\frac{\partial}{\partial v_{\text{ext}}(\boldsymbol{r})} E = \rho(\boldsymbol{r}) \tag{C.2c}$$

$$\frac{\partial}{\partial \mu} E = -\rho(\boldsymbol{r}). \tag{C.2d}$$

The second order functional derivatives exlcuding symmetric contributions, as $\partial^2/(\partial\rho\partial u) E = \partial^2/(\partial u \partial \rho) E$ and so forth, are

$$\frac{\partial^2}{\partial \rho(\boldsymbol{r}) \partial \rho(\boldsymbol{r}')} E = \frac{\partial^2}{\partial \rho(\boldsymbol{r}) \partial \rho(\boldsymbol{r}')} E^{\text{exp}}[\rho] \tag{C.3a}$$

$$\frac{\partial^2}{\partial u(\boldsymbol{r}) \partial u(\boldsymbol{r}')} E = \frac{\partial^2}{\partial u(\boldsymbol{r}) \partial u(\boldsymbol{r}')} E^{\text{o}}[u] \tag{C.3b}$$

$$\frac{\partial^2}{\partial v_{\text{ext}}(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}')} E = 0 \tag{C.3c}$$

$$\frac{\partial^2}{\partial \mu \partial \mu'} E = 0 \tag{C.3d}$$

$$\frac{\partial^2}{\partial \rho(\boldsymbol{r}) \partial u(\boldsymbol{r}')} E = -1 \tag{C.3e}$$

$$\frac{\partial^2}{\partial \rho(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}')} E = 1 \tag{C.3f}$$

$$\frac{\partial^2}{\partial \rho(\boldsymbol{r}') \partial \lambda} E = -1 \tag{C.3g}$$

$$\frac{\partial^2}{\partial u(\boldsymbol{r}) \partial v_{\text{ext}}(\boldsymbol{r}')} E = 0 \tag{C.3h}$$

$$\frac{\partial^2}{\partial u(\boldsymbol{r}) \partial \mu} E = 0 \tag{C.3i}$$

$$\frac{\partial^2}{\partial v_{\text{ext}}(\boldsymbol{r}') \partial \mu} E = 0. \tag{C.3j}$$

C.2. Second order energy expansion of electronic response.

The total electronic energy of a system is transformed to a linearly expansion series around an arbitrary unperturbed reference state, $E_0{=}E[\rho_0, u_0, v_{\text{ext}}, \mu_0]$, up to second order in each variable explicitly. Note, inclusion of the expansion with respect to changes in the external potential allows polarization due to the presence of perturbation potentials. The first order derivatives of the total energy with respect to the electron density and auxiliary potential (not the external potential) are approximately zero close to the self-consistent electronic ground state, ${}^{\partial E_0}/_{\partial\rho}{=}0$ and ${}^{\partial E_0}/_{\partial u}{=}0$ (${}^{\partial E_0}/_{\partial v_{\text{ext}}} \neq 0$), which for example is employed in the derivation of density functional tight binding methods. The general second order energy expansion is

$$\Delta L^{\text{resp.}} = E[\rho_0 + \Delta\rho, u_0 + \Delta u, v_{\text{ext},0} + \Delta v_{\text{ext}}, \mu_0 + \Delta\mu] - E[\rho_0, u_0, v_{\text{ext},0}\mu_0] \tag{C.4a}$$

$$\approx E^{\text{2nd}}[\rho_0 + \Delta\rho, u_0 + \Delta u, v_{\text{ext},0} + \Delta v_{\text{ext}}] - E[\rho_0, u_0, v_{\text{ext},0}\mu_0] \tag{C.4b}$$

$$\begin{aligned}
&= \int \frac{\partial E}{\partial \rho(\boldsymbol{r})} \partial\rho(\boldsymbol{r}) d\boldsymbol{r} + \int \frac{\partial E}{\partial u(\boldsymbol{r})} \partial u(\boldsymbol{r}) d\boldsymbol{r} \\
&+ \int \frac{\partial E}{\partial v_{\text{ext}}(\boldsymbol{r})} \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} + \int \frac{\partial E}{\partial \mu} \partial\mu d\boldsymbol{r} \\
&+ \iint \frac{1}{2} \frac{\partial^2 E}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')} \partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{1}{2} \frac{\partial^2 E}{\partial u(\boldsymbol{r})\partial u(\boldsymbol{r}')} \partial u(\boldsymbol{r})\partial u(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{1}{2} \frac{\partial^2 E}{\partial v_{\text{ext}}(\boldsymbol{r})\partial v_{\text{ext}}(\boldsymbol{r}')} \Delta v_{\text{ext}}(\boldsymbol{r}) \Delta v_{\text{ext}}(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{1}{2} \frac{\partial^2 E}{\partial\mu\partial\mu'} (\Delta\mu)^2 d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{\partial^2 E}{\partial\rho(\boldsymbol{r})\partial u(\boldsymbol{r}')} \partial\rho(\boldsymbol{r})\, \partial u(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{\partial^2 E}{\partial\rho(\boldsymbol{r})\partial v_{\text{ext}}(\boldsymbol{r}')} \partial\rho(\boldsymbol{r})\, \Delta v_{\text{ext}}(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{\partial^2 E}{\partial\rho(\boldsymbol{r})\partial\mu} \partial\rho(\boldsymbol{r})\, \Delta\mu d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{\partial^2 E}{\partial u(\boldsymbol{r})\partial v_{\text{ext}}(\boldsymbol{r})'} \partial u(\boldsymbol{r})\, \Delta v_{\text{ext}}(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{\partial^2 E}{\partial u(\boldsymbol{r})\partial\mu} \partial u(\boldsymbol{r})\, \Delta\mu d\boldsymbol{r} d\boldsymbol{r}' \\
&+ \iint \frac{\partial^2 E}{\partial v_{\text{ext}}(\boldsymbol{r})\partial\mu} \partial v_{\text{ext}}(\boldsymbol{r})\, \Delta\mu d\boldsymbol{r} d\boldsymbol{r}'
\end{aligned} \tag{C.4c}$$

$$= \int \left(\frac{\partial E^{\text{exp}}[\rho]}{\partial \rho(\boldsymbol{r})} + v_{\text{ext}}(\boldsymbol{r}) - u(\boldsymbol{r}) - \mu \right)_{[\rho_0, v_{\text{KS},0}, u_{,0}, \mu_0]} \Delta\rho(\boldsymbol{r}) d\boldsymbol{r} \tag{C.4d}$$
$$+ \int \left(\frac{\partial E^{\text{o}}[u]}{\partial u(\boldsymbol{r})} - \rho(\boldsymbol{r}) \right)_{\rho_0, u_0} \Delta u(\boldsymbol{r}) d\boldsymbol{r}$$
$$+ \int \left(\rho(\boldsymbol{r})\right)_{\rho_0} \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r}$$
$$- \int \left(\rho(\boldsymbol{r})\right)_{\rho_0} \Delta\mu d\boldsymbol{r}$$
$$+ \iint \frac{1}{2} \Delta\rho(\boldsymbol{r}) \left(\frac{\partial^2 E^{\text{exp}}[\rho]}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')} \right)_{\rho_0} \Delta\rho(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}'$$
$$- \int \Delta u(\boldsymbol{r}) \Delta\rho(\boldsymbol{r}) d\boldsymbol{r}$$
$$+ \int \Delta\rho(\boldsymbol{r}) \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r}$$
$$+ \iint \frac{1}{2} \Delta u(\boldsymbol{r}) \left(\frac{\partial^2 E^{\text{o}}[u]}{\partial u(\boldsymbol{r})\partial u(\boldsymbol{r}')} \right)_{u_0} \Delta u(\boldsymbol{r}') d\boldsymbol{r} d\boldsymbol{r}'$$

The stationary point (ground state) of the response in an Euler-Lagrange formulism—defined by $^{\partial \Delta L^{\text{resp.}}}/_{\partial \Delta\rho}$, $^{\partial \Delta L^{\text{resp.}}}/_{\partial \Delta u}$ and constraint of total charge conservation enforced by $^{\partial \Delta L^{\text{resp.}}}/_{\partial \Delta\mu}$—to a change in the external potential can now, with help of the second order energy expansion, be expressed solely in terms of non-perturbed reference state contributions. Any derivatives of order greater than two are disregarded.

$$\int \frac{\partial^2 E^{\text{exp}}[\rho_0]}{\partial\rho(\boldsymbol{r})\partial\rho(\boldsymbol{r}')} \Delta\rho(\boldsymbol{r}') d\boldsymbol{r}' + \Delta v_{\text{ext}}(\boldsymbol{r}) - \Delta u(\boldsymbol{r}) = \Delta\mu \tag{C.5a}$$
$$\int \frac{\partial^2 E^{\text{o}}[u_0]}{\partial u(\boldsymbol{r})\partial u(\boldsymbol{r}')} \Delta u(\boldsymbol{r}') d\boldsymbol{r}' - \Delta\rho(\boldsymbol{r}) = 0 \tag{C.5b}$$
$$\int \Delta\rho(\boldsymbol{r}) = 0 \tag{C.5c}$$

C.3. Atomic forces in ACKS2.

Direct computation of atomic forces due to polarization is a valuable tool in molecular dynamics simulations. This section is dedicated to a general derivation of ACKS2 electronic polarization contributions to atomic forces, performed in more detail for a Cartesian Gaussian basis representation.

In the limit of linear electronic response, typically valid for small to medium homogeneous electric field perturbations, the energetic cost to rearrange the electron density from equilibrium is half, but opposite sign to the energy gain by the interaction of the induced electronic density changes $\Delta\rho$ with the external potential [128]. By construction, the ACKS2 model enforces validity of the linear limit of its dielectric response, which I consider an useful sanity check for any future basis set options. Thereby, the total ACKS2 polarization can be simplified [115] to

$$E^{\text{pol.}}_{\text{lin.limit}} \approx \frac{1}{2} \sum_i c_i V_i = \frac{1}{2} \sum_i c_i \int g_i(\boldsymbol{r}) \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r}. \tag{C.6}$$

The general polarization-contribution to the atomic force follows

$$\boldsymbol{F}_A^{\text{pol.}} = -\boldsymbol{\nabla}_A E_{\text{lin.limit}}^{\text{pol.}} \tag{C.7}$$

$$= -\frac{1}{2}\sum_i c_i \int \left[\boldsymbol{\nabla}_A g_i(\boldsymbol{r})\right] \Delta v_{\text{ext}}(\boldsymbol{r}) + g_i(\boldsymbol{r}) \left[\boldsymbol{\nabla}_A \Delta v_{\text{ext}}(\boldsymbol{r})\right] d\boldsymbol{r}. \tag{C.8}$$

This force expression exhibits general validity, independent of the choice of basis set representation for the ACKS2 dielectric response or interaction potential. In a non-additive approach, i.e. non-self-consistent eletronic polarization due to electrostatic interactions with the fixed (multipole) charge potential only, the external potential v_{ext} is independent of the atomic coordinates $\boldsymbol{r}_A$ and hence equal to zero. This assumption is not valid in the scf-ACKS2 approach introduced in publication #2, but I expect the contribution by the iteration cycle to be small and recommend to disregard the second term in the general force expression. Here, the electron density basis function g_i refers to a Gaussian function centered at atom A with individual angular momentum $\{k, l, m\}$ along the $\{x, y, z\}$ Cartesian coordinates, $g_i{=}g^{\{k,l,m\}}(\boldsymbol{r}, \boldsymbol{r}_A)$. Derivation of the first term will be illustrated in some detail for the x-component, the y- and z- components straightforwardly,

$$\frac{\partial}{\partial x_A} g^{\{k\}}(x, x_A) = \frac{\partial}{\partial x_A}[(x - x_A)^k \exp^{-\alpha(x-x_A)^2}] \tag{C.9}$$

$$= (-k) * (x - x_A)^{k-1} \exp^{-\alpha(x-x_A)^2} + 2\alpha(x - x_A)^{k+1} \exp^{(-\alpha(x-x_A)^2)} \tag{C.10}$$

$$= 2\alpha g_i^{\{k+1\}}(x, x_A) - k g_i^{\{k-1\}}(x, x_A). \tag{C.11}$$

The derivative of an one-dimensional Cartesian Gaussian function is a simple superposition of two functions of decreased/increased (by +1) angular momentum. Note, the second term vanishes for s-type Gaussian functions (k=0). In conclusion, a Gaussian function ACKS2 implementation allows analytical evaluation of polarization induced atomic force contributions,

$$F_{A,x}^{\text{pol.}} = \frac{1}{2}\sum_i c_i k \int g_i^{\{k-1,l,m\}}(\boldsymbol{r}) \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} - \sum_i c_i \alpha \int g_i^{\{k+1,l,m\}}(\boldsymbol{r}) \Delta v_{\text{ext}}(\boldsymbol{r}) d\boldsymbol{r} \tag{C.12}$$

The current python implementation of Gaussian-based ACKS2 does not yet support the calculation of atomic forces for two reasons. First, there is currently now way to extract these forces from DFT simulations available and hence, validation of these forces to the parent method is impossible. Second, the study of ACKS2 executed within this research thesis is still in a bottom-up proof-of-principle stage and neglects application in actual molecular dynamics simulations.

D. Bibliography

(1) Gütlein, P.; Lang, L.; Reuter, K.; Blumberger, J.; Oberhofer, H. *Journal of Chemical Theory and Computation* **2019**, *15*, 4516–4525 (cit. on pp. v, 17, 29).

(2) Gütlein, P.; Blumberger, J.; Oberhofer, H. *Journal of Chemical Theory and Computation* **2020**, *(accepted)* (cit. on pp. v, 18).

(3) Hänsel-Hertsch, R.; Di Antonio, M.; Balasubramanian, S. *Nature Reviews Molecular Cell Biology* **2017**, *18*, 279–284 (cit. on p. 1).

(4) Kuhlman, B.; Bradley, P. *Nature Reviews Molecular Cell Biology* **2019**, *20*, 681–697 (cit. on p. 1).

(5) Bruix, A.; Margraf, J. T.; Andersen, M.; Reuter, K. *Nature Catalysis* **2019**, *2*, 659–670 (cit. on p. 1).

(6) Plaisance, C. P.; Beinlich, S. D.; Reuter, K. *The Journal of Physical Chemistry C* **2019**, *123*, 8287–8303 (cit. on p. 1).

(7) Kendrick, B. K.; Alden Mead, C.; Truhlar, D. G. *Chemical Physics* **2002**, *277*, 31–41 (cit. on p. 1).

(8) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. *Journal of the American Chemical Society* **1992**, *114*, 10024–10035 (cit. on pp. 2, 3).

(9) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *Journal of the American Chemical Society* **1996**, *118*, 11225–11236 (cit. on pp. 2, 3).

(10) Halgren, T. A. *Journal of Computational Chemistry* **1996**, *17*, 490–519 (cit. on pp. 2, 3).

(11) Sun, H. *The Journal of Physical Chemistry B* **1998**, *102*, 7338–7364 (cit. on pp. 2, 3).

(12) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *Journal of Computational Chemistry* **2004**, *25*, 1157–1174 (cit. on pp. 2, 3).

(13) Brooks, B. R. et al. *Journal of Computational Chemistry* **2009**, *30*, 1545–1614 (cit. on pp. 2, 3).

(14) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *The Journal of Physical Chemistry. B* **2010**, *114*, 2549–2564 (cit. on pp. 2, 3).

(15) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. *Journal of Physics: Condensed Matter* **2009**, *21*, 333102 (cit. on pp. 2, 5).

(16) Lopes, P. E. M.; Roux, B.; MacKerell, A. D. *Theoretical Chemistry Accounts* **2009**, *124*, 11–28 (cit. on pp. 2, 5).

(17) Baker, C. M. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2015**, *5*, 241–254 (cit. on pp. 2, 5).

(18) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. *Annual Review of Biophysics* **2019**, *48*, 371–394 (cit. on pp. 2, 5, 9).

(19) Wang, J.; Cieplak, P.; Li, J.; Cai, Q.; Hsieh, M.-J.; Luo, R.; Duan, Y. *The Journal of Physical Chemistry B* **2012**, *116*, 7088–7101 (cit. on p. 2).

(20) Mohebifar, M.; Johnson, E. R.; Rowley, C. N. *Journal of Chemical Theory and Computation* **2017**, *13*, 6146–6157 (cit. on p. 2).

(21) Cerezo, J.; Prampolini, G.; Cacelli, I. *Theoretical Chemistry Accounts* **2018**, *137*, 80 (cit. on p. 2).

(22) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; van Duin, A. C. T. *npj Computational Materials* **2016**, *2*, 15011 (cit. on p. 2).

(23) Behler, J.; Parrinello, M. *Physical Review Letters* **2007**, *98*, 146401 (cit. on p. 2).

(24) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. *Physical Review Letters* **2010**, *104*, 136403 (cit. on p. 2).

(25) Behler, J. *International Journal of Quantum Chemistry* **2015**, *115*, 1032–1050 (cit. on p. 2).

(26) Bartók, A. P.; Csányi, G. *International Journal of Quantum Chemistry* **2015**, *115*, 1051–1057 (cit. on p. 2).

(27) Mackerell, A. D. *Journal of Computational Chemistry* **2004**, *25*, 1584–1604 (cit. on p. 2).

(28) Mueller, J. E.; van Duin, A. C. T.; Goddard, W. A. *The Journal of Physical Chemistry C* **2010**, *114*, 4939–4949 (cit. on p. 2).

(29) Yun, K.-S.; Pai, S. J.; Yeo, B. C.; Lee, K.-R.; Kim, S.-J.; Han, S. S. *The Journal of Physical Chemistry Letters* **2017**, *8*, 2812–2818 (cit. on p. 2).

(30) Heenen, H. H.; Voss, J.; Scheurer, C.; Reuter, K.; Luntz, A. C. *The Journal of Physical Chemistry Letters* **2019**, *10*, 2264–2269 (cit. on p. 3).

(31) Giannini, S.; Carof, A.; Blumberger, J. *The Journal of Physical Chemistry Letters* **2018**, *9*, 3116–3123 (cit. on p. 3).

(32) Giannini, S.; Carof, A.; Ellis, M.; Yang, H.; Ziogos, O. G.; Ghosh, S.; Blumberger, J. *Nature Communications* **2019**, *10*, 3843 (cit. on p. 3).

(33) Futera, Z.; Blumberger, J. *Journal of Chemical Theory and Computation* **2019**, *15*, 613–624 (cit. on p. 3).

(34) Stone, A. J. *The Journal of Physical Chemistry A* **2011**, *115*, 7017–7027 (cit. on pp. 3, 35).

(35) Wang, Q.; Rackers, J. A.; He, C.; Qi, R.; Narth, C.; Lagardere, L.; Gresh, N.; Ponder, J. W.; Piquemal, J.-P.; Ren, P. *Journal of Chemical Theory and Computation* **2015**, *11*, 2609–2618 (cit. on p. 3).

(36) Stone, A.; Alderton, M. *Molecular Physics* **1985**, *56*, 1047–1064 (cit. on pp. 3, 35).

(37) Biegler-könig, F. W.; Bader, R. F. W.; Tang, T.-H. *Journal of Computational Chemistry* **1982**, *3*, 317–328 (cit. on p. 3).

(38) Hirshfeld, F. L. *Theoretica Chimica Acta* **1977**, *44*, 129–138 (cit. on p. 3).

(39) Mulliken, R. S. *The Journal of Chemical Physics* **1955**, *23*, 1833–1840 (cit. on p. 3).

(40) Momany, F. A. *The Journal of Physical Chemistry* **1978**, *82*, 592–601 (cit. on p. 3).

(41) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *The Journal of Physical Chemistry* **1993**, *97*, 10269–10280 (cit. on p. 3).

(42) De Proft, F.; Van Alsenoy, C.; Peeters, A.; Langenaeker, W.; Geerlings, P. *Journal of Computational Chemistry* **2002**, *23*, 1198–1209 (cit. on p. 3).

(43) Mobley, D. L.; Dumont, É.; Chodera, J. D.; Dill, K. A. *The Journal of Physical Chemistry B* **2007**, *111*, 2242–2254 (cit. on pp. 3, 4).

(44) Verstraelen, T.; Ayers, P. W.; Van Speybroeck, V.; Waroquier, M. *Journal of Chemical Theory and Computation* **2013**, *9*, 2221–2225 (cit. on p. 3).

(45) Lu, Z.; Zhou, N.; Wu, Q.; Zhang, Y. *Journal of Chemical Theory and Computation* **2011**, *7*, 4038–4049 (cit. on pp. 3, 4).

(46) Kramer, C.; Spinn, A.; Liedl, K. R. *Journal of Chemical Theory and Computation* **2014**, *10*, 4488–4496 (cit. on p. 3).

(47) Cardamone, S.; Hughes, T. J.; Popelier, P. L. A. *Physical Chemistry Chemical Physics* **2014**, *16*, 10367 (cit. on p. 3).

(48) Jakobsen, S.; Bereau, T.; Meuwly, M. *The Journal of Physical Chemistry B* **2015**, *119*, 3034–3045 (cit. on p. 3).

(49) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. *Journal of Chemical Theory and Computation* **2007**, *3*, 1960–1986 (cit. on p. 3).

(50) Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. *Annual Review of Physical Chemistry* **2013**, *64*, 553–578 (cit. on p. 3).

(51) Verstraelen, T.; Van Speybroeck, V.; Waroquier, M. *The Journal of Chemical Physics* **2009**, *131*, 044127 (cit. on pp. 4, 8, 9).

(52) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. *Journal of Chemical Theory and Computation* **2013**, *9*, 4046–4063 (cit. on pp. 4, 6).

(53) Demerdash, O.; Mao, Y.; Liu, T.; Head-Gordon, M.; Head-Gordon, T. *The Journal of Chemical Physics* **2017**, *147*, 161721 (cit. on p. 4).

(54) Van der Vaart, A.; Merz, K. M. *The Journal of Physical Chemistry A* **1999**, *103*, 3321–3329 (cit. on p. 4).

(55) Duan, L.; Zhu, T.; Ji, C.; Zhang, Q.; Zhang, J. Z. H. *Physical Chemistry Chemical Physics* **2017**, *19*, 15273–15284 (cit. on p. 4).

(56) Ryno, S. M.; Lee, S. R.; Sears, J. S.; Risko, C.; Brédas, J.-L. *The Journal of Physical Chemistry C* **2013**, *117*, 13853–13860 (cit. on p. 4).

(57) Tu, Z.; Han, G.; Yi, Y. *The Journal of Physical Chemistry Letters* **2020**, *11*, 2585–2591 (cit. on p. 4).

(58) Caldwell, J. W.; Kollman, P. A. *Journal of the American Chemical Society* **1995**, *117*, 4177–4178 (cit. on p. 4).

(59) Halgren, T. A.; Damm, W. *Current Opinion in Structural Biology* **2001**, *11*, 236–242 (cit. on p. 4).

(60) Warshel, A.; Kato, M.; Pisliakov, A. V. *Journal of Chemical Theory and Computation* **2007**, *3*, 2034–2045 (cit. on p. 4).

(61) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D. *Journal of Chemical Theory and Computation* **2013**, *9*, 5430–5449 (cit. on p. 4).

(62) Geerke, D. P.; van Gunsteren, W. F. *The Journal of Physical Chemistry B* **2007**, *111*, 6425–6436 (cit. on p. 4).

(63) Leontyev, I.; Stuchebrukhov, A. *Physical Chemistry Chemical Physics* **2011**, *13*, 2613 (cit. on p. 4).

(64) Lamoureux, G.; Roux, B. *The Journal of Chemical Physics* **2003**, *119*, 3025–3039 (cit. on p. 5).

(65) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. *Chemical Reviews* **2016**, *116*, 4983–5013 (cit. on p. 5).

(66) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. *Journal of Chemical Theory and Computation* **2005**, *1*, 153–168 (cit. on p. 5).

(67) Vorobyov, I. V.; Anisimov, V. M.; MacKerell, A. D. *The Journal of Physical Chemistry B* **2005**, *109*, 18988–18999 (cit. on p. 5).

(68) Harder, E.; Anisimov, V. M.; Vorobyov, I. V.; Lopes, P. E. M.; Noskov, S. Y.; Roux, B.; Roux, B. *Journal of Chemical Theory and Computation* **2006**, *2*, 1587–1597 (cit. on p. 5).

(69) Baker, C. M.; Anisimov, V. M.; MacKerell, A. D. *The Journal of Physical Chemistry B* **2011**, *115*, 580–596 (cit. on p. 5).

(70) Huang, J.; Simmonett, A. C.; Pickard, F. C.; MacKerell, A. D.; Brooks, B. R. *The Journal of Chemical Physics* **2017**, *147*, 161702 (cit. on pp. 5, 6).

(71) Lamoureux, G.; MacKerell, A. D.; Roux, B. *The Journal of Chemical Physics* **2003**, *119*, 5185–5197 (cit. on p. 5).

(72) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D. *Chemical Physics Letters* **2006**, *418*, 245–249 (cit. on p. 5).

(73) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D. *The Journal of Chemical Physics* **2013**, *138*, 034508 (cit. on p. 5).

(74) Rick, S. W. *Journal of Computational Chemistry* **2016**, *37*, 2060–2066 (cit. on p. 5).

(75) Applequist, J. *Accounts of Chemical Research* **1977**, *10*, 79–85 (cit. on p. 6).

(76) Kaminski, G. A.; Friesner, R. A.; Zhou, R. *Journal of Computational Chemistry* **2003**, *24*, 267–276 (cit. on p. 6).

(77) Elking, D.; Darden, T.; Woods, R. J. *Journal of Computational Chemistry* **2007**, *28*, 1261–1274 (cit. on p. 6).

(78) Alcaraz, O.; Bitrián, V.; Trullàs, J. *The Journal of Chemical Physics* **2007**, *127*, 154508 (cit. on p. 6).

(79) Sala, J.; Guàrdia, E.; Masia, M. *The Journal of Chemical Physics* **2010**, *133*, 234101 (cit. on p. 6).

(80) Kiss, P. T.; Baranyai, A. *The Journal of Chemical Physics* **2009**, *131*, 204310 (cit. on p. 6).

(81) Schmollngruber, M.; Lesch, V.; Schröder, C.; Heuer, A.; Steinhauser, O. *Physical Chemistry Chemical Physics* **2015**, *17*, 14297–14306 (cit. on p. 6).

(82) Holt, A.; Karlström, G. *Journal of Computational Chemistry* **2008**, *29*, 2033–2038 (cit. on p. 6).

(83) Holt, A.; Boström, J.; Karlström, G.; Lindh, R. *Journal of Computational Chemistry* **2010**, NA–NA (cit. on p. 6).

(84) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. *Journal of Computational Chemistry* **2002**, *23*, 1515–1531 (cit. on pp. 6, 9).

(85) Sanderson, R. T. *Science* **1955**, *121*, 207–208 (cit. on p. 7).

(86) Iczkowski, R. P.; Margrave, J. L. *Journal of the American Chemical Society* **1961**, *83*, 3547–3551 (cit. on p. 7).

(87) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *The Journal of Chemical Physics* **1978**, *68*, 3801–3807 (cit. on pp. 7, 39).

(88) Mortier, W. J.; Van Genechten, K.; Gasteiger, J. *Journal of the American Chemical Society* **1985**, *107*, 829–835 (cit. on pp. 7, 8, 38).

(89) Mortier, W. J.; Ghosh, S. K.; Shankar, S. *Journal of the American Chemical Society* **1986**, *108*, 4315–4320 (cit. on pp. 7, 8, 38).

(90) Van Genechten, K. A.; Mortier, W. J.; Geerlings, P. *The Journal of Chemical Physics* **1987**, *86*, 5063–5071 (cit. on p. 7).

(91) York, D. M.; Yang, W. *The Journal of Chemical Physics* **1996**, *104*, 159–172 (cit. on pp. 7–9, 37).

(92) Rappe, A. K.; Goddard, W. A. *The Journal of Physical Chemistry* **1991**, *95*, 3358–3363 (cit. on pp. 7, 8, 38).

(93) Itskowitz, P.; Berkowitz, M. L. *The Journal of Physical Chemistry A* **1997**, *101*, 5687–5691 (cit. on p. 8).

(94) Chelli, R.; Procacci, P.; Righini, R.; Califano, S. *The Journal of Chemical Physics* **1999**, *111*, 8569–8575 (cit. on pp. 8, 9).

(95) Chelli, R.; Procacci, P. *The Journal of Chemical Physics* **2002**, *117*, 9175–9189 (cit. on pp. 8, 9).

(96) Njo, S. L.; Fan, J.; van de Graaf, B. *Journal of Molecular Catalysis A: Chemical* **1998**, *134*, 79–88 (cit. on p. 8).

(97) Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *The Journal of Chemical Physics* **1999**, *110*, 741–754 (cit. on pp. 8, 9).

(98) Chaves, J.; Barroso, J. M.; Bultinck, P.; Carbó-Dorca, R. *Journal of Chemical Information and Modeling* **2006**, *46*, 1657–1665 (cit. on p. 8).

(99) Rick, S. W.; Stuart, S. J.; Berne, B. J. *The Journal of Chemical Physics* **1994**, *101*, 6141–6156 (cit. on pp. 8, 10).

(100) Van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *The Journal of Physical Chemistry A* **2001**, *105*, 9396–9409 (cit. on p. 8).

(101) Tabacchi, G.; Mundy, C. J.; Hutter, J.; Parrinello, M. *The Journal of Chemical Physics* **2002**, *117*, 1416–1433 (cit. on p. 8).

(102) Bultinck, P.; Langenaeker, W.; Lahorte, P.; De Proft, F.; Geerlings, P.; Van Alsenoy, C.; Tollenaere, J. P. *The Journal of Physical Chemistry A* **2002**, *106*, 7895–7901 (cit. on p. 8).

(103) Kaminski, S.; Giese, T. J.; Gaus, M.; York, D. M.; Elstner, M. *The Journal of Physical Chemistry A* **2012**, *116*, 9131–9141 (cit. on p. 8).

(104) Wilmer, C. E.; Kim, K. C.; Snurr, R. Q. *The Journal of Physical Chemistry Letters* **2012**, *3*, 2506–2511 (cit. on p. 8).

(105) Chelli, R.; Righini, R.; Califano, S.; Procacci, P. *Journal of Molecular Liquids* **2002**, *96-97*, 87–100 (cit. on p. 9).

(106) Lee Warren, G.; Davis, J. E.; Patel, S. *The Journal of Chemical Physics* **2008**, *128*, 144110 (cit. on p. 9).

(107) Chen, J.; Martínez, T. J. *Chemical Physics Letters* **2007**, *438*, 315–320 (cit. on p. 9).

(108) Nistor, R. A.; Polihronov, J. G.; Müser, M. H.; Mosey, N. J. *The Journal of Chemical Physics* **2006**, *125*, 094108 (cit. on p. 9).

(109) Mathieu, D. *The Journal of Chemical Physics* **2007**, *127*, 224103 (cit. on p. 9).

(110) Chen, J.; Hundertmark, D.; Martínez, T. J. *The Journal of Chemical Physics* **2008**, *129*, 214113 (cit. on p. 9).

(111) Lipparini, F.; Lagardère, L.; Stamm, B.; Cancès, E.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J.-P. *Journal of Chemical Theory and Computation* **2014**, *10*, 1638–1651 (cit. on p. 10).

(112) Nocito, D.; Beran, G. J. O. *Journal of Chemical Theory and Computation* **2018**, *14*, 3633–3642 (cit. on p. 10).

(113) Vitale, V.; Dziedzic, J.; Albaugh, A.; Niklasson, A. M. N.; Head-Gordon, T.; Skylaris, C.-K. *The Journal of Chemical Physics* **2017**, *146*, 124115 (cit. on p. 10).

(114) Verstraelen, T.; Ayers, P. W.; Van Speybroeck, V.; Waroquier, M. *The Journal of Chemical Physics* **2013**, *138*, 074108 (cit. on pp. 11, 13, 25).

(115) Verstraelen, T.; Vandenbrande, S.; Ayers, P. W. *The Journal of Chemical Physics* **2014**, *141*, 194114 (cit. on pp. 13, 25, 43).

(116) Kutzelnigg, W. *Journal of Molecular Structure: THEOCHEM* **2006**, *768*, 163–173 (cit. on p. 13).

(117) Hammond, J. R.; Govind, N.; Kowalski, K.; Autschbach, J.; Xantheas, S. S. *The Journal of Chemical Physics* **2009**, *131*, 214103 (cit. on p. 16).

(118) Wu, T.; Kalugina, Y. N.; Thakkar, A. J. *Chemical Physics Letters* **2015**, *635*, 257–261 (cit. on p. 16).

(119) Hait, D.; Head-Gordon, M. *Physical Chemistry Chemical Physics* **2018**, *20*, 19800–19810 (cit. on p. 16).

(120) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1996**, *77*, 3865–3868 (cit. on p. 16).

(121) De Proft, F.; Langenaeker, W.; Geerlings, P. *Journal of Molecular Structure: THEOCHEM* **1995**, *339*, 45–55 (cit. on p. 19).

(122) Foulkes, W. M. C.; Haydock, R. *Physical Review B* **1989**, *39*, 12520–12536 (cit. on pp. 21, 22).

(123) Frauenheim, T.; Seifert, G.; Elsterner, M.; Hajnal, Z.; Jungnickel, G.; Porezag, D.; Suhai, S.; Scholz, R. *physica status solidi (b)* **2000**, *217*, 41–62 (cit. on pp. 21, 22).

(124) Koskinen, P.; Mäkinen, V. *Computational Materials Science* **2009**, *47*, 237–253 (cit. on p. 21).

(125) Gaus, M.; Cui, Q.; Elstner, M. *Journal of Chemical Theory and Computation* **2011**, *7*, 931–948 (cit. on p. 21).

(126) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Physical Review B* **1998**, *58*, 7260–7268 (cit. on p. 22).

(127) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. *Computer Physics Communications* **2009**, *180*, 2175–2196 (cit. on p. 26).

(128) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *The Journal of Physical Chemistry* **1987**, *91*, 6269–6271 (cit. on p. 43).

E. Full excerpts of the peer-reviewed scientific publications.

E.1. Publication #1

Toward First-Principles-Level Polarization Energies in Force Fields: A Gaussian Basis for the Atom-Condensed Kohn-Sham Method.

P. Gütlein, L. Lang, K. Reuter, J. Blumberger and H. Oberhofer
J. Chem. Theory Comput. **15**, 8, 4516-4525 (2019)
DOI 10.1021/acs.jctc.9b00415

JCTC Journal of Chemical Theory and Computation

Article

Cite This: J. Chem. Theory Comput. 2019, 15, 4516−4525

pubs.acs.org/JCTC

Toward First-Principles-Level Polarization Energies in Force Fields: A Gaussian Basis for the Atom-Condensed Kohn−Sham Method

Patrick Gütlein,[†] Lucas Lang,[†] Karsten Reuter,[†] Jochen Blumberger,[‡,¶] and Harald Oberhofer*[,†]

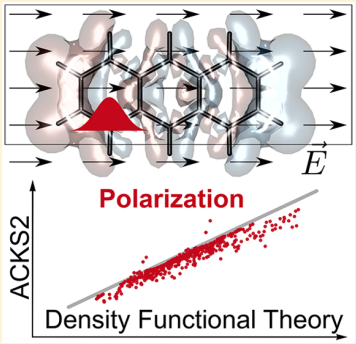
[†]Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, Lichtenbergstrasse 4, D-85747 Garching, Germany

[‡]Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.

[¶]Institute for Advanced Study, Technische Universität München, Lichtenbergstrasse 2 a, D-85748 Garching, Germany

Supporting Information

ABSTRACT: The last 20 years of force field development have shown that even well parametrized classical models need to at least approximate the dielectric response of molecular systems—based, e.g., on atomic polarizabilities—in order to correctly render their structural and dynamic properties. Yet, despite great advances most approaches tend to be based on *ad hoc* assumptions and often insufficiently capture the dielectric response of the system to external perturbations, such as, e.g., charge carriers in semiconducting materials. A possible remedy was recently introduced with the atom-condensed Kohn−Sham density-functional theory approximated to second order (ACKS2), which is fully derived from first principles. Unfortunately, specifically its reliance on first-principles derived parameters so far precluded the widespread adoption of ACKS2. Opening up ACKS2 for general use, we here present a reformulation of the method in terms of Gaussian basis functions, which allows us to determine many of the ACKS2 parameters analytically. Two sets of parameters depending on exchange-correlation interactions are still calculated numerically, but we show that they could be straightforwardly parametrized owing to the smoothness of the new basis. Our approach exhibits three crucial benefits for future applications in force fields: *i*) efficiency, *ii*) accuracy, and *iii*) transferability. We numerically validate our Gaussian augmented ACKS2 model for a set of small hydrocarbons which shows a very good agreement with density-functional theory reference calculations. To further demonstrate the method's accuracy and transferability for realistic systems, we calculate polarization responses and energies of anthracene and tetracene, two major building blocks in organic semiconductors.



1. INTRODUCTION

Owing to their novel materials properties and the vast chemical space they span, molecular organic semiconductors (OS) have attracted the interest of both science and industry for many years.[1−5] In parallel to applied OS research, there is a vibrant community aiming to understand the intrinsic electronic properties and fundamental mechanisms involved in charge and energy transfer processes.[6−12] To the theorist, an extensive toolbox of methods is available to model these compounds, often involving molecular dynamics simulations as an integral component to sample structures and responses to external perturbations.[9,10,13]

Since the dawn of molecular simulation, classical force fields (FF) have been used to efficiently sample the phase space of systems too large to be tackled from first principles.[14−17] Unfortunately, a correct dielectric response to external perturbation poses a severe challenge for an effective FF treatment where electrostatic interactions are usually treated in terms of fixed charges or multipoles.[18−22] A number of approaches have been put forward to tackle this problem, typically based on an atom-centered expansion of classical multipoles.[23−25] An early and successful example of this is the electronegativity equalization method (EEM),[26,27] where a population of fluctuating atomic charges is applied to model the formation of molecular bonds as well as rearrangements of the charge density upon perturbation (cf. Figure 1).

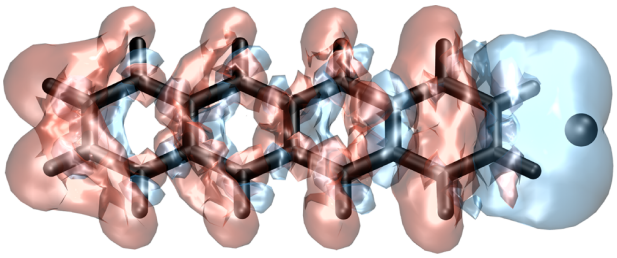


Figure 1. Illustration of the perturbation of the charge density of a tetracene molecule (black scaffold) upon responding to a point charge of $+1e$ on the right (black sphere). Charge accumulation vs ground state is depicted in blue ($1 \times 10^{-3} e\, a_0^{-3}$) and reduction ($-1 \times 10^{-3} e\, a_0^{-3}$) is depicted in red density difference isosurfaces: calculated with FHI-aims and the settings described in the text.

Received: April 30, 2019
Published: June 21, 2019

This simple representation is further refined in subsequently developed methods, like the chemical potential equalization method[28,29] or charge equilibration model.[30] Later models, like the atom−atom charge transfer theory[31] and the split-charge equilibration scheme,[32] tackle the more fundamental shortcomings of EEM, related to the overestimation of long-range charge transfer in dielectrics and vacuum. Yet, so far only *ad hoc* corrections employing interatomic charge constraints based on topological considerations alleviated these limitations, while still manacled to daunting calibration procedures due to the nature of the split-charge variable.[33,34]

Aspiring to eliminate the fundamental issues of previous methods, the atom-condensed Kohn−Sham density-functional theory approximated to second order (ACKS2)[34,35] has been developed on the foundations of EEM. It is a linear electronic response theory derived entirely from first principles and hence provides an exact *ab initio* many-body polarization correction to semiempirical interaction potentials. Thereby, ACKS2 rests on an atom-projected description of changes to charge density and Kohn−Sham potential due to external perturbations such as applied electric fields or other molecules. It thus also bears some conceptional similarity to distributed polarizability approaches pioneered, e.g., by Misquitta and Stone,[36,37] where perturbative responses of molecules are separated into individual atomic contributions. However, in terms of practical applications, the ACKS2 method has so far only been used to complement more traditional EEM models,[38] due to the fact that full ACKS2 response calculations still rely on system dependent parameters derived from density functional theory (DFT) calculations. For the same reasons, ACKS2 has to date never been applied to the calculation of electronic polarization energies, to the best of our knowledge.

In this work we thus present an efficient implementation of the ACKS2 method allowing, for the first time, the calculation of electronic polarization energies on top of the polarization of the electronic density. In order to demonstrate its efficacy, we then apply our implementation to calculate the charge responses and energies of small and medium sized organic molecules to several external fields. Following the original parametrization recipe,[35] we choose a Gaussian representation of the electronic linear response, i.e., changes to the electron density and Kohn−Sham potential, in which many of the ACKS2 energy terms can be calculated analytically. This results in a formally exact description (up to second order) of the electronic polarization response without the need for time-consuming self-consistent iteration of the electron density for each external perturbation, as, e.g., in DFT. Yet, a few parameters depending on exchange and correlation terms of the KS DFT remain. These need to be parametrized from a single, unperturbed DFT calculation, which is facilitated by the smoothness of our new basis. We demonstrate the accuracy of our approach by calculating the electronic polarization energy of a set of hydrocarbons and π-conjugated organic molecules forming crystalline semiconductors. Compared to DFT reference calculations, we obtain very small errors, while the computational cost currently exceeds standard force field calculations—i.e., with fixed bonding topologies and either fixed- or variable charge representations—due to the few parameters still dependent on electronic structure calculations. These will be discussed in detail below. Hence, this work lays the foundation for MD simulations with rigorous and efficient ACKS2-based polarization responses.

1.1. Theoretical Background of the ACKS2 Method. In order to determine the electronic response to an external perturbation, regular Kohn−Sham (KS) DFT offers two distinct approaches. First, one could perform self-consistent DFT calculations of both perturbed and unperturbed system states, to take their difference to compute the changes to the electronic density, the KS potential, and thus the energy. Second, for small perturbations, one could employ a linear response scheme, using density functional perturbation theory (DFPT),[39,40] to determine these properties in a single calculation.

The ACKS2 method follows the latter approach in that it is a simplified linear electronic response model derived from regular KS DFT. At its heart lies a linear equation, derived using the Euler−Lagrange formalism, which describes the response of the electron density to the external perturbation v_{ext}, while a second coupled linear equation accounts for changes of the auxiliary Kohn−Sham potential induced by the density changes. A third and final equation adds a constraint on the total number of electrons for any density rearrangements. Note that this formulation only considers induced changes of the electron density and KS potential and hence cannot describe properties of the unperturbed system state. In order to provide an efficient electronic response scheme for force field simulations, two crucial approximations are introduced to the full KS linear response formalism to arrive at the ACKS2 method. First, in analogy to linear DFPT,[39] the functional derivatives of the energy terms are linearized. This is a particularly important simplification of the KS DFT equations, as—in the limit of a linear electronic response—it removes the necessity of a self-consistent iteration scheme. Second, representations of density and potential response are chosen to be atom-condensed, i.e., all basis sets and parameters are mapped to atom centers similar to previous equilibration schemes like EEM, in order to provide transferability between different geometries.

Considering polarization models, the significant innovation of the ACKS2 approach over, e.g., the EEM method is now to not only account for the density change $\Delta\rho$ but also to explicitly treat all induced changes to the auxiliary KS potential Δv_{KS}, both in mean-field as well as xc components. To make them computationally accessible, both are expanded in a series of atom-centered basis functions

$$\Delta\rho(\mathbf{r}) = \sum_{i}^{N} c_i\, g_i(\mathbf{r}) \tag{1a}$$

$$\Delta v_{\text{KS}}(\mathbf{r}) = \sum_{j}^{M} d_j h_j(\mathbf{r}) \tag{1b}$$

where g_i (h_j) denote the N (M) basis functions, while c_i (d_j) denote the expansion coefficients of the induced density (potential) changes $\Delta\rho$ (Δv_{KS}).

Given the atom-centered basis set expansions of eqs 1, the parametrized form of the KS-DFT linear response, i.e., the working equations of ACKS2, reads

$$\begin{bmatrix} [\eta_{i,i}]_{N,N} & [O_{i,j}]_{N,M} & [D_i]_N \\ [O_{j,i}]_{M,N} & [\chi_{j,j}]_{M,M} & 0 \\ [D_i]_N & 0 & 0 \end{bmatrix} \begin{bmatrix} [c_i]_N \\ [d_j]_M \\ \Delta\tilde{\mu} \end{bmatrix} = \begin{bmatrix} [V_i]_N \\ 0 \\ 0 \end{bmatrix} \tag{2}$$

The matrix on the left-hand side of eq 2 contains the entire KS electronic structure-dependent linear response information on the system and assigns parameters to each basis function. Contrary to earlier schemes,[28−30] the ACKS2 hardness submatrix $\eta_{N,N}$ accounts on a DFT level for both, changes of

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

the Coulomb energy—which was accounted for in earlier schemes—and of the exchange-correlation energies—which was either estimated empirically or neglected. Both energy contributions described by $\eta_{N,N}$ are thereby due to rearrangements of the electron density. The noninteracting linear response kernel parameters $\chi_{M,M}$, on the other hand, incorporate changes of the KS energy induced by changes of the KS potential.

$$\eta_{kl} = \int\int g_k(\mathbf{r})\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\partial^2 E^{\mathrm{xc}}[\rho]}{\partial\rho(\mathbf{r})\partial\rho(\mathbf{r}')}\right)g_l(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{3a}$$

$$\chi_{kl} = \int\int h_k(\mathbf{r})\left(\frac{\partial^2 E^{\mathrm{KS}}[v_{\mathrm{KS}}]}{\partial v_{\mathrm{KS}}(\mathbf{r})\partial v_{\mathrm{KS}}(\mathbf{r}')}\right)h_l(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{3b}$$

Energy changes due to concurrent perturbations of the charge density and KS potential are mediated through the basis set overlap integrals, $O_{kl} = \int g_k(\mathbf{r})h_l(\mathbf{r})d\mathbf{r}$. Additionally, the last row of the left-hand side working matrix (eq 2) contains the integrals of the density basis functions, $D_k = \int g_k(\mathbf{r})d\mathbf{r}$, to establish a constraint on the total charge of the system. In analogy to KS DFT the change in chemical potential $\Delta\tilde{\mu}$ thereby mediates the charge constraint in the form of a Lagrange multiplier. The right-hand side of eq 2 introduces external perturbations acting on the charge density, $V_k = \int g_k(\mathbf{r})\Delta v_{\mathrm{ext}}(\mathbf{r})d\mathbf{r}$. Finally, the ACKS2 based linear electronic response, represented by the expansion coefficients c_i and d_j, is calculated by solving eq 2.

In our reformulation of the ACKS2 approach, a subset of parameters (O, D, and the Coulomb integral in η) can be calculated analytically, i.e., independently of the KS-DFT based electronic structure, while the remainders (χ and the exchange and correlation contributions to η) are evaluated exactly from a reference KS-DFT calculation. Hence, the ACKS2 method theoretically reproduces the exact electronic response of the underlying DFT reference up to second order in the limit of linear response and a complete basis set representation. While the limitation to second order effects at first may seem overly restrictive, recent work, e.g., by Giese and York,[41,42] has shown that carefully conducted second order expansions can yield polarization responses and even geometries nearly indistinguishable from the full KS DFT case. In section 2 we show ACKS2 to go even one step further, yielding excellent agreement not only for the induced molecular dipole moments but also the polarization energy.

1.1.1. ACKS2 Energy Expression. To be useful in actual applications, ACKS2 needs to be able to reproduce a molecule's electronic response to an external potential. Here we focus on the induced dipole moment and the energetic cost of that response known as the polarization energy. While the induced dipole moment follows straightforwardly from the density response (eq 1a), the expression for the energy response expanded in a Taylor-series up to second order is more involved:[35]

$$\begin{aligned}\Delta E &= E_{v_0+\Delta v_{\mathrm{ext}}}[\rho_0+\Delta\rho, v_{\mathrm{KS},0}+\Delta v_{\mathrm{KS}}] - E_{v_0}[\rho_0, v_{\mathrm{KS},0}] \\ &\approx \int\rho_0(\mathbf{r})v_{\mathrm{ext}}(\mathbf{r})d\mathbf{r} + \frac{1}{2}\sum_{i,i'}^{N,N} c_i\,\eta_{i,i'}c_{i'} + \frac{1}{2}\sum_{j,j'}^{M,M} d_j\,\chi_{j,j'}d_{j'} \\ &\quad - \sum_{i,j'}^{N,M} c_i O_{i,j}d_j + \sum_{i}^{N} c_i V_i\end{aligned} \tag{4a}$$

$$\Delta E^{\mathrm{pol.}}_{\mathrm{ACKS2}} = \frac{1}{2}\sum_{i,i'}^{N,N} c_i\,\eta_{i,i'}c_{i'} + \frac{1}{2}\sum_{j,j'}^{M,M} d_j\,\chi_{j,j'}d_{j'} - \sum_{i,j'}^{N,M} c_i O_{i,j}d_j + \sum_{i}^{N} c_i V_i \tag{4b}$$

$$\Delta E^{\mathrm{intra}}_{\mathrm{ACKS2}} = \frac{1}{2}\sum_{i,i'}^{N,N} c_i\,\eta_{i,i'}c_{i'} + \frac{1}{2}\sum_{j,j'}^{M,M} d_j\,\chi_{j,j'}d_{j'} - \sum_{i,j'}^{N,M} c_i O_{i,j}d_j \tag{4c}$$

$$\Delta E^{\mathrm{inter}}_{\mathrm{ACKS2}} = \sum_{i}^{N} c_i V_i \tag{4d}$$

The first term on the right-hand side of eq 4a represents the interaction energy of the unperturbed reference charge density (of both electrons and nuclei), ρ_0, with the external perturbation v_{ext}. The ACKS2 methodology does not account for this term, as it operates in the space of the *induced* charge density $\Delta\rho$ and KS-potential Δv_{KS} only. Yet, it could easily and efficiently be calculated, e.g., from a static force field's multipole representation of the unperturbed charge density.[19,21,22] The energy of induced charge and KS potential polarization $\Delta E^{\mathrm{pol.}}_{\mathrm{ACKS2}}$ provide a static response correction accounting for induction effects. The term $\Delta E^{\mathrm{intra}}_{\mathrm{ACKS2}}$ describes the intramolecular polarization energy necessary to rearrange the charge density of the reference state, while $\Delta E^{\mathrm{inter}}_{\mathrm{ACKS2}}$ is the interaction energy of the induced charge density with the external perturbation. Evaluation of the ACKS2 polarization energy and hence the computational expense to be expected in its application are determined by the vector–matrix–vector multiplications and a single vector–vector scalar product, cf. eq 4b. The induced dipole moment is then a simple scalar product of the density expansion coefficients and density basis function times dipole operator integrals, which can simply be tabulated for a specific Gaussian type density basis set

$$\Delta\mu_{\mathrm{ACKS2}} = \sum_i c_i \int g_i(\mathbf{r})\mathbf{r}d\mathbf{r} \tag{5}$$

1.2. A Gaussian Basis for ACKS2. Inspired by the point charges in classical FFs, earlier studies expanded the response of the electron density and KS potential in Fukui and Hirshfeld functions.[34,35] While these implementations show a relatively high accuracy, they produce a large computational overhead as the basis functions directly depend on the electronic structure of the molecule. In contrast, our implementation employs a set of primitive Cartesian Gaussian functions (GF) centered at the atomic positions. Next to a greatly increased computational efficiency, this choice of representation has the added advantages of a tunable accuracy and—even more importantly—basis set transferability. Thereby, the accuracy of the ACKS2 representation can straightforwardly be increased by adding higher angular momentum basis functions. As will be made clear below, the (atom-centered) basis functions are solely element specific, which ensures the transferability among different molecular structures.

By derivation, the ACKS2 parameters $\chi_{M,M}$, $\eta_{M,M}$, D_N, $O_{M,N}$, and V_N in eq 2 are KS-DFT expectation values and need to be obtained from DFT calculations. In contrast to earlier incarnations of ACKS2, our Gaussian representation allows many of the parameters to be calculated analytically. Nevertheless, the evaluation of $\chi_{M,M}$ as well as exchange-correlation contributions to $\eta_{M,M}$ currently necessitates a KS-DFT electronic structure, whereas Coulomb-interaction portions of $\eta_{M,M}$ as well as all terms in D_N, $O_{M,N}$, and V_N are obtained directly

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516–4525

by recursive GF integration schemes. Hence, computational cost of the analytically derived parameters depends on the specific software implementation of the recursion relations introduced in the Supporting Information. The software thereby used is an in-house proof-of-concept Python package to be made available upon reasonable request. With a view on the use of ACKS2 for force fields, preliminary studies indicate that the ACKS2 parameters that require explicit DFT calculation depend very smoothly on nuclear geometry, see below. Hence, DFT calculations on only one or a few molecular geometries should be sufficient for parametrization of an ACKS2-based force field and drastically improve the computational cost, which in this case would then solely be due to the necessary inversion of the ACKS2 response matrices.

2. RESULTS

2.1. Fitting the New Basis Set. We illustrate our Gaussian basis set implementation of ACKS2 for carbon and hydrogen. We choose a minimal basis of Cartesian s- and p-type Gaussian functions for both the electron density and Kohn−Sham potential responses to balance good accuracy with high computational efficiency of the ACKS2 method as well as a manageable procedure for basis set optimization. The s-type functions describe atom−atom charge transfer, whereas the p-type functions account for intra-atomic dipole polarization. The latter are particularly important for planar or linear molecular fragments often encountered, e.g., in organic semiconductor materials and biochemical systems. Together, both types of functions can be expected to cover the basic electronic response to external electrostatic perturbations, both in-plane and perpendicular to the molecular plane.

An atomic species-dependent sp-GF basis set is generated independently for both carbon and hydrogen by optimizing the width σ of the Gaussian functions

$$f_{a,b,c}(\mathbf{r}, \sigma, \mathbf{R}_A) = (x - X_A)^a (y - Y_A)^b (z - Z_A)^c \times \exp\left(-\frac{1}{2\sigma^2}(\mathbf{r} - \mathbf{R}_A)^2\right) \tag{6}$$

where $\mathbf{R}_A = \{X_A, Y_A, Z_A\}$ denotes the position of atom A, and $\mathbf{r} = \{x, y, z\}$ is the position vector. To find the optimal σ for each elementary s- and p-function, atomic dimers of H and C are subjected individually to a set of external electrostatic perturbations. The basis sets are then optimized for the ACKS2 response to match the full KS-DFT reference by variation of the width (or radial decay) σ of the Gaussian functions separately for the representations of electron density and KS-potential response, respectively. The space of possible perturbations is sampled by a radial and angular distribution of point charges (PC) of +1e placed between 1.3 times and up to 3.5 times the van der Waals radii, complemented by homogeneous electric field (HEF) potentials of similar strengths, yielding a total of 110 and 130 different perturbation potentials for carbon and hydrogen, respectively. A more detailed description of our fitting set can be found in the Supporting Information.

The quality of the basis set is measured by the relative mean error (RME) of the ACKS2 response from the respective full KS-DFT reference

$$\mathrm{RME} = \frac{1}{N_{\mathrm{data}}} \sum_{i}^{N_{\mathrm{data}}} \left| \frac{x_i^{\mathrm{ACKS2}} - x_i^{\mathrm{DFT}}}{x_i^{\mathrm{DFT}}} \right| \tag{7}$$

where x represents the property of interest, e.g., induced dipole moment, intramolecular rearrangement energy $\Delta E^{\mathrm{intra}}_{\mathrm{ACKS2}}$ (eq 4c), and intermolecular interaction energy $\Delta E^{\mathrm{inter}}_{\mathrm{ACKS2}}$ (eq 4d), respectively. Screening of the four-dimensional parameter space (σ_s and σ_p for the electron density and KS potential response functions, respectively) exposed a great number of local minima. Therefore, we applied a global particle swarm optimizer[43] to select the global minimum among several local minima with similar fitness scores, as shown in Table 1. For the present basis set representation, the radial decay matches the order of magnitude of atomic valence shells. This is not surprising, given that the basis needs to represent changes to the density and the potential, respectively. However, this being an atom-condensed approximation of these quantities, likely dependent on the angular momentum order of the expansion, one should not expect a physical relevance of our σ values beyond the correct order of magnitude.

Table 1. Optimized Radial Parameters, i.e., Gaussian Widths σ [a_0] for an sp-Type ACKS2 Basis Set for Hydrogen and Carbon[a]

	$\sigma^{(s)}_{\Delta\rho}$	$\sigma^{(p)}_{\Delta\rho}$	$\sigma^{(s)}_{\Delta v_{KS}}$	$\sigma^{(p)}_{\Delta v_{KS}}$
H	0.608	0.863	0.490	1.000
C	0.446	1.174	0.637	5.652

[a]Note the different parameters for the density and the potential basis, respectively.

The results obtained with the best basis set are shown in Figure 2. For H_2 the relative mean error of induced dipole moments and polarization energy are 4.4% and 4.1% with respect to KS-DFT, while they are 2.5% and 6.5% for C_2, respectively. The optimized sp-GF basis shows excellent agreement with DFT references for induced dipole moments (A, D) and for the polarization energy (B, E) over 3 orders of

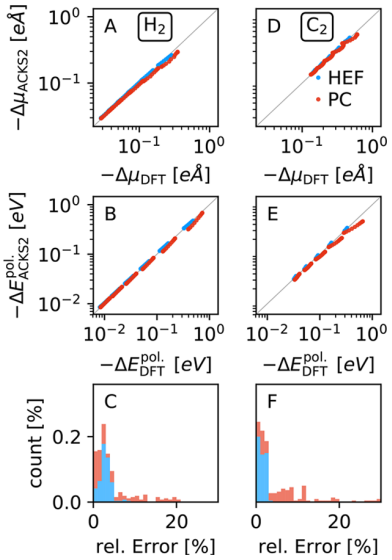


Figure 2. Performance of the fitted sp-GF minimal basis for the hydrogen (A, B, C) and carbon (D, E, F) training sets. Panels (A) and (D) illustrate the induced dipole moment, and panels (B) and (E) depict the polarization energy of the ACKS2 method, cf. eq 4a, compared to the DFT reference. The relative error is summarized in panels (C) and (F).

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

magnitude, from 0.01 to 1.0 eV. In general, ACKS2 predicts the response to HEF perturbations (blue) better than to PC perturbations (red), in particular for the carbon dimer. A more detailed view of the differences in panels (C, F) shows deviations for HEF perturbations to follow a rather narrow distribution up to 5%, while distributions for PC perturbations exhibit a flat-bottomed tail up to 20% error.

An exemplary illustration of the density changes due to external perturbations is shown in Figure 3. Here, a one-

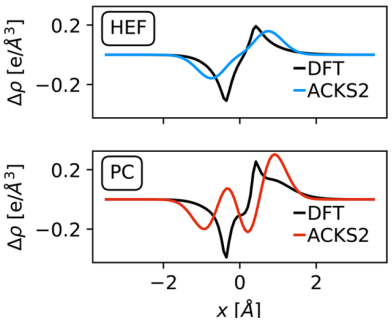


Figure 3. One-dimensional cuts of the induced electron density rearrangements in H_2. The external perturbations are aligned with the molecular bond axis. The response to a PC with center-of-mass distance of 2.0 Å and single elemental charge and a homogeneous electric field matching the PC potential are illustrated.

dimensional cut of the electronic response to a single HEF or PC along the bond of a hydrogen dimer reveals the simplicity of the presented approach based on a sp-GF minimal basis set. Evidently, the fine details of the density response around the nuclei are not and cannot be resolved as the basis functions are not flexible enough. Nevertheless, the very good agreement of integrated properties—i.e., energies and induced dipoles—with respect to DFT demonstrate that only a sufficiently accurate averaged ACKS2 description of charge and potential response is required. Our effective basis thus follows the spirit of other effective methods such as, e.g., pseudopotentials, which are designed to reproduce orbital energies without reproducing the nodal features of the orbitals close to the nucleus. A more in-depth study of the convergence of the density response near the cores with basis set size is beyond the scope of this work.

2.2. Small Molecule Test Set. To test the transferability of the Gaussian basis from H_2 and C_2 to more general CH-containing molecules, we curated a set of ten small hydrocarbons. They were selected for their different chemical features, i.e., four linear alkanes, three compounds containing single, double, and triple bonds $C_2H_{\{2,4,6\}}$, and four unsaturated cyclic hydrocarbons, cf. Table 2. Similar to the training procedure, each molecule is separately exposed to a set of external perturbations sampling the three-dimensional space by a radial and angular distribution of HEF and PC potentials, yielding 30 data points per molecule and 300 data points in total for the test set.

Again, our ACKS2 implementation is capable of predicting the DFT reference electronic response properties over an order of magnitude with good accuracy, as illustrated in Figure 4. The induced dipole moment, panel (A), and polarization energy, panel (B), match the DFT reference data very well. The slight underestimation of both induced dipole and polarization energy can be attributed to a slight overestimation of the intramolecular polarization energy contributions compared to the intermo-

Table 2. Linear Response Prediction for HEF and PC Perturbations of a Hydrocarbon Test Set[a]

alkanes	RME($\Delta E^{\text{pol.}}$)	RME($\Delta\mu$)
methane	10.2	10.5
ethane	8.5	9.0
propane	10.9	9.9
n-butane	11.9	10.7
$C_2H_{\{2,4,6\}}$		
ethane	8.5	9.0
ethylene	13.1	11.3
acetylene	21.4	21.8
cyclic		
cyclopropanene	12.4	11.2
cyclobutadiene	16.4	14.6
cyclopentadiene	14.8	13.2
benzene	18.4	16.4

[a]The relative errors of the ACKS2 polarization energy (RME($\Delta E^{\text{pol.}}$) [%]) and induced dipole moment (RME($\Delta\mu$) [%]) compared to the DFT reference are listed.

lecular interactions. We believe this to be a side effect of the minimal basis set representation. To our knowledge, this is for the first time that the polarization energy $\Delta E^{\text{pol}}_{\text{ACKS2}}$ has been determined solely within the ACKS2 framework. Our results further show that atom-centered basis sets, though optimized for simple diatomics, are transferable to other molecules with different chemical features (bonds between heteroatoms, double, triple bond, etc.). Unfortunately, simple stacking of more basis functions does not necessarily improve the basis set representation, as the ACKS2 method does not follow the variational principle of KS-DFT due to the perturbative treatment of the KS response matrix χ.[34,35] Again, deviations are generally smaller for HEF perturbations (blue) compared to external PCs (red) as illustrated by the distributions of relative errors given in Figure 4C−E.

The subset of linear alkanes exhibits narrow distributions in the 7−10% range, with a slight tail toward larger errors for PC perturbations. Hence, accuracy slightly falls off compared to the training set data averaging at relative errors of about 4.5%. Nevertheless, such small errors provide evidence for the transferability of the simple GF basis set representation. A more detailed view, given in Table 2, indicates small increases in error with increasing chain size except ethane, as it resembles the geometric structure of the carbon training dimer C_2, the closest in the present test set.

The $C_2H_{\{2,4,6\}}$ compounds mimic the effect of different chemical bonding types and local electronic structure. They display broader distributions of errors, Figure 4, with shallow tails up to the 30% range, but are again centered at the 10% mark. A closer look reveals an increased error with larger bond orders, Table 2, due to enhanced nonlinear response effects in π-conjugated organic systems, in particular triple bonded carbon.

Finally, the set of cyclic unsaturated hydrocarbons, relevant for application in OS material modeling, reproduces the DFT reference reasonably well. The distribution of errors, cf. Figure 4, is again centered around the 10% mark. It closely resembles the deviations of the linear alkane chains but exhibits a slightly enhanced tail up to the 30% mark, similar to the $C_2H_{\{2,4,6\}}$ molecules.

2.3. Acenes. While the test set comprised of simple molecules already demonstrates the efficacy of ACKS2 for different chemical species and bonding situations, we now turn

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

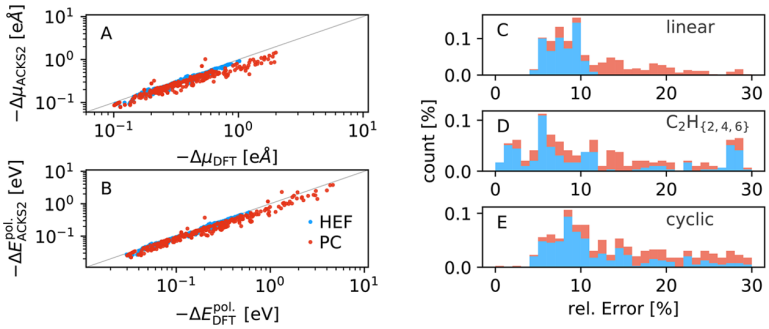


Figure 4. Performance of the ACKS2 method in predicting induced dipole moment (A) and polarization energy (B). The left column depicts the absolute induced dipole moment and polarization energy, while the right column displays the relative errors of the linear one-dim. alkanes (C), $C_2H_{\{2,4,6\}}$ compounds (D), and cyclic molecules subsets (E).

to two larger molecules that form semiconducting organic crystals, anthracene and tetracene.[44] The electronic response properties of these acenes are not trivial and exhibit complex effects as illustrated in Figure 1. For a number of HEF and PC perturbations the ACKS2 response again matches the DFT reference properties very well as depicted in Figure 5. Hence, the present work provides a crucial first step toward an ACKS2-based electronic response function for large, application-relevant molecules.

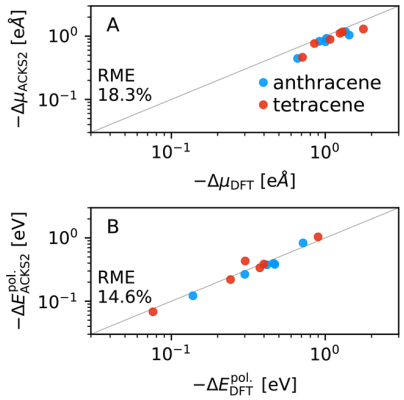


Figure 5. Numerical validation for anthracene and tetracene polarization properties. In panel (A), the induced dipole moment of the ACKS2 method is depicted vs the DFT reference, and panel (B) shows the polarization energy defined in eq 4a. RME gives the relative mean error for each property.

2.4. Polarizability and Raman Intensity. We now turn to the calculation of polarizability and the nuclear derivative of polarizability determining the intensity of Raman spectra, where other effective polarization models often struggle to yield accurate results.[45,46]

To this end, we compute the isotropic polarizability—i.e., the average of the diagonal of the polarizability tensor, $\alpha^{\mathrm{iso}} = \frac{1}{3}(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})$—and its dependence on the molecular structure, by perturbing the equilibrium geometry of benzene along the symmetric ring-breathing mode.

Here, the polarizability is calculated numerically by a finite electric field differentiation of the total molecular dipole moment with a field strength of 0.01 VÅ^{-1}. Figure 6 depicts the isotropic polarizability for small displacements along the benzene breathing mode. Separately, Figure 6 highlights changes of α^{iso} with respect to the polarizability in the vibrational

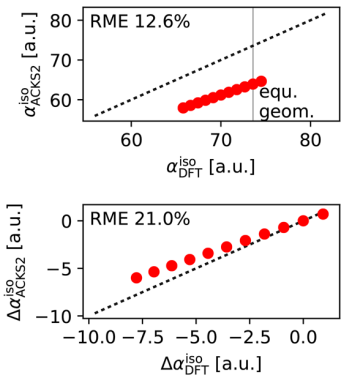


Figure 6. Illustration of the isotropic polarizability and its relative changes for a benzene symmetric breathing mode.

equilibrium geometry—$\Delta\alpha^{\mathrm{iso}} = \alpha^{\mathrm{iso}}_{\mathrm{vib}} - \alpha^{\mathrm{iso}}_{\mathrm{equ}}$. The overall trend of total isotropic molecular polarizabilities of benzene, relative mean error of 12.6%, matches the accuracy of previously introduced molecular test sets of small organic molecules and acenes, see Figures 4 and 5. However, a closer look reveals a slight, albeit systematic growth of the relative error of the isotropic molecular polarizabilities outside the equilibrium configuration. This is most likely a consequence of the small basis set, fitted to yield a reasonable approximation of the overall electronic response properties but potentially too insensitive to capture quantitatively the small changes in polarization response with nuclear displacements.

Finally, we point out that each of the distorted geometries still necessitated a full DFT calculation to extract the ACKS2 response matrices as discussed in detail in the next section.

2.5. Nonanalytic Parameters of ACKS2. Envisioning the ACKS2 method as polarization correction to force fields, the current bottleneck remains the evaluation of the Kohn−Sham DFT electronic structure dependent parameters, namely the exchange-correlation contributions to the hardness—$\eta^{\mathrm{xc}}_{i,j}$—as well as the noninteracting linear response kernel—$\chi_{i,j}$. Our preliminary studies of those ACKS2 parameters indicate a smooth dependency on nuclear geometry due to our smooth Gaussian representation of the response properties. We exemplify this on benzene, cf. Figure 7a, where we symmetrically alter the equilibrium distance of all C−H bonds ($\Delta d^{\mathrm{bond}}_{\mathrm{C,H}} = 1.09$ Å) over a range of −0.075 Å to +0.075 Å by equally displacing the C atoms inward (outward) and the H atoms outward (inward) for a bond expansion (contraction). This then also changes the C−C distances in the ring, where the respective

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

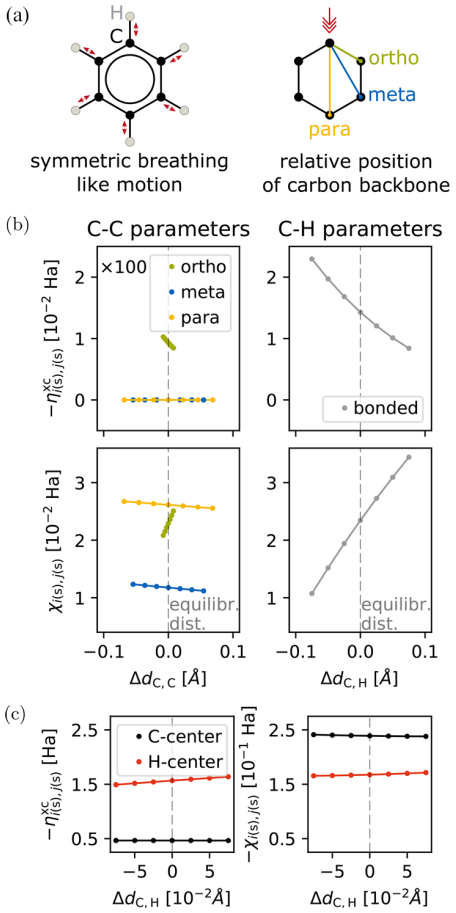


Figure 7. Illustration of pairwise distance dependency of KS-DFT derived ACKS2 parameters for atomic s-type Gaussian basis functions for breathing mode of benzene. (a) Benzene molecular motion and relative position of carbon atoms. The pairwise interatomic distances are applied as metric for the ACKS2 parameters following the color coding scheme. (b) Off-center KS-DFT parameters for symmetric breathing mode. (c) Same-center KS-DFT parameters for symmetric breathing mode.

change of the pairwise distance between two adjacent C atoms (ortho), two second-nearest C atoms (meta), and two-third-nearest C atoms (para) on the ring follows simple geometric dependencies on the C−H bond length ($\Delta d^{\mathrm{ortho}}_{\mathrm{C,C}}$ = 1.40 Å, $\Delta d^{\mathrm{meta}}_{\mathrm{C,C}}$ = 2.42 Å, $\Delta d^{\mathrm{para}}_{\mathrm{C,C}}$ = 2.79 Å). As illustrated in Figure 7b, the corresponding response matrix elements of the exchange-correlation contributions as well as the noninteracting linear response kernel vary smoothly near the equilibrium interatomic positions. For the sake of simplicity we here focus on a pure s-basis for both density and potential and plot carbon−carbon as well as bonded carbon−hydrogen interaction elements. Hydrogen−hydrogen terms as well as hydrogen−carbon terms of atoms, that are not directly bonded, are not shown as we find them negligibly small. The exchange-correlation contributions decay to zero rapidly with distance, as shown in the top graph of Figure 7b. Hence, same-center contributions, depicted in Figure 7c, and off-center matrix elements of atoms partaking in a bond are the most important features of $\eta^{\mathrm{xc}}_{i,j}$ and could be tabulated or interpolated smoothly from a DFT reference. The elements of the noninteracting linear response kernel $\chi_{i,j}$ generally decay less rapidly and involve considerable nonlocal contributions due to the long-range effects of the Kohn−Sham potential changes. Nevertheless, they vary smoothly with interatomic distances, which should allow for interpolation from a few DFT reference calculations. Considering basis functions of higher angular momentum like p-type GF, the ACKS2 model requires descriptors that are able to account for an angular resolution of the latter response parameters as well compared to a simple interatomic distance metric of an s-type representation, as illustrated here. One candidate for such a metric could, e.g., be a simple GF basis overlap, which naturally encodes distances and relative orientations of basis functions. We found that both $\eta^{\mathrm{xc}}_{i,j}$ and $\chi_{i,j}$ still vary smoothly with the their respective overlap matrix elements (see the Supporting Information). Hence, for a given molecule or set of molecules, a careful parametrization of an ACKS2 model from a set of a few DFT reference calculations should straightforwardly be possible and greatly facilitate the method's application, e.g., as a polarization correction in molecular dynamics calculations.

2.6. Comparison to Other Polarizeable Methods. Setting our results into perspective, even state of the art approximations to DFT such as the semiempirical modified neglect of differential overlap (MNDO) tend to struggle to accurately reproduce polarization responses with relative mean errors (RME) for the induced dipoles of 25%,[47] compared to hybrid-level DFT. Similar failures can be found for straightforward applications of density functional tight binding DFTB2.[42] Both of these can be substantially improved though, by coupling the respective ground state electronic structures to carefully parametrized implementations of the chemical potential equalization (CPE). A single postprocessing step of CPE, for example, yields an RME of 12% and 3% for induced dipole moments and molecular polarizabilties, respectively, for MNDO[47] and 5% for molecular polarizabilities in DFTB2.[42] Using a careful parametrization, similar results—RME of 12% for induced dipole moments and RME of 4% for molecular isotropic polarizabilities—can also be found for a direct application of CPE[29] to a set of small organic molecules, very similar to the set used here (cf. section 2.2). On the more empirical side, the popular Drude oscillator model, for example, has been shown to yield polarizabilities down to 3%,[48] if carefully parametrized, albeit at the potential cost of a loss of transferability. Furthermore, relying on fixed charges, such an approach could potentially struggle with polarizations along conjugated bonds such as in the acenes discussed in section 2.3 above.

Additionally, we point out that these studies did not include any correction to the polarization energy terms, which are conceptionally more involved (see section 1.1.1) to accurately estimate and typically would require even more sophisticated parameter training. In ACKS2, on the other hand, density response and polarization energy can be computed from the same set of matrices. Finally, given that ACKS2 is parametrized directly from DFT, currently without any empirical fitting, the accuracy of the model rests on a single parameter, the quality of the basis set.

3. DISCUSSION

In conclusion, we presented a highly efficient Gaussian basis set implementation of the ACKS2 method, that is transferable from diatomics to general molecules and that shows good accuracy when compared to reference DFT calculations. A new minimal GF basis set has been developed for carbon and hydrogen species, and consequent numerical validation illustrated the

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

power of the ACKS2 method despite the rather simplistic training procedure.

Employing the new Gaussian basis set for ACKS2 responses, we found it to yield admirable accuracies compared to DFT reference calculations. The electronic response properties of small molecules such as alkanes and unsaturated cyclic hydrocarbons, as well as single, double, and triple bonded compounds, demonstrated the accuracy of ACKS2 for a large number of bonding perturbation scenarios. As observables, we thereby calculated induced dipole moments and presented the first evaluation of the polarization energy within the ACKS2 formalism, which matches DFT energies to a remarkable degree. Although one-dimensional cuts through the ACKS2 density and potential response show great differences to their DFT counterparts, the observables nevertheless agree well due to error cancellations when integrated over all space. Finally, we demonstrated the good performance of ACKS2 on application-relevant molecules in organic semiconductor research, tetracene and anthracene.

Future work will focus on the improvement of basis sets, extension to further compounds, and in particular the development of strategies for electronic structure free incorporation of ACKS2 polarization energy in force field based molecular dynamics simulation. Due to the Gaussian basis introduced in this work, most of the ACKS2 parameters can be determined analytically, yet a few of them still rely on a full electronic structure evaluation. In the current state, the exchange-correlation contributions to the hardness η and the KS noninteracting linear response kernel χ are evaluated from the KS orbitals of the reference system. For an application of the ACKS2 model in a force field this computational bottleneck needs to be removed. We envision the elimination of the computationally expensive DFT parametrization process in a second step by geometric scaling or interpolation schemes, facilitated by the smoothness of the ACKS2 parameters with respect to nuclear displacements, to finally obtain linear electronic responses that only depend on nuclear geometry and hence are ideally suited for force fields. In such a framework, conventional interatomic potentials would account for the intramolecular interactions, permanent electrostatic interactions, and dispersion corrections, while the ACKS2 model provides static electronic polarization.

4. METHODS

Throughout this study, DFT reference calculations as well as the KS electronic structure dependent ACKS2 parameter evaluations were carried out with the FHI-aims full potential all electron DFT simulation package.[49] We applied the PBE generalized gradient approximated density functional. Integrations were conducted using "tight" integration grids with wave functions expanded in a *tier 3* numeric atomic orbital basis, to ensure the numerical convergence of our results.

While PBE certainly is not the perfect functional for both the C_2 dimer used in the basis set parametrization, as well as the polarizability of acenes, the exact choice of functional is not relevant. As shown earlier,[49] even an imperfect description of the electronic structure of elementary dimers seems sufficient to yield accurate and transferable basis sets (cf. section 2.1). On the other hand, in our accuracy tests, ACKS2 shows itself able to very well reproduce the DFT reference. We have no reason to suspect this not to hold for other, modern DFT functionals such as the promising SCAN meta-GGA,[50] that can be expected to yield better polarization properties of molecules.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.9b00415.

Theoretical foundations of ACKS2 approach and technical details of evaluation of new Gaussian basis and description of training and test sets used in this work, as well as description of how to extract reference data from DFT calculations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: harald.oberhofer@tum.de.

ORCID

Karsten Reuter: 0000-0001-8473-8659
Jochen Blumberger: 0000-0002-1546-6765
Harald Oberhofer: 0000-0002-5791-6736

Author Contributions

L.L. initially implemented the Gaussian-based ACKS2 method. P.G. refined and extended this implementation. Furthermore, P.G. implemented the ACKS2 energy terms and performed all training and testing calculations. The initial idea to the project was conceived by H.O. and K.R. H.O. and J.B. derived the ACKS2 energy expressions, designed, and supervised the project. All authors contributed to the writing of the manuscript.

Notes

The authors declare no competing financial interest.

Supporting training data for the Gaussian parametrization as well as raw data of the test sets and all software used in this work present as an in-house proof-of-concept Python package for this article are available from the corresponding author upon reasonable request.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the Solar Technologies Go Hybrid Initiative of the State of Bavaria. P.G. further acknowledges the support of the Technische Universität München − Institute for Advanced Study, funded by the German Excellence Initiative (and the European Union Seventh Framework Programme under Grant Agreement No. 291763). H.O. acknowledges support from the German Research Foundation (DFG), priority program 1928 COORNETS, Grant OB425/3-1. J.B. acknowledges TUM-IAS for the award of a generous Hans-Fischer Fellowship.

■ REFERENCES

(1) Forrest, S. R. The Path to Ubiquitous and Low-Cost Organic Electronic Appliances on Plastic. *Nature* **2004**, *428*, 911−918.

(2) Brédas, J.-L.; Norton, J. E.; Cornil, J.; Coropceanu, V. Molecular understanding of organic solar cells: the challenges. *Acc. Chem. Res.* **2009**, *42*, 1691−1699.

(3) Wang, L.; Nan, G.; Yang, X.; Peng, Q.; Li, Q.; Shuai, Z. Computational methods for design of organic materials with high charge mobility. *Chem. Soc. Rev.* **2010**, *39*, 423−434.

(4) Minemawari, H.; Yamada, T.; Matsui, H.; Tsutsumi, J.; Haas, S.; Chiba, R.; Kumai, R.; Hasegawa, T. Inkjet Printing of Single-Crystal Films. *Nature* **2011**, *475*, 364−367.

(5) Stavrinidou, E.; Gabrielsson, R.; Gomez, E.; Crispin, X.; Nilsson, O.; Simon, D. T.; Berggren, M. Electronic Plants. *Sci. Adv.* **2015**, *1*, No. e1501136.

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

(6) Coropceanu, V.; Cornil, J.; da Silva Filho, D. A.; Olivier, Y.; Silbey, R.; Brédas, J.-L. Charge Transport in Organic Semiconductors. *Chem. Rev.* **2007**, *107*, 926−952.
(7) Jailaubekov, A. E.; Willard, A. P.; Tritsch, J. R.; Chan, W.-L.; Sai, N.; Gearba, R.; Kaake, L. G.; Williams, K. J.; Leung, K.; Rossky, P. J.; Zhu, X.-Y. Hot charge-transfer excitons set the time limit for charge separation at donor/acceptor interfaces in organic photovoltaics. *Nat. Mater.* **2013**, *12*, 66−73.
(8) Oberhofer, H.; Reuter, K.; Blumberger, J. Charge Transport in Molecular Materials: An Assessment of Computational Methods. *Chem. Rev.* **2017**, *117*, 10319−10357.
(9) Nelson, J.; Kwiatkowski, J. J.; Kirkpatrick, J.; Frost, J. M. Modeling Charge Transport in Organic Photovoltaic Materials. *Acc. Chem. Res.* **2009**, *42*, 1768−1778.
(10) Troisi, A. Charge transport in high mobility molecular semiconductors: classical models and new theories. *Chem. Soc. Rev.* **2011**, *40*, 2347−2358.
(11) Giannini, S.; Carof, A.; Blumberger, J. Crossover from Hopping to Band-Like Charge Transport in an Organic Semiconductor Model: Atomistic Nonadiabatic Molecular Dynamics Simulation. *J. Phys. Chem. Lett.* **2018**, *9*, 3116−3123.
(12) Spencer, J.; Gajdos, F.; Blumberger, J. FOB-SH: Fragment orbital-based surface hopping for charge carrier transport in organic and biological molecules and materials. *J. Chem. Phys.* **2016**, *145*, 064102.
(13) Rühle, V.; Lukyanov, A.; May, F.; Schrader, M.; Vehoff, T.; Kirkpatrick, J.; Baumeier, B.; Andrienko, D. Microscopic Simulations of Charge Transport in Disordered Organic Semiconductors. *J. Chem. Theory Comput.* **2011**, *7*, 3335−3345.
(14) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(15) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(16) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; van Duin, A. C. T. The ReaxFF reactive force-field: development, applications and future directions. *Npj Comput. Mater.* **2016**, *2*, 15011.
(17) Xu, P.; Guidez, E. B.; Bertoni, C.; Gordon, M. S. Perspective: Ab initio force field methods derived from quantum mechanics. *J. Chem. Phys.* **2018**, *148*, 090901.
(18) Rahman, A.; Stillinger, F. H. Molecular Dynamics Study of Liquid Water. *J. Chem. Phys.* **1971**, *55*, 3336−3359.
(19) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(20) Haughney, M.; Ferrario, M.; McDonald, I. R. Molecular-dynamics simulation of liquid methanol. *J. Phys. Chem.* **1987**, *91*, 4934−4940.
(21) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(22) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krger, P.; van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(23) Wang, H.; Yang, W. Determining polarizable force fields with electrostatic potentials from quantum mechanical linear response theory. *J. Chem. Phys.* **2016**, *144*, 224107.
(24) Nyman, J.; Pundyke, O. S.; Day, G. M. Accurate force fields and methods for modelling organic molecular crystals at finite temperatures. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15828−15837.
(25) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. Polarization effects in molecular mechanical force fields. *J. Phys.: Condens. Matter* **2009**, *21*, 333102.
(26) Mortier, W. J.; Van Genechten, K.; Gasteiger, J. Electronegativity equalization: application and parametrization. *J. Am. Chem. Soc.* **1985**, *107*, 829−835.
(27) Mortier, W. J.; Ghosh, S. K.; Shankar, S. Electronegativity-equalization method for the calculation of atomic charges in molecules. *J. Am. Chem. Soc.* **1986**, *108*, 4315−4320.
(28) York, D. M.; Yang, W. A chemical potential equalization method for molecular simulations. *J. Chem. Phys.* **1996**, *104*, 159−172.
(29) Chelli, R.; Procacci, P. A transferable polarizable electrostatic force field for molecular mechanics based on the chemical potential equalization principle. *J. Chem. Phys.* **2002**, *117*, 9175−9189.
(30) Rappe, A. K.; Goddard, W. A. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **1991**, *95*, 3358−3363.
(31) Chelli, R.; Procacci, P.; Righini, R.; Califano, S. Electrical response in chemical potential equalization schemes. *J. Chem. Phys.* **1999**, *111*, 8569−8575.
(32) Nistor, R. A.; Polihronov, J. G.; Müser, M. H.; Mosey, N. J. A generalization of the charge equilibration method for nonmetallic materials. *J. Chem. Phys.* **2006**, *125*, 094108.
(33) Lee Warren, G.; Davis, J. E.; Patel, S. Origin and control of superlinear polarizability scaling in chemical potential equalization methods. *J. Chem. Phys.* **2008**, *128*, 144110.
(34) Verstraelen, T.; Ayers, P. W.; Speybroeck, V. V.; Waroquier, M. ACKS2: Atom-condensed Kohn-Sham DFT approximated to second order. *J. Chem. Phys.* **2013**, *138*, 074108.
(35) Verstraelen, T.; Vandenbrande, S.; Ayers, P. W. Direct computation of parameters for accurate polarizable force fields. *J. Chem. Phys.* **2014**, *141*, 194114.
(36) Misquitta, A. J.; Stone, A. J. Distributed polarizabilities obtained using a constrained density-fitting algorithm. *J. Chem. Phys.* **2006**, *124*, 024111.
(37) Misquitta, A. J.; Stone, A. J. Accurate induction energies for small organic molecules: 1. Theory. *J. Chem. Theory Comput.* **2008**, *4*, 7−18.
(38) Islam, M. M.; Kolesov, G.; Verstraelen, T.; Kaxiras, E.; van Duin, A. C. T. eReaxFF: A Pseudoclassical Treatment of Explicit Electrons within Reactive Force Field Simulations. *J. Chem. Theory Comput.h* **2016**, *12*, 3463−3472.
(39) Baroni, S.; Giannozzi, P.; Testa, A. Green's-function approach to linear response in solids. *Phys. Rev. Lett.* **1987**, *58*, 1861.
(40) Gonze, X. Adiabatic density-functional perturbation theory. *Phys. Rev. A: At., Mol., Opt. Phys.* **1995**, *52*, 1096.
(41) Giese, T. J.; York, D. M. Density-functional expansion methods: Evaluation of LDA, GGA, and meta-GGA functionals and different integral approximations. *J. Chem. Phys.* **2010**, *133*, 244107.
(42) Giese, T. J.; York, D. M. Density-functional expansion methods: grand challenges. *Theor. Chem. Acc.* **2012**, *131*, 1145.
(43) Garrett, A. *inspyred: Bio-inspired Algorithms in Python*; 2017. http://aarongarrett.github.io/inspyred/ (accessed June 28, 2019).
(44) Anthony, J. E. The Larger Acenes: Versatile Organic Semiconductors. *Angew. Chem., Int. Ed.* **2008**, *47*, 452−483.
(45) Kaminski, S.; Giese, T. J.; Gaus, M.; York, D. M.; Elstner, M. Extended Polarization in Third-Order SCC-DFTB from Chemical-Potential Equalization. *J. Phys. Chem. A* **2012**, *116*, 9131−9141.
(46) Kaminski, S.; Gaus, M.; Elstner, M. Improved Electronic Properties from Third-Order SCC-DFTB with Cost Efficient Post-SCF Extensions. *J. Phys. Chem. A* **2012**, *116*, 11927−11937.
(47) Giese, T. J.; York, D. M. Improvement of semiempirical response properties with charge-dependent response density. *J. Chem. Phys.* **2005**, *123*, 164108.
(48) Baker, C. M.; Anisimov, V. M.; MacKerell, A. D., Jr Development of CHARMM polarizable force field for nucleic acid bases based on the classical Drude oscillator model. *J. Phys. Chem. B* **2011**, *115*, 580−596.
(49) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175−2196.

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

(50) Sun, J.; Ruzsinszky, A.; Perdew, J. P. Strongly Constrained and Appropriately Normed Semilocal Density Functional. *Phys. Rev. Lett.* **2015**, *115*, 036402.

DOI: 10.1021/acs.jctc.9b00415
J. Chem. Theory Comput. 2019, 15, 4516−4525

E.2. Publication #2

An iterative fragment-scheme for the ACKS2 electronic polarization model: Application to molecular dimers and chains.

P. Gütlein, J. Blumberger and H. Oberhofer
J. Chem. Theory Comput. (accepted)
DOI 10.1021/acs.jctc.0c00151

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

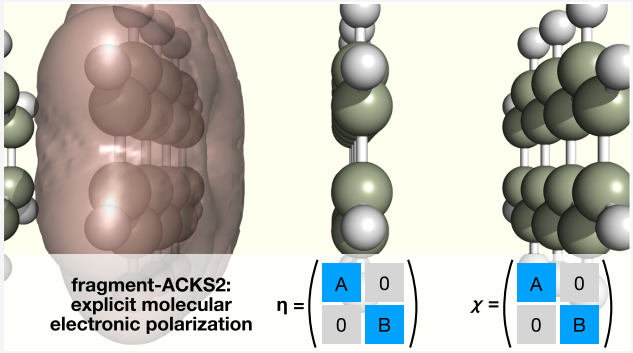
An Iterative Fragment Scheme for the ACKS2 Electronic Polarization Model: Application to Molecular Dimers and Chains

Patrick Gütlein, Jochen Blumberger, and Harald Oberhofer*

Cite This: https://dx.doi.org/10.1021/acs.jctc.0c00151 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The treatment of electrostatic interactions is a key ingredient in the force field-based simulation of condensed phase systems. Most approaches used fixed, site-specific point charges. Yet, it is now clear that many applications of force fields (FFs) demand more sophisticated treatments, prompting the implementation of charge equilibration methods in polarizable FFs to allow the redistribution of charge within the system. One approach allowing both, charge redistribution and site-specific polarization, while at the same time solving methodological shortcomings of earlier methods, is the first-principles-derived atom-condensed Kohn−Sham density functional theory method approximated to the second order (ACKS2). In this work, we present two fragment approaches to ACKS2, termed f-ACKS2 and a self-consistent version, scf-ACKS2, that treat condensed phase systems as a collection of electronically polarizable molecular fragments. The fragmentation approach to ACKS2 not only leads to a more transferable and less system-specific collection of electronic response parameters but also opens up the method to large condensed phase systems. We validate the accuracies of f-ACKS2 and scf-ACKS2 by comparing polarization energies and induced dipole moments for a number of charged hydrocarbon dimers against DFT reference calculations. Finally, we also apply both fragmented ACKS2 variants to calculate the polarization energy for electron−hole pair separation along a chain of anthracene molecules and find excellent agreement with reference DFT calculations.



1. INTRODUCTION

Nowadays, atomistic or coarse-grained force fields are an essential tool in the materials' modeling community, enabling an efficient sampling of the phase space of condensed phase systems.[1−4] Whether based on training by ab initio references or experimental properties, or, recently, machine learned models, force fields allow molecular simulations of systems that are too large for treatment with explicit electronic structure methods. They have been used very successfully for the simulation of biological systems,[5−7] catalysis,[8,9] and energy conversion materials like batteries[10,11] or organic semiconductors.[12,13]

In all cases, the accuracy of force field approaches rests critically on their ability to represent the different types of interactions between atoms or molecules.[4,14,15] One key ingredient is the accurate description of electrostatic interactions, especially in systems where small changes in the electrostatic potential strongly influence local properties, as, e.g., in protein folding[16] or excitation energy and charge transport.[17,18] Therefore, earlier molecular mechanics simulations frequently employed an inexpensive but crude classical point charge representation of the mean-field electrostatic interactions.[19,20] In recent years, though, the static, isotropic nature and lack of directionality of the simple point charge model were highlighted,[21−24] prompting the development of improved electrostatic models by either adding off-center point charges or higher angular momentum terms like dipole and quadrupole moments.[25,26] Going further, incorporating the electronic many-body polarization induced by local electrostatic potential changes, e.g., due to dynamical fluctuations of other molecules, the presence of excess charge carriers, or external voltages, in a classical interatomic model framework is a difficult task.[21,27] Indeed, polarization corrections so far exhibited mixed successes compared to well-tuned non-polarizable force field parametrizations.[28−31] Popular empirical models for the explicit polarization in force fields are the Drude oscillator (or core−shell)[32,33] and the atomic inducible dipole (or distributed polarizabilities) approach,[34,35] which employ a charge attached to the atomic nuclei by a harmonic spring or atom-centered inducible point dipoles, respectively, to capture intra-atomic charge rearrangements. While they are very efficient and easy to implement, they exclude specific atom−

Received: February 14, 2020
Published: July 23, 2020

atom charge transfer terms and are difficult to parametrize thoroughly, raising the problem of system transferability.[36,37] On the contrary, the X-Pol[38] (formerly MODEL[39,40]) method combines a simplified semi-empirical quantum mechanical description of a molecule, including its polarization, with the efficient force field representation of the intermolecular interaction potential, to approximately maintain the non-classical nature of electrons. Along similar lines, force field techniques have been combined with quantum chemical calculations of molecular fragments to obtain atomic partial charges and molecular fragment dipole moments on the fly (i.e., every few picoseconds for protein simulations).[41,42] While these approaches show promising accuracy and transferability, e.g., protein solvation in water,[43,44] they suffer from a much increased computational cost compared to standard force fields.[45] Finally, in an attempt to capture charge reorganization and polarization yet at the cost of standard force fields, a population of fluctuating atomic charges, initially described by the electronegativity equilibration method (EEM),[46] has become a widespread choice for the description of electronic polarization. In the three and a half decades since EEM's inception, it has undergone many refinements in the form of the charge equilibration (QEq),[47] chemical potential equilibration (CPE),[48] atom−atom charge transfer (AACT),[49] and split-charge equilibration (SQE)[50] schemes. However, many of these models suffer from overestimation of the long-range charge transfer due to incorrect kinetic energy contributions to the electronic many-body response. A recently suggested method for the calculation of electronic polarization, the atom-condensed Kohn−Sham density functional theory approximated to the second order (ACKS2),[51,52] could alleviate these methodological drawbacks.

ACKS2 is a linear electronic response technique derived from first principles, based on an atom-projected description of the perturbative Kohn−Sham (KS) density functional theory (DFT) electron density rearrangements induced by an external potential. In a recent study, we developed a transferable Cartesian Gaussian basis set representation as the first step toward applying ACKS2 as a general electronic polarization contribution in force fields.[53] We found that ACKS2-calculated induced dipoles and polarization energies of organic molecules due to simple external electrostatic perturbations were in excellent agreement with the results of full DFT calculations. This suggests that ACKS2 could be a promising method for the description of electronic polarization between molecules in the condensed phase. However, so far, ACKS2 has only been used to calculate the polarization response to a static external electric field, not between two mutually polarizable entities like molecules. Therefore, in order to make ACKS2 amenable to force field simulations, it is necessary to develop an approach that describes the mutual polarization of molecules within the ACKS2 framework.

In this study, we introduce a fragment approach to ACKS2 and treat the simulation system as an assembly of polarizable molecular subunits. In this approach, the intramolecular and intermolecular polarization energies are accounted for at the ACKS2 level under the constraint that the total charge of each fragment is equal to the charge of the unperturbed fragment. Hence, similar to other polarization models, the method accounts for intramolecular but not intermolecular charge redistribution. In section 2, we briefly summarize the main concepts of ACKS2 before we describe in detail our new ACKS2 fragmentation approach. The latter is validated against the perturbative KS-DFT parent method in section 3 on a set of neutral and charged organic dimers and illustrated further for a chain of organic molecules carrying an excess electron and electron hole. Finally, concluding remarks are given in section 4.

2. THEORY

2.1. Atom-Condensed Kohn−Sham DFT Approximated to the 2nd Order. The atom-condensed Kohn−Sham density functional theory approximated to the second order,[51,52] a recently developed extension to more traditional charge equilibration schemes,[48] was demonstrated to be an accurate yet transferable first-principles-based method for the calculation of response properties, like the electronic polarization energies and induced dipole moments.[53] It describes electronic polarization in the presence of an external potential using a simplified perturbative Kohn−Sham density functional theory approach within the limit of linear response. The first-principles nature of this method has a number of distinct benefits compared to empirical polarization models in that all parameters of the ACKS2 model have a direct physical meaning and can be derived exactly from DFT calculations.[51,53] In the interest of brevity, here, we only give a brief overview of the method, necessary for the derivation of the fragment scheme. For a more detailed description of the method, we refer the reader to references[51,52] or our previous work.[53]

In ACKS2, the response of the electron density, $\Delta\rho$, and the Kohn−Sham potential, Δv_{KS}, to an external potential, Δv_{ext}, is expanded in terms of atom-centered basis functions g_n (h_m) with expansion coefficients c_n (d_m)

$$\Delta\rho(\mathbf{r}) = \sum_{n}^{N} c_n\, g_n(\mathbf{r}) \tag{1a}$$

$$\Delta v_{KS}(\mathbf{r}) = \sum_{m}^{M} d_m h_m(\mathbf{r}) \tag{1b}$$

This atom-condensed representation of the electronic structure response is in principle exact in the limit of a complete basis set. Given the atom-centered basis set expansions of eq 1, the parametrized matrix form of the KS-DFT linear response equation, i.e., the working equation of ACKS2, reads

$$\begin{bmatrix} [\eta_{n,n'}]_{N,N} & -[O_{n,m}]_{N,M} & [D_{n,1}]_N \\ -[O_{m,n}]_{M,N} & [\chi_{m,m'}]_{M,M} & 0 \\ [D_{1,n}]_N & 0 & 0 \end{bmatrix} \begin{bmatrix} [c_{n,1}]_N \\ [d_{m,1}]_N \\ \Delta\mu \end{bmatrix} = \begin{bmatrix} [-V_{n,1}]_N \\ [0_{m,1}]_M \\ 0 \end{bmatrix} \tag{2a}$$

$$\mathbf{Px} = -\mathbf{V} \tag{2b}$$

The right hand side vector with elements $[\mathbf{V}]_n = \int d\mathbf{r} g_n(\mathbf{r}) \Delta v_{ext(\mathbf{r})}$ represents the potential Δv_{ext} of an external perturbation acting on the system, such as an external field. The solution vector $\mathbf{x} = \{c_n, d_m\}$, collects the expansion series coefficients for the change in electron density and KS potential. The matrix $\mathbf{P}$ on the left hand side of eq 3 encodes the reduced KS-DFT ground-state response information, condensed according to the choice of basis set in eq 1. It contains four different parameters, $\{\eta_{i,j}, \chi_{i,j}, O_{i,j}, D_i\}$, which are all well-

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

defined expectation values of the KS-DFT ground-state electronic structure.

Two parameters, $\eta_{i,j}$ and $\chi_{i,j}$, require KS-DFT ground-state orbitals and energies for their calculation

$$\eta_{i,j} = \iint g_i(\mathbf{r})\left(\frac{1}{|\mathbf{r}-\mathbf{r}'|} + \frac{\partial^2 E^{\mathrm{xc}[\rho]}}{\partial\rho(\mathbf{r})\partial\rho(\mathbf{r}')}\right) g_j(\mathbf{r}')\,\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \tag{3}$$

$$\chi_{i,j} = \iint h_i(\mathbf{r})\left(\frac{\partial^2 E^{\mathrm{KS}}[v_{\mathrm{KS}}]}{\partial v_{\mathrm{KS}}(\mathbf{r})\partial v_{\mathrm{KS}}(\mathbf{r}')}\right) h_j(\mathbf{r}')\,\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \tag{4}$$

The hardness kernel $\eta_{i,j}$ accounts for the energetic cost of electronic polarization due to changes in the electron density $\Delta\rho$. The non-interacting linear response kernel $\chi_{i,j}$ accounts for the energetic cost of electronic polarization due to changes in the Kohn−Sham potential Δv_{KS}. The remaining two ACKS2 parameters, $O_{i,j}$ and D_i, are independent of the KS electronic structure and depend only on the choice of the basis functions in eq 1. $O_{i,j}$ is the overlap integral of basis functions for density and KS potential response: $O_{i,j} = \int g_i(\mathbf{r}) h_j(\mathbf{r})\,\mathrm{d}\mathbf{r}$, accounting for the energetic contribution of concerted changes of electron density and KS potential. $D_i = \int g_i(\mathbf{r})\,\mathrm{d}\mathbf{r}$ and ensures the conservation of total charge.

With these parameters, the total ACKS2 polarization energy of the system can be expressed as[53]

$$\Delta E^{\mathrm{pol.}} = \frac{1}{2}\sum_{n,n'}^{N,N} c_n \eta_{n,n'} c_{n'} + \frac{1}{2}\sum_{m,m'}^{M,M} d_m \chi_{m,m'} d_{m'} - \sum_{n,m}^{N,M} c_n O_{n,m} d_m + \sum_{n}^{N} c_n D_n \tag{5}$$

2.2. Fragment-ACKS2. We here introduce a fragment approach to ACKS2, termed fragment-ACKS2 or f-ACKS2. The key idea here is to divide a large condensed phase system into smaller polarizable subunits or fragments, typically molecules, and account for the total electronic polarization in a simulation cell by means of the individual fragment responses. Specifically, in this study, the f-ACKS2 technique is applied to evaluate the dielectric response contributions to the intermolecular interactions in dense phase media. Therefore, as illustrated for a benzene molecular dimer in Figure 1, fragments are chosen to represent the entire simulation cell and consecutively parametrized in a vacuum reference framework. This means that the f-ACKS2 matrices for each fragment are calculated individually from the KS-DFT electronic ground state of an isolated molecule in vacuum. Additionally, an effective interaction potential of the isolated ground-state charge density of each fragment is developed, which is used as external perturbation to simulate the effect of neighboring molecules in dense phase media. Solution of the f-ACKS2 equations then yields the electronic polarization of a vacuum-level parametrized fragment embedded in a dense phase environment represented by vacuum-level charge potentials. This way, the f-ACKS2 approach follows the spirit of force field techniques to treat contributions to the total energy on different footings based on, e.g., chemical connectivity, bonded or non-bonded, or distance measures, long-range vs. short-range. It provides a computationally efficient tool to account for the electronic polarization between fragments, like molecules in the condensed phase, and hence is ideally suited for use in molecular mechanics simulations. In

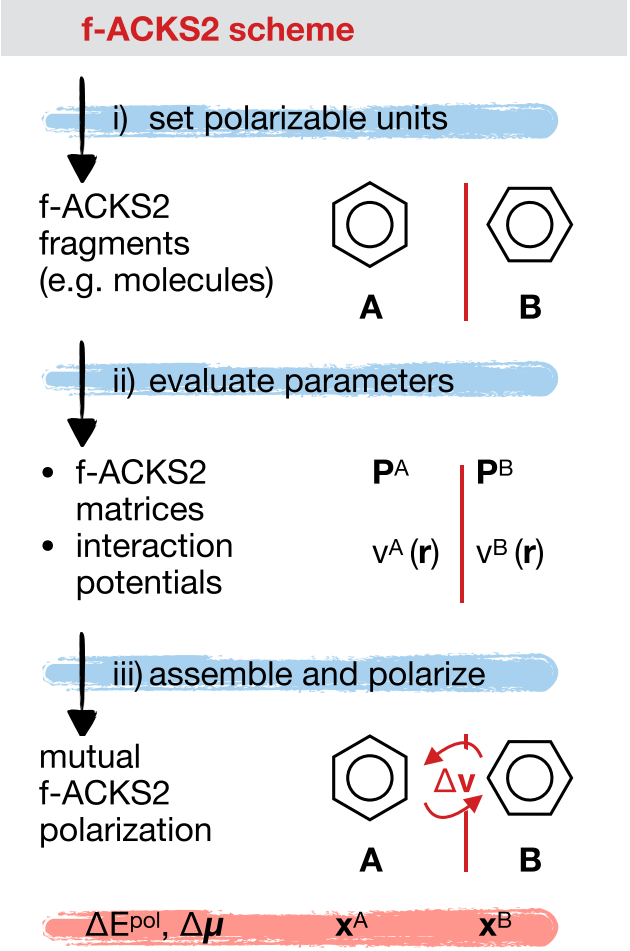


Figure 1. Illustration of the methodological steps of the f-ACKS2 approach for the mutual polarization of two benzene molecules.

the following, the ACKS2 foundations of the three practical steps, (i) fragmentation, (ii) parametrization, and (iii) mutual polarization, depicted in Figure 1 are assessed, starting from the viewpoint of an ACKS2 matrix formulation of the entire simulation cell.

In a nutshell, the system cell fragmentation step is justified by theoretical arguments to decouple the response between different fragments (sparsification), to change the ACKS2 framework from an embedded molecule to a vacuum-level reference electronic structure and to set a suitable charge constraint. The parametrization step is concerned with the calculation of the f-ACKS2 matrix elements and an efficient yet accurate interfragment interaction potential. The fragment polarization step involves the solution of the f-ACKS2 matrix equations and numerical implementations to allow possible coupling between different fragments, as well as correct evaluation of the response properties.

In step (i), the partitioning of a simulation system in κ polarizable fragments is a new concept to the ACKS2 theory, inspired by other constraint and fragmentation techniques like fragment-orbital DFT.[54] Similar schemes have been reported as ad hoc corrections in charge equilibration schemes, where a molecule has been partitioned topologically in small subunits to avoid the erroneous superlinear scaling of molecular polarizabilities with increasing size.[49,55]

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

The choice of fragments is in principle arbitrary, e.g., following atom proximity, chemical connectivity, or functional grouping. The use of molecules as polarizable subunits arises naturally in a molecular condensed phase system like organic crystals or solvated proteins.

The f-ACKS2 technique represents a sparsification approach to the total electronic response, which in the present study contains three conceptual approximations (fragmentation of the ACKS2 equations and methodological implications are derived in detail in the Supporting Information). First, all ACKS2 matrix elements ($O_{i,j}$, $\eta_{i,j}$, and $\chi_{i,j}$) related to basis functions located on different fragments are set to zero, which prohibits the coupling of the electronic response between different fragments. Note, any intrafragment matrix elements are unchanged. The overlap matrix elements $O_{i,j}$ (between density basis functions and KS potential functions) generally depend on the radial decay of each function and the distance between them, which leads to small overlaps for well-separated molecular fragments. Following the theorem of Unsöld,[56] ACKS2 matrix elements of the non-interacting linear response kernel $\chi_{i,j}$ related to different molecular fragments, i.e., at a large distance and small electronic overlap between different atomic sites, are approximately zero.[51] Similarly, the exchange-correlation contributions to the hardness kernel $\eta_{i,j}$ for semi-local functionals such as PBE[57] rapidly decay with increasing interatomic distances, yielding negligible intermolecular contributions to the ACKS2 matrix. Only the classical electrostatic interaction contribution to the hardness kernel is long-range in nature and goes well-beyond typical intermolecular distances in dense matter. However, the electronic response in dielectric media (like an organic molecule) to an external potential and especially its induced change of the potential at other fragments are reasonably small, and setting these contributions to zero is a valid first approximation (depending on the system polarizability). The fragmentation (sparsification) and its practical impact on the ACKS2 (intermolecular vs. intramolecular) matrix elements are illustrated for a molecular dimer of benzene, see section 1. As model refinement to f-ACKS2, mutual fragment response interactions due to classical electrostatic contributions are implemented in an iterative, self-consistent field approach introduced in section 2.3.

The second approximation in our fragmentation approach concerns the evaluation of the ACKS2 parameters for the individual fragments. Executing and postprocessing a DFT calculation for the entire simulation cell at every molecular dynamics or Monte Carlo simulation step (and then setting all interfragment matrix elements to zero anyways) are computationally prohibitive in force field methods. One route to circumvent this is to sample a reasonable number of relevant structures of the molecule embedded in the dense phase environment and extract an estimate of their exactly calculated ACKS2 parameters (e.g., via averaging, interpolation, or even machine learning). In this study, we opt for an even more transferable approach, where the parameters of each fragment are derived from a reference DFT calculation of a single molecule in vacuum. Other fragments otherwise present in dense matter are introduced by an approximate Coulomb response interaction potential (like a simple point charge model), very much in the spirit of other force field techniques. Furthermore, a vacuum reference parametrized f-ACKS2 model allows the evaluation of electronic polarization contributions to the intermolecular interactions, which will be studied for a set of small (partially charged) aromatics below.

The third, and last, approximation concerns the constraint of the total charge due to electronic rearrangements present in the ACKS2 method, c.f. the last row in eq 2, which necessitates a translation to fragment contributions in the f-ACKS2. Practically, this yields two distinct choices. In a straightforward approach, a constraint is imposed on the overall polarization charge, i.e., summation over all fragment charges is constant, which allows charge transfer between different sites. Thereby, charge equilibration is achieved in the entire system, and the responses of the individual fragments are effectively coupled by charge transfer energetics (any implementation requires a numerical solution of a coupled set of linear equations). Alternatively, a charge constraint is imposed for each individual fragment in f-ACKS2, preventing charge transfer between different sites. This is the approach taken in this work. While it is expected to be a good approximation for the (adiabatic) ground state of non-polar or weakly polar systems, it is the natural choice for the modeling of the charge localized (or (quasi)-diabatic) electronic states of electron transfer reactions[58] and charge transport in the condensed phase.[59−61]

In step (ii), the parametrization process of the f-ACKS2 response matrices follows the standard ACKS2 model for each of the κ fragments in vacuum. The perturbation of a given fragment A by the other fragments B is modeled by a sum of fragment potentials v^A_{frag}

$$\Delta v^A_{\mathrm{ext}}(\mathbf{r}) = \sum_{B\neq A}^{\kappa} v^B_{\mathrm{frag}}(\mathbf{r}) \tag{6}$$

The idea of an effective potential representation v^B_{frag}here is a general concept and can be tuned to specific numerical needs and applications. In order to avoid costly real space integrals in the determination of the perturbation vector $\mathbf{V}$, we chose a discrete representation based on fixed atom-centered point charges (fc), i.e., $v^B_{\mathrm{frag}} = v^B_{\mathrm{fc}}$. The latter represents a simple Coulomb interaction term but neglects higher fixed multipole terms or exchange-correlation contributions between the fragments. While it would also be possible to introduce, e.g., atom-centered multipoles, to describe fragment−fragment interactions, our benchmark results presented in section 3 already show that such a simple representation is able to capture most of the induced polarization effects.

In step (iii), computing the (non-self-consistent) response for each individual fragment again follows standard ACKS2 methodology.

$$[\mathbf{V}^A]_n = \sum_{B\neq A}^{\kappa} \int g_n^A(\mathbf{r}) v^B_{\mathrm{fc}}(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{7a}$$

$$\mathbf{P}^A \mathbf{x}^A = -\mathbf{V}^A \tag{7b}$$

The solution vector $\mathbf{x}^A$ contains the expansion coefficients for electron density and Kohn−Sham potential changes, $c^A_n$and d^A_m, of basis functions located at fragment A, c.f. eqs 1a and 1b. Observables like polarization energy and dipole moment of the individual fragments follow straightforwardly, see ref 53. The overall response is finally obtained by simple superposition of the individual fragment electronic polarization

$$\Delta\mu = \sum_{A}^{\kappa} \Delta\mu_A \tag{8a}$$

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$\Delta E^{\text{pol.}} = \sum_{A}^{\kappa} \Delta E_A^{\text{pol.}} \tag{8b}$$

2.3. Self-Consistent Fragment-ACKS2. With the introduction of a fixed atom-centered point charge representation v_{fc}^B, we provide an effective interaction potential between f-ACKS2 fragments based on a simplified electronic ground-state charge distribution. However, the influence of mutual electronic polarization interactions between different fragments has not yet been accounted for. Illustrated in Figure 2, the presence of fragment A and its effective ground-state KS potential polarizes fragment B, which in turn induces local potential changes and a response of fragment A, in turn changing the potential on fragment A and so on. In total, this yields a contribution v_{resp}^B to the local potential of a fragment due to the polarization of all other fragments in the system

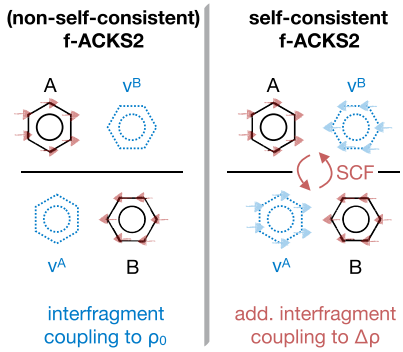


Figure 2. Illustration of the effective interaction treatment in f-ACKS2 and scf-ACKS2.

$$\Delta v_{\text{ext}}^A = \sum_{B\neq A}^{\kappa} v_{\text{fc}}^B(\mathbf{r}) + \sum_{B\neq A}^{\kappa} v_{\text{resp}}^B(\mathbf{r}) \tag{9}$$

A straightforward way to introduce coupling of the mutual polarization response interactions is the addition of a Coulomb potential due to the response density,

$$v_{\text{resp}}^B = \sum_{n'}^{N'} c_{n'} \int \frac{1}{|\mathbf{r}-\mathbf{r}'|} g_{n'}(\mathbf{r}')\mathrm{d}\mathbf{r}'$$

$$\Delta v_{\text{ext}}^A = \sum_{B\neq A}^{\kappa} v_{\text{fc}}^B(\mathbf{r}) + \sum_{B\neq A}^{\kappa} \sum_{n'}^{N'} c_{n'}^B \int \frac{1}{|\mathbf{r}-\mathbf{r}'|} g_{n'}^B(\mathbf{r}')\mathrm{d}\mathbf{r}' \tag{10}$$

Note that the mutual response interaction term shown here for fragment A includes the f-ACKS2 expansion coefficients $c_{n'}^B$ of the other fragments B. Thus, the new working equations cannot simply be solved by a single matrix inversion of κ response-independent fragments, as in f-ACKS2, because now, all fragment perturbations include the polarization of each other, and the individual fragment coefficients are interdependent. The interdependence of the (atomic) inducible multipole moments leads to a large yet sparse ACKS2 set of linear equations for the total system (all fragments). The number of equations scales with the number of basis functions and would introduce a considerable bottleneck in calculations on large clusters/condensed phase systems. The blocks on the diagonal represent individual fragments (here parametrized from molecules in vacuum). All interfragment off-diagonal elements introduce an approximate coupling, see second term in Eq 9, which is long-range in nature due to the $1r$ and $1r^3$ decay of the electrostatic potential of (point) charges and dipoles, respectively, and somewhat reduces the actual sparsity of the ACKS2 equations. The difference between a fully fragmented f-ACKS2 and the coupled working equations is illustrated for a two-fragment system in Figure 3.

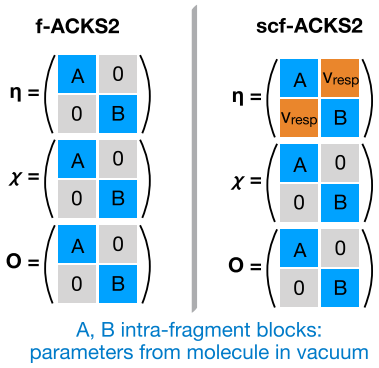


Figure 3. Illustration of the different parameter elements in a direct matrix approach to f-ACKS2 excluding and including the response coupling v_{resp} (as a function of the charge rearrangements $\Delta\rho$ only) between different fragments for a benzene molecular dimer.

An exact solution of the sparse polarization matrix is computationally very expensive, and instead, it is often approximated to arbitrary accuracy following two general concepts (or combinations thereof). In iterative solver techniques, the simulation cell is divided into smaller subunits (with the smallest being an individual site) for which the individual polarization linear equations are solved, similar to the f-ACKS2 approach. The response of the individual sites to an initially estimated charge distribution updates the local electrostatic potential, prompting a new response and so forth until self-consistency is reached.[62,63] Alternatively, in time-dependent molecular dynamics simulations, the explicit polarization degrees of freedom can be propagated directly from an initial solution by Lagrangian methods, which introduces an additional thermostat for the response.[64,65] Here, we employ an iterative self-consistency scheme, termed scf-ACKS2, where the response of each fragment is fed stepwise into the external potential of other fragments until self-consistency is reached

$$\mathbf{V}^{A,\{i\}} = \mathbf{V}_{\text{fc}}^A + \mathbf{V}_{\text{response}}^A(\mathbf{x}^{B,\{i-1\}}) \tag{11a}$$

$$[V_{\text{fc}}^A]_n = \sum_{B\neq A}^{\kappa} \int g_n^A(\mathbf{r}) v_{\text{fc}}^B(\mathbf{r})\mathrm{d}\mathbf{r} \tag{11b}$$

$$[V_{\text{response}}^A]_n(\mathbf{x}^{B,\{i-1\}}) = \sum_{B\neq A}^{\kappa} \sum_{n'}^{N'} c_{n'}^{B,\{i-1\}} \int g_n^A(\mathbf{r}) \frac{1}{|\mathbf{r}-\mathbf{r}'|} g_{n'}^{B'}(\mathbf{r}')\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \tag{11c}$$

The additional superscript $\{i\}$ indicates the step number within the iterative scf-ACKS2 scheme. In the scf-ACKS2 scheme, all interaction integral vectors in $\mathbf{V}_{\text{fc}}$ and $\mathbf{V}_{\text{response}}$ are constant and can be stored in memory, whereas the expansion prefactors $\mathbf{x}^A$ ($\{c_n^A\}$ and $\{d_m^A\}$) change with each iteration step,

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

accounting for the mutual polarization response interaction. Following the update of the perturbation vectors, solution of the ACKS2 matrix equation for each fragment gives the individual expansion coefficient vectors, $\mathbf{P}^A\mathbf{x}^{A,\{i\}} = \mathbf{V}^{A,\{i\}}\ \forall\ A \in \kappa$. Note, the solutions to the ACKS2 matrix equation of each fragment are interdependent, as defined in eq 11c. Efficient implementation strategies of the iterative solvers, including initial guesses to the solution, preconditioning, and propagation algorithms, e.g., variants of Jacobi and conjugate gradient methods, to compute reliable explicit polarization properties at reasonable computational cost is a matter of ongoing research.[62,63,66] Focusing on a proof of principle in the present study, we include an intermediate linear mixing step for an update of the ACKS2 density and KS potential expansion coefficients $\mathbf{x}^A$ mediated by a constant mixing parameter $\tau_{\text{mix}} < 1$

$$\mathbf{x}^{A,\{i\}} \rightarrow \tau_{\text{mix}}\mathbf{x}^{A,\{i\}} + (1 - \tau_{\text{mix}})\mathbf{x}^{A,\{i-1\}} \tag{12}$$

The iteration of the mutual electronic polarization interaction between different fragments is continued until self-consistency is reached, ensured by some termination criterion like the maximum change of coefficients to be smaller than some threshold,

$$(\max_n(\{|x_n^{A,\{i\}} - x_n^{A,\{i-1\}}|\}) < \tau_{\text{threshold}})\ \forall\ A.$$

The scf-ACKS2 scheme is initiated by solution of the f-ACKS2 equations for each fragment where the external perturbation includes the fixed-charge distribution interaction potential v_{fc}^B. That is, the expansion coefficients in the potential (eq 10) are set to zero for the first step of the iterative scheme, $c_n^{B,\{0\}} = 0$. Hence, the initial polarization of each fragment contains the electronic response due to the interactions with the fixed point charge distribution of all other fragments. In principle, SCF convergence can in some cases further be improved by using a different mixing constant, $\tau_{\text{init}} \neq \tau_{\text{mix}}$, for the first few steps. In practice, though, for most of the systems considered here, this proved unnecessary, cf. the Supporting Information.

The electronic response of the total system is obtained by summing up the polarization energies of the individual fragments but corrected for the double counting of the mutual response polarization interaction contributions, where $\mathbf{x}$ represents the self-consistent solution in the scf-ACKS2 scheme

$$\Delta E_{\text{scf-ACKS2}}^{\text{pol.}} = \sum_A^{\kappa} \Delta E_A^{\text{pol.}} - \frac{1}{2}\sum_A^{\kappa}\sum_{B\neq A}^{\kappa} \mathbf{x}^A\mathbf{V}_{\text{response}}^A(\mathbf{x}^B) \tag{13}$$

3. RESULTS

3.1. Parametrization of f-ACKS2. First, we highlight the impact of the above discussed approximate treatment of the f-ACKS2 parameters for a molecular dimer of benzene taken from a single crystal (Pbca),[67] where each molecule represents one fragment. In Figure 4, the distributions and distance dependence of the various matrix elements of the total dimer are color-coded to illustrate the relative importance of the individual contributions. Panel A displays the large difference of the exchange-correlation contributions to the hardness, η^{xc}, for a semi-local functional (PBE) and fast decay with interatomic distance. In Panel B, the Coulomb hardness

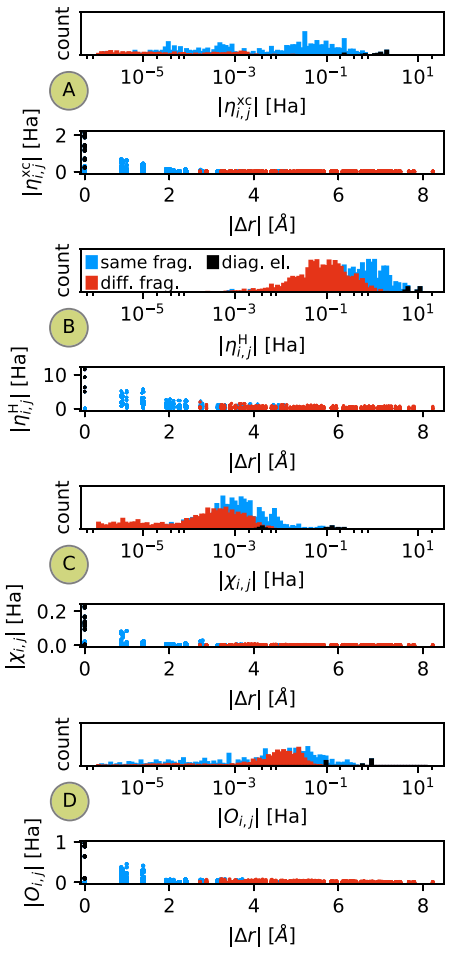


Figure 4. Comparison of interfragment and intrafragment ACKS2 matrix parameters (defined in section 2.1) and their dependence on the distances between the center of the basis functions for the nearest neighbor benzene molecular dimer of a single crystal.[67]

distributions, η^{H}, of inter- and intrafragments are of a similar shape but slightly shifted by about half an order of magnitude. The slow relative decay of these matrix elements reflects the (non-negligible) long-range nature of the classical electrostatic interaction contribution of the response. Panel C demonstrates a similar behavior of the distributions of the non-interacting linear response kernel, χ, with respect to intra- and interfragment contributions. However, there is an extended tail in the intrafragment distribution about two orders of magnitudes larger than the interfragment counterpart, which can be expected to dominate the response contributions. The same is reflected in the rather rapid relative decay of the matrix elements with respect to interatomic distances. A similar but less pronounced effect is visible in the distribution and distance plot of the overlap matrix elements O. In Figure 5, the difference of same-fragment ACKS2 matrix elements parametrized from the DFT electronic structure of a molecule embedded in the simulation cell versus the molecule in vacuum is illustrated. While both f-ACKS2 matrix elements exhibit broad distributions, they are at least one order of magnitude smaller than the largest (dominating) values displayed in Figure 5. Note that the slightly different distributions for the matrix elements of fragment A and fragment B are a consequence of the effective overlap of basis functions due to their relative orientation.

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

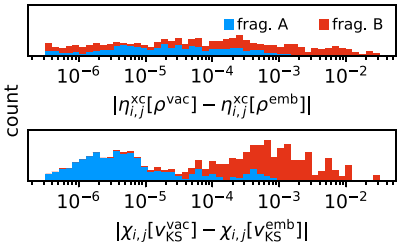


Figure 5. Comparison of intrafragment f-ACKS2 matrix parameters calculated from the molecule embedded in a dimer (emb) or isolated in vacuum (vac) for a benzene molecular dimer.

3.2. Electronic Polarization in Molecular Dimers A-B$^+$.

As a first demonstration, we apply the f-ACKS2 model for the electronic polarization between two charge-constrained molecules. Illustratively, we consider dimers of fairly small organic molecules such as benzene (Pbca),[67] naphthalene ($P2_1/a$),[68] anthracene ($P2_1/a$),[69] and tetracene ($P\bar{1}$),[70] where the first molecule (fragment A) is charge neutral, while the second molecule (fragment B) carries a charge of +1.0e. The dimer geometries are extracted from their respective single crystal structures, taking into account two or three different next-nearest-neighbor site dimers. With this, we aim for a realistic test case of the electronic polarization of a molecule by an electron−hole constrained to another molecule (and vice versa), as, e.g., encountered during charge transport in organic semiconductors.

In the first step, we account for fixed-charge distribution interactions between the fragments and exclude any mutual polarization response interactions. The approximate fixed-charged interaction is represented by atom-centered point charges derived from an electrostatic potential fit of the KS-DFT ground-state charge distribution, c.f. fixed point charge potential in eq 6. The respective data are labeled ACKS2(PC) in Figure 6 and thereafter. The full electronic response reference data to validate the ACKS2(PC) simulation setup are obtained from a perturbed KS-DFT calculation of the charge-constrained molecules. Therein, the approximate intermolecular interaction potential comes in two variants: a full Coulomb potential due to electron density and nuclear cores (DFT(FC)) and an atomic point charge approximation of the latter (DFT(PC)) via electrostatic potential fitting. The former is based on a fragment-orbital (FO) DFT implementation[54] in the FHI-AIMS code,[71] which provides a tool to get a self-consistent KS-DFT calculation of one molecule in the presence of the exact Coulomb potential of the self-consistent charge distribution of another molecule. Further methodological and numerical details of the DFT validation setup and properties are given in the Supporting Information. A comparison between the different DFT models, DFT(PC) and DFT(FC), allows us to gauge the error made by approximating the full Coulomb electrostatic interaction potential by point charges. On the other hand, the difference between DFT(PC) and ACKS2(PC) is the true test for the accuracy and validity of the ACKS2 electronic polarization model with respect to the DFT parent method.

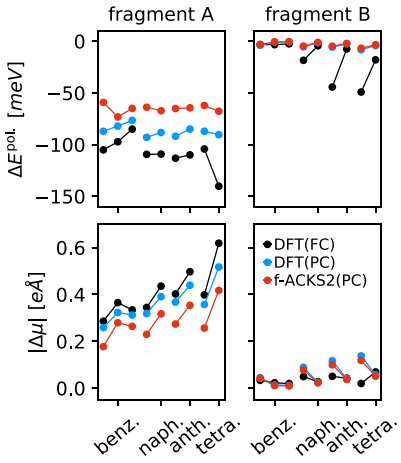


Figure 6. Electronic response for different approximations within the f-ACKS2 approach for a set of small molecular dimers (benzene, naphthalene, anthracene, and tetracene), where each data point represents a different nearest neighbor geometry extracted from the molecular single crystal structure.

Figure 6 illustrates the polarization energies and induced dipole moments for each fragment and provides comparison of the f-ACKS2 response model to the DFT data. Note that the response of fragment B to the non-charged non-polar fragment A is very small and hence more prone to small numerical inaccuracies due to the f-ACKS2 approximations. Assessment of the polarization of fragment A by the positively charged molecule B is less subject to numerical issues and hence of greater significance here. Generally, the trends of the polarization energy and the dipole moment agree well for fragment A despite the different levels of interaction potentials and electronic polarization response representations applied here. Following the individual trends, we can decompose the influence of the different levels of electronic response treatment for these molecular dimer systems. DFT(PC) introduces a point charge simplification to the full Coulomb intermolecular interaction potential in DFT(FC). Hence, the differences between the DFT(FC) and DFT(PC) trends, black and blue in Figure 6, represent an estimate of the validity of the first-order multipole analysis fragment interaction representation of the full Coulomb potential. For the small organic molecule dimers in Figure 6, the mean unsigned relative errors for the response, polarization energy and induced dipole moment, of fragment A by the DFT(PC) model are 19 and 10%. Extension of the multipole expansion series by including higher angular momentum terms like dipole moments and quadrupole moments could provide a simple conceptual approach to improve the interaction potential for even higher accuracies.[72] The ACKS2(PC) model introduces an approximate electronic response representation to the full KS-DFT polarization DFT(PC), both considering a point charge intermolecular interaction potential. The trends of the polarization properties from both methods agree well with a slight offset due to underestimation by the ACKS2 approach, yielding relative mean unsigned errors of 24 and 22% for the polarization energies and induced dipole moments, respectively. Introduction of a larger or more well-trained basis set provides a pathway to improve the current ACKS2 model. Considering fragment B, the ACKS2(PC) response matches the DFT(PC) polarization quite well, while both partly show significant deviation from DFT(FC). Indeed, condensing the Coulomb interaction potential into atom-centered point charge representations is quite inaccurate for large non-polar molecules and can yield tremendous differences in the local perturbation potentials.[73,74]

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

So far, we disregarded the mutual interaction between the response densities on each fragment; however, it can be crucial in a non-polar system with strong electrostatic fields such as a (partly) localized charge on a neighboring fragment. Figure 7

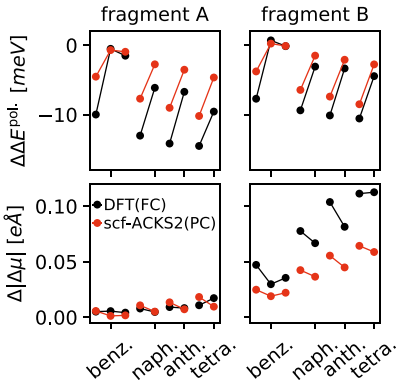


Figure 7. Change of the response by addition of the mutual polarization response interaction potential, $\Delta\Delta E^{pol.} = \Delta E^{pol.}_{scf-ACKS2} - \Delta E^{pol.}_{f-ACKS2}$ and $\Delta\Delta\mu = |\Delta\mu_{scf-ACKS2} - \Delta\mu_{f-ACKS2}|$, where $\Delta E^{pol.}_{f-ACKS2}$ is calculated according to eq 8b and $\Delta E^{pol.}_{scf-ACKS2}$ is calculated according to eq 13. The scf-corrections in fragment-based ACKS2 approaches are illustrated for a set of small molecular dimers (benzene, naphthalene, anthracene, and tetracene), where each data point represents a different nearest neighbor geometry extracted from the molecular single crystal structure.

illustrates the change of electronic polarization by addition of the iterative polarization response interaction of scfACKS2 and its validation reference KS-DFT based on a full Coulomb perturbation potential. Note, the scf-ACKS2 model includes different levels of approximations to the electrostatic intermolecular interaction treatment, a simple point charge representation for the fixed-charge distribution and an ACKS2 basis set-dependent representation of the mutual polarization response, c.f. eq 10. The response property trends of scf-ACKS2 generally exhibit good agreement with the DFT validation data. Considering that a self-consistent polarization response induces small changes of the polarization energies for fragment A and B, leading to further energetic stabilization of the charge, which is about an order of magnitude smaller compared to the polarization energy due to the fixed-charge contribution in f-ACKS2, illustrated earlier in Figure 6. The induced dipole moment of fragment B also increases notably, while it remains approximately constant for fragment A. Fragment A does not undergo significant electronic rearrangements during the iteration process because the response density on the charged fragment B is very small. In contrast, the local potential changes drastically for fragment B compared to the simple fixed-charge contribution by the neutral fragment A and hence increases its electronic response significantly compared to that in Figure 6

The f-ACKS2 polarization energy is calculated according to eq 8b (see main text and the Supporting Information for details on DFT calculations).

3.3. Electronic Polarization in A^+-B^- Anthracene Dimers. We investigate the mutual electronic polarization of two anthracene molecules in different relative orientations and distances, where anthracene molecular fragments A and B have a charge of +1e and −1e, respectively. This is a simple model for electronic polarization in an adjacent electron−hole pair or charge transfer exciton. Both anthracene molecules are aligned with their long axes in the *y* direction, with the second anthracene placed next to the first, shifted along the *x*-, *y*-, or *z*-axis to sample a number of different dimer configurations (cf. Figure 8). The non-self-consistent f-ACKS2 response to a fixed

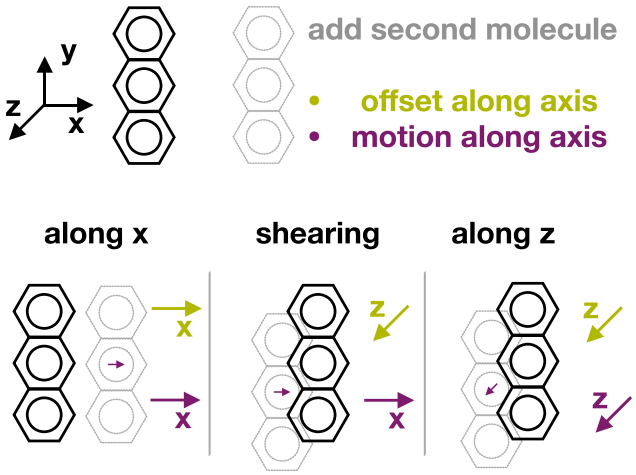


Figure 8. Structure and stacking geometry of anthracene molecular dimers.

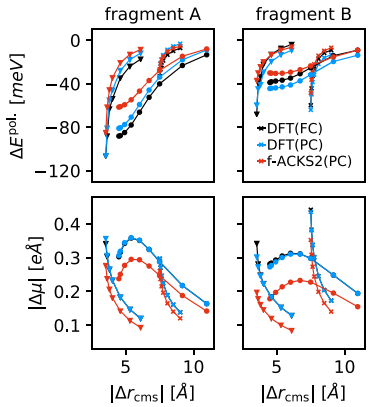
point charge intermolecular interaction model is again validated by DFT(PC) and DFT(FC). The response properties, induced dipole moment and polarization energy, are illustrated in Figure 9 for different relative arrangements of the anthracene molecular dimer.

The simulation models for the electronic polarization show good agreement across all different relative arrangements of the anthracene molecular dimer. In Figure 9a, the KS-DFT electronic polarization by the approximate fixed point charge interaction potential, DFT(PC), slightly underestimates (overestimates) the polarization energy of fragment A (fragment B) compared to a full Coulomb representation, DFT(FC), while induced dipole moments exhibit near-perfect coincidence for both methods. Similar to the previous section, the f-ACKS2 model tends to underestimate the response properties for all geometries. A more detailed list of the relative deviations, see Table 1, reveals a trend of increasing relative mean unsigned error in the ACKS2 model with increasing polarization perpendicular to the molecular bonding plane of the anthracene fragments. We expect the latter two effects to be a consequence of the small ACKS2 basis set representation employed.[53]

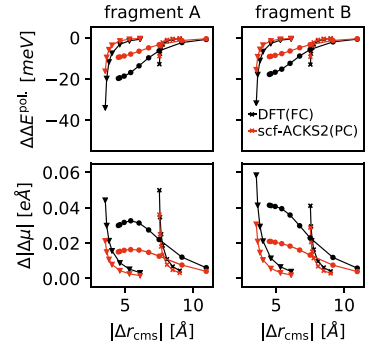
We include the mutual polarization interactions between the response of the different fragments to obtain a self-consistent electronic polarization and illustrate the respective change of induced dipole moments and polarization energies in Figure 9b. The trend of the scf-ACKS2 response changes matches those of the DFT reference very well for the parallel motion along the *x*-axis, while f-ACKS2, and in consequence the scf-ACKS2, still struggles to represent the motion along the *z*-axis and shearing of the molecular dimer. We expect the latter to be a consequence of the underestimation of the response to the fixed-charge interaction in Figure 9a, which accounts for the mutual polarization interaction contributions.

3.4. Polarization Energy of Electron−Hole Pair in 1D Chain. Finally, we study the stabilization of a separated electron−hole pair, located at two different molecular fragments, by the electronic response of its environment. To

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX



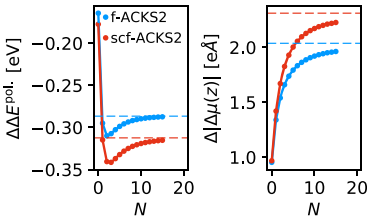
(a) Electronic response to fixed-charge distribution intermolecular interaction.



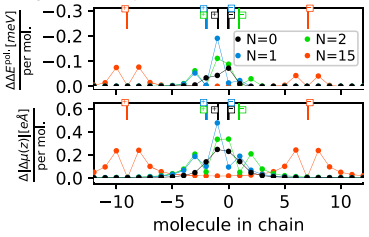
(b) Change of electronic response by addition of the mutual polarization response interaction potential.

Figure 9. Illustration of electronic response for different relative orientations (along the *x*-axis: star, shearing: circle, along the *z*-axis: triangle, displayed in Figure 8) and center-of-mass distances Δcms for an A^+-B^- anthracene molecular dimer. $\Delta E^{pol.}_{f-ACKS2(PC)}$ is calculated according to eq 8b and $\Delta E^{pol.}_{f-ACKS2(PC)}$ according to eq 13.

this end, we consider a chain of 50 anthracene molecules with a stacking pattern as shown in Figure 8, "along z", and a stacking distance of 4.0. One molecule of the chain is positively charged, and another one is negatively charged; all other molecules are neutral, with the number of neutral molecules between the two charged molecules denoted by *N*. Hence, we consider chains of the type $(A)_M$-$(A)^{+1}$-$(A)_N$-$(A)^{-1}$-$(A)_M$, where (A) denotes an anthracene molecule and $M = \frac{50-N-2}{2}$. The neutral fixed-charge distribution of the chain is again represented by fixed atomic point charges, fitted to the KS-DFT-based electrostatic potential of the free molecules. In order to show screening effects stabilizing the charged molecules, we place them at different separations along the chain, starting with a configuration where they are directly adjacent to each other, in analogy to section 3.3. The total electronic response of the anthracene molecule chain is depicted in Figure 10a as a function of *N*. To focus on the



(a) Total electronic response properties of the whole molecular chain. The dashed horizontal lines demonstrate the response properties of two independent chains containing a single independent charge carrier only.



(b) Electronic response properties of the anthracene chain resolved per molecule, where the positions of the charged molecular fragments are indicated individually by the vertical bars on top.

Figure 10. Illustration of the electronic polarization of a linear chain of anthracene molecules containing an electron−hole pair as a function of the number of separating molecules.

polarization effect solely due to the electron−hole pair, we subtract the polarization energy of the chain where all 50 molecules are neutral

$$\Delta\Delta E^{pol.} = \Delta E^{pol.}[\text{electron} - \text{hole}] - \Delta E^{pol.}[\text{neutral}] \tag{14}$$

$$\Delta|\Delta\mu(z)| = \sum_A^{\kappa} \Delta|\mu(z)[\text{electron} - \text{hole}]| - \sum_A^{\kappa} \Delta|\mu(z)[\text{neutral}]| \tag{15}$$

The total polarization energy as a function of *N* exhibits a rather unexpected trend. At first, it strongly decreases with

Table 1. Summary of the Relative Mean Unsigned Errors for the Electronic Polarization Properties of the A^+-B^- Anthracene Dimers Displayed in Figure 9a[a]

rmue [%]		$\Delta\mu_{DFT(PC)}$	$\Delta E^{pol.}_{DFT(PC)}$	$\Delta\mu_{f-ACKS2(PC)}$	$\Delta E^{pol.}_{f-ACKS2(PC)}$
along the *x*-axis	fragment A	0	32	13	19
	fragment B	0	24	19	33
shearing	fragment A	0	14	17	20
	fragment B	1	24	29	27
along the *z*-axis	fragment A	1	16	22	24
	fragment B	2	50	35	34

[a]Calculated following $\text{rmue}_{DFT(PC)} = \frac{1}{N}\sum_i^N \left|\frac{x_i^{DFT(PC)}}{x_i^{DFT(FC)}} - 1\right|$ and $\text{rmue}_{ACKS2(PC)} = \frac{1}{N}\sum_i^N \left|\frac{x_i^{ACKS2(PC)}}{x_i^{DFT(PC)}} - 1\right|$.

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

increasing N, goes through an energy minimum at $N = 3$, and then increases to reach a plateau at about $N = 12-15$. We explain this observation by two opposing trends. At an initially large distance, both the positively and negatively charged molecules lead to polarization of neutral molecules only, as displayed in Figure 10b for $N = 15$, and effectively act like independent charge carriers in a large one-dimensional chain of non-charged anthracene molecules. Once the electron−hole pair separation decreases, here at $N = 12-15$, the charge carriers start to polarize each other, while at the same time the polarizability of a negatively (positively) charged anthracene is greater (lesser) than a neutral molecule. In the present one-dimensional chain, the gain in polarization by the electron carrier outperforms loss in polarization by the hole carrier and yields a net growth in polarization energy with decreasing N. On the other hand, several neutral molecules in the chain are exposed to the presence of two charge carriers for small separations of the electron−hole pair. Thereby, the dielectric response of the N intermediate neutral molecules is bumped up linearly due to a simple amplifying superposition of opposite point charge interaction potentials. However, the anthracene molecules outside the electron−hole pair perceive the presence of the latter as a finite dipole moment, similar to an ideal point dipole, which exhibits stronger locality than an ideal point charge. As N decreases, the dipole perturbation potential contribution increases, and polarization energy decreases, which starts to dominate eventually at $N = 3$ in the one-dimensional chain here. On the contrary, we find that the absolute value of the induced dipole moment steadily grows with increasing electron−hole separation. This effect levels off until it reaches a plateau at $N = 15-20$. For large N, the polarization energy approaches the polarization energy of two non-interacting charges (dashed line), as it should. The induced dipole moments per anthracene molecule display a larger effective range dependency, c.f. $N = 15$ plot in Figure 10b. Hence, the point-dipole-like contributions to the intermolecular perturbation potential dominate the trend of the total dipole moment and overshadow any polarizability effects of the charged anthracene molecules. Figure 10b thus explains the mismatching trends of the polarization energy (dip at $N = 3$) and the induced dipole moment (no dip) observed in Figure 10a.

Overall, only a few molecules adjacent to the electron or hole carriers significantly contribute to the $\Delta\Delta E^{\mathrm{pol.}}$ and $\Delta|\Delta\mu(z)|$. This observed locality of the response in even a low dielectric system such as the here treated organic semiconductor chain opens up the f-ACKS2 method for use in efficient force field calculations based, e.g., on neighbor lists.

Figure 10a further illustrates the impact of polarization response self-consistency introduced in the scf-ACKS2 approach. As expected, the energetic stabilization and induced dipole moments increase compared to the non-self-consistent treatment. Despite a rather small difference, the non-iterative f-ACKS2 approach predicts the largest polarization stabilization for a spacing of two molecules, while the scf-ACKS2 yields three spacing fragments.

Finally, in order to gauge the importance of f-ACKS2/scf-ACKS2 polarization compared to the total electrostatic energy, we also need to account for the interaction between the two charge carriers themselves, not currently included in our approach. Instead, we compute them with a simple Coulomb model based on our point charge representation of the electrostatic potential. A more detailed description of this approximation can be found in the Supporting Information. The total electrostatic energy of the molecular chain calculated using this model is depicted in Figure 11. While the total electrostatic interaction energy in this system is dominated by the contributions of the (unpolarized) fixed-charge contribution, the electronic polarization energy as obtained by ACKS2 calculations yields a notable energetic stabilization, particularly at medium to large electron−hole separations. On the other hand, the mutual polarization response as described by our scf-ACKS2 scheme is almost negligible on the energy scales in section 3.4 and may therefore be regarded as a high-accuracy option for specific simulation targets.

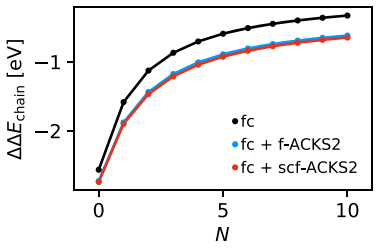


Figure 11. Illustration of the total electrostatic interaction energy of a one-dimensional chain of anthracene molecules based on a simple fixed-charge interaction model (fc) and f-ACKS2/scf-ACKS2 electronic polarization.

4. CONCLUSIONS

In summary, here, we have introduced a novel fragment-based approach to ACKS2 that allows us to calculate the electron density response, electronic polarization energy, and induced dipole moments of molecules in condensed phase molecular systems. We presented a detailed discussion of the fundamental approximations within the f-ACKS2 approach, with particular focus on the applied perturbation treatment, based on fixed-charge distribution interactions. These can optionally be augmented by a self-consistent mutual polarization response contribution. Validation of the electronic polarization contribution from f-ACKS2 and scf-ACKS2 shows reasonable accuracy compared to full KS-DFT for different singly charged organic molecular dimers as well as various relative orientations of an electron−hole dimer. An already simple point charge representation of the intermolecular perturbation potential was shown to account for the majority of the molecular polarization. A more accurate representation improved the f-ACKS2 electronic response. Additionally, future development of more flexible response basis sets beyond the simple s-p Gaussian basis used here should further improve the accuracy of the ACKS2 method compared to the parent-DFT method.

Finally, a 1D chain of anthracene molecules showcased the ability of the f-ACKS2 and scf-ACKS2 approaches to include first-principles-based electronic polarization in large molecular systems at a negligible cost compared to explicit electronic structure calculations. The present work is an important step toward force field development with ACKS2-based electronic polarization.

5. METHODS

Throughout this study, DFT reference calculations as well as the KS electronic structure-dependent ACKS2 parameter evaluations were carried out with the FHI-aims full potential

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

all electron DFT simulation package.[71] We applied the PBE generalized gradient approximated density functional, and electron spin was treated explicitly. Integrations were conducted using "tight" integration grids with wave functions expanded in a tier 2 numeric atomic orbital basis to ensure the numerical convergence of our results.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00151.

Detailed description of the KS-DFT origins of (f)ACKS2 as well as the energy references used to benchmark the method, lists of the dimers used in the testing procedure as well as convergence tests to fix the meta parameters of the anthracene chain model, and the method used to estimate the ground-state energies of the anthracene chain model (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Harald Oberhofer − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, D-85747 Garching, Germany;* orcid.org/0000-0002-5791-6736; Email: harald.oberhofer@tum.de

Authors

Patrick Gütlein − *Chair for Theoretical Chemistry and Catalysis Research Center, Technische Universität München, D-85747 Garching, Germany*

Jochen Blumberger − *Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.; Institute for Advanced Study, Technische Universität München, D-85748 Garching, Germany;* orcid.org/0000-0002-1546-6765

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00151

Author Contributions

The initial idea of the project was conceived by J.B. and H.O.; all authors contributed to the derivation of the (sc)f-ACKS2 method. P.G. conducted all implementation work and all calculations presented here. P.G. wrote the first draft of the manuscript, later corrected by J.B. and H.O.

Notes

The authors declare no competing financial interest.

Training data for the Gaussian parametrization as well as raw data of the test sets are available for free at the media server of the Technical University of Munich at the following link https://mediatum.ub.tum.de/1546236.

All software used in this work are present as an in-house proof-of-concept Python package to be made available upon reasonable request.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge support from the Solar Technologies Go Hybrid Initiative of the State of Bavaria. P.G. further acknowledges the support of the Technische Universität München − Institute for Advanced Study, funded by the German Excellence Initiative (and the European Union Seventh Framework Programme under Grant Agreement 291763). H.O. acknowledges support from the German Research Foundation (DFG), priority program 1928 COORNETS, Grant OB425/3-1. J.B. acknowledges TUM-IAS for the award of a generous Hans Fischer Fellowship.

■ REFERENCES

(1) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(2) Sun, H. COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase ApplicationsOverview with Details on Alkane and Benzene Compounds. *J. Phys. Chem.B* **1998**, *102*, 7338−7364.

(3) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(4) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; van Duin, A. C. T. The ReaxFF reactive force-field: development, applications and future directions. *npj Comput. Mater.* **2016**, *2*, 15011.

(5) Mackerell, A. D. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem.* **2004**, *25*, 1584−1604.

(6) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2009**, 671.

(7) Futera, Z.; Blumberger, J. Adsorption of Amino Acids on Gold: Assessing the Accuracy of the GolP-CHARMM Force Field and Parametrization of Au−S Bonds. *J. Chem. Theory Comput.* **2019**, *15*, 613−624.

(8) Mueller, J. E.; van Duin, A. C. T.; Goddard, W. A., III Development and Validation of ReaxFF Reactive Force Field for Hydrocarbon Chemistry Catalyzed by Nickel. *J. Phys. Chem. C* **2010**, *114*, 4939−4949.

(9) Thiel, W. Computational Catalysis—Past, Present, and Future. *Angew. Chem. Int. Ed.* **2014**, *53*, 8605−8613.

(10) Yun, K.-S.; Pai, S. J.; Yeo, B. C.; Lee, K.-R.; Kim, S.-J.; Han, S. S. Simulation Protocol for Prediction of a Solid-Electrolyte Interphase on the Silicon-based Anodes of a Lithium-Ion Battery: ReaxFF Reactive Force Field. *J. Phys. Chem. Lett.* **2017**, *8*, 2812−2818.

(11) Heenen, H. H.; Voss, J.; Scheurer, C.; Reuter, K.; Luntz, A. C. Multi-ion Conduction in Li3OCl Glass Electrolytes. *J. Phys. Chem. Lett.* **2019**, *10*, 2264−2269.

(12) Rühle, V.; Lukyanov, A.; May, F.; Schrader, M.; Vehoff, T.; Kirkpatrick, J.; Baumeier, B.; Andrienko, D. Microscopic Simulations of Charge Transport in Disordered Organic Semiconductors. *J. Chem. Theory Comput.* **2011**, *7*, 3335−3345.

(13) Giannini, S.; Carof, A.; Blumberger, J. Crossover from Hopping to Band-Like Charge Transport in an Organic Semiconductor Model: Atomistic Nonadiabatic Molecular Dynamics Simulation. *J. Phys. Chem. Lett.* **2018**, *9*, 3116−3123.

(14) Engkvist, O.; Åstrand, P.-O.; Karlström, G. Accurate Intermolecular Potentials Obtained from Molecular Wave Functions: Bridging the Gap between Quantum Chemistry and Molecular Simulations. *Chem. Rev.* **2000**, *100*, 4087−4108.

(15) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(16) Vizcarra, C. L.; Mayo, S. L. Electrostatics in computational protein design. *Curr. Opin. Chem. Biol.* **2005**, *9*, 622−626.

(17) Wen, S.; Beran, G. J. O. Accurate Molecular Crystal Lattice Energies from a Fragment QM/MM Approach with On-the-Fly Ab Initio Force Field Parametrization. *J. Chem. Theory Comput.* **2011**, *7*, 3733−3742.

(18) Kordt, P.; van der Holst, J. J. M.; Al Helwi, M.; Kowalsky, W.; May, F.; Badinski, A.; Lennartz, C.; Andrienko, D. Modeling of Organic Light Emitting Diodes: From Molecular to Device Properties. *Adv. Funct. Mater.* **2015**, *25*, 1955−1971.
(19) Sherrill, C. D.; Sumpter, B. G.; Sinnokrot, M. O.; Marshall, M. S.; Hohenstein, E. G.; Walker, R. C.; Gould, I. R. Assessment of standard force field models against high-quality ab initio potential curves for prototypes of π-π, CH/π, and SH/π interactions. *J. Comput. Chem.* **2009**, *30*, 2187−2193.
(20) Leontyev, I.; Stuchebrukhov, A. Accounting for electronic polarization in non-polarizable force fields. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2613−2626.
(21) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A., Jr.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(22) Kramer, C.; Spinn, A.; Liedl, K. R. Charge Anisotropy: Where Atomic Multipoles Matter Most. *J. Chem. Theory Comput.* **2014**, *10*, 4488−4496.
(23) Jakobsen, S.; Bereau, T.; Meuwly, M. Multipolar Force Fields and Their Effects on Solvent Dynamics around Simple Solutes. *J. Phys. Chem.B* **2015**, *119*, 3034−3045.
(24) Cardamone, S.; Hughes, T. J.; Popelier, P. L. A. Multipolar electrostatics. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10367−10387.
(25) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramolecular Interactions and Ligand-Macromolecule Complexes. A Bottom-Up Strategy. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.
(26) Gordon, M. S.; Smith, Q. A.; Xu, P.; Slipchenko, L. V. Accurate First Principles Model Potentials for Intermolecular Interactions. *Annu. Rev. Phys. Chem.* **2013**, *64*, 553−578.
(27) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046−4063.
(28) Halgren, T. A.; Damm, W. Polarizable force fields. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236−242.
(29) Warshel, A.; Kato, M.; Pisliakov, A. V. Polarizable Force Fields: History, Test Cases, and Prospects. *J. Chem. Theory Comput.* **2007**, *3*, 2034−2045.
(30) Lopes, P. E. M.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell, A. D., Jr. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2013**, *9*, 5430−5449.
(31) Liu, Z.; Timmermann, J.; Reuter, K.; Scheurer, C. Benchmarks and Dielectric Constants for Reparametrized OPLS and Polarizable Force Field Models of Chlorinated Hydrocarbons. *J. Phys. Chem. B* **2018**, *122*, 770−779.
(32) Lamoureux, G.; Roux, B. Modeling induced polarization with classical Drude oscillators: Theory and molecular dynamics simulation algorithm. *J. Chem. Phys.* **2003**, *119*, 3025−3039.
(33) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D., Jr. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116*, 4983−5013.
(34) Applequist, J. An atom dipole interaction model for molecular optical properties. *Acc. Chem. Res.* **2002**, *10*, 79−85.
(35) Davis, M. E.; McCammon, J. A. Electrostatics in biomolecular structure and dynamics. *Chem. Rev.* **1990**, *90*, 509−521.
(36) Anisimov, V. M.; Vorobyov, I. V.; Roux, B.; MacKerell, A. D. Polarizable Empirical Force Field for the Primary and Secondary Alcohol Series Based on the Classical Drude Model. *J. Chem. Theory Comput.* **2007**, *3*, 1927−1946.
(37) Yu, W.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D., Jr. Six-site polarizable model of water based on the classical Drude oscillator. *J. Chem. Phys.* **2013**, *138*, No. 034508.
(38) Xie, W.; Gao, J. Design of a Next Generation Force Field: The X-POL Potential. *J. Chem. Theory Comput.* **2007**, *3*, 1890−1900.
(39) Gao, J. Toward a Molecular Orbital Derived Empirical Potential for Liquid Simulations. *J. Phys. Chem. B* **1997**, *101*, 657−663.
(40) Gao, J. A molecular-orbital derived polarization potential for liquid water. *J. Chem. Phys.* **1998**, *109*, 2346−2354.
(41) Jiang, N.; Ma, J. Conformational Simulations of Aqueous Solvated alpha-Conotoxin GI and Its Single Disulfide Analogues Using a Polarizable Force Field Model. *J. Phys. Chem. A* **2008**, *112*, 9854−9867.
(42) Zhu, Q.; Lu, Y.; He, X.; Liu, T.; Chen, H.; Wang, F.; Zheng, D.; Dong, H.; Ma, J. Entropy and polarity control the partition and transportation of drug-like molecules in biological membrane. *Sci. Rep.* **2017**, *7*, 17749.
(43) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. X-Pol Potential: An Electronic Structure-Based Force Field for Molecular Dynamics Simulation of a Solvated Protein in Water. *J. Chem. Theory Comput.* **2009**, *5*, 459−467.
(44) Jiang, N.; Ma, J. Multi-layer coarse-graining polarization model for treating electrostatic interactions of solvated α-conotoxin peptides. *J. Chem. Phys.* **2012**, *136*, 134105.
(45) Wang, X.; Yan, T.; Ma, J. Polarizable force fields based on physical models and quantum chemical calculations. *Int. J. Quantum Chem.* **2015**, *115*, 545−549.
(46) Mortier, W. J.; Van Genechten, K.; Gasteiger, J. Electronegativity equalization: application and parametrization. *J. Am. Chem. Soc.* **1985**, *107*, 829−835.
(47) Rappe, A. K.; Goddard, W. A. Charge equilibration for molecular dynamics simulations. *J. Phys. Chem.* **1991**, *95*, 3358−3363.
(48) York, D. M.; Yang, W. A chemical potential equalization method for molecular simulations. *J. Chem. Phys.* **1996**, *104*, 159−172.
(49) Chelli, R.; Procacci, P.; Righini, R.; Califano, S. Electrical response in chemical potential equalization schemes. *J. Chem. Phys.* **1999**, *111*, 8569−8575.
(50) Nistor, R. A.; Polihronov, J. G.; Müser, M. H.; Mosey, N. J. A generalization of the charge equilibration method for nonmetallic materials. *J. Chem. Phys.* **2006**, *125*, No. 094108.
(51) Verstraelen, T.; Ayers, P. W.; van Speybroeck, V.; Waroquier, M. ACKS2: Atom-condensed Kohn-Sham DFT approximated to second order. *J. Chem. Phys.* **2013**, *138*, No. 074108.
(52) Verstraelen, T.; Vandenbrande, S.; Ayers, P. W. Direct computation of parameters for accurate polarizable force fields. *J. Chem. Phys.* **2014**, *141*, 194114.
(53) Gütlein, P.; Lang, L.; Reuter, K.; Blumberger, J.; Oberhofer, H. Toward First-Principles-Level Polarization Energies in Force Fields: A Gaussian Basis for the Atom-Condensed Kohn−Sham Method. *J. Chem. Theory Comput.* **2019**, *15*, 4516−4525.
(54) Schober, C.; Reuter, K.; Oberhofer, H. Critical analysis of fragment-orbital DFT schemes for the calculation of electronic coupling values. *J. Chem. Phys.* **2016**, *144*, No. 054103.
(55) Lee Warren, G.; Davis, J. E.; Patel, S. Origin and control of superlinear polarizability scaling in chemical potential equalization methods. *J. Chem. Phys.* **2008**, *128*, 144110.
(56) Ángyán, J. G. Correlation of bond orders and softnesses. *J. Mol. Struct.: THEOCHEM* **2000**, *501-502*, 379−388.
(57) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(58) Oberhofer, H.; Blumberger, J. Insight into the Mechanism of the Ru2+−Ru3+ Electron Self-Exchange Reaction from Quantitative Rate Calculations. *Angew. Chem. Int. Ed.* **2010**, *49*, 3631−3634.
(59) Blumberger, J.; McKenna, K. P. Constrained density functional theory applied to electron tunnelling between defects in MgO. *Phys. Chem. Chem. Phys.* **2013**, *15*, 2184−2196.
(60) Spencer, J.; Gajdos, F.; Blumberger, J. FOB-SH: Fragment orbital-based surface hopping for charge carrier transport in organic and biological molecules and materials. *J. Chem. Phys.* **2016**, *145*, 064102.
(61) Giannini, S.; Carof, A.; Ellis, M.; Yang, H.; Ziogos, O. G.; Ghosh, S.; Blumberger, J. Quantum localization and delocalization of

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

charge carriers in organic semiconducting crystals. *Nat. Commun.* **2019**, *10*, 3843.
(62) Lagardère, L.; Lipparini, F.; Polack, É.; Stamm, B.; Cancès, É.; Schnieders, M.; Ren, P.; Maday, Y.; Piquemal, J. P. Scalable Evaluation of Polarization Energy and Associated Forces in Polarizable Molecular Dynamics: I. Toward Massively Parallel Direct Space Computations. *J. Chem. Theory Comput.* **2014**, *10*, 1638−1651.
(63) Nocito, D.; Beran, G. J. O. Massively Parallel Implementation of Divide-and-Conquer Jacobi Iterations Using Particle-Mesh Ewald for Force Field Polarization. *J. Chem. Theory Comput.* **2018**, *14*, 3633−3642.
(64) Albaugh, A.; Demerdash, O.; Head-Gordon, T. An efficient and stable hybrid extended Lagrangian/self-consistent field scheme for solving classical mutual induction. *J. Chem. Phys.* **2015**, *143*, 174104.
(65) Vitale, V.; Dziedzic, J.; Albaugh, A.; Niklasson, A. M. N.; Head-Gordon, T.; Skylaris, C.-K. Performance of extended Lagrangian schemes for molecular dynamics simulations with classical polarizable force fields and density functional theory. *J. Chem. Phys.* **2017**, *146*, 124115.
(66) Aviat, F.; Lagardère, L.; Piquemal, J.-P. The truncated conjugate gradient (TCG), a non-iterative/fixed-cost strategy for computing polarization in molecular dynamics: Fast evaluation of analytical forces. *J. Chem. Phys.* **2017**, *147*, 161724.
(67) Budzianowski, A.; Katrusiak, A. Pressure-frozen benzene I revisited. *Acta Crystallogr., Sect. B* **2006**, *62*, 94−101.
(68) Ponomarev, V. I.; Filipenko, O. S.; Atovmyan, L. O. *Kristallografiya* **1976**, *21*, 392.
(69) Brock, C. P.; Dunitz, J. D. Temperature dependence of thermal motion in crystalline anthracene. *Acta Crystallogr., Sect. B* **1990**, *46*, 795−806.
(70) Holmes, D.; Kumaraswamy, S.; Matzger, A. J.; Vollhardt, K. P. C. On the Nature of Nonplanarity in the [N]Phenylenes. *Chem.- Eur. J.* **1999**, *5*, 3399−3412.
(71) Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M. Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* **2009**, *180*, 2175−2196.
(72) Stone, A. J. Distributed multipole analysis, or how to describe a molecular charge distribution. *Chem. Phys. Lett.* **1981**, *83*, 233−239.
(73) Price, S. L. A distributed multipole analysis of the charge densities of some aromatic hydrocarbons. *Chem. Phys. Lett.* **1985**, *114*, 359−364.
(74) Fowler, P. W.; Buckingham, A. D. Central or distributed multipole moments? Electrostatic models of aromatic dimers. *Chem. Phys. Lett.* **1991**, *176*, 11−18.

https://dx.doi.org/10.1021/acs.jctc.0c00151
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX