UNIVERSITY OF CALIFORNIA,
IRVINE

Developing More Accurate Methods of Warm Dense Matter Simulation via Finite-Temperature Density Functional Theory

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

John Kozlowski

Dissertation Committee:
Professor Kieron Burke, Chair
Professor Steven R. White
Professor Craig Martens

2023

DEDICATION

To my parents, my grandmother, and Ricky.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Reflecting on my journey, I cannot help but be overcome with a deep sense of gratitude for those that have supported me along the way. Graduate school has been difficult to say the least, and I am sincerely indebted to those that have stood by me to help celebrate successes, overcome failures, lament misfortunes, and awaken perseverance. The sincere appreciation I feel for your generosity over the years cannot be expressed in words (but I will try).

I want to wholeheartedly acknowledge the honor it has been to work closely with my advisor, Kieron Burke. From the very beginning (i.e. when he asked me to take out his office trash upon first meeting me), I knew there were *many* things he could teach me. In earnest, his guidance over the last few years has been some of the most meaningful in my life, and I thank him sincerely for making me a better scientist, person, and coffee-addict.

I couldn't have asked for a better group of colleagues in the Burke group. You all provided me with a second family, full of fellow students and postdocs alike. In no particular order, I would like to thank Ryan Pederson, Bhupalee Kalita, Thais Scott, Suhwan Song, Pavel Okun, Steven Crisostomo, Brandon Momanyi, Natali Fisher, and Francisca Sagredo for the infinite support, perspective, discussion, and fun you have provided over the years. I am especially indebted to Natali, who I have had the pleasure of mentoring, for she has surely made me a better teacher, co-worker, and friend.

There exists a very large group ($N \to \infty$) of friends and family members that have inspired me, all of which deserve to be acknowledged here. This includes (but is not limited to) *my friends:* Ryan & Lauren Pederson, Claire Butler, Alissa Matus, Cynthia Wong, Chantel Mao, Dayra Lopez, and Ernesto Barraza-Valdez; *my family:* the Kozlowski's, the Jenkin's, all of our beloved companions, and Jaeger Kozlowski-Lui; *my chosen family:* Evelyn Huerta, Angelica Castro, and Mayra & Eric Preciado-Gwynn. You all have encouraged me along this journey, each in their own way. Thank you for everything.

To my parents: The unwavering devotion you have shown our family is truly remarkable. You have always believed in me, from the very beginning, with no exceptions. I cannot thank you enough, and I love you both dearly.

To my partner, Ricky: I could have never imagined meeting someone like you, particularly not while lost in the chaos of graduate school, during a global pandemic, on the verge of 5 (give or take) equally terrifying things. The level of love and compassion you command is humbling, and it inspires me everyday. Decades from now, I have no idea where I will be, or what I will be doing, but there is nobody else I would rather share that uncertainty with. Thank you from the bottom of my heart.

I carry you all with me, everyday.

My work as a graduate student was funded by the Department of Energy (Award No. DE-FG02-08ER46496), the National Science Foundation (Award No. CHE-2154371), and the Brython Davis Fellowship at the University of California, Irvine.

This dissertation is a collection of selected journal articles and a book chapter. I acknowledge the following copyright holders:

⟨Ch. 2⟩ Reproduced with permission from Kieron Burke and John Kozlowski; "Lies My Teacher Told Me About Density Functional Theory: Seeing Through Them with the Hubbard Dimer," in *Simulating Correlations with Computers*, edited by E. Pavarini and E. Koch (Forschungszentrum Jülich, 2021) Chap. 3, pp. 65-96. Copyright 2021 Forschungszentrum Jülich GmbH.

⟨Ch. 3⟩ Reproduced with permission from Stefan Vuckovic, Suhwan Song, John Kozlowski, Eunji Sim, and Kieron Burke; "Density Functional Analysis: The Theory of Density-Corrected DFT." *Journal of Chemical Theory and Computation* 15(12), 6636-6646 (2019). Copyright 2019 American Chemical Society.

⟨Ch. 4⟩ Reproduced with permission from Thais Scott, John Kozlowski, Steven Crisostomo, Aurora Pribram-Jones, and Kieron Burke; "Exact Conditions for Ensemble Density Functional Theory." *Under Review with Physical Review Letters* (2023). Copyright 2023 American Physical Society, if accepted.

⟨Ch. 5⟩ Reproduced with permission from John Kozlowski, Dennis Perchak, and Kieron Burke; "Generalized Gradient Approximation Made Thermal." *Under Review with Physical Review Letters* (2023). Copyright 2023 American Physical Society, if accepted.

⟨Ch. 6⟩ Reproduced with permission from Ryan Pederson, John Kozlowski, Ruyi Song, Jackson Beall, Martin Ganahl, Markus Hauru, Adam GM Lewis, Yi Yao, Shrestha Basu Mallick, Volker Blum, and Guifre Vidal; "Large Scale Quantum Chemistry with Tensor Processing Units." *Journal of Chemical Theory and Computation* 19(1), 25-32 (2022). Copyright 2022 American Chemical Society.

VITA

John Kozlowski

EDUCATION

Doctor of Philosophy in Chemistry **2018–2023**
University of California, Irvine
Advisor: Kieron Burke

Master of Science in Chemistry **2018–2022**
University of California, Irvine
Advisor: Kieron Burke

Bachelor of Science in Chemistry **2014–2018**
University of California, Los Angeles
Departmental Honors

RESEARCH EXPERIENCE

Graduate Student Researcher **2018–2023**
University of California, Irvine

Resident Research Scientist **2021–2022**
X, the moonshot factory (formerly Google X)

Undergraduate Student Researcher **2017–2018**
University of California, Los Angeles

TEACHING EXPERIENCE

Course Co-Creator: Machine Learning in Physical Sciences **2021–2023**
University of California, Irvine

Summer Undergraduate Research Program (SURP) Mentor **2020–2023**
University of California, Irvine

Science Olympiad Event Writer: Water Quality B/C **2020–2021**
University of California, Irvine

Graduate Teaching Assistant **2018–2020**
University of California, Irvine

PUBLICATIONS

Generalized Gradient Approximation Made Thermal **2023**
J. Kozlowski, D. Perchak, and K. Burke; *Under Review with Physical Review Letters.* ArXiv: 2308.03319

Exact Conditions for Ensemble Density Functional Theory **2023**
T. Scott, J. Kozlowski, S. Crisostomo, A. Pribram-Jones, and K. Burke; *Under Review with Physical Review Letters.* ArXiv: 2307.00187

Seven Useful Questions in Density Functional Theory **2023**
S. Crisostomo, R. Pederson, J. Kozlowski, B. Kalita, A. Cancio, K. Datchev, A. Wasserman, S. Song, and K. Burke; *Letters in Mathematical Physics* 113, 42.

Large Scale Quantum Chemistry with Tensor Processing Units **2022**
R. Pederson, J. Kozlowski, R. Song, J. Beall, M. Ganahl, M. Hauru, A.G. Lewis, S.B. Mallick, V. Blum, Y. Yao, and G. Vidal; *Journal of Chemical Theory and Computation* 19(1), 25-32.

Lies My Teacher Told Me About Density Functional Theory: Seeing Through Them with the Hubbard Dimer **2021**
J. Kozlowski and K. Burke; Appears in *Simulating Correlations with Computers*, edited by E. Pavarini and E. Koch (Forschungszentrum Jülich, 2021) Chap. 3, pp. 65-96.

Density Functional Analysis: The Theory of Density-Corrected DFT **2019**
S. Vuckovic, S. Song, J. Kozlowski, E. Sim, and K. Burke; *Journal of Chemical Theory and Computation* 15(12), 6636-6646.

HONORS, AWARDS, & FELLOWSHIPS

Wolfsberg Award in Theoretical Chemistry **2023**

Brython Davis Fellowship **2020-2023**

NSF Graduate Research Fellowship Honorable Mention **2020**

Alumni Summer Undergraduate Research Fellowship **2017**

Alpha Chi Sigma Summer Research Fellowship **2016**

ABSTRACT OF THE DISSERTATION

Developing More Accurate Methods of Warm Dense Matter Simulation via Finite-Temperature Density Functional Theory

By

John Kozlowski

Doctor of Philosophy in Chemistry

University of California, Irvine, 2023

Professor Kieron Burke, Chair

Warm Dense Matter (WDM) can be found in a myriad of useful and worthwhile applications, whether it be in the ongoing study of inertial confinement fusion, novel shock experiments, or planetary interiors. One crucial tool that has revolutionized our understanding of WDM over the last few decades has been Finite-Temperature Density Functional Theory (FT-DFT), which has enabled various groundbreaking simulations in the field (often when experimental studies were too difficult and costly to be performed). Although successful, modern implementations of FT-DFT suffer from significant drawbacks, namely the large computational burden inherent to studying high-temperature materials, and the lack of sophisticated approximations describing temperature-dependent exchange-correlation interactions. The research contained in this dissertation works to address these fundamental challenges both directly and indirectly.

Chapter 2 serves as an introduction to the underlying theorems of ground-state DFT, all of which extend to nonzero temperatures, and are depicted through illustrative examples using the Hubbard dimer. Chapter 3 introduces Density-Corrected DFT, a novel theory studying the origin of errors in functional approximations, and a means of analysis for approximations in both the ground state and at finite temperatures.

Chapter 4 derives exact conditions for Ensemble Density Functional Theory, all of which provide insight into the behavior of temperature-dependent exact conditions for FT-DFT. Chapter 5 proposes a novel temperature-dependent generalized gradient approximation, the locally thermal Perdew–Burke–Ernzerhof (ltPBE) approximation, for FT-DFT calculations of WDM. This temperature-dependent approximation is the first of its kind, with its origin rooted entirely in a model of the PBE gradient-dependent exchange-correlation hole density at nonzero temperatures.

Lastly, Chapter 6 discusses how Tensor Processing Units (hardware that was originally designed to perform machine-learning tasks) has been repurposed to perform highly efficient Kohn-Sham DFT calculations. This in particular has significant impact on future studies of WDM, where high temperatures greatly increase the computational cost of calculations.

Chapter 1

Introduction

Due to the extreme (and often expensive) conditions required to study Warm Dense Matter (WDM) experimentally, much of the groundbreaking research in this field over the past few decades has been performed computationally [55, 123, 147, 260, 207, 341]. However, there exist various difficulties inherent to theoretical descriptions of WDM, many of which have long been noted [99, 223]. In fact the WDM regime, commonly described as a subset of matter with characteristic densities and temperatures between condensed matter and plasma, is commonly referred to as the "malfunction junction," referring to the notorious difficulty scientists from each of these fields have experienced while studying this complicated state [223]. These challenges are to be expected, as any fully theoretical description of WDM must simultaneously address phenomena such as quantum effects, strong correlation, and partial ionization [223].

Over the years, many of the tremendous leaps in our ability to understand WDM have taken place due to clever applications of *ab initio* Molecular Dynamics (MD), this being a mixed quantum-classical simulation method scientists have utilized to predict characteristic properties of various WDM systems. In particular, this approach has been most successful

when Finite-Temperature Density Functional Theory (FT-DFT) is employed to generate forces during each iteration of a classical MD simulation [128, 202]. Here, the states of a system are thermally occupied and solved self-consistently within the Mermin-Kohn-Sham scheme [157, 206], where all terms contributing to the Helmholtz free energy are known other than the exchange-correlation (XC) component:

$$A_{\mathrm{XC}}[n] = A[n] - T_{\mathrm{s}}[n] - V_{\mathrm{ext}}[n] - U[n] + \tau S_{\mathrm{s}}[n], \tag{1.1}$$

where A is the Helmholtz free energy, T_{s} is the non-interacting kinetic energy, V_{ext} is the external potential energy, U is the classical electrostatic interaction energy, and S_{s} is the non-interacting entropic contribution at temperature τ. This is the underlying approximation of all modern FT-DFT calculations, many of which still make use of ground-state XC functionals (due to the inherent difficulty of developing systematic temperature-dependent approximations). If the exact temperature-dependent XC expression were known, FT-DFT would yield the exact free energy of a system, which in turn could be utilized (in conjunction with MD) to better describe the properties of WDM.

Thus, the work discussed throughout this dissertation focuses on the ongoing development of FT-DFT, addressing many of the issues plaguing its implementation in modern WDM calculations. Below I present a selection of self-contained chapters centered on research performed with my advisor and collaborators to enhance its implementation, either through direct development of novel approximations, or analysis of core/surrounding levels of theory related to FT-DFT. I have organized this work in the following manner:

Chapter 2 introduces the theory of ground-state DFT through the lens of the simplest possible system (with correlated electrons) available, the asymmetric Hubbard dimer. This system serves as a valuable tool for the development of novel theorems and approximations in DFT; we use it here to illustrate various concepts at the core of all DFT calculations (all

of which generalize to nonzero temperature).

Chapter 3 proposes the novel theory of Density-Corrected DFT (DC-DFT), which is a manner of understanding approximations in DFT, and the origin of their errors.

Chapter 4 proves various exact conditions for Ensemble DFT (EDFT), which is utilized to describe ensembles of both ground- and excited-state systems. The development of exact conditions in this field is important not only for the benefit of EDFT applications (such as the calculation of accurate excitation energies), but for the insight it provides into exact conditions for FT-DFT. This is because EDFT, where states are not necessarily occupied by a Fermi-Dirac distribution, can be thought of as a more general version of FT-DFT.

Chapter 5 proposes a novel temperature-dependent Generalized Gradient Approximation (GGA) stemming from a recently-invented theory known as Conditional-Probability DFT. Here we derive a new and extremely simple ansatz for the study of WDM at the thermal GGA level, referred to as the locally thermal Perdew–Burke–Ernzerhof (ltPBE) approximation. Extending PBE to nonzero temperatures, this approximation unlocks a new level of accuracy for modern WDM simulations, and thus may have significant consequences for the field.

Lastly, Chapter 6 explores the use of Google's cloud-based Tensor Processing Units (TPU's) to both accelerate and scale-up conventional DFT calculations by re-purposing hardware designed for machine learning to perform highly efficient Kohn-Sham DFT calculations. This work addresses the large computational burden required to perform FT-DFT calculations of WDM, which often involve temperatures where a multitude of high-energy states are occupied, and thus would greatly benefit from more efficient KS-DFT calculations.

As each of these chapters originate from different pieces of work, much of the notation changes throughout. To avoid confusion, please refer to the notation given in each self-contained chapter.

Chapter 2

Lies My Teacher Told Me About Density Functional Theory: Seeing Through Them with the Hubbard Dimer

This chapter is a reproduction of Ref. [27], which I co-authored with Kieron Burke.

2.1 Abstract

Most realistic calculations of moderately correlated materials begin with a ground-state density functional theory (DFT) calculation. While Kohn-Sham DFT is used in about 40,000 scientific papers each year, the fundamental underpinnings are not widely appreciated. In this chapter, we analyze the inherent characteristics of DFT in their simplest form, using the asymmetric Hubbard dimer as an illustrative model. We begin by working through the

core tenets of DFT, explaining what the exact ground-state density functional yields and does not yield. Given the relative simplicity of the system, almost all properties of the exact exchange-correlation functional are readily visualized and plotted. Key concepts include the Kohn-Sham scheme, the behavior of the XC potential as correlations become very strong, the derivative discontinuity and the difference between KS gaps and true charge gaps, and how to extract optical excitations using time-dependent DFT. By the end of this text and accompanying exercises, the reader will improve their ability to both explain and visualize the concepts of DFT, as well as better understand where others may go wrong. This chapter appears in the book Autumn School on Correlated Electrons: Simulating Correlations with Computers (2021) prepared by Forschungszentrum Jülich.

2.2 Introduction

Density functional theory (DFT) is an extremely sophisticated approach to many-body problems [194, 16]. It must be among the most used and least understood of all successful theories in physics. Currently, about 50,000 papers each year report results of Kohn-Sham (KS) DFT calculations [256], including room temperature superconductors under high pressure [251], heterogeneous catalysis at metal surfaces and for nanoparticles [220], understanding the interior of Jupiter and exoplanets [340], studying how ocean acidification affects the seabream population [311], and even which water to use when making coffee [121].

But much of modern condensed matter physics involves using model Hamiltonians to study strongly correlated systems, where understanding new phenomena is considered far more important than generating accurate materials-specific properties [166, 282]. In fact, our standard diagrammatic approach (expansions in the strength of the electron-electron coupling) is hard-wired into all our descriptions of such many-body phenomena, be it the fractional quantum Hall effect [302] or the Kondo effect (even when perturbation theory fails, we still

think of resummed diagrams) [155].

Because DFT is *logically* subtle, without requiring much mathematical gymnastics (although they are available for those that enjoy them [183]) or skill with summing Feynman diagrams, and because DFT is entirely different from the standard approach, most of what you may have learned is hopelessly confused or simply downright untrue. Hence the title of this article, taken from a popular book on history [187]. For example, any conflation of the KS scheme with traditional mean-field theory is a dire mistake, and should be avoided at all costs.

This chapter is primarily designed to explain essential concepts of DFT to theorists more familiar with standard many-body theory and perhaps more experienced in dealing with strongly correlated systems. It should also prove useful for anyone performing DFT calculations on weakly correlated systems, who might be wondering where things go wrong as correlations grow stronger. Additionally, the Hubbard dimer is a wonderful teaching tool for basic concepts, as so many of its exact results can be derived analytically.

The first use of this material came in a conversation between KB and Duncan Haldane at a meeting sponsored by the US Department of Energy. Duncan asked KB to explain this DFT business, and he suggested the dimer as the minimal relevant model. After 45 minutes of tough argument, Haldane said "That's the first time I've ever really understood this Kohn-Sham scheme. Thanks." Within 2 years, he was awarded a share in a Nobel Prize in physics [113]. While correlation is not causation, Haldane did not win his share until *after* he understood KS-DFT with the aid of this simple model!

However, it is important to note that the benefits of this type of analysis are not solely limited to those working in theoretical physics. In the fields of theoretical chemistry and material science, for instance, where ground-state electronic energies are often required to be extremely accurate [168, 65, 317], there has been growing technological interest in the

study of both chemically complex and strongly correlated materials [112, 227]. This chapter was partly designed with these fields in mind, serving as a resource for any computational scientist who wishes to better comprehend the limitations of their computational methods. Throughout this text, there will be various highlighted sections dedicated to examples, exercises, and key concepts to aid the reader in applying what is learned in this study to their own endeavors.

There are now a huge number of diverse introductions to DFT, with many different perspectives. These include a simple tutorial for anyone with knowledge of quantum mechanics [29], a very long online textbook with lots of nasty problems [22], a many-body introduction [56], and even video lectures [50]. But this chapter is specifically aimed at explaining the most essential concepts, and why strongly correlated systems are more challenging in DFT. All the Hubbard material appears in two long review articles, one on the ground state theory [35] and a second on linear-response TDDFT [33]. The Hubbard dimer has been recently used to explore effects in other aspects of DFT, such as magnetic DFT [305], ensemble DFT [267], and thermal DFT [269].

Takeaway: DFT appears deceptively simple to understand. It is much trickier than people realize. This chapter provides a unique explanation of basic ideas using a simple model.

2.2.1 Background

We work in the non-relativistic non-magnetic Born-Oppenheimer approximation, using Hartree atomic units ($e^2 = \hbar = m_e = 1$). The Hamiltonian for the electrons is simple and known exactly

$$\hat{H} = \hat{T} + \hat{V}_{\text{ee}} + \hat{V}, \tag{2.1}$$

where $\hat{T}$ is their kinetic energy, $\hat{V}_{\rm ee}$ is the electron-electron Coulomb repulsion, and $\hat{V}$ is the one-body potential, equal to a sum of Coulomb attractions to the ions in an isolated molecule or solid. We let N be the number of electrons.

A first-principles approach to this problem is to feed a computer a list of nuclear types and positions and, following a recipe, it spits out various properties of the electronic system. In quantum chemistry [290], the recipe is called a model chemistry [222, 9] if both the method (e.g. Hartree-Fock) and the basis set are specified.

We contrast this with traditional approaches in condensed matter [80]. Often a model Hamiltonian is written down, hoping that it describes the dominant physical effects. For most interesting problems, standard approaches to solving this Hamiltonian will fail, i.e., be hopelessly inadequate or require near-infinite computer resources. An inspired approximation may be found that works well enough, and so the underlying physics can be explained. Well enough will usually mean that with good estimates of the model parameters, qualitative and even semi-quantitative agreement is found with key properties of interest.

Each of these are excellent approaches, especially for the purposes they were designed for. Modern DFT calculations of weakly correlated materials (and molecules) are of the first-principles type, and often yield atomic positions within 1-2 hundredths of an Ångstrom and phonon frequencies within 10%, without any materials-specific input, an impossibility with a simple model Hamiltonian. On the other hand, with standard approximations, DFT calculations always fail whenever a bond such as H_2 is stretched, and correlations become strong [41]. Even simple Mott-Hubbard physics is beyond such methods (and we shall see why in this chapter), or Kondo physics (but see Reference [135]).

But more and more of modern materials research requires the intelligent application of both approaches, and many methods, such as DFT+U [125] or dynamical mean field theory (DMFT) [5, 161, 313, 10] are being developed to bridge the gap. Many of the materials

of greatest practical interest to energy research (such as for batteries [112] or photovoltaics [227]) include a moderate level of correlation that require a pure DFT approach to be enhanced, by adding vital missing ingredients of the physics.

The US and Britain are friends 'separated by a common language' [330]. This is essentially true of the mass of confusion between traditional many-body theory and DFT. In DFT, we use the same words as in MBT, but giving them different meanings, simply because we enjoy confusing folks.

Finally, we mention an intermediate Hamiltonian between the dazzling complexity of the real physical and chemical world and the beautiful simplicity of the Hubbard model. A great challenge to studying the effects of strong correlation has been the difficulty in producing highly accurate benchmark data. Molecular electronic structure calculations are much simpler than materials calculations, and quantum chemistry has long been able to provide highly accurate answers for many small molecules at or near equilibrium [9], as well as the complete binding energy curves of others [116]. But this is much harder to do for materials. Recent illustrations of this difficulty are the careful bench-marking of model Hamiltonians (such as an 8×8 Hubbard lattice) using highly accurate many-body solvers [154], the amount of computation needed to find an accurate cohesive energy of the benzene crystal [333], and the celebration of merely being able to agree on approximate DFT results with a variety of solid-state codes [171].

To overcome this difficulty, about 10 years ago, a mimic of realistic electronic structure calculations was established [285]. This mimic uses potentials that are defined continuously in space (i.e., not a lattice model) but are one-dimensional. In fact, ultimately, a single exponential was chosen [8], whose details mimic those of the popular soft Coulomb potential. With about 20 grid points per 'atom', standard density-matrix renormalization (DMRG) methods [328, 329] could then rapidly produce extremely accurate ground-state energies and densities for chains of up to about 100 atoms [285]. By living in 1D, not only is DMRG very

efficient, but the thermodynamic limit (of the number of atoms going to infinity with fixed interatomic spacing) is also reached much more quickly than in 3D. Moreover, the parameters were chosen so that standard density functional approximations, such as the local density approximation [157], succeeded and failed in ways that were qualitatively similar to those in the real world [321]. We will refer to this 1D laboratory for further demonstration of some of the simple results shown in this chapter.

Takeaway: DFT is ideally suited to produce useful accuracy for ground-state energetics of realistic Hamiltonians. Many-body theory is more often used to produce approximate answers to model Hamiltonians, and often focuses on response properties. Both are useful in their own fields and, increasingly, interesting problems require input from both.

2.2.2 Hubbard dimer

The Hubbard model (in 1, 2 or 3D) [130] is the standard model for studying the effects of strong correlation on electrons. By default, it implies an infinite periodic array of sites. For our demonstration, we simply need two sites. We have $N = 2$ and the ground-state is always a singlet. The Hamiltonian (in 2nd quantization) is

$$\hat{H} = -t \sum_{\sigma} \left(\hat{c}^{\dagger}_{1\sigma} \hat{c}_{2\sigma} + h.c. \right) + U \sum_{i} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} + \sum_{i} v_i \hat{n}_i \, . \tag{2.2}$$

The kinetic term is just hopping between the sites, and is the discretization of the kinetic operator on the lattice, with the diagonal elements set to 0. The electron-electron repulsion is just an onsite U, while the one-body operator is just an on-site potential, v_1 and v_2.

In this chapter, we imagine a world in which Eq. (2.1) is replaced by Eq. (2.2), i.e., as if the many-body problem to be solved is simply that of Eq. (2.2). So, for us, the Hubbard dimer

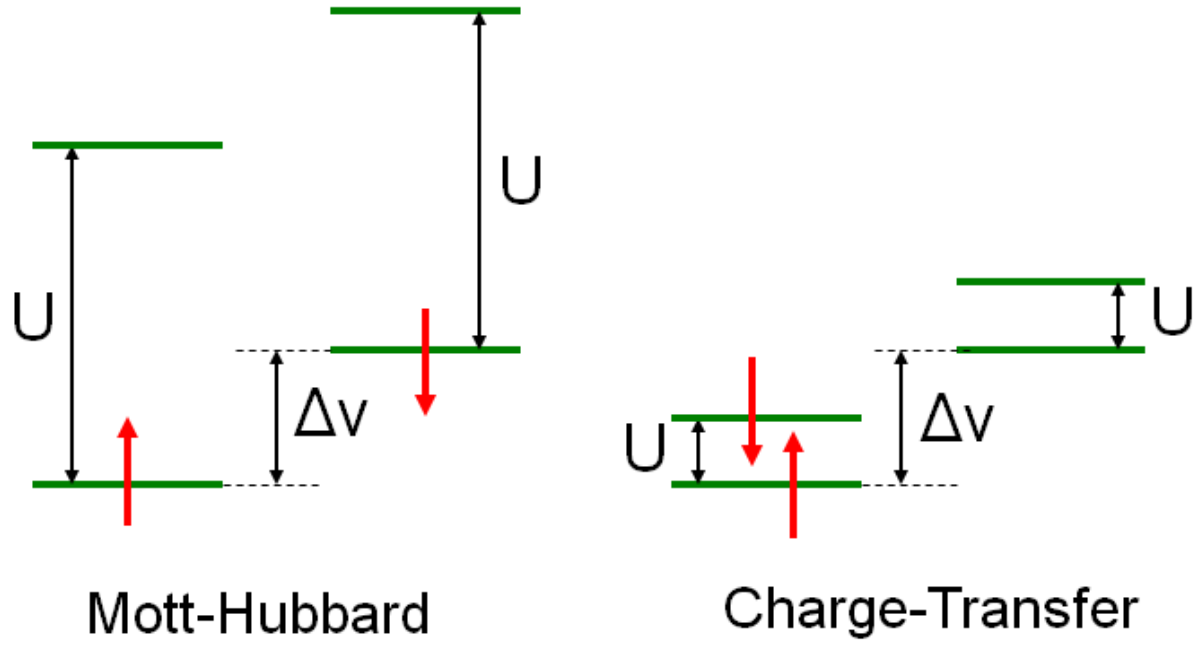

Figure 2.1: Two distinct regimes of the asymmetric Hubbard dimer. On the left, the charging energy is much greater than the difference in on-site potentials, and the left- and right-occupation numbers are similar. On the right, the situation is reversed, and the occupation on the left is much greater than that of the right. Reproduced from Ref. [35].

is *not* an approximation to anything. We will choose the values of U, t, and v_i as we wish, to explore various regimes in the model. Any question concerning the origins of these values in terms of realistic orbitals and matrix elements is irrelevant to our work here.

Since a constant in the potential is just a shift in the energy, we set $v_2 = -v_1$ and use the parameter $\Delta v = v_2 - v_1$ as the sole determinant of the potential of our system. Similarly, with $N = 2$, $n_2 = N - n_1$, and we use $\Delta n = n_2 - n_1$ as the single parameter characterizing the ground-state density. Thus ground-state DFT in this model is simply site-occupation function theory (SOFT) and density functionals are replaced by simple functions of a single variable, Δn. Finally, we choose $t = 1/2$ and report all variables in units of $2t$, as one can scale all energies by a constant.

Different physics appears depending on the ratio of U to Δv, i.e., on-site repulsion versus inhomogeneity, see Fig. 2.1. When $U \gg \Delta v$, the system is strongly correlated, with both

site occupations close to 1, despite any inhomogeneity. For $\Delta v \gg U$, the system is weakly correlated, and the on-site U is insufficient to stop one occupation becoming much greater than the other.

For those with a chemical inclination, this is a minimal basis model for a diatomic with 2 electrons (with some matrix elements and orbital overlap ignored). For H_2, $\Delta v = 0$, but t decreases as the separation between the nuclei is increased, so that U (in units of $2t$) grows exponentially. The ground-state is close to a single Slater determinant near equilibrium ($U \ll 1$), so that Hartree-Fock (HF) is a reasonable approximation. But $U \gg 1$ when very stretched, so that the ground-state is now a Heitler-London wavefunction, and (restricted) HF is very poor. The highly unsymmetric case corresponds to HeH^+, where both electrons reside on the He side, as long as Δv remains larger than U as the bond is stretched.

There are well-known analytic solutions for all states of the 2-site Hubbard model and the behavior of the ground-state energy [35] is shown in Fig. 2.2. Simple limits include the symmetric case

$$E = -\sqrt{1 + (U/2)^2} + U/2, \qquad \Delta n = 0 \qquad \text{SYM} \tag{2.3}$$

An expansion of the square root in the symmetric case in powers of U has a radius of convergence of 2, while the opposite expansion in $1/U$ has a radius of $1/2$. Thus there is a well-defined critical point at $U = 2$, below which perturbation in the electron-electron coupling strength converges, i.e., the system is weakly correlated, and above which it is strongly correlated. Another simple limit is the non-interacting (tight-binding) case ($U = 0$)

$$E = -\sqrt{1 + \Delta v^2}, \qquad \Delta n = -2\,\frac{\Delta v}{\sqrt{1 + \Delta v^2}} \qquad (U{=}0) \tag{2.4}$$

which is given by the blue curve in the figure. We see from the figure that, on a broad

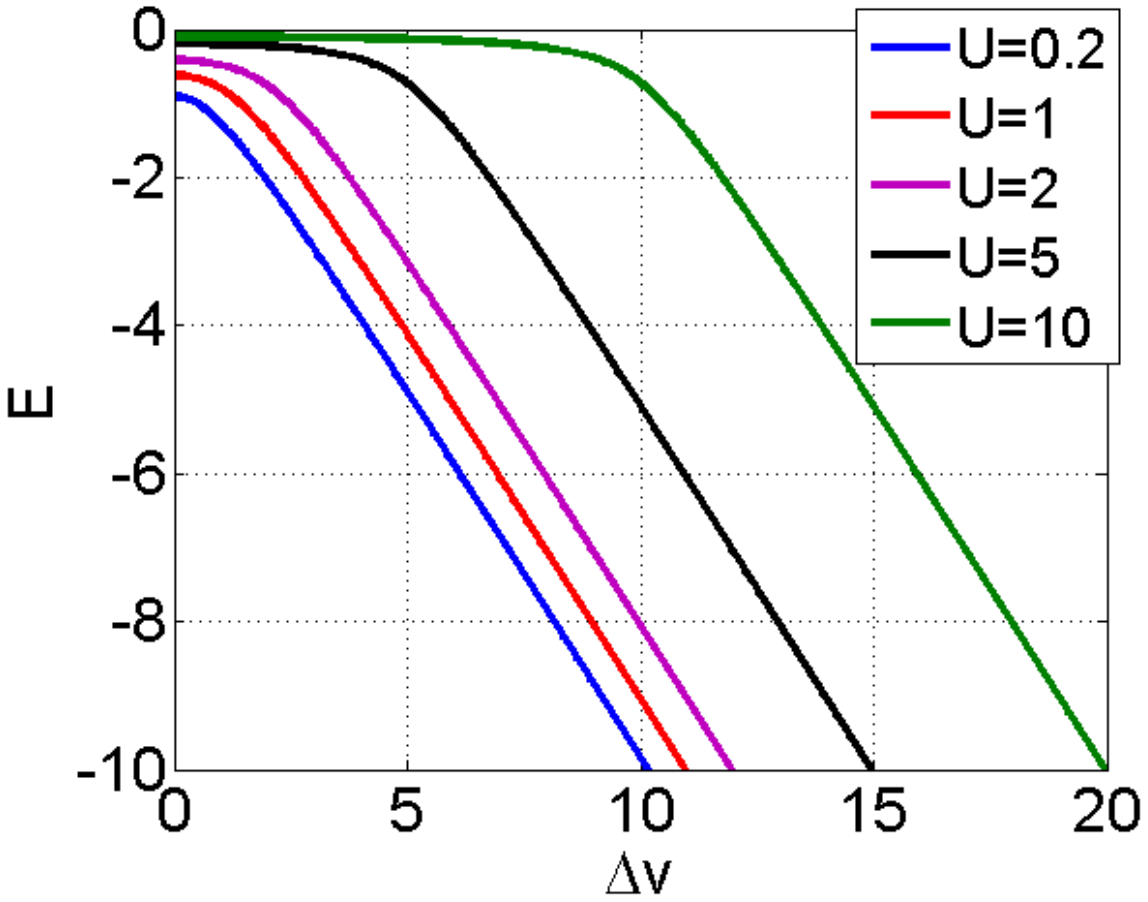

Figure 2.2: Exact ground-state energy of the Hubbard dimer as a function of Δv for several values of U. The qualitative behavior changes as Δv passes through U. Reproduced from Ref. [35].

scale, $E \approx -(\Delta v - U)\,\Theta(\Delta v - U)$. Explicit formulas exist for all the excited-state energies, wavefunctions, and densities also. Approximations in many different limits are given in the many appendices of Reference [35].

We can also extract any other property we wish from the analytic solution, such as the one-electron density (here the occupations). Fig. 2.3 shows the ground-state density as a function of Δv for several values of U. For any U, $n_2 = n_1$ when $\Delta v = 0$. The blue line is essentially the tight-binding solution. In that case, as Δv increases, the occupation difference rapidly increases towards 2. Then, as we turn on U, this increase becomes less and less rapid. By the time U reaches 10, the occupations remain close to balanced until Δv becomes close to

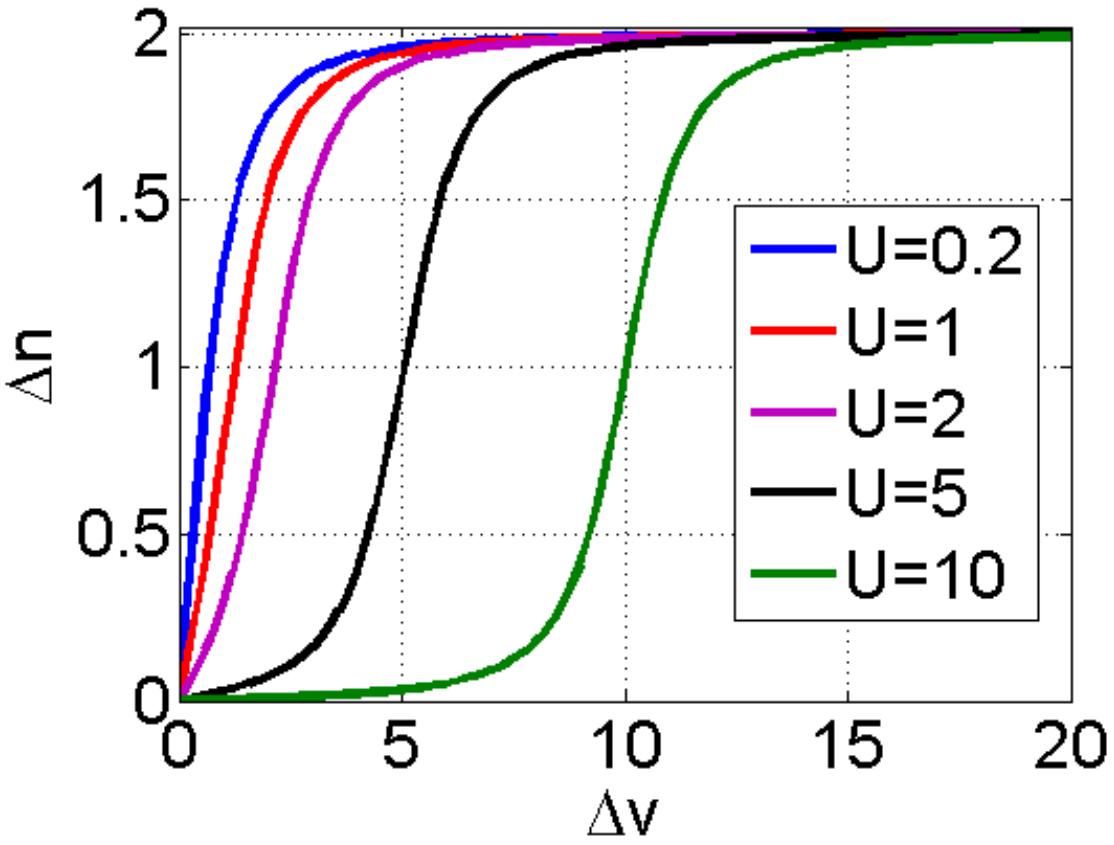

Figure 2.3: Ground-state occupation of the Hubbard dimer as function of Δv for several values of U. Reproduced from Ref. [35].

10, when (on the scale of Δv), it rapidly flips to close to 2.

Takeaway: We take the 2-site Hubbard model as our Hamiltonian, and apply DFT concepts directly to it. Here, it is not a simple model for a more realistic Hamiltonian. Analytic solutions are trivial, and we can plot any properties we wish.

2.3 Density functional theory

We have now defined the machinery required to understand the central theorems of DFT through the lens of the Hubbard dimer. The theorems discussed in this section, like their real-space counterparts, are exact and apply directly to ground-state calculations (we will

cover time-dependent DFT later). Most DFT calculations are used to determine the ground-state electronic energy of a system, or more specifically, determine the energy of a system as a function of nuclear coordinates. In this section, we will discuss the underlying principles of these calculations by examining their role at the most fundamental level, in their simplest form.

The Hohenberg-Kohn theorem[126] is actually three theorems in sequence. These were proved in a simple proof-by-contradiction argument based on the Rayleigh-Ritz variational principle for the wavefunction. Later, the more direct and more general constrained search approach was given by Levy [175] and Lieb [183].

2.3.1 Hohenberg-Kohn I

HKI proves that the (usual) map of $\Delta v \rightarrow \Delta n$ is invertible, i.e., Δn is a single-valued function of Δv for a given U. This is obvious from Fig. 2.3 (and its inversion, Fig. 2.4), and in the TB case

$$\Delta v = \frac{\Delta n}{\sqrt{4 - \Delta n^2}} \qquad (U{=}0). \tag{2.5}$$

Fig. 2.4 is simply Fig. 2.3 drawn sideways, i.e., with x and y axes reversed. Clearly, for any given value of U, there is a unique Δv.

A much-stated (but often out of context) corollary of this is that *all* properties of the system are (implicitly) functionals of n_1. While this is true, almost all research in DFT focuses on the ground-state energy functional, because it is so useful, and we have few useful approximations for others (e.g., for the first excited-state energy, but see discussion in TDDFT section). Recently, machine learning methods have been trained to find some of these other functionals [208, 209].

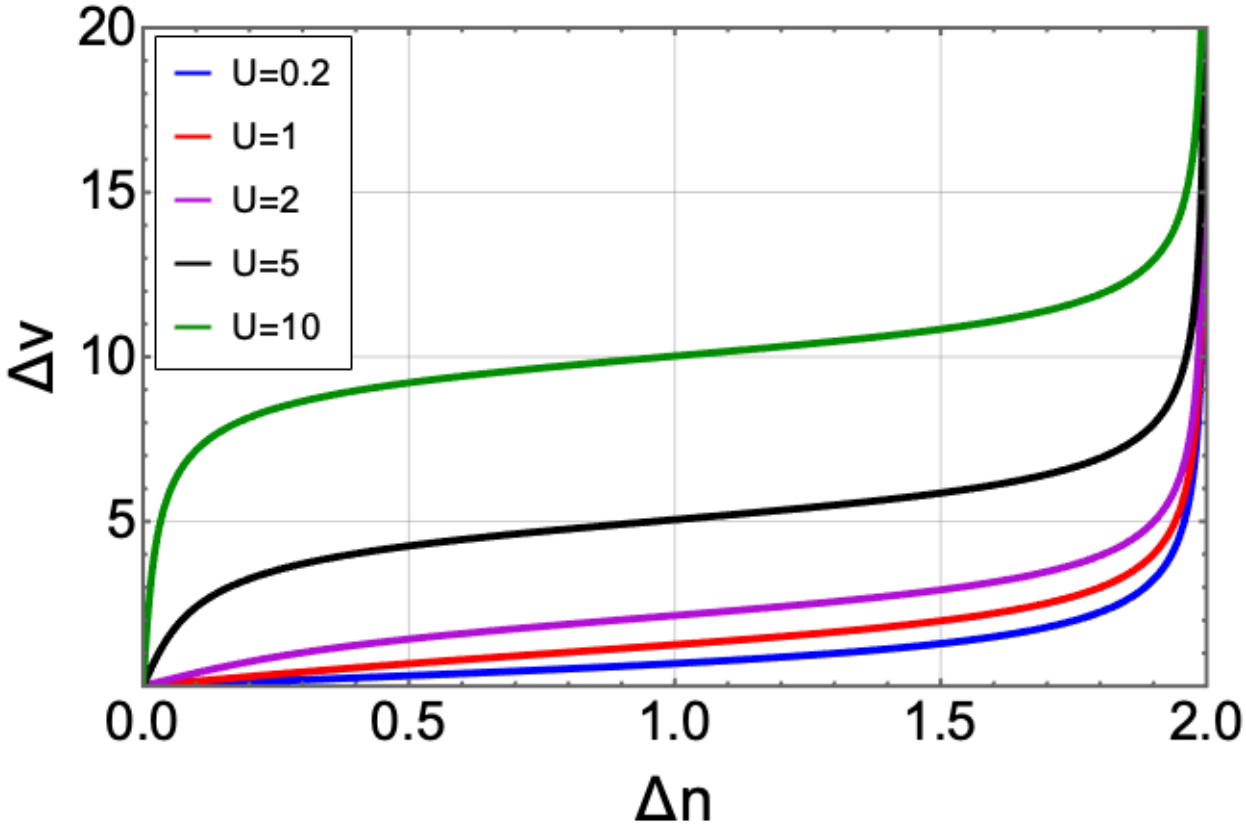

Figure 2.4: Ground-state potential difference as a function of Δn for several values of U.

2.3.2 Hohenberg-Kohn II

HKII states that the function below exists and is independent of Δv:

$$F_{\rm U}(n_1) = \min_{\Psi \to n_1} \langle \Psi | \hat{T} + \hat{V}_{\rm ee} | \Psi \rangle = \max_{\Delta v} \Big\{ E(\Delta v) - \Delta v \Delta n/2 \Big\}. \tag{2.6}$$

where the minimum is over all antisymmetrized normalized 2-electron wavefunctions whose occupation of site 1 is n_1. The middle expression is the constrained search definition due to Levy [174]. The rightmost form is due to Lieb [183]. Either definition works here. This $F_{\rm U}$ functional was termed universal by HK, by which they simply meant that it does not depend on the Δv of your given system, i.e., it is a pure density functional. The phrase, often appearing in the literature, that F is a universal functional, is not meaningful.

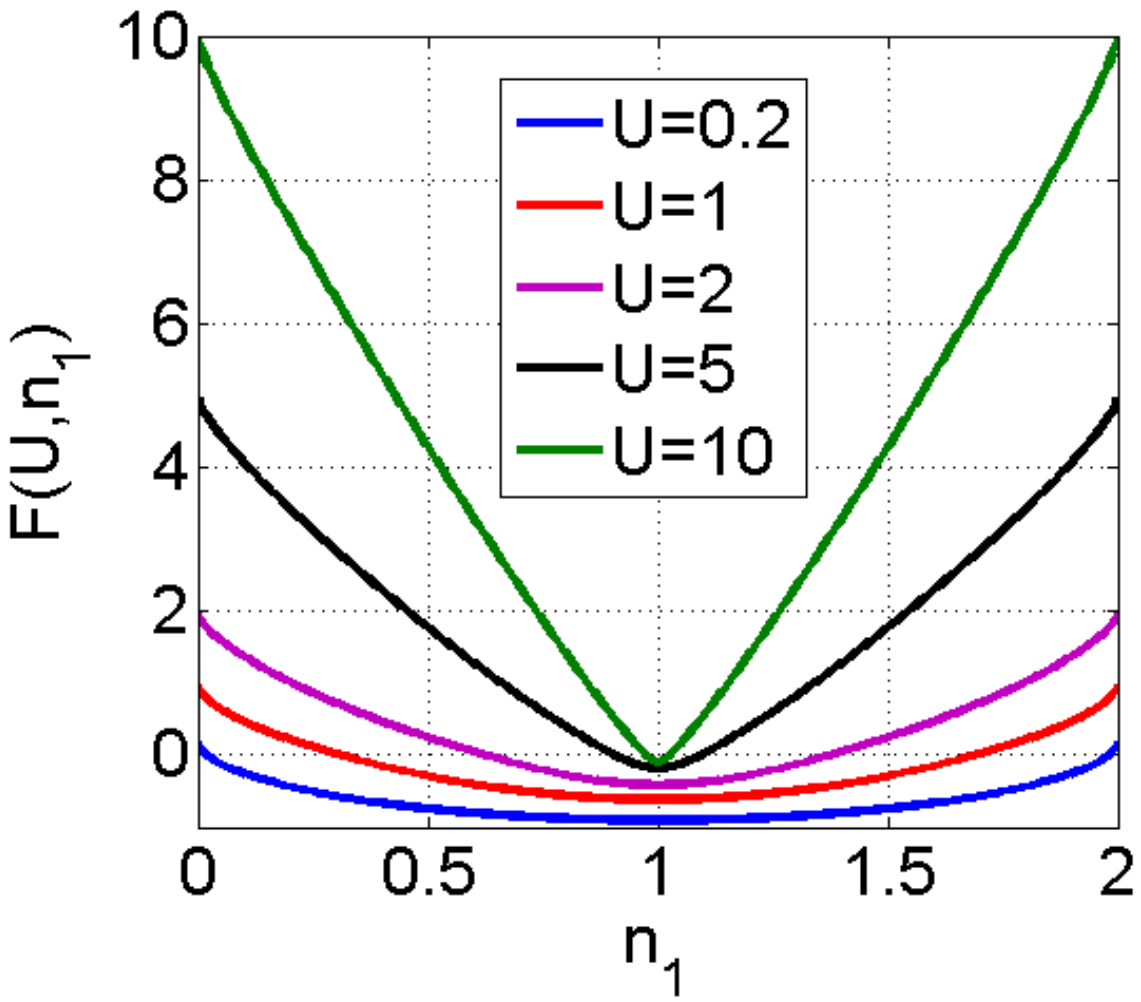

Figure 2.5: Universal part of the energy function(al) of a Hubbard dimer as a function of n_1 for several values of U. As U increases, F tends to $U|1-n_1|$. Reproduced from Ref. [35].

Although one can write analytic formulas for the ground-state energy for the dimer, there is no explicit analytic formula for F. It is trivial to calculate F numerically and F is shown in the Fig. 2.5. In the special case of $U = 0$, it is easy,

$$F_{U=0}(n_1) = T_{\rm s}(n_1) = -\sqrt{n_1(2-n_1)}. \tag{2.7}$$

Here we have attached the subscript S to remind us that $U = 0$, so this is the kinetic energy function for a single Slater determinant, and is indistinguishable from the blue line of Fig. 2.5.

2.3.3 Hohenberg-Kohn III

HKIII states that there is a variational principle for the ground-state energy directly in terms of the density alone:

$$E(\Delta v) = \min_{n_1} \left\{ F_{\mathrm{U}}(n_1) + \Delta v \Delta n/2 \right\}. \tag{2.8}$$

This bypasses all the difficulties of approximating the wavefunction (but of course buries them in the definition of F_{U}). Usually, the minimum can be found from the Euler equation

$$\frac{dF_{\mathrm{U}}(n_1)}{dn_1} - \frac{\Delta v}{2} = 0, \tag{2.9}$$

and the unique $n_1(\Delta v)$ is the one that satisfies this equation.

This allows us to find a solution to the many-body problem, without ever calculating the wavefunction. Given an expression for $F_{\mathrm{U}}(n_1)$, either exact or approximate, for any value of Δv, one can solve Eq. (2.9) above to find the corresponding Δv (exact or approximate) and insert into Eq. (2.8) to find the energy. Any approximation to $F(n_1)$ provides approximate solutions to all many body problems (every value of Δv).

Takeaway: The HK theorems prove the existence of an exact variational principle for the ground-state energy based on the density, not the wavefunction, but give no information on how to approximate it. This is an (almost) useless statement in practice. But to any unbeliever in DFT, one can always tell them (to go look at) F_{U}.

2.4 Kohn-Sham DFT

The original DFT, called Thomas-Fermi theory [296, 67], tried to approximate $F_{\rm U}(n_1)$ directly, but such direct approximations have never been accurate enough for most electronic structure calculations. A tremendous step forward occurred when Kohn and Sham considered a fictitious system of non-interacting fermions with the same ground-state density as the true many-body one [23]. In our case, this is just the TB problem, for which we already have explicit solutions.

They wrote the F function in terms of quantities that could easily be calculated in such a system:

$$F_{\rm U}(n_1) = T_{\rm S}(n_1) + U_{\rm H}(n_1) + E_{\rm XC}(n_1). \tag{2.10}$$

Here, $T_{\rm S}$ is just the TB hopping energy of Eq. (2.4), and the Hartree energy is just the mean-field electron-electron repulsion

$$U_{\rm H} = \frac{U}{2}\big(n_1^2 + n_2^2\big), \tag{2.11}$$

which is an explicit function of the occupations. Then $E_{\rm XC}$, the exchange-correlation (XC) energy (about which, much more, later) is simply everything else, i.e., $E_{\rm XC}$ is defined by Eq. (2.10). It is then trivial to show, from the Euler equation, that the TB potentials that will reproduce the exact occupations are

$$v_{\rm S,i} = v_i + Un_i + \frac{\partial E_{\rm XC}}{\partial n_i}\,. \tag{2.12}$$

The first correction to v_i is the Hartree potential, while the second is the XC potential. These KS TB equations must be solved self-consistently, as the potentials depend on the occupations. Once converged, the final densities can be used to extract the total energy of

the MB system, via

$$E = T_{\rm S} + U_{\rm H} + E_{\rm XC} + V = \epsilon - U_{\rm H} + E_{\rm XC} - \Delta v_{\rm XC}\Delta n/2\,, \tag{2.13}$$

where ϵ is the eigenvalue in the TB KS calculation. Again, just like in the HK case, once $E_{\rm XC}(n_1)$ is given (either approximate or exact), the KS equations can be solved for any electronic system and a ground-state energy and occupation extracted.

The wondrous improvement due to the KS scheme is that only a small fraction of the total energy (the XC part) need be approximated. Many of the most important quantum effects, such as screening, shell structure, binding energies, etc. are mostly accounted for by the quantum effects of the one-body system. Finally, a very simple, intuitive approximation suggested by KS themselves (the local density approximation (LDA) [54, 157]) produced far better results than they expected (but with binding energy errors too large for quantum chemistry taste).

Fig. 2.6 gives us some sense of how this works, for $\Delta v = 1$. Then, if $U = 0$, most occupation is on the left. For $U = 2$, the repulsion makes the occupations more equal. The KS potential is simply that TB potential that produces those (many-body) occupations. So it must be a smaller potential difference than the real potential. One can see that the Hartree potential will typically overestimate repulsion, while XC corrects that to give the exact answer. Finally, when U is ramped up to 5, the occupations become very close to equal, and the KS potential difference becomes very small.

Traditionally, $E_{\rm XC}$ is separated into an exchange and a correlation contribution. The exchange contribution is then defined as

$$E_{\rm X} = \langle\Phi_{\rm S}|\hat{V}_{\rm ee}|\Phi_{\rm S}\rangle - U_{\rm H}, \tag{2.14}$$

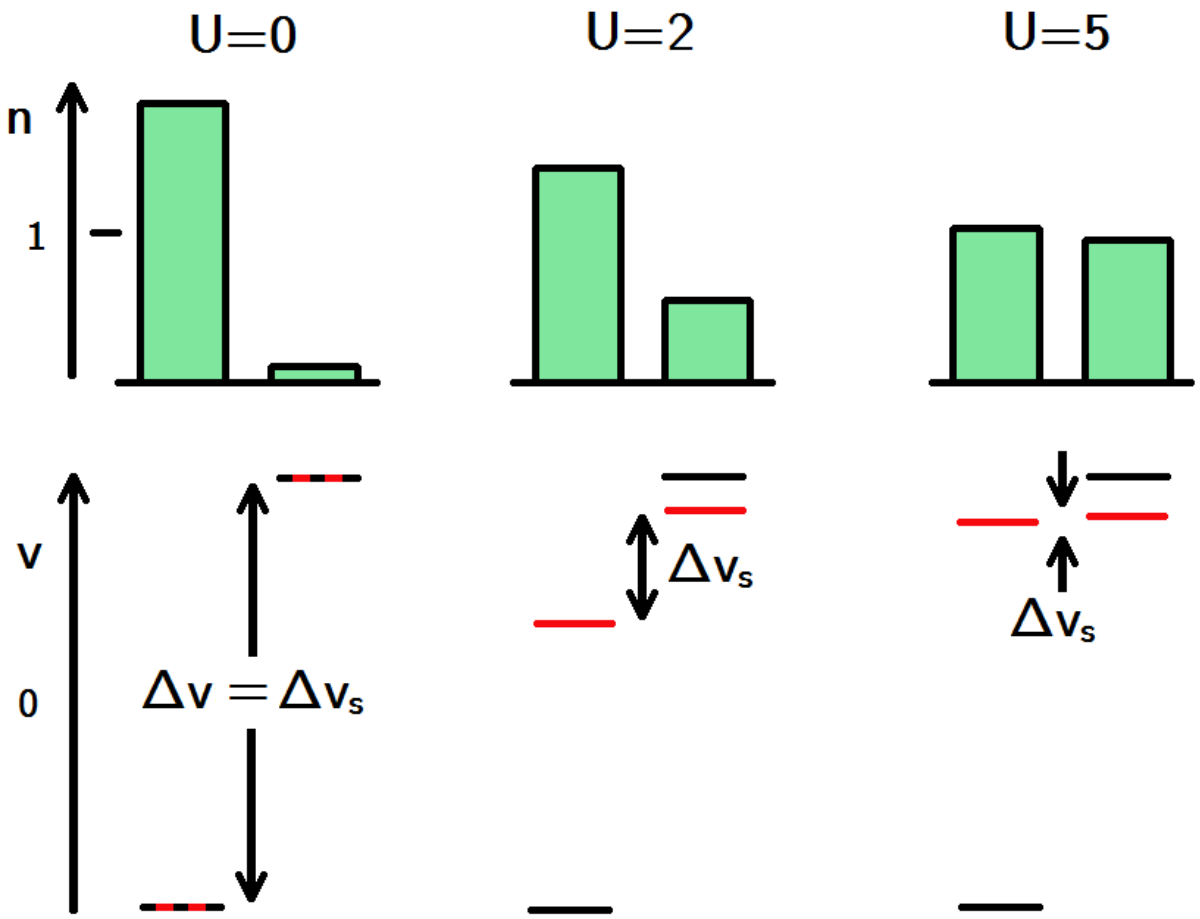

Figure 2.6: KS DFT view of an asymmetric half-filled Hubbard dimer as a function of U. The on-site potential difference Δv is shown in black and the KS on-site potential difference $\Delta v_{\rm s}$ is in red. Reproduced from Ref. [35].

where $\Phi_{\rm s}$ is the KS wavefunction, and $E_{\rm x}$ is always negative. Then one can show correlation is just

$$E_{\rm c} = \langle\Psi|\hat{H}|\Psi\rangle - \langle\Phi_{\rm s}|\hat{H}|\Phi_{\rm s}\rangle \tag{2.15}$$

and, by the variational principle, is also never positive. These definitions (almost) match those of quantum chemistry [306], except that in KS-DFT, all orbitals come from a single potential, while in HF orbitals are freely chosen to minimize the HF energy. But there are some surprises relative to the traditional many-body expansion. For example, because of the definitions, $E_{\rm x}$ includes some 'self-exchange', i.e., it is non-zero even for a single electron

(where $E_{\rm X}$ is $-U_{\rm H}$ and $E_{\rm C} = 0$). DFT approximations which do not satisfy these conditions for all one-electron densities are said to have self-interaction errors [249]. Moreover, 'higher-order exchange effects' are all lumped into the correlation energy. In any event, for our 2-electron problem, in a spin singlet, $E_{\rm X} = -U_{\rm H}/2$, but no simple relation exists for larger N.

The traditional Hartree-Fock approximation comes from expanding the electron-electron interaction to first order, which means neglecting $E_{\rm C}$, and then minimizing the energy. In full DFT terms, for our 2-electron system,

$$F^{\rm HF} = T_{\rm S} + \frac{1}{2}U_{\rm H}, \tag{2.16}$$

or in KS-DFT terms

$$E_{\rm XC}^{\rm HF} = -U_{\rm H}/2\,. \tag{2.17}$$

Thus, solving the TB equation self-consistently with Eq. (2.17) produces the minimum for the total energy using $F^{\rm HF}$ of Eq. (2.16).

In Fig. 2.7, we show the contributions to the KS potential for a sequence of different U values, as a function of the occupation. The effect of repulsion is to always oppose the potential difference, making the KS potential difference smaller. In the first, U is small, and correlation is of order U^2 (see Reference [35]). Thus the correlation contribution is negligible (red and green overlap) and HF is an excellent approximation. In the middle, $U = 1$ is moderate, and now we begin to see the difference correlation makes in the potential. Moreover, its effect is to make $\Delta v_{\rm HXC}$ deviate from a straight line. Finally, for strong correlation, the HXC potential (almost) exactly is equal and opposite to the one-body potential. Again, the HX contribution has much curvature, but now correlation wipes that out (almost) entirely. Clearly, the HF approximation will be terrible for the potential in this case, and yield entirely incorrect

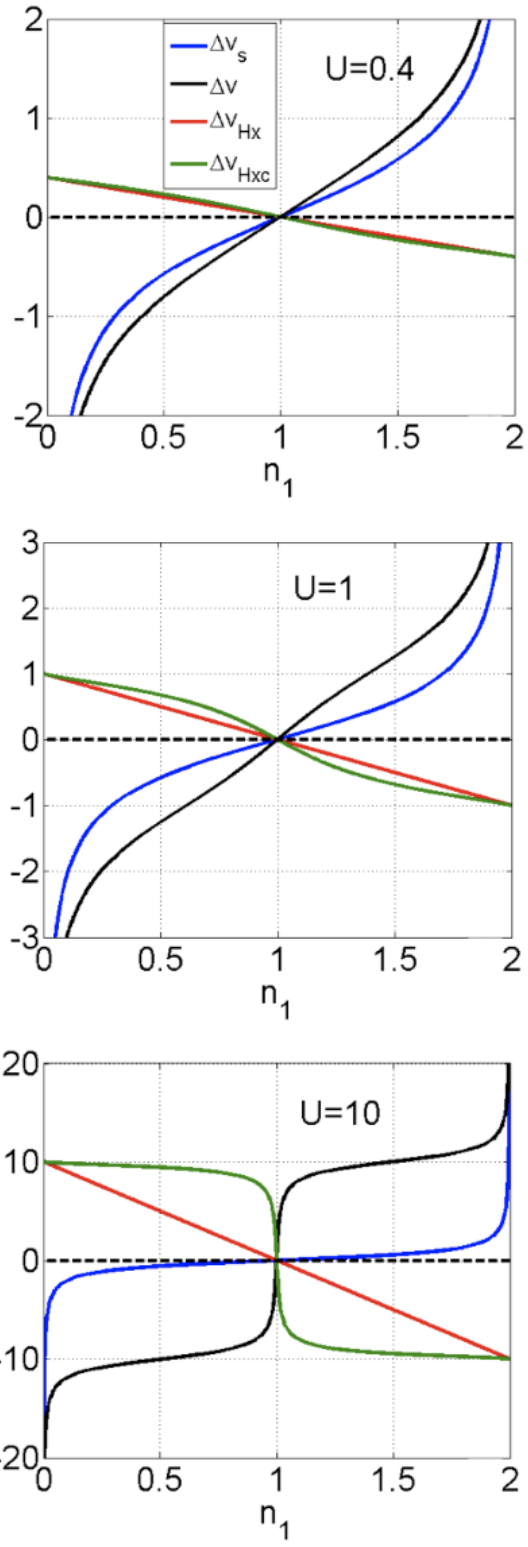

Figure 2.7: Plots of $\Delta v_{\rm S}$ (blue) and its components, the one-body potential Δv (black), the Hartree plus exchange potentials, $U\Delta n/2$ (red), and the same with correlation added, $U\Delta n/2+\Delta v_{\rm C}$ (green) plotted against n_1 for various values of U. Reproduced from Ref. [35].

densities. In fact, a lower-energy solution appears if one allows spin symmetry breaking [246].

It is now relatively routine to calculate accurate KS potentials from highly accurate densities found, e.g., via quantum chemical methods [74]. In an insanely demanding calculation, it is even possible to solve the KS equations using the exact XC functional [320]. Convergence becomes more difficult as correlations grow stronger, but remains possible [323].

Takeaway: The KS scheme is exact meaning that, if we only knew the exact exchange-correlation functional, we could determine the ground-state energy exactly, of every electronic problem. There are many existing calculations of the exact XC potential. In practice, we must approximate XC, but because XC is a small fraction of the total energy, standard KS calculations are usefully accurate for ground-state energies and densities.

2.4.1 KS spectral function

There is a *pernicious superstition* [292] that the KS spectrum is related to the physical response properties of the real system. This false belief has arisen because, for weakly correlated systems, this is approximately true, apart from the fundamental gap of a semiconductor. From a practical viewpoint, the KS bands are marvelously useful as a starting point for Green function calculations of real spectral functions. Moreover, long ago, when the local density approximation ruled supreme, there was no way to know if differences between the KS and exact response properties was due to the crudeness of this approximation [245, 276]. These days, there are simple exact answers to such speculations, if we only have the patience to read them.

2.4.2 The ionization potential theorem

As a simple example of the mysterious workings of the exact functional, we state an important exact result

$$I(N) = E(N-1) - E(N) = -\epsilon^{\mathrm{HO}}(N). \tag{2.18}$$

Here $E(N)$ is the ground-state of the N-electron system, and ϵ^{HO} is the energy of the highest-occupied KS orbital. (For those with some chemistry leaning, Koopmans' theorem is an approximate version of this for HF calculations [230]). This illustrates some of the power of KS-DFT. You might think that, with the exact functional, all one can extract is the ground-state energy and density of our system. But the above result shows that the HO of the KS scheme also tells you the ionization energy. One can also extract all static response functions exactly by turning on weak external perturbations, and applying the exact functional to the perturbed systems. In practice, standard DFT approximations tend to violate this exact condition very badly [245, 276, 244]. Nonetheless, they often still yield usefully accurate ground-state energies, thus performing their primary function. (On the other hand, returning to the discussion of HKI, knowing the exact XC does *not*, in general, give you access to, say, the first excited state energy. It is a functional of n_1 alone, but we cannot deduce that functional from $E_{\mathrm{XC}}(n_1)$.)

Increasing N by 1 in Eq. (2.18) yields

$$A(N) = E(N) - E(N+1) = -\epsilon^{\mathrm{HO}}(N+1) \neq \epsilon^{\mathrm{LU}}(N), \tag{2.19}$$

where A is called the electron affinity of the system in chemistry. The difference between the KS HO of the $N+1$ electron system and the lowest unoccupied (LU) level of the N-electron system is called Δ_{XC}, where the Δ indicates its origin from the infamous derivative

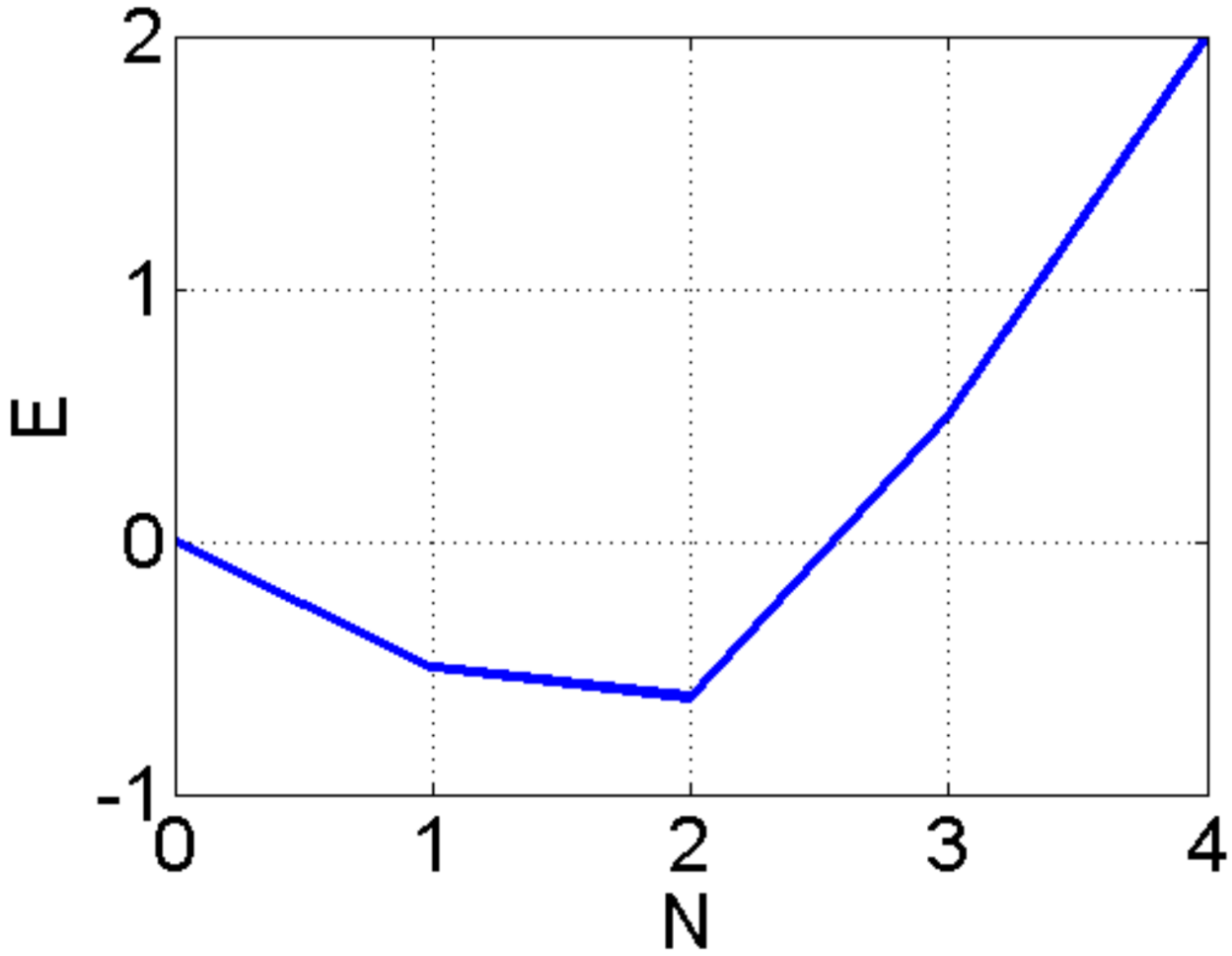

Figure 2.8: Plot of $E(N)$ for $U = 1$ and $\Delta v = 0$. Reproduced from Ref. [35].

discontinuity of DFT [245]. This simply means, that at zero temperature, the energy of the system consists of straight line segments between integer values, as shown below in Fig. 2.8. The energy itself is continuous, but its derivative, the chemical potential, is not. For a neutral system, the chemical potential is $-I$ below the integer and $-A$ above. This discontinuous jump in μ shifts the KS HO eigenvalue by the same amount, producing the difference with the KS LU of the neutral. (Realistic electronic systems do not have an upward pointing portion of the curve in Fig. 2.8. This occurs for the dimer because electrons cannot escape to outside the system.)

2.4.3 Mind the gap

We are now ready to see the relevance of this to solids. Even for a finite system, we define the charge (or fundamental) gap as

$$E_g = I - A\,. \tag{2.20}$$

As the size of the system grows toward a bulk material, this quantity tends to the fundamental charge (or transport) gap of the system (at least for ordered systems [156]). But, because of Eqs. (2.18) and (2.19) above, we find

$$E_g = E_{s,g} + \Delta_{\rm XC}\,, \tag{2.21}$$

where$E_{s,g}$ is the KS gap (i.e., the difference between the LU and HO level, or the gap between the KS valence and conduction bands in a solid). Thus, with the *exact* XC functional of ground-state DFT, we do not get the true gap by looking at its KS value for the neutral system.

Fig. 2.9 shows the spectral function (projected onto the left-hand site) in a weakly correlated case [33], the symmetric dimer with $U = 1$. We can see the sense in which the KS spectral function (red) resembles the blue exact one: the significant KS peaks are of about the same height and position as their blue counterparts, and the blue peaks without KS counterparts are relatively small. The KS gap is smaller than the true gap, but not by much. Because both the KS and the exact spectral functions satisfy the same sum rule (even with an approximate XC), if the dominant peaks are reproduced (even with the wrong gap), only small peaks are missed in the KS spectrum.

On the other hand, Fig. 2.10 shows the same system with a larger U value. Now the strong KS peaks are not in the right place and are noticeably too large. Moreover, the blue peaks

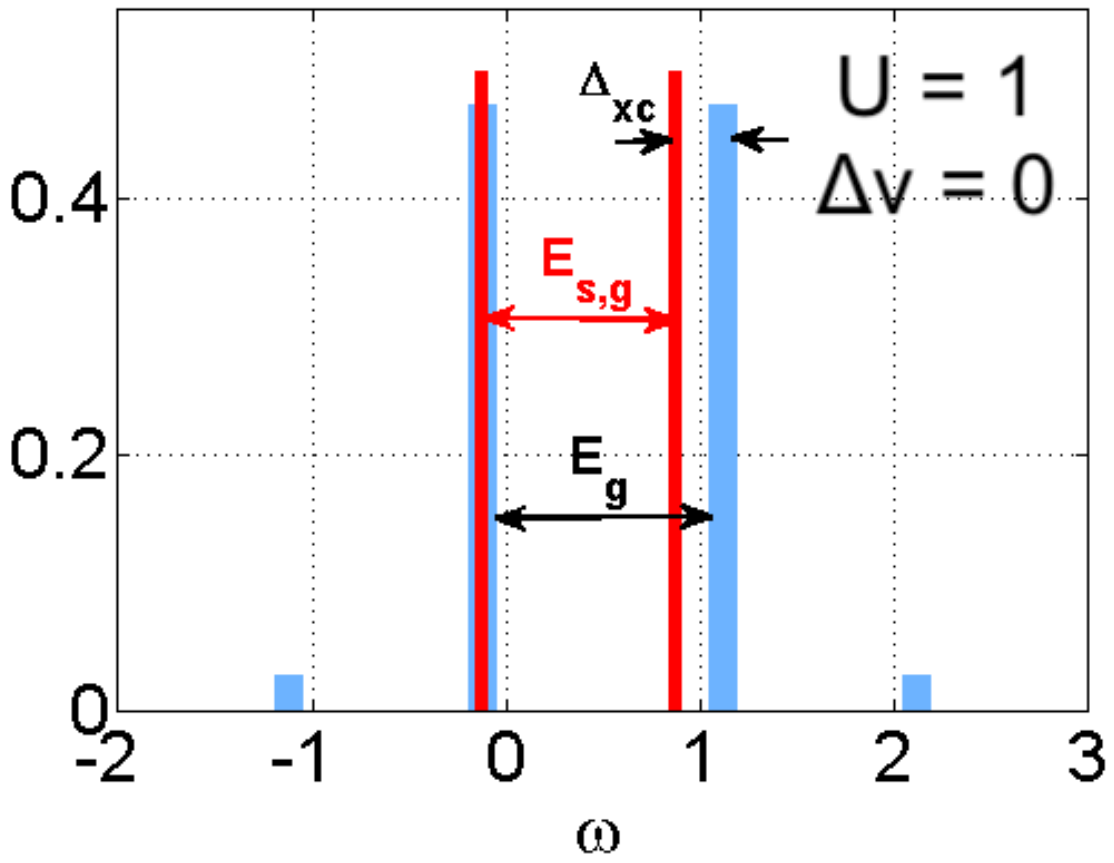

Figure 2.9: Spectral function of the symmetric dimer for $U = 1$ and $\Delta v = 0$. The physical MB peaks are plotted in blue, the KS in red. Here $I = 0.1$, $A = -1.1$, and $\epsilon^{\mathrm{LU}} = 0.9$. Reproduced from Ref. [35].

with no KS analogs are a substantial contribution. Finally, in the inhomogeneous case, the potential asymmetry overcomes the effects of the Hubbard U. In Fig. 2.11, we see that for $\Delta v = 2$ and $U = 1$, the KS spectral function is almost identical to the true one.

Lastly, we finish this section illustrating the relevance of this discussion to the thermodynamic limit. The canonical example of the Mott-Hubbard transition is a chain (or lattice) of H atoms. Each atom has one electron, so the bands of the KS potential are always half-filled, with no gap at the Fermi energy. Thus the gap is always zero and the KS band structure suggests it's a metal. This may be true at moderate separations of the atoms, but as the separation is increased, the electrons must localize on atoms, and it must become a Mott

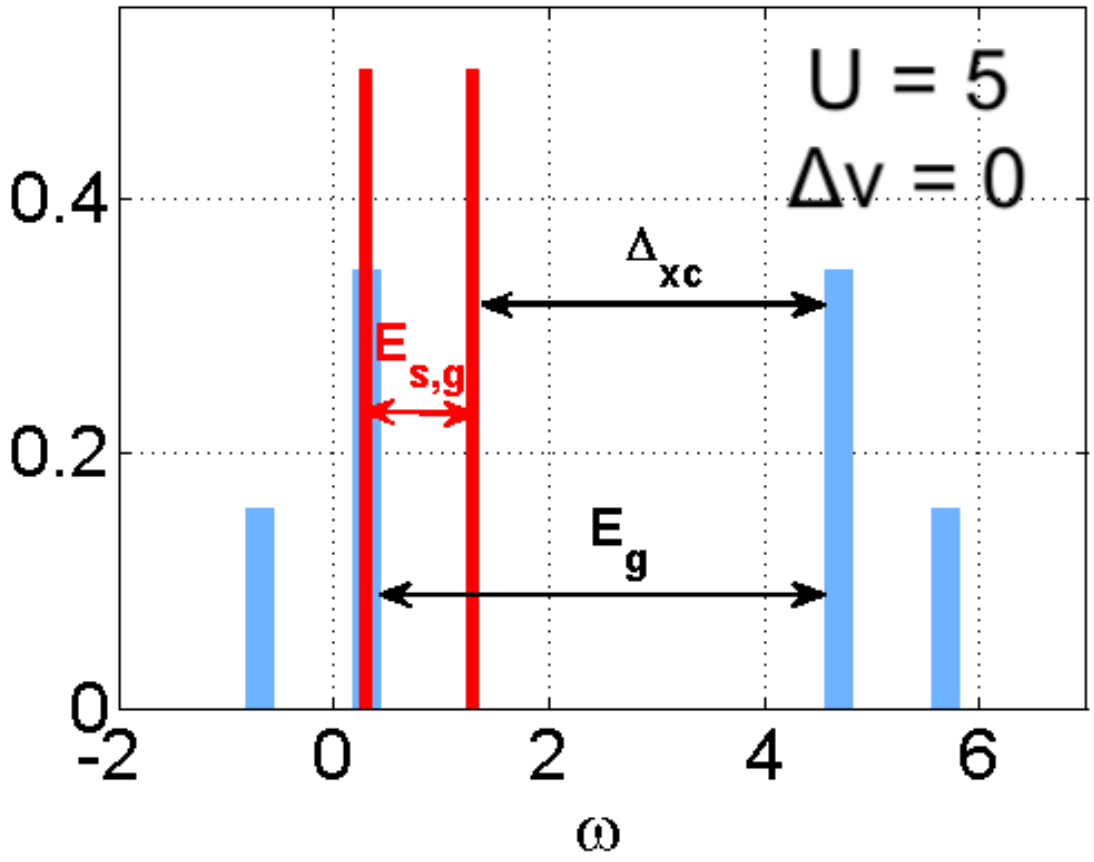

Figure 2.10: Same as Fig. 2.9, but now $U = 5$. Here $I = -0.3$, $A = -4.7$, and $\epsilon^{\mathrm{LU}} = 1.3$. Note that the KS gap remains unchanged by the alteration of U because $\Delta n = 0$ in both cases. Reproduced from Ref. [35].

insulator.

Fig. 2.12 shows the gap, calculated for chains of well-separated 1D H atoms of increasing length [285]. By performing the calculation with finite systems, i.e., without periodic boundary conditions, we calculate the gap for each N by adding and removing electrons, as in Eq. (2.20), and then take the limit as $N \to \infty$. On the other hand, we extract the exact ground-state density from our DMRG calculation at each N, and find the corresponding exact KS potential for each N. We could then as easily extrapolate the KS gap, from the HO and LU, showing that indeed the KS gap vanishes in the thermodynamic limit – exactly the same as if we had calculated the KS band structure, in which the Fermi energy

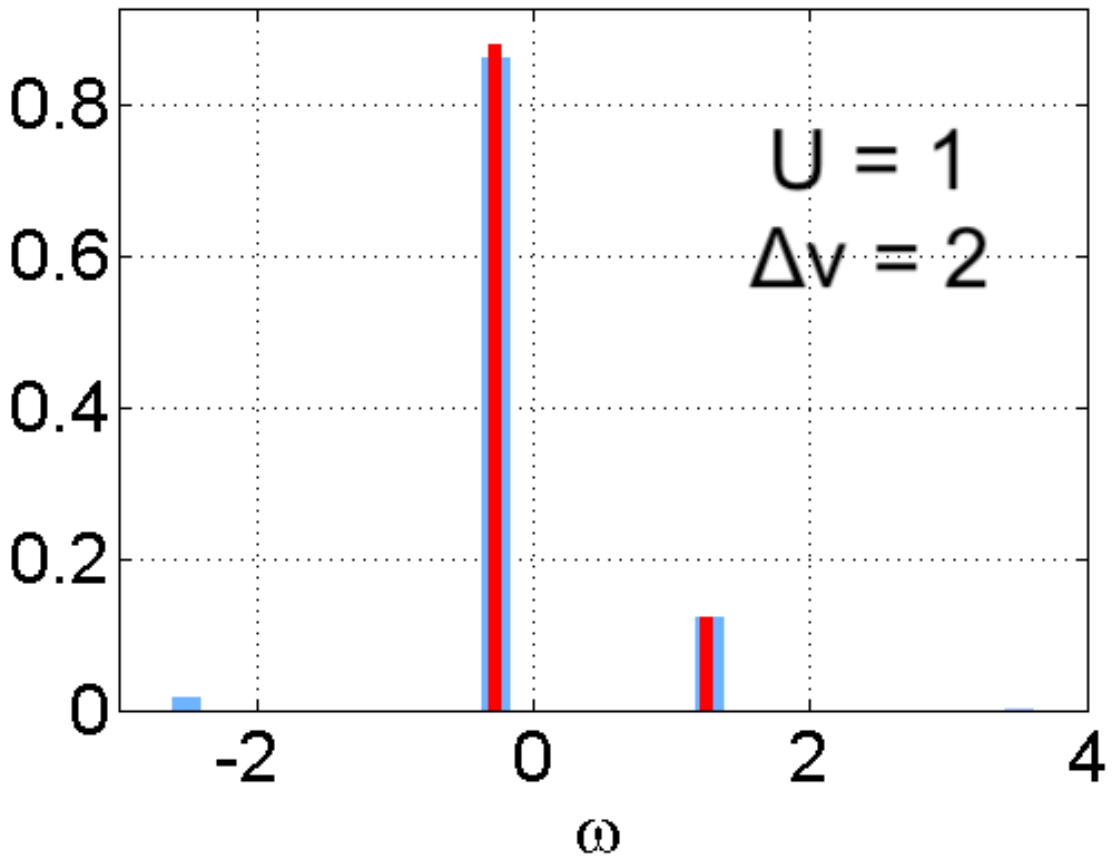

Figure 2.11: Same as Fig. 2.9, but now $U = 1$, $\Delta v = 2$. Here $I = 0.27$, $A = -1.27$, and $\epsilon^{\mathrm{LU}} = 1.25$. Reproduced from Ref. [35].

would be right in the middle of the band. This provides a dramatic illustration of the KS underestimate of the true gap, even when using the exact XC functional.

Takeaway: The KS Green function does not match the true Green function. If correlation is weak, it may be a good approximation to it, with its main deficiency being an underestimate of the gap. For stronger correlations, there can be huge differences, and there are always more features in the real Green function. In the thermodynamic limit, the exact KS gap can vanish for a simple Mott insulator.

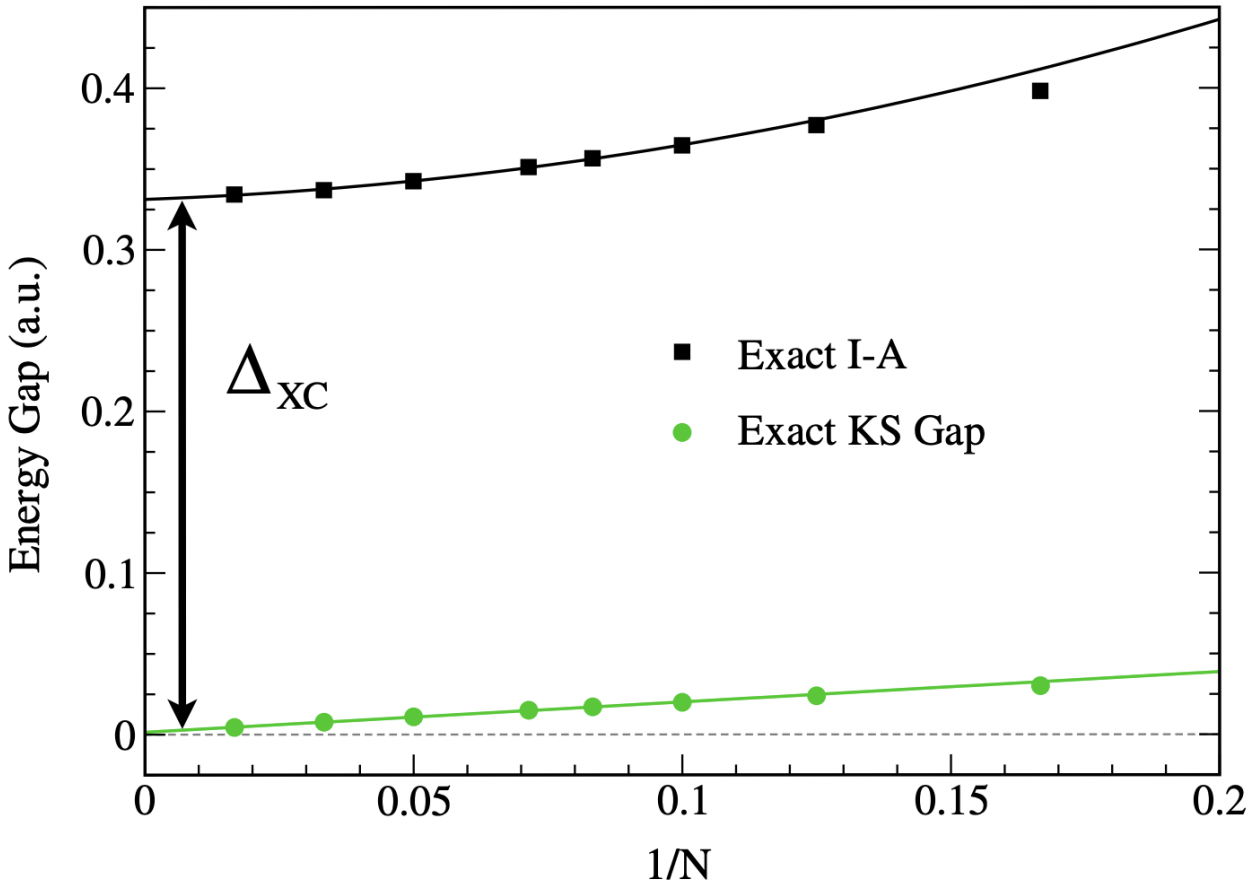

Figure 2.12: Exact gaps for chains of N soft hydrogen atoms with atomic separation $b = 4$ (error bars are less than symbol sizes). The upper curve is a quadratic fit of exact gaps of the largest six systems and extrapolates to a finite value $E_g \approx 0.33$. The exact Kohn-Sham gaps, in contrast, extrapolate to zero showing that for $N \to \infty$ the true KS system is metallic (lower curve is a linear fit of exact KS gaps of the largest six systems). Reproduced from Reference [285].

2.4.4 Talking about ground-state DFT

First, we review our crucial formal points.

1. In general, the KS scheme with the exact functional yields ground-state energy and density, and any other quantities that can be teased from them, such as static response properties and ionization potentials.

2. There is no formal meaning for most KS eigenvalues in ground-state DFT, despite the

fact that many practitioners treat them as if there were. Of course, they do provide tremendous physical and intuitive insight, especially for weakly correlated systems, where they are good approximations to the excitations (either quasi-particle or optical). But when correlations are strong, explicit methods are needed to correct them [342].

3. The strongest manifestation of point 2 above is that the exact KS gap is typically smaller than the true gap, and can vanish in cases where the true gap is finite (Mott insulator).

4. Moreover, there is an exact formula relating the total energy to the sum of the KS eigenvalues, which contains finite corrections for double counting. There is no ambiguity about these corrections, they are derived from the formal theory, and yield the exact many-body energy. But when correlated methods are used for a *subset* of the orbitals, ambiguities can arise that affect occupancies [203].

5. Although in principle, all properties are functionals of the ground-state density, knowledge of the exact ground-state energy functional (via $E_{\rm XC}$) does not provide a way to calculate these other functionals. As we see later, TDDFT is a way to do precisely this.

Next, we discuss how these points show up in practical DFT calculations of solids, where XC approximations must be made.

1. The steady progress within quantum chemistry and materials in functional development is almost entirely focused on improvements in accuracy and reliability of the total energy for weakly correlated systems [319, 114]. This is by far the most important use of DFT in modern electronic structure. Such improvements are often not particularly relevant to the response properties of greatest interest in strongly correlated materials. For example, the KS eigenvalues are often not improved significantly by functionals

yielding better energies [286, 299]. Although the KS eigenvalues cannot be directly interpreted in general, they are uniquely defined (up to a constant). Thus the exact KS Hamiltonian is a well-defined starting point for many-body methods.

2. The KS scheme is not a mean-field scheme in the traditional sense of the word, and it can be extremely difficult to relate its features to those of traditional many-body theory. The KS wavefunction is typically a single Slater determinant, but yields the exact many-body energy via its density.

3. Standard approximations, such as LDA and generalized gradient approximations (GGA), by construction produce total energies that are smooth and continuous at integer N, unlike the exact $E(N)$. Thus their corresponding Δ_{xc} is zero[245]. According to Sec. 2.4.3, the KS band gap in such approximations *is* their prediction for the fundamental gap. In fact, it has been found that their KS gaps are likely a good approximation to the exact KS gap [193], but their lack of discontinuous behavior means they miss the correction to turn it into the true gap.

4. On the other hand, the range-separated hybrid functional HSE06 is well-known to produce reasonable gaps for moderate gap semiconductors. This is because, instead of performing a true pure KS calculation, most codes (like VASP) perform a *generalized* KS calculation [273] when a functional is orbital-dependent[111, 248]. They treat the orbital-dependent part of the potential as if it were a many-body potential, just as is done in HF. (A similar but smaller effect occurs in meta-GGA's that depend on the kinetic energy density, such as SCAN [287]). And in fact clever tricks may be used to extract the true gap, even from a periodic code [86].

Takeaway: Even with the exact functional, the KS band gap does not equal the true transport gap of the system. Likely, semilocal functionals yield accurate KS gaps but, because they lack a discontinuous behavior at integer particle numbers, cannot yield accurate transport gaps. Modern hybrid functionals that depend explicitly on KS orbitals yield band gaps closer to fundamental gaps, but only when treated with generalized KS theory.

2.5 Time-dependent DFT (TDDFT)

Our last main section is about time-dependent density functional theory (TDDFT) [30, 200, 304, 195]. While this uses many of the forms and conventions of ground-state DFT, it is in fact based on a very different theorem from the HK theorems. When applied to the linear response of a system to a dynamic electric field, it yields the optical transitions (and oscillator strengths) of that system. It has become the standard method for extracting low-level excitations in molecules, where traditional quantum chemical calculations are even more demanding than those for the ground state.

The Runge-Gross theorem [264] states that, for a given initial wavefunction, statistics, and interaction, the *time-dependent* density uniquely determines the one-body potential. In principle, this can be used for any many-electron time-dependent problem, including those in strong laser fields [30]. In practice, such calculations are limited by the accuracy of the approximations and whether the observable of interest can be extracted directly from the one-electron density. One constructs TD KS equations, defined to yield the exact time-dependent one-electron density. Because TDDFT applies to the time-dependent Schrödinger equation, the XC functional differs from that of ground-state DFT in general, and has a time-dependence.

Our interest will be only in the linear-response regime. In that case, one can derive a crucial result, which we give in operator form, called the Gross-Kohn equation [102]

$$\chi(\omega) = \chi_{\rm S}(\omega) + \chi_{\rm S}(\omega) * (f_{\rm H} + f_{\rm XC}(\omega)) * \chi(\omega), \tag{2.22}$$

where $\chi(\omega)$ is the dynamic density-density response function of the system, and $\chi_{\rm S}$ is its KS counterpart. The kernel, f, is the functional derivative of the time-dependent potential. Thus, $f_{\rm H}$ is the Hartree contribution, while $f_{\rm XC}(\omega)$ is the XC correction.

Eq. (2.22) is a Dyson-like equation for the polarization. If we set $f_{\rm XC} = 0$, it is the standard random-phase approximation, the Coulomb interaction simply dressing the bare interaction, and producing all the bubble diagrams. But things get a little weird when we assert that inclusion of $f_{\rm XC}(\omega)$ produces the *exact* response of the system, for all frequencies. From a many-body viewpoint, this is suspicious, as these are a closed set of equations without coupling to 4-point functions. But the logic is sound and exactly analogous to the ground-state: there exists such a function that could be considered as defined by Eq. (2.22).

The excitations of a system are given by poles of its response function. Simple analysis (exactly that of RPA) yields a matrix equation that corrects KS transition frequencies to the true transition frequencies, where the matrix elements involve $f_{\rm H} + f_{\rm XC}$. With standard ground-state approximations, folks have merrily calculated mostly low-lying valence transitions from the ground-state of many molecules [3], finding accuracies a little lower than those of ground-state DFT [136], and computational costs that are comparable. This has been invaluable for larger molecules, where many excitations of the same symmetry may overlap, and so TDDFT yields a semiquantitative signature that can be easily matched with experiment [259].

However, not all is well in paradise. Almost immediately, it was noticed that the use of a ground-state approximation is simply the static limit of the corresponding kernel, and can be

easily shown to produce only single excitations. While useful workarounds were created for some cases, it was also found that going to higher-order response does not solve the problem. And many of the most exciting transitions in biochemistry are double excitations.

Takeaway: Time-dependent DFT applies DFT methods to time-dependent problems. Within linear response, this yields exact expressions for the dynamic polarization, but at the cost of introducing a new functional, the frequency-dependent XC kernel. Ignoring its frequency dependence yields useful accuracy for low-lying molecular excitations with standard functionals. TDDFT is now standard for calculating optical response of molecules and materials.

2.5.1 Hubbard dimer

Happily we care only about Hubbard dimers, where everything is much simpler. First, we note our Hubbard dimer, in the singlet space, has just three states: the ground-state, the first excited state, which has a single excitation, and the second excited state, which is a double excitation out of the ground-state. Since there are no spatial degrees of freedom, our $\chi(\omega)$ is the Fourier transform of $\Delta n(t)/\Delta v(0)$, which is just a scalar, with ω-dependence

$$\chi_n(\omega) = \frac{a_1}{\omega^2 - \omega_1^2} + \frac{a_2}{\omega^2 - \omega_2^2}, \tag{2.23}$$

where ω_i denotes the transition frequency and a_i is related to its oscillator strength [33]. Thus χ has poles at each of the transition frequencies. Fig. 2.13 shows the value of each of these transitions as a function of Δv for $U = 1$. The double excitation is a little above the single for the symmetric case, but grows linearly with Δv. The single remains about the same, and even dips, until $\Delta v = U$, and then begins to grow itself. Here we can use our model system to examine one of the key mysteries of practical TDDFT: Where did all the higher excitations go?

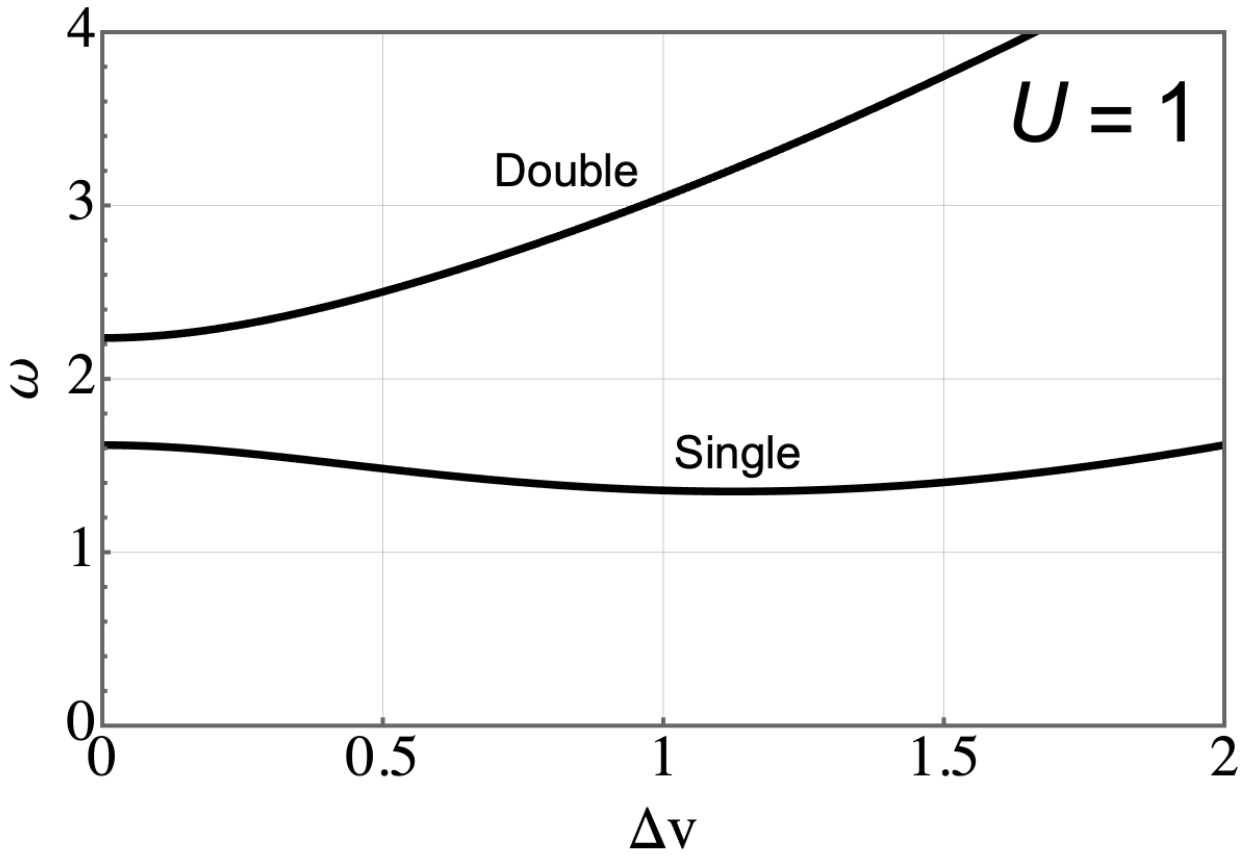

Figure 2.13: Transition frequencies of the first and second excitations as a function of Δv for $U = 1$.

First we do an exact ground-state KS calculation, as in the previous sections. Thus the exact KS system is a tight-binding problem with effective potential, $\Delta v_{\rm s}$, defined to yield the exact ground state Δn. This yields two eigenvalues, the lower symmetric combination and the higher asymmetric combination. The KS ground-state has the lower one doubly occupied. There do exist KS analogs of the many-body states. The single excitation has one electron excited to the higher level, the double has both. Fig. 2.14 adds the KS transitions to Fig. 2.13, showing that they loosely follow the accurate transitions, but are significantly different.

In the KS response function, $\chi_{\rm s}$, the matrix elements of the density operator between ground and double excitation are zero, since both KS orbitals are different, so the Slater determinants are not coupled by a single density operator. Hence, such states have no numerator,

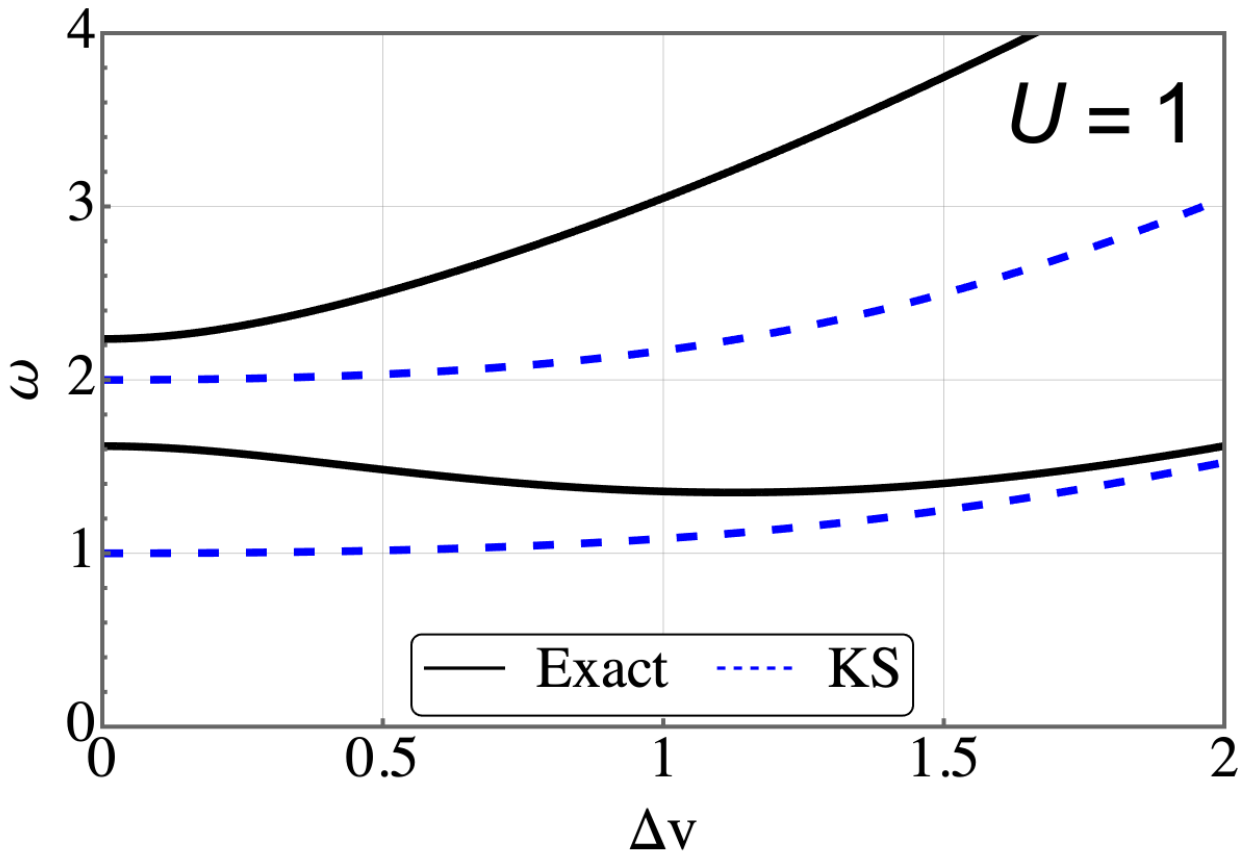

Figure 2.14: Same as Fig. 2.13, but with KS transitions (depicted in blue). For $\Delta v > U$, the KS transition is a very good approximation to the true transition.

eliminating any poles that might have arisen in the denominator, i.e.,

$$\chi_{\rm S}(\omega) = \frac{a_s}{\omega^2 - \omega_s^2}\,. \tag{2.24}$$

Thus the second KS transition, the double, does not appear at all in the response function! It's position is correctly marked in Fig. 2.14, but cannot be seen in $\chi_{\rm S}$.

By requiring the poles occur at the right places, one finds (in general) a matrix equation in the space of single excitations for the true transitions, whose elements are determined by the kernel. Here, this is one dimensional, yielding

$$\omega^2 = \omega_{\rm S}^2 + 2\omega_{\rm S}\, f_{\rm HXC}(\omega)\, \frac{2}{1 + \Delta v_{\rm S}^2}\,. \tag{2.25}$$

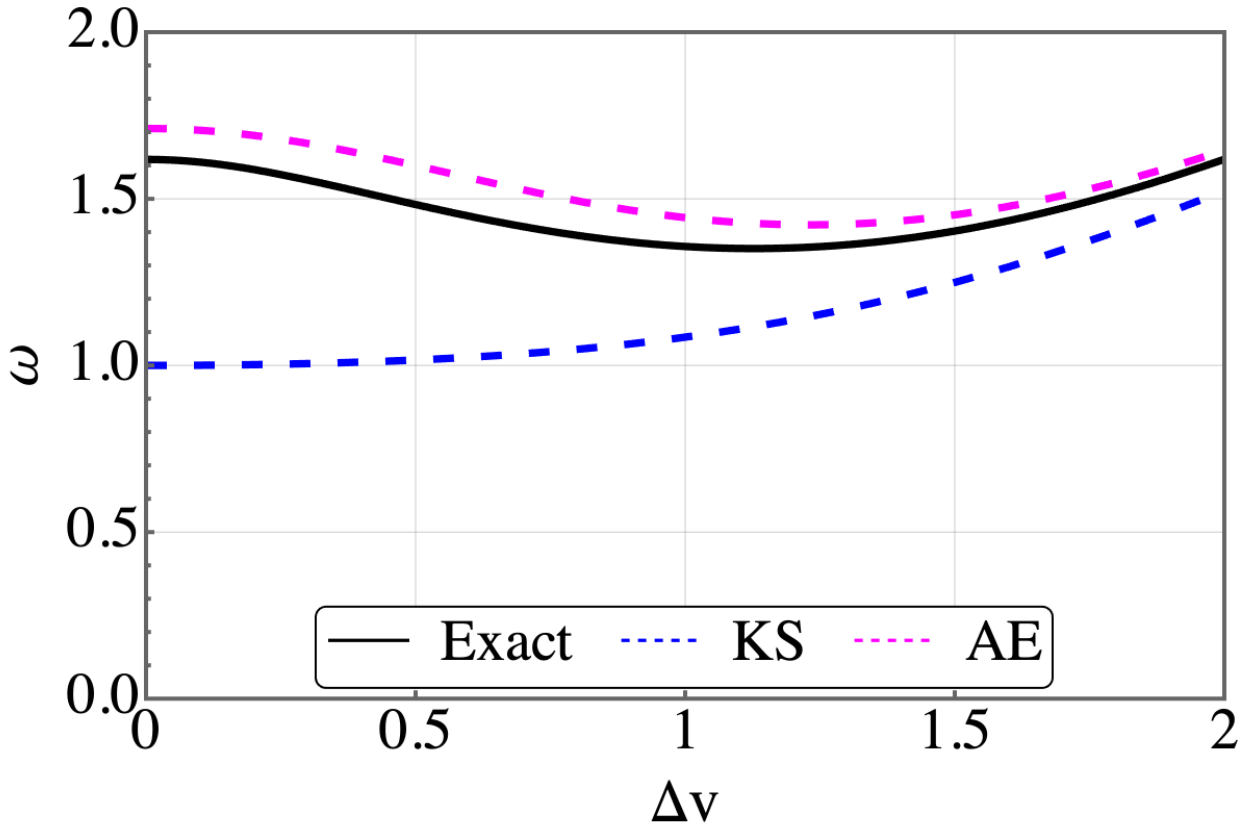

Figure 2.15: Same as Fig. 2.13, but with the adiabatically exact approximation (AE, pink dashes).

The adiabatically exact approximation (AE) is to use the exact ground-state functional here to calculate $f_{\rm HXC}$. This corrects the single KS transition and is shown in Fig. 2.15. This works extremely well to capture almost all the difference with the KS transition, yielding very accurate excitations. This becomes even better for Δv greater than U, where the corrections virtually vanish (just as in Fig. 2.11 for the spectral function).

But Eq. (2.25) just has one solution if the ω-dependence in the kernel is neglected. On the other hand, if there is strong frequency dependence in the kernel, new transitions, not in the KS system, may appear. In fact, we know that is precisely what happens, as the physical system *does* have a double excitation. To understand how standard TDDFT fails, we note that we can calculate the exact kernel by finding $\chi(\omega)$ from many-body calculations, $\chi_{\rm S}(\omega)$

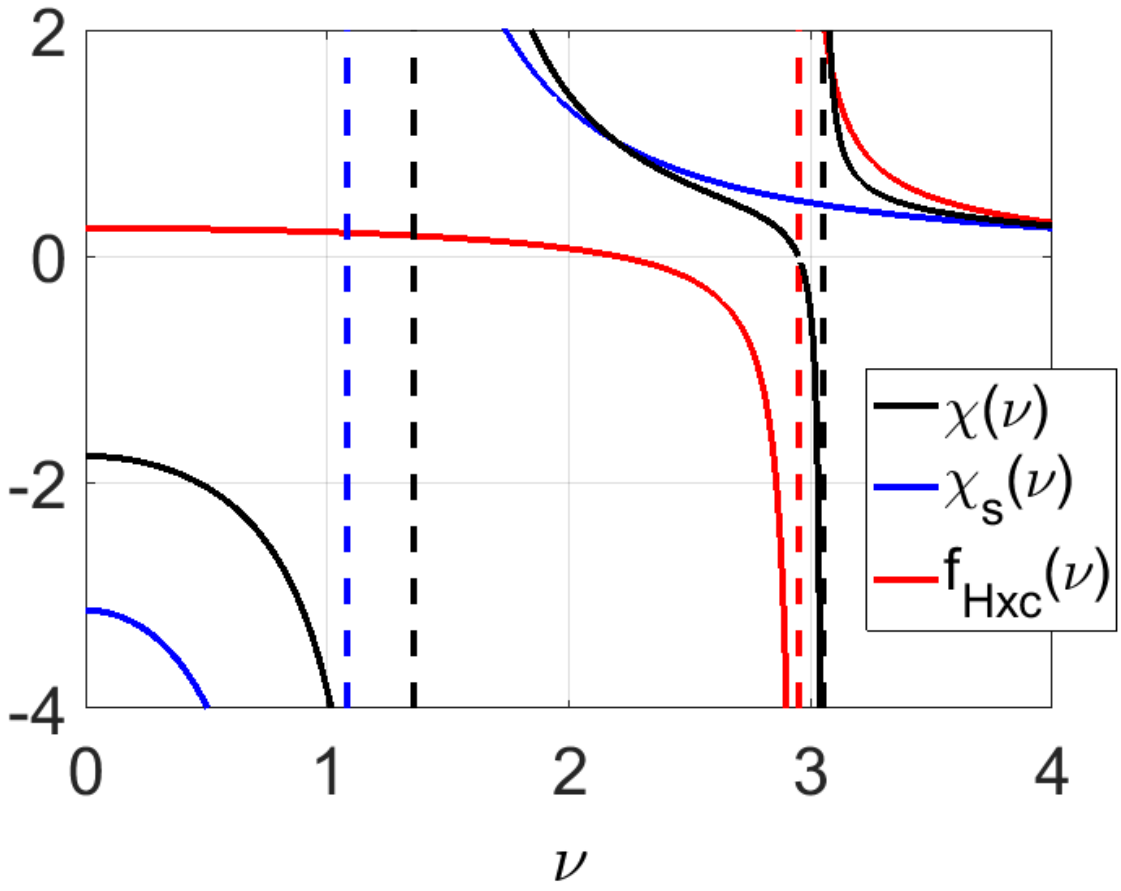

Figure 2.16: Frequency dependence of exact (black) and Kohn-Sham susceptibilities (blue) and exchange-correlation kernel (red) for $U = \Delta v = 1$. Poles marked by dashed vertical lines, as a function of frequency ν. The red line shows the exchange-correlation kernel. Reproduced from Ref. [33].

by the techniques of the earlier section, inverting and subtracting

$$f_{\rm HXC}(\omega) = \chi_{\rm S}^{-1}(\omega) - \chi^{-1}(\omega)\,. \tag{2.26}$$

Fig. 2.16 shows the singular frequency-dependence of the kernel from Eq. (2.26), which allows Eq. (2.25) to have an additional solution.

However, while all this provides insight into how the exact functional performs its magic, it does not tell us directly how to create a general purpose model, which would build this frequency-dependence into an explicit density functional sufficiently accurately to capture

double excitations [195].

Takeaway: The Hubbard dimer demonstrates the accuracy of the adiabatic approximation in TDDFT for single excitations, and also the missing frequency dependence needed to generate the double excitations missing in adiabatic TDDFT.

2.5.2 Talking about TDDFT

We saw in the earlier sections how the KS eigenvalues did not have a formal meaning in pure ground-state KS-DFT. We have seen here that, with the advent of TDDFT, they form the starting point of a scheme which produces the *optical* excitations. These are not the quasiparticle excitations associated with the Green function, which involve a change in particle number.

While the primary function of approximate ground-state DFT is to find energies, it usually also produces reasonably accurate densities, but rather erroneous XC potentials. In fact, this feat is achieved by having all the occupied orbitals shifted (higher) than their exact KS counterparts. A constant shift has no effect on the density. But if the unoccupied levels (at least, the low-lying valence excitations) suffer the same shift, then KS transition frequencies are unaffected, and the adiabatic approximation (usually applied to the same XC approximation as the ground-state calculation) is reasonably accurate for many weakly correlated molecules.

Linear-response TDDFT has been less used in solids, because in the case of insulators, it became clear early on [226] that there is a long-range contribution to the XC kernel (as long-ranged as the Hartree contribution is) that is missed when using a semilocal ground-state approximation adiabatically. There are now many ways around this difficulty [277], some based on modelling the kernel using many-body techniques.

There have been many other approaches suggested for extracting optical excitations from DFT. An old simple one is called Δ-SCF [291], which involves simply using excited-state occupation numbers in a KS calculation, and finding the energy the usual way. Another, which has seen considerable recent interest [94, 335], is to use ensemble DFT [105].

Takeaway: TDDFT can be considered an algorithm for finding the functional (of the ground-state density) for optical excitations.

2.6 Summary

This short review is aimed at broadening understanding of the basic differences between a density functional viewpoint and that of traditional many-body theory. The emphasis here has been on the exact theory, which we have illustrated on the 2-site Hubbard model. We have shown it is confusing to consider KS theory as any kind of traditional mean-field theory, and how the addition of TDDFT allows one to consider the KS eigenvalues as zero-order approximations to the optical excitations, not the quasiparticle excitations.

However, the only reason that anyone cares about the exact theory of DFT is because, in practice, it is extremely useful with relatively unsophisticated approximations. These begin with the famous local density approximation, in which the XC energy per electron at each point in a system is approximated by that of a uniform gas matching the density at this point. This was introduced already in the KS paper (where the statement of exactness appears as a mere footnote), thereby totally muddying the waters between exact and approximate statements. Walter Kohn told KB that he simply noticed the exact nature of the KS scheme after submitting the paper. From about 1990 onwards, many users began using more sophisticated functionals, whose primary effect was to improve total energies and energy differences.

This article has said little or nothing about how to understand such approximations. This is because local (and semilocal) approximations capture a universal limit of all electronic systems, by yielding relatively exact XC energies in this limit [182, 60, 31, 224]. Traditional many-body theory generally considers a power series expansion in the electron-electron interaction. The alternative limit simultaneously increases the number of particles, in a way that the total electron-electron repulsion remains a finite fraction of the total energy even as interactions become weaker. The simplest example of this is that the LDA for exchange, whose formula can be derived by hand, has a percentage error that vanishes for atoms as $Z = N \to \infty$ [60].

This limit is hard-wired into the last term of the real-space Hamiltonian of Eq. (2.1), which is the integral of the density times the one-body potential. This is why the density is the basic variable in DFT. Even if formal theorems can be proven using other variables, this is why density functional theory has been so successful. It is also the case that the one-body potentials to which we apply DFT are diagonal in coordinate space, which is related to why the LDA is a universal limit.

Thus, key aspects of DFT approximations that are crucial to its success are missing from lattice models like the Hubbard model. There is no corresponding universal limit in which LDA becomes exact, even if one uses an approximation based on the uniform case [186, 69]. Again, this is why we created our 1D real-space mimic of 3D reality, instead of just solving lattice models.

Takeaway: This chapter has illustrated a variety of key conceptual points about DFT on a simple model system. Anyone who can answer the exercises will have absorbed 90% of the material, and should be well-qualified to understand exactly what a DFT calculation does, and does not, tell you. In the twenty-first century, with so many DFT calculations being performed in so many different fields, the phrase "Oh, that's just mean-field theory" should no longer have any place in scientific discussions about DFT results.

2.7 Acknowledgments

K.B. acknowledges support from the Department of Energy, Award No. DOE DE-SC0008696.

J.K. acknowledges support from the Department of Energy, Award No. DE-FG02-08ER46496.

We thank Eva Pavarini for suggesting this chapter, and making us write it.

Chapter 3

Density Functional Analysis: The Theory of Density-Corrected DFT

This chapter is a reproduction of Ref. [318], which I co-authored with Stefan Vuckovic, Suhwan Song, Eunji Sim, and Kieron Burke.

3.1 Abstract

Density-corrected density functional theory (DC-DFT) is enjoying substantial success in improving semilocal DFT calculations in a wide variety of chemical problems. This paper provides the formal theoretical framework and assumptions for the analysis of any functional minimization with an approximate functional. We generalize DC-DFT to allow comparison of any two functionals, not just comparison with the exact functional. We introduce a linear interpolation between any two approximations, and use the results to analyze global hybrid density functionals. We define the basins of density-space in which this analysis should apply, and give quantitative criteria for when DC-DFT should apply. We also discuss the effects of

strong correlation on density-driven error, utilizing the restricted HF Hubbard dimer as an illustrative example.

3.2 Introduction and background

Kohn-Sham density functional theory (KS DFT) [158] is widely popular as an electronic structure method [257]. Despite the proliferation of choices of approximate functionals, most calculations use one of a few standard approximations that have been available for the past twenty years, namely generalized gradient approximations (GGAs) or global hybrids, with some enhancements, such as van der Waals corrections [100] or range separation [270]. While moderately accurate for many useful properties, these functionals suffer from well-known deficiencies, including unbound anions, poorly positioned eigenvalues, incorrect molecular dissociation curves, underestimation of reaction barriers, and many others [42, 43]. Thus the never-ending search for improved functionals.

Over the years, many pioneers have shown in specific cases that use of approximate functionals on Hartree-Fock (HF) densities can yield surprisingly accurate results. This includes the early work of Gordon and Kim for weak forces [82], Janesko and Scuseria for reaction barriers and other properties [266, 138], and the original works of Gill *et al.* testing GGA's and hybrids for main group chemistry that led to the adoption of DFT for widespread use in chemistry [79]. Even the prototype of KS-DFT, the X$-\alpha$-method of Slater [279], was designed to yield approximations to HF potentials, which led to an inconsistency between the associated energy functional and its derivative, the potential (see ref 101 for a recent discussion on this topic). Analysis of this difficulty was part of the impetus for the KS paper.

The errors made in DFT calculations were formally separated into two contributions, a functional error and a density-driven error, thereby yielding a formal framework in which

the two errors could be analyzed independently [150]. This led to the theory of density-corrected DFT (DC-DFT), which explains the success of the early work, and has provided a simple procedure for significantly improving the results of semi-local DFT calculations in many situations. For example, for Halogen and Chalcogen weak bonds, which have been used in databases to train van der Waals functionals, the errors are dominated by density-driven errors in the semilocal functional, so such databases cannot be used for that purpose without a correction [153]. In addition to the standard semilocal functionals, it has been recently shown that in specific situations, the energetic accuracy of other density functionals, such as the nonlocal functionals based on adiabatic connection models [316], can be greatly improved by using the HF density and orbitals [63, 314].

Thus, DC-DFT, especially in the form of HF-DFT, in which the Hartree-Fock density is used in place of the exact density, is an extremely practical procedure for improving energetics of abnormal DFT calculations, i.e., those dominated by density-driven errors, but in which the approximate functional is still highly accurate.

Here, we give a detailed formal analysis of the differences that arise between the self-consistent solutions of two distinct density functionals. We consider any two functionals, including the possibility of two different approximations. Thus, DC-DFT is a special case of this more general analysis. We also consider other special cases, including the one-electron case, for which we can calculate all the quantities arising from our analysis that require access to the exact functional and the exact density. The accuracy of PBE for the H-atom is due to a spurious cancellation of both density and functional errors, as well as exchange and correlation errors. We extend our analysis to energy differences, that are of the key importance in chemistry. We also construct measures for the abnormality character of DFT calculations.

3.3 Density functional analysis

In DFT [158, 257, 43, 24, 12], the ground state energy and density of a system with an external potential v are given by:

$$E_v = \min_n E_v[n]. \tag{3.1}$$

The total energy functional $E_v[n]$ is given by:

$$E_v[n] = F[n] + n \cdot v, \tag{3.2}$$

where $n \cdot v = \int d^3r\, n(\mathbf{r})v(\mathbf{r})$, and where $F[n]$ is the universal part of the functional commonly partitioned as:

$$F[n] = T_{\rm S}[n] + U_{\rm H}[n] + E_{\rm XC}[n], \tag{3.3}$$

$T_{\rm S}[n]$ is the KS noninteracting kinetic energy functional, $U_{\rm H}[n]$ the Hartree energy and $E_{\rm XC}[n]$ is the exchange-correlation (XC) functional, which in practical calculations must be approximated. Starting from a given approximate or exact XC functional $E_{\rm XC}[n]$, we can write the corresponding approximate universal functional as:

$$F[n] = F_{\rm SH}[n] + E_{\rm XC}[n], \tag{3.4}$$

where $F_{\rm SH}[n]$ is the universal functional within the Hartree approximation, which neglects exchange and correlation effects: $F_{\rm SH}[n] = T_{\rm S}[n] + U_{\rm H}[n]$. The minimization in eq 3.1 is performed over all N-representable densities, and the density that achieves this minimum we denote by n_v. We define an energetic measure of any arbitrary density difference from

n_v as:

$$D_v[\Delta n] = E_v[n_v + \Delta n] - E_v \geq 0, \tag{3.5}$$

where eq 3.1 ensures that $D_v[\Delta n] \geq 0$ for any isoelectronic change in n_v (i.e. $\int d^3r\, \Delta n(\mathbf{r}) = 0$). We refer to this as the energetic distance from the minimum. We can use this measure to say that n is sufficiently close to n_v, if:

$$D_v[n - n_v] \leq \Delta_c, \tag{3.6}$$

provided that Δ_c is sufficiently small.

Throughout this work, we encounter simple quadratic density functionals, which correspond to normal forms in algebra, and we write:

$$A[\Delta n] = \int d^3r' \int d^3r\, A(\mathbf{r}, \mathbf{r}')\Delta n(\mathbf{r})\Delta n(\mathbf{r}'). \tag{3.7}$$

To gain more insight into the $D_v[\Delta n]$ functional, we can expand $E_v[n]$ around its minimum in a Taylor series: [61]

$$E_v[n_v + \Delta n] = E_v + \frac{1}{2}K_v[\Delta n] + \mathcal{O}(\Delta n^3), \tag{3.8}$$

where $\Delta n(\mathbf{r}) = n(\mathbf{r}) - n_v(\mathbf{r})$, and $K_v(\mathbf{r}, \mathbf{r}')$ is given by:

$$K_v(\mathbf{r}, \mathbf{r}') = \left.\frac{\delta^2 E_v[n]}{\delta n(\mathbf{r}')\delta n(\mathbf{r})}\right|_{n=n_v}. \tag{3.9}$$

Combining eqs 3.4, 3.2 and 3.9, we can write $K_v(\mathbf{r}, \mathbf{r}')$ as:

$$K_v(\mathbf{r}, \mathbf{r}') = f_{\text{SH}}[n_v](\mathbf{r}, \mathbf{r}') + f_{\text{XC}}[n_v](\mathbf{r}, \mathbf{r}'), \tag{3.10}$$

where

$$f_{\mathrm{SH}}[n](\mathbf{r},\mathbf{r}') = \underbrace{\frac{\delta^2 T_{\mathrm{S}}[n]}{\delta n(\mathbf{r}')\delta n(\mathbf{r})}}_{f_{\mathrm{S}}(\mathbf{r},\mathbf{r}')}, + \frac{1}{|\mathbf{r}-\mathbf{r}'|} \tag{3.11}$$

and

$$f_{\mathrm{XC}}[n](\mathbf{r},\mathbf{r}') = \frac{\delta^2 E_{\mathrm{XC}}[n]}{\delta n(\mathbf{r}')\delta n(\mathbf{r})} \tag{3.12}$$

is the static XC kernel. Combining now eqs 3.5 and 3.8, for arbitrary and sufficiently small density difference (i.e. $D_v[\Delta n] \leq \Delta_c$), we can write $D_v[\Delta n]$ as:

$$D_v[\Delta n] \approx \frac{1}{2} K_v[\Delta n]. \tag{3.13}$$

For any $\Delta n(\mathbf{r})$, satisfying $\int d^3r\, \Delta n(\mathbf{r}) = 0$, we define:

$$\Delta n_\beta(\mathbf{r}) = \beta \Delta n(\mathbf{r}). \tag{3.14}$$

In this way, we can see *how far one can go away from* $n_v(\mathbf{r})$, in the $\Delta n(\mathbf{r})$ 'direction', and yet *stay within* Δ_c energetically. Plugging eq 3.14 into eq 3.13, we can easily find the $\beta = \beta_c$ parameter at the boundary (i.e., the one satisfying: $D_v[\Delta n_\beta] = \Delta_c$), and it is given by:

$$\beta_c = \sqrt{\frac{2\Delta_c}{K_v[\Delta n]}}. \tag{3.15}$$

Note again that we use the notation of eq 3.7 for the $K_v[\Delta n]$ quantity. The principal goal of the theory outlined here is to carefully analyze the origin of the energy difference that arises between a pair of different density functionals, when applied to the same system/process. For a given pair of approximate (or one approximate and the other exact) XC functionals:

$E_{\rm XC}^{(0)}$ and $E_{\rm XC}^{(1)}$, we define their difference as:

$$\Delta E_{\rm XC}[n] = E_{\rm XC}^{(1)}[n] - E_{\rm XC}^{(0)}[n]. \tag{3.16}$$

For a given $v(\mathbf{r})$, the difference in the two ground state energies arising from a pair of different functionals is:

$$\Delta E_v = E_v^{(1)}[n_v^{(1)}] - E_v^{(0)}[n_v^{(0)}]. \tag{3.17}$$

By simply adding and subtracting $E_v^{(1)}[n_v^{(0)}]$ from the r.h.s. of eq 3.17, we find:

$$\Delta E_v = -D_v^{(1)}[-\Delta n_v] + \Delta E_{\rm XC}[n_v^{(0)}] \tag{3.18}$$

where $\Delta n_v(\mathbf{r}) = n_v^{(1)}(\mathbf{r}) - n_v^{(0)}(\mathbf{r})$. Reversing the choice of 1 and 0, we also find:

$$\Delta E_v = \Delta E_{\rm XC}[n_v^{(1)}] + D_v^{(0)}[\Delta n_v]. \tag{3.19}$$

Given that $D_v^{(j)} \geq 0$, eqs 3.18 and 3.19 dictate the following chain of inequalities:

$$\Delta E_{\rm XC}[n_v^{(1)}] \leq \Delta E_v \leq \Delta E_{\rm XC}[n_v^{(0)}]. \tag{3.20}$$

By virtue of eq 3.16, the $\Delta E_{\rm XC}[n_v^{(i)}]$ quantity represents the difference between the two functionals evaluated on each density. Therefore, we can identify $\Delta E_{\rm XC}[n_v^{(1)}]$ and $\Delta E_{\rm XC}[n_v^{(0)}]$ of eqs 3.18 and 3.19 as functional-driven terms. On the other hand, $D_v^{(0)}[n_v^{(1)}]$ and $D_v^{(1)}[n_v^{(0)}]$ are the density-driven terms, as they are given by the difference between the same energy functional evaluated on different densities. Generalizing the ideas of DC-DFT [150, 153, 152, 149, 325, 283, 278], for any pair of density functionals, we classify a ΔE_v energy difference as *energy-* or *density*-driven. We consider energy-driven ΔE_v as ones whose functional-driven

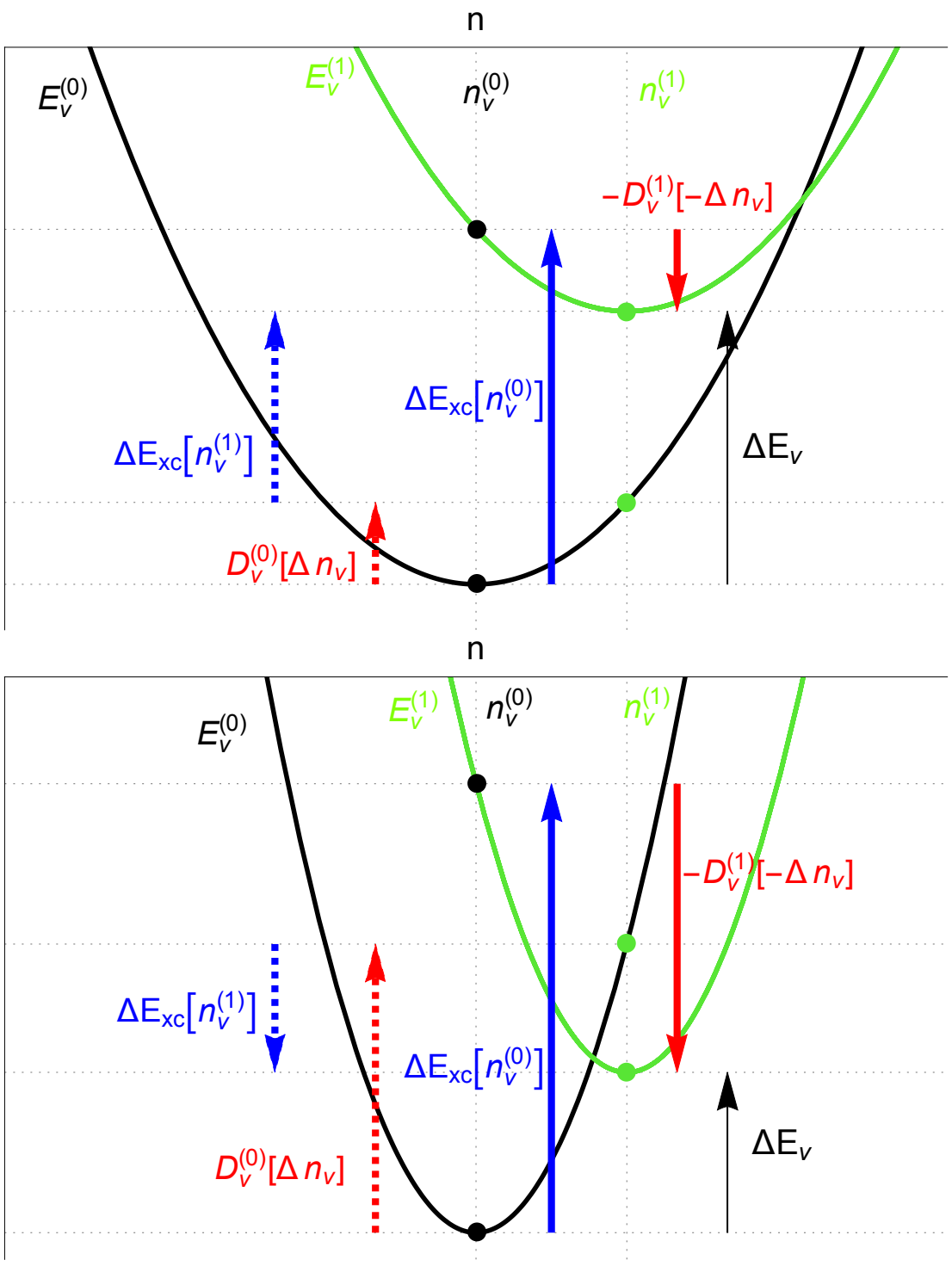

Figure 3.1: Cartoon showing the density-driven and functional-driven contributions to ΔE_v (eqs 3.18 and 3.19) in an energy-driven difference (top panel) and a density-driven difference (bottom panel)

terms in eqs 3.18 and 3.19 strongly dominate the density-driven terms (i.e. $\left|\Delta E_{\rm XC}[n_v^{(0)}]\right| >> D_v^{(1)}[-\Delta n_v]$ and $\left|\Delta E_{\rm XC}[n_v^{(1)}]\right| >> D_v^{(0)}[\Delta n_v]$). On the other hand, in *density-driven* cases the density-driven terms are no longer negligible. In Figure 3.1, we show the two density-driven and the two functional-driven contributions to their energy difference in a cartoon representing an energy-driven difference (top panel) and a density-driven difference (bottom panel). Measures that quantify density-driven character in a given system (again for a given pair of functionals) will be introduced and discussed in Section 3.6.

3.4 Density functional interpolation

To derive an exact expression for ΔE_v by smoothly connecting $E_{\rm XC}^{(0)}[n]$ to $E_{\rm XC}^{(1)}[n]$, we introduce the α-parameter dependent XC functional:

$$E_{\rm XC}^{(\alpha)}[n] = E_{\rm XC}^{(0)}[n] + \alpha \Delta E_{\rm XC}[n]. \tag{3.21}$$

The corresponding total energy functional reads as:

$$E_v^{(\alpha)}[n] = \underbrace{F_{\rm SH}[n] + n \cdot v + E_{\rm XC}^{(0)}[n]}_{E_v^{(0)}[n]} + \alpha \Delta E_{\rm XC}[n] \tag{3.22}$$

and achieves its minimum at $n_v^{(\alpha)}(\mathbf{r})$. Thus its ground state energy is given by $E_v^{(\alpha)} = E_v^{(\alpha)}[n_v^{(\alpha)}]$. More generally, we consider the following energy difference:

$$\Delta E_v^{(\alpha)} = E_v^{(\alpha)} - E_v^{(0)} = E_v^{(\alpha)}[n_v^{(\alpha)}] - E_v^{0}[n_v^{(0)}]. \tag{3.23}$$

Writing:

$$\Delta E_v^{(\alpha)} = \int_0^\alpha d\alpha' \, \frac{\partial E_v^{(\alpha')}}{\partial \alpha'}, \tag{3.24}$$

from eq 3.22, via the Hellmann-Feynman theorem, it follows:

$$\frac{\partial E_v^{(\alpha)}}{\partial \alpha} = \Delta E_{\mathrm{XC}}[n_v^{(\alpha)}]. \tag{3.25}$$

Plugging eq 3.25 into eq 3.24, we find:

$$\Delta E_v^{(\alpha)} = \int_0^{\alpha} d\alpha' \, \Delta E_{\mathrm{XC}}[n_v^{(\alpha')}]. \tag{3.26}$$

Equation 3.26 is analogous to, but different from, the adiabatic connection formula for the correlation energy in DFT [118, 164, 108]. When $\alpha = 1$, eq 3.26 becomes:

$$\Delta E_v = \int_0^{1} d\alpha \, \Delta E_{\mathrm{XC}}[n_v^{(\alpha)}]. \tag{3.27}$$

This shows that the energy difference between two KS calculations with different XC functionals can be found knowing only the difference functional and the interpolating ground-state density. Obtaining ΔE_v from eq 3.27 requires knowledge of $n_v^{(\alpha)}(\mathbf{r})$ for all α values between 0 and 1. To find $n_v^{(\alpha)}(\mathbf{r})$, we write the corresponding Euler equation:

$$v_{\mathrm{SHXC}}^{(0)}[n](\mathbf{r}) + \alpha \Delta v_{\mathrm{XC}}[n](\mathbf{r}) + v(\mathbf{r}) \doteq 0. \tag{3.28}$$

where $v_{\mathrm{SHXC}}^{(0)}[n](\mathbf{r}) = \delta F^{(0)}[n]/\delta n(\mathbf{r})$. Because the number of electrons does not change throughout this work, the potentials in eq 3.28 or any subsequent equations are only determined up to a constant, as denoted by the dot over the equals sign. The density that satisfies eq 3.28 is $n_v^{(\alpha)}(\mathbf{r})$, and by expanding it around $n_v^{(0)}(\mathbf{r})$: $n_v^{(\alpha)}(\mathbf{r}) = n_v^{(0)}(\mathbf{r}) + \alpha \Delta n_v^{(\alpha)}(\mathbf{r})$

($\int d^3r\, n_v^{(\alpha)}(\mathbf{r}) = 0$), we can write eq 3.28 as:

$$
\begin{aligned}
&v_{\mathrm{SHXC}}^{(0)}[n_v^{(0)}](\mathbf{r}) + \alpha \left[K_v^{(0)} \cdot \Delta n_v^{(\alpha)}\right](\mathbf{r}) \\
&+ \alpha \left(\Delta v_{\mathrm{XC}}[n_v^{(0)}](\mathbf{r}) + \alpha \left[\Delta f_{\mathrm{XC}}\left[n_v^{(0)}\right] \cdot \Delta n_v^{(\alpha)}\right](\mathbf{r})\right) \\
&+ v(\mathbf{r}) \doteq 0,
\end{aligned} \tag{3.29}
$$

where we simplified the notation for the following integral:

$$
[A \cdot \Delta n](\mathbf{r}) = \int d^3r'\, A(\mathbf{r}, \mathbf{r}')\Delta n(\mathbf{r}'). \tag{3.30}
$$

At $\alpha = 0$, eq 3.28 becomes:

$$
v_{\mathrm{SHXC}}^{(0)}[n_v^{(0)}](\mathbf{r}) + v(\mathbf{r}) \doteq 0. \tag{3.31}
$$

Plugging eq 3.31 into eq 3.29 gives:

$$
\begin{aligned}
&\left[K_v^{(0)}[n_v^{(0)}] \cdot \Delta n_v^{(\alpha)}\right](\mathbf{r}) + \Delta v_{\mathrm{XC}}[n_v^{(0)}](\mathbf{r}) \\
&+ \alpha \left[\Delta f_{\mathrm{XC}}[n_v^{(0)}] \cdot \Delta n_v^{(\alpha)}\right](\mathbf{r}) \doteq 0,
\end{aligned} \tag{3.32}
$$

Also plugging $K_v^{(\alpha)}(\mathbf{r}, \mathbf{r}') = K_v^{(0)}(\mathbf{r}, \mathbf{r}') + \alpha\Delta f_{\mathrm{XC}}(\mathbf{r}, \mathbf{r}')$ (eqs 3.10 and 3.22) into eq 3.32, we obtain:

$$
\begin{aligned}
&\int d^3r'\, K_v^{(\alpha)}[n_v^{(0)}](\mathbf{r}, \mathbf{r}')\, \Delta n_v^{(\alpha)}(\mathbf{r}') \\
&\doteq -\Delta v_{\mathrm{XC}}[n_v^{(0)}](\mathbf{r})
\end{aligned} \tag{3.33}
$$

In principle, $\Delta n_v^{\alpha}(\mathbf{r}')$ can be obtained by solving eq 3.33, and we can write the solution in terms of the inverse of $K_v^{(\alpha)}[n_v^{(0)}]$:

$$\begin{aligned}&\Delta n_v^{(\alpha)}(\mathbf{r}) \\ &= -\int d^3r' \left\{K_v^{(\alpha)}\right\}^{-1} [n_v^{(0)}](\mathbf{r},\mathbf{r}')\, \Delta v_{\text{XC}}[n_v^{(0)}](\mathbf{r}').\end{aligned} \tag{3.34}$$

We expect that $\Delta n_v^{(\alpha)}(\mathbf{r})$ of eq 3.34 can be fairly approximated by Δn_v and this is equivalent to approximating $n_v^{(\alpha)}$ via the following linear interpolation:

$$n_v^{(\alpha)}(\mathbf{r}) \approx n_v^{(0)}(\mathbf{r}) + \alpha \Delta n_v(\mathbf{r}). \tag{3.35}$$

To explore in what situation the approximation of eq 3.35 becomes exact, we now write $n_v^{(\alpha)}$ as : $n_v^{(\alpha)}(\mathbf{r}) = n_v^{(1)}(\mathbf{r}) - \bar{\alpha}\Delta n_v^{(\alpha)'}$, where $\bar{\alpha} = 1 - \alpha$. Repeating the steps given by eqs 3.28 to 3.34, we find:

$$\begin{aligned}&\Delta n_v^{(\alpha)'}(\mathbf{r}) \\ &= -\int d^3r' \left\{K_v^{(\alpha)}\right\}^{-1} [n_v^{(1)}](\mathbf{r},\mathbf{r}')\, \Delta v_{\text{XC}}[n_v^{(1)}](\mathbf{r}').\end{aligned} \tag{3.36}$$

In this way, when $\Delta n_v^{(\alpha)'}(\mathbf{r})$ is equal to $\Delta n(\mathbf{r})$ of eq 3.34, then the exact $n_v^{(\alpha)}$ is indeed given by the r.h.s of eq 3.35.

To obtain the leading order of $E_v^{(\alpha)}$ in powers of α, we set again: $n_v^{\alpha}(\mathbf{r}) = n_v^{(0)}(\mathbf{r}) + \alpha \Delta n_v^{(\alpha)}(\mathbf{r})$. Then, $E_v^{(\alpha)}$ becomes:

$$\begin{aligned}E_v^{(\alpha)}[n_v^{(\alpha)}] =& F_{\text{SH}}[n_v^{(0)} + \alpha\Delta n_v^{(\alpha)}] + n_v^{(0)} \cdot v \\ &+ \alpha \Delta n_v^{(\alpha)} \cdot v + E_{\text{XC}}^{(0)}[n_v^{(0)} + \alpha\Delta n_v^{(\alpha)}] \\ &+ \alpha \Delta E_{\text{XC}}[n_v^{(0)} + \alpha\Delta n_v^{(\alpha)}].\end{aligned} \tag{3.37}$$

We can expand $E_v^{(\alpha)}[n_v^{(0)} + \alpha\Delta n_v^{(\alpha)}]$ around $n_v^{(0)}(\mathbf{r})$, and write $E_v^{(\alpha)}$ in powers of α:

$$\begin{aligned} E_v^{(\alpha)} &= E_v^{(0)}[n_v^{(0)}] + \alpha\Delta E_{\text{XC}}[n_v^{(0)}] \\ &+ \alpha\left(v_{\text{SHXC}}^{(0)}[n_v^{(0)}] + v\right)\cdot\Delta n_v^{(\alpha)} \\ &+ \alpha^2\left(\Delta v_{\text{XC}}[n_v^{(0)}]\cdot\Delta n_v^{(\alpha)} + \frac{1}{2}K_v^{(0)}\left[\Delta n_v^{(\alpha)}\right]\right) \\ &+ \mathcal{O}(\alpha^3) \end{aligned} \tag{3.38}$$

Combining eqs 3.31 and 3.38, the third term on the r.h.s of eq 3.38 vanishes:

$$\begin{aligned} E_v^{(\alpha)} &= E_v^{(0)}[n_v^{(0)}] + \alpha\Delta E_{\text{XC}}[n_v^{(0)}] \\ &+ \alpha^2\left(\frac{1}{2}K_v^{(0)}\left[\Delta n_v^{(\alpha)}\right] + \Delta v_{\text{XC}}[n_v^{(0)}]\cdot\Delta n_v^{(\alpha)}\right) \\ &+ \mathcal{O}(\alpha^3). \end{aligned} \tag{3.39}$$

Using eq 3.33, we can further simplify eq 3.39:

$$\begin{aligned} E_v^{(\alpha)} &= E_v^{(0)}[n_v^{(0)}] + \alpha\Delta E_{\text{XC}}[n_v^{(0)}] \\ &+ \frac{\alpha^2}{2}\Delta v_{\text{XC}}[n_v^{(0)}]\cdot\Delta n_v^{(\alpha)} + \mathcal{O}(\alpha^3). \end{aligned} \tag{3.40}$$

From eq 3.40, we can see that $E_v^{(\alpha)}$ truncated at second order in α can be obtained from the xc energies and xc potentials at the endpoints. We can also use the following functional expansion:

$$\begin{aligned} &\Delta E_{\text{XC}}\left[n_v^{(0)} + \frac{\alpha}{2}\Delta n_v^{(\alpha)}\right] = \\ &\Delta E_{\text{XC}}\left[n_v^{(0)}\right] + \frac{\alpha}{2}\Delta v_{\text{XC}}\left[n_v^{(0)}\right]\cdot\Delta n_v^{(\alpha)} + \mathcal{O}(\alpha^2), \end{aligned} \tag{3.41}$$

and plug it into eq 3.40 to finally obtain:

$$E_v^{(\alpha)} = E_v^{(0)}\left[n_v^{(0)}\right] + \alpha\Delta E_{\mathrm{XC}}\left[n_v^{(0)} + \frac{\alpha}{2}\Delta n_v^{(\alpha)}\right] + \ldots \tag{3.42}$$

Similarly, writing $n_v^{(\alpha)} = n_v^{(1)}(\mathbf{r}) - \bar{\alpha}\Delta n_v^{(\alpha)}(\mathbf{r})$, we can expand $E_v^{(\alpha)}$ around $n_v^{(1)}$ to obtain:

$$E_v^{(\alpha)} = E_v^{(1)}\left[n_v^{(1)}\right] - \bar{\alpha}\Delta E_{\mathrm{XC}}\left[n_v^{(1)} - \frac{\bar{\alpha}}{2}\Delta n_v^{(\alpha)}\right] + \ldots \tag{3.43}$$

Both results are consistent with applying eq 3.26 and expanding only the density to first order in α. In Section 3.5, we will illustrate the usefulness of the formalism developed in this section for connecting a global hybrid functional to its parent GGA.

3.5 Specific cases

3.5.1 Quantifying errors with DC-DFT

There has been recent great interest in quantifying the errors in densities in DFT [205, 146, 87, 115, 160]. However, in ref 278, it was shown that the theory of DC-DFT provides a natural and unambiguous measure of density error. With that measure, it was not possible to distinguish the densities of empirical and non-empirical functionals based on their self-consistent densities alone.

To understand the background to this, we first must distinguish ground-state KS DFT from other areas of electronic structure. The primary purpose of such calculations is to produce the ground-state energy as a function of nuclear coordinates. Indeed, in principle, one can

deduce the density (and hence any integral over it) from a sequence of such calculations, via the functional derivative with respect to the potential. Of course, such calculations produce KS potentials, orbitals, and eigenvalues, as well as densities and ground-state energies, and all such quantities can be compared (for systems for which the calculation is feasible) to their exact counterparts extracted from a more accurate quantum solver [307, 343, 98, 265]. These are of great interest as inputs to response calculations, such as in linear-response TDDFT or GW methods, and such procedures might be extremely sensitive to such inputs. But in ground-state DFT, the main prediction is the energy of the many-body system, for which the KS scheme is simply a brilliant construct that balances efficiency and accuracy.

Intuitively, one feels that a 'better' XC potential must yield a 'better' density, and in turn, a 'better' density must yield a better energy. After all, the Hohenberg-Kohn (HK) theorem tells us that we reach the ground-state energy only with the exact density and exact KS potential. But such formal statements give no measure of the quality of a density or a potential. Even a well-defined mathematical norm between two densities that vanishes as the exact density is approached does not really provide what we wish for, as an infinity of arbitrarily different norms can be constructed. All can tell us when we have found the exact density, but give differing results for how far away we are from it. A deep part of the problem is that both potentials and densities are functions of $\mathbf{r}$, and so are not characterized by a single number.

As mentioned in ref 278, the basic theorems of DFT give us an ideal solution to this dilemma, via DC-DFT. To write this measure in the language of the density functional analysis, we consider now the specific case where $E_{\rm XC}^{(0)}[n] = E_{\rm XC}[n]$ is the exact XC functional and where $E_{\rm XC}^{(1)}[n] = \tilde{E}_{\rm XC}[n]$ is an approximate functional. Note also that this is the basis of all DC-DFT applications. For a given $v(\mathbf{r})$, the difference between the two corresponding ground state

energies becomes the error in the approximate ground-state energy:

$$\Delta\tilde{E}_v = \tilde{E}_v - E_v. \tag{3.44}$$

For this special case, eq 3.19 becomes:

$$\Delta\tilde{E}_v = \underbrace{\tilde{E}_v[\tilde{n}_v] - E_v[\tilde{n}_v]}_{\Delta\tilde{E}_{\mathrm{XC}}[\tilde{n}_v]} + \underbrace{E_v[\tilde{n}_v] - E_v[n_v]}_{\Delta E^{ideal}[\Delta\tilde{n}_v]=D_v[\Delta\tilde{n}_v]} . \tag{3.45}$$

where $\Delta\tilde{n}_v = \tilde{n}_v - n_v$. For any system, the variational principle (eq 3.1) ensures that $\Delta E^{ideal}[\Delta\tilde{n}_v]$ is always a positive energy for any $\tilde{n}_v(\mathbf{r})$ and vanishes only for the exact density. Another advantage of $\Delta E^{ideal}[\Delta\tilde{n}_v]$ over other metrics is that this measurement of density error is given terms of the energetic consequences. A single number, instead of a function that depends on $\mathbf{r}$, determines how good a given approximate density is. As such, it even provides a useful scale for density differences. For example, if this measure yields results in the microHartree range, why would one even care about errors in the density? Given that it is very difficult in practice to evaluate the exact functional on an approximate density (but see e.g., refs 322 and 315), DC-DFT procedures use the following equation instead:

$$\Delta\tilde{E}_v = \underbrace{\tilde{E}_v[\tilde{n}_v] - \tilde{E}_v[n_v]}_{\Delta E_{\mathrm{D}}=-\tilde{D}_v[-\Delta\tilde{n}_v]} + \underbrace{\Delta\tilde{E}_{\mathrm{XC}}[n_v]}_{\Delta E_{\mathrm{F}}}. \tag{3.46}$$

Equation 3.46 allows us to decompose $\Delta\tilde{E}_v$, the total error made by $\tilde{E}_{\mathrm{XC}}[n]$ and $\tilde{n}_v$, into the functional error $\Delta E_{\mathrm{F}} = \Delta\tilde{E}_{\mathrm{XC}}[n_v]$ and the density-driven error $\Delta E_{\mathrm{D}} = -\tilde{D}_v[-\Delta\tilde{n}_v]$, which is much more practical than the ideal, as it needs only be evaluated on the approximate functional. We can in fact expect ΔE_{D} to be a practical proxy for the intractable $E^{ideal}[\Delta\tilde{n}_v]$ measure. When the approximate density is sufficiently close to its exact counterpart (more precisely, when the two inequalities hold: $D_v[\Delta\tilde{n}_v] \leq \Delta_c$ and $\tilde{D}_v[-\Delta\tilde{n}_v] \leq \Delta_c$), we can write

$\Delta E^{ideal}[\Delta\tilde{n}_v]$ as:

$$\Delta E^{ideal}[\Delta\tilde{n}_v] \approx \frac{1}{2}K_v[\Delta\tilde{n}_v], \tag{3.47}$$

and $\tilde{D}_v[-\Delta n_v]$ as:

$$\tilde{D}_v[-\Delta\tilde{n}_v] \approx \frac{1}{2}\tilde{K}_v[\Delta\tilde{n}_v] \tag{3.48}$$

From eqs 3.47 and 3.48, we can see that if the approximate functional has accurate curvature, $\tilde{D}_v[-\Delta\tilde{n}_v] = -\Delta E_{\mathrm{D}}$ measure is very similar to $\Delta E^{ideal}[\Delta\tilde{n}_v]$.

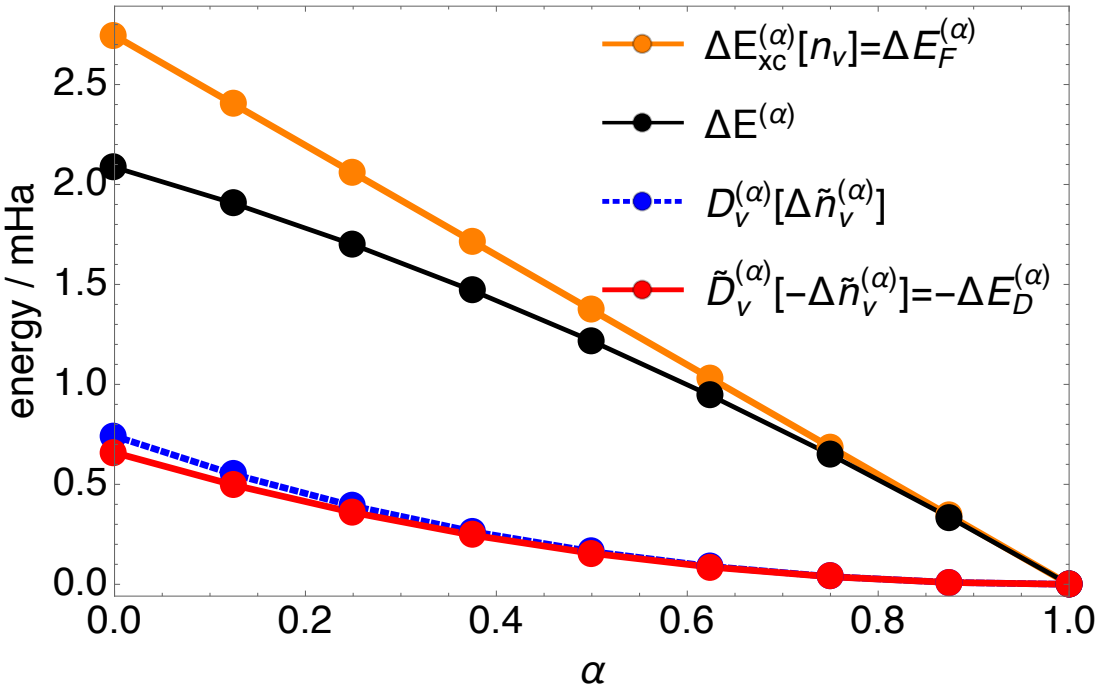

Figure 3.2: Various errors of eqs 3.51 and 3.52 for the α-BLYP calculations of the hydrogen atom as a function of amount of exact exchange mixing

3.5.2 Illustration

Figure 3.2 illustrates many aspects of the analysis described so far. Here we consider the hydrogen atom and the BLYP GGA [11, 167]. We choose this example carefully, because

(a) we have easy access to the exact density, since this is a one-electron case, and (b) our functional correctly has no correlation energy (as LYP correlation vanishes for all fully spin-polarized systems). Thus, when we interpolate between BLYP and HF, we create a global hybrid with a fraction α of exact exchange (EXX):

$$\tilde{E}_{\mathrm{XC}}^{(\alpha)}[n] = \tilde{E}_{\mathrm{XC}}^{\mathrm{GGA}}[n] + \alpha \left(E_{\mathrm{X}}[n] - \tilde{E}_{\mathrm{X}}^{\mathrm{GGA}}[n] \right). \tag{3.49}$$

For this specific cases $\tilde{E}_{\mathrm{XC}}^{(\alpha)}[n]$ reduces to:

$$\tilde{E}_{\mathrm{XC}}^{(\alpha)}[n] = \tilde{E}_{\mathrm{X}}^{\mathrm{B88}}[n] + \alpha \left(E_{\mathrm{X}}[n] - \tilde{E}_{\mathrm{X}}^{\mathrm{B88}}[n] \right), \tag{3.50}$$

where $E_{\mathrm{X}}^{\mathrm{B88}}[n]$ stands for the exchange functional of Becke [11]. For the $\Delta\tilde{E}_v^{(\alpha)} = \tilde{E}_v^{(\alpha)} - E_v$ energy difference (i.e. the error of the hybrids functional), we can rewrite eq 3.46 as:

$$\Delta\tilde{E}_v^{(\alpha)} = \underbrace{-\tilde{D}_v^{(\alpha)}[-\Delta\tilde{n}_v^{(\alpha)}]}_{\Delta E_{\mathrm{D}}^{(\alpha)}} + \underbrace{\Delta\tilde{E}_{\mathrm{XC}}^{(\alpha)}[n_v]}_{\Delta E_{\mathrm{F}}^{(\alpha)}}, \tag{3.51}$$

where $\Delta\tilde{n}_v^{(\alpha)} = \tilde{n}_v^{(\alpha)} - n_v$. In eq 3.51, we can recognize α-dependent density-driven and functional errors. In the same manner, we can re-write eq 3.45:

$$\Delta\tilde{E}_v^{(\alpha)} = \Delta\tilde{E}_{\mathrm{XC}}^{(\alpha)}[\tilde{n}_v^{(\alpha)}] + D_v[\Delta\tilde{n}_v^{(\alpha)}]. \tag{3.52}$$

In this case, all error contributions (eqs 3.51 and 3.52), shown in Figure 3.2, vanish at $\alpha = 1$. On the extreme left ($\alpha = 0$), we see the functional error exceeds the self-consistent error, and the density-driven error is negative, as it should be (note that $-\Delta E_{\mathrm{D}}$ is plotted in Figure 3.2). The functional error is exactly linear, going to zero as $\alpha \to 1$. Notice that the density-driven error must always behave parabolically around $\alpha = 1$.

We also compare the two choices of reference for D_v (blue and red), finding that they are

almost identical. This is telling us that the BLYP density is sufficiently close to the exact density that the expansion to second-order is fine. Moreover, note that as $\alpha \to 1$, the blue and red merge datapoints, and both are on top of a perfect parabola whose curvature is given by $K_v[\Delta n]$ (again as $\alpha \to 1$). Since the self-consistent error is the sum of the functional and density-driven errors, we can deduce its curve just from the values at $\alpha = 0$. Thus the black line is always a parabola if the densities are sufficiently similar, as is the case here. Note that we can see that the energetic difference between the D_v values (blue and red plots) is rather low, showing that they are indeed sufficiently close that there are no significant energetic consequences to approximating all such curves as parabolas.

Finally, we note that, relative to DC-DFT, we have chosen the two functionals in the reverse order: 0 denotes the approximate functional, 1 the exact answer. This is to make these results more readily comparable to other results for hybrids. Simply replace α by $1 - \alpha$ to make Fig. 3.2 in the form for DC-DFT.

3.5.3 Self-interaction error and one-electron systems

For one-electron systems:

$$F[n] = T_{\mathrm{S}}[n],\ E_{\mathrm{X}}[n] = -U_{\mathrm{H}}[n],\ E_{\mathrm{C}} = 0 \qquad (N = 1). \tag{3.53}$$

Standard DFT approximations typically do not satisfy all the conditions of eq 3.53, and for this reason, they suffer from one-electron self-interaction error (SIE) [43, 250]. On the other hand, the HF method is exact for one-electron systems, and thus we can use the HF method to calculate the functional- and density-driven term of SIE. This has been already done in Figure 3.2 for the BLYP hybrids, and in Figure 3.3 we apply the same analysis to hybrids from the PBE functional [239]. The PBE correlation energy, unlike that of LYP, does not

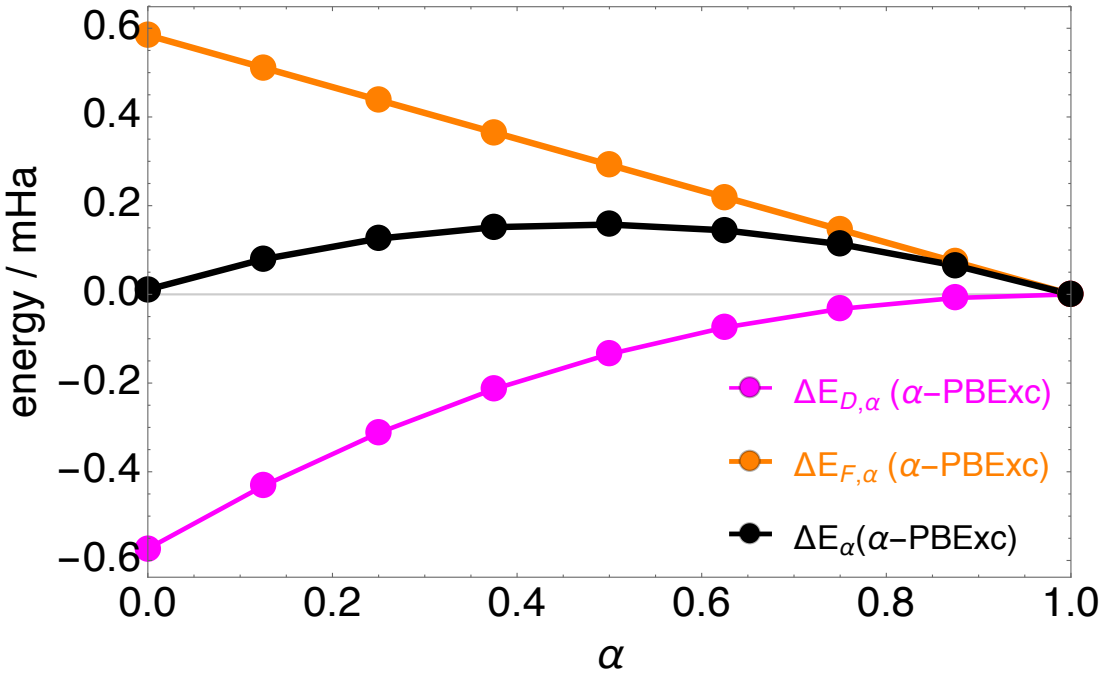

Figure 3.3: Density-driven and functional errors for the α-PBExc (eq 3.54) calculations of the hydrogen atom as a function of amount of exact exchange mixing

vanish for one-electron systems. For this reason, in the case of the PBE functional we modify eq 3.49:

$$\tilde{E}_{\mathrm{XC}}^{(\alpha)}[n] = \tilde{E}_{\mathrm{XC}}^{\mathrm{PBE}}[n] + \alpha\ (E_{\mathrm{X}}[n] - \tilde{E}_{\mathrm{XC}}^{\mathrm{PBE}}[n]), \tag{3.54}$$

In this way, we ensure that the error of the PBE hybrid of eq 3.54 (hereinafter α-PBExc) vanishes at $\alpha = 1$. Note that this does *not* include PBE0 [243], as this PBExc has only 0.75 of PBE correlation at $\alpha = 1/4$. In Figure 3.3, we show the density-driven and functional error of α-PBExc for the hydrogen atom. First, note that the scale of the errors is smaller than that in Figure 3.2. We can also see that both $|\Delta E_{\mathrm{F}}^{(\alpha)}|$ and $|\Delta E_{\mathrm{D}}^{(\alpha)}|$ errors of α-PBExc decrease with α. Nonetheless, its total $\Delta E^{(\alpha)}$ error peaks at $\alpha \approx 0.5$ and nearly vanishes for the PBE functional (the $\alpha = 0$ case). The fact that the PBE gives almost exact energy for the hydrogen atom relies on a cancellation between the functional and density-driven errors (as well as a cancellation between exchange and correlation errors).

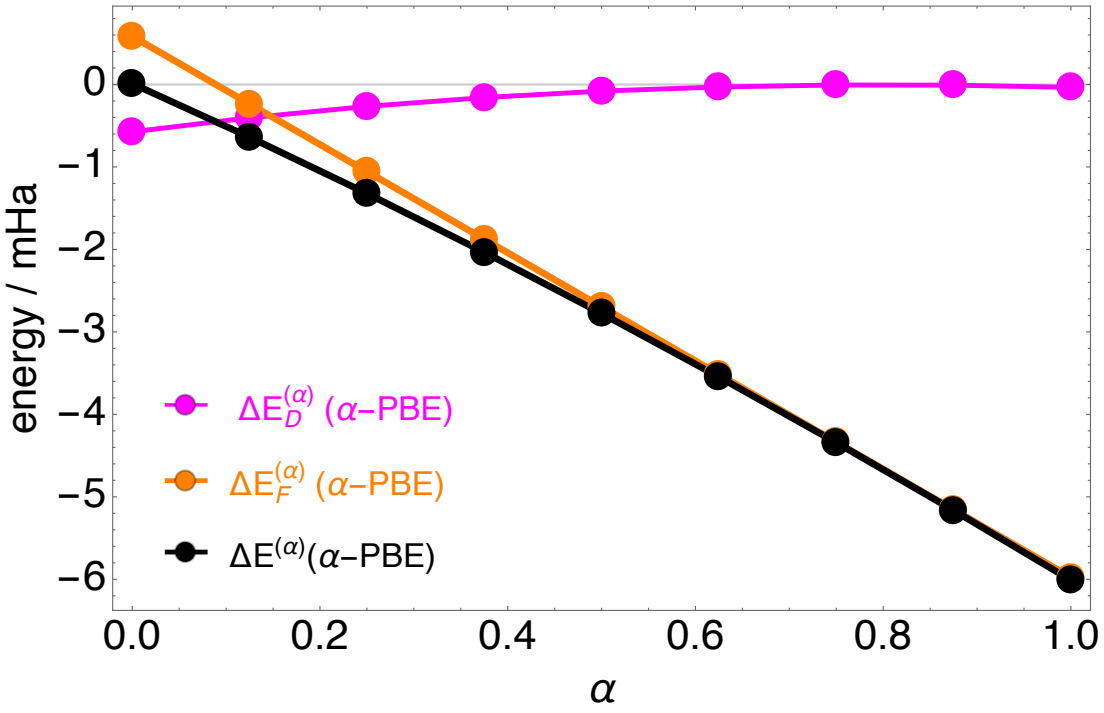

Figure 3.4: Density-driven and functional errors for the α-PBE (eq 3.49) calculations of the hydrogen atom as a function of amount of exact exchange mixing

The same plot for the hydrogen atom obtained with the regular α-PBE hybrid (eq 3.49) is shown in Figure 3.4. Note that now the $\alpha = 1/4$ point represents the PBE0 functional. We can see from Figure 3.4 that as α approaches 1, the functional error strongly dominates its density-driven counterpart. Note here the much larger scale: The self-interaction error in the PBE correlation functional error yields much larger total energy errors than in the previous figure, illustrating the increased error when semilocal correlation functionals are combined with exact exchange. We can also note that ΔE_{D} gets very close to 0 as α approaches 1, although the PBE correlation potential does not vanish for $N = 1$ systems.

3.5.4 The Hartree approximation

Another special case is the Hartree approximation, i.e., solution of the KS equations with XC set to zero. In the Hartree approximation the functional error is just the XC energy

itself, while the density-driven error is just: $\Delta E_{\rm D} = -\frac{1}{2} f_{\rm SH}[n_v^{(0)}, \Delta n]$. To give an illustration, here we consider again the hydrogen atom and the following functional: $\tilde{E}_{\rm XC}^{(\alpha)}[n] = \alpha E_{\rm x}[n]$. The $\tilde{E}_{\rm XC}^{(\alpha)}[n]$ functional allows us to analyze the errors along the path that connects the Hartree approximation ($\alpha = 0$) and the exact functional ($\alpha = 1$) for the hydrogen atom. In Figure 3.5, we show the errors of this functional as a function of α for the hydrogen atom. As expected, the scale of errors is much larger than those shown in Figures 3.2-3.4. For the Hartree approximation, one would expect $E_{\rm D}$ to be different from the ideal, which includes the XC contributions. However, we can see in Figure 3.5 interesting that the $-\Delta E^{ideal}$ datapoints in Figure 3.5 are hardly distinguishable from their $\Delta E_{\rm D}$ counterparts. This implies that at $\alpha = 0$ (the Hartree approximation) we have:

$$\underbrace{f_{\rm S}[n_v, \Delta n]}_{-2\Delta E^{ideal}} \sim \underbrace{f_{\rm SH}[n_v^{(0)}, \Delta n]}_{-2\Delta E_{\rm D}} \quad \text{(H atom)}, \tag{3.55}$$

where $f_{\rm S}(\mathbf{r}, \mathbf{r}')$ is the kinetic component of $f_{\rm SH}(\mathbf{r}, \mathbf{r}')$ (see eq 3.11).

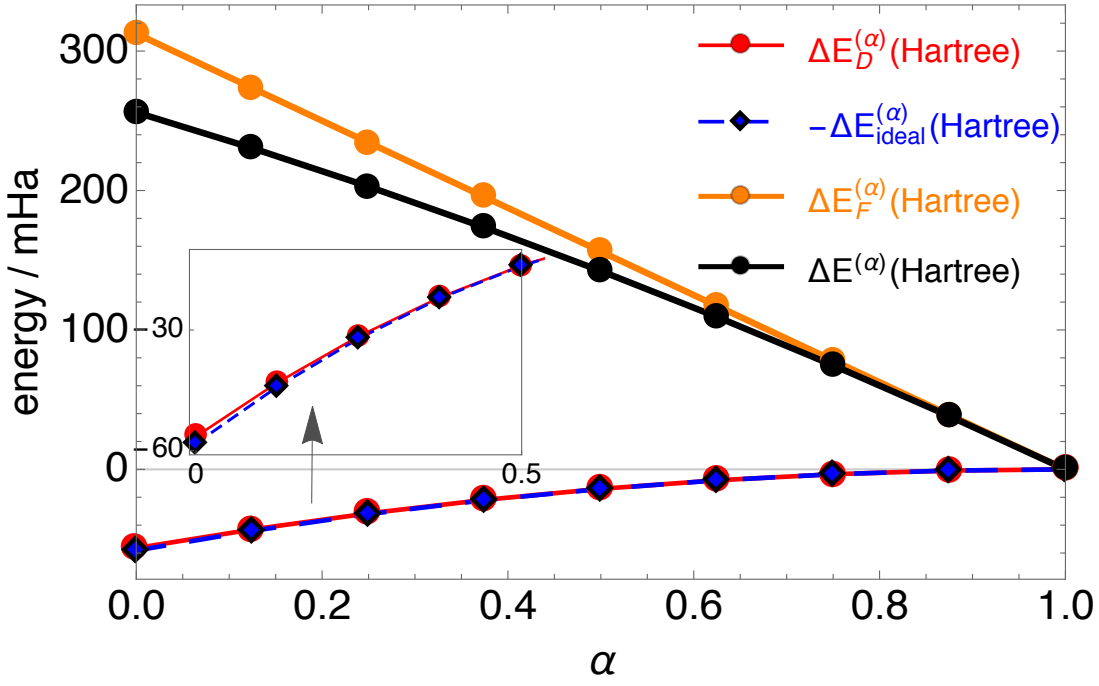

Figure 3.5: Density-driven, the ideal, and functional errors for the hydrogen atom calculation with the $\tilde{E}_{\rm XC}^{(\alpha)}[n] = \alpha E_{\rm x}[n]$ functional.

While Hartree calculations are certainly too inaccurate for chemical purposes [247, 238], one would expect them to have the greatest delocalization error of any approximate functional, since not even LDA exchange is opposing the Hartree energy. A Hartree calculation might thus prove useful in creating a non-empirical measure of delocalization to be used in DC-DFT, as surely no sensible XC approximation should produce a larger density-driven error.

3.5.5 Pure density functionals

So far, we have considered only approximations to XC within the KS scheme, as this is the most common DFT calculation today by far. However, there is much interest in developing orbital-free functionals, especially in contexts when the KS scheme becomes too cumbersome.

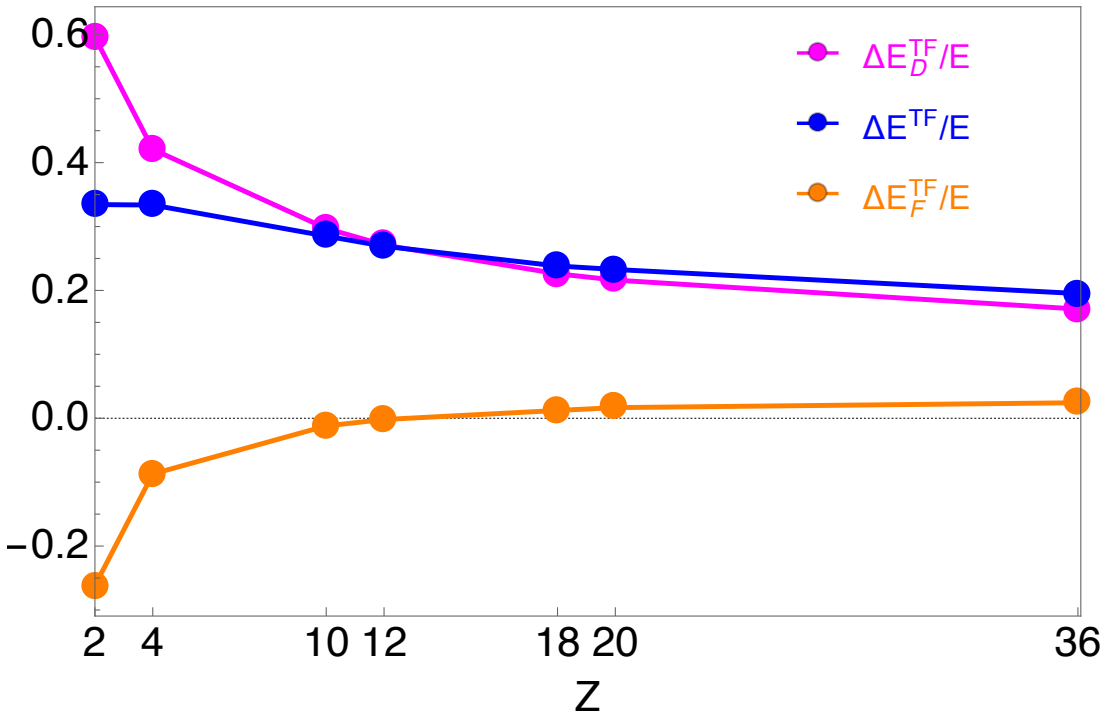

Figure 3.6: Plots showing quantities that involve the density-driven and functional errors of the Thomas-Fermi method (eq 3.57) with Z for a range of small atoms.

Since the entire functional $F[n]$ is approximated in such a scheme, the density is often much poorer than in a KS calculation. In fact, estimates suggest that simple orbital-free approxi-

mations, such as those used in Thomas-Fermi theory [297, 66, 185], produce sufficiently poor densities that their errors are *dominated* by errors in the density, i.e., the density-driven error is much larger than the functional error in most calculations. This is seen in total energy calculations of atoms and of one-dimensional fermions in a flat box [325]. The simplest DC-DFT in orbital-free DFT is to apply the approximation on the exact density to eliminate the density-driven error.

To exemplify the error analysis of orbital-free functionals, we consider here the TF energy functional, whose universal part reads as:

$$F^{\text{TF}}[n] = \underbrace{T_{\text{S}}^{\text{TF}}[n]}_{A^{\text{TF}} \int d^3r\, n(\mathbf{r})^{5/3}} + U[n], \tag{3.56}$$

with $A^{\text{TF}} = \frac{3}{10}\left(3\pi^2\right)^{\frac{2}{3}}$. The total TF error can be, analogously to eq 3.46, partitioned as:

$$\Delta E_v^{\text{TF}} = \underbrace{E_v^{\text{TF}}[n_v^{\text{TF}}] - E_v^{\text{TF}}[n_v]}_{\Delta E_{\text{D}}^{\text{TF}}} + \underbrace{\Delta F^{\text{TF}}[n_v]}_{\Delta E_{\text{F}}^{\text{TF}}}, \tag{3.57}$$

where $\Delta F^{\text{TF}}[n] = \Delta E_v^{\text{TF}}[n] = \Delta T_{\text{S}}^{\text{TF}}[n] - E_{\text{XC}}[n]$. Here we calculate $\Delta E_{\text{D}}^{\text{TF}}$ and $\Delta E_{\text{F}}^{\text{TF}}$ for alkaline earth metals and noble gases up to krypton (Z=36). They are computed by utilizing that for neutral atoms: $E_Z^{\text{TF}} \sim -0.7687\, Z^{7/3}$ [231]. Highly accurate energies and densities, i.e., E_v and n_v entering eq 3.57 have been obtained from the PySCF software [15] at the CCSD level within the aug-cc-pVmZ basis set [59], with the largest m available for each of the atoms. From the plots shown in Figure 3.6, we can see that the density-driven component strongly dominates the total ΔE^{TF} error. For example, in the case of the neon atom ($Z = 10$) most of the TF error is practically density-driven, with $\Delta E_{\text{D}}^{\text{TF}}/E \sim 28.4\%$ and $\Delta E_{\text{F}}^{\text{TF}}/E$ being -1.3%. We remember that for neutral atoms, as $Z \to \infty$, the TF theory

becomes relatively exact, in the sense that it satisfies: [184, 26, 32]

$$\lim_{Z\to\infty} \frac{\Delta E_Z^{\mathrm{TF}}}{E_Z} \to 0. \tag{3.58}$$

Thus, as $Z \to \infty$, our blue curve in Figure 3.6 should vanish. Nevertheless, in Figure 3.6 we are still far from this limit, as at our largest Z value ($Z = 36$), $\Delta E_Z^{\mathrm{TF}}/E_Z$ is around 1/5.

This suggests several important points regarding these functionals. First, they must always be tested self-consistently, as in, e.g., refs 170, 45, 46. This is because tests of new orbital-free approximations on KS densities does not tell us much about their overall accuracy, given that the density-driven errors can be very large. At the same time, comparison of the functional on the KS density then provides enough information to separate functional- from density-driven errors, and we expect that even the KS densities obtained from the (semi)local XC approximations are sufficiently accurate for this purpose. Second, reports of failures of TF theory and its extensions should be revisited to determine if these are density-driven or functional-driven. If the former, one should focus on improving the densities rather than the total energies alone. Third, this supports efforts [68] to approximate the Pauli potential [197, 177] directly as a density functional, without requiring that the KS potential be a functional derivative.

In the context of the present paper, it should prove useful when comparing two orbital-free approximations to decompose their differences into functional- and density-driven contributions. If two different approximations differ in both contributions, this would suggest that good aspects of both might be combined to separately minimize each error.

3.5.6 Strong correlation

In our last example, we show that density-driven errors can become large when systems are strongly correlated, but need not be. Generating a simple example is not so easy, as one needs essentially exact densities upon which to make evaluations and comparisons. Fortunately, the two-site two-fermion Hubbard model is an example where all quantities can be determined analytically. Many relevant KS-DFT quantities have been calculated exactly and summarized in two recent reviews, one on the ground-state[36] and one on linear-response TDDFT.[34]

For any spin-unpolarized two-electron system (in the absence of magnetic fields), the restricted HF functional (RHF) is:

$$F^{RHF}[n] = T_{\rm S}[n] + U_{\rm H}[n]/2, \tag{3.59}$$

as half the Hartree is canceled by exchange, and correlation is ignored. In fact, the traditional definition of correlation energy in quantum chemistry, is:

$$E_{\rm C}^{QC} = E_v - E_v^{RHF}[n^{RHF}]. \tag{3.60}$$

In the RHF case, we have $\tilde{E}_{\rm XC}[n] = E_{\rm X}[n]$, and thus:

$$\Delta\tilde{E}_{xc}[n] = E_{\rm X}[n] - E_{\rm XC}[n] = -E_{\rm C}[n]. \tag{3.61}$$

From eqs 3.46 and 3.61, we obtain the functional error of RHF:

$$\Delta E_{\rm F} = -E_{\rm C}[n_v] = -E_{\rm C}. \tag{3.62}$$

By subtracting $\Delta E_{\rm F}$ from the total error made by RHF, which is equal to $-E_{\rm C}^{QC}$ (eq. 3.60),

we obtain:

$$\Delta E_{\mathrm{D}} = E_{\mathrm{C}} - E_{\mathrm{C}}^{QC}. \tag{3.63}$$

For our 2-site model, the functional reduces to a simple function. The onsite occupations are n_1 and n_2, i.e., the density is just two non-negative real numbers. Moreover, since they always sum to 2, the density is fully represented by their difference. Likewise, we can choose the average potential to be zero and represent the inhomogeneity in the potential by a single number, Δv, the on-site potential difference. If we choose the hopping parameter $t = 1/2$, the only other parameter is U, the energy cost of double occupation of a site.

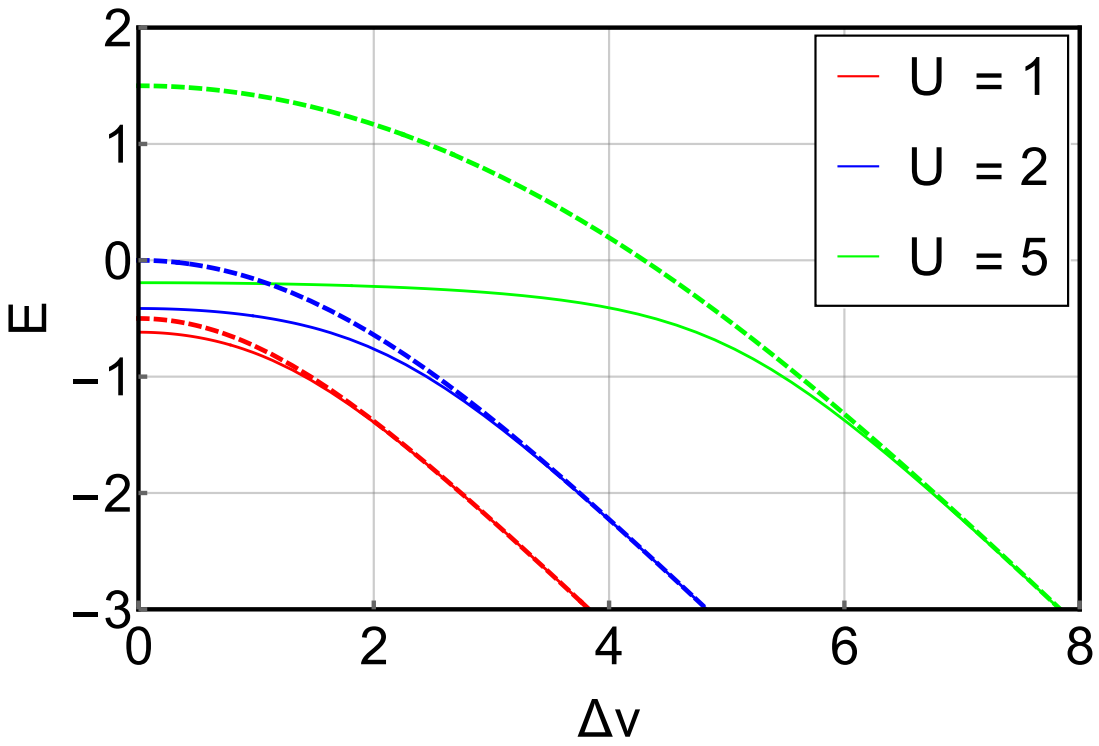

Figure 3.7: Restricted Hartree-Fock Hubbard dimer ground-state energy (dashed line) and exact Hubbard dimer ground-state (solid line) as functions of Δv for varying values of U (see ref 36

The error in RHF and the exact ground state energy is explored in Figure 3.7 for varying levels of correlation and inhomogeneity. The absolute error increases with U, as expected.

The energy error in energy for each level of correlation is most prominent in the symmetric dimer ($\Delta v = 0$), and diminishes and rapidly vanishes beyond Δv larger than U, where the energy becomes linearly correlated with the on-site potential difference. Thus the system becomes weakly correlated for $\Delta v > U$.

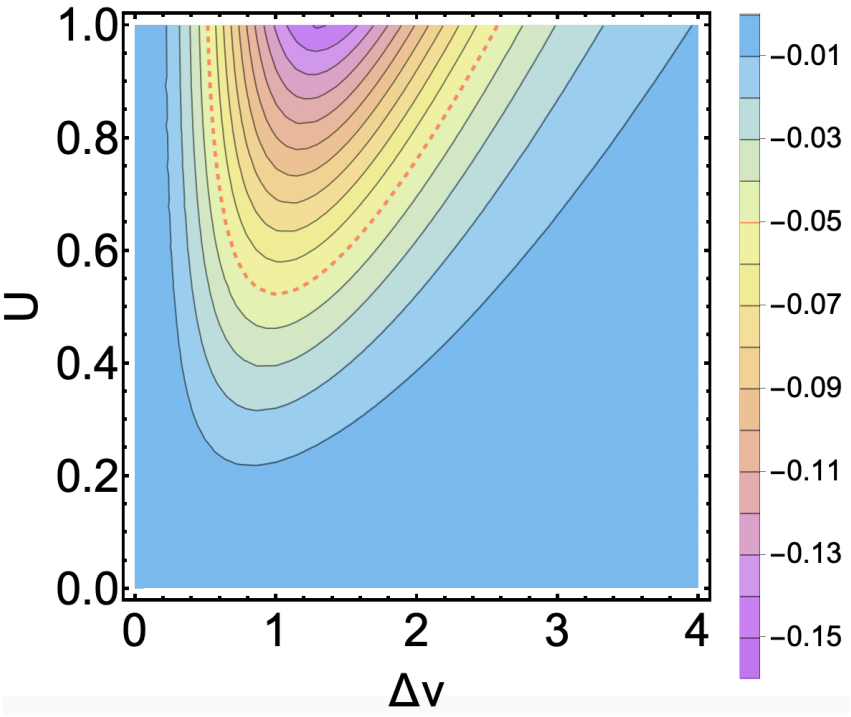

Figure 3.8: Fraction of error that is density-driven for moderate values of U, with the 5% contribution contour marked by a red, dashed line.

The functional-driven and density-driven contributions to this HF error were then isolated through the use of eqs 3.60, 3.62 and 3.63. The fraction of the total error attributed to the density-driven component ($\Delta E_{\text{D}}/E_C^{QC}$) is shown in Figure 3.8 for weakly correlated dimers with values of U up to 1. As U decreases in size, both the magnitude of total error and its density-driven contribution decrease substantially. For $U < 0.5$, there is no Δv for which there is a density-driven contribution greater than 5% of the total error. Of course, as $U \to 0$, this ratio must vanish, so this is not unexpected. But we also see that the density-driven error vanishes at $\Delta v = 0$ for any value of U, no matter how large, for symmetry reasons. Thus even a strongly correlated system might have no density-driven error. Moreover, for $\Delta v > 1+2U$,

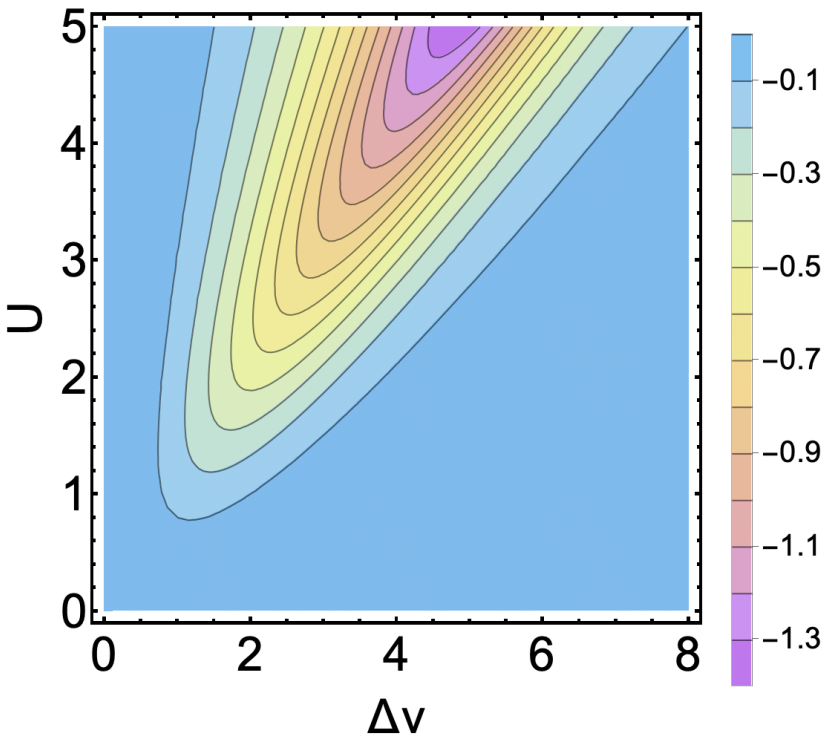

Figure 3.9: Same as Fig. 3.8, but zoomed out in U and Δv. Note the change in contour/color scale.

again the error is less than 5%, due to correlation being weakened by inhomogeneity. So for any given U, there is a maximum in the fraction of density-driven error as a function of Δv, and it is at non-zero Δv.

The density-driven error ratio for more strongly correlated dimers is shown in Figure 3.9, and has characteristics identical to the weak correlation plot, but on a larger scale. Clearly the maximum fractional density-driven error becomes much larger with U and can even exceed -1. We also see that for $U > 1$, the region of small density-driven error around $\Delta v = 0$ can even increase with U. For fixed Δv, the fraction of density-driven error goes down with sufficiently large U!. The relation between RHF density-driven error and strong correlation is clearly not trivial.

To avoid confusion, we note that this section has focused on the density-driven error in RHF. In the more realistic calculations of weakly-correlated systems in the rest of this work, we

often assume that error is much smaller than the density-driven error of a semilocal DFT calculation, and hence HF-DFT yields more accurate energies in such cases. Because the Hubbard dimer is a site model, there is no genuine correspondence with semilocal DFT approximations to test on here.

3.6 Energy differences

Key chemical concepts are determined by energy differences (e.g., atomization energies, ionization energies, barrier heights, reaction energies, etc.). So far we have focused on total energies and in this section we extend our analysis to energy differences. If, for example, we use an approximate functional to calculate an energy difference, what is the density-driven error that pertains to that energy difference? The following analysis answers this question.

For simplicity, we first look at the energy difference of systems A and B, whose external potentials are v_A and v_B, respectively. This energy difference obtained from a total energy functional that corresponds to a given $E_{\rm XC}^{(j)}$ is given by:

$$E_{AB}^{(j)} = E_A^{(j)}[n_A^{(j)}] - E_B^{(j)}[n_B^{(j)}]. \tag{3.64}$$

When two functionals are involved, we can also define the difference between $E_{AB}^{(1)}$ and $E_{AB}^{(0)}$:

$$\Delta E_{AB} = E_{AB}^{(1)} - E_{AB}^{(0)}. \tag{3.65}$$

Plugging eq 3.64 into eq 3.65 gives:

$$\begin{aligned}\Delta E_{AB} =& E_A^{(1)}[n_A^{(1)}] - E_A^{(0)}[n_A^{(0)}] \\ & - \Big(E_B^{(1)}[n_B^{(1)}] - E_B^{(0)}[n_B^{(0)}]\Big).\end{aligned} \tag{3.66}$$

Plugging eq 3.18 into eq 3.66, we can obtain the counterpart of eq 3.18 for the energy differences between systems A and B:

$$\begin{aligned}\Delta E_{AB} =& \Delta E_{\text{XC}}[n_A^{(0)}] - \Delta E_{\text{XC}}[n_B^{(0)}] \\ & \underbrace{-D_A^{(1)}[-\Delta n_A] + D_B^{(1)}[-\Delta n_B]}_{-D_{AB}^{(1)}}.\end{aligned} \tag{3.67}$$

Similarly, we can also plug eq 3.19 into eq 3.66, to obtain the counterpart of eq 3.19 for the energy differences between systems A and B:

$$\begin{aligned}\Delta E_{AB} =& \Delta E_{\text{XC}}[n_A^{(0)}] - \Delta E_{\text{XC}}[n_B^{(0)}] \\ & + \underbrace{D_A^{(0)}[\Delta n_A] - D_B^{(0)}[\Delta n_B]}_{D_{AB}^{(0)}}.\end{aligned} \tag{3.68}$$

In eqs 3.67 and 3.68, we recognize $D_{AB}^{(1)}$ and $D_{AB}^{(0)}$ as the density-driven terms, pertinent to the energy differences between systems A and B. While $D_v^{(j)}$ of eq 3.5, which corresponds to the total energies is always greater or equal to 0, its counterpart that pertains to the energy differences (eqs 3.67 and 3.68) does not have a definite sign. Furthermore, if we look at $D_{AB}^{(0)} = D_A^{(0)}[\Delta n_A] - D_B^{(0)}[\Delta n_B]$ (eq 3.68), where $D_A^{(0)}[\Delta n_A] \geq 0$ and $D_B^{(0)}[\Delta n_B] \geq 0$ we can see that $D_{AB}^{(0)}$ can easily vanish when $D_A^{(0)}[n_A^{(1)}] \sim D_B^{(0)}[n_B^{(1)}]$. Therefore, $D_{AB}^{(0)}$ and $D_{AB}^{(1)}$ can vanish even when $n_A^{(1)}$ and $n_B^{(1)}$ are drastically different from $n_A^{(0)}$ and $n_B^{(0)}$, respectively.

The shown example that involves the energy difference between systems A and B, can be easily generalized to any energy difference of interest. For instance, consider the following

chemical reaction:

$$\sum_{l=1}^{L} R_l \rightarrow \sum_{m=1}^{M} P_m, \tag{3.69}$$

where $\{R_l\}$ is a set of reactants and $\{P_m\}$ is a set of products. Then the energy of this reaction obtained from the $E_v^{(j)}[n]$ functional is:

$$E_{\mathrm{ED}}^{(j)} = \sum_{m=1}^{M} E_{P,m}^{(j)}[n_{P,m}^{(j)}] - \sum_{l=1}^{L} E_{R,l}^{(j)}[n_{R,l}^{(j)}]. \tag{3.70}$$

The corresponding difference in $E_{\mathrm{ED}}^{(j)}$ between $j = 0$ and $j = 1$ functional is :

$$\Delta E_{\mathrm{ED}} = E_{\mathrm{ED}}^{(1)} - E_{\mathrm{ED}}^{(0)}. \tag{3.71}$$

Then $D_{\mathrm{ED}}^{(0)}$ that corresponds to ΔE_{ED} is given by:

$$D_{\mathrm{ED}}^{(0)} = \sum_{m=1}^{M} D_{P,m}^{(0)}[\Delta n_{P,m}] - \sum_{l=1}^{L} D_{R,l}^{(0)}[\Delta n_{R,l}], \tag{3.72}$$

and its $D_{\mathrm{ED}}^{(1)}$ counterpart is given by:

$$D_{\mathrm{ED}}^{(1)} = \sum_{m=1}^{M} D_{P,m}^{(1)}[-\Delta n_{P,m}] - \sum_{l=1}^{L} D_{R,l}^{(1)}[-\Delta n_{R,l}]. \tag{3.73}$$

In this section we have shown that once we know the density-driven terms that pertain to total energies, it is easy to calculate these terms that pertain to energy differences. If we are again interested in the energy difference between system A and B obtained with an approximate functional, first we can use eq 3.46 to obtain ΔE_{D} of that functional pertaining to the total energies of systems A and B. Then, following eqs 3.67 and 3.68, we can easily calculate ΔE_{D} that pertains to the desired energy difference. It will be just: $\Delta E_{\mathrm{D}} = \Delta E_{\mathrm{D}}[A] - \Delta E_{\mathrm{D}}[B]$. The same analysis can be extended to differentials, such as those determining equilibrium

bond lengths or transition state structures.

3.7 Conclusion

We have given a detailed account of the considerations that led to the recent successes of density-corrected DFT. We have also generalized that theory to allow different approximate functionals to be compared in the same way as DC-DFT allows one approximation to be compared with exact results. We have shown that typical density differences between reasonably accurate functionals can be shown quantitatively to be close enough to allow treatment via density functional analysis expansions truncated at second order. We consider different special cases of our analysis, and note that DC-DFT is just one of these. For a given pair of density functional approximations, we also construct measures for their *relative abnormality* (i.e. the situation when the density-driven terms strongly dominate the energy difference obtained with two different approximations).

We have noted many pioneering efforts in the chemistry literature in which density-corrected calculations were performed, usually based on intuition. We also point out that, as long ago as 1996, Levy and Görling advocated the use of self-consistent exact-exchange calculations, with a correction defined to produce the exact ground-state energy as a functional of the 'wrong' EXX density [176]. For purposes of calculating energies, the approach of Levy and Görling is nearly equivalent to HF-DFT, given that the EXX and HF densities are probably indistinguishable. Of course, we only advocate this procedure for abnormal systems, as in normal systems we expect the density from the standard semilocal approximations to be more accurate than that of HF.

Finally, we note that of course it is highly unsettling to run KS calculations in this non-self-consistent fashion. Many advantages that are often taken for granted, such as the exactness

of the Hellmann-Feynman theorem in the basis-set limit, are no longer true, and many corrections need to be coded. But, in fact, all information about the density can be extracted from a sequence of total energy calculations, since:

$$n(\mathbf{r}) = \frac{\delta E_v}{\delta v(\mathbf{r})}. \tag{3.74}$$

Treating HF as EXX, this leads to a predicted change in the density of a HF-DFT calculation relative to a self-consistent DFT density:

$$\Delta n(\mathbf{r}) = \frac{\delta(E^{DFT}[n^{\mathrm{HF}}] - E^{DFT}[n^{DFT}])}{\delta v(\mathbf{r})}. \tag{3.75}$$

Thus, in principle, one could calculate the improvement to the density predicted by HF-DFT, and any other properties depending only on the density.

3.8 Acknowledgements

KB and JK acknowledge funding from NSF (CHE 1856165). SV acknowledges funding from the Rubicon project (019.181EN.026), which is financed by the Netherlands Organisation for Scientific Research (NWO). SS and ES were supported by the grant from the Korean Research Foundation (2017R1A2B2003552).

Chapter 4

Exact Conditions for Ensemble Density Functional Theory

This chapter is a reproduction of Ref. [271], which I co-authored with Thais Scott, Steven Crisostomo, Aurora Pribram-Jones, and Kieron Burke.

4.1 Abstract

Ensemble density functional theory (EDFT) is a promising alternative to time-dependent density functional theory for computing electronic excitation energies. Using coordinate scaling, we prove several fundamental exact conditions in EDFT, and illustrate them on the exact singlet bi-ensemble of the Hubbard dimer. Several approximations violate these conditions and some ground-state conditions from quantum chemistry do not generalize to EDFT. The strong-correlation limit is derived for the dimer, revealing weight-dependent derivative discontinuities in EDFT.

4.2 Introduction

Sophisticated functional approximations and a relatively low computational cost have made density functional theory [127, 159] (DFT) the prevailing method used in electronic structure calculations. [25, 13, 140, 210, 303] Currently, the most popular way to access excited states in the DFT formalism is through time-dependent DFT (TDDFT), [263, 38, 199, 303, 196] which has been used to predict electronic excitation spectra among other properties. Although TDDFT has been incredibly successful, [196] standard approximations fail to replicate charge-transfer excitation energies [300], correctly locate conical intersections [173] or recover double excitations [196] without an *ad hoc* dressing. [133]

A less well-known but comparably rigorous alternative to TDDFT is ensemble density functional theory [295, 104, 103] (EDFT), which is currently experiencing a renaissance. [337, 258, 336, 93, 51, 52, 95, 71, 96, 97, 91, 88, 188, 201, 92, 334, 40, 268] As the EDFT field is revived, it is important to find exact conditions that can be enforced on newly developed EDFT approximations. This is especially important in EDFT, where the choice of ensemble weights is unlimited (assuming they are normalized and are monotonically non-increasing with energy) and can significantly impact the accuracy of the energies. Exact conditions have been essential in the development of accurate functionals in ground-state DFT, and we expect them to be more critical in EDFT. [240, 178, 233]

Here, several exact conditions for EDFT are proven and illustrated. We generalize coordinate scaling inequalities and equalities of the exchange and correlation energies and the concavity condition to ensembles. Using the Hubbard dimer, we show examples of each foundational condition and examine approximations in EDFT, finding examples of compliance and violation. Fig 4.1 illustrates some of these conditions nicely. [131] It shows the limits (red) one can place on the $U = 5$ dimer (black) from results for $U = 4$ (blue), using one of our inequalities. The rest of this paper explains the behavior of these curves, including

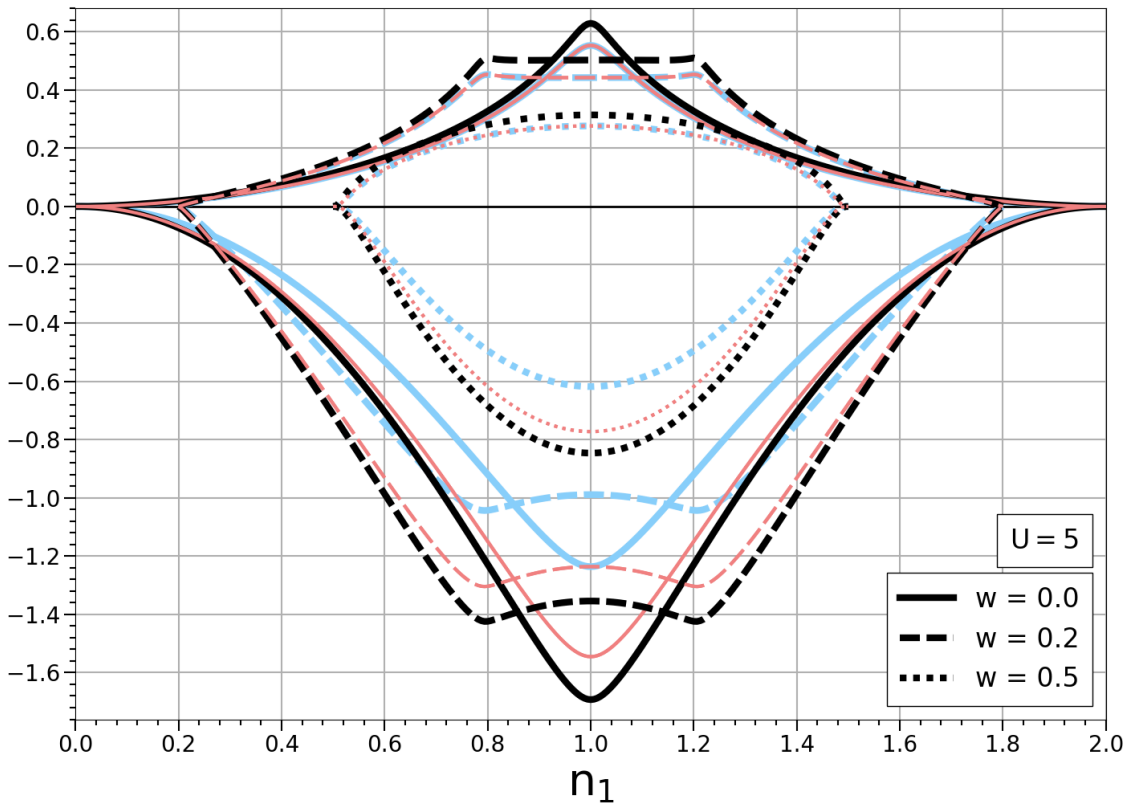

Figure 4.1: The Hubbard dimer singlet bi-ensemble correlation energies (negative values) and kinetic contribution (positive values) for $U = 4$ (light blue) and $U = 5$ (black) as a function of site-occupation and different weights. Red curves deduced from $U = 4$ constrain the $U = 5$ curve via Eq. 4.21.

non-monotonicity with weight and their shapes for large U. These exact results provide examples of the many ways in which EDFT can differ from ground-state DFT.

4.3 Background

EDFT is a formally exact generalization of ground-state KS-DFT, where the ensemble consists of several eigenstates of an N-electron system. Consider any ensemble density matrix, $\hat{\Gamma}_w$, of the form

$$\hat{\Gamma}_w(\mathbf{r}_1...\mathbf{r}_N, \mathbf{r}'_1...\mathbf{r}'_N) = \sum_{m=0}^{M} w_m \left|\Psi_m(\mathbf{r}_1...\mathbf{r}_N)\right\rangle \left\langle\Psi_m(\mathbf{r}'_1...\mathbf{r}'_N)\right|, \tag{4.1}$$

where Ψ_m are any orthonormal wave functions, and w_m are positive monotonically non-increasing weights that are normalized. The expectation value of any operator $\hat{A}$ is then

$$A[\hat{\Gamma}_w] = \mathrm{Tr}\{\hat{\Gamma}_w \hat{A}\} = \sum_{m=0}^{M} w_m \langle \Psi_m | \hat{A} | \Psi_m \rangle . \tag{4.2}$$

An ensemble energy is then the variational minimum of the Hamiltonian, yielding

$$E_w = \min_{\Gamma_w} \mathrm{Tr}\{\hat{\Gamma}_w \hat{H}\}, \tag{4.3}$$

where m labels the eigenstates, in order of increasing energy, and E_m are the eigenvalues. Transition energies can be deduced from differences between ensemble calculations of differing weights. [225] EDFT tells us that there exists a w-dependent density functional

$$F_w[n] = \min_{\Gamma_w \to n} \mathrm{Tr}\{\hat{\Gamma}_w (\hat{T} + \hat{V}_{\mathrm{ee}})\}, \tag{4.4}$$

where $\hat{T}$ is the kinetic energy operator and $\hat{V}_{\mathrm{ee}}$ is the electron-electron repulsion. We denote the minimizer by $\Gamma_w[n]$. Then

$$E_w = \min_{n} \left\{ F_w[n] + \int n(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} \right\}, \tag{4.5}$$

where $v(\mathbf{r})$ is the external potential. Any expectation value can be converted into a density functional via $A[n] = A[\Gamma_w[n]]$. The minimizing density is

$$n_w(\mathbf{r}) = \sum_{m=0}^{M} w_m n_m(\mathbf{r}), \tag{4.6}$$

where $n_m(\mathbf{r})$ is the density of the m-th level.

A key facet of EDFT is that the equivalence between the exact density and the non-

interacting KS density is only true for the ensemble average, and it is not necessarily true for the individual densities within the weighted sum. The following conditions are true only for the ensemble energy, not the individual excited-state energies.

Uniform coordinate scaling has been responsible for multiple advances in DFT. However, coordinate scaling investigations in EDFT have thus far only been used to define the adiabatic connection formula for the exchange-correlation energy [212] or examining the behavior of EDFT in the low-density and high-density regimes, without formal theorems based on scaling. [90] Additional work on foundational theorems include the virial theorem for EDFT by Nagy [215, 213, 214] and the signs of correlation energy components, by Pribram-Jones *et al.* [258] We build on this foundation by deriving uniform scaling inequalities based on the variational definition of the ensemble functional. [178, 252] We also provide numerical verification and proofs of the basic principles and some additional exact conditions.

4.4 Results

We use norm-preserving homogeneous scaling of the coordinate $\mathbf{r} \rightarrow \gamma\mathbf{r}$ with $0 < \gamma < \infty$. The scaled density matrix is defined as

$$\Gamma_{w,\gamma}(\mathbf{r}_1 ... \mathbf{r}'_N) := \gamma^{3N}\, \Gamma_w(\gamma\mathbf{r}_1 ... \gamma\mathbf{r}'_N), \tag{4.7}$$

and a scaled density is $n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r})$. Trivially,

$$T[\Gamma_{w,\gamma}] = \gamma^2 T[\Gamma_w], \quad V_{\rm ee}[\Gamma_{w,\gamma}] = \gamma\, V_{\rm ee}[\Gamma_w]. \tag{4.8}$$

Because these scale differently, $\Gamma_{w,\gamma}[n] \neq \Gamma_w[n_\gamma]$. By the variational principle, $F[n_{w,\gamma}] \leq F[\hat{\Gamma}_{w,\gamma}[n]]$, which gives the fundamental inequality of scaling,

$$T_w[n_\gamma] + V_{\mathrm{ee},w}[n_\gamma] \leq \gamma^2 T_w[n] + \gamma V_{\mathrm{ee},w}[n]. \tag{4.9}$$

Manipulation of this formula yields, for $\gamma \geq 1$, [178]

$$T_w[n_\gamma] \leq \gamma^2 T_w[n], \quad V_{\mathrm{ee},w}[n_\gamma] \geq \gamma V_{\mathrm{ee},w}[n], \quad \gamma \geq 1 \tag{4.10}$$

and setting $\gamma \to 1/\gamma$ yields results for $\gamma \leq 1$.

Next, we turn to the KS scheme, used in modern EDFT approaches. Here $F_w[n] = T_{\mathrm{S},w}[n] + E_{\mathrm{HXC},w}[n]$ where $T_{\mathrm{S},w}$ is the KS kinetic energy and $E_{\mathrm{HXC},w}$ is the Hartree-exchange-correlation. Because there is no interaction,

$$T_{\mathrm{S},w}[n_\gamma] = \gamma^2 T_{\mathrm{S},w}[n]. \tag{4.11}$$

Moreover, because the Hartree-exchange is linear in the scaling parameter:

$$E_{\mathrm{HX},w}[n_\gamma] = \gamma E_{\mathrm{HX},w}[n]. \tag{4.12}$$

In EDFT, separation of Hartree from exchange is more complicated than in ground-state DFT. [90, 334, 95] Subtracting these larger energies following the usual procedure from ground-state DFT [178] yields, for $\gamma \geq 1$,

$$T_{\mathrm{C},w}[n_\gamma] \leq \gamma^2 T_{\mathrm{C},w}[n], \quad E_{\mathrm{C},w}[n_\gamma] \geq \gamma E_{\mathrm{C},w}[n], \quad \gamma \geq 1 \tag{4.13}$$

where $E_{\mathrm{C},w}[n]$ is the correlation energy, and $T_{\mathrm{C},w} = T_w - T_{\mathrm{S},w}$ is its kinetic contribution.

Considering $\gamma = 1 + \epsilon$ in Eq.4.13, and taking $\epsilon \to 0$, yields differential versions of Eq. 4.13:

$$\frac{d}{d\gamma}\left\{\frac{T_{\mathrm{C},w}[n_\gamma]}{\gamma^2}\right\} \leq 0, \qquad \frac{d}{d\gamma}\left\{\frac{E_{\mathrm{C},w}[n_\gamma]}{\gamma}\right\} \geq 0 \tag{4.14}$$

Combining these using Nagy's generalization (Eq. 24 of Ref. 213) of the ground-state equality

$$\left.\frac{dE_{\mathrm{C},w}[n_\gamma]}{d\gamma}\right|_{\gamma=1} = E_{\mathrm{C},w}[n] + T_{\mathrm{C},w}[n], \tag{4.15}$$

we find

$$\left(2 - 2\gamma\frac{d}{d\gamma} + \gamma^2\frac{d^2}{d\gamma^2}\right) E_{\mathrm{C},w}[n_\gamma] \leq 0, \tag{4.16}$$

the condition for concavity in the ensemble correlation energy. This is the ensemble form of Eq. 40 in Ref. 179. Eqs. 4.9, 4.13, and 4.16 are primary results of the current work, being the ensemble generalizations of their ground-state analogs.

An immediate application of Eq. 4.12 is to extract the HX component from any HXC approximation. As the conditions limit growth with γ,

$$E_{\mathrm{HX},w}[n] = \lim_{\gamma\to\infty} E_{\mathrm{HXC},w}[n_\gamma]/\gamma, \tag{4.17}$$

an exact condition which can prove useful for separating HX from C components. [93, 90]

To conclude this section, we use the pioneering relationship between coupling constant and coordinate scaling. Defining λ dependence via

$$F_w^\lambda[n] = \min_{\Gamma_w\to n} \mathrm{Tr}\{\Gamma_w(\hat{T} + \lambda\hat{V}_{\mathrm{ee}})\}, \tag{4.18}$$

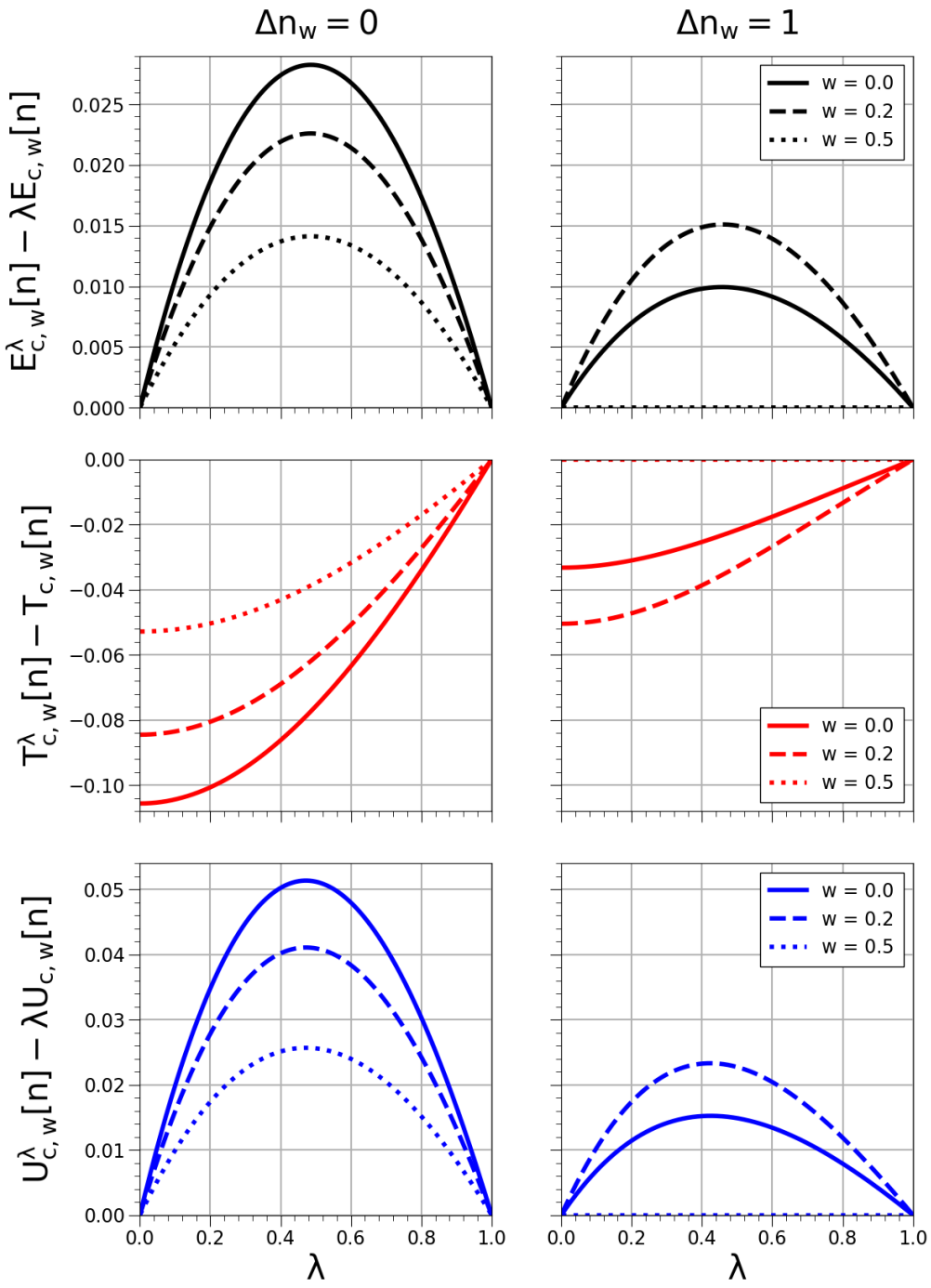

Figure 4.2: Correlation inequalities (Eq. 4.21) for the total (top), kinetic (middle), and potential (bottom) correlation energies, depicted by varying λ in the Hubbard dimer bi-ensemble with $U = 1$. More cases are provided in Figs. S6-S7 of the supplemental material.

Nagy showed [212]

$$E^{\lambda}_{\mathrm{HXC},w}[n] = \lambda^2 E_{\mathrm{HXC},w}[n_{1/\lambda}]. \tag{4.19}$$

Using Eq. 4.19, it is possible to rewrite all results given in terms of scaled densities as λ-dependent relations. Such relations are well known and much used in ground-state DFT, via the adiabatic connection formalism. [117, 62] For real-space Hamiltonians, these relations are simply a rewriting of the scaling relations in a more popular form, but they also apply to lattice Hamiltonians, where scaling is not possible. Converting from scaling in Eq. 4.15 gives

$$T^{\lambda}_{\mathrm{C},w}[n] = E^{\lambda}_{\mathrm{C},w}[n] - \lambda \frac{dE^{\lambda}_{\mathrm{C},w}[n]}{d\lambda}. \tag{4.20}$$

The scaling inequalities (Eqs. 4.13) become

$$T^{\lambda}_{\mathrm{C},w}[n] \leq T_{\mathrm{C},w}[n], \quad E^{\lambda}_{\mathrm{C},w}[n] \geq \lambda E_{\mathrm{C},w}[n], \quad \lambda \leq 1, \tag{4.21}$$

with differential versions

$$\frac{dT^{\lambda}_{\mathrm{C},w}[n]}{d\lambda} \leq 0, \quad E^{\lambda}_{\mathrm{C},w}[n] \geq \lambda \frac{dE^{\lambda}_{\mathrm{C},w}[n]}{d\lambda}, \tag{4.22}$$

while Eq. 4.16 becomes quite simply:

$$\frac{d^2E^{\lambda}_{\mathrm{C},w}[n]}{d\lambda^2} \leq 0. \tag{4.23}$$

Note that all inequalities for $E_{\mathrm{C},w}$, both coordinate-scaled (Eqs. 4.13, 4.14) and λ-dependent (Eqs. 4.21, 4.22), are also true for $U_{\mathrm{C},w} = E_{\mathrm{C},w} - T_{\mathrm{C},w}$, the potential contribution to correlation. The HX energy (Eq. 4.17) may be extracted via

$$E_{\mathrm{HX},w}[n] = \lim_{\lambda\to 0} E^{\lambda}_{\mathrm{HXC},w}[n]/\lambda. \tag{4.24}$$

Our last condition concerns the relationship between DFT and traditional approaches to quantum chemistry. In the ground state, it has long been known [47, 189] that $0 \geq E_{\mathrm{C}}^{\mathrm{HF}} \geq E_{\mathrm{C}}$, where $E_{\mathrm{C}}^{\mathrm{HF}}$ is the traditional definition of the correlation energy, i.e., relative to the Hartree-Fock (HF) energy (we treat only restricted HF here, RHF). Given the complications of EDFT, we discuss here only the case of the first singlet bi-ensemble for two electrons. In this case, we equate EHF with an EDFT EXX calculation ('exact exchange only'). The only difference between EHF and EDFT is that the EHF quantities are evaluated on the approximate EHF density, while EDFT quantities are evaluated on the exact density. Exactly the same variational reasoning leads us to

$$0 \geq E_{\mathrm{C},w}^{\mathrm{HF}} \geq E_{\mathrm{C},w}[n] \tag{4.25}$$

where $E_{\mathrm{C},w}^{\mathrm{HF}} = E_w - E_w^{\mathrm{HF}}$, and E_w^{HF} minimizes $F_w = T_{\mathrm{S},w} + E_{\mathrm{HX},w}$. We leave the more general case to braver souls.

The Hamiltonian of the Hubbard dimer is

$$\hat{H} = -t\sum_{\sigma}(\hat{c}^{\dagger}_{1\sigma}\hat{c}_{2\sigma} + h.c.) + U\sum_{i}\hat{n}_{i\uparrow}\hat{n}_{i\downarrow} + \sum_{i} v_i\hat{n}_i, \tag{4.26}$$

where t is the hopping parameter, U the on-site electrostatic self-repulsion, and v_i the on-site potential (which controls the asymmetry of the dimer). For this lattice system, with $N = 2$, the electronic density is characterized by a single number, the difference between occupations of the two sites, $\Delta n = n_2 - n_1$. The λ-dependence of any quantity is found by replacing U by λU, keeping Δn fixed. We choose $t = 1/2$ everywhere.

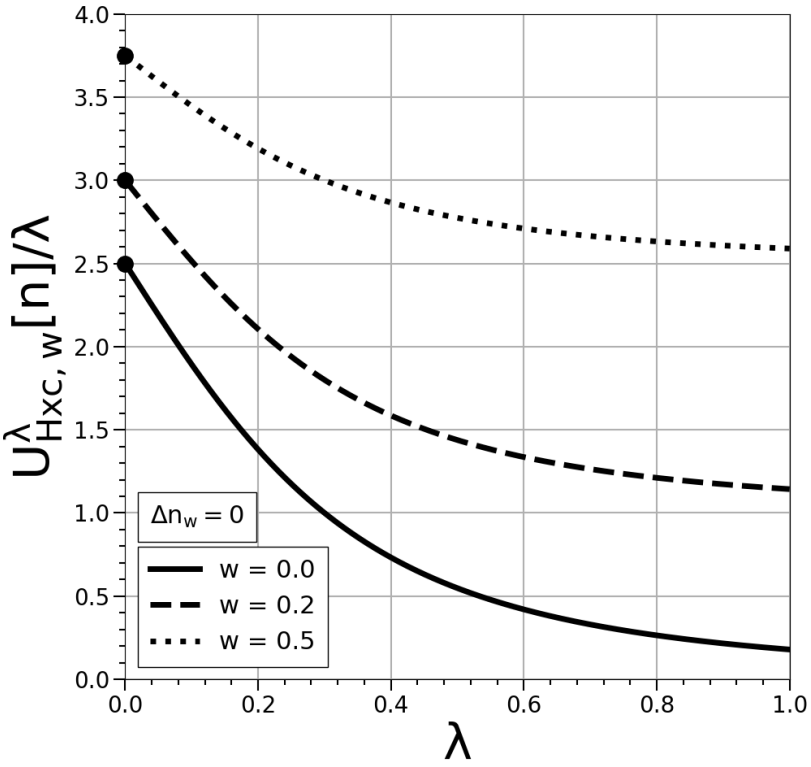

Figure 4.3: Ensemble adiabatic connection with $\Delta n_w = 0$ and $U = 5$; circles represent the weight-dependent HX energy, which the HXC expression approaches as $\lambda \to 0$ (Eq. 4.24). More cases are provided in Fig. S8 of the supplemental material.

We consider the simplest bi-ensemble, a mixture of the ground-state with the first excited singlet. Full analytic expressions of $|\Psi_0\rangle$ and $|\Psi_1\rangle$, as well as plots of various bi-ensemble quantities, are given in the supplemental material in section 1. The value of Δn_w is constrained by w:

$$|\Delta n_w| \leq 2\overline{w}, \tag{4.27}$$

where $\overline{w} = 1 - w$, i.e. is smaller than that of the ground state ($w = 0$). Densities are shown in Fig. S1 of the supplemental material. The total energy of the ensemble is defined as

$$E_w = \overline{w} \langle \Psi_0 | \hat{H} | \Psi_0 \rangle + w \langle \Psi_1 | \hat{H} | \Psi_1 \rangle . \tag{4.28}$$

Plots of E_w are depicted in Figs. S2-S3 of the supplemental material, showing the quantity both as a function of Δv and Δn_w. We also show analogous plots of $F_w = E_w - \Delta v \Delta n_w/2$ in Figs. S4-S5. For this bi-ensemble, the exact HX energy has the simple analytical form [52]:

$$E_{\mathrm{HX},w} = \frac{U}{2}\left[1 + w + \frac{(1-3w)}{\overline{w}^2}\frac{\Delta n_w^2}{4}\right]. \tag{4.29}$$

4.4.1 Inequalitites

We plot λ-dependencies in Fig. 4.2 that have a definite sign according to Eq. 4.13. We show several values of w for two densities for $U = 1$ (moderate correlation). Scanning over all U and Δn_w, these inequalities are always satisfied. For the symmetric dimer ($\Delta n_w = 0$), $w = 0.0$ has the largest maximum and $w = 0.5$ has the smallest. As Δn_w is increased, the trend disappears and the curves are not monotonic in w. For $w = 0.5$, the inequality becomes an equality for $\Delta n_w = 1$, the maximum representable value of Δn for the ensemble. In Fig. 4.3, we show that all $U_{\mathrm{HXC},w}$ curves approach their corresponding HX value as $\lambda \to 0$, in accordance with Eq. 4.24. More plots of Figs. 4.2 and 4.3 for various combinations of w and Δn_w are provided in Figs. S6-S8 of the supplemental material.

The non-monotonic behavior in Fig. 4.1 can be easily understood. By definition, $E_w(\Delta v)$ is linear in w, as is F_w. But, when converted to density functionals, and with KS quantities subtracted, these become highly non-monotonic, as shown in Figs. S2 and S3 in the supplemental material.

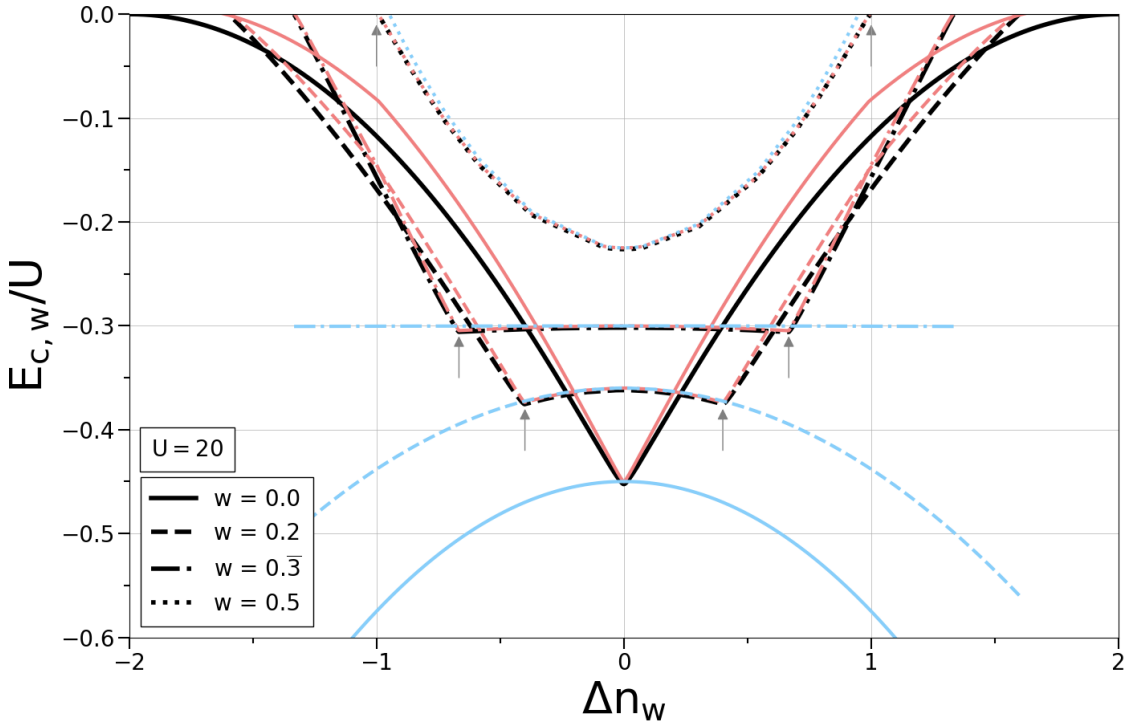

Figure 4.4: Exact correlation energy (black), leading-order expansion in large U (red) and the expansion in the symmetric limit (blue) for the correlation energy are all plotted as a function of the exact density. Small arrows indicate the region between $2w$ and $-2w$ where the symmetric expansion matches the exact.

4.4.2 Strong Correlation

In Fig. 4.4 we plot the exact correlation energy, our approximation (Eq. S.16 of the supplemental material), and the symmetric limit expansion of Deur *et al.* [53], each evaluated at the exact density. The last yields the strongly correlated correlation energy only for $|\Delta n_w| \leq 2w$, but our expansion yields the correct limit for all allowed Δn_w, including the slope discontinuity at $|\Delta n_w| = 2w$. Such w-dependent derivative discontinuities occur only in EDFT. The approximate weight-dependent strongly correlated correlation energy is derived in the supplemental material along with further analysis of the energy components and approximation of the density. For the strong-interaction limit of the dimer, the correlation energy contains a non-trivial weight-dependence. This differs from real space [90] where the energies were found to be weight-independent. This is not a counter example, because the

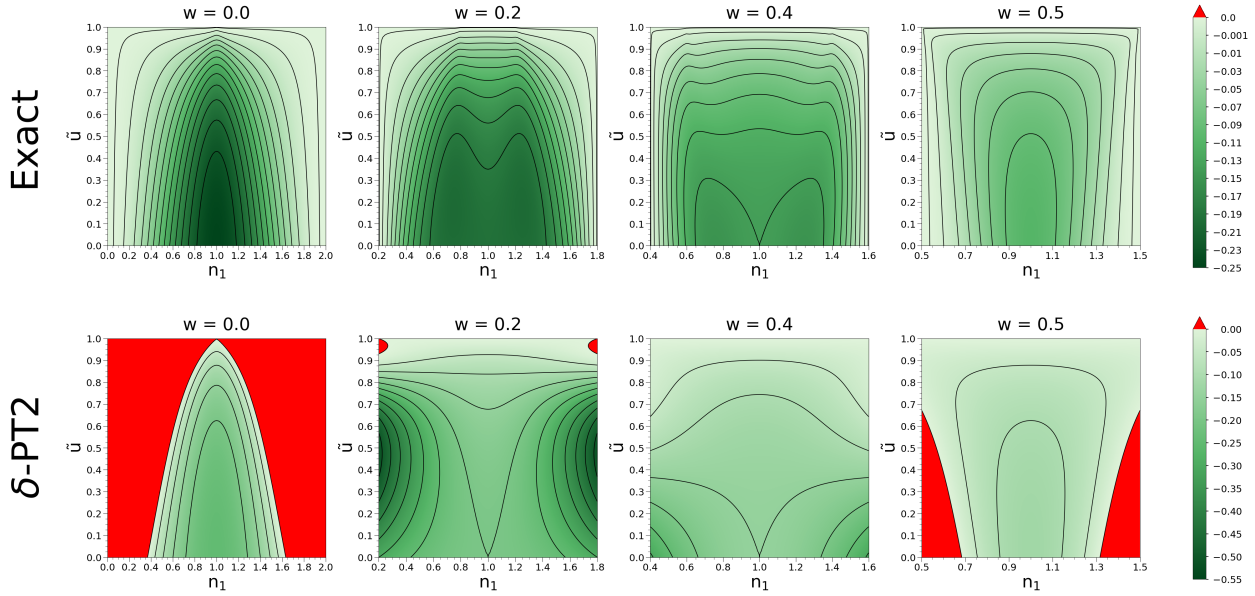

Figure 4.5: The second derivative of the Hubbard dimer bi-ensemble correlation energy with respect to U for all values of the reduced variable $\tilde{u} = U/\sqrt{1+U^2}$. The concavity condition is satisfied, but is violated by the δ-PT2 approximation in certain regimes (red denotes positive). But U-PT2 automatically satisfies it, by construction.

dimer is a site-model. This difference manifests in the expansion of the strongly correlated energies in powers of the coupling-constant. Our first correction, relative to the leading term, is $\mathcal{O}(\lambda^{-2})$ and differs qualitatively from the $\mathcal{O}(\lambda^{-1/2})$ behavior found by Gould and coauthors.

4.4.3 Concavity Condition of the Correlation Energy

We illustrate the concavity condition of Eq. 4.23 using contour plots depicting all possible combinations of U and n_1, making use of the reduced variable $\tilde{u} = U/\sqrt{1+U^2}$. Illustrated by Fig. 4.5, the second derivative is negative for all values of U and thus satisfies the concavity condition for all electronic correlation strengths.

The standard use of exact conditions in DFT is to ensure that approximate functionals satisfy them. [233] We illustrate our conditions by applying them to existing approximations on the

Hubbard model. The first is the standard many-body expansion in powers of the interaction, U, which we perform up to 2nd-order, i.e., the analog of Møller-Plesset perturbation theory, denoted U-PT2. The second is less familiar: an expansion in powers of Δn around the symmetric case, $\Delta n = 0$, called δ-PT2. [53] This can be considered a (tortured) analog of the gradient expansion of DFT, [240] as it is an expansion around the uniform limit. Fig. 4.5 shows that the δ-PT2 approximation violates the concavity condition, even for $w = 0$, while U-PT2 never does. The violations are not monotonic with increasing weights, as $w = 0.4$ has none. Deur *et al.* reported that, compared to U-PT2, δ-PT2 produced more accurate equi-ensemble energies and densities. Likely, the accuracy of δ-PT2 could be further improved by imposing concavity. Recent advances in EDFT, such as the direct ensemble correction [336] and the perturbative EDFT method, [89] are explicitly computed in the perturbative limit, $w \to 0^+$. If an approximation is derived *before* such a limit is taken, and its ground-state approximation satisfies concavity; the resulting approximation should satisfy concavity also.

4.4.4 Quantum Chemistry

Finally, we examine in detail the difference between the DFT and HF correlation energies and their components in Figs. S5, S6, and S7 in the supplemental material. We provide plots of the exact/EHF total correlation energies for the dimer bi-ensemble, where we show the ground-state inequalities ($E_{\rm C}^{\rm HF} \geq E_{\rm C}$) holds for any w-value (Figs. S3 and S4). It also is known that $T_{\rm C}^{\rm HF}$ can become negative in the ground state of the Hubbard dimer, [78] and we find this is also true when $w \neq 0$, but this is likely an artifact of lattice Hamiltonians that cannot occur in the real-space analog.[48, 47]

4.5 Conclusion

This work provides new exact conditions for EDFT which can be used to analyze and/or improve new approximations in EDFT. Further work is being performed to improve approximations and provide a pathway to accurate EDFT functionals.

We provide analytical expressions (and plots) for various weight-dependent quantities of interest for the Hubbard dimer bi-ensemble, giving both the exact/EHF solution, in the supplemental material.

4.6 Acknowledgements

T.R.S. thanks the Molecular Software Science Institute for funding, award no. 480718-19905B, and Sina Mostafanejad for fruitful discussions. K.B., J.K., and S.C. acknowledge support from NSF award number CHE-2154371. A.P.J. acknowledges financial support for this publication from Cottrell Scholar Award #28281, sponsored by Research Corporation for Science Advancement.

Chapter 5

Generalized Gradient Approximation Made Thermal

This chapter is a reproduction of Ref. [162], which I co-authored with Dennis Perchak and Kieron Burke.

5.1 Abstract

Using the methodology of conditional-probability density functional theory, and several mild assumptions, we calculate the temperature-dependence of the Perdew–Burke–Ernzerhof (PBE) generalized gradient approximation (GGA). This numerically-defined thermal GGA reduces to the local approximation in the uniform limit and PBE at zero temperature, and can be fit reasonably accurately (within 8%) assuming the temperature-dependent enhancement is independent of the gradient. This locally thermal PBE satisfies both the coordinate-scaled correlation inequality and the concavity condition, which we prove for finite temperatures. The temperature dependence differs markedly from existing thermal

GGA's.

5.2 Introduction

Density functional theory (DFT) is utilized throughout modern science, wherever electronic structure is important, and has enormous impact in materials simulations [137, 107, 251, 298] and quantum chemistry [293, 221, 332]. Almost all such calculations employ the Kohn-Sham (KS) scheme [157], in which only the exchange-correlation (XC) energy need be approximated as a functional of electronic (spin) densities [57]. The quality of results generated depends crucially on that approximation, and many hundreds of such approximations are readily available in modern codes [169].

Warm dense matter (WDM) includes nuclear fusion at the national ignition facility, interiors of gas giant planets, and matter under extreme shock conditions at the Sandia Z-machine or SLAC's free electron laser [165, 217, 216, 326]. Simulations of WDM have been greatly improved by the use of KS-DFT at finite temperature [128, 147, 260, 207, 55, 123]. Mermin showed that an equilibrium grand canonical density and free energy can be found this way, but the unknown XC free energy is now also a function of temperature [206]. Most successful simulations performed today, however, use existing ground-state XC approximations. While their success can be understood, the missing thermal dependence is an uncontrolled error that may significantly impact their results.

At the local density approximation (LDA) level [157], there is a long history of parameterizing the XC free energy of a uniform electron gas [144, 106, 143, 145] as a function of temperature. But modern materials simulations require at least a generalized gradient approximation (GGA) level of accuracy to achieve chemical specificity, which is why the Perdew–Burke–Ernzerhof (PBE) approximation is used in most such simulations. The un-

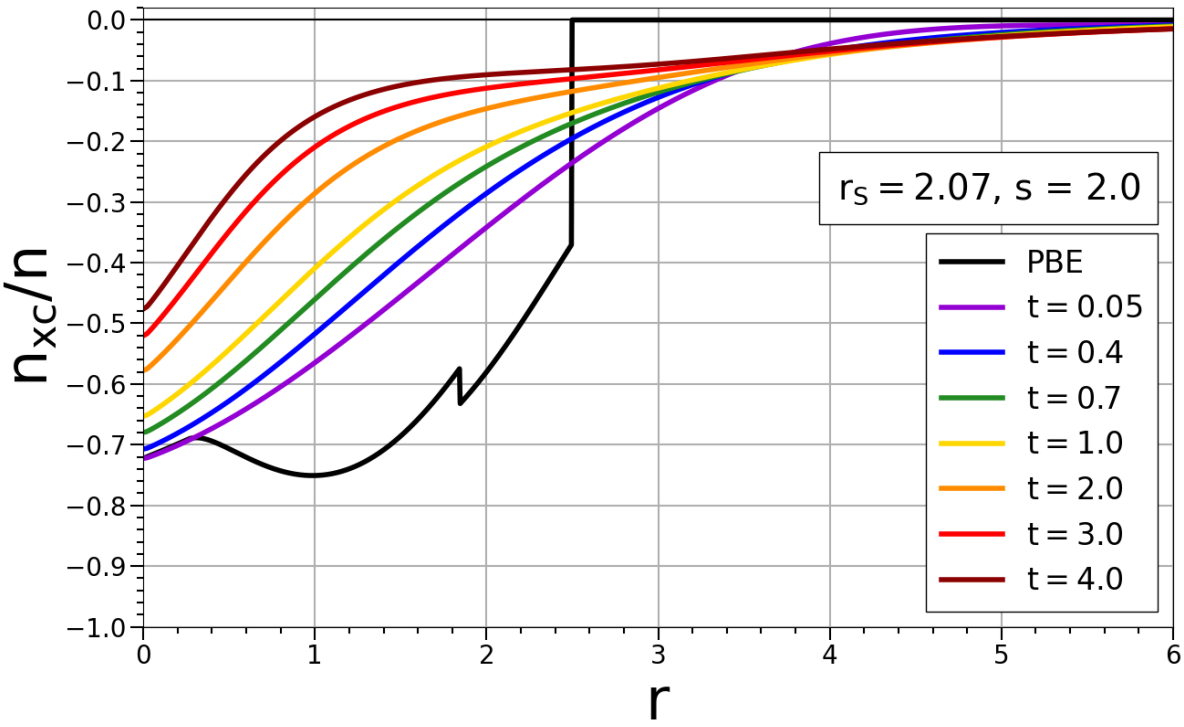

Figure 5.1: Typical temperature-dependent XC hole densities. Black denotes the ground-state real-space cutoff GGA hole density, while the CP-DFT XC hole at different reduced temperatures is depicted in various colors. As $t \to 0$, CP-DFT yields an XC hole density with the same on-top value and energy as PBE. See text for definitions.

derlying rationale for the PBE functional is a detailed model of the XC hole, and the exact conditions used in its derivation were chosen to (approximately) recover the numerical results of imposing exact conditions on the hole, not the energy directly.

The present work gives the result of calculating the XC hole of PBE as a function of temperature, using the recently-invented conditional-probability (CP) density functional theory [235, 204, 237] (see Refs. 83, 84, 85 for its precursors). In this procedure one first finds, at zero temperature, the CP potential that generates the PBE hole at a given density and gradient. Then, by ignoring the temperature dependence of the CP potential, one can calculate the XC hole using finite-temperature KS-DFT. Throughout this work, the Wigner-Seitz radius $r_s = (3/4\pi n)^{1/3}$, the dimensionless density gradient $s = |\nabla n|/2k_F n$ with Fermi wavevector $k_F = (3\pi^2 n)^{1/3}$, and the reduced temperature $t = T/T_F$ with Fermi temperature

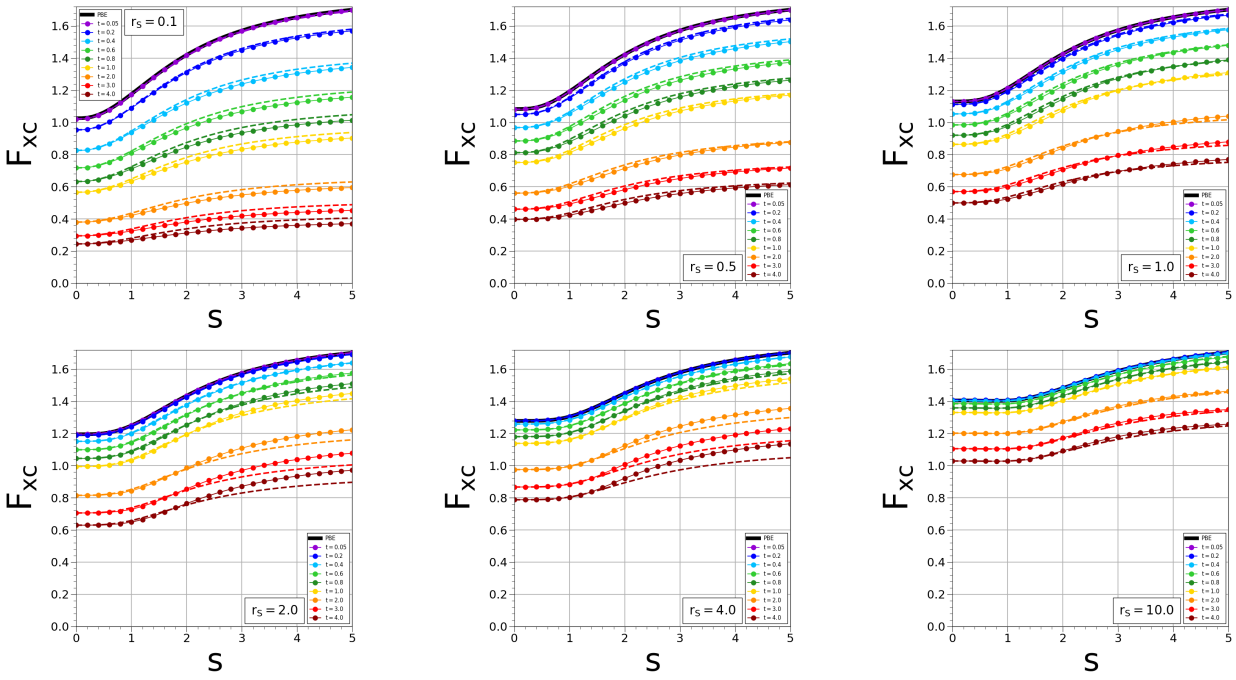

Figure 5.2: Temperature-dependent XC enhancement factors plotted as functions of the dimensionless gradient s for unpolarized systems of various r_s values. Here we denote the results of CP-DFT calculations at different temperatures using data points of various colors, while the ltPBE approximation (Eq. 5.1) is represented by dashed curves. The ground-state PBE XC enhancement factor is shown in black, which the CP-DFT results approach as $t \to 0$. More plots (including those for fully polarized systems) are shown in Figs. S1 and S2 of the supplemental material.

$T_F = k_F^2/2k_B$. Fig. 5.1 illustrates our work, showing the XC hole of aluminum ($r_\mathrm{s} = 2.07$) with a nonzero gradient at various temperatures. This result not only demonstrates the behavior of PBE's XC hole at WDM conditions, but also graphically depicts the diminishing effect temperature has on the XC free energy, a quantity directly related to the XC hole via integration. By construction, the resulting (numerically-defined) free energy GGA reduces to PBE at zero temperature. Moreover, very little error is introduced by ignoring temperature dependence in the CP potential, as has already been shown for the uniform gas [204].

An important second step is to create a simple parameterization of the resulting functional, so that it can be easily implemented and utilized. We find that the simplest possibly physical

realistic case (referred to throughout this work as *locally thermal PBE*)

$$F_{\mathrm{XC}}^{\mathrm{ltPBE}}(r_{\mathrm{S}}, s, t) = \frac{F_{\mathrm{XC}}^{\mathrm{unif}}(r_{\mathrm{S}}, t)}{F_{\mathrm{XC}}^{\mathrm{unif}}(r_{\mathrm{S}})} \times F_{\mathrm{XC}}^{\mathrm{PBE}}(r_{\mathrm{S}}, s), \tag{5.1}$$

matches the output from our CP-DFT calculation to within a few percent, examples of which are shown in Fig. 5.2 for unpolarized systems of various r_{S} values. Such a simple parameterization automatically satisfies several key exact conditions, and is trivial to implement in existing codes. Thus, the effects of these new thermal corrections can be immediately tested in any WDM simulation currently making use of PBE.

A prior attempt to incorporate finite temperature-dependence makes use of exact conditions [143], but yields a strikingly different temperature dependence.

5.3 Background

CP-DFT makes use of an approximate potential (referred to as the CP potential [235]) at each point in space to calculate the conditional probability density, which can be integrated to yield the XC energy via

$$E_{\mathrm{XC}} = \frac{1}{2}\int_0^1 d\lambda \int d^3r \int d^3r' \, \frac{n(\mathbf{r})\left[\tilde{n}_{\mathbf{r}}^{\lambda}(\mathbf{r}') - n(\mathbf{r})\right]}{|\mathbf{r}-\mathbf{r}'|}, \tag{5.2}$$

where $\tilde{n}_{\mathbf{r}}^{\lambda}(\mathbf{r}')$ is the conditional probability density (the probability of finding an electron at $\mathbf{r}'$, given already having found one at $\mathbf{r}$), and $n_{\mathrm{XC}}(\mathbf{r},\mathbf{r}') = \int_0^1 d\lambda \left[\tilde{n}_{\mathbf{r}}^{\lambda}(\mathbf{r}') - n(\mathbf{r})\right]$ is the coupling-constant averaged XC hole density. Thus, the XC hole of a system may be written as the difference between the CP density and standard electronic density, each of which are determined self-consistently using independent KS-DFT calculations. The reliability of this approach is readily explained through this feature - the extraction of XC free energies via Eq. 5.2 eliminates functional error from our energies, as no approximate XC density

functional need be evaluated [317]. The quality of our result is directly attributed to the accuracy of the two self-consistent densities, which have previously been shown to be highly accurate (even when approximate functionals are utilized) [151]. In this work, we use this methodology to generate a thermal GGA.

The CP potential for an unpolarized uniform electron gas of $N-1$ electrons may be written as [204]

$$v_{\mathrm{S}}(\mathbf{r}) = \Delta\tilde{v}(r) + \int d^3r' \, \frac{\tilde{n}(r') - n_0}{|\mathbf{r} - \mathbf{r}'|} + v_{\mathrm{XC}}^{\mathrm{LDA}}[\tilde{n}](r), \tag{5.3}$$

where $n_0 = N/V$ is a constant, and

$$\Delta\tilde{v}(r) = \frac{1}{2r}\left[1 + \mathrm{erf}\left(\frac{r}{r_{\mathrm{S}}}\right)\right] + A(r_{\mathrm{S}}, s)e^{-r^2/2\sigma(r_{\mathrm{S}},s)^2} \tag{5.4}$$

approximates the effect of removing one electron from the system. The first term of this approximation [204, 83] classically approximates the effect of the missing electron, while the second term works to recreate the correct high-density limit for uniform ($s = 0$) systems.

In our procedure, we modify this Gaussian term to approximate the CP potential of non-uniform systems, using it as a means of recreating the characteristic qualities of the PBE XC hole. The amplitude and width of a given Gaussian is numerically fit for each r_{S} and s value, such that the result of a CP-DFT calculation yields an XC hole with the same (1) on-top value and (2) XC energy as PBE's hole as the reduced temperature t approaches zero. For fitting purposes, we generate PBE's XC hole using the damped, numerical procedure in Refs. 242 & 28. This makes use of the second-order gradient expansion of the XC hole, but enforces numerical cut-offs to ensure satisfaction of the exchange/correlation sum rules. In principle, the exact CP potential is temperature-dependent, but we expect this temperature dependence to have little effect, vanishing in the high temperature limit. In the uniform limit [204], this approximation yields errors of order 5%.

There exists no equivalent CP potential approximation (Eq. 5.4) for polarized systems, and the derivation of one is beyond the scope of this work. In its place, we generate data for fully polarized systems by making use of the exchange hole's spin-scaling relation $n_{\rm x}^{\rm pol}[n](\mathbf{r},\mathbf{r}') = n_{\rm x}^{\rm unpol}[2n](\mathbf{r},\mathbf{r}')$. Thus, we are able to recover the polarized exchange hole exactly through scaling, but miss the correlation hole, which must be approximated with an added corrective potential such that Eq. 5.4 becomes:

$$\Delta\tilde{v}(r) = \frac{1}{2r}\left[1 + \mathrm{erf}\left(\frac{r}{r_{\rm s}}\right)\right] + A_1 e^{-r^2/2\sigma_1^2} + A_2 e^{-r^{3/4}/2\sigma_2^2} \tag{5.5}$$

We note that A_1 and σ_1 are found first, in the same manner as before (but this time using a scaled density). The last term of Eq. 5.5 is then introduced (via a second numerical search to find A_2 and σ_2) so that the characteristics of the polarized version of PBE's XC hole are recreated. All numerical parameters (A_1, A_2, σ_1, σ_2) are functions of both $r_{\rm s}$ and s, and given as .csv files in the supplemental material. Note the Fermi temperature of a fully polarized system $T_F^{\rm pol} = 2^{2/3} T_F^{\rm unpol}$.

Fig. 5.2 depicts CP-DFT results for unpolarized systems of $r_{\rm s} = 0.1$, 0.5, 1, 2, 4, 10 and $0.05 \leq t \leq 4$; equivalent calculations have been performed for fully polarized systems of $r_{\rm s} = 0.1$, 0.5, 1, 2, 4. The resulting XC hole densities and XC enhancement factors for all CP-DFT calculations are plotted in the first section of the supplemental material, the latter also as .csv files.

The output of the CP-DFT procedure, applied to the numerically-defined cut-off GGA hole, yields a well-defined thermal GGA, which we refer to as CPTGGA (conditional probability thermal GGA). The results depend weakly on the specific choices made here. We anticipate both refining those choices and carefully parameterizing the output in the future. For the present, the locally thermal PBE defined by Eq. 5.1 captures most of the temperature-dependence in CPTGGA. Moreover, this simple prescription can be applied to any approxi-

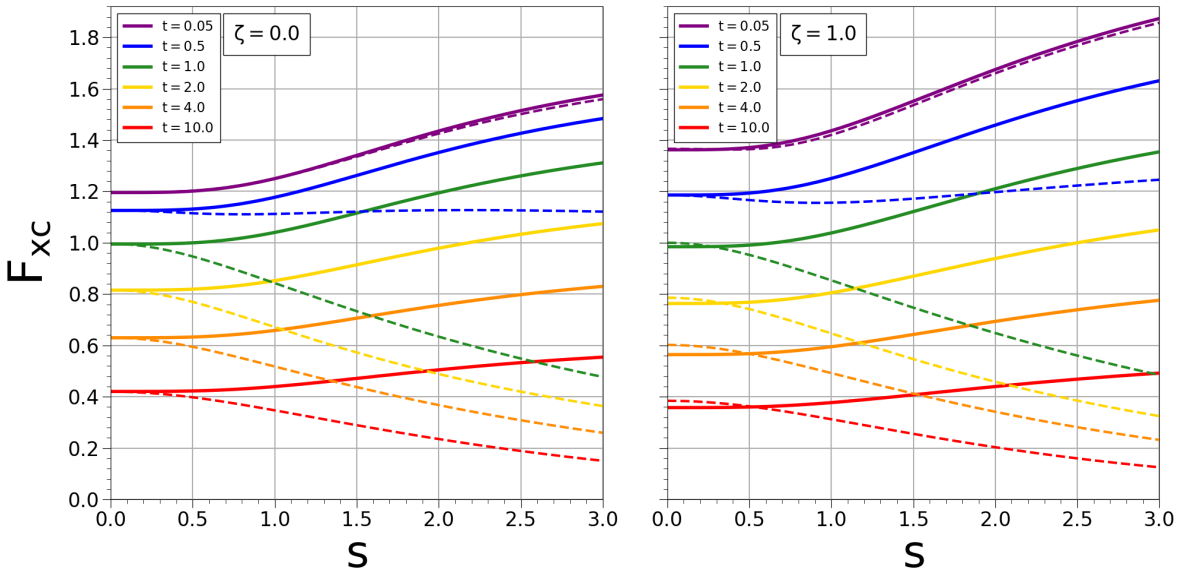

Figure 5.3: Temperature-dependent XC enhancement factors plotted as functions of the dimensionless gradient s. Here we set $r_\mathrm{S} = 2$, showing the results for both unpolarized (left) and fully polarized (right) systems. Solid curves represent $F_\mathrm{XC}^\mathrm{ltPBE}$, while dashed curves represent $F_\mathrm{XC}^\mathrm{KDT16}$.

mate ground-state functional, as a suggested temperature dependence, which could then be tested.

5.4 Analysis

We begin by emphasizing again that any conclusions drawn from this work are based on a temperature-dependent model for PBE's XC hole, which is shown in Fig. 5.1 for an unpolarized system with $r_\mathrm{S} = 2.07$ and $s = 2$. As t is increased, it is clear that the negative on-top value of the hole (and its surrounding density) monotonically increases, resulting in smaller XC free energies. Although the hole shrinks with temperature, the sum rule $\int n_\mathrm{XC}(\mathbf{r}, \mathbf{r}+\mathbf{r}')\, d^3r' = -1$ is satisfied by the CP-DFT XC hole for all temperatures by construction.

When discussing GGA XC free energies, it is useful to define the temperature-dependent

XC enhancement factor F_{XC} as a multiple of the ground-state uniform electron gas exchange energy:

$$A_{\mathrm{XC}}(r_{\mathrm{s}},\zeta,s,t) = \int d^3r\, \epsilon_{\mathrm{X}}^{\mathrm{unif}}(n)\, F_{\mathrm{XC}}(r_{\mathrm{s}},\zeta,s,t) \tag{5.6}$$

Here A_{XC} is the XC free energy, $\zeta = (n_\uparrow - n_\downarrow)/n$ is the relative spin polarization, and $\epsilon_{\mathrm{X}}^{\mathrm{unif}}(n) = -3k_F/4\pi$. This allows us to not only perform direct comparisons to the ground-state PBE approximation, but also to the thermal GGA previously proposed by Karasiev *et al.* (referred to here as KDT16) [143]. Throughout this work, we implement ltPBE via Eq. 5.1 using the thermal LDA parameterization proposed by Groth *et al.* [106], which is given for all ζ.

In Fig. 5.3 we plot XC enhancement factors as functions of the dimensionless gradient s, comparing ltPBE to KDT16 at various temperatures. First, note that both GGA's converge to PBE as $t \to 0$, with negligible differences for $t = 0.05$; in the low temperature limit, both GGA's predict F_{XC} to monotonically increase with respect to the gradient. Similarly, both GGA's approach nearly the same F_{XC} value as $s \to 0$, as both reduce to their corresponding thermal LDA (which have noticeable differences for $\zeta = 1$) in this limit. The two approximations however predict quite different behaviors for nonzero temperature/gradient values. Although KDT16 recreates the monotonically increasing behavior of PBE at $t = 0$, this trend flips and becomes monotonically decreasing with density gradient for warm/hot temperatures. In contrast, ltPBE by definition is a simple temperature-dependent multiple of the ground-state $F_{\mathrm{XC}}^{\mathrm{PBE}}$ curve, and thus inherits its curvature. As the temperature is increased further, the importance of the gradient on the XC free energy becomes less prominent, so both approximations vary less with s. Thus, there are significant differences between the approximations in the warm dense matter regime, particularly for systems with large gradients.

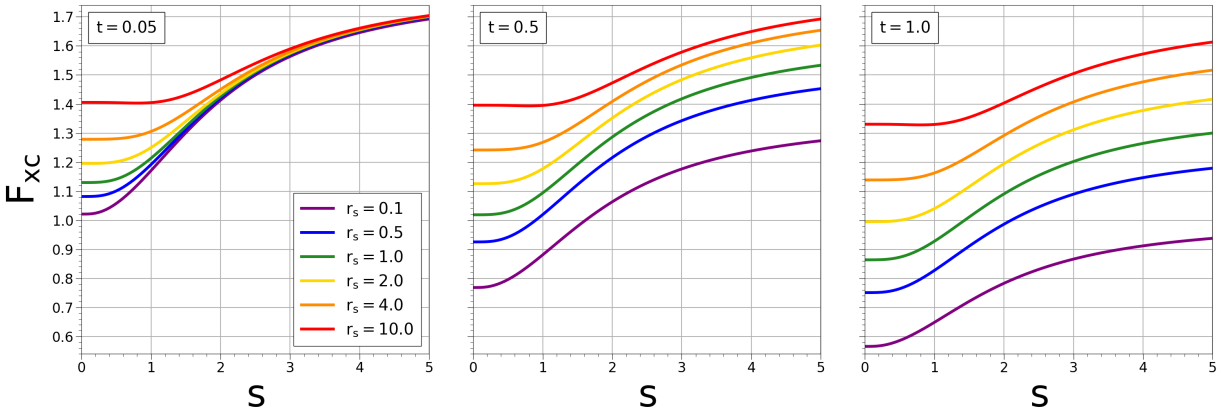

Figure 5.4: Several ltPBE XC enhancement factors plotted as functions of the dimensionless gradient s for unpolarized systems of various r_s values, with each plot having a different fixed reduced temperature t. The thermal coordinate scaling inequality (Eq. 5.8) mandates these curves not cross for any t, which the ltPBE approximation is shown to satisfy.

We now turn our attention to two exact conditions, both of which were omitted in the formulation of the "Strongly Constrained and Appropriately Normed" (SCAN) semilocal functional [288, 234], and are readily generalizable to finite temperatures. Using uniform coordinate scaling inequalities at nonzero temperature, the XC free energy has been shown to scale according to the relation [253]

$$A_{\mathrm{XC}}^{T}[n_\gamma] \geq \gamma A_{\mathrm{XC}}^{T/\gamma^2}[n], \qquad (\gamma \geq 1) \tag{5.7}$$

where γ represents the coordinate scaling parameter, and the scaled density $n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r})$. This results in an exact condition for thermal generalized gradient approximations:

$$F_{\mathrm{XC}}(r_\mathrm{s},\, s,\, t) \geq F_{\mathrm{XC}}(r'_\mathrm{s},\, s,\, t) \qquad (r'_\mathrm{s} \leq r_\mathrm{s}) \tag{5.8}$$

Note that the reduced temperature remains constant here, since $t = T/T_F = 2k_BT/(3\pi^2 n)^{2/3}$ is scaled equally on both sides of Eq. 5.7. This signifies that the XC enhancement factors of different densities should not cross for any chosen reduced temperature. In Fig. 5.4 we plot F_{XC} curves at fixed values of t for the ltPBE approximation, illustrating its satisfaction of

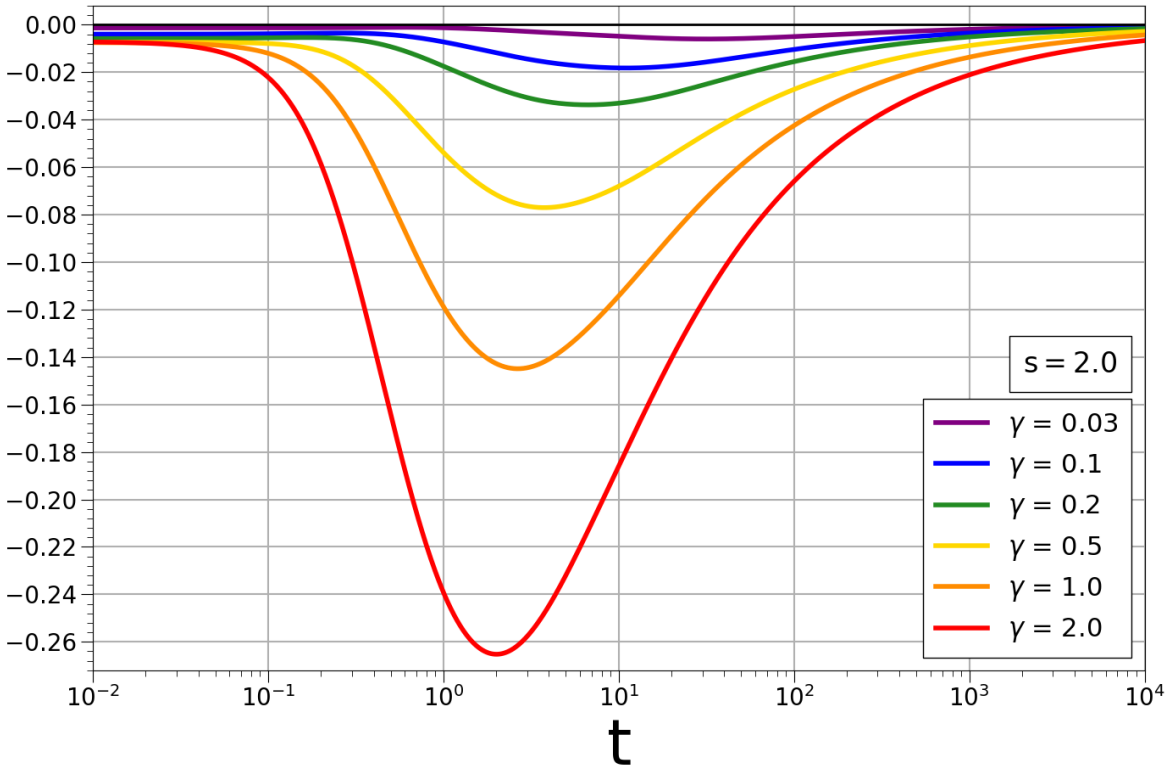

Figure 5.5: The thermal concavity condition inequality (Eq. 5.11) plotted for ltPBE with initial density $n = 1$ and various γ. All curves are negative, indicating the exact condition is satisfied.

this condition. The ratio $F_{\mathrm{XC}}^{\mathrm{unif}}(r_{\mathrm{S}}, t)/F_{\mathrm{XC}}^{\mathrm{unif}}(r_{\mathrm{S}})$ decreases with r_{S}, by inspection. This fact, combined with PBE's satisfaction of Eq. 5.8 at zero temperature [234], guarantees ltPBE satisfies it by construction for all temperatures. Equivalent plots and analysis of KDT16, which is shown to violate this condition, are given in Fig. S5 of the supplemental material.

The second exact condition we test involves the second derivative of the correlation free energy (the concavity condition) with respect to the coordinate scaling parameter γ. First derived as an exact condition of the ground state correlation energy [180], it has recently been extended to ensemble DFT by Scott *et al.* [272]. Here we show the extension of this condition to the correlation free energy of thermal ensembles, where states are occupied using temperature-dependent Fermi weights. To achieve this, we make note of the coordinate

scaling relations [253]

$$K_{\mathrm{C}}^{T}[n_\gamma] \leq \gamma^2 K_{\mathrm{C}}^{T/\gamma^2}[n], \quad A_{\mathrm{C}}^{T}[n_\gamma] \geq \gamma A_{\mathrm{C}}^{T/\gamma^2}[n], \quad (\gamma \geq 1) \tag{5.9}$$

where $A_{\mathrm{C}}^{T}[n]$ is the correlation free energy, and $K_{\mathrm{C}}^{T}[n] = T_{\mathrm{C}}^{T}[n] - TS_{\mathrm{C}}^{T}[n]$ is its kentropic contribution. Considering $\gamma = 1 + \epsilon$ and taking $\epsilon \to 0$, we find differential versions of Eq. 5.9 to be

$$\frac{d}{d\gamma}\left\{\frac{K_{\mathrm{C}}^{T}[n_\gamma]}{\gamma^2}\right\} \leq 0, \quad \frac{d}{d\gamma}\left\{\frac{A_{\mathrm{C}}^{T}[n_\gamma]}{\gamma}\right\} \geq 0. \tag{5.10}$$

Taking these relations in conjunction with the thermal virial theorem (Eq. 23 of Ref. 255), we find the concavity constraint on the correlation free energy

$$\left(2 - 2\gamma\frac{d}{d\gamma} + \gamma^2\frac{d^2}{d\gamma^2}\right) A_{\mathrm{C}}^{T}[n_\gamma] \leq 0. \tag{5.11}$$

In Fig. 5.5, we plot an example of Eq. 5.11 for ltPBE with initial density $n = 1$ and $s = 2$. All curves, regardless of the value of γ, are negative and thus satisfy Eq. 5.11. Contour plots showing more examples for both ltPBE and KDT16 are given in Fig. S6 of the supplemental material. While ltPBE is shown to satisy Eq. 5.11 by inspection, KDT16 is found to only satisfy the condition as $s \to 0$.

While both exact conditions are found to be satisfied by ltPBE, we have checked the first only by inspection for CPTGGA. Noting that the ltPBE free exchange energy is found using a modified version of Eq. 5.1

$$F_{\mathrm{X}}^{\mathrm{ltPBE}}(s,t) = F_{\mathrm{X}}^{\mathrm{unif}}(t) \times F_{\mathrm{X}}^{\mathrm{PBE}}(s), \tag{5.12}$$

where by definition the ground-state $F_{\mathrm{X}}^{\mathrm{unif}} = 1$, we find that ltPBE satisfies all exact conditions built into KDT16, other than the weakly inhomogeneous electron gas limit [143]. Since

PBE violates the gradient expansion, it follows that CPTGGA (and therefore ltPBE) does as well.

5.5 Conclusion

In this work we have generated the temperature-dependence of the PBE ground-state functional through a series of CP-DFT calculations capturing the characteristics of the PBE XC hole at zero temperature. We have provided a simple ansatz (Eq. 5.1) that allows for our temperature-dependent GGA to be readily implemented in standard DFT codes. Analysis of our ltPBE approximation reveals striking differences with previously proposed thermal GGA's. We have shown that this simple approximation satisfies crucial exact constraints of the XC free energy, namely the temperature-dependent XC coordinate scaling inequality (Eq. 5.8) and concavity condition (Eq. 5.11). WDM simulations are ongoing to determine the importance of ltPBE thermal gradient effects, as there may be significant improvements from incorporating the temperature-dependence of the XC energy explicitly. Furthermore, locally thermalized versions of other ground-state functionals may offer significant improvements to the accuracy of WDM simulations, and should be tested in future work.

5.6 Acknowledgements

J.K. and K.B. acknowledge support from NSF award number CHE-2154371. We thank Attila Cangi and Tobias Dornheim for helpful discussions.

Chapter 6

Large scale quantum chemistry with tensor processing units

This chapter is a reproduction of Ref. [236], which I co-authored with Ryan Pederson, Ruyi Song, Jackson Beall, Martin Ganahl, Markus Hauru, Adam GM Lewis, Yi Yao, Shrestha Basu Mallick, Volker Blum, and Guifre Vidal.

6.1 Abstract

We demonstrate the use of Google's cloud-based Tensor Processing Units (TPUs) to accelerate and scale up conventional (cubic-scaling) density functional theory (DFT) calculations. Utilizing 512 TPU cores, we accomplish the largest such DFT computation to date, with 247848 orbitals, corresponding to a cluster of 10327 water molecules with 103270 electrons, all treated explicitly. Our work thus paves the way towards accessible and systematic use of conventional DFT, free of any system-specific constraints, at unprecedented scales.

6.2 Introduction

Computational methods for quantum chemistry and quantum physics have proven to be invaluable tools in modern scientific research and technological innovation. The application space of such methods is vast, ranging from the prediction of novel high-temperature superconductors [58] to the acceleration of drug discovery [39]; from the study of catalytic processes for e.g. CO_2 sequestration [338] and plastic recycling [139] to the design of nanomaterials [19], solar cells [72], and batteries [308].

In the landscape of quantum-based computational methods, density functional theory (DFT) especially stands out due to its ability to produce accurate results for a wide range of systems at a relatively low computational cost [140]. Accordingly, an impressive amount of computational research utilizes DFT calculations each year. For instance, the US National Energy Research Scientific Computing Center (NERSC) reported that nearly 30% of their supercomputer resources in 2018 were spent on DFT calculations alone [7]. Widespread research and development effort is continuously devoted towards optimizing the performance and accuracy of DFT calculations, giving rise to a plethora of open-source and commercial DFT software packages [294]. Several packages can leverage specialized hardware, such as general-purpose graphics processing units (GPUs), for most of the workload [132, 275, 6, 70, 81, 77, 110]. However, in conventional DFT implementations, i.e., without specific sparsity assumptions for the density matrix or Hamiltonian matrix, the computational cost scales as the third power of the number N of orbitals used to describe the system (referred to $O(N^3)$ DFT throughout this work), and this cubic scaling often makes simulating large systems, such as protein-ligand complexes or metal-organic frameworks [331], prohibitively expensive.

Google's Tensor Processing Units (TPUs) are application-specific integrated circuits originally designed to accelerate large-scale machine learning workloads [141, 142, 21, 73, 2]. By leveraging the JAX library [21, 73, 2], it is nevertheless possible to repurpose TPUs

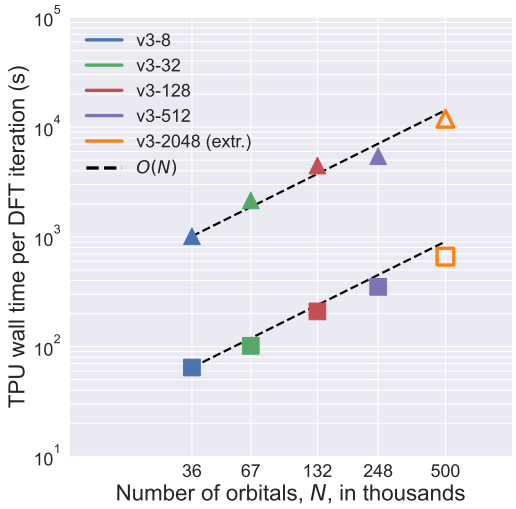

Figure 6.1: TPU v3 wall time for $O(N^3)$ density matrix purification, Eqs. (6.5)-(6.7), as a function of the number N of orbitals, for clusters of water molecules, both in single (squares) and double (triangles) precision. A full TPU v3 pod with 2048 cores and 32 TB of memory is expected to handle $N \sim 500\,000$ orbitals in our current implementation (extrapolated results necessitated by temporary resource unavailability).

for other computational tasks [14, 324, 229, 191, 192, 190, 134, 211, 120, 181, 109, 75]. In this work, we demonstrate the use of TPUs as quantum chemistry supercomputers by accelerating the $O(N^3)$ computational bottleneck of DFT approaches which use an auxiliary single-particle kinetic energy approximation, such as Kohn-Sham (KS) [126, 157] and generalized KS (gKS) [274] DFT, where gKS admits hybrid DFT functionals. This enables the systematic study of quantum chemistry problems at unprecedented scales. As a concrete demonstration, we performed end-to-end $O(N^3)$ DFT calculations on large clusters of water molecules, reaching a total of $N = 247\,848$ DFT orbitals, corresponding to 10 327 water molecules with 103 270 electrons, see Fig. 6.1 and Table 6.1. To our knowledge, this is the largest $O(N^3)$ DFT calculation to date, with the previously largest computation consisting of a single $O(N^3)$ DFT iteration with $N \approx 230\,000$ orbitals on Fujitsu's K computer [119].

Some variants of DFT, most notably linear-scaling DFT [254, 20, 281, 310], avoid the $O(N^3)$ bottleneck altogether and can thus reach an even larger number of orbitals. However, these variants rely on additional approximations and conditions, such as truncating density matrix

Number of orbitals	Number of atoms	Number of electrons	TPU configuration	TPU wall time (s)		Relative energy per molecule (mHa)
				FP32	FP64	
35 544	4 443	14 810	v3-8	65	1 012	0.934
65 668	8 211	27 370	v3-32	102	2 150	0.531
131 544	16 443	54 810	v3-128	209	4 465	0.291
247 848	30 981	103 270	v3-512	350	5 434	0

Table 6.1: Tabulated results in Fig. 6.1, including also number of atoms and electrons. Wall times for the matrix purification step are shown both for single (FP32) and double (FP64) precision. Energies are relative to the largest calculation, $E[(H_2O)_{N_{\rm mol}}]/N_{\rm mol} - E[(H_2O)_{10327}]/10327$, where $N_{\rm mol}$ is the total number of water molecules. In this sequence, we used a number of TPU cores that grows roughly as N^2. As a result, walltimes are seen to roughly scale linearly in N, instead of the expected $O(N^3)$ scaling.

elements [44], or on special properties of only a subset of density functionals (such as semilocal density functional approximations). In turn, this results in restricted applicability, with e.g. linear-scaling DFT being suitable for insulating systems but not for metals or systems with a small energy gap [254]. In practice, conventional $O(N^3)$ DFT is a more preferable choice since it alleviates technical complexity and problem space restrictions associated with current lower-scaling methods, greatly extending the domain of problems to which DFT can be applied reliably and with relative ease.

There are many aspects that go into an $O(N^3)$ DFT calculation. Throughout we focus on atom-centered basis sets with all electrons treated explicitly, that is we do not consider e.g. plane waves or pseudopotentials. At a high level, one can identify two main computational steps: (a) building the DFT Hamiltonian matrix (with cost $O(N)$ to $O(N^2)$) and (b) computing the ground state density matrix ($O(N^3)$), see Fig. 6.2.

(a) *DFT Hamiltonian build*: Given a choice of N atom-centered basis functions $\chi_i(\mathbf{r})$, one needs to compute the DFT Hamiltonian matrix H and the overlap matrix S, with coefficients given by

$$H_{ij} = \langle \chi_i | \, \mathcal{H} \, | \chi_j \rangle \,, \qquad S_{ij} = \langle \chi_i | \chi_j \rangle \,, \tag{6.1}$$

where $\mathcal{H}$ represents the DFT Hamiltonian in the continuum and each matrix coefficient requires computing one or several integrals. Over the past few decades much effort has been devoted to optimizing the build of the $N \times N$ matrix H. Naively, the computational time here scales as $O(N^4)$, however, in many implementations the scaling is effectively reduced to $O(N^2)$ due to two-electron integral screening methods. The scaling can be further reduced to almost $O(N)$ if other strategies, such as fast multipole methods [327] or fast fourier transform based methods [310], such as the Ewald method for periodic systems, are employed. In this work we do *not* attempt to accelerate the Hamiltonian or overlap matrix build times with TPUs. Instead, we simply use a well-established all-electron DFT package, the *Fritz Haber Institute ab initio molecular simulations package* (FHI-aims) [132, 17, 18], which we run using CPUs.

(b) *Density matrix purification:* The pair of matrices H and S define a generalized eigenvalue problem, the so-called KS equations,

$$H \left|\phi_\alpha\right\rangle = e_\alpha S \left|\phi_\alpha\right\rangle, \tag{6.2}$$

with $|\phi_\alpha\rangle$ and e_α the KS orbitals and energies. Our goal is to compute the ground state density matrix

$$D \equiv \sum_{\alpha=1}^{N} \theta(\mu - e_\alpha) \left|\phi_\alpha\right\rangle \left\langle\phi_\alpha\right|, \tag{6.3}$$

where $\theta(x)$ is the step function and μ is the chemical potential, chosen such that $\sum_{\alpha=1}^{N} \theta(\mu - e_\alpha) = N_e$, for N_e the number of electrons in the system. The density matrix D can be obtained by solving the KS equations (6.2) using standard linear algebra libraries, such as LAPACK [4] or Intel MKL [1]. An alternative route, which we follow in this work, is to use a density matrix purification scheme [301, 148]. First, by computing the inverse square root

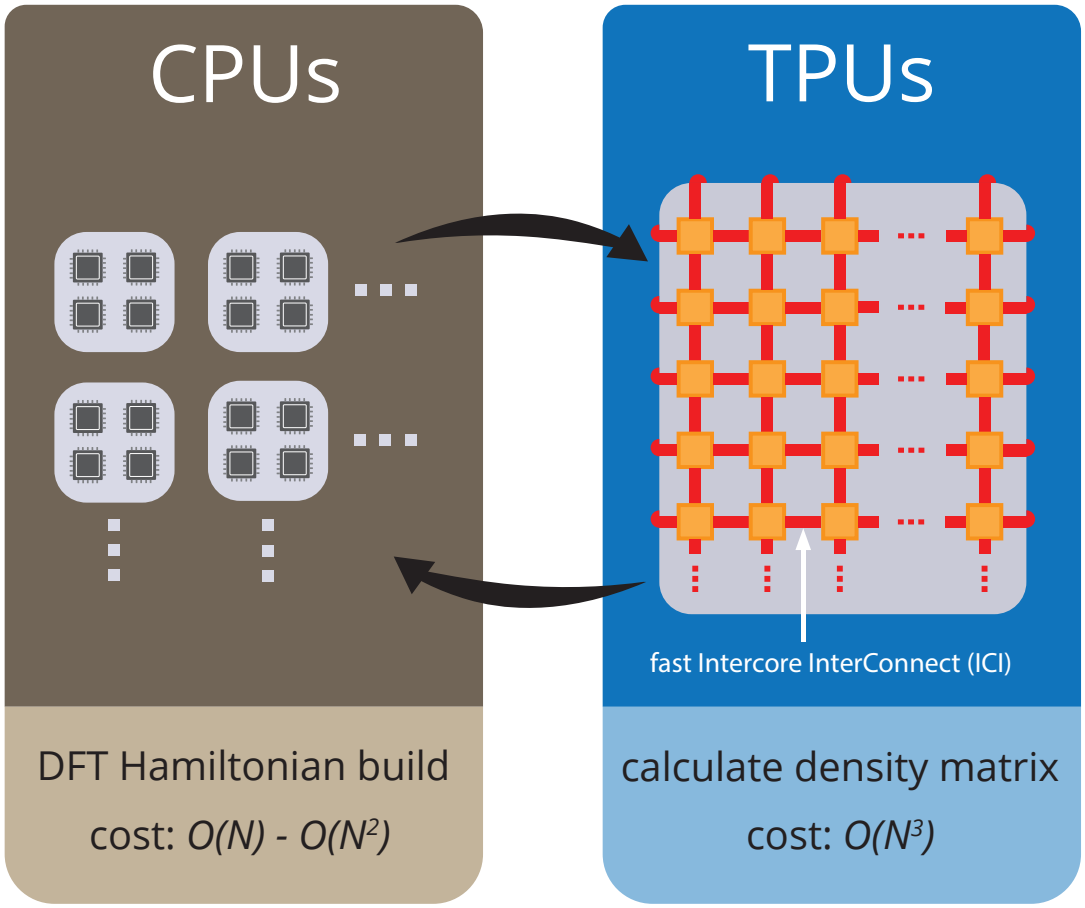

Figure 6.2: The two main steps of our implementation of an $O(N^3)$ DFT computation, the Hamiltonian build and computing the ground state density matrix, which we run on CPUs and TPUs, respectively. The DFT code FHI-aims [17, 18] is used to set up the Hamiltonian and the ELSI library [339, 344] is used to facilitate the integration of the TPU-based solver to FHI-aims.

of S,

$$S \mapsto S^{-\frac{1}{2}} \tag{6.4}$$

we can write the Hamiltonian in an orthonormal basis,

$$H \mapsto \tilde{H} \equiv S^{-\frac{1}{2}} H S^{-\frac{1}{2}}. \tag{6.5}$$

and re-express (6.2) as a standard eigenvalue problem $\tilde{H}\,|\tilde{\phi}_\alpha\rangle = e_\alpha\,|\tilde{\phi}_\alpha\rangle$, where $|\tilde{\phi}_\alpha\rangle \equiv S^{1/2}\,|\phi_\alpha\rangle$. Next we compute the density matrix $\tilde{D}$,

$$\tilde{H} \to \tilde{D} \equiv \theta(\mu I - \tilde{H}) = \sum_{\alpha=1}^{N_e} \theta(\mu - e_\alpha)\,|\tilde{\phi}_\alpha\rangle\,\langle\tilde{\phi}_\alpha|\,, \tag{6.6}$$

and finally re-express it in the original basis,

$$\tilde{D} \mapsto D \equiv S^{-\frac{1}{2}} \tilde{D} S^{-\frac{1}{2}}. \tag{6.7}$$

The transformation in Eq. (6.6) is obtained using a standard density matrix purification scheme that is suitable for TPUs, namely the hole-particle canonical purification scheme [301], which we elaborate on later in the paper.

If no further modifications are made (e.g., density matrix truncations in linear-scaling DFT), then the cost of computing D, whether by solving Eq. (6.2) or performing the four matrix transformations in Eqs. (6.4)-(6.7), scales as $O(N^3)$. This constitutes what is known as the *cubic wall* of DFT.

The density matrix D is used to derive several important quantities. The real-space electron

density $n(\mathbf{r})$ is given by

$$n(\mathbf{r}) = \sum_{i,j}^{N} \chi_i(\mathbf{r}) D_{ij} \chi_j(\mathbf{r}) \,, \tag{6.8}$$

which can be computed on a real-space grid [17]. The sum of occupied KS eigenvalues, given by $\mathrm{Tr}(H\,D) = \mathrm{Tr}(\tilde{H}\,\tilde{D})$, is also used to compute the total ground-state energy. Additionally, the *energy weighted density matrix* Q,

$$Q = D\,H\,D \,, \tag{6.9}$$

is also useful to compute atomic forces analytically [17].

6.3 Results

6.3.1 DFT with TPUs

The main result of our work is the successful use of TPUs to perform the four matrix transformations (6.4)-(6.7), thereby tackling the $O(N^3)$ computational bottleneck of DFT. We employed TPUs of the third generation, denoted v3. A single TPU v3 core contains two matrix multiply units (MXUs) to formidably accelerate matrix-matrix multiplication (matmul), resulting in about 10 teraFLOPS (floating point operations per second) of measured single-core matmul performance in single precision. Importantly for our purposes, matmuls are also available in double precision using a software-emulated 57-bit floating point format. In this approach, utilized algorithms require many more single precision floating point operations when operating in our emulated double precision than in single precision, and as a result matmuls in double precision take $\sim 11\times$ longer than in single precision.

The smallest available TPU configuration consists of 8 TPU v3 cores with a total of 128 GB of dedicated high bandwidth memory (HBM), controlled by a single host with 48 CPU cores. The largest configuration is a pod with 2048 TPU v3 cores and 32 TB of HBM, controlled by 256 hosts. Given a choice of configuration, the available TPU cores are directly connected to nearest neighbors in a 2D torus network through fast inter-core interconnects (ICIs), see Fig. 6.2. The ICIs are critical to maintaining high performance when distributing matmuls and other dense linear algebra operations over all available TPU cores. In this work we used the JAX library [21, 73, 2] to write *single program multiple data* (SPMD) code and executed it on configurations made of p TPU cores, denoted v3-p, for $p = 8, 32, 128$ and 512.

The TPU hardware architecture is especially suited for dense large-scale matmuls, which we perform in distributed form using the SUMMA algorithm [309], as recently demonstrated in Ref. [181]. Here it was shown that for sufficiently large matrices a v3-512 TPU can perform dense matmuls at near-optimal efficiency: the performance per TPU core (measured in single-precision FLOPS) is maintained at roughly 93% of the single TPU core maximum performance [181]. It is important to emphasize that TPUs are often ill-suited for other tasks, and hence the algorithms utilized in this work and those in Ref. [181] had to be picked carefully and may differ from more conventional choices used in CPUs or GPUs. The use of DM purification algorithms, rather than direct diagonalization, is especially attractive for TPUs since all steps can be evaluated from a series of matmuls. Clearly, transformations (6.5) and (6.7) require large-scale matmuls. Transformations (6.4) and (6.6) are implemented by an iteration involving matrix polynomials of small degree, where each polynomial requires a short sequence of matrix additions and multiplications. Specifically, the matrix inverse square root in (6.4) is implemented using a standard Newton-Schulz iteration [124], whereas for the density matrix purification in (6.6) we implemented the hole-particle canonical purification scheme [301]. Further algorithm details can be found in the Supporting Information.

For benchmarking purposes, we have performed end-to-end DFT computations on a sequence

of increasingly large water clusters with geometries obtained from standard molecular dynamics simulations (see Supporting Information). We leverage the DFT code FHI-aims [17, 18] to set up and drive calculations using CPUs, then use the TPUs to tackle the $O(N^3)$ dense linear algebra bottlenecks (6.4)-(6.7). We also utilize the *ELectronic Structure Infrastructure* (ELSI) library [339, 344] to facilitate the integration of FHI-aims and the TPU solver. In particular, the DFT Hamiltonian build time and associated computational scaling and parallelization are dictated exclusively by the FHI-aims code, which uses numeric atom-centered orbitals (NAOs) with an explicit finite spatial extent, and a truncated multipole expansion to accomplish low prefactor and efficient scaling of the Hamiltonian matrix build. While the computational time required to build the DFT Hamiltonian may vary greatly between different systems with the same total number of orbitals N (due to possible differences in the resulting sparsity in the systems), the computational time required for the $O(N^3)$ DM purification step performed on the TPU has much less variability since dense matrix operations are assumed, which do not utilize any sparsity present (see Supporting Information for more discussion). Thus, we emphasize that the TPU wall times, which are reported throughout only for water clusters, are fairly robust with respect to different systems with the same total number of orbitals.

Throughout this work we perform all-electron calculations using the PBE exchange-correlation functional and utilize an NAO basis set such that each H_2O molecule contributes 10 electrons, represented by 24 orbitals (5 for each hydrogen atom and 14 for the oxygen atom). First we consider a single TPU board with 8 TPU v3 cores controlled by a host with 48 CPU cores, and we run FHI-aims on the host. Fig. 6.3 shows the wall time for a single DFT iteration (including both Hamiltonian build on CPUs and density matrix purification on TPUs) as a function of the size of the water cluster, which ranges from a few thousand to $N \approx 50\,000$ orbitals. When using the TPU solver in single precision (green curve) we see that the $O(N)$ Hamiltonian build on 48 CPU cores takes longer than the $O(N^3)$ density matrix purification run on 8 TPU cores, thus shifting the bottleneck. Using the TPU solver in double precision

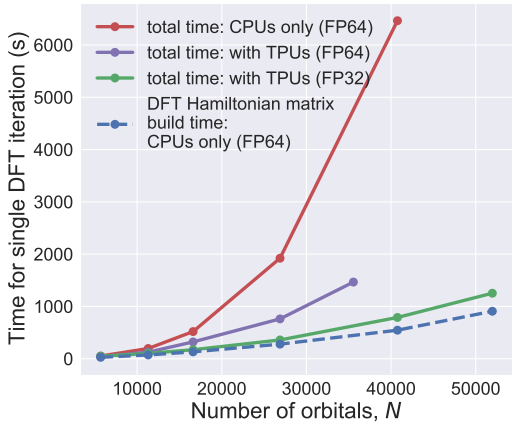

Figure 6.3: Wall times for a single DFT iteration on water clusters, using a TPU board composed of a CPU host (48 CPU cores) and 8 TPU cores with 128 GB of HBM. Green and purple curves correspond to using single and double precision in the TPU solver, respectively. The dashed blue curve corresponds to the CPU time spent on FHI-aims (always in double precision), and should be subtracted from the other curves in order to obtain the time spent on the TPU solvers. For reference, in red we also plot the time required for a CPU-only computation using the *Eigenvalue soLvers for Petaflop Applications* (ELPA), a highly parallelized eigensolver library [163, 198], run on 48 CPU cores.

is an order of magnitude slower and saturates the TPU's HBM for $N \approx 36\,000$ orbitals.

Then we consider larger TPU configurations, of up to 512 TPU v3 cores, to perform end-to-end DFT computations on larger clusters, of up to 10 327 water molecules (or $N = 247\,848$ orbitals). Fig. 6.1 shows the TPU wall time for the $O(N^3)$ density matrix purification for one DFT iteration. These include 350 (5 434) seconds for a density matrix purification in single (double) precision on the largest cluster, demonstrating feasibility of DFT computations at that scale of a quarter of a million orbitals.

6.3.2 Dynamic precision on TPUs

In our implementation, early DFT iterations are treated with single precision and later ones in double precision. This *dynamic precision* approach allows us to cut down on the use of double precision matmuls on TPUs (which are significantly slower than single precision

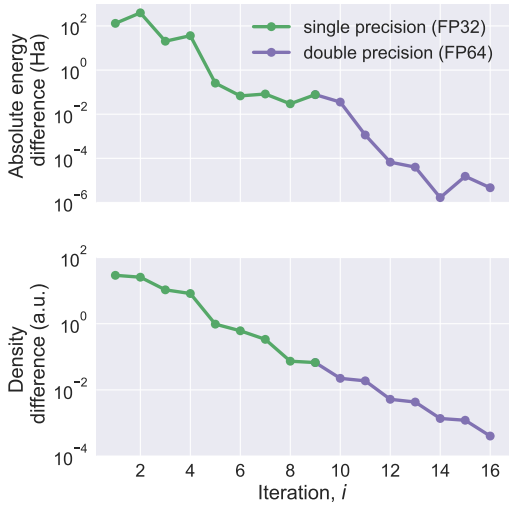

Figure 6.4: Convergence trajectory of an end-to-end dynamic precision DFT calculation on a $(H_2O)_{10327}$ cluster. The absolute total energy differences between subsequent DFT iterations, i and $i-1$, are plotted (top). The corresponding difference in real-space densities within the L^1 norm is plotted (bottom).

ones) without sacrificing accuracy of the final converged DFT result. Our criteria to switch precision is based on relative density changes, using the L^1 norm, defined as $L^1[f(\mathbf{r})] \equiv \int d^3r\, |f(\mathbf{r})|$, and relative energy changes:

$$\frac{1}{N_e} L^1[n^{[i]}(\mathbf{r}) - n^{[i-1]}(\mathbf{r})] < \epsilon \qquad \text{and} \tag{6.10}$$

$$|E^{[i]} - E^{[i-1]}|/|E^{[i]}| < \epsilon\,, \tag{6.11}$$

where $n^{[i]}(\mathbf{r})$ is the real-space density at DFT iteration i, $E^{[i]}$ is the corresponding total ground-state energy, and we use $\epsilon = 5 \times 10^{-7}$ for single precision.

Fig. 6.4 shows the convergence trajectory of a dynamic precision DFT calculation for the largest cluster we have considered. We are able to converge such a DFT calculation to a fairly tight convergence threshold using first 9 single precision DFT iterations, followed by 7 double precision DFT iterations. In a smaller cluster, $(H_2O)_{1481}$ with $N = 35\,544$ orbitals, a smaller number of double precision iterations are required for convergence, resulting in

an overall DFT calculation time that is under 5 hours on a single TPU board (v3-8), see Supporting Information.

6.4 Conclusion

This work has successfully demonstrated that TPUs can both *accelerate* and *scale up* DFT computations. Significant *acceleration* is already achieved using only a single TPU board with 8 TPU v3 cores, see Fig. 3. For instance, an end-to-end dynamic precision DFT computation with N = 35 544 orbitals consisting of 12 iterations in single precision and 4 iterations in double precision yields converged results in under 5 hours. For context, using double precision only and the highly optimized ELPA $O(N^3)$ solver with 48 CPUs, the same water cluster calculation required 20 hours to achieve 16 DFT iterations.

In order to *scale up* the size of DFT computations while retaining high performance two main ingredients are involved: (i) a larger amount of high bandwidth memory, scaling as $O(N^2)$, to be able to store dense $N \times N$ matrices; (ii) a larger number of cores, with state-of-the-art inter-core connectivity, to more effectively execute the $O(N^3)$ floating point operations involved in the required distributed matrix transformations. As shown in Fig. 1, by using a number of cores that scales as $O(N^2)$ and commensurate amounts of HBM, we can scale up to $N = 500\,000$ orbitals with wall times that only grow proportional to N.

Once the main ingredients (i) and (ii) are satisfied, the Input/Output (IO) time, i.e. the end-to-end communication time between the TPU and CPU, can become an important (and possibly limiting) factor. This IO step is not exclusive to TPUs, for instance, it is also relevant and analogous in GPU setups that require communication with CPUs. This step is parallelizable, such that by using a number of TPU and CPU cores that scales as $O(N^2)$, the total IO time remains constant $O(1)$. However, in this context, obtaining

such optimality will depend on specific implementation details of the DFT code utilized, such as the matrix distribution pattern on the CPU processor grid. Our current prototype implementation is modular and generalizeable and IO times scale unfavorably as $O(N^2)$, even becoming the rate-limiting step in some cases (see Supporting Information). Less-general (but straightforward) engineering approaches are expected to be much more optimal, but are beyond the scope of this work which aims to demonstrate the use of TPUs in a more general context.

We emphasize that other hardware accelerators, most notably GPUs, can also accelerate and scale up DFT computations in a similar manner as discussed above, and Ref. [49] recently presented a useful and positive development in this direction. Modern distributed GPU configurations are expected to achieve similar performance to TPUs in this regard, however, a direct comparison is complicated by the highly diverse nature of distributed GPU configurations found in practice. On the Summit supercomputer configuration, it has been reported that 432 distributed Nvidia V100 GPUs can perform matmuls for dense $N > 500\,000$ matrices with a performance per GPU (measured in FLOPS) that is roughly 85% of the single V100 GPU maximum performance [122].

Here we have focused, for simplicity, on applying DFT to clusters of water molecules. More complicated systems may present additional difficulties. For instance, protein-ligand complexes often require more elaborate schemes, such as including solvation to facilitate the convergence of the DFT iteration [172, 261]. Work in progress shows that our TPU-based large-scale DFT computations can also successfully address protein-ligand complexes with explicit solvents, as well as in a variety of other large systems, including DNA segments, carbon nanotubes, and graphene surfaces. In addition to single-point DFT energy calculations, analytical forces can also be extracted from the TPU-calculated *energy-weighted density matrix*, enabling large-scale geometry optimization or Ab initio molecular dynamics calculations.

DFT is a highly successful quantum-based method, but it is ultimately a consistent-field approximation, which may not be accurate enough for certain applications. Fortunately, TPUs can also accelerate and scale up other, more accurate quantum chemistry approaches where the computational bottleneck is again given by dense linear algebra operations. For example, in density matrix renormalization group (DMRG) [328] calculations, TPUs can be used to reach an unprecedentedly large bond dimension $D = 65\,536$ [75]. Similarly, we anticipate that TPUs will thrive in other methods such as coupled cluster [64] and Møller–Plesset perturbation theory [312]. Even when applying such higher-level methods, large-scale DFT may still be a crucial piece in simulations that require a quantum-mechanically treated region that embeds a subsystem treated with higher-level correlated methods [289].

To conclude, in this work we have successfully repurposed TPUs as quantum chemistry supercomputers by tackling the $O(N^3)$ computational bottleneck of density functional theory. We demonstrated performance and scalability with a water cluster with $N = 247\,848$ orbitals, which to our knowledge is the largest $O(N^3)$ DFT computation to date. We remark that cloud-based TPUs, and other hardware accelerators such as GPUs, are more accessible and affordable than traditional supercomputer resources. Our work thus paves the way towards accessible and straightforward use of quantum chemistry computational methods for much larger systems than were previously possible.

Competing Interests: V.B. received compensation as an advisor from Google during part of this work. V.B. is also a board member of MS1P e.V., the non-profit organization that licenses the FHI-aims electronic structure code used in this work. He does not receive any financial gains from this position.

6.5 Acknowledgments

The authors would like to thank Toru Shiozaki and Garnet Kin-Lic Chan for suggesting to investigate the use of TPUs to accelerate mean-field quantum chemistry methods (by accelerating matrix multiplications for a diagonalization-free construction of the density matrix, as in the density matrix purification used in this paper), and Toru Shiozaki, Garnet Kin-Lic Chan, Chase Roberts and Stefan Leichenauer for previous exploratory work in this direction. Also, special thanks to Xing Zhang and Garnet Kin-Lic Chan for work adjusting PySCF, as part of an on-going integration of our TPU solver, to be described elsewhere, and to David Bowler, Tsuyoshi Miyazaki, and Jun-ichi Iwata for help documenting the largest DFT computation run on the K computer. Fianlly, the authors would also like to thank Giuseppe M. J. Barca, Anudhyan Boral, Michael Brenner, Kieron Burke, Rafael Gomez-Bombarelli, JW Feng, Filipp Furche, Andreas Goeller, Stephan Hoyer, Olivier Lacombe, Stefan Leichenauer, Lin Lin, Ruben Martin Romo, Todd Martinez, Anders M. N. Niklasson, Nicholas Rubin, Zak Stone, Matthias Tan, Keiran Thompson, Edward Valeev, and Jae Yoo for useful discussions and comments. Research supported with Cloud TPUs from Google's TPU Research Cloud (TRC). Sandbox is a team within the Alphabet family of companies, which includes Google, Verily, Waymo, X, and others. G.V. is a CIFAR fellow in the Quantum Information Science Program, a Distinguished Invited Professor at the Institute of Photonic Sciences (ICFO), and a Distinguished Visiting Research Chair at Perimeter Institute. Research at Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research, Innovation, and Science. R.S. and Y.Y. were partially supported by the National Science Foundation (NSF), USA under Award No. 1450280.

Bibliography

[1] *Intel Math Kernel Library. Reference Manual*. Intel Corporation, Santa Clara, USA, 2009. ISBN 630813-054US.

[2] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng. Tensorflow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, OSDI'16, page 265–283, USA, 2016. USENIX Association.

[3] C. Adamo and D. Jacquemin. The calculations of excited-state properties with time-dependent density functional theory. *Chem. Soc. Rev.*, 42:845–856, 2013.

[4] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, Philadelphia, PA, third edition, 1999.

[5] V. I. Anisimov, A. I. Poteryaev, M. A. Korotin, A. O. Anokhin, and G. Kotliar. First-principles calculations of the electronic structure and spectra of strongly correlated systems: dynamical mean-field theory. *Journal of Physics: Condensed Matter*, 9(35):7359, 1997.

[6] E. Aprà, E. J. Bylaska, W. A. de Jong, N. Govind, K. Kowalski, T. P. Straatsma, M. Valiev, H. J. J. van Dam, Y. Alexeev, J. Anchell, V. Anisimov, F. W. Aquino, R. Atta-Fynn, J. Autschbach, N. P. Bauman, J. C. Becca, D. E. Bernholdt, K. Bhaskaran-Nair, S. Bogatko, P. Borowski, J. Boschen, J. Brabec, A. Bruner, E. Cauët, Y. Chen, G. N. Chuev, C. J. Cramer, J. Daily, M. J. O. Deegan, T. H. Dunning, M. Dupuis, K. G. Dyall, G. I. Fann, S. A. Fischer, A. Fonari, H. Früchtl, L. Gagliardi, J. Garza, N. Gawande, S. Ghosh, K. Glaesemann, A. W. Götz, J. Hammond, V. Helms, E. D. Hermes, K. Hirao, S. Hirata, M. Jacquelin, L. Jensen, B. G. Johnson, H. Jónsson, R. A. Kendall, M. Klemm, R. Kobayashi, V. Konkov, S. Krishnamoorthy, M. Krishnan, Z. Lin, R. D. Lins, R. J. Littlefield, A. J. Logsdail, K. Lopata, W. Ma, A. V. Marenich, J. Martin del Campo, D. Mejia-Rodriguez, J. E. Moore, J. M. Mullin, T. Nakajima, D. R. Nascimento, J. A. Nichols, P. J. Nichols, J. Nieplocha, A. Otero-de-la Roza, B. Palmer, A. Panyala, T. Pirojsirikul, B. Peng, R. Peverati,

J. Pittner, L. Pollack, R. M. Richard, P. Sadayappan, G. C. Schatz, W. A. Shelton, D. W. Silverstein, D. M. A. Smith, T. A. Soares, D. Song, M. Swart, H. L. Taylor, G. S. Thomas, V. Tipparaju, D. G. Truhlar, K. Tsemekhman, T. Van Voorhis, A. Vázquez-Mayagoitia, P. Verma, O. Villa, A. Vishnu, K. D. Vogiatzis, D. Wang, J. H. Weare, M. J. Williamson, T. L. Windus, K. Woliński, A. T. Wong, Q. Wu, C. Yang, Q. Yu, M. Zacharias, Z. Zhang, Y. Zhao, and R. J. Harrison. Nwchem: Past, present, and future. *The Journal of Chemical Physics*, 152(18):184102, 2020.

[7] B. Austin. Nersc-10 workload analysis (data from 2018). `https://portal.nersc.gov/project/m888/nersc10/workload/N10_Workload_Analysis.latest.pdf`. Accessed: 11-20-2022.

[8] T. E. Baker, E. M. Stoudenmire, L. O. Wagner, K. Burke, and S. R. White. One-dimensional mimicking of electronic structure: The case for exponentials. *Phys. Rev. B*, 91:235141, Jun 2015.

[9] R. J. Bartlett and M. Musial. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.*, 79:291–352, Feb 2007.

[10] P. D. Beale, R.K. Pathria. *Statistical Mechanics*. Academic Press, third edition, 2011.

[11] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098, 1988.

[12] A. D. Becke. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.*, 140:18A301, 2014.

[13] A. D. Becke. Perspective: Fifty years of density-functional theory in chemical physics. *J. Chem. Phys.*, 140(18):18A301, 2014.

[14] F. Belletti, D. King, K. Yang, R. Nelet, Y. Shafi, Y.-F. Shen, and J. Anderson. Tensor processing units for financial monte carlo. In *Proceedings of the 2020 SIAM Conference on Parallel Processing for Scientific Computing*, pages 12–23, 2020.

[15] T. Berkelbach, G. Chan, S. Sharma, A. Sokolov, Q. Sun, et al. PySCF. `https://pyscf.org/`. Accessed: 01-17-2022.

[16] P. Blöchl. Theory and practice of density-functional theory. In Pavarini et al. [232].

[17] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler. Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications*, 180(11):2175–2196, 2009.

[18] V. Blum, S. Kokott, M. Rossi, and M. Scheffler. FHI-AIMS. https://fhi-aims.org/. Accessed: 11-20-2022.

[19] G. Boschetto, T. Xu, M. Yehya, J. Thireau, A. Lacampagne, B. Charlot, T. Gil, and A. Todri-Sanial. Exploring 1d and 2d nanomaterials for health monitoring wearable devices. In *2021 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS)*, pages 1–4. IEEE, 2021.

[20] D. R. Bowler and T. Miyazaki. Calculations for millions of atoms with density functional theory: linear scaling shows its potential. *Journal of Physics: Condensed Matter*, 22(7):074207, 2010.

[21] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, and S. Wanderman-Milne. JAX: composable transformations of Python+NumPy programs. `http://github.com/google/jax`, 2018.

[22] K. Burke. *The ABC of DFT*. 2007.

[23] K. Burke. Perspective on density functional theory. *J. Chem. Phys.*, 136:150901, 2012.

[24] K. Burke. Perspective on density functional theory. *J. Chem. Phys.*, 136(15):150901, 2012.

[25] K. Burke. Perspective on density functional theory. *J. Chem. Phys.*, 136(15):150901, 2012.

[26] K. Burke, A. Cancio, T. Gould, and S. Pittalis. Locality of correlation in density functional theory. *The Journal of chemical physics*, 145(5):054112, 2016.

[27] K. Burke and J. Kozlowski. *Lies My Teacher Told Me About Density Functional Theory: Seeing Through Them with the Hubbard Dimer*, pages 65–96. Forschungszentrum Jülich GmbH Institute for Advanced Simulation, 2021.

[28] K. Burke, J. P. Perdew, and Y. Wang. *Derivation of a Generalized Gradient Approximation: The PW91 Density Functional*, pages 81–111. Springer US, Boston, MA, 1998.

[29] K. Burke and L. O. Wagner. Dft in a nutshell. *Int. J. Quant. Chem.*, 113:96–101, 2013.

[30] K. Burke, J. Werschnik, and E. K. U. Gross. Time-dependent density functional theory: Past, present, and future. *The Journal of Chemical Physics*, 123(6):062206, 2005.

[31] A. Cancio, G. P. Chen, B. T. Krull, and K. Burke. Fitting a round peg into a round hole: asympotically correcting the generalized gradient approximation for correlation. *The Journal of Chemical Physics*, 149:084116, 8 2018.

[32] A. Cancio, G. P. Chen, B. T. Krull, and K. Burke. Fitting a round peg into a round hole: Asymptotically correcting the generalized gradient approximation for correlation. *The Journal of chemical physics*, 149(8):084116, 2018.

[33] D. J. Carrascal, J. Ferrer, N. Maitra, and K. Burke. Linear response time-dependent density functional theory of the hubbard dimer. *Accepted arXiv:1802.09988*, 2018.

[34] D. J. Carrascal, J. Ferrer, N. Maitra, and K. Burke. Linear response time-dependent density functional theory of the hubbard dimer. *The European Physical Journal B*, 91(7):142, 07 2018.

[35] D. J. Carrascal, J. Ferrer, J. C. Smith, and K. Burke. The hubbard dimer: a density functional case study of a many-body problem. *Journal of Physics: Condensed Matter*, 27(39):393001, 2015.

[36] D. J. Carrascal, J. Ferrer, J. C. Smith, and K. Burke. The hubbard dimer: a density functional case study of a many-body problem. *Journal of Physics: Condensed Matter*, 27(39):393001, 2015.

[37] D. J. Carrascal, J. Ferrer, J. C. Smith, and K. Burke. The hubbard dimer: a density functional case study of a many-body problem. *Journal of Physics: Condensed Matter*, 27(39):393001, sep 2015.

[38] M. Casida. *Time-dependent density functional response theory of molecular systems: Theory, computational methods, and functionals*, volume 4. Elsevier, 1996.

[39] C. N. Cavasotto, N. S. Adler, and M. G. Aucar. Quantum chemical approaches in structure-based virtual screening and lead optimization. *Frontiers in chemistry*, 6:188, 2018.

[40] F. Cernatic, B. Senjean, V. Robert, and E. Fromager. Ensemble density functional theory of neutral and charged excitations. *Top. Curr. Chem.*, 380(1):1–80, 2022.

[41] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Insights into current limitations of density functional theory. *Science*, 321(5890):792–794, 2008.

[42] A. J. Cohen, P. Mori-Sanchez, and W. Yang. Insights into current limitations of density functional theory. *Science*, 321:792, 2008.

[43] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Challenges for density functional theory. *Chem. Rev.*, 112:289, 2012.

[44] D. Cole, C.-K. Skylaris, E. Rajendra, A. Venkitaraman, and M. Payne. Protein-protein interactions from linear-scaling first-principles quantum-mechanical calculations. *Europhysics Letters*, 91(3):37004, 2010.

[45] L. A. Constantin, E. Fabiano, and F. Della Sala. Semilocal pauli–gaussian kinetic functionals for orbital-free density functional theory calculations of solids. *The journal of physical chemistry letters*, 9(15):4385–4390, 2018.

[46] L. A. Constantin, E. Fabiano, and F. Della Sala. Performance of semilocal kinetic energy functionals for orbital-free density functional theory. *Journal of chemical theory and computation*, 15(5):3044–3055, 2019.

[47] S. Crisostomo, M. Levy, and K. Burke. Can the hartree–fock kinetic energy exceed the exact kinetic energy? *J. Chem. Phys.*, 157(15):154106, 2022.

[48] S. Crisostomo, R. Pederson, J. Kozlowski, B. Kalita, A. C. Cancio, K. Datchev, A. Wasserman, S. Song, and K. Burke. Seven useful questions in density functional theory. *Lett. Math Phys.*, 113(2):42, 2023.

[49] S. Das, P. Motamarri, V. Subramanian, D. M. Rogers, and V. Gavini. Dft-fe 1.0: A massively parallel hybrid cpu-gpu density functional theory code using finite-element discretization. *arXiv preprint arXiv:2203.07820*, 2022.

[50] C. E. de Calcul Atomique et Moléculaire. Teaching the theory in density functional theory.

[51] K. Deur and E. Fromager. Ground and excited energy levels can be extracted exactly from a single ensemble density-functional theory calculation. *J. Chem. Phys.*, 150(9):094106, 2019.

[52] K. Deur, L. Mazouin, and E. Fromager. Exact ensemble density functional theory for excited states in a model system: Investigating the weight dependence of the correlation energy. *Phys. Rev. B*, 95(3):035120, 2017.

[53] K. Deur, L. Mazouin, B. Senjean, and E. Fromager. Exploring weight-dependent density-functional approximations for ensembles in the hubbard dimer. *The European Physical Journal B*, 91(7):1–18, 2018.

[54] P. A. M. Dirac. Note on exchange phenomena in the Thomas atom. *Mathematical Proceedings of the Cambridge Philosophical Society*, 26(03):376–385, 1930.

[55] T. Dornheim, Z. A. Moldabekov, K. Ramakrishna, P. Tolias, A. D. Baczewski, D. Kraus, T. R. Preston, D. A. Chapman, M. P. Böhme, T. Döppner, F. Graziani, M. Bonitz, A. Cangi, and J. Vorberger. Electronic density response of warm dense matter. *Physics of Plasmas*, 30(3):032705, 03 2023.

[56] R. Dreizler and E. Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Springer Berlin Heidelberg, 1989.

[57] R. M. Dreizler and E. K. U. Gross. *Density Functional Theory: An Approach to the Quantum Many-Body Problem*. Springer–Verlag, Berlin, 1990.

[58] D. Duan, H. Yu, H. Xie, and T. Cui. Ab initio approach and its impact on superconductivity. *Journal of Superconductivity and Novel Magnetism*, 32(1):53–60, 2019.

[59] T. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90:1007, 1989.

[60] P. Elliott and K. Burke. Non-empirical derivation of the parameter in the b88 exchange functional. *Can. J. Chem. Ecol.*, 87(10):1485–1491, Oct 2009.

[61] M. Ernzerhof. Taylor-series expansion of density functionals. *Physical Review A*, 50(6):4593, 1994.

[62] M. Ernzerhof. Construction of the adiabatic connection. *Chem. Phys. Lett.*, 263(3-4):499–506, 1996.

[63] E. Fabiano, P. Gori-Giorgi, M. Seidl, and F. Della Sala. Interaction-strength interpolation method for main-group chemistry: Benchmarking, limitations, and perspectives. *J. Chem. Theory Comput*, 12(10):4885–4896, 2016.

[64] B. S. Fales, E. R. Curtis, K. G. Johnson, D. Lahana, S. Seritan, Y. Wang, H. Weir, T. J. Martínez, and E. G. Hohenstein. Performance of coupled-cluster singles and doubles on modern stream processing architectures. *Journal of Chemical Theory and Computation*, 16(7):4021–4028, 2020.

[65] D. Feller and K. A. Peterson. Probing the limits of accuracy in electronic structure calculations: Is theory capable of results uniformly better than "chemical accuracy"? *The Journal of Chemical Physics*, 126(11):–, 2007.

[66] E. Fermi. A statistical method for the determination of some priorieta dell'atome. *Rend. Accad. Nat. Lincei*, 6(602-607):32, 1927.

[67] E. Fermi. Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente (a statistical method for the determination of some atomic properties and the application of this method to the theory of the periodic system of elements). *Zeitschrift für Physik A Hadrons and Nuclei*, 48:73–79, 1928.

[68] K. Finzel. Local conditions for the pauli potential in order to yield self-consistent electron densities exhibiting proper atomic shell structure. *The Journal of chemical physics*, 144(3):034108, 2016.

[69] V. V. Franca, D. Vieira, and K. Capelle. Analytical parametrization for the ground-state energy of the one-dimensional hubbard model. *arXiv:1102.5018v1*, 2011.

[70] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 Revision A.1, 2016. Gaussian Inc. Wallingford CT 2009.

[71] E. Fromager. Individual correlations in ensemble density functional theory: State- and density-driven decompositions without additional kohn-sham systems. *Phys. Rev. Lett.*, 124(24):243001, 2020.

[72] J. M. Frost, K. T. Butler, F. Brivio, C. H. Hendon, M. Van Schilfgaarde, and A. Walsh. Atomistic origins of high-performance in hybrid halide perovskite solar cells. *Nano letters*, 14(5):2584–2590, 2014.

[73] R. Frostig, M. Johnson, and C. Leary. Compiling machine learning programs via high-level tracing. 2018.

[74] A. P. Gaiduk, D. S. Firaha, and V. N. Staroverov. Improved electronic excitation energies from shape-corrected semilocal kohn-sham potentials. *Phys. Rev. Lett.*, 108:253005, Jun 2012.

[75] M. Ganahl, J. Beall, M. Hauru, A. G. Lewis, J. H. Yoo, Y. Zou, and G. Vidal. Density matrix renormalization group with tensor processing units. *arXiv preprint arXiv:2204.05693*, 2022.

[76] A. García, N. Papior, A. Akhtar, E. Artacho, V. Blum, E. Bosoni, P. Brandimarte, M. Brandbyge, J. I. Cerdá, F. Corsetti, R. Cuadrado, V. Dikan, J. Ferrer, J. Gale, P. García-Fernández, V. M. García-Suárez, S. García, G. Huhs, S. Illera, R. Korytár, P. Koval, I. Lebedeva, L. Lin, P. López-Tarifa, S. G. Mayo, S. Mohr, P. Ordejón, A. Postnikov, Y. Pouillon, M. Pruneda, R. Robles, D. Sánchez-Portal, J. M. Soler, R. Ullah, V. W.-z. Yu, and J. Junquera. Siesta: Recent developments and applications. *The Journal of Chemical Physics*, 152(20):204108, 2020.

[77] L. Genovese, B. Videau, M. Ospici, T. Deutsch, S. Goedecker, and J.-F. Méhaut. Daubechies wavelets for high performance electronic structure calculations: The bigdft project. *Comptes Rendus Mécanique*, 339(2):149–164, 2011.

[78] S. Giarrusso and A. Pribram-Jones. Comparing correlation components and approximations in hartree–fock and kohn–sham theories via an analytical test case study. *J. Chem. Phys.*, 157(5):054102, 2022.

[79] P. M. Gill, B. G. Johnson, J. A. Pople, and M. J. Frisch. The performance of the becke—lee—yang—parr (b—lyp) density functional theory with various basis sets. *Chemical Physics Letters*, 197(4-5):499–505, 1992.

[80] S. Girvin and K. Yang. *Modern Condensed Matter Physics*. Cambridge University Press, 2019.

[81] X. Gonze, F. Jollet, F. Abreu Araujo, D. Adams, B. Amadon, T. Applencourt, C. Audouze, J.-M. Beuken, J. Bieder, A. Bokhanchuk, E. Bousquet, F. Bruneval, D. Caliste, M. Côté, F. Dahm, F. Da Pieve, M. Delaveau, M. Di Gennaro, B. Dorado, C. Espejo, G. Geneste, L. Genovese, A. Gerossier, M. Giantomassi, Y. Gillet, D. Hamann, L. He, G. Jomard, J. Laflamme Janssen, S. Le Roux, A. Levitt, A. Lherbier, F. Liu, I. Lukačević, A. Martin, C. Martins, M. Oliveira, S. Poncé, Y. Pouillon, T. Rangel, G.-M. Rignanese, A. Romero, B. Rousseau, O. Rubel, A. Shukri, M. Stankovski, M. Torrent, M. Van Setten, B. Van Troeye, M. Verstraete, D. Waroquiers, J. Wiktor, B. Xu, A. Zhou, and J. Zwanziger. Recent developments in the abinit software package. *Computer Physics Communications*, 205:106–131, 2016.

[82] R. G. Gordon and Y. S. Kim. Theory for the forces between closed-shell atoms and molecules. *The Journal of Chemical Physics*, 56(6):3122–3133, 1972.

[83] P. Gori-Giorgi and J. P. Perdew. Short-range correlation in the uniform electron gas: Extended overhauser model. *Phys. Rev. B*, 64:155102, Sep 2001.

[84] P. Gori-Giorgi and A. Savin. Simple model for the spherically and system-averaged pair density: Results for two-electron atoms. *Phys. Rev. A*, 71:032513, Mar 2005.

[85] P. Gori-Giorgi and A. Savin. Kohn-sham calculations combined with an average pair-density functional theory. *International Journal of Modern Physics B*, 21(13n14):2449–2459, 2007.

[86] A. Görling. Exchange-correlation potentials with proper discontinuities for physically meaningful kohn-sham eigenvalues and band structures. *Phys. Rev. B*, 91:245120, Jun 2015.

[87] T. Gould. What makes a density functional approximation good? insights from the left fukui function. *J. Chem. Theory Comput.*, 13(6):2373–2377, 2017.

[88] T. Gould. Approximately self-consistent ensemble density functional theory: toward inclusion of all correlations. *J. Phys. Chem. Lett.*, 11(22):9907–9912, 2020.

[89] T. Gould, Z. Hashimi, L. Kronik, and S. G. Dale. Single excitation energies obtained from the ensemble "homo–lumo gap": Exact results and approximations. *J. Phys. Chem. Lett.*, 13(10):2452–2458, 2022.

[90] T. Gould, D. P. Kooi, P. Gori-Giorgi, and S. Pittalis. Electronic excited states in extreme limits via ensemble density functionals. *Phys. Rev. Lett.*, 130(10):106401, 2023.

[91] T. Gould and L. Kronik. Ensemble generalized kohn–sham theory: The good, the bad, and the ugly. *J. Chem. Phys.*, 154(9):094125, 2021.

[92] T. Gould, L. Kronik, and S. Pittalis. Double excitations in molecules from ensemble density functionals: Theory and approximations. *Phys. Rev. A*, 104(2):022803, 2021.

[93] T. Gould and S. Pittalis. Hartree and exchange in ensemble density functional theory: Avoiding the nonuniqueness disaster. *Phys. Rev. Lett.*, 119(24):243001, 2017.

[94] T. Gould and S. Pittalis. Density-driven correlations in many-electron ensembles: Theory and application for excited states. *Phys. Rev. Lett.*, 123:016401, Jul 2019.

[95] T. Gould and S. Pittalis. Density-driven correlations in many-electron ensembles: Theory and application for excited states. *Phys. Rev. Lett.*, 123(1):016401, 2019.

[96] T. Gould and S. Pittalis. Density-driven correlations in ensemble density functional theory: Insights from simple excitations in atoms. *Aust. J. Chem.*, 73(8):714–723, 2020.

[97] T. Gould, G. Stefanucci, and S. Pittalis. Ensemble density functional theory: Insight from the fluctuation-dissipation theorem. *Phys. Rev. Lett.*, 125(23):233001, 2020.

[98] T. Gould and J. Toulouse. Kohn-sham potentials in exact density-functional theory at noninteger electron numbers. *Physical Review A*, 90(5):050502, 2014.

[99] F. Graziani, M. P. Desjarlais, R. Redmer, and S. B. Trickey, editors. *Frontiers and Challenges in Warm Dense Matter*, volume 96 of *Lecture Notes in Computational Science and Engineering*. Springer International Publishing, 2014.

[100] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of chemical physics*, 132(15):154104, 2010.

[101] O. Gritsenko, Ł. Mentel, and E. Baerends. On the errors of local density (lda) and generalized gradient (gga) approximations to the kohn-sham potential and orbital energies. *J. Chem. Phys.*, 144(20):204114, 2016.

[102] E. Gross and W. Kohn. Local density-functional theory of frequency-dependent linear response. *Phys. Rev. Lett.*, 55:2850, 1985.

[103] E. K. Gross, L. N. Oliveira, and W. Kohn. Density-functional theory for ensembles of fractionally occupied states. i. basic formalism. *Phys. Rev. A*, 37(8):2809, 1988.

[104] E. K. Gross, L. N. Oliveira, and W. Kohn. Rayleigh-ritz variational principle for ensembles of fractionally occupied states. *Phys. Rev. A*, 37(8):2805, 1988.

[105] E. K. U. Gross, L. N. Oliveira, and W. Kohn. Density-functional theory for ensembles of fractionally occupied states. i. basic formalism. *Phys. Rev. A*, 37:2809–2820, Apr 1988.

[106] S. Groth, T. Dornheim, T. Sjostrom, F. D. Malone, W. M. C. Foulkes, and M. Bonitz. Ab initio exchange-correlation free energy of the uniform electron gas at warm dense matter conditions. *Phys. Rev. Lett.*, 119:135001, Sep 2017.

[107] R. J. Gummow, G. Vamvounis, M. B. Kannan, and Y. He. Calcium-ion batteries: Current state-of-the-art and future perspectives. *Advanced Materials*, 30(39):1801702, 2018.

[108] O. Gunnarsson and B. I. Lundqvist. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B*, 13:4274, 1976.

[109] E. Gustafson, B. Holzman, J. Kowalkowski, H. Lamm, A. Y. Li, G. Perdue, S. V. Isakov, O. Martin, R. Thomson, J. Beall, M. Ganahl, G. Vidal, and E. Peters. Large scale multi-node simulations of z2 gauge theory quantum circuits using google cloud platform. In *2021 IEEE/ACM Second International Workshop on Quantum Computing Software (QCS)*, pages 72–79, Los Alamitos, CA, USA, nov 2021. IEEE Computer Society.

[110] M. Hacene, A. Anciaux-Sedrakian, X. Rozanska, D. Klahr, T. Guignon, and P. Fleurat-Lessard. Accelerating vasp electronic structure calculations using graphic processing units. *Journal of computational chemistry*, 33(32):2581–2589, 2012.

[111] J. Hafner. Ab-initio simulations of materials using vasp: Density-functional theory and beyond. *Journal of Computational Chemistry*, 29(13):2044–2078, 2008.

[112] J. Hafner, C. Wolverton, and G. Ceder. Toward Computational Materials Design: The Impact of Density Functional Theory on Materials Research. *MRS Bulletin*, 31(09):659–668, Jan. 2011.

[113] F. D. M. Haldane. Nobel lecture: Topological quantum matter. *Rev. Mod. Phys.*, 89:040502, Oct 2017.

[114] D. Hamann. *Phys. Rev. Lett.*, 76:660, 1996.

[115] S. Hammes-Schiffer. A conundrum for density functional theory. *Science*, 355(6320):28–29, 2017.

[116] J. E. Harriman. Densities, operators, and basis sets. *Phys. Rev. A*, 34:29–39, Jul 1986.

[117] J. Harris. Adiabatic-connection approach to kohn-sham theory. *Phys. Rev. A*, 29(4):1648, 1984.

[118] J. Harris and R. Jones. The surface energy of a bounded electron gas. *J. Phys. F*, 4:1170, 1974.

[119] Y. Hasegawa, J.-I. Iwata, M. Tsuji, D. Takahashi, A. Oshiyama, K. Minami, T. Boku, H. Inoue, Y. Kitazawa, I. Miyoshi, and M. Yokokawa. Performance evaluation of ultra-large-scale first-principles electronic structure calculation code on the k computer. *Int. J. High Perform. Comput. Appl.*, 28(3):335–355, aug 2014.

[120] M. Hauru, A. Morningstar, J. Beall, M. Ganahl, A. Lewis, and G. Vidal. Simulation of quantum physics with tensor processing units: brute-force computation of ground states and time evolution. *arXiv preprint arXiv:2111.10466*, 2021.

[121] C. H. Hendon, L. Colonna-Dashwood, and M. Colonna-Dashwood. The role of dissolved cations in coffee extraction. *Journal of Agricultural and Food Chemistry*, 62(21):4947–4950, 05 2014.

[122] T. Herault, Y. Robert, G. Bosilca, and J. Dongarra. Generic matrix multiplication for multi-gpu accelerated distributed-memory platforms over parsec. In *2019 IEEE/ACM 10th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems (ScalA)*, pages 33–41. IEEE, 2019.

[123] J. A. Hernandez, M. Bethkenhagen, S. Ninet, M. French, A. Benuzzi-Mounaix, F. Datchi, M. Guarguaglini, F. Lefevre, F. Occelli, R. Redmer, T. Vinci, and A. Ravasio. Melting curve of superionic ammonia at planetary interior conditions. *Nature Physics*, 2023.

[124] N. J. Higham. Stable iterations for the matrix square root. *Numerical Algorithms*, 15(2):227–242, 1997.

[125] B. Himmetoglu, A. Floris, S. de Gironcoli, and M. Cococcioni. Hubbard-corrected dft energy functionals: The lda+u description of correlated systems. *International Journal of Quantum Chemistry*, 114(1):14–49, 2014.

[126] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864–B871, Nov 1964.

[127] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864, 1964.

[128] B. Holst, R. Redmer, and M. P. Desjarlais. Thermophysical properties of warm dense hydrogen using quantum molecular dynamics simulations. *Phys. Rev. B*, 77:184201, May 2008.

[129] B. Hourahine, B. Aradi, V. Blum, F. Bonafé, A. Buccheri, C. Camacho, C. Cevallos, M. Y. Deshaye, T. Dumitrică, A. Dominguez, S. Ehlert, M. Elstner, T. van der Heide, J. Hermann, S. Irle, J. J. Kranz, C. Köhler, T. Kowalczyk, T. Kubař, I. S. Lee, V. Lutsker, R. J. Maurer, S. K. Min, I. Mitchell, C. Negre, T. A. Niehaus, A. M. N. Niklasson, A. J. Page, A. Pecchia, G. Penazzi, M. P. Persson, J. Řezáč, C. G. Sánchez, M. Sternberg, M. Stöhr, F. Stuckenberg, A. Tkatchenko, V. W.-z. Yu, and T. Frauenheim. Dftb+, a software package for efficient approximate density functional theory based atomistic simulations. *The Journal of Chemical Physics*, 152(12):124101, 2020.

[130] J. Hubbard. Electron correlations in narrow energy bands. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 276(1365):238–257, 1963.

[131] J. Hubbard. Electron correlations in narrow energy bands. *Proc. R. Soc. Lond. USA A*, 276(1365):238–257, 1963.

[132] W. P. Huhn, B. Lange, V. W.-z. Yu, M. Yoon, and V. Blum. Gpu acceleration of all-electron electronic structure theory using localized numeric atom-centered basis functions. *Computer Physics Communications*, 254:107314, 2020.

[133] M. Huix-Rotllant, A. Ipatov, A. Rubio, and M. E. Casida. Assessment of dressed time-dependent density-functional theory for the low-lying valence states of 28 organic chromophores. *Chem. Phys.*, 391(1):120–129, 2011.

[134] F. Huot, Y.-F. Chen, R. Clapp, C. Boneti, and J. Anderson. High-resolution imaging on tpus. *arXiv preprint arXiv:1912.08063*, 2019.

[135] D. Jacob, G. Stefanucci, and S. Kurth. Mott metal-insulator transition from steady-state density functional theory. *Phys. Rev. Lett.*, 125:216401, Nov 2020.

[136] D. Jacquemin, V. Wathelet, E. A. Perpète, and C. Adamo. Extensive td-dft benchmark: Singlet-excited states of organic molecules. *Journal of Chemical Theory and Computation*, 5(9):2420–2435, 2009.

[137] A. Jain, Y. Shin, and K. A. Persson. Computational predictions of energy materials using density functional theory. *Nature Reviews Materials*, 1(1):15004, 2016.

[138] B. G. Janesko and G. E. Scuseria. Hartree–fock orbitals significantly improve the reaction barrier heights predicted by semilocal density functionals. *The Journal of chemical physics*, 128(24):244112, 2008.

[139] G. O. Jones, A. Yuen, R. J. Wojtecki, J. L. Hedrick, and J. M. Garcia. Computational and experimental investigations of one-step conversion of poly (carbonate) s into value-added poly (aryl ether sulfone) s. *Proceedings of the National Academy of Sciences*, 113(28):7722–7726, 2016.

[140] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, 87(3):897, 2015.

[141] N. Jouppi, D. Yoon, G. Kurian, S. Li, N. Patil, J. Laudon, C. Young, and D. Patterson. A domain-specific supercomputer for training deep neural networks. *Communications of the ACM*, 63:67–78, 06 2020.

[142] N. P. Jouppi, C. Young, N. Patil, D. Patterson, G. Agrawal, R. Bajwa, S. Bates, S. Bhatia, N. Boden, A. Borchers, et al. In-datacenter performance analysis of a tensor processing unit. In *Proceedings of the 44th Annual International Symposium on Computer Architecture*, ISCA '17, pages 1–12, New York, NY, USA, 2017. Association for Computing Machinery.

[143] V. V. Karasiev, J. W. Dufty, and S. B. Trickey. Nonempirical semilocal free-energy density functional for matter under extreme conditions. *Phys. Rev. Lett.*, 120:076401, Feb 2018.

[144] V. V. Karasiev, T. Sjostrom, J. Dufty, and S. B. Trickey. Accurate homogeneous electron gas exchange-correlation free energy for local spin-density calculations. *Phys. Rev. Lett.*, 112:076403, Feb 2014.

[145] V. V. Karasiev, S. B. Trickey, and J. W. Dufty. Status of free-energy representations for the homogeneous electron gas. *Phys. Rev. B*, 99:195134, May 2019.

[146] K. P. Kepp. Comment on "density functional theory is straying from the path toward the exact functional". *Science*, 356(6337):496–496, 2017.

[147] A. Kietzmann, R. Redmer, M. P. Desjarlais, and T. R. Mattsson. Complex behavior of fluid lithium under extreme conditions. *Phys. Rev. Lett.*, 101:070401, Aug 2008.

[148] J. Kim and Y. Jung. A perspective on the density matrix purification for linear scaling electronic structure calculations. *Int. J. Quantum Chem*, 2015.

[149] M.-C. Kim, H. Park, S. Son, E. Sim, and K. Burke. Improved dft potential energy surfaces via improved densities. *J. Phys. Chem. Lett.*, 6(19):3802–3807, 2015.

[150] M.-C. Kim, E. Sim, and K. Burke. Understanding and reducing errors in density functional calculations. *Phys. Rev. Lett.*, 111:073003, Aug 2013.

[151] M.-C. Kim, E. Sim, and K. Burke. Understanding and reducing errors in density functional calculations. *Phys. Rev. Lett.*, 111:073003, Aug 2013.

[152] M.-C. Kim, E. Sim, and K. Burke. Ions in solution: Density corrected density functional theory (dc-dft). *J. Chem. Phys.*, 140(18):18A528, 2014.

[153] Y. Kim, S. Song, E. Sim, and K. Burke. Halogen and chalcogen binding dominated by density-driven errors. *The journal of physical chemistry letters*, 10(2):295–301, 2018.

[154] G. Knizia and G. K.-L. Chan. Density matrix embedding: A simple alternative to dynamical mean-field theory. *Phys. Rev. Lett.*, 109:186404, Nov 2012.

[155] W. Kohn. Theory of the insulating state. *Phys. Rev.*, 133:A171, 1964.

[156] W. Kohn. Theory of the insulating state. *Phys. Rev.*, 133:A171–A181, Jan 1964.

[157] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.

[158] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133, 1965.

[159] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133, 1965.

[160] M. Korth. Density functional theory: Not quite the right answer for the right reason yet. *Angewandte Chemie International Edition*, 56(20):5396–5398, 2017.

[161] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti. Electronic structure calculations with dynamical mean-field theory. *Rev. Mod. Phys.*, 78:865–951, Aug 2006.

[162] J. Kozlowski, D. Perchak, and K. Burke. Generalized gradient approximation made thermal. 2023.

[163] P. Kůs, A. Marek, S. Köcher, H.-H. Kowalski, C. Carbogno, C. Scheurer, K. Reuter, M. Scheffler, and H. Lederer. Optimizations of the eigensolvers in the elpa library. *Parallel Computing*, 85:167–177, 2019.

[164] D. C. Langreth and J. P. Perdew. The exchange-correlation energy of a metallic surface. *Solid State Communications*, 17(11):1425–1429, 1975.

[165] S. Le Pape, L. F. Berzak Hopkins, L. Divol, A. Pak, E. L. Dewald, S. Bhandarkar, L. R. Bennedetti, T. Bunn, J. Biener, J. Crippen, D. Casey, D. Edgell, D. N. Fittinghoff, M. Gatu-Johnson, C. Goyon, S. Haan, R. Hatarik, M. Havre, D. D.-M. Ho, N. Izumi, J. Jaquez, S. F. Khan, G. A. Kyrala, T. Ma, A. J. Mackinnon, A. G. MacPhee, B. J. MacGowan, N. B. Meezan, J. Milovich, M. Millot, P. Michel, S. R. Nagel, A. Nikroo, P. Patel, J. Ralph, J. S. Ross, N. G. Rice, D. Strozzi, M. Stadermann, P. Volegov, C. Yeamans, C. Weber, C. Wild, D. Callahan, and O. A. Hurricane. Fusion energy output greater than the kinetic energy of an imploding shell at the national ignition facility. *Phys. Rev. Lett.*, 120:245003, Jun 2018.

[166] F. Lechermann. Model hamiltonians and basic techniques. In Pavarini et al. [232].

[167] C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.

[168] T. J. Lee and G. E. Scuseria. Achieving chemical accuracy with coupled-cluster theory. In S. Langhoff, editor, *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, volume 13 of *Understanding Chemical Reactivity*, pages 47–108. Springer Netherlands, 1995.

[169] S. Lehtola, C. Steigemann, M. J. Oliveira, and M. A. Marques. Recent developments in libxc — a comprehensive library of functionals for density functional theory. *SoftwareX*, 7:1–5, 2018.

[170] J. Lehtomäki, I. Makkonen, M. A. Caro, A. Harju, and O. Lopez-Acevedo. Orbital-free density functional theory implementation with the projector augmented-wave method. *The Journal of chemical physics*, 141(23):234102, 2014.

[171] K. Lejaeghere *et al.* Reproducibility in density functional theory calculations of solids. *Science*, 351(6280), 2016.

[172] G. Lever, D. J. Cole, N. D. Hine, P. D. Haynes, and M. C. Payne. Electrostatic considerations affecting the calculated homo–lumo gap in protein molecules. *Journal of Physics: Condensed Matter*, 25(15):152101, 2013.

[173] B. G. Levine, C. Ko, J. Quenneville, and T. J. MartÍnez. Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.*, 104(5-7):1039–1051, 2006.

[174] M. Levy. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences of the United States of America*, 76(12):6062–6065, 1979.

[175] M. Levy. Electron densities in search of hamiltonians. *Phys. Rev. A*, 26:1200–1208, Sep 1982.

[176] M. Levy and A. Görling. Correlation-energy density-functional formulas from correlating first-order density matrices. *Phys. Rev. A*, 52(3):R1808, 1995.

[177] M. Levy and H. Ou-Yang. Exact properties of the pauli potential for the square root of the electron density and the kinetic energy functional. *Physical Review A*, 38(2):625, 1988.

[178] M. Levy and J. P. Perdew. Hellmann-feynman, virial, and scaling requisites for the exact universal density functionals. shape of the correlation potential and diamagnetic susceptibility for atoms. *Phys. Rev. A*, 32(4):2010, 1985.

[179] M. Levy and J. P. Perdew. Tight bound and convexity constraint on the exchange-correlation-energy functional in the low-density limit, and other formal tests of generalized-gradient approximations. *Phys. Rev. B*, 48(16):11638, 1993.

[180] M. Levy and J. P. Perdew. Tight bound and convexity constraint on the exchange-correlation-energy functional in the low-density limit, and other formal tests of generalized-gradient approximations. *Phys. Rev. B*, 48(16):11638, 1993.

[181] A. G. Lewis, J. Beall, M. Ganahl, M. Hauru, S. B. Mallick, and G. Vidal. Large-scale distributed linear algebra with tensor processing units. *Proceedings of the National Academy of Sciences*, 119(33):e2122762119, 2022.

[182] E. Lieb and B. Simon. Thomas-fermi theory revisited. *Phys. Rev. Lett.*, 31:681, 1973.

[183] E. H. Lieb. Density functionals for coulomb systems. *Int. J. Quantum Chem.*, 24(3):243–277, 1983.

[184] E. H. Lieb and B. Simon. Thomas-fermi theory revisited. *Physical Review Letters*, 31(11):681, 1973.

[185] E. H. Lieb and B. Simon. The thomas-fermi theory of atoms, molecules and solids. *Advances in mathematics*, 23(1):22–116, 1977.

[186] N. A. Lima, L. N. Oliveira, and K. Capelle. Density-functional study of the mott gap in the hubbard model. *EPL (Europhysics Letters)*, 60(4):601, 2002.

[187] J. Loewen. *Lies My Teacher Told Me: Everything Your American History Textbook Got Wrong*. New Press, 2008.

[188] P.-F. Loos and E. Fromager. A weight-dependent local correlation density-functional approximation for ensembles. *J. Chem. Phys.*, 152(21):214101, 2020.

[189] P.-O. Löwdin. Quantum theory of many-particle systems. iii. extension of the hartree-fock scheme to include degenerate systems and correlation effects. *Phys. Rev.*, 97(6):1509, 1955.

[190] T. Lu, Y.-F. Chen, B. Hechtman, T. Wang, and J. Anderson. Large-scale discrete fourier transform on tpus. *IEEE Access*, 9:93422–93432, 2021.

[191] T. Lu, T. Marin, Y. Zhuo, Y.-F. Chen, and C. Ma. Accelerating mri reconstruction on tpus. In *2020 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–9. IEEE, 2020.

[192] T. Lu, T. Marin, Y. Zhuo, Y.-F. Chen, and C. Ma. Nonuniform fast fourier transform on tpus. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 783–787. IEEE, 2021.

[193] A. R. M. Grüning, A. Marini. Density functionals from many-body perturbation theory: The band gap for semiconductors and insulators. *J. Chem. Phys.*, 124:154108, 2006.

[194] G. D. Mahan. *Many-Particle Physics*. Springer, 3rd edition, New York, 2000.

[195] N. T. Maitra. Perspective: Fundamental aspects of time-dependent density functional theory. *The Journal of Chemical Physics*, 144(22):220901, 2016.

[196] N. T. Maitra. Perspective: Fundamental aspects of time-dependent density functional theory. *J. Chem. Phys.*, 144(22):220901, 2016.

[197] N. March. The local potential determining the square root of the ground-state electron density of atoms and molecules from the schrödinger equation. *Physics Letters A*, 113(9):476–478, 1986.

[198] A. Marek, V. Blum, R. Johanni, V. Havu, B. Lang, T. Auckenthaler, A. Heinecke, H.-J. Bungartz, and H. Lederer. The ELPA library: scalable parallel eigenvalue solutions for electronic structure theory and computational science. *Journal of Physics: Condensed Matter*, 26(21):213201, 2014.

[199] M. A. Marques, N. T. Maitra, F. M. Nogueira, E. K. Gross, and A. Rubio. *Fundamentals of time-dependent density functional theory*, volume 837. Springer, 2012.

[200] M. A. L. Marques, N. T. Maitra, F. M. S. Nogueira, E. K. U. Gross, and A. Rubio, editors. *Fundamentals of Time-Dependent Density Functional Theory*. Number 837 in Lecture Notes in Physics. Springer, Heidelberg, 2012.

[201] C. Marut, B. Senjean, E. Fromager, and P.-F. Loos. Weight dependence of local exchange–correlation functionals in ensemble density-functional theory: double excitations in two-electron systems. *Faraday Discuss.*, 224:402–423, 2020.

[202] T. R. Mattsson and M. P. Desjarlais. Phase diagram and electrical conductivity of high energy-density water from density functional theory. *Phys. Rev. Lett.*, 97:017801, Jul 2006.

[203] D. A. Mazziotti. Significant conditions for the two-electron reduced density matrix from the constructive solution of n representability. *Phys. Rev. A*, 85:062507, Jun 2012.

[204] R. J. McCarty, D. Perchak, R. Pederson, R. Evans, Y. Qiu, S. R. White, and K. Burke. Bypassing the energy functional in density functional theory: Direct calculation of electronic energies from conditional probability densities. *Phys. Rev. Lett.*, 125:266401, Dec 2020.

[205] M. G. Medvedev, I. S. Bushmarinov, J. Sun, J. P. Perdew, and K. A. Lyssenko. Density functional theory is straying from the path toward the exact functional. *Science*, 355(6320):49–52, 2017.

[206] N. D. Mermin. Thermal properties of the inhomogenous electron gas. *Phys. Rev.*, 137:A: 1441, 1965.

[207] Z. A. Moldabekov, T. Dornheim, and A. Cangi. Thermal excitation signals in the inhomogeneous warm dense electron gas. *Scientific Reports*, 12:1093, Jan 2022.

[208] J. R. Moreno, G. Carleo, and A. Georges. Deep learning the hohenberg-kohn maps of density functional theory. *Phys. Rev. Lett.*, 125:076402, Aug 2020.

[209] J. R. Moreno, J. Flick, and A. Georges. Machine learning band gaps from the electron density, 2021.

[210] P. Morgante and R. Peverati. The devil in the details: A tutorial review on some undervalued aspects of density functional theory calculations. *Int. J. Quantum. Chem.*, 120(18):e26332, 2020.

[211] A. Morningstar, M. Hauru, J. Beall, M. Ganahl, A. G. Lewis, V. Khemani, and G. Vidal. Simulation of quantum many-body dynamics with tensor processing units: Floquet prethermalization. *PRX Quantum*, 3(2):020331, 2022.

[212] Á. Nagy. Coordinate scaling and adiabatic connection formula for ensembles of fractionally occupied excited states. *Int. J. Quantum Chem.*, 56(4):225–228, 1995.

[213] Á. Nagy. Theory for a single excited state differential virial theorem. In *Recent Advances In Density Functional Methods: Part III*, pages 247–256. World Scientific, 2002.

[214] A. Nagy. Functional derivative of the kinetic energy functional for spherically symmetric systems. *J. Chem. Phys.*, 135(4):044106, 2011.

[215] Á. Nagy. *Local virial theorem for ensembles of excited states*, page 135. CRC Press, 2013.

[216] N. Nettelmann, B. Holst, A. Kietzmann, M. French, R. Redmer, and D. Blaschke. Ab initio equation of state data for hydrogen, helium, and water and the internal structure of jupiter. *The Astrophysical Journal*, 683(2):1217–1228, aug 2008.

[217] N. Nettelmann, R. Redmer, and D. Blaschke. Warm dense matter in giant planets and exoplanets. *Physics of Particles and Nuclei*, 39:1122–1127, 01 2008.

[218] A. M. N. Niklasson. Expansion algorithm for the density matrix. *Phys. Rev. B*, 66(15):155115, 2002.

[219] A. M. N. Niklasson, C. J. Tymczak, and M. Challacombe. Trace resetting density matrix purification in O(N) self-consistent-field theory. *The Journal of Chemical Physics*, 118(19):8611–8620, 2003.

[220] J. K. Nørskov, F. Abild-Pedersen, F. Studt, and T. Bligaard. Density functional theory in surface chemistry and catalysis. *Proceedings of the National Academy of Sciences*, 108(3):937, 01 2011.

[221] J. K. Nørskov, F. Abild-Pedersen, F. Studt, and T. Bligaard. Density functional theory in surface chemistry and catalysis. *Proceedings of the National Academy of Sciences*, 108(3):937–943, 2011.

[222] J. W. Ochterski, G. A. Petersson, and K. B. Wiberg. A comparison of model chemistries. *Journal of the American Chemical Society*, 117(45):11299–11308, 1995.

[223] U. D. of Energy. Basic research needs for high energy density laboratory physics: Report of the workshop on high energy density laboratory physics research needs. Technical report, Office of Science and National Nuclear Security Administration, 2009.

[224] P. Okun and K. Burke. Semiclassics: The hidden theory behind the success of DFT, 2021.

[225] L. N. d. Oliveira, E. Gross, and W. Kohn. Density-functional theory for ensembles of fractionally occupied states. ii. application to the he atom. *Phys. Rev. A*, 37(8):2821, 1988.

[226] G. Onida, L. Reining, and A. Rubio. Electronic excitations: density-functional versus many-body green's-function approaches. *Rev. Mod. Phys.*, 74(2):601–659, Jun 2002.

[227] B. O'Regan and M. Graetzel. A low-cost, high-efficiency solar cell based on dye-sensitized colloidal tio2 films. *Nature*, 353:737–740, 1991.

[228] A. H. R. Palser and D. E. Manolopoulos. Canonical purification of the density matrix in electronic-structure theory. *Phys. Rev. B*, 58(19):12704–12711, 1998.

[229] Z. Pan and P. Mishra. Hardware acceleration of explainable machine learning using tensor processing units. *arXiv preprint arXiv:2103.11927*, 2021.

[230] R. G. Parr and W. Yang. *Density Functional Theory of Atoms and Molecules*. Oxford University Press, 1989.

[231] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1989.

[232] E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein, editors. Forschungszentrum Jülich, 2011.

[233] R. Pederson and K. Burke. Reassessing the role of exact conditions in density functional theory. *arXiv preprint arXiv:2303.01766*, 2023.

[234] R. Pederson and K. Burke. Reassessing the role of exact conditions in density functional theory, 2023.

[235] R. Pederson, J. Chen, S. R. White, and K. Burke. Conditional probability density functional theory. *Phys. Rev. B*, 105:245138, Jun 2022.

[236] R. Pederson, J. Kozlowski, R. Song, J. Beall, M. Ganahl, M. Hauru, A. G. M. Lewis, Y. Yao, S. B. Mallick, V. Blum, and G. Vidal. Large scale quantum chemistry with tensor processing units. *Journal of Chemical Theory and Computation*, 19(1):25–32, 2023.

[237] D. Perchak, R. J. McCarty, and K. Burke. Correlation energy of the uniform electron gas determined by ground-state conditional probability density functional theory. *Phys. Rev. B*, 105:165143, Apr 2022.

[238] J. P. Perdew. Climbing the ladder of density functional approximations. *MRS bulletin*, 38(9):743–750, 2013.

[239] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865, 1996.

[240] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865, 1996.

[241] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865–3868, Oct 1996. *ibid.* **78**, 1396(E) (1997).

[242] J. P. Perdew, K. Burke, and Y. Wang. Generalized gradient approximation for the exchange-correlation hole of a many-electron system. *Phys. Rev. B*, 54:16533–16539, Dec 1996.

[243] J. P. Perdew, M. Ernzerhof, and K. Burke. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.*, 105(22):9982–9985, 1996.

[244] J. P. Perdew and M. Levy. Physical content of the exact kohn-sham orbital energies: Band gaps and derivative discontinuities. *Phys. Rev. Lett.*, 51:1884–1887, Nov 1983.

[245] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz. Density-functional theory for fractional particle number: Derivative discontinuities of the energy. *Phys. Rev. Lett.*, 49:1691–1694, Dec 1982.

[246] J. P. Perdew, A. Savin, and K. Burke. Escaping the symmetry dilemma through a pair-density interpretation of spin-density functional theory. *Phys. Rev. A*, 51:4531–4541, Jun 1995.

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

[248] J. P. Perdew, W. Yang, K. Burke, Z. Yang, E. K. U. Gross, M. Scheffler, G. E. Scuseria, T. M. Henderson, I. Y. Zhang, A. Ruzsinszky, H. Peng, J. Sun, E. Trushin, and A. Görling. Understanding band gaps of solids in generalized kohn–sham theory. *Proceedings of the National Academy of Sciences*, 114(11):2801–2806, 2017.

[249] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23(10):5048–5079, May 1981.

[250] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Physical Review B*, 23(10):5048, 1981.

[251] C. J. Pickard, I. Errea, and M. I. Eremets. Superconducting hydrides under pressure. *Annual Review of Condensed Matter Physics*, 11(1):57–76, 2020.

[252] S. Pittalis, C. Proetto, A. Floris, A. Sanna, C. Bersier, K. Burke, and E. K. Gross. Exact conditions in finite-temperature density-functional theory. *Phys. Rev. Lett.*, 107(16):163001, 2011.

[253] S. Pittalis, C. R. Proetto, A. Floris, A. Sanna, C. Bersier, K. Burke, and E. K. U. Gross. Exact conditions in finite-temperature density-functional theory. *Phys. Rev. Lett.*, 107:163001, Oct 2011.

[254] J. C. A. Prentice, J. Aarons, J. C. Womack, A. E. A. Allen, L. Andrinopoulos, L. Anton, R. A. Bell, A. Bhandari, G. A. Bramley, R. J. Charlton, R. J. Clements, D. J. Cole, G. Constantinescu, F. Corsetti, S. M.-M. Dubois, K. K. B. Duff, J. M. Escartín, A. Greco, Q. Hill, L. P. Lee, E. Linscott, D. D. O'Regan, M. J. S. Phipps, L. E. Ratcliff, A. R. Serrano, E. W. Tait, G. Teobaldi, V. Vitale, N. Yeung, T. J. Zuehlsdorff, J. Dziedzic, P. D. Haynes, N. D. M. Hine, A. A. Mostofi, M. C. Payne, and C.-K. Skylaris. The onetep linear-scaling density functional theory program. *The Journal of Chemical Physics*, 152(17):174111, 2020.

[255] A. Pribram-Jones and K. Burke. Connection formulas for thermal density functional theory. *Phys. Rev. B*, 93:205140, May 2016.

[256] A. Pribram-Jones, D. A. Gross, and K. Burke. Density functional theory: A theory full of holes? *Annual Review of Physical Chemistry*, 66(1), 2015.

[257] A. Pribram-Jones, D. A. Gross, and K. Burke. Dft: A theory full of holes? *Annual review of physical chemistry*, 66:283–304, 2015.

[258] A. Pribram-Jones, Z.-h. Yang, J. R. Trail, K. Burke, R. J. Needs, and C. A. Ullrich. Excitations and benchmark ensemble density functional theory for two electrons. *J. Chem. Phys.*, 140(18):18A541, 2014.

[259] F. H. H. R. Bauernschmitt, R. Ahlrichs and M. M. Kappes. Experiment versus time dependent density functional theory prediction of fullerene electronic absorption. *J. Am. Chem. Soc.*, 120:5052, 1998.

[260] S. Root, R. J. Magyar, J. H. Carpenter, D. L. Hanson, and T. R. Mattsson. Shock compression of a fifth period element: Liquid xenon to 840 gpa. *Phys. Rev. Lett.*, 105:085501, Aug 2010.

[261] E. Rudberg. Difficulties in applying pure Kohn–Sham density functional theory electronic structure methods to protein molecules. *Journal of Physics: Condensed Matter*, 24(7):072202, 2012.

[262] E. Rudberg, E. H. Rubensson, P. Salek, and A. Kruchinina. Ergo: An open-source program for linear-scaling electronic structure calculations. *SoftwareX*, 7:107–111, 2018.

[263] E. Runge and E. K. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52(12):997, 1984.

[264] E. Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52(12):997, Mar 1984.

[265] I. G. Ryabinkin, S. V. Kohut, and V. N. Staroverov. Reduction of electronic wave functions to kohn-sham effective potentials. *Phys. Rev. Lett.*, 115:083001, Aug 2015.

[266] H. R. Sadeghpour. Resonant electron-hydrogen atom scattering using hyperspherical coordinate method (letter to the editor). *J. Phys. B: At. Mol. Opt. Phys.*, 25:L29, 1992.

[267] F. Sagredo and K. Burke. Accurate double excitations from ensemble density functional calculations. *The Journal of Chemical Physics*, 149(13):134103, 2018.

[268] F. Sagredo and K. Burke. Accurate double excitations from ensemble density functional calculations. *J. Chem. Phys.*, 149(13):134103, 2018.

[269] F. Sagredo and K. Burke. Confirmation of the pplb derivative discontinuity: Exact chemical potential at finite temperatures of a model system. *Journal of Chemical Theory and Computation*, 16(12):7225–7231, 12 2020.

[270] A. Savin. On degeneracy, near degeneracy and density functional theory. In J. M. Seminario, editor, *Recent Developments of Modern Density Functional Theory*, pages 327–357. Elsevier, Amsterdam, 1996.

[271] T. R. Scott, J. Kozlowski, S. Crisostomo, A. Pribram-Jones, and K. Burke. Exact conditions for ensemble density functional theory. 2023.

[272] T. R. Scott, J. Kozlowski, S. Crisostomo, A. Pribram-Jones, and K. Burke. Exact conditions for ensemble density functional theory, 2023.

[273] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy. Generalized kohn-sham schemes and the band-gap problem. *Phys. Rev. B*, 53:3764–3774, Feb 1996.

[274] A. Seidl, A. Görling, P. Vogl, J. A. Majewski, and M. Levy. Generalized kohn-sham schemes and the band-gap problem. *Physical Review B*, 53(7):3764, 1996.

[275] S. Seritan, C. Bannwarth, B. S. Fales, E. G. Hohenstein, C. M. Isborn, S. I. L. Kokkila-Schumacher, X. Li, F. Liu, N. Luehr, J. W. Snyder Jr., C. Song, A. V. Titov, I. S.

Ufimtsev, L.-P. Wang, and T. J. Martínez. Terachem: A graphical processing unit-accelerated electronic structure package for large-scale ab initio molecular dynamics. *WIREs Computational Molecular Science*, 11(2):e1494, 2021.

[276] L. J. Sham and M. Schlüter. Density-functional theory of the energy gap. *Phys. Rev. Lett.*, 51:1888–1891, Nov 1983.

[277] S. Sharma, J. K. Dewhurst, A. Sanna, and E. K. U. Gross. Bootstrap approximation for the exchange-correlation kernel of time-dependent density-functional theory. *Phys. Rev. Lett.*, 107:186401, Oct 2011.

[278] E. Sim, S. Song, and K. Burke. Quantifying density errors in dft. *The journal of physical chemistry letters*, 9(22):6385–6392, 2018.

[279] J. C. Slater. A simplification of the hartree-fock method. *Phys. Rev.*, 81(3):385, 1951.

[280] J. C. Smith, A. Pribram-Jones, and K. Burke. Exact thermal density functional theory for a model system: Correlation components and accuracy of the zero-temperature exchange-correlation approximation. *Phys. Rev. B*, 93:245131, Jun 2016.

[281] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal. The SIESTA method for ab initio order-n materials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745, 2002.

[282] I. V. Solovyev. Combining dft and many-body methods to understand correlated materials. *Journal of Physics: Condensed Matter*, 20(29):293201, 2008.

[283] S. Song, M.-C. Kim, E. Sim, A. Benali, O. Heinonen, and K. Burke. Benchmarks and reliable dft results for spin gaps of small ligand fe(ii) complexes. *Journal of Chemical Theory and Computation*, 14(5):2304–2311, 2018.

[284] D. Spaangberg. Ergoscf xyz - water clusters. `http://www.ergoscf.org/xyz/h2o.php`. Accessed: 11-20-2022.

[285] E. M. Stoudenmire, L. O. Wagner, S. R. White, and K. Burke. One-dimensional continuum electronic structure with the density-matrix renormalization group and its implications for density-functional theory. *Phys. Rev. Lett.*, 109:056402, Aug 2012.

[286] R. Stowasser and R. Hoffmann. What do the kohn-sham orbitals and eigenvalues mean? *Journal of the American Chemical Society*, 121(14):3414–3420, 04 1999.

[287] J. Sun, R. C. Remsing, Y. Zhang, Z. Sun, A. Ruzsinszky, H. Peng, Z. Yang, A. Paul, U. Waghmare, X. Wu, M. L. Klein, and J. P. Perdew. Accurate first-principles structures and energies of diversely bonded systems from an efficient density functional. *Nature Chemistry*, 8(9):831–836, 2016.

[288] J. Sun, A. Ruzsinszky, and J. P. Perdew. Strongly constrained and appropriately normed semilocal density functional. *Phys. Rev. Lett.*, 115:036402, Jul 2015.

[289] Q. Sun and G. K.-L. Chan. Quantum embedding theories. *Accounts of chemical research*, 49(12):2705–2712, 2016.

[290] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. Dover Publishing, Mineola, New York, 1996.

[291] E. J. B. T. Ziegler, A. Rauk. On the calculation of multiplet energies by the hartree-fock-slater method. *Theoret. Chim. Acta*, 43:261, 1977.

[292] P. C. Tacitus. *Annals 15.44*. 109.

[293] T. Tajima. Laser acceleration in novel media. *The European Physical Journal Special Topics*, 223(6):1037–1044, 2014.

[294] L. Talirz, L. M. Ghiringhelli, and B. Smit. Trends in atomistic simulation software usage. *Living Journal of Computational Molecular Science*, 3(1):1–12, 2021.

[295] A. K. Theophilou. The energy density functional formalism for excited states. *J. Phys. C: Solid State Phys.*, 12(24):5419, 1979.

[296] L. H. Thomas. The calculation of atomic fields. *Math. Proc. Camb. Phil. Soc.*, 23(05):542–548, 1927.

[297] L. H. Thomas. The calculation of atomic fields. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 23, pages 542–548. Cambridge Univ Press, 1927.

[298] G. Tirimbò and B. Baumeier. Ab initio modeling of excitons: from perfect crystals to biomaterials. *Advances in Physics: X*, 6(1):1912638, 2021.

[299] D. Tozer and N. Handy. The development of new exchange-correlation functionals. *J. Chem. Phys.*, 108:2545, 1998.

[300] D. J. Tozer. Relationship between long-range charge-transfer excitation energy error and integer discontinuity in kohn–sham theory. *J. Chem. Phys.*, 119(24):12697–12699, 2003.

[301] L. A. Truflandier, R. M. Dianzinga, and D. R. Bowler. Communication: Generalized canonical purification for density matrix minimization. *The Journal of Chemical Physics*, 144(9):091102, 2016.

[302] D. C. Tsui, H. L. Stormer, and A. C. Gossard. Two-dimensional magnetotransport in the extreme quantum limit. *Phys. Rev. Lett.*, 48:1559–1562, May 1982.

[303] C. A. Ullrich. *Time-dependent density-functional theory: concepts and applications*. OUP Oxford, 2011.

[304] C. A. Ullrich. *Time-Dependent Density-Functional Theory*. Oxford University Press, Oxford, 2012.

[305] C. A. Ullrich. Density-functional theory for systems with noncollinear spin: orbital-dependent exchange-correlation functionals and their application to the hubbard dimer. *Arxiv:1805.06417*, May 2018.

[306] C. J. Umrigar and X. Gonze. Accurate exchange-correlation potentials and total-energy components for the helium isoelectronic series. *Phys. Rev. A*, 50(5):3827–3837, Nov 1994.

[307] C. J. Umrigar and X. Gonze. Accurate exchange-correlation potentials and total-energy components for the helium isoelectronic series. *Phys. Rev. A*, 50:3827–3837, Nov 1994.

[308] A. Urban, D.-H. Seo, and G. Ceder. Computational understanding of li-ion batteries. *npj Computational Materials*, 2(1):1–13, 2016.

[309] R. A. Van De Geijn and J. Watts. SUMMA: scalable universal matrix multiplication algorithm. *Concurrency: Practice and Experience*, 9(4):255–274, 1997.

[310] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, and J. Hutter. Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Computer Physics Communications*, 167(2):103–128, 2005.

[311] Z. Velez, C. C. Roggatz, D. M. Benoit, J. D. Hardege, and P. C. Hubbard. Short- and medium-term exposure to ocean acidification reduces olfactory sensitivity in gilthead seabream. *Frontiers in Physiology*, 10:731, 2019.

[312] L. Vogt, R. Olivares-Amaya, S. Kermes, Y. Shao, C. Amador-Bedolla, and A. Aspuru-Guzik. Accelerating resolution-of-the-identity second-order Møller-Plesset quantum chemistry calculations with graphical processing units. *The Journal of Physical Chemistry A*, 112(10):2049–2057, 2008.

[313] D. Vollhardt. Dynamical mean-field approach for strongly correlated materials. In Pavarini et al. [232].

[314] S. Vuckovic, P. Gori-Giorgi, F. Della Sala, and E. Fabiano. Restoring size consistency of approximate functionals constructed from the adiabatic connection. *J. Phys. Chem. Lett.*, 2018.

[315] S. Vuckovic, T. J. Irons, A. Savin, A. M. Teale, and P. Gori-Giorgi. Exchange–correlation functionals via local interpolation along the adiabatic connection. *J. Chem. Theory Comput.*, 12(6):2598–2610, 2016.

[316] S. Vuckovic, T. J. P. Irons, L. O. Wagner, A. M. Teale, and P. Gori-Giorgi. Interpolated energy densities, correlation indicators and lower bounds from approximations to the strong coupling limit of dft. *Phys. Chem. Chem. Phys.*, 19:6169–6183, 2017.

[317] S. Vuckovic, S. Song, J. Kozlowski, E. Sim, and K. Burke. Density functional analysis: The theory of density-corrected dft. *Journal of Chemical Theory and Computation*, 15(12):6636–6646, 2019. PMID: 31682433.

[318] S. Vuckovic, S. Song, J. Kozlowski, E. Sim, and K. Burke. Density functional analysis: The theory of density-corrected dft. *Journal of Chemical Theory and Computation*, 15(12):6636–6646, November 2019. PMID: 31682433.

[319] R. P. W. Kohn, A. Becke. Density functional theory of electronic structure. *J. Phys. Chem.*, 100:12974, 1996.

[320] L. O. Wagner, T. E. Baker, M. Stoudenmire, E., K. Burke, and S. R. White. Kohn-sham calculations with the exact functional. *Phys. Rev. B*, 90:045109, Jul 2014.

[321] L. O. Wagner, E. Stoudenmire, K. Burke, and S. R. White. Reference electronic structure calculations in one dimension. *Phys. Chem. Chem. Phys.*, 14(24):8581–8590, 2012.

[322] L. O. Wagner, E. Stoudenmire, K. Burke, and S. R. White. Guaranteed convergence of the kohn-sham equations. *Physical review letters*, 111(9):093003, 2013.

[323] L. O. Wagner, E. M. Stoudenmire, K. Burke, and S. R. White. Guaranteed convergence of the kohn-sham equations. *Phys. Rev. Lett.*, 111:093003, Aug 2013.

[324] Q. Wang, M. Ihme, Y.-F. Chen, and J. Anderson. A tensorflow simulation framework for scientific computing of fluid flows on tensor processing units. *Computer Physics Communications*, page 108292, 2022.

[325] A. Wasserman, J. Nafziger, K. Jiang, M.-C. Kim, E. Sim, and K. Burke. The importance of being inconsistent. *Annual Review of Physical Chemistry*, 68(1):555–581, 2017.

[326] P. F. Weck, P. E. Kalita, T. Ao, S. D. Crockett, S. Root, and K. R. Cochrane. Shock compression of vanadium at extremes: Theory and experiment. *Phys. Rev. B*, 102:184109, Nov 2020.

[327] C. A. White, B. G. Johnson, P. M. Gill, and M. Head-Gordon. Linear scaling density functional calculations via the continuous fast multipole method. *Chemical Physics Letters*, 253(3-4):268–278, 1996.

[328] S. R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69(19):2863–2866, Nov 1992.

[329] S. R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, Oct 1993.

[330] O. Wilde. *The Canterville Ghost*. 1887.

[331] H. B. Wu and X. W. D. Lou. Metal-organic frameworks and their derived materials for electrochemical energy storage and conversion: Promises and challenges. *Science Advances*, 3(12):eaap9252, 2017.

[332] L. Xi, M. Wang, Y. Liang, Y. Zhao, and Z. Shi. Tunably strained metallacycles enable modular differentiation of aza-arene c–h bonds. *Nature Communications*, 2023.

[333] J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz, and G. K.-L. Chan. Ab initio determination of the crystalline benzene lattice energy to sub-kilojoule/mole accuracy. *Science*, 345(6197):640–643, 2014.

[334] Z.-h. Yang. Second-order perturbative correlation energy functional in the ensemble density-functional theory. *Phys. Rev. A*, 104(5):052806, 2021.

[335] Z.-h. Yang, A. Pribram-Jones, K. Burke, and C. A. Ullrich. Direct extraction of excitation energies from ensemble density-functional theory. *Phys. Rev. Lett.*, 119:033003, Jul 2017.

[336] Z.-h. Yang, A. Pribram-Jones, K. Burke, and C. A. Ullrich. Direct extraction of excitation energies from ensemble density-functional theory. *Phys. Rev. Lett.*, 119(3):033003, 2017.

[337] Z.-h. Yang, J. R. Trail, A. Pribram-Jones, K. Burke, R. J. Needs, and C. A. Ullrich. Exact and approximate kohn-sham potentials in ensemble density-functional theory. *Phys. Rev. A*, 90(4):042501, 2014.

[338] J. Yu, L.-H. Xie, J.-R. Li, Y. Ma, J. M. Seminario, and P. B. Balbuena. Co2 capture and separations using MOFs: computational and experimental studies. *Chemical reviews*, 117(14):9674–9754, 2017.

[339] V. W.-z. Yu, C. Campos, W. Dawson, A. Garcia, V. Havu, B. Hourahine, W. P. Huhn, M. Jacquelin, W. Jia, M. Keceli, et al. Elsi—an open infrastructure for electronic structure solvers. *Computer Physics Communications*, 256:107459, 2020.

[340] L. Zeng, S. B. Jacobsen, D. D. Sasselov, M. I. Petaev, A. Vanderburg, M. Lopez-Morales, J. Perez-Mercader, T. R. Mattsson, G. Li, M. Z. Heising, A. S. Bonomo, M. Damasso, T. A. Berger, H. Cao, A. Levi, and R. D. Wordsworth. Growth model interpretation of planet size distribution. *Proceedings of the National Academy of Sciences*, 116(20):9723–9728, 2019.

[341] L. Zeng, S. B. Jacobsen, D. D. Sasselov, M. I. Petaev, A. Vanderburg, M. Lopez-Morales, J. Perez-Mercader, T. R. Mattsson, G. Li, M. Z. Heising, et al. Growth model interpretation of planet size distribution. *Proceedings of the National Academy of Sciences*, 116(20):9723–9728, 2019.

[342] G. Zhang and E. Pavarini. Optical conductivity, fermi surface, and spin-orbit coupling effects in sr_2rho_4. *Phys. Rev. B*, 99:125102, Mar 2019.

[343] Q. Zhao, R. C. Morrison, and R. G. Parr. From electron densities to kohn-sham kinetic energies, orbital energies, exchange-correlation potentials, and exchange-correlation energies. *Phys. Rev. A*, 50:2138–2142, Sep 1994.

[344] V. W. zhe Yu, F. Corsetti, A. García, W. P. Huhn, M. Jacquelin, W. Jia, B. Lange, L. Lin, J. Lu, W. Mi, A. Seifitokaldani, Álvaro Vázquez-Mayagoitia, C. Yang, H. Yang, and V. Blum. Elsi: A unified software interface for kohn–sham electronic structure solvers. *Computer Physics Communications*, 222:267–285, 2018.

Appendix A

Supplemental Material for Chapter 4

This chapter is a reproduction of the supplemental material for Ref. [271].

A.1 Abstract

Below we list the analytical expressions defining the Hubbard dimer bi-ensemble, and plot various weight-dependent quantities of interest. In all plots we set $t = 1/2$.

A.2 Exact Solution

We begin with the analytic solutions of the Hubbard dimer that were used to create a bi-ensemble of the ground and first-excited singlet states. Solving the dimer Hamiltonian (Eq. 4.26), one obtains the energies:

$$E_i = \frac{2U}{3} + \frac{2r}{3}\cos\left[\theta + \frac{2\pi(i+1)}{3}\right], \qquad i = 0, 1. \tag{A.1}$$

Here we have defined

$$r = \sqrt{3\left(4t^2 + \Delta v^2\right) + U^2}, \qquad \cos\left(3\theta\right) = \frac{9U\left(\Delta v^2 - 2t^2\right) - U^3}{r^3},$$

where t represents the hopping parameter, U the on-site electrostatic self-repulsion, and $\Delta v = v_2 - v_1$ the on-site potential difference. Furthermore, the wavefunction of each state may be written as

$$\left|\Psi_i\right\rangle = \alpha_i\left(\left|12\right\rangle + \left|21\right\rangle\right) + \beta_i^+\left|11\right\rangle + \beta_i^-\left|22\right\rangle, \tag{A.2}$$

with coefficients:

$$\alpha_i = \frac{2t\left(E_i - U\right)}{c_i E_i}, \qquad \beta_i^{\pm} = \frac{U - E_i \pm \Delta v}{c_i},$$

$$c_i = \sqrt{2\left[\Delta v^2 + \left(E_i - U\right)^2\left(1 + 4t^2/E_i^2\right)\right]}.$$

Here $\left|ij\right\rangle$ represents a state in which an electron is present at sites i and j. These analytical expressions (both of the energy and wavefunction) may also be found in the appendices of References 52 and 280.

A.2.1 Densities

We use the same bi-ensemble as described in the main text; the two lowest lying singlet states in the Hubbard dimer. The density of each state is trivially $\Delta n_i = 2\left[\left(\beta_i^-\right)^2 - \left(\beta_i^+\right)^2\right]$, meaning that the ensemble density is

$$\Delta n_w = 2\overline{w}\left[\left(\beta_0^-\right)^2 - \left(\beta_0^+\right)^2\right] + 2w\left[\left(\beta_1^-\right)^2 - \left(\beta_1^+\right)^2\right], \tag{A.3}$$

where $\overline{w} = 1 - w$. Plots of this difference are included below in Fig. A.1 for three interaction strengths, $U = 0$, 1, & 5.

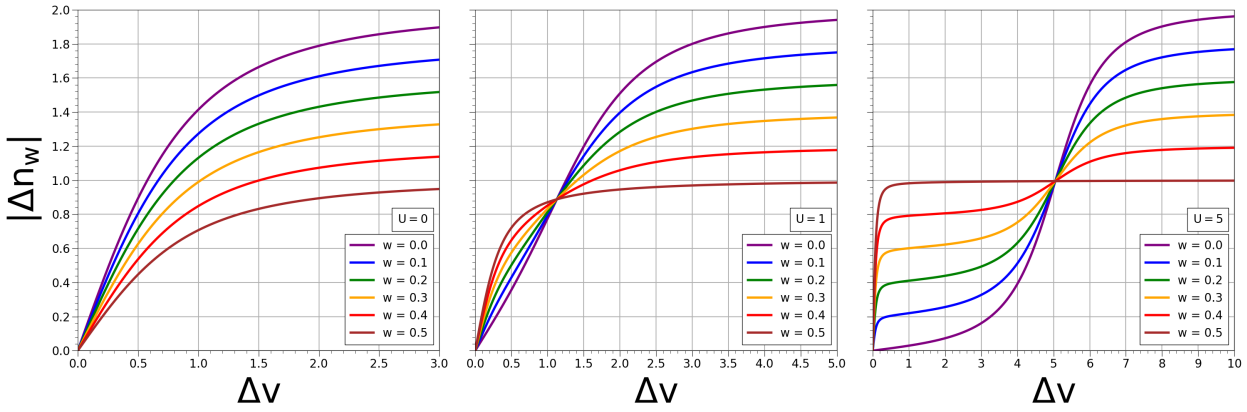

Figure A.1: Absolute value of the density of the Hubbard dimer bi-ensemble, plotted for various weight values as a function of Δv. Here we set $U = 0$ (left), $U = 1$ (center), and $U = 5$ (right).

Analyzing this figure, it is clear that there are vast differences in the behavior of Δn_w with respect to the value of U, illustrating the importance of developing weight-dependent approximations for electronic correlation. Two characteristics are present in all plots of Δn_w, these being an adherence to the symmetric limit ($\Delta v = \Delta n_w = 0$) and a maximum value constraint imposed by the representability condition $|\Delta n_w| \leq 2\overline{w}$, which each curve approaches as $\Delta v \to \infty$.

As predicted by EDFT (eq X), the density is linear in w. For sufficiently large U, there exists a value of Δv at which the initial slope of the first excited state density difference is negative, but it always becomes positive for sufficiently large Δv. Thus there is a specific value of Δv at which the first excited state density vanishes, and all curves meet at that point, independent of w. This point tends to $\Delta v = U$ as U becomes large.

For finite values of U, a trend in curve steepness with respect to the weight is evident for small Δv; the steepness of each curve is directly proportional to the value of w, signifying that ensembles with larger w values approach their maximum value more quickly. The severity of steepness increases drastically as U is increased, as shown by the behavior of the

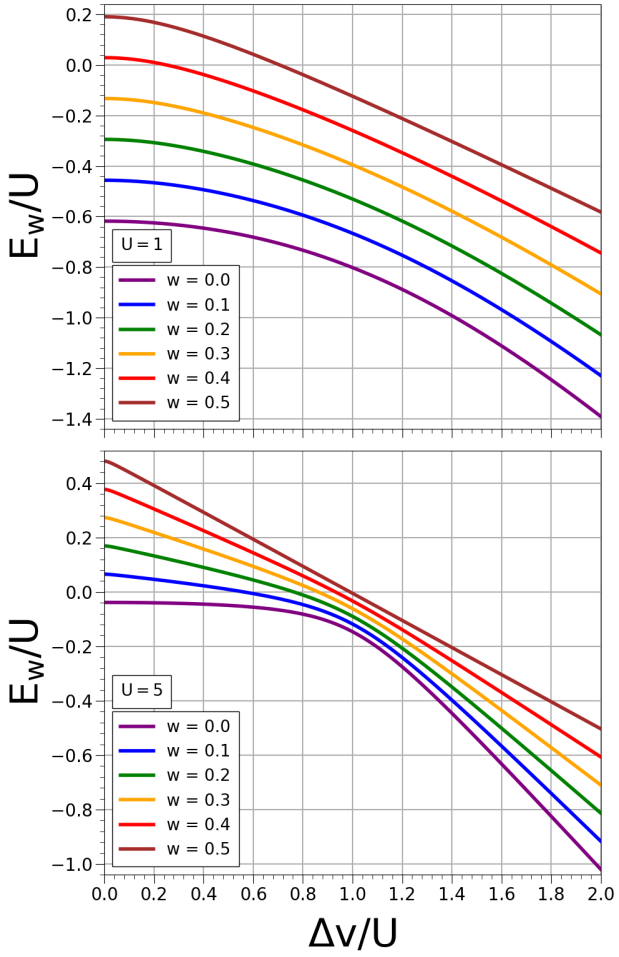

Figure A.2: Total energy of the Hubbard dimer bi-ensemble plotted as a function of Δv for various w values. Here we set $U = 1$ (top) and $U = 5$ (bottom).

$U = 5$ curves as $\Delta v \to 0$. Here, the densities increase rapidly to $|\Delta n_w| \approx 2w$, becoming nearly perfectly anti-symmetric around $\Delta v = U$, with a very sharp dive to 0 for very small Δv.It also appears that all Δn_w curves approach the same value at $\Delta v \approx U$ as $U \to \infty$. As $U \to \infty$ the density forms a step function, flipping from $2w$ to $2\overline{w}$ at $\Delta v = U$ (see Fig. A.13).

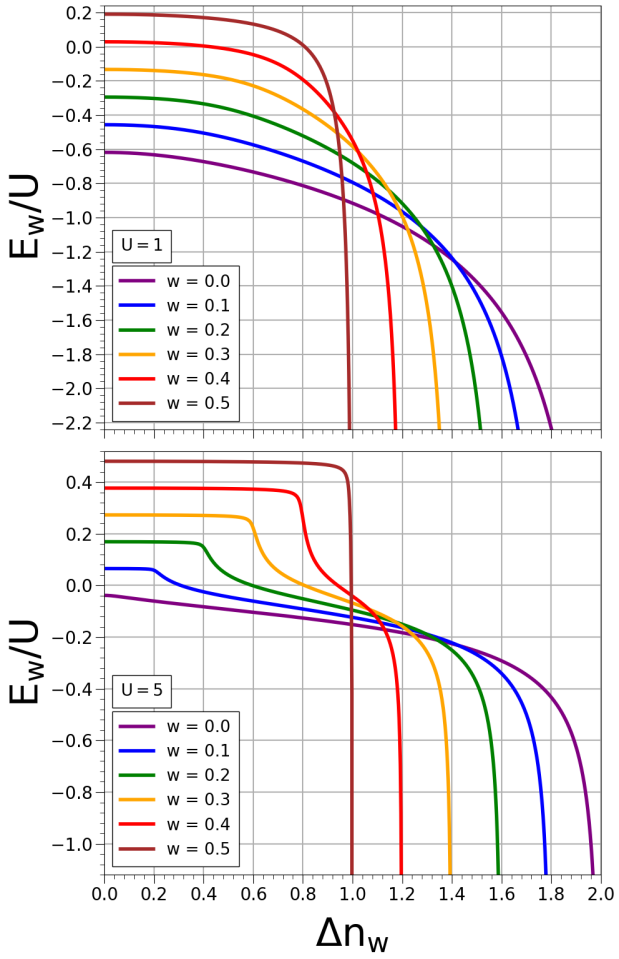

Figure A.3: Total energy of the Hubbard dimer bi-ensemble plotted as a function of Δn_w for various w values. Here we set $U = 1$ (top) and $U = 5$ (bottom).

A.2.2 Total Energies

In this section we depict plots of the total bi-ensemble energy, E_w, defined as the weighted sum of Eq. 4.28. Here we choose to depict two interaction strengths ($U = 1$ and $U = 5$), plotting both as a function of Δv and Δn_w below. We make use of the derivation put forth by Deur *et al.* [52] to define:

$$T_{s,w} = -2t\sqrt{\overline{w}^2 - \Delta n_w^2/4} \tag{A.4}$$

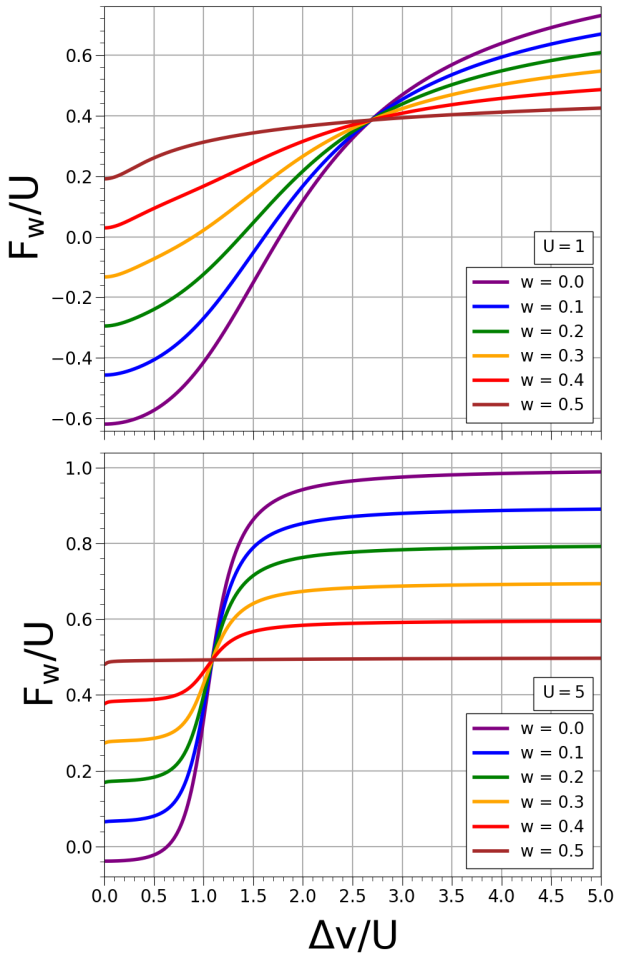

Figure A.4: F_w/U of the Hubbard dimer bi-ensemble plotted as a function of Δv for various w values. Here we set $U = 1$ (top) and $U = 5$ (bottom).

$$E_{Hx,w} = \frac{U}{2}\left(1 + w - \frac{(3w-1)}{\overline{w}^2}\frac{\Delta n_w^2}{4}\right). \tag{A.5}$$

The correlation energies are then found by using the exact expressions:

$$T_{c,w} = T_w - T_{s,w}, \tag{A.6}$$

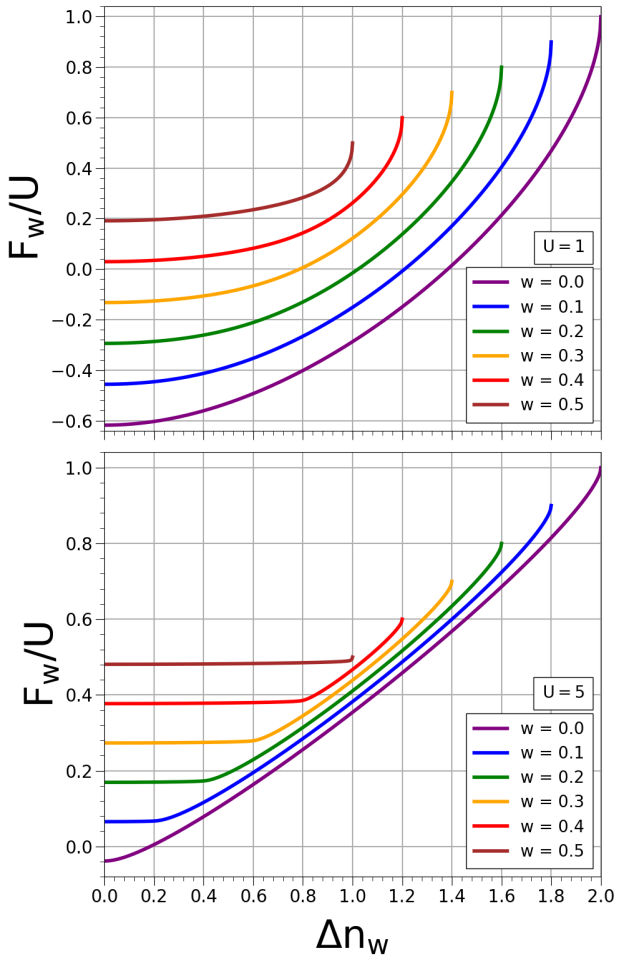

Figure A.5: F_w/U of the Hubbard dimer bi-ensemble plotted as a function of Δn_w for various w values. Here we set $U = 1$ (top) and $U = 5$ (bottom).

$$U_{c,w} = V_{\mathrm{ee},w} - E_{Hx,w}, \tag{A.7}$$

$$E_{c,w} = T_{c,w} + U_{c,w}. \tag{A.8}$$

For fixed Δv, Fig. A.2 illustrates that E_w is correctly linear in w. The curves are rather boring for $U = 1$, but develop a pinch around $\Delta v = U$ as U grows larger (see Sec X).

However, Fig. A.3 shows that, as a functional of Δn, the curves are no longer linear in w. They are not even monotonic. Moreover as $\Delta n_w \to 2\overline{w}$, $E_w \to -\infty$, ensuring the curves cross. There exists interesting behavior as U is increased, with E_w becoming ever more slowly varying with density for $\Delta n_w < 2w$. This behavior may be explained through the relationship between Δn_w and Δv shown in the right panel of Fig. A.1, where there is drastic change in Δn_w for $\Delta v \approx 0$.

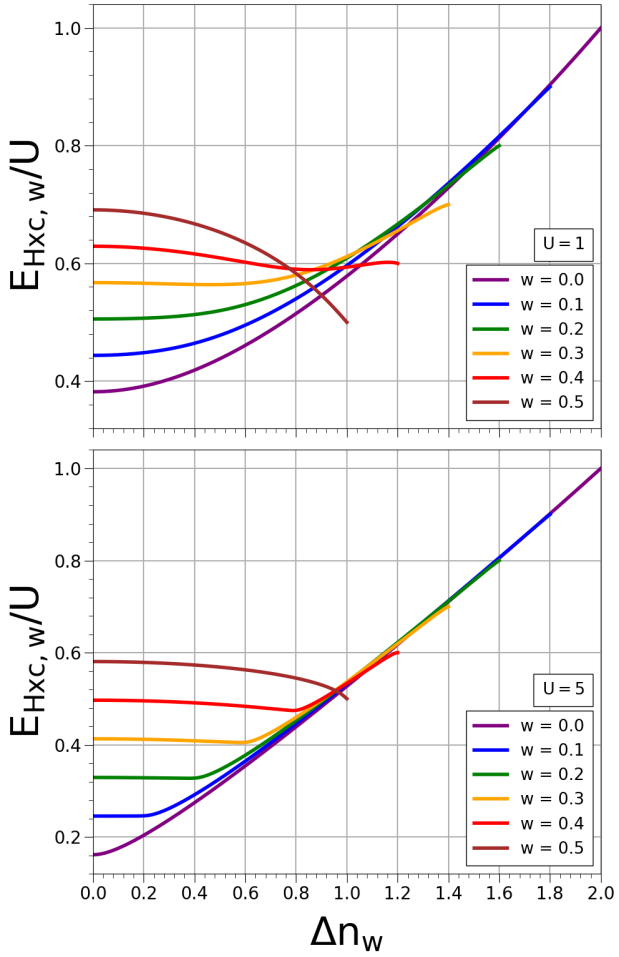

Figure A.6: $E_{\mathrm{HXC},w}/U$ of the Hubbard dimer bi-ensemble plotted as a function of Δn_w for various w values. Here we set $U = 1$ (top) and $U = 5$ (bottom).

We also examine the properties of the Hubbard dimer equivalent of the universal part of the density functional, $F_w = E_w - \Delta v \Delta n_w/2$, plotting again as a function of Δv and Δn_w in Figs. A.4 and A.5 respectively. In Fig. A.4, one can see that F_w is linear with respect to

bi-ensemble weight when plotted as a function of Δv. This characteristic is not present in Fig. A.5, where monotonicity is broken as $\Delta n_w \to 2\overline{w}$. Furthermore, as U is increased, one can see the appearance of regimes around $\Delta n_w = 2w$, with F_w being nearly independent of Δn_w for $\Delta n_w < 2w$ and linearly increasing as $\Delta n_w > 2w$. Additionally, the curves depicting F_w tend to flatten as w increases.

In contrast, we plot $E_{\text{HXC},w}$ in Fig. A.6 and see it is non-monotonic in w. The curvature of $E_{\text{HXC},w}$ changes from convex to concave as the weight is increased and the $E_{\text{HXC},w}$ curves cross each other at various points; with the most noticeable crossings happening at $U = 1$. However, the curves cross at all the values of U plotted.

We conclude that curves become non-monotonic when plotted as a functional of the density instead of the potential, as the curves of Fig. A.1 are definitely not monotonic.

A.2.3 Correlation Inequalities

In this section we depict the correlation inequalities (Eq. 4.21), showing more cases of Fig. 4.2 of the main text. We highlight in Fig. A.7 the definite signs of the correlation energy and its components. This validates results initially introduced by Pribram-Jones *et al.* [258]

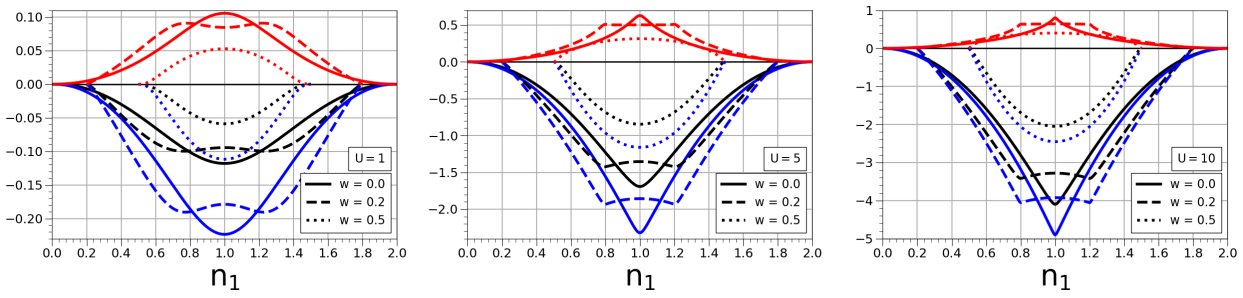

Figure A.7: Variation of the potential (blue), kinetic (red), and total (black) correlation energies in the Hubbard dimer bi-ensemble, plotted as functions of site-occupation of the first site for various weights. We set $U = 1$ in the left, $U = 5$ in the center, and $U = 10$ in the right.

Looking at Figs. A.8 and A.9, one can see that the correlation inequalities of Eq. 4.13 are

satisfied for all values of Δn_w. As noted in the main text, there exists a clear trend with respect to weight for the symmetric dimer, with the ground-state having the maximum magnitude in each plot, and decreasing in magnitude with w. This trend no longer holds for any nonzero value of Δn_w. Alternative approaches were implemented in which Δv was held fixed, again showing no clear trend for asymmetric dimers. Furthermore, it is clear that the inequalities of Eq. 4.13 become equalities as $\Delta n_w \to 2\overline{w}$, explaining the flat behavior of the $w = 0.5$ curves for $\Delta n_w = 1$. It also appears that increasing U morphs the shape of each curve, breaking symmetry around $\lambda = 0.5$ for E_{C} and U_{C}.

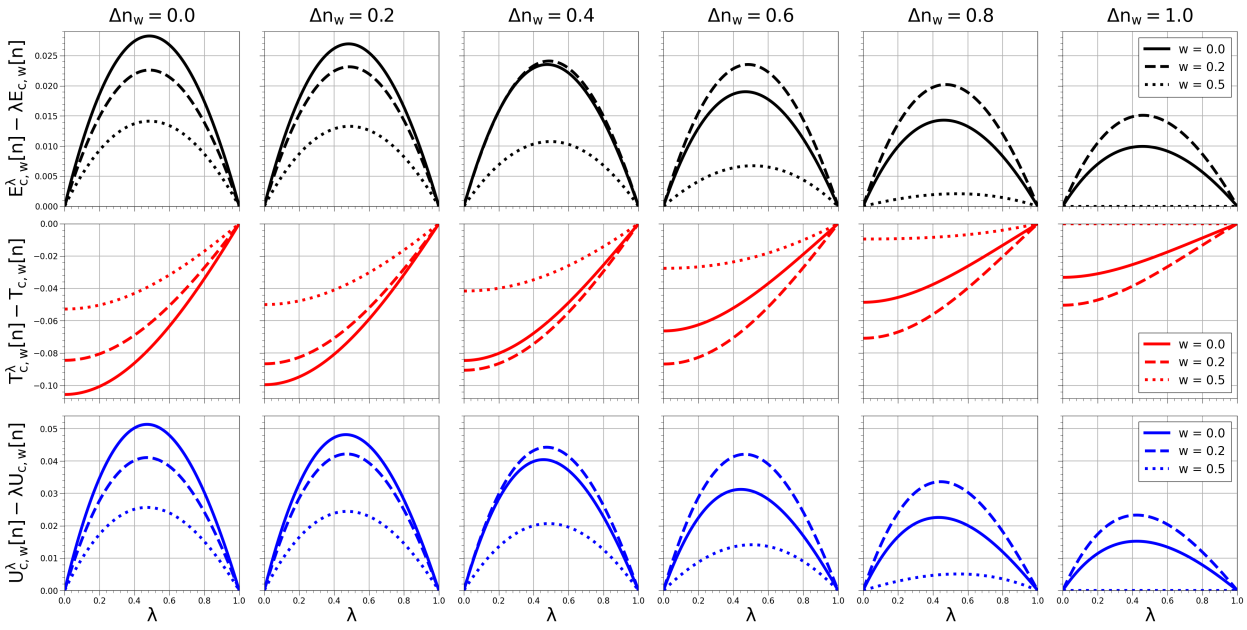

Figure A.8: Correlation inequalities (Eq. 4.21) for the total (top), kinetic (middle), and potential (bottom) correlation energies, depicted by varying λ in the Hubbard dimer bi-ensemble with $U = 1$.

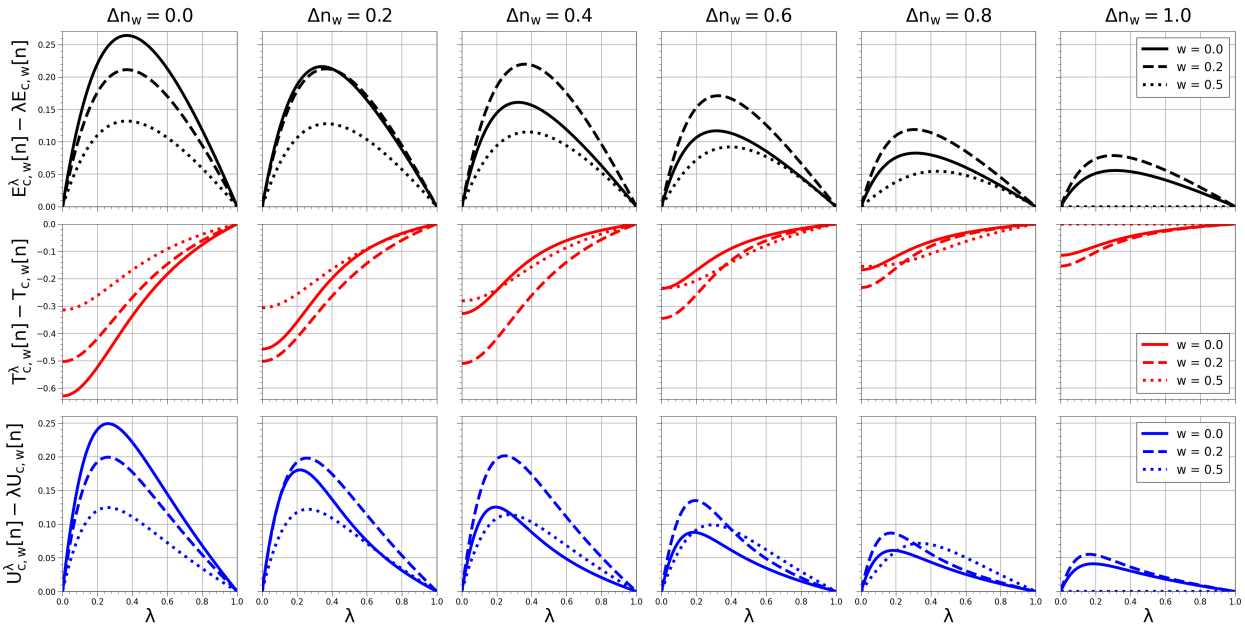

Figure A.9: Correlation inequalities (Eq. 4.21) for the total (top), kinetic (middle), and potential (bottom) correlation energies, depicted by varying λ in the Hubbard dimer bi-ensemble with $U = 5$.

A.2.4 Adiabatic Connection

In this section we include plots depicting more adiabatic connection curves than just that of Fig. 4.3 of the main text.

We first note that all curves of a given weight look similar to ground-state DFT curves. They are monotonically decreasing and are convex. Looking at Fig. A.10 one can observe an interesting change in ordering of the HX energy values based on the weights. For $\Delta n_w < 0.6$ the HX energy monotonicly decreases in value as the weights increase and the spacing between the values shrinks. However, after that point the ordering shifts and the $w = 0.5$ weight has the maximum HX value. Additionally, as the $\Delta n_w \to 1$ the asymptotic value of the HXC expression is same for all weights.

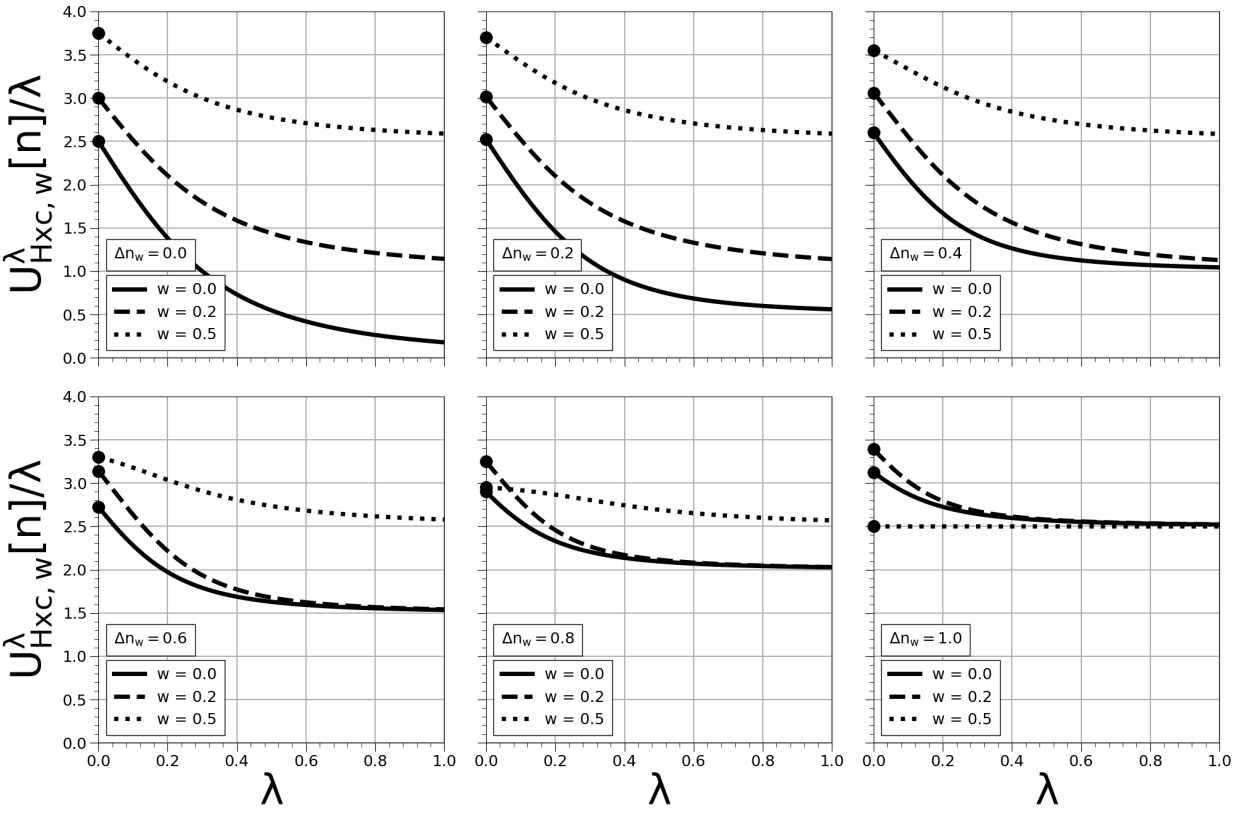

Figure A.10: Ensemble adiabatic connection for $U = 5$ and various Δn_w; circles represent the weight-dependent HX energy, which the HXC expression approaches as $\lambda \to 0$.

A.3 Strong Correlation Limits of Energy Components

A.3.1 Large U Expansions

For fixed w, as U becomes large, one can either keep Δv fixed, or $\Delta v/U$ fixed. The former was explored by Deur et al. [53], and produces the pale blue curves of Fig. 5 of the main text. As is clear from the figure, the blue curves yield the correct answer only for $|\Delta n_w| \leq 2w$, which shrinks to a point as $w \to 0$.

The appropriate expansion to find $E_{\mathrm{C},w}(\Delta n)$ for large U is a different one. We take $U \to \infty$, but keep $\Delta v/U$ fixed. This must be done to include values of Δn away from $\Delta n \approx 0$, while including all allowed values of Δn. A careful expansion yields the total energy as a function of $x = \Delta v/U$:

$$E_w(x) \to U\left(g_w^{(0)}(x) + \frac{g_w^{(2)}(x)}{U^2} + \dots\right), \tag{A.9}$$

where

$$g_w^{(0)}(x) = \frac{1}{3}\left(2 - (c - \sqrt{3}s)h\right) - \frac{2shw}{\sqrt{3}}, \tag{A.10}$$

and

$$g_w^{(2)}(x) = \frac{1}{2h}\left(\frac{(\alpha-3)s}{\sqrt{3}} - (\alpha+1)c\right) + \frac{1}{h}\left(\alpha c + \sqrt{3}s\right)w, \tag{A.11}$$

where $\alpha = |4x/(x^2-1)|$, $c = \cos(\phi)$, and $s = \sin(\phi)$ with,

$$\phi = \frac{1}{3}\cos^{-1}\left(\frac{3h^2-4}{h^3}\right), \qquad h = \sqrt{3x^2+1}. \tag{A.12}$$

The angle ϕ is positive for all values of x, where it takes its maximal value of $\pi/3$ as $x \to 0$

and minimal value of 0 as $x \to \pm 1$, and in the limit $\phi(x \to \pm\infty) = \pi/6$. Because the angle is constrained to $0 \leq \phi \leq \pi/3$, the sine and cosine must always satisfy $0 \leq c, s \leq \sqrt{3}/2$.

From this we have the corresponding density via the exact expression $\Delta n_w = 2dE_w/d(\Delta v)$,

$$\Delta n_w(x) = 2\left(g_w^{(0)\prime}(x) + \frac{g_w^{(2)\prime}(x)}{U^2} + \dots\right), \tag{A.13}$$

where primes denote derivatives with respect to x. Retaining only zero-order terms yields,

$$g_w^{(0)\prime}(x) = \frac{1}{h}\left(\frac{(\gamma - 3x)s}{\sqrt{3}} - (\gamma + x)c\right) + \frac{2}{h}\left(\gamma c + \sqrt{3}xs\right)w, \tag{A.14}$$

with $\gamma = sgn(x(1-x^2))$. Because the expansion in U is singular near $x = 0$ and $x = \pm 1$, $g_w^{(2)}(x)$ diverges at $|x| = 1$, and even $n_w^{(0)}(x) = 2g_w^{(0)\prime}(x)$ contains discontinuous steps. While formally correct in the limit $U \to \infty$, the exact density cannot have such steps, due to the Hohenberg-Kohn theorems. We therefore smooth Eq. A.14 with exponentials that become infinitely sharp as $U \to \infty$:

$$\Delta n_w \approx \frac{x}{|x|}(f(|x|) - 1)\left(1 + (1 - 2w)\tanh\left(\frac{\beta(|x| - 1)}{2}\right)\right), \tag{A.15}$$

where $f(x) = (\exp(\beta x) + 1)^{-1}$ is the Fermi-Dirac distribution with $\beta = 5U$. This is plotted in Fig. A.13 and is compared with the exact density. As $U \to \infty$, Eq. A.15 matches Eq. A.14.

Finally we subtract the remaining components to find the correlation energy:

$$E_{C,w}(\Delta n_w(x)) \approx U\left(g_w^{(0)}(x) - \frac{x\Delta n_w(x)}{2} - e_{Hx,w}(\Delta n_w(x))\right) - T_{s,w}(\Delta n_w(x)) + \frac{g_w^{(2)}(x)}{U}, \tag{A.16}$$

where $e_{Hx,w}(\Delta n_w) = E_{Hx,w}(\Delta n_w)/U$. Including only lowest order, and inserting the smooth density, Eq. A.15, yields the plot in figure Fig. A.11, and the curves in Fig. 5 of the main

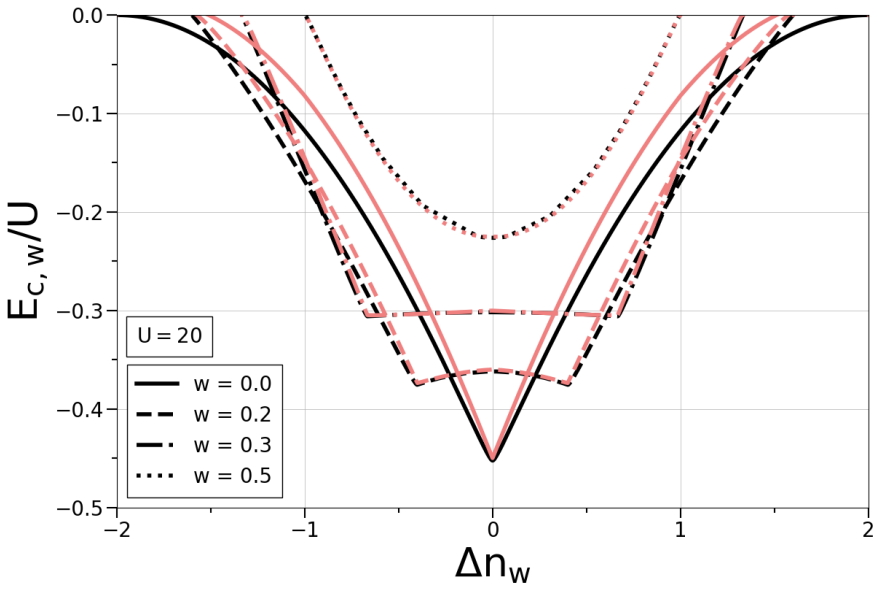

Figure A.11: The exact (black) correlation energy as a function of the exact density and the leading-order expansion in large U (light red) for the correlation energy, using A.10, plotted as a function of the smooth approximation for the density, A.15.

text. We also plot the the next order contribution to the correlation energy, Eq. A.16, in Fig. A.12. To avoid divergences from α at $x = \pm 1$ we use a smooth approximation to the absolute values that appear in the denominator of Eq. A.11.

A.3.2 Correlation Energy on the Adiabatic Connection

We produce an expression for $E^{\lambda}_{C,w}(\Delta n_w)$ where Δn_w is kept fixed for each λ along the adiabatic connection. For sufficiently large U we neglect all terms of $\mathcal{O}(1/U^2)$ and lower. In this limit, $\Delta n_w(x) \to \Delta n^{(0)}_w(x)$, and by the adiabatic connection construction we have the requirement that $\Delta v/U \approx \Delta v^{\lambda}/(\lambda U)$, where Δv^{λ} is the λ-dependent potential that keeps Δn_w fixed along the connection. As a consequence, to leading-order,

$$\Delta v^{\lambda}(\Delta n_w) \approx \lambda \Delta v^{(0)}(\Delta n_w), \tag{A.17}$$

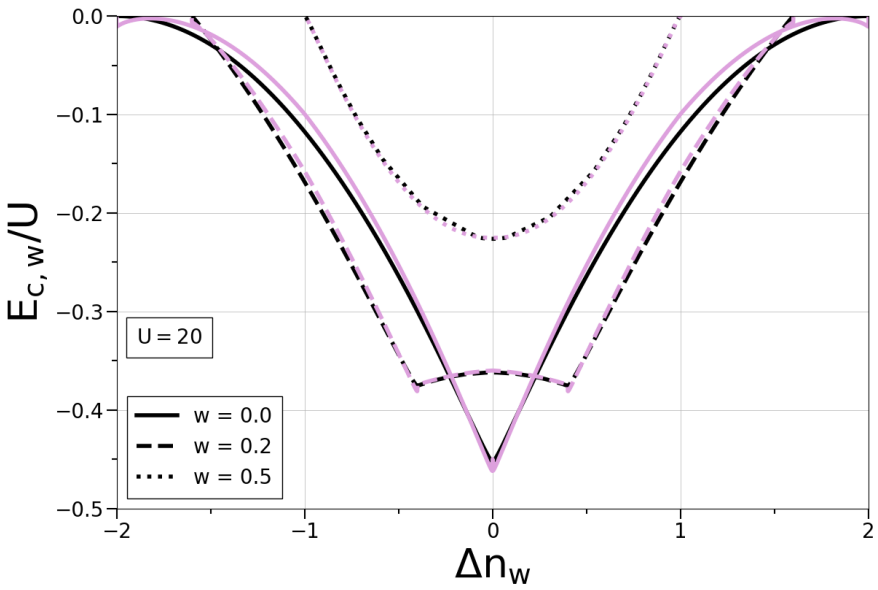

Figure A.12: The exact (black) correlation energy as a function of the exact density and our higher-order expansion in large U (light purple) for the correlation energy, A.16, plotted as a function of the smooth approximation for the density, A.15.

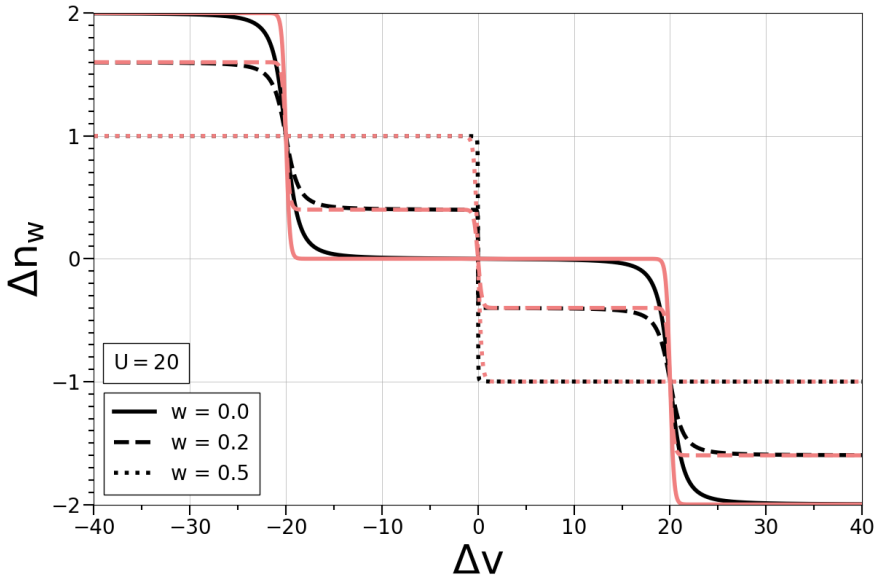

Figure A.13: Smooth approximation for the density A.15 (light red) and the exact density (black) A.3 plotted against Δv.

and thus,

$$E_w^\lambda(\Delta n_w) \approx \lambda U g_w^{(0)}(x^{(0)}(\Delta n_w)) + \frac{g_w^{(2)}(x^{(0)}(\Delta n_w))}{\lambda U}, \tag{A.18}$$

where $x^{(0)}(\Delta n_w) = \Delta v^{(0)}(\Delta n_w)/U$ is the inversion of the leading-order density-potential map Eq. A.17. To produce the correlation energy we subtract from Eq. A.18 the external potential energy (along the connection curve), the KS kinetic energy Eq. A.4, and the HX energy Eq. A.5,

$$E_{\mathrm{C},w}^\lambda(\Delta n_w) \approx \lambda U \Big(g_w^{(0)}(x^{(0)}(\Delta n_w)) - \frac{x^{(0)}(\Delta n_w)\Delta n_w}{2} - e_{Hx,w}(\Delta n_w) \Big) - T_{s,w}(\Delta n_w) + \frac{g_w^{(2)}(x^{(0)}(\Delta n_w))}{\lambda U}, \tag{A.19}$$

where Δn_w remains fixed for all λ.

A.3.3 Contributions to the Energy

Equation 20 of the main text yields expressions for the separate kinetic and potential contributions to the correlation energy,

$$T_{\mathrm{C},w}(\Delta n_w) \approx \frac{2 g_w^{(2)}(x^{(0)}(\Delta n_w))}{U} - T_{s,w}(\Delta n_w), \tag{A.20}$$

$$U_{\mathrm{C},w}(\Delta n_w) \approx U \Big(g_w^{(0)}(x^{(0)}(\Delta n_w)) - \frac{x^{(0)}(\Delta n_w)\Delta n_w}{2} - e_{Hx,w}(\Delta n_w) \Big) - \frac{g_w^{(2)}(x^{(0)}(\Delta n_w))}{U}. \tag{A.21}$$

From the separate contributions of the correlation energy we deduce that,

$$T_w(\Delta n_w) \approx \frac{2g_w^{(2)}(x^{(0)}(\Delta n_w))}{U}, \tag{A.22}$$

$$V_{\mathrm{ee},w}(\Delta n_w) \approx U\left(g_w^{(0)}(x^{(0)}(\Delta n_w)) - \frac{x^{(0)}(\Delta n_w)\Delta n_w}{2}\right) - \frac{g_w^{(2)}(x^{(0)}(\Delta n_w))}{U}. \tag{A.23}$$

We plot Eqs. A.20-A.23 in Fig. A.14 with the exact expressions in black and the approximate expressions evaluated with the smooth density in purple. In all plots we use the same smooth approximation to the absolute values in the denominators of $g_w^{(2)}(x)$, Eq. A.11. There are errors in the plots of the correlation kinetic energy, but these vanish as $U \to \infty$.

We compare our result in Eq. A.16, for $|\Delta n_w| \leq 2w$, with the symmetric limit expansion of Deur et al. [53]. To properly compare our approximate correlation energy to the previously reported expansion in the symmetric limit, we produce the weight-dependent constant that vanishes in the limit $U \to \infty$,

$$\frac{E_{\mathrm{C},w}(\Delta n_w)}{U} \approx \frac{\overline{w}}{U} - \frac{1}{2}\left(\overline{w} - \frac{(3w-1)}{\overline{w}^2}\frac{\Delta n_w^2}{4}\right), \tag{A.24}$$

which is derived following the procedure in Ref [53]. In Fig. 5 of the main text we plot our approximate leading-order correlation energy Eq. A.16 evaluated with smooth density along with Eq. A.24.

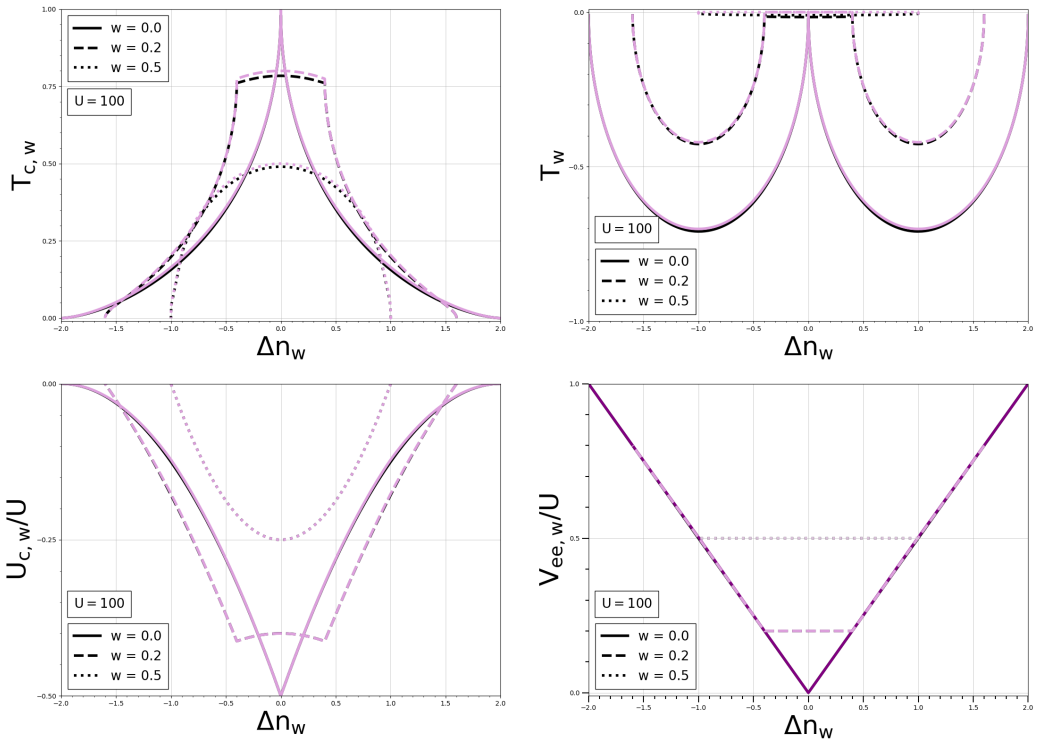

Figure A.14: $T_{c,w}$ (A.20), $U_{c,w}$ (A.21), T_w (A.22), and $V_{\mathrm{ee},w}$ (A.23) for various values of w and $U = 100$. The exact values in the strongly correlated limit are represented by the black curves, which are exact as $U \to \infty$.

A.4 Ensemble Hartree-Fock Approximation

We now turn our attention to the ensemble generalization of Hartree-Fock. We define the Hartree-Fock solution of each state (analogous to Eq. A.2) to be

$$|\Phi_i\rangle = \alpha_i^{\mathrm{HF}} \left(|12\rangle + |21\rangle\right) + \beta_i^{+\mathrm{HF}} |11\rangle + \beta_i^{-\mathrm{HF}} |22\rangle \,, \tag{A.25}$$

with coefficients determined to first order in U:

$$\begin{aligned}
\alpha_0^{\mathrm{HF}} &= 2t/c_0^{\mathrm{HF}} & \alpha_1^{\mathrm{HF}} &= -\Delta v_{\mathrm{eff},w}^2/2tc_1^{\mathrm{HF}} \\
\beta_0^{\pm\mathrm{HF}} &= 1/2 \pm \Delta v_{\mathrm{eff},w}/c_0^{\mathrm{HF}} & \beta_1^{\pm\mathrm{HF}} &= \pm\Delta v_{\mathrm{eff},w}/c_1^{\mathrm{HF}} \\
c_0^{\mathrm{HF}} &= 2\sqrt{4t^2 + \Delta v_{\mathrm{eff},w}^2} & c_1^{\mathrm{HF}} &= \Delta v_{\mathrm{eff},w}\sqrt{2 + \Delta v_{\mathrm{eff},w}^2/2t^2}
\end{aligned}$$

Here the weight-dependent effective mean-field potential $\Delta v_{\mathrm{eff},w}$ takes the form [53]

$$\Delta v_{\mathrm{eff},w} = \Delta v + \frac{(1-3w)}{\overline{w}^2}\frac{U\Delta n_w^{\mathrm{HF}}}{2} \tag{A.26}$$

where Δn_w^{HF} is found from (A.3) with coefficients as above.

A.4.1 Densities and Total Energy

The self-consistent EHF density is found numerically throughout this work by solving (A.26). Plots of the exact/EHF self-consistent site-occupation differences are included below in Fig. A.15 for various interaction strengths, $U = 0$, 1, & 5. Here (and for the remainder of this section) we denote the exact solution using solid curves, and the EHF approximation using dashed curves. Looking at the left panel ($U = 0$) of Fig. A.15, the EHF approximation matches the exact density exactly, as expected in the limit of weak correlation. One can see that the exact/EHF densities (regardless of weight) begin to differ as U is increased, but must

always match at the origin (where $\Delta v = \Delta n_w = 0$) and as $\Delta v \to \infty$ (where $|\Delta n_w| = 2\overline{w}$). This behavior would be expected to hold for larger U values, although as noted previously in Ref. 53, there exists a critical interaction strength U_{crit} at which nonphysical behavior is observed for symmetric dimers with bi-ensemble weight $w \geq 1/3$. This critical interaction strength is

$$U_{\text{crit}} = \frac{\overline{w}}{3w - 1}. \tag{A.27}$$

For $w \to \frac{1}{2}$, $U_{\text{crit}} \to 1$. At this point the energy expression for dimers with $U > U_{\text{crit}}$ have multiple degenerate minima. This explains the deviation from expected behavior for $w = 0.4, 0.5$ in the right panel of Fig. A.15, where both curves approach a finite value as $\Delta v \to 0$.

We also plot the exact/EHF bi-ensemble total energy below, using two interaction strengths ($U = 1$ and $U = 5$) to examine the EHF approximation in more detail. We depict this quantity as a function of Δv below in Fig. A.16 and separate each w value in new plots to better illustrate the weight-dependence of E_w and E_w^{HF}.

Analyzing Fig. A.16, one can see that the EHF approximation obeys the variational principle for all viable weight values (where $w \leq 0.5$); as the weight of the first-excited state is increased, both the exact/EHF energy becomes more positive for all Δv. We note that the EHF approximation always approaches the exact energy as $\Delta v \to \infty$, as shown previously for the ground state. [37]

A.4.2 Correlation Energy

Below we provide plots of the total weight-dependent correlation energy, as well as its kinetic/potential contributions, in Figs. A.17, A.18, and A.19. Note that the definition of the

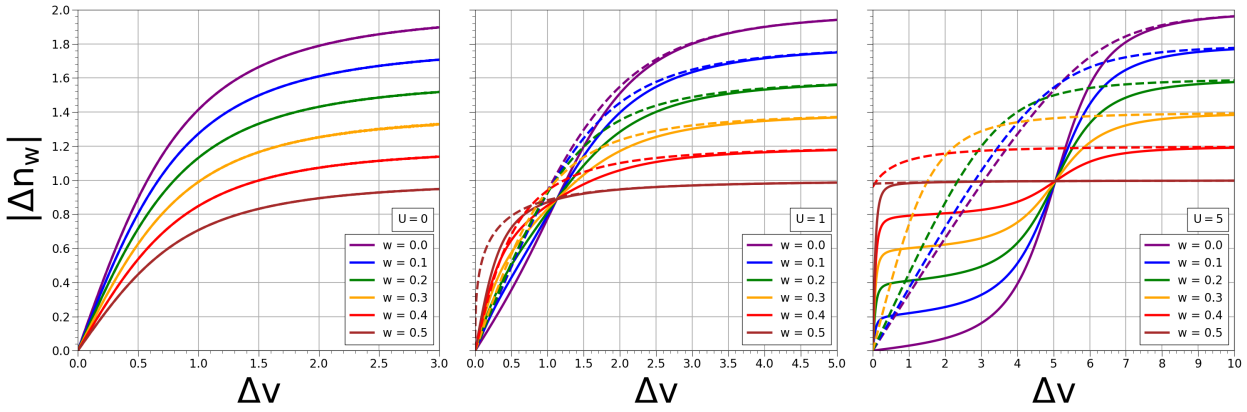

Figure A.15: Absolute value of the density of the Hubbard dimer bi-ensemble, plotted for various weight values as a function of Δv. Here we set $U = 0$ (left), $U = 1$ (center), and $U = 5$ (right). Dashed curves represent the ensemble HF approximation, and the solid curves are exact.

EHF correlation energy is

$$E_{\mathrm{C},w}^{\mathrm{HF}} = E_w[n_w] - E_w^{\mathrm{HF}}[n_w^{\mathrm{HF}}], \tag{A.28}$$

where each energy functional has been minimized by its respective weight-dependent self-consistent density.

Looking at Fig. A.17, it is evident that the behavior of the correlation energy greatly depends upon the value of w. We find that the inequality relating the exact/approximate correlation energy holds for all ensembles (i.e. $E_{\mathrm{C}}^{\mathrm{HF}} \geq E_{\mathrm{C}}$ for all weights). We show that a different trend exists for the kinetic/potential correlation components, as the inequalities describing the ground state ($T_{\mathrm{C}}^{\mathrm{HF}} \leq T_{\mathrm{C}}$ and $U_{\mathrm{C}}^{\mathrm{HF}} \leq U_{\mathrm{C}}$) no longer apply for ensembles with $w \neq 0$.

Note that for each of these quantities, the EHF approximate correlation energy matches the exact solution at $\Delta v = 0$, except for strongly correlated systems with $w \geq 1/3$ (due to the nonphysical behavior in this regime discussed previously).

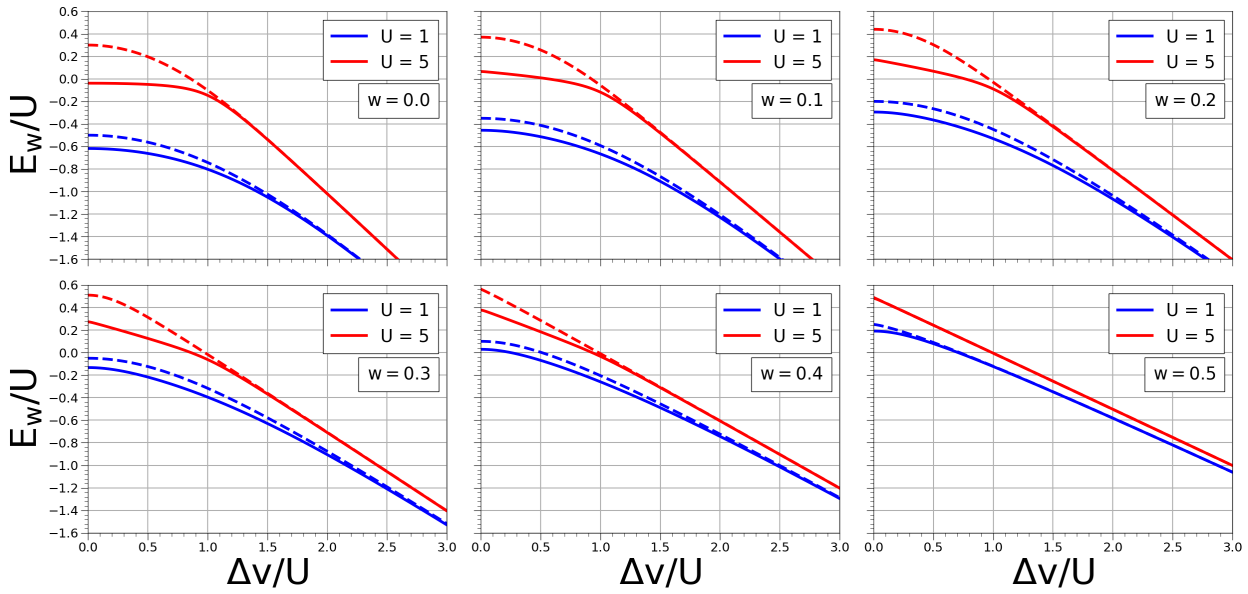

Figure A.16: Total energy of the Hubbard dimer bi-ensemble plotted as a function of Δv for various w values. Dashed lines are the HF approximation and the solid lines are exact.

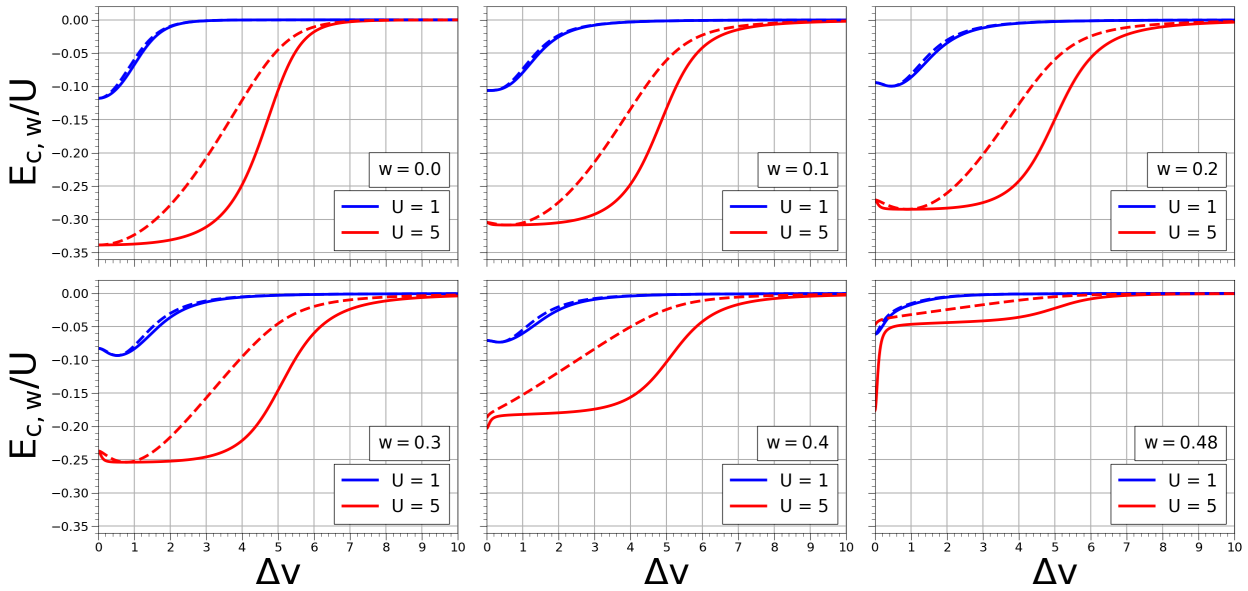

Figure A.17: Total correlation energy of the Hubbard dimer bi-ensemble plotted as a function of Δv for various w values. Dashed lines are the HF approximation and the solid lines are exact.

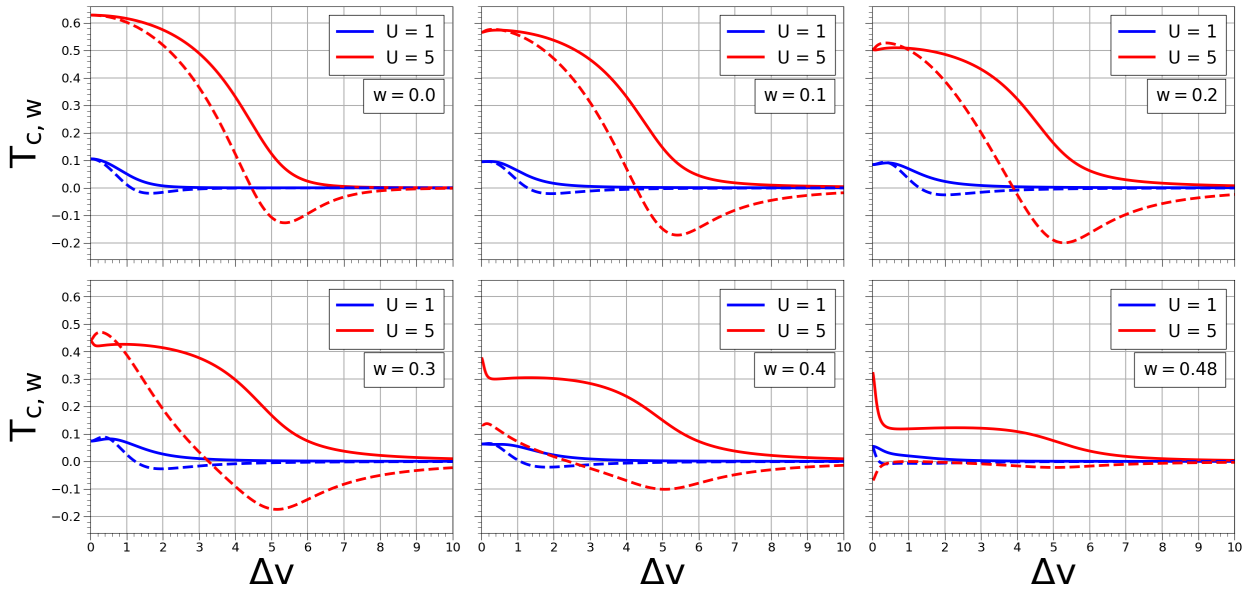

Figure A.18: Kinetic correlation energy of the Hubbard dimer bi-ensemble plotted as a function of Δv for various w values. Dashed lines are the HF approximation and the solid lines are exact.

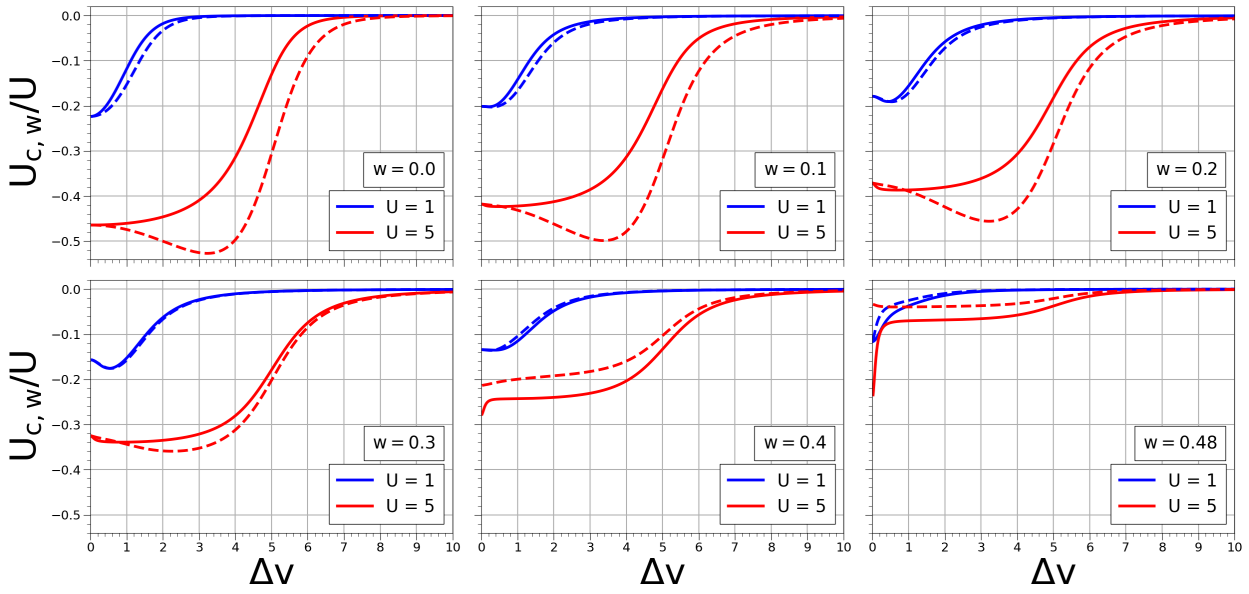

Figure A.19: Potential correlation energy of the Hubbard dimer bi-ensemble plotted as a function of Δv for various w values. Dashed lines are the HF approximation and the solid lines are exact.

Appendix B

Supplemental Material for Chapter 5

This chapter is a reproduction of the supplemental material for Ref. [162].

B.1 Abstract

Below we provide supplemental figures accompanying the main text, including various plots depicting the accuracy of the locally-thermal Perdew–Burke–Ernzerhof (ltPBE) ansatz relative to conditional-probability density functional theory (CP-DFT) data, and its satisfaction of two exact conditions. When useful, we make comparisons to other thermal generalized gradient approximations in the literature, namely the one proposed by Karasiev *et al.* (referred to here as KDT16).

B.2 CP-DFT Results

B.2.1 XC Enhancement Factors

All CP-DFT XC enhancement factor data has been included as .csv files for systems of various densities, gradients, and temperatures. We include files for unpolarized systems: $r_\mathrm{s} = 0.1, 0.5, 1, 2, 4, 10$ for $0 \leq s \leq 5$ and $0.05 \leq t \leq 4$; and for fully polarized systems: $r_\mathrm{s} = 0.1, 0.5, 1, 2, 4$ for $0 \leq s \leq 3$ and $0.05 \leq t \leq 4$.

The ltPBE approximation is depicted alongside CP-DFT data in the following figures, showing its performance for each density at select reduced temperatures. Fig. B.1 depicts the results for unpolarized data, while Fig. B.2 depicts the results for fully polarized data.

B.2.2 XC Hole Densities

Below we depict various plots illustrating the temperature-dependent XC hole densities output from the CP-DFT procedure outlined in the main text. Fig. B.3 depicts the results for unpolarized systems, while Fig. B.4 depicts the results for fully polarized systems. In all cases, the CP-DFT procedure produces an XC hole density with the same on-top value and energy as PBE as $t \to 0$.

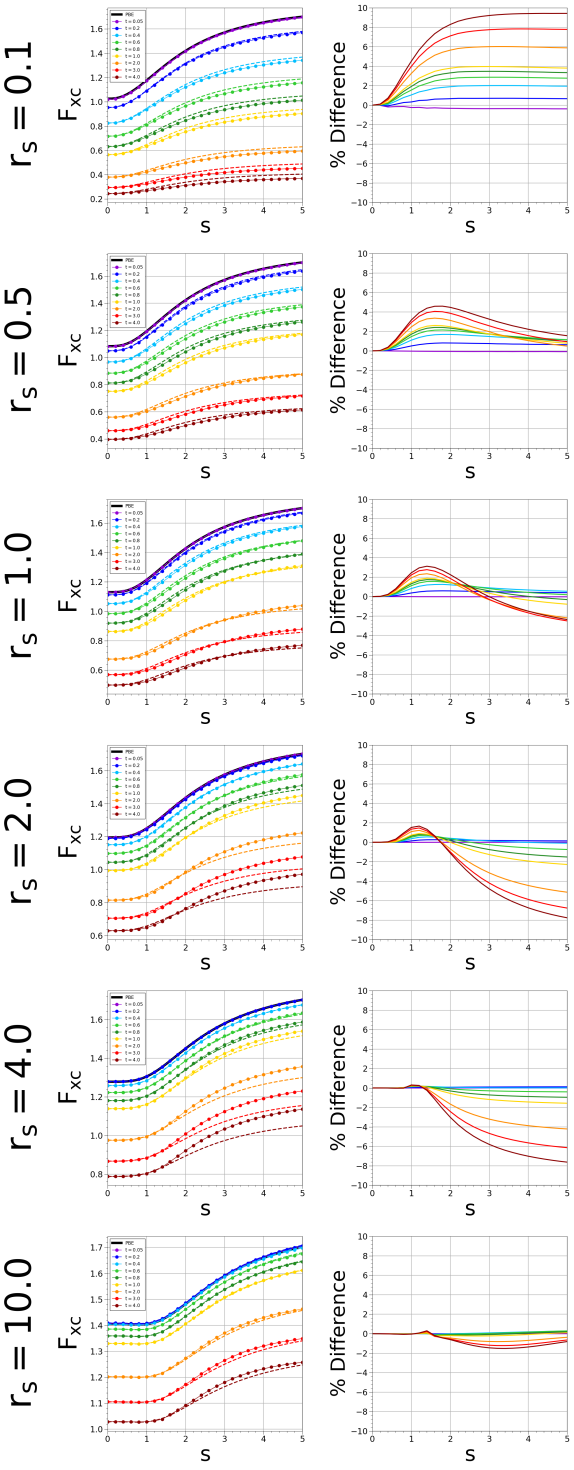

Figure B.1: Grid of plots depicting temperature-dependent XC enhancement factors as functions of the dimensionless gradient s for unpolarized systems of various r_S values. The data points represent CP-DFT calculations connected by solid lines, while the ltPBE approximation (Eq. 1) is represented by the dashed curves. The corresponding percent deviation of the ltPBE ansatz is included, largely found to be $\leq 8\%$ (but often much less) in the tested temperature range.

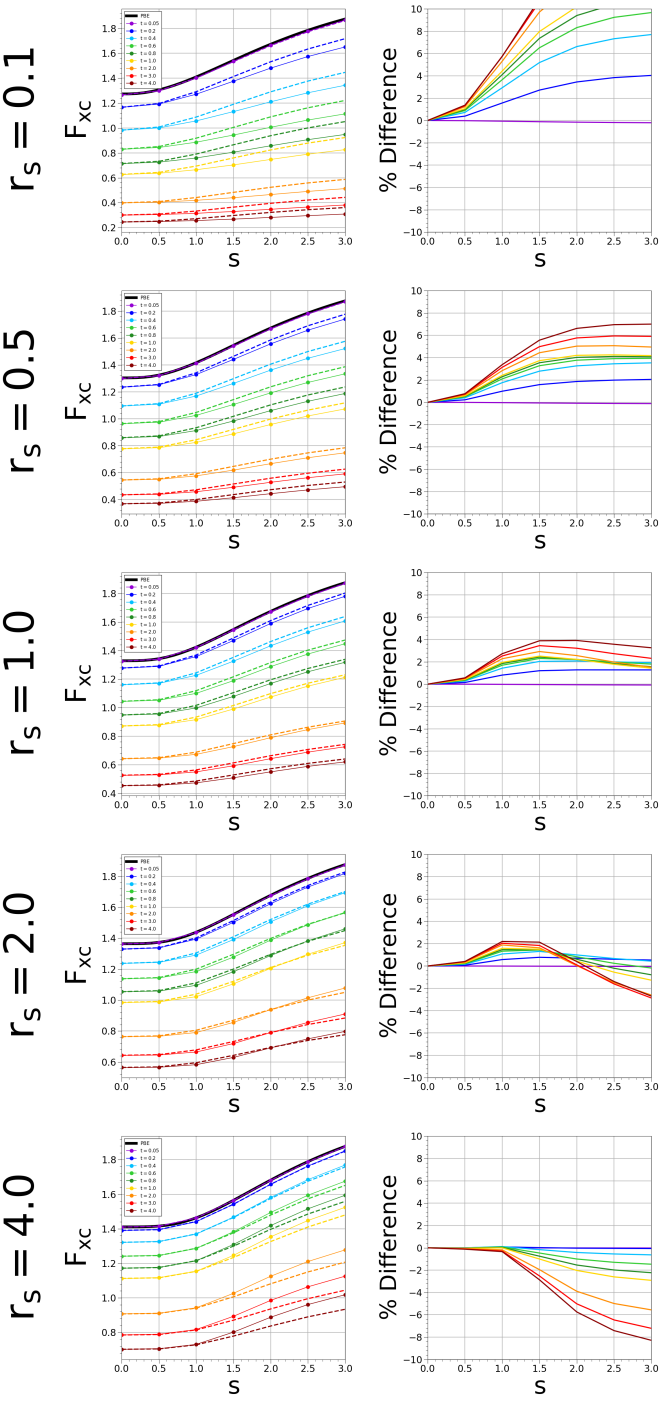

Figure B.2: Grid of plots depicting temperature-dependent XC enhancement factors as functions of the dimensionless gradient s for fully polarized systems of various r_s values. The data points represent CP-DFT calculations connected by solid lines, while the ltPBE approximation (Eq. 1) is represented by the dashed curves. The corresponding percent deviation of the ltPBE ansatz is included, largely found to be $\leq 8\%$ (but often much less) in the tested temperature range.

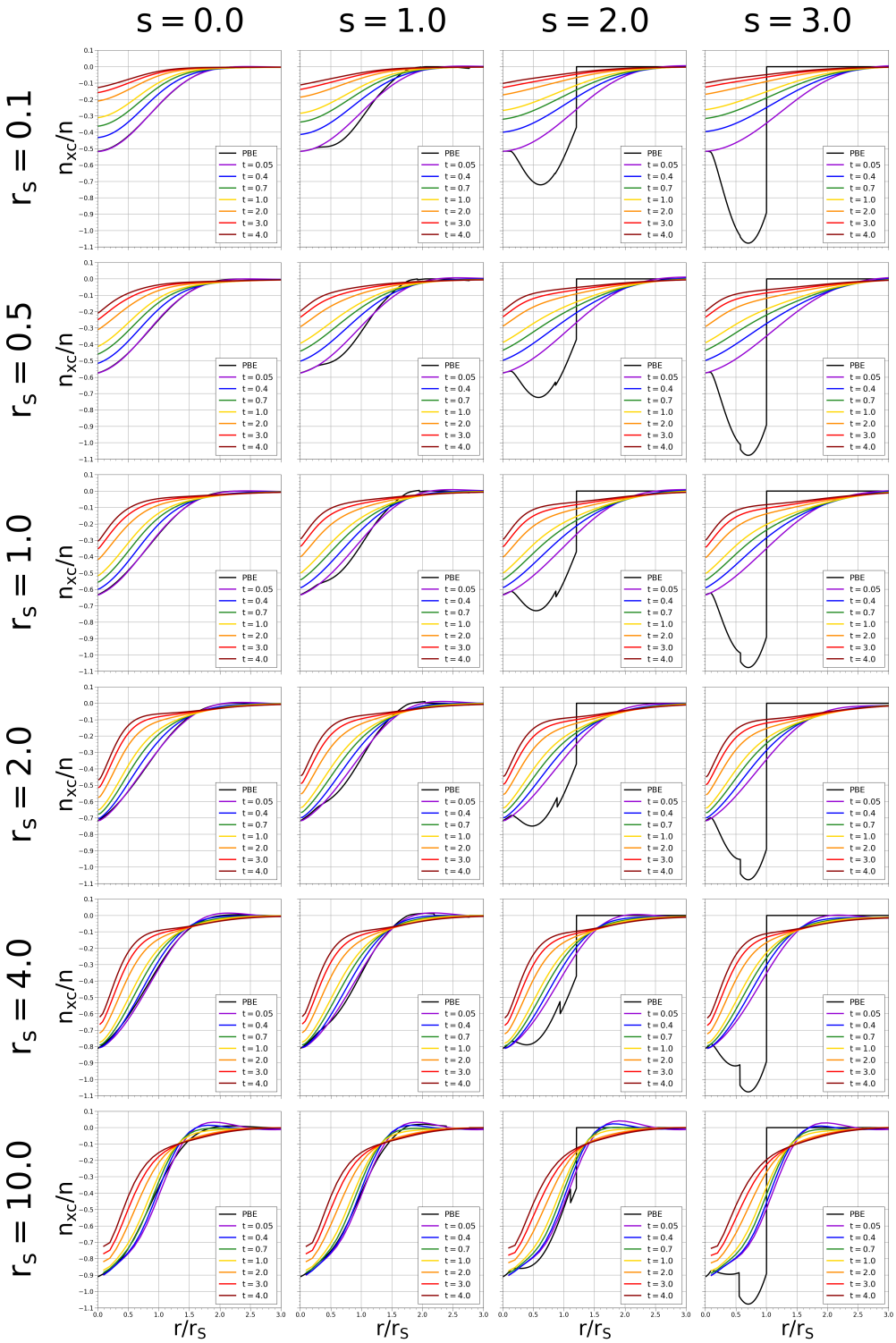

Figure B.3: Grid of plots depicting temperature-dependent XC hole densities for unpolarized systems of various r_s, s, and t. We use black to denote the ground-state PBE XC hole density, while the CP-DFT XC hole at different reduced temperatures is depicted in various colors.

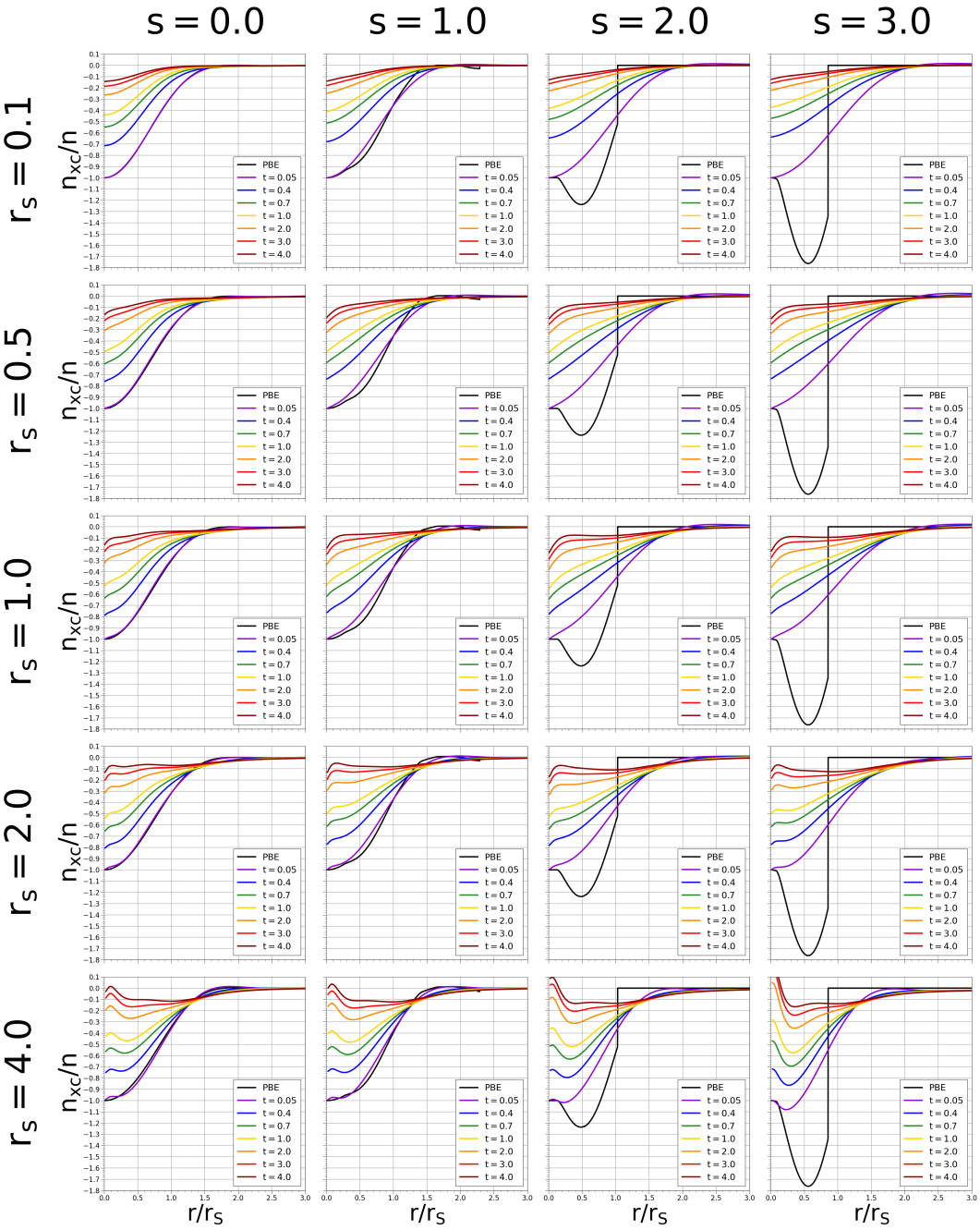

Figure B.4: Grid of plots depicting temperature-dependent XC hole densities for fully polarized systems of various r_s, s, and t. We use black to denote the ground-state PBE XC hole density, while the CP-DFT XC hole at different reduced temperatures is depicted in various colors.

B.3 Exact Conditions at Nonzero Temperature

B.3.1 Coordinate Scaling Inequality

Here we provide additional plots depicting the satisfaction of the thermal coordinate scaling inequality (Eq. 8 of the main text) by ltPBE, and violation by KDT16. In Fig. B.5 we plot F_{XC} curves at fixed values of t for ltPBE and KDT16, illustrating the differences between the two approximations. The ltPBE enhancement factor, having inherited the curvature of the ground-state $F_{\text{XC}}^{\text{PBE}}$, naturally satisfies this exact condition at all temperatures. In contrast the KDT16 approximation only satisfies this exact condition when $t \to 0$, where it recreates PBE, but violates it as the temperature is increased into the warm dense matter regime.

B.3.2 Concavity Condition

Here we show various contour plots depicting the thermal concavity condition (Eq. 11 of the main text) for ltPBE and KDT16 at different s values. Looking at Fig. B.6, it is clear that the ltPBE approximation satisfies this condition, regardless of the value of s, due to its expression always being negative. It is also clear that KDT16 satisfies the concavity condition as $s \to 0$ (where it reduces to thermal LDA), but violates the condition as s is increased. Noting that the density of a system scales according to $n_\gamma(\mathbf{r}) = \gamma^3 n(\gamma\mathbf{r})$, we observe violations for KDT16 for both moderate/high densities with nonzero s in the warm dense matter regime.

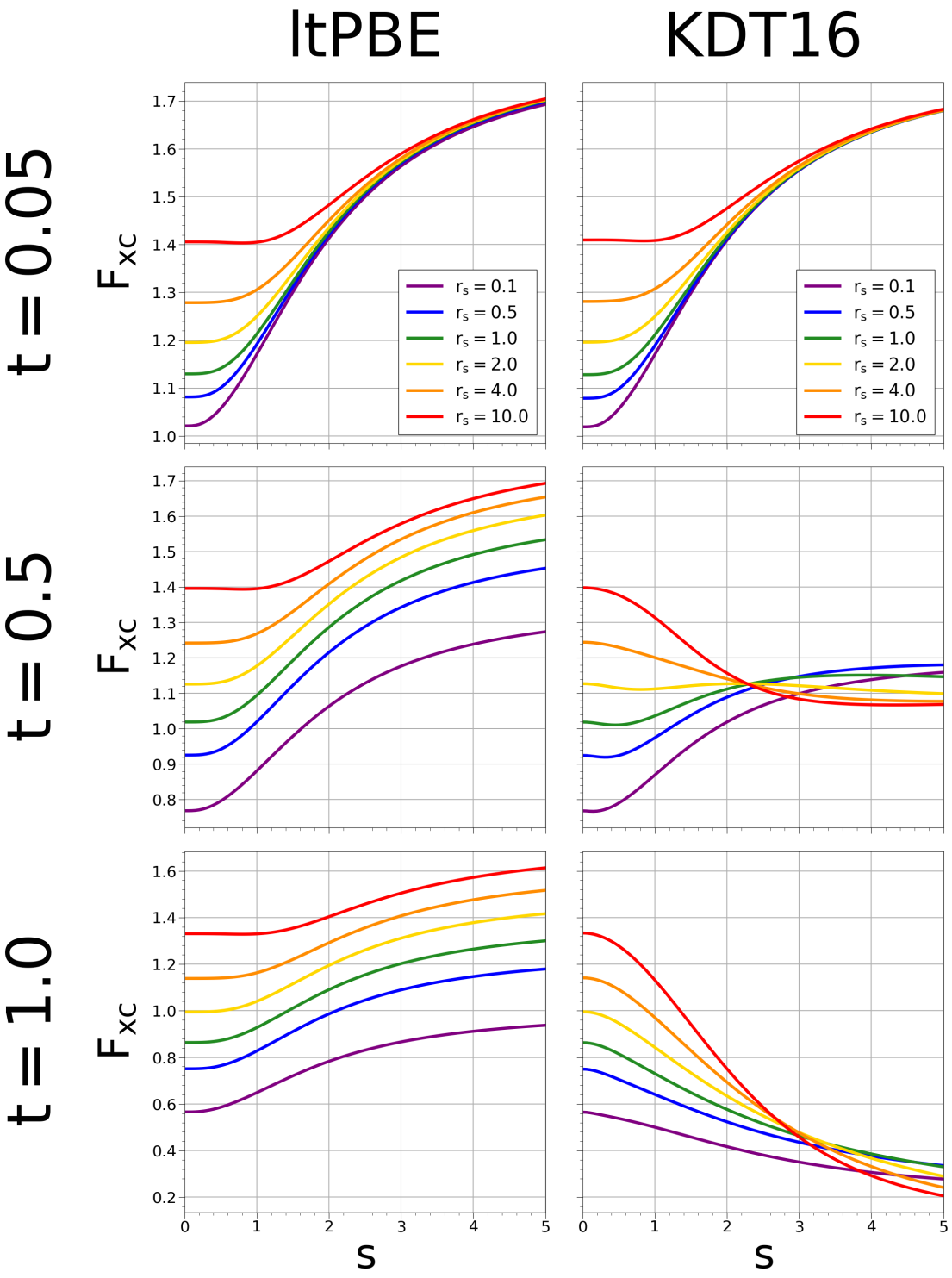

Figure B.5: Temperature-dependent XC enhancement factors plotted as functions of the dimensionless gradient s for unpolarized systems of various r_s values, with each row having a fixed reduced temperature t. The left column depicts the ltPBE approximation, while the right column depicts KDT16. The thermal coordinate scaling inequality mandates these curves not cross for any t.

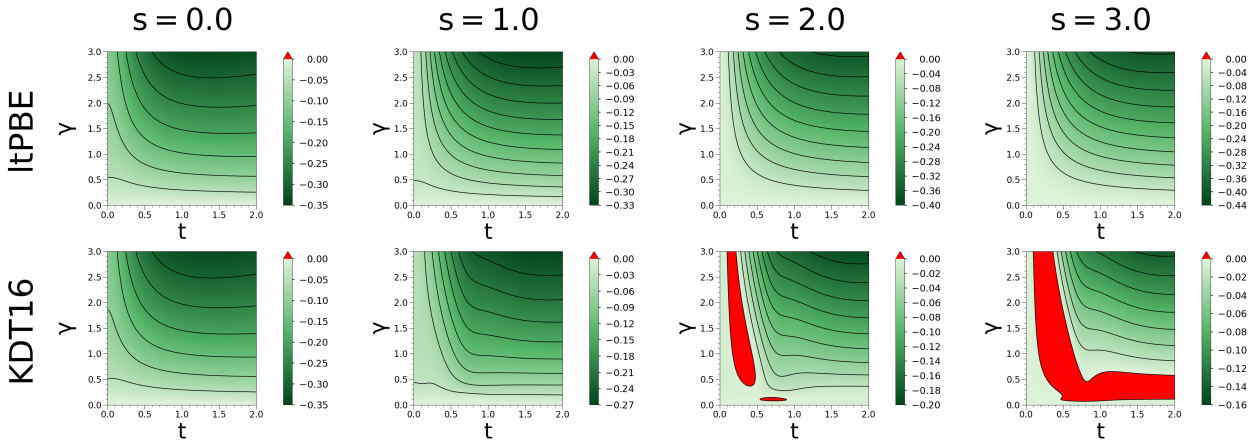

Figure B.6: Contours of the thermal concavity condition inequality (Eq. 11) with initial density $n = 1$. Here we show both the ltPBE (top) and KDT16 (bottom) approximations, confirming that the condition is satisfied by ltPBE, but is violated by KDT16 (red denotes positive).

Appendix C

Supplemental Material for Chapter 6

This chapter is a reproduction of the supplemental material for Ref. [236].

C.1 orthogonalization details

The original set of N basis functions $\chi_i(\mathbf{r})$, corresponding in our current FHI-aims implementation to numeric atom-centered orbitals (NAOs), are not orthogonal, in the sense that the overlap matrix S, with coefficients

$$S_{ij} = \langle \chi_i | \chi_j \rangle = \int d^3r \; \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) \tag{C.1}$$

is not the identity matrix, but some non-trivial, positive definite matrix. To identify the linear combinations of orbitals that are occupied in the ground state of the system, we need to account for overlaps between the orbitals. This can be done using the Löwdin

decomposition, where the Hamiltonian matrix H, with coefficients

$$H_{ij} = \langle \chi_i | \mathcal{H} | \chi_j \rangle = \int d^3r \; \chi_i(\mathbf{r}) \mathcal{H}(\mathbf{r}) \chi_j(\mathbf{r}), \tag{C.2}$$

is transformed into an orthonormal basis as $H \mapsto \tilde{H} = S^{-\frac{1}{2}} H S^{-\frac{1}{2}}$. The transformed Hamiltonian $\tilde{H}$ can then be purified, as described in the next section, to yield the density matrix $\tilde{D}$, which is then transformed back into the original orbitals basis,

$$\tilde{D} \mapsto D = S^{-\frac{1}{2}} \tilde{D} S^{-\frac{1}{2}}. \tag{C.3}$$

For this purpose, the inverse square root $S^{-\frac{1}{2}}$ of the overlap matrix is needed. To compute it efficiently on TPUs, we need an algorithm for computing the matrix inverse square root for which the computational bottleneck reduces to repeated matrix multiplications. As discussed in [124], this can be achieved using Newton-Schulz iterations. The iteration

$$X_{[n+1]} = \frac{1}{2} X_{[n]} (3I - X_{[n]}^2), \quad X_{[0]} = A, \tag{C.4}$$

converges to the matrix sign function of A, which turns positive eigenvalues to $+1$ and negative eigenvalues to -1. Notice that two matrix multiplications are needed for each iteration. These matrix multiplications can be executed very quickly when distributed over a set of TPUs. The inverse square root can in turn be cast as a sign function as

$$\operatorname{sgn}\left(\begin{bmatrix} 0 & S \\ I & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & S^{\frac{1}{2}} \\ S^{-\frac{1}{2}} & 0 \end{bmatrix}. \tag{C.5}$$

Applying the iteration (C.4) to the block matrix in (C.5) results in the Denman-Beavers iteration for computing the inverse square root.

In single (double) precision, the above procedure typically converges in about 35-50 (65-90)

iterations, depending on how small the absolute value of the smallest (in absolute value) eigenvalue of matrix A is. In order to further accelerate this computation, we introduce the pre-conditioning polynomial iteration (which we described and justified in Sect. III.D of [181] in the related context of the matrix sign function for singular values),

$$X_{[n+1]} = aX_{[n]}(I - \frac{4}{27}a^2X_{[n]}^2), \quad X_{[0]} = A, \tag{C.6}$$

where $a = \frac{3}{2}\sqrt{3} - s_-$ for some choice of small $s_- > 0$. [Notice that for $a = 3/2$, that is $s_- = 3(\sqrt{3}-1)/2$ we recover (C.4).] This pre-conditioning polynomial accelerates the growth of small eigenvalues of A, until they become of size at least s_-. From then on, the regular Newton-Schulz iteration is used to bring all the positive eigenvalues to 1, with quadratic convergence. For $s_- = 0.1$, in single (double) precision we need 15-20 (35) iterations of the pre-conditioning polynomial and 10 (10) iterations of the regular polynomial, for a total of 25-30 (45) iterations.

The inverse of S is of course very sensitive to poor conditioning of S, which for the overlap matrix corresponds to (nearly) linearly dependent orbitals, a common occurrence when dealing with large molecules. The danger of instability and poor accuracy due to small eigenvalues of S is especially pressing if operating in low numerical precision. Consequently we always compute $S^{-\frac{1}{2}}$ in double precision. Because TPUs do not operate natively in double precision but rather rely on software emulation, this incurs a significant time cost. However, in a full DFT simulation that cost gets amortized: the overlap matrix does not change between DFT iterations (only the Hamiltonian does), and thus we only need to compute $S^{-\frac{1}{2}}$ once at the first iteration, write the result to disk, and reread and use it at all the successive iterations with negligible cost.

C.2 purification details

By *density matrix purification* we mean the map from a Hermitian matrix $\tilde{H}$ of linear size N to a certain density matrix $\tilde{D}$, itself a projector into the subspace corresponding to the N_e smallest (or most negative) eigenvalues of $\tilde{H}$, where N_e is the number of electrons in the system. [As in the rest of the paper, the tilde in the Hamiltonian $\tilde{H}$ and density matrix $\tilde{D}^{(k)}$ denotes that these matrices are expressed in an orthonormalized basis of orbitals, as described in the previous section.] In symbols, let

$$\tilde{H} = V\Sigma V^H \tag{C.7}$$

be the ascendingly sorted eigendecomposition of $\tilde{H}$, and let $\rho \equiv \mathrm{diag}(1_1, 1_2, \ldots, 1_{N_e}, 0_{N_e+1}, 0_{N_e+2}, \ldots, 0_N)$ be a diagonal matrix with N_e 1's followed by $N - N_e$ 0's on the main diagonal. Then $\tilde{D}$ is defined by

$$\tilde{D} \equiv V\rho V^H. \tag{C.8}$$

It follows identically that $\tilde{D}^2 = \tilde{D}$, so that $\tilde{D}$ is indeed a projector.

With the decomposition (C.7) in hand, $\tilde{D}$ is trivially computed by the manual substitution $\Sigma \to \rho$. Unfortunately an efficient algorithm for Hermitian eigendecomposition is not presently available in a distributed-TPU context. Instead, we turn to matrix-multiplication based purification algorithms originally developed in the context of linear scaling methods (see [148] for a review; note that since our matrices are dense and are not truncated, our implementations scale as N^3 despite the name).

Density matrix purification algorithms can be divided into two classes [228] by the manner in which N_e is specified. In *grand canonical purification*, a so-called *chemical potential* μ is given, and $\tilde{D}$ found so that the $N_e + 1$'th most negative entry of Σ is the first to exceed μ.

This can be achieved by shifting the spectrum of $\tilde{H}$ by μ so that the latter divides negative from positive eigenvalues, and then computing a polar decomposition using the methods of [181].

In the *canonical purification* used in this work, N_e is instead specified directly. Compared to the grand canonical case this is more directly relevant to computations of molecular electronic structure, where μ is unknown but the number N_e of electrons is provided.

Various algorithms for canonical purification have been proposed in the literature. The original scheme is presented in [228], and is variously referred to as *canonical purification* (in which case other algorithms are given a different name), *trace-preserving purification*, or the *Palser and Manolopoulos scheme*. The *trace-resetting* schemes proposed in [218, 219] are probably most common in practical use. We use the *generalized* or *hole-particle* scheme presented in [301]. In our TPU experiments this iteration yields performance comparable to that of [218, 219], but avoids certain branching conditionals which are awkward to phrase efficiently on the TPU.

All such schemes work by first mapping the input $\tilde{H}$ to some initial $X_{[0]}$ with eigenvectors unchanged but eigenvalues bound in $[0, 1]$, and then repeatedly applying a matrix-multiplication based iteration which also preserves eigenvectors. This iteration is chosen so that the eigenvalues of its fixed point $X_{[\infty]}$ are *exactly* either 0 or 1 with $\mathrm{Tr}(X_{[\infty]}) = N_e$; $X_{[\infty]}$ then satisfies (C.8) and it is thus equal to $\tilde{D}$, up to numerical error. In practice, the number of purification iterations required for convergence varies across different Hamiltonians and the numerical precision desired. Relevant factors include the size of the energy gap and the fraction of occupied to unoccupied states. Typically, less than 50 purification iterations are required [301]. Calculations performed to double precision tend to require more purification iterations, up to twice as many as the same calculation performed to single precision.

Details of the specific iteration we use are given in [301]. It can be reproduced by the

initialization

$$X_{[0]} = \beta_1 I + \beta_2(\mu I - \tilde{H}), \tag{C.9a}$$

$$\mu = \frac{\mathrm{Tr}\tilde{H}}{N}, \tag{C.9b}$$

$$\beta_1 = \frac{k/N}{e_+ - \mu}, \tag{C.9c}$$

$$\beta_2 = \frac{1 - k/N}{\mu - e_-}, \tag{C.9d}$$

where e_+ and e_- are estimates of the largest and smallest eigenvalues of $\tilde{H}$ obtained by e.g. the Gershgorin circle theorem. Note that in practice we use the slightly more complicated initialization referred to as *HPCP+* in [301], which gives moderately improved performance when N_e is far from $N/2$. In either case, the iterate $X_{[n+1]}$ is found from its predecessor $X_{[n]}$ via

$$X'_{[n]} = I - X_{[n]}, \tag{C.10a}$$

$$X_{[n+1]} = X_{[n]} + 2\left(X^2_{[n]}X'_{[n]} - \frac{\mathrm{Tr}(X^2_{[n]}X'_{[n]})}{\mathrm{Tr}(X_{[n]}X'_{[n]})}X_{[n]}X'_{[n]}\right). \tag{C.10b}$$

Once $\tilde{D}$ is found, its counterpart in the non-orthogonal basis, D, is found by applying (C.3).

In Fig. C.1 we demonstrate the computational scaling of density matrix purification on TPUs using dense random Hermitian matrices and single precision. In this benchmark we scale both the system size (dimension of the matrix, N) and the number of TPU v3 cores used. Starting with a single TPU board, consisting of 8 TPU v3 cores, we can systematically scale up to hundreds (or thousands) of TPU v3 cores. Using a full TPU v3 pod (consisting of 2048 TPU v3 cores), we project that we can address dense systems of $N = 500\,000$ orbitals within 30 minutes using single precision.

For dense linear algebra, the computational scaling here is cubic. If suitable sparsity is assumed, and the density matrix is correspondingly truncated, sparse linear algebra can be

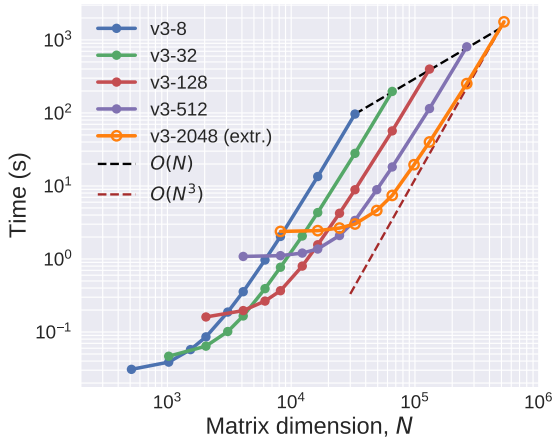

Figure C.1: Average wall times for TPU density matrix purification (normalized to a total of 50 purification iterations) in single precision using dense random Hermitian matrices of dimension N. Open-circle data points (v3-2048 results) are a linear extrapolation from v3-512 results.

used to obtain linear scaling. Such linear scaling approaches have been implemented and used in practice within computational quantum chemistry packages, however their practical application is limited to systems whose density matrix has a sufficient sparsity structure to ensure accurate results.

C.3 FHI-aims TPU integration details

We outline the practical details of integrating a TPU-based density matrix solver with the CPU-based DFT package FHI-aims. The platform and integration described here is a prototype. Its main purpose is to illustrate, in actual end-to-end DFT computations, the viability of accelerating the $O(N^3)$ bottleneck using TPUs.

The software ELSI [339, 344] is used to facilitate the connection between FHI-aims and the TPU by providing an interface and abstraction in which FHI-aims, or other codes, such as Siesta [76] and DFTB+ [129], can utilize external eigensolvers launched within ELSI. We implement in-house routines to launch the TPU-based density matrix purification (instead of

an eigensolver) using the ELSI standard. In all calculations, we use an off-the-shelf FHI-aims code with *no* modifications (version 210226).

In the course of a DFT calculation, FHI-aims utilizes the following matrices: the overlap matrix S, the DFT Hamiltonian H, and the density matrix D. FHI-aims distributes each matrix across CPU processes and memory using a 2D block-cyclic distribution pattern. On the other hand, in our current TPU implementation which utilizes SUMMA (Scalable Universal Matrix Multiplication Algorithm), our TPU-based solver requires matrices to be distributed across a TPU processor grid as 2D blocks in a *checkerboard* distribution, see [181] for more details. This poses a practical matrix communication challenge between the CPU-based and TPU-based schemes since their matrix distribution patterns differ in both cyclicity and the number of processors. A simple solution to communicate such matrices between CPU and TPU is to serialize and transfer the respective matrices and deserialize and redistribute them in the desired scheme. Specifically, we utilize available MPI processes on the CPU to serialize (and deserialize) to a network disk using the ELSI IO module and compressed sparse column (CSC) format with no cyclicity. Double precision is used throughout. We note that each process (CPU and TPU) calculates where its data should be within the serialized CSC representation of the whole matrix and reads (writes) to (from) only that portion of the matrix representation on the centralized network drive. That is, our implementation incurs some algorithmic overhead but only communicates the data that is needed.

This is not the most performant solution, however, it is generalizable, makes use of existing tooling within ELSI, and avoids the complexity of the various distribution patterns. Due to the use of the CSC format, serializing (writing) dense matrices to disk is especially costly and dominates the total CPU-TPU communication time for large system sizes. For transparency, in Fig. C.2 we plot the average observed total end-to-end CPU-TPU communication time (excluding the TPU density matrix purification time) incurred in our current implementation. There are, however, several ways to further optimize the current integra-

tion. For instance, we are currently using an off-the-shelf network file system (NFS) share, and replacing it with a different implementation of a portable operating system interface (POSIX) compliant distributed file system (one designed for high-performance applications) would result in a much higher throughput and would not require any changes to our code. In addition, further algorithmic optimizations are likely possible.

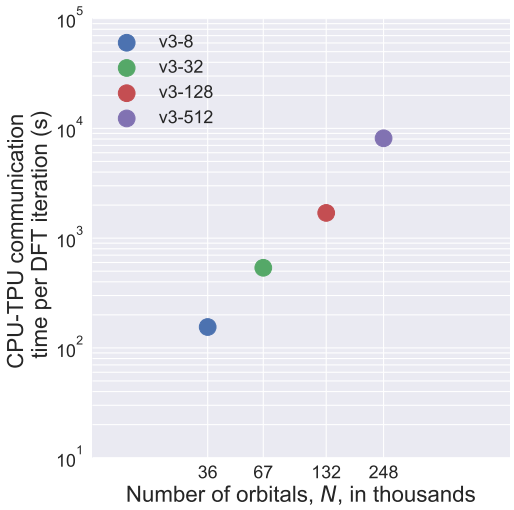

Figure C.2: Total end-to-end CPU-TPU communication time per DFT iteration (excluding the TPU density matrix purification time) incurred in current integration with FHI-aims.

Within FHI-aims all calculations are performed using the all-electron "light defaults" numeric atom-centered basis set [17]. This results in 5 basis functions per H atom and 14 basis functions per O atom in the water cluster calculations. All calculations are non-periodic with open boundary conditions and utilize the PBE [241] XC functional. The geometries of water clusters are directly obtained from Ref. [284] which were generated by taking spherical cutouts of varying radii from a large molecular dynamics simulation of bulk water at standard pressure and an average temperature of 300K (further details can be found in Ref. [262]).

In our implementation with FHI-aims, hybrid functionals can also be used without any modification, but simply result in longer DFT Hamiltonian build times on CPUs. Analytical forces are also available from FHI-aims using TPU-computed energy-weighted density

matrices, which are also communicated using the above scheme and facilitated using ELSI.

C.4 dynamic precision on smaller water clusters

The dynamic precision approach illustrated in Fig. 4 of the main text for 10 327 water molecules, when applied to smaller systems, allows for a larger part of the computation to be performed in single precision. For instance, an end-to-end converged DFT calculation on $(H_2O)_{1481}$ cluster with $N = 35\,544$ orbitals required 11 iterations in single precision and 4 iterations in double precision, with an overall time of under 5 hours on a single TPU (v3-8) board, see Fig. C.3.

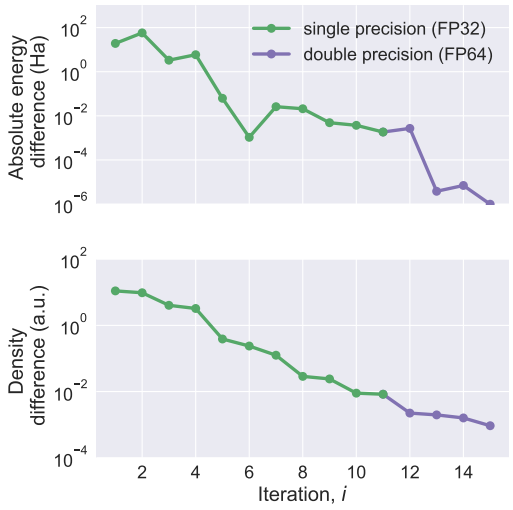

Figure C.3: Convergence trajectory of an end-to-end dynamic precision DFT calculation on a $(H_2O)_{1481}$ cluster with $N = 35\,544$ orbitals. The absolute total energy differences between subsequent DFT iterations, i and $i-1$, are plotted (top). The corresponding difference in real-space densities within the L^1 norm is plotted (bottom).