Real-space descriptors : a window into the properties of hydrogen-based superconductors

Trinidad Antonia Novoa Aguirre

Trinidad Antonia Novoa Aguirre

Real-space descriptors: A window into the properties of hydrogen-based superconductors

A thesis submitted to Sorbonne Université
to obtain the degree of

Doctor of Philosophy in
Physical Chemistry

January 2024

Dr. Julia Contreras-García	–	Thesis advisor
Prof. Yvon Maday	–	Thesis co-advisor
Prof. Eva Zurek	–	Thesis evaluator
Prof. Lilia Boeri	–	Thesis evaluator
Prof. Carlos Cárdenas	–	President of the jury
Dr. Gilles Frapper	–	Member of the jury
Dr. Ion Errea	–	Member of the jury

Abstract

In the last years, hydrogen-rich materials have proven their tremendous potential as high-temperature superconductors, showing critical temperatures over 200 K for various systems. Although the microscopic theory describing the properties of these conventional superconductors is well-known, computing their critical temperature is a computationally expensive task, limiting its use in the prediction of new promising compounds. Moreover, the chemical features that characterize high-temperature hydrogen-based superconductors are not yet fully understood, further impeding the rational design of those materials. In this thesis, we analyze the real-space properties of hydrogen-based systems with the aim of finding the fingerprints of high-temperature superconductivity. This is mainly done through the study of the electron localization function (ELF), both in a one-dimensional model of a hydrogen-chain and in real hydrides. The networking value of the hydrides, a topological descriptor of the ELF in the crystal lattice, shows to correlate with the critical temperature, T_c. An explanation for this is found in the model, where a superconducting ELF is defined, showing a similar topology to that of the ELF in the normal state. This justifies the use of the normal state properties to infer those of the superconducting state. We make practical use of the observed correlation and present TcESTIME, a program for the fast estimation of the networking value, ϕ, and thus of T_c. That correlation can be improved by considering other descriptors. We introduce the molecularity index, ϕ^*, that represents the degree of molecularity of the hydrogens in the system, and allows to filter the most promising hydrides, where H_2 molecular units are not common. This is further supported by the study of dimerization in the one-dimensional chain, where it is observed that molecularity hinders superconductivity as a result of the tendency of the normal state to become an insulator. In this way, the combination of ϕ and ϕ^* give us a good overall description of the superconducting hydrides, and permit to anticipate some of its properties. Finally, spatial fluctuations of the potential as a result of the lattice vibrations are studied. Preliminary results indicate some trends among the magnitude of the fluctuations around the hydrogen positions, and a thorough topological analysis of the spatial functions shows promise as a means of understanding the effect of chemical precompression in those systems.

Résumé

Ces dernières années, les matériaux riches en hydrogène ont démontré leur potentiel en tant que supraconducteurs à haute température, affichant des températures critiques supérieures à 200 K pour divers systèmes. Bien que la théorie microscopique décrivant les propriétés de ces supraconducteurs conventionnels soit bien connue, le calcul de leur température critique est une tâche informatiquement coûteuse, ce qui limite son utilisation dans la prédiction de nouveaux composés prometteurs. De plus, les traits chimiques qui caractérisent les supraconducteurs à haute température à base d'hydrogène ne sont pas encore totalement comprises, ce qui entrave encore plus la conception rationnelle de ces matériaux. Dans cette thèse, nous analysons les propriétés dans l'espace réel des systèmes à base d'hydrogène dans le but de trouver les empreintes de la supraconductivité à haute température. Cela se fait principalement par l'étude de la fonction de localisation électronique (ELF), à la fois dans un modèle unidimensionnel d'une chaîne d'hydrogène et dans des hydrures réels. La valeur de réseau des hydrures, un descripteur topologique de l'ELF dans le réseau cristallin, montre une corrélation avec la température critique, T_c. Une explication est trouvée dans le modèle, où une ELF supraconductrice est définie, montrant une topologie similaire à celle de l'ELF dans l'état normal. Cela justifie l'utilisation des propriétés de l'état normal pour déduire celles de l'état supraconducteur. On fait un usage pratique de la corrélation observée et présente TcESTIME, un programme pour l'estimation rapide de la valeur du réseau, ϕ, et donc de T_c. Cette corrélation peut être améliorée en prenant en compte d'autres descripteurs. On introduit l'indice de molécularité, ϕ^*, qui représente le degré de molécularité des hydrogènes dans le système et permet de filtrer les hydrures les plus prometteurs, où les unités moléculaires H_2 ne sont pas courantes. L'étude de la dimérisation dans la chaîne unidimensionnelle confirme cette hypothèse, en montrant que la molécularité entrave la supraconductivité en raison de la tendance de l'état normal à devenir un isolant. Ainsi, la combinaison de ϕ et ϕ^* nous donne une bonne description globale des hydrures supraconducteurs, et permet d'anticiper certaines de leurs propriétés. Enfin, les fluctuations spatiales du potentiel résultant des vibrations du réseau sont étudiées. Les résultats préliminaires indiquent quelques tendances dans l'ampleur des fluctuations autour des positions de l'hydrogène, et une analyse topologique approfondie des fonctions spatiales semble prometteuse pour comprendre l'effet de la précompression chimique dans ces systèmes.

To my family, who are always there despite the kilometers.

Acknowledgments

I would like to begin by thanking the members of the jury, for taking their time to read this manuscript, and for their dedication in the evaluation of this thesis.

To my supervisors, Julia Contreras-García and Yvon Maday, thank you for your guidance and support during the three years of my thesis. It has been a life-changing experience and I greatly appreciate that you gave me the opportunity to come to Paris and work with you and your team. Thank you, also, for supporting me and encouraging me to work with other teams, attend seminars and conferences, and expand my scientific network.

I would like to further thank Julia for everything that she has done for me during the past three years. Thank you for your trust when you first invited me to come here, and throughout the PhD. It was that trust that many times gave me the confidence I needed to face the challenges that seemed scary at first but that surely made me grow as a scientist. I am grateful for everything that I have learned from you scientifically, and I can only hope to, one day, inherit a small part of that gift of yours to turn a scientific work upside down, transforming it into a much better version of itself, and highlighting what nobody else could previously see. You have also taught me so much about this world of science of which I too am now an inhabitant, that I continue to learn how to navigate. But, most importantly, I would like to thank you for being so human, for your support in those times when I felt overwhelmed with the thesis, and for the authentic love that you put into taking care of each one of your students. It is this lesson that I want to carry with me for the rest of my career, and my life.

During the PhD, I had the chance to spend around four months in San Sebastian, in the Basque Country, doing a collaboration with Ion Errea and his group at the CFM. I would like to thank Ion for this opportunity, which was very gratifying both professionally and personally. I felt at home within the group, and I learned so much from the scientific collaboration. It was a pleasure to work alongside Francesco Belli, who taught me so much and became a dear friend. I would also like to thank the other people of the office at the CFM for their warm welcome and for creating an amazing work environment; thank you Giulia, Antonella, and Stefano. And thanks to the many friends that I made at the CFM and in Donosti. I know I will miss mentioning many of them, but I would like to specially thank my dear girls Ebtisam, Paula, and Alaa; you are the best and I appreciate your friendship so much, you made my stay at San Sebastian an unforgettable one.

Continuing with the international note, I would like to extend a special thanks to the Theoretical Chemical Physics group back in Chile that, to my pleasure, keeps considering me one of them and welcomes me with open arms whenever I come back. Special thanks to Carlos Cárdenas, who guided me in the first steps of my scientific life, without whom I would not be here in Paris today.

Back in Paris, I would like to thank the LCT crew, for making every day at the lab a fun one. Some people complain about their daily routine, but I have to say that mine has been quite fulfilling these last years, thanks to you. The lunch breaks at tower 33 saved me from the monotony of my first months here in Paris, when the curfew was at 18h and the halls of LCT were quite deserted. Later, when life came back to normal, the monthly student seminars (and specially the beers after that) became a great opportunity for us to know each other better and to share great moments. I would like to thank everyone that has been part of that, with a special mention to Chiara, Timothée, Federica, Théo, Diata, Pier, Marina, and Adam; you have become great friends and I hope that this remains so in the future.

I could not finish the LCT acknowledgements without a very special word of appreciation for everyone who has been part of our aquarium. This office that has been my home for three years has been continuously populated by the funniest Spanish speakers, who have lighten up my daily life (which, as an informed reader may already know, is specially important in this windowless office). Thank you Bruno, Victoria, Matías, Esaú and Francisca, *son los mejores!* A little piece of our chaotic Latin America has been here with me thanks to you, and it has been priceless. And to my dear friend, Andrea, the other long-term inhabitant of this work place, thank you so much. You have been a great companion and an amazing friend. I have enjoyed our endless laughter, I have appreciated your advice, and I have had a great time working together, too. I am certain that we will continue to find each other in this scientific life of ours.

To Pame, Cris, Nati and Erik, my chilean family here in Paris, who have been my home away from home. None of this would have been the same without you, and I am so grateful for your friendship and your support. I have had the best of times in your company, and I am happy to have shared these *good old days* with you.

To my friends in Chile, who I keep in my heart wherever I go. To the girls from Viña, thank you for always keeping in touch and considering me in your plans even when you know I cannot come. And to my friends from Santiago: seeing you once a year has been a great way of recharging my batteries, of staying connected to my life in Chile. The same goes to my large family, cousins, uncles, and aunts; thank you for being the home that I can always come back to.

And last but definitely not least, thank you to my closest family, for their infinite support in this transatlantic adventure of mine. Although the distance has been hard sometimes, you make it easier with your loving ways, with your constant presence in my life. To my parents and step parents, you are the most important pillar of my life, and I would not be anywhere near here if it was not because of you. Thank you for always believing in me and encouraging me to pursue my wildest dreams. A most special thanks to my beloved siblings, for making me laugh at every occasion and for letting me be your accomplice even at a distance, you do not know how much that has meant to me during this journey.

Scientific production

Thesis output

Publications:

- Trinidad Novoa, Matías E. di Mauro, Benoît Braïda, Ion Errea, Julia Contreras-García. Molecularity: a quick probe for superconductivity (submitted).
- Trinidad Novoa, Matías E. di Mauro, Kaoutar El Haloui, Nicolas Sisourat, Yvon Maday, Julia Contreras-García. TcESTIME: predicting high-temperature superconductors (to be submitted).
- Wilver A. Muriel, Trinidad Novoa, Carlos Cárdenas, Julia Contreras-García. Introducing electron correlation in solid state calculations for superconducting states (submitted to Faraday Discussions Journal).
- Francesco Belli, Trinidad Novoa, Julia Contreras-García, and Ion Errea. Strong correlation between electronic bonding network and critical temperature in hydrogen-based superconductors. Nature Communications, 12(5381), 2021.

Software:

- TcESTIME: https://github.com/juliacontrerasgarcia/Tcestime

Other output

Publications:

- Trinidad Novoa, Julia Contreras-García, and Patrick Chaquin. Conformational preference analysis in C2H6 using Orbital Forces and Non-Covalent Interactions; comparison with related systems. Phys. Chem. Chem. Phys., 25(4276), 2023.
- Trinidad Novoa, Rubén Laplaza, Francesca Peccati, Franck Fuster, and Julia Contreras-García. The NCIWEB Server: A Novel Implementation of the Non-covalent Interactions Index for Biomolecular Systems. Journal of Chemical Information and Modeling, 63(15): 4483-4489, 2023.

Book chapter:

- Trinidad Novoa, Francesca Peccati, Mercedes Alonso, David Arias Olivares, Hugo Bohorquez, and Julia Contreras-García. New developments in the Non-Covalent Interactions (NCI) index. In Paul Popelier, editor, Comprehensive Computational Chemistry. Elsevier. (to be published soon).

Software:

- NCIWEB server: https://nciweb.dsi.upmc.fr/index.php
- NCICLUSTER: https://github.com/Non-Covalent-Interactions/NCIcluster

Contents

List of Abbreviations

BCP Bond Critical Point
BCS Bardeen-Cooper-Schrieffer
BZ Brillouin Zone
CP Critical Point
CSP Crystal Structure Prediction
DFT Density Functional Theory
DI Delocalization Index
DOS Density of States
ELF Electron Localization Function
GGA Generalized Gradient Approximation
KED Kinetic Energy Density
KS Kohn-Sham
LI Localization Index
ML Machine Learning
NNA Non-Nuclear Attractor
NUC Nuclear attractor
NS Normal State
PAW Plane-Augmented-Wave
PBC Periodic Boundary Conditions
PBE Perdew–Burke-Ernzerhof
PP Pseudopotential
QCT Quantum Chemical Topology
QTAIM Quantum Theory of Atoms in Molecules
RDM Reduced Density Matrix
SC Superconducting
SCDFT Superconducting Density Functional Theory
SCF Self-Consistent Field
SR Symbolic Regression

SSCHA Stochastic Self-Consistent Harmonic Approximation
TF Thomas-Fermi
US Ultra-Soft
VB Valence Bond
WHA Wolfsberg-Helmholtz Approximation

1

Introduction

The phenomenon of superconductivity was first observed by Kamerlingh Onnes in 1911, when he noticed that the electrical resistivity of mercury at low temperatures abruptly dropped to zero, becoming a perfect conductor [1]. Later on, more materials showing these features were observed, and their characterization led scientists to notice another fundamental property of the superconductors: the magnetic field in their interior is always null, even when exposed to an external field. This phenomenon is called the *Meissner effect*, and it was first described in 1933 [2]. In other words, besides of being perfect conductors, superconductors are also perfect diamagnets [3, 4, 5]. These features make them great candidates for a plethora of technological applications. To name a few, these materials can be exploited for energy storage and transport [6, 7, 8], for strong magnetic levitation [9, 10], for applications on medical imaging [11, 12], and for their use in quantum computers [13, 14], among others.

Unfortunately, in the more than a hundred years since the discovery of superconductors, the scientific community has realized that actually exploiting those properties is quite a difficult task, as the superconducting phase usually appears under very low critical temperatures, T_c. In fact, the first superconductors that were discovered, consisting on metallic phases of single-element materials, have critical temperatures of only a few Kelvin. For example, at ambient pressure, bulk aluminum has a critical temperature of 1.18 K [15], and for mercury this quantity barely rises to 4.15 K [1, 16].

Many years were necessary for the scientific community to produce a microscopic theory that would properly account for the phenomenon of superconductivity. In 1957, Bardeen, Cooper and Schrieffer finally published a theory that successfully predicted the critical temperature of many superconductors [17], also describing the drop of resistance and the Meissner effect in those materials, earning them the Nobel prize in 1972. Moreover, the theory reproduces the so-called *isotope effect*, which consists of a dependence of the critical temperature on the mass of the atom conforming the material, M, when it is replaced by its isotope, that had experimentally demonstrated to scale as $T_c \propto M^{-1/2}$ [18]. Bardeen-Cooper-Schrieffer (BCS) theory is based on the assumption that electrons around the Fermi energy couple into what are now known as Cooper pairs. The formation of those bound states of electron

pairs stabilizes the system, and opens up a gap in the energy spectrum. Based on the dependence of T_c on the ionic mass, BCS theory assumes that the origin of the effective attraction between the electrons near the Fermi energy is phonon-mediated, but does not delve into a detailed description of that interaction.

Following this same assumption, the later advent of the Migdal-Eliashberg formalism meant an extension to the theory of BCS, where the electron-phonon interaction was treated explicitly [19, 20, 21]. With this, a wider range of superconducting materials had their critical temperatures properly predicted. These superconductors were categorized as conventional or phonon-mediated. It is this formalism, along with the modern implementation of computational tools that permit to properly describe the electron-phonon spectra from first principles, that allow us nowadays to theoretically characterize conventional superconductors.

In 1972, when those computational tools were not yet available, an upper limit for the critical temperature of conventional superconductors, of barely 25 K, was proposed by Cohen and Anderson [22]. Seven years later, the record holder was Nb_3Ge with a T_c of 23K, still below the so-called maximum T_c [23]. For a long time, no conventional superconductor was able to overcome that barrier, discouraging the research done on those materials and shifting the focus towards unconventional superconductors, i.e. those that could not be explained by BCS theory.

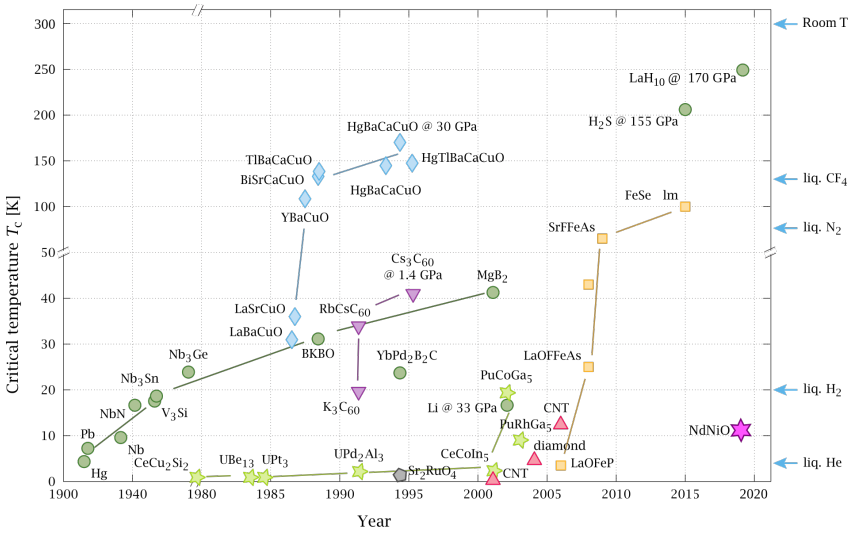


Fig. 1.1: Timeline of superconductors observed between 1900 and 2015, and their critical temperatures, separated by type: conventional superconductors are depicted in green circles, the cuprates are represented by blue diamonds, and the rest of the data correspond to other different types of unconventional superconductors. By PJRay - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=46193149.

A particular set of unconventional superconductors, the *cuprates*, had the record of highest T_c's for many years. Their name comes from the common pattern in the crystal structure of such materials, consisting on layers of CuO_2. For a long time they were also referred to as high-temperature superconductors, owing to the sudden jump that was made within a few years in the field, where critical temperatures above 100 K were obtained for a few cuprates, as it can be seen in the timeline of Figure 1.1. The record within the cuprates is of 133 K for HgBaCaCuO at ambient conditions [24], rising to 164 K at high pressures [25]. This is well above the boiling point of nitrogen, at 77 K, which makes them very convenient from an experimental viewpoint. However, the drawback of this set of superconductors is that we do not actually understand the mechanism that enables the superconducting transitions, thus impeding the possibility of prediction by theoretical means, and slowing down the evolution of the field.

In the early 2000's, the experimental observation of MgB_2 with a T_c of 39 K proved Cohen and Anderson's hypothesis wrong [26], and the interest of the scientific community on conventional superconductors was reborn. Although the superconducting phase presented itself at much lower temperatures than in cuprates, the simplicity of this compound compared to those high-temperature superconductors was undoubtedly enticing. The fact that T_c in conventional superconductors had been increased by a material composed by very light atoms did not come as a surprise, as it was known that the presence of light elements should enhance the electron-phonon coupling, and thus favor the transition. Indeed, as early as 1968, Ashcroft had proposed that metallic hydrogen, the lightest of elements, would become a superconductor at a very high temperature [27]. Testing this proved problematic, as hydrogen is found in gas phase at ambient conditions, and synthesizing it in a metallic phase requires to increase pressure enormously, to a predicted value of 400 GPa [28]. Such a pressure is even larger than that found at the center of our planet [29], and it is very difficult to achieve experimentally. Ashcroft himself postulated in 2004 that this difficulty could be overcome by alloying hydrogen with other elements, which would act as chemical precompressors, i.e. they would aid in the stabilization of the hydrogen-based crystal at lower pressures [30].

The scene changed abruptly in 2015 when Drozdov and coworkers successfully synthesized a superconducting sulfur hydride, likely to correspond to a stechiometry H_3S, with T_c = 203 K at high pressures [31], opening up a new era of hydrogen-rich superconductors. This was succeeded by the even higher temperature of 250 K for which superconductivity was observed in LaH_{10} just a few years later [32]. This trend of near room-temperature superconducting hydrides prevailed, and up to now it has been observed in ThH_{10} at 161 K [33], YH_9 at 243 K [34], and YH_6 at 224 K [35], among others; all of them at very high pressures of 175 GPa, 201 GPa, and 166 GPa, respectively. Observations of room-temperature in hydrides have also been claimed twice, and then retracted: first for a C-S-H compound [36], and then for a nitrogen-doped lutetium hydride [37].

A particularity of this era of high-temperature hydrogen-based superconductors research is that experimental breakthroughs have been accompanied by numerous theoretical predictions, that have helped in guiding those findings. This has been

possible thanks to the accurate predictions of T_c from first principles that are available today, accompanied by an ever-growing community performing crystal structure predictions [38, 39, 40, 41, 42], with the aid of complex algorithms that have been implemented for that purpose [43, 44, 45, 46]. The noticeable presence of hydrogen-based superconductors in the scientific literature is evidenced in the more than five thousand publications in the topic that appeared in the 2010-2020 period [16].

A feature of hydrogen-rich compounds that is worth mentioning, and that has an important effect on the crystal structure prediction of hydrides and on the estimation of their phonon spectra, is the quantum fluctuations suffered by the hydrogen due to its light mass. It has been shown that a harmonic approximation does not properly describe the dynamic behavior of the hydrogen nuclei, and that an anharmonic treatment of the systems is necessary. This is the case of LaH_{10}, for which a quantum description is required to stabilize the experimentally observed structure [47].

The clear drawback to hydrogen-rich type of superconductors is the need to achieve such extreme pressures, usually well above 100 GPa. Just as hydrogen-rich binary compounds proved that adding other atoms to a hydrogen lattice allows to diminish the pressure for which the material is stable, the search for superconductors at ambient conditions has now shifted towards hydrogen-rich ternaries. The hope is that the new, much vaster space of ternaries will allow to further reduce the pressure. However, although having a huge search space means having more possibilities, it also means that it is much more difficult to explore extensively. In this way, theoretical predictions become even more important in guiding experimental research.

Although the theoretical tools to predict T_c have evolved enormously through the years, they are still a bottleneck in the prediction of new structures. In fact, to compute this value one must optimize the structure and get its energy and phonon spectra, and then proceed to get the electron-phonon coupling properties. For a system like LaH_{10}, this can take up to $\sim$ 30000 CPU hours. If the stoichiometry of a compound is not known, one must do this for several compositions and candidate structures, further increasing the computational cost.

In this way, it seems interesting to study hydrogen-rich systems from another perspective, to get new insights on their properties that could help narrow down the search space. A path that has not been fully explored yet is the employment of quantum chemistry tools to study these compounds. In fact, the microscopic theory of conventional superconductivity is described entirely in terms of quantities defined in momentum space. Studying them in real-space should bring in a brand new perspective. In particular, learning about the spatial behavior of the electrons seems interesting as it should unveil the chemistry that enables and enhances superconductivity in hydrides. Hopefully, the newly-gained intuition should help in the faster prediction of future promising high-T_c superconductors.

Another reason to study the real-space properties of hydrogen-based superconductors is because, being under such extreme pressures, the way that elements interact within those structures will be different from their behaviors at ambient conditions. Certainly, it is well known that pressure can induce phase transitions that alter the conducting properties of a system. An example of this is given in Figure 1.2 [48], where pressure induces a transition from a metallic to an insulating phase in

a sodium crystal [49]. This is visualized by the changes of the electron localization function in the compound, that is rather flat for the metallic state, whereas it shows clear electron localization sites in interstitial sites for the insulating state.

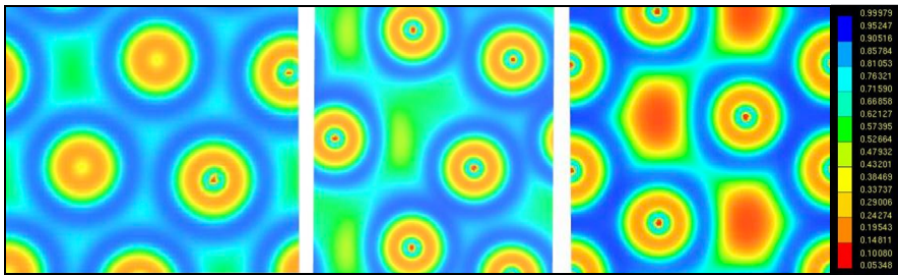


Fig. 1.2: From left to right: Electron localization function (ELF) of low pressure phase of Na-cI16, intermediate-pressure phase of Na-oP8, and higher pressure phase Na-hP4. Higher values of the ELF are depicted in red, and lower ones in blue. Adapted with permission from ref. [48].

In summary, the synthesis and prediction of hydrogen-based superconductors pose three major difficulties for scientists. The first one is that they are stable at very high pressures; the second one, that there is currently no way of knowing systematically which system will be a good superconductor. Finally, their critical temperature is very expensive to compute. In this thesis, our aim is to contribute in the understanding of the properties that characterize a high-temperature hydrogen-based superconductor. The approach will be that of Quantum Chemistry, and in particular of Quantum Chemical Topology (QCT). This formalism provides us with tools to analyze the localization and delocalization of electrons in real-space, and thus reveal their chemical structure. In the first part of this manuscript, some of the theoretical background needed for the development of this work will be presented. This comprehends some general introduction of the quantum mechanics of many-electron systems, as well as a look into the theory behind conventional superconductivity, and the introduction of some of the concepts inherent to QCT that have been used throughout this thesis, with particular emphasis on the Electron Localization Function (ELF). Part II introduces a one-dimensional model, where we define topological descriptors on the superconducting state using concepts from Superconducting Density Functional Theory (SCDFT). Finally, Part III focuses on real systems, in which we analyze some patterns of electronic localization that characterize hydrogen-based superconductors, as well as a real-space study of the anharmonic fluctuations of the effective potential in superconducting hydrides, with an outlook on assessing its effects on the electron-phonon coupling in those materials.

Part I

Theoretical background

2

The many-electron problem

Many of the properties of the matter that surrounds us are a consequence of the electronic structure of the materials that make it up. Hence, it becomes very important to be able to solve the many-electron problem in quantum mechanics by finding the electronic wavefunction, $\Psi(\mathbf{r}_1, \dots \mathbf{r}_N)$, with N the number of electrons, with which we can evaluate the expectation value of many of its properties. However, this problem is a complex one, and we will need to introduce a few approximations. In this chapter, we will begin by presenting the Schrödinger equation in Section 2.1, and the Born-Oppenheimer approximation in Section 2.2, where the degrees of freedom of the nuclei and the electrons will be separated to simplify the problem. Then, the basics of density functional theory will be layed out in Section 2.3. In Section 2.4 we will introduce the concept of electron-phonon coupling, which describes the effect of the vibration of the nuclei on the electronic structure. Finally, we will describe the tight-binding approximation in Section 2.5, where we treat the electronic problem with a model Hamiltonian.

2.1 The Schrödinger equation

In non-relativistic quantum mechanics, all of the properties of a given system can be obtained by solving the Schrödinger equation. In its time-independent form, it reads

$$\hat{H}\Psi_i = E_i\Psi_i\,. \tag{2.1}$$

Here, E_i is the energy of the i-th state of the system, Ψ_i its many-body wavefunction, and $\hat{H}$ is the molecular Hamiltonian (in a.u.),

$$\hat{H} = -\sum_i \frac{1}{2}\nabla^2_{\mathbf{r}_i} - \sum_A \frac{1}{2M_A}\nabla^2_{\mathbf{R}_A} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,A}\frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} + \sum_{A=1}^{M}\sum_{B=1}^{M}\frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|}\,, \tag{2.2}$$

where N and M are the number of electrons and nuclei, respectively, Z_A is the atomic number of nucleus A, M_A its mass, and the electronic and atomic coordinates are

denoted by $\mathbf{r}_i$ and $\mathbf{R}_A$, respectively. The first two terms of the Hamiltonian correspond to the electron and nuclei kinetic contribution, $\hat{T}_e$ and $\hat{T}_n$; the third term to electron-electron repulsion, $\hat{V}_{ee}$; the fourth to electron-nuclei attraction, $\hat{V}_{ne}$; and the last one to the repulsion between the nuclei, $\hat{V}_{nn}$.

We will generally be interested in finding the ground state of the system, Ψ_0, that has the lowest energy, E_0. However, solving 2.2 becomes practically impossible, as there are $3N+3M$ degrees of freedom. Therefore, it becomes necessary to introduce some approximations in order to solve the Schödinger's equation.

2.2 The Born-Oppenheimer approximation

The wavefunction $\Psi(\mathbf{r},\mathbf{R})$ depends on all electronic and atomic coordinates, $\mathbf{r} = \{\mathbf{r}_i\}$ and $\mathbf{R} = \{\mathbf{R}_A\}$, but the difference in mass between electrons and nuclei, $m_e \ll M_A$ ($m_e = 1$ in a.u.), allows to separate the motions of the two. In fact, electrons will move much faster, and we can consider them to interact instantly to changes in the potential generated by the comparatively slow movement of the nuclei. Therefore, we can write $\Psi(\mathbf{r},\mathbf{R}) = \Phi(\mathbf{R})\psi(\mathbf{r};\mathbf{R})$, where $\Phi(\mathbf{R})$ describes the movement of the nuclei, and $\psi(\mathbf{r};\mathbf{R})$ is the electronic wavefunction, that depends on the atomic coordinates only parametrically. The latter satisfies

$$\hat{H}_e\,\psi_i(\mathbf{r};\mathbf{R}) = \mathcal{E}_i(\mathbf{R})\,\psi_i(\mathbf{r};\mathbf{R})\,, \tag{2.3}$$

with

$$\hat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}\,. \tag{2.4}$$

In this way, the structural parameters $\mathbf{R}$ only appear in the electron-nuclei interaction term, which becomes an external potential acting on the electrons (the term $\hat{V}_{nn}$ now only adds a constant to the energy).

Once again, even after introducing this approximation, solving 2.3 can be quite challenging. Plenty of other different approximations can be introduced to finally find $\psi(\mathbf{r};\mathbf{R})$, leading to a plethora of methods that quantum chemists and physicists use nowadays to tackle the electronic many-body problem. Hartree-Fock (HF) and Density Functional Theory (DFT), for example, approximate the electronic wavefunction with a Slater determinant, leading to a mean-field treatment of the electron-electron repulsion. These methods are well-known and we will not describe them thoroughly in these thesis other than a brief introduction to DFT. Great text book references can be found in [50, 51, 52, 53].

2.3 Basics of Density Functional Theory

Density Functional Theory (DFT) dates back to 1964, when Hohenberg and Kohn demonstrated two theorems that stated (i) that there is a one-to-one relationship between the external potential created by the nuclei, $v_{\text{ext}}(\mathbf{r})$, and the electron density,

$\rho(\mathbf{r})$; and (ii) that the electronic energy, now written as the functional $E[\rho(\mathbf{r})]$, follows a variational principle with respect to the density [54]. The energy functional can be separated as,

$$E[\rho(\mathbf{r})] = T_e[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] + V_{ne}[\rho(\mathbf{r})], \tag{2.5}$$

$$= F[\rho(\mathbf{r})] + \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})\, d\mathbf{r}, \tag{2.6}$$

where $F[\rho(\mathbf{r})]$ is a universal functional that is not system-dependent and contains the electronic kinetic energy and electron-electron interaction terms coming from the Hamiltonian in equation (2.4), while the electron-nuclei interaction is captured in the external potential $v_{\text{ext}}(\mathbf{r})$. The analytical form of the universal functional not being known, the DFT formalism was not widespread until Kohn and Sham introduced a method in which a system of non-interacting electrons is used as reference [55]. The many-body wavefunction of such a system can be written exactly as a Slater determinant,

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \cdots & \phi_1(\mathbf{r}_N) \\ \vdots & \ddots & \vdots \\ \phi_N(\mathbf{r}_1) & \cdots & \phi_N(\mathbf{r}_N) \end{vmatrix}, \tag{2.7}$$

where $\{\phi_i\}$ is a set of N one-electron wavefunctions, or *orbitals*. Therefore, the kinetic energy of the Kohn-Sham (KS) system can be written directly as

$$T_s[\{\phi_i\}] = \frac{1}{2} \int \sum_i |\nabla \phi_i(\mathbf{r})|^2 \, d\mathbf{r}. \tag{2.8}$$

In the same way, the density is also a functional of the orbitals, $\rho(\mathbf{r}) = \sum_i |\phi_i(\mathbf{r})|^2$. Another simplification is to consider the electron-electron interaction as the classical Coulomb repulsion between averaged electron distributions,

$$V_H[\rho(\mathbf{r})] = \int \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r}\, d\mathbf{r}'. \tag{2.9}$$

In this way, the exact energy of the system can be rewritten as

$$E[\{\phi_i(\mathbf{r})\}] = T_s[\{\phi_i(\mathbf{r})\}] + V_H[\rho(\mathbf{r})] + E_{xc}[\rho(\mathbf{r})] + \int v_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})\, d\mathbf{r}, \tag{2.10}$$

where E_{xc} is the exchange-correlation energy, containing all the corrections to the approximations made by introducing T_s and V_H,

$$E_{xc} = T - T_s + V_{ee} - V_H. \tag{2.11}$$

As to now, the expression for the energy in (2.10) is exact. However, because the functional form of the exchange-correlation is not known, we need to do some

approximations. The first attempts in that direction used the formalism of the homogeneous electron gas, leading to a local-density approximation [55], and yielding surprisingly good results in metals considering the simplicity of the model. Further approximations to E_{xc} include the gradient of the density, grouped under the name of Generalized Gradient Approximation (GGA) (generalized-gradient approx.) functionals, which are very broadly used nowadays in solids state calculations [56, 57, 58, 59]. A plethora of other more sophisticated and specialized functionals are available nowadays, see ref. 60.

Once the level of theory (i.e. the E_{xc} functional) is chosen, one must minimize the energy in eq. (2.10) with respect to the orbitals. This leads to the Kohn-Sham (KS) equations

$$\left[-\frac{1}{2}\nabla^2 + v_{KS}(\mathbf{r})\right]\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}), \tag{2.12}$$

describing the behavior of a single electron in the effective potential

$$v_{KS}(\mathbf{r}) = v_{\text{ext}}(\mathbf{r}) + \frac{\delta V_H[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})}. \tag{2.13}$$

Because this potential depends on the density and therefore on the orbitals, the equations (2.12) must be solved self-consistently.

In a periodic system of lattice vector $\mathbf{T}$, the KS potential must satisfy

$$v_{KS}(\mathbf{r}+\mathbf{T}) = v_{KS}(\mathbf{r}), \tag{2.14}$$

and Bloch's theorem states that the solutions to (2.12) must be of the form [61]

$$\phi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n\mathbf{k}}(\mathbf{r}), \tag{2.15}$$

where the function $u_{n\mathbf{k}}$ has the periodicity of the lattice,

$$u_{n\mathbf{k}}(\mathbf{r}+\mathbf{T}) = u_{n\mathbf{k}}(\mathbf{r}). \tag{2.16}$$

In this way, the KS orbitals have the following property

$$\phi_{n\mathbf{k}}(\mathbf{r}+\mathbf{T}) = e^{i\mathbf{k}\cdot\mathbf{T}}\phi_{n\mathbf{k}}(\mathbf{r}). \tag{2.17}$$

Notice that the index i of the KS equations have unfolded into the indices n and $\mathbf{k}$. The first one is the band index, and the second corresponds to the plane wave vector of (2.15), and it is a vector in the reciprocal space of $\mathbf{r}$. Imposing periodic boundary conditions (Born-von Karman boundary conditions) on $\phi_{n\mathbf{k}}(\mathbf{r})$ restricts the possible values of $\mathbf{k}$, that have a periodicity determined by a vector $\mathbf{G}$ such that $e^{\mathbf{G}\cdot\mathbf{T}} = 1$. Such choice of boundary conditions allow us to write the wavefunction as

$$\phi_{n\mathbf{k}} = \frac{1}{\sqrt{\Omega}}\sum_{\mathbf{G}} c_{n\mathbf{k}+\mathbf{G}}\, e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \tag{2.18}$$

with Ω the volume of the lattice and the sum being over all possible reciprocal space vectors (in practice, this sum must be truncated, see ref. 53). This plane wave formulation of DFT allows us to solve the electronic problem in a crystal lattice by introducing equation (2.18) into (2.12), and solving self-consistently for the coefficients $c_{n\mathbf{k}+\mathbf{G}}$.

2.4 Electron-phonon coupling

The Born-Oppenheimer approximation allowed us to separate the equations for electrons and nuclei, considering that the latter move so slow that the former feel the effect in the potential instantly. However, the nuclei are in fact moving, vibrating around their equilibrium positions. We will assume that these deviations are small, and consider nuclei as harmonic oscillators. This is the simplest approximation that we can make, and it turns out to work rather well for many situations. The quasiparticles describing these harmonic vibrations are called phonons, and they can actually interact with the electrons. In this section, we will first introduce the formalism to understand phonon modes in a solid, and then we will see how those couple to the electronic wavefunctions.

2.4.1 Lattice dynamics and phonons

Once the electronic ground state of energy $\mathcal{E}_0(\mathbf{r})$ has been found by solving eq. (2.3), the potential energy to be considered is [62]

$$\Omega(\mathbf{R}) = V_{ii}(\mathbf{R}) + \mathcal{E}_0(\mathbf{R}) . \tag{2.19}$$

Here, we have included the ion-ion interaction potential, V_{ii}, instead of the nucleus-nucleus term of section 2.1, V_{nn}. In this way, we consider that electrons of the inner core move along with the nuclei, screening the potential for the rest of the electrons. The deviations of the position of atom A with respect to equilibrium, $\mathbf{R}_A^0$, is given by

$$\mathbf{R}_A = \mathbf{R}_A^0 + \mathbf{u}_A . \tag{2.20}$$

A Taylor expansion of $\Omega(\mathbf{R})$ around equilibrium to first order will yield terms containing the derivative $\frac{\partial \Omega}{\partial R_{A\alpha}}\Big|_{\mathbf{u}=0}$, where α corresponds to a cartesian index. Those terms actually correspond to the negative of the α component of the force acting over ion A at the equilibrium position, which is zero by definition. In this way, the first non-trivial order of the expansion is the square term, leading to

$$\Omega(\mathbf{R}) = \Omega(\mathbf{R}^0) + \sum_{A\alpha B\beta} \Phi_{\alpha\beta}(A,B)\, u_{A\alpha}\, u_{B\beta} + \cdots , \tag{2.21}$$

where we have introduced

$$\Phi_{\alpha\beta}(A,B) = \left. \frac{\partial^2 \Omega}{\partial R_{A\alpha} \partial R_{B\beta}} \right|_{\mathbf{u}=0} . \tag{2.22}$$

We identify this second derivative of Ω as a three-dimensional analogue of a spring constant. In fact, we can write the β component of the force acting on ion B due to the displacement of ion A by $u_{A\alpha}$ as

$$F_{B\beta} = -\sum_{A\alpha} \Phi_{\alpha\beta}(A,B)\, u_{A\alpha} . \tag{2.23}$$

In this way, the hessian of Ω describes a linear dependence of the force with respect to the displacements. This is a characteristic of the harmonic approximation, which we imposed when we truncated the expression in (2.21). Higher-order terms of the expansion of $\Omega(\mathbf{R})$ are said to be *anharmonic*.

We hereby consider that the ions form a periodic lattice, and we shall therefore use $A = (l\kappa)$, with l the index of unit cell, and κ the index of the ion inside that cell. We define the *dynamical matrix* as the Fourier transform of the force constant $\Phi_{\alpha\beta}(A, B)$,

$$D_{\kappa\alpha\kappa'\beta}(\mathbf{q}) = \frac{1}{\sqrt{M_\kappa M_{\kappa'}}} \sum_l \Phi_{\alpha\beta}(l\kappa, 0\kappa')\, e^{-i\mathbf{q}\cdot\left(\mathbf{R}^0_{l\kappa} - \mathbf{R}^0_{0\kappa'}\right)}. \tag{2.24}$$

The eigenvectors and eigenvalues of this matrix correspond to the normal modes $\eta_\kappa(\mathbf{q}j)$ and the square of their frequencies, $\omega_{\mathbf{q}j}$, respectively. Here, j is the branch index describing the different possible modes for a given momentum $\mathbf{q}$,

$$\sum_{\kappa'\beta} D_{\kappa\alpha\kappa'\beta}(\mathbf{q})\, \eta_{\kappa'\beta}(\mathbf{q}j) = \omega^2_{\mathbf{q}j}\, \eta_{\kappa\alpha}(\mathbf{q}j). \tag{2.25}$$

We can now write a Hamiltonian for the nuclei [63],

$$\hat{H}_p = -\sum_{l\kappa\alpha} \frac{\hbar}{2M_\kappa} \frac{\partial^2}{\partial R^2_{l\kappa\alpha}} + \frac{1}{2} \sum_{l\kappa\alpha} \sum_{l'\kappa'\beta} \Phi_{\alpha\beta}(l\kappa, l'\kappa')\, u_{l\kappa\alpha}\, u_{l'\kappa'\beta} + \Omega(\mathbf{R}^0) \tag{2.26}$$

where the ground state-energy $\Omega(\mathbf{R}^0)$ is a constant with respect to ionic displacements. Rewriting the Hamiltonian in eq. (2.26) in second quantization can be useful, as it can be diagonalized with respect to the ladder operators $\hat{b}_{\mathbf{q}j}$ and $\hat{b}^\dagger_{-\mathbf{q}j}$ such that

$$u_{l\kappa\alpha} = \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}j} e^{i\mathbf{q}\cdot\mathbf{R}_l} \eta_{\kappa\alpha}(\mathbf{q}j) \sqrt{\frac{\hbar}{2M_\kappa \omega_{\mathbf{q}j}}} \left(\hat{b}_{\mathbf{q}j} + \hat{b}^\dagger_{-\mathbf{q}j}\right), \tag{2.27}$$

and

$$P_{l\kappa\alpha} = i\, \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}j} e^{i\mathbf{q}\cdot\mathbf{R}_l} \eta_{\kappa\alpha}(\mathbf{q}j) \sqrt{\frac{\hbar M_\kappa \omega_{\mathbf{q}j}}{2}} \left(\hat{b}^\dagger_{-\mathbf{q}j} - \hat{b}_{\mathbf{q}j}\right), \tag{2.28}$$

where N_l is the number of unit cells, $\mathbf{R}_l$ the origin of the l-th unit cell, and $\mathbf{P}_{l\kappa}$ the momentum operator of the ion of index $(l\kappa)$, corresponding to the derivative of the ionic kinetic energy of equation (2.26). With this, we obtain the diagonalized Hamiltonian

$$\hat{H}_p = \sum_{\mathbf{q}j} \hbar\omega_{\mathbf{q}j} \left(\hat{b}^\dagger_{\mathbf{q}j} \hat{b}_{\mathbf{q}j} + \frac{1}{2}\right). \tag{2.29}$$

In this way, we can identify $\hat{b}^\dagger_{\mathbf{q}j}$ and $\hat{b}_{-\mathbf{q}j}$ as the creation an annihilation operators of quasi-particles that represent the quanta of vibrations, which we call *phonons*. The commutation rules of those operators are

$$\left[\hat{b}_{\mathbf{q}j}, \hat{b}^\dagger_{\mathbf{q}'j'}\right] = \delta_{\mathbf{q}\mathbf{q}'}\delta_{jj'}, \quad \left[\hat{b}_{\mathbf{q}j}, \hat{b}_{\mathbf{q}'j'}\right] = \left[\hat{b}^\dagger_{\mathbf{q}j}, \hat{b}^\dagger_{\mathbf{q}'j'}\right] = 0, \tag{2.30}$$

i.e. phonons follow Bose-Einstein statistics.

2.4.2 Electron-phonon coupling Hamiltonian

If we consider the ionic vibrations in the context of density functional theory (see Section 2.3), and the KS potential of eq. (2.13) now depending explicitly on the ionic coordinates $\{\mathbf{R}_{l\kappa}\}$, its expansion to first order with respect to ionic displacements would be

$$v_{KS}(\mathbf{r},\{\mathbf{R}_{l\kappa}\}) = v_{KS}(\mathbf{r},\{\mathbf{R}^0_{l\kappa}\}) + \sum_{l\kappa\alpha} \frac{\partial v_{KS}(\mathbf{r},\{\mathbf{R}_{l\kappa}\})}{\partial R_{l\kappa_\alpha}} u_{l\kappa\alpha}\,. \tag{2.31}$$

Replacing the expression for $u_{l\kappa\alpha}$ in second quantization of equation (2.27) into (2.31), we obtain

$$v_{KS}(\mathbf{r},\{\mathbf{R}_{l\kappa}\}) = v_{KS}(\mathbf{r},\{\mathbf{R}^0_{l\kappa}\}) + \frac{1}{\sqrt{N_l}} \sum_{\mathbf{q}j} e^{i\mathbf{q}\cdot\mathbf{r}} \Delta_{\mathbf{q}j} v_{KS} \left(\hat{b}_{\mathbf{q}j} + \hat{b}^\dagger_{-\mathbf{q}j}\right), \tag{2.32}$$

where we have defined the lattice-periodic quantity

$$\Delta_{\mathbf{q}j} v_{KS} = \sum_{\kappa\alpha} \sqrt{\frac{\hbar}{2M_\kappa \omega_{\mathbf{q}j}}} \eta_{\kappa\alpha}(\mathbf{q}j) \sum_l e^{-i\mathbf{q}\cdot(\mathbf{r}-\mathbf{R}_l)} \frac{\partial v_{KS}(\mathbf{r}-\mathbf{R}_l,\{\mathbf{R}_\kappa\})}{\partial R_{\kappa\alpha}}, \tag{2.33}$$

that is a measure of the fluctuation of the KS effective potential when the lattice has been distorted by a phonon of wavevector $\mathbf{q}$ and branch index j.

Considering that the difference between the perturbed and the unperturbed KS potentials is what coupled the electrons to the ionic vibrations, one finds the electron-phonon coupling Hamiltonian in second quantization,

$$\hat{H}_{e-ph} = \sum_{n\mathbf{k},n'K'} \langle \phi_{n\mathbf{k}} | v_{KS}(\mathbf{r},\{\mathbf{R}_{l\kappa}\}) - v_{KS}(\mathbf{r},\{\mathbf{R}^0_{l\kappa}\}) | \phi_{n'\mathbf{k}'} \rangle \hat{c}^\dagger_{n\mathbf{k}} \hat{c}_{n'\mathbf{k}'}\,, \tag{2.34}$$

with $|\phi_{n\mathbf{k}}\rangle$ the ket form of the Bloch functions of (2.15), and $\hat{c}^\dagger_{n\mathbf{k}}$, $\hat{c}_{n\mathbf{k}}$ the creation and annihilation operators that diagonalize the KS all-electron Hamiltonian:

$$\hat{H}^{KS}_e = \sum_{n\mathbf{k}} \varepsilon_{n\mathbf{k}} \hat{c}^\dagger_{n\mathbf{k}} \hat{c}_{n\mathbf{k}}\,, \tag{2.35}$$

with $\varepsilon_{n\mathbf{k}}$ the eigenvalues in eq. (2.12). In this sense, those operators represent the creation and annihilation of non-interacting electrons in the KS framework, which correspond to quasi-particles in the interacting picture.

Replacing equation (2.32) into (2.34), we obtain

$$\hat{H}_{e-ph} = \frac{1}{\sqrt{N_l}} \sum_{\mathbf{k},\mathbf{q}} \sum_{mnj} g_{mnj}(\mathbf{k},\mathbf{q})\, \hat{c}^\dagger_{m\mathbf{k}+\mathbf{q}} \hat{c}_{n\mathbf{k}} \left(\hat{b}_{\mathbf{q}j} + \hat{b}^\dagger_{-\mathbf{q}j}\right), \tag{2.36}$$

where we have introduced the electron-phonon matrix element

$$g_{mnj}(\mathbf{k},\mathbf{q}) = \langle u_{m\mathbf{k}+\mathbf{q}} | \Delta_{\mathbf{q}j} v_{KS} \, | u_{n\mathbf{k}} \rangle_{uc}\,, \tag{2.37}$$

where the integral is inside a single unit cell (hence the subscript *uc*). Notice that only the periodic part of the Bloch functions survives, $|u_{n\mathbf{k}}\rangle$, as the exponentials cancel each other out. The electron-phonon Hamiltonian in equation (2.36) determines the energy involved in the scattering of an electron from state n with momentum $\mathbf{k}$ into a state m with momentum $\mathbf{k}+\mathbf{q}$ by the absorption of a phonon of momentum $\mathbf{q}$ (or emission of phonon of momentum $-\mathbf{q}$). Then, $g_{mnj}(\mathbf{k},\mathbf{q})$ gives us the magnitude of that energy, dictating how important the electron-phonon term is inside the Hamiltonian and giving us a measure of the strength of the coupling.

In order to compute $g_{mnj}(\mathbf{k},\mathbf{q})$, one must evaluate the derivative of the KS potential with respect to phonon-induced ionic displacements. There is more than one approach to do so, the simplest one being the *frozen phonon* method, where the derivatives of the potential are performed numerically by evaluating

$$\left.\frac{\partial v_{KS}}{\partial R_{l\kappa\alpha}}\right|_{\mathbf{R}^0_{lk}} \approx \frac{v_{KS}(\mathbf{r}, R^0_{l\kappa\alpha}+\epsilon_{l\kappa\alpha}(\mathbf{q}j)) - v_{KS}(\mathbf{r}, R^0_{l\kappa\alpha})}{\epsilon_{l\kappa\alpha}(\mathbf{q}j)}, \tag{2.38}$$

with $\epsilon_{l\kappa\alpha}(\mathbf{q}j)$ a small displacement of the ionic coordinate $R_{l\kappa\alpha}$ with respect to equilibrium, which is characteristic of a given phonon of wavevector $\mathbf{q}$ and branch index j. In this sense, it is necessary to already know the normal modes of the system, either by infering them from the symmetry, which is only possible in very simple systems, or by using other type of methods. Moreover, if this method is used, it is often necessary to use a big supercell, in order to fully capture the effect of the fluctuations at all ranges in the periodic lattice. Hence, the frozen phonon method, although very useful considering its simplicity, can become very expensive, and more modern calculation are done using Density Functional Perturbation Theory (DFPT), where it is possible to do the integrals inside the unit cell [64, 65, 66].

Finally, it is worth mentioning that it is possible to compute the electron-phonon matrix elements using localized orbitals instead of plane waves. These localized orbitals correspond to Wannier functions [63, 67],

$$w_{ml}(\mathbf{r}) = \frac{1}{N_l}\sum_{n\mathbf{k}} e^{-i\mathbf{k}\cdot(\mathbf{r}-\mathbf{R}_l)} U_{nm\mathbf{k}} u_{n\mathbf{k}}(\mathbf{r}), \tag{2.39}$$

where $U_{nm\mathbf{k}}$ is a unitary matrix that allows for band mixing, and $w_{ml}(\mathbf{r})$ the spatial representation of the m-th Wannier function in the l-th unit cell. Then, the matrix element in spatial representation is

$$g_{mn\kappa\alpha}(\mathbf{R}_l,\mathbf{R}_{l'}) = \left\langle w_{m0}(\mathbf{r}) \left| \frac{\partial v_{KS}(\mathbf{r}-\mathbf{R}_{l'})}{\partial R_{\kappa\alpha}} \right| w_{n0}(\mathbf{r}-\mathbf{R}_l) \right\rangle_{sc}, \tag{2.40}$$

where the subscript 0 indicates that the functions are in the unit cell at the origin, and *sc* that the integral is taken over the entire supercell. The relationship between this and the expression for $g_{mnj}(\mathbf{k},\mathbf{q})$ in equation (2.37) is

$$g_{mnj}(\mathbf{k},\mathbf{q}) = \sum_{ll'} e^{i(\mathbf{k}\cdot\mathbf{R}_l+\mathbf{q}\cdot\mathbf{R}_{l'})} \sum_{m'n'\kappa\alpha} \sqrt{\frac{\hbar}{2M_\kappa\omega_{\mathbf{q}j}}} U_{mm'\mathbf{k}+\mathbf{q}} g_{m'n'\kappa\alpha}(\mathbf{R}_l,\mathbf{R}_{l'}) U^\dagger_{n'n\mathbf{k}} \eta_{\kappa\alpha}(\mathbf{q}j). \tag{2.41}$$

Because in these thesis we are mainly interested in descriptors in real space that can account for the superconducting properties of materials, the expression in (2.40) is of special interest to us, as we shall see in Section 7.

2.4.3 SSCHA: beyond the harmonic approximation

Everything that we have reviewed so far about ionic vibrations has been done within the harmonic approximation. However, the Born-Oppenheimer landscape, $\Omega(\mathbf{R})$, is not a collection of harmonic potentials, even if it resembles it around the ionic equilibrium positions. In some cases, the ionic behavior deviates considerably from that description, for example when temperature-driven fluctuations are important, or when very light atoms are present and their quantum nature begins to play a more important role. The latter is exactly the case of hydrogen-rich materials, for which it has been shown that quantum fluctuations should be considered to properly describe their crystal structure, which in turn will affect their predicted superconducting properties [68, 69, 70]. One way of describing anharmonicity would be to include higher-order terms in equation (2.21), but it would mean to assume that the deviations are small, which is not necessarily the case for quantum fluctuations.

The Stochastic Self-Consistent Harmonic Approximation (SSCHA) allows to account for both thermodynamic and quantum anharmonic effects [71, 72, 73]. Its main use is to optimize a crystal structure based on the principle that it must minimize the free energy,

$$\mathcal{F}[\tilde{\rho}] = E[\tilde{\rho}] - TS[\tilde{\rho}], \tag{2.42}$$

with respect to the (quantum) density matrix $\tilde{\rho}$. Here, T is the temperature, and S is the entropy of the ions, that accounts for the disorder in the different microstates of the ensemble. The energy E is

$$E[\tilde{\rho}] = \langle T_i + \Omega(\mathbf{R}) \rangle_{\tilde{\rho}}, \tag{2.43}$$

where T_i is the kinetic energy of the ions. The trial densities,

$$\tilde{\rho}_{\mathcal{R},\Phi}(\mathbf{R}) = \langle \mathbf{R} | \tilde{\rho}_{\mathcal{R},\Phi} | \mathbf{R} \rangle, \tag{2.44}$$

represent the probability of finding ions at the positions $\mathbf{R}$. They will be constrained to a Gaussian form, with centroids $\mathcal{R}$, and mean squared displacements Φ.

The choice of a Gaussian probability distribution allows to write the problem as that given by the harmonic Hamiltonian,

$$H_{\mathcal{R},\Phi} = T_i + \frac{1}{2}\sum_{ab}(R_a - \mathcal{R}_a)\Phi_{ab}(R_b - \mathcal{R}_b), \tag{2.45}$$

where the subindices a, b run over all ionic and Cartesian coordinates. Here, we identify Φ as an auxiliary force constant matrix. The physical meaning of that quantity, however, is not that of the real force constants, and one must be careful not to pretend to extract actual phonon frequencies from it. It is, however, quite convenient, allowing to obtain T_i and S easily from the Gaussian distribution.

In this way, once an initial guess for the SSCHA parameters, $\mathcal{R}$ and Φ, are made, a trial function is constructed as

$$\tilde{\rho}_{\mathcal{R},\Phi}(\mathbf{R}) = \sqrt{\det\left(\frac{\Psi^{-1}}{2\pi}\right)}\exp\left[-\frac{1}{2}\sum_{a,b}(R-\mathcal{R}_a)\Psi^{-1}_{ab}(R-\mathcal{R}_b)\right]. \tag{2.46}$$

Here, we have used the displacement-displacement correlation matrix

$$\Psi_{ab} = \langle u_a u_b\rangle_{\mathcal{R},\Phi} = \frac{1}{\sqrt{M_a M_b}}\sum_{\mu}\frac{\hbar(2n_\mu+1)}{2\omega_\mu}\eta_a(\mu)\eta_b(\mu), \tag{2.47}$$

that depends on the eigenvalues and eigenvectors of $\Phi_{\mathcal{R},\Phi}$, ω_μ and $\eta(\mu)$, respectively; and on the Bose-Einstein occupation numbers, n_μ.

Having a trial $\tilde{\rho}_{\mathcal{R},\Phi}(\mathbf{R})$, an ensemble of random configurations is obtained according to the probability density. At this point, an *ab initio* code must be employed to obtain all the energies and forces for each configuration. Then, the free energy $\mathcal{F}[\tilde{\rho}_{\mathcal{R},\Phi}]$ of the ensemble and its derivatives with respect to the optimization parameters are performed. This allows to update $\mathcal{R}$ and Φ, and to start the process with a new trial density, until convergence, giving the self-consistent part of the name to the method. The stochasticity, on the other hand, comes from the fact that a Monte Carlo method is used to evaluate the statistical averages in the computation of $\mathcal{F}[\tilde{\rho}_{\mathcal{R},\Phi}]$ and its derivatives.

A great illustration of the method can be made by considering the anharmonic one-dimensional energy surface of Fig. 2.1a, originally presented in ref. 73. There, despite the fact the SSCHA density is constrained to being a Gaussian, deviating from its exact shape, it very successfully predicts the position an energy of the average ionic position, greatly outperforming the harmonic approach.

The mechanism that we have laid out is the basic idea behind the SSCHA, although it can become more complex if one wishes to perform optimizations at given pressures, or take into account the electronic entropy, for example. Further, one can obtain physical quantities, like the real force constant matrix, from the formalism of the SSCHA. This allows to actually characterize the system and makes it a great tool to include quantum-thermal fluctuations in a crystal structure optimization. We shall not dig deeper into the generosities of the method in this thesis, but the interested reader can refer to ref. 73 for more information.

2.5 The tight-binding model

Although the ideal would be to always approach the many-electron problem from first principles, i.e. solving Schrödinger's equation for the molecular Hamiltonian in (2.2), it is sometimes desirable to have a simpler way of representing our systems. This permits to reduce the variables and assess how their modification affects the properties of the system. The tight-binding model allows to describe the electrons in a crystal lattice considering that the electronic wavefunction is very similar to

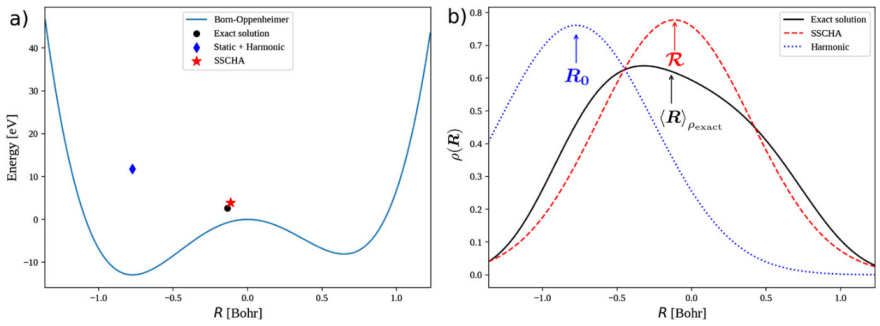


Fig. 2.1: a) The continuous line represents a one-dimensional anharmonic energy surface. The exact solution for the average position and energy of an electron in such a potential is marked by a black circle. The same quantities, as evaluated using the harmonic approximation and the SSCHA are marked by a blue diamond and a red star, respectively. b) The exact probability density, as well as its approximations are displayed using the same color coding. Reprinted from Ref. [70].

the atomic wavefunctions in the proximity of the ions. In this way, the one-particle Hamiltonian looks as follows [61, 74],

$$\hat{H} = \hat{H}_{at} + \Delta U\,, \tag{2.48}$$

where $\hat{H}_{at}$ is the Hamiltonian of a single atom, satisfying

$$\hat{H}_{at}\chi_i(\mathbf{r}) = \epsilon_i \chi_i(\mathbf{r})\,, \tag{2.49}$$

with $\chi_i(\mathbf{r})$ the atomic orbitals and ϵ_i its corresponding energies. In this way, ΔU is the deviation of $\hat{H}$ from the isolated atom picture, which we assume to be small and to go to zero when $\mathbf{r} \to \mathbf{R}$. The Schrödinger's equation for $\hat{H}$ is

$$\hat{H}\,\phi_{n\mathbf{k}} = E_{n\mathbf{k}}\,\phi_{n\mathbf{k}}\,. \tag{2.50}$$

Here, $\phi_{n\mathbf{k}}$ and $E_{n\mathbf{k}}$ are the system's orbitals and their energies, respectively; and they are labelled by the band index n and a wave-vector $\mathbf{k}$ of the first Brillouin Zone (BZ). Due to the periodicity of the crystal, the orbitals must follow Bloch's theorem and abide by eq. (2.17). This can be done by writing $\phi_{n\mathbf{k}}$ in terms of the atomic orbitals.

In this thesis, we are interested in using the tight-binding model to describe a one-dimensional hydrogen chain. This is a well-known model that is able to explain the metal-to-insulator transition, and the phenomenon of Peierls distortion, where the symmetry of a hydrogen chain is spontaneously broken towards a dimerized lattice [75, 76].

In order to be able to distort the lattice later on, we will consider a unit cell containing two atoms, as represented in Figure 2.2. We divide the lattice into sublatices A and B, whose atoms have the same atomic wavefunction, $\chi(\mathbf{r})$, but centered at different positions. Atoms in sublattice A are located at

$$R_A = la\,, \quad l = 0, \pm 1, \pm 2 \ldots \tag{2.51}$$

and in sublattice B, atoms are centered at positions

$$R_B = R_{AB} + la\,, \quad l = 0, \pm 1, \pm 2 \ldots \tag{2.52}$$

We have dropped the vector notation because of the one-dimensionality of the system. The distance a is the length of the unit cell, and R_{AB} is the distance between two A and B atoms inside the same cell.

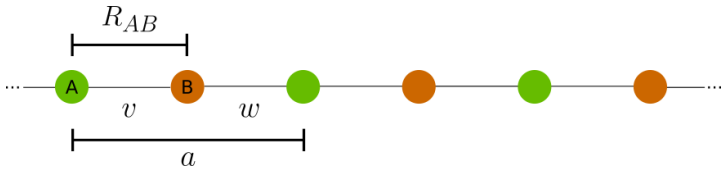


Fig. 2.2: Tight-binding infinite chain formed by one-dimensional sublattices A and B. The A nuclei are placed at multiple integers of the unit cell length, a, and the B nuclei at a distance R_{AB} to the right of A.

The orbitals satisfying Bloch's theorem can be written as linear combination of the atomic orbitals, that are used as a basis set,

$$\varphi_{ik}(r) = \frac{1}{\sqrt{N_l}} \sum_{R_i} e^{ikR_i} \chi_i(r - R_i)\,, \tag{2.53}$$

with N_l the number of lattice sites, and R_i iterates over the position of the atom i ($R_A = la$ or $R_B = R_{AB} + la$ for $l = 0, \pm 1, \ldots$) the lattice (one-dimensional) vector. Then, the crystal wavefunctions are

$$\phi_{nk}(r) = \sum_i c_{ik}^n \varphi_{ik}(r) = \frac{1}{\sqrt{N_l}} \sum_i c_{ik}^n \sum_{R_i} e^{ikR_i} \chi_i(r - R_i)\,, \tag{2.54}$$

where the coefficients c_{ik}^n must be chosen so that they satisfy

$$\sum_i \left(H_{ij} + \delta_{ij} E_{nk} \right) c_{jk}^n = 0\,, \tag{2.55}$$

with $H_{ij} = \langle \varphi_{ik} | \hat{H} | \varphi_{jk} \rangle$ the elements of the Hamiltonian in matrix form, $\mathbf{H}$. This equation has a solution if

$$|\mathbf{H} - E_{nk}\mathbf{I}| = 0\,, \tag{2.56}$$

with $\mathbf{I}$ the identity matrix. For a two-atom basis this means

$$E_{nk} = E_{\pm}(k) = \frac{1}{2}(H_{AA} + H_{BB}) \pm \sqrt{\frac{1}{4}(H_{AA} - H_{BB})^2 + |H_{AB}|^2}\,. \tag{2.57}$$

The matrix element for $i = j = A$ yields

$$H_{AA} = \langle \varphi_{Ak} | \hat{H} | \varphi_{Ak} \rangle = \epsilon_A + \sum_{\tau} e^{ik\tau} \int \chi^*(x) \hat{H} \chi(x-\tau)\, dx \tag{2.58}$$

where ϵ_A is the eigenvalue from eq. (2.49) and τ is the distance between any two A atoms. We will neglect any overlap between atoms centered at a distance $\tau \geq a$, restricting the sums to first neighbors only. This leads to $H_{AA} = \epsilon_A$. In an analogue manner, we arrive at $H_{BB} = \epsilon_B$. Because the atomic functions in both sublattices are the same, we will actually have $H_{AA} = H_{BB} = \epsilon$.

For the terms outside the diagonal we have

$$H_{AB} = \sum_{\tau} e^{ik\tau} \gamma(|\tau|)\,, \tag{2.59}$$

where now τ represents the distance between atoms A and B, and we have defined the *hopping* γ as a function of the distance,

$$\gamma(|\tau|) = \int \chi^*(x) \hat{H} \chi(x-\tau)\, dx\,. \tag{2.60}$$

The value of this overlap integral is usually taken as a parameter, or evaluated empirically. It is also possible to find empirical estimations for it that allow to scale it with respect to distances in an approximate manner, to deal with deformations in crystal, e.g. $\gamma(|\tau|) = Ae^{-\alpha|\tau|^2}/|\tau|^2$ used in silicon [74].

Taking only first neighbors in eq. (2.59), we can have $\tau = R_{AB}$ or $\tau = R_{AB} - a$, leading to two different hoppings,

$$v \equiv \gamma(|R_{AB}|)\,, \quad w \equiv \gamma(|R_{AB} - a|)\,. \tag{2.61}$$

Then, the matrix element becomes

$$H_{AB} = v\, e^{ikR_{AB}} + w\, e^{ik(R_{AB}-a)}\,. \tag{2.62}$$

Considering that $H_{BA} = H_{AB}^*$, we get an expression for the energies in (2.57):

$$E_{\pm}(k) = \epsilon \pm \sqrt{v^2 + w^2 + 2vw\cos(ka)}\,. \tag{2.63}$$

If the hopping parameters are the same, the model can represent an unperturbed hydrogen chain. Changing them, i.e. letting $v \neq w$, can be used to model other types of chains, namely those where we find alternating double bonds [77]. In this case, we will use different hoppings to represent changes in the interatomic distances, in order to evaluate how the energy of a hydrogen chain changes for different distances R_{AB}, leading to different hopping parameters v and w. Further, finding the coefficients using eq. (2.55) and writing the eigenfunctions $\phi_{nk}(r)$ allows us to analyze the spatial properties of the electrons distribution, permitting us to compute quantities as the electron density, kinetic energy density, electron localization function (see Section 4.1), among others; for this simple model system, and to understand how they change.

3

Conventional superconductivity

Depending on the origin of their properties, we can identify two classes of superconductors: conventional and unconventional. The first category corresponds to materials in which the vibrations of the lattice play an important role in the emergence of the phenomenon of superconductivity. This is evidenced by the isotope effect, where the critical temperature, T_c, changes with the magnitude of the mass of the ions, M, as $T_c \propto M^{-1/2}$. Contrary to unconventional superconductors, we understand the microscopic mechanism that takes place in conventional ones, and we are able to predict their properties with certain accuracy. The BCS theory was the first breakthrough in this direction – proposing that superconductivity is enabled by the formation of pairs of electrons that we call *Cooper pairs* – and we shall delve into in Section 3.1. Later on, Migdal and Eliashberg developed a deeper and more accurate description of the theory, but based on the same physical principles. Although we will not discuss their theory thoroughly, we do present the most modern and accurate expressions to compute the critical temperature stemming from it in Section 3.2. Then, in Section 3.3, we present a spatial description of the Cooper pair; and in Section 3.4 we describe the formalism of superconducting density functional theory, a variant of DFT (see Section 2.3) specially crafted for superconducting materials, leading to an *ab initio* description of the system.

3.1 Bardeen-Cooper-Schrieffer theory

In 1957, Bardeen, Cooper and Schrieffer published an article containing for the first time a microscopic theory for the physics of conventional superconductors [17]. This happened almost 50 years after the first observation of superconductivity, and the long awaited theory was awarded the Nobel prize in 1972. This BCS theory, named after the three authors, starts from a model Hamiltonian in which there is an effective attractive interaction between opposite-spin and opposite-momentum electrons. This interaction enables the formation of Cooper pairs around the Fermi energy, a bound state of two electrons which appears to be key for the superconducting transition. Although the origin of this effective interaction is not treated explicitly in BCS

theory, it was known to come from the coupling between electronic and ionic degrees of freedom, as evidenced by the isotope effect. A simple explanation of this is represented in the scheme in Figure 3.1, where vibration of the nuclei induce local excesses of positive charges around an electron, resulting in an effective attractive interaction between that and another electron.

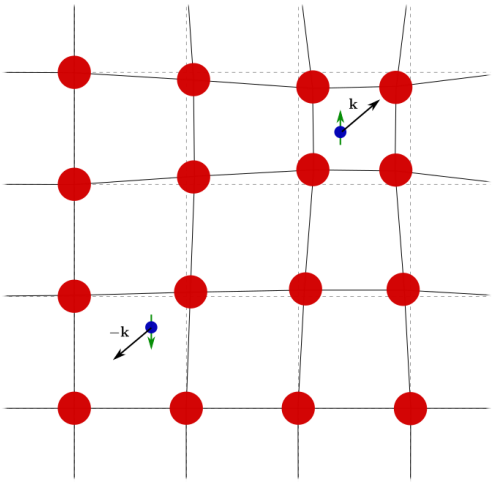


Fig. 3.1: Simple representation of the origin of the effective attractive interaction between opposite-momentum and spin electrons: the deformations of the ionic lattice (red) induce a local positive charge around one of the electrons (blue), which attracts the other electron to it.

The BCS effective Hamiltonian, in second quantization, is

$$\hat{H}_{BCS} = \sum_{\mathbf{k}\sigma} \xi_{\mathbf{k}} c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma} + \sum_{\mathbf{k}\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow} c_{\mathbf{k}'\uparrow} c_{-\mathbf{k}'\downarrow} , \tag{3.1}$$

where $c^{\dagger}_{\mathbf{k}\sigma}$ and $c_{\mathbf{k}\sigma}$ are the creation and annihilation operators of a σ electron in state $\mathbf{k}$, $\xi_{\mathbf{k}} = \epsilon_{\mathbf{k}} - \mu$ are the free-electron energies shifted by the chemical potential, and $V_{\mathbf{k}\mathbf{k}'}$ an effective potential for the electron-electron interaction [78]. We will solve this problem in a mean-field approach, making use of the approximation,

$$\langle c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow} c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \rangle \approx \langle c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow} \rangle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} + c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow} \langle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \rangle - \langle c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow} \rangle \langle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \rangle . \tag{3.2}$$

For reasons that will become evident later on, we define the *gap* function as

$$\Delta_{\mathbf{k}} = \frac{1}{N} \sum_{\mathbf{k}'} V_{\mathbf{k}\mathbf{k}'} \langle c_{-\mathbf{k}'\downarrow} c_{\mathbf{k}'\uparrow} \rangle . \tag{3.3}$$

We also introduce the Bogoliubov transformation

$$c_{\mathbf{k}\uparrow} = u_{\mathbf{k}}^{*}\gamma_{\mathbf{k}\uparrow} + v_{\mathbf{k}}\gamma_{-\mathbf{k}\downarrow}^{\dagger}\,, \tag{3.4}$$

$$c_{-\mathbf{k}\downarrow}^{\dagger} = u_{\mathbf{k}}\gamma_{-\mathbf{k}\downarrow}^{\dagger} - v_{\mathbf{k}}^{*}\gamma_{\mathbf{k}\uparrow}\,, \tag{3.5}$$

where the fermionic commutation relations of the electronic operators yield the constraint $|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1$. Using these definitions in the model Hamiltonian of eq. (3.1), we obtain that $\hat{H}_{BCS}$ can be diagonalized by $\gamma_{\mathbf{k}\sigma}^{\dagger}$ and $\gamma_{\mathbf{k}\sigma}$ when

$$|u_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 + \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}^2|}}\right); \quad |v_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 - \frac{\xi_{\mathbf{k}}}{\sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}^2|}}\right). \tag{3.6}$$

This leads to the final expression for the BCS Hamiltonian,

$$\hat{H}_{BCS} = \sum_{\mathbf{k}\sigma} E_{\mathbf{k}}\gamma_{\mathbf{k}\sigma}^{\dagger}\gamma_{\mathbf{k}\sigma} + E_0\,, \tag{3.7}$$

where we have defined the ground-state energy as

$$E_0 = \sum_{\mathbf{k}} \xi_{\mathbf{k}} - E_{\mathbf{k}} + \Delta_{\mathbf{k}}\langle c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger}\rangle\,, \tag{3.8}$$

and the excitation energies

$$E_{\mathbf{k}} = \sqrt{\xi_{\mathbf{k}}^2 + |\Delta_{\mathbf{k}}|^2}\,. \tag{3.9}$$

The expression in equation (3.9) justifies the name of the gap function, $\Delta_{\mathbf{k}}$. We see that the energy needed to excite the quasi-particles of the system, created by $\gamma_{\mathbf{k}\sigma}^{\dagger}$ and called *Bogoliubons*, is now finite, of at least $2|\Delta_{\mathbf{k}}|$.

The BCS ground-state wavefunction corresponds to that of energy E_0, where there are no excitations, i.e. $\gamma_{\mathbf{k}\sigma}|\psi_{BCS}\rangle = 0$. One can show that $|\psi_{BCS}\rangle$ can be written as a combination of Cooper pairs,

$$|\psi_{BCS}\rangle = \prod_{\mathbf{k}}\left(u_{\mathbf{k}} + v_{\mathbf{k}} c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger}\right)|0\rangle \tag{3.10}$$

where $|0\rangle$ is the state with no electrons (vacuum).

What is left now is to determine the form of the gap function, which we do using once again the transformation in equations (3.4) and (3.5), and considering that Bogoliubons follow the Fermi-Dirac distribution,

$$\left\langle \gamma_{\mathbf{k}\uparrow}^{\dagger}\gamma_{\mathbf{k}\uparrow}\right\rangle = \left\langle \gamma_{-\mathbf{k}\downarrow}^{\dagger}\gamma_{-\mathbf{k}\downarrow}\right\rangle = \frac{1}{e^{\beta E_{\mathbf{k}}} + 1}\,, \tag{3.11}$$

with $\beta = \frac{1}{k_B T}$, T the temperature and k_B the Boltzmann constant. This leads to an equation for the gap,

$$\Delta_{\mathbf{k}} = -\frac{1}{N}\sum_{\mathbf{k}'}\frac{V_{\mathbf{k}\mathbf{k}'}\Delta_{\mathbf{k}'}}{2E_{\mathbf{k}'}}\tanh\left(\frac{E_{\mathbf{k}'}}{2k_BT}\right). \tag{3.12}$$

At this point, some assumptions for the potential $V_{\mathbf{k}\mathbf{k}'}$ must be made. We will consider that it is zero except for $|\xi_{\mathbf{k}}|, |\xi_{\mathbf{k}'}| < \hbar\omega_D$, with ω_D the Debye frequency, where $V_{\mathbf{k}\mathbf{k}'} = -V_0$. Since ω_D is the highest possible phonon frequency, the attractive potential only applies to electrons in a similar energy window to that of phonons, around the Fermi level. Because the potential does not depend on $\mathbf{k}, \mathbf{k}'$, neither will the gap, $\Delta_{\mathbf{k}} = \Delta$, that we call an *s-wave* gap due to its symmetry.

With this, and considering the density of states at the Fermi level, $N(\epsilon_F)$, the equation for the gap becomes,

$$1 = V_0 N(\epsilon_F)\int_0^{\hbar\omega_D}\frac{d\epsilon}{\sqrt{\epsilon^2+\Delta^2}}\tanh\left(\frac{\sqrt{\epsilon^2+\Delta^2}}{2k_BT}\right). \tag{3.13}$$

Here, the gap is actually a function of the temperature. Taking $T \to 0$, we obtain the value of the gap at that temperature,

$$\Delta_0 = 2\hbar\omega_D e^{-\frac{1}{V_0 N(\epsilon_F)}}. \tag{3.14}$$

Meanwhile, the limit of $\Delta \to 0$ yields the critical temperature of the superconductor-metal transition,

$$T_c = \frac{2e^{\gamma_E}}{\pi}\frac{\hbar\omega_D}{k_B}e^{-\frac{1}{V_0 N(\epsilon_F)}}, \tag{3.15}$$

where $\gamma_E \approx 0.577$ is the Euler constant. One can approximate the temperature-dependent expression for the gap as [3, 79]

$$\Delta(T) = \Delta_0 \tanh\left(1.74\sqrt{\frac{T_c - T}{T}}\right). \tag{3.16}$$

Because in (3.15) we showed that $T_c \propto \omega_D \propto M^{-1/2}$, with M the mass of the atoms in the lattice, the BCS theory succeeds to explain the isotope effect. Further, it accounts for the experimental observation of the exponential behavior of the electronic specific heat at temperatures well below T_c and its discontinuity around the transition temperature [3]. Most importantly, from the BCS formalism one can recover the Meissner effect, which is a key property of a superconductor.

It is worth noting that BCS theory does present some disadvantages. The first one is that it is not *ab initio*, as the potential $V_{\mathbf{k}\mathbf{k}'}$ cannot be obtained from first principles. The origin of that interaction is not well established within the theory, even though it is known to be a result of electron-phonon interaction. An explicit dependence of the potential with respect to the phonon modes would thus be ideal. Moreover, this electron-electron interaction mediated by the lattice should not be instantaneous, as it is considered in BCS. Both of these considerations are made in more modern formulations of the microscopic theory of (conventional) superconductivity. Although it is not our aim to describe those approaches deeply, we will briefly mention them in the following section.

3.2 Modern superconducting theories and methods to estimate T_c

The Migdal-Eliashberg theory is a many-body perturbative approach to the problem of superconductivity [19, 20]. Both electron and phonon Green's functions are taken into account, leading to a more rigorous description of the origin of the Cooper pairs than in BCS. This formalism leads to a set of self-consistent equations that must be solved to find the form of the gap and the critical temperature, see Appendix A. To simplify the estimation of T_c, it is possible to derive an analytical expression for it, but some approximations must be made. Before delving into them, however, we must introduce some concepts of Migdal-Eliashberg theory.

In general, the contribution of the Coulomb interaction is taken as an averaged quantity around the Fermi energy, represented by the brackets $\langle\cdot\rangle_{\epsilon_F}$,

$$\mu_c = N(\epsilon_F)\langle V_{n\mathbf{k},m\mathbf{k}'}\rangle_{\epsilon_F}\,, \tag{3.17}$$

with $V_{n\mathbf{k},m\mathbf{k}'}$ the screened Coulomb interaction matrix element. Then, we define the electron-phonon coupling constant as,

$$\lambda = 2\int \frac{\alpha^2F(\omega)}{\omega^2}\,d\omega\,. \tag{3.18}$$

where $\alpha^2F(\omega)$ is the electron-phonon coupling spectral function, also known as the Eliashberg function,

$$\alpha^2F(\omega) = \frac{1}{N(E_F)}\sum_{\mathbf{k}\mathbf{q}\nu}|g_{mn\nu}(\mathbf{k},\mathbf{q})|^2\,\delta(\xi_{n\mathbf{k}})\delta(\xi_{m\mathbf{k}+\mathbf{q}})\delta(\omega-\omega_{\mathbf{q}\nu})\,. \tag{3.19}$$

With these definitions, and fitting to empirical results for the spectral function of Nb to improve accuracy, McMillan derived an analytical expression of T_c [80], that was later further improved by Allen and Dynes [21], leading to

$$T_c = \frac{\omega_{log}}{1.2}\exp\left[-\frac{1.04(1+\lambda)}{\lambda-\mu_c^*(1+0.62\lambda)}\right]. \tag{3.20}$$

where we have introduced the logarithmic-averaged frequency

$$\omega_{log} = \exp\left[\frac{2}{\lambda}\int\log(\omega)\frac{\alpha^2F(\omega)}{\omega}\,d\omega\right], \tag{3.21}$$

and the pseudo-Coulomb potential,

$$\mu_c^* = \frac{\mu_c}{1+\mu_c\log(\epsilon_F/\omega_c)}\,, \tag{3.22}$$

where ω_c is the characteristic phonon frequency. The value of μ_c^* is usually taken to be small, between 0.1 and 0.15.

This expression in equation (3.20) yields results very close to those obtained by solving Eliashberg's equations directly, which is why it is the most broadly used

equation to estimate the critical temperature. However, it starts to fail when the electron-phonon coupling, measured by λ, becomes too strong (e.g. $\lambda > 1.5$). In those cases, one can resort to the extended McMillan-Allen-Dynes formula [21],

$$T_c = f\frac{\omega_{log}}{1.2}\exp\left[-\frac{1.04(1+\lambda)}{\lambda-\mu_c^*(1+0.62\lambda)}\right], \tag{3.23}$$

where we have introduced a correcting factor,

$$f = \left[1+\left(\frac{\lambda}{2.46(1+3.8\mu_c^*)}\right)^{3/2}\right]^{1/3}\left[1+\frac{(\omega_2/\omega_{log}-1)\lambda^2}{\lambda^2+3.31(1+6.3\mu_c^*)(\omega_2/\omega_{log})}\right], \tag{3.24}$$

with

$$\omega_2 = \left(\frac{2}{\lambda}\int \omega\alpha^2F(\omega)\,d\omega\right)^{1/2}. \tag{3.25}$$

The temperatures obtained with equation (3.23) are in very good agreement with those obtained by solving the Eliashberg equations, even for large electron-phonon coupling.

3.3 Singlet-pair wavefunction

One of the most important – and most intriguing – properties of a superconductor is its ability to carry a supercurrent, where electrons can mantain their energy by avoiding any scattering with the lattice formed by the nuclei. This phenomenon can be compared to the superfluidity of systems like liquid ^{4}He, which flows without loosing any energy due to its null viscosity [81]. The key property of such systems is its *long-range phase coherence*, which means that the density matrix can be written as the product

$$\langle\psi^\dagger(\mathbf{r}')\psi(\mathbf{r})\rangle = \Psi^*(\mathbf{r}')\Psi(\mathbf{r}), \tag{3.26}$$

where $\psi^\dagger(\mathbf{r})$ ($\psi(\mathbf{r})$) is the creation (annihilation) field operator for electrons. The function $\Psi(\mathbf{r})$ is an effective wavefunction for the system and it gives a macroscopic description of it [82].

Let us think of a particular case of bosons forming a Bose-Einstein condensate. If $a_{\mathbf{k}}^\dagger$ and $a_{\mathbf{k}}$ are the operators creating and destroying bosons in the plane-wave state of wavevector $\mathbf{k}$, then the density matrix becomes

$$\langle\psi^\dagger(\mathbf{r}')\psi(\mathbf{r})\rangle = \frac{1}{\Omega}\left\langle\sum_{\mathbf{k}'}a_{\mathbf{k}'}^\dagger e^{-i\mathbf{k}'\cdot\mathbf{r}'}\sum_{\mathbf{k}}a_{\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{r}}\right\rangle, \tag{3.27}$$

$$= \frac{1}{\Omega}\sum_{\mathbf{k}}n_{\mathbf{k}}e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}, \tag{3.28}$$

where Ω is the volume of the lattice in k-space (see eq. (2.18), in second quantization form), and $n_{\mathbf{k}} = a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}}$ the occupation number of the state $\mathbf{k}$. Because in the case of the Bose-Einstein condensate all particles live in the same state, say $\mathbf{k} = \mathbf{s}$, only one term of the sum survives and the macroscopic wavefunction becomes

$$\Psi(\mathbf{r}) = \sqrt{\frac{n_s}{\Omega}}\, e^{\frac{i\mathbf{s}\cdot\mathbf{r}}{2}} . \tag{3.29}$$

The density matrix resulting from this macroscopic wavefunction yields finite values even when $\mathbf{r}'$ and $\mathbf{r}$ are far apart. This long-range behavior is a key feature of a superfluid.

In a more general case, we consider particles that are correlated, and they do not all fall into the same state purely for statistical reasons as they did in the Bose-Einstein condensate. There might be, however, many N-particle states for slightly different N values, with very similar energies. Let us consider such a superfluid state, $|N\rangle$. Then, we have

$$\langle N|\psi^{\dagger}(\mathbf{r}')\psi(\mathbf{r})|N\rangle = \sum_i \langle N|\psi^{\dagger}(\mathbf{r}')|i\rangle\langle i|\psi(\mathbf{r})|N\rangle \tag{3.30}$$

where the sum is over intermediate states of the superfluid [82]. We can find an effective function $\Psi(\mathbf{r})$ if one of the terms in this sum is dominant, which will necessarily contain $N-1$ particles,

$$\Psi(\mathbf{r}) = \langle N-1|\psi(\mathbf{r})|N\rangle . \tag{3.31}$$

The expression in eq. (3.31) gives to this phenomenon the name of *off-diagonal long-range order*.

The phase of the superfluid state $|N\rangle$ is indeterminate, but we can fix it to be equal to 0 by defining

$$|\tilde{0}\rangle = \sum_N \alpha_N |N\rangle , \tag{3.32}$$

and we can show that there are certain α_N coefficients for which the effective wavefunction can be written as

$$\Psi(\mathbf{r}) = \langle\tilde{0}|\psi(\mathbf{r})|\tilde{0}\rangle , \tag{3.33}$$

in which case we have a diagonal order. The price to pay is that now we do not have a determinate number of particles. This is exactly the case of the BCS wavefunction in (3.10). The effective wavefunction for a singlet-pair (SP) can thus be written as

$$\Psi_{SP}(\mathbf{r}_2,\mathbf{r}_1) = \left\langle \Psi_{BCS} \left| \psi_{\downarrow}(\mathbf{r}_2)\psi_{\uparrow}(\mathbf{r}_1) \right| \Psi_{BCS} \right\rangle , \tag{3.34}$$

$$= \frac{1}{\Omega}\sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{r}_1-\mathbf{r}_2)} , \tag{3.35}$$

which can be obtained considering $\psi_\sigma(\mathbf{r}) = \sum_{\mathbf{k}} c_{\mathbf{k}\sigma} e^{i\mathbf{k}\cdot\mathbf{r}}$ and the Bogoliubov transformation of equations (3.4) and (3.5). $\Psi_{SP}(\mathbf{r})$ is also called the Gor'kov function, and it is actually the order parameter of the superconducting transition, becoming zero above the critical temperature, T_c.

Kadin used this definition to study the spatial structure of the Cooper pair [83]. Defining $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ and using the definitions of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ from BCS, he wrote

$$\Psi_{SP}(\mathbf{r}) \propto \sum_{\mathbf{k}} \frac{\cos(\mathbf{k}\cdot\mathbf{r})}{\sqrt{\xi_{\mathbf{k}}^2 + \Delta^2}}. \tag{3.36}$$

Doing an approximation, only the radial part of the momentum is included, i.e. $\mathbf{k}\cdot\mathbf{r} \approx kr$. This lets us change from a sum over $\mathbf{k}$ to an integral over energies, $\sum_{\mathbf{k}} \to \int_0^{\hbar\omega_D} N(\epsilon) d\epsilon$, where we have included that only the terms near the Fermi energy contribute to the integral. Here, $N(\epsilon)$ is the density of states at ϵ, and we will approximate it by $N(\epsilon_F)$, i.e. its value at the Fermi energy. We thus obtain

$$\Psi_{SP}(r) \propto N(\epsilon_F) \int_0^{\hbar\omega_D} \frac{\cos(kr)}{\sqrt{\epsilon^2 + \Delta^2}} d\epsilon. \tag{3.37}$$

Then, doing a Taylor expansion to first order around $k - k_F$, we have that

$$\epsilon = \frac{\hbar^2}{2m}(k-k_F)^2 \approx \frac{\hbar^2 k_F}{m_e}(k-k_F) = \pi\xi_0\Delta(k-k_F), \tag{3.38}$$

where we have defined the BCS *coherence length* $\xi_0 = \hbar v_F/\pi\Delta$, with $v_F = \hbar k_F/m_e$. Using this and defining the new integration variable $\epsilon' = \epsilon/\Delta$, the wavefunction becomes

$$\Psi_{SP}(r) \propto \int_0^{\frac{\hbar\omega_D}{\Delta}} \frac{\cos\left(k_F r + \frac{\epsilon' r}{\pi\xi_0}\right)}{\sqrt{\epsilon'^2 + 1}} d\epsilon', \tag{3.39}$$

$$\approx \cos(k_F r) \int_0^{\frac{\hbar\omega_D}{\Delta}} \frac{\cos\left(\frac{\epsilon' r}{\pi\xi_0}\right)}{\sqrt{\epsilon'^2 + 1}} d\epsilon'. \tag{3.40}$$

Finally, the singlet-pair wavefunction derived by Kadin depends on the zero-th order modified Bessel function,

$$\Psi_{SP}(r) \propto \cos(k_F r) K_0\left(\frac{r}{\pi\xi_0}\right). \tag{3.41}$$

Equation (3.41) thus gives a macroscopic spatial description of a Cooper pair as a function of the distance r between the electrons.

3.4 Superconducting Density Functional Theory

The widespread use of DFT for electronic structure calculations, due to its great compromise between accuracy and computational time, has served as a motivation

to extend it to a wider variety of systems. The case of superconductors is particular in the sense that it cannot be solved in a perturbative fashion. In fact, this is so because in these systems the phase symmetry is broken, which implies that the number of particles will not be conserved. Another key property of superconductors is that the coupling between electrons and phonons becomes important, and that it is in fact this what permits the superconductivity. The ionic displacements around their equilibrium positions must therefore be taken into account in the Hamiltonian. This is why, in Superconducting Density Functional Theory (SCDFT) [84, 85], we must consider the nuclear degrees of freedom, as in equation (2.2). The purely ionic part is written in terms of the ionic field operators, $\Phi(\mathbf{R})$, and contains the kinetic and interaction terms,

$$\hat{H}_i = -\int \Phi^\dagger(\mathbf{R})\frac{\nabla^2}{2M}\Phi(\mathbf{R})d\mathbf{R} + \frac{1}{2}\int \Phi^\dagger(\mathbf{R})\Phi^\dagger(\mathbf{R}')\frac{Z}{|\mathbf{R}-\mathbf{R}'|}\Phi(\mathbf{R}')\Phi(\mathbf{R})\,d\mathbf{R}\,d\mathbf{R}'\,. \tag{3.42}$$

Meanwhile, the interaction between electrons and ions is

$$\hat{H}_{ie} = -\frac{1}{2}\sum_\sigma \int \psi^\dagger_\sigma(\mathbf{r})\Phi^\dagger(\mathbf{R})\frac{Z}{|\mathbf{R}-\mathbf{r}|}\Phi(\mathbf{R})\psi_\sigma(\mathbf{r})\,d\mathbf{R}\,d\mathbf{r}\,. \tag{3.43}$$

For the electronic part the Hamiltonian takes the form,

$$\hat{H}_e = \sum_\sigma \int \psi^\dagger_\sigma(\mathbf{r})\left[-\frac{\nabla^2}{2}-\mu\right]\psi_\sigma(\mathbf{r})\,d\mathbf{r} + \frac{1}{2}\sum_{\sigma\sigma'}\int \psi^\dagger_\sigma(\mathbf{r})\psi^\dagger_{\sigma'}(\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_\sigma(\mathbf{r})\psi_{\sigma'}(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}'\,. \tag{3.44}$$

Finally, we must consider three different external potentials: one that couples to electrons, $v_{ext}(\mathbf{r})$, one that couples to ions, $W_{ext}(\mathbf{R})$, and an anomalous potential, $\Delta_{ext}(\mathbf{r},\mathbf{r}')$, that is responsible for the symmetry breaking, allowing Cooper pairs to tunnel in and out of the system,

$$\hat{H}_{\Delta_{ext}} = \int \Delta^*_{ext}(\mathbf{r},\mathbf{r}')\psi_\uparrow(\mathbf{r})\psi_\downarrow(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}' + h.c. \tag{3.45}$$

If we let this quantity go to zero, the Hamiltonian converges to a non-superconducting one, i.e. that of a normal state system. In this way, SCDFT allows to *turn on and off* superconductivity, and to compare the system's properties when it is in the normal or in the superconducting state.

As in usual DFTs, there will be a Hohenberg-Kohn theorem establishing a one-to-one mapping between, in this case, those three external potentials and corresponding densities. The inclusion of the anomalous potential, leading to a fluctuation of the particle number, forces us to work in the grand-canonical ensemble. The expectation values are thus taken considering the grand-canonical density matrix,

$$\hat{\varrho}_0 = \frac{e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathrm{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}]}\,, \tag{3.46}$$

where $\hat{N}$ is the number operator. Then, the electron density is

$$\rho^{SC}(\mathbf{r}) = \left\langle \sum_{\sigma} \psi^{\dagger}_{\sigma}(\mathbf{r})\psi_{\sigma}(\mathbf{r}) \right\rangle = \mathrm{Tr}\left[\hat{\varrho}_0 \sum_{\sigma} \psi^{\dagger}_{\sigma}(\mathbf{r})\psi_{\sigma}(\mathbf{r}) \right], \tag{3.47}$$

where we have used the *SC* superscript to differentiate it from the normal state electron density.

In a similar way, we define the ionic density,

$$\Gamma(\{\mathbf{R}_i\}) = \mathrm{Tr}\left[\hat{\varrho}_0 \prod_j \Phi^{\dagger}(\mathbf{R}_j)\Phi(\mathbf{R}_j) \right]. \tag{3.48}$$

Just as did the ionic external potential, this density depends on *all* ionic coordinates, thus corresponding to an N-particle density matrix. This will be convenient when we start treating the ionic deviations from equilibrium positions as collective vibrations (phonons). Finally, we define the anomalous density,

$$\chi(\mathbf{r},\mathbf{r}') = \mathrm{Tr}\left[\hat{\varrho}_0 \psi_{\uparrow}(\mathbf{r})\psi_{\downarrow}(\mathbf{r}') \right], \tag{3.49}$$

which is a two-body object and the order parameter of the transition. Notice that $\chi(\mathbf{r},\mathbf{r}')$ is analogue to the singlet-pair function of equation (3.34).

A second Hohenberg-Kohn-like theorem proves that the grand-canonical potential follows a variational principle as a function of the densities in (3.47), (3.48) and (3.49). This potential is defined as

$$\begin{aligned}
\Omega\left[\rho^{SC},\chi,\Gamma\right] =& T_e\left[\rho^{SC},\chi,\Gamma\right] + T_n\left[\rho^{SC},\chi,\Gamma\right] - \frac{1}{\beta}S\left[\rho^{SC},\chi,\Gamma\right] + V_e e\left[\rho^{SC},\chi,\Gamma\right] + V_n n\left[\rho^{SC},\chi,\Gamma\right] \\
&+ \int v_{ext}(\mathbf{r})\rho^{SC}(\mathbf{r})\,d\mathbf{r} + \int W_{ext}(\{\mathbf{R}_i\})\Gamma(\{\mathbf{R}_i\}) \prod_j d\mathbf{R}_j \\
&+ \int \Delta^{*}_{ext}(\mathbf{r},\mathbf{r}')\chi(\mathbf{r},\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}', \qquad (3.50) \\
\equiv& F\left[\rho^{SC},\chi,\Gamma\right] + \int v_{ext}(\mathbf{r})\rho^{SC}(\mathbf{r})\,d\mathbf{r} + \int W_{ext}(\{\mathbf{R}_i\})\Gamma(\{\mathbf{R}_i\}) \prod_j d\mathbf{R}_j \\
&+ \int \Delta^{*}_{ext}(\mathbf{r},\mathbf{r}')\chi(\mathbf{r},\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}', \qquad (3.51)
\end{aligned}$$

where T corresponds to kinetic energy functionals and S to the entropy. We have also introduced the system-independent universal functional, $F[\rho,\chi,\Gamma]$, analogue to the DFT functional of eq. (2.6). Then, the variational principle ensures $\Omega[\rho_0^{SC},\chi_0,\Gamma_0] < \Omega[\rho^{SC},\chi,\Gamma]$ for any ρ^{SC},χ,Γ different from the ground state ones, $\rho_0^{SC},\chi_0,\Gamma_0$. The proof can be found elsewhere [86].

As usual, we do not know how $F[\rho^{SC},\chi,\Gamma]$ looks in terms of the densities, so we will have to introduce a reference (Kohn-Sham) system with non-interacting electrons whose grand-canonical potential is minimized by the same densities. The resulting equation for the ions is analogue to (2.12), but considering the corresponding

ionic potential. Meanwhile, the Kohn-Sham electronic equations yield a Hamiltonian of the form

$$\hat{H}_s = \sum_\sigma \int d\mathbf{r}\, \psi^\dagger_\sigma(\mathbf{r}) \left[-\frac{\nabla^2}{2} + v_{KS}(\mathbf{r}) - \mu \right] \psi_\sigma(\mathbf{r}) + \int d\mathbf{r} d\mathbf{r}' \left[\Delta^*_{KS}(\mathbf{r},\mathbf{r}') \psi_\uparrow(\mathbf{r}) \psi_\downarrow(\mathbf{r}') + h.c. \right], \tag{3.52}$$

where $v_{KS}(\mathbf{r})$ is the usual electronic KS potential (see eq. (2.13)) and the effective anomalous potential takes the form

$$\Delta_{KS}(\mathbf{r},\mathbf{r}') = \Delta_{ext}(\mathbf{r},\mathbf{r}') + \frac{\delta F_{xc}[\rho^{SC},\chi,\Gamma]}{\delta\chi}, \tag{3.53}$$

with $F_{xc}[\rho^{SC},\chi,\Gamma]$ the exchange-correlation contribution to the universal functional (see ref. 84).

The Hamiltonian in eq. (3.52) can be diagonalized by introducing the Bogoliubov transformation [87], just as we did in BCS (see eqs. (3.4) and (3.5)),

$$\psi_\sigma(\mathbf{r}) = \sum_i \left[u_i(\mathbf{r}) \gamma_{i\sigma} - \mathrm{sgn}(\sigma) v^*_i(\mathbf{r}) \gamma^\dagger_{i,-\sigma} \right], \tag{3.54}$$

$$\psi^\dagger_\sigma(\mathbf{r}) = \sum_i \left[u^*_i(\mathbf{r}) \gamma^\dagger_{i\sigma} - \mathrm{sgn}(\sigma) v_i(\mathbf{r}) \gamma_{i,-\sigma} \right]. \tag{3.55}$$

Using it in (3.52), we arrive at set of Bogoliubov-deGennes (BdG) self-consistent equations,

$$\left[-\frac{\nabla^2}{2} + v_{KS}(\mathbf{r}) - \mu \right] u_i(\mathbf{r}) + \int \Delta_{KS}(\mathbf{r},\mathbf{r}') v_i(\mathbf{r}')\, d\mathbf{r}' = E_i u_i(\mathbf{r}), \tag{3.56}$$

$$-\left[-\frac{\nabla^2}{2} + v_{KS}(\mathbf{r}) - \mu \right] v_i(\mathbf{r}) + \int \Delta_{KS}(\mathbf{r},\mathbf{r}') u_i(\mathbf{r}')\, d\mathbf{r}' = E_i v_i(\mathbf{r}), \tag{3.57}$$

which shall determine $u_i(\mathbf{r})$ and $v_i(\mathbf{r})$.

The transformation in eqs. (3.54) and (3.55) permits to rewrite the electron density of eq. (3.47) as

$$\rho^{SC}(\mathbf{r}) = 2 \sum_i \left[|u_i(\mathbf{r})|^2 f(E_i) + |v_i(\mathbf{r})|^2 f(-E_i) \right]. \tag{3.58}$$

with $f(E_i) = (1 + e^{\beta E_i})^{-1}$ the Fermi-Dirac distribution. However, the form of $u_i(\mathbf{r})$ and $v_i(\mathbf{r})$ is still unknown and to solve equations (3.56) and (3.57) we must resort to some approximations.

Let $\varphi_{n\mathbf{k}}(\mathbf{r})$ be the solutions of equation (3.56) when $\Delta_{KS}(\mathbf{r},\mathbf{r}') = 0$, and $\xi_{n\mathbf{k}}$ the corresponding eigenvalues. Because we keep $\chi(\mathbf{r},\mathbf{r}') \neq 0$ in $v_{KS}[\rho,\chi,\Gamma]$, $\{\varphi_{n\mathbf{k}}(\mathbf{r})\}$ are not the usual KS orbitals, but they will be very similar to them. We will use them as a basis set,

$$u_i(\mathbf{r}) = \sum_{n\mathbf{k}} u_{i,n\mathbf{k}} \varphi_{n\mathbf{k}}(\mathbf{r}), \quad v_i(\mathbf{r}) = \sum_{n\mathbf{k}} v_{i,n\mathbf{k}} \varphi_{n\mathbf{k}}(\mathbf{r}). \tag{3.59}$$

Now, to simplify this, we introduce an approximation assuming that the inclusion of $\Delta_{KS}(\mathbf{r},\mathbf{r}')$ does not induce a hybridization of bands, i.e.

$$u_i(\mathbf{r}) = u_{n\mathbf{k}}\varphi_{n\mathbf{k}}(\mathbf{r}), \quad v_i(\mathbf{r}) = v_{n\mathbf{k}}\varphi_{n\mathbf{k}}(\mathbf{r}). \tag{3.60}$$

With this in the BdG equations (3.56) and (3.57) we obtain

$$u_{n\mathbf{k}} = \frac{1}{\sqrt{2}}\,\mathrm{sgn}(E_{n\mathbf{k}})e^{\phi_{n\mathbf{k}}}\sqrt{1+\frac{\xi_{n\mathbf{k}}}{|E_{n\mathbf{k}}|}}, \tag{3.61}$$

$$v_{n\mathbf{k}} = \frac{1}{\sqrt{2}}\sqrt{1-\frac{\xi_{n\mathbf{k}}}{|E_{n\mathbf{k}}|}}, \tag{3.62}$$

with $e^{\phi_{n\mathbf{k}}} = \frac{\Delta_{KS}(n\mathbf{k})}{|\Delta_{KS}(n\mathbf{k})|}$ and $E_{n\mathbf{k}} = \pm\sqrt{\xi^2_{n\mathbf{k}} + \Delta_{KS}(n\mathbf{k})^2}$. These results are reminiscent of those obtained in BCS theory, see eqs. (3.6) and (3.9). Under the previous approximations, the electron density in the superconducting state becomes

$$\rho^{SC}(\mathbf{r}) = \sum_{n\mathbf{k}}\left[1-\frac{\xi_{n\mathbf{k}}}{|E_{n\mathbf{k}}|}\tanh\left(\frac{\beta|E_{n\mathbf{k}}|}{2}\right)\right]|\varphi_{n\mathbf{k}}(\mathbf{r})|^2\,. \tag{3.63}$$

Notice that in the normal state limit, $\Delta_s(n\mathbf{k}) \to 0$, we recover the normal state density,

$$\rho^{\mathrm{NS}}(\mathbf{r}) = 2\sum_{n\mathbf{k}} f(\xi_{n\mathbf{k}})|\varphi_{n\mathbf{k}}(\mathbf{r})|^2\,. \tag{3.64}$$

In the same way, one can compute the superconducting order parameter,

$$\chi(\mathbf{r}_1,\mathbf{r}_2) = [\varrho_0\psi_\uparrow(\mathbf{r}_1)\psi_\downarrow(\mathbf{r}_2)] = \frac{1}{2}\sum_{n\mathbf{k}}\frac{\Delta_{KS}(n\mathbf{k})}{|E_{n\mathbf{k}}|}\tanh\left(\frac{\beta|E_{n\mathbf{k}}|}{2}\right)\varphi^*_{n\mathbf{k}}(\mathbf{r}_1)\varphi_{n\mathbf{k}}(\mathbf{r}_2). \tag{3.65}$$

This quantity holds information about the spatial distribution of Cooper pairs. Although it is not exactly a density, one can see from its definition in eq. (3.49) that is represents some sort of averaged Cooper-pair wavefunction.

Evaluating the spatial functions in (3.63) and (3.65) requires an explicit expression for the gap, $\Delta_{KS}(n\mathbf{k})$. Notice that we can rewrite the exchange-correlation contribution in eq. (3.53) as

$$\Delta_{xc} = \frac{\delta F_{xc}\left[\rho^{SC},\chi[\Delta_{KS},\rho^{SC},\Gamma],\Gamma\right]}{\delta\chi}. \tag{3.66}$$

This is usually called the gap equation of SCDFT , and it must be solved self-consistently. Notice that the dependence on Γ means that the ionic degrees of freedom play a role in the electronic equations.

In practice, the gap will be determined by bridging SCDFT with the theory of Eliashberg. There, the ionic contribution to the Hamiltonian considers harmonic oscillations around equilibrium positions and electron-phonon coupling. A detailed description of the derivation is out of the scope of this thesis, but we provide one of the

simplest examples of proposed functionals for illustration. For the LM2005 functional [88], the gap equation reads

$$\Delta_{KS}(n\mathbf{k}) = \mathcal{Z}(n\mathbf{k})\Delta_{xc}(n\mathbf{k}) + \frac{1}{2}\sum_{n'\mathbf{k}'}\mathcal{K}(n\mathbf{k},n'\mathbf{k}')\frac{\tanh\left(\frac{\beta}{2}E_{n'\mathbf{k}'}\right)}{E_{n'\mathbf{k}'}}\Delta_{xc}(n'\mathbf{k}'), \tag{3.67}$$

where $\mathcal{Z}(n\mathbf{k})$ and $\mathcal{K}(n\mathbf{k},n'\mathbf{k}')$ are two kernels that depend explicitly on the electron-phonon matrix elements $g_{mnj}(\mathbf{k},\mathbf{q})$ of equation (2.37). The gap equation for this functional resembles that of the BCS theory, see eq. (3.12), but it has the advantage of having been found from first principles.

An advantage of SCDFT is the possibility of doing a real-space analysis of the descriptors that characterize the superconducting state. Their study in phonon-mediated superconductors, such as MgB_2, CaC_6 and hole-doped graphene, has been performed in ref. [89], showing a clear correlation between chemical bonding patterns and the spatial distribution of both the order-parameter and the superconducting gap.

A clear drawback of SCDFT is that, although it has been implemented by some groups and calculations on real systems have been done, the code is not available to the public at this moment. Moreover, the mathematical complexity of the formalism hinders the possibility of a self-made code.

4

Electronic localization and delocalization

The question of where electrons are and how they are moving in a molecule or in a periodic system is not one that we can actually answer, at least not precisely. The uncertainty principle of Heisenberg tells us that it is in fact impossible to know simultaneously the position and the momentum of a quantum particle with infinite precision. However, we do have access to different quantum observables that are evaluated considering the system's probability density, that can be obtained from its wavefunction, Ψ. In this way, we can learn about the behavior of the electrons by looking at averaged quantities, like the electronic density, $\rho(\mathbf{r})$, or the kinetic energy density, $\tau(\mathbf{r})$, for example. As a consequence of the averaged nature of these quantities, instead of asking where electrons are and what is their momentum, it becomes more useful to talk about regions where electrons are more localized, delocalized, and how their populations fluctuate among those regions.

Localization and delocalization of electrons in real space is certainly a topic of interest for quantum chemists and it has been studied from many different angles. From a macroscopic point of view, describing the degree of localization within a crystal lattice can help determining the transport properties of the material. In fact, an alternative definition to what an insulator and a conductor are, that is independent from the concept of a conduction gap, is based on the localized nature of the electrons in an insulating material (leading to characteristic properties such as the ability to have a non-zero polarization), and the delocalized behavior of electrons in a conductor [90, 91].

On the other hand, determining where electrons are localized within a molecule or in a crystal lattice reinforces the idea of chemical structure. For Bader, a descriptor like the charge density can actually bridge the gap between quantum mechanics and the historical picture of a chemical atom inside a molecule [92, 93]. This can be done, in the context of the Quantum Theory of Atoms in Molecules (QTAIM), by analyzing the gradient field $\nabla\rho(\mathbf{r})$ and identifying atomic regions separated by isosurfaces of zero charge flux. The maximum within such a region reveals the position of the core, and the gradient paths connecting two maxima and passing through a zero-flux isosurface and a saddle point are identified with the chemical bond. Descriptors other than the electron density can also be used to separate the space into regions

that we can identify with well-known chemical concepts. This is done through the study of the *topology* of the given spatial function, allowing us to get insights on the chemistry of the system by analyzing its gradient field, finding and classifying its critical points, and dividing the space into *basins*, as we shall see on Section 4.2.

The Electron Localization Function (ELF) is a scalar field widely used in quantum chemistry. Similar to $\rho(\mathbf{r})$ in the context of QTAIM, the study of the topology of the ELF allows to recognize the structure of a molecule or a solid [94]. Moreover, it is possible to identify different bonding patterns using the ELF and, for example, infer the metallic nature of a system, therefore learning something about its conduction properties. This function is the main tool that we use to study the localization of electrons throughout this thesis, and we will begin this chapter by introducing its definition in Section 4.1. Then, in Section 4.2, we aim to introduce key concepts in quantum chemical topology, employing the ELF as an illustrative example. There, we also emphasize the ELF's significant potential for elucidating various chemical properties inherent to the system under examination. Finally, in Section 4.3 we introduce localization and delocalization indices, a very useful instrument that can further contribute to characterize the spatial behavior of electrons in a material.

4.1 Electron localization function

The electron localization function was first introduced by Edgecome and Becke to identify regions of localized same-spin electron pairs, or groups of them, in atomic and molecular systems [95]. It is based on the same-spin pair probability as approximated in Hartree-Fock,

$$P_2^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2) = \rho_\sigma(\mathbf{r}_1)\rho_\sigma(\mathbf{r}_2) - \left|\rho_1^\sigma(\mathbf{r}_1,\mathbf{r}_2)\right|^2 , \tag{4.1}$$

which is the probability of finding simultaneously one electron with spin σ at $\mathbf{r}_1$, and another σ electron at $\mathbf{r}_2$. Here, $\rho_1^\sigma(\mathbf{r}_1,\mathbf{r}_2)$ is the spin-resolved one-electron reduced density matrix (1-RDM), and $\rho^\sigma(\mathbf{r}_1)$ its diagonal, which corresponds to the electron density for σ-spin. In eq. (4.1), we recognize the first term as the product of the individual probabilities of finding each electron at the said positions, to which we need to substract a term arising from the correlation introduced in Hartree-Fock by the Pauli principle (exchange interaction).

Assuming that there is one σ electron at $\mathbf{r}_1$, the probability of finding another electron with the same spin at $\mathbf{r}_2$ can be expressed as

$$P_{cond}^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2) = \frac{P_2^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2)}{\rho_\sigma(\mathbf{r}_1)} = \rho_\sigma(\mathbf{r}_2) - \frac{\left|\rho_1^\sigma(\mathbf{r}_1,\mathbf{r}_2)\right|^2}{\rho_\sigma(\mathbf{r}_1)} , \tag{4.2}$$

which is called the conditional same-spin pair probability. Fixing one of the electrons permits to study the behavior of this probability when $\mathbf{r}_2 \to \mathbf{r}_1$. Changing to the spherically averaged version of $P_{cond}^{\sigma\sigma}$, that depends on the coordinates $(\mathbf{r}, s)$, where $\mathbf{r}$ is the reference point and s a distance from it, and doing a Taylor expansion, one can show that

$$P^{\sigma\sigma}_{cond}(\mathbf{r}, s) = \frac{1}{3}\left[\sum_i^{\sigma} |\nabla\phi_i(\mathbf{r})|^2 - \frac{1}{4}\frac{|\nabla\rho_\sigma(\mathbf{r})|}{\rho_\sigma(\mathbf{r})}\right] s^2 + \dots \tag{4.3}$$

Here, $\phi_i(\mathbf{r})$ are the HF orbitals, and the sum $\sum_i^\sigma$ means that only the orbitals containing electrons of spin σ are being considered.

The term in brackets in eq. (4.3),

$$D_\sigma(\mathbf{r}) = \sum_i^{\sigma} |\nabla\phi_i(\mathbf{r})|^2 - \frac{1}{4}\frac{|\nabla\rho_\sigma(\mathbf{r})|}{\rho_\sigma(\mathbf{r})}, \tag{4.4}$$

can be recognized as a measure of localization, being the leading term for small distances s between the electrons. This is, when D_σ is small, the probability of finding a σ electron very close to the reference one is also small. This means that the reference electron is very localized, and so is the Fermi hole that comes with it, not allowing a same-spin electron to come near. The opposite is true when the reference electron is delocalized: it is more likely that the other electron comes close to the reference position, and D_σ will be high.

All of this has been introduced for electrons of the same spin. However, if the reference electron is very localized, it is likely than an opposite electron is also very localized in that region, as it will also be avoiding its own same-spin electron, thus forming a singlet pair. This is a consequence of the exchange introduced in the same-spin probability in eq. (4.1).

The function $D_\sigma(\mathbf{r})$ being opposite to localization, we introduce the ELF,

$$\eta(\mathbf{r}) = \left[1 + \left(\frac{D_\sigma(\mathbf{r})}{D^0_\sigma(\mathbf{r})}\right)^2\right]^{-1}, \tag{4.5}$$

where

$$D^0_\sigma(\mathbf{r}) = \frac{3}{5}(6\pi^2)^{2/3}\rho_\sigma(\mathbf{r})^{5/3} \tag{4.6}$$

is the term D_σ as evaluated for a uniform electron gas. This normalization allows to compare the values of the kernel $\chi_\sigma(\mathbf{r}) = D_\sigma(\mathbf{r})/D^0_\sigma(\mathbf{r})$ of different systems. Further, the Lorentzian transformation applied on that kernel in the definition of $\eta(\mathbf{r})$ in (4.5) restricts the ELF to values between 0 and 1. In this way, the ELF is high in the regions of high localization ($\eta \to 1$), and it conserves the topology of $\chi_\sigma(\mathbf{r})$. This, as we shall see, will be very important in the description of electron localization.

Although the ELF was first introduced in the context of the HF approximation, a close inspection of eq. (4.4) allows us to identify the first and second terms as the kinetic energy density of the system, $\tau(\mathbf{r})$, and its form in the von-Weizsacker approximation, $\tau_{vW}(\mathbf{r})$ [96]. In fact, in a closed-shell system we have $\rho(\mathbf{r}) = 2\rho_\sigma(\mathbf{r})$, and therefore we can define the spinless quantity

$$D(\mathbf{r}) = \frac{1}{2}\sum_i |\nabla\phi_i(\mathbf{r})|^2 - \frac{1}{8}\frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}, \tag{4.7}$$

$$= \tau(\mathbf{r}) - \tau_{vW}(\mathbf{r}), \tag{4.8}$$

and

$$\eta(\mathbf{r}) = \left[1 + \left(\frac{D(\mathbf{r})}{D_0(\mathbf{r})}\right)^2\right]^{-1}, \tag{4.9}$$

with $D_0(\mathbf{r}) = \frac{3}{10}(3\pi^2)^{2/3}\rho(\mathbf{r})^{5/3}$. This form of the ELF written in terms of kinetic energy densities was first introduced by Savin and coworkers [97], and allows to compute it beyond the HF approximation. Further, it introduces a new interpretation: because the von-Weizsacker kinetic energy is exact for a bosonic system of the same density $\rho(\mathbf{r})$, the term $\tau(\mathbf{r}) - \tau_{vW}(\mathbf{r})$ is a local measure of the excess kinetic energy due to the fermionic nature of the electrons, or what we call the *Pauli kinetic energy*. If this is high, it means that electron pairs are delocalized in that region, and the ELF will be small. If the kinetic energy density is not locally increased as an effect of the exclusion principle, in that case we say that electrons are localized, which will be reflected on a high value of the ELF.

4.2 Topology of the ELF

The topological analysis of scalar fields in the context of chemistry was pioneered by Richard F. W. Bader, who extensively studied the electron density and its derivatives in real-space [93]. However, the formalism can be applied to any scalar function, $g(\mathbf{r})$, for which it is possible to define the vector field $\nabla g(\mathbf{r})$. In this section, we will expose some concepts of quantum chemical topology by focusing on the ELF, i.e. $g(\mathbf{r}) = \eta(\mathbf{r})$. Nonetheless, the techniques employed in the analysis are general, and in some of the following chapters we will use them for other functions, too.

Within the dynamical system framework, a formal analogy is made between a vector field bounded on a manifold and a velocity field [48, 93]. In the present case, the vector field is the gradient field of the ELF, $\nabla\eta(\mathbf{r})$, and the manifold the three-dimensional geometrical space. In this way, each point on the manifold has exactly one gradient vector assigned to it, unless it corresponds to a critical point of the ELF. That vector is pointing towards the direction of steepest ascent, and following it allows to define what we call *field trajectories*. These lines are perpendicular to the isosurfaces of the ELF, and they are also unique for each non-critical point in space. At the Critical Points (CPs), i.e. $\mathbf{r}_c$ such that $\nabla\eta(\mathbf{r}_c) = 0$, more that one field trajectories intersect, and the type of CP determines the nature of that intersection: local maxima become points of convergence of the trajectories, and we call them *attractors*; local minima, or *repellors*, are where the trajectories diverge, and saddle points show different behaviors depending on whether it is a maximum or a minimum on a given direction. An example of this is given in Figure 4.1, where one can see the gradient field and critical points of a simple two-dimensional scalar field.

Classifying the CPs of the ELF is a key step in the topological analysis of that field. The standard procedure involves finding the Hessian of the ELF, and its eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$. The signs of those eigenvalues reveal the curvatures of the ELF in the three Cartesian dimensions, and this in turn reveals their nature. It is usual to

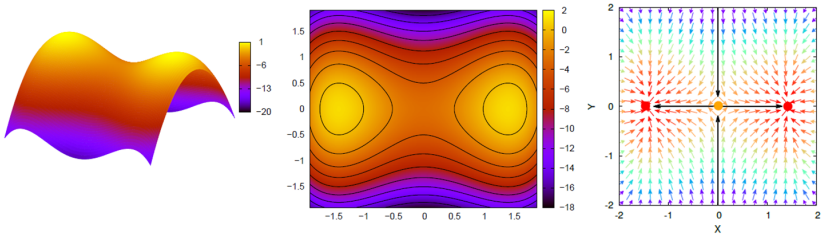


Fig. 4.1: Left: Example two dimensional function $f(x,y)$. Middle: Isolines of f on the (x,y) plane. Right: Vector field $\nabla f(x,y)$, where the directions of the vectors converge at the attractors of f, marked in red. The orange dot marks a saddlepoint of f. Reproduced from R. Boto's Ph.D. thesis with permission [98].

denote the CPs as (r,s), where r is the total number of non-zero eigenvalues of the Hessian, and s the sum of the signs of λ_1, λ_2, and λ_3. Using this notation, one can list the most common types of ELF CPs as inferred from the signs of the eigenvalues:

1. $(3,+3)$, or $\lambda_1,\lambda_2,\lambda_3 < 0$: corresponds to local maxima.
2. $(3,-3)$, or $\lambda_1,\lambda_2,\lambda_3 > 0$: corresponds to local minima.
3. $(3,-1)$, or $\lambda_1,\lambda_2 < 0$ and $\lambda_3 > 0$: corresponds to saddle points of first kind.
4. $(3,+1)$, or $\lambda_1 < 0$ and $\lambda_2,\lambda_3 > 0$: corresponds to saddle points of second kind.

This formalism provides a partition of the direct geometrical space into regions called basins, consisting of all the points whose field trajectories finish in the same attractor.

In the case of the ELF, the basins can be thought of as electronic domains corresponding to the chemical entities of the Lewis picture. Some of those basins will correspond to the valence, and are denoted by V(A, ...), where A is the atom of the neighboring atomic core basin, C(A) (which can be constituted by K, L, ... shells). In agreement with Lewis's picture, V(A,...) may belong to several atomic shells. The separation of space into ELF basins can be performed in molecules as well as in periodic systems, as exhibited in Figures 4.2 and 4.3, respectively.

According to the number of core basins that are neighbors to a given valence basin, the latter can be classified as asynaptic, monosynaptic, disynaptic and polysynaptic. Asynaptic basins do not share a border with any core basin, they are rare and not the topic of this thesis. Monosynaptic basins correspond to lone pairs, bisynaptic to a typical valence shared by two cores, and polysynaptic to multicenter bonding regions. The case of hydrogen is special, as it does not have a core and its nucleus will be placed in a valence basin. The synaptic order of such a basin will correspond to the number of neighboring core basins plus one [99].

In general, the population N_i of a given basin can be obtained by integrating the electron density inside its volume, Ω_i,

$$N_i = \bar{N}(\Omega_i) = \int_{\Omega_i} \rho(\mathbf{r})\, d\mathbf{r}\,. \tag{4.10}$$

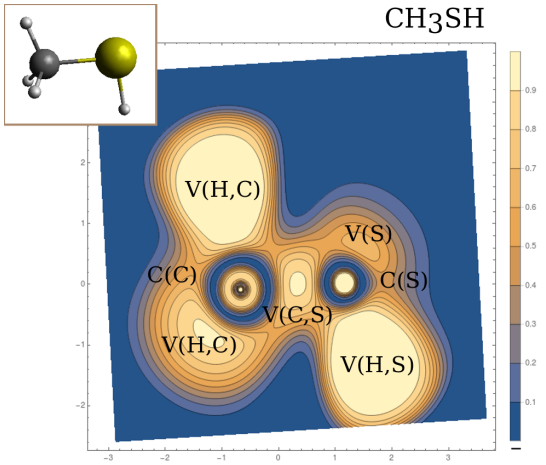


Fig. 4.2: ELF contour plot for the methanethiol molecule (CH_3SH, drawn in the inset), on the plane containing the C-S and S-H bonds. The core and valence basins, that contain the ELF attractors (maxima, in light colors) and are limited by the saddle points, are labeled and let us identify the cores (C(C) and C(S)), the bonds (V(H,C), V(C,S) and V(S,H)), and the lone pair regions (V(S)).

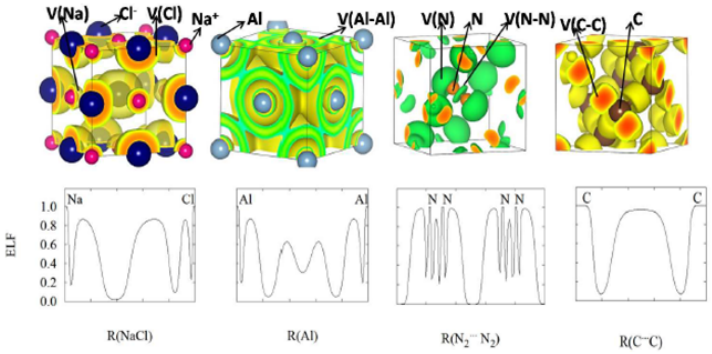


Fig. 4.3: Top: ELF isosurfaces revealing bonding patterns in solids. From left to right: NaCl, Al, N_2 and diamond. Bottom: ELF profile between atomic centers for the same systems. Reprinted with permission from Ref. [48].

This will correspond to the average number of electrons that are in said basin, and if we want to know the uncertainty of N_i we can assess its variance

$$\sigma^2(\bar{N};\Omega_i) = \langle N^2\rangle_{\Omega_i} - \langle N\rangle^2_{\Omega_i}\,. \tag{4.11}$$

It is also interesting to look at the fluctuation between two basins i, j,

$$\sigma^2(\bar{N};\Omega_i,\Omega_j) = \sigma^2(\bar{N};\Omega_i) + \sigma^2(\bar{N};\Omega_j) - 2B(\Omega_i,\Omega_j)\,, \tag{4.12}$$

where $B(\Omega_i, \Omega_j)$ is the *interbasin integrated exchange density* [100],

$$B(\Omega_i, \Omega_j) = \sum_{ij} (n_i^\alpha n_j^\alpha + n_i^\beta n_j^\beta) \int_{\Omega_i} \phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r} \int_{\Omega_j} \phi_j^*(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r}, \tag{4.13}$$

with n_i^σ the occupation number for σ electrons of the natural orbital $\phi_i(\mathbf{r})$. We also introduce the relative fluctuation,

$$F_{rel}(\Omega) = \frac{\sigma^2(\bar{N}; \Omega)}{\bar{N}(\Omega)}, \tag{4.14}$$

which is a normalized quantity and can thus be used for comparison. There, the volume Ω could correspond to one basin or to the union of two basins, in which case σ^2 would correspond to that of equation (4.12). While the separation of the ELF into basins allow us to identify regions of highest probability of finding electron pairs, the quantities (4.10), (4.11) and (4.12) provide statistical information on the electronic behavior within those basins.

Other useful information that can be obtained from the topology of the ELF is the interpretation derived from the f-domain concept [101], that enables to recover chemical units in the system, as well as to characterize the basins according to common chemical knowledge. Introduced by Mezey [102] within the QTAIM framework, the concept of an f-domain accounts for the volume enclosed by an isosurface of a certain value of $\eta(\mathbf{r}) = f$. As the value f increases, the isosurface suffers from topological changes as successive splitting of the initial domains take place, until all of them contain one, and only one, attractor. These final domains are called irreducible, and the order in which they appear with increasing f reveals the nature of the interactions taking place in the system and the relationship between basins. The turning points of the splitting corresponds to that of the highest (3,-1) point of the separatrix connecting the basins. According to the value of η at these nodes, a bifurcation tree can be constructed that reveals the basin hierarchy at a glance. To illustrate this, Figure 4.4 presents the bifurcation tree of the ELF of the molecular solid N_2, whose isosurfaces and one-dimensional internuclear profile are displayed in Fig. 4.3.

It has been observed that there is a correlation between fluctuation between two basins and the value of ELF at the separatrix between them. In this way, a persistence of a reducible domain at values of ELF close to 1 is a signature of a delocalized bond [100]. The opposite is true, for example, between two molecular units in the N_2 molecular solid, as shown in the bifurcation tree in Fig. 4.4. The value of the ELF close to zero in that region is a stamp of the lack of delocalization between those units.

4.3 Localization and delocalization indices

Another very useful tool that can be employed to understand the behavior of electrons in a system are localization and delocalization indices. A general expression for the (spinless) pair probability in terms of the electronic density is

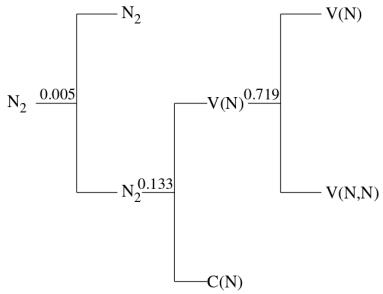


Fig. 4.4: Bifurcation tree of the N_2 molecular crystal. Adapted with permission from Ref. [48].

$$\rho_2(\mathbf{r}_1,\mathbf{r}_2) = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) + \rho_{XC}(\mathbf{r}_1,\mathbf{r}_2)\,. \tag{4.15}$$

This is a similar expression to the same-spin pair probability of eq. (4.1), but it considers all spin combinations and all exchange-correlation effects, as it is not contrained to a HF approximation. The term $\rho_{XC}(\mathbf{r}_1,\mathbf{r}_2)$ is the exchange-correlation density, and contains all the information that we need to understand the pairing of electrons. It is related to the Fermi hole, $\rho_{XC}^{\sigma\sigma}(\mathbf{r}_1,\mathbf{r}_2)$, which is the same-spin component of the conditional density $\rho_{XC}(\mathbf{r}_1,\mathbf{r}_2)/\rho(\mathbf{r}_1)$. When $\mathbf{r}_2 \to \mathbf{r}_1$, the Fermi-hole approaches $-\rho(\mathbf{r}_2)$ and the same-spin pair probability goes to zero. However, because the reference electron is not really completely localized in $\mathbf{r}_1$, the persistence of the Fermi-hole values around $\mathbf{r}_1$ gives us an idea of the localization of the reference electron.

Based on these ideas, one can define the Localization Index (LI) for a given spatial basin, and the Delocalization Index (DI) between two basins. These indices were first introduced in the framework of QTAIM [103], and the basins where considered to be those coming from the electron density. However, the definition can be extended to other partitions of space, for example using the one that is derived from the ELF.

The localization index for a basin Ω_A is defined as

$$\lambda(\Omega_A) = -\int_{\Omega_A} \rho_{XC}(\mathbf{r}_1,\mathbf{r}_2)\,d\mathbf{r}_1 d\mathbf{r}_2\,, \tag{4.16}$$

and it represents the number of electrons that are localized in the basin. Its relationship to the variance of (4.11) and the average population of (4.10) is

$$\sigma^2(\bar{N};\Omega_A) = \bar{N}(\Omega_A) - \lambda(\Omega_A)\,. \tag{4.17}$$

Here we see that, when the electrons are fully localized, the $\sigma^2(\bar{N};\Omega_A)$ goes to zero and $\lambda(\Omega_A)$ takes its maximum value of $\bar{N}(\Omega_A)$.

On the other hand, the delocalization index between basins Ω_A and Ω_B is defined as

$$\delta(\Omega_A,\Omega_B) = -2\int_{\Omega_A}\int_{\Omega_B}\rho_{XC}(\mathbf{r}_1,\mathbf{r}_2)\,d\mathbf{r}_1 d\mathbf{r}_2\,, \tag{4.18}$$

and it yields the number of electrons that fluctuate between the two basins. In this way, considering a total number of N electrons in the system, the relationship between the LIs and DIs of all basins must be

$$N = \sum_{\Omega_A}\left[\lambda(\Omega_A) + \frac{1}{2}\sum_{\Omega_B\neq\Omega_A}\delta(\Omega_A,\Omega_B)\right]. \tag{4.19}$$

In practice, one can compute these indices making use of molecular orbitals, ϕ_i, by expressing the pair density as

$$\rho_2(\mathbf{r}_1,\mathbf{r}_2) = \sum_{ijkl}{}^2D^{ij}_{kl}\phi^*_i(\mathbf{r}_1)\phi^*_j(\mathbf{r}_2)\phi_k(\mathbf{r}_1)\phi_l(\mathbf{r}_2)\,, \tag{4.20}$$

where ${}^2D^{ij}_{kl}$ are the matrix elements of the 2-RDM. Then, the DIs and LIs take the form

$$\lambda(\Omega_A) = -\sum_{ijkl}{}^2D^{ij}_{kl}S_{ik}(\Omega_A)S_{jl}(\Omega_A) + \bar{N}(\Omega_A)^2\,, \tag{4.21}$$

$$\delta(\Omega_A,\Omega_B) = -2\sum_{ijkl}{}^2D^{ij}_{kl}S_{ik}(\Omega_A)S_{jl}(\Omega_B) + 2\bar{N}(\Omega_A)\bar{N}(\Omega_B)\,, \tag{4.22}$$

with

$$S_{ik}(\Omega_A) = \int_{\Omega_A}\phi^*_i(\mathbf{r})\phi_k(\mathbf{r})\,d\mathbf{r}\,. \tag{4.23}$$

Finally, one can extend this formalism to solid-state by considering the crystal orbitals $\psi_{n,\mathbf{k}}(\mathbf{r})$, defined in a given $\mathbf{k}$-mesh [104]. In such a case, integrals over the first BZ become summations and we have

$$\delta(\Omega_A,\Omega_B) = \frac{2}{K_{BZ}}\sum_{n,n'}\sum_{\mathbf{k},\mathbf{k}'}S_{n\mathbf{k},n'\mathbf{k}'}(\Omega_A)S_{n\mathbf{k},n'\mathbf{k}'}(\Omega_B)\,n_{n\mathbf{k}}\,n_{n'\mathbf{k}'}\,, \tag{4.24}$$

where K_{BZ} is a normalization corresponding to the number of $\mathbf{k}$-points in the BZ, $n_{n\mathbf{k}}$ are occupation numbers of $\psi_{n,\mathbf{k}}(\mathbf{r})$, that we now allow to be different from 0 or 1, and the overlap integral is defined as

$$S_{n\mathbf{k},n'\mathbf{k}'}(\Omega) = \int_{\Omega}\psi^*_{n,\mathbf{k}}(\mathbf{r})\psi_{n',\mathbf{k}'}(\mathbf{r})\,d\mathbf{r}\,. \tag{4.25}$$

Estimating the localization index is straightforward from $\delta(\Omega_A,\Omega_B)$ using 4.19.

Part II

Real-space descriptors in a model system

5

One-dimensional model of a superconducting hydride

Although the microscopic BCS and Migdal-Eliashberg theories successfully explain the mechanism that enables the superconducting transition in conventional materials, those theories are formulated entirely in reciprocal space, and therefore they prevent us from understanding how electrons behave in real-space. The topological analysis of real-space quantities, like the electron localization function introduced in Section 4.1, is a very useful tool used in Quantum Chemistry, and a lot of information about the system can be extracted from such an analysis. It would thus be interesting to carry out a similar examination in the superconducting state, and study the localization and delocalization of electrons there.

This chapter focuses on studying model systems. These usually depend on just a few variables, which can be changed at will, allowing a better understanding of the problem at hand. This feature probes particularly useful in this kind of studies, where the descriptors that we define in real-space have not been previously studied in the literature, and similar approaches are scarce. In this way, we center our attention on one-dimensional chains of atoms. In Section 5.1, we present a preliminary model of a superconducting ELF, derived from the singlet-pair function of Section 3.3. The lack of an energy scale in that model encourage us to change to a different framework, and propose a set of real-space descriptors derived from SCDFT, that are presented in Section 5.2. Finally, Section 5.3 is entirely dedicated to the study of those descriptors in hydrogen chains (H-chains). This choice is based on two assumptions: i) that the properties of interest in a superconducting hydride should be inherent of the hydrogen sublattice, and ii) that it is possible to construct a superconducting state from the metallic state of the H-chain (without stating the conditions for its existence). The first assumption can be justified because, in a great amount of hydrogen-based superconductors, the hydrogens form a weakly-bonded sublattice that is practically independent from the other atoms, that only act as chemical precompressors (see Chapter 6).

5.1 Singlet-pair function orbital model

In this section, we present our first attempt at studying electron localization and delocalization in the superconducting state. We start from a spatial representation of a Cooper pair, that is inherited from the Gor'kov wavefunction describing the macroscopic behavior of singlet pairs [82], and which was later approximated by Kadin [83]. After introducing a local density approximation, we use the resulting spatial description of the Cooper pair to represent a valence orbital of a superconducting system. The one-dimensional model that will be presented allows for the definition of an electron localization function, both in the normal and in the superconducting states. This allows to analyze and compare the localization patterns between the two phases.

5.1.1 Spatial representation of the Cooper pair

The singlet pair function proposed by Kadin [83], displayed in equation (3.41) and reprinted here for convenience,

$$\Psi_{SP}(r) \propto \cos(k_F r) K_0\left(\frac{r}{\pi\xi_0}\right), \tag{5.1}$$

is an approximation to the wavefunction of a Cooper pair in one dimension. It is a function of the distance between the electrons that form the pair, $r = |r_1 - r_2|$, with r_1 and r_2 the positions of the electrons. It also depends on the BCS coherence length, $\xi_0 = \frac{k_F}{\pi\Delta}$, that is in turn determined by the gap, approximated here by a constant, $\Delta_{\mathbf{k}} = \Delta$.

In order to use the expression in eq. 5.1 to construct a spatial model of the superconducting state, it would be necessary to know the dependence of Ψ_{SP} with respect to the spatial position of the pair, i.e. $R = \frac{1}{2}(r_1 + r_2)$. To do that, we introduce a local density approximation, that assumes that the density is locally constant, as in the case of the homogeneous electron gas. Under this approximation, the Fermi momentum can be written in terms of the electron density as $k_F = k_F(R) = (3\pi^2\rho(R))^{1/3}$. Replacing this into eq. (5.1) one obtains,

$$\Psi_{SP}(r,R) \propto \cos(k_F(R) r)\, K_0\left(\frac{r}{\pi\xi_0}\right), \tag{5.2}$$

$$= \cos\left(\frac{(9\pi\rho(R)^2)^{1/3}}{\Delta}\right) K_0\left(\frac{r}{\pi\xi_0}\right). \tag{5.3}$$

The dependence on r remains an inconvenience, as it is not the internal structure of the Cooper pair that we wish to analyze. Therefore, we fix this distance to the coherence length ξ_0, arriving at

$$\Psi_{SP}(R) \equiv \Psi_{SP}(\xi_0(R), R) \propto \cos\left(\frac{(9\pi\rho(R)^2)^{1/3}}{\Delta}\right) K_0\left(\frac{1}{\pi}\right), \tag{5.4}$$

The expression in equation (5.4) provides a real-space representation of the Cooper pair wavefunction. In the following, we will use it to describe the valence electrons of a system in the superconducting state.

5.1.2 The one-dimensional orbital model

The one-dimensional orbital model aims to represent a chain containing M atoms. First, we introduce the model for the normal state. The unit cells consist of two core orbitals, approximated by two Gaussian functions centered at 0 and $2R_0$, and a diffuse function that represents the metallic valence. The two core orbitals for the n-th unit cell are

$$\chi_1(r) = N_1 e^{-\alpha(r-4nR_0)^2}, \tag{5.5}$$

$$\chi_2(r) = N_2 e^{-\alpha(r-2R_0(1+2n))^2}. \tag{5.6}$$

Meanwhile, the valence is depicted by a very broad Gaussian,

$$\chi_3(r) = N_3 e^{-\beta(r-R_0(1+2n))^2}, \tag{5.7}$$

by taking $\beta \ll \alpha$. Figure 5.1 shows the spatial form of the three orbitals of the unit cell when taking the numerical values of $\alpha = 1.0$ a.u. and $\beta = 0.001$ a.u., and computing the normalization constants, $N_A = N_B$ and N_C. Although we use atomic units throughout the description of this model, they are actually quite arbitrary and what matters in this case is the relationship between the chosen parameters. In the case of the choice of R_0 and α, because the Gaussian functions represent core electrons, it is important that the overlap is minimal, as it can be seen in Fig. 5.1 for $R_0 = 2.0$ a.u.

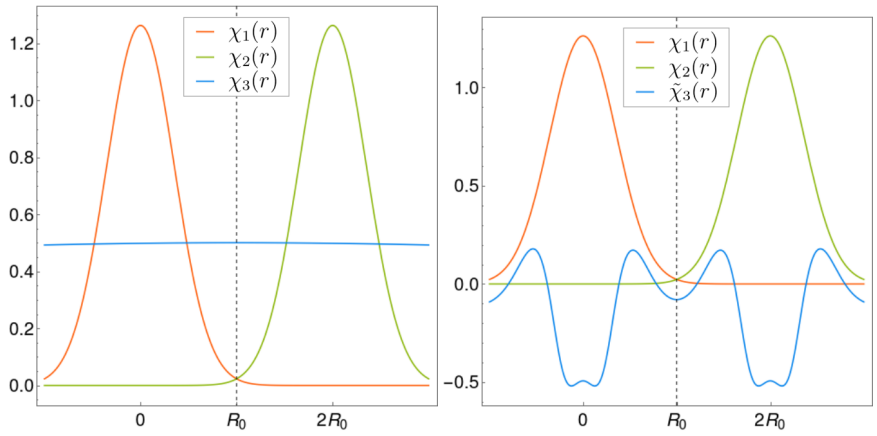


Fig. 5.1: To the left, spatial form of gaussian core and valence function of diffuse, metallic state, for $R_0 = 2$ a.u. On the right, the valence wavefunction is replaced by the Kadin Cooper-pair wavefunction, with $\Delta = 2$ a.u.

The molecular orbitals for the normal state, $\{\phi_1^{NS}(r), \phi_2^{NS}(r), \phi_3^{NS}(r)\}$, are obtained by orthonormalizing the basis set $\{\chi_1(r), \chi_2(r), \chi_3(r)\}$. The new basis can be used to find the normal state electron density,

$$\rho^{NS}(r) = 2\sum_{i=1}^{3}\left|\phi_i^{NS}(r)\right|^2, \tag{5.8}$$

and kinetic energy density,

$$\tau^{NS}(r) = \sum_{i=1}^{3}\left|\frac{d\phi_i^{NS}(r)}{dr}\right|^2. \tag{5.9}$$

Having those two quantities, it is possible to define the ELF for this system, $\eta^{NS}(r)$, using eq. (4.9). The profile of the ELF in the normal state for this model is offered in Fig. 5.2, for different iteratomic distances $2R_0$. It shows the expected shape, clearly revealing the position of the cores and the bonding electrons. It can be appreciated that the value of the ELF at the minima between the nuclei increases when the atoms approach each other. This is a well-known and expected behavior, that reproduces the effect of pressure on the ELF profile.

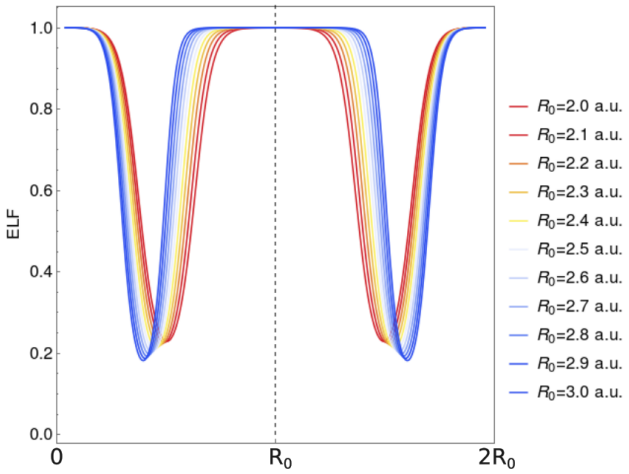


Fig. 5.2: ELF of the normal state model for different interatomic distances ($2R_0$). Atoms are positioned at the extremes of the plot.

For the superconducting state, the valence orbital function has been substituted by the real space representation of the Cooper pair wavefunction, as expressed in equation (5.4),

$$\tilde{\chi}_3(r) = \cos\left[\frac{\sqrt[3]{9\pi}(\rho^{NS}(r))^{2/3}}{\Delta}\right]K_0\left(\frac{1}{\pi}\right), \tag{5.10}$$

where the normal state density has been used as a reference upon which the Cooper pairs first form. This is represented in Fig. 5.1 for a value of $\Delta = 2.0$ a.u. In analogy with the normal state model, the basis set $\{\chi_1(r), \chi_2(r), \tilde{\chi}_3(r)\}$ is diagonalized to

obtain the molecular orbital orthonormalized basis set, $\{\phi_1^{SC}(r), \phi_2^{SC}(r), \phi_3^{SC}(r)\}$. The molecular orbitals are used to define the superconducting electron density, kinetic energy density and ELF, just as it was done for the normal state. Those quantities are denoted by $\rho^{SC}(r)$, $\tau^{SC}(r)$ and $\eta^{SC}(r)$, respectively.

The comparison between the ELF profiles for the normal and superconducting states is shown in Fig. 5.3, for $R_0 = 2.0$ a.u. and $R_0 = 3.0$ a.u., and different values of the superconducting gap. One can see that for both states the ELF shares a similar envelope. Further, for the larger values of Δ, a higher value of the ELF at the saddle points seems to coincide with more delocalized electrons in the superconducting state. However, the superconducting ELF is highly dependent on the value of the gap, and the lack of a Hamiltonian for this system prevents us from having a proper energy scale to choose an appropriate value for Δ.

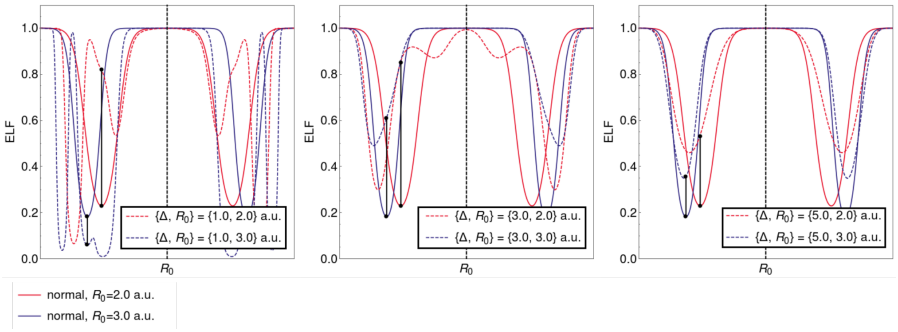


Fig. 5.3: ELF profiles of the normal state (continuous line) and superconducting state (dashed line) for $R_0 = 2.0$ a.u. (red) and $R_0 = 3.0$ a.u. (blue). The values of the gap are of $\Delta = 1.0$ a.u. (left), $\Delta = 3.0$ a.u. (middle), $\Delta = 5.0$ a.u.(right). In all of the pannels, the atoms are located at the extremes of the plot.

We observe that the superconducting ELF has a more oscillatory behavior than that of the normal state. In fact, it is the Kadin wavefunction that becomes highly-oscillatory for lower values of the gap, which is a direct consequence of the cosine in eq. (5.10). For $\Delta \to \infty$, the Cooper-pair wavefunction becomes constant and both systems resemble each other. This is not the expected limit of the system, on the contrary, both phases should converge at $\Delta \to 0$, and they should deviate from each other at higher values of Δ. For these reasons, we have decided to put aside this first model, and start from scratch using a completely different framework, that we shall present in the following sections.

5.2 Electronic localization in Superconducting DFT

Superconducting DFT, introduced in Section 3.4, provides a description of the superconducting state that depends on the anomalous potential, $\Delta_{ext}(\mathbf{r}, \mathbf{r}')$, that allows

Cooper pairs to tunnel in and out of the system. It is related to the BCS superconducting gap, and it is an order parameter of the transition: if the anomalous potential goes to zero, one recovers the normal state of the system. This is specially interesting to us, as it allows to properly assess the changes on the electron localization when the system undergoes the superconducting transition, i.e. when Δ_{ext} becomes non-zero.

In SCDFT, it is possible to obtain an expression for the electron density of the superconducting state in terms of the (normal state) KS orbitals, their energies, the gap, and the temperature; see eq. (3.63). This is already an advantage of SCDFT, providing a first look at the spatial distribution of the electrons in a superconductor. In this section, we will introduce other spatial functions that can be derived in the SCDFT framework, that will allow us to go deeper into the analysis of the real-space properties of superconductors.

5.2.1 Generalized superconducting ELF

The first real-space function that we define in SCDFT is the superconducting ELF (SC-ELF). To do so, it is necessary to first obtain an expression for the superconducting 1-RDM, $\rho_1^{SC}(\mathbf{r},\mathbf{r}')$. As it has been done in Section 3.4, we shall use the grand canonical ensemble, leading to

$$\rho_1^{SC}(\mathbf{r},\mathbf{r}') = \mathrm{Tr}\left[\varrho_0 \sum_\sigma \psi_\sigma^\dagger(\mathbf{r})\psi_\sigma(\mathbf{r}')\right]. \tag{5.11}$$

The procedure to obtain a final expression for $\rho_1^{SC}(\mathbf{r},\mathbf{r}')$ is analogue to the one followed in Section 3.4 for the Superconducting (SC) electron density. Just as in eq. (3.58), the SC 1-RDM can be expressed in terms of the functions $u_i(\mathbf{r})$ and $v_i(\mathbf{r})$, and of the Fermi-Dirac distribution, $f(E)$,

$$\rho_1^{SC}(\mathbf{r},\mathbf{r}') = 2\sum_i \left[u_i^*(\mathbf{r})u_i(\mathbf{r}')f(E_i) + v_i^*(\mathbf{r})v_i(\mathbf{r}')f(-E_i)\right]. \tag{5.12}$$

Then, using the expressions in equations 3.62 and 3.61 for $u_i(\mathbf{r})$ and $v_i(\mathbf{r})$, one obtains

$$\rho_1^{SC}(\mathbf{r},\mathbf{r}') = \sum_{n\mathbf{k}} n_{n\mathbf{k}}^{SC}\varphi_{n\mathbf{k}}^*(\mathbf{r})\varphi_{n\mathbf{k}}(\mathbf{r}'), \tag{5.13}$$

where

$$n_{n\mathbf{k}}^{SC} = 1 - \frac{\xi_{n\mathbf{k}}}{|E_{n\mathbf{k}}|}\tanh\left(\frac{\beta|E_{n\mathbf{k}}|}{2}\right). \tag{5.14}$$

Using the density matrix in (5.13), it is possible to compute the kinetic energy of the system,

$$
\begin{aligned}
T^{SC} &= -\frac{1}{2}\int \nabla^2_{\mathbf{r}'}\rho_1^{SC}(\mathbf{r},\mathbf{r}')\big|_{\mathbf{r}'=\mathbf{r}}\, d\mathbf{r}\,, \\
&= -\frac{1}{2}\sum_{n\mathbf{k}} n^{SC}_{n\mathbf{k}} \int \varphi^*_{n\mathbf{k}}(\mathbf{r})\nabla^2_{\mathbf{r}}\varphi_{n\mathbf{k}}(\mathbf{r})\, d\mathbf{r}\,, \\
&= \frac{1}{2}\sum_{n\mathbf{k}} n^{SC}_{n\mathbf{k}} \int |\nabla_{\mathbf{r}}\varphi_{n\mathbf{k}}(\mathbf{r})|^2\, d\mathbf{r}\,.
\end{aligned}
$$

Then, the positive definite Kinetic Energy Density (KED) of the superconducting state can be defined as,

$$
\tau^{SC}(\mathbf{r}) = \frac{1}{2}\sum_{n\mathbf{k}} n^{SC}_{n\mathbf{k}} |\nabla_{\mathbf{r}}\varphi_{n\mathbf{k}}(\mathbf{r})|^2\,. \tag{5.15}
$$

The von Weizsacker and Thomas-Fermi approximations to the KEDs are obtained using $\rho^{SC}(\mathbf{r})$ from eq. (3.63),

$$
\tau^{SC}_{vW}(\mathbf{r}) = \frac{1}{8}\frac{\left|\nabla_{\mathbf{r}}\rho^{SC}(\mathbf{r})\right|^2}{\rho(\mathbf{r})}\,, \tag{5.16}
$$

and

$$
\tau^{SC}_{TF}(\mathbf{r}) = \frac{3}{10}(3\pi^2)^{2/3}\rho^{SC}(\mathbf{r})^{5/3}\,, \tag{5.17}
$$

respectively. Finally, equations (5.15), (5.16), and (5.17) enable us to define a superconducting ELF,

$$
\eta^{SC}(\mathbf{r}) = \left[1+\left(\frac{\tau^{SC}(\mathbf{r})-\tau^{SC}_{vW}(\mathbf{r})}{\tau^{SC}_{TF}(\mathbf{r})}\right)^2\right]^{-1}\,. \tag{5.18}
$$

In the same way as in the normal state, this function describes the localization patterns of the electrons in the superconducting state. The topology of the superconducting ELF introduces a partition of real-space corresponding to those regions where it is more likely to find electron pairs. However, one must not confuse those with Cooper pairs. First of all, because not all of the electrons in a superconductor form Cooper pairs, only those near the Fermi energy do (see Section 3.1); and secondly, because Cooper pairs are known to span large regions of space, evidenced by the values of the coherence length in hydrogen-based superconductors, $\xi_0 \sim 15-30\,$Å[1] [105]. Nonetheless, the superconducting ELF should help shine light on how the electrons arrange in that state, its similarities and differences with respect to the normal state, and how that changes from one system to another.

[1] In non-hydrogen-based conventional superconductors this distance is even larger, of the order of magnitude of $\sim 1\mu$m. In the case of hydrides, which tend to have a much larger critical temperatures, this diminishes abruptly. A first explanation to this is that, from BCS, $\xi_0 \propto \frac{1}{T_c}$.

It is interesting to comment on the fact that the difference between the spatial properties of the normal and the superconducting state lie exclusively on the occupation numbers of the KS orbitals, see eq. 5.13. This is an effect of the approximations that are introduced in SCDFT, and it proves particularly useful during implementation, which can be done without having to find orbitals again through an expensive self-consistent procedure. The introduction of fractional occupations in the superconducting state is a consequence of the highly-correlated nature of the superconducting electrons, which is well-known. The new SC occupation numbers thus allows to study this correlation, as we shall see in the following sections.

5.2.2 Other descriptors

Just as in the normal state, the SC ELF will be high in those places where the probability of finding singlet pairs of electrons due to the Pauli principle is higher. This is because the kernel of $\eta^{SC}(\mathbf{r})$ considers the excess kinetic energy density of the fermionic system compared to a bosonic one with the same density (see Section 4.1). Considering that not all the electrons in a superconductor form Cooper pairs, even if the exact values around the maxima of the ELF change when going from the normal to the superconducting state, the SC ELF should remain high in regions where there are covalent bonds, for example. This is so because we have assumed that the crystal structure does not suffer changes in the phase transition. In contrast, if we would like to study the localization/delocalization of those pairs due to the formation of Cooper pairs, it would be necessary to replace the kernel of the ELF by the difference between the kinetic energies in the superconducting and normal states. In this way, we can define a one-body Cooper pair localization function (1-CPLF),

$$\eta_1^{CP}(\mathbf{r}) = \left[1 + \left(\frac{\tau^{SC}(\mathbf{r}) - \tau^{NS}(\mathbf{r})}{\tau_{TF}^{SC}(\mathbf{r})}\right)^2\right]^{-1}, \tag{5.19}$$

where $\tau^{NS}(\mathbf{r})$ is the normal state KED. The reference Thomas-Fermi (TF) KED is taken in the superconducting state just as in eq. 5.18, to focus on assessing the changes coming only from the reference in the difference between the KEDs.

Although the expression in eq. 5.19 should help us learn about the behavior of Cooper pairs in real space, it is a one-body object, and it will therefore instrinsically lack information about the two-body nature of the pairs. It is also possible to define a two-body Cooper pair localization function, or 2-CPLF. To do so we make use of the spin-resolved up-down 2-RDM,

$$\rho_2^{\uparrow\downarrow}(\mathbf{r}_1\mathbf{r}_2, \mathbf{r}_1'\mathbf{r}_2') = \left\langle \psi_\uparrow^\dagger(\mathbf{r}_1)\psi_\downarrow^\dagger(\mathbf{r}_2)\psi_\downarrow(\mathbf{r}_2')\psi_\uparrow(\mathbf{r}_1')\right\rangle. \tag{5.20}$$

This is one of the components of the total 2-RDM, such that $\rho_2(\mathbf{r}_1\mathbf{r}_2, \mathbf{r}_1'\mathbf{r}_2') = \sum_{\sigma,\sigma'}\rho_2^{\sigma\sigma'}(\mathbf{r}_1\mathbf{r}_2, \mathbf{r}_1'\mathbf{r}_2')$. It is possible to think about $\rho_2^{\uparrow\downarrow}(\mathbf{r}_1\mathbf{r}_2, \mathbf{r}_1'\mathbf{r}_2')$ as a 1-RDM for up-down pairs,

$$\rho_2^{\uparrow\downarrow}(\mathbf{r}_1\mathbf{r}_2, \mathbf{r}_1'\mathbf{r}_2') = \left\langle \psi_{\uparrow\downarrow}^\dagger(\mathbf{r}_1, \mathbf{r}_2)\psi_{\uparrow\downarrow}(\mathbf{r}_1', \mathbf{r}_2')\right\rangle \equiv \rho_1^{\uparrow\downarrow}(\mathbf{R}\mathbf{s}, \mathbf{R}'\mathbf{s}'). \tag{5.21}$$

changing to the new set of coordinates that represent the average position of the pair and the distance between the electrons, $\mathbf{R} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)$ and $\mathbf{s} = |\mathbf{r}_1 - \mathbf{r}_2|$, respectively; and considering the average to be over the ensemble, $\langle A \rangle = \mathrm{Tr}[\varrho_0 A]$. In order to evaluate (5.21), the following approximation from ref. [106] can be used,

$$\begin{aligned}
\left\langle \psi^\dagger_\uparrow(\mathbf{r}_1)\psi^\dagger_\downarrow(\mathbf{r}_2)\psi_\downarrow(\mathbf{r}'_2)\psi_\uparrow(\mathbf{r}'_1)\right\rangle \approx & -\langle\psi^\dagger_\uparrow(\mathbf{r}_1)\psi_\downarrow(\mathbf{r}'_2)\rangle\langle\psi^\dagger_\downarrow(\mathbf{r}_2)\psi_\uparrow(\mathbf{r}'_1)\rangle && (5.22)\\
& + \langle\psi^\dagger_\downarrow(\mathbf{r}_2)\psi_\downarrow(\mathbf{r}'_2)\rangle\langle\psi^\dagger_\uparrow(\mathbf{r}_1)\psi_\uparrow(\mathbf{r}'_1)\rangle \\
& + \langle\psi_\downarrow(\mathbf{r}'_2)\psi_\uparrow(\mathbf{r}'_1)\rangle\langle\psi^\dagger_\uparrow(\mathbf{r}_1)\psi^\dagger_\downarrow(\mathbf{r}_2)\rangle, && (5.23)\\
= & \langle\psi^\dagger_\downarrow(\mathbf{r}_2)\psi_\downarrow(\mathbf{r}'_2)\rangle\langle\psi^\dagger_\uparrow(\mathbf{r}_1)\psi_\uparrow(\mathbf{r}'_1)\rangle \\
& + \langle\psi_\downarrow(\mathbf{r}'_2)\psi_\uparrow(\mathbf{r}'_1)\rangle\langle\psi^\dagger_\uparrow(\mathbf{r}_1)\psi^\dagger_\downarrow(\mathbf{r}_2)\rangle, && (5.24)\\
= & \rho^\downarrow_1(\mathbf{r}_2,\mathbf{r}'_2)\rho^\uparrow_1(\mathbf{r}_1,\mathbf{r}'_1) + \chi(\mathbf{r}'_2,\mathbf{r}'_1)\chi^*(\mathbf{r}_2,\mathbf{r}_1). && (5.25)
\end{aligned}$$

Here we have used the fact that both factors of the first term of eq. 5.22 are zero. The spin 1-RDM, $\rho^\sigma_1(\mathbf{r},\mathbf{r}') \equiv \langle\psi^\dagger_\sigma(\mathbf{r})\psi_\sigma(\mathbf{r}')\rangle$, is such that $\rho_1(\mathbf{r},\mathbf{r}') = \rho^\uparrow_1(\mathbf{r},\mathbf{r}') + \rho^\downarrow_1(\mathbf{r},\mathbf{r}')$, and therefore for a closed-shell system we have $\rho^\sigma_1(\mathbf{r},\mathbf{r}') = \frac{1}{2}\rho_1(\mathbf{r},\mathbf{r}')$, and

$$\rho^{\uparrow\downarrow}_2(\mathbf{r}_1\mathbf{r}_2,\mathbf{r}'_1\mathbf{r}'_2) \approx \frac{1}{4}\rho_1(\mathbf{r}_2,\mathbf{r}'_2)\rho_1(\mathbf{r}_1,\mathbf{r}'_1) + \chi(\mathbf{r}'_2,\mathbf{r}'_1)\chi^*(\mathbf{r}_2,\mathbf{r}_1). \tag{5.26}$$

The expression in (5.26) can thus be computed using equations (5.13) and (3.65). In terms of the pair coordinates $(\mathbf{R},\mathbf{s})$, it becomes,

$$\begin{aligned}
\rho^{\uparrow\downarrow}_1(\mathbf{R}\mathbf{s},\mathbf{R}'\mathbf{s}') \approx & \frac{1}{4}\rho_1\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}'-\frac{\mathbf{s}'}{2}\right)\rho_1\left(\mathbf{R}+\frac{\mathbf{s}}{2},\mathbf{R}'+\frac{\mathbf{s}'}{2}\right) \\
& + \chi\left(\mathbf{R}'-\frac{\mathbf{s}'}{2},\mathbf{R}'+\frac{\mathbf{s}'}{2}\right)\chi^*\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}+\frac{\mathbf{s}}{2}\right). && (5.27)
\end{aligned}$$

With this, a Normal State (NS) and SC KED of oppposite-spin pairs can be defined, $\tau^{\uparrow\downarrow}_{NS}(\mathbf{R},\mathbf{s})$ and $\tau^{\uparrow\downarrow}_{SC}(\mathbf{R},\mathbf{s})$, respectively. In both cases, this is done as follows,

$$\begin{aligned}
\tau^{\uparrow\downarrow}(\mathbf{R},\mathbf{s}) = & -\frac{1}{2}\nabla^2_{\mathbf{R}'}\rho^{\uparrow\downarrow}_1(\mathbf{R}\mathbf{s},\mathbf{R}'\mathbf{s}')\Big|_{\mathbf{R}'=\mathbf{R}}, \\
= & -\frac{1}{8}\Big[\rho_1\left(\mathbf{R}+\frac{\mathbf{s}}{2},\mathbf{R}'+\frac{\mathbf{s}}{2}\right)\nabla^2_{\mathbf{R}'}\rho_1\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}'-\frac{\mathbf{s}}{2}\right) + \\
& + \rho_1\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}'-\frac{\mathbf{s}}{2}\right)\nabla^2_{\mathbf{R}'}\rho_1\left(\mathbf{R}+\frac{\mathbf{s}}{2},\mathbf{R}'+\frac{\mathbf{s}}{2}\right) \\
& + 2\nabla_{\mathbf{R}'}\rho_1\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}'-\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}'}\rho_1\left(\mathbf{R}+\frac{\mathbf{s}}{2},\mathbf{R}'+\frac{\mathbf{s}}{2}\right)\Big]_{\mathbf{R}'=\mathbf{R}} \\
& - \frac{1}{2}\chi^*\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}+\frac{\mathbf{s}}{2}\right)\nabla^2_{\mathbf{R}}\chi\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}+\frac{\mathbf{s}}{2}\right). && (5.28)
\end{aligned}$$

Entering the explicit expression of the 1-RDM, its laplacian becomes

$$\nabla^2_{\mathbf{R}'}\rho_1\left(\mathbf{R}\pm\frac{\mathbf{s}}{2},\mathbf{R}'\pm\frac{\mathbf{s}}{2}\right)\Big|_{\mathbf{R}'=\mathbf{R}} = \sum_{n\mathbf{k}} n_{n\mathbf{k}}\varphi^*_{n\mathbf{k}}\left(\mathbf{R}\pm\frac{\mathbf{s}}{2}\right)\nabla^2_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}\pm\frac{\mathbf{s}}{2}\right),$$
$$= -\sum_{n\mathbf{k}} n_{n\mathbf{k}}\left|\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}\pm\frac{\mathbf{s}}{2}\right)\right|^2, \tag{5.29}$$

where $n_{n\mathbf{k}}$ corresponds to either $n^{NS}_{n\mathbf{k}}$ or $n^{SC}_{n\mathbf{k}}$. The gradient, on the other hand, is harder to simplify:

$$\nabla_{\mathbf{R}'}\rho_1\left(\mathbf{R}\pm\frac{\mathbf{s}}{2},\mathbf{R}'\pm\frac{\mathbf{s}}{2}\right)\Big|_{\mathbf{R}'=\mathbf{R}} = \sum_{n\mathbf{k}} n_{n\mathbf{k}}\varphi^*_{n\mathbf{k}}\left(\mathbf{R}\pm\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}\pm\frac{\mathbf{s}}{2}\right), \tag{5.30}$$

leading to

$$\nabla_{\mathbf{R}'}\rho_1\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}'-\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}'}\rho_1\left(\mathbf{R}+\frac{\mathbf{s}}{2},\mathbf{R}'+\frac{\mathbf{s}}{2}\right)\Big|_{\mathbf{R}'=\mathbf{R}} = \sum_{nn'\mathbf{k}\mathbf{k}'} n_{n\mathbf{k}}n_{n'\mathbf{k}'}\varphi^*_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)$$
$$\times\varphi^*_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)$$
$$\times\nabla_{\mathbf{R}}\varphi_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right). \tag{5.31}$$

In the normal state, the order parameter vanishes and, using eqs. (5.29) and (5.31), one obtains the KED,

$$\tau^{\uparrow\downarrow}_{NS}(\mathbf{R},\mathbf{s}) = \frac{1}{8}\sum_{nn'\mathbf{k}\mathbf{k}'} n^{NS}_{n\mathbf{k}}n^{NS}_{n'\mathbf{k}'}\left[\left|\varphi_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\right|^2\left|\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\right|^2\right.$$
$$+\left|\varphi_{n'\mathbf{k}'}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\right|^2\left|\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\right|^2$$
$$\left.-2\varphi^*_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\varphi^*_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}}\varphi_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\right]. \tag{5.32}$$

Whereas for the SC state we must consider the order parameter, obtaining

$$\tau^{\uparrow\downarrow}_{SC}(\mathbf{R},\mathbf{s}) = \frac{1}{8}\sum_{nn'\mathbf{k}\mathbf{k}'} n^{SC}_{n\mathbf{k}}n^{SC}_{n'\mathbf{k}'}\left[\left|\varphi_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\right|^2\left|\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\right|^2\right.$$
$$+\left|\varphi_{n'\mathbf{k}'}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\right|^2\left|\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\right|^2$$
$$\left.-2\varphi^*_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\varphi^*_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}}\varphi_{n\mathbf{k}}\left(\mathbf{R}-\frac{\mathbf{s}}{2}\right)\nabla_{\mathbf{R}}\varphi_{n'\mathbf{k}'}\left(\mathbf{R}+\frac{\mathbf{s}}{2}\right)\right]$$
$$-\frac{1}{2}\chi^*\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}+\frac{\mathbf{s}}{2}\right)\nabla^2_{\mathbf{R}}\chi\left(\mathbf{R}-\frac{\mathbf{s}}{2},\mathbf{R}+\frac{\mathbf{s}}{2}\right). \tag{5.33}$$

Finally, the 2-body Cooper pair localization function can be defined as

$$\eta^{CP}_2(\mathbf{R},\mathbf{s}) = \left[1+\left(\frac{\tau^{\uparrow\downarrow}_{SC}(\mathbf{R},\mathbf{s})-\tau^{\uparrow\downarrow}_{NS}(\mathbf{R},\mathbf{s})}{\tau^{SC}_{TF}(\mathbf{R})}\right)^2\right]^{-1}, \tag{5.34}$$

It is also possible to integrate the variable **s** to obtain a local function, as is the original ELF,

$$\eta_2^{CP}(\mathbf{R}) = \left[1 + \left(\int d\mathbf{s}\, \frac{\tau_{SC}^{\uparrow\downarrow}(\mathbf{R},\mathbf{s}) - \tau_{NS}^{\uparrow\downarrow}(\mathbf{R},\mathbf{s})}{\tau_{TF}^{SC}(\mathbf{R})}\right)^2\right]^{-1}. \tag{5.35}$$

Although the expression in (5.34) is more physically meaningful, as it allows to consider all possible values of **s**, it could be more difficult to implement computationally. Alternatively, one can analyze eq. (5.35) for set values of **s**. For example, following what has been done for the spatial analysis of $\chi(\mathbf{R},\mathbf{s})$ in ref. 89, one can choose to look at the values of the functions at $s = |\mathbf{s}| = 0$. Given the typically oscillating behavior of the order parameter with respect to s, it makes sense to evaluate its value at its maximum, located at $s = 0$.

5.2.3 Approximation of the superconducting gap

In order use the expression in eq. 5.18 to examine the SC ELF in a model or in a real system, it is necessary to have an expression for the gap. In SCDFT, this is done making use of the Green's functions formalism, and solving a self-consistent equation, see Section 3.4. Unfortunately, this procedure is very complex and there are no computational tools developed for this purpose available to the public at this moment. It is possible, however, to introduce an approximation and represent the dependence of the gap at zero Kelvin with respect to the energies ξ as an isotropic Lorentzian function,

$$\Delta_0(\xi) = \frac{\Delta_0}{N_0\pi}\frac{\omega/2}{\xi^2 + (\omega/2)^2}, \tag{5.36}$$

where ω is a parameter that adjusts the width of the peak, and N_0 is a normalization such that the height of the peak at $\xi = 0$ is Δ_0. The latter is the constant of the gap at $T = 0$ K in BCS [3], that depends on the critical temperature:

$$\Delta_0 = 1.76 k_B T_c, \tag{5.37}$$

with k_B being Boltzmann's constant. Then, considering the dependence of the gap with respect to the temperature from eq. 3.16, we shall use

$$\Delta(\xi;T) = \Delta_0(\xi)\tanh\left(1.74\sqrt{\frac{T_c - T}{T}}\right). \tag{5.38}$$

In this way, since eq. (5.37) dictated the dependence of the gap at $T = 0$ with respect to the critical temperature, eq. (5.38) specifies the gap for any temperature and energy, for a given T_c.

Figure 5.4 offers the profile of the gap function for $\omega = 0.2$ eV, considering different critical temperatures, T_c, and at different temperatures, T. This functions represent the width of the window around the Fermi energy where electrons will form

Cooper pairs. One can see how the gap goes to zero everywhere when the temperature reaches $T = T_c$, where the transition occurs. There, the properties of the normal and superconducting state become the same, as expected from SCDFT. A similar thing occurs to the anomalous density,

$$\chi(\xi;T) = \frac{\Delta(\xi;T)}{2\sqrt{\xi^2+\Delta(\xi;T)^2}} \tanh\left(\frac{\sqrt{\xi^2+\Delta(\xi;T)^2}}{2k_BT}\right), \tag{5.39}$$

as displayed in Fig. 5.5.

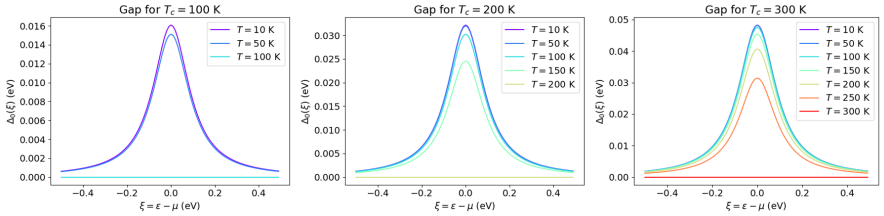


Fig. 5.4: Gap function around the Fermi energy, considering different critical temperatures: T_c = 100 K (left), T_c = 200 K (middle), and T_c = 300 K (right). In each case, the dependence on the temperature T is shown by different colored lines.

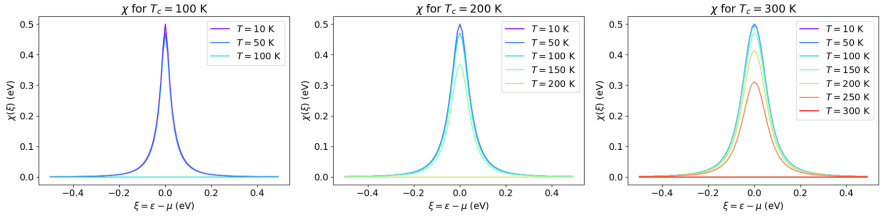


Fig. 5.5: Anomalous density as a function of the energy, for three different critical temperatures: T_c = 100 K (left), T_c = 200 K (middle), and T_c = 300 K (right). In each case, the dependence on the temperature T is shown by different colored lines.

Taking this into account, one can compare the occupation numbers for the metallic and the superconducting state, as can be seen in Figure 5.6. For a fixed critical temperature, it can be seen that the occupancies of the normal state deviate from the step function as the temperature increases, softening the transition around the Fermi energy, as expected. On the other hand, the superconducting occupation numbers do not suffer great alterations with the temperature. In fact, they tend to resemble the occupations at the critical temperature (in both states, as they are the same), showing

the larger correlation of the superconducting state in comparison with the normal state below T_c.

For the highest critical temperature, $T_c = 300$ K, Fig. 5.6 shows that a larger range of occupancies is spanned. Because all of the real-space functions that interest us depend only on the superconducting parameters through the occupation numbers, we will hereby take that value of T_c for the analysis, and seize the effect of the changes in those functions with respect to the temperature, T.

5.3 Tight-binding approach to the hydrogen chain

In this Section, we focus on the analysis of real-space functions in a one-dimensional hydrogen chain, constructed using a tight-binding model. A comparison is made between the symmetric and dimerized chains, that have a metallic and insulating character, respectively. Further, we construct the superconducting functions in those systems, using the formalism presented in the previous Section, and compare the results to those of the normal state. In order to do that, it is necessary to assume a superconducting critical temperature, T_c. This might seem arbitrary and a lost of generality of the results, but the analysis of the occupation numbers in the SC state performed in Section 5.2.3 shows that this choice should not affect the validity of the results.

5.3.1 The tight-binding model in real space

In order to make use of the tight-binding formalism to evaluate the real-space properties of the hydrogen chain, as reviewed in Section 2.5, we begin by building a basis set of Gaussian atomic functions:

$$\chi_{lA}(x) = e^{-\alpha(x-la)^2}, \tag{5.40}$$

$$\chi_{lB}(x) = e^{-\alpha(x-(R_{AB}+la))^2}, \tag{5.41}$$

that depends on the unit cell index, l. In this way, every unit cell contains two basis functions, just as in the case presented in Section 2.5. Fig. 5.7 is a spatial representation of the two atomic orbitals in the unit cell. In general, we consider the length of the unit cell, a, to be equal to $1.0\,u$, with u some arbitrary units. In the next section, we shall see which values of a properly represent a hydrogen chain. The same is done for the Gaussian exponent α.

With this, the atomic orbitals fulfilling Bloch's theorem, or Bloch sums, are

$$\varphi_{Ak}(x) = \frac{1}{\sqrt{N_l}} \sum_l e^{ikla} e^{-\alpha(x-la)^2}, \tag{5.42}$$

$$\varphi_{Bk}(x) = \frac{1}{\sqrt{N_l}} \sum_l e^{ik(R_{AB}+la)} e^{-\alpha(x-(R_{AB}+la))^2}, \tag{5.43}$$

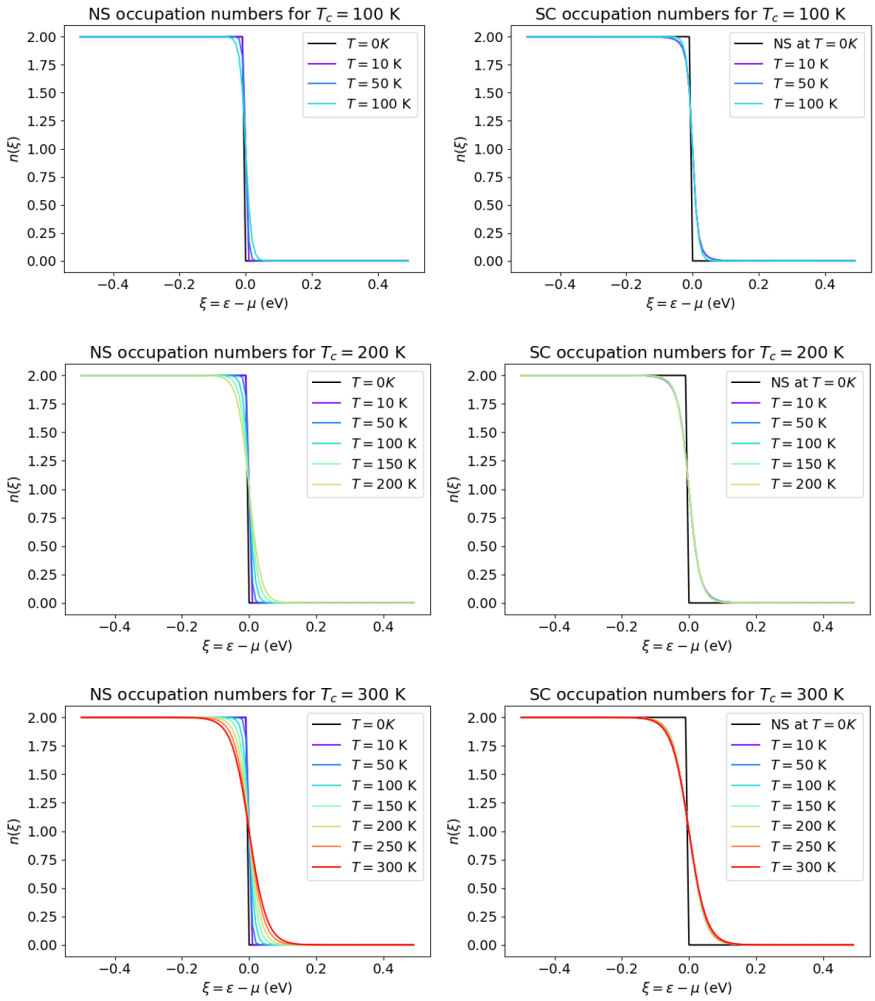


Fig. 5.6: Occupation numbers at different energies. In all panels, that of the normal state at $T = 0$ K is depicted in black. To the left, each figure shows the occupation numbers of the normal state at different temperatures, for a given T_c. To the right, the same is displayed for the superconducting state.

with N_l the number of unit cells considered in the model. Here, k is selected to take values inside the first BZ, $k \in \left[-\frac{\pi}{a}, \frac{\pi}{a}\right]$, with $k = \frac{2\pi}{N_l a} l$ and l taking integer values in the range $\left[-\frac{N_l}{2}, \frac{N_l}{2}\right]$. For simplicity, we will generally refer to k in units of $\frac{2\pi}{a}$. A graphical representation of these orbitals for $k = 0.02$ (i.e. $k = 0.02 \times \frac{2\pi}{a}$) is offered in Figure 5.8, for a total number of $N_l = 101$.

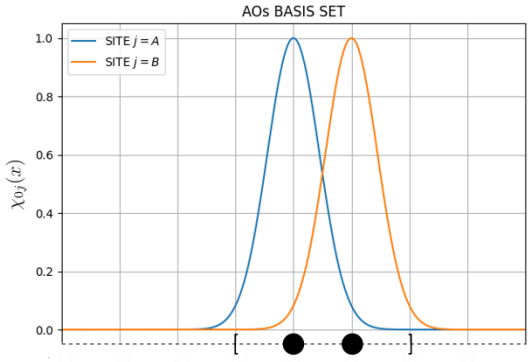


Fig. 5.7: Atomic orbitals used in the unit cell of the hydrogen chain tight-binding model, represented by two identical Gaussian functions. The black spheres mark the positions of the two nuclei inside the unit cell, delimited by squared brackets. Here, the unit cell length is taken as $1\,u$, and the Gaussian exponent as $\alpha = 10\,u^{-2}$

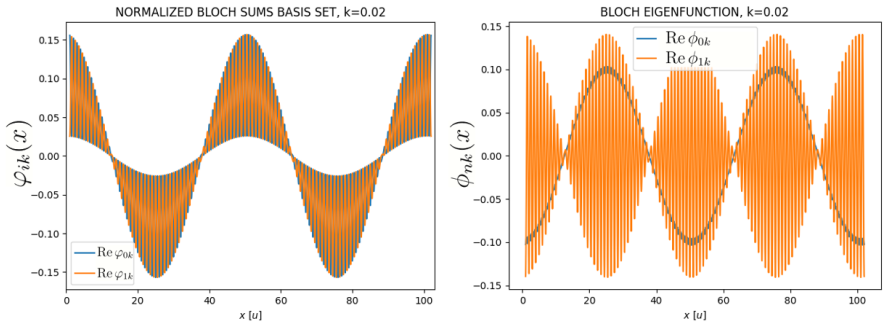


Fig. 5.8: Left: Real part of the Bloch sums of each band index, $i = 0$ and $i = 1$, for $k = 0.02 \times \frac{2\pi}{a}$. Right: Real part of the Bloch wavefunctions at the same k point for the symmetric chain.

Finally, the Bloch wavefunctions that diagonalize the Hamiltonian of the hydrogen chain can be written as a linear combination of the Bloch sums in (5.42) and (5.43), as in eq. 2.54,

$$\phi_{nk}(x) = \frac{1}{\sqrt{N_l}} \sum_i \sum_{R_i} c_{ik}^n e^{ikR_i} e^{-\alpha(x-R_i)^2}. \tag{5.44}$$

In this case, the band index n can take the values 0 and 1. The diagonalization determines the coefficients c_{ik}^n, ant it is performed using the Gaussian elimination method [107], as implemented in NumPy [108]. The result of the diagonalization will necessarily depend on the hopping parameters. Let us recall that the hopping parameter

between site A and the site B to its right is v, whereas between the same site A and the atom B on its left, the hopping parameter is w (see Fig. 2.2). In the special case in which $v = w$, the resulting coefficients are $c^0_{Ak} = c^0_{Bk} = \frac{1}{\sqrt{2}}$ and $c^1_{Ak} = -c^1_{Bk} = \frac{1}{\sqrt{2}}$, corresponding to a bonding and antibonding orbitals, respectively. Figure 5.8 shows an example of such a Bloch wavefunction.

Once the Bloch wavefunctions have been defined and computed, the next step is to evaluate the real-space functions of interest. In order to do so, we will generally assume that the Bloch wavefunctions are orthonormal, which allows expressions such as

$$\rho(x) = \sum_{nk} n_{nk} \left|\phi_{nk}(x)\right|^2 , \tag{5.45}$$

for the electron density, and similar ones for the other functions.

To evaluate the orthogonality, we first resort to the nearest-neighbors approximation, in which the only non-zero overlaps between the basis functions are

$$\int \chi^*_{lA}(x)\chi_{lA}(x)dx = \int \chi^*_{lB}(x)\chi_{lB}(x)dx = C\,, \tag{5.46}$$

and that between two adjacent atomic sites,

$$\int \chi^*_{lA}(x)\chi_{lB}(x)\,dx = \int \chi^*_{l-1B}(x)\chi_{lA}(x)\,dx = \int \chi^*_{lB}(x)\chi_{l+1A}(x)\,dx = \mathcal{S} \tag{5.47}$$

because we consider Periodic Boundary Conditions (PBC), we must also include

$$\int \chi^*_{N_lB}(x)\chi_{1A}(x)\,dx = \mathcal{S}\,. \tag{5.48}$$

Equation (5.46) allows to normalize the basis set. Once this is done, the Bloch overlap will yield

$$\int \varphi^*_{ik}(x)\varphi_{jk'}(x)\,dx = \delta_{k,k'} c^{k,k'}_{i,j}\,, \tag{5.49}$$

where it can be shown that $c^{k,k}_{A,A} = c^{k,k}_{B,B} = 1$, and $c^{k,k}_{i,j}$ for $i \neq j$ is dependent on $\mathcal{S}$. In this way, equation (5.49) shows that the Bloch sums are orthogonal with respect to the wave vector index k, but not with respect to the site. This changes when we diagonalize the Hamiltonian and find the Bloch eigenfunctions, which are orthonormal

$$\int \phi^*_{nk}(x)\phi_{n'k'}(x)\,dx = \delta_{n,n'}\delta_{k,k'}\,. \tag{5.50}$$

Equation (5.50) is true analytically, but in practice there is a numerical error that comes from the integration of the Bloch sums in (5.49): the overlap of two Bloch sums of different k is not exactly zero. Given the importance of the orthogonality condition in the computation of the density and the other real-space descriptors, we define the numerical error associated to those computations as

$$ERROR = \max_{k \neq k'} \left\{ \int \varphi^*_{ik}(x) \varphi_{jk'}(x)\, dx \right\}. \tag{5.51}$$

This quantity sets the precision of the results presented in the following, and it will be of special interest to minimize it when choosing the parameters of the model.

As a final remark concerning the computation of the real-space quantities in this model, it is important to note that the expression of the ELF in eq. (4.9) is not valid for one-dimensional systems, as the expression of the TFKED changes in those cases to [109]

$$\tau_{TF}(x) = \frac{\pi^2}{24} \rho(x)^3 . \tag{5.52}$$

It is this one-dimensional version of the TFKED and the corresponding ELF, that we consider for all of the calculations of this Section.

5.3.2 The model parameters

The construction of the hydrogen chain using the tight-binding formalism relies on a set of parameters that need to be chosen carefully for the model to properly represent the physical system. Firstly, we choose the lattice parameter, a, for which we hereby adopt the value of $a = 2.64\,\text{Å}$. This determines the value of the arbitrary units, $1\,u = 2.64$ Å. This value is in accordance with what has been observed in the literature for high-temperature hydrogen-based superconductors of interest [110, 111]. In this way, for the symmetric chain we shall set the hydrogen-hydrogen distance to $d_{HH} = 0.5\,u = 1.32$ Å.

The Gaussian exponent, α, is optimized using Valence Bond (VB) theory on a similar hydrogen chain, using a 6-31G* basis, and fitting the proposed Gaussian orbitals to those results [112]. The outcome of such a calculation indicates that α should live in the range of $[8, 15]\,u^{-2}$ or, equivalently, $\alpha \in [1.15, 2.15]$ Å^{-2}. In general, we will use $\alpha = 10\,u^{-2} = 1.43$ Å^{-2}.

In theory, the hopping parameters, v and w, could be evaluated analytically through the integral of equation 2.60. In practice this is not possible, as divergences arise from the Coulomb potential of hydrogen in one dimension,

$$V_0(x) = \frac{1}{4\pi\epsilon_0} \frac{1}{x}, \tag{5.53}$$

with ϵ_0 the permittivity of vacuum. This issue is avoided by using the Wolfsberg-Helmholtz Approximation (WHA), that has been proposed for the Hückel model, that is, in principle, physically equivalent to the tight-binding formalism. For one-dimensional systems, the WHA allows to establish the relationship

$$t = \kappa \cdot \epsilon \cdot \mathcal{S}, \tag{5.54}$$

for the hopping t, with respect to the onsite energy ϵ and the overlap $\mathcal{S}$ of eq. 5.47. There, κ is a constant that is set to 0.787 [2] and the onsite energy for the hydrogen's $1s$

[2] Here, the value of κ is scaled from the typical value of $\kappa = 1.75$ for covalent bonds, to yield the correct values for the hopping when the on site energy is shifted from $\epsilon = 0$ eV to $\epsilon = -13.6$ eV.

orbital is considered, $\epsilon = -13.6$ eV. With this, the hopping parameters are restrained to the range $t \in [-3.0, -1.5]$ eV. Moreover, equation (5.54) is particularly useful when atomic distances are variable, because it allows to scale the hopping parameters, simply by assessing the new overlap integrals, $\mathcal{S}$.

In the case of the symmetric chain, when only one hopping parameter is present, we will use the value $t = v = w = -1.5$ eV. It is interesting to note that, actually, in that case the only effect of changing t is that the energy scale varies, resulting in flatter bands for a higher $|t|$, affecting the Density of States (DOS) at the Fermi energy.

Another two parameters that arise from the discretization of the problem in real and reciprocal space must be determined, namely the real-space grid step, Δx, and the number of k points, N_k (that is equal to the number of unit cells, N_l). To ensure the reliability of our results, we perform a convergence test on those parameters, in order to minimize the numerical error arising from such an approximation. First, we fix $N_k = 101$ and vary the grid step Δx from $10^{-1}\,u$ to $10^{-4}\,u$. We evaluate the error of the overlap with respect to that distance, as defined in equation (5.51), as it is displayed in Fig. 5.9. In order to check the convergence of the localization properties, we also evaluate progress with respect to Δx of the minimum value of the ELF, baptized here as ϕ, for different temperatures and models (NS and SC). Because temperature does not seem to have much effect on the convergence, only the results for $T = 10$ K are shown in Fig 5.10. Setting $\Delta x = 10^{-3}\,u$ to ensure an error below 10^{-5}, the same analysis is performed by varying the number of k-points. Only odd numbers are considered to ensure a proper sampling of the BZ, that incorporates the Γ-point, i.e. $x = 0\,u$. Considering the results obtained in this convergence test, the value of $N_k = 101$ is chosen for the rest of the calculations.

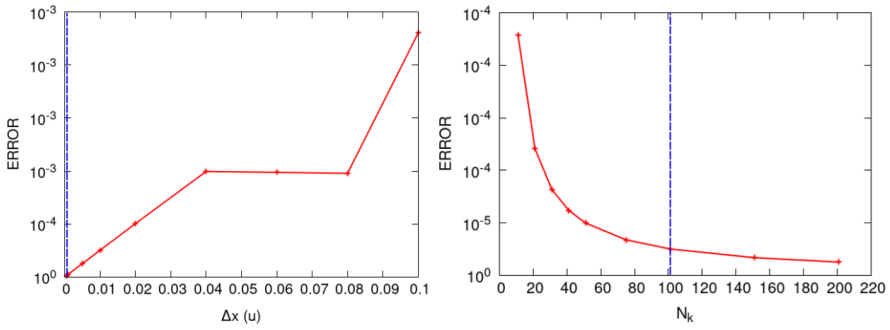


Fig. 5.9: Error in the orthogonality of the Bloch sums with respect to: the grid step considered in the discretization of the spatial functions, $\Delta x\,(u)$ (left); and the number of k-points considered in the discretization of reciprocal space, N_k. The dashed vertical line marks the chosen value of Δx abd N_k.

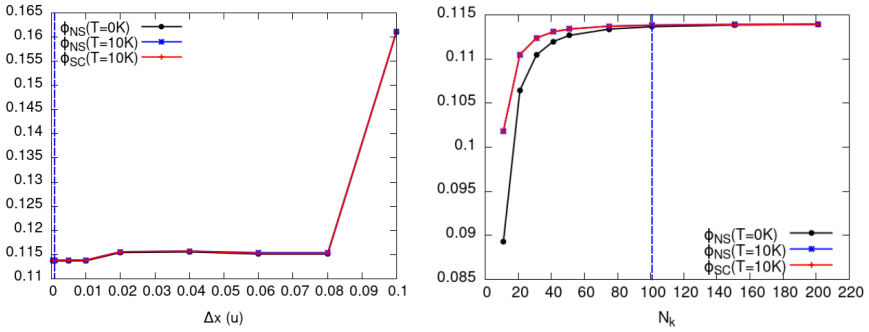


Fig. 5.10: Value of the minima of the ELF of the NS at zero temperature (red), the NS at $T = 10$ K (green) and the SC state at $T = 10$ (K) (blue), for $T_c = 300$ K; with respect to: the grid step considered in the discretization of the spatial functions, $\Delta x(u)$ (left); and the number of k-points considered in the discretization of reciprocal space, N_k. The dashed vertical line marks the chosen value of Δx abd N_k.

5.3.3 Results in the symmetric chain

Here we present the results obtained for a symmetric hydrogen chain, i.e. when all the interatomic distances and hoppings are considered to be the same, set to $t = -1.5$ eV and $d_{HH} = 1.32\,\text{Å}$, respectively. The energy bands of the system is shown in Fig. 5.11, where the absence of a gap signals the metallic nature of the chain.

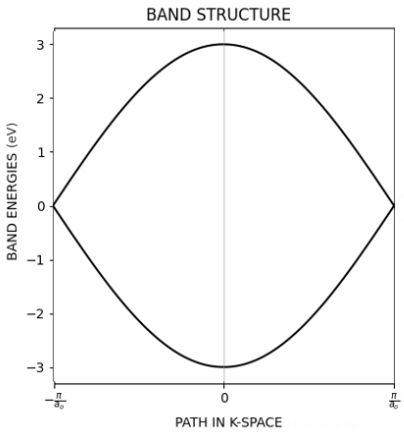


Fig. 5.11: Energy spectrum of the two-band model of the tight-binding symmetric hydrogen chain in the BZ.

Figure 5.12 shows the spatial form of the electron density and KED in the symmetric chain. In each case, the NS quantities at $T = 0$ K, at $T = 10$ K, and the SC functions at $T = 10$ K are considered. However, it is clear that they do not differ by much, as both functions seem to be the same to the eye. The same is observed for the ELF in Figure 5.13. In this way, the differences between the occupation numbers in the three cases do not seem to change the spatial descriptors significantly. Considering that at low temperatures like $T = 10$ K the occupation numbers are the most different in the NS and SC cases, the former remark is true for all higher temperatures, too.

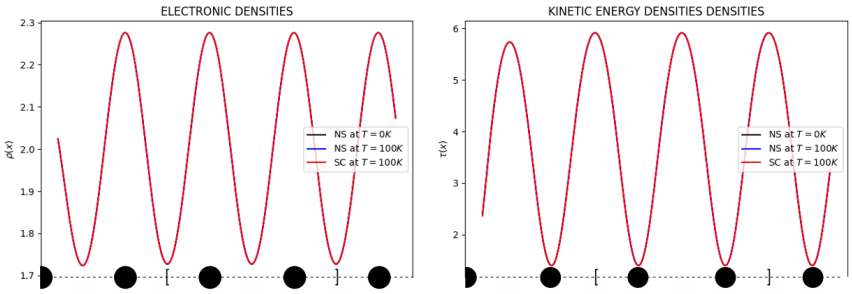


Fig. 5.12: Profiles of the electron density (left) and the kinetic energy density (right) in the symmetric hydrogen chain for $T_c = 300$ K. In both panels, the three overlapping lines represent the three considered models: normal state at $T = 0$ K and $T = 100$ K, and superconducting state at $T = 100$ K. Atomic positions are marked by black circles, and the unit cell is delimited by squared brackets.

Although we expected to gain more information about the spatial properties in the superconducting state, the similarity of the descriptors in both states is actually quite practical. Indeed, this allows to infer localization and delocalization properties of the SC state from those of the NS, that are much cheaper to reproduce. In this sense, if a link between the SC spatial properties and the usual observables of interest is to be found, that link can immediately be inherited to the normal state properties. This is exactly what is observed in real systems, as we shall see in Chapter 6, where the delocalization properties in the NS are shown to correlate with the critical temperature. This model permits to understand why the NS properties are able to carry such information.

From the similarities in the results of the KED in both states, one can immediately infer that the 1- and 2-CPLF of equations (5.19) and (5.35) will be equal to 1 everywhere. This does not provide much information about the system, and we thus abstain from drawing any conclusions. On the other hand, the spatial profile of the diagonal of the anomalous density of eq. (3.65), $\chi(x, x)$, offers some information about the spatial properties of superconductors. Fig. 5.14 shows how the anomalous density diminishes as the temperature reaches T_c, as expected. The spatial maxima

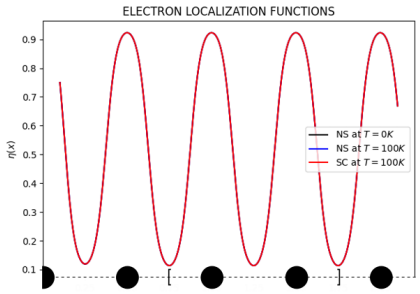


Fig. 5.13: Profile of the ELF in the symmetric hydrogen chain for T_c = 300 K. In both panels, the three overlapping lines represent the three considered models: normal state at T = 0 K and T = 100 K, and superconducting state at T = 100 K. Atomic positions are marked by black circles, and the unit cell is delimited by squared brackets.

of $\chi(x,x)$ is located at the atomic positions, but even there they present a very small magnitude, specially in comparison with the values of the electronic density (see Fig. 5.12). This large difference in scale between the two densities explains why the ELFs in the different states are so similar. Although from eq. (3.65) is it clear that the magnitude and the width of the gap play a key role in the magnitude of the anomalous density, the sum over $n\mathbf{k}$ states indicates that the DOS at the Fermi energy is also important in determining the value of $\chi(x,x)$.

The localization and delocalization indices, introduced in Section 4.3, were evaluated for the symmetric chain. Figure 5.15 shows how the LIs, λ, evolve with respect to the temperature, taking values of ca. 0.46. Taking into account that $\lambda = 1$ is the highest possible localization index for a hydrogen chain (no fluctuations), and that $\lambda = 0.5$ is the expected index for non-interacting electrons, one can say that the symmetric chain shows important fluctuations in the ELF basins, similar to the non-interacting case. Although the results show that in the normal state the localization diminishes when the temperature increases, as expected, those changes are very small (of the order of 10^{-6}), and it is not possible to rule out that they might be a product of numerical error. The same holds for the comparison between λ for the NS and SC state: although the SC case seems more localized, no conclusion can be drawn at this scale.

The DIs in the hydrogen chain are denoted by $\delta^{1,j}$, to indicate that the overlap integrals of equation (4.24) are taken for a reference ELF basin and its j-th neighbor. The indices in Figure 5.15 show that the charge fluctuations are maximal between consecutive basins ($j = 2$), and then they decrease in an alternate fashion. In fact, in analytical tight-binding models, it is expected that the values of odd numbers of j would go to zero [113], and the deviations from that result in our model are a consequence of the numerical error in the overlap of the orbitals. In general, the alternation can be explained by the phenomenon of Friedel oscillations that arises

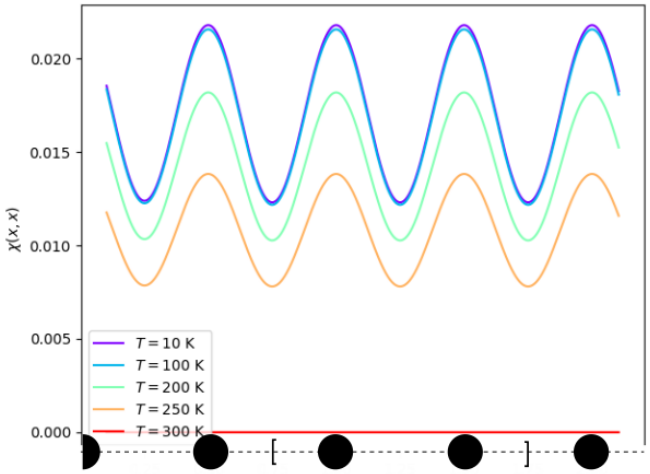


Fig. 5.14: Anomalous SC density, $\chi(x)$, along the hydrogen chain for $T_c = 300$ K and different temperatures, T. The atomic positions are marked by the black circles, with the squared brackets delimiting the unit cell.

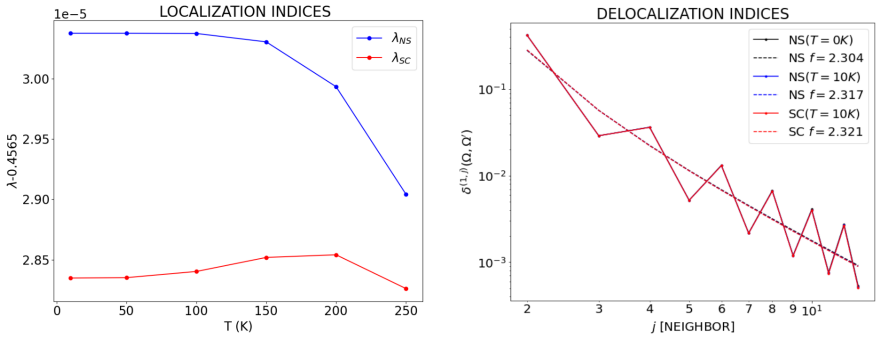


Fig. 5.15: Left: Localization indices of the normal and superconducting states in blue and red, respectively, for different temperatures between 10 K and $T_c = 300$ K. Right: Delocalization indices for the normal state at $T = 0$ K (black) and $T = 100$ K (blue), and for the SC state at $T = 100$ K (red); for increasing order of neighbors, j. The dashed lines show the algebraic fit to the decay of the indices.

in metallic systems, or with Pauling resonance structures in conjugated molecules. The effect of this is that the fluctuation of the electrons is preferred within the same sublattice (A or B), even when the system is symmetric [114].

It has been observed that for a metallic system, the decay of the DIs is algebraic, and it is slower than that of an insulators, that decreases exponentially [113]. In the

first case, evolution of the DIs can be modeled by

$$\delta^{1,j} = (j-1)^{-f} \times \delta^{1,2}, \tag{5.55}$$

where the value of the exponent f measures the rate of the decay. This is what can be observed in Fig. 5.15, where all three curves are modeled by an algebraic fit of similar decay rate, f, although showing a slightly faster decay in the superconducting case.

5.3.4 Results in a dimerized chain

When the distance between a hydrogen in sublattice A with respect to the B atom to its right is different to that with respect to the B atom on its left, we encounter ourselves with a dimerized chain. We will analyze the changes of its properties for different distances, considering the smallest hydrogen-hydrogen distance, hereby referred to as d_{HH}, to range between 0.74 Å and 1.32 Å, the latter corresponding to the symmetric case. The hopping parameters v and w will vary accordingly, following the scaling in eq. 5.54.

The energy spectrum of this chain is substantially different from its symmetric analogue, characterized by the presence of a gap between the two bands, distinctive of an insulating state (see Fig. 5.16).

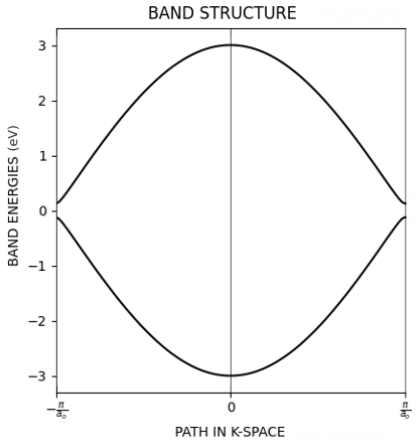


Fig. 5.16: Energy spectrum of the two-band model of the tight-binding asymmetric hydrogen chain in the BZ, for $d_{HH} = 1.3$ Åand a unit cell of size $a = 2.64$ Å.

Figure 5.17 displays the electron density, the KED, and the ELF in the dimerized case, for $d_{HH} = 1.3$ Å. Once again, the real-space functions of the NS and SC cases overlap. This is true for the whole range of temperatures and interatomic distances. The most evident changes of the topology of the ELF are with respect to the variation

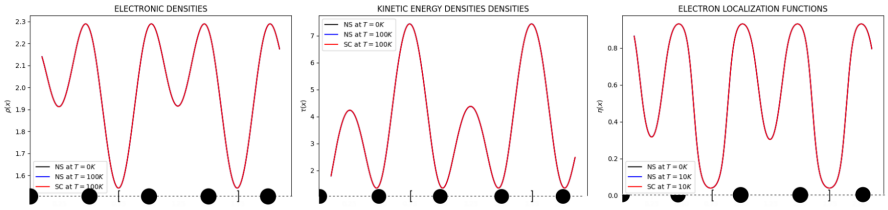


Fig. 5.17: Profile of the electron density (left), kinetic energy density (middle), and ELF (right), in the dimerized chain where $d_{HH} = 1.3$ Å, at $T = 10$ K. The atomic positions are marked by the black circles, with the squared brackets delimiting the unit cell.

of the interatomic distances as a result of the symmetry loss. In fact, the minimum value of the ELF, ϕ, has dropped to nearly zero when slightly shifting the distances from $d_{HH} = 1.32$ Å to $d_{HH} = 1.3$ Å. The profile of the ELF reveals the molecularization of the system as two of the hydrogens approach each other. The value of the ELF at the higher local minima between the hydrogens in the same unit cell, on the other hand, increases with respect to the symmetric case. This is characteristic of a more delocalized behavior of electrons in the intramolecular region. The value of the ELF at this local minima is hereby denoted by ϕ^*.

Further decreasing the minimum interatomic distance accentuates those changes in the topology of the function, as it can be noticed in Figure 5.18. In fact, when d_{HH} decreases sufficiently, the atoms in the lattice form units resembling H_2 molecules, characterized by the flat ELF profiles. We thus refer to this distance as intramolecular.

The analysis of the evolution of the topological descriptors ϕ and ϕ^* with respect to the intramolecular distance, and to the energy gap (see Fig. 5.19), show that a small value of ϕ and a large one for ϕ^* are a feature of an insulating state in the hydrogen chain, where intramolecular distances are shortened. On the other hand, both topological descriptors will show a similar value in a metallic state, where distances are more elongated. This is further supported by the increase of the localization in the ELF basins with the decreasing of the intramolecular distance, as it can be seen in Figure 5.21.

The anomalous density in the dimerized case shows a clear tendency: it diminishes abruptly with the interatomic distance, as evidenced by the logarithmic scale in Fig. 5.20. This is in accordance with what has been observed in hydrogen-based superconductors, as we shall see in Section 6.2, where systems presenting molecular-like units of H_2 tend to be worse superconductors. In this way, the topological descriptors ϕ and ϕ^*, that correlate with d_{HH}, also hold information about the anomalous density. Interestingly, the evolution of the topology of $\chi(x, x)$ seems to be similar to that of the ELF in Fig. 5.18.

The DIs in the dimerized chain show a larger decreasing rate, as measured by the larger value of f (see Fig. 5.21), which is expected for the insulating state. Indeed, the charge fluctuations in an insulator are more local, as the electrons are more tightly

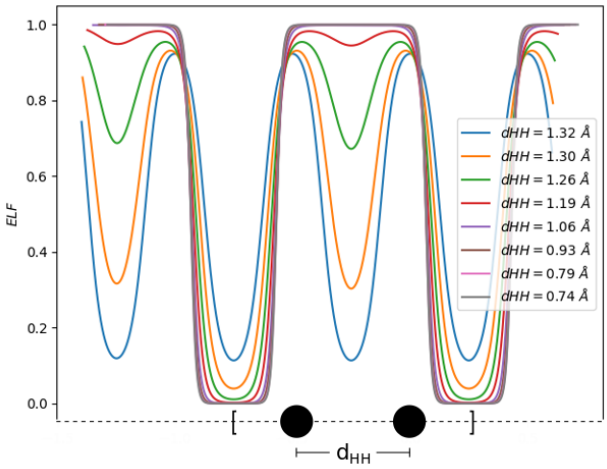


Fig. 5.18: Normal state ELF profile for the dimerized hydrogen chain at $T = 100$ K, considering different minimum interatomic distances, d_{HH}.

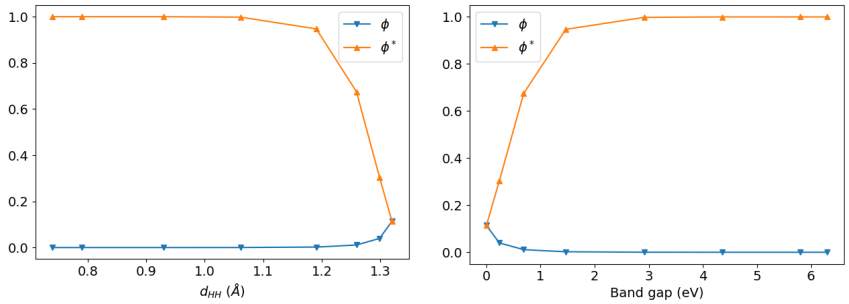


Fig. 5.19: Evolution of the topological descriptors ϕ and ϕ^* with respect to the intramolecular distance (left) and the band gap (right).

bound to their basins. This is also in agreement with the nearsightedness of electronic matter in gapped systems, where the fluctuations of the density away from a given point where the density is perturbed decays exponentially with the distance to that point [115].

5.3.5 High-correlation limit

The comparison of the superconducting spatial properties using the descriptors proposed in Section 5.2 yielded results that were indistinguishable from those in the normal state. As it was shown in Fig. 5.15, however, there does seem to be a difference between the two states, but at those values of the superconducting gap they

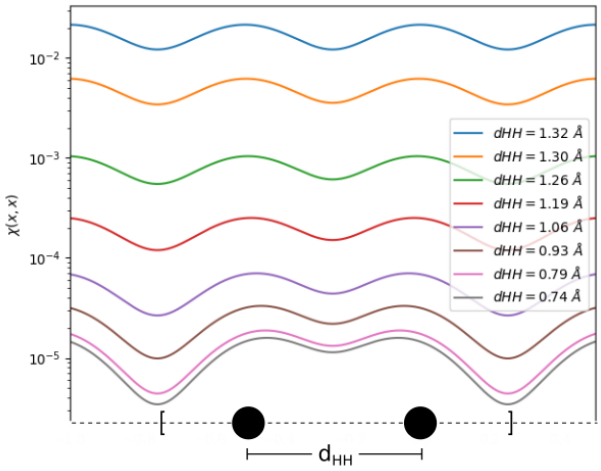


Fig. 5.20: Anomalous SC density, $\chi(x)$, for the dimerized hydrogen chain at $T = 100$ K, considering different minimum interatomic distances, d_{HH}.

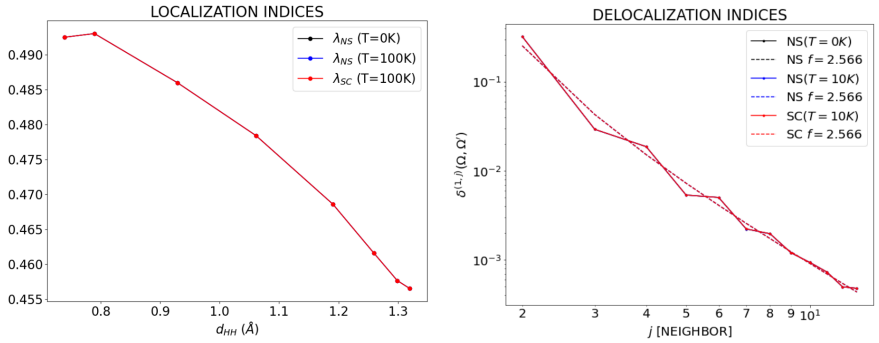


Fig. 5.21: Left: Localization indices of the dimerized chain with $d_{HH} = 1.3$ Å for the normal state at 0 K in black, and of normal and superconducting states at $T = 100$ K in blue and red, respectively; with respect to different intramolecular distances. Right: Delocalization indices for the normal state at $T = 0$ K (black) and $T = 100$ K (blue), and for the SC state at $T = 100$ K (red); for increasing order of neighbors, j. The dashed lines show the algebraic fit to the decay of the indices.

were not considerable, and were of the order of magnitude of the numerical error. Having this in mind, it becomes interesting to study the limiting case in which the superconducting properties are *amplified*, to assess the effect of the highly correlated state in the spatial properties.

The limit case of highest correlation in the superconducting state corresponds to when $\Delta \to \infty$ and $T \to 0$. The first condition can also be reformulated as $T_c \to \infty$. From eq. (5.14), one can already infer that the occupancies will go to 1 in those places where $\Delta(\xi) \neq 0$. Taking the very large value of $T_c = 5000$ K and a low temperature of $T = 10$ K, we obtain the occupation numbers shown in the inset of Figure 5.23.

Surprisingly, even at this limiting case the profile of the real-space functions in the symmetric chain in the SC state, namely the electron density, the KED, and the ELF seem indistinguishable from those of the normal state. We do not display them here to avoid redundancies with respect to Figures 5.12 and 5.13. The anomalous density does increase in this limiting case, but its magnitude remains far below that of the electron density, see Fig. 5.22.

The analysis of the LIs and DIs, on the other hand, shows a more clear tendency. For the LIs, we obtain the values of $\lambda_{NS}(T = 0K) = 0.45654$, $\lambda_{NS}(T = 10K) = 0.45653$, and $\lambda_{SC}(T = 10K) = 0.45645$. Although the differences remain very small, there seems to be a tendency for higher localization in the superconducting state. This is further supported by the higher decay rate of the SC DIs, as it becomes clear from Figure 5.23.

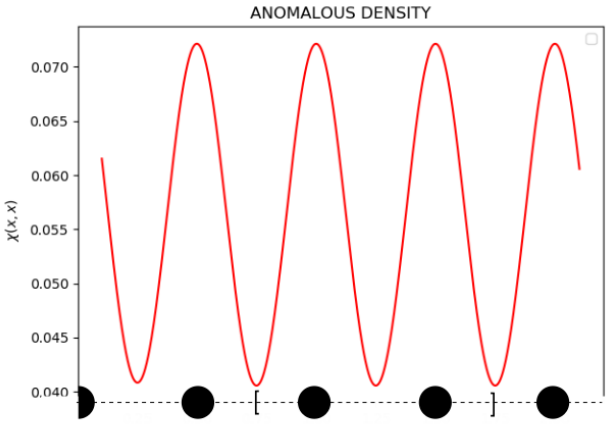


Fig. 5.22: Anomalous SC density, $\chi(x)$, along the hydrogen chain for the amplified gap ($T_c = 5000$ K). The atomic positions are marked by the black circles, with the squared brackets delimiting the unit cell.

Even if there seems to be a tendency for more localized ELF basins in the superconducting state, it is very weak considering that the chosen values for the parameters

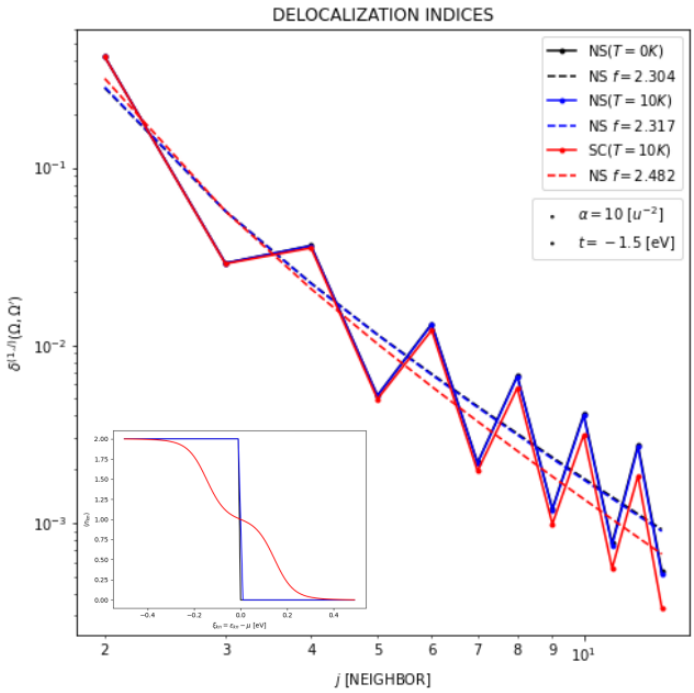


Fig. 5.23: Delocalization indices in the NS at 0 K (black) and at 10 K (blue), and for the SC state at 10 K (red). The dashed lines shows the best algebraic fit for each curve, and the inset the occupation numbers of each state, in the same color coding.

were greatly amplified. In this sense, it is not clear if is a product of numerical error, of the approximations made in the tight-binding model, or if in fact it is a physical result. At this stage, it becomes important to migrate from this model and to explore other methods for the evaluation of the real-space SC quantities. As a perspective, it would be particularly interesting to apply the formalism introduced in Section 5.2 to real hydrogen-based superconductors.

Part III

Real-space descriptors in hydrogen-based superconductors

6

Insights into superconducting properties using the ELF

This chapter focuses on the topological analysis of the ELF in real hydrogen-based superconductors. In Section 6.1, we briefly present the results of this analysis for a large database of binary hydrides, leading to the recognition of different bonding families and the definition of the *networking value*, ϕ, that correlates with the superconducting temperature. Then, in Section 6.2, we delve into the details of the implementation of an algorithm to automatize the computation of ϕ, and hence of T_c; and compare the results to those previously obtained by visual inspection. Finally, in Section 6.3 we present the results obtained by applying the program to a new set of ternary systems. Further, the analysis of the ELF in those structures conduct us to define the *molecularity index*, another descriptor that allows for a better characterization of the bonding patterns in hydrogen-based superconductors. We finish by proposing some machine-learning models that serve as an example of the perspective uses and expansions of this work.

6.1 ELF in superconducting hydrides

In this section, we briefly present the results and conclusions of the first-principles study of the ELF in a dataset of 132 hydrides, as they are presented in Ref. 110. A detail of the systems considered for this analysis can be found in Appendix B.

6.1.1 Chemical bonding patterns

The visual inspection of the ELF isosurfaces of the superconducting hydrides reveals certain chemical bonding patterns. This permits to classify the structures into six different bonding families, that are mainly characterized by the shape of the ELF or its value at the CPs between two hydrogen atoms. The classification criteria can be summarized as follows:

1. Molecular: the value of the ELF at the critical points between two hydrogens is above 0.8 [1], indicating the formation of H_2-like molecular units.
2. Covalent: the dominant type of bond is covalent, characterized by hydrogen basins elongated towards the other atom.
3. Weak H-H: hydrogens connect to each other at ELF values between 0.45 and 0.8, forming weak covalent bonds between them.
4. Ionic: hydrogens remain isolated at low values of the ELF, even below 0.5, and there is an elevated mean extra charge per hydrogen. The latter should be at least above 0.3 a.u. [2].
5. Electride: electrons are highly localized in the interstitial regions, playing the role of a pseudo-anion, with maximum ELF values between 0.35 and 0.7.
6. Isolated: contains all the remaining systems, where the hydrogens are completely isolated, with ELF values at the critical point below 0.25.

An example of a system from each of those families is provided in Fig. 6.1.

Comparison with the value of the critical temperature of each structure, as obtained from the literature, shows that the Covalent and Weak H-H bonding patterns are more likely to lead to a high-T_c compound (see Fig. 6.1). Further, the chemical origin of the different bonding patterns is clear when the non-hydrogen (or host) atom location in the periodic table is taken into account, as proven in Fig. 6.1. This is a clear step forward in the understanding of the chemistry of superconducting hydrides, as it allows us to propose new promising superconductors by choosing the correct host atoms from the periodic table.

6.1.2 The networking value

After the careful visual study of the ELF isosurfaces of the superconducting hydrides in the database, it became clear that there was a trend between the critical temperature and some of the topological features of the ELF. To make this correlation more evident, and useful, we introduced the *networking value*, ϕ. This quantity is defined as the maximum value of the ELF for which the f-domain remains periodic in three dimensions, which can be revealed by constructing a bifurcation tree for the three-dimensional periodic system of interest.

Figure 6.2 illustrates this principle for YH_9. At $f = 0.72$, some ELF basins are connected leading to H_2 molecules (Fig. 6.2-top-left). This is schematized in Fig. 6.2-top-right. As the ELF value diminishes, more connections appear. At $f = 0.57$ (Fig. 6.2-bottom) the volume in this isosurface (associated with all points such that $f \geq 0.57$) is connected. This is the networking value, $\phi = 0.57$.

In contrast to other descriptors obtained from the ELF that were introduced in Section 4.2, as the average population of a basin or the fluctuation between basins, that capture the behavior of electrons locally, the networking value describes a global picture of electron localization in the lattice. In fact, the only way of finding a high

[1] This value was set to 0.85 in the original paper, but in this thesis we lower it to 0.8 for reasons that will be explained in the chapters to come.

[2] This value was set to 0.5 a.u. in the original reference.

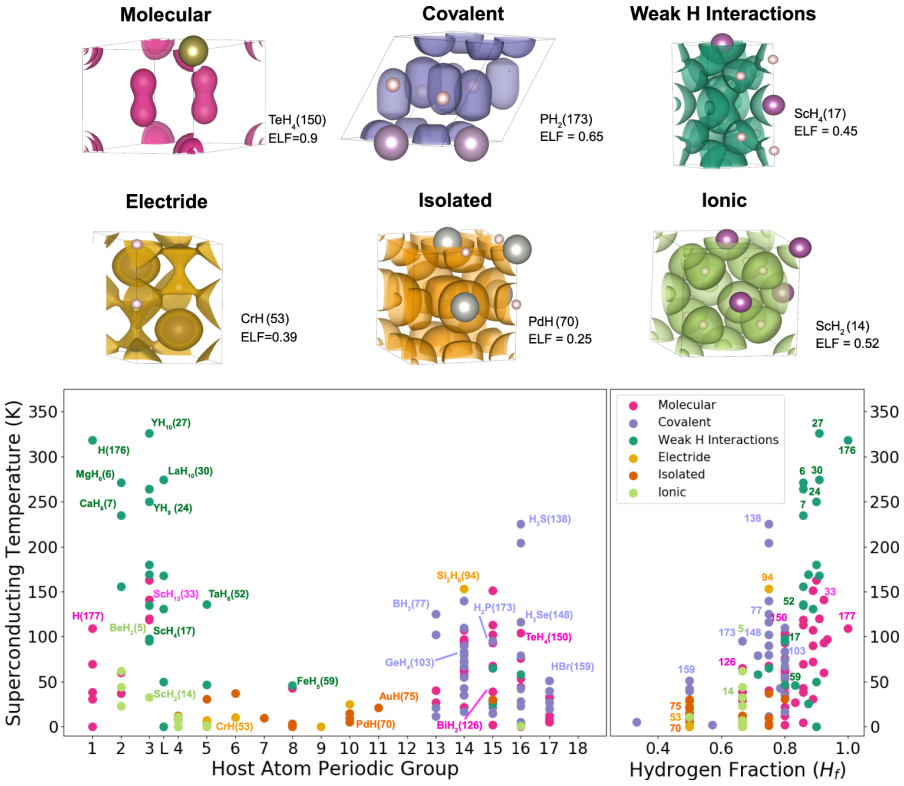


Fig. 6.1: Top: ELF isosurface for one structure belonging to each of the six bonding families. Bottom: Superconducting temperature, T_c, with respect to the periodic group of the host atom (left); and with respect to the fraction of atoms in the unit cell that are hydrogens, H_f. Reprinted with permission from Ref. [110].

networking value is if the value of the ELF at the bifurcation points connecting the hydrogens are all high. In such a case, the basins associated to the connected hydrogens would show high ELF values at the separatrix, and one would expect the delocalization inside the f-domain to be high, too.

The importance of the networking value comes from its relation to the critical temperature of hydrogen-based superconductors. A strong correlation between the two is observed [110], as evidenced in Fig. 6.3. This provides a much more efficient way of estimating critical temperatures, as the ELF can be obtained practically for free after a SCF calculation is performed, unlike the usual methods for the estimation of T_c (see Section 3.2). This is true even if the correlation is not very sharp, as it can be used for predicting T_c for large amounts of systems. It can however be improved if two key quantities are taken into account: the fraction of atoms in the unit cell and the fraction of the density of states (DOS) at the Fermi level that correspond to hydrogen, H_f and H_{DOS}, respectively. Including these quantities is not trivial, as

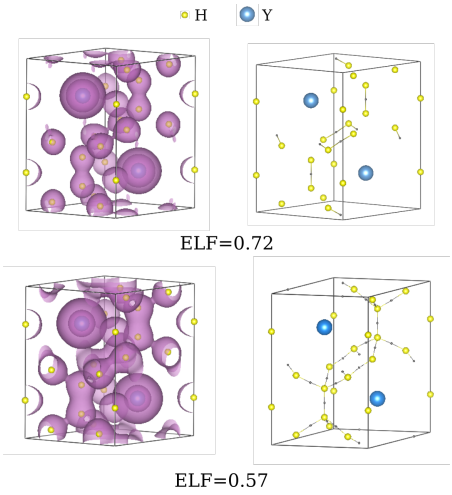


Fig. 6.2: To the left, the electron localization function surfaces for the isovalues 0.72 (top) and 0.57 (bottom) for $P6_3/mmc$-YH_9. The connections between the Hydrogen atoms basins on the left are sketched on the right pannel, where the saddle points that mediate those connections are marked in dark grey, for each isovalue.

it is typically observed that hydrogen-rich systems with a high DOS at the Fermi level with H-character tend to be better superconductors. In this way, the correlation improves significantly by defining

$$\Phi_{DOS} = \phi H_f \sqrt[3]{H_{DOS}}, \tag{6.1}$$

as can be observed in Fig. 6.3. This allows for a quick estimation of the critical temperature through the linear relationship

$$T_c = (750\Phi_{DOS} - 85)K, \tag{6.2}$$

that minimizes the error in the database for all the range of temperatures.

6.2 TcESTIME: automatic computation of T_c

Automatizing the search of the networking value is the missing piece for an efficient computation of the critical temperature. The TcESTIME algorithm has been developed for this purpose. The key feature of the networking value ϕ is that, for $f = \phi$ the associated ELF f-domain is connected (in the topological sense i.e. has no disjoint components), and thus expand in all three dimensions. To evaluate this connectedness, we build an associated complex network for each candidate networking value, $\tilde{\phi}$, where the nodes are the attractors of the ELF in the unit cell, and the edges represent a gradient path passing through a saddle point and connecting two attractors. Then, we define the periodic connectivity of the network as follows:

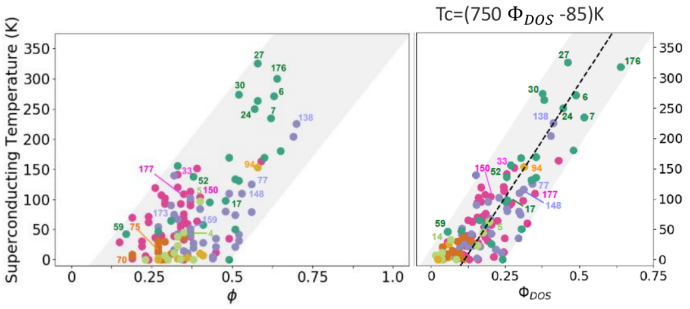


Fig. 6.3: Correlation between the superconducting critical temperature, T_c, and the networking value ϕ (left), for a database of 132 hydrogen-based superconductors. The correlation is improved by defining $\Phi_{DOS} = \phi H_f \sqrt[3]{H_{DOS}}$ (right), with H_f the fraction of hydrogen atoms in the unit cell, and H_{DOS} the fraction of the density of states at the Fermi level of hydrogen character. Reprinted with permission from Ref. [110].

1. In 1D : there is at least one path that connects a node to one of its translated analogues in another unit cell.
2. In 3D : there must be three such paths in three different, linearly independent, directions.

In section 6.2.1 we will begin by showing how to build the complex networks. Then, because the crystal lattice is periodic and our computational resources are finite, taking into account all the nuclear attractors is impossible and we will show how the the analysis can be reduced to the unit cell. In section 6.2.2, we will compare the results of TcESTIME with those obtained by evaluating ϕ by visual inspection in ref. 110, and provide a new correlation between Φ_{DOS} and T_c. Then, in Section 6.3, we will show how the algorithm works in an unseen dataset of ternary systems. There, we will see how we can improve our predictions by looking at other information contained in the ELF, and using a Machine Learning approach for the fitting. Finally, an interface between TcESTIME and a structure prediction algorithm is presented in Section 6.2.3.

6.2.1 TcESTIME: the algorithm

The networking value will be determined through the topological analysis of the ELF, thanks to the program `critic2` [116, 117]. This code allows reading grid data as the ones obtained from Quantum ESPRESSO [118, 119] (used along this application). Using a Newton-Raphson algorithm, it allows the localization of the ELF CPs: maxima located in the nuclei (nuclear attractors), other maxima corresponding to the atomic shell structure and valence (non-nuclear attractors), and saddle points corresponding to bonding regions, i.e. CPs that are minima in one direction and maxima in the other two (bond critical points) [116, 117]. Each Bond Critical Point (BCP)

is connected to two attractors, be it Nuclear attractor (NUC) or Non-Nuclear Attractor (NNA), inside an isosurface of isovalue that is at most equal to the value of the ELF in the BCP (f-domain, see Section 4.2). These attractors can be either in the same unit cell, or in a neighboring one. With this information, it is possible to transform the typically analyzed ELF isosurfaces (Fig. 6.4-top-left) into graphs (Fig. 6.4-bottom). Note how the NNAs (in blue) surrounding the S atoms (in green) connect the neighbouring H atoms (in red) for the $f = 0.69$ domain.

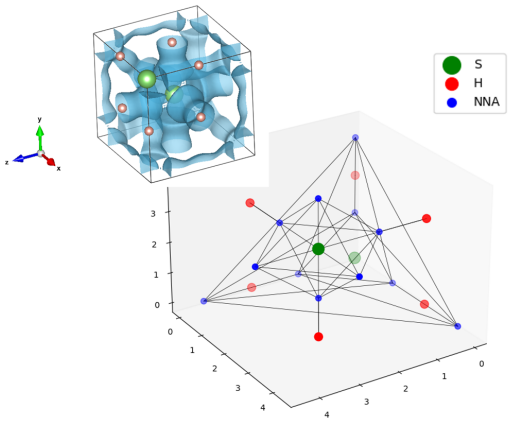


Fig. 6.4: Unit cell of $(H_2S)2H_2$ in the $Im\bar{3}m$ crystal structure, with S atoms in green and H atoms in red. On the top-left corner, the ELF isosurface for the networking value $\phi = 0.69$ is shown in blue. At the bottom, the representation of the same system in the complex networks created by TcESTIME that connect the atoms and the NNAs in the unit cell.

Two main issues were observed when constructing the network, mainly arising from the use of grids:

1. Due to the rapid variations of the ELF in the core shells, some connections between NNA's in the cores and the corresponding nuclear attractor are missed by `critic2`.
2. The edges of the graphs are very sensitive to the tuning of the `critic2` parameters, namely the distance thresholds for which two CPs are considered to be the same one.

Effective core radii

Although shell ELF attractors are spherically degenerated, they become non degenerated when the atoms are coordinated [94]. An example is shown in Fig. 6.5a, where the maxima around the Cr are highly concentrated in the direction of the Hydrogens. This great variety and amount of critical points per unit cell sets a big algorithmic

difference between solids and molecules. Whereas the valence region has been found to be exhaustively tracked by algorithms analogous to those employed for the analysis of the topology induced by the electron density, missing connections between NUC's and core NNA's are a common feature in the networks that are built from the `critic2` and Quantum ESPRESSO output data. Hence, the complete characterization of the ELF topology in the solid requires a hybrid method, with a combination of core-valence approaches.

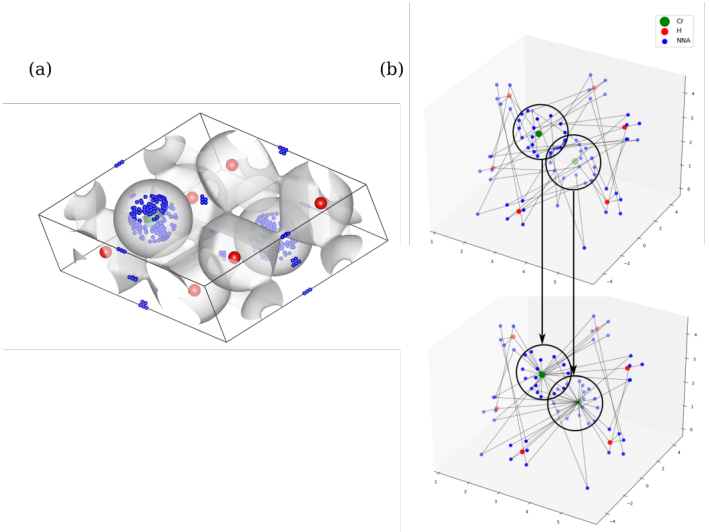


Fig. 6.5: (a) ELF isosurface at $\phi_{\text{ref}} = 0.35$ of the CrH_3 $P6_3/mmc$ crystal at 81 GPa. ELF non-nuclear attractors as found with the default `critic2` parameters are marked in blue, while the atomic positions of Cr and H are in green and red, respectively. (b) Graphs of the same system as represented by TcESTIME, with Cr and H nuclear attractors in green and red, respectively, and NNA's depicted in blue (computed with optimized parameters). The original graph on top is transformed into the one on the bottom when all NNA's inside an effective atomic radius are considered to be connected.

We use chemical intuition to avoid this problem and presuppose that all ELF maxima that correspond to the same core are connected, as it is in fact the core the one that needs to be part of the three-dimensional lattice. This requires the definition of an effective core radius, which can be done from the ELF of isolated atoms, as it has been shown that the values in the core region are quite rigid, staying practically unchanged upon bonding and induced pressure.[120] Core radii values, r_A, are tabulated for elements A with $Z \in [3,38] \cup \{48\}$ in ref. 121 (also available in Table C.1 of the Appendix). Hence, all critical points around a nuclear maximum of atom A and

within a distance r_A of it will be linked together, as shown by the black circles in Fig. 6.5b.

As it is usual, the case of Hydrogen is a special one, where we cannot define a core shell because its only (valence) electron participates in the bonds when the atom forms part of a crystal. Defining a nuclear attractor for H is not chemically correct then, and we should only find NNAs close to the nucleus. However, for the sake of simplicity, `critic2` does define a NUC for H, and this misconception can sometimes lead to having many NNAs close to the Hydrogens, and the connections between them and the bond CPs are easier to miss. Therefore, even though it does not have a chemical meaning to assign a core radius to H, doing so can be useful as it allows to group all those CPs that are found very close to the nucleus. We choose $r_H = 0.7a_0$, which corresponds to half the bonding distance of the H_2 molecule.

The definition of this radius is particularly convenient for systems where molecular hydrogen is present, as the profile of the ELF along the bond axis is very flat and it is more common that the algorithm fails to find the BCPs or to properly connect them to the NUCs and NNAs.

Tuning `critic2` parameters

The algorithm implemented in `critic2` for the search of the critical points of the ELF requires a set of thresholds that let the program identify two critical points as being the same or not. The default minimum distance to consider two CPs to be equivalent is $0.2a_0$. We have noticed that networking value is more accurately estimated when this threshold is raised to $0.3a_0$, which we have set as default in TcESTIME. This bypasses the need to get a finer grid when the algorithm oscillates and does not converge properly, and it avoids the unnecessary computation of too many equivalent CPs in highly symmetric systems (e.g. spherically symmetric maxima and saddle points in the cores).

Another parameter that can be tuned is the minimum distance between the nuclear coordinates of hydrogen and an ELF attractor for the latter to be considered nuclear. To avoid a large amount of NNAs to form around the hydrogens, we have raised this quantity to $0.6a_0$, which has shown to optimize the search of the networking value in the reference set. An example input file for `critic2` using these optimized parameters can be found in Appendix C.1.

These parameters can be changed by the user, as well as other more specific ones, as TcESTIME also allows to run `critic2` externally with the preferred options, using the resulting output file as input.

Reducing the analysis to the unit cell

Considering the definition of periodic three-dimensional connectivity given before, to determine ϕ we must search for paths that connect a node to itself in another unit cell. However, for each trial f this means using supercells composed of at least 8 cells of unit size, in order to check for all possible connectivities outside the unit

cell. We will try to simplify this picture thanks to periodicity. This is done in two steps: translating the graph to the unit cell, and then checking for closed paths.

To reduce the network to the unit cell, we associate a direction and weight to the edges, that will correspond to the translation vector of the connection. For example, if node 1 is directed towards node 2 with translation $\mathbf{t}_{12} = (-1,0,0)$ it means that, in every unit cell, 1 is connected to 2 in a neighboring cell to its left (see Fig. 6.6). It is worth noting that these directions are simply an attribute of the connection, and it does not mean that these are directed graphs in the complex network sense. (In the previous example, we do consider that node 2 is connected to 1, but the translation would change sign. This would not be true in a classic directed graph.)

With this "labelled" graph, we can translate all the information of the periodic connectivity to the unit cell while keeping the translational information.

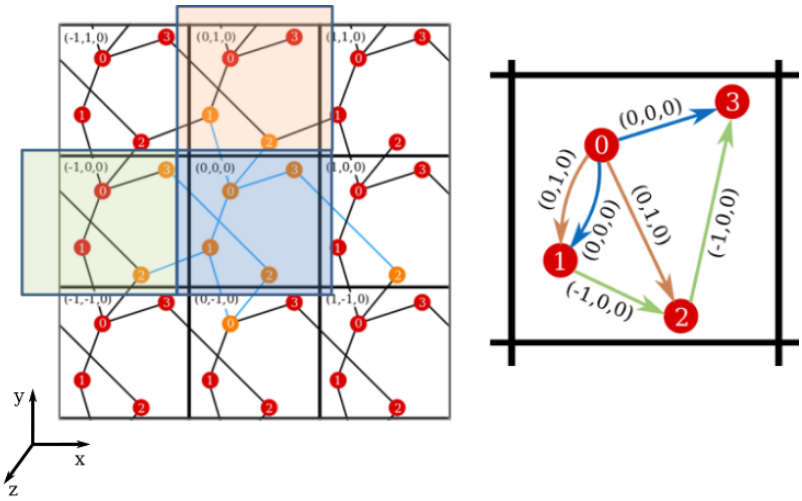


Fig. 6.6: Left: Connections between nodes in a single unit cell, (0,0,0), and those surrounding it. The nodes in orange are those in or connected to the (0,0,0) unit cell, forming the minimal network that captures periodicity. Right: Simplified representation of all the connections in the lattice in one unit cell. The network edges are labeled with the translation between two nodes. The color of the arrows represents the translation to the respective neighboring cell on the left panel.

Finding closed paths in the unit cell

Having defined the network in the unit cell, the connectivity along a direction is reduced to having closed paths P_i connecting a node i to itself within the unit cell. Considering all edges $k \in P_i$, the net translation of P_i is defined as

$$\mathbf{t}_{ii} = \sum_k s_k \mathbf{t}_k \,, \tag{6.3}$$

with $\mathbf{t}_k$ the translation vector of k and s_k the sign assigned that the edge within P_i, i.e. $s_k = 1$ if the path goes in the original direction of the edge, and $s_k = -1$ otherwise. Note that the first and the last edges must contain i: $k^1 = i, j$ and $k^{n_i} = l, i$, with n_i the number of edges in P_i. For the system in Fig. 6.6, we can identify a closed path $3 \to 2 \to 0 \to 3$, with a net translation of $\mathbf{t}_{33} = (1, -1, 0)$.

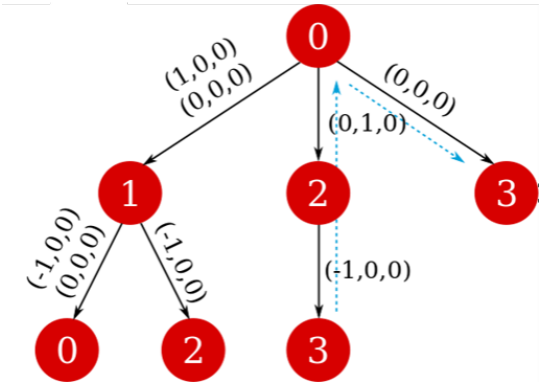


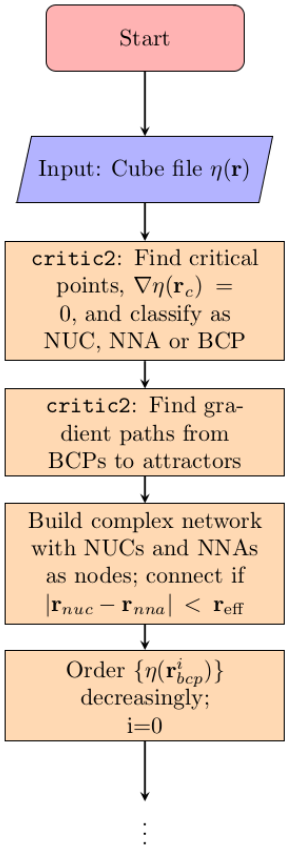
Fig. 6.7: Tree data structure of the network starting from node 0. The blue dashed arrows in each panel mark the closed path $3 \to 2 \to 0 \to 3$.

While it is rather easy to visually identify all possible closed paths for small systems, it is not straightforward to do so for larger ones. We propose an algorithm that is based in the breadth-first search algorithm [122], that builds the tree data structure of connections in a network starting from one node, that we call n_0. In the original algorithm, this is done by checking all the connections of the initial node, and creating a new branch for each one of them, which are put in a queue, while the initial node is marked as visited. The process is repeated for each of the nodes in the queue, which are then taken off of it when they are visited, and finishes when it is empty. Because in our networks there could be more than one connection between the nodes (see for example the double connection between nodes 0 and 1 in Fig. 6.6), we use the adjacency matrix, A, instead of the visited flag, where $A_{ij} = A_{ji}$ is equal to the number of edges connecting nodes i and j. The matrix is updated by subtracting 1 to A_{ij} (and A_{ji}) when a connection between said nodes is added to the tree. The translations of these connections are also taken into account and stored, taking care of its sign, as can be seen in the example in Fig. 6.7.

Once the tree starting from a node n_0 is found, the nodes n_r that appear more than once are identified. For those there will be a closed path in the network, and its net translation can be found by subtracting the net translation of each of the branches, up to n_r (see Fig. 6.7). If there is more than one connection between two nodes, we consider a different path for each of them, with different net translations.

TcESTIME workflow

The algorithm to find the networking value consists, in the end, of building complex networks for each candidate networking value, $\tilde{\phi}$, and assessing its connectivity. All of the most important considerations when building such complex network have been described in the previous sections. There are, of course, other details to take into account, like the possibility of having a disconnected network for a given $\tilde{\phi}$, composed of two or mode fully connected networks, that should be treated separately. The workflow of TcESTIME for the estimation of the networking value is thus summarized in the flowchart in Figure 6.1.



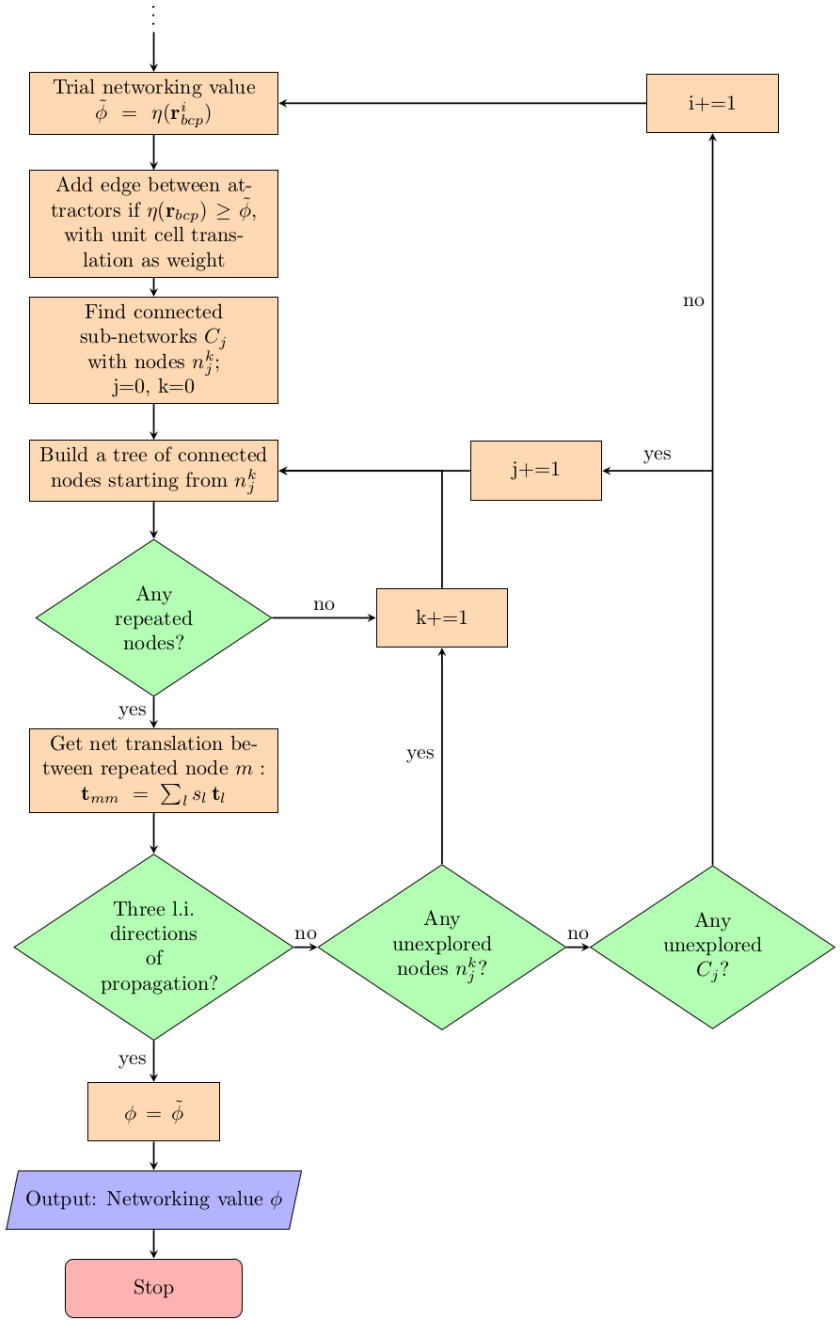


Table 6.1: Workflow of the program TcESTIME for the determination of the networking value.

6.2.2 Results for reference binary systems

The networking value was computed for an ensemble of 132 systems, corresponding to the hydrogen-based superconductors of ref. 110, from where the reference values ϕ_{ref} were taken. These values were originally visually determined, analyzing different ELF isosurfaces and finding the one for which a 3D network formed.

TcESTIME was able to find a networking value for 129 systems, with a mean absolute error of 0.04 with respect to the reference values ϕ_{ref}, a standard deviation of 0.08, and an average computation time of ca. 2s per system. A distribution of the difference $\phi_{\text{ref}} - \phi$ is reported in Figure 6.9, where we compare the estimations made with and without the considerations that were made concerning the cores: the core radius for non-H and H atoms, and the nuclear distance parameter for hydrogen, too. For simplicity, we call all these considerations the effective core radius r_{eff}. The $Im\bar{3}m$–$(H_2S)2H_2$ structure of Fig. 6.4 is a successful example for which TcESTIME found the exact expected networking value, $\phi = \phi_{\text{ref}} = 0.69$. TcESTIME was not able to find a networking value for a total of 3 systems, as reported in Table 6.2. A more detailed look into the ELF graph for the $C2/m$–LaH_8 crystal at the reference value $\phi_{ref} = 0.52$ (see Fig. 6.8), shows that it is likely that two unconnected bond critical points should have been considered as equivalent. A hand tuning of `critic2`'s parameters solves the issue for these problematic systems in most cases.

Table 6.2: Systems for which TcESTIME fails to compute a networking value, as indicated by their chemical formula and space group; and their expected networking value, ϕ_{ref}, and critical temperature, T_c.

Chemical formula	Space group	ϕ_{ref}	T_c (K)	Reference
LaH_8	$C2/m$	0.52	131	[123]
PoH	$P6_3/mmc$	0.34	0.65	[124]
H_2I	$Pnma$	0.36	5.3	[125]

Using the same values of H_{DOS} as in ref. 110, it is straightforward to compute Φ_{DOS}. The correlation between Φ_{DOS}, as computed by TcESTIME, and the reported critical temperature of the reference systems is recovered. Fig. 6.10 shows the linear fitting for the new data, yielding

$$T_c = (429.3\,\Phi_{DOS} - 10.4)K\,, \tag{6.4}$$

with a standard deviation of 46.1K. Note that both approaches, TcESTIME and visual search of ϕ, yield the most similar results for the structures with higher predicted T_c, which interest us the most.

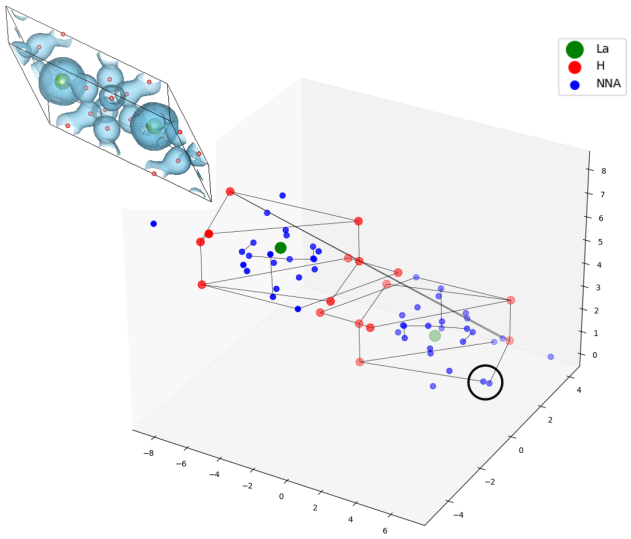


Fig. 6.8: ELF isosurface at the reference networking value, $\phi_{ref} = 0.52$, of the LaH_8 $C2/m$ crystal (top-left corner) and its graph representation as constructed by TcESTIME, with the La and H atoms in green and red, respectively, and the NNA's in blue. The black circle marks one of the regions where critical points fail to get connected.

6.2.3 Perspectives: interface with XtalOpt

TcESTIME can be used to quickly estimate the critical temperature of a given material. This is very useful when we already have a dataset of candidate superconductors, as it allows to efficiently identify the promising systems. However, as it is, the program does not allow to actually predict new high-temperature superconductors. Predicting new stable (and metastable) materials is usually done with crystal structure prediction (CSP) programs, that optimize the enthalpy of those systems for a given stoichiometry. XtalOpt [43] is one of them, and it consists of an evolutionary algorithm where individuals are the crystal structures, which can be selected or mutated from one generation to another, in each one of which the structures are relaxed with some *ab initio* program, e.g. Quantum ESPRESSO [118, 119] or VASP [126, 127, 128]. As the evolutionary algorithm passes from one generation to another, a fitness function is optimized. Usually, this function corresponds to the enthalpy of the stuctures, that is minimized to ensure thermodynamical stability. However, one can choose to optimize a multiobjective fitness function, allowing a simultaneous minimization of the enthalpy and maximization of the critical temperature. In this way, interfacing XtalOpt with TcESTIME permits to use the latter on every step of the evolutionary algorithm, and have an estimation of T_c that will be included in that multiobjective fitness function. The ability to define such fitness functions, and the

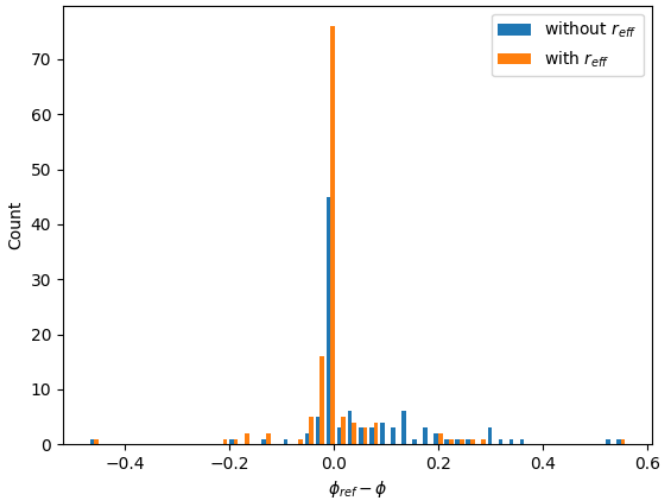


Fig. 6.9: Histogram of the error in the estimation of the networking value ϕ with respect to the reference value ϕ_{ref}. Results with (orange) and without (blue) the inclusion of effective core radii are presented.

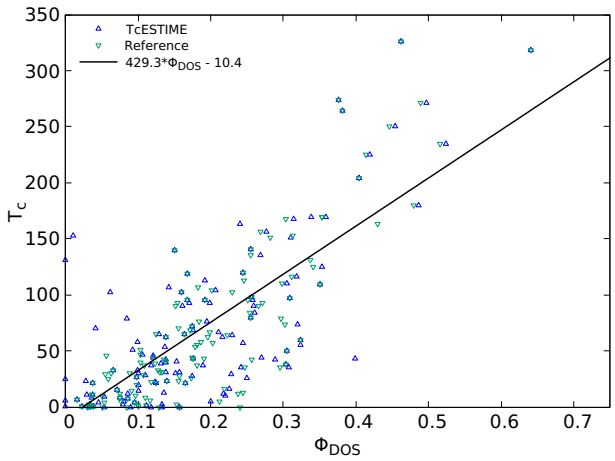


Fig. 6.10: Correlation between Φ_{DOS} and the superconducting critical temperature, T_c for the reference values of the networking value and those computed with TcESTIME. The black line represents the linear dependence of T_c on Φ_{DOS} according to TcESTIME.

fact that it is an open-source program, are some of the reasons why we prefer XtalOpt over other CSP algorithms.

Because the users of XtalOpt seem to prefer VASP for the *ab initio* calculations, the interface will be done using that package. Once the relaxation of an individual at a given generation is performed, and if the multiobjective option is activated in XtalOpt, an external script will be launched to estimate T_c. That script consists of a series of steps, summarized as follows:

1. A new single-point VASP calculation is launched to get the ELF field, using the atomic coordinates of the relaxed structures from the POSCAR that resulted from the VASP run performed by XtalOpt. This is stored in the ELFCAR file.
2. A non-self-consistent calculation is launched in VASP with a higher number of k-points (k-spacing of 0.1 Å^{-1}), to compute the DOS in a small window around the Fermi energy, E_F, as evaluated in step 1. This is stored in the DOSCAR file.
3. The DOS at the Fermi level for each atom is read from the DOSCAR and used to compute H_{DOS}, which is the fraction of the DOS at E_F that corresponds to hydrogen atoms.
4. The critical points and gradient paths of the ELF are found using `critic2` and the ELFCAR file.
5. TcESTIME reads the H_{DOS} value obtained in step 3 and the output of `critic2` to compute T_c.

With this, we hope to soon have access to a series of new predicted high-temperature hydrogen-based superconductors. Specially in the case of ternaries, it can be difficult to search for such materials, given the large amount of possibilities. In this way, TcESTIME opens up the possibility to exploit CSP programs for this aim, which was not possible before due to the cumbersome computations that had to be done to obtain T_c.

6.3 Application to ternary systems

6.3.1 General results

It has become evident that the search for high-temperature hydrogen-based superconductors has to expand towards ternary systems, a much broader space that has not yet been explored exhaustively. Because the correlation between Φ_{DOS} and the critical temperature was first observed in a large set of binary systems, the evident next step is to test it in ternaries. We have chosen a set of 21 X-RE-H hydrogen-rich compounds from the literature, with X and RE *s*-block and rare-earth elements, respectively. A list of these systems is provided in Table 6.3.

We have carried out the calculations for the ELF and DOS of these systems using DFT within the Kohn-Sham framework and the Plane-Waves Pseudo Potentials method as implemented in Quantum ESPRESSO[118, 119]. We chose the PBE scheme for the exchange-correlation functional,[59] including scalar relativistic effects. We set 80 Ry and 800 Ry for the wavefunction and density energy cut-offs, and sampled the Brillouin Zone with an unshifted 12x12x12 regular grid.

Table 6.3: List of X-RE-H systems (X is *s*-block element, RE is rare earth) in the dataset, including their chemical formula, pressure (GPa), space group, and superconducting critical temperature T_c(K), networking value ϕ, H_{DOS}, bonding type, and reference from where it was taken. The systems classified as Ionic and Molecular with an asterisk use a different definition to that of Belli *et al.* (2021), see Sec. 6.1.

Chem. formula	Pressure (GPa)	Space group	T_c (K)	ϕ	H_{DOS}	Bonding type	Ref.
$LiScH_{10}$	300	$R\bar{3}m$	52	0.30	0.28044	Molecular	[39]
Li_2ScH_{20}	300	$Immm$	242	0.36	0.74533	Molecular	[39]
Li_2ScH_{16}	300	$Fd\bar{3}m$	262	0.63	0.6838	Weak H-H	[39]
Li_2ScH_{16}	230	$Fd\bar{3}m$	281	0.63	0.68616	Weak H-H	[39]
Li_2ScH_{17}	300	$Fd\bar{3}m$	94	0.57	0.59984	Molecular*	[39]
Li_2LaH_{17}	300	$Fd\bar{3}m$	118	0.50	0.67135	Weak H-H	[39]
Li_2YH_{16}	300	$Fd\bar{3}m$	251	0.59	0.76919	Weak H-H	[39]
Li_2YH_{17}	300	$Fd\bar{3}m$	64	0.55	0.79807	Molecular*	[39]
$CaYH_{20}$	600	$P4/mmm$	250	0.62	0.81869	Weak H-H	[41]
Ca_2YH_{18}	200	$P\bar{3}m1$	217	0.59	0.75185	Weak H-H	[41]
Ca_3YH_{24}	200	$Fm\bar{3}m$	225	0.58	0.80000	Weak H-H	[41]
CaY_3H_{24}	200	$Fm\bar{3}m$	252	0.55	0.55800	Weak H-H	[41]
$CaScH_2$	250	$Fm\bar{3}m$	31	0.19	0.0237	Ionic	[129]
$CaScH_4$	200	$P6/mmm$	2	0.17	0.0978	Ionic*	[129]
$CaScH_6$	200	$P4/mmm$	57	0.36	0.60139	Ionic*	[129]
$CaScH_8$	200	$P4/mmm$	212	0.42	0.5569	Weak H-H	[129]
$CaScH_{12}$	160	$Pm\bar{3}m$	175	0.55	0.48617	Weak H-H	[129]
$BeLaH_8$	50	$Fm\bar{3}m$	191	0.29	0.66781	Weak H-H	[130]
$BeYH_8$	100	$Fm\bar{3}m$	249	0.12	0.69714	Weak H-H	[130]
$BeCeH_8$	30	$Fm\bar{3}m$	201	0.29	0.77857	Weak H-H	[40]
$BeThH_7$	20	$P6/mmc$	70	0.28	0.50172	Ionic	[40]

The values of the networking value according to TcESTIME have a mean absolute error of 0.02 with respect to the ones estimated visually, showing similar trends to those observed for the binary compounds. As can be seen in Fig. 6.11 the correlation between Φ_{DOS} and T_c still holds for the ternaries, although some of them do seem to further deviate from the tendency. This is normal considering that the definition of Φ_{DOS} was proposed considering only binary compounds, and that the former are more complex systems than the latter.

It is also useful to classify the systems according to the bonding patterns of the hydrogens, as was originally done by Belli *et al.*[110]. This is done by visual inspection of the ELF and using the same criteria as the reference (except for some small changes in ELF values for classification, see Sec. 6.1). In that work, two families were identified as being more prone to show high T_c's, namely the Covalent family, where the hydrogens form covalent bonds with the other atoms; and the Weak H-H family, with hydrogens forming a rather independent lattice where weak covalent interactions are dominant, and the other atoms act as chemical precompressors. The former is not present in our ternary dataset, which is in agreement with what was observed for binaries, where *s*-block elements and rare-earth with high critical tem-

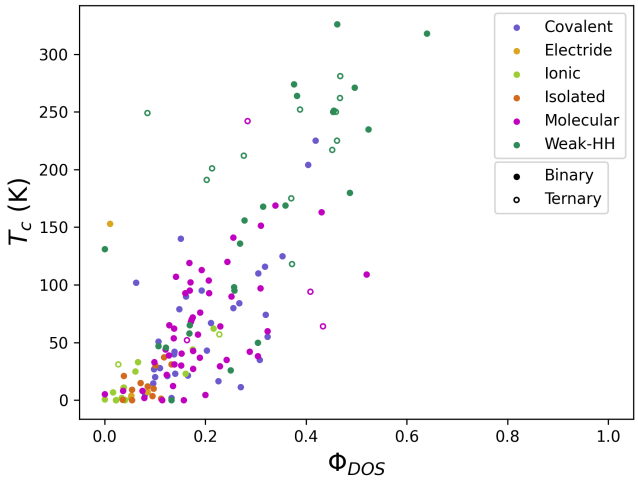


Fig. 6.11: Critical temperature, T_c, as reported in the literature, with respect to the corrected networking value, $\Phi_{DOS} = \phi H_f \sqrt[3]{H_{DOS}}$ for binary and ternary systems. The systems are classified according to the bonding type that is dominant in the system.

peratures belong to the Weak H-H family. This is confirmed for our ternary systems, as can be seen in Table 6.3 and Fig. 6.11.

A more careful inspection of the new systems can help us elucidate what kind of information the fit is lacking. In fact, we noticed that there are some systems that have the same symmetry and similar stoichiometry, yielding a very similar networking value, but with completely different critical temperatures. This is the case of Li_2ScH_{16} at 300 GPa, with $\phi = 0.63$ and $T_c = 281$K, and Li_2ScH_{17} at the same pressure, with $\phi = 0.57$ and $T_c = 94$K. The inclusion of H_f and H_{DOS} does not seem to help, as they give $\Phi_{DOS} = 0.47$ and $\Phi_{DOS} = 0.41$, respectively. The severe drop of the critical temperature upon inclusion of one hydrogen is understood when one examines the ELF isosurfaces: the hydrogens rearrange shortening the minimum distance between them, and as a result they form molecular units, which has been shown to be detrimental for high-temperature superconductivity [110, 111] (see Fig. 6.12).

Although the former conclusion is evident from the visual inspection of the ELF, it is not very practical if we would like to do a high-throughput screening of potential superconductors. We are thus in need of a descriptor that can account for this observed tendency to form molecular units.

6.3.2 Molecularity Index

We define the molecularity index, ϕ^*, as the maximum value of the ELF for which hydrogen atoms connect inside the same isosurface. In molecular systems, this will

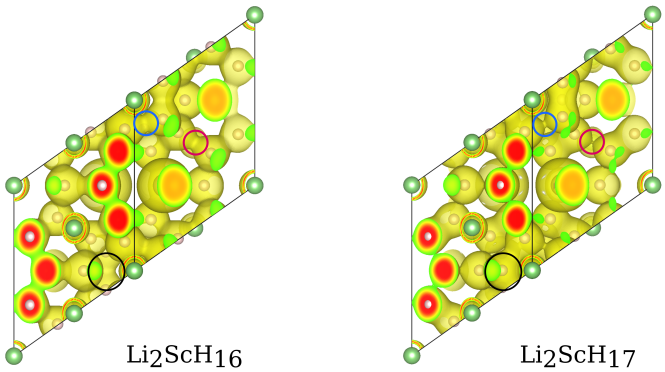
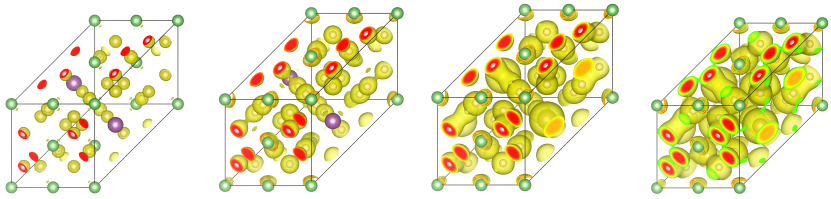


Fig. 6.12: Isosurface of ELF= 0.57 for Li_2ScH_{16} (left, $\phi = 0.63$) and Li_2ScH_{17} (right, $\phi = 0.57$) at 300 GPa. The black circle marks the place where the extra H is added. The blue and red circles show how the values of the ELF at the critical points change with this addition, becoming more prone to form molecular units.

necessarily correspond to the value of the ELF for which molecular units appear. Figure 6.13 illustrates the visual search for both ϕ^* and ϕ. In TcESTIME, the molecularity index corresponds to the highest ELF isovalue that gives a connected complex network containing more than one hydrogen nuclear attractor; or containing just one of them, along with a path that connects it to itself in another unit cell.



(a) $ELF = 0.99$ (b) $ELF = 0.83 = \phi^*$ (c) $ELF = 0.69$ (d) $ELF = 0.57 = \phi$

Fig. 6.13: ELF isosurfaces of Li_2ScH_{17} at isovalues (a) 0.99, (b) $\phi^* = 0.83$, (c) 0.69, and (d) $\phi = 0.57$.

Results for ϕ^* in binary and ternary systems can be seen in Fig. 6.14. As expected, ϕ^* is high for molecular systems, ranging between 0.8 and 1.0, where no other type of bonding family is present. We notice that the two bond categories that are most likely to have a high T_c, Covalent and Weak H-H, are dominant in the region where $\phi^* \in [0.45, 0.8]$, meanwhile all other types of bonding show $\phi^* < 0.45$ and generally low

critical temperatures. Hence, the molecularity index is a quantitative tool to separate the families of interest.

One can also look at the difference between the molecularity index and the networking value, $\Delta\phi = \phi^* - \phi$. A large $\Delta\phi$ means that the formation of the first H-H connections happens at a very different scale of ELF values than the formation of the three-dimensional periodic ELF network. We observe in Fig. 6.14 that $\Delta\phi$ shows a correlation with T_c, with a small $\Delta\phi$ being necessary but not sufficient for high T_c, similar to what was observed for H_{DOS} and H_f in binaries.

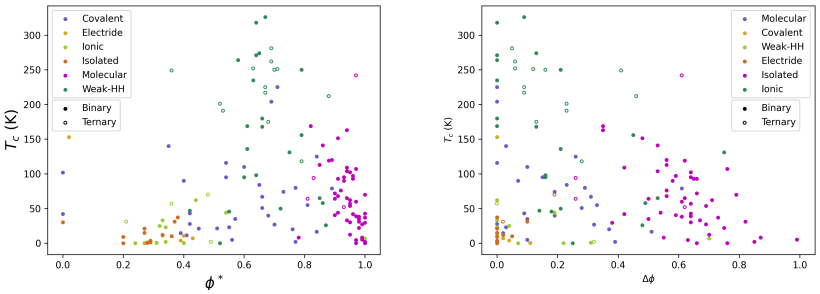


Fig. 6.14: Reference critical temperature T_c (K) with respect to the molecularity index ϕ^* (left), and to the different between that and the networking value, $\Delta\phi = \phi^* - \phi$ (center), for all binary and ternary data.

6.3.3 Machine learned model

In order to improve the prediction of new superconductors with the use of TcESTIME, we use Symbolic Regression (SR), an evolutionary algorithm where individuals correspond to mathematical expressions, as implemented in PySR [131]. These are optimized through generations in order to minimize the mean squared error of the evaluated expressions with respect to reference data. In this way, the output models are formulas for T_c. Symbolic Regression thus provides a way to do Machine Learning (ML) while retaining the scientific insight that a mathematical formula bears.

We have chosen five input quantities to be considered to compute T_c, for binary and ternary systems: ϕ, H_f, H_{DOS}, Φ_{DOS} and $1-\Delta\phi$, as they have all shown to be high in high-critical temperature compounds. Although the inclusion of Φ_{DOS} might be redundant, we can accelerate the search by adding it, since we already know that it correlates rather well with T_c. The output expressions have different accuracy and complexity, the latter being a measure of how many nested operators are present. Eligible unary operators were x^2, x^3, $\sqrt{x}$, $\sqrt[3]{x}$, and e^x; while binary operators could be chosen between $x+y$, $x-y$, and $x*y$. Three of the fits are displayed in Figure 6.15, where we notice that the simplest corresponds to a linear correlation between Φ_{DOS} and T_c, as expected. However, higher accuracy can be obtained from considering

expressions that include $1-\Delta\phi$, namely $T_c^{SR2} = 418.37(1-\Delta\phi)\sqrt{H_{DOS}}H_f^3$ or $T_c^{SR3} = 226.42H_f^3(-H_{DOS}^3 + e^{(1-\Delta\phi)})$. This inclusion aids in lowering the mean absolute error (MAE) also in the regions of high-T_c, where there tends to be a worse performance due to the less amount of data.

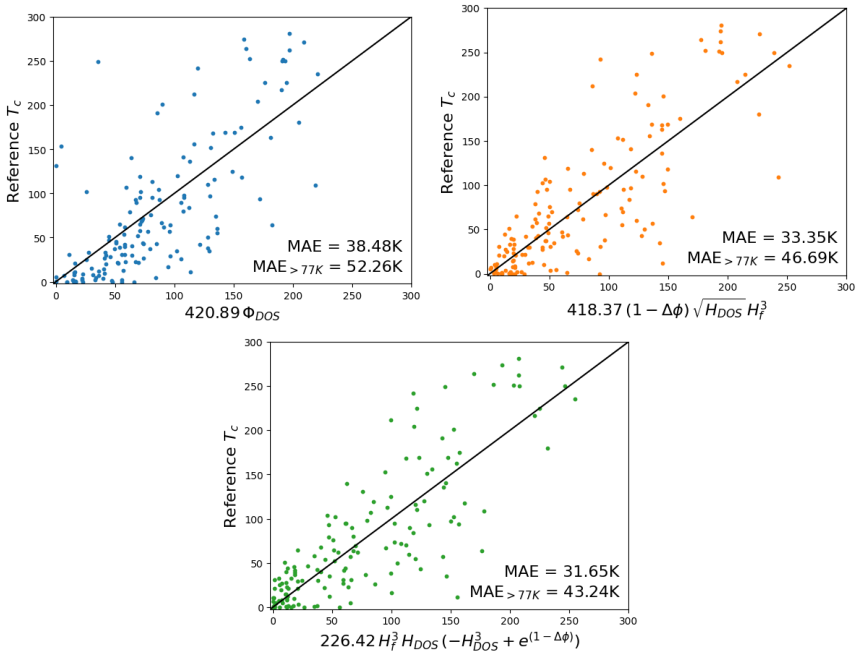


Fig. 6.15: Reference T_c (K) values with respect to its predicted values according to three different models trained with Symbolic Regression. The inset shows mean absolute errors for all considered data points (MAE) and for those with $T_c > 77K$ (MAE$_{>}77K$).

One thing that can be done to avoid this problem, is to filter the data to include a similar amount of high and low-T_c systems. Instead of doing it in a random way, we have chosen to filter keeping only the systems with $\phi^* \in [0.45, 0.8]$, which correspond to the bonding families that interest us. A much sharper correlation between Φ_{DOS} and T_c is observed with the remaining data (see Fig. 6.16). We thus trained a model using this data and the magnitudes ϕ, H_f, H_{DOS}, and Φ_{DOS}. Again, the simplest model found is a linear function of Φ_{DOS}, but there are also more promising expressions like $T_c^{SR4} = 249H_f(\phi^2 + H_{DOS})$, $T_c^{SR5} = 784\,\phi^3 H_f^3 + 175H_{DOS}$, and $T_c^{SR6} = 625\,\phi^3 H_f^2 + 219H_f H_{DOS}$. The overall errors in all of these models are much more consistent with those in high-T_c regions, thus being more reliable for our purposes. Although it is difficult to choose only one model, we do observe a similar trend among them, that confirms what was first proposed for binaries: the most im-

portant factor is always the networking value, ϕ, then the hydrogen fraction, H_f, leaving H_{DOS} as the magnitude with less weight. However, all three of them must be present.

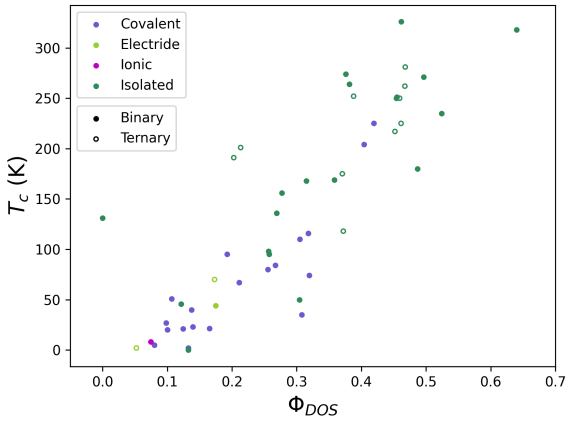


Fig. 6.16: Reference T_c (K) with respect to Φ_{DOS} for filtered systems with $\phi^* \in [0.45, 0.8]$.

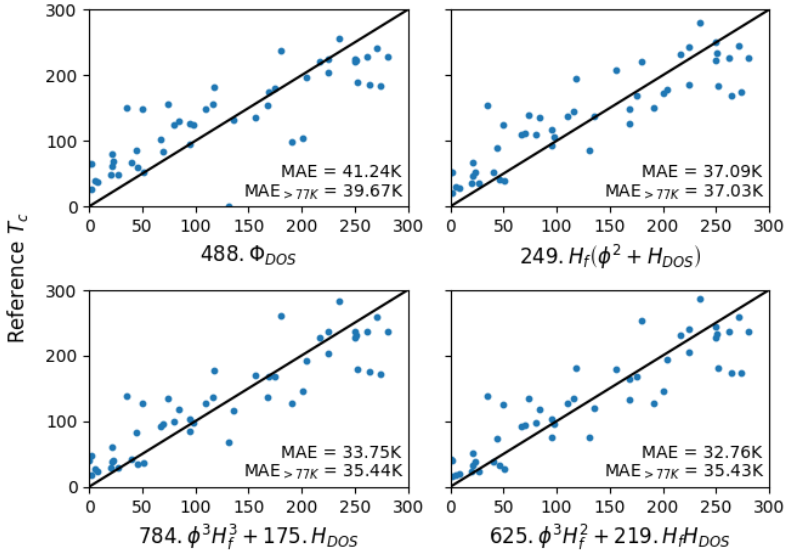


Fig. 6.17: Reference T_c (K) for systems with $\phi^* \in [0.45, 0.8]$, with respect to four different expressions found with Symbolic Regression for its prediction. Mean absolute errors for all considered data points (MAE) and for those with $T_c > 77K$ (MAE$_{>}77K$) are displayed for each model.

7

Fluctuations of effective potential for anharmonic perturbations

In Section 2.4, we showed that electron-phonon coupling, responsible for enabling the phenomenon of superconductivity in hydrides, can be measured through the screened electron-phonon coupling matrix elements, given by equation (2.37), that we recall here

$$g_{mnj}(\mathbf{k},\mathbf{q}) = \langle u_{m\mathbf{k}+\mathbf{q}} | \Delta_{\mathbf{q}j} v_{KS} \, | u_{n\mathbf{k}} \rangle_{uc} \, . \tag{7.1}$$

Broadly speaking, from looking at the expression for $g_{mnj}(\mathbf{k},\mathbf{q})$, we can say that there are two main factors that have a role in the magnitude of the electron-phonon coupling: the overlap between the electronic wavefunctions involved and the fluctuations of the potential that result from the ionic motion induced by a phonon. A spatial analysis of those quantities in real systems should help us understand where the phenomenon takes place and how that changes for different types of bonds, shedding light on the chemistry that is needed to conceive a good superconductor. This is exactly what has been done for the ELF, allowing to establish patterns among families of different bonding type; and to define the networking value, that showed to correlate with T_c. In fact, the analysis of the ELF is related to the overlap of the wavefunctions in the electron-phonon coupling matrix, which can be more clearly visualized when we express it in terms of Wannier functions instead of periodic ones, as in equation (2.40), i.e.

$$g_{mn\kappa\alpha}(\mathbf{R}_l,\mathbf{R}_{l'}) = \left\langle w_{m0}(\mathbf{r}) \left| \frac{\partial v_{KS}(\mathbf{r}-\mathbf{R}_{l'})}{\partial R_{\kappa\alpha}} \right| w_{n0}(\mathbf{r}-\mathbf{R}_l) \right\rangle_{sc} . \tag{7.2}$$

In that case, the overlap is between wavefunctions that are localized around ionic positions, similar to a molecular orbital, and the magnitude of such an overlap is related to the localization/delocalization of electrons in interstitial regions. In that sense, it contains similar information to that of the ELF, that depends on the electron density and, in the case if hydrogen, even shows a similar topology (see Section 5.3). It would therefore be interesting to study the other part of the electron-phonon coupling, i.e. the fluctuations of the KS potential, in real-space. In Section 7.1, we will present a simplified mechanism to evaluate the fluctuations of $v_{KS}(\mathbf{r},\mathbf{R})$ for a

large set of perturbed potentials that are obtained using the SSCHA. Then, in Section 7.1.1, we will show how the symmetries of the crystal structures can be exploited to reach convergence more quickly. We will present the computational details for computing the effective potential on each perturbed configuration in Section 7.2.2. Finally, in Section 7.3 we will present the results and analysis of the anharmonic fluctuations of the potential for five superconducting hydrides, and compare it to what is obtained for MgB_2, a well-known conventional superconductor that is not hydrogen-based.

7.1 Anharmonic fluctuations of effective potential

When a crystal structure is optimized using the stochastic self-consistent harmonic approximation, we have access to the Gaussian ionic probability distribution, $\tilde{\rho}_{\mathcal{R},\Phi}(\mathbf{R})$, that minimizes the free energy (see Section 2.4.3). With this, an ensemble of N_c perturbed configurations can be obtained, where the ionic positions are deviated from their equilibrium, $\mathbf{R}_0$, that in this case corresponds to the Gaussian centroids, $\mathcal{R}$. With $\mathbf{u}_I$ a random deviation obtained from the probability $\tilde{\rho}_{\mathcal{R},\Phi}(\mathbf{R})$, the ionic positions of the perturbed configurations become $\mathbf{R} = \mathbf{R}_0 + \mathbf{u}_I$. Then, for each perturbed system, labeled by $I = 1, \ldots, N_c$, it is possible to perform a DFT calculation and estimate the KS effective potential, $v_{KS}(\mathbf{r}; \mathbf{R}^0 + \mathbf{u}_I)$ (see Section 2.3). Because it is our interest to characterize the real-space properties of superconductors, we shall analyze the effect of the anharmonic fluctuations on the potentials by evaluating the spatial scalar field of the difference between $v_{KS}(\mathbf{r}; \mathbf{R}^0 + \mathbf{u}_I)$ and the equilibrium potential,

$$\Delta v^I_{KS}(\mathbf{r}) \equiv v_{KS}(\mathbf{r}; \mathbf{R}^0) - v_{KS}(\mathbf{r}; \mathbf{R}^0 + \mathbf{u}_I). \tag{7.3}$$

Then, to assess the effect of the anharmonic fluctuations on the potential for the ensemble of perturbed configurations, the standard deviation of $v_{KS}(\mathbf{r})$ will be used,

$$\sigma(v_{KS})(\mathbf{r}) = \sqrt{\frac{1}{N_c} \sum_{I=1}^{N_c} \Delta v^I_{KS}(\mathbf{r})^2}. \tag{7.4}$$

We expect this quantity to converge with increasing N_c, as the infinite ensemble limit is approached. Notice that $\sigma(v_{KS})(\mathbf{r})$ has the same units as the effective potential.

Care must be taken when choosing a supercell, as a large one is usually required to provide a correct description of the normal modes, and thus of the perturbed configurations. Using large supercells can become quite expensive if the number of configurations is high, considering that we must compute the effective potential for each one of them. It is thus important to assess the effect of the size of the supercell, and find a balance between that and N_c.

The standard deviation in eq. (7.4) was computed for a series of well-known hydrogen-based superconductors, namely PdH, H_3S, LaH_{10}, ScH_6 and $LaBH_8$. As it was mentioned, the same was done for a non-hydrogen-based superconductor, MgB_2, for comparison.

7.1.1 Symmetrization of the perturbed potentials

Hydrogen-based superconductors usually have crystal structures exhibiting high symmetry. The operations of symmetry, S, of a given system can be exploited to converge $\sigma(v_{KS})(\mathbf{r})$ much faster. In the limit of infinite perturbed configurations, if the number of symmetry operations is N_{symm}, for each configuration $\mathbf{R}_I$ we should find in the ensemble the other $N_{symm}-1$ rotated configurations, $S\mathbf{R}_I$. In this way, the standard deviation of the KS potential will have the same symmetry of the crystal. We can impose this symmetry, for example by creating every symmetry-equivalent configuration for each of those in the ensemble. However, even if the SSCHA code allows to do this directly, this can become very expensive, specially for large supercells. This is a consequence of the large number of Self-Consistent Field (SCF) calculations that would have to be performed, a total of $N_c * N_{symm}$ (one per symmetry operation, per initial configuration).

Another, less expensive, approach would be to take advantage of the symmetries of the system to build the scalar field $v_{KS}(\mathbf{r}; S\mathbf{R}_I)$ directly from the values of the non-rotated potential $v_{KS}(\mathbf{r}; \mathbf{R}_I)$, for an original perturbation $\mathbf{u}_I$. Notice that if a function only depends on the distance $|\mathbf{r}-\mathbf{R}_I|$, as is the case of the effective potential, rotating the positions $\mathbf{R}_I$ by S is the same as anti-rotating the coordinates $\mathbf{r}$, i.e.

$$v_{KS}(\mathbf{r}; S\mathbf{R}_I) = v_{KS}(S^{-1}\mathbf{r}; \mathbf{R}_I). \tag{7.5}$$

Equation (7.5) lets us recover the potential in all the unit cell for another perturbed configuration, $S\mathbf{R}_I$, by anti-rotating the grid in the original system.

Because the grid is discrete, making use of equation (7.5) is not straightforward, as one must match the rotated point $S^{-1}\mathbf{r}_i$ with a point in the original grid, $\mathbf{r}_j$, to access the value of the potential at that point. Also, care must be taken when rotating, as some of them will fall outside of the unit cell, and periodic boundary conditions must be applied to do the match. In what follows, we shall present a code that was developed to do this in an efficient way.

7.2 Computational details

7.2.1 Algorithm for symmetrization

The algorithm to find the symmetrical replicas of each perturbed configuration, $\mathbf{R}_I$, relies on the `spglib` Python package [132], which was used to find the matrices representing the symmetry operations of the unperturbed system. Then, a Python custom-made code was developed and employed to extract the grid and the values of the potential at each grid-point, $\{\mathbf{r}_i\}$ and $v_{KS}(\{\mathbf{r}_i\}, \mathbf{R}_I)$, respectively; and then write the N_{symm} new scalar fields, $v_{KS}(\{\mathbf{r}_i\}, S\mathbf{R}_I)$.

The input data of the code is provided by the program that is used to compute the potential, in this case Quantum ESPRESSO [118, 119], and is in a cube file format. The grid-points, $\{\mathbf{r}_i = (x_i, y_i, z_i)\}$, are written in crystallographic coordinates, i.e.

$$\mathbf{r}_i = x_i\mathbf{a} + y_i\mathbf{b} + z_i\mathbf{c}\,, \tag{7.6}$$

with **a**, **b** and **c** the lattice vectors. Moreover, we know that the points of the grid are placed along the lattice vectors, so that $x_i, y_i, z_i \in [0,1)$ (i.e. the grid is not cubic in Cartesian coordinates, unless the unit cell is itself cubic). Considering that the axes contain a number of N_a, N_b and N_c equidistant points, respectively, we define $\Delta\mathbf{a} = \frac{\mathbf{a}}{N_a}$, $\Delta\mathbf{b} = \frac{\mathbf{b}}{N_b}$ and $\Delta\mathbf{c} = \frac{\mathbf{c}}{N_c}$, so that

$$\mathbf{r}_i = k_i\Delta\mathbf{a} + l_i\Delta\mathbf{b} + m_i\Delta\mathbf{c}\,, \tag{7.7}$$

where k_i, l_i and m_i are integer numbers going from 0 to $N_a - 1$, to $N_b - 1$ and to $N_c - 1$, respectively. Analogously, we can write $\mathbf{r}_i = \left(\frac{k_i}{N_a}, \frac{l_i}{N_b}, \frac{m_i}{N_c}\right)$.

Finding a logic way to order the grid-points in a list is essential to facilitate the matching between the original and the rotated grid. We arrange them in such a way that the last component is the fastest, i.e. two consecutive points along the **c** axis will be consecutive in the list. After each line along said axis is finished, the list continues by summing one to the index l_i, corresponding to the axis along **b**, making the **a** axis the slowest one. In this way, the indices i in the list will be given by

$$i = N_b * N_c * k_i + N_c * l_i + m_i\,. \tag{7.8}$$

Now, let us consider a symmetry operation S of the unperturbed crystal. If there is a perturbation in the position of the atoms, that are now given by $\mathbf{R}_I$, this symmetry is lost, as portrayed in Fig. 7.1(a). Given a potential field evaluated in the grid over this perturbed configuration, $v_{KS}(\{\mathbf{r}_i\}, \mathbf{R}_I)$, we would like to infer the potential field for a rotated atomic configuration, $v_{KS}(\{\mathbf{r}_i\}, S\mathbf{R}_I)$, as in Fig. 7.1(c). To do so, we make use of the relation in eq. (7.5), and focus on computing $v_{KS}(\{S^{-1}\mathbf{r}_i\}; \mathbf{R}_I)$, as shown on Figure 7.1(b). For every point in the grid, $\mathbf{r}_i$, we can find its rotation $\mathbf{r}'_i = S^{-1}\mathbf{r}_i = \left(\frac{k'_i}{N_a}, \frac{l'_i}{N_b}, \frac{m'_i}{N_c}\right)$. Finding the index of this point in the list is straightforward using eq. (7.8):

$$i' = N_b * N_c * k'_i + N_c * l'_i + m'_i\,. \tag{7.9}$$

Once the grid-point has been found, its corresponding potential value $v_{KS}(S^{-1}\mathbf{r}_i; \mathbf{R}_I)$ is known, too.

In this way, one can recover the potential of a new perturbed configuration, $S\mathbf{R}_I$, without the need to do a self-consistent calculation. In a system like the superconducting PdH at 5 GPa, for example, where the space group is $Fm\bar{3}m$, the number of symmetry operations is 48. Thus, this symmetrization procedure allows to add 48 perturbed configurations to the ensemble per SCF calculation, representing a huge gain in computational time.

7.2.2 Methods: *ab initio* calculations

The *ab initio* calculations performed to find the effective potential of the perturbed configurations were done within DFT at the GGA level, using the Perdew–Burke-Ernzerhof (PBE) exchange-correlation functional [59] as implemented in the Quantum ESPRESSO package [118, 119]. The plane-wave basis cutoff for the energy was

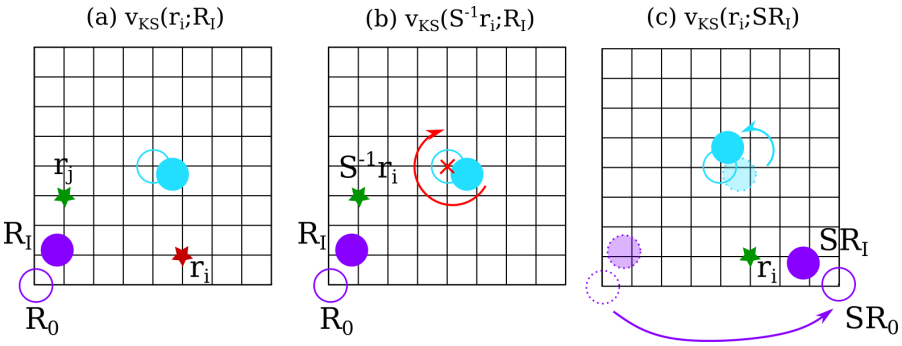


Fig. 7.1: Representation of the use of symmetry operations, S, to find $v_{KS}(\mathbf{r}_i; S\mathbf{R}_I)$ from $v_{KS}(\mathbf{r}_i; \mathbf{R}_I)$. In the three panels, there is a cubic unit cell containing two atoms. The unperturbed positions of those atoms, $\mathbf{R}_0$, are represented by empty circles, whereas the perturbed configuration positions, $\mathbf{R}_I$, are represented by filled circles. Grid points yielding different potential values are marked by stars of different colors. In (a), the original potential for $\mathbf{R}_I$ is known at grid points $\mathbf{r}_i$ and $\mathbf{r}_j$. In (b), the grid is rotated in $-90°$ by S^{-1} (red arrow) and the value of v_{KS} at $\mathbf{r}'_i = S^{-1}\mathbf{r}_i$ is the same as the value at $\mathbf{r}_j$ in the original grid, both in green. In (c), transparent and dotted circles mark the atomic perturbed and unperturbed positions before rotation, and solid and continuous circles are those after rotation by S (i.e. 90°). The value of the potential in the original grid at $\mathbf{r}_i$ takes the same value as $v_{KS}(\mathbf{r}_j, \mathbf{R}_I)$, in green.

set to 70 Ry for PdH, H_3S and ScH_6; and to 700 Ry for the density in the same systems. For LaH_{10}, on the other hand, lower values of 50 Ry and 500 Ry were needed, respectively; meanwhile for $LaBH_8$ they were of 90 Ry and 900 Ry. A comparative study of the effect of Pseudopotentials (PPs) on the scalar field $\sigma(v_{KS})(\mathbf{r})$ was performed for PdH and LaH_{10}, where both the Plane-Augmented-Wave (PAW) and Ultra-Soft (US) approximations were used. The choice of the PP for the rest of the compounds was done following the literature articles from where T_c values were taken. In this way, PAW pseudopotentials were used for H_3S, ScH_6 and MgB_2, whereas the ultra-soft approximation was employed in the study of $LaBH_8$.

7.3 Results for hydrogen-based superconductors

The standard deviation in eq. (7.4) was computed for a series of hydrogen-based superconductors, namely $Fm\bar{3}m$-PdH at 5 GPa [72], $Im\bar{3}m$-H_3S at 150 GPa, $Fm\bar{3}m$-LaH_{10} at 150 GPa [47], $P6_3/mmc$-ScH_6 at 100 GPa [133] and $Fm\bar{3}m$-$LaBH_8$ at 100 GPa [69]. For each of those systems, 1x1x1 and 2x2x2 supercells were considered, for comparison. A total of $N_c = 50$ perturbed configurations per system, as obtained from the SSCHA, were taken into consideration. After symmetrization, the number of total configurations thus became $50N_{symm}$, depending on the number of symmetry operators of each system.

7.3.1 Convergence analysis

The convergence of the scalar field $\sigma(v_{KS})(\mathbf{r})$ with respect to an increasing number of configurations, N_c, allows to visualize the effect of symmetrization on $\sigma(v_{KS})(\mathbf{r})$: a symmetrized version of the latter converges more rapidly, quickly reaching values that before were only attainable by increasing N_c to a much higher number. This is so because adding symmetric replicas allows to have a much more representative ensemble of perturbations. This is clear for the example case of the 1x1x1 cell of PdH in Fig. 7.2, where the isosurfaces of large deviation of v_{KS} ($\sigma(v_{KS})(\mathbf{r}) = 0.9$ Ry, located around the hydrogens) become more symmetrical after the procedure. In fact, before the procedure the isosurface looks quite irregular and it is easy to distinguish the directions of the perturbations that were used to compute $\sigma(v_{KS})$, see Fig. 7.2(c). After symmetrization, the isosurface in Fig. 7.2(d) is more uniform, and more accurately represents the average behavior of the fluctuations of the potential.

The evolution of the maximum of $\sigma(v_{KS})$ for increasing N_c is used to assess the convergence of the procedure. The meaning and importance of those maxima is further discussed in Sec. 7.3.2. The faster convergence for the symmetrized case is displayed in Fig. 7.3, where it is possible to see that large supercells suffer a more pronounced change upon symmetrization. On the other hand, the size of the supercell does not seem to change the results greatly, which is surprising considering that not having a large enough supercell can hinder the proper computation of phonon modes. It is, however, a very useful result, as it allows us to considerably reduce the computational cost of these calculations. Similar results are observed for convergence of H_3S, as shown in Figures 7.2 and 7.3. The convergence of the maxima of $\sigma(v_{KS})(\mathbf{r})$ for all the systems and supercell sizes are gathered in Figure 7.4 and 7.5.

The comparative study of the fluctuations of $v_{KS}(\mathbf{r})$ using different PPs yields similar results for both cases, in PdH and LaH_{10}. Although the maximum standard deviation changes slightly with the PP, it stays in a similar range, see Figs. 7.4 and 7.5. Other quantities, like the deviations around the non-hydrogen nuclei, also remain similar (see Table 7.1). One explanation for this would be that, since $\sigma(v_{KS})(\mathbf{r})$ is a measure of how the potentials change with the perturbations, it is differences and not absolute values of v_{KS} that are computed, which should be less prone to be affected by the chosen PP approximation. These observations justify the comparison of results that have been indistinctly obtained using either of the two approximations.

7.3.2 Topological analysis

The maxima of $\sigma(v_{KS})(\mathbf{r})$ are located at those places where the effective potential suffers the biggest changes as a consequence of the ionic vibrations. For a sufficiently large ensemble, it is expected that this maxima occurs at the equilibrium positions of the ions, or very close to them. In particular, because the hydrogen atom is so light, the amplitude of its vibrations should be larger, and the same goes for the potential fluctuations around it. This is exactly what we observe for the 2x2x2 supercell of H_3S in Figure (7.6). As the isovalue σ of the standard deviation increases, the isosurfaces begin to get localized solely around the hydrogen atoms. In fact, all of the studied

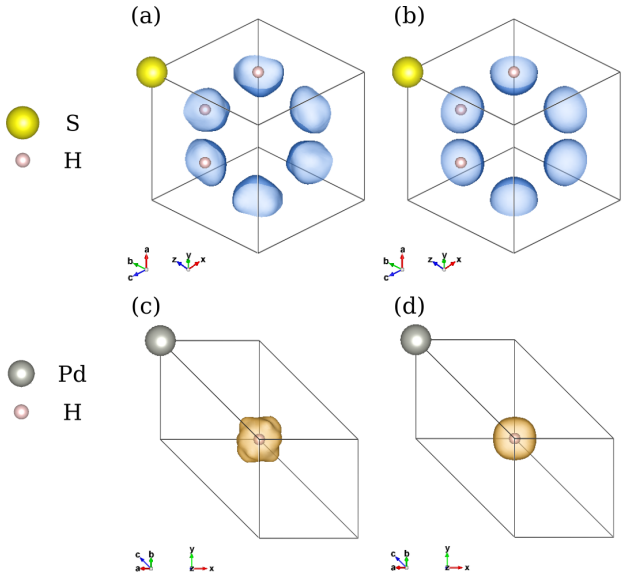


Fig. 7.2: Top: Isosurface of $\sigma(v_{KS})(\mathbf{r}) = 0.6$ Ry. in the 1x1x1 supercell of H_3S, for a number of (a) $N_c = 100$ perturbed configurations in the non-symmetrized case, and (b) $N_c = 50$ in the symmetrized case. Bottom: The same number of configurations are used for the (c) non-symmetrized and (d) symmetrized cases of PdH, respectively, for an isovalue of 0.9 Ry.

hydrides show this same tendency. This makes the maximum of $\sigma(v_{KS})(\mathbf{r})$ a good quantity to use for comparison between systems, to characterize the magnitude of the fluctuations at the hydrogen positions.

Table 7.1 summarizes the results obtained for all the hydrides, and for the 1x1x1 supercell of $P6/mmm$-MgB_2 [1]. We observe that almost all of the hydrides converge to a maximum $\sigma(v_{KS})(\mathbf{r})$ around 2.5 Ry. PdH is the exception, having a maximum deviation from the equilibrium potential of roughly 3.5 Ry in all cases. In contrast, a lower value of 1.77 Ry is found for this maximum in MgB_2, located on top of the magnesium. It is in fact expected that the deviations of the potential from its equilibrium value should be higher for hydrogen, the lightest of atoms, so the lower value obtained for MgB_2 is not surprising. In this sense, considering the low T_c value of 39 K of MgB_2, and keeping in mind the expression for $g_{mnj}(\mathbf{k},\mathbf{q})$ in eq. (2.40), one would expect that a higher maximum of $\sigma(v_{KS})(\mathbf{r})$ would be a fingerprint of high-T_c. However, PdH falses the hypothesis, with the highest deviations and a critical temperature of 5 K. The particularity of this system with respect to the others is its high degree of anharmonicity, as evidenced by the large anharmonic correction to the harmonic spectra found in ref. 72. The large fluctuations of the KS potential that we observe are consistent with this.

[1] Computations for MgB_2 in a larger supercell are in progress.

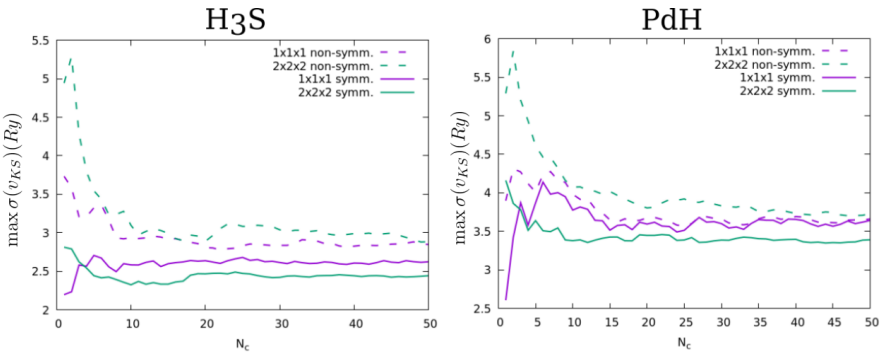


Fig. 7.3: Maximum of the standard deviation $\sigma(V_{KS})(\mathbf{r})$, located at the hydrogen atom positions, with respect to the number of perturbed configurations, N_c, for H_3S (left) and PdH (right). The non-symmetrized (dashed line) and symmetrized (continuous line) potentials are compared for two different supercell sizes are considered, 1x1x1 (violet) and 2x2x2 (green). In the symmetrized cases, each N_c actually corresponds to $N_{symm} * N_c$ perturbed configurations.

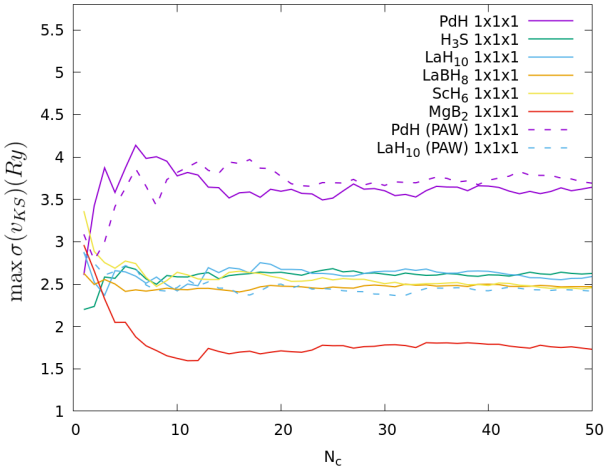


Fig. 7.4: Convergence of the maximum of the standard deviation of the KS potential, $\sigma(v_{KS})(R)$ with respect to the considered number of symmetrized perturbed configurations, N_c, for different systems in a 1x1x1 supercell.

The local maxima surrounding the other atoms in the unit cell was also evaluated, as shown in Table 7.1 (in the case of inequivalent atoms in the unit cell, where different values are found in atoms of the same element, the highest one was taken). The choice of PP does not affect the value of the atomic maxima significantly, the

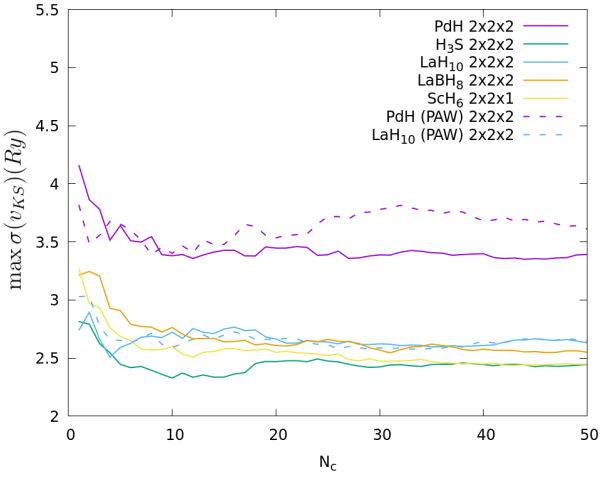


Fig. 7.5: Convergence of the maximum of the standard deviation of the KS potential, $\sigma(v_{KS})(R)$ with respect to the considered number of symmetrized perturbed configurations, N_c, for different systems in a 2x2x2 or 2x2x1 supercell.

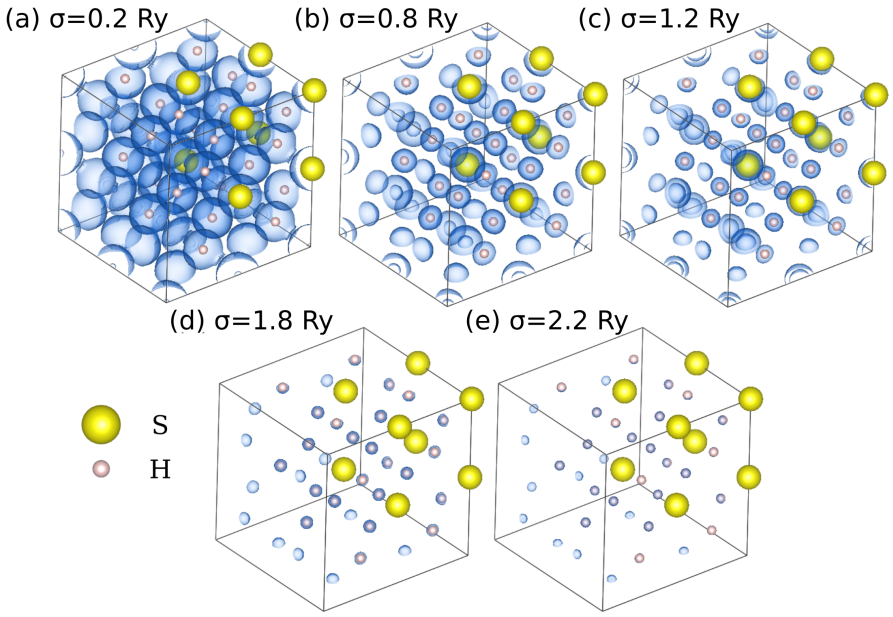


Fig. 7.6: Isosurfaces of $\sigma(v_{KS})$, depicted in transparent blue, in the $Im\bar{3}m$-H_3S lattice at 150 GPa for increasing isovalues: (a) σ = 0.2 Ry, (b) σ = 0.8 Ry, (c) σ = 1.2 Ry, (d) σ = 1.8 Ry, and (e) σ = 2.2 Ry.

difference being in fact negligeable. On the other hand, it becomes clear that the supercell does have a great influence on that maxima, which increases with the size of the supercell. In this sense, it only makes sense to compare results of similar supercell size, and we choose to take into account the larger one, as it better represents the phonon modes of the bulk. We notice that the maximum values around lanthanum and boron in $LaBH_8$ (0.43 and 0.79 Ry) are similar to those obtained for LaH_{10} (0.39 Ry) and MgB_2 (0.7 Ry), respectively. This, along with the repeated value of 2.5 Ry obtained on top of the hydrogens in almost all of the systems, suggests that the value of $\sigma(v_{KS})$ really close to the atoms has a high dependence on the element, which is expected from the dependence of the KS potential on the atomic number. However, this value is also determined by the amplitude of vibrations, and therefore has to depend on the atomic environment.

Table 7.1: Summary of critical temperature, supercell size, networking value, and topological descriptors related to the standard deviation $\sigma(v_{KS})$ for $N_{symm} = 50$ symmetrized configurations for different superconductors.

System (PP)	$T_c\,(K)$	Supercell	ϕ	$\max\sigma(v_{ks})$ (Ry)	Other at. maxima (Ry)
$Fm\bar{3}m$-PdH (US)	5	1x1x1	0.21	3.64	–
		2x2x2	0.21	3.39	0.28
$Fm\bar{3}m$-PdH (PAW)	5	1x1x1	0.21	3.69	–
		2x2x2	0.21	3.61	0.27
$Im\bar{3}m$-H_3S (PAW)	225	1x1x1	0.71	2.63	0.28
		2x2x2	0.71	2.43	1.40
$Fm\bar{3}m$-LaH_{10} (US)	274	1x1x1	0.56	2.59	–
		2x2x2	0.56	2.63	0.39
$Fm\bar{3}m$-LaH_{10} (PAW)	274	1x1x1	0.56	2.42	–
		2x2x2	0.56	2.65	0.38
$P6_3/mmc$-ScH_6 (PAW)	119	1x1x1	0.26	2.46	1.25
		2x2x1	0.26	2.45	1.57
$Fm\bar{3}m$-$LaBH_8$ (US)	100	1x1x1	0.33	2.47	– (La)/0.73(B)
		2x2x2	0.58	2.55	0.43(La)/0.79(B)
$P6/mmm$-MgB_2 (PAW)	39	1x1x1	0.22	1.77 (Mg)	0.70 (B)

It is also interesting to study the topology of $\sigma(v_{KS})(\mathbf{r})$ beyond its maximum values around atomic positions. In the larger supercells, one can separate the bifurcation diagrams of the hydrides into two categories. The first one contains PdH and H_3S, for which the isosurfaces go from fully connected to fully disconnected at a single value of $\sigma(v_{KS})$. For H_3S this happens at the isovalue of 0.2 Ry, as one can see from Fig. 7.6(a). On the other hand, bifurcation diagrams for LaH_{10}, $LaBH_8$, and ScH_6 have the common trait that one of the non-hydrogen atoms separates from the rest at very low values of $\sigma(v_{KS})$, whereas the isosurfaces surrounding the hydrogen (or the hydrogen plus the boron in $LaHB_8$) slowly break down into separated units. In the case of LaH_{10}, the lowest bifurcation point is at 0.10 Ry, above which the lanthanum is separated from the rest of the atoms, as can be seen in Figure 7.7(a) and in the

bifurcation diagram of Fig. 7.8. Then, above an isovalue of 0.37 Ry, the hydrogens separate into units of five atoms (see Fig. 7.7(b)). Finally, above 0.5 Ry, the isosurfaces of $\sigma(v_{KS})$ are separated and each one of them surrounds a single hydrogen atom (see Fig. 7.7(c)). The case of MgB_2 ressembles those of the second category, with an bifurctation value of 0.06 Ry separating the magnesium, and another at 0.16 Ry for the borons, which take longer to disconnect from each other.

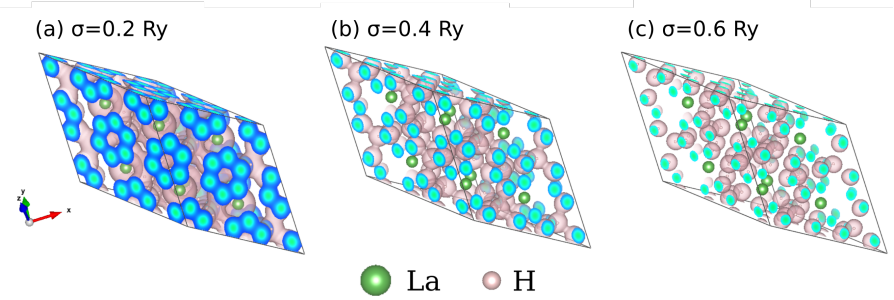


Fig. 7.7: In pink, sosurfaces of $\sigma(v_{KS})$ in the $Fm\bar{3}m$-LaH_{10} lattice at 150 GPa for increasing isovalues: (a) $\sigma = 0.2$ Ry, (b) $\sigma = 0.4$ Ry, and (c) $\sigma = 0.6$ Ry. The sections of the isosurfaces are colored according to the value of $\sigma(v_{KS})$ using an B-G-R color map, with the blue representing the lowest values.

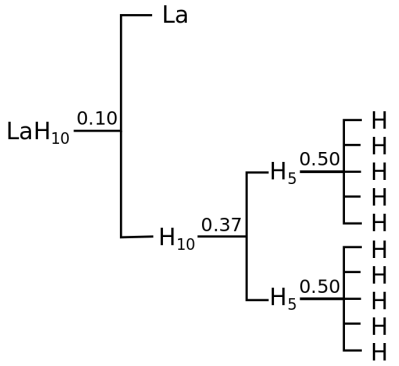


Fig. 7.8: Bifurcation diagram of $\sigma(v_{KS})(\mathbf{r})$ for LaH_{10}.

Understanding the trends observed in the bifurcation diagrams requires that we assign a meaning to the basins of $\sigma(v_{KS})(\mathbf{r})$. Unlike those of the electron density or of the ELF, here the basins do not necessarily coincide with regions of core and valence

electrons. For a point in space corresponding to a particular basin of $\sigma(v_{KS})(\mathbf{r})$, one can say that the changes of the potential at that point are due to the fluctuations of the atom in the basin. We hereby refer to the bifurcation point of $\sigma(v_{KS})$ at which an atom is separated from the rest as its *isolation value*. If, for a given atom, the isolation value is low, it means that its vibrations do not have a significant influence on the effective potential far from the nucleus. This is the case of lanthanum in LaH_{10} and $LaBH_8$, and scandium in ScH_6. However, if the isolation value is high, the fluctuations of the potential for those atoms cannot be separated into single-atom contributions so easily. This is the case of the hydrogens in the three previously mentioned hydrides.

Another trend in this category is that those three systems show clathrate-like structures, where the hydrogens form a cage around the heavier atoms, which probably serve mostly as chemical precompressors. What the bifurcation diagrams are putting into evidence is exactly this: the vibrations of the heavier atoms do not affect the fluctuations of the KS potential around the hydrogens, which in turn does not influence the electron-phonon matrix elements, $g_{mnj}(\mathbf{k},\mathbf{q})$. This is why they can be included without losing the superconducting properties of the system. This is not true for PdH and H_3S, where the hydrogen and the heavier atom share the same isolation value.

7.4 Perspectives

Although some things can be learned from the analysis of the fluctuation of the KS potential, the studies that have been performed so far are until now exploratory and work is still needed to draw more solid conclusions. In this line, considering the trends observed for the maxima of $\sigma(v_{KS})$ in the hydrides, and the deviation of PdH from that trend, we are particularly interested in verifying the hypothesis that this deviation is caused by the highly-anharmonic nature of the PdH crystal. To do that, we plan to do a comparative analysis of a harmonic and anharmonic version of $\sigma(v_{KS})$, where we expect that the highest differences among the two in the hydrides will be for PdH, from where its separation from the other hydrides' tendency should be understood.

On the other hand, further studying the relationship between the topology of the scalar field $\sigma(v_{KS})(\mathbf{r})$ and the tendency of certain elements to act as chemical precompressors seems to show some promise. The interpretation of the isolation value can be studied in more depth by assessing its magnitude in a given structure at different pressures, or by evaluating local pressures for different host (non-hydrogen) atoms.

Finally, actually computing the integrals of the electron-phonon matrix elements in real space, $g_{mn\kappa\alpha}$ in eq. 7.2, should help us understand the link between the networking value, the standard deviation of v_{KS} with respect to fluctuations, and the magnitude of the electron-phonon coupling in hydrogen-based superconductors.

8

Conclusions

The main goal of this thesis has been to provide a real-space description of hydrogen-based superconductors, with the aim of adding to the common knowledge of what makes up a good superconductor. We believe that our modest contribution to that knowledge has its strength in that it proposes to study superconducting materials from a different viewpoint, assessing their properties in the framework of QCT.

The first approach presented in this work is the study of the real-space properties of a model system. After discarding an initial model based on the Gor'kov wavefunction, that did not provide an energy scale nor reproduce the expected limit cases, we focused on a tight-binding one-dimensional hydrogen chain. Using the normal state of that system as a starting point, we constructed a hypothetical superconducting state on top of it, using the framework of SCDFT and a simplified expression for the superconducting gap.

Even if for temperatures well below T_c the occupation numbers of the normal and superconducting states were very different, the effect of this in the real-space functions was negligible. In fact, the topology of the electron density, the kinetic energy density, and the ELF did not change much when going from one state to the other, as it was observed both in the symmetric and the dimerized case. Surprisingly, this similarity persisted even at the limiting case of an amplified T_c (and Δ), where the superconducting state seemed to be only slightly more localized. On the other hand, the differences of the real-space properties between the symmetric and the dimerized hydrogen-chains allowed to characterize the respective metallic and insulating states better. It was not only the gap that granted a description of those conducting properties, but also the LIs, that increased when the intramolecular distance decreased; and the DIs, that showed a tendency to decrease faster for farther away basins, as it is expected for insulators.

The evolution of the topology of the ELF in the dimerized chain with respect to the intramolecular distances was assessed through ϕ and ϕ^*, that correspond to the values of the ELF at the minima in the intermolecular and intramolecular regions, respectively. A metallic state was described by similar values of ϕ and ϕ^*, whereas for the insulating case a bigger gap was measured when the intramolecular distances shortended, leading to a more evident pattern of molecular units, in which cases the

value of ϕ^* was very high with respect to that of ϕ. The differences in the superconducting states of those systems was better understood through the anomalous density, which is some sort of averaged Cooper pair wavefunction and an order parameter of the transition. Having its maxima on top of the atomic positions, and its minima in the interstitial regions, the average values of the anomalous density decreased with the intramolecular distance, showing that molecularization has a negative effect on the superconducting properties.

The following part of this thesis focused on the spatial properties of real structures of predicted or experimentally observed superconductors. The work began with the analysis of the ELF in binary superconductors, led by Francesco Belli as part of our collaboration with him and Ion Errea from the Materials Physics Center in San Sebastian. This work, that can be found in ref. [110], allowed to define the networking value, a topological descriptor that enriched our intuition on what makes a good superconductor; and to separate the hydrides into families according to their bonding types. It is not arbitrary that the symbol used to denote the networking value, ϕ, is the same than that used to name the minima of the ELF in the hydrogen chain. The latter is actually the one-dimensional version of the former, and its definition in the hydrogen chain helps understanding why it is possible to gain information about the superconducting state from the normal state ELF in hydrides, as it was shown that the ELF in both states is practically the same. Even if the negligible differences between the two were amplified, as a consequence of increasing the dimensions or using an approach other than tight-binding, it seems very likely that the correlation between them persists.

The networking value can also be used as a tool to get a faster prediction of the critical temperature of hydrogen-rich superconductors, leading to an immense potential for its use in the context of Crystal Structure Prediction (CSP). In order to make that possible, it was necessary to provide a computational tool to automatize the evaluation of the networking value. The TcESTIME program, developed for that purpose, allows to do such a computation in a very efficient manner, thanks to the complex-network-based algorithm that has been proposed in Section 6.2. The use of this computational tool to accelerate the predictions of hydrogen-based superconductors is now close, as the interface between TcESTIME and XtalOpt is under development, in the context of an ongoing collaboration with the groups of Eva Zurek (University at Buffalo) and Gilles Frapper (Université de Poitiers).

The analysis of the topological properties of the ELF beyond the networking value in ternary superconductors was possible thanks to TcESTIME. Even if the predictions of T_c in those systems using the networking value, H_{DOS}, and H_f, followed the trend observed for binaries, the deviations from the expected values were enlarged in ternaries. A close inspection of the ELF in the ternary systems allowed to identify how the molecularization of the hydrogens in the lattice led to a huge reduction of T_c, but not of the networking value. The molecularity index was thus defined, in analogy to the quantity ϕ^* from the one-dimensional model. The conclusions obtained in the hydrogen chain are in agreement with the observations in hydrides, and lead us to propose that the appearance of the H_2 molecular units hinders the superconducting transition because the normal state starts to resemble that of an insulator

in terms of electron delocalization, while it is known that a metallic state is needed as a precursor to conventional superconductivity.

The molecularity index also became a quantitative way to classify the systems of interest according to the bonding type, and to filter those that have potential as high-T_c superconductors. In this way, it is possible to use it to obtain a better fit for the prediction of the critical temperatures. The ML models obtained using SR presented in this work are an example of what can be done to improve the estimation of T_c for its use in CSP.

In general, the study of the ELF in the hydrogen chain and in hydrides reinforces a description that has already been presented in the literature: that a hydride containing a negatively charged hydrogen sublattice with elongated covalent H-H bonds enhances the superconducting transition [111]. In this thesis, we have shown some new approaches for the characterization of such a hydrogen sublattice. Although in the hydrogen chain a descriptor such as the shortest distance between hydrogens is able to fully characterize it and determine its potential for superconductivity, in more complex systems like hydrides this simple relationship is lost. In contrast, the combination of topological descriptors such as the networking value and the molecularity index seems to properly and fully describe such a hydrogen sublattice, in a global and in local manner, respectively.

Although the analysis of the ELF in binary and ternary superconductors allowed us to better understand the real-space properties of high-T_c hydrides, it is well-known that the critical temperature of those conventional superconductors should depend on other quantities that the ELF does not depend on, namely those related to the electron-phonon coupling. The effect of the anharmonic fluctuations in hydrides in the KS potential was thus analyzed using the standard deviation $\sigma(v_{KS})(\mathbf{r})$ for an ensemble of perturbed atomic configurations. The symmetries of the crystal structures were exploited to optimize the procedure and reduce the amount of SCF calculations that were needed for convergence. The analysis of the topology of $\sigma(v_{KS})$ on a set of hydrides showed some interesting trends. We observed that the potentials change the most at the position of the hydrogens, as expected, and that the values of the standard deviation at those points was almost always around 2.5 Ry, with PdH being the only exception. The possibility that this is a consequence of the high anharmonicity of the PdH structure has been proposed, and a mechanism to test this hypothesis by comparing the results to a harmonic version of $\sigma(v_{KS})$ has been offered in the perspectives.

The comparison of the values of those maxima in the hydrides with respect to the much lower values obtained for MgB_2 led us to propose the following hypothesis. If the hydrides were to always have a higher and almost constant value of the derivative of v_{KS} with respect to ionic displacements (that we find in the expression for $g_{mn\kappa\alpha}$), one could explain: (i) that the information contained in the part of $g_{mn\kappa\alpha}$ related to the electronic orbitals is sufficient to describe the hydride's superconducting properties, and (ii) that the magnitude of the electron-phonon coupling matrix element is always higher for the hydrides, explaining their tendency for higher T_c. Nonetheless, this remains a hypothesis, and it is important to pursue the lines of work proposed in the perspectives to get closer to a verification (or refutation) of that hypothesis.

The analysis of the bifurcation trees of the scalar field $\sigma(v_{KS})(\mathbf{r})$ for the different hydrogen-rich superconductors allowed to distinguish two types of systems: those where the basins around the atoms separate at once, and those where the topology is more persistent and the so-called isolation value of the hydrogens is much higher than that of the other atom(s). We propose that this could be a signature of systems where the host atom acts as a chemical precompressor, and does not take part on the superconducting transition enabled by the hydrogen sublattice. More work is needed to make sure that the previous affirmation is in fact correct.

All in all, the study of the fluctuations of the KS potential with respect to anharmonic perturbations shows a lot of promise. Although by the end of this thesis it was not possible to finish everything that we wanted to do in this line, and the scientific contribution is yet to be polished, I would like to highlight that it has contributed to my personal formation in more than one aspect. Firstly, having had to work with the symmetries of the crystals and to write the code helped me to feel more at ease with some concepts of crystal structures that I had not fully explored in the other parts of this work, and that I think are very important for my future scientific endeavors. Moreover, starting to learn about modern methods that allow to describe quantum fluctuations in solids, like the SSCHA, has been of special interest to me. Last but not least, proposing a real-space descriptor from zero, without fully understanding how to interpret its topology has been quite a challenging and interesting task, that shall be continued as we finish this part of the work.

Finally, I permit myself to comment on a few reflections that accompany me as a PhD student in the finishing stage of her thesis. A general conclusion that I will carry with me in the coming years, and that I would like the reader to take home with them, is that gaining intuition is crucial if we want to do prediction of new materials. Even if the properties of hydrogen-based superconductors can be predicted up to reasonable accuracy nowadays, it is our jobs as scientists to guide where we focus our research, and in order to do that it is necessary to have certain tools at hand that give us a broad picture of the problem in question, specially in cases like this where the computation costs are so high. QCT allows to infer some properties of the materials by analyzing electronic functions in the lattice environment. This is in fact a very powerful instrument, as those functions are not very expensive compared to more specific computations that might be needed to get exact quantities of those properties. For this to be of practical use, however, it is necessary to learn how to interpret the topology of the electronic functions in real-space. In this sense, this thesis falls perfectly within the scope of what we do in the Chemical Interpretation team at the Laboratoire de Chimie Théorique. A common thread throughout the different parts of this thesis is that we have had to describe and interpret what we observed in the results of our simulations. In other words, we have had to ask ourselves: what do these results mean? Of course, it has been our interest to answer this from a scientific viewpoint, and the work that I have been able to undertake throughout my thesis has been carried out with this in mind.

Appendix

A

Basics of Migdal-Eliashberg formalism

The effect of the electron-phonon coupling in a many-electron metallic system was approached by Migdal making use of Green functions in 1968 [19]. It was Eliashberg who later extended that formalism to superconducting materials [20]. In it, he considered both the electron and the phonon Green functions.

The formalism can be formulated as many-body perturbation theory [16]. It considers the zero order Hamiltonian

$$H_0 = H_s + H_{ext} \tag{A.1}$$

where the first term corresponds to the KS contribution

$$H_s = \sum_\sigma \int \psi^\dagger_\sigma(\mathbf{r}) \left[-\frac{\nabla^2}{2} + v_{KS}(\mathbf{r}) - \mu \right] \psi_\sigma(\mathbf{r})\, d\mathbf{r}\,, \tag{A.2}$$

and the second one considers an external potential, Δ_{ext}, that allows Cooper pairs to tunnel in and out of the system, thus breaking particle-conservation,

$$H_{ext} = \int \Delta^*_{ext}(\mathbf{r},\mathbf{r}')\psi_\uparrow(\mathbf{r})\psi_\downarrow(\mathbf{r}')\, d\mathbf{r} d\mathbf{r}' + h.c. \tag{A.3}$$

The first-order Hamiltonian, on the other hand, accounts for the interactions,

$$H_I = H_{ee} + H_{e-ph} - H_{DC}\,, \tag{A.4}$$

with H_{ee} the usual electron-electron interaction term, H_{e-ph} the electron-phonon interaction Hamiltonian of equation (2.36) (that we use to approximate the real electron-ion interaction), and

$$H_{DC} = \sum_\sigma \int \psi^\dagger_\sigma(\mathbf{r}) v_s(\mathbf{r}) \psi_\sigma(\mathbf{r})\, d\mathbf{r}\,, \tag{A.5}$$

which avoids double counting of the exchange-correlation already taken into consideration in H_s.

Perturbation theory cannot be used for the total Hamiltonian $H = H_0 + H_I$ as it is, because the loss of particle conservation does not allow to diagonalize H_0 by the field operators $\psi^\dagger_\sigma, \psi_\sigma$. The Nambu-Gor'kov formalism is a workaround for this problem, defining the new operators

$$\bar{\psi}(\mathbf{r}) = \begin{pmatrix} \psi_\uparrow(\mathbf{r}) \\ \psi^\dagger_\downarrow(\mathbf{r}), \end{pmatrix}, \quad \bar{\psi}^\dagger(\mathbf{r}) = \left(\psi^\dagger_\uparrow(\mathbf{r})\ \psi_\downarrow(\mathbf{r})\right). \tag{A.6}$$

This allows to rewrite the Hamiltonian in a matrix form, and to use perturbation theory and Green's functions in a standard manner.

The details of how the self-energy is estimated and how the system's Dyson's equation is solved are out of the scope of this thesis. However, it is worth mentioning that, since both electron and phonon propagators are taken into account, the origin of the attractive interaction is explicit, contrary to BCS theory. A few approximations can be made to simplify the procedure, e.g. the isotropic approximation, where one assumes that the gap in Fourier space does not depend anymore on the index n nor on the wavevector $\mathbf{k}$, i.e. $\Delta(n\mathbf{k}) = \Delta$. Further, if the gap is considered to be small, which is true around T_c, one obtains the linearized Eliashberg equations

$$\Delta(i\omega_i)Z(i\omega_i) = \frac{\pi}{\beta}\sum_{\omega_j}\frac{\Delta(i\omega_j)}{|\omega_j|}\left[\lambda(i\omega_j - i\omega_i) - \mu_c\right], \tag{A.7}$$

where μ_c is the averaged Coulomb interaction at the Fermi surface weighted by the density of states at the Fermi energy, $N(\epsilon_F)$, and

$$Z(i\omega_i) = 1 + \frac{\pi}{\omega_i\beta}\sum_{\omega_j}\lambda(i\omega_j - i\omega_i)\frac{\omega_j}{|\omega_j|}, \tag{A.8}$$

$$\lambda(i\omega_i - i\omega_j) = \int \frac{2\omega\alpha^2F(\omega)}{(\omega_i - \omega_j)^2 + \omega^2}\,d\omega, \tag{A.9}$$

$$\alpha^2F(\omega) = \frac{1}{N(E_F)}\sum_{\mathbf{kq}\nu}|g_{mn\nu}(\mathbf{k},\mathbf{q})|^2\,\delta(\xi_{n\mathbf{k}})\delta(\xi_{m\mathbf{k}+\mathbf{q}})\delta(\omega - \omega_{\mathbf{q}\nu}). \tag{A.10}$$

Equation (A.7) must be solved self-consistently to find the superconducting gap and the transition temperature, T_c.

B

Additional information of the ELF in hydrides

B.1 Networking value of binary systems

Table B.1: Reference value of the networking value, ϕ_{ref}, according to F. Belli *et al.* (2021), along with the value ϕ computed in this work, the relative error for each one of the 132 hydrogen-based systems, and the molecularity index, ϕ^*.

System	Chemical formula	ϕ_{ref}	ϕ	$\phi_{\text{ref}} - \phi$	ϕ^*
0	LiH_2	0.24	0.24	0	1.0
1	LiH_2	0.37	0.37	0	0.98
2	LiH_8	0.15	0.16	-0.01	0.98
3	KH_6	0.19	0.21	-0.02	1.0
4	BeH_2	0.35	0.35	0	0.54
5	BeH_2	0.4	0.44	-0.04	0.44
6	MgH_6	0.63	0.64	-0.01	0.64
7	CaH_6	0.62	0.63	-0.01	0.63
8	SrH_6	0.33	0.34	-0.01	0.79
14	ScH_2	0.33	0.33	0	0.33
15	YH_4	0.43	0.44	-0.01	0.6
16	YH_6	0.58	0.58	0	0.58
17	ScH_4	0.48	0.48	0	0.64
18	ScH_6	0.6	0.61	-0.01	0.61
19	PrH_9	0.49	0.27	0.22	0.52
20	CeH_{10}	0.51	0.53	-0.02	0.66
21	CeH_9	0.51	0.51	0	0.72
22	LaH_9	0.13	0.31	-0.18	0.95
24	YH_9	0.57	0.58	-0.01	0.79
25	ScH_9	0.65	0.66	-0.01	0.66
27	YH_{10}	0.58	0.58	0	0.67
29	LaH_8	0.52	-	-	0.75
30	LaH_{10}	0.52	0.52	0	0.65

31	ScH_9	0.59	0.59	0	0.94
32	ScH_{10}	0.33	0.33	0	0.89
33	ScH_{12}	0.33	0.33	0	0.86
34	ScH_7	0.49	0.47	0.02	0.82
37	ScH_6	0.26	0.26	0	0.88
39	TiH_2	0.29	0.3	-0.01	1.0
40	TiH_2	0.38	0.18	0.2	0.4
41	ZrH	0.32	0.33	-0.01	0.33
42	HfH_2	0.23	0.17	0.06	0.24
43	HfH_2	0.23	0.23	0	0.78
44	HfH_2	0.27	0.31	-0.04	0.33
45	NbH_2	0.29	0.29	0	0.4
46	VH_2	0.3	0.31	-0.01	0.31
47	NbH_2	0.27	0.28	-0.01	0.28
48	NbH_2	0.36	0.35	0.01	0.39
49	NbH_4	0.27	0.28	-0.01	0.42
50	TaH_2	0.41	0.41	0	0.43
51	TaH_4	0.27	0.27	0	0.37
52	TaH_6	0.38	0.4	-0.02	0.61
53	CrH	0.4	0.4	0	0.4
54	CrH_3	0.35	0.38	-0.03	0.38
57	TcH_3	0.31	0.31	0	0.36
58	FeH_6	0.36	0.36	0	1.0
59	FeH_5	0.17	0.37	-0.2	0.55
61	RuH	0.28	0.28	0	0.28
62	RuH_3	0.29	0.29	0	0.29
63	RuH_3	0.29	0.29	0	0.29
64	OsH	0.33	0.34	-0.01	0.34
65	CoH	0.27	0.27	0	0.27
68	PdH	0.19	0.2	-0.01	0.2
70	PdH	0.19	0.2	-0.01	0.2
73	PtH	0.27	0.27	0	0.27
74	PtH	0.27	0.28	-0.01	0.32
75	AuH	0.27	0.27	0	0.27
76	BH	0.51	0.51	0	0.51
77	BH_3	0.56	0.58	-0.02	0.84
78	AlH_3	0.36	0.41	-0.05	0.41
80	GaH_3	0.37	0.1	0.27	0.0
81	InH_5	0.27	0.3	-0.03	0.99
82	InH_3	0.28	0.31	-0.03	0.9
83	SiH_4	0.27	0.48	-0.21	0.81
84	SiH_4	0.51	0.54	-0.03	0.73
87	SiH_4	0.45	0.47	-0.02	0.57
88	SiH_4	0.49	0.5	-0.01	0.6
89	SiH_4	0.49	0.48	0.01	0.98

91	SiH_4	0.32	0.33	-0.01	0.84
93	Si_2H_6	0.48	0.48	0	0.77
94	Si_2H_6	0.58	0.02	0.56	0.02
95	Si_2H_6	0.48	0.54	-0.06	0.96
97	GeH_4	0.39	0.4	-0.01	0.9
99	GeH_4	0.34	0.31	0.03	0.95
100	GeH_4	0.4	0.42	-0.02	0.65
101	$GeH_4(H_2)2$	0.37	0.35	0.02	0.93
103	GeH_4	0.33	0.35	-0.02	0.66
104	$Ge3H_{11}$	0.31	0.33	-0.02	0.42
105	GeH_3	0.31	0.33	-0.02	0.4
106	GeH_3	0.32	0.32	0	0.35
107	SnH_8	0.24	0.24	0	0.9
108	SnH_4	0.22	0.23	-0.01	0.98
109	SnH_4	0.24	0.24	0	0.95
111	SnH_4	0.29	0.29	0	0.93
113	SnH_{12}	0.34	0.26	0.08	0.91
114	SnH_{14}	0.37	0.37	0	0.96
115	$PbH_4(H_2)2$	0.27	0.21	0.06	0.97
119	PH_4	0.19	0.15	0.04	1.0
120	PH_2	0.15	0.62	-0.47	1.0
121	AsH	0.45	0.45	0	0.45
122	AsH_8	0.39	0.43	-0.04	0.91
123	SbH	0.35	0.37	-0.02	0.39
124	SbH_3	0.27	0.39	-0.12	0.87
125	SbH_4	0.29	0.31	-0.02	0.95
126	BiH_2	0.32	0.33	-0.01	0.97
127	BiH_3	0.31	0.32	-0.01	0.85
128	SbH_3	0.35	0.35	0	0.91
129	BiH	0.29	0.33	-0.04	0.0
130	BiH_2	0.31	0.32	-0.01	0.94
131	BiH_4	0.29	0.3	-0.01	0.96
133	BiH_6	0.37	0.29	0.08	0.85
136	$(H_2S)2H_2$	0.69	0.69	0	0.69
138	H_3S	0.7	0.71	-0.01	0.71
140	H_4S_3	0.44	0.38	0.06	0.77
141	H_5S_2	0.56	0.28	0.28	0.89
146	HSe_2	0.44	0.46	-0.02	0.56
147	HSe	0.52	0.53	-0.01	0.0
148	H_3Se	0.53	0.54	-0.01	0.54
149	HSe	0.52	0.52	0	0.55
150	H_4Te	0.4	0.41	-0.01	0.94
151	H_5Te_2	0.41	0.37	0.04	0.86
152	HTe	0.41	0.42	-0.01	0.42
153	H_4Te	0.35	0.36	-0.01	0.92

154	HTe	0.37	0.37	0	0.91
155	PoH_4	0.34	0.26	0.08	0.95
156	PoH	0.34	-	-	0.31
157	PoH_2	0.35	0.31	0.04	0.97
158	PoH_6	0.36	0.34	0.02	0.97
159	HBr	0.37	0.38	-0.01	0.66
160	HCl	0.37	0.49	-0.12	0.68
161	HBr	0.38	0.38	0	0.7
162	HCl	0.38	0.39	-0.01	0.76
163	H_2I	0.36	0.12	0.24	0.99
164	H_4I	0.45	0.25	0.2	0.95
165	H_2I	0.36	-	-	0.99
166	H_2I	0.35	0.27	0.08	0.94
169	MgH_2	0.33	0.34	-0.01	0.34
170	MgH_4	0.26	0.27	-0.01	1.0
171	MgH_{12}	0.36	0.36	0	0.97
173	PH_2	0.39	0.39	0	0.54
176	H	0.64	0.64	0	0.64
177	H	0.35	0.52	-0.17	0.94

C

TcESTIME supplementary information

C.1 `critic2` input file

A typical input file for `critic2` used for TcESTIME, where the ELF cube file is `example.ELF.cube`, located in the directory `/path_to/cube_dir`, reads

```
crystal /path_to/cube_dir/example.ELF.cube
load /path_to/cube_dir/example.ELF.cube
AUTO CPEPS 0.3 NUCEPSH 0.6
```

The parameters CPEPS and NUCEPSH, here equal to 0.3 and 0.6, respectively, are those who optimize the performance in the binary dataset. However, they can be manually tuned if the used desires it (see the reference manual at https://aoterodelaroza.github.io/critic2/manual/).

C.2 Core radii

Table C.1: Effective radii, r_{eff} of the core of different chemical elements, as obtained from the atomic ELF in Kohout and Savin (1996).

Element	Z	r_{eff}
Li	3	1.53
Be	4	1.02
B	5	0.75
C	6	0.58
N	7	0.47
O	8	0.40
F	9	0.34
Ne	10	0.30
Na	11	2.135
Mg	12	1.686

Al	13	1.395
Si	14	1.183
P	15	1.024
S	16	0.912
Cl	17	0.817
Ar	18	0.739
K	19	3.07
Ca	20	2.55
Sc	21	2.39
Ti	22	2.29
V	23	2.20
Cr	24	2.37
Mn	25	2.07
Fe	26	2.01
Co	27	1.95
Ni	28	1.89
Cu	29	2.17
Zn	30	1.81
Ga	31	1.59
Ge	32	1.42
As	33	1.28
Se	34	1.18
Br	35	1.10
Kr	36	1.02
Rb	37	3.51
Sr	38	2.95
Cd	48	2.34

References

[1] H. Kamerlingh Onnes. The resistance of pure mercury at helium temperatures. *Commun. Phys. Lab. Univ. Leiden*, 12, 1911.

[2] W. Meissner and R. Ochsenfeld. Ein neuer effekt bei eintritt der supraleitfähigkeit. *Naturwissenschaften*, 21:787–788, 1933.

[3] M. Tinkham. *Introduction to Superconductivity*. International series in pure and applied physics. McGraw Hill, 1996.

[4] A.M. Portis. *Electrodynamics of High-temperature Superconductors*. Lecture notes in physics. World Scientific, 1993.

[5] John Boyd Ketterson and Shengnian N Song. *Superconductivity*. Cambridge university press, 1999.

[6] W. V. Hassenzahl, D. W. Hazelton, B. K. Johnson, P. Komarek, M. Noe, and C. Reis. Electric power applications of superconductivity. In *Proceedings of the IEEE*, volume 92, pages 1655–1674, 2004.

[7] W.V. Hassenzahl. Applications of superconductivity to electric power systems. *IEEE Power Engineering Review*, 20(5):4–7, 2000.

[8] Harold Weinstock. *Applications of superconductivity*, volume 365. Springer Science & Business Media, 2000.

[9] Alvaro Sanchez, Nuria Del Valle, Enric Pardo, Du-Xing Chen, and Carles Navau. Magnetic levitation of superconducting bars. *Journal of applied physics*, 99(11), 2006.

[10] P Bernstein and J Noudem. Superconducting magnetic levitation: principle, materials, physics and models. *Superconductor Science and Technology*, 33(3):033001, 2020.

[11] Ali Raza and SS Ali. Superconductors for magnetic imaging resonance applications. *Materials Research Foundations*, 132:230–255, 2022.

[12] Ya S Greenberg. Application of superconducting quantum interference devices to nuclear magnetic resonance. *Reviews of Modern Physics*, 70(1):175, 1998.

[13] John Clarke and Frank K Wilhelm. Superconducting quantum bits. *Nature*, 453(7198):1031–1042, 2008.

[14] He-Liang Huang, Dachao Wu, Daojin Fan, and Xiaobo Zhu. Superconducting quantum computing: a review. *Science China Information Sciences*, 63:1–32, 2020.

[15] D. U. Gubser and A. W. Webb. High-pressure effects on the superconducting transition temperature of aluminum. *Phys. Rev. Lett.*, 35:104–107, Jul 1975.

[16] José A. Flores-Livas, Lilia Boeri, Antonio Sanna, Gianni Profeta, Ryotaro Arita, and Mikhail Eremets. A perspective on conventional high-temperature superconductors at high pressure: Methods and materials. *Physics Reports*, 856:1–78, 2020. A perspective on conventional high-temperature superconductors at high pressure: Methods and materials.

[17] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Phys. Rev.*, 108:1175–1204, Dec 1957.

[18] Emanuel Maxwell. Isotope effect in the superconductivity of mercury. *Physical Review*, 78(4):477, 1950.

[19] A. B. Migdal. Interactions between electrons and lattice vibrations in a normal metal. *Sov. Phys. - JETP (Engl. Transl.)*, 34(7), 1958.

[20] G. M. Eliashberg. Interactions between electrons and lattice vibrations in a superconductor. *Sov. Phys. - JETP (Engl. Transl.)*, 11(3), 1960.

[21] P. B. Allen and R. C. Dynes. Transition temperature of strong-coupled superconductors reanalyzed. *Phys. Rev. B*, 12:905–922, Aug 1975.

[22] Marvin L Cohen and PW Anderson. Comments on the maximum superconducting transition temperature. In *AIP Conference Proceedings*, volume 4, pages 17–27. American Institute of Physics, 1972.

[23] Gi ichiro Oya and E.J. Saur. Preparation of nb_3ge films by chemical transport reaction and their critical properties. *J. Low Temp. Phys.*, 34:569–583, 1979.

[24] Andreas Schilling, M Cantoni, JD Guo, and HR Ott. Superconductivity above 130 k in the hg–ba–ca–cu–o system. *Nature*, 363(6424):56–58, 1993.

[25] L. Gao, Y. Y. Xue, F. Chen, Q. Xiong, R. L. Meng, D. Ramirez, C. W. Chu, J. H. Eggert, and H. K. Mao. Superconductivity up to 164 k in $hgba_2ca_{m-1}cu_mo_{2m+2+\delta}$ (m=1, 2, and 3) under quasihydrostatic pressures. *Phys. Rev. B*, 50:4260–4263, Aug 1994.

[26] Jun Nagamatsu, Norimasa Nakagawa, Takahiro Muranaka, Yuji Zenitani, and Jun Akimitsu. Superconductivity at 39 k in magnesium diboride. *Nature*, 410(6824):63–64, 2001.

[27] N. W. Ashcroft. Metallic hydrogen: A high-temperature superconductor? *Phys. Rev. Lett.*, 21:1748–1749, Dec 1968.

[28] Jeffrey M. McMahon, Miguel A. Morales, Carlo Pierleoni, and David M. Ceperley. The properties of hydrogen and helium under extreme conditions. *Rev. Mod. Phys.*, 84:1607–1653, Nov 2012.

[29] D Alfè, M.J Gillan, and G.D Price. Composition and temperature of the earth's core constrained by combining ab initio calculations and seismic data. *Earth and Planetary Science Letters*, 195(1):91–98, 2002.

[30] NW Ashcroft. Hydrogen dominant metallic alloys: high temperature superconductors? *Physical Review Letters*, 92(18):187002, 2004.

[31] A. P. Drozdov, M. I. Eremets, I. A. Troyan, V. Ksenofontov, and S. I. Shylin. Conventional superconductivity at 203 kelvin at high pressures in the sulfur hydride system. *Nature*, 525:73, ags 2015.

[32] AP Drozdov, PP Kong, VS Minkov, SP Besedin, MA Kuzovnikov, S Mozaffari, L Balicas, FF Balakirev, DE Graf, VB Prakapenka, et al. Superconductivity at 250 k in lanthanum hydride under high pressures. *Nature*, 569(7757):528–531, 2019.

[33] Dmitry V. Semenok, Alexander G. Kvashnin, Anna G. Ivanova, Volodymyr Svitlyk, Vyacheslav Yu. Fominski, Andrey V. Sadakov, Oleg A. Sobolevskiy, Vladimir M. Pudalov, Ivan A. Troyan, and Artem R. Oganov. Superconductivity at 161 k in thorium hydride thh10: Synthesis and properties. *Materials Today*, 33:36–44, 2020.

[34] Panpan Kong, Vasily S. Minkov, Mikhail A. Kuzovnikov, Alexander P. Drozdov, Stanislav P. Besedin, Shirin Mozaffari, Luis Balicas, Fedor Fedorovich Balakirev, Vitali B. Prakapenka, Eran Greenberg, and Mikhail I. Eremets. Superconductivity up to 243 k in yttrium hydrides under high pressure. *Nature Communications*, 12(5075), 2021.

[35] Ivan A. Troyan, Dmitrii V. Semenok, Alexander G. Kvashnin, Andrey V. Sadakov, Oleg A. Sobolevskiy, Vladimir M. Pudalov, Anna G. Ivanova, Vitali B. Prakapenka, Eran Greenberg, Alexander G. Gavriliuk, Igor S. Lyubutin, Viktor V. Struzhkin, Aitor Bergara, Ion Errea, Raffaello Bianco, Matteo Calandra, Francesco Mauri, Lorenzo Monacelli, Ryosuke Akashi, and Artem R. Oganov. Anomalous high-temperature superconductivity in yh6. *Advanced Materials*, 33(15):2006832, 2021.

[36] Elliot Snider, Nathan Dasenbrock-Gammon, Raymond McBride, Xiaoyu Wang, Noah Meyers, Keith V. Lawler, Eva Zurek, Ashkan Salamat, and Ranga Dias. Superconductivity to 262 kelvin via catalyzed hydrogenation of yttrium at high pressures, 2020.

[37] Nathan Dasenbrock-Gammon, Elliot Snider, Raymond McBride, Hiranya Pasan, Dylan Durkee, Nugzari Khalvashi-Sutter, Sasanka Munasinghe, Sachith E Dissanayake, Keith V Lawler, Ashkan Salamat, et al. Evidence of near-ambient superconductivity in a n-doped lutetium hydride. *Nature*, 615(7951):244–250, 2023.

[38] Xudong Wei, Xiaokuan Hao, Aitor Bergara, Eva Zurek, Xiaowei Liang, Linyan Wang, Xiaoxu Song, Peifang Li, Lin Wang, Guoying Gao, et al. Designing ternary superconducting hydrides with a15-type structure at moderate pressures. *Materials Today Physics*, 34:101086, 2023.

[39] Ying Sun, Yanchao Wang, Xin Zhong, Yu Xie, and Hanyu Liu. High-temperature superconducting ternary Li-R-H superhydrides at high pressures (R=Sc, Y, La). *Phys. Rev. B*, page 24519, 2022.

[40] Yao Sun, Shuai Sun, Xin Zhong, and Hanyu Liu. Prediction for high superconducting ternary hydrides below megabar pressure. *J. Phys. Condens. Matter*, page 505404, 2022.

[41] Wendi Zhao, Defang Duan, Xianxiang Yao, Zihao Huo, Qiwen Jiang, and Tian Cui. Pressure-induced high-t_c superconductivity in the ternary clathrate system Y-Ca-H. *Phys. Rev. B*, page 14521, 2022.

[42] Kai Hu, Qingjun Tong, Li-Min Guan, Dahui Wang, and Jinqing Yu. Crystal structure evolution and superconductivity of the ternary hydride csh 3 under pressure. *Physical Review B*, 105(9):094108, 2022.

[43] David C. Lonie and Eva Zurek. Xtalopt: An open-source evolutionary algorithm for crystal structure prediction. *Computer Physics Communications*, 182:372–387, 2011.

[44] Artem R Oganov and Colin W Glass. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *The Journal of chemical physics*, 124(24), 2006.

[45] Yanchao Wang, Jian Lv, Li Zhu, and Yanming Ma. Crystal structure prediction via particle-swarm optimization. *Phys. Rev. B*, 82:094116, Sep 2010.

[46] Chris J Pickard and RJ Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

[47] Ion Errea, Francesco Belli, Lorenzo Monacelli, Antonio Sanna, Takashi Koretsune, Terumasa Tadano, Raffaello Bianco, Matteo Calandra, Ryotaro Arita, Francesco Mauri, and José A. Flores-Livas. Quantum crystal structure in the 250-kelvin superconducting lanthanum hydride. *Nature*, 578:66–69, 2019.

[48] Ángel Martín-Pendás and Julia Contreras-García. *Topological Approaches to the Chemical Bond*. Springer Cham, 2023.

[49] Yanming Ma, Mikhail Eremets, Artem R Oganov, Yu Xie, Ivan Trojan, Sergey Medvedev, Andriy O Lyakhov, Mario Valle, and Vitali Prakapenka. Transparent dense sodium. *Nature*, 458(7235):182–185, 2009.

[50] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic Structure Theory*. John Wiley & Sons, LTD, Chichester, 2000.

[51] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, Inc., Mineola, first edition, 1996.

[52] Robert G. Parr and Weitao Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, USA, 1994.

[53] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2 edition, 2020.

[54] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical Review*, 136(3B), 1964.

[55] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A), 1965.

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

[57] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, Jan 1988.

[58] Kieron Burke, John P. Perdew, and Yue Wang. *Derivation of a Generalized Gradient Approximation: The PW91 Density Functional*, pages 81–111. Springer US, Boston, MA, 1998.

[59] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, Oct 1996.

[60] Julien Toulouse. *Review of Approximations for the Exchange-Correlation Energy in Density-Functional Theory*, pages 1–90. Springer International Publishing, Cham, 2023.

[61] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt-Saunders, 1976.

[62] Rolf Heid. *Emergent Phenomena in Correlated Matter*, chapter Density Functional Perturbation Theory and Electron Phonon Coupling. Forschungszentrum Jülich GmbH, 2013.

[63] Feliciano Giustino. Electron-phonon interactions from first principles. *Rev. Mod. Phys.*, 89:015003, Feb 2017.

[64] Stefano Baroni, Paolo Giannozzi, and Andrea Testa. Green's-function approach to linear response in solids. *Phys. Rev. Lett.*, 58:1861–1864, May 1987.

[65] Paolo Giannozzi, Stefano de Gironcoli, Pasquale Pavone, and Stefano Baroni. Ab initio calculation of phonon dispersions in semiconductors. *Phys. Rev. B*, 43:7231–7242, Mar 1991.

[66] Stefano Baroni, Stefano de Gironcoli, Andrea Dal Corso, and Paolo Giannozzi. Phonons and related crystal properties from density-functional perturbation theory. *Rev. Mod. Phys.*, 73:515–562, Jul 2001.

[67] Feliciano Giustino, Marvin L. Cohen, and Steven G. Louie. *Physical Review B*, 76:165108, 2007.

[68] Ion Errea, Francesco Belli, Lorenzo Monacelli, Antonio Sanna, Takashi Koretsune, Terumasa Tadano, Raffaello Bianco, Matteo Calandra, Ryotaro Arita, Francesco Mauri, and José A. Flores-Livas. Quantum crystal structure in the 250-kelvin superconducting lanthanum hydride. *Nature*, 578:66–69, 2020.

[69] Francesco Belli and Ion Errea. Impact of ionic quantum fluctuations on the thermodynamic stability and superconductivity of $labh_8$. *Phys. Rev. B*, 106:134509, Oct 2022.

[70] Lorenzo Monacelli, Ion Errea, Matteo Calandra, and Francesco Mauri. Black metal hydrogen above 360 GPa driven by proton quantum fluctuations. *Nat. Phys.*, 17:63–67, 2021.

[71] Ion Errea, Matteo Calandra, and Francesco Mauri. First-principles theory of anharmonicity and the inverse isotope effect in superconducting palladium-hydride compounds. *Phys. Rev. Lett.*, 111:177002, Oct 2013.

[72] Ion Errea, Matteo Calandra, and Francesco Mauri. Anharmonic free energies and phonon dispersions from the stochastic self-consistent harmonic approximation: Application to platinum and palladium hydrides. *Phys. Rev. B*, 89:064302, Feb 2014.

[73] Lorenzo Monacelli, Raffaello Bianco, Marco Cherubini, Matteo Calandra, Ion Errea, and Francesco Mauri. The stochastic self-consistent harmonic approximation: calculating vibrational properties of materials with full quantum and anharmonic effects. *J. Phys.: Condens. Matter*, 33:363001, 2021.

[74] Mervyn Roy. The tight binding method. *Rutgers University*, 5:57, 2015.

[75] Roald Hoffmann. How chemistry and physics meet in the solid state. *Angewandte Chemie*, 26:846–878, 1987.

[76] Thomas A Albright, Jeremy K Burdett, and Myung-Hwan Whangbo. *Orbital interactions in chemistry*. John Wiley & Sons, 2013.

[77] Angel Martín Pendás, Julia Contreras-García, Fernanda Pinilla, José D. Mella, Carlos Cardenas, and Francisco Muñoz. A chemical theory of topological insulators. *Chem. Commun.*, 55:12281–12287, 2019.

[78] R. Fernandes. *Lecture notes: BCS theory of superconductivity*. 2015.

[79] Koji Kajimura and Hisao Hayakawa, editors. *Advances in Superconductivity III*. Springer Tokyo, 1991.

[80] W. L. McMillan. Transition temperature of strong-coupled superconductors. *Phys. Rev.*, 167:331–344, Mar 1968.

[81] Andreas Schmitt. *Introduction to superfluidity – Field-theoretical approach and applications*. Springer Cham, 2014.

[82] J. R. Waldran. *Superconductivity of metals and cuprates*. Institute of Physics, 1996.
[83] A. Kadin. Spatial structure of the cooper pair. *J. Supercond. Nov. Magn.*, 20:285–292, 2007.
[84] L. N. Oliveira, E. K. U. Gross, and W. Kohn. Density-functional theory for superconductors. *Phys. Rev. Lett.*, 60:2430–2433, Jun 1988.
[85] Antonio Sanna. Introduction to superconducting density functional theory. volume 7 of *The Physics of Correlated Insulators, Metals, and Superconductors*. Forschungszentrum Jülich, 2017.
[86] N. David Mermin. Thermal properties of the inhomogeneous electron gas. *Phys. Rev.*, 137:A1441–A1443, Mar 1965.
[87] P.G. de Gennes. *Superconductivity of Metals and Alloys*. Westview Press, 1966.
[88] M. Lüders, M. A. L. Marques, N. N. Lathiotakis, A. Floris, G. Profeta, L. Fast, A. Continenza, S. Massidda, and E. K. U. Gross. Ab initio theory of superconductivity. i. density functional formalism and approximate functionals. *Phys. Rev. B*, 72:024545, Jul 2005.
[89] A. Linscheid, A. Sanna, A. Floris, and E. K. U. Gross. First-principles calculation of the real-space order parameter and condensation energy density in phonon-mediated superconductors. *Phys. Rev. Lett.*, 115:097002, Aug 2015.
[90] Walter Kohn. *Physical Review*, 133:A171–A181, 1964.
[91] R. Resta and S. Sorella. Electron localization in the insulating state. *Physical Review Letters*, 82:370–373, 1999.
[92] Richard F. W. Bader. A quantum theory of molecular structure and its applications. *Chem. Rev.*, 91:893–928, 1991.
[93] R. F. W. Bader. *Atoms in Molecules: A Quantum Theory*. Oxford Univ. Press, 1990.
[94] B. Silvi and A. Savin. Classification of chemical bonds based on topological analysis of electron localization functions. *Nature*, 371:683, 1994.
[95] K. E. Becke, A. D.and Edgecombe. *J. Chem. Phys.*, 92:5397, 1990.
[96] C. F. von Weizsäcker. Zur theorie de kernmassen. *Z. Phys.*, 96:431, 1935.
[97] A. Savin, R. Nesper, and S. Wengert. *Angew. Chem. Int. Ed. Engl.*, 36:1808–1832, 1997.
[98] Roberto A. Boto. *Development of the NCI method : high performance optimization and visualization*. Phd thesis, Sorbonne Université, 2017. Available at https://theses.hal.science/tel-01490183.
[99] Bernard Silvi. The synaptic order: a key concept to understand multicenter bonding. *Journal of Molecular Structure*, 614:3–10, 2002.
[100] A. Savin, B. Silvi, and F. Colonna. Topological analysis of the electron localization function applied to delocalized bonds. *Can. J. Chem*, 74:1088–1096, 1996.
[101] M. Calatayud, J. Andrés, Beltrán, and B. A.; Silvi. *Theor. Chem. Acc.*, 112:453, 2004.
[102] P. G. Mezey. *Can. J. Chem.*, 72:928, 1993.

[103] R. F. W. Bader and M. E. Stephens. Spatial localization of the electronic pair and number distributions in molecules. *Journal of the American Chemical Society*, 97:7391–7399, 1975.

[104] Alexey I. Baranov and Miroslav Kohout. Electron localization and delocalization indices for solids. *Journal of Computational Chemistry*, 32:2064–2076, 2011.

[105] R. Abd-Shukor. Coherence length versus transition temperature of hydride-based and room temperature superconductors. *Results in Physics*, 25:104219, 2021.

[106] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski. *Methods of Quantum Field Theory in Statistical Physics*, page 295. Pergamon Press, 1965.

[107] K. F. Riley, M. P. Hobson, and S. J. Bence. *Mathematical Methods for Physics and Engineering: A Comprehensive Guide*. Cambridge University Press, 3 edition, 2006.

[108] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020.

[109] Martin-Isbjörn Trappe, Jun Hao Hue, and Berthold-Georg Englert. Density-potential functional theory for fermions in one dimension. In *Density Functionals For Many-particle Systems: Mathematical Theory And Physical Applications Of Effective Equations*, pages 251–267. World Scientific, 2023.

[110] Francesco Belli, Trinidad Novoa, J. Contreras-García, and Ion Errea. Strong correlation between electronic bonding network and critical temperature in hydrogen-based superconductors. *Nature Communications*, 12:1–11, 2021.

[111] Santanu Saha, Simone Di Cataldo, Federico Giannessi, Alessio Cucciari, Wolfgang Von Der Linden, and Lilia Boeri. mapping superconductivity in high-pressure hydrides: The superhydra project. *Physical Review Materials*, 7(5):054806, 2023.

[112] Sason S Shaik and Philippe C Hiberty. *A chemist's guide to valence bond theory*. John Wiley & Sons, 2007.

[113] Alfonso Gallo Bueno. *topological indices for the characterization of electronic localization in molecules and solids*. Phd thesis, Universidad de Oviedo, 2016. Available at https://digibuo.uniovi.es/dspace/bitstream/handle/10651/38790/TD_AlfonsoGallo.pdf?sequence=1&isAllowed=y.

[114] Angel Martín Pendás, Francisco Muñoz, Carlos Cardenas, and Julia Contreras-García. Understanding topological insulators in real space. *Molecules*, 26(10):2965, 2021.

[115] Emil Prodan and Walter Kohn. Nearsightedness of electronic matter. *Proceedings of the National Academy of Sciences*, 102(33):11635–11638, 2005.

[116] A. Otero-de-la-Roza, M. A. Blanco, A. Martín Pendás, and V. Luaña. Critic: a new program for the topological analysis of solid-state electron densities. *Comput. Phys. Commun.*, 180:157–166, 2009.

[117] A. Otero-de-la-Roza, E. R. Johnson, and V. Luaña. Critic2: A program for real-space analysis of quantum chemical interactions in solids. *Comput. Phys. Commun.*, 185:1007–1018, 2014.

[118] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. Fabris, G. Fratesi, S. de Gironcoli, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch. Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter*, 21:395502, 2009.

[119] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J Jia, M. Kawamura, H.-Y. Ko, A. Kokalj, E. Küçükbenli, M .Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero de-la Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni. Advanced capabilities for materials modelling with quantum ESPRESSO. *J. Phys.: Condens. Matter*, 29:465901, 2017.

[120] Julia Contreras-García, P. Mori-Sánchez, B. Silvi, and J. M. Recio. A quantum chemical interpretation of compressibility in solids. *J. Chem. Theory Comput.*, 5:2108–2114, 2009.

[121] M. Kohout and A. Savin. Atomic shell structure and electron numbers. *Int. J. Quantum Chem.*, 60:875–882, 1996.

[122] Thomas H. Cormen, Charles E. Leiserson, Ronals L. Rivest, and Clifford Stein. *Introduction to algorithms*, page 594. The MIT Press, Cambridge, Massachusetts, 3rd edition, 2009.

[123] Hanyu Liu, Ivan I. Naumov, Roald Hoffmann, N. W. Ashcroft, , and Russell J. Hemley. Potential high-tc superconducting lanthanum and yttrium hydrides at high pressure. *Proceedings of the National Academy of Sciences*, 114:6990–6995, 2017.

[124] Gang Wu, Yanbin Ma, Hongyu Yu, Da Li, Bingbing Liu, and Tian Cui. Prediction of stoichiometric poh n compounds: crystal structures and properties. *RSC Adv.*, 5:103445–103450, 2015.

[125] Defang Duan, Fubo Tian, Yunxian Liu, Xiaoli Huang, Da Li, Hongyu Yu, Yanbin Ma, Bingbing Liu, and Tian Cui. Enhancement of t_c in the atomic phase of iodine-doped hydrogen at high pressures. *Phys. Chem. Chem. Phys.*, 17:32335–32340, 2015.

[126] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47:558–561, Jan 1993.

[127] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.

[128] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, Oct 1996.

[129] Lan-Ting Shi, Yong-Kai Wei, A-Kun Liang, Robin Turnbull, Cai Cheng, Xiang-Rong Chen, and Guang-Fu Ji. Prediction of pressure-induced superconductivity in the novel ternary system $ScCah_{2n}$ (n = 1–6). *J. Mater. Chem. C*, pages 7284–7288, 2021.

[130] Zihan Zhang, Tian Cui, Michael J. Hutcheon, Alice M. Shipley, Hao Song, Mingyang Du, Vladimir Z. Kresin, Defang Duan, Chris J. Pickard, and Yansun Yao. Design principles for high-temperature superconductors with a hydrogen-based alloy backbone at moderate pressure. *Phys. Rev. Lett.*, page 47001, 2022.

[131] Miles Cranmer. Interpretable machine learning for science with pysr and symbolicregression.jl. 2023.

[132] Atsushi Togo and Isao Tanaka. Spglib: a software library for crystal symmetry search. *arXiv preprint arXiv:1808.01590*, 2018.

[133] Pugeng Hou, Francesco Belli, Raffaello Bianco, and Ion Errea. Quantum anharmonic enhancement of superconductivity in P63/mmc ScH6 at high pressures: A first-principles study. *Journal of Applied Physics*, 130(17):175902, 2021.