A Theoretical and Numerical Analysis of a Kohn-Sham Equation and Related Control Problems

vorgelegt von

Martin Sprengel

Würzburg 2017

Acknowledgements

First, I want to thank my supervisor Prof. Alfio Borzì who had the idea for this challenging Ph.D. project which gave me the possibility to combine both my physics and mathematics background in one research project. He carefully introduced me into the art of paper-writing and supported me in getting our papers published even under the particular circumstances we faced. Furthermore, he gave me the exceptional possibility to co-author the comprehensive textbook "Formulation and Numerical Solution of Quantum Control Problems" with him and Gabriele — an opportunity not many Ph.D. students have.

A special thanks goes to Gabriele with whom I did several research projects together. It was always nice to work with you and we had many long discussions — both in person and on Skype — in which we clarified many issues. You were always very precise which helped to produce nice mathematical results. But you were not just a colleague but also a friend and I thank you for the encouragement and the nice time spend together in Würzburg but also in Bad Honnef, Nottingham, Nice and Geneva! We should go for a walk again. . .

Furthermore, I want to thank my office mates Juri and Andrea for the mathematical insights and nice time in the office but also in our leisure activities where Stephan ("Fly save!") often joined us for nice discussions and funny games.

I also want to thank Petra who has done perfectly in keeping the bureaucracy out of our way. She managed everything easily and fast and at the same time she always had friendly words for us. This is really something special at our chair!

The atmosphere at our chair was very pleasant; for this I want to thank my fellow Ph.D. colleges (in random order) Tim, Suttida, Duncan, Andreas, Bea, Tanvir, Masoumeh, Christian. Additional thanks goes to Jan, Nadja, Gabriele and Andrea for proofreading my thesis.

Last but not least I want to thank my parents, my brother, and Gerlinde who have always supported in every possible way.

Abstract

In this work, multi-particle quantum optimal control problems are studied in the framework of time-dependent density functional theory (TDDFT). Quantum control problems are of great importance in both fundamental research and application of atomic and molecular systems. Typical applications are laser induced chemical reactions, nuclear magnetic resonance experiments, and quantum computing. Theoretically, the problem of how to describe a non-relativistic system of multiple particles is solved by the Schrödinger equation (SE). However, due to the exponential increase in numerical complexity with the number of particles, it is impossible to directly solve the Schrödinger equation for large systems of interest. An efficient and successful approach to overcome this difficulty is the framework of TDDFT and the use of the time-dependent Kohn-Sham (TDKS) equations therein. This is done by replacing the multi-particle SE with a set of nonlinear single-particle Schrödinger equations that are coupled through an additional potential.

Despite the fact that TDDFT is widely used for physical and quantum chemical calculation and software packages for its use are readily available, its mathematical foundation is still under active development and even fundamental issues remain unproven today. The main purpose of this thesis is to provide a consistent and rigorous setting for the TDKS equations and of the related optimal control problems.

In the first part of the thesis, the framework of density functional theory (DFT) and TDDFT are introduced. This includes a detailed presentation of the different functional sets forming DFT. Furthermore, the known equivalence of the TDKS system to the original SE problem is further discussed.

To implement the TDDFT framework for multi-particle computations, the TDKS equations provide one of the most successful approaches nowadays. However, only few mathematical results concerning these equations are available and these results do not cover all issues that arise in the formulation of optimal control problems governed by the TDKS model. It is the purpose of the second part of this thesis to address these issues such as higher regularity of TDKS solutions and the case of weaker requirements on external (control) potentials that are instrumental for the formulation of well-posed TDKS control problems. For this purpose, in this work, existence and uniqueness of TDKS solutions are investigated in the Galerkin framework and using energy estimates for the nonlinear TDKS equations.

In the third part of this thesis, optimal control problems governed by the TDKS model are formulated and investigated. For this purpose, relevant cost functionals that model the purpose of the control are discussed. Henceforth, TDKS control problems result from the requirement of optimising the given cost functionals subject to the differential constraint given by the TDKS equations. The analysis of these problems is novel and represents one of the main contributions of the present thesis. In particular, existence of minimizers is proved and their characterization by TDKS optimality systems is discussed in detail. To this end, Fréchet differentiability of the TDKS model and of the cost functionals is addressed considering H^1 cost of the control. This part is concluded by deriving the reduced gradient in the L^2 and H^1 inner product. While the L^2 optimization is widespread in the literature, the choice of the H^1 gradient is motivated in this work by theoretical consideration and by resulting numerical advantages.

The last part of the thesis is devoted to the numerical approximation of the TDKS optimality systems and to their solution by gradient-based optimization techniques. For the former purpose, Strang time-splitting pseudo-spectral schemes are discussed including a review of some recent theoretical estimates for these schemes and a numerical validation of these estimates. For the latter purpose, nonlinear (projected) conjugate gradient methods are implemented and are used to validate the theoretical analysis of this thesis with results of numerical experiments with different cost functional settings.

Zusammenfassung

In dieser Arbeit werden quantenmechanische Vielteilchen-Optimalsteuerungsprobleme im Rahmen der zeitabhängigen Dichtefunktionaltheorie (TDDFT) untersucht. Quantenmechanische Optimalsteuerungsprobleme sind sowohl in der Grundlagenforschung atomarer und molekularer Systeme als auch in entsprechenden Anwendungen von großer Bedeutung. Typische Anwendungen sind laserinduzierte chemische Reaktionen, Kernspinresonanzexperimente und Quantencomputer. Theoretisch ist das Problem einer nichtrelativistischen Beschreibung von Vielteilchensystemen mit der Schrödingergleichung (SG) gelöst. Tatsächlich ist es aber wegen des exponentiellen Anstiegs der numerischen Komplexität mit der Teilchenzahl unmöglich, die Schrödingergleichung für große Systeme von Interesse direkt zu lösen. Ein effizienter und erfolgreicher Ansatz diese Schwierigkeit zu überwinden ist die TDDFT und die Verwendung der zeitabhängigen Kohn-Sham-Gleichungen (TDKS) im Rahmen der TDDFT. Diese ersetzen die Vielteichlchen-SG durch ein System nichtlinearer Einteilchen-SGn, die mittels eines zusätzlichen Potentials gekoppelt sind.

Obwohl die TDDFT für physikalische und quantenchemische Rechungen weit verbreitet ist und Softwarepakete zur direkten Verwendung zur Verfügung stehen, sind die mathematischen Grundlagen der TDDFT noch in der Entwicklung und grundlegende Vermutungen sind noch immer unbewiesen. Das Hauptanliegen der vorliegenden Arbeit ist es, einen konsistenten und mathematisch präzisen Rahmen für die TDKS-Gleichungen und verwandte Optimalsteuerungsprobleme zu liefern.

Im ersten Teil der Arbeit wird die Dichtefunktionaltheorie (DFT) und die TDDFT eingeführt. Diese Einführung enthält eine detaillierte Darstellung der für die DFT relevanten Funktionenmengen. Außerdem wird die bereits bekannte Äquivalenz zwischen dem ursprünglichen Schrödingerproblem und dem TDKS-System mathematisch weitergehend diskutiert.

Der derzeit erfolgreichste Ansatz, Vielteichenrechnungen im Rahmen der TDDFT umzusetzen, sind die TDKS-Gleichungen. Es sind jedoch bisher nur wenige mathematische Resultate über diese Gleichungen verfügbar und diese Ergebnisse behandeln nicht alle Probleme, die bei der Formulierung von Optimalsteuerungsproblemen bei TDKS-Gleichungen auftreten. Es ist das Ziel des zweiten Teils dieser Arbeit, diese für die Wohldefiniertheit der Formulierung der Optimalsteuerungsaufgabe maßgeblichen Probleme, wie die höhere Regularität der Lösungen der TDKS-Gleichungen und schwächere Voraussetzungen an das externe Kontrollpotential, zu behandeln. Dazu wird die Existenz und Eindeutigkeit von Lösungen der nichtlinearen TDKS-Gleichungen mit dem Galerkin-Ansatz und Energieabschätzungen untersucht.

Im dritten Teil dieser Arbeit werden Probleme optimaler Steuerung bei TDKS-Gleichungen formuliert und untersucht. Dafür werden relevante Kostenfunktionale, die das Ziel der Steuerung modellieren, diskutiert. Die Optimalsteuerungsprobleme ergeben sich aus der Optimierung dieser Kosten unter der Nebenbedingung der TDKS-Gleichungen. Die Analyse dieser Probleme ist neu und stellt eines der Hauptergebnisse der vorliegenden Arbeit dar. Insbesondere wird die Existenz einer optimalen Steuerung bewiesen und ihre Charakterisierung mittels eines TDKS-Optimalitätssystem im Detail diskutiert. Dazu wird die Fréchet-Differenzierbarkeit des TDKS-Models und des Kostenfunktionals mit

H^1-Steuerungskosten betrachtet. Abschließend wird der reduzierte Gradient im L^2- und im H^1-Skalarprodukt hergeleitet. Während die L^2-Optimierung in der Literatur weit verbreitet ist, wird in dieser Arbeit die Verwendung des H^1-Gradienten mit theoretischen Argumenten und resultierenden numerischen Vorteilen motiviert.

Der letzte Teil dieser Arbeit ist der numerischen Approximation des TDKS-Optimalitätssystems und seiner Lösung mittels gradientenbasierter Optimierungsmethoden gewidmet. Für ersteres wird die Strang Zeitsplitting-Pseudospektralmethode diskutiert, eine Zusammenfassung einiger aktueller theoretischer Abschätzungen für dieses Schema angegeben und diese Abschätzungen numerisch überprüft. Für letzteres wird das (projizierte) nichtlineare Verfahren der konjugierten Gradienten (NCG) implementiert und verwendet um die theoretische Analyse dieser Arbeit mit den Ergebnissen numerischer Rechnungen für verschiedene Kostenfunktionale zu validieren.

Contents

Contents

1. Introduction

Many models of interest in quantum physics and chemistry consist of multi-particle systems that can be modelled by the multi-particle Schrödinger equation (SE). However, the space dimensionality of this equation increases linearly with the number of particles involved and thus the corresponding computational cost increases exponentially, making the use of the multi-particle SE prohibitive. This fact has motivated a great research effort towards an alternative formulation to the SE description that allows to compute the observables of a multi-particle quantum system. This is achieved in the framework of density functional theory (DFT) by using the particle density instead of the wavefunction as the main variable. The development of DFT, which starts with the Thomas-Fermi theory in 1927, reached a decisive point with the works of Hohenberg, Kohn, and Sham [HK64, KS65] in the sixties. They proposed an appropriate way to replace a system of N interacting particles in an external potential V_{ext} with another system of non-interacting particles, called the Kohn-Sham (KS) system, in an external potential $V_{ext} + V_{Hxc}$, such that both models have the same electronic density.

A topical introduction to DFT from the physical perspective is given in [ED11]. For a good mathematical introduction to ground state DFT and the KS theory, we refer to the original work of Lieb [Lie83, Lie85] and the more recent paper [AC09]. A discussion of norm-conserving semilocal pseudopotentials for Kohn-Sham models can be found in [CM16] where most results were proved for the Hartree model but extensions to the Kohn-Sham model were also discussed.

The DFT is focused on the computation of stationary (ground) states and successfully provides results that motivated the extension of the DFT to time-dependent phenomena. This extension was first proposed by Runge and Gross in [RG84], and further investigated from a more mathematical point of view in the work of van Leeuwen [vL99]. We also refer to [MUN+06] and [MMN+12] for an introduction to time-dependent DFT (TDDFT) where the latter reference contains also a chapter on the formulation of optimal control within the framework of TDDFT. The extension to TDDFT is necessary not only to investigate time-dependent phenomena but also to accurately calculate atomic and molecular excitation spectra.

We remark that the previously mentioned works on TDDFT are more concerned about the physical ideas behind the model including some calculation to motivate them. However, these works are less focused on mathematical issues. On the other hand, we highlight the work of Ruggenthaler et al. where a mathematical accurate derivation of TDDFT is attempted. An excellent overview can be found in the review article [RPvL15]. In this detailed review, starting from the theory of classical and non-classical solutions of the Schrödinger equation, the density-potential mapping is considered. Furthermore, the time-dependent Kohn-Sham (TDKS) system is introduced, and the central theorem of [vL99] is revised and two proofs are given, one based on a Taylor expansion and one based on a fixed point argument. A discussion on the invertibility of the so called Sturm-Liouville operator is given including a reference to [PR11] where the details are elaborated. Finally, additional topics like memory and initial state dependence of the KS system as well as lattice systems, vector potentials and photons are discussed. Based on

these results, we present the first mathematically rigorous statement of the van Leeuwen's theorem in the literature.

The TDKS equations consist of the following coupled system of PDEs that is nonlinear through the density $\rho = \sum_{i=1}^{N} |\psi_i(x,t)|^2$:

$$
\begin{aligned}
i\frac{\partial \psi_j}{\partial t}(x,t) &= \left(-\nabla^2 + V_{ext}(x,t) + V_{Hxc}(x,\rho)\right)\psi_j(x,t), \\
\psi_j(x,t) &= \psi_j^0(x), \qquad j = 1,\ldots,N.
\end{aligned}
\tag{1.1}
$$

The theory of the KS equations is very recent. Some studies neglect parts of the KS term V_{Hxc} and consider only the Hartree equation. The Hartree equation is analysed for example in [CL99, JP14]. To the best of our knowledge the only publication studying the full TDKS equations is [Jer15]. However, the assumptions made in [Jer15], namely $V_{ext} \in C^1(0,T;C^1(\Omega))$ and $V_{ext} \geq 0$, are not suitable for our optimization problem. Therefore, we prove existence and uniqueness results for the solution of (1.1) with regularity requirements on V_{ext} that are suitable for the optimization framework presented in Chapter 3, see also [SCB17c]. Furthermore, we show improved regularity of the solution. This is essential for the analysis of optimal control problems governed by TDKS equations.

Recent research interest in the KS model focuses on the dependence of the KS potential V_{Hxc} on the initial condition of the original problem and memory effects. In [FNRM16], it is suggested that splitting V_{Hxc} in an exchange and a correlation potential does not make sense if the initial state is not a Slater determinant and propose as an alternative starting point a single-particle contribution plus a remainder term. In [RNvL13], analytic expressions for the exchange-correlation potential and initial state for simple interactions between two electrons are investigated and memory effects are studied. This leads to a comparison of adiabatic and single-pole approximations.

Other recent developments aim at extending TDKS beyond the classical Schrödinger equation; for example in [FRAR15], the TDKS scheme is extended to quantum electro-dynamics to correctly describe electron-photon interaction.

Since the early development of quantum mechanics, controlling quantum mechanical processes has been a goal in quantum physics and chemistry; see, e.g., [KRG$^+$89, PDR88] for early approaches on the mathematical framework of optimal control theory and [DP10, GBC$^+$15] for recent review articles. Research on this subject dates back more than three decades as it is important for many applications like the control of electrons in quantum chemistry and spin control for nuclear magnetic resonance (NMR) and magnetic resonance imaging (MRI) [CBDW15, CB16, GBC$^+$15]. Better understanding of quantum control is also important for the development of quantum computing [BMC15]. Much progress has been made in the recent years due to better experimental methods and larger computers available. Active control of chemical reactions at a molecular level like the breaking and forming of specific bonds is the ultimate goal of chemical quantum control [BG03]. To this end, different strategies have been developed.

On the one hand, there are experimental techniques which are mostly closed-loop feedback strategies where a control is applied to a system, the response is measured, and depending on the response, the control is updated in an iterative procedure; see, e.g., the review [BG03]. The obvious advantages are that no approximations are made as the real quantum system is measured and that a single iteration is fast as the control pulses are on a femtosecond time scale (10^{-15}s–10^{-12}s).

On the other hand, theoretical methods try to determine optimal controls off-line by calculations such that these controls can later directly be implemented in laboratories

(open-loop methods). For quantum systems that are not too complex, geometric optimal control techniques can be used to determine the control analytically, see [GBC$^+$15] and references therein. More involved systems, however, need to be solved by numerical methods. For quantum optimal control both time sequential Krotov-type methods, see, e.g., [MT03] and the monotonic scheme in [MST06], and time simultaneous methods like gradient and Newton schemes are used [vWBV10]. In this thesis we use the latter approach to determine optimal controls using nonlinear conjugate gradient (NCG) methods in the framework of Lagrange multipliers.

The theory of optimal control problems governed by PDE is elaborated in detail in the distinguished works of Lions [Lio71] and Tröltzsch [Trö10]. To characterize stationary points in this framework, first-order optimality conditions are used, see [Trö10, Bor12, BS12]. Various quantum mechanical optimal control problems governed by the SE using the Lagrange framework have been studied in the literature, e.g., [BSH02, HMMS13]. We remark especially the works [vWB08, vWB10, vWBV10] where H^1 quantum control problems are discussed.

In modern quantum chemistry, a major goal is to control chemical reactions by optimally shaped laser pulses [BG03]. However, even for small molecules many electrons are involved. Hence, an approach capable of dealing with the high-dimensional multi-electron problem is needed. This leads to the emergence of optimal control using TDDFT models. This approach is already widely used in applications [Cas13, RCW$^+$08a, WLAWLL15] and software packages of are readily available, e.g. OCTOPUS [CAO$^+$06]. The formulas used in the OCTOPUS software package are explained in detail in [CWG12, Wer06]. Further numerical experiments using optimal control with TDDFT include coherent quantum switches of a single electron in a quantum ring (magnetic switch) and a double quantum dot (charge switch), [RCW$^+$08b], and ultra-fast manipulation of a single electron spin in 2D semiconductor quantum dots through spin-orbit coupling, [BMC15], aiming at fast and decoherence-free devices used for quantum computing.

However, as for the theory of the governing KS equations, the research work on optimal control with the KS equations is in its infancy. We are only aware of a study analysing optimal control with the Hartree equation, [FW16]. Therefore, we investigated optimal control using the full KS equations (1.1) as constraint, see Chapter 4 and [SCB17b]. Together with the analysis of the TDKS equation, we provided the first comprehensive theoretical investigation of TDKS optimal control problems which provides the mathematical foundations for the numerous applications currently in use.

We use the framework of Lagrange multiplier to derive an optimality system for the following optimization problem:

$$\min_{(\Psi, u)} J(\Psi, u) \qquad \text{subject to (1.1)}, \tag{1.2}$$

where

$$J(\Psi, u) = \frac{\beta}{2} \int_0^T \int_\Omega (\rho(x, t) - \rho_d(x, t))^2 \mathrm{d}x \mathrm{d}t + \frac{\eta}{2} \int_\Omega \chi_A(x) \rho(x, T) \mathrm{d}x + \frac{\nu}{2} \|u\|_{H^1}^2 \tag{1.3}$$

and u is a control function which models, for example, the amplitude of a laser field.

The bilinear control structure $V_{ext} \Psi = V_0 \Psi + u V_u \Psi$ in (1.1) makes the problem already a nonlinear control problem. However, the complicated nonlinearity in the Kohn-Sham potential turns out to be the main difficulty in the analysis of (1.2) subject to (1.1). For this reason, standard results for bilinear control problems cannot be directly applied. The

nonlinearity V_{Hxc} reduces the regularity of the KS equations. Moreover, as a real-valued function of a complex variable, V_{Hxc} is only differentiable in the sense of the Wirtinger calculus [Rem91, Kre09]. Therefore, it is hard to prove Fréchet differentiability of the constraint (1.1) which is essential for the analysis of the optimal control problem and for the characterization of its solution using the first-order optimality conditions as well as the development of numerical methods for the solution of (1.2).

To use optimal control with TDDFT in practical application, one needs to employ numerical methods both for the optimization and for the solution of the PDEs therein. To solve the optimization problem, we use a nonlinear conjugate gradient (NCG) method, which is a classical gradient-based method, and the quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno (BFGS).

To solve the PDEs, we use the Strang splitting method, an operator splitting technique, in time and evaluate the Laplacian using a Fourier spectral method, see [BJM02, FOS15, BH08]. This combination results in second-order convergence in time and analytic convergence in space This is proved for constant potentials in [BJM02]. In [BJM02] it is further shown that the method is norm preserving and gauge invariant also for non-constant potentials. The numerical analysis of this method applied to the Gross-Pitaevskii equation can be found in [Tha12] and the splitting analysis for the Hartree equation in [Lub08]. Splitting methods applied to quantum models with general Lipschitz nonlinearities are studied in [BBD02]. For the splitting method applied to the full KS equations there exists no theoretical analysis of the convergence properties. However, we investigate the convergence properties in numerical studies and confirm the results obtain for the other models mentioned above.

In the physical and quantum chemical community, a large interest in optimal control governed by the TDKS equations has lead to many applications. However, the theoretical foundations were still in its infancy and no consistent framework encompassing theory and applications was available. This work aims at filling this gap. To this end, the main novelties achieved in this thesis are the following.

1. We present a clear and precise introduction to the time-dependent Kohn-Sham framework and a mathematical rigorous presentation of the extended Runge-Gross theorem, originally published in [vL99].

2. We prove existence and uniqueness of solutions to KS equations for both forward and adjoint equations with requirements for the potential that are suitable for optimization. Additionally, we show improved regularity of the solution of the KS equations which is necessary for the analysis of the optimization problems.

3. We investigate optimal control problems governed by the KS equations. We prove existence of an optimal control and derive the corresponding optimality system using the framework of Lagrange multipliers.

4. We study the convergence properties of the splitting method for the KS and the adjoint equations numerically showing that the obtained accuracy is in good agreement with the convergence theory for similar nonlinear SE.

5. We successfully validate our optimal control framework with several numerical experiments.

This thesis is organized as follows. In Chapter 2, we introduce the TDKS model. We begin with presenting the multi-particle Schrödinger equation (SE). However, the computational effort for solving the SE increases exponentially with the number of particles. Therefore, we introduce the concept of the density functional theory (DFT). We give a clear account on the different maps and function sets used in DFT. We continue by presenting the Kohn-Sham approach that is essential for numerical application. Finally, we discuss the theoretical justification of time-dependent density functional theory (TDDFT) and the time-dependent Kohn-Sham (TDKS) framework. This includes — to the best of our knowledge — the first presentation of the fundamental van Leeuwen's theorem in a mathematically rigorous way.

In Chapter 3, we analyse the time-dependent Kohn-Sham equations which are a set of nonlinearly coupled Schrödinger equations. To this end, we introduce the weak form of the TDKS equations and make some reasonable assumptions on the nonlinear potentials and some preliminary estimates are presented. Then we introduce the Galerkin framework and show the pivotal energy estimates in this framework. These estimates allow us to prove existence and uniqueness of the solution of the TDKS equations. Further, by using stronger assumptions on the boundary of the domain and on the initial data, we show improved regularity results for the solution.

In Chapter 4, with the results of this analysis of the TDKS equations, we study optimization problems using the TDKS equations as PDE constraints. A general cost functional containing several different kinds of targets, such as a tracking type target and a final time target, are discussed. As we use the Lagrange framework for constraint optimization, we introduce the corresponding Lagrange functional. Existence of a minimizer is proved. One of the main contributions of this work is the analysis of the optimal control problem. In particular, we show differentiability properties of the TDKS constraint equations that allow us to derive a necessary optimality condition and present the optimality system. The fact that the TDKS equations depend nonlinearly on the density raises the issue that the TDKS equations are not complex differentiable. This is considered here; a more detailed discussion and a second equivalent approach to solve this issue are presented in the appendix.

In Chapter 5, we first discuss the discretization scheme we use to solve the forward and adjoint PDE. We implement a time-splitting pseudo-spectral method, in particular the Strang splitting, that achieves second-order accuracy in time and spectral convergence in space. An algorithm to efficiently calculate the global Hartree potential is also presented. Second, we present gradient-based optimization methods, notably the NCG and the BFGS methods. These algorithms are then used in the third part to solve quantum optimal control problems. We consider four different numerical examples using different kinds of targets covered in the theoretical analysis of Chapter 4. Furthermore, we compare L^2 and H^1 optimization and study the scaling of the computational complexity as a function of the particle number.

A summary in Chapter 6 concludes this work. In the appendix, the complex differentiability of the TDKS model is discussed in more detail and another approach — equivalent to the one presented in Chapter 4 — is presented. Furthermore, we collect some important results from functional analysis and some exact solutions for one- and two-particle (interacting) SEs to be used in numerical test cases. For the convenience of the reader, we include a nomenclature overview that contains the most frequently used notation on page 122. Each chapter contains a conclusion at the end to summarize the most important questions and facts.

Parts of the results presented in this thesis are published in the following literature.

- Martin Sprengel, Gabriele Ciaramella, Alfio Borzì, *A theoretical investigation of time-dependent Kohn–Sham equations*, SIAM Journal on Mathematical Analysis **49** (2017), 1681–1704.

- Martin Sprengel, Gabriele Ciaramella, Alfio Borzì, *Investigation of optimal control problems governed by a time-dependent Kohn-Sham model*, arXiv:1701.02679 [math.OC], 2017.

- Martin Sprengel, Gabriele Ciaramella, and Alfio Borzì, *A COKOSNUT code for the control of the time-dependent Kohn–Sham model*, Computer Physics Communications **214** (2017), 231–238.

- Martin Sprengel, Gabriele Ciaramella, Alfio Borzì, *Formulation and Numerical Solution of Quantum Control Problems*, Society for Industrial and Applied Mathematics, Philadelphia, 2017.

2. The framework of time-dependent density functional theory

In this chapter, we introduce the quantum mechanical framework of the time-dependent density functional theory. After a brief overview of the multi-particle Schrödinger equation in Section 2.1, we introduce the density functional theory (DFT) in Section 2.2. We discuss the abstract setting including the maps between potentials, wavefunctions and densities that form the core of DFT along with the Hohenberg-Kohn theorem whose proof is presented for non-degenerate and degenerate ground states. Moreover, we introduce the Kohn-Sham approach to DFT that is essential for applications of DFT.

In Section 2.3, we illustrate the extension of the DFT to the time-dependent DFT (TDDFT). This is done by proving the van Leeuwen's theorem, an extension of the Runge-Gross theorem that includes the usage of the Kohn-Sham approach in TDDFT.

As a specific implementation of TDDFT, we consider the time-dependent Kohn-Sham (TDKS) equations that represent the governing model of our control problem throughout this work.

2.1. The multi-particle Schrödinger equation

As opposed to classical mechanics where the state of a physical system is characterized by the positions and momenta of all particles involved, in quantum mechanics, the complete description of the system is given by a complex-valued wavefunction $\psi : \Omega \to \mathbb{C}$, $\Omega \subset \mathbb{R}^3$. The square of the absolute value of the wavefunction $|\psi(x,t)|^2$ can be interpreted as the probability density of finding a particle at this position. For systems of more than one particle, the system is described by a single wavefunction depending on all coordinates of all particles. We analyse problems of N interacting identical particles of mass m. The wavefunction is then a function defined on $3N$ space coordinates $x_1, \ldots, x_N \in \mathbb{R}^3$ and time t, $\Psi : \mathbb{R}^{3N+1} \to \mathbb{C}$.

The evolution of this wavefunction is governed by the Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(x_1, \ldots, x_N, t) = H\Psi(x_1, \ldots, x_N, t). \tag{2.1}$$

The Hamilton operator $H = \hat{T} + V + W$ consists of the kinetic operator term $\hat{T} = -\frac{\hbar^2}{2m} \sum_{i=1}^{N} \nabla_i^2$ where ∇_i is the derivative with respect to the coordinates of the ith particle, V denotes a (sum of) single-particle potential, and in this chapter W represents a particle-particle interaction potential.

In this setting, the issue arises of how to encode the fact that the particles are identical. In fact, to be physically meaningful, the probability distribution $|\Psi(x_1, \ldots, x_N, t)|^2$ must be symmetric to permutation of pairs of coordinates, i.e. particles. This condition is satisfied if the wavefunction itself is symmetric (bosons) or antisymmetric (fermions) when the coordinates of its constituent particles are exchanged. To illustrate this fact in more detail, consider the case of a quantum mechanical system with two interacting particles, subject to an external potential. The corresponding Hamiltonian is given by

$$H = -\frac{\hbar^2}{2m} [\nabla_1^2 + \nabla_2^2] + V(x_1) + V(x_2) + W(x_1 - x_2). \tag{2.2}$$

If the interaction term is zero, $W = 0$, then the wavefunction describing the state of this system is given by the (symmetrized or anti-symmetrized) product of the single particle wavefunctions, that is, the variables corresponding to the two particles separate.

More generally, we can consider a parity operator $\mathcal{P}$ which acts to permute the positions of the two particles as follows:

$$\mathcal{P}\Psi(x_1, x_2) = \Psi(x_2, x_1).$$

If we permute the two particles twice, we obtain the original wavefunction back $\mathcal{P}^2\Psi(x_1, x_2) = \Psi(x_1, x_2)$. This implies that the wavefunction is an eigenfunction of the parity operator and that the corresponding eigenvalue satisfies $\lambda^2 = 1$, which implies that $\lambda = \pm 1$. In the case where $\lambda = +1$, the particles are bosons and if $\lambda = -1$ the particles are fermions. Therefore, the wavefunction describing a two-particle system can be written as follows:

$$\Psi(x_1, x_2) = \psi_1(x_1)\psi_2(x_2) \pm \psi_2(x_1)\psi_1(x_2), \tag{2.3}$$

where the symmetric case applies to bosons and the antisymmetric case applies to fermions. The expression in (2.3) can be represented in terms of matrices with function-valued arguments. For fermions, the wavefunction Ψ_f is obtained as the following determinant:

$$\Psi_f(x_1, x_2) = \begin{vmatrix} \psi_1(x_1) & \psi_2(x_1) \\ \psi_1(x_2) & \psi_2(x_2) \end{vmatrix}.$$

For bosons, the wavefunction Ψ_b can be expressed as a matrix permanent as follows:

$$\Psi_b(x_1, x_2) = \operatorname{perm} \begin{pmatrix} \psi_1(x_1) & \psi_2(x_1) \\ \psi_1(x_2) & \psi_2(x_2) \end{pmatrix}.$$

These constructions are called Slater determinants and Slater permanents. In the case of N particles, we have

$$\Psi_f(x_1, \ldots, x_N) = \det\big((\psi_j(x_i))_{ij}\big) := \begin{vmatrix} \psi_1(x_1) & \cdots & \psi_N(x_1) \\ \vdots & \ddots & \vdots \\ \psi_1(x_N) & \cdots & \psi_N(x_N) \end{vmatrix},$$

$$\Psi_b(x_1, \ldots, x_N) = \operatorname{perm} \begin{pmatrix} \psi_1(x_1) & \cdots & \psi_N(x_1) \\ \vdots & \ddots & \vdots \\ \psi_1(x_N) & \cdots & \psi_N(x_N) \end{pmatrix}.$$

Consider again a system consisting of two electrons. The wavefunction can be separated into a space-dependent wavefunction ψ_x and a spin part ψ_s. The possible spin states are any combination of the following four unique configurations:

$$| \uparrow\uparrow \rangle, \quad | \uparrow\downarrow \rangle, \quad | \downarrow\uparrow \rangle, \quad | \downarrow\downarrow \rangle.$$

These spin states can be combined to create a new basis with either symmetric or antisymmetric components. These states are denoted as follows:

$$
\begin{aligned}
\text{triplet states} \quad & \begin{cases} |1,1\rangle &= | \uparrow\uparrow \rangle \\ |1,0\rangle &= \frac{1}{\sqrt{2}}(| \uparrow\downarrow \rangle + | \downarrow\uparrow \rangle), \\ |1,-1\rangle &= | \downarrow\downarrow \rangle \end{cases} \\
\text{singlet state} \quad & |0,0\rangle = \frac{1}{\sqrt{2}}(| \uparrow\downarrow \rangle - | \downarrow\uparrow \rangle).
\end{aligned}
$$

Notice that the product $\psi_x \psi_s$ must be antisymmetric by the Pauli exclusion principle. Therefore, if ψ_s is a triplet spin state, ψ_x must be a Slater determinant for the spatial component, whereas if ψ_s is a singlet state, ψ_x must correspondingly be a Slater permanent. In the presence of an interaction potential, the two-particle wavefunction will not be a single Slater determinant or permanent. However, it can be expressed as an infinite sum of determinants or permanents. Although in practice the true wavefunction may be reasonably well approximated by truncating the sum after only a few terms. Based on this consideration, powerful methods have been developed that allow to bypass, at least in part, the prohibitive exponential computational complexity of solving (2.1). In particular, we refer to the multi-configuration time-dependent Hartree method; see, e.g., [BJWM00, CL10, Koc06] and the density functional theory discussed in this work.

As the description of spin in the framework of TDDFT is quite involved, we do not explicitly consider the spin of the particles. However, the repulsion due to the Pauli exclusion principle for fermions is included in the single-particle potential under consideration.

2.2. The density functional theory

Many models of interest in quantum physics and chemistry consist of multi-particle systems, that can be modelled by the multi-particle Schrödinger equation (SE). For example, consider a molecule of N electrons. Since the nuclei are much heavier than the electrons (nearly by a factor of 2000 per proton or neutron), the nuclei are moving much slower and can be considered as static such that it is reasonable to study the system of the N interacting electrons in the static potential of the nuclei. This is the Born-Oppenheimer approximation. This system is fully characterized by the N-particle wavefunction Ψ depending on the $3N$ coordinates of all electrons. For $N > 2$, no analytic solution exists for such a system. Therefore, for the solution one has to rely on numerical methods. As the space dimensionality of this equation increases linearly with the number of particles involved N, the corresponding computational cost increases exponentially, thus making the use of the multi-particle SE prohibitive.

To illustrate the exponential increase, consider a simple molecule like water with $N = 10$ electrons. This is already a complicated challenge that needs sophisticated methods because even storing the wavefunction using a standard discretization with 40 grid points in each coordinate direction results in about 10^{48} numbers to store which is approximately the number of atoms in the earth.

This curse of dimensionality has motivated great research efforts towards an alternative to the SE description that allows to compute the observables of a quantum multi-particle system using particle density functions. This development, which started with the Thomas-Fermi theory in 1927 (named after Llewellyn H. Thomas and Enrico Fermi), reached a decisive point with the works of Pierre C. Hohenberg, Walter Kohn, and Lu Jeu Sham [HK64, KS65] in 1964 and 1965 who proposed an appropriate way to replace a system of N interacting particles in an external potential V_{ext} by another system of non-interacting particles with an external potential $V_{ext} + V_{Hxc}$, such that the two models have the same electronic density. Hence, all observables formulated in terms of the density can be determined from both models where the second is numerically feasible because the task scales linearly instead of exponentially with the number of particles.

The density associated with the N-particle wavefunction $\Psi(x_1, \dots, x_N)$ is defined as

follows:

$$\rho(x) := N \int_{\Omega^{N-1}} |\Psi(x, \ldots, x_N)|^2 \mathrm{d}x_2 \cdots \mathrm{d}x_N, \tag{2.4}$$

where the integral is understood as to include a sum over all N spin values ($\uparrow$ or $\downarrow$ for electrons). We want to draw attention to the fact that independently of the number of particles N, the density ρ depends on just 3 space coordinates. As we consider electrons, the wavefuncion is antisymmetric in the sense that, only changing space and not spin variables, we have

$$\Psi(x, \ldots, x_k, \ldots, x_N) = -\Psi(x_k, \ldots, x, \ldots, x_N), \quad \forall k = 2, \ldots, N,$$

and hence any coordinate can be chosen instead of x_1 as the free coordinate in (2.4). An early comprehensive work on the mathematical aspects of density function theory (DFT) is given by Elliott H. Lieb in [Lie85].

These works and many following ones focused on the computation of stationary (ground) states and obtained successful results that motivated the extension of the DFT to include time-dependent phenomena. This extension was first proposed by Erich Runge and Eberhard K. U. Gross in [RG84] in 1984, and further investigated from a mathematical point of view in the work of van Leeuwen [vL99]. We refer to [ED11] for a modern introduction to DFT and to [MUN+06] for an introduction to time-dependent DFT (TDDFT).

Similar to the stationary case, the Runge–Gross theorem proves that, given an initial wavefunction configuration, there exists a one-to-one mapping between the potential in which the system evolves and the density of the system. Therefore, under appropriate assumptions, given a SE for a system of interacting particles in an external potential, there exists another SE model, unique up to a purely time-dependent function in the potential [vL99], of a non-interacting system with an augmented potential whose solution provides the same density as the solution to the original SE problem. We refer to this TDDFT model as the time-dependent Kohn-Sham (TDKS) equations. These TDKS equations are discussed in detail in Section 2.3.

2.2.1. Foundations of density functional theory

In this and the next subsection, we consider the stationary case before these results are extended to the time-dependent case in Section 2.3. The central idea of DFT is to consider the particle density instead of the wavefunction of a quantum system. One might ask, whether the density still contains the full information about the system despite its much simpler structure. Astonishingly, this is indeed the case for some systems. To answer this question in detail, the following maps are considered.

A potential V gives rise to a ground state wavefunction Ψ by means of the solution of the stationary SE

$$M_{VW} : V \mapsto \Psi, \qquad \Psi \text{ is solution to } \left(-\frac{\hbar^2}{2m} \nabla^2 + V \right) \Psi = E\Psi. \tag{2.5}$$

The second map that we consider is $M_{WA} : \Psi \mapsto \rho$ defined by (2.4).

Now, the issue is to specify when the inverse map $\rho \mapsto V(\rho)$ exists. This depends crucially on the domain and on the range on which the maps are considered. For this reason, we start defining the set of N-particle wavefunctions as follows:

$$\mathbb{W} := \left\{ \Psi \in H^1(\Omega^N; \mathbb{C}) \,\middle|\, \int_{\Omega^N} |\Psi(x_1, \ldots, x_N)|^2 \mathrm{d}x_1 \cdots \mathrm{d}x_N = 1 \right\},$$

	Potential	Wavefunction	Density
all N particle	$\mathbb{V}$	$\mathbb{W}$	$\mathbb{A}$
with a ground state	V_N	W_N	A_N
with a non-degenerate ground state	$\mathcal{V}$	$\mathcal{W}$	$\mathcal{A}$

Table 2.1.: DFT's function sets.

where $\Omega = \mathbb{R}^3$ in this section. Further, we define the following set of N-particle densities:

$$\mathbb{A} := \{\rho \in L^1(\Omega; \mathbb{R}) | \rho(x) \geq 0, \sqrt{\rho} \in H^1(\Omega; \mathbb{R}), \|\rho\|_{L^1} = N\}.$$

The condition $\sqrt{\rho} \in H^1(\Omega; \mathbb{R})$ ensures that the kinetic energy of the density is finite. It can be shown, that $\mathbb{A} \subset L^1 \cap L^3$; see, e.g., [Lie85].

Since we want to study potentials with finite expectation values, the term

$$\int_\Omega \rho(x)|v(x)|\mathrm{d}x = \|\rho v\|_{L^1(\Omega)}$$

of the Hamiltonian has to be finite. If we use the fact that $\rho \in L^1(\Omega)$, we find that $v \in L^\infty(\Omega)$ is admissible. On the other hand, if we use $\rho \in L^3(\Omega)$, V can be in $L^{3/2}(\Omega)$ by Hölder's inequality. Therefore, we define the space of potentials as the set of functions that can be decomposed into a sum of a $L^\infty(\Omega)$ and a $L^{3/2}(\Omega)$ function as follows:

$$\mathbb{V} = L^\infty(\Omega) + L^{3/2}(\Omega), \qquad \|v\|_{\mathbb{V}} = \inf_{g \in L^{3/2}, h \in L^\infty, g+h=v} \|g\|_{L^{3/2}(\Omega)} + \|h\|_{L^\infty(\Omega)}.$$

This excludes potentials that diverge for large x. However, it includes the Coulomb potential, which is one of the most important potentials in quantum physics and chemistry. In fact, we have the following decomposition:

$$\frac{1}{|x|} = \frac{e^{-|x|}}{|x|} + \frac{1 - e^{-|x|}}{|x|},$$

where the first summand is integrable in $L^{3/2}(\Omega)$ and the second one is bounded.

The global phase of the wavefunction is determined by the initial condition and does not influence the value of any observable. Therefore, it has no physical meaning, in the sense that wavefunctions differing only by a global phase are considered equivalent. As a constant in the potential changes only the global phase of the wavefunctions, potentials that differ only by a constant are also considered equivalent.

DFT is a theory about the ground state, that is, the wavefunction associated with the lowest possible energy. We introduce the potentials that have at least one ground state, which define the set $V_N \subset \mathbb{V}$, and the potentials having a non-degenerate ground state, i.e. potentials for which a unique solution to (2.5) with lowest eigenvalue exists. These potentials define the set $\mathcal{V} \subset V_N$. Correspondingly, the ground state wavefunctions are denoted by $W_N \subset \mathbb{W}$, and the ground state densities with A_N. The non-degenerate ground state wavefunctions are denoted by $\mathcal{W} \subset W_N$. All subsets are proper subsets. These definitions are summarized in Table 2.1.

Now, we study the maps M_{VW} and M_{WA} between the different sets, see Figure 2.1 for an overview. After discussing their properties, we prove most of them in the following lemmas; see [Lie85] for the remaining proofs.

Since not all potentials have a ground state and not all wavefunctions are solutions of the SE (2.5), the map M_{VW} can only be discussed from V_N to W_N and not from $\mathbb{V}$ to
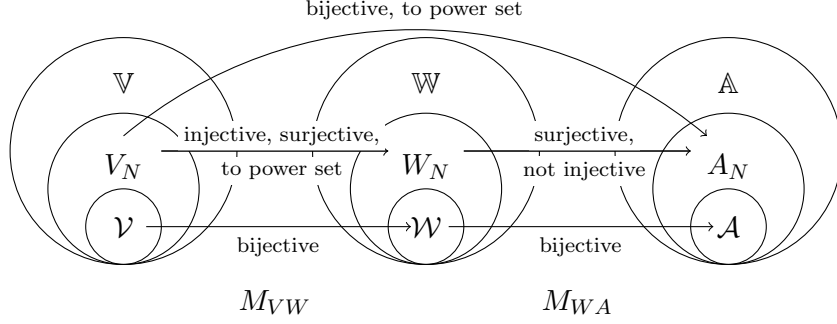
Figure 2.1.: The different subsets of potentials, wavefunctions and densities and the properties of the maps M_{VW} and M_{WA} between them.

$\mathbb{W}$. The map $M_{VW} : V_N \to W_N$ is not well-defined as one potential can have multiple ground states. More precisely, it maps not to W_N but to the power set of W_N. With this adjustment M_{VW} is surjective to W_N by definition of W_N and by Lemma 2.1 it is injective. For this reason, one defines the sets $\mathcal{V}$ and $\mathcal{W}$ such that $M_{VW} : \mathcal{V} \to \mathcal{W}$ is well-defined and bijective.

The map M_{WA} is always well-defined for all $\Psi \in \mathbb{W}$ by (2.4) and surjective to $\mathbb{A}$, see Lemma 2.3, but not injective, see Lemma 2.4. $M_{WA} : W_N \to A_N$ is surjective by definition, but not injective, whereas $M_{WA} : \mathcal{W} \to \mathcal{A}$ is surjective by definition and injective. There exists extensions using density matrices instead of ground state wavefunctions to define M_{WA} for a larger class of densities.

The original theorem of Hohenberg and Kohn (HK), stated in Theorem 2.5 considers the situation of non-degenerate ground states and proves the bijectivity of the maps $M_{VW} : \mathcal{V} \to \mathcal{W}$ and $M_{WA} : \mathcal{W} \to \mathcal{A}$. This theorem can be generalized to show that the combined map $M_{WA} \circ M_{VW} : V_N \to A_N$ is bijective even though $M_{VW} : V_N \to W_N$ maps to the power set and $M_{WA} : W_N \to A_N$ is not injective, see Theorem 2.6. This is because the injectivity of M_{WA} is only violated for wavefunctions being different ground states of the same potential. Before proving the HK theorem, we study the maps M_{VW} and M_{WA} separately.

Lemma 2.1. *The map $M_{VW} : V_N \to W_N$ is injective, if potentials that differ only by a constant are summarized in an equivalence class.*

Proof. We want to prove the statement by contradiction. Given two potentials V_1 and V_2 that differ by more than a constant with corresponding ground states Ψ_1 and Ψ_2. If the potentials are degenerate, take any of their ground states. Now, assume that $\Psi_1 = \Psi_2 = \Psi$. Subtracting the corresponding eigenvalue equations gives the following:

$$(E_1 - E_2)\Psi = (V_1 - V_2)\Psi \qquad \Leftrightarrow \qquad 0 = (V_1 - V_2)\Psi,$$

because V_1, V_2 are equivalence classes of potentials differing by more than a constant like $E_1 - E_2$. By the unique continuation theorem, [RS78], the wavefunction cannot be zero on a set of positive measure. Otherwise it is zero almost everywhere and hence not normalizable. Therefore, we have $V_1 = V_2$ which contradicts our assumption. $\qquad\square$

We remark that Lieb [Lie85] comments about the unique continuation theorem, that "Strictly speaking, this theorem is only known to hold for L_{loc}^3, but it is believed to hold for $L^{3/2} + L^{\infty}$".

Before showing the results on M_{WA}, we have to study Slater determinants in more detail in the following lemma.

Lemma 2.2. *Given N orthonormal wavefunctions $\{\psi_i\}_i$, $\psi_i \in H^1(\Omega; \mathbb{C})$. Their Slater determinant $\Psi(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}} \det\big((\psi_i(x_j))_{ij}\big)$ is a normed wavefunction in $\mathbb{W}$ and has the density $\rho(x) = \sum_{i=1}^{N} |\psi_i(x)|^2 \in \mathbb{A}$.*

Proof. Using the Leibniz formula for the determinant and denoting the permutations of $1, \ldots, N$ with π and σ, direct calculation shows

$$
\begin{aligned}
\rho(x) &= N \int |\Psi(x, x_2 \ldots, x_N)|^2 \mathrm{d}x_2 \cdots \mathrm{d}x_N \\
&= N \int \frac{1}{N!} \left| \det\big((\psi_i(x_j))_{ij}\big) \right|^2 \mathrm{d}x_2 \cdots \mathrm{d}x_N \\
&= \frac{N}{N!} \int \sum_\pi \prod_{i=1}^{N} (-1)^\pi \psi_i(x_{\pi(i)}) \sum_\sigma \prod_{j=1}^{N} (-1)^\sigma \overline{\psi_j(x_{\sigma(j)})} \mathrm{d}x_2 \cdots \mathrm{d}x_N \\
&= \frac{N}{N!} \sum_\pi \sum_\sigma (-1)^{\pi+\sigma} \int \prod_{i=1}^{N} \psi_i(x_{\pi(i)}) \overline{\psi_i(x_{\sigma(i)})} \mathrm{d}x_2 \cdots \mathrm{d}x_N \\
&= \frac{N}{N!} \sum_\pi \sum_\sigma (-1)^{\pi+\sigma} \int \prod_{i=1}^{N} \psi_{\pi^{-1}(i)}(x_i) \overline{\psi_{\sigma^{-1}(i)}(x_i)} \mathrm{d}x_2 \cdots \mathrm{d}x_N \\
&= \frac{N}{N!} \sum_\pi \sum_\sigma \underbrace{(-1)^{\pi+\sigma}}_{=1, \text{ if } \pi=\sigma} \prod_{i \neq 1} \underbrace{\int \psi_{\pi^{-1}(i)}(x_i) \overline{\psi_{\sigma^{-1}(i)}(x_i)} \mathrm{d}x_i}_{\delta_{\pi^{-1}(i)\sigma^{-1}(i)}} \cdot \left(\psi_{\pi^{-1}(1)}(x) \overline{\psi_{\sigma^{-1}(1)}(x)} \right) \\
&= \frac{N}{N!} \sum_\pi \psi_{\pi^{-1}(1)}(x) \overline{\psi_{\pi^{-1}(1)}(x)} = \sum_{i=1}^{N} |\psi_i(x)|^2,
\end{aligned}
$$

because for all $i = 1, \ldots N$ there are $(N-1)!$ permutations π, such that $\pi^{-1}(1) = i$. Together with the factor N this cancels the $N!$. Analogously, one shows that the kinetic energy of the Slater determinant is given by

$$
\hat{T}(\Psi) = \frac{\hbar^2}{2m} \sum_{i=1}^{N} \int |\nabla \psi_i(x)|^2 \mathrm{d}x.
$$

The norm of Ψ is given by $\|\Psi\|_{L^2(\Omega^N)}^2 = \frac{1}{N} \int_\Omega \rho(x) \mathrm{d}x = \frac{1}{N} \sum_{i=1}^{N} \int_\Omega |\Psi_i(x)|^2 \mathrm{d}x = 1$. From $\hat{T}(\Psi) < \infty$ one has that $\sqrt{\rho} \in H^1(\Omega)$, hence $\rho \in \mathbb{A}$. $\qquad\square$

Lemma 2.3. *Consider the map $M_{WA} : \mathbb{W} \to \mathbb{A}$. We have that M_{WA} is well-defined and surjective.*

Proof. By definition of $\mathbb{W}$ and $\mathbb{A}$, for every wavefunction $\Psi \in \mathbb{W}$ there exists a density $\rho \in \mathbb{A}$. Hence, M_{WA} is well-defined. Next, we show surjectivity of M_{WA}. To do so, we prove that for every density $\rho \in \mathbb{A}$ there exists a wavefunction $\Psi \in \mathbb{W}$ (as a Slater determinant), such that

$$
\rho(x) = N \int_{\Omega^{N-1}} |\Psi(x, \ldots, x_N)|^2 \mathrm{d}x_2 \cdots \mathrm{d}x_N
$$

holds.

Given a density ρ, we explicitly construct a wavefunction whose density is given by ρ. To this end, let $x = (x_1, x_2, x_3)$ and define

$$
f(x_1) := \frac{2\pi}{N} \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho(s, x_2, x_3) \mathrm{d}x_2 \mathrm{d}x_3 \mathrm{d}s.
$$

Then f is monotonically increasing from 0 to 2π. Now, define the following functions:

$$\phi^k(x) = \sqrt{\frac{\rho(x)}{N}} \exp\left(ikf(x_1)\right), \quad k = 0, \ldots, N-1.$$

These functions are orthogonal in $L^2(\mathbb{R}^3)$, as the following calculation shows:

$$\left\langle \phi^k, \phi^l \right\rangle_{\mathbb{R}^3} = \int_{\mathbb{R}^3} \phi^k(x)\overline{\phi^l(x)}\mathrm{d}x = \int_{\mathbb{R}^3} \frac{\rho(x)}{N} e^{if(x_1)(k-l)}\mathrm{d}x$$

$$= \int_{\mathbb{R}} \frac{1}{2\pi} \frac{\mathrm{d}f(y)}{\mathrm{d}y}\bigg|_{x_1} e^{if(x_1)(k-l)}\mathrm{d}x_1 = \begin{cases} \frac{1}{2\pi(k-l)} e^{if(x_1)(k-l)}|_{-\infty}^{\infty} = 0 & k \neq l, \\ \frac{1}{2\pi} f(x_1)|_{-\infty}^{\infty} = 1 & k = l. \end{cases}$$

By Lemma 2.2, $\Psi(x_1, \ldots, x_N) := \frac{1}{\sqrt{N!}} \det\left((\phi_i(x_j))_{ij}\right) \in \mathbb{W}$ and has the density

$$\sum_{i=1}^{N} |\phi_i(x)|^2 = \sum_{i=1}^{N} \frac{\rho(x)}{N} = \rho(x).$$

$\square$

Lemma 2.4. *The map $M_{WA} : \mathbb{W} \to \mathbb{A}$ is not injective, even if one considers equivalence classes of wavefunctions that only differ by a global phase.*

Proof. We present a counterexample for $N = 1$ particle and in $n = 1$ dimensions. Let $\Psi(x) = \sqrt{\frac{5}{2}} e^{-|x|} \sin(2x) \in \mathbb{W}$. Then $\Psi_1(x) = \Psi(x)e^{i\sin(x)}$ and $\Psi_2(x) = \Psi(x)e^{i\sin(2x)}$ are in $\mathbb{W}$, too, and Ψ_1 and Ψ_2 differ by more than a global phase. However, the densities are obviously identical, as they depend only on $|\Psi|^2$, hence $\rho_1(x) = \rho_2(x)$. $\square$

With this preparation, we can now discuss the theorem of Hohenberg and Kohn, published in [HK64].

Theorem 2.5 (Hohenberg and Kohn)**.** *Given two potentials V_1 and V_2 that have non-degenerate ground states and differ by more than a constant, then their ground state densities ρ_1 and ρ_2 differ. In other words, $M_{WA} \circ M_{VW} : \mathcal{V} \to \mathcal{A}$ is a bijection.*

Proof. The proof uses the variational property of the ground state. To obtain a contradiction, we assume that $\rho_1 = \rho_2 = \rho$. However, the corresponding ground states differ, $\Psi_1 \neq \Psi_2$, as discussed in Lemma 2.1. As the ground states are non-degenerate, we have for the ground state energies E_1 and E_2 corresponding to the ground states $\Psi_1 \neq \Psi_2$

$$E_1 < \langle \Psi_2, H_1\Psi_2 \rangle_{L^2(\mathbb{R}^3;\mathbb{C}^N)} = \langle \Psi_2, H_2\Psi_2 \rangle_{L^2(\mathbb{R}^3;\mathbb{C}^N)} + \langle \Psi_2, (H_1 - H_2)\Psi_2 \rangle_{L^2(\mathbb{R}^3;\mathbb{C}^N)}$$

$$= E_2 + \int_{\Omega} \rho(x)(V_1 - V_2)(x)\mathrm{d}x,$$

and similarly

$$E_2 < E_1 + \int_{\Omega} \rho(x)(V_2 - V_1)(x)\mathrm{d}x.$$

Adding both lines yields the contradiction to conclude the proof. $\square$

This theorem can be generalized to the case of degenerate ground states as follows.

Theorem 2.6 (Hohenberg and Kohn 2)**.** *Theorem 2.5 holds also for degenerate systems, i.e. the map $M_{WA} \circ M_{VW} : V_N \to A_N$ is bijective.*

Proof. The map $M_{WA} \circ M_{VW}$ is surjective by the definition of the sets V_N and A_N.

Let $A_1 = \{\Psi_1, \ldots, \Psi_K\}$ be the ground states of V_1 and $A_2 = \{\Phi_1, \ldots, \Phi_L\}$ the ones of V_2. Then by Lemma 2.1 $\Psi_k \neq \Phi_l \, \forall k = 1, \ldots, K, l = 1, \ldots, L$. Taking linear combinations of the wavefunctions one finds $\Psi \neq \Phi \, \forall \Psi \in \text{span} \, A_1, \Phi \in \text{span} \, A_2$; in other words $\text{span} \, A_1 \cap \text{span} \, A_2 = \emptyset$. Therefore, $\forall k = 1, \ldots, K, l = 1, \ldots, L$, we find for the ground state energies E_1 of V_1 and E_2 of V_2

$$
\begin{aligned}
E_1 &= \langle \Psi_l, H_1 \Psi_l \rangle_{L^2(\mathbb{R}^3; \mathbb{C}^N)} < \langle \Phi_k, H_1 \Phi_k \rangle_{L^2(\mathbb{R}^3; \mathbb{C}^N)} \\
&= \langle \Phi_k, H_2 \Phi_k \rangle_{L^2(\mathbb{R}^3; \mathbb{C}^N)} + \int_\Omega \rho_k(x)(V_2(x) - V_1(x)) \mathrm{d}x \\
&= E_2 + \int_\Omega \rho_k(x)(V_1(x) - V_2(x)) \mathrm{d}x
\end{aligned}
$$

and analogously for E_2. Summing both inequalities yields

$$
\int_\Omega (\rho_k(x) - \rho_l(x))(V_2(x) - V_1(x)) \mathrm{d}x < 0.
$$

Hence, the densities coming from different potentials have to differ in the Lebesgue sense. $\square$

The HK theorems are important to DFT because they state that given a density $\rho \in A_N$, it is theoretically possible to determine the potential $V \in V_N$ it is generated from. With this potential, the SE (2.5) is known and the wavefunction $\Psi \in W_N$ could be calculated. Hence, the density contains the full information about the system and the ground state energy and — in the case of $\rho \in \mathcal{A}$ — all ground state observables can be calculated from ρ as follows:

$$
O(\rho) = \left\langle \Psi(\rho), \hat{O} \Psi(\rho) \right\rangle_{L^2(\mathbb{R}^3; \mathbb{C}^N)}.
$$

If the system has a degenerate ground state, only the observables that can be written in terms of the density can be determined.

The HK theorem states a bijection between potentials and densities for the sets $\mathcal{V}$ and $\mathcal{A}$ and V_N and A_N. We remark that the sets $\mathcal{A}$ and A_N are defined such that this bijection holds and there is no other characterization of them. Only a few properties of them are proved and their exact structure remains unknown.

The reason for asking for a non-degenerate ground state in Theorem 2.5 is that in this case it is possible to single out exactly one state with a variational principle. Energy is chosen as a variable because it is a very important quantity. However, it would also be possible to define a density theory based on, e.g., angular momentum, if one is interested in the state of lowest angular momentum.

Furthermore, we remark that the HK theorem states that Schrödinger equations with different potentials have different solutions. This is of course stronger than asking for a single PDE to have a unique solution.

2.2.2. The Kohn-Sham approach

The HK theorem states that in principle it is completely equivalent to describe a quantum system by the density instead of the wavefunction. This equivalence is obtained by showing that it is possible to find the potential giving rise to the density and then solve the SE with this potential. However, this does not provide a way to escape the problem of solving the high-dimensional SE. The breakthrough for a successful application of

DFT was developed by Kohn and Sham (KS) in [KS65]. They studied the SE of N interacting particles

$$\left(-\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N}V_{ext}(x_i) + W\right)\Psi = E\,\Psi,$$

where W is the interaction between the particles. For electrons, it is given by the Coulomb repulsion $W = \sum_{i,j=1}^{N}\frac{1}{4\pi\epsilon_0}\frac{e^2}{2|x_i-x_j|}$, where ϵ_0 is the vacuum permittivity.

Kohn and Sham recognized that it is possible to obtain the same density ρ from a non-interacting SE with an additional single-particle potential $V_{KS} = V_{ext} + V_{Hxc}$, which embodies the interaction between particles. For this purpose, the lowest N eigenvalues and corresponding orthogonal eigenfunctions of the following problem are determined:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{ext}(x) + V_{Hxc}(x,\rho)\right)\psi_i = E_i\,\psi_i. \tag{2.6}$$

Forming the Slater determinant of these orthogonal eigenfunctions, $\tilde{\Psi} = \det(\psi_i(x_j))$, $i,j = 1,\ldots,N$, a solution to

$$\sum_{i=1}^{N}\left(-\frac{\hbar^2}{2m}\nabla_i^2 + V_{ext}(x_i) + V_{Hxc}(x_i,\rho)\right)\tilde{\Psi} = \left(\sum_{i=1}^{N}E_i\right)\tilde{\Psi} \tag{2.7}$$

is found. We remark that for Slater determinants Lemma 2.2 applies and the density is given by $\rho(x) = \sum_{i=1}^{N}|\psi_i(x)|^2$.

The key problem is therefore to choose V_{KS} such that the density corresponding to $\tilde{\Psi}$ is the same as the density ρ associated to Ψ. If this is true, then it is possible to calculate the value of all observables, which can be written in terms of the density, by solving N single-particle problems. In particular, since the energy is a functional of the density alone, we have $E = \sum_{i=1}^{N}E_i$. So a linear growth of the complexity of the problem with the number of particles can be achieved instead of an exponential one as in the interacting case.

The KS potential V_{KS} is not explicitly known. A complete analytic form of V_{KS} would solve the multi-particle SE, which cannot be expected. Finding a good approximation of the true V_{KS} is one of the main challenges in DFT. Since the emergence of the KS approach, there has been intense research on approximations to V_{KS} that have proved to be successful in applications.

In this quest, the prevalent approach is the so-called local density approximation (LDA). It assumes that the quantum mechanical part of V_{Hxc} depends pointwise on the density alone. To find an explicit formula, the KS potential is decomposed into the following three terms:

$$V_{Hxc}(x,\rho) = V_H(x,\rho) + V_x(\rho) + V_c(\rho). \tag{2.8}$$

In (2.8), the Hartree potential

$$V_H(x,\rho) = \frac{e^2}{4\pi\epsilon_0}\int_{\Omega}\frac{\rho(y)}{|x-y|}\mathrm{d}y \tag{2.9}$$

is defined as the convolution of the density with the Coulomb potential. It represents the classical electromagnetic interaction with a fixed system of independent electrons. In LDA, this is the potential term that depends globally on the density. The other two potential terms are local and represent quantum mechanical corrections.

In quantum mechanics the Pauli exclusion principle states that two electrons cannot share the same quantum state; see Section 2.1. This results in an additional repulsive interaction between the electrons. In density functional theory, this phenomenon is modelled by two terms: the exchange potential V_x and the correlation potential V_c.

The exchange potential contains the Pauli principle for a non-interacting homogeneous electron gas, and it can be calculated explicitly in the case of LDA (see [Con08, PY89]) as follows:

$$V_x^{2D}(x, \rho) = -\frac{e^2}{4\pi\epsilon_0}\sqrt{\frac{8}{\pi}}\sqrt{\rho(x)}, \qquad V_x^{3D}(x, \rho) = -\frac{e^2}{4\pi\epsilon_0}\sqrt[3]{\frac{3}{\pi}}\sqrt[3]{\rho(x)},$$

where V_x^{2D} refers to the two-dimensional case and V_x^{3D} to the three-dimensional case.

The remaining part of the interaction is about one order of magnitude smaller and it is called the correlation potential. No analytic expression is known for V_c. However, it is possible to derive the limits for high and low densities and to calculate the potential at single intermediate values of the density using Quantum Monte Carlo methods. These results are then interpolated to give a function of ρ; see, e.g., [AMGGB02]. In Figure 2.2, we plot the shape of V_c used in the numerical experiments in Section 5.3.
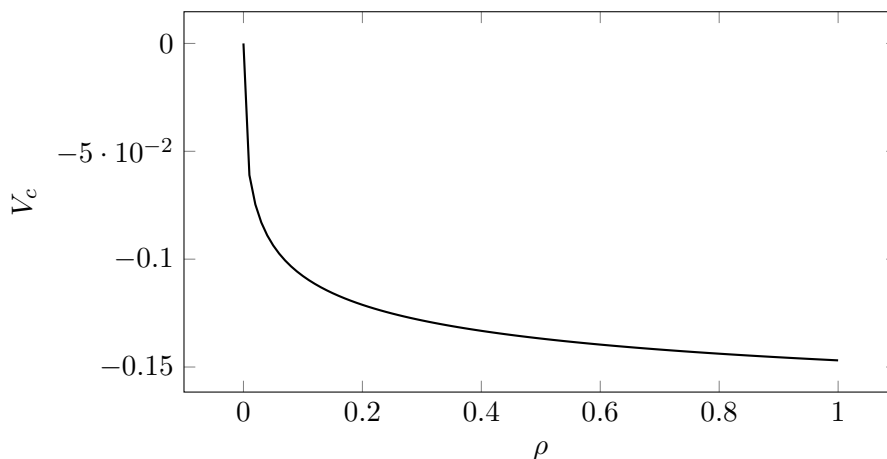
Figure 2.2.: The Quantum Monte Carlo fit used in the numerical experiments of this work from [AMGGB02]. The values for the limits are $\lim_{\rho\to 0} V_c(\rho) = 0$ and $\lim_{\rho\to\infty} V_c(\rho) = -0.1925$.

All correlation potentials commonly used are of similar structure; see, e.g., [MOB12]. They are zero for zero density and otherwise negative, while having a convex shape. Furthermore, they are bounded by V_x in the sense that $|V_c(\rho)| < |V_x(\rho)|$ for all $\rho \in \mathbb{R}^+$.

There are extensions of LDA using potentials V_{Hxc} that depend, in addition to ρ, also on $\nabla\rho$ (generalized gradient approximation, GGA), and also on ρ, $\nabla\rho$, $\nabla^2\rho$ (meta-GGA) and so forth; see, e.g., [ED11].

Furthermore, LDA can be understood as an expansion around the limit of non-interacting particles because V_H, which describes the interaction with a given electronic density, is the dominant term. It is also possible to pursue the approach to interpolate between no and infinitely strong interaction, which is the so-called strictly correlated electrons (SCE) approach; see, e.g., [SGGS07].

2.3. Time-dependent density functional theory

The DFT is a stationary ground state theory. To study excited states and the time evolution of a multi-particle quantum system, DFT was extended to time-dependent DFT (TDDFT) by Runge and Gross in 1984.

The original theorem stated in [RG84] is the following. Consider every single-particle potential $V_{ext}(x, t)$ which can be expanded into a Taylor series with respect to the time coordinate around $t = t_0$. Consider the map $G : V_{ext} \mapsto \rho$ defined as

$$G(V_{ext}) := \rho(\Psi(V_{ext})),$$

where $\Psi(V_{ext})$ is the solution to the time-dependent Schrödinger equation

$$i\frac{\partial \Psi}{\partial t} = \left(-\frac{1}{2}\nabla^2 + W + V_{ext}\right)\Psi, \quad \Psi(0) = \Psi^0,$$

and ρ is the density corresponding to $\Psi(V_{ext})$. The map G is injective up to an additive merely time-dependent function in the potential. A rigorous theorem that also specifies the spaces between which the map is invertible is given later in Theorem 2.10.

Remark. We remark that there are different units used in physics. The most widespread system of units is the "International System of Units" (SI). In experimental physics and chemistry, these units are used nearly exclusively. However, in theoretical physics it is common to use so called "natural units". In this setting, natural constants describing the characteristic physical scale are defined to be 1 which is the mathematical unity (multiplicative identity). Obviously, this procedure does not only change the numbers but also the dimension of the quantities. In this section, we follow the literature on TDDFT which uses exclusively natural units in the form of $\hbar = m = \frac{e^2}{4\pi\epsilon_0} = 1$. If not explicitly stated otherwise, we use natural units throughout the rest this thesis.

Since there is no variational principle available for the time-dependent case, some of the DFT concepts have to be adapted. However, it is still possible to introduce a time-dependent Kohn-Sham (TDKS) scheme which allows to transform the $3N$ dimensional interacting time-dependent SE to a system of N single-particle SEs. The theoretical justification for this formulation is given by an extension of the Runge-Gross theorem due to van Leeuwen, see [vL99, Boi].

We remark that the investigation of TDDFT is an active research topic. In particular, there are only few results on the theory of existence and uniqueness of solutions to the time-dependent Kohn-Sham equations, [Jer15, SCB17c]. Taking into account only the Hartree potential, some results can also be found in [CL99]. However, in applications, TDDFT is already widely used in quantum physics and chemistry and there are several software packages available for this purpose; see, e.g., the code Octopus [CAO+06].

Now, in order to continue our discussion and state the van Leeuwen's theorem, we make the following assumption.

Assumption (TDDFT 1)**.** We assume the following:

(a) $\Omega \subset \mathbb{R}^3$ is a bounded domain with sufficiently smooth boundary, e.g. $\partial\Omega \in C^2$; the boundary condition is only necessary for Assumption (e);

(b) $\Psi \in C^1([0, T]; C^2(\Omega^N))$ and $\frac{\partial^2 \Psi}{\partial x \partial t} = \frac{\partial^2 \Psi}{\partial t \partial x}$;

(c) the density $\rho : [0, T] \rightarrow C^1(\Omega)$ is analytic at $t = t_0$ and $\frac{\partial^k \rho}{\partial t^k}(t_0) \in C^1(\Omega) \; \forall k \in \mathbb{N}_0$;

(d) $\Psi(x,t) = 0$ for all $t \in [0,T]$ and $x \in \partial\Omega$.

Further, we define the following short-hand notation: We omit the arguments in the wavefunction, $\Psi := \Psi(x_1, \ldots, x_N, t)$ and denote the gradient with respect to the coordinates of the j-th particle x_j by ∇_j. For the first particle, the index is omitted, i.e. $x_1 = x$ and $\nabla_1 = \nabla$. Moreover, we define the probability current by

$$\mathcal{J}(x,t) := -i\frac{N}{2}\int_{\Omega^{N-1}} \overline{\Psi(x, x_2, \ldots, x_N, t)}\nabla\Psi(x, x_2, \ldots, x_N, t)$$
$$- \Psi(x, x_2, \ldots, x_N, t)\overline{\nabla\Psi(x, x_2, \ldots, x_N, t)}\mathrm{d}x_2 \cdots \mathrm{d}x_N. \tag{2.10}$$

Lemma 2.7 (Continuity equation)**.** *The continuity equation describes the time evolution of the probability current and is given by*

$$\frac{\partial\rho(x,t)}{\partial t} = -\nabla \cdot \mathcal{J}(x,t).$$

Proof. From the definition (2.4) of the density, we have

$$\frac{\partial\rho(x,t)}{\partial t} = N\int_{\Omega^{N-1}} \Psi\frac{\partial\overline{\Psi}}{\partial t} + \overline{\Psi}\frac{\partial\Psi}{\partial t}\mathrm{d}x_2 \cdots \mathrm{d}x_N.$$

Using the Schrödinger equation $\frac{\partial\Psi}{\partial t} = -iH\Psi$ and its conjugate, we find

$$\frac{\partial\rho(x,t)}{\partial t} = iN\int_{\Omega^{N-1}} \Psi H\overline{\Psi} - \overline{\Psi}H\Psi\mathrm{d}x_2 \cdots \mathrm{d}x_N$$

$$= -\frac{iN}{2}\int_{\Omega^{N-1}} \Psi\sum_{j=1}^{N}\nabla_j^2\overline{\Psi} - \overline{\Psi}\sum_{j=1}^{N}\nabla_j^2\Psi\mathrm{d}x_2 \cdots \mathrm{d}x_N$$

$$= -\frac{iN}{2}\int_{\Omega^{N-1}} \sum_{j=1}^{N}\nabla_j \cdot \left(\Psi\nabla_j\overline{\Psi} - \overline{\Psi}\nabla_j\Psi\right)\mathrm{d}x_2 \cdots \mathrm{d}x_N.$$

Now, we can use the divergence theorem to turn every integral of a divergence into a surface integral. Due to the boundary conditions, Assumption (d), these integrals vanish. The only term left is for $j = 1$. Since the integration is not over x_1, we can exchange divergence and integration to obtain

$$\frac{\partial\rho(x,t)}{\partial t} = -\nabla \cdot \frac{iN}{2}\int_{\Omega^{N-1}} \Psi\nabla\overline{\Psi} - \overline{\Psi}\nabla\Psi\mathrm{d}x_2 \cdots \mathrm{d}x_N = -\nabla \cdot \mathcal{J}(x,t).$$

$\square$

The following lemma discusses an equation that is central for the proof of the van Leeuwen's theorem as it provides a relation between the density and the external potential.

Lemma 2.8. *Consider the Hamiltonian $H = \hat{T} + W + V_{ext}$ with $\hat{T} = -\frac{1}{2}\sum_j \nabla_j^2$, $W = \sum_{i,j=1}^{N}\frac{1}{2|x_i - x_j|}$, and $V_{ext} \in C(0,T;C^2(\Omega))$, then the following relation between the density ρ and the potential V_{ext} holds*

$$\frac{\partial^2}{\partial t^2}\rho(x,t) = \nabla \cdot (\rho(x,t)\nabla V_{ext}(x,t)) + q(x,t), \tag{2.11}$$

where q is defined as

$$q(x,t) = \sum_{\alpha,\beta} \frac{\partial^2}{\partial x_\beta \partial x_\alpha} T_{\beta\alpha}(x,t) + \sum_\alpha \frac{\partial}{\partial x_\alpha} W_\alpha(x,t), \qquad (2.12)$$

x_α, $\alpha = 1,2,3$ *denotes the components of* $x \in \mathbb{R}^3$, *the momentum-stress tensor* $T_{\alpha\beta}$ *is given by*

$$T_{\beta\alpha}(x,t) := \frac{N}{2} \int_{\Omega^{N-1}} \frac{\partial \overline{\Psi}}{\partial x_\beta} \frac{\partial \Psi}{\partial x_\alpha} + \frac{\partial \Psi}{\partial x_\beta} \frac{\partial \overline{\Psi}}{\partial x_\alpha} - \frac{1}{2} \left(\frac{\partial^2}{\partial x_\beta \partial x_\alpha} (\overline{\Psi}\Psi) \right) dx_2 \cdots dx_N,$$

and the Coulomb term W_α *is defined as follows*

$$W_\alpha(x,t) := N \int_{\Omega^{N-1}} |\Psi|^2 \frac{\partial}{\partial x_\alpha} \sum_{j=2}^{N} \frac{1}{|x - x_j|} dx_2 \cdots dx_N.$$

Proof. We denote the components of x and $\mathcal{J}$ with x_α and $\mathcal{J}_\alpha$, $\alpha = 1, 2, 3$ and consider the time evolution of the component α of the probability current. We use the SE to determine the time derivatives as follows.

$$\frac{\partial}{\partial t}\mathcal{J}_\alpha = -\frac{iN}{2} \int_{\Omega^{N-1}} \frac{\partial \overline{\Psi}}{\partial t} \frac{\partial \Psi}{\partial x_\alpha} + \overline{\Psi} \frac{\partial}{\partial t} \frac{\partial}{\partial x_\alpha} \Psi - \frac{\partial \Psi}{\partial t} \frac{\partial \overline{\Psi}}{\partial x_\alpha} - \Psi \frac{\partial}{\partial t} \frac{\partial}{\partial x_\alpha} \overline{\Psi} dx_2 \cdots dx_N$$

$$= \frac{N}{2} \int_{\Omega^{N-1}} (H\overline{\Psi}) \frac{\partial \Psi}{\partial x_\alpha} - \overline{\Psi} \frac{\partial}{\partial x_\alpha} (H\Psi) + (H\Psi) \frac{\partial \overline{\Psi}}{\partial x_\alpha} - \Psi \frac{\partial}{\partial x_\alpha} (H\overline{\Psi}) dx_2 \cdots dx_N$$

$$=: \frac{\partial}{\partial t}\mathcal{J}_\alpha^T + \frac{\partial}{\partial t}\mathcal{J}_\alpha^V + \frac{\partial}{\partial t}\mathcal{J}_\alpha^W,$$

where we used Assumption (b). Now, we look at each term of the Hamiltonian $\hat{T}$, V_{ext}, and W separately. We start with the kinetic energy $\hat{T}$,

$$\left(\sum_{j=1}^{N} \nabla_j^2 \overline{\Psi} \right) \frac{\partial}{\partial x_\alpha} \Psi = \sum_{j=1}^{N} \nabla_j \cdot \left((\nabla_j \overline{\Psi}) \frac{\partial \Psi}{\partial x_\alpha} \right) - \nabla_j \overline{\Psi} \cdot \nabla_j \frac{\partial \Psi}{\partial x_\alpha},$$

$$-\overline{\Psi} \sum_{j=1}^{N} \nabla_j^2 \frac{\partial \Psi}{\partial x_\alpha} = -\sum_{j=1}^{N} \nabla_j \cdot \left(\overline{\Psi} \nabla_j \frac{\partial \Psi}{\partial x_\alpha} \right) + \nabla_j \overline{\Psi} \cdot \nabla_j \frac{\partial \Psi}{\partial x_\alpha}.$$

Adding the two lines cancels the last term in each line and only the divergence remains.

Thus, similar to the proof of Lemma 2.7, we have sums of integrals containing only divergences. By the divergence theorem all the integrals for $j = 2, \ldots, N$ are zero due to the boundary condition Assumption (d). What remains is the divergence for $j = 1$.

$$\frac{\partial}{\partial t}\mathcal{J}_\alpha^T = \frac{N}{2} \nabla \cdot \frac{1}{2} \int_{\Omega^{N-1}} (\nabla \overline{\Psi}) \frac{\partial \Psi}{\partial x_\alpha} - \overline{\Psi} \nabla \frac{\partial \Psi}{\partial x_\alpha} + (\nabla \Psi) \frac{\partial \overline{\Psi}}{\partial x_\alpha} - \Psi \nabla \frac{\partial \overline{\Psi}}{\partial x_\alpha} dx_2 \cdots dx_N.$$

The β-th summand of the divergence is given by

$$\frac{\partial}{\partial x_\beta} \frac{1}{2} \left(\frac{\partial \overline{\Psi}}{\partial x_\beta} \frac{\partial \Psi}{\partial x_\alpha} - \overline{\Psi} \frac{\partial^2 \Psi}{\partial x_\beta \partial x_\alpha} + \frac{\partial \Psi}{\partial x_\beta} \frac{\partial \overline{\Psi}}{\partial x_\alpha} - \Psi \frac{\partial^2 \overline{\Psi}}{\partial x_\beta \partial x_\alpha} \right)$$

$$= \frac{\partial}{\partial x_\beta} \left(\frac{\partial \overline{\Psi}}{\partial x_\beta} \frac{\partial \Psi}{\partial x_\alpha} + \frac{\partial \Psi}{\partial x_\beta} \frac{\partial \overline{\Psi}}{\partial x_\alpha} - \frac{1}{2} \left(\overline{\Psi} \frac{\partial^2 \Psi}{\partial x_\beta \partial x_\alpha} + \frac{\partial \overline{\Psi}}{\partial x_\beta} \frac{\partial \Psi}{\partial x_\alpha} + \frac{\partial \Psi}{\partial x_\beta} \frac{\partial \overline{\Psi}}{\partial x_\alpha} + \Psi \frac{\partial^2 \overline{\Psi}}{\partial x_\beta \partial x_\alpha} \right) \right)$$

$$= \frac{\partial}{\partial x_\beta} \left(\frac{\partial \overline{\Psi}}{\partial x_\beta} \frac{\partial \Psi}{\partial x_\alpha} + \frac{\partial \Psi}{\partial x_\beta} \frac{\partial \overline{\Psi}}{\partial x_\alpha} - \frac{1}{2} \left(\frac{\partial^2}{\partial x_\beta \partial x_\alpha} (\overline{\Psi}\Psi) \right) \right).$$

Using the definition of the momentum-stress tensor, we find

$$\frac{\partial}{\partial t}\mathcal{J}_\alpha^T = \sum_{\beta=1}^{N}\frac{\partial T_{\alpha\beta}}{\partial x_\beta}. \tag{2.13}$$

Next, we consider the terms involving the potential V_{ext},

$$
\begin{aligned}
\hat{V} &:= V_{ext}\overline{\Psi}\frac{\partial\Psi}{\partial x_\alpha} - \overline{\Psi}\frac{\partial}{\partial x_\alpha}(V_{ext}\Psi) + V_{ext}\Psi\frac{\partial\overline{\Psi}}{\partial x_\alpha} - \Psi\frac{\partial}{\partial x_\alpha}(V_{ext}\overline{\Psi}) \\
&= V_{ext}\overline{\Psi}\frac{\partial\Psi}{\partial x_\alpha} - \overline{\Psi}\frac{\partial V_{ext}}{\partial x_\alpha}\Psi - \overline{\Psi}\frac{\partial\Psi}{\partial x_\alpha}V_{ext} + V_{ext}\Psi\frac{\partial\overline{\Psi}}{\partial x_\alpha} - \Psi\frac{\partial V_{ext}}{\partial x_\alpha}\overline{\Psi} - \Psi\frac{\partial\overline{\Psi}}{\partial x_\alpha}V_{ext} \\
&= -2|\Psi|^2\frac{\partial V_{ext}}{\partial x_\alpha}.
\end{aligned}
$$

Since the integration is only over $x_2,\dots,x_N$ and V_{ext} is assumed to depend only on x, we can push $\frac{\partial V_{ext}}{\partial x_\alpha}$ out of the integral and obtain

$$\frac{\partial}{\partial t}\mathcal{J}_\alpha^V = \frac{N}{2}\int_{\Omega^{N-1}}\hat{V}\mathrm{d}x_2\cdots\mathrm{d}x_N = -\rho(x)\frac{\partial V_{ext}}{\partial x_\alpha}.$$

Finally, we study the electron-electron interaction W which is a two-particle interaction in contrast to V_{ext}. As in the case of V_{ext}, we obtain

$$
\begin{aligned}
\frac{\partial}{\partial t}\mathcal{J}_\alpha^W &= -\frac{N}{2}\int|\Psi|^2\frac{\partial}{\partial x_\alpha}\sum_{i\neq j}\frac{1}{|x_i-x_j|}\mathrm{d}x_2\cdots\mathrm{d}x_N \\
&= -N\int|\Psi|^2\frac{\partial}{\partial x_\alpha}\sum_{j=2}^{N}\frac{1}{|x_1-x_j|}\mathrm{d}x_2\cdots\mathrm{d}x_N =: -W_\alpha,
\end{aligned}
$$

where we used the fact, that x_α is one component of x_1. In contrast to the case of V_{ext}, we cannot simplify this term further as now the derivative of the potential depends not only on x_1 but also on $x_2,\dots x_N$.

Summing the three terms $\frac{\partial}{\partial t}\mathcal{J}_\alpha^T$, $\frac{\partial}{\partial t}\mathcal{J}_\alpha^V$, and $\frac{\partial}{\partial t}\mathcal{J}_\alpha^W$, we obtain

$$\frac{\partial}{\partial t}\mathcal{J}_\alpha(x,t) = -\rho(x,t)\frac{\partial V_{ext}(x,t)}{\partial x_\alpha} - \sum_\beta\frac{\partial}{\partial x_\beta}T_{\beta\alpha} - W_\alpha.$$

Taking the divergence and using continuity (Lemma 2.7) one finds

$$\frac{\partial^2}{\partial t^2}\rho(x,t) = \nabla\cdot(\rho(x,t)\nabla V_{ext}(x,t)) + q(x,t)$$

$$\text{with } q(x,t) = \sum_{\alpha,\beta}\frac{\partial^2}{\partial x_\beta\partial x_\alpha}T_{\beta\alpha}(x,t) + \sum_\alpha\frac{\partial}{\partial x_\alpha}W_\alpha(x,t). \qquad \square$$

With these preparations, we can present a constructive proof of existence and uniqueness of a KS potential, see also [RPvL15, Pen16]. The original Runge-Gross theorem states that the map $V\mapsto\rho$ is injective. The following theorem is more general as the second two-particle interaction W' can be chosen different from W and not only uniqueness but also existence of the single-particle potential is shown. The latter fact makes the proof more involved. However, as it is possible to choose $W' = 0$ – which is the KS case – the theorem also shows the existence of a KS potential. Further, we want to

remark that no restriction is given on W, apart from the fact that the resulting PDE has to have a sufficiently regular solution.

In [RPvL15, Pen16], a proof of the so-called van Leeuwen's theorem was proposed, however the convergence proof of the Taylor series and the precise formulation of the theorem cannot be found in these references. It was our care to write the precise statements by inspecting the steps of their proof and all related conditions. Using their calculations and different assumptions, we completed the proof of the statement.

The constructive proof of the following theorem is based on the idea that, for a given density ρ, all time derivatives of the external potential at $t = t_0$, $V_{ext}'^{(k)}$, can be obtained from (2.11), where we introduce the notation

$$f^{(k)}(x) = \left. \frac{\partial^k}{\partial t^k} f(x,t) \right|_{t=t_0}.$$

With these derivatives, the time dependence of V_{ext}' can be determined as its Taylor series at t_0, requiring that the density is an analytic function of t. To show convergence of the resulting Taylor series, we use the following theorem from standard calculus.

Lemma 2.9. *Let B denote any Banach space, e.g. $H^2(\Omega)$. Given a sequence $(M_k)_k$ with $0 < M_k \in \mathbb{R}$ such that $\sum_{k=1}^{\infty} M_k$ converges and a sequence $(v_k)_k$ with $v_k \in B$ such that $\|v_k\|_B \le M_k$ for all $k \in \mathbb{N}$. Then the series $\sum_{k=1}^{\infty} v_k$ converges (absolutely) in B.*

Assumption (TDDFT 2). We assume the following:

(e) For $0 \le \rho(t_0) \in C^1(\overline{\Omega})$, $\rho(x,t_0) \ne 0$ for $x \in \Omega$, $\rho(t_0)^{-2} \in L^1(\Omega)$, and $f \in L^2(\Omega)$, the degenerate elliptic PDE

$$\langle \rho(t_0)\nabla v, \, \nabla\varphi \rangle_{L^2} = \langle f, \, \varphi \rangle_{L^2} \quad \forall \varphi \in H_0^1(\Omega), \tag{2.14}$$

possess a unique weak solution $v \in H_0^1(\Omega) \cap H^2(\Omega)$. Furthermore, we assume the bound

$$\|v\|_{H^2} \le K\|f\|_{L^2} \tag{2.15}$$

where the constant K is allowed to depend on $\rho(t_0)$. See [PR11, Pen16] for similar statements.

(f) ρ is a polynomial in time of degree n_ρ,

$$\rho(x,t) = \sum_{l=0}^{n_\rho} \frac{r_l(x)}{l!}(t-t_0)^l, \tag{2.16}$$

and $\rho^{(k)} \in C^1(\overline{\Omega})$; this assumptions replaces Assumption (c);

(g) There exists a $\tilde{r} < \frac{1}{Kn_\rho}$, such that $\|\rho^{(k)}\|_{C(\overline{\Omega})} + \|\nabla\rho^{(k)}\|_{C(\overline{\Omega})} \le \tilde{r}$ for $k \ge 1$ (no restriction for ρ itself);

(h) let W' be such that $q'^{(k)} \in L^2(\Omega)$ for all $k \ge 0$;

(i) q' is a polynomial in time of degree n_q,

$$q'(x,t) = \sum_{l=0}^{n_q} \frac{q_l(x)}{l!}(t-t_0)^l. \tag{2.17}$$

Next, consider the set U of potentials V that are analytic in time for $t \in [t_0, T]$, $T > t_0$ with $V^{(k)} \in H_0^1(\Omega) \cap H^2(\Omega) \ \forall k \in \mathbb{N}_0$. We remark that by prescribing the zero boundary conditions, we remove the ambiguity from the external potential that is commonly found in the TDDFT literature where potentials are considered "up to a purely time-dependent function" that only changes the global phase of the wavefunction.

Theorem 2.10 (van Leeuwen's Theorem). *Consider a Hamiltonian $H = \hat{T} + W + V_{ext}$ and an initial state $\Psi^0 \in C^4(\Omega)$ whose resulting density ρ satisfies Assumption (c) with $\rho(x, t) \neq 0 \ \forall x \in \Omega$, $\forall t \in [0, T]$ and $\rho^{-2}(\cdot, t) \in L^1(\Omega)$.*

Then, given another two-particle interaction potential W' and a second initial state $\Psi^{0\prime}$ given such that $\rho(\Psi^{0\prime}) = \rho(\Psi^0)$. Furthermore, Assumptions (a)–(i) hold.

Then there exists a unique single-particle potential V_{ext}' in U such that the densities $\rho(x, t)$ and $\rho'(x, t)$ of the two systems coincide.

Proof. We have two systems, described by the Hamiltonians $H(t)$ and $H'(t)$, where $H(t) = \hat{T} + W + V_{ext}(t)$ is given and $H'(t) = \hat{T} + W' + V_{ext}'$ is prescribed up to $V_{ext}'(t)$. Furthermore, the initial states Ψ^0 and $\Psi^{0\prime}$ are given. The aim is to find a single-particle potential V_{ext}' such that the densities $\rho(r, t)$, $\rho'(r, t)$, which are analytic in $t = t_0$, are the same for all times, $\rho(r, t) = \rho'(r, t)$. This is achieved by constructing V_{ext}' from its Taylor series around $t = t_0$ in such a way that $\rho^{(k)} = \rho'^{(k)}$.

We recall the definition of q from (2.12). Requiring $\rho'(t_0) = \rho(t_0)$ and $\rho'^{(2)} = \rho^{(2)}$, we obtain from (2.11)

$$\rho^{(2)}(x) = q'(x, t_0) + \nabla \cdot (\rho(x, t_0) \nabla V_{ext}'(x, t_0)). \tag{2.18}$$

$q'(x, t_0)$ depends only on the initial condition $\Psi^{0\prime}$ which is known and the new interaction W' respectively their spacial derivatives. As we assume that $\rho^{-2} \in L^1(\Omega)$ and $\rho \neq 0$ in the interior of Ω, we can apply the result of [PR11], stating that (2.18) has a unique solution $V_{ext}'(\cdot, t_0) \in H_0^1(\Omega)$. By Assumption (e), the solution is even in $H_0^1(\Omega) \cap H^2(\Omega)$.

Now, we determine the higher derivatives of V_{ext}' recursively. To this end, we take the k-th time derivative of (2.11) in Lemma 2.8 and obtain

$$\rho'^{(k+2)}(x) = q'^{(k)}(x) + \sum_{l=0}^{k} \binom{k}{l} \nabla \cdot \left(\rho'^{(k-l)}(x) \nabla V_{ext}'^{(l)}(x) \right).$$

Demanding $\rho'^{(k+2)} = \rho^{(k+2)}$, this gives the PDE

$$\nabla \cdot \left(\rho^{(0)}(x) \nabla V_{ext}'^{(k)}(x) \right) = \rho^{(k+2)}(x) - q'^{(k)}(x) - \sum_{l=0}^{k-1} \binom{k}{l} \nabla \cdot \left(\rho^{(k-l)}(x) \nabla V_{ext}'^{(l)}(x) \right),$$

$$V_{ext}'^{(k)}(x) = 0 \text{ for } x \in \partial \Omega. \tag{2.19}$$

In step k, $V_{ext}'^{(l)}$ is already known for $l = 0, \ldots, k-1$, the power series of ρ is also known. However, it is not possible to determine $q'^{(k)}$ by directly differentiating (2.11) because Ψ' is only known at t_0 and hence its derivatives cannot be calculated. Furthermore, from the known $\rho'^{(k)}$ it is also not possible to establish $\Psi'^{(k)}$. Therefore, we use the $(k-1)$-th derivative of the SE to determine $\Psi'^{(k)}$ and this in turn gives us $q^{(k)}$,

$$\Psi'^{(k)} = (-iH'(t_0)\Psi)^{(k-1)}.$$

Therefore, all terms on the right hand side of (2.19) are given and in $L^2(\Omega)$ for all $t \in [t_0, T]$; with the boundary condition $V_{ext}'^{(k)} = 0$ on $\partial\Omega$, by Assumption (e), the PDE has a unique solution $V_{ext}'^{(k)}(t) \in H_0^1(\Omega) \cap H^2(\Omega)$.

Now, we have $V_{ext}'^{(k)}$ for all $k \in \mathbb{N}$, and we define the potential to be the power series

$$V_{ext}'(x,t) := \sum_{k=0}^{\infty} \frac{V_{ext}'^{(k)}(x)}{k!}(t - t_0)^k. \tag{2.20}$$

To conclude this proof, we need to show convergence of this power series. This is done by bounding $V_{ext}'^{(k)}$ as a solution of (2.19) by the right hand side and then use Lemma 2.9. The weak form of (2.19) is given by

$$
\begin{aligned}
&- \left\langle \rho^{(0)}(x) \nabla V_{ext}'^{(k)}(x), \, \nabla\varphi \right\rangle_{L^2} \\
&= \left\langle \rho^{(k+2)}(x) - q'^{(k)}(x) - \sum_{l=0}^{k-1} \binom{k}{l} \nabla \cdot \left(\rho^{(k-l)}(x) \nabla V_{ext}'^{(l)}(x) \right), \, \varphi \right\rangle_{L^2}, \quad \forall \varphi \in H_0^1(\Omega).
\end{aligned}
\tag{2.21}
$$

To (2.21), we apply Assumption (e) and obtain the existence of a unique solution $V_{ext}'^{(k)} \in H_0^1(\Omega) \cap H^2(\Omega)$ and the following bound:

$$\|V_{ext}'^{(k)}\|_{H^2} \leq K \left\| \rho^{(k+2)}(x) - q'^{(k)}(x) - \sum_{l=0}^{k-1} \binom{k}{l} \nabla \cdot \left(\rho^{(k-l)}(x) \nabla V_{ext}'^{(l)}(x) \right) \right\|_{L^2}. \tag{2.22}$$

As ρ and q are assumed to be polynomial in time, we can use the series expansions (2.16) and (2.17) to obtain $\|\rho^{(k+2)} + q'^{(k)}\|_{L^2} \leq \|r_{k+2}\|_{L^2} + \|q_k\|_{L^2}$ and

$$\left\| \nabla \cdot \left(\rho^{(k-l)} \nabla V_{ext}'^{(l)} \right) \right\|_{L^2} \leq \left(\|r_{k-l}\|_C + \|\nabla r_{k-l}\|_C \right) \|V_{ext}'^{(l)}\|_{H^2}. \tag{2.23}$$

Altogether, we obtain the following bounds for $\|V_{ext}'^{(k)}\|_{H^2}$, $k \geq 1$,

$$\left\| V_{ext}'^{(k)} \right\|_{H^2} \leq K \left(\|r_{k+2}\|_{L^2} + \|q_k\|_{L^2} + \sum_{l=0}^{k-1} \binom{k}{l} \left(\|r_{k-l}\|_C + \|\nabla r_{k-l}\|_C \right) \left\| V_{ext}'^{(l)} \right\|_{H^2} \right). \tag{2.24}$$

To apply Lemma 2.9 for $B = H_0^1(\Omega) \cap H^2(\Omega)$, we need to bound the right hand side in such a way that the sum over the bounds divided by $k!$ still converges, i.e.

$$\sum_{k=0}^{\infty} \frac{\|r_{k+2}\|_{L^2} + \|q_k\|_{L^2} + \sum_{l=0}^{k-1} \binom{k}{l} \left(\|r_{k-l}\|_C + \|\nabla r_{k-l}\|_C \right) \|V_{ext}'^{(l)}\|_{H^2}}{k!} < \infty. \tag{2.25}$$

By Assumptions (f) and (i) the sum $\sum_{k=0}^{\infty} \frac{\|r_{k+2}\|_{L^2} + \|q_k\|_{L^2}}{k!} = \sum_{k=0}^{n} \frac{\|r_{k+2}\|_{L^2} + \|q_k\|_{L^2}}{k!}$ is a finite sum and hence converges. Therefore, we need to consider the double sum in the following. Recall that ρ and q have series expansions of finitely many terms n_ρ respectively n_q. Define $n := \max\{n_\rho, n_q\}$. Further recall that $\|r_k\|_C + \|\nabla r_k\|_C \leq \tilde{r} < \frac{1}{n_\rho}$ by Assumption (g). As there are only finitely many $q_k \neq 0$, there exists a $\tilde{q}$ such that $\|q_k\|_{L^2} \leq \tilde{q}$.

To get a better understanding of the double sum, we first consider the case $n_\rho = 1$ which means that the sum in (2.22) only has the term $l = k - 1$. Using (2.23) in (2.22), by induction, we have

$$\left\| V_{ext}'^{(k)} \right\|_{H^2} \leq K \left(\|r_{k+2}\|_{L^2} + \|q_k\|_{L^2} + k\tilde{r} \left\| V_{ext}'^{(k-1)} \right\|_{H^2} \right) \tag{2.26}$$

$$\leq K \left(\|r_{k+2}\|_{L^2} + \|q_k\|_{L^2} + k\tilde{r}K \left(\|r_{k+1}\|_{L^2} + \|q_{k-1}\|_{L^2} + (k-1)\tilde{r} \left\| V_{ext}'^{(k-2)} \right\|_{H^2} \right) \right) \tag{2.27}$$

$$\leq \sum_{l=0}^{k} \left(\|r_{k-l+2}\|_{L^2} + \|q_{k-l}\|_{L^2} \right) \frac{k!}{(k-l)!} \tilde{r}^l K^{l+1} + k!\tilde{r}^k K^k \left\| V_{ext}'^{(0)} \right\|_{H^2}. \tag{2.28}$$

Using Assumption (g) and the geometric series, the series

$$\sum_{k=1}^{\infty} \frac{1}{k!} k! \tilde{r}^k K^k \| V_{ext}'^{(0)} \|_{H^2} \tag{2.29}$$

converges. For the first term and sufficiently large k, we have

$$\sum_{l=0}^{k} (\|r_{k-l+2}\|_{L^2} + \|q_{k-l}\|_{L^2}) \frac{k!}{(k-l)!} \tilde{r}^l K^{l+1} = \sum_{j=0}^{k} (\|r_{j+2}\|_{L^2} + \|q_j\|_{L^2}) \frac{k!}{j!} \tilde{r}^{k-j} K^{k-j+1}$$

$$\leq (\tilde{r} + \tilde{q}) \sum_{j=0}^{n} \frac{k!}{j!} \tilde{r}^{k-j} K^{k-j+1} \leq (\tilde{r} + \tilde{q}) \frac{nk!}{\tilde{r}^n} \tilde{r}^k \max\{1, K^{k+1}\},$$

where we used the fact that both $r_k = 0$ and $q_k = 0$ for $k > n$. Using this for the double sum in (2.25), we have

$$\sum_{k=1}^{\infty} \frac{1}{k!} \sum_{l=0}^{k} (\|r_{k-l+2}\|_{L^2} + \|q_{k-l}\|_{L^2}) \frac{k!}{(k-l)!} \tilde{r}^l K^{k-l+1}$$

$$\leq (\tilde{r} + \tilde{q}) \frac{n}{\tilde{r}^n} \sum_{k=1}^{\infty} \tilde{r}^k \max\{1, K^{k+1}\} < \infty, \tag{2.30}$$

where the convergence is given by the geometric series and Assumption (g).

For $n_\rho > 1$, the estimate goes similar although there are more terms appearing in (2.26)–(2.28). However, $\tilde{r}$ appears with the same exponent and the terms also have a prefactor of $k!$ and can therefore be treated in the same way. In each iteration the number of terms is multiplying at most with n_ρ, so, in total, we get an additional factor of n_ρ^k in (2.28). Due to the bound $\tilde{r} < \frac{1}{n_\rho K}$ the series

$$\sum_{k=1}^{\infty} \tilde{r}^k n_\rho^k K^k \| V_{ext}'^{(0)} \|_{H^2} < \infty \tag{2.31}$$

$$\text{and } (\tilde{r} + \tilde{q}) \frac{n}{\tilde{r}^n} \sum_{k=1}^{\infty} \tilde{r}^k n_\rho^k \max\{1, K^{k+1}\} < \infty \tag{2.32}$$

replacing (2.29) and (2.30), respectively, still converge and we have

$$\left\| V_{ext}'^{(k)} \right\|_{H^2} \leq K \left(\|r_{k+2}\|_{L^2} + \|q_k\|_{L^2} + \sum_{j=1}^{n_\rho} \binom{k}{k-j} \tilde{r} \left\| V_{ext}'^{(k-j)} \right\|_{H^2} \right) \leq M_k$$

$$\text{with } M_k := (\tilde{r} + \tilde{q}) \frac{n}{\tilde{r}^n} k! \tilde{r}^k n_\rho^k \max\{1, K^{k+1}\} + k!\tilde{r}^k n_\rho^k K^k \| V_{ext}'^{(0)} \|_{H^2}$$

and the series $\sum_{k=1}^{\infty} \frac{M_k}{k!} < \infty$ converges.

Altogether, we have shown the convergence of the power series of V'_{ext}, and, therefore, the desired V'_{ext} exists in U. As all functions in U are analytic and uniquely determined by their Taylor coefficients and these are uniquely determined by the solutions of (2.21), V'_{ext} is also unique in U. This completes the proof of the van Leeuwen's theorem. $\qquad\square$

We want to conclude this theorem with a discussion on several issues in the proof. The argument was originally presented by R. van Leeuwen in [vL99]. While we still use his creative main idea, there were two issues in his calculations.

First, in [vL99], existence and uniqueness of the solutions of (2.18) and (2.19) were not discussed as the PDEs was not recognized to be degenerate which prohibits the use of standard theory of elliptic PDEs. For (2.18), existence and uniqueness of a solution in $H_0^1(\Omega)$ was proved in [PR11] using weighted Sobolev spaces; see also [GU09] and references therein for a more general analysis of degenerate elliptic PDEs and weighted Sobolev spaces. However, for the right hand side of (2.21) to be well-defined in $L^2(\Omega)$, we need to assume higher regularity of the solution in Assumption (e).

The second question is the convergence of the Taylor series. Contrary to the classical Runge-Gross theorem, see [RG84], the van Leeuwen's theorem aims at showing the existence of V'_{ext}. Hence, convergence of the series has to be proved directly and properties of the presumed V'_{ext} cannot be used for this.

A simple solution to this issue is the assumption made in [BCS17] that there exists an $M > 0$ such that $\max_{k\in\mathbb{N}} \|V'^{(k)}_{ext}\|_{C(\overline{\Omega})} < M$. We have proved the same result using different assumptions: First, no assumption is made beforehand on the values of $V'^{(k)}_{ext}$ that are to be determined during the proof and are not known in advance. Second, we did not require a uniform global bound on the solution of (2.21) but allow for the bound to depend on the right hand side and use the growth Assumption (g).

Furthermore, we want to remark that the uniqueness of the potential V'_{ext} is a delicate issue. We refer to [Pen16, p. 153] and reference therein for a discussion on the issue of uniqueness for analytic external potentials with finite convergence radii. In the recent work of Fournais et al. [FLLS16], the interplay of the regularity of initial condition and the regularity requirements of the external potential in the uniqueness proof (i.e. the original Runge-Gross theorem) is discussed in detail; however the existence proof (van Leeuwen's theorem) is not discussed therein.

To complete our discussion on the general TDKS framework, we remark that there are two (open) issues that are specific to the time-dependent KS approach. First, as illustrated in the discussion above, the KS potential depends on the initial states of both the interacting N-particle wavefunction Ψ^0 and the KS Slater determinant $\Psi^{0\prime}$. This is a serious complication compared to ground-state DFT, as for every initial state a new KS potential is needed. Our second remark is that the exact Kohn-Sham potential at time t actually depends on the density at all times and not just on the density at time t. Specifically, since the Coulomb interaction is instantaneous and the Hartree potential has no memory, the history dependence is contained in the exchange-correlation potential. The memory effect is due to the fact that most of the degrees of freedom of the wavefunction are traced out when calculating the density. This is similar to open quantum systems, where a memory effect is introduced by tracing out the bath degrees of freedom.

In applications, the prevalent approach is to use an adiabatic ansatz. This means that the KS potential at time t is assumed to depend only on the density at time t and the history is completely neglected. The initial state has still to be chosen, e.g., by using the ground state. For most applications, this is a sufficiently good approximation. For

a detailed discussion on the initial state and history dependence of TDKS we refer to [MUN$^+$06, MB01] and the recent work of [FNRM16, RNvL13].

Now, we present a specific TDKS model, which is also our focus in the development of a control strategy in the TDDFT framework. This model is given by the following system of coupled single-particle SE:

$$
\begin{aligned}
i\frac{\partial \psi_j}{\partial t}(x,t) &= \left(-\frac{1}{2}\nabla^2 + V_{ext}(x,t) + V_{Hxc}(x, \rho(x,t)) \right) \psi_j(x,t), \\
\psi_j(x,0) &= \psi_j^0(x), \qquad j = 1\ldots, N, \\
\rho(x,t) &= \sum_{j=1}^{N} |\psi_j(x,t)|^2,
\end{aligned}
\tag{2.33}
$$

where V_{Hxc} is given by (2.8). As already mentioned, the choice of the initial wavefunctions ψ_j^0 may depend on the specific application and on the a priori knowledge on the system. A reasonable choice is to take ψ_j^0, $j = 1, \ldots, N$, equal to the N orthogonal eigenfunctions of the eigenproblem (2.6) with the lowest energy. Choosing the initial wavefunctions ψ_j^0 orthogonal in $L^2(\Omega; \mathbb{C})$, it is guaranteed that they remain so during the evolution governed by (2.33). Therefore, it is common practice to consider ρ given as in (2.33); see [MUN$^+$06].

In the following chapter, we will analyse the TDKS equations (2.33) in the adiabatic framework and prove existence and uniqueness of their solution under appropriate assumptions on the domain and on V_{ext} and V_{Hxc}. In Chapter 4, these results will be used to study optimal control problems with TDKS constraints where the control mechanism is included in V_{ext}.

For further discussion on various extensions of TDDFT that include vector potentials, spin, and relativistic effects see, e.g., [Vig04, ED11, MUN$^+$06].

2.4. Conclusion

In this chapter, we discussed the difficulties arising from the exponential complexity of the multi-particle Schrödinger equation. We introduced the density functional theory as an efficient way to overcome this curse of dimensionality and to deal with multi-particle quantum systems. The central idea of DFT is to consider the density ρ instead of the wavefunction Ψ as the main variable. The mappings between potentials, corresponding ground-state wavefunctions and densities were discussed. Restricting ourselves to a subset of suitable quantum problems, an equivalence between the wavefunction and the density description could be proved, thus justifying the DFT model as a general ab-initio method in quantum mechanics.

To take full advantage of the DFT, the Kohn-Sham approach was introduced. In this framework, the interacting multi-particle SE is replaced by a system of non-interacting SEs, the KS equations. To ensure the equivalence of the two descriptions an additional potential is introduced in the KS equations that depends on the density and thus couples the KS equations in a nonlinear way. This potential is chosen such that the two systems result in the same density and thus all observables depending only on the density can be obtained from the non-interacting system.

Subsequently, we extended DFT to the time-dependent DFT framework of Runge and Gross. A rigorous foundation was provided by presenting the van Leeuwen's theorem. This allows to study excited states and the time evolution of non-stationary problems.

Finally, using adiabatic local density approximation (LDA), we introduced the time-dependent Kohn-Sham equations (2.33). These are the basic equations to describe the

quantum models investigated in this thesis. They allow the solution of general time-dependent problems of interacting quantum particles in a very efficient way as their complexity scales only linearly with the number of particles involved. Although already widely used in applied quantum physics and chemistry the theoretical understanding of the TDKS equations is still an active research topic and the focus of the next chapter.

3. The time-dependent Kohn-Sham equations

3.1. Introduction

In this chapter, we analyse the time-dependent Kohn-Sham equations that govern the evolution of N single-particle wavefunctions $\Psi = (\psi_1, \ldots, \psi_N)$, $\psi_i = \psi_i(x, t)$, $x \in \mathbb{R}^n$, $t \in \mathbb{R}$. As introduced in (2.33), the TDKS system can be written as follows:

$$i\partial_t \Psi(x, t) = I_N \otimes \left[-\nabla^2 + V_{ext}(x, t; u) + V_{Hxc}(x, t; \Psi) \right] \Psi(x, t), \quad \Psi(x, 0) = \Psi^0(x), \quad (3.1)$$

where ∇^2 is the Laplacian, V_{ext} is an external potential that includes the confining potential, e.g., the surrounding walls or the Coulomb potential of the nuclei of a molecule, and, possibly, a control potential. The time interval is fixed as $[0, T]$. V_{Hxc} denotes the coupling KS potential introduced in Section 2.2.2 and further discussed in Section 3.2.

The purpose of this chapter is to theoretically investigate (3.1), where $u \in H^1(0, T)$ may represent a given control function, and to analyse an adjoint version of (3.1) as it appears in the first-order optimality system of the following optimal control problem:

$$\min_{(\Psi, u) \in (W, H^1(0,T))} J_1(\Psi) + J_2(\Psi(T)) + \nu \|u\|_{H^1(0,T)}^2 \quad \text{s.t. } \Psi \text{ solves (3.1)}, \qquad (3.2)$$

where $\nu > 0$ is a weight parameter, J_1 depends on the solution Ψ at all times in $[0, T]$, while J_2 depends on the wavefunction at the final time $\Psi(T)$ only. The space W is introduced in Definition 3.1 on page 31. The functionals J_1, J_2 are assumed to be lower semicontinuous and Fréchet differentiable with respect to Ψ.

To characterize the solutions to (3.2) using the adjoint method [Trö10, BS12, Bor12, SCB17b], the following adjoint equation is considered:

$$i\frac{\partial \Psi}{\partial t} = I_N \otimes \left(-\nabla^2 + V_{ext}(x, t, u) + V_{Hxc}(\Lambda) \right) \Psi$$
$$+ I_N \otimes \left(V_H(2 \operatorname{Re} (\Psi, \Lambda)_{\mathbb{C}}) + 2\frac{\partial V_{xc}}{\partial \rho}(\Lambda) \operatorname{Re} (\Psi, \Lambda)_{\mathbb{C}} + D_\psi J_1(\Lambda) \right) \Lambda, \qquad (3.3)$$
$$\Psi(T) = -D_\Psi J_2(\Lambda(T)),$$

where we also denote by Ψ the adjoint variable while, in this case, Λ denotes the solution to (3.1).

We remark that (3.3) has a similar structure to (3.1) with an additional inhomogeneity resulting from the Fréchet derivative $D_\psi J_1(\Lambda)\Lambda$ of J_1 with respect to the wavefunction, as well as additional terms resulting from the linearization of the Kohn-Sham potential. On the other hand, V_{Hxc} now depends on Λ and is no longer a function of the unknown variables. The derivative of J_2 gives a terminal condition for (3.3) that evolves backwards in time.

In this chapter, we theoretically analyse (3.1) and (3.3) as two particular instances of a generalized TDKS equation, proving existence and uniqueness of solutions. We remark that the results of [Caz03] cannot be applied to our problem as we consider explicitly time-dependent potentials that are not covered in [Caz03].

To the best of our knowledge, the TDKS problem (3.1) is only addressed in [Jer15] as follows: Assuming that $V_{ext} \in C^1([0,T]; C^1(\overline{\Omega}; \mathbb{R}))$ and $V_{ext} \geq 0$, and a Lipschitz condition on V_{xc} and a continuity assumption on V_{xc}, then the weak form of (3.1) with $\Psi^0 \in H_0^1(\Omega; \mathbb{C}^N)$ has a unique solution in $C([0,T]; H_0^1(\Omega; \mathbb{C}^N)) \cap C^1([0,T]; H^{-1}(\Omega; \mathbb{C}^N))$.

In this reference, the author proves existence and uniqueness of solutions assuming that the Hamiltonian is continuously differentiable in time. We improve these results in such a way to accommodate TDKS optimal control problems. In particular, existence and uniqueness of solutions with similar regularity as in [Jer15] are proved also in the case when the external potential is only H^1 and not C^1. These results are achieved in the Galerkin framework. We remark that by this approach, we address the TDKS equation (3.1) and its adjoint (3.3) in a unified framework. Genuinely new are our analysis of the adjoint equation and the improved regularity results in Section 3.8.

Notice that, taking into account only the Hartree potential but not the exchange-correlation potential, existence of a unique solution of the forward equation (3.1) in $C([0,\infty); H^2(\mathbb{R}^3; \mathbb{C})) \cap C^1([0,\infty); L^2(\mathbb{R}^3; \mathbb{C}))$ is shown in [CL99] using semigroup theory. Notice that, especially for the adjoint problem, it can be difficult to use semigroup theory, as in [Jer15, CL99], due to the explicit time-dependence of the potential; see, e.g., [Eva10, p. 422].

For optimal control application using the TDKS equation as differential constraints (cf. Chapter 4), it is important to have a unique solution of this equation. This is necessary to have a well-posed optimization problem and to guarantee a well-defined control-to-state map. The higher regularity of the forward equation and the unique solvability of the adjoint equation are needed to characterize the optimal control using the Lagrange framework.

This chapter is organized as follows. In Section 3.2, we discuss the KS potential V_{Hxc} and the external potential V_{ext}. Further, we formulate our evolution problem in a weak form that embodies both (3.1) and (3.3). Also in this section, we discuss the initial and boundary conditions, and provide specific assumptions on the potentials and the spatial domain Ω where the KS problem is considered. In Section 3.3, we investigate some properties of the KS potential and discuss continuity of the bilinear form resulting from the weak formulation. In Section 3.4, we use the Galerkin framework to obtain a finite-dimensional approximation of our weak problem. In Section 3.5, we present energy estimates for the finite-dimensional representation and their extension to the infinite-dimensional case. In Sections 3.6 and 3.7, we prove existence and uniqueness of solutions to our weak problem. First, we prove convergence of the Galerkin approximation to the infinite-dimensional solution and then use our results and assumptions on the Lipschitz properties of the potential to prove uniqueness of this solution. In Section 3.8, assuming higher regularity of the data, we prove that the solution to our problem has higher regularity. The Sections 3.6, 3.7, and 3.8 present the main results of this chapter.

3.2. The model description

In this section, we introduce the weak formulation of our evolution problem, define the potentials, and discuss our assumptions. To introduce the weak formulation of the evolution problem, we define the following function spaces. We use $L^2 := L^2(\Omega; \mathbb{C}^N)$ endowed with the inner product $\langle \cdot, \cdot \rangle_{L^2}$ defined as

$$\langle \Psi, \Phi \rangle_{L^2} := \int_\Omega (\Psi, \Phi)_\mathbb{C} \, \mathrm{d}x,$$

and $\| \cdot \|_{L^2}$ denotes the corresponding norm. Further, we denote by $(\cdot, \cdot)_{\mathbb{C}}$ the scalar product for $\mathbb{C}^N$ and $|\cdot|$ is the corresponding norm. The scalar product of the Sobolev space $H^1 := H^1(\Omega; \mathbb{C}^N)$ is given by

$$\langle \Psi, \Phi \rangle_{H^1} := \langle \Psi, \Phi \rangle_{L^2} + \sum_{j=1}^{N} \langle \nabla \psi_j, \nabla \varphi_j \rangle_{L^2},$$

for $\Psi = (\psi_1, \ldots, \psi_N)^T$ and $\Phi = (\varphi_1, \ldots \varphi_N)$, and $\| \cdot \|_{H^1}$ is the corresponding norm.

Definition 3.1. *We define the following spaces of functions of time and space with values in $\mathbb{C}^N$ and their norms:*

$$Y := L^2(0, T; L^2(\Omega; \mathbb{C}^N)), \qquad \|u\|_Y^2 = \int_0^T \|u(t)\|_{L^2}^2 \mathrm{d}t,$$

$$X := L^2(0, T; H_0^1(\Omega; \mathbb{C}^N)), \qquad \|u\|_X^2 = \int_0^T \|u(t)\|_{H^1}^2 \mathrm{d}t,$$

$$X^* = L^2(0, T; H^{-1}(\Omega; \mathbb{C}^N)), \qquad \|u\|_{X^*} = \sup_{v \in X \setminus \{0\}} \frac{|u(v)|}{\|v\|_X},$$

$$W := \{u \in X \text{ such that } u' \in X^*\}, \qquad \|u\|_W^2 = \|u\|_X^2 + \|u'\|_{X^*}^2,$$

$$Z_\star := L^\infty(0, T; H^2(\Omega; \mathbb{C}^N) \cap H_0^1(\Omega; \mathbb{C}^N)), \qquad \|u\|_{Z_\star} = \operatorname*{ess\,sup}_{t \in [0,T]} \|u(t)\|_{H^2(\Omega; \mathbb{C}^N)},$$

$$Z := \{u \in Z_\star \text{ such that } u' \in X^*\}, \qquad \|u\|_Z^2 = \|u\|_{Z_\star}^2 + \|u'\|_{X^*}^2.$$

We prove the existence of a solution to the controlled Kohn-Sham model (3.1) and at the same time to its adjoint (3.3) on a bounded domain $\Omega \subset \mathbb{R}^n$, $n = 3$, for a finite time interval $[0, T]$ and with homogeneous Dirichlet boundary conditions. For this purpose, we denote by $\Psi \in X$ the vector of the wavefunctions corresponding to N particles

$$\Psi := (\psi_1, \ldots, \psi_N)^T, \tag{3.4}$$

and assume that $\psi_j(x, t) = 0$ for $x \in \partial\Omega$ and consider the initial condition $\psi_j(x, 0) = \psi_j^0(x)$ with $\psi_j^0 \in L^2(\Omega; \mathbb{C})$. Moreover, to include a possible inhomogeneity of the PDE, we consider the function $F \in Y$ defined as follows:

$$F := (f_1, \ldots, f_N)^T, \tag{3.5}$$

where $f_j \in L^2(0, T; L^2(\Omega; \mathbb{C}))$.

The wavefunction Ψ gives rise to the density ρ defined as follows:

$$\rho(x, t) := \sum_j |\psi_j(x, t)|^2 = |\Psi(x, t)|^2, \tag{3.6}$$

which is used to characterize the nonlinear potential $V_{Hxc}(x, t; \Psi)$. The dependence of V_{Hxc} on Ψ is always through the density ρ, so we may also write $V_{Hxc}(x, t; \rho)$. In the local density approach (LDA) framework, V_{Hxc} is given by the sum of the Hartree, the exchange, and the correlation potentials, see Section 2.2.2. We have

$$V_{Hxc}(x, t; \Psi) = V_H(x, \Psi(t)) + V_{xc}(\Psi(x, t)) = V_H(x, \Psi(t)) + V_x(\Psi(x, t)) + V_c(\Psi(x, t)),$$

$$V_H = \int_\Omega \frac{\rho(y, t)}{|x - y|} \mathrm{d}y, \quad V_x = V_x(\rho(x, t)), \quad V_c = V_c(\rho(x, t)).$$

$$\tag{3.7}$$

Recall that V_x is often derived from an approximation called the homogeneous electron gas [PY89] and then given by $V_x = c\rho(x,t)^\beta$, where c is a negative constant and $0 < \beta < 1$ depends on the dimension n. For the correlation potential, V_c, only numerical approximations exist. In the course of the years, physicists and quantum chemists have developed a collection of different V_c functions. Similar to Jerome [Jer15], who uses a Lipschitz assumption on $V_x + V_c$, we make some general assumptions on the structure of the potentials rather than using an explicit form for one of the approximations used in applications.

For our analysis of the exchange potential, we slightly change the definition of V_x in a way that does not change its application. Quantum mechanics is not applicable for "very short" distances (i.e. at or below the Planck scale of about 10^{-35}m) and for "very high" mass or energy densities such that relativistic effects start to play a role. Therefore, we can modify the density on these short length scales such that the density is bounded without changing anything that quantum mechanics can resolve. This corresponds to introducing a cut-off of the potential at unphysically large densities, while preserving all required properties in the range of validity of the DFT framework. Remark 3.8 on page 51 further supports this approach as the a-posteriori estimate shows that the density is globally bounded.

We will use this fact for an estimate of an upper bound of the density. Therefore, we estimate the radius of a spherical black hole of the mass of an electron by the Schwarzschild radius $r_s = \frac{2mG}{c^2} \approx 10^{-57}$meters, where m is the mass of the electron, G is the gravitational constant and c is the speed of light in vacuum [Wal09]. Thus, we introduce the cut-off density $R = \frac{m}{V_s} \approx \left(10^{57}\frac{1}{\text{m}}\right)^n m$, where V_s is the volume of the sphere of radius r_s. This allows us to modify the exchange potential in the following way while preserving all properties in the range where DFT is applicable. We have

$$V_x : [0,\infty) \to [0, p(2R)], \qquad V_x(\rho) = \alpha_n \begin{cases} \sqrt[n]{\rho} & \rho \le R \\ p(\rho) & R < \rho < 2R \\ p(2R) & \rho \ge 2R \end{cases}, \qquad (3.8)$$

where

$$p(\rho) = \frac{(n+1)R^{\frac{1}{n}-4}}{4n^2}\rho^4 - \frac{(4n+5)R^{\frac{1}{n}-3}}{3n^2}\rho^3 + \frac{2(n+2)R^{\frac{1}{n}-2}}{n^2}\rho^2 - \frac{4R^{\frac{1}{n}-1}}{n^2}\rho + \frac{(12n^2-11n+17)R^{\frac{1}{n}}}{12n^2}, \qquad (3.9)$$

with $\alpha_2 = -\sqrt{\frac{8}{\pi}}$ and $\alpha_3 = -\sqrt[3]{\frac{3}{\pi}}$ and R sufficiently large; e.g., $R \approx (10^{57}\frac{1}{\text{m}})^3 N$. This potential is twice continuously differentiable and globally bounded.

Lemma 3.2. *The exchange potential term $\Psi \mapsto V_x(\Psi)\Psi$, $\Psi \in W$ is Lipschitz continuous, i.e.*

$$\|V_x(\Psi(t))\Psi(t) - V_x(\Upsilon(t))\Upsilon(t)\|_{L^2(\Omega;\mathbb{C}^N)} \le L_1 \|\Psi(t) - \Upsilon(t)\|_{L^2(\Omega;\mathbb{C}^N)}, \text{ for a.a. } t \in [0,T],$$

$$\|V_x(\Psi)\Psi - V_x(\Upsilon)\Upsilon\|_{X^*} \le c\|V_x(\Psi)\Psi - V_x(\Upsilon)\Upsilon\|_Y \le L_2\|\Psi - \Upsilon\|_X.$$

The Lipschitz continuity holds also in zero: $\|V_x(\Psi(t))\Psi(t)\|_{L^2(\Omega;\mathbb{C}^N)} \le L_1 \|\Psi(t)\|_{L^2(\Omega;\mathbb{C}^N)}.$

Proof. The function $f : \mathbb{C}^N \to \mathbb{C}^N$, $f(z) = V_x(z)z$ is continuously differentiable with bounded derivative, hence Lipschitz continuous with Lipschitz constant L_1 from $\mathbb{C}^N$ to $\mathbb{C}^N$. With this preparation, we have Lipschitz continuity from L^2 to L^2 as follows:

$$\int_\Omega |V_x(\Psi(x,t))\Psi(x,t) - V_x(\Upsilon(x,t))\Upsilon(x,t)|^2 \text{d}x \le \int_\Omega L_1^2 |\Psi(x,t) - \Upsilon(x,t)|^2 \text{d}x.$$

Similarly, we have Lipschitz continuity from X to X^* as follows:

$$\|V_x(\Psi)\Psi - V_x(\Upsilon)\Upsilon\|^2_{X^*} \leq c\|V_x(\Psi)\Psi - V_x(\Upsilon)\Upsilon\|^2_Y$$

$$= c\int_0^T \|V_x(\Psi(x,t))\Psi(x,t) - V_x(\Upsilon(x,t))\Upsilon(x,t)\|^2_{L^2(\Omega;\mathbb{C}^N)}\mathrm{d}t$$

$$\leq c\int_0^T L_1^2\|\Psi(x,t) - \Upsilon(x,t)\|^2_{L^2(\Omega;\mathbb{C}^N)}\mathrm{d}t = cL_1^2\|\Psi - \Upsilon\|^2_Y \leq L_2^2\|\Psi - \Upsilon\|^2_X,$$

where we use the Gelfand triple $X \hookrightarrow Y \hookrightarrow X^*$ and the fact that $V_x(\Psi)\Psi \in Y$ as $V_x(\Psi) \in L^\infty(0,T;L^\infty(\Omega;\mathbb{R}))$. $\qquad\square$

The external potential is given by

$$V_{ext}(x,t;u) = V_0(x) + V_u(x)u(t), \tag{3.10}$$

where V_0 models a confinement potential, e.g., a harmonic trap in a solid state system or a molecule. The control potential $V_u(x)u(t)$ may represent a gate voltage applied to the solid state system or a laser pulse to the molecule.

We consider problems (3.1) and (3.3) in a unified framework by introducing a parameter α that indicates the case (3.1) by $\alpha = 1$ and (3.3) by $\alpha = 0$. The inhomogeneity F can be zero as in (3.1) or given as in (3.3). For the purpose of our work, we introduce the following general weak TDKS-type equation that includes the TDKS equation and its adjoint counterpart as particular cases. Our weak formulation is essential to prove existence of solutions in Sobolev spaces.

Find a wavefunction $\Psi \in X$ with $\Psi' \in X^*$, that is $\Psi \in W$, such that

$$i\,\langle\partial_t\Psi(t),\,\Phi\rangle_{L^2} = B(\Psi(t),\Phi;u(t)) + \alpha\,\langle V_{Hxc}(\Psi(t))\Psi(t),\,\Phi\rangle_{L^2} + \langle F(t),\,\Phi\rangle_{L^2}$$

a.e. in $(0,T)$ and $\forall \Phi \in H_0^1(\Omega;\mathbb{C}^N)$, $\tag{3.11}$

$$\Psi(0) = \Psi^0 \in L^2(\Omega;\mathbb{C}^N),$$

where the bilinear form $B(\Psi,\Phi;u)$ is defined as follows:

$$B(\Psi,\Phi;u) := \langle\nabla\Psi,\,\nabla\Phi\rangle_{L^2} + \langle V_{ext}(\cdot,t;u)\Psi,\,\Phi\rangle_{L^2}$$
$$+ (1-\alpha)\,\langle V_{Hxc}(\cdot,\Lambda)\Psi,\,\Phi\rangle_{L^2} + (1-\alpha)D(\Psi,\Phi). \tag{3.12}$$

The additional terms of the adjoint equation are given by

$$D(\Psi,\Phi) = D_H(\Psi,\Phi) + D_{xc}(\Psi,\Phi), \tag{3.13}$$

where

$$D_H(\Psi,\Phi) = \langle V_H(2\,\mathrm{Re}\,(\Psi,\Lambda)_{\mathbb{C}})\Lambda,\,\Phi\rangle_{L^2},$$

$$D_{xc}(\Psi,\Phi) = \left\langle 2\frac{\partial V_{xc}}{\partial\rho}(\Lambda)\,\mathrm{Re}\,(\Psi,\Lambda)_{\mathbb{C}}\,\Lambda,\,\Phi\right\rangle_{L^2}.$$

We remark that when studying the adjoint equation, the adjoint variable is also denoted with Ψ, and $\Lambda = (\lambda_1,\ldots,\lambda_N)$ corresponds to the solution of the forward equation (3.1). As we later prove in Theorem 3.16, the solution Λ of the forward equation is in $L^2(0,T;H^2(\Omega;\mathbb{C}^N))$ and the embedding $H^2(\Omega;\mathbb{C}^N) \hookrightarrow C(\overline{\Omega};\mathbb{C}^N)$ guarantees that $\Lambda(t)$ is bounded a.e. in $(0,T)$; see, e.g., [Cia13, p. 332]. Here, $C(\overline{\Omega};\mathbb{C}^N)$ is the space of continuous functions endowed with the norm $\|f\|_C = \max_{k=1,\ldots,N}\sup_{x\in\overline{\Omega}}|f_k(x)|$.

In a quantum control setting, the inhomogeneity F is zero in the forward equation and contains the derivative of J_1 with respect to the wavefunction in the adjoint equation. However, for generality we allow a non-zero F when studying (3.1). As in the argument above, Λ and $J_1(\Lambda)$ are continuous functions of x and hence in $L^2(\Omega)$. To incorporate a final condition $\Psi(T)$ instead of an initial condition $\Psi(0)$, we substitute $t \mapsto T - t$.

Now, we summarize the assumptions that we make throughout this thesis.

Assumption. We consider the following:

1) A bounded domain $\Omega \subset \mathbb{R}^n$ with $n = 3$ and a Lipschitz boundary;

2) For the improved regularity in Theorem 3.16 and 3.17 the boundary is taken to be even $\partial\Omega \in C^2$.

3) We use a cut-off at unphysically large densities for the exchange potential defined in (3.8). For all applications this does not change anything. Furthermore, this can be verified a posteriori; see also the remark on page 51.

4) The correlation potential V_c is uniformly bounded in the sense that $|V_c(\Psi(x,t))| \leq K \ \forall x \in \Omega, \ t \in [0,T], \ \Psi \in Y$; this corresponds to the correlation potentials from the Libxc library [MOB12] applying a similar approach to (3.8).

5) V_c is continuous from $L^2(\Omega; \mathbb{C}^N)$ to $L^2(\Omega)$.

6) The confining potential and the spacial dependence of the control potential are bounded, i.e. $V_0, V_u \in L^\infty(\Omega; \mathbb{R})$, where $\|f\|_{L^\infty} := \operatorname{ess\,sup}_{x \in \Omega} |f(x)|$ is the norm for $L^\infty(\Omega; \mathbb{R})$; as we consider a finite domain, this is equivalent to excluding divergent external potentials.

7) The control is $u \in H^1(0,T)$. This is a classical assumption in optimal control; see, e.g. [vWB08].

8) $\Psi^0 \in L^2$ for existence and uniqueness of the forward and adjoint equations.

9) $\Psi^0 \in H_0^1(\Omega; \mathbb{C}^N)$ and $F \in H^1(0,T; L^2(\Omega))$ for Theorems 3.16 and 3.17.

10) Only for Theorem 3.17: $\Psi^0 \in H^2(\Omega; \mathbb{C}^N) \cap H_0^1(\Omega; \mathbb{C}^N)$.

11) For the adjoint equation, we assume that the solution of the forward problem Λ is in $L^2(0,T; H^2(\Omega; \mathbb{C}^N))$; this can be shown by applying Theorem 3.16 to the forward problem.

We remark that the analysis provided in this work holds also for time-dependent external potentials $V_0 \in C([0,T]; L^\infty(\Omega))$.

3.3. Preliminary estimates

In this section, we study continuity properties of the KS potential and of the bilinear form. We begin with a general result on the Coulomb potential $w(x) = \frac{1}{|x|}$. Then we investigate the continuity of the Hartree potential that is defined as the convolution of w with the density ρ, and of the KS potential in more detail. Finally, we prove some estimates for the bilinear forms B and D defined in (3.12) and (3.13).

Lemma 3.3. *Given a bounded domain $\Omega \subset \mathbb{R}^n$ containing the origin, it holds that the Coulomb potential $w \in L^p(\Omega)$ if and only if $n > p$.*

Proof. By $B_R(0) := \{x \in \mathbb{R}^m : |x| < R\}$, we denote the open ball of radius $R \in \mathbb{R}^+$ around the origin. Consider now a ball $B_R(0) \subset \Omega$. Then by using spherical coordinates and the fact that $|x|$ does not depend on the orientation of x, we get (see e.g. [Eva10])

$$
\int_{B_R(0)} \frac{1}{|x|^p} \mathrm{d}x = \frac{n\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} \int_0^R \frac{1}{r^p} r^{n-1} \mathrm{d}r
$$

$$
= \frac{n\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} \left[\frac{r^{n-p}}{n-p} \right]_0^R = \begin{cases} \frac{n\pi^{n/2}}{\Gamma(\frac{n}{2}+1)} \frac{R^{n-p}}{n-p} < \infty, & n > p, \\ \infty, & n \le p, \end{cases}
$$

where Γ is the Γ-function. Outside this ball, $\frac{1}{|x|}$ is globally bounded. $\qquad\square$

Lemma 3.4. *For $\Phi, \Psi \in H^1(\Omega; \mathbb{C}^N)$ there exists a positive constant C_u such that*

$$
\|V_H(\Phi)\Phi - V_H(\Psi)\Psi\|_{L^2} \le C_u \left(\|\Phi\|_{H^1}^2 + \|\Psi\|_{H^1}^2 \right) \|\Phi - \Psi\|_{L^2}. \tag{3.14}
$$

Proof. We adapt Lemma 5 in [CL99] to our case of vector valued functions. To this end, we define $g^k(\Phi, \Psi)(x) := \int_\Omega \frac{\phi_k(y)\overline{\psi_k(y)}}{|x-y|} \mathrm{d}y$, $\tilde{g}(\Phi, \Psi)(x) := \sum_{k=1}^N g^k(\Phi, \Psi)(x)$. Then Lemma 3 in [CL99] gives

$$
|\tilde{g}(\Phi_1, \Phi_2)(x)| \le \sum_{k=1}^N |g^k(\Phi_1, \Phi_2)(x)| \le \sum_{k=1}^N \|\phi_{1,k}\|_{L^2} \|\nabla\phi_{2,k}\|_{L^2}.
$$

Using this fact and setting $\Upsilon = (v_1, \ldots, v_N)$, we have

$$
\begin{aligned}
\|\tilde{g}(\Phi_1, \Phi_2)\Upsilon\|_{L^2}^2 &= \sum_{k=1}^N \|\tilde{g}v_k\|_{L^2}^2 = \sum_{k=1}^N \int_\Omega |\tilde{g}(x)v_k(x)|^2 \mathrm{d}x \\
&\le \sum_{k=1}^N \left(\sum_{l=1}^N \|\phi_{1,l}\|_{L^2} \|\nabla\phi_{2,l}\|_{L^2} \right)^2 \int_\Omega |v_k(x)|^2 \mathrm{d}x \\
&= \left(\sum_{l=1}^N \|\phi_{1,l}\|_{L^2} \|\nabla\phi_{2,l}\|_{L^2} \right)^2 \|\Upsilon\|_{L^2}^2 \\
&\le N \sum_{l=1}^N \left(\|\phi_{1,l}\|_{L^2} \|\nabla\phi_{2,l}\|_{L^2} \right)^2 \|\Upsilon\|_{L^2}^2 \\
&\le N \sum_{l=1}^N \|\phi_{1,l}\|_{L^2}^2 \sum_{j=1}^N \|\nabla\phi_{2,j}\|_{L^2}^2 \|\Upsilon\|_{L^2}^2 \\
&= N \|\Phi_1\|_{L^2}^2 \|\nabla\Phi_2\|_{L^2}^2 \|\Upsilon\|_{L^2}^2.
\end{aligned} \tag{3.15}
$$

To prove (3.14), we add and subtract $V_H(\Phi)\Psi$ and use the triangle inequality as follows:

$$
\|V_H(\Phi)\Phi - V_H(\Psi)\Psi\|_{L^2} \le \|V_H(\Phi)(\Phi - \Psi)\|_{L^2} + \|(V_H(\Phi) - V_H(\Psi))\Psi\|_{L^2}. \tag{3.16}
$$

For the first term on the right-hand side of (3.16), we apply (3.15) with $\Phi_1 = \Phi_2 = \Phi$, $\Upsilon = \Phi - \Psi$ to obtain

$$
\|(|\Phi|^2 \star w)(\Phi - \Psi)\|_{L^2} \le \sqrt{N} \|\Phi\|_{L^2} \|\nabla\Phi\|_{L^2} \|\Phi - \Psi\|_{L^2}. \tag{3.17}
$$

For the second term on the right-hand side of (3.16), we introduce the decomposition

$$\sum_{k=1}^{N} |\phi_k|^2 - \sum_{k=1}^{N} |\psi_k|^2 = \sum_{k=1}^{N} \left(\phi_k \overline{(\phi_k - \psi_k)} + \overline{\psi_k}(\phi_k - \psi_k) \right).$$

Applying the triangle inequality to this decomposition and using (3.15) with $\Upsilon = \Psi$, $\Phi_1 = \Phi - \Psi$, $\Phi_2 = \Phi$ for the first term, and $\Upsilon = \Psi$, $\Phi_1 = \Phi - \Psi$, $\Phi_2 = \Psi$ for the second term, we find

$$\|((|\Phi|^2 - |\Psi|^2) \star w)\Psi\|_{L^2} \leq \|\tilde{g}(\Phi - \Psi, \Phi)\Psi\|_{L^2} + \|\tilde{g}(\Phi - \Psi, \Psi)\Psi\|_{L^2}$$
$$\leq \sqrt{N}\|\Psi\|_{L^2}(\|\nabla\Phi\|_{L^2} + \|\nabla\Psi\|_{L^2})\|\Phi - \Psi\|_{L^2}.$$

With this result, the proof of Lemma 5 in [CL99] extends to the vector case. $\qquad\square$

Lemma 3.5. *The nonlinear KS potential V_{Hxc} is a continuous function from $L^2(\Omega; \mathbb{C}^N)$ to $L^2(\Omega)$.*

Proof. First, we show that $\rho(\Psi)$ is a continuous mapping from $L^2(\Omega)$ to $L^1(\Omega)$ in the sense that from $\Psi^n \xrightarrow{L^2} \hat{\Psi}$ follows $\rho^n \xrightarrow{L^1} \hat{\rho}$. Using the Cauchy-Schwarz inequality, we have

$$\|\hat{\rho} - \rho^n\|_{L^1} = \int_{\Omega} \left| \sum_j |\hat{\psi}_j|^2 - \sum_j |\psi_j^n + \hat{\psi}_j - \hat{\psi}_j|^2 \right| \mathrm{d}x$$
$$\leq \int_{\Omega} \sum_j |\hat{\psi}_j - \psi_j^n|^2 + 2|\hat{\psi}_j - \psi_j^n||\hat{\psi}_j|\mathrm{d}x$$
$$\leq \|\hat{\psi} - \psi^n\|_{L^2}^2 + 2\|\hat{\psi} - \psi^n\|_{L^2}\|\hat{\psi}\|_{L^2}$$
$$\leq \left(1 + 2\|\hat{\psi}\|_{L^2}\right)\|\hat{\psi} - \psi^n\|_{L^2}$$

for n large enough.

Second, V_H is a continuous function of ρ as follows:

$$\|V_H(\rho_1) - V_H(\rho_2)\|_{L^2} = \|w \star (\rho_1 - \rho_2)\|_{L^2} \leq \|w\|_{L^2}\|\rho_1 - \rho_2\|_{L^1},$$

where Lemma 3.3 and Young's inequality [Sch05, Theorem 14.6] were used.

Finally, since V_x is Hölder continuous from $L^2(\Omega; \mathbb{C}^N)$ to $L^2(\Omega)$ it is also continuous between these spaces. The continuity of V_c is given by Assumption 5. $\qquad\square$

We remark that Lemma 3.4 holds for space dimension $n = 3$ as the Hardy's inequality used in [CL99]. Furthermore, for Young's inequality to hold in Lemma 3.5, the dimension needs to be larger than two. Therefore the analysis of the cases $n = 1$ and $n = 2$ requires a different approach that is not pursued in this work.

We continue with some estimates for the bilinear form B for arbitrary wavefunctions.

Lemma 3.6. *There exist positive constants c_0, c_1, and c_3 such that for any $\Psi, \Phi \in H^1(\Omega; \mathbb{C}^N)$ the following estimates hold:*

$$|D(\Psi, \Phi)| \leq c_0\|\Psi\|_{L^2}\|\Phi\|_{L^2}, \tag{3.18}$$
$$\operatorname{Re} B(\Psi, \Phi; u) \leq |B(\Psi, \Phi; u)| \leq c_1\|\Psi\|_{H^1}\|\Phi\|_{H^1}, \tag{3.19}$$
$$|\operatorname{Im} B(\Psi, \Psi; u)| \leq c_0\|\Psi\|_{L^2}^2, \tag{3.20}$$
$$\|\Psi\|_{H^1}^2 \leq \operatorname{Re} B(\Psi, \Psi; u) + c_3\|\Psi\|_{L^2}^2. \tag{3.21}$$

Proof. For $D(\Psi, \Phi)$ given by (3.13), we use the fact that $\Lambda(\cdot, t) \in L^\infty(\Omega)$ and Assumption 5 to get

$$
\begin{aligned}
|D_{xc}(\Psi, \Phi)| &= \left| \sum_{j=1}^{N} \sum_{k=1}^{N} \int_\Omega \frac{\partial V_{xc}}{\partial \rho}(\Lambda)(\psi_j \overline{\lambda_j} + \overline{\psi_j}\lambda_j)\lambda_k \overline{\phi_k} \mathrm{d}x \right| \\
&\leq 2 \sum_{j=1}^{N} \sum_{k=1}^{N} \int_\Omega \left| \frac{\partial V_{xc}}{\partial \rho}(\Lambda) \right| |\psi_j| \left(\sum_{l=1}^{N} |\lambda_l| \right)^2 |\phi_k| \mathrm{d}x \\
&\leq 2N \sum_{j=1}^{N} \sum_{k=1}^{N} \int_\Omega \left| \frac{\partial V_{xc}}{\partial \rho}(\Lambda) \right| \left(\sum_{l=1}^{N} |\lambda_l|^2 \right) |\psi_j||\phi_k| \mathrm{d}x \\
&\leq 2N \sum_{j=1}^{N} \sum_{k=1}^{N} \left\| \frac{\partial V_{xc}}{\partial \rho}(\Lambda) \sum_{l=1}^{N} |\lambda_l|^2 \right\|_{L^\infty} \int_\Omega |\psi_j||\overline{\phi_k}| \mathrm{d}x \\
&\leq 2N \left\| \frac{\partial V_{xc}}{\partial \rho}(\Lambda) \sum_{l=1}^{N} |\lambda_l|^2 \right\|_{L^\infty} \sum_{j=1}^{N} \sum_{k=1}^{N} \|\psi_j\|_{L^2} \|\phi_k\|_{L^2} \\
&\leq 2N^2 \left\| \frac{\partial V_{xc}}{\partial \rho}(\Lambda) \sum_{l=1}^{N} |\lambda_l|^2 \right\|_{L^\infty} \|\Psi\|_{L^2} \|\Phi\|_{L^2} \\
&\leq c_0' \|\Psi\|_{L^2} \|\Phi\|_{L^2}.
\end{aligned}
\tag{3.22}
$$

Similarly, defining $\theta := \sum_{j=1}^{N} (|\overline{\psi_j}| + |\psi_j|) \star w$, $\Theta := (\theta, \dots, \theta)^T$, we have

$$
\begin{aligned}
|D_H(\Psi, \Phi)| &= \left| \sum_{k=1}^{N} \int_\Omega \int_\Omega \frac{\sum_{j=1}^{N}(\psi_j(y)\overline{\lambda_j(y)} + \overline{\psi_j(y)}\lambda_j(y))}{|x - y|} \lambda_k(x)\overline{\phi_k(x)} \mathrm{d}y \mathrm{d}x \right| \\
&\leq \|\Lambda\|_{L^\infty}^2 \langle \Theta, |\Phi| \rangle_{L^2} \leq 2\|\Lambda\|_{L^\infty}^2 \sqrt{N} \left\| \sum_{j=1}^{N} |\psi_j| \star w \right\|_{L^2(\Omega;\mathbb{C})} \|\Phi\|_{L^2} \\
&\leq 2N^{\frac{3}{2}} \|w\|_{L^1} \|\Lambda\|_{L^\infty}^2 \|\Psi\|_{L^2} \|\Phi\|_{L^2} \leq c_0'' \|\Psi\|_{L^2} \|\Phi\|_{L^2},
\end{aligned}
\tag{3.23}
$$

where Young's inequality is used; see, e.g., [Sch05, Theorem 14.6]. Using (3.22) and (3.23), we have the desired bound on D with $c_0 = c_0' + c_0''$.

For the second estimate, we first recall that the embedding $H^1(0, T) \hookrightarrow C[0, T]$ is continuous and compact (see, e.g., [Cia13]), hence, there exists a positive constant K such that $\|u\|_{C[0,T]} \leq K\|u\|_{H^1(0,T)}$ for any $u \in H^1(0, T)$; this is used for the control function u. Consequently, recalling (3.12) and using (3.18), we obtain the following estimate:

$$
\begin{aligned}
|B(\Psi, \Phi; u)| \\
&\leq \|\nabla\Psi\|_{L^2}\|\nabla\Phi\|_{L^2} + (1-\alpha)\big(|D(\Psi, \Phi)| + \|V_{Hxc}(\Lambda)\Psi\|_{L^2}\|\Phi\|_{L^2}\big) \\
&\quad + \left| \int_\Omega (V_0(x)\Psi(x), \Phi(x))_\mathbb{C} \, \mathrm{d}x \right| + \left| \int_\Omega (V_u(x)u(t)\Psi(x), \Phi(x))_\mathbb{C} \, \mathrm{d}x \right| \\
&\leq \|\Psi\|_{H^1}\|\Phi\|_{H^1} + (c_0 + \|V_{Hxc}(\Lambda)\|_{L^\infty})\|\Psi\|_{H^1}\|\Phi\|_{H^1} \\
&\quad + \|V_0\|_{L^\infty}\|\Psi\|_{L^2}\|\Phi\|_{L^2} + K\|u\|_{H^1(0,T)}\|V_u\|_{L^\infty}\|\Psi\|_{L^2}\|\Phi\|_{L^2} \\
&= \Big(1 + c_0 + \|V_{Hxc}(\Lambda)\|_{L^\infty} + \|V_0\|_{L^\infty} + K\|u\|_{H^1(0,T)}\|V_u\|_{L^\infty}\Big)\|\Psi\|_{H^1}\|\Phi\|_{H^1}.
\end{aligned}
\tag{3.24}
$$

Hence there exists a constant c_1 such that (3.19) holds.

The estimate (3.20) is easily verified with the above estimates for D, as $\langle \nabla \Psi, \nabla \Psi \rangle_{L^2}$ and $\langle V_{ext} \Psi, \Psi \rangle_{L^2}$ in $B(\Psi, \Psi; u)$ are real.

To prove the last statement, we write as follows:

$$\langle \nabla \Psi, \nabla \Psi \rangle_{L^2} = B(\Psi, \Psi; u) - \langle V_{ext} \Psi, \Psi \rangle_{L^2} - (1-\alpha)D(\Psi, \Psi) - (1-\alpha)\langle V_{Hxc}(\Lambda)\Psi, \Psi \rangle_{L^2}.$$

Taking the real part of this equation and estimating as in (3.24) results in

$$\langle \nabla \Psi, \nabla \Psi \rangle_{L^2} \leq \operatorname{Re} B(\Psi, \Psi; u) + \Big(\|V_0\|_{L^\infty} + K\|u\|_{H^1(0,T)}\|V_u\|_{L^\infty}$$
$$+ (c_0 + \|V_{Hxc}(\Lambda)\|_{L^\infty})\Big)\|\Psi\|_{L^2}^2.$$

Adding $\|\Psi\|_{L^2}^2$ on both sides, we obtain

$$\|\Psi\|_{H^1}^2 \leq \operatorname{Re} B(\Psi, \Psi; u) + c_3\|\Psi\|_{L^2}^2, \tag{3.25}$$

where $c_3 = \Big(\|V_0\|_{L^\infty} + K\|u\|_{H^1(0,T)}\|V_u\|_{L^\infty} + c_0 + \|V_{Hxc}(\Lambda)\|_{L^\infty} + 1 \Big)$ and, hence, (3.21) holds. $\qquad\square$

3.4. A Galerkin approach

In this section, we introduce a finite-dimensional subspace P_m of $H_0^1(\Omega; \mathbb{C}^N)$, and show existence of a unique solution of (3.11) in this subspace. To this end, we take smooth functions $\phi_k = \phi_k(x) \in C_0^\infty(\Omega; \mathbb{C})$ for $k = 1, 2, \ldots$, such that $\{\phi_k\}_k$ is an orthogonal basis for $H_0^1(\Omega; \mathbb{C})$ and an orthonormal basis for $L^2(\Omega; \mathbb{C})$. Further, we construct a basis $\{\Phi_{k,j}\}_{k,j}$ that is orthogonal for $H_0^1(\Omega; \mathbb{C}^N)$ and orthonormal for $L^2(\Omega; \mathbb{C}^N)$ by defining

$$\Phi_{k,j}(x) := e_j \phi_k(x), \tag{3.26}$$

where $e_j = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^N$ is the jth canonical basis vector.

For a fixed positive integer m, we define a function Ψ_m as follows:

$$\Psi_m(x,t) := \sum_{k=1}^{m} \sum_{j=1}^{N} d_{k,j}^m(t) \Phi_{k,j}(x), \tag{3.27}$$

where the coefficients $d_{k,j}^m : [0,T] \to \mathbb{C}$ are such that

$$d_{k,j}^m(0) = \langle \Psi^0, \Phi_{k,j} \rangle_{L^2} \tag{3.28}$$

for $k = 1, \ldots, m$, $j = 1, \ldots, N$. The space spanned by the first mN basis functions is called $P_m = \operatorname{span}_{k=1,\ldots,m; j=1,\ldots,N}\{\Phi_{kj}\}$.

Moreover, by testing (3.11) for $\Phi = \Phi_{k,j}$, we obtain the following:

$$i \langle \partial_t \Psi_m, \Phi_{k,j} \rangle_{L^2} = B(\Psi_m, \Phi_{k,j}; u) + \alpha \langle V_{Hxc}(\Psi_m)\Psi_m, \Phi_{k,j} \rangle_{L^2} + \langle F, \Phi_{k,j} \rangle_{L^2}, \tag{3.29}$$

for almost all $0 \leq t \leq T$ and all $k = 1, \ldots, m$, $j = 1, \ldots, N$. Thus we seek a solution Ψ_m in the form (3.27) that satisfies the projection (3.29) of problem (3.11) onto the finite dimensional subspace $W_m = L^2(0, T; P_m)$.

Lemma 3.7. *Recall* Ψ_m *as in* (3.27) *and define* $G_{k,j} : \mathbb{C}^{mN} \to \mathbb{C}$ *as* $d^m \mapsto G_{k,j}(d^m) := \langle V_{Hxc}(\Psi_m)\Psi_m, \Phi_{k,j} \rangle_{L^2}$, $d^m = (d_{1,1}^m, \ldots, d_{m,N}^m)$. *The map* $d^m \mapsto G_{k,j}(d^m)$ *is locally Lipschitz continuous.*

Proof. We want to show local Lipschitz continuity in d_m, i.e. that for every $\epsilon > 0$ there exists a positive constant L such that $|G_{k,j}(d^m) - G_{k,j}(b^m)| \leq L|d^m - b^m|$ for all $d^m, b^m \in B_\epsilon(0)$. For a wavefunction Ψ_m in the Galerkin subspace W_m with $d_m \in B_\epsilon(0)$, we have the following bounds:

$$\|\Psi_m\|_{L^2}^2 = \int_\Omega \left| \sum_{k=1}^m \sum_{j=1}^N d_{k,j}^m \Phi_{k,j}(x) \right|^2 \mathrm{d}x \leq m \sum_{k=1}^m \sum_{j=1}^N |d_{k,j}^m|^2 \int_\Omega |\Phi_{k,j}(x)|^2 \mathrm{d}x$$

$$= m \sum_{k=1}^m \sum_{j=1}^N |d_{k,j}^m|^2 \leq m^2 N \epsilon^2,$$

$$\|\nabla \Psi_m\|_{L^2}^2 = \int_\Omega \left| \sum_{k=1}^m \sum_{j=1}^N d_{k,j}^m \nabla\Phi_{k,j}(x) \right|^2 \mathrm{d}x$$

$$\leq m \sum_{k=1}^m \sum_{j=1}^N |d_{k,j}^m|^2 \int_\Omega |\nabla\Phi_{k,j}(x)|^2 \mathrm{d}x = m \sum_{k=1}^m \sum_{j=1}^N |d_{k,j}^m|^2 C_m \leq m^2 N \epsilon^2 C_m$$

with $C_m = \max_{k=1,\ldots,m, j=1,\ldots,N} \|\nabla\Phi_{k,j}\|_{L^2}^2$. From these two bounds, we obtain $\|\Psi_m\|_{H^1}^2 \leq (C_m+1)m^2 N \epsilon^2$. Now, we prove the local Lipschitz property for the different potentials. Consider Ψ_m and Υ_m with coefficients d_m, b_m in $B_\epsilon(0)$. We obtain the following:

$$\begin{aligned} &| \langle V_H(\Psi_m)\Psi_m - V_H(\Upsilon_m)\Upsilon_m, \Phi_{k,j} \rangle_{L^2} | \\ &\leq \|V_H(\Psi_m)\Psi_m - V_H(\Upsilon_m)\Upsilon_m\|_{L^2} \|\Phi_{k,j}\|_{L^2} \\ &\leq C_u \left(\|\Psi_m\|_{H^1}^2 + \|\Upsilon_m\|_{H^1}^2 \right) \|\Psi_m - \Upsilon\|_{L^2} \leq \tilde{L} \|\Psi_m - \Upsilon_m\|_{L^2}, \end{aligned} \tag{3.30}$$

where the constant L depends on the dimension of the Galerkin space m, the norm of the derivatives of the basis functions C_m, and ϵ.

For the exchange-correlation potential, we have from Assumption 4 and Lemma 3.2 the following estimates:

$$\begin{aligned} &| \langle V_c(\Psi_m)\Psi_m - V_c(\Upsilon_m)\Upsilon_m, \Phi_{k,j} \rangle_{L^2} | \leq \|V_c\|_{L^\infty} \|\Psi_m - \Upsilon_m\|_{L^2}, \\ &| \langle V_x(\Psi_m)\Psi_m - V_x(\Upsilon_m)\Upsilon_m, \Phi_{k,j} \rangle_{L^2} | \leq L \|\Psi_m - \Upsilon_m\|_{L^2}. \end{aligned} \tag{3.31}$$

Using the estimates (3.30) and (3.31), we have

$$\begin{aligned} |G_{k,j}(d^m) - G_{k,j}(b^m)| &= | \langle V_H(\Psi_m)\Psi_m + V_x(\Psi_m)\Psi_m + V_c(\Psi_m)\Psi_m \\ &\quad -V_H(\Upsilon_m)\Upsilon_m - V_x(\Upsilon_m)\Upsilon_m - V_c(\Upsilon_m)\Upsilon_m, \Phi_k \rangle_{L^2} | \\ &\leq | \langle V_H(\Psi_m)\Psi_m - V_H(\Upsilon_m)\Upsilon_m, \Phi_k \rangle_{L^2} | \\ &\quad + | \langle V_c(\Psi_m)\Psi_m + V_x(\Psi_m)\Psi_m - V_c(\Upsilon_m)\Upsilon_m - V_x(\Upsilon_m)\Upsilon_m, \Phi_k \rangle_{L^2} | \\ &\leq L' \|\Psi_m - \Upsilon_m\|_{L^2}. \end{aligned}$$

Further, we have

$$\|\Psi_m - \Upsilon_m\|_{L^2}^2 = \int_\Omega \left| \sum_{k=1}^m \sum_{j=1}^N (d_{k,j}^m - b_{k,j}^m)\Phi_{k,j}(x) \right|^2 \mathrm{d}x$$

$$\leq mN \int_\Omega \sum_{k=1}^m \sum_{j=1}^N |d_{k,j}^m - b_{k,j}^m|^2 |\Phi_{k,j}(x)|^2 \mathrm{d}x$$

$$= mN \sum_{k=1}^m \sum_{j=1}^N |d_{k,j}^m - b_{k,j}^m|^2 \int_\Omega |\Phi_{k,j}(x)|^2 \mathrm{d}x \leq mN |d^m - b^m|^2.$$

All together, we have that $d^m \mapsto G_{k,j}(d^m)$ is locally Lipschitz continuous. $\qquad\square$

To show existence of a unique solution in the finite-dimensional Galerkin space, we use the Carathéodory theorem, see, e.g., [Wal98], because the time-dependent coefficients satisfy our differential equation only almost everywhere.

Theorem 3.8 (Carathéodory). *Consider the following initial value problem:*

$$\partial_t y(t) = f(t, y), \quad y(0) = \eta. \tag{3.32}$$

Let $S = [0, T] \times \mathbb{R}^m$ and assume f satisfies $f(\cdot, y) \in L^1(0, T)$ for fixed y and a generalized Lipschitz condition

$$|f(t, y) - f(t, \bar{y})| \le l(t)|y - \bar{y}| \quad in \ S, \ where \ l(t) \in L^1(0, T). \tag{3.33}$$

Then there exists a unique absolutely continuous solution satisfying (3.32) a.e. in $[0, T]$.

Theorem 3.9 (Construction of approximate solutions). *For each integer $m = 1, 2, \dots$ there exists a unique function $\Psi_m \in W_m$ of the form (3.27) satisfying (3.28) and (3.29).*

Proof. Assuming Ψ_m has the structure (3.27), we note that from the fact that the Φ_k are an orthonormal basis follows

$$\langle \partial_t \Psi_m(t), \Phi_{k,j} \rangle_{L^2} = \partial_t d_{k,j}^m(t). \tag{3.34}$$

Furthermore,

$$B(\Psi_m, \Phi_{k,j}; u) - (1 - \alpha)D(\Psi_m, \Phi_{k,j}) = \sum_{k'=1}^{m} \sum_{j'=1}^{N} e^{k'j'kj}(t) d_{k',j'}^m(t),$$

$$D(\Psi_m, \Phi_{k,j}) = \sum_{k'=1}^{m} \sum_{j'=1}^{N} \tilde{e}_r^{k'j'kj} \operatorname{Re}(d_{k',j'}^m(t)) + \tilde{e}_i^{k'j'kj} \operatorname{Im}(d_{k',j'}^m(t))$$

for $e^{k'j'kj} := B(\Phi_{k',j'}, \Phi_{k,j}; u) - (1 - \alpha)D(\Phi_{k',j'}, \Phi_{k,j})$, $\tilde{e}_r^{k'j'kj} := D(\Phi_{k',j'}, \Phi_{k,j})$, and $\tilde{e}_i^{k'j'kj} := D(i\Phi_{k',j'}, \Phi_{k,j})$, $k, k' = 1, \dots, m$, $j, j' = 1, \dots, N$. Define $f_{k,j}(t) := \langle F(t), \Phi_{k,j} \rangle_{L^2}$. Then (3.29) becomes a nonlinear system of ODEs as follows:

$$
\begin{aligned}
i\partial_t d_{k,j}^m(t) = f_{k,j}(t) + \alpha G_{k,j}(d^m(t)) + \sum_{k'=1}^{m} \sum_{j'=1}^{N} \Big(& e^{kjk'j'}(t) d_{k',j'}^m(t) \\
+ (1 - \alpha) \Big(& \tilde{e}_r^{kjk'j'} \operatorname{Re}(d_{k',j'}^m(t)) + \tilde{e}_i^{kjk'j'} \operatorname{Im}(d_{k',j'}^m(t)) \Big) \Big)
\end{aligned}
\tag{3.35}
$$

for $k = 1, \dots, m$, $j = 1, \dots, N$ with the initial conditions (3.28).

Notice that f is constant with respect to d^m, the first term in the brackets in (3.35) is linear, and the second and third terms are globally Lipschitz continuous with Lipschitz constant 1. By Lemma 3.7, $G_{k,j}$ is locally Lipschitz continuous in d^m on every ball $B_r(0)$, so the right-hand side of (3.35) is locally Lipschitz in d^m. As $G_{k,j}(d^m(t))$ depends on t only through $d_{k,j}^m(t)$ and $f_{k,j} \in L^2(0, T; \mathbb{C})$ and $e^{kjk'j'} \in H^1(0, T; \mathbb{C})$ through u, the right hand side is also in $L^2(0, T; \mathbb{C})$ and therefore the required $L^1(0, T; \mathbb{C})$-bound exists. Hence, we can invoke the Carathéodory theorem to show that (3.35) has a unique solution in the sense of Theorem 3.8. $\qquad\square$

3.5. Energy estimates

In this section, we prove energy estimates concerning our evolution problem that are used to prove existence and uniqueness of a solution in W.

Theorem 3.10. *Let $\Psi \in W_m$ be a solution of*

$$i \left\langle \partial_t \Psi(t), \Phi \right\rangle_{L^2} = B(\Psi(t), \Phi; u(t)) + \alpha \left\langle V_{Hxc}(\Psi(t)) \Psi(t), \Phi \right\rangle_{L^2} + \left\langle F(t), \Phi \right\rangle_{L^2} \quad (3.36)$$

$\forall \Phi \in P_m$, a.e. in $(0, T)$. Then there exist positive constants C, C_0, C_1, C', and C'' such that the following estimates hold:

$$\max_{0 \leq t \leq T} \|\Psi(t)\|_{L^2}^2 \leq C \left(\|\Psi^0\|_{L^2}^2 + \|F\|_Y^2 \right), \quad (3.37)$$

$$|B(\Psi(t), \Psi(t); u)| \leq C_0 (\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2) \text{ for a.a. } t \in [0, T], \quad (3.38)$$

$$\|\Psi(t)\|_{H^1}^2 \leq C_1 (\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2 + \|F\|_Y^2) \text{ for a.a. } t \in [0, T], \quad (3.39)$$

$$\|\Psi\|_X^2 \leq C', \quad (3.40)$$

$$\|\Psi'\|_{X^*}^2 \leq C''. \quad (3.41)$$

The same estimates hold for $\Psi \in W$ solving (3.11).

Proof. <u>Estimate 1:</u> Testing (3.36) with $\Psi(\cdot, t)$, we obtain

$$i \left\langle \partial_t \Psi(t), \Psi(t) \right\rangle_{L^2} = B(\Psi(t), \Psi(t); u) + \alpha \left\langle V_{Hxc}(\Psi(t)) \Psi(t), \Psi(t) \right\rangle_{L^2} + \left\langle F(t), \Psi(t) \right\rangle_{L^2}, \quad (3.42)$$

a.e. in $(0, T)$. This equation is equivalent to (see e.g. [Eva10])

$$i \frac{1}{2} \frac{d}{dt} \|\Psi(t)\|_{L^2}^2 = B(\Psi(t), \Psi(t); u) + \alpha \left\langle V_{Hxc}(\Psi(t)) \Psi(t), \Psi(t) \right\rangle_{L^2} + \left\langle F(t), \Psi(t) \right\rangle_{L^2}. \quad (3.43)$$

Now, we notice that the left-hand side is purely imaginary, while the terms $\left\langle V_{Hxc}(\Psi) \Psi, \Psi \right\rangle_{L^2}$ and $B(\Psi, \Psi; u)$ apart from $D(\Psi, \Psi)$ are purely real. Therefore, by splitting (3.43) into real and imaginary parts, we obtain the following:

$$\frac{1}{2} \frac{d}{dt} \|\Psi(t)\|_{L^2}^2 = \text{Im} \left(\left\langle F(t), \Psi(t) \right\rangle_{L^2} + (1 - \alpha) D(\Psi(t), \Psi(t)) \right) \quad (3.44)$$

and

$$\text{Re} \, B(\Psi(t), \Psi(t); u) + \alpha \left\langle V_{Hxc}(\Psi(t)) \Psi(t), \Psi(t) \right\rangle_{L^2} + \text{Re} \left(\left\langle F(t), \Psi(t) \right\rangle_{L^2} \right) = 0. \quad (3.45)$$

Now, using Lemma 3.6 and defining $\tilde{c}_0 := (1 - \alpha) c_0$, equation (3.44) becomes as follows:

$$\frac{d}{dt} \|\Psi(t)\|_{L^2}^2 \leq 2 \|F(t)\|_{L^2} \|\Psi(t)\|_{L^2} + 2 \tilde{c}_0 \|\Psi(t)\|_{L^2}^2 \leq \|F(t)\|_{L^2}^2 + (1 + 2\tilde{c}_0) \|\Psi(t)\|_{L^2}^2.$$

By defining $\eta(t) := \|\Psi(t)\|_{L^2}^2$ and $\xi(t) := \|F(t)\|_{L^2}^2$ the previous inequality becomes as follows:

$$\eta'(t) \leq (1 + 2\tilde{c}_0) \eta(t) + \xi(t), \quad (3.46)$$

a.e. in $(0, T)$. Thus, by applying the Gronwall inequality [Eva10] in the differential form, we obtain that

$$\eta(t) \leq e^{(1 + 2\tilde{c}_0)t} \left(\eta(0) + \int_0^t \xi(s) ds \right). \quad (3.47)$$

Notice that because of (3.28), it holds that $\eta(0) = \|\Psi(0)\|_{L^2}^2 = \|\Psi^0\|_{L^2}^2$. Hence, by using (3.47), we know that there exists a positive constant C such that the following estimate holds:

$$\max_{0 \leq t \leq T} \|\Psi(t)\|_{L^2}^2 \leq C \left(\|\Psi^0\|_{L^2}^2 + \|F\|_Y^2 \right). \quad (3.48)$$

For $\Psi \in W$, we have the continuous embedding $W \hookrightarrow C([0,T]; L^2(\Omega; \mathbb{C}^N))$; see, e.g. [Eva10, p. 287]. With this, we can evaluate Ψ at each time $t \in [0,T]$ and find the same estimate in the case that $\Psi \in W$ solves (3.11).

Estimate 2: Taking the real part of (3.43), we find

$$\operatorname{Re} B(\Psi(t), \Psi(t); u) + \alpha \langle V_{Hxc}(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2} + \operatorname{Re}\langle F(t),\, \Psi(t)\rangle_{L^2} = 0.$$

Using that $V_H \geq 0$ and $V_c \in L^\infty(\Omega)$, we get

$$\begin{aligned}
\operatorname{Re} B(\Psi(t), \Psi(t); u) = &-\operatorname{Re}\langle F(t),\, \Psi(t)\rangle_{L^2} \\
&+ \alpha\big(-\langle V_H(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2} - \langle V_x(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2} - \langle V_c(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2}\big) \\
\leq &\,|\langle V_x(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2}| + |\langle V_c(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2}| + |\operatorname{Re}\langle F(t),\, \Psi(t)\rangle_{L^2}| \\
\leq &\,|\langle V_x(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2}| + C_{V_c}\|\Psi(t)\|_{L^2}^2 + \|F(t)\|_{L^2}^2 + \|\Psi(t)\|_{L^2}^2.
\end{aligned}$$

From Lemma 3.2 and using (3.37), we obtain the following:

$$\operatorname{Re} B(\Psi(t), \Psi(t); u) \leq C\|\Psi(t)\|_{L^2}^2 + \|F(t)\|_{L^2}^2 \leq C_0'(\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2).$$

By Lemma 3.6, it holds that $\operatorname{Im} B(\Psi(t), \Psi(t); u) \leq c_0\|\Psi(t)\|_{L^2}^2$. Combining these two estimates one concludes (3.38).

As for the first estimate, the same applies in the case when $\Psi(t) \in W$ solves (3.11).

Estimate 3: We simply combine Lemma 3.6 with (3.38) and (3.37) to obtain

$$\begin{aligned}
\|\Psi(t)\|_{H^1}^2 &\leq \operatorname{Re} B(\Psi(t), \Psi(t); u(t)) + c_3\|\Psi(t)\|_{L^2}^2 \\
&\leq C_0(\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2) + c_3\|\Psi(t)\|_{L^2}^2 \\
&\leq (C_0 + c_3 C)(\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2 + \|F\|_Y^2).
\end{aligned}$$

Estimate 4: First, we need an adequate bound for the term $\langle V_{Hxc}(\Psi)\Psi,\, \Phi\rangle_{L^2}$ for any $\Phi \in L^2(\Omega; \mathbb{C}^N)$. For this reason, we write the following:

$$\langle V_{Hxc}(\Psi)\Psi,\, \Psi\rangle_{L^2} = \langle V_H(\Psi)\Psi,\, \Psi\rangle_{L^2} + \langle V_x(\Psi)\Psi,\, \Psi\rangle_{L^2} + \langle V_c(\Psi)\Psi,\, \Psi\rangle_{L^2}. \tag{3.49}$$

To bound V_H, we use the Cauchy-Schwarz inequality, Lemma 3.4, and (3.39) to arrive at

$$\begin{aligned}
\langle V_H(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2} &\leq C_u\|\Psi(t)\|_{H^1}^2\|\Psi(t)\|_{L^2}^2 \\
&\leq C_1(\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2 + \|F\|_Y^2)\|\Psi(t)\|_{L^2}^2.
\end{aligned} \tag{3.50}$$

Next, we recall that V_c is uniformly bounded (Assumption 4) and V_x is Lipschitz continuous from $L^2(\Omega; \mathbb{C}^N)$ to $L^2(\Omega; \mathbb{C}^N)$ (Lemma 3.2). Consequently, from (3.49), it follows that there exists a positive constant K' such that

$$\begin{aligned}
\langle V_{Hxc}(\Psi(t))\Psi(t),\, \Psi(t)\rangle_{L^2} &\leq \big(C_1(\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2 + \|F\|_Y^2) + K + L\big)\|\Psi(t)\|_{L^2}^2 \\
&\leq \big(C_1(\|\Psi^0\|_{L^2}^2 + \|F(t)\|_{L^2}^2 + \|F\|_Y^2) + K + L\big)C\big(\|\Psi^0\|_{L^2}^2 + \|F\|_Y^2\big) \\
&\leq K'\|F(t)\|_{L^2}^2 + K'',
\end{aligned} \tag{3.51}$$

where we used (3.37). By summing term by term (3.44) with (3.45), we get the following:

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\Psi(t)\|_{L^2}^2 + \operatorname{Re} B(\Psi(t), \Psi(t); u) = &\operatorname{Im}\langle F(t), \Psi(t)\rangle_{L^2} + (1-\alpha)\operatorname{Im} D(\Psi(t), \Psi(t)) \\
&- \alpha\langle V_{Hxc}(\Psi(t))\Psi(t), \Psi(t)\rangle_{L^2} - \operatorname{Re}\langle F(t), \Psi(t)\rangle_{L^2}.
\end{aligned} \tag{3.52}$$

Adding to both sides the term $c_3\|\Psi\|_{L^2}^2$, where c_3 is the same as in Lemma 3.6, we obtain the following:

$$
\begin{aligned}
\frac{1}{2}&\frac{\mathrm{d}}{\mathrm{d}t}\|\Psi(t)\|_{L^2}^2 + \operatorname{Re} B(\Psi(t),\Psi(t);u) + c_3\|\Psi(t)\|_{L^2}^2 \\
&= \operatorname{Im}\langle F(t),\,\Psi(t)\rangle_{L^2} + (1-\alpha)\operatorname{Im} D(\Psi(t),\Psi(t)) + c_3\|\Psi(t)\|_{L^2}^2 \\
&\quad - \alpha\langle V_{Hxc}(\Psi(t))\Psi(t),\,\Psi(t)\rangle_{L^2} - \operatorname{Re}\langle F(t),\,\Psi(t)\rangle_{L^2}.
\end{aligned}
\tag{3.53}
$$

Next, by applying Lemma 3.6 and using (3.51), (3.37), and the definition of K', we get the following:

$$
\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\Psi(t)\|_{L^2}^2 + \|\Psi(t)\|_{H^1}^2 &\le 2\|F(t)\|_{L^2}^2 + (3+c_3)\|\Psi(t)\|_{L^2}^2 + K'\|F(t)\|_{L^2}^2 + K'' \\
&\le c_4\|F(t)\|_{L^2}^2 + c_5
\end{aligned}
\tag{3.54}
$$

for some constants c_4, c_5 depending on $\|\Psi^0\|_{L^2}$ and $\|F\|_Y$. By manipulating (3.54) and integrating over $(0,T)$, we have

$$
\int_0^T \|\Psi(t)\|_{H^1}^2\,\mathrm{d}t \le \int_0^T c_4\|F(t)\|_{L^2}^2 + c_5\,\mathrm{d}t - \int_0^T \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\Psi(t)\|_{L^2}^2\,\mathrm{d}t,
\tag{3.55}
$$

which implies that

$$
\|\Psi\|_X^2 \le c_4\|F\|_Y^2 + c_5 T + \frac{1}{2}\left(\|\Psi^0\|_{L^2}^2 - \|\Psi(T)\|_{L^2}^2\right) \le c_4\|F\|_Y^2 + c_5 T + \frac{1}{2}\|\Psi^0\|_{L^2}^2.
\tag{3.56}
$$

This implies that there exists a positive constant C' depending only on T, $\|\Psi^0\|_{L^2}$, and $\|F\|_Y$ such that $\|\Psi\|_X^2 \le C'$. The same calculation can be done for $\Psi \in W$ being a solution of (3.11).

<u>Estimate 5:</u> Fix any $v \in H_0^1(\Omega;\mathbb{C}^N)$ with $\|v\|_{H^1} \le 1$. Write $v = v_1 + v_2$, where $v_1 \in \operatorname{span}\{\Phi_{k,j}\}_{k=1,j=1}^{m,N}$ and $\langle v_2, \phi_k\rangle_{L^2} = 0$ for $k = 1,\ldots,m$. Since the functions $\{\Phi_{k,j}\}_{k=1,j=1}^{\infty,N}$ are orthogonal in $H_0^1(\Omega;\mathbb{C}^N)$, we have

$$
1 \ge \|v\|_{H^1}^2 = \langle v_1 + v_2,\, v_1 + v_2\rangle_{H^1} = \|v_1\|_{H^1}^2 + \|v_2\|_{H^1}^2 \ge \|v_1\|_{H^1}^2.
\tag{3.57}
$$

Next, utilizing (3.11) with $\Psi \in W_m$ and testing with v_1, we obtain

$$
i\langle \partial_t \Psi(t),\, v_1\rangle_{L^2} = B(\Psi(t),v_1;u) + \alpha\langle V_{Hxc}(\Psi(t))\Psi(t),\, v_1\rangle_{L^2} + \langle F(t),\, v_1\rangle_{L^2}
\tag{3.58}
$$

a.e. in $[0,T]$. Using the decomposition of v, Lemma 3.6, and Young's inequality this implies that

$$
\begin{aligned}
|\langle \Psi'(t), v\rangle| &= |\langle \partial_t\Psi(t), v\rangle_{L^2}| = |\langle \partial_t\Psi(t), v_1\rangle_{L^2}| \\
&= |B(\Psi(t),v_1;u) + \alpha\langle V_{Hxc}(\Psi(t))\Psi(t),\, v_1\rangle_{L^2} + \langle F(t),\, v_1\rangle_{L^2}| \\
&\le c_1\left(\|\Psi(t)\|_{H^1}^2 + \|v_1\|_{H^1}^2\right) + \|V_{Hxc}(\Psi(t))\Psi(t)\|_{L^2}\|v_1\|_{L^2} + \|F(t)\|_{L^2}^2 + \|v_1\|_{L^2}^2,
\end{aligned}
\tag{3.59}
$$

where $\partial_t\Psi \in L^2(0,T;H_0^1(\Omega;\mathbb{C}^N))$ is the Riesz representative of $\Psi' \in L^2(0,T;H^{-1}(\Omega;\mathbb{C}^N))$ and $\langle\cdot,\,\cdot\rangle : H^{-1}(\Omega;\mathbb{C}^N) \times H_0^1(\Omega;\mathbb{C}^N) \to \mathbb{C}$ denotes the dual pairing for $H_0^1(\Omega;\mathbb{C}^N)$ and its dual $H^{-1}(\Omega;\mathbb{C}^N)$.

By using the Cauchy-Schwarz inequality, Assumptions 4, Lemma 3.2, and $\|v_1\|_{H^1} \le 1$, we have that there exists a positive constant $\tilde{K}$ such that

$$
|\langle (V_x(\Psi(t)) + V_c(\Psi(t)))\Psi(t),\, v_1\rangle_{L^2}| \le \|(V_x(\Psi(t)) + V_c(\Psi(t)))\Psi(t)\|_{L^2}\|v_1\|_{L^2} \le \tilde{K}\|\Psi(t)\|_{L^2}.
$$

By recalling (3.39), (3.51), and $\|v\|_{H^1} \leq 1$, we obtain that there exist positive constants $\tilde{C}_1$, $\tilde{C}_2$ depending on T, $\|\Psi^0\|_{L^2}$, and $\|F\|_Y$ such that

$$|\langle \Psi'(t), v \rangle| \leq \tilde{C}_1 \|F(t)\|_{L^2} + \tilde{C}_2, \tag{3.60}$$

and from (3.60), we have the following:

$$\|\Psi'(t)\|_{H^{-1}} = \sup_{0 \neq v \in H_0^1(\Omega)} \frac{|\langle \Psi'(t), v \rangle|}{\|v\|_{H^1}} \leq \tilde{C}_1 \|F(t)\|_{L^2} + \tilde{C}_2. \tag{3.61}$$

This implies that

$$\|\Psi'(t)\|_{H^{-1}}^2 \leq 2\tilde{C}_1^2 \|F(t)\|_{L^2}^2 + 2\tilde{C}_2^2.$$

By integrating over $(0, T)$, we obtain that there exists a positive constant C'' depending on T, $\|\Psi^0\|_{L^2}$, and $\|F\|_Y$ such that the following estimate holds:

$$\|\Psi'\|_{X^*}^2 \leq C'', \tag{3.62}$$

where $X^* = L^2(0, T; H^{-1}(\Omega))$ and the proof for $\Psi \in W_m$ is completed.

For $\Psi \in W$, no decomposition is necessary in (3.59), so we can use $v_1 = v, v_2 = 0$ and apply the same estimates to conclude our proof. $\qquad\square$

From (3.44), we have the following useful corollary for the solutions of the forward equation.

Corollary 3.11. *The $L^2(\Omega; \mathbb{C})$-norm of the solution to (3.11) for $F = 0$ and $\alpha = 1$, i.e. the forward equation, is conserved in the sense that $\|\Psi(\cdot, t)\|_{L^2(\Omega; \mathbb{C}^N)} = \|\Psi^0\|_{L^2(\Omega; \mathbb{C}^N)}$ for all $t \in [0, T]$.*

3.6. Existence of a weak solution

In the preceding section, we have proved the estimates in Theorem 3.10 for solutions $\Psi_m \in W_m$ in the Galerkin subspace. In this section, we use these estimates to show the existence of a solution in the full Sobolev space W. To this end, we make use of the embedding theorem by Aubin and Lions; see Theorem A.12.

Theorem 3.12. *Under Assumptions 1–8, problem (3.11) admits a weak solution for $\Psi^0 \in L^2(\Omega; \mathbb{C}^N)$, i.e. there exists a $\Psi \in W$ such that*

$$
\begin{aligned}
&i \langle \partial_t \Psi, \Phi \rangle_{L^2} = B(\Psi, \Phi; u) + \alpha \langle V_{Hxc}(\rho)\Psi, \Phi \rangle_{L^2} + \langle F, \Phi \rangle_{L^2} \\
&\text{a.e. in } (0, T) \text{ and } \forall \Phi \in H_0^1(\Omega; \mathbb{C}^N).
\end{aligned}
\tag{3.63}
$$

Proof. Consider a sequence $\{\Psi_m\}_{m=1}^\infty$ of solutions of the Galerkin problem (3.36), then according to the estimates (3.40) and (3.41) in Theorem 3.10, the sequence is bounded in X and $\{\Psi'_m\}_{m=1}^\infty$ is bounded in X^*. Consequently, there exists a subsequence $\{\Psi_{m_l}\}_{l=1}^\infty$ and a function $\Psi \in X$ with $\Psi' \in X^*$ such that $\Psi_{m_l} \rightharpoonup \Psi$ in X and $\Psi'_{m_l} \rightharpoonup \Psi'$ in X^*; see, e.g., [Eva10]. Moreover, by the Aubin-Lions Theorem A.12 we know that W is compactly embedded in $Y := L^2(0, T; L^2(\Omega))$. Consequently, we have strong convergence of the subsequence $\Psi_{m_l} \to \Psi$ in Y.

Next, we fix a positive integer M and construct a test function $\Phi \in C^1([0, T]; H_0^1(\Omega; \mathbb{C}^N))$ as follows:

$$\Phi(x, t) := \sum_{k=1}^M \sum_{j=1}^N d_{k,j}^m(t) \Phi_{k,j}(x), \tag{3.64}$$

where $\{d_m^k\}_{k=1,j=1}^{M,N}$ are given smooth functions. We choose $m \geq M$, multiply (3.29) by $d_{k,j}^m(t)$, sum over $k = 1, \ldots, M$, $j = 1, \ldots, N$, and integrate with respect to t to obtain the following:

$$\int_0^T i\langle \Psi_m', \Phi\rangle \mathrm{d}t = \int_0^T B(\Psi_m, \Phi; u) + \alpha \langle V_{Hxc}(\Psi_m)\Psi_m, \Phi\rangle_{L^2} + \langle F, \Phi\rangle_{L^2} \,\mathrm{d}t. \quad (3.65)$$

By now setting $m = m_l$ and by recalling continuity of $V_{Hxc}(\Psi)$ from Lemma 3.5 and strong convergence $\Psi_{m_l} \to \Psi$ in Y, we can pass to the limit to obtain

$$\int_0^T i\langle \Psi', \Phi\rangle \mathrm{d}t = \int_0^T B(\Psi, \Phi; u) + \alpha \langle V_{Hxc}(\Psi)\Psi, \Phi\rangle_{L^2} + \langle F, \Phi\rangle_{L^2} \,\mathrm{d}t. \quad (3.66)$$

This equality holds for all $\Phi \in X$ as functions of the form (3.64) are dense in X. Hence, in particular,

$$i\langle \Psi', v\rangle = B(\Psi, \Phi; u) + \alpha \langle V_{Hxc}(\Psi)\Psi, v\rangle_{L^2} + \langle F, v\rangle_{L^2} \quad (3.67)$$

for any $v \in H_0^1(\Omega; \mathbb{C}^N)$ and a.e. in $[0, T]$. From [Eva10, Theorem 3, p. 287], we also know that $\Psi \in C([0, T]; L^2(\Omega; \mathbb{C}^N))$.

It remains to prove that $\Psi(\cdot, 0) = \Psi^0$. To do so, we first notice from (3.66) that the following holds:

$$\int_0^T -i\langle \Phi', \Psi\rangle \mathrm{d}t = \int_0^T B(\Psi, \Phi; u) + \alpha \langle V_{Hxc}(\Psi)\Psi, \Phi\rangle_{L^2} + \langle F, \Phi\rangle_{L^2} \mathrm{d}t + \langle \Psi(0), \Phi(0)\rangle_{L^2} \quad (3.68)$$

for any $\Phi \in C^1([0, T]; H_0^1(\Omega; \mathbb{C}^N))$ with $\Phi(T) = 0$. Similarly, from (3.65) we get

$$\begin{aligned}\int_0^T -i\langle \Phi', \Psi_m\rangle \mathrm{d}t = &\int_0^T B(\Psi_m, \Phi; u) + \alpha \langle V_{Hxc}(\Psi_m)\Psi_m, \Phi\rangle_{L^2} + \langle F, \Phi\rangle_{L^2} \,\mathrm{d}t \\ &+ \langle \Psi_m(0), \Phi(0)\rangle_{L^2}.\end{aligned} \quad (3.69)$$

We set $m = m_l$ and use again the considered convergences to find

$$\int_0^T -i\langle \Phi', \Psi\rangle \mathrm{d}t = \int_0^T B(\Psi, \Phi; u) + \alpha \langle V_{Hxc}(\Psi)\Psi, \Phi\rangle_{L^2} + \langle F, \Phi\rangle_{L^2} \,\mathrm{d}t + \langle \Psi^0, \Phi(0)\rangle_{L^2}, \quad (3.70)$$

because $\Psi_{m_l}(0) \to \Psi^0$. As $\Phi(0)$ is arbitrary, by comparing (3.68) and (3.70) we conclude that $\Psi(0) = \Psi^0$. $\qquad\square$

3.7. Uniqueness of a weak solution

We have shown that there exists at least one solution $\Psi \in W$ of (3.11). Now, we can apply Theorem 3.10 to the space W and use the Lipschitz properties of the potentials to show that the solution is indeed unique.

Theorem 3.13. *The weak form of the Kohn-Sham equations* (3.11) *is uniquely solvable.*

Proof. Seeking a contradiction, we assume that there exist two distinct weak solutions of (3.11), Ψ and Υ, in W with $\|\Psi - \Upsilon\|_X > 0$. Therefore, we have

$$i\langle \partial_t \Psi, \Phi\rangle_{L^2} = B(\Psi, \Phi; u) + \alpha \langle V_{Hxc}(\Psi)\Psi, \Phi\rangle_{L^2} + \langle F, \Phi\rangle_{L^2} \quad (3.71)$$

and

$$i \langle \partial_t \Upsilon, \, \Phi \rangle_{L^2} = B(\Upsilon, \Phi; u) + \alpha \langle V_{Hxc}(\Upsilon)\Upsilon, \, \Phi \rangle_{L^2} + \langle F, \, \Phi \rangle_{L^2} \tag{3.72}$$

for all test functions $\Phi \in H_0^1(\Omega; \mathbb{C}^N)$. Subtracting term-by-term (3.72) from (3.71) and defining $\hat{\Psi} := \Psi - \Upsilon$, we obtain the following:

$$i \left\langle \partial_t \hat{\Psi}, \, \Phi \right\rangle_{L^2} = B(\hat{\Psi}, \Phi; u) + \alpha \langle V_{Hxc}(\Psi)\Psi - V_{Hxc}(\Upsilon)\Upsilon, \, \Phi \rangle_{L^2}. \tag{3.73}$$

By testing the previous identity (3.73) with $\Phi = \hat{\Psi}(t)$, we obtain

$$i \left\langle \partial_t \hat{\Psi}, \, \hat{\Psi} \right\rangle_{L^2} = B(\hat{\Psi}, \hat{\Psi}; u) + \alpha \left\langle V_{Hxc}(\Psi)\Psi - V_{Hxc}(\Upsilon)\Upsilon, \, \hat{\Psi} \right\rangle_{L^2}. \tag{3.74}$$

Similarly, as for (3.43), we have

$$i \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\hat{\Psi}\|_{L^2}^2 = B(\hat{\Psi}, \hat{\Psi}; u) + \alpha \left\langle V_{Hxc}(\Psi)\Psi - V_{Hxc}(\Upsilon)\Upsilon, \, \hat{\Psi} \right\rangle_{L^2}. \tag{3.75}$$

Now, we notice that the left-hand side is purely imaginary. Consequently, by taking the imaginary part of (3.75), we obtain the following:

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\hat{\Psi}\|_{L^2}^2 = \alpha \operatorname{Im} \left(\left\langle V_{Hxc}(\Psi)\Psi - V_{Hxc}(\Upsilon)\Upsilon, \, \hat{\Psi} \right\rangle_{L^2} + (1-\alpha) D(\hat{\Psi}, \hat{\Psi}) \right). \tag{3.76}$$

From (3.76) and (3.18) in Lemma 3.6, we get

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\hat{\Psi}\|_{L^2}^2 = 2\alpha \operatorname{Im} \left(\left\langle V_{Hxc}(\Psi)\Psi - V_{Hxc}(\Upsilon)\Upsilon, \, \hat{\Psi} \right\rangle_{L^2} \right) + 2(1-\alpha) \operatorname{Im} D(\hat{\Psi}, \hat{\Psi})$$
$$\leq \|V_{Hxc}(\Psi)\Psi - V_{Hxc}(\Upsilon)\Upsilon\|_{L^2} \|\hat{\Psi}\|_{L^2} + 2c_0 \|\hat{\Psi}\|_{L^2}^2.$$

Using Lemma 3.4, Assumption 4, Lemma 3.2, (3.37), and Theorem 3.10, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\hat{\Psi}\|_{L^2}^2 \leq (C_u(\|\Psi\|_{H^1}^2 + \|\Upsilon\|_{H^1}^2) + K + L) \|\hat{\Psi}\|_{L^2}^2 + 2c_0 \|\hat{\Psi}\|_{L^2}^2$$
$$\leq c^{\#} \left(L + K + \|F\|_{L^2}^2 + \|F\|_Y^2 + \|\Psi^0\|_{L^2}^2 + c_0 \right) \|\hat{\Psi}\|_{L^2}^2.$$

By defining $\eta(t) := \|\hat{\Psi}\|_{L^2}^2$ and $\vartheta(t) := c^{\#} \left(L + K + \|F(t)\|_{L^2}^2 + \|F\|_Y^2 + \|\Psi^0\|_{L^2}^2 + c_0 \right)$, we obtain the following inequality:

$$\eta'(t) \leq \vartheta(t)\eta(t). \tag{3.77}$$

By applying Gronwall's inequality, we obtain the following:

$$\eta(t) \leq \exp\left(\int_0^t \vartheta(s) ds \right) \eta(0). \tag{3.78}$$

By noticing that

$$\int_0^t \vartheta(s) ds = \int_0^t c^{\#} \left(L + K + \|F(t)\|_{L^2}^2 + \|F\|_Y^2 + \|\Psi^0\|_{L^2}^2 + c_0 \right) ds$$
$$\leq \int_0^T c^{\#} \left(L + K + \|F(t)\|_{L^2}^2 + \|F\|_Y^2 + \|\Psi^0\|_{L^2}^2 + c_0 \right) ds \tag{3.79}$$
$$\leq c^{\#} \left(\|F\|_Y^2 + T \|\Psi^0\|_{L^2}^2 + T(\|F\|_Y^2 + c_0 + L + K) \right),$$

and by recalling that $\eta(0) = \|\hat{\Psi}(0)\|_{L^2} = 0$, we obtain that $\|\hat{\Psi}\|_{L^2} \leq 0$ a.e. in $(0, T)$, and the claim follows. $\square$

3.8. Improved regularity

We have established the existence and uniqueness of a solution to (3.11) in W. So far, our result is similar to those in [Jer15], who has shown the existence and uniqueness of the solution of a KS equation in $C([0,T]; H_0^1(\Omega;\mathbb{C}^N)) \cap C^1([0,T]; H^{-1}(\Omega;\mathbb{C}^N))$; however, we used a different methodology that allowed us to derive these results using less strong assumptions. This is crucial for applications to optimal control problems. Now, we improve these results in the case $\Psi^0 \in H_0^1(\Omega)$. In particular, we prove that the solution to (3.11) is twice weakly differentiable in space and its first spatial derivative is bounded. To do so, we need the following preliminary lemmas.

Lemma 3.14 (Difference quotients). *Assume that for fixed $v \in L^p(V)$, $1 < p < \infty$, $V \subset\subset \Omega$, there exists a constant C such that $\|D^h v\|_{L^p(V)} \leq C$ for all $0 < |h| < \frac{1}{2}\operatorname{dist}(V, \partial\Omega)$, where*

$$D_i^h v(x) = \frac{v(x + e_i h) - v(x)}{h}, \quad D^h v = (D_1^h v, \ldots, D_n^h v).$$

Then

$$v \in H^{1,p}(V) \text{ with } \|Dv\|_{L^p(V)} \leq C,$$

where C may depend on u, e.g. on $\|v\|_{L^p(\Omega)}$. Furthermore, the statement holds for the case of two half-balls $\Omega = \{|x| < R\} \cap \{x_n > 0\}$ and $V = \{|x| < \frac{R}{2}\} \cap \{x_n > 0\}$.

Proof. For the proof, see [Eva10, §5.8.2, Theorem 3] and the remark after the proof. $\square$

Next, we extend the result in [Eva10, §6.3.2, Theorem 4] for linear elliptic problems to the case of a specific nonlinear problem.

Lemma 3.15. *Let $\varphi \in H_0^1(\Omega;\mathbb{C}^N)$ be a weak solution of the elliptic boundary value problem*

$$B(\varphi, v; u) + \langle V_{Hxc}(\varphi)\varphi, v \rangle_{L^2} = \langle A, v \rangle_{L^2} \quad \forall v \in H_0^1(\Omega;\mathbb{C}^N), \ A \in L^2(\Omega;\mathbb{C}^N).$$

such that $\|\varphi\|_{L^2}^2 \leq \gamma \|A\|_{L^2}^2$ holds for a constant γ independent on A. Furthermore, assume $\partial\Omega \in C^2$. Then $\varphi \in H^2(\Omega;\mathbb{C}^N)$ and

$$\|\varphi\|_{H^2} \leq c \left(\|A\|_{L^2} + \|\varphi\|_{L^2} \right),$$

where $c = \max\left\{ 1, \|V_0\|_{L^\infty} + \|u\|_{C(0,T)}\|V_u\|_{L^\infty} + \|\varphi\|_{H^1}^2 + L + K \right\}$.

Proof. To extend the results in [Eva10, §6.3.2, Theorem 4] to our case, two issues have to be treated carefully. First, the nonlinear potential has to be bounded in a suitable way and, second, extra care has to be taken when changing the coordinates.

The nonlinear potential has to be bounded in such a way that Lemma 3.14 can be applied. Therefore, we need to find a constant c such that $\|V_{Hxc}(\varphi)\varphi\|_{L^2}^2 \leq c\|\varphi\|_{L^2}^2$, where c is allowed to depend on φ. This can be done using Lemma 3.4 as follows:

$$\|V_H(\varphi)\varphi\|_{L^2} \leq C_u \|\varphi\|_{H^1}^2 \|\varphi\|_{L^2},$$

and using Lemma 3.2 and Assumption 4, we obtain

$$\|V_{Hxc}(\varphi)\varphi\|_{L^2}^2 \leq (C_u^2 \|\varphi\|_{H^1}^4 + L^2 + K^2)\|\varphi\|_{L^2}^2.$$

Now, we can apply Lemma 3.14 and the same argument as in [Eva10, §6.3.2, Theorem 4, 1.–5.] to obtain that the solution is in $H^2(B)$ for a half-ball B.

Furthermore, in the proof it is necessary to locally flatten out the boundary. This is done by a C^2-map that keeps all the coordinates apart from one dimension which is transformed onto a line. This ensures that the determinant of the Jacobian is equal to one.

The coordinate transformation of the Laplacian and the linear external potential is as for standard parabolic PDEs. The exchange and correlation potentials do not explicitly depend on space and time but only pointwise on the wavefunction. Hence a change of coordinates does not change the potential. For the Hartree potential, however, more care is needed. Let the change of coordinates be given by

$$x = k(\hat{x}), \qquad\qquad \hat{\Psi}(\hat{x}) = \Psi(k(\hat{x})).$$

Regarding the Hartree potential, one has to account for the fact that the transformation k is only locally defined as a C^2 map, so the transform to a global integral operator is not well-defined. However, it is possible to evaluate $V_H(\hat{\Psi})(\hat{x})$ as $V_H(\Psi)(x)$ in $x = k(\hat{x})$.

With this preparation, let B be the image of a half-ball under k. Then we bound Ψ by $\|\Psi\|_{H^1(B)} \leq C(\|F\|_{L^2(\Omega)} + \|\Psi\|_{L^2(\Omega)})$. As $\overline{\Omega}$ is compact, it can be covered with finitely many sets B_i, so we find

$$\|\Psi\|_{H^1(\Omega)} \leq \sum_i \|\Psi\|_{H^1(B_i)} \leq \sum_i C\left(\|F\|_{L^2(\Omega)} + \|\Psi\|_{L^2(\Omega)}\right).$$

Now the standard proof for elliptic equations based on difference quotients can be applied, e.g., [Eva10, §6.3.2, Theorem 4]. $\qquad\square$

Theorem 3.16. *Assume $\Psi^0 \in H_0^1(\Omega)$, $F \in H^1(0,T;L^2(\Omega))$ and $\partial\Omega \in C^2$, that is Assumptions 2 and 9. Suppose $\Psi \in W$ is the solution to (3.11). Then*

$$\Psi \in L^2(0,T;H^2(\Omega;\mathbb{C}^N)) \cap L^\infty(0,T;H_0^1(\Omega;\mathbb{C}^N)), \quad \Psi' \in L^2(0,T;L^2(\Omega;\mathbb{C}^N)).$$

Furthermore, the following estimate holds:

$$\operatorname*{ess\,sup}_{0\leq t\leq T} \|\Psi(t)\|_{H^1} + \|\Psi\|_{L^2(0,T;H^2(\Omega))} + \|\Psi'\|_Y \leq C\left(\|\Psi^0\|_{H^1} + \|F\|_{C([0,T],L^2(\Omega))}\right). \quad (3.80)$$

Proof. We remark that due to the embedding $H^1(0,T;L^2(\Omega;\mathbb{C}^N)) \hookrightarrow C([0,T];L^2(\Omega;\mathbb{C}^N))$ the function F is bounded by a constant $\hat{f} := \max_{t\in[0,T]} \|F(t)\|_{L^2}$. This means, we can take the essential supremum of (3.39) and obtain

$$\operatorname*{ess\,sup}_{0\leq t\leq T} \|\Psi(t)\|_{H^1(\Omega)}^2 \leq C(\|\Psi^0\|_{L^2(\Omega)}^2 + \hat{f}^2), \quad (3.81)$$

which means that $\Psi \in L^\infty(0,T;H_0^1(\Omega))$.

For $\partial_t\Psi$, we consider the Galerkin space W_m and take a fixed m, multiply (3.29) by $\partial_t d_{k,j}^m(t)$, and sum for $k = 1,\ldots,m$, $j = 1,\ldots,N$ to obtain the following:

$$\begin{aligned}
\langle \partial_t\Psi_m(t), \partial_t\Psi_m(t) \rangle_{L^2} &= B(\Psi_m(t), \partial_t\Psi_m(t); u(t)) \\
&\quad + \alpha \langle V_{Hxc}(\Psi(t))\Psi(t), \partial_t\Psi_m(t) \rangle_{L^2} + \langle F(t), \partial_t\Psi_m(t) \rangle_{L^2}
\end{aligned} \quad (3.82)$$

a.e. in $(0,T)$. For $D(\Psi_m(t), \partial_t\Psi_m(t))$, we have

$$|D_{xc}(\Psi_m(t), \partial_t\Psi_m(t))| = \left|2\left\langle \frac{\partial V_{xc}}{\partial\rho}(\Lambda(t))\operatorname{Re}(\Psi_m(t), \Lambda(t))_{\mathbb{C}} \Lambda(t), \partial_t\Psi_m(t) \right\rangle_{L^2}\right|.$$

Because $\Lambda(t) \in L^\infty(\Omega)$ and using Assumption 5, we have

$$
|D_{xc}(\Psi_m(t), \partial_t \Psi_m(t))| \leq C \int_\Omega \left| \sum_{i=1}^N \operatorname{Re} \psi_{i,m}(t) \sum_{j=1}^N \overline{\partial_t \psi_{j,m}(t)} \right| \mathrm{d}x
$$

$$
\leq C \int_\Omega \sum_{i,j=1}^N \frac{1}{\epsilon} |\psi_{i,m}(t)|^2 + \epsilon |\partial_t \psi_{j,m}(t)|^2 \mathrm{d}x \leq CN \left(\frac{1}{\epsilon} \|\Psi_m(t)\|_{L^2}^2 + \epsilon \|\partial_t \Psi_m(t)\|_{L^2}^2 \right),
$$

$$(3.83)$$

where we used Young's inequality for products. For D_H, we use Young's inequality for convolutions [Sch05, Theorem 14.6] and the fact that $\Lambda \in L^\infty(\Omega)$. We have

$$
|D_H(\Psi_m(t), \partial_t \Psi_m(t))| = \left| \sum_{j=1}^N \int_\Omega (2 \operatorname{Re} (\Psi_m(t), \Lambda(t))_{\mathbb{C}} \star w)(x) \Lambda_j(x,t) \overline{\partial_t \Psi_{m,j}(x,t)} \mathrm{d}x \right|
$$

$$
\leq C' \langle |\Psi_m(t)| \star w, \, |\partial_t \Psi_m(t)| \rangle_{L^2} \leq C' \| |\Psi_m(t)| \star w \|_{L^2} \|\partial_t \Psi_m(t)\|_{L^2}
$$

$$
\leq C' \|\Psi_m(t)\|_{L^2} \|w\|_{L^1} \|\partial_t \Psi_m(t)\|_{L^2} \leq C' \|w\|_{L^1} \left(\frac{1}{\epsilon} \|\Psi_m(t)\|_{L^2}^2 + \epsilon \|\partial_t \Psi_m(t)\|_{L^2}^2 \right),
$$

$$(3.84)$$

where w is the Coulomb potential. Consequently, by (3.83) and (3.84), we get the following:

$$
|D(\Psi_m(t), \partial_t \Psi_m(t))| \leq c \left(\frac{1}{\epsilon} \|\Psi_m(t)\|_{L^2}^2 + \epsilon \|\partial_t \Psi_m\|_{L^2}^2 \right) =: \tilde{D} \tag{3.85}
$$

for some constant c. Estimate (3.85) is used together with (3.82) to obtain the following:

$$
\|\partial_t \Psi_m(t)\|_{L^2}^2
$$

$$
\leq \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \langle \nabla \Psi_m(t), \nabla \Psi_m(t) \rangle_{L^2} + \langle V_{ext}(t) \Psi_m(t), \partial_t \Psi_m(t) \rangle_{L^2} + D(\Psi_m(t), \partial_t \Psi_m(t))
$$

$$
+ \langle V_{Hxc}(\Psi_m(t)) \Psi_m(t), \partial_t \Psi_m \rangle_{L^2} + \langle F(t), \partial_t \Psi_m(t) \rangle_{L^2}
$$

$$
\leq \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \langle \nabla \Psi_m(t), \nabla \Psi_m(t) \rangle_{L^2} + \|V_{ext}(t)\|_{L^\infty} \|\Psi_m(t)\|_{L^2} \|\partial_t \Psi_m(t)\|_{L^2} + \tilde{D}
$$

$$
+ C_u \|\Psi_m(t)\|_{L^2} \|\Psi_m(t)\|_{H^1}^2 \|\partial_t \Psi_m(t)\|_{L^2} + (K+L) \|\Psi_m(t)\|_{L^2} \|\partial_t \Psi_m(t)\|_{L^2}
$$

$$
+ \|F(t)\|_{L^2} \|\partial_t \Psi_m(t)\|_{L^2},
$$

where we use Lemma 3.4 and 3.2 and Assumption 4 to estimate V_{Hxc}. Next, by using the Cauchy-Schwarz inequality, (3.81), and Young's inequality with an arbitrary positive ϵ, we get

$$
\|\partial_t \Psi_m(t)\|_{L^2}^2
$$

$$
\leq \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\nabla \Psi_m(t)\|_{L^2}^2 + \frac{1}{\epsilon} \|F(t)\|_{L^2}^2 + \epsilon \|\partial_t \Psi_m(t)\|_{L^2}^2 + \tilde{D}
$$

$$
+ (\|V_{ext}(t)\|_{L^\infty} + K + L + C_1(\|\Psi^0\|_{L^2}^2 + \hat{f}^2)) \left(\frac{1}{\epsilon} \|\Psi_m(t)\|_{L^2}^2 + \epsilon \|\partial_t \Psi_m(t)\|_{L^2}^2 \right)
$$

$$
\leq \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \|\nabla \Psi_m(t)\|_{L^2}^2 + \frac{\Gamma}{\epsilon} \left(\|\Psi_m(t)\|_{L^2}^2 + \hat{f}^2 \right) + \epsilon \Gamma \|\partial_t \Psi_m(t)\|_{L^2}^2,
$$

where Γ is a constant depending only on $\|\Psi^0\|_{L^2}$, $\|F\|_Y$, $\max_{t \in [0,T]} \|V_{ext}(t)\|_{L^\infty}$, and K. Now, we choose ϵ small enough, that is, $\epsilon < \frac{1}{\Gamma}$ and integrate from 0 to T. We obtain

$$
\int_0^T \|\partial_t \Psi_m(t)\|_{L^2}^2 \mathrm{d}t \leq \frac{1}{1 - \epsilon \Gamma} \left(\operatorname*{ess\,sup}_{0 \leq t \leq T} \|\Psi_m(t)\|_{H^1}^2 + \frac{\Gamma}{\epsilon} \int_0^T \|\Psi_m(t)\|_{L^2}^2 + \hat{f}^2 \mathrm{d}t \right).
$$

Using (3.37) and (3.81), this gives

$$\|\partial_t \Psi_m\|_Y^2 \leq \Gamma'(\|\Psi^0\|_{L^2}^2 + \hat{f}^2). \tag{3.86}$$

Passing to the limit as $m \to \infty$ we find $\Psi' \in Y$.

Now, we rewrite (3.11) for a fixed time t as follows:

$$B(\Psi(t), \Phi; u(t)) + \alpha \langle V_{Hxc}(\Psi(t))\Psi(t), \Phi\rangle_{L^2} = \langle -F(t) + i\partial_t \Psi(t), \Phi\rangle_{L^2}, \tag{3.87}$$

where Ψ is the solution to (3.11). Using Theorem 3.10 we have that the solution is bounded in $L^2(\Omega; \mathbb{C}^N)$. We have

$$\|\Psi(t)\|_{H^2} \leq c(\|A(t)\|_{L^2} + \|\Psi(t)\|_{L^2}) \leq c(\|F(t)\|_{L^2} + \|\Psi'(t)\|_{L^2} + \|\Psi(t)\|_{L^2}), \tag{3.88}$$

where $A(t) = -F(t) + i\partial_t \Psi(t)$. Next, we integrate (3.88) from 0 to T, and use (3.37) and (3.86) to obtain

$$\|\Psi\|_{L^2(0,T;H^2(\Omega))}^2 \leq C(\|\Psi^0\|_{L^2}^2 + \|F\|_Y^2).$$

All together, we have shown the estimate (3.80). $\qquad\square$

Theorem 3.17. *If Assumptions 2, 9, and 10 hold, then for the solution of* (3.11), *we have*

$$\Psi' \in L^\infty(0,T; L^2(\Omega; \mathbb{C}^N)) \qquad and \qquad \Psi \in L^\infty(0,T; H^2(\Omega; \mathbb{C}^N)). \tag{3.89}$$

Recalling the Sobolev embedding, we also have $\Psi \in L^\infty(0,T; C(\overline{\Omega}; \mathbb{C}^N))$.

Proof. Take a fixed $m \geq 1$. Differentiate (3.29) with respect to t, multiply this equation with $\partial_t d_{k,j}^m(t)$, sum over k, j, and integrate over t to obtain

$$\int_0^t i\left\langle \partial_t^2 \Psi_m, \partial_t \Psi_m\right\rangle_{L^2} dt = \int_0^t B(\partial_t \Psi_m, \partial_t \Psi_m; u) + \left\langle V_u \frac{\partial}{\partial t}(u\Psi_m), \partial_t \Psi_m\right\rangle_{L^2}$$
$$+ \langle \partial_t F, \partial_t \Psi_m\rangle_{L^2} + \alpha\left\langle V_{Hxc}(\Psi_m)\partial_t \Psi_m + \frac{\partial V_{Hxc}(\Psi_m)}{\partial t}\Psi_m, \partial_t \Psi_m\right\rangle_{L^2} dt. \tag{3.90}$$

For the left-hand side, we have

$$i\left\langle \partial_t^2 \Psi_m, \partial_t \Psi_m\right\rangle_{L^2} = i\frac{1}{2}\frac{d}{dt}\|\partial_t \Psi_m\|_{L^2}^2. \tag{3.91}$$

We remark that for any $f(\Psi, x, t) \in \mathbb{R}$, we have

$$\langle f(\Psi, \cdot, t)\Psi(t), \partial_t \Psi(t)\rangle_{L^2} = \langle f(\Psi, \cdot, t), (\Psi, \partial_t \Psi)_{\mathbb{C}}\rangle_{L^2} = \left\langle f(\Psi, \cdot, t), \frac{1}{2}\frac{d}{dt}\|\Psi(t)\|_{\mathbb{C}}^2\right\rangle_{L^2} \in \mathbb{R}.$$

Hence, using this result for the product terms in (3.90), we get

$$\left\langle V_u \frac{\partial}{\partial t}(u\Psi_m(t)), \partial_t \Psi_m\right\rangle_{L^2} \in \mathbb{R} \quad \text{and} \quad \left\langle \frac{\partial}{\partial t}(V_{Hxc}(\Psi_m)\Psi_m), \partial_t \Psi_m\right\rangle_{L^2} \in \mathbb{R}. \tag{3.92}$$

Taking the imaginary part of (3.90) and using (3.91) and (3.92) gives us the following:

$$\frac{1}{2}(\|\partial_t \Psi_m(t)\|_{L^2}^2 - \|\partial_t \Psi_m(0)\|_{L^2}^2) = \int_0^t (1-\alpha)\operatorname{Im} D(\partial_t \Psi_m, \partial_t \Psi_m) + \operatorname{Im}\langle \partial_t F, \partial_t \Psi_m\rangle_{L^2} dt.$$

From this, using (3.18), we obtain the following:

$$
\begin{aligned}
\sup_{0 \le t \le T} \|\partial_t \Psi_m(t)\|_{L^2}^2 &\le \|\partial_t \Psi_m(0)\|_{L^2}^2 + 2 \int_0^T (1-\alpha)|\operatorname{Im}\langle \partial_t \Psi,\, \partial_t \Psi_m \rangle_{L^2}| \\
&\quad + |\operatorname{Im}\langle \partial_t F,\, \partial_t \Psi_m \rangle_{L^2}|\mathrm{d}t \\
&\le \|\partial_t \Psi_m(0)\|_{L^2}^2 + 2\int_0^T (1-\alpha)c_0\|\partial_t \Psi_m\|_{L^2}^2 + \|F\|_{L^2}^2 + \|\partial_t \Psi_m\|_{L^2}^2 \mathrm{d}t \\
&\le \|\partial_t \Psi_m(0)\|_{L^2}^2 + 2(c_0+1)\|\partial_t \Psi_m\|_Y^2 + 2\|F\|_Y^2.
\end{aligned}
$$

By (3.80), $\|\partial_t \Psi_m\|_Y$ is bounded by F and Ψ^0. Hence, there exists a constant c_6 depending only on T, $\|\Psi^0\|_{L^2}$, $\|F\|_Y$, and $\|u\|_{H^1(0,T)}$, such that the following holds:

$$
\sup_{0 \le t \le T} \|\partial_t \Psi_m(t)\|_{L^2}^2 \le \|\partial_t \Psi_m(0)\|_{L^2}^2 + c_6. \tag{3.93}
$$

To bound $\|\partial_t \Psi_m(0)\|_{L^2}^2$, we test (3.29) with $\partial_t \Psi_m(0)$ to obtain

$$
\begin{aligned}
i\langle \partial_t \Psi_m(0),\, \partial_t \Psi_m(0)\rangle_{L^2} &= B(\Psi_m(0), \partial_t \Psi_m(0); u) \\
&\quad + \alpha\langle V_{Hxc}(\Psi_m(0))\Psi_m(0),\, \partial_t \Psi_m(0)\rangle_{L^2} + \langle F(0),\, \partial_t \Psi_m(0)\rangle_{L^2},
\end{aligned}
$$

$$
\begin{aligned}
\|\partial_t \Psi_m(0)\|_{L^2}^2 &\le |B(\Psi_m(0), \partial_t \Psi_m(0); u)| + |\langle V_{Hxc}(\Psi_m(0))\Psi_m(0),\, \partial_t \Psi_m(0)\rangle_{L^2}| \\
&\quad + \|F(0)\|_{L^2}\|\partial_t \Psi_m(0)\|_{L^2} \\
&\le c_1'\|\Psi_m(0)\|_{H^2}\|\partial_t \Psi_m(0)\|_{L^2} + K'\|\Psi_m(0)\|_{L^2}\|\partial_t \Psi_m(0)\|_{L^2} \\
&\quad + \|F(0)\|_{L^2}\|\partial_t \Psi_m(0)\|_{L^2}.
\end{aligned} \tag{3.94}
$$

Here, we used (3.51) for the nonlinear potential and we use the modified proof of Lemma 3.6 by replacing $\langle \nabla\Psi, \nabla\Phi \rangle_{L^2}$ by $\langle \nabla^2\Psi, \Phi \rangle_{L^2}$ using integration by parts. Dividing by $\|\partial_t \Psi_m(0)\|_{L^2}$ gives

$$
\|\partial_t \Psi_m(0)\|_{L^2} \le (c_1' + K)\|\Psi_m(0)\|_{H^2} + \|F(0)\|_{L^2}. \tag{3.95}
$$

Furthermore, we have $\|\Psi_m(0)\|_{H^2} \le C\|\Psi^0\|_{H^2}$; see, e.g., [Eva10, p. 363]. Using this in (3.95) gives the following:

$$
\|\partial_t \Psi_m(0)\|_{L^2} \le (c_1' + 1)C\|\Psi^0\|_{H^2} + \|F(0)\|_{L^2}. \tag{3.96}
$$

Therefore, using (3.96) in (3.93), we obtain the following:

$$
\sup_{0 \le t \le T} \|\partial_t \Psi_m(t)\|_{L^2}^2 \le c_7\left(\|\Psi^0\|_{H^2}^2 + \|F(0)\|_{L^2}^2\right) + c_6.
$$

Taking the limit $m \to \infty$, we find $\Psi' \in L^\infty(0, T; L^2(\Omega))$.

Using this result in (3.88), we have that $\Psi \in L^\infty(0, T; H^2(\Omega))$. By the Sobolev embedding, we have $\Psi \in L^\infty(0, T; C(\overline{\Omega}))$ and there exists a constant c_8 depending on T, $\|\Psi^0\|_{L^2}$, $\|F\|_Y$, and $\|u\|_{H^1(0,T)}$ such that

$$
\operatorname*{ess\,sup}_{0 \le t \le T} \max_{x \in \Omega} |\Psi(x,t)| \le c_8. \tag{3.97}
$$

$\square$

Remark. By (3.97), the solution of (3.11) is everywhere and for almost all times bounded by a constant. Assumption 3 is, hence, a reasonable assumption as it holds for all solutions.

3.9. Conclusion

In this chapter, the existence, uniqueness, and improved regularity properties of solutions to the time-dependent Kohn-Sham (KS) equations and related adjoint equations were proved. We want to emphasize especially the results about higher regularity in Section 3.8 that are genuinely new. This work is instrumental for investigating optimal control problems governed by the KS equations as done in next chapter.

We summarize the main results of this chapter in the following theorem.

Theorem 3.18. *The weak formulation of (3.1), with $V_{ext} \in H^1(0, T; L^\infty(\Omega; \mathbb{R}))$, $\Psi(0) = \Psi^0 \in L^2(\Omega; \mathbb{C}^N)$, admits a unique solution in W, that is, there exists $\Psi \in X$, $\Psi' \in X^*$, such that*

$$i \langle \partial_t \Psi(t), \, \Phi \rangle_{L^2} = \langle \nabla \Psi(t), \, \nabla \Phi \rangle_{L^2} + \langle V_{ext}(\cdot, t, u)\Psi(t), \, \Phi \rangle_{L^2}$$
$$+ \langle V_{Hxc}(\Psi(t))\Psi(t), \, \Phi \rangle_{L^2} , \tag{3.98}$$

for all $\Phi \in H_0^1(\Omega; \mathbb{C}^N)$ and a.e. in $(0, T)$.
By the continuous embedding $W \hookrightarrow C([0, T]; L^2(\Omega; \mathbb{C}^N))$, see e.g. [Eva10, p. 287], the solution is continuous in time.

Furthermore, if $\Psi^0 \in H_0^1(\Omega; \mathbb{C}^N)$ and $\partial \Omega \in C^2$, then the unique solution to (3.98) is as follows:

$$\Psi \in L^2(0, T; H^2(\Omega; \mathbb{C}^N)) \cap L^\infty(0, T; H_0^1(\Omega; \mathbb{C}^N));$$

if in addition $\Psi^0 \in H^2(\Omega; \mathbb{C}^N)$, we have

$$\Psi \in L^\infty(0, T; H^2(\Omega; \mathbb{C}^N) \cap H_0^1(\Omega; \mathbb{C}^N)).$$

4. Optimal control of the TDKS model

4.1. Introduction

Many applications in quantum physics and chemistry require to control multi-electron systems in order to achieve a desired target configuration. Recently, various quantum mechanical optimal control problems governed by the SE have been studied in the literature, see for example [vWB08], [vWBV10] and [MST06].

We remark that the external potential V_{ext} modelling the interaction of the particles (in particular, electrons) with an external (electric) field enters without modification in both the multi-particle SE model and the TDKS model. This latter fact is important in the design of control strategies for multi-particle quantum systems because the control functions usually enter in the SE model as external time-varying potentials. Therefore, control mechanisms can be determined in the TDDFT framework that are also valid for the original multi-particle SE system.

Therefore, the challenging task of quantum control appears possible in the framework of time-dependent density functional theory (TDDFT). TDDFT allows to describe these systems while avoiding the high dimensionality resulting from the multi-particle Schrödinger equation. However, due to the high nonlinearity of the problem, a strong effort is required to deal with this system.

Indeed, quantum control problems governed by TDDFT models have already been investigated; see, e.g., [CWG12] and have been implemented in TDDFT codes as the well-known Octopus [CAO$^+$06]. However, the available optimization schemes are mainly based on less competitive Krotov's method and consider only finite-dimensional parametrized controls. Furthermore, not much is known about the theory of the TDDFT optimal control framework and about the use and analysis of more efficient optimization schemes that allow to compute control functions belonging to a much larger function space.

We remark that the functional analysis of optimization problems governed by the TDKS equations and the investigation of optimization schemes requires the mathematical foundation of the governing model. To the best of our knowledge, only few contributions addressing this issue are available; we refer to [RPvL15, Jer15, SCB17c] for results concerning the existence and the uniqueness of solutions to the TDKS equations. However, the only existence theorem in a framework suitable for control problems is given by Theorem 3.18; see also, [SCB17c].

To formulate our optimal control problems, we start by recalling the TDKS model as follows (see Section 2.2.2):

$$i\frac{\partial \psi_j}{\partial t}(x,t) = \left(-\nabla^2 + V_{ext}(x,t,u) + V_{Hxc}(x,\rho)\right)\psi_j(x,t), \tag{4.1}$$

$$\psi_j(x,0) = \psi_j^0(x), \qquad j = 1,\ldots,N. \tag{4.2}$$

Its weak formulation is given in Theorem 3.18: there exists a unique $\Psi \in W$, such that $\Psi(0) = \Psi^0$ and

$$i\left\langle \partial_t \Psi(t),\, \Phi \right\rangle_{L^2} = \left\langle \nabla \Psi(t),\, \nabla \Phi \right\rangle_{L^2} + \left\langle V_{ext}(\cdot,t,u)\Psi(t),\, \Phi \right\rangle_{L^2}$$
$$+ \left\langle V_{Hxc}(\Psi(t))\Psi(t),\, \Phi \right\rangle_{L^2} \tag{4.3}$$

53

for all $\Phi \in H_0^1(\Omega; \mathbb{C}^N)$ and a.e. in $(0, T)$. In this chapter, we choose $\Psi^0 \in H_0^1(\Omega; \mathbb{C}^N)$. The KS potential in the LDA framework is given by

$$V_{Hxc}(x, \rho(x, t)) = V_H(x, \rho(x, t)) + V_x(\rho(x, t)) + V_c(\rho(x, t)). \tag{4.4}$$

Notice that the TDKS system is formulated in n spatial dimensions and consists of N coupled Schrödinger equations. With the appropriate choice of V_{Hxc}, which contains the coupling through the dependence on ρ, the solution to (4.1)–(4.2) provides the correct density of the original system, so that all observables which can be formulated in terms of the density can be determined by this method.

There are several approaches in the TDDFT framework how to construct the KS potential. A common choice in DFT is the local density approximation (LDA) where the KS potential at x is assumed to depend only on $\rho(x)$; see, e.g., [ED11]. We use the adiabatic LDA, which means that LDA is applied at every time separately such that $V_{Hxc}(x, t) = V_{Hxc}(\rho(x, t))$. Notice that, if one allows $V_{Hxc}(t)$ to depend on the whole history $V_{Hxc}(\tau), 0 \leq \tau \leq t$, the resulting adjoint equation would be an integro-differential equation, which would be much more involved to solve.

It is clear that, in application, confined electron systems subject to external control are of paramount importance. The confinement is obtained considering external potentials such that Ψ is non-zero only on a bounded domain Ω. For this reason, we denote by V_0 a confining potential that may represent the attracting potential of the nuclei of a molecule or the walls of a quantum dot. A typical model is the harmonic oscillator potential, $V_0(x) = u_0 \, x^2$.

A control potential aims at steering the quantum system to change its configuration towards a target state or to optimize the value of a given observable. In most cases, this results in a change of energy that necessarily requires a time-dependent interaction of the electrons with an external electro-magnetic force. For this purpose, we introduce a control potential with the following structure $u(t) \, V_u(x)$, where $u(t)$ has the role of a modulating amplitude. A specific case is the dipole control potential, $u(t) \, x$.

In our TDKS system, we consider the following external potential:

$$V_{ext}(x, t, u) = u(t) V_u(x) + V_0(x).$$

In particular, we consider the control of a quantum dot by a changing gate voltage modelled by a variable quadratic potential, $V_u(x) = x^2$, and a laser control in dipole approximation, $V_u(x) = x \cdot p$, with a polarization vector p.

The notation used is the one introduced in Section 3.2. To improve readability of the analysis that follows, we write the potentials in (4.4) as functions of Ψ instead of $\rho = |\Psi|^2$. We shall also omit the explicit dependence of V_H on x if no confusion may arise.

Assumption. We require the Assumptions 1–11 to hold (see page 34). Furthermore, we add the following assumption for this chapter:

12) The correlation potential term $\Psi \mapsto V_c(\Psi)\Psi$ is continuously real-Fréchet differentiable (see Definition 4.1) with bounded derivative from Z to Y; c.f. Lemma 4.4 for the same result on the exchange potential.

In this chapter, we investigate the optimal control theory for multi-particle quantum systems presenting a theoretical analysis of optimal control problems governed by the TDKS equations. For this purpose, the theory and numerical solution of optimal control
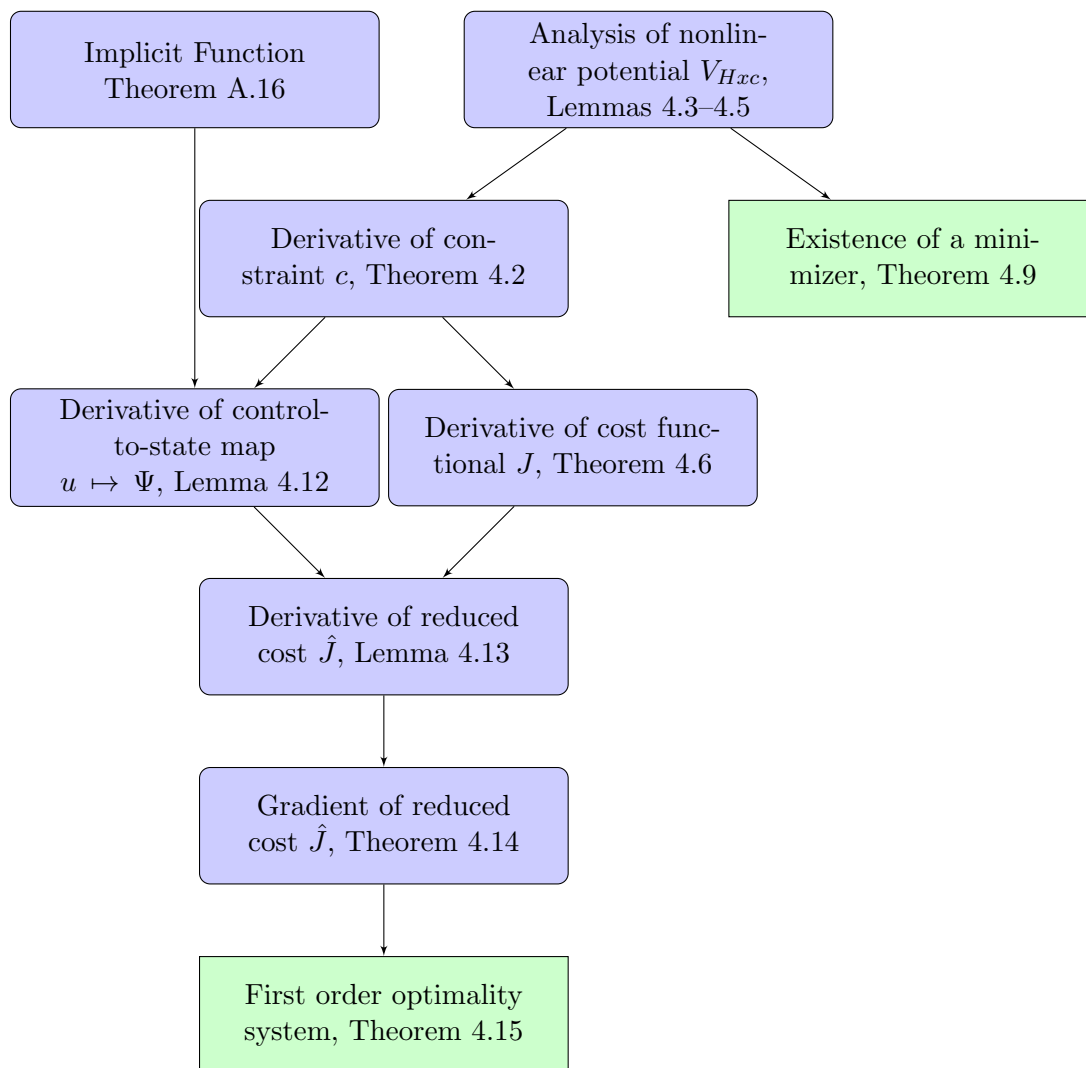
Figure 4.1.: Dependence of the lemmas and the theorems in Section 4.3. This chart can be used as a guide how to achieve the two main goals, namely the existence of minimizers and their characterization through the first order optimality system.

problems governed by a Kohn-Sham TDDFT model are investigated, considering different objectives and a bilinear control mechanism. Existence of optimal control solutions and their characterization as solutions to Kohn-Sham optimality systems are discussed.

The discussion below is organized as follows. We begin with the formulation of the optimization problem in Section 4.2. In Section 4.3.1, the nonlinear potential and the constraint equation are analysed. Section 4.3.2 is devoted to the proof of existence of a minimizer of the optimal control problem specified in (4.6). The first-order optimality system is discussed in Sections 4.3.3 and 4.4. In particular, in Section 4.3.3 we discuss and derive the optimality system considering the optimization problem in the reduce form (4.10). Figure 4.1 clarifies the dependences (work flow) of the lemmas and theorems in Section 4.3. It is given as a guide through the section and shows how the two main goals of this section, namely the existence of minimizers and their characterization using the first order optimality system, are achieved. In Section 4.4, the optimality system is derived (again for the sake of clarity) by differentiating the Lagrange function.

We remark, that in Appendix A.1, the differentiability issues arising from the Kohn-Sham potential are discussed in more detail and an alternative approach to the one considered in Section 4.3 is presented. In order to validate the proposed framework, we implement an efficient approximation and optimization scheme for these problems in Chapter 5.

4.2. Formulation of TDKS optimal control problems

Optimal control of quantum systems is of fundamental importance in quantum mechanics applications. The objectives of the control may be of different nature ranging from the breaking of a chemical bond in a molecule by an optimally shaped laser pulse to the manipulation of electrons in two-dimensional quantum dots by a gate voltage. In this framework, the objective of the control is modelled by a cost functional to be optimized under the differential constraints represented by the quantum model (in our case the TDKS equations) including the control mechanism.

We consider an objective J that includes different target functionals and a control cost as follows:

$$J(\Psi, u) = \underbrace{\frac{\beta}{2} \int_0^T \int_\Omega (\rho(x,t) - \rho_d(x,t))^2 \mathrm{d}x\mathrm{d}t}_{J_\beta} + \underbrace{\frac{\eta}{2} \int_\Omega \chi_A(x)\rho(x,T)\mathrm{d}x}_{J_\eta} + \underbrace{\frac{\nu}{2}\|u\|_{H^1}^2}_{J_\nu}. \quad (4.5)$$

The first term J_β models the requirement that the electron density evolves following as close as possible a given target trajectory ρ_d. We remark that J_β is only well-defined if Ψ is at least in $L^4(0,T;L^4(\Omega;\mathbb{C}^N))$. This is guaranteed by the improved regularity from Theorem 3.17 for $\Psi^0 \in H^2(\Omega;\mathbb{C}^N)$. The term J_η aims at locating the density outside of a certain region A. The characteristic function of A is given by $\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases}$.
The term J_ν penalizes the cost of the control. We remark that the regularization term J_ν can be any weighted $H^1(0,T;\mathbb{R})$ norm, e.g. $\|u'\|_{L^2(0,T;\mathbb{R})}^2 + a\|u\|_{L^2(0,T;\mathbb{R})}^2$ for $a > 0$. We assume that the target weights are all non-negative $\beta, \eta \geq 0$, with $\beta + \eta > 0$, and the regularization weight $\nu > 0$.

As in many applications, our purpose is to find an optimal control function u, which modulates a dipole or a quadratic potential, such that $J(\Psi, u)$ is minimized subject to the constraint that Ψ satisfies the TDKS equations. This problem is formulated as follows:

$$\min_{(\Psi,u)\in(Z,U)} J(\Psi, u) \qquad \text{subject to (4.3)}, \quad (4.6)$$

where U is either $H^1(0,T;\mathbb{R})$ or $H_0^1(0,T;\mathbb{R})$. From now on, we assume that, if $\beta = 0$, then $\Psi^0 \in H_0^1(\Omega;\mathbb{C}^N)$, and if $\beta > 0$, then $\Psi^0 \in H_0^1(\Omega;\mathbb{C}^N) \cap H^2(\Omega;\mathbb{C}^N)$.

The solutions to this PDE-constrained optimization problem are characterized as solutions to the corresponding first-order optimality conditions [BS12, Trö10]. These conditions for (4.6) can be formally obtained by setting to zero the gradient of the following Lagrange function:

$$L(\Psi, u, \Lambda) = J(\Psi, u) + L_1(\Psi, u, \Lambda), \qquad \text{where} \quad (4.7)$$

$$L_1 = \mathrm{Re}\sum_{j=1}^N \int_0^T \int_\Omega \left(i\frac{\partial \psi_j(x,t)}{\partial t} - \left(-\nabla^2 + V_{ext}(x,t,u) + V_{Hxc}(x,t,\rho)\right)\psi_j(x,t)\right)\overline{\lambda_j(x,t)}\mathrm{d}x\mathrm{d}t.$$

The function $\Lambda = (\lambda_1, \dots, \lambda_N)$, where $\lambda_j \in L^2(0, T; H_0^1(\Omega; \mathbb{C}))$, $j = 1, \dots, N$, represents the adjoint variable. In the Lagrange formalism, a solution to (4.6) corresponds to a stationary point of L, where the derivatives of L with respect to ψ_j, λ_j and u must be zero along any directions $\delta\psi$, $\delta\lambda$, and δu. A detailed calculation of these derivatives can be found in Section 4.4. The main difficulty in the derivation is the complex and non-analytic dependence of the Kohn-Sham potential on the wavefunction, which results in the terms $\nabla_\psi L_H$, $\nabla_\psi L_{xc}$ in (4.8c) below.

The first-order optimality conditions are given by the following optimality system:

$$i\frac{\partial \psi_m(x,t)}{\partial t} = \left(-\nabla^2 + V_{ext}(x,t,u) + V_{Hxc}(x,t,\rho)\right)\psi_m(x,t), \tag{4.8a}$$

$$\psi_m(x,0) = \psi_m^0(x), \tag{4.8b}$$

$$i\frac{\partial \lambda_m(x,t)}{\partial t} = \left(-\nabla^2 + V_{ext}(x,t,u)\right)\lambda_m(x,t) + \nabla_\psi L_H + \nabla_\psi L_{xc} - 2\beta(\rho - \rho_d)\psi_m(x,t), \tag{4.8c}$$

$$\lambda_m(x,T) = -i\eta\chi_A(x)\psi_m(x,T), \tag{4.8d}$$

$$\nu u(t) + \mu(t) = 0, \tag{4.8e}$$

where $m = 1, \dots, N$, and

$$\nabla_\psi L_H := \sum_{j=1}^N V_H\left(2\operatorname{Re}(\psi_j(y,t)\overline{\lambda_j(y,t)})\right)(x,t)\psi_m(x,t) + V_H(\rho)(x,t)\lambda_m(x,t),$$

$$\nabla_\psi L_{xc} := 2\sum_{j=1}^N \frac{\partial V_{xc}}{\partial \rho}(\rho(x,t))\psi_m(x,t)\operatorname{Re}(\psi_j(x,t)\overline{\lambda_j(x,t)}) + V_{xc}(\rho(x,t))\lambda_m(x,t).$$

Further, in (4.8e), μ is the H^1-Riesz representative of the continuous linear functional $\langle -\operatorname{Re}\langle \Lambda, V_u\Psi\rangle_{L^2}, \cdot\rangle_{L^2(0,T)}$; see Theorem 4.14 below. Assuming that $u \in H_0^1(0,T;\mathbb{R})$, μ can be computed by solving the equation

$$\left(-\frac{d^2}{dt^2} + 1\right)\mu = -\operatorname{Re}\langle \Lambda, V_u\Psi\rangle_{L^2}, \quad \mu(0) = 0, \mu(T) = 0, \tag{4.9}$$

which is understood in a weak sense. For more details, see, e.g., [vWB08].

Theorem 3.13 guarantees that (4.8a)–(4.8b) is uniquely solvable and hence the reduced cost functional

$$\hat{J} : U \to \mathbb{R}, \quad \hat{J}(u) := J(\Psi(u), u), \tag{4.10}$$

is well-defined.

In order to ensure the correct regularity properties of the adjoint variables, we do not use the Lagrange formalism explicitly in this work. Instead, we use the existence of a unique solution of the adjoint equation (4.8c)–(4.8d) from Theorem 3.18 directly, and derive the gradient of the reduced cost functional $\hat{J}$ from these solutions in Theorem 4.14 below. Later, we use this gradient to construct a numerical optimization scheme to minimize $\hat{J}$.

4.3. Theoretical analysis of TDKS optimal control problems

4.3.1. Analysis of the constraint equation

In this section, we present a mathematical analysis of the optimal control problem (4.6). To this end, we first show that both the constraint given by the TDKS equations and the cost functional J are continuously real-Fréchet differentiable.

The Kohn-Sham potential depends on the density ρ. The density is a real-valued function of the complex wavefunction and, therefore, can not be holomorphic. As complex differentiability is a stronger property than what we need in the following, we introduce the following weaker notion of real-differentiability. This is discussed in more detail in Section A.1 where also an alternative approach is presented.

Definition 4.1. *Let V_1, V_2 be complex Banach spaces. A map $f : V_1 \to V_2$ is called* real-linear *if and only if*

1. *$f(x) + f(y) = f(x + y) \; \forall x, y \in V_1$ and*

2. *$f(\alpha x) = \alpha f(x) \; \forall \alpha \in \mathbb{R}$ and $\forall x \in V_1$.*

The space of bounded real-linear maps from V_1 to V_2 is denoted with $\mathfrak{L}(V_1; V_2)$. This is a Banach space, the norm is the same as for $\mathcal{L}(V_1; V_2)$.

We call a map $f : V_1 \to V_2$ real-Gâteaux (real-Fréchet) *differentiable if the standard definition of Gâteaux (Fréchet) differentiability holds for a real-linear derivative operator.*

Remark. In complex spaces, the notion of real-Gâteaux (real-Fréchet) differentiability is weaker than Gâteaux (Fréchet) differentiability. However, all theorems for differentiable functions in $\mathbb{R}^2$ also hold in $\mathbb{C}$ for functions that are just real-differentiable. This is the case for all theorems that we will make use of as the chain rule and the implicit function theorem. Therefore, it is enough to show real-Fréchet differentiability of the constraint.

Recall the definition of the spaces Y, X, X^*, W, and Z from Definition 3.1.

Theorem 4.2. *The map $c : Z \times U \to X^*$, defined as*

$$c(\Psi, u) := \tilde{c}(\Psi, u) - V_{Hxc}(\Psi)\Psi, \; \text{where} \; \tilde{c}(\Psi, u) := i\frac{\partial \Psi}{\partial t} - \left(-\nabla^2 + V_0 + V_u u(t)\right)\Psi,$$
(4.11)

is continuously real-Fréchet differentiable.

Notice that $c(\Psi, u) = 0$ represents the TDKS equations.

Remark. We remark that $V_x(\Psi)\Psi$, $V_H(\Psi)\Psi$, and $V_u u\Psi$ are in Y. With the Gelfand triple $X \hookrightarrow Y \hookrightarrow X^*$, the operator norm of their derivatives in $\mathcal{L}(W, X^*)$ can be bounded in $\mathcal{L}(W, Y)$ as follows:

$$\|B(\Psi)\|_{\mathcal{L}(W,X^*)} = \sup_{\delta\Psi \in W \setminus \{0\}} \frac{\|B(\Psi)\delta\Psi\|_{X^*}}{\|\delta\Psi\|_W} \leq c \sup_{\delta\Psi \in W \setminus \{0\}} \frac{\|B(\Psi)\delta\Psi\|_Y}{\|\delta\Psi\|_W}$$

$$\leq c\|B(\Psi)\|_{\mathcal{L}(W,Y)} \sup_{\delta\Psi \in W \setminus \{0\}} \frac{\|\delta\Psi\|_Y}{\|\delta\Psi\|_W} \leq c\,\|B(\Psi)\|_{\mathcal{L}(W,Y)},$$

where B denotes the derivatives of $V_x(\Psi)\Psi$, $V_H(\Psi)\Psi$, or $V_u u\Psi$, respectively.

Further, notice that using the embedding $Z \hookrightarrow W$, it follows that a function f that is real-Fréchet differentiability from W to Y at $a \in Z$ is also real-Fréchet differentiable from Z to Y at a as follows:

$$\lim_{\|h\|_Z \to 0} \frac{\|f(a+h) - f(a) - Df(a)(h)\|_Y}{\|h\|_Z} \leq \lim_{\|h\|_Z \to 0} \frac{\|f(a+h) - f(a) - Df(a)(h)\|_Y}{\|h\|_W} = 0.$$

Further, to prove Theorem 4.2, we need the following lemmas. We begin with studying the nonlinear exchange potential term.

Lemma 4.3. *The map $W \ni \Psi \mapsto V_x(\Psi)\Psi \in Y$ is real-Gâteaux differentiable for all $\Psi \in W$ and its real-Gâteaux derivative, denoted by $A(\Psi)$, is specified as follows:*

$$A(\Psi) \in \mathcal{L}(W, Y), \qquad A(\Psi)(\delta\Psi) = A_1(\Psi)(\delta\Psi) + A_2(\Psi)(\delta\Psi),$$

$$A_1(\Psi)(\delta\Psi) := V_x(\Psi)\delta\Psi, \qquad A_2(\Psi)(\delta\Psi) := \frac{\partial V_x}{\partial \rho} 2 \operatorname{Re}(\Psi, \delta\Psi)_{\mathbb{C}} \Psi.$$

Proof. First, we need the derivative of the density ρ which is a non-holomorphic function of a complex variable. Differentiation of ρ can be done using the Wirtinger calculus [Rem91] where Ψ and $\overline{\Psi}$ are treated as independent variables. By using these calculus rules, the real-Fréchet derivatives of $\rho = \sum_{m=1}^{M} \psi_m \overline{\psi_m} = (\Psi, \Psi)_{\mathbb{C}}$ are given by

$$\frac{\partial \rho}{\partial \Psi}(\delta\Psi) = 2 \operatorname{Re}(\Psi, \delta\Psi)_{\mathbb{C}},$$

$$\frac{\partial^2 \rho}{\partial \Psi^2}(\delta\Psi, \delta\Phi) = 2 \operatorname{Re}(\delta\Psi, \delta\Phi)_{\mathbb{C}}.$$

The directional derivative of $V_x(\Psi)\Psi$ along $\delta\Psi$ is given by $A(\Psi) : W \to X^*$,

$$A(\Psi)(\delta\Psi) = \lim_{t \searrow 0} \frac{1}{t} \left(V_x(\Psi + t\delta\Psi)(\Psi + t\delta\Psi) - V_x(\Psi)\Psi \right)$$

$$= \lim_{t \searrow 0} \left(V_x(\Psi)\delta\Psi + \frac{\partial V_x}{\partial \rho}\frac{\partial \rho}{\partial \Psi}(\delta\Psi)\Psi + \mathcal{O}(t) \right)$$

$$= A_1(\Psi)(\delta\Psi) + A_2(\Psi)(\delta\Psi).$$

Using the definition of V_x in (3.8) on page 32, we have

$$A(\Psi)(\delta\Psi) = V_x(\Psi)\delta\Psi + \alpha_n \begin{cases} \frac{|\Psi|^{2/n-2}}{n} 2 \operatorname{Re}(\Psi, \delta\Psi)_{\mathbb{C}} \Psi & |\Psi|^2 \leq R, \\ \frac{\partial p}{\partial \rho}(|\Psi|^2) 2 \operatorname{Re}(\Psi, \delta\Psi)_{\mathbb{C}} \Psi & R < |\Psi|^2 < 2R, \\ 0 & |\Psi|^2 \geq 2R \end{cases}$$

and $\delta\Psi \mapsto A(\Psi)(\delta\Psi)$ is obviously real-linear in $\delta\Psi$ and $A(\Psi)(\delta\Psi) \in Y$ for $\delta\Psi \in W$. We are left to show that $A(\Psi)$ is a bounded operator.

$$\|A(\Psi)\|_{\mathcal{L}(W,Y)} \leq \sup_{\delta\Psi \in W \setminus \{0\}} \frac{\|A_1(\Psi)\delta\Psi\|_Y}{\|\delta\Psi\|_W} + \sup_{\delta\Psi \in W \setminus \{0\}} \frac{\|A_2(\Psi)(\delta\Psi)\|_Y}{\|\delta\Psi\|_W}.$$

For the first term, we use $V_x(\Psi) \in L^\infty(0, T; L^\infty(\Omega))$ as well as $\|\Psi\|_Y \leq \|\Psi\|_W$ to obtain

$$\sup_{\delta\Psi \in W \setminus \{0\}} \frac{\|A_1(\Psi)\delta\Psi\|_Y}{\|\delta\Psi\|_W} \leq \sup_{\delta\Psi \in W \setminus \{0\}} \frac{\|V_x(\Psi)\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}))}\|\delta\Psi\|_Y}{\|\delta\Psi\|_W}$$

$$\leq \|V_x(\Psi)\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}))}.$$

For the second term, we decompose the domain into $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3$ depending on the size of ρ: $\Omega_1 := \{x \in \Omega \,|\, |\Psi|^2 \leq R\}$, $\Omega_2 := \{x \in \Omega \,|\, R < |\Psi|^2 < 2R\}$, and $\Omega_3 := \{x \in \Omega \,|\, |\Psi|^2 \geq 2R\}$. Using the fact that $|\Psi|^2 = \rho$ is bounded by R in Ω_1 and by $2R$ in Ω_2 as well as that $A_2(\Psi) = 0$ in Ω_3, and that $\frac{\partial p}{\partial \rho}$ is monotonically decreasing

between R and $2R$, we obtain

$$
\begin{aligned}
\|A_2(\Psi)(\delta\Psi)\|_Y^2 &= \int_0^T \int_{\Omega_1} |A_2(\Psi)(\delta\Psi)|^2 \mathrm{d}x\mathrm{d}t + \int_0^T \int_{\Omega_2} |A_2(\Psi)(\delta\Psi)|^2 \mathrm{d}x\mathrm{d}t \\
&= \int_0^T \int_{\Omega_1} \left| \frac{\alpha_n}{n}|\Psi|^{2/n-2} 2\operatorname{Re}(\Psi,\,\delta\Psi)_{\mathbb{C}}\,\Psi \right|^2 \mathrm{d}x\mathrm{d}t \\
&\quad + \int_0^T \int_{\Omega_2} \left| \alpha_n \frac{\partial p}{\partial \rho}(\rho) 2\operatorname{Re}(\Psi,\,\delta\Psi)_{\mathbb{C}}\,\Psi \right|^2 \mathrm{d}x\mathrm{d}t \\
&\leq \int_0^T \int_{\Omega_1} 4\frac{\alpha_n^2}{n^2}|\Psi|^{4/n}|\delta\Psi|^2 \mathrm{d}x\mathrm{d}t + 4\alpha_n^2 \left(\frac{\partial p}{\partial \rho}(R) \right)^2 \int_0^T \int_{\Omega_2} |\Psi|^4 |\delta\Psi|^2 \mathrm{d}x\mathrm{d}t \\
&\leq 4\frac{\alpha_n^2}{n^2} \left(R^{2/n} + R^{2/n-2} \right) \|\delta\Psi\|_Y^2.
\end{aligned}
$$

We have shown that the directional derivative $\delta\Psi \mapsto A(\Psi)(\delta\Psi)$ is a bounded linear map for all $\Psi \in W$, hence $\Psi \mapsto V_x(\Psi)\Psi$ is real-Gâteaux differentiable. $\qquad\square$

Lemma 4.4. *The map $Z \ni \Psi \mapsto V_x(\Psi)\Psi \in Y$ is continuously real-Fréchet differentiable with derivative $\mathrm{D}(V_x(\Psi)\Psi) = A(\Psi)$, where $A(\Psi) \in \mathfrak{L}(Z,Y)$ is given in Lemma 4.3.*

Proof. We prove that the real-Gâteaux derivative $A(\Psi)$ of $V_x(\Psi)\Psi$ at Ψ is continuous from Z to $\mathfrak{L}(Z,Y)$. Then the real-Fréchet differentiability follows immediately from [AH11, Proposition A.3].

Once it is proved that $V_x(\Psi)\Psi$ is real-Fréchet differentiable, the real-Gâteaux and the real-Fréchet derivatives coincide, and the continuity of the real-Gâteaux derivative carries over to the real-Fréchet derivative. Hence, we have to show the following:

$$
\forall \epsilon > 0,\ \exists \delta > 0,\ \text{such that } \|A(\Psi) - A(\Phi)\|_{\mathcal{L}(Z,Y)} < \epsilon,\ \forall \|\Psi - \Phi\|_Z < \delta.
$$

For this purpose, to ease our discussion, we consider $\tilde{A}_2(\Psi)\psi_j\psi_m$, $j, m = 1, \ldots, N$, where

$$
\tilde{A}_2(\Psi) = \alpha_n
\begin{cases}
\frac{|\Psi|^{2/n-2}}{n} & |\Psi|^2 \leq R, \\
\frac{\partial p}{\partial \rho}(|\Psi|^2) & R < |\Psi|^2 < 2R, \\
0 & |\Psi|^2 \geq 2R,
\end{cases}
$$

such that $\tilde{A}_2(\Psi)\psi_m \sum_{j=1}^N 2\operatorname{Re}(\psi_j\overline{\delta\psi_j}) = \big(A_2(\Psi)(\delta\Psi)\big)_m$. For all $\epsilon > 0$, $\Psi \mapsto \tilde{A}_2(\Psi)$ is continuously real differentiable with bounded derivative for $|\Psi| \geq \epsilon$, hence the same holds for $\tilde{A}_2(\Psi)\psi_j\psi_m$, which is therefore Lipschitz continuous. In zero, $\tilde{A}_2(\Psi)\psi_j\psi_m$ is Hölder continuous with exponent $\alpha = \frac{2}{n}$ as the following calculation shows:

$$
|\tilde{A}_2(\Psi)\psi_j\psi_m| = \frac{\alpha_n}{n}|\Psi|^{2/n-2}|\psi_j\psi_m| \leq \frac{\alpha_n}{n}|\Psi|^{2/n-2}|\Psi||\Psi| = \frac{\alpha_n}{n}|\Psi|^{2/n}.
$$

Together, we have Hölder continuity for all $\Psi \in Z$,

$$
|\tilde{A}_2(\Psi)\psi_j\psi_m - \tilde{A}_2(\Phi)\phi_j\phi_m| \leq c|\Psi - \Phi|_{\mathbb{C}^N}^\alpha
$$

and similarly

$$
|\tilde{A}_2(\Psi)\overline{\psi_j}\psi_m - \tilde{A}_2(\Phi)\overline{\phi_j}\phi_m| \leq c|\Psi - \Phi|_{\mathbb{C}^N}^\alpha.
$$

We remark that, if a function $f : \mathbb{C}^N \to \mathbb{C}^N$ is Hölder continuous with exponent α, i.e. $|f(x) - f(y)| < c|x - y|^\alpha$, then $f : L^2(\Omega; \mathbb{C}^N) \to L^2(\Omega; \mathbb{C}^N)$ is also Hölder continuous with the same exponent as follows:

$$
\begin{aligned}
\|f(x) - f(y)\|_{L^2(\Omega;\mathbb{C}^N)}^2 &= \int_\Omega |f(x(\xi)) - f(y(\xi))|^2 \mathrm{d}\xi \leq c^2 \int_\Omega |x(\xi) - y(\xi)|^{2\alpha} \mathrm{d}\xi \\
&= c^2 \|x - y\|_{L^{2\alpha}(\Omega;\mathbb{C}^N)}^{2\alpha} \leq c'^2 \|x - y\|_{L^2(\Omega;\mathbb{C}^N)}^{2\alpha}.
\end{aligned}
\tag{4.12}
$$

Define $Z_1 := L^\infty(0, T; H^2(\Omega; \mathbb{C}))$. Now, we turn our attention to $A_2(\Psi)$ and use the Hölder continuity of $\tilde{A}_2(\Psi)\psi_j\psi_m$ to obtain the following estimate. We have

$$
\|A_2(\Psi)(\delta\Psi) - A_2(\Phi)(\delta\Psi)\|_Y^2
$$

$$
= \sum_{m=1}^N \int_0^T \int_\Omega \left| \tilde{A}_2(\Psi)\psi_m \sum_{j=1}^N (\psi_j \overline{\delta\psi_j} + \overline{\psi_j}\delta\psi_j) - \tilde{A}_2(\Phi)\phi_m \sum_{j=1}^N (\phi_j\overline{\delta\psi_j} + \overline{\phi_j}\delta\psi_j) \right|^2 \mathrm{d}x\mathrm{d}t
$$

$$
\leq \sum_{m=1}^N \int_0^T \int_\Omega \sum_{j=1}^N \left| \overline{\delta\psi_j}(\tilde{A}_2(\Psi)\psi_j\psi_m - \tilde{A}_2(\Phi)\phi_j\phi_m) \right|^2 \mathrm{d}x\mathrm{d}t
$$

$$
+ \sum_{m=1}^N \int_0^T \int_\Omega \sum_{j=1}^N \left| \delta\psi_j(\tilde{A}_2(\Psi)\overline{\psi_j}\psi_m - \tilde{A}_2(\Phi)\overline{\phi_j}\phi_m) \right|^2 \mathrm{d}x\mathrm{d}t
$$

$$
\leq c \sum_{j=1}^N \|\delta\psi_j\|_{Z_1}^2 \sum_{m=1}^N \left(\|\tilde{A}_2(\Psi)\psi_j\psi_m - \tilde{A}_2(\Phi)\phi_j\phi_m)\|_{Y_1}^2 + \|\tilde{A}_2(\Psi)\overline{\psi_j}\psi_m - \tilde{A}_2(\Phi)\overline{\phi_j}\phi_m\|_{Y_1}^2 \right)
$$

$$
\leq c_1 \sum_{j=1}^N \|\delta\psi_j\|_{Z_1}^2 \sum_{m=1}^N \|\Psi - \Phi\|_Y^{2\alpha}
$$

$$
= c' \|\delta\Psi\|_{Z_\star}^2 N \|\Psi - \Phi\|_Y^{2\alpha} < c' \|\delta\Psi\|_{Z_\star}^2 N \delta^{2\alpha}.
$$

Furthermore, by the Hölder continuity of V_x and the embedding $Z \hookrightarrow Y$, we have

$$
\begin{aligned}
\|A_1(\Psi)(\delta\Psi) - A_1(\Phi)(\delta\Psi)\|_Y &\leq c_2 \|A_1(\Psi) - A_1(\Phi)\|_{Y_1} \|\delta\Psi\|_{Z_\star} \\
&\leq c_3 \|\Psi - \Phi\|_Y^{2/n} \|\delta\Psi\|_{Z_\star} \leq c'' \delta^\alpha \|\delta\Psi\|_{Z_\star}.
\end{aligned}
$$

Now, we have the following:

$$
\begin{aligned}
\|A(\Psi) - A(\Phi)\|_{\mathcal{L}(Z,Y)} &= \sup_{\delta\Psi \in Z \setminus \{0\}} \frac{\|A(\Psi)(\delta\Psi) - A(\Phi)(\delta\Psi)\|_Y}{\|\delta\Psi\|_Z} \\
&\leq (\sqrt{c'N} + c'')\delta^\alpha \sup_{\delta\Psi \in Z \setminus \{0\}} \frac{\|\delta\Psi\|_{Z_\star}}{\|\delta\Psi\|_Z} \leq (\sqrt{c'N} + c'')\delta^\alpha =: \epsilon.
\end{aligned}
$$

This completes the proof of the continuous real-Fréchet differentiability of $V_x(\Psi)\Psi$. $\qquad\square$

Next, we study the Hartree potential and show that the map $\Psi \mapsto V_H(\Psi)\Psi$ is continuously real-Fréchet differentiable.

Lemma 4.5. *The map $\Psi \mapsto V_H(\Psi)\Psi$ is continuously real-Fréchet differentiable from Z to Y and from Z to X^* with derivative $\mathrm{D}(V_H(\Psi)\Psi)(\delta\Psi) = V_H(\Psi)\delta\Psi + \int_\Omega \frac{2\operatorname{Re}(\Psi, \delta\Psi)_\mathbb{C}}{|x-y|}\mathrm{d}y\Psi$.*

Proof. By Lemma 3.4, $V_H(\Psi)\Psi \in Y$. The following expansion holds:

$$
V_H(\Psi + \delta\Psi) = V_H(\Psi) + \int_\Omega \frac{2\operatorname{Re}(\Psi, \delta\Psi)_\mathbb{C}}{|x-y|}\mathrm{d}y + \int_\Omega \frac{(\delta\Psi, \delta\Psi)_\mathbb{C}}{|x-y|}\mathrm{d}y.
$$

Hence, we get

$$V_H(\Psi + \delta\Psi)(\Psi + \delta\Psi) = V_H(\Psi)\Psi + V_H(\Psi)\delta\Psi + \int_\Omega \frac{2\operatorname{Re}(\Psi, \delta\Psi)_\mathbb{C}}{|x-y|} \mathrm{d}y \Psi$$
$$+ \int_\Omega \frac{2\operatorname{Re}(\Psi, \delta\Psi)_\mathbb{C}}{|x-y|} \mathrm{d}y \delta\Psi + \int_\Omega \frac{(\delta\Psi, \delta\Psi)_\mathbb{C}}{|x-y|} \mathrm{d}y \Psi + \int_\Omega \frac{(\delta\Psi, \delta\Psi)_\mathbb{C}}{|x-y|} \mathrm{d}y \delta\Psi.$$

By proof of Lemma 3.4, the $L^2(\Omega; \mathbb{C}^N)$ norms of three terms in the previous line are a.e. in $[0, T]$ bounded by

$$C_N \|\delta\Psi(t)\|_{L^2(\Omega;\mathbb{C}^N)} \|\Psi(t)\|_{H^1(\Omega;\mathbb{C}^N)} \|\delta\Psi(t)\|_{L^2(\Omega;\mathbb{C}^N)},$$
$$C_N \|\delta\Psi(t)\|_{L^2(\Omega;\mathbb{C}^N)} \|\delta\Psi(t)\|_{H^1(\Omega;\mathbb{C}^N)} \|\Psi(t)\|_{L^2(\Omega;\mathbb{C}^N)},$$
$$C_N \|\delta\Psi(t)\|_{H^1(\Omega;\mathbb{C}^N)} \|\delta\Psi(t)\|^2_{L^2(\Omega;\mathbb{C}^N)},$$

respectively. Defining $\mathrm{D}(V_H(\Psi)\Psi)(\delta\Psi) := V_H(\Psi)\delta\Psi + \int_\Omega \frac{2\operatorname{Re}(\Psi, \delta\Psi)_\mathbb{C}}{|x-y|} \mathrm{d}y \Psi$, and using the embeddings $Y \hookrightarrow X^*$, $W \hookrightarrow C([0,T]; L^2(\Omega; \mathbb{C}^N))$, and $Z \hookrightarrow W$, we obtain

$$\frac{\|V_H(\Psi + \delta\Psi)(\Psi + \delta\Psi) - V_H(\Psi)\Psi - \mathrm{D}(V_H(\Psi)\Psi)(\delta\Psi)\|^2_{X^*}}{\|\delta\Psi\|^2_Z}$$

$$\leq c_1 \frac{\|V_H(\Psi + \delta\Psi)(\Psi + \delta\Psi) - V_H(\Psi)\Psi - \mathrm{D}(V_H(\Psi)\Psi)(\delta\Psi)\|^2_Y}{\|\delta\Psi\|^2_Z}$$

$$\leq \frac{c_1 C_N}{\|\delta\Psi\|^2_Z} \left(\int_0^T \|\delta\Psi\|^2_{L^2(\Omega;\mathbb{C}^N)} \|\Psi\|^2_{H^1(\Omega;\mathbb{C}^N)} \|\delta\Psi\|^2_{L^2(\Omega;\mathbb{C}^N)} \right.$$

$$\left. + \|\delta\Psi\|^2_{L^2(\Omega;\mathbb{C}^N)} \|\delta\Psi\|^2_{H^1(\Omega;\mathbb{C}^N)} \|\Psi\|^2_{L^2(\Omega;\mathbb{C}^N)} + \|\delta\Psi\|^2_{H^1(\Omega;\mathbb{C}^N)} \|\delta\Psi\|^4_{L^2(\Omega;\mathbb{C}^N)} \mathrm{d}t \right)$$

$$\leq \frac{c_1 C_N}{\|\delta\Psi\|^2_Z} \left(\max_{0 \leq t \leq T} \|\delta\Psi\|^4_{L^2(\Omega;\mathbb{C}^N)} \left(\|\Psi\|^2_X + \|\delta\Psi\|^2_X \right) \right.$$

$$\left. + \max_{0 \leq t \leq T} \|\delta\Psi\|^2_{L^2(\Omega;\mathbb{C}^N)} \max_{0 \leq t \leq T} \|\Psi\|^2_{L^2(\Omega;\mathbb{C}^N)} \|\delta\Psi\|^2_X \right)$$

$$\leq c_2 C_N \frac{\|\delta\Psi\|^4_Z \left(\|\Psi\|^2_X + \|\delta\Psi\|^2_X \right) + \|\delta\Psi\|^2_Z \|\Psi\|^2_W \|\delta\Psi\|^2_X}{\|\delta\Psi\|^2_Z}$$

$$\leq c_2 C_N \left(\|\delta\Psi\|^2_Z \left(\|\Psi\|^2_X + \|\delta\Psi\|^2_X \right) + \|\delta\Psi\|^2_Z \|\Psi\|^2_W \right) \to 0 \text{ for } \|\delta\Psi\|_Z \to 0.$$

Hence, $V_H(\Psi)\Psi$ is real-Fréchet differentiable with derivative $V_H(\Psi)\delta\Psi + \int_\Omega \frac{2\operatorname{Re}(\Psi, \delta\Psi)_\mathbb{C}}{|x-y|} \mathrm{d}y \Psi$ and the derivative is continuous from Z to $\mathfrak{L}(Z, Y)$ and from Z to $\mathfrak{L}(Z, X^*)$ by (3.15) in Lemma 3.4. □

Using these results, we can now prove Theorem 4.2 that states the real-Fréchet differentiability of the map c.

Proof of Theorem 4.2. We prove that c is a real-Fréchet differentiable function of Ψ and u. For this purpose, we first consider the map $\tilde{c}$ defined in (4.11). Simple algebraic manipulation results in the following:

$$\tilde{c}(\Psi + \delta\Psi, u + \delta u) = \tilde{c}(\Psi, u) + \tilde{c}(\delta\Psi, u) - V_u \delta u \Psi - V_u \delta u \delta\Psi.$$

Next, we show that the real-Fréchet derivative of $\tilde{c}$ is given by

$$\mathrm{D}\,\tilde{c}(\Psi, u)(\delta\Psi, \delta u) = \tilde{c}(\delta\Psi, u) - V_u \delta u \Psi = i \frac{\partial \delta\Psi}{\partial t} - \left(-\nabla^2 + V_0 + V_u u \right) \delta\Psi - V_u \delta u \Psi.$$

$$(4.13)$$

In order to bound the reminder $\tilde{c}(\Psi + \delta\Psi, u + \delta u) - \tilde{c}(\Psi, u) - D\,\tilde{c}(\Psi, u)(\delta\Psi, \delta u)$, we consider

$$\|V_u \delta u \delta\Psi\|_Y^2 \leq \|V_u\|_{L^\infty(\Omega;\mathbb{R})}^2 \|\delta u\|_{C[0,T]}^2 \|\delta\Psi\|_X^2,$$

where we use the embedding $U \hookrightarrow C[0,T]$. Using the estimate $\|\delta u\|_{C[0,T]} \leq c\|\delta u\|_{H^1(0,T;\mathbb{R})}$ and $\|\delta\Psi\|_X \leq \|\delta\Psi\|_W$, we can further estimate as follows:

$$\|V_u \delta u \delta\Psi\|_Y \leq c'\|\delta u\|_{H^1(0,T;\mathbb{R})}\|\delta\Psi\|_W \leq c'\left(\|\delta u\|_{H^1(0,T;\mathbb{R})}^2 + \|\delta\Psi\|_W^2\right)$$

$$\leq c'\left(\|\delta u\|_{H^1(0,T;\mathbb{R})} + \|\delta\Psi\|_W\right)^2.$$

With this result, we can show the Fréchet differentiability as follows:

$$\frac{\|V_u \delta u \delta\Psi\|_Y}{\|\delta u\|_{H^1(0,T;\mathbb{R})} + \|\delta\Psi\|_W} \leq c'\left(\|\delta u\|_{H^1(0,T;\mathbb{R})} + \|\delta\Psi\|_W\right) \to 0$$

for $\|\delta u\|_{H^1(0,T;\mathbb{R})} + \|\delta\Psi\|_W \to 0$. Hence, $\tilde{c}$ is Fréchet differentiable, and $D\,\tilde{c}$ given in (4.13) represents its derivative. This derivative is continuous from $W \times U$ to X^*. By the remark on page 58 this means $\tilde{c}$ is continously differentiable from $Z \times U$ to X^*.

The exchange potential is continuously real-Fréchet differentiable from Z to Y by Lemma 4.4, the correlation potential from Z to Y by Assumption 12), and the Hartree potential from Z to Y by Lemma 4.5. To summarize, we have the following real-Fréchet derivative of c:

$$Dc(\Psi, u)(\delta\Psi, \delta u)$$

$$= c(\delta\Psi, u) - V_u \delta u \Psi - \frac{\partial V_{xc}}{\partial\rho} 2\,\mathrm{Re}\,(\Psi, \delta\Psi)_{\mathbb{C}} - \int_\Omega \frac{2\,\mathrm{Re}\,(\Psi, \delta\Psi)_{\mathbb{C}}}{|x - y|}\mathrm{d}y\Psi$$

$$= i\frac{\partial\delta\Psi}{\partial t} - \left(-\nabla^2 + V_{ext}(x, t, u) + V_{Hxc}(\Psi)\right)\delta\Psi - V_u \delta u \Psi \qquad (4.14)$$

$$- \frac{\partial V_{xc}}{\partial\rho} 2\,\mathrm{Re}\,(\Psi, \delta\Psi)_{\mathbb{C}}\,\Psi - \int_\Omega \frac{2\,\mathrm{Re}\,(\Psi, \delta\Psi)_{\mathbb{C}}}{|x - y|}\mathrm{d}y\Psi.$$

$\square$

We have discussed the differentiability properties of the differential constraint c that are required below to prove the existence of a minimizer and to establish the gradient of the reduced cost functional. Next, we study the cost functional J.

Theorem 4.6. *The cost functional $J : Z \times U \to \mathbb{R}$ defined in (4.5) is continuously real-Fréchet differentiable.*

Proof. The norm $\|u\|_{H^1(0,T;\mathbb{R})}^2$ is differentiable by standard results with $D\left(\frac{\nu}{2}\|u\|_{H^1(0,T;\mathbb{R})}^2\right)(\delta u) = \nu\,\langle u, \delta u\rangle_{H^1(0,T)}$.

The tracking term $J_\beta = \frac{\beta}{2}\int_0^T \int_\Omega (\rho(x,t) - \rho_d(x,t))^2 \mathrm{d}x\mathrm{d}t$ is a quadratic functional and hence real-Fréchet differentiable with derivative

$$DJ_\beta(\Psi)(\delta\Psi) = 2\beta \int_0^T \int_\Omega (\rho(x,t) - \rho_d(x,t))\,\mathrm{Re}\,(\Psi(x,t), \delta\Psi(x,t))_{\mathbb{C}}\,\mathrm{d}x\mathrm{d}t.$$

$J_\eta = \frac{\eta}{2}\int_\Omega \chi_A(x)\rho(x,T)\mathrm{d}x$ is well-defined, because $\Psi \in C([0,T]; L^2(\Omega;\mathbb{C}^N))$. The directional derivative

$$D\,J_\eta(\Psi)(\delta\Psi) = \eta \int_\Omega \chi_A(x)\,\mathrm{Re}\,(\Psi(x,T), \delta\Psi(x,T))_{\mathbb{C}}\,\mathrm{d}x$$

is obviously real-linear and continuous in $\delta\Psi$, hence it is the real-Gâteaux derivative. Furthermore, we have

$$\frac{|J_\eta(\Psi + \delta\Psi) - J_\eta(\Psi) - \mathrm{D}\, J_\eta(\Psi)(\delta\Psi)|}{\|\delta\Psi\|_W} = \frac{\left|\int_\Omega \chi_A(x)|\delta\Psi(T)|^2 \mathrm{d}x\right|}{\|\delta\Psi\|_W} \leq \frac{\|\delta\Psi(T)\|^2_{L^2(\Omega;\mathbb{C}^N)}}{\|\delta\Psi\|_W}$$

$$\leq \frac{\max_{0 \leq t \leq T} \|\delta\Psi(t)\|^2_{L^2(\Omega;\mathbb{C}^N)}}{\|\delta\Psi\|_W} \leq \frac{c\|\delta\Psi\|^2_W}{\|\delta\Psi\|_W} = c\|\delta\Psi\|_W \to 0 \text{ for } \|\delta\Psi\|_W \to 0.$$

Therefore, J_η is real-Fréchet differentiable from W to $\mathbb{R}$. Its Fréchet derivative depends real-linearly on Ψ and is bounded by

$$\|\mathrm{D}\, J_\eta(\Psi)\|_{\mathcal{L}(W,\mathbb{R})} = \eta \sup_{\delta\Psi \in W \setminus \{0\}} \frac{\left|\int_\Omega \chi_A(x) \operatorname{Re}\left(\Psi(x,T),\, \delta\Psi(x,T)\right)_\mathbb{C} \mathrm{d}x\right|}{\|\delta\Psi\|_W}$$

$$\leq \eta \frac{\|\Psi(T)\|_{L^2(\Omega;\mathbb{C}^N)}\|\delta\Psi(T)\|_{L^2(\Omega;\mathbb{C}^N)}}{\|\delta\Psi\|_W} \leq \eta c^2 \frac{\|\Psi\|_W \|\delta\Psi\|_W}{\|\delta\Psi\|_W} = \eta c^2 \|\Psi\|_W.$$

Hence, this derivative is continuous from W to $\mathcal{L}(W,\mathbb{R})$. By the remark on page 58 J_η is continuously real-Fréchet differentiable from Z to $\mathbb{R}$. $\qquad\square$

4.3.2. Existence of a minimizer

In this section, we discuss existence of a minimizer of the optimization problem (4.6). Uniqueness cannot be expected because the phase of the wavefunction does not appear in the cost functional J.

We start by collecting some known facts. We use the following version of the Arzelà-Ascoli theorem which holds also for general Banach spaces, see, e.g., [Cia13, Theorem 3.10-2] and remark below.

Theorem 4.7 (Arzelà-Ascoli). *Let $\mathcal{B}$ be a Banach space and $K = [0,T]$. Given a sequence $(f_n)_n$ of functions $f_n \in C(K;\mathcal{B})$. If the sequence is*

1. *uniformly bounded, i.e. $\exists M$, such that $\|f_n\|_{C(K;\mathcal{B})} \leq M \; \forall n \in \mathbb{N}$, and*

2. *equicontinuous, i.e. given any $\epsilon > 0$, there exists $\delta(\epsilon) > 0$, such that $\|f_n(t) - f_n(s)\|_\mathcal{B} < \epsilon$ for all $t, s \in K$ with $|t - s| < \delta(\epsilon)$ and all $n \in \mathbb{N}$;*

then there exists a subsequence $(f_{n_l})_l$ and a function $f \in C(K;\mathcal{B})$ such that

$$\lim_{l \to \infty} \|f_{n_l} - f\|_{C(K;\mathcal{B})} = 0.$$

For the purpose of our discussion, notice that the semigroup generated by $H_0 = -\nabla^2 + V_0$ is denoted by $\mathcal{U}(t) = e^{-iH_0 t}$. For self-adjoint operators H_0, as in our case, $\mathcal{U}(t)$ is strongly continuous by the Stone's theorem; see [Sto32] and [Yse10, p. 34].

We need the following lemma; see also [RPvL15].

Lemma 4.8. *The Duhamel form of the TDKS equation* (4.1) *is given by*

$$\Psi(t) = e^{-iH_0 t}\Psi(0) - ie^{-iH_0 t}\int_0^t g(s)\mathrm{d}s \quad \text{with } g(s) = e^{iH_0 s}(u(s)V_u\Psi(s) + V_{Hxc}(\Psi(s))\Psi(s)).$$

$$\tag{4.15}$$

Proof. We follow the approach in [Sal05] and write

$$i\partial_t \Psi(t) = \left(-\nabla^2 + V_0 + uV_u + V_{Hxc}(\Psi(t))\right)\Psi(t)$$

$$= H_0\Psi(t) + uV_u\Psi(t) + V_{Hxc}(\Psi(t))\Psi(t)$$

$$\Leftrightarrow ie^{iH_0t}(\partial_t\Psi(t) + iH_0\Psi(t)) = e^{iH_0t}\left(uV_u\Psi(t) + V_{Hxc}(\Psi(t))\Psi(t)\right)$$

$$\Leftrightarrow \frac{\mathrm{d}}{\mathrm{d}t}\left(ie^{iH_0t}\Psi(t)\right) = e^{iH_0t}\left(uV_u\Psi(t) + V_{Hxc}(\Psi(t))\Psi(t)\right)$$

$$\Leftrightarrow ie^{iH_0t}\Psi(t) - i\Psi(0) = \int_0^t e^{iH_0s}\left(uV_u\Psi(s) + V_{Hxc}(\Psi(s))\Psi(s)\right)\mathrm{d}s,$$

which proves the lemma. $\qquad\square$

Now, we can prove the existence of a solution to (4.6).

Theorem 4.9. *The optimal control problem (4.6) with $\Psi^0 \in H_0^1(\Omega;\mathbb{C}^N)$ and $\beta = 0$ admits at least one solution in $(\Psi, u) \in W \times U$. In the case $\beta \neq 0$, an optimal solution exists in $Z \times U$ assuming that $\Psi^0 \in H^2(\Omega;\mathbb{C}^N) \cap H_0^1(\Omega;\mathbb{C}^N)$.*

Proof. For a given control $u \in U$, we define $\Psi(u)$ as the unique solution to (4.3). Let $(\Psi_n, u_n) := (\Psi(u_n), u_n)$ be a minimizing sequence of J, i.e.

$$\lim_{n\to\infty} J(\Psi(u_n), u_n) = \inf_{u\in U} J(\Psi(u), u). \tag{4.16}$$

As J is coercive with respect to u in the H^1 norm, the sequence $(u_n)_n$ is bounded in U. Hence, we can extract a weakly convergent subsequence again denoted by $(u_n)_n$, $u_n \rightharpoonup \hat{u}$ in U. By the Rellich-Kondrachov Theorem A.8, we then have $u_n \to \hat{u}$ in $C[0,T]$.

As the controls u_n in the sequence above are globally bounded in U, by Theorem 3.10 we have $\|\Psi_n\|_X \leq K$ and $\|\Psi'_n\|_{X^*} \leq K'$, where the constants K, K' can be chosen to be independent of u_n. Hence, we can extract weakly convergent subsequences, again denoted by Ψ_n, Ψ'_n, as follows:

$$\Psi_n \xrightarrow{X} \hat{\Psi}, \qquad\qquad \Psi'_n \xrightarrow{X^*} \hat{\Psi}'.$$

By the Rellich-Kondrachov theorem, we have $H^1(\Omega;\mathbb{C}^N) \Subset L^2(\Omega;\mathbb{C}^N)$, and by Aubin-Lions's theorem A.12 we have $W \Subset Y$, hence

$$\Psi_n \xrightarrow{Y} \hat{\Psi}. \tag{4.17}$$

By Lemma 4.4 and 4.5 and Assumption 12), $\Psi \mapsto V_{Hxc}(\Psi)\Psi$ is real-Fréchet differentiable, hence continuous from W to Y. Every (Ψ_n, u_n) solves the Schrödinger equation (4.3). Using the strong convergence of u_n and Ψ_n, we can employ [Cia13, p. 291] for the products $u_n\Psi_n$ and $V_{Hxc}(\Psi_n)\Psi_n$. By standard results [ADPM11, Proposition 3.6], a sequence converging in the $L^2(0,T)$-norm contains a subsequence that converges a.e. in $[0,T]$. Consequently, we can extract a subsequence, again denoted by (Ψ_n) such that we can pass to the limit. Then, the equation

$$\lim_{n\to\infty} -\left\langle i\frac{\partial\Psi_n}{\partial t}, \Phi\right\rangle_{L^2} + \langle\nabla\Psi_n, \nabla\Phi\rangle_{L^2} + \langle V_0\Psi_n + V_u u_n\Psi_n, \Phi\rangle_{L^2} + \langle V_{Hxc}(\Psi_n)\Psi_n, \Phi\rangle_{L^2}$$

$$= -\left\langle i\frac{\partial\hat{\Psi}}{\partial t}, \Phi\right\rangle_{L^2} + \left\langle\nabla\hat{\Psi}, \nabla\Phi\right\rangle_{L^2} + \left\langle V_0\hat{\Psi} + V_u\hat{u}\hat{\Psi}, \Phi\right\rangle_{L^2} + \left\langle V_{Hxc}(\hat{\Psi})\hat{\Psi}, \Phi\right\rangle_{L^2}$$

holds a.e. in $[0,T]$ and for all $\Phi \in H_0^1(\Omega;\mathbb{C}^N)$. Hence, $\Psi(\hat{u}) = \hat{\Psi}$ and $(\hat{\Psi}, \hat{u})$ solves (4.3).

If $\eta \neq 0$, then the evaluation of ρ at the final time T in the cost J is required, such that strong convergence in $C([0,T]; L^2(\Omega; \mathbb{C}^N))$ is needed. So far, we only have weak convergence in W, which is continuously embedded into $C([0,T]; L^2(\Omega; \mathbb{C}^N))$, but not compactly embedded. To overcome this problem, we improve our convergence result by using Theorem 4.7. The required uniform bound is given by Lemma 3.11, we have $\|\Psi_n(t)\|_{L^2} = 1$, for all n and all $t \in [0,T]$. It remains to show the equicontinuity of Ψ_n for the Arzelà-Ascoli theorem, i.e. find an $\delta(\epsilon)$ that does not depend on the n of the sequence. To this end, we take a fixed but arbitrary $\epsilon > 0$. With $\|\Psi_n\|_{L^2} = 1$, the Lipschitz continuity of V_{xc}, Lemma 3.2 and Assumption 4, and estimates in Lemma 3.4 and Theorem 3.10, we have

$$
\begin{aligned}
\left\| \Psi^0 - i \int_0^t g(s) \mathrm{d}s \right\|_{L^2(\Omega; \mathbb{C}^N)} &\leq \|\Psi^0\|_{L^2} + \|g\|_Y \\
&\leq \|\Psi^0\|_{L^2} + \|V_u \Psi_n\|_Y + \|V_{Hxc}(\Psi_n)\Psi_n\|_Y \\
&\leq \|\Psi^0\|_{L^2} + \left(\|u_n\|_{C[0,T]} \|V_u\|_{L^\infty(\Omega;\mathbb{C}^N)} + L + c\|\Psi_n\|_X^2 \right) \|\Psi_n\|_Y \leq C,
\end{aligned}
$$

where g is defined in Lemma 4.8. As $\mathcal{U}(t) = e^{-iH_0 t}$ is continuous by Stone's theorem, there exists a δ_1 such that

$$
\|\mathcal{U}(t) - \mathcal{U}(t')\|_{\mathcal{L}(L^2, L^2)} < \frac{\epsilon}{2C} \quad \text{for } |t - t'| < \delta_1.
$$

Using the Duhamel formula from Lemma 4.8, we find for two different times t, t', with $|t - t'| < \delta$, the following:

$$
\begin{aligned}
\|\Psi_n(t) - \Psi_n(t')\|_{L^2} &\leq \left\| \left(e^{-iH_0 t} - e^{iH_0 t'} \right) \left(\Psi^0 - i \int_0^t g(s) \mathrm{d}s \right) \right\|_{L^2} \\
&\quad + \left\| e^{-iH_0 t'} \int_{t'}^t g(s) \mathrm{d}s \right\|_{L^2}.
\end{aligned}
$$

The first term is bounded by

$$
\|\mathcal{U}(t) - \mathcal{U}(t')\|_{\mathcal{L}(L^2, L^2)} \left\| \Psi^0 - i \int_0^t g(s)\mathrm{d}s \right\|_{L^2} \leq \frac{\epsilon}{2C} C = \frac{\epsilon}{2}.
$$

For the second term, we have

$$
\begin{aligned}
\left\| \int_{t'}^t g(s)\mathrm{d}s \right\|_{L^2}^2 &= \int_\Omega \left| \int_{t'}^t g(s,x)\mathrm{d}s \right|^2 \mathrm{d}x \leq \int_\Omega \left(\int_{t'}^t |g(s,x)|\mathrm{d}s \right)^2 \mathrm{d}x \\
&= \int_\Omega \|g(\cdot, x)\|_{L^1(t',t;\mathbb{C}^N)}^2 \mathrm{d}x \leq \int_\Omega \sqrt{|t-t'|}^2 \|g(\cdot, x)\|_{L^2(t',t;\mathbb{C}^N)}^2 \mathrm{d}x \\
&\leq |t - t'| \int_\Omega \int_{t'}^t |g(s,x)|^2 \mathrm{d}s \mathrm{d}x \\
&= |t - t'| \int_{t'}^t \|e^{iH_0 s}\left(u V_u \Psi_n(s) + V_{Hxc}(\Psi_n(s))\Psi_n(s) \right)\|_{L^2}^2 \mathrm{d}s.
\end{aligned}
$$

Using the fact that $e^{iH_0 s}$ is unitary, we obtain

$$
\left\| \int_{t'}^{t} g(s) \mathrm{d}s \right\|_{L^2}^2
$$

$$
\leq |t - t'| \int_{t'}^{t} \| \left(V_u \Psi_n(s) + V_{xc}(\Psi_n(s))\Psi_n(s) + V_H(\Psi_n(s))\Psi_n(s) \right) \|_{L^2}^2 \mathrm{d}s
$$

$$
\leq 3|t - t'| \int_{t'}^{t} C^2 \|u_n\|_{C[0,T]}^2 \|\Psi_n(s)\|_{L^2}^2 + \|V_{xc}(\Psi_n(s))\Psi_n(s)\|_{L^2}^2
$$

$$
+ \|V_H(\Psi_n(s))\Psi_n(s)\|_{L^2}^2 \mathrm{d}s
$$

$$
\leq 3|t - t'| \int_{t'}^{t} C' + L\|\Psi_n(s)\|_{L^2}^2 + C''\|\Psi_n(s)\|_{H^1}^4 \|\Psi_n(s)\|_{L^2}^2 \mathrm{d}s
$$

$$
\leq K|t - t'|^2,
$$

where we use the Lipschitz continuity of $V_{xc} = V_x + V_c$ and Lemma 3.10 for the estimate of V_H. Furthermore, we used Corollary 3.11, $\|\Psi_n\|_{L^2} = 1$, and according to Theorem 3.10 we have $\|\Psi(t)\|_{H^1} \leq C(\Psi^0)$. Moreover, since the u_n are from a bounded sequence, we have that $\|u_n\|_{C[0,T]}$ is bounded by a global constant.

Now, we define $\delta := \min\{\delta_1, \frac{\epsilon}{2\sqrt{K}}\}$. Then

$$
\|\Psi_n(t) - \Psi_n(t')\|_{L^2(\Omega; \mathbb{C}^N)} < \frac{\epsilon}{2} + \sqrt{K}\delta \leq \epsilon \quad \forall |t - t'| < \delta \text{ and } \forall n \geq 1. \tag{4.18}
$$

This means that the sequence $(\Psi_n)_n$ is equicontinuous. As it is also uniformly bounded, we can invoke the Arzelà-Ascoli theorem (Theorem 4.7) to conclude that there exists a subsequence Ψ_{n_l} and a function $\hat{\Psi} \in C([0,T]; L^2(\Omega; \mathbb{C}^N))$, such that $\lim_{l \to \infty} \Psi_{n_l} = \hat{\Psi}$ in $C([0,T]; L^2(\Omega; \mathbb{C}^N))$. With the strong convergence in $C([0,T]; L^2(\Omega; \mathbb{C}^N))$, we have

$$
\rho(\hat{u})(\cdot, T) = \sum_{j=1}^{N} |\Psi_j(\hat{u})(\cdot, T)|^2 = \sum_{j=1}^{N} |\hat{\Psi}_j(\cdot, T)|^2 = \hat{\rho}(\cdot, T).
$$

With this result, we have the convergence of (4.16) also for a target depending only on the final time.

For the target J_β to be well-defined, we need to assume higher regularity. Assuming $\Psi^0 \in H^2(\Omega; \mathbb{C}^N)$, we get $\Psi \in L^\infty(0,T; H^2(\Omega; \mathbb{C}^N))$ from Theorem 3.18. Due to the Sobolev embedding $L^\infty(0,T; H^2(\Omega; \mathbb{C}^N)) \hookrightarrow L^\infty(0,T; L^\infty(\Omega; \mathbb{C}))$, the wavefunction is globally bounded in space and time by a constant depending on the initial condition and $\|u\|_{H^1(0,T;\mathbb{R})}$. As $\|u_n\|_{H^1(0,T;\mathbb{R})}$ is bounded, we have $\|\Psi_n\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{C}))} < K$, $\|\hat{\Psi}\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{C}))} < K$ and the same holds for the squares ρ_n, $\hat{\rho}$.

As $\Psi_n \xrightarrow{Y} \hat{\Psi}$, so does $\overline{\Psi}_n \xrightarrow{Y} \overline{\hat{\Psi}}$. Hence the product converges in the L^1-norm, $\rho_n \xrightarrow{L^1(0,T;L^1(\Omega;\mathbb{R}))} \hat{\rho}$. Furthermore, $(\rho_n)_n$ converges in the Y-norm as follows:

$$
\|\rho_n - \hat{\rho}\|_Y^2 = \|\rho_n(\rho_n - \hat{\rho}) - \hat{\rho}(\rho_n - \hat{\rho})\|_{L^1(0,T;L^1(\Omega;\mathbb{R}))}
$$

$$
\leq \|\rho_n\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}))} \|\rho_n - \hat{\rho}\|_{L^1(0,T;L^1(\Omega;\mathbb{R}))}
$$

$$
+ \|\hat{\rho}\|_{L^\infty(0,T;L^\infty(\Omega;\mathbb{R}))} \|\rho_n - \hat{\rho}\|_{L^1(0,T;L^1(\Omega;\mathbb{R}))}.
$$

As $J_\beta(\Psi) = \frac{\beta}{2}\|\rho - \rho_d\|_Y^2$ and the norm is continuous, we can pass to the limit.

The norm $\|u\|_{H^1(0,T;\mathbb{R})}$ is weakly lower semicontinuous. Hence, $(\hat{\Psi}, \hat{u})$ minimizes (4.6) as follows:

$$
J(\hat{\Psi}, \hat{u}) \leq \liminf_{n \to \infty} J(\Psi_n, u_n) = \inf_{u \in U} J(\Psi(u), u). \tag{4.19}
$$

$\square$

4.3.3. Necessary optimality conditions

In this section, we discuss the Lagrange multiplier Λ appearing in (4.7) and (4.8) and state the first-order optimality condition for a minimum.

We start by showing the existence of a Lagrange multiplier in the Lagrange framework.

Theorem 4.10. *Given a control u and a corresponding state Ψ, there exists a Lagrange multiplier $\Lambda \in X$ associated with (Ψ, u).*

Proof. The constraint c and the objective J are continuously real-Fréchet differentiable by Lemma 4.6 and Theorem 4.2, and the derivative $D\,c(\Psi, u) : Z \times U \to X^*$ is surjective by Theorem 3.13, where the results can readily be extended to a non-zero right-hand side $F \in X^*$. Hence, the constraint qualification of Zowe and Kurcyusz is fulfilled [ZK79]. Therefore, we have the existence of a Lagrange multiplier $\Lambda \in X$; see, e.g., [Trö10, Section 6.1]. □

Since we need higher regularity, namely $\Lambda \in W$, we do not make further use of this result. Instead, we proceed in a different way using the existence of a unique solution of the adjoint equation in W from Theorem 3.13.

As $c(\Psi, u) = 0$ is uniquely solvable, we have the following equivalent formulation of the optimization problem.

Lemma 4.11. *The minimization problem* (4.6) *is equivalent to the unconstrained minimization of the reduced cost functional $\hat{J}(u) := J(\Psi(u), u)$ as follows:*

$$\min_{(\Psi, u) \in Z \times U} J(\Psi, u), \ s.t. \ c(\Psi, u) = 0 \quad \Leftrightarrow \quad \min_{u \in U} \hat{J}(u). \tag{4.20}$$

To calculate the gradient of the reduced cost functional, we make use of the implicit function theorem; see, e.g., [Cia13, p. 548]. We apply the chain rule for the real-Fréchet derivative to the cost functional and remark again that the chain rule and the implicit function theorem hold for both the Fréchet derivative and the real-Fréchet derivative introduced in Definition 4.1.

Lemma 4.12. *The control-to-state map $U \ni u \mapsto \Psi(u) \in Z$ is real-Fréchet differentiable and the derivative $D_u \Psi(u) = \delta\Psi$ is given by the solution $\delta\Psi$ of the linearised constraint $\big(Dc(\Psi, u)\big)(\delta\Psi, \delta u) = 0$.*

Proof. In order to show that the control-to-state map $u \mapsto \Psi$ is differentiable and the term $D_u\,\Psi(u)$ is well-defined, we use the implicit function Theorem A.16, which ensures real-differentiability of the map $u \mapsto \Psi(u)$, $U \to Z$. To this end, we show that the real-Fréchet derivative $D_\Psi\,c(\Psi, u) : Z \to X^*$ is a bijection at any $(\Psi, u) \in Z \times U$.

We consider the real-derivative of c with respect to $\delta\Psi$ from Theorem 4.2. This is given by

$$D_\Psi\,c(\Psi, u)(\delta\Psi) = i\frac{\partial \delta\Psi}{\partial t} - \big(-\nabla^2 + V_{ext}(x, t, u) + V_{Hxc}(\Psi)\big)\delta\Psi - \frac{\partial V_{xc}}{\partial \rho}2\,\mathrm{Re}\,(\Psi,\, \delta\Psi)_{\mathbb{C}}\,\Psi$$

$$- \int_\Omega \frac{2\,\mathrm{Re}\,(\Psi,\, \delta\Psi)_{\mathbb{C}}}{|x - y|}\mathrm{d}y\,\Psi.$$

As shown in Theorem 3.17, the equation $D_\Psi\,c(\Psi, u)(\delta\Psi) = F$ with the initial condition $\delta\Psi(t = 0) = 0$ is uniquely solvable for any right-hand side $F \in X^*$. Hence $D_\Psi\,c(\hat{\Psi}, \hat{u})$ is a bijection.

Now, from $c(\Psi, u) = 0$, we have

$$\big(\mathrm{D}\, c(\Psi, u)\big)(\delta\Psi, \delta u) = \mathrm{D}_\Psi\, c(\Psi, u)(\delta\Psi) + \mathrm{D}_u\, c(\Psi, u)(\delta u) = 0.$$

Solving this equation for $\delta\Psi$ results in

$$\delta\Psi = -\big(\mathrm{D}_\Psi\, c(\Psi, u)\big)^{-1}\big(\mathrm{D}_u\, c(\Psi, u)\big)(\delta u) = \mathrm{D}_u\, \Psi(u)(\delta u), \qquad (4.21)$$

where the last equality is due to the implicit function Theorem A.16. This means that the solution $\delta\Psi$ of the linearized equation $\mathrm{D}\, c = 0$ is in fact $\mathrm{D}_u\, \Psi(u)$. $\qquad \square$

Lemma 4.13. *The derivative of the reduced cost functional is given by*

$$\mathrm{D}\,\hat{J}(u)(\delta u) = \mathrm{D}_\Psi\, J(\Psi(u), u)\big(\mathrm{D}_u\, \Psi(u)(\delta u)\big) + \mathrm{D}_u\, J(\Psi(u), u)(\delta u), \qquad (4.22)$$

where $\mathrm{D}_u\, \Psi(u)(\delta u) = \delta\Psi$ *is the solution to the linearized equation*

$$\big(\mathrm{D}\, c(\Psi, u)\big)(\delta\Psi, \delta u) = \mathrm{D}_\Psi\, c(\Psi, u)(\delta\Psi) + \mathrm{D}_u\, c(\Psi, u)(\delta u) = 0. \qquad (4.23)$$

Proof. The chain rule for the real-Fréchet derivative is given by

$$\mathrm{D}\,\hat{J}(u) = \mathrm{D}_\Psi\, J(\Psi(u), u)\big(\mathrm{D}_u\, \Psi(u)\big) + \mathrm{D}_u\, J(\Psi(u), u).$$

The real-derivatives $\mathrm{D}_\Psi\, J(\Psi(u), u)$ and $\mathrm{D}_u\, J(\Psi(u), u)$ are given in Theorem 4.6. The derivative of the control-to-state map is given in the previous Lemma 4.12. Altogether, we find

$$\begin{aligned}
\mathrm{D}\,\hat{J}(u)(\delta u) &= \mathrm{D}_\Psi\, J(\Psi(u), u)\big(\mathrm{D}_u\, \Psi(u)(\delta u)\big) + \mathrm{D}_u\, J(\Psi(u), u)(\delta u) \\
&= \mathrm{D}_\Psi\, J(\Psi(u), u)(\delta\Psi) + \mathrm{D}_u\, J(\Psi(u), u)(\delta u),
\end{aligned}$$

where $\delta\Psi$ is given by (4.21). $\qquad \square$

Next, we determine the Riesz representative of $\mathrm{D}\,\hat{J}(u)$ with respect to the H^1-scalar product $\langle v, w\rangle_{H^1(0,T)} = \langle v, w\rangle_{L^2(0,T)} + b\,\langle \dot{v}, \dot{w}\rangle_{L^2(0,T)}$ for any $v, w \in H^1(0,T;\mathbb{R})$ and $b > 0$. This is obtained in the next theorem by introducing an adjoint variable Λ and the corresponding adjoint equation $a(\Psi, u, \Lambda) = 0$.

Theorem 4.14. *The U-gradient (so the H^1- or H_0^1-gradient) with respect to the scalar product $\langle \cdot, \cdot\rangle_{H^1(0,T)}$ of the reduced optimization problem is given by*

$$\nabla \hat{J}(t) = \nu u(t) + \mu(t), \qquad (4.24)$$

where μ is the U-Riesz representative of the continuous linear functional $\langle -\operatorname{Re}\langle \Lambda, V_u\Psi\rangle_{L^2}, \cdot\rangle_{L^2(0,T)}$, Ψ is the unique solution of $c(\Psi, u) = 0$, and Λ is the unique solution of $a(\Psi, u, \Lambda) = 0$, where

$$\begin{aligned}
a(\Psi, u, \Lambda) := {}& i\frac{\partial\Lambda}{\partial t} - \big(-\nabla^2 + V_{ext}(x, t, u) + V_{Hxc}(\Psi)\big)\Lambda - \frac{\partial V_{xc}}{\partial\rho} 2\operatorname{Re}(\Psi, \Lambda)_\mathbb{C}\,\Psi \\
& - \int_\Omega \frac{2\operatorname{Re}(\Psi, \Lambda)_\mathbb{C}}{|y - x|}\mathrm{d}y\,\Psi + 2\beta(\rho - \rho_d)\Psi,
\end{aligned} \qquad (4.25)$$

with the terminal condition

$$i\Lambda(T) = \eta\,\chi_A\,\Psi(T). \qquad (4.26)$$

Proof. We want to calculate the Riesz representative of $\mathrm{D}\,\hat{J}(u)(\delta u)$. By Lemma 4.13, we have

$$\mathrm{D}\,\hat{J}(u)(\delta u) = \mathrm{D}_\Psi\, J(\Psi(u), u)\big(\mathrm{D}_u\,\Psi(u)(\delta u)\big) + \mathrm{D}_u\,J(\Psi(u), u)(\delta u).$$

We have $\mathrm{D}_u\,J(\Psi(u), u)(\delta u) = \nu\,\langle u,\,\delta u\rangle_{H^1(0,T)}$, hence, the gradient of the second term is given by

$$\langle\nabla_u J(\Psi, u),\,\delta u\rangle_{H^1(0,T)} = \nu\,\langle u,\,\delta u\rangle_{H^1(0,T)}. \tag{4.27}$$

To express the first term as an operator acting on δu, we need to use the fact from Lemma 4.13 that $\mathrm{D}_u\,\Psi(u)(\delta u) = \delta\Psi$ is the solution of the linearized equation. Using the directional derivative calculated in the Section 4.4, we obtain for the first term

$$\begin{aligned}
&\mathrm{D}_\Psi\,J(\Psi, u)(\delta\Psi) \\
&= \beta\int_0^T\int_\Omega (\rho - \rho_d)2\,\mathrm{Re}\,(\Psi,\,\delta\Psi)_{\mathbb{C}}\,\mathrm{d}x\mathrm{d}t + \int_\Omega \eta\chi_A\,\mathrm{Re}\,(\Psi(T),\,\delta\Psi(T))_{\mathbb{C}}\,\mathrm{d}x \\
&= \beta 2\,\mathrm{Re}\,\langle(\rho - \rho_d)\Psi,\,\delta\Psi\rangle_Y + \mathrm{Re}\,\langle\eta\chi_A\Psi(T),\,\delta\Psi(T)\rangle_{L^2}.
\end{aligned} \tag{4.28}$$

To simplify the equation, we focus on the last term of the right hand side in (4.28). By the continuous embedding $W \hookrightarrow C([0,T]; L^2(\Omega; \mathbb{C}^N))$, we can invoke the fundamental theorem of calculus in time.

$$\begin{aligned}
\langle\Lambda(T),\,\delta\Psi(T)\rangle_{L^2} &= \langle\Lambda(0),\,\delta\Psi(0)\rangle_{L^2} + \int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\langle\Lambda(t),\,\delta\Psi(t)\rangle_{L^2}\,\mathrm{d}t \\
&= \langle\Lambda(0),\,\delta\Psi(0)\rangle_{L^2} + \int_0^T \left\langle\frac{\partial\Lambda(t)}{\partial t},\,\delta\Psi(t)\right\rangle_{L^2} + \left\langle\Lambda(t),\,\frac{\partial\delta\Psi(t)}{\partial t}\right\rangle_{L^2}\,\mathrm{d}t.
\end{aligned}$$

Observing that $\delta\Psi(0) = 0$ and using equations (4.25), (4.26) for Λ and the fact that $\delta\Psi$ is the solution of the linearized equation (4.14), as well as the previous result, we obtain

$$\begin{aligned}
&\int_\Omega \eta\chi_A\,\mathrm{Re}\,(\Psi(T),\,\delta\Psi(T))_{\mathbb{C}}\,\mathrm{d}x = \mathrm{Re}\,\langle\eta\chi_A\Psi(T),\,\delta\Psi(T)\rangle_{L^2} = \mathrm{Re}\,\langle i\Lambda(T),\,\delta\Psi(T)\rangle_{L^2} \\
&= \mathrm{Re}\int_0^T \left\langle(-\nabla^2 + V_{ext}(u) + V_{Hxc}(\Psi))\Lambda + \frac{\partial V_{xc}}{\partial\rho}2\,\mathrm{Re}\,(\Psi,\,\Lambda)_{\mathbb{C}}\,\Psi + \int_\Omega \frac{2\,\mathrm{Re}\,(\Psi,\,\Lambda)_{\mathbb{C}}}{|y - x|}\mathrm{d}y\,\Psi\right. \\
&\qquad\left. - 2\beta(\rho - \rho_d)\Psi,\,\delta\Psi\right\rangle_{L^2}\,\mathrm{d}t \\
&\quad - \mathrm{Re}\int_0^T \left\langle\Lambda(t),\,\big(-\nabla^2 + V_{ext}(u) + V_{Hxc}(\Psi)\big)\delta\Psi + V_u\delta u\Psi + \frac{\partial V_{xc}}{\partial\rho}2\,\mathrm{Re}\,(\Psi,\,\delta\Psi)_{\mathbb{C}}\,\Psi\right. \\
&\qquad\left. + \int_\Omega \frac{2\,\mathrm{Re}\,(\Psi,\,\delta\Psi)_{\mathbb{C}}}{|x - y|}\mathrm{d}y\,\Psi\right\rangle_{L^2}\,\mathrm{d}t.
\end{aligned}$$

Using integration by parts with the zero boundary condition on $\partial\Omega$, this simplifies to

$$\int_\Omega \eta\chi_A\,\mathrm{Re}\,(\Psi(T),\,\delta\Psi(T))_{\mathbb{C}}\,\mathrm{d}x = 2\,\mathrm{Re}\,\langle-\beta(\rho - \rho_d)\Psi,\,\delta\Psi\rangle_Y - \mathrm{Re}\,\langle\Lambda,\,V_u\delta u\Psi\rangle_Y.$$

Using this result in (4.28), we obtain

$$\mathrm{D}_\Psi\,J(\Psi, u)(\delta\Psi) = -\,\mathrm{Re}\,\langle\Lambda,\,V_u\delta u\Psi\rangle_Y = \langle\mu,\,\delta u\rangle_{H^1(0,T)}, \tag{4.29}$$

since, by definition of μ, we have $\langle-\,\mathrm{Re}\,\langle\Lambda,\,V_u\Psi\rangle_{L^2},\,\delta u\rangle_{L^2(0,T)} = \langle\mu,\,\delta u\rangle_{H^1(0,T)}$.

Adding (4.29) and (4.27) together, we obtain (4.24). $\qquad\square$

Theorem 4.15. *Given a local solution $(\Psi, u) \in Z \times U$ of the minimization problem (4.6), i.e. $\hat{J}(u) \leq \hat{J}(\tilde{u})$ for all $\|u - \tilde{u}\|_{H^1(0,T;\mathbb{R})} < \epsilon$ for some fixed $\epsilon > 0$. Then there exists a unique Lagrange multiplier $\Lambda \in W$, such that the following first-order optimality system is fulfilled:*

$$c(\Psi, u) = 0, \qquad\qquad \Psi(0) = \Psi^0, \qquad\qquad (4.30\text{a})$$

$$a(\Psi, u, \Lambda) = 0, \qquad\qquad i\Lambda(T) = \eta \chi_A \Psi(T), \qquad\qquad (4.30\text{b})$$

$$\nabla \hat{J}(u) = 0, \qquad\qquad\qquad\qquad (4.30\text{c})$$

where $\nabla \hat{J}(u)$ is given by Theorem 4.14.

Proof. For all admissible pairs (Ψ, u) that satisfy (4.30a), the adjoint problem (4.30b) has a unique solution in W by Theorem 3.13. As $\hat{J}$ is differentiable, a local minimum is characterized by a zero gradient; see, e.g., [BS12], hence (4.30c) holds. $\qquad\square$

4.4. Another derivation of the first-order optimality system

Another way to derive the optimality system is given by calculating the directional derivatives of the Lagrange functional (4.7) with respect to the states ψ_j, the adjoint variables λ_j, and the control u.

Since J does not depend on Λ and L_1 depends on λ_j linearly, so the derivative with respect to λ_j gives the TDKS equations for ψ_j, $j = 1, \ldots, N$:

$$i\frac{\partial \psi_j(x,t)}{\partial t} = \left(-\nabla^2 + V_{ext}(x,t,u) + V_{Hxc}(x,t,\rho)\right) \psi_j(x,t). \qquad (4.31)$$

To derive the adjoint equations, we consider the derivative of (4.7) with respect to ψ_j. Both L_1 and the objective J depend on Ψ. We start with L_1 which is split into two terms:

$$L_1(\Psi, u, \Lambda) = L_2(\Psi, u, \Lambda) + L_3(\Psi, u, \Lambda),$$

where L_2 represents the linear part given by

$$L_2(\Psi, u, \Lambda) = \text{Re}\left(\sum_{j=1}^{N} \int_0^T \int_\Omega \left(i\frac{\partial \psi_j(x,t)}{\partial t} - \left(-\nabla^2 + V_{ext}(x,t,u)\right)\psi_j(x,t)\right)\overline{\lambda_j(x,t)}\mathrm{d}x\mathrm{d}t\right)$$

and L_3 is the nonlinear part containing the Kohn-Sham potential

$$L_3(\Psi, u, \Lambda) = \text{Re}\left(\sum_{j=1}^{N} \int_0^T \int_\Omega V_{Hxc}(x,t,\rho)\psi_j(x,t)\overline{\lambda_j(x,t)}\mathrm{d}x\mathrm{d}t\right).$$

We begin with the linear part L_2. Differentiating L_2 with respect to ψ_j gives the TDKS equation for λ_j as it is linear in ψ_j and the sign from integration by parts is cancelled by the complex conjugation. One boundary term B_1 remains.

$$\nabla_{\psi_j} L_2 = i\frac{\partial \lambda_j(x,t)}{\partial t} - \left(-\nabla^2 + V_{ext}(x,t,u)\right)\lambda_j(x,t) + B_1. \qquad (4.32)$$

For the calculation details, we fix one particle index j and we denote by $\delta\Psi^j$ a vector with all zero components up to the j-th component which is $\delta\psi_j$:

$$\delta\Psi^j := \left(0, \ldots 0, \delta\psi_j, 0, \ldots 0, \right),$$

and we notice that each component of $\delta\Psi^j$ can be written as $(\delta\Psi^j)_k = \delta_{kj}\delta\psi_j$ where δ_{jk} denotes the Kronecker delta. Then, we compute the derivative of L_2 with respect to ψ_j along $\delta\psi_j$ as follows:

$$
\begin{aligned}
\left\langle \nabla_{\psi_j} L_2,\, \delta\psi_j \right\rangle_Y &= \lim_{\alpha\searrow 0} \frac{L_2(\{\psi_k + \alpha\delta_{kj}\delta\psi_j\}, u) - L_2(\{\psi_k\}, u)}{\alpha} \\
&= \lim_{\alpha\searrow 0} \frac{1}{\alpha} \operatorname{Re} \sum_{k=1}^{N} \int_0^T \int_\Omega \left(i\frac{\partial(\psi_k + \delta_{kj}\alpha\delta\psi_k)}{\partial t} - \left(-\nabla^2 + V_{ext}\right)(\psi_k + \delta_{kj}\alpha\delta\psi_k) \right) \overline{\lambda_k}\,\mathrm{d}x\mathrm{d}t \\
&\quad - \operatorname{Re} \sum_{k=1}^{N} \int_0^T \int_\Omega \left(i\frac{\partial\psi_k}{\partial t} - \left(-\nabla^2 + V_{ext}\right)\psi_k \right) \overline{\lambda_k}\,\mathrm{d}x\mathrm{d}t \\
&= \operatorname{Re} \int_0^T \int_\Omega \left(i\frac{\partial\delta\psi_j}{\partial t} - \left(-\nabla^2 + V_{ext}\right)\delta\psi_j \right) \overline{\lambda_j}\,\mathrm{d}x\mathrm{d}t.
\end{aligned}
$$

Now, we use integration by parts and use the fact that λ_j and ψ_j are zero on the boundary. We obtain

$$
\begin{aligned}
\left\langle \nabla_{\psi_j} L_2,\, \delta\psi_j \right\rangle_Y &= \operatorname{Re} \int_0^T \int_\Omega \left(i\frac{\partial\lambda_j}{\partial t} - \left(-\nabla^2 + V_{ext}\right)\lambda_j \right) \overline{\delta\psi_j}\,\mathrm{d}x\mathrm{d}t \\
&\quad + \operatorname{Re} \int_\Omega -i\big(\lambda_j(x,T)\overline{\delta\psi_j(x,T)} - \lambda_j(x,0)\overline{\delta\psi_j(x,0)}\big)\,\mathrm{d}x.
\end{aligned}
$$

At time $t = 0$, all wavefunctions have to fulfil the initial condition, hence $\delta\psi_j(x,0) = 0$. We are left with one boundary term

$$
B_1 = -\operatorname{Re} \int_\Omega i\lambda_j(x,T)\overline{\delta\psi_j(x,T)}\,\mathrm{d}x, \tag{4.33}
$$

which can be removed by prescribing the terminal condition $\lambda_j(x,T) = 0$ in the case of $\eta = 0$, otherwise by the terminal condition derived below in (4.35). We obtain

$$
\left\langle \nabla_{\psi_j} L_2,\, \delta\psi_j \right\rangle_Y = \operatorname{Re} \int_0^T \int_\Omega \left(i\frac{\partial\lambda_j}{\partial t} - \left(-\nabla^2 + V_{ext}\right)\lambda_j \right) \overline{\delta\psi_j}\,\mathrm{d}x\mathrm{d}t + B_1. \tag{4.34}
$$

The derivative of the nonlinear part L_3 is given in the Lemmas 4.3 and 4.5.

Derivative of the target functional

Consider the tracking term $J_\beta = \frac{\beta}{2}\int_0^T \int_\Omega (\rho(x,t) - \rho_d(x,t))^2\mathrm{d}x\mathrm{d}t$, and notice the following relations:

$$
\begin{aligned}
|\psi_j + \alpha\delta\psi_j|^2 &= |\psi_j|^2 + 2\operatorname{Re}(\overline{\psi_j}\alpha\delta\psi_j) + |\alpha\delta\psi_j|^2, \\
|\psi_j + \alpha\delta\psi_j|^4 &= |\psi_j|^4 + 4|\psi_j|^2 \operatorname{Re}(\overline{\psi_j}\alpha\delta\psi_j) + \mathcal{O}(\alpha^2).
\end{aligned}
$$

With this preparation, we have

$$\left\langle \nabla_{\psi_j} J_\beta,\, \delta\psi_j \right\rangle_Y = \lim_{\alpha \searrow 0} \frac{\beta}{2\alpha} \int_0^T \int_\Omega \left(|\psi_j + \alpha\delta\psi_j|^2 + (\sum_{i\neq j} |\psi_i|^2 - \rho_d) \right)^2 - (\rho - \rho_d)^2 \mathrm{d}x\mathrm{d}t$$

$$= \lim_{\alpha \searrow 0} \frac{\beta}{2\alpha} \int_0^T \int_\Omega \left(|\psi_j + \alpha\delta\psi_j|^4 + (\sum_{i\neq j} |\psi_i|^2 - \rho_d)^2 + 2|\psi_j + \alpha\delta\psi_j|^2 (\sum_{i\neq j} |\psi_i|^2 - \rho_d) \right.$$

$$\left. -(\rho - \rho_d)^2 \right) \mathrm{d}x\mathrm{d}t$$

$$= 2\beta \int_0^T \int_\Omega |\psi_j|^2 \operatorname{Re}(\overline{\psi_j}\delta\psi_j) + \operatorname{Re}(\overline{\psi_j}\delta\psi_j)(\sum_{i\neq j} |\psi_i|^2 - \rho_d)\mathrm{d}x\mathrm{d}t$$

$$= 2\beta \operatorname{Re} \int_0^T \int_\Omega (\rho - \rho_d)\psi_j\overline{\delta\psi_j}\mathrm{d}x\mathrm{d}t. \tag{b2}$$

Similarly, we find for the terminal term $J_\eta = \int_\Omega \rho(x,T)\chi_A(x)\mathrm{d}x$ the following:

$$\left\langle \nabla_{\psi_j} J_\eta,\, \delta\psi_j \right\rangle_{L^2} = \lim_{\alpha \searrow 0} \frac{\eta}{2\alpha} \int_\Omega \chi_A \left(\sum_{i=1}^N |\psi_i(x,T) + \alpha\delta_{ij}\delta\psi_j(x,T)|^2 - \sum_{i=1}^N |\psi_i(x,T)|^2 \right) \mathrm{d}x$$

$$= \eta \operatorname{Re} \int_\Omega \chi_A(x)\psi_j(x,T)\overline{\delta\psi_j(x,T)}\mathrm{d}x. \tag{4.35}$$

Combining this with (4.33) determines the terminal condition for λ_j, see (4.8d).

The reduced gradient

Finally, we want to explicitly calculate the reduced gradient. The gradient is given in Theorem 4.14; however, we still have to find an explicit expression for μ. μ is defined to be the U-Riesz representative of the continuous linear functional $\langle -\operatorname{Re} \langle \Lambda,\, V_u\Psi \rangle_{L^2},\, \cdot \rangle_{L^2(0,T)}$. Therefore, $\mu \in U$ is the solution of the weak ODE

$$\langle -\operatorname{Re} \langle \Lambda,\, V_u\Psi \rangle_{L^2},\, \delta u \rangle_{L^2(0,T)} = \langle \mu,\, \delta u \rangle_{H^1(0,T)} = \int_0^T \mu(t)\delta u(t) + \dot\mu(t)\dot{\delta u}(t)\mathrm{d}t \quad \forall \delta u \in U.$$

Next, we derive the strong form of this ODE. Integration by parts gives

$$\langle \mu,\, \delta u \rangle_{H^1(0,T)} = \int_0^T \mu(t)\delta u(t) - \ddot\mu(t)\delta u(t)\mathrm{d}t + \dot\mu(t)\delta u(t)\Big|_{t=0}^T. \tag{4.36}$$

Considering the case $U = H_0^1(0,T;\mathbb{R})$, the boundary term vanishes because of $\delta u(0) = \delta u(T) = 0$ and $\mu \in H_0^1(0,T;\mathbb{R})$ is given by the solution of the ODE

$$-\operatorname{Re} \langle \Lambda,\, V_u\Psi \rangle_{L^2} = \left(1 - \frac{\mathrm{d}^2}{\mathrm{d}t^2} \right)\mu, \quad \mu(0) = \mu(T) = 0,$$

see also (4.9). We remark that here the boundary condition is due to the fact that μ is sought in the space $U = H_0^1(0,T;\mathbb{R})$.

In the case $U = H^1(0,T;\mathbb{R})$, we need to ensure that the boundary term in (4.36) vanishes for all $\delta u \in H^1(0,T;\mathbb{R})$. From this, we obtain Neumann boundary conditions and $\mu \in H^1(0,T;\mathbb{R})$ is given by the solution of

$$-\operatorname{Re} \langle \Lambda,\, V_u\Psi \rangle_{L^2} = \left(1 - \frac{\mathrm{d}^2}{\mathrm{d}t^2} \right)\mu, \quad \dot\mu(0) = \dot\mu(T) = 0.$$

4.5. Conclusion

In this chapter, an optimal control framework for the time-dependent Kohn-Sham model using the Lagrange formalism was presented and analysed. As standard results on the optimal control of the Schrödinger equation cannot be applied due to the complicated nature of the Kohn-Sham interaction, a new mathematical rigorous proof of the existence of optimal controls and their characterization as solutions to the corresponding optimality system was presented.

For this purpose, the differentiability properties of the nonlinear Kohn-Sham potential and of the TDKS equations were investigated. The two main issues were, on the one hand the complicated nonlinearity of the KS equations demanding an extensive analysis of the differentiability properties, and on the other hand, the fact that the KS potential is a real function of a complex variable hence not complex differentiable in the classical sense.

Based on this extensive analysis, a proof of the existence of a minimizer was given and the solutions of the optimal control problem were characterized by means of the first-order optimality system.

5. Numerical implementation and experiments

This chapter focuses on the numerical solution of TDKS optimization problems and consists of three parts: First, we discuss the discretization scheme used in our applications. We present an introduction to splitting methods and spectral methods. For selected models, stability and convergence results are provided. In Section 5.2, we introduce gradient based optimization schemes. The last section of this chapter contains various numerical experiments. We perform optimization experiments inspired from physical questions with different kinds of targets. Furthermore, we discuss L^2 versus H^1 optimization and analyse the dependence of the computational time on number of particles.

5.1. Discretization scheme

The Hamilton operator in the Schrödinger equation consists of two parts: the kinetic operator and the potential operator. These two terms are quite different. In fact, the kinetic operator is diagonal in the Fourier space which motivates the use of spectral methods. On the other hand, the potentials used in physical applications are usually local operators, hence represented by diagonal operators in real space.

The idea of operator splitting methods is to split the Hamilton operator into several parts, e.g., the kinetic and the potential operators, and calculate the time evolution caused by each part individually. In this section, we first describe the operator splitting methods in general and prove error estimates for simpler potentials. Then, we extend the splitting method to the nonlinear, and in the case of the adjoint equation inhomogeneous, TDKS equations.

In the first part, we consider the following one-dimensional Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi + V(x,t)\psi, \quad \psi(x,0) = \psi^0. \tag{5.1}$$

We denote the kinetic operator by $\hat{T} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$ and the potential by $V = V(x,t)$. The extension to several dimensions pursues the same arguments that follow. We remark that in this section we use SI units to clarify how the mass and the quantum scale $\hbar$ enters the discretization. The formal solution to (5.1) is given by

$$\psi(t) = e^{-\frac{i}{\hbar}(\hat{T}+V)t} \psi^0. \tag{5.2}$$

The idea of the splitting method is to split the exponential for a time step δt into s products of $e^{-\frac{i}{\hbar}\hat{T}\delta t}$ and $e^{-\frac{i}{\hbar}V\delta t}$ as follows:

$$\psi(x, t + \delta t) \approx \prod_{i=1}^{s} e^{-\frac{i}{\hbar}a_i \hat{T}\delta t} e^{-\frac{i}{\hbar}b_i V\delta t} \psi(x,t),$$

with appropriate coefficients $a_i, b_i \in \mathbb{R}$, $i = 1, \ldots, s$. The parameter s is called the number of compositions of the method. As $\hat{T}$ and V do not commute, a splitting error arises from this step.

The simplest splitting method is the Lie-Trotter splitting given by

$$\psi(x, t + \delta t) \approx e^{-\frac{i}{\hbar} V(x, t_n) \delta t} e^{-\frac{i}{\hbar} \hat{T} \delta t} \psi(x, t), \tag{5.3}$$

which is first-order accurate in time. The symmetric Lie-Trotter splitting or Strang splitting is given by

$$\psi(x, t + \delta t) \approx e^{-\frac{i}{\hbar} V(x, t_{n+1}) \frac{\delta t}{2}} e^{-\frac{i}{\hbar} \hat{T} \delta t} e^{-\frac{i}{\hbar} V(x, t_n) \frac{\delta t}{2}} \psi(x, t). \tag{5.4}$$

This scheme is second-order accurate in time.

Clearly, a larger s requires more evaluation of exponentials and transformations between real space and Fourier space and is thus more expensive. However, it can result in a higher order of accuracy p.

Operator splitting methods of different order and number of compositions can be derived using the Baker–Campbell–Hausdorff formula to expand the exponential of the sum $\hat{T} + V$ into a product of high order commutators, see [TCN09] for an overview of splitting methods with different order and number of compositions.

5.1.1. Time splitting spectral methods

In this section, we discuss the so-called time splitting spectral (TSSP) methods that result from the combination of the operator splitting schemes illustrated in the previous section and a spectral approximation of the kinetic operator $\hat{T}$, while the action of the potential is evaluated in real space.

The Schrödinger equation (5.1) is studied on $(x, t) \in \Omega \times (0, T)$, $\Omega = (0, L)$ with periodic boundary conditions in Ω. This domain is divided into N subintervals of length $h = \frac{L}{N}$ with end points $x_j = jh$, $j = 0, \ldots, N$ and the time is discretized as $t_n = n\delta t$, $n = 0, \ldots, M$, $\delta t = \frac{T}{M}$. For a given continuous periodic function u, consider the trigonometric polynomial

$$I_N u(x) = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{u}_k \, e^{i k \frac{2\pi x}{L}}, \tag{5.5}$$

where

$$\tilde{u}_k = \frac{1}{N} \sum_{j=1}^{N} u(x_j) \, e^{-i k \frac{2\pi x_j}{L}} \qquad \text{with} \quad x_j = (j-1) \, h. \tag{5.6}$$

The function $I_N u(x)$ is the $\frac{N}{2}$-degree trigonometric interpolant of u at the nodes x_j, i.e.

$$I_N u(x_j) = u(x_j) \qquad j = 1, \ldots, N. \tag{5.7}$$

This polynomial is the discrete Fourier series of u.

The interpolation operator I_N can be regarded as an orthogonal projection upon the space

$$S_N = \text{span} \left\{ e^{i k \frac{2\pi x}{L}} \, \middle| \, -\frac{N}{2} + 1 \leq k \leq \frac{N}{2} \right\}, \tag{5.8}$$

with respect to the discrete inner product

$$(u, v)_N = h \sum_{j=1}^{N} u(x_j) \, \overline{v(x_j)}, \tag{5.9}$$

where the overline denotes complex conjugation. In the following $\|\cdot\|_N = (\cdot,\cdot)_N^{1/2}$ denotes the norm associated to (5.9). Notice that for all $u, v \in S_N$, we have $(u, v)_N = \langle u, v \rangle_{L^2(\Omega;\mathbb{C})}$.

The Fourier pseudospectral derivative of u is defined by $D_N u = \frac{\mathrm{d}}{\mathrm{d}x}(I_N u)$. We have

$$D_N u(x) = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{u}_k' \, e^{i\,k\,\frac{2\pi x}{L}}$$

where

$$\tilde{u}_k' = \frac{i\,k}{N}\frac{2\pi}{L} \sum_{j=1}^{N} u(x_j)e^{-i\,k\,\frac{2\pi x_j}{L}}.$$

If $u \in S_N$ then $D_N u = \frac{\mathrm{d}u}{\mathrm{d}x}$. The operator D_N is skew-symmetric, in the sense that

$$(D_N u, v)_N = -(u, D_N v)_N \qquad \text{for all } u, v \in S_N.$$

Further, the operator D_N satisfies the product rule for differentiation. If $z = u\,v$ with $u, v \in S_N$, then $z \in S_N$ and

$$D_N z = D_N(u\,v) = v\,D_N u + u\,D_N v. \tag{5.10}$$

Applying D_N several times leads to the following formula for higher derivatives:

$$D_N^l u(x) = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{u}_k^l e^{ik\frac{2\pi}{L}x},$$

$$\tilde{u}_k^l = \frac{1}{N}\left(ik\frac{2\pi}{L}\right)^l \sum_{j=1}^{N} u(x_j)e^{-ik\frac{2\pi}{L}x_j}. \tag{5.11}$$

such that

$$D_N^l u = \frac{\mathrm{d}^l u}{\mathrm{d}x^l} \quad \text{for } u \in S_N. \tag{5.12}$$

In particular, the Laplacian is approximated by D_N^2. Notice that $-D_N^2$ is a symmetric positive definite operator on S_N, that is, $-D_N^2$ has positive eigenvalues and

$$\left\langle D_N^2 u, v \right\rangle_{L^2(\Omega;\mathbb{C})} = \left\langle u, D_N^2 v \right\rangle_{L^2(\Omega;\mathbb{C})},$$

and we have the bound $\|D_N^2\| \le c_D\, N^2$ for some constant c_D.

Stability of the splitting method

The rest of this subsection is based on [BJM02]. The first-order TSSP method to solve the Schrödinger equation implements the Lie-Trotter splitting with the following two steps:

1. Solve $i\hbar\frac{\partial\hat{\psi}}{\partial t} = -\frac{\hbar^2}{2m}\partial_{xx}^2\hat{\psi}$, with initial condition $\hat{\psi}(t_n) = \psi(t_n)$ in $(t_n, t_{n+1}]$.

2. Solve $i\hbar\frac{\partial\hat{\hat{\psi}}}{\partial t} = V(x,t_n)\hat{\hat{\psi}}$, with initial condition $\hat{\hat{\psi}}(t_n) = \hat{\psi}(t_{n+1})$ in $(t_n, t_{n+1}]$. Set $\psi(t_{n+1}) = \hat{\hat{\psi}}(t_{n+1})$.

For Step 1, we use the pseudo-spectral scheme to obtain the following ODE in Fourier space:

$$i\hbar \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \frac{\mathrm{d}\tilde{\tilde{\psi}}_k}{\mathrm{d}t} e^{ik\frac{2\pi}{L}x} = -\frac{\hbar^2}{2m} \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \left(-k^2 \frac{4\pi^2}{L^2} \tilde{\tilde{\psi}}_k(t)\right) e^{ik\frac{2\pi}{L}x}$$

$$\Leftrightarrow \quad i\frac{\mathrm{d}\tilde{\tilde{\psi}}_k}{\mathrm{d}t} = \frac{4\pi^2\hbar k^2}{2mL^2} \tilde{\tilde{\psi}}_k(t), \quad \forall k \in \left[-\left(\frac{N}{2}-1\right), \frac{N}{2}\right].$$

This ODE can be integrated exactly, and we obtain the following solution, first in the Fourier space and then in the real space:

$$\tilde{\tilde{\psi}}_k(t_{n+1}) = \tilde{\tilde{\psi}}_k(t_n) e^{-i\frac{4\pi^2\hbar k^2}{2mL^2}\delta t},$$

$$\hat{\psi}(x, t_{n+1}) = \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{\tilde{\psi}}_k(t_n) e^{-i\frac{4\pi^2\hbar k^2}{2mL^2}\delta t} e^{ik\frac{2\pi}{L}x}.$$

Step 2 is done in real space. For a time-independent potential, we have

$$\psi(x, t_{n+1}) = \hat{\hat{\psi}}(x, t_{n+1}) = \hat{\psi}(x, t_{n+1}) e^{-\frac{i}{\hbar}V(x)\delta t}.$$

This TSSP scheme is unitary as stated in the following lemma.

Lemma 5.1. *The Lie-Trotter* (5.3) *and the Strang* (5.4) *TSSP schemes are norm preserving, and thus unconditionally stable for a time-dependent potential* $V(x,t)$, *in the sense that*

$$\|\psi^n(\cdot)\|_N = \|\psi^0(\cdot)\|_N, \quad \forall n \in \mathbb{N},$$

$$\|I_N\psi^n(\cdot)\|_{L^2(\Omega;\mathbb{C})} = \|I_N\psi^0(\cdot)\|_{L^2(\Omega;\mathbb{C})}, \quad \forall n \in \mathbb{N}.$$

Proof. The proof is based on the fact that $|e^{iV(x,t)}| = 1$ for any real potential $V(x,t)$. We have

$$\|\psi(\cdot, t_{n+1})\|_N^2 = h\sum_{j=1}^{N} |\psi(x_j, t_{n+1})|^2 = h\sum_{j=1}^{N} |e^{-\frac{i}{\hbar}V(x_j,t_n)\delta t}\hat{\psi}(x_j, t_{n+1})|^2$$

$$= h\sum_{j=1}^{N} |\hat{\psi}(x_j, t_{n+1})|^2 = h\sum_{j=1}^{N} \left| \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \tilde{\tilde{\psi}}_k(t_n) e^{ik\frac{2\pi}{L}x_j} e^{-i\frac{4\pi^2\hbar k^2}{2mL^2}\delta t} \right|^2.$$

Now, expanding the square and using the fact that

$$\sum_{j=1}^{N} e^{i(l_1-l_2)\frac{2\pi}{L}x_j} = \begin{cases} 0 & l_1 - l_2 \neq sN, \\ N & l_1 - l_2 = sN, \end{cases} \quad s \in \mathbb{Z},$$

we obtain the following:

$$\|\psi(\cdot, t_{n+1})\|_N^2 = hN \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \left| \tilde{\tilde{\psi}}_k(t_n) e^{-i\frac{4\pi^2\hbar k^2}{2mL^2}\delta t} \right|^2 = hN \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \left| \tilde{\tilde{\psi}}_k(t_n) \right|^2$$

$$= hN \sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} \left| \frac{1}{N} \sum_{j=1}^{N} \hat{\psi}(x_j, t_n) e^{-ik\frac{2\pi}{L}x_j} \right|^2.$$

Again expanding the square and using

$$\sum_{k=-\left(\frac{N}{2}-1\right)}^{\frac{N}{2}} e^{ik(x_{l_1}-x_{l_2})\frac{2\pi}{L}} = \begin{cases} 0 & l_1 - l_2 \neq sN, \\ N & l_1 - l_2 = sN, \end{cases} \quad s \in \mathbb{Z},$$

we arrive at

$$\|\psi(\cdot, t_{n+1})\|_N^2 = h\sum_{j=1}^{N} \left|\hat{\psi}(x_j, t_n)\right|^2 = \|\psi(\cdot, t_n)\|_N^2.$$

Since $\left|\hat{\psi}(x_j, t_n)\right|^2 = \left|\hat{\psi}(x_j, t_n)e^{-\frac{i}{\hbar}V(x_j, t_{n+1})\frac{\delta t}{2}}\right|^2$, the same result holds for the Strang splitting. $\qquad \square$

Error estimate for a constant potential

To obtain an error estimate for the TSSP scheme, we first consider the case of a constant potential $V(x, t) = V$. Furthermore, we assume that the exact solution ψ of the SE is in $C([0, T]; C^k(\mathbb{R}))$, $k \geq 2$, and periodic in $\Omega = (0, L)$. Since the initial condition is a C^k-function, all its m-order derivatives up to order k are bounded functions in the bounded interval Ω; for ease of notation, we denote this bound C_m as follows:

$$\left\|\frac{\mathrm{d}^j}{\mathrm{d}x^j}\psi^0(x)\right\|_{L^2(\Omega;\mathbb{C})} \leq C_j, \quad j = 0, \ldots, k. \tag{5.13}$$

With this preparation, we can state the following theorem.

Theorem 5.2. *Let $\psi \in C([0, T]; C^k(\mathbb{R}))$, $k \geq 2$, be the exact solution to the Schrödinger equation (5.1) for constant potential $V(x, t) = V$ and initial condition $\psi^0 \in C^k(\mathbb{R})$ with ψ and ψ^0 periodic in Ω. Denote by ψ^n the numerical solution obtained with the Lie-Trotter or Strang TSSP methods. Then using the notation of (5.13), there exists a constant $C > 0$ such that the following holds:*

$$\|I_N\psi^n - \psi(\cdot, t_n)\|_{L^2(\Omega;\mathbb{C})} \leq CC_k\left(\frac{h}{L}\right)^k.$$

Proof. For the proof, we need the following interpolation estimate from [Pas80, Theorem 3]: Given any $k > \frac{d}{2}$ (where $d = 1$ in our case), then there exists a constant C, such that for all $\phi \in H^k$ the following estimate holds for all $0 \leq j \leq k$:

$$\|I_N\phi - \phi\|_{H^j(\Omega;\mathbb{C})} \leq CN^{j-k}\left\|\frac{\mathrm{d}^k\phi}{\mathrm{d}x^k}\right\|_{L^2(\Omega;\mathbb{C})}. \tag{5.14}$$

We apply this estimate to the initial data ψ^0, with $j = 0$, to obtain

$$\|I_N\psi(\cdot, t_0) - \psi(\cdot, t_0)\|_{L^2(\Omega;\mathbb{C})} \leq C\frac{1}{N^k}\left\|\frac{\mathrm{d}^k}{\mathrm{d}x^k}\psi(\cdot, t_0)\right\|_{L^2(\Omega;\mathbb{C})}.$$

Using (5.13) and $\frac{1}{N} = \frac{h}{L}$, we have

$$\|I_N\psi^0 - \psi^0\|_{L^2(\Omega;\mathbb{C})} \leq CC_k\left(\frac{h}{L}\right)^k.$$

Since $I_N\psi^n$ is an exact solution to the SE (5.1) subject to the initial condition $I_N\psi^0$ and $\psi(t_n)$ is the exact solution to the SE with initial data ψ^0 and the SE generates a unitary group, see Lemma 5.1, the above estimate holds for all t_n. $\qquad \square$

Error estimates for time-independent space-variable potentials

In the case of the Lie-Trotter splitting, we show an estimate similar to the one above for the case of a time-independent periodic $C^\infty(\mathbb{R})$ potential $V = V(x)$. To this end, we assume that for all $j \in \mathbb{N}$ there exists positive constants C_j, F_j independent on x and t such that the following bounds on the solution of the SE hold:

$$\left\| \frac{\partial^{j_1+j_2}}{\partial x^{j_1} \partial t^{j_2}} \psi \right\|_{C([0,T];L^2(\Omega;\mathbb{C}))} \leq C_{j_1+j_2}, \qquad \left\| \frac{\mathrm{d}^j V}{\mathrm{d}x^j} \right\|_{L^\infty(\Omega;\mathbb{C})} \leq F_j. \qquad (5.15)$$

Theorem 5.3. *Let $\psi \in C((0,T); C^k(\mathbb{R}))$ be the exact solution of the SE (5.1) with periodic $V \in C^\infty(\mathbb{R})$, and let ψ^n represent the numerical Lie-Trotter TSSP solution. Under the assumption that (5.15) is satisfied, there exists a positive constant C_t independent on h, δt, and j and constants E_j independent on h and δt such that for all positive integers $1 \leq j \leq k$ and $t_n \in [0,T]$ the following holds:*

$$\|\psi(\cdot, t_n) - I_N \psi^n\|_{L^2(\Omega)} \leq E_j \frac{T}{\delta t N^j} + C_t T \delta t.$$

Proof. First, we want to estimate the time splitting error. For this purpose, we define the operators

$$A = i\frac{\hbar}{2m}\partial_{xx}^2, \qquad\qquad B = -\frac{i}{\hbar}V(x).$$

Then the exact and the splitting evolution are given by

$$\psi(x, t_{n+1}) = e^{(B+A)\delta t}\psi(x, t_n),$$
$$\psi_{SP}(x, t_{n+1}) = e^{B\delta t}e^{A\delta t}\psi(x, t_n).$$

Notice that ψ_{SP} contains the splitting error, while the differential operator is implemented exactly. On the other hand, ψ^n is the solution obtained using the splitting and the approximated Laplacian from (5.11).

Using the exponential series up to third order for $e^{(B+A)\delta t}$ and $e^{B\delta t}e^{A\delta t}$, we have

$$\psi_{SP}(x, t_{n+1}) - \psi(x, t_{n+1}) = \left(e^{B\delta t}e^{A\delta t} - e^{(B+A)\delta t}\right)\psi(x, t_n)$$
$$= \frac{\delta t^2}{2}(BA - AB)\psi(x, t_n) + \mathcal{O}(\delta t^3)$$
$$\text{with } \frac{\delta t^2}{2}(BA - AB)\psi(x, t) = \frac{\delta t^2}{4m}\left(\psi(x,t)\partial_{xx}^2 V + 2\partial_x\psi(x,t)\partial_x V\right).$$

With (5.15), we conclude for all $t \in [0, T]$

$$\|\psi_{SP}(\cdot, t) - \psi(\cdot, t)\|_{L^2(\Omega;\mathbb{C})} \leq C_t \delta t^2. \qquad (5.16)$$

Now, we estimate the difference between the exact solution and the interpolation of the approximated solution after one time step at the new time t_{n+1} using the triangle inequality as follows:

$$\|\psi(\cdot, t_{n+1}) - I_N \psi^{n+1}\|_{L^2(\Omega)}$$
$$\leq \|\psi(\cdot, t_{n+1}) - \psi_{SP}(\cdot, t_{n+1})\|_{L^2(\Omega)} + \|\psi_{SP}(\cdot, t_{n+1}) - I_N \psi_{SP}(\cdot, t_{n+1})\|_{L^2(\Omega)} \qquad (5.17)$$
$$+ \|I_N \psi_{SP}(\cdot, t_{n+1}) - I_N \psi^{n+1}\|_{L^2(\Omega)}.$$

For the last term, we recall the definition of ψ_{SP}. From (5.12), we have that the action of A and of $\frac{i\hbar}{2m}D_N^2$ coincide for all $\psi \in S_N$. Further, recall the norm preservation property given in Lemma 5.1. We use all these properties in the following computation.

$$
\begin{aligned}
\|I_N \psi_{SP}(\cdot, t_{n+1}) - I_N \psi^{n+1}\|_{L^2(\Omega;\mathbb{C})} &= \|\psi_{SP}(\cdot, t_{n+1}) - \psi^{n+1}\|_N \\
&= \left\| e^{\delta t B}\left(e^{\delta t A}\psi(t_n) - e^{\delta t \frac{i\hbar}{2m}D_N^2}\psi^n \right) \right\|_N \\
&= \left\| e^{\delta t B}\left(e^{\delta t A}\psi(t_n) - e^{\delta t A}\psi^n \right) \right\|_N \\
&= \|\psi(t_n) - \psi^n\|_N \\
&= \|I_N \psi(t_n) - I_N \psi^n\|_{L^2(\Omega;\mathbb{C})} \\
&\leq \|I_N \psi(t_n) - \psi(t_n)\|_{L^2(\Omega;\mathbb{C})} + \|\psi(t_n) - I_N \psi^n\|_{L^2(\Omega;\mathbb{C})}.
\end{aligned}
$$

This gives

$$
\begin{aligned}
\|\psi(\cdot, t_{n+1}) - I_N \psi^{n+1}\|_{L^2(\Omega;\mathbb{C})} &\leq \|\psi(\cdot, t_{n+1}) - \psi_{SP}(\cdot, t_{n+1})\|_{L^2(\Omega;\mathbb{C})} \\
&\quad + \|\psi_{SP}(\cdot, t_{n+1}) - I_N \psi_{SP}(\cdot, t_{n+1})\|_{L^2(\Omega;\mathbb{C})} + \|I_N \psi(t_n) - \psi(t_n)\|_{L^2(\Omega;\mathbb{C})} \\
&\quad + \|\psi(t_n) - I_N \psi^n\|_{L^2(\Omega;\mathbb{C})}.
\end{aligned}
\tag{5.18}
$$

The first term is the splitting error estimated in (5.16). Introducing the constant E_j, the third term can be bounded using the interpolation estimate (5.14) and the assumption (5.15),

$$
\|I_N \psi(t_n) - \psi(t_n)\|_{L^2(\Omega;\mathbb{C})} \leq CC_j N^{-j} \leq \frac{1}{2}E_j N^{-j}.
\tag{5.19}
$$

The second term on the right hand side of (5.18) can be reduced to the same bound via (5.14) as follows:

$$
\begin{aligned}
\|I_N \psi_{SP}(t_n) - \psi_{SP}(t_n)\| &\leq CN^{-j}\left\| \frac{\mathrm{d}^j}{\mathrm{d}x^j}\psi_{SP}(t_{n+1}) \right\|_{L^2(\Omega;\mathbb{C})} \\
&= CN^{-j}\left\| \sum_{l=0}^{j} \binom{j}{l} \frac{\mathrm{d}^l e^B}{\mathrm{d}x^l} \frac{\mathrm{d}^{j-l} e^A \psi(t_n)}{\mathrm{d}x^{j-l}} \right\|_{L^2(\Omega;\mathbb{C})} \\
&\leq CN^{-j} \sum_{l=0}^{j} \binom{j}{l} \left\| \frac{\mathrm{d}^l e^B}{\mathrm{d}x^l} \right\|_{L^\infty} \left\| \frac{\mathrm{d}^{j-l} e^A \psi(t_n)}{\mathrm{d}x^{j-l}} \right\|_{L^2(\Omega;\mathbb{C})} \\
&\leq \frac{1}{2}E_j N^{-j}.
\end{aligned}
\tag{5.20}
$$

Finally, combining (5.16), (5.19), and (5.20) in (5.18), we obtain

$$
\|\psi(\cdot, t_{n+1}) - I_N \psi^{n+1}\|_{L^2(\Omega;\mathbb{C})} \leq E_j \frac{1}{N^j} + C_t \delta t^2 + \|\psi(\cdot, t_n) - I_N \psi^n\|_{L^2(\Omega;\mathbb{C})}.
$$

By induction over the time steps the claim follows. $\qquad\square$

5.1.2. Application to nonlinear Schrödinger equations

We have already remarked that TSSP schemes allow to apply different efficient approximation schemes to the different parts of a quantum Hamiltonian. In particular, the action of potentials is easily implemented in the real space and this is also true for implementing nonlinearities.

To show this fact, we discuss the application of TSSP schemes for solving the following nonlinear Schrödinger equation, which is called the Gross-Pitaevskii equation, see [Pit61, HRBS07, BCS17]. We have

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\partial^2_{xx}\psi + V(x,\psi)\psi, \quad \psi(x,0) = \psi^0(x),$$

$$\text{with } V(x,\psi) = V_0(x) + g|\psi(x,t)|^2, \quad g \in \mathbb{R}, \tag{5.21}$$

where the nonlinearity is interpreted as an additional potential. We remark that (5.21) is similar to a TDKS equation insofar as the nonlinearity is given by a function of the density $\rho = |\psi|^2$.

For nonlinear equations, we use an intermediate wavefunction ψ^* to evaluate the nonlinear potential in the last step of the TSSP procedure. In the case of the Lie-Trotter splitting, this technique is illustrated as follows:

$$\psi^{n+1} = e^{-\frac{i}{\hbar}V(x,\psi^*)\delta t} \underbrace{e^{\frac{i\hbar}{2m}D_N^2\delta t}\psi^n}_{=:\psi^*}. \tag{5.22}$$

In the case of the Strang splitting, we have

$$\psi^{n+1} = e^{-\frac{i}{\hbar}V(x,\psi^*)\frac{\delta t}{2}} \underbrace{e^{\frac{i\hbar}{2m}D_N^2\delta t}e^{-\frac{i}{\hbar}V(x,\psi^n)\frac{\delta t}{2}}\psi^n}_{=:\psi^*}. \tag{5.23}$$

The detailed procedure for the Strang splitting method applied to a nonlinear SE is presented in Algorithm 1 on page 86.

For the analysis of TSSP schemes applied to a nonlinear SE, we refer to [Tha12]. In particular, we mention a theorem from [Tha12] concerning the Lie-Trotter and Strang TSSP schemes applied to (5.21). For this purpose, we make the following preparation. Consider the eigenvalue representation of the Laplacian

$$-\nabla^2 B_\mu = \lambda_\mu B_\mu, \quad \mu \in \mathbb{Z}^n,$$

with the eigenvalues λ_μ and the eigenfunctions $B_\mu \in L^2(\Omega)$. Then, we define the application of the α-th power of $-\nabla^2$ on a function v as follows:

$$(-\nabla^2)^\alpha v := \sum_{\mu\in\mathbb{Z}^n} \langle v, B_\mu\rangle_{L^2} \lambda_\mu^\alpha B_\mu.$$

We define the Hilbert space X_α given by

$$X_\alpha = \left\{ v = \sum_{\mu\in\mathbb{Z}^n} \langle v, B_\mu\rangle_{L^2} B_\mu \in L^2(\Omega) \,\middle|\, \|(-\nabla^2)^\alpha v\|_{L^2} < \infty \right\},$$

and the following norm:

$$\|\cdot\|_{X_\alpha} := \|(-\nabla^2)^\alpha \cdot\|_{L^2}.$$

With this preparation, we recall the following theorem from [Tha12].

Theorem 5.4. *Consider a bounded domain $\Omega \subset \mathbb{R}^d$ in $d \geq 1$ dimensions and let $p = 1$ for the Lie-Trotter splitting and $p = 2$ for the Strang splitting. If the potential V and the analytic solution ψ of (5.21) are bounded in X_α for $\alpha \geq \max\{\frac{\delta}{2}, p\}$, where $\delta \in \mathbb{N}$ is such that the Sobolev embedding $H^\delta(\Omega) \hookrightarrow C(\Omega)$ holds, then the global error is given by*

$$\|\psi^n - \psi(t_n)\|_{L^2(\Omega)} \leq C\left(\|\psi^0 - \psi(0)\|_{L^2(\Omega)} + \delta t^p + N^{-2(\alpha-1)}\right).$$

The constant C depends on d, g, $\|V\|_{X_\alpha}$, and $\sup_{0\leq t\leq T}\|\psi(t)\|_{X_\alpha}$.

This result shows that the Lie-Trotter splitting provides first-order accuracy in time and the Strang splitting is second-order accurate in time.

Both methods profit of the Fourier spectral method to achieve spectral convergence in space as the spacial error is bounded by the regularity of the exact solution in an exponential way.

We remark that the order of accuracy in time of the Lie-Trotter and Strang methods is independent of the order of application of the kinetic and spatial operators. However, a smaller error can be achieved if the nonlinear part in V is evaluated last. This is explained in [BBD02] where general SE nonlinearities that are Lipschitz continuous with respect to ψ are discussed.

Considering the Hartree equation, which is a nonlinear SE with Hartree interaction but without exchange and correlation potential, see also (2.9) in Section 2.2.2, given by

$$
i\frac{\partial \psi}{\partial t}(x,t) = -\nabla^2 \psi(x,t) + \int_{\mathbb{R}^3} \frac{|\psi(y)|^2}{|x-y|} \mathrm{d}y \psi(x,t),
$$
$$
\psi(x,0) = \psi^0(x) \in \mathbb{C}, \quad x \in \mathbb{R}^3, \quad \lim_{|x|\to\infty} \psi(x,t) = 0 \ \forall t > 0,
$$
(5.24)

we recall the following theorem from [Lub08] for the Strang splitting applied to (5.24).

Theorem 5.5. *Suppose that the exact solution $\psi(t)$ to the Hartree equation (5.24) is in $H^4(\mathbb{R}^3;\mathbb{C})$ for $0 \le t \le T$. Then, the numerical solution ψ^n given by the Strang splitting scheme (5.23) with step size $\delta t > 0$ has a first-order error bound in H^1 and a second-order error bound in L^2 as follows:*

$$
\|\psi^n - \psi(t_n)\|_{H^1(\mathbb{R}^3)} \le C(m_3, T)\delta t,
$$
$$
\|\psi^n - \psi(t_n)\|_{L^2(\mathbb{R}^3)} \le C(m_4, T)\delta t^2,
$$

for $t_n = n\delta t$, and $m_k = \max_{0 \le t \le T} \|\psi(t)\|_{H^k(\mathbb{R}^3;\mathbb{C})}$.

5.1.3. Inhomogeneous Schrödinger equations

In the optimal control framework discussed in Chapter 4, adjoint equations appear that contain inhomogeneities. In this case, we need to extend the TSSP schemes to accommodate these additional terms. To discuss this extension, we consider the following SE problem:

$$
i\hbar\frac{\partial \psi}{\partial t} = \frac{\hbar^2}{2m}\partial_{xx}^2 \psi + V(x,\psi)\psi + f(x,t), \quad \psi(x,0) = \psi^0(x), \tag{5.25}
$$

where $f \in C^2([0,T], L^2(\Omega;\mathbb{C}))$.

A natural extension of the Strang TSSP scheme is given by

$$
\psi^{n+1} = e^{-\frac{i}{\hbar}V(x,t_{n+1},\psi^*)\frac{\delta t}{2}} \underbrace{e^{\frac{i\hbar}{2m}D_N^2\frac{\delta t}{2}}\left(e^{\frac{i\hbar}{2m}D_N^2\frac{\delta t}{2}}e^{-\frac{i}{\hbar}V(x,t_n,\psi^n)\frac{\delta t}{2}}\psi^n - i\delta t f\left(t + \frac{\delta t}{2}\right)\right)}_{=\psi^*}.
$$
(5.26)

This formula is motivated by the following discussion.

Consider the inhomogeneous SE equation (5.25) in the form

$$
i\frac{\mathrm{d}\psi(t)}{\mathrm{d}t} = (A + B)\psi(t) + f(t).
$$

Using the formula for the variation of constants, the solution to this equation is given by

$$\psi(t) = e^{-i(A+B)t}\psi(0) - i\int_0^t e^{-i(t-\tau)(A+B)}f(\tau)\mathrm{d}\tau$$

$$= e^{-i(A+B)t}\psi(0) - ie^{-it(A+B)}\int_0^t e^{i\tau(A+B)}f(\tau)\mathrm{d}\tau.$$

Now, to evaluate the integral, we apply the midpoint rule as follows:

$$\int_a^b g(\tau)\mathrm{d}\tau = (b-a)g\left(a+\frac{b-a}{2}\right) + \mathcal{O}\left((b-a)^3\right).$$

In this way, we obtain

$$\psi(t_{n+1}) \approx e^{-i(A+B)\delta t}\psi(t_n) - i\delta t\left(e^{-i\frac{\delta t}{2}(A+B)}f(t_{n+1/2})\right)$$

$$= e^{-i\frac{\delta t}{2}(A+B)}\left(e^{-i\frac{\delta t}{2}(B+A)}\psi(t_n) - i\delta t f(t_{n+1/2})\right).$$

Because the function f is evaluated in the midpoint, this splitting method (5.26) is still second-order accurate in time and has analytic convergence in space if the spectral derivatives are used. A more detailed discussion on the Lie-Trotter and Strang splitting schemes in the case of inhomogeneous parabolic equations is given in [FOS15].

5.1.4. Discretization scheme for TDKS equations

Following the general discussion above, we now focus on the Strang splitting applied to the TDKS equations. To solve the TDKS equations (4.8a), we use the Strang splitting (5.23) given by

$$\psi_j^* = e^{i\delta t\nabla^2}e^{-i\frac{\delta t}{2}V(\Psi(t),t)}\psi_j(t),$$

$$\psi_j(t+\delta t) = e^{-i\frac{\delta t}{2}V(\Psi^*,t+\delta t)}\psi_j^*, \tag{5.27}$$

where $V = V_{ext} + V_{Hxc}$ and $j = 1,\ldots,N$. To solve the adjoint TDKS equations (4.8c) we have to include the inhomogeneous right-hand side as follows:

$$\psi_j^* = e^{-i\frac{\delta t}{2}\nabla^2}\left(e^{-i\frac{\delta t}{2}\nabla^2}e^{i\frac{\delta t}{2}V(\Psi(t),t)}\psi_j(t) + i\delta t g_j(t-\delta t/2)\right),$$

$$\psi_j(t-\delta t) = e^{i\frac{\delta t}{2}V(\Psi^*,t-\delta t)}\psi_j^*, \tag{5.28}$$

with a right-hand side $g(t)$ given by

$$g_j = 2\beta(\rho - \rho_d)\psi_j + \sum_{l=1}^N \left(V_H\left(2\operatorname{Re}(\psi_l\overline{\lambda_j})\right) + 2\frac{\partial V_{xc}}{\partial\rho}(\rho)\operatorname{Re}(\psi_l\overline{\lambda_l})\right)\psi_j.$$

The Laplacian $-\nabla^2$ is evaluated spectrally as described in section 5.1.1. With this setting, we obtain a discretization scheme that provides second-order convergence in time and analytic convergence in space; see the Algorithms 1 and 2. Furthermore, this time-splitting scheme is unconditional stable and norm preserving as well as time reversible and gauge invariant. The latter means that adding a constant to the potential changes

only the phase of the wavefunction in such a way that discrete quadratic observables are not changed, see also, e.g., [BJM02].

The convergence analysis discussed in the previous Subsections 5.1.1 and 5.1.2 is for problems similar to the TDKS equations. Even though these results cannot be applied directly to our TDKS problem, they suggest that similar accuracy can be expected in our case. Therefore, we study this accuracy issue numerically. To this end, we solve the TDKS equation for two interacting particles in a harmonic trap $V_{ext} = \left(40 + 20\frac{t}{T}\right) x^2$. The used initial condition is given by the coherent states of two non-interacting particles in the harmonic oscillator, see Appendix A.3.1. We set $\Omega = (0, L)^2$, $L = 7$, and the time interval is $[0, 0.001]$. Since an analytic solution is not available, we consider a reference solution Ψ_{ref} obtained solving the problem on a very fine mesh ($dt_{\text{ref}} = 1.3021 \cdot 10^{-7}$ for time convergence and $h_{\text{ref}} = 2.9167 \cdot 10^{-2}$ for space convergence). In Figure 5.1, we report the results for the forward equation and in Figure 5.2 for the adjoint equation. The results confirm the hypothesis of second-order convergence in space (slope factors of 1.97 and 2.01) and spectral convergence in space. For the smallest values of h in Figures 5.1b and 5.2b the error is dominated by the time error (compare to the values of the error in the figures on the left) which explains why the analytic behaviour is hidden there.
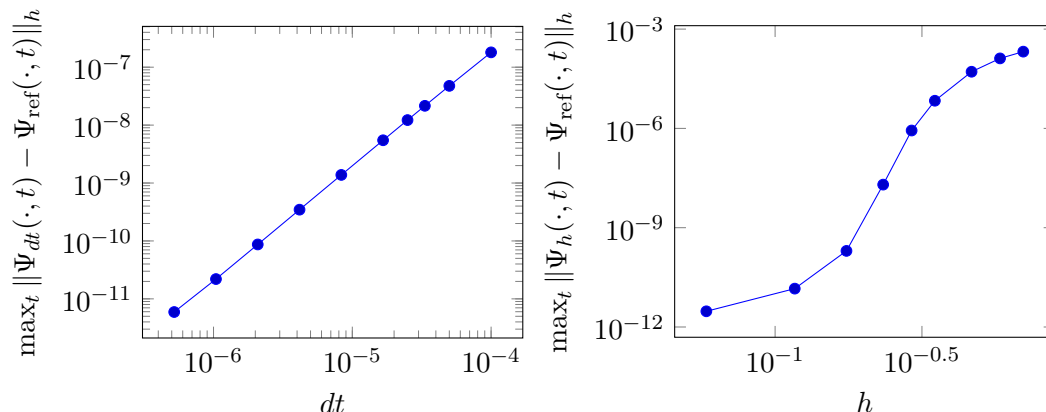
(a) The error for different time steps dt and fixed space steps $h = 0.1167$. The interpolation gives a slope factor of 1.97 confirming the theoretical convergence order of $p = 2$.

(b) The error for varying mesh size h and fixed time step $dt = 8.33 \cdot 10^{-7}$. The analytical convergence can clearly be seen apart from the smallest values of h where the time error dominates.

Figure 5.1.: Accuracy of the numerical solution of the forward TDKS equation.

Now, we present the implementation of the splitting method in the Algorithms 1 and 2. We consider a space grid of $N_x \times N_y$ points at the times $0 = t_1 < t_2 < \cdots < t_{M+1} = T$. The wavefunction Ψ is then stored in a 4-dimensional array of size $N_x \times N_y \times N \times (M + 1)$, $\Psi_j = \Psi(t_j) \in \mathbb{C}^{N_x \times N_y \times N}$, and $\psi_{k,j}$ is the matrix representing the k-th DFT wavefunction at time t_j.
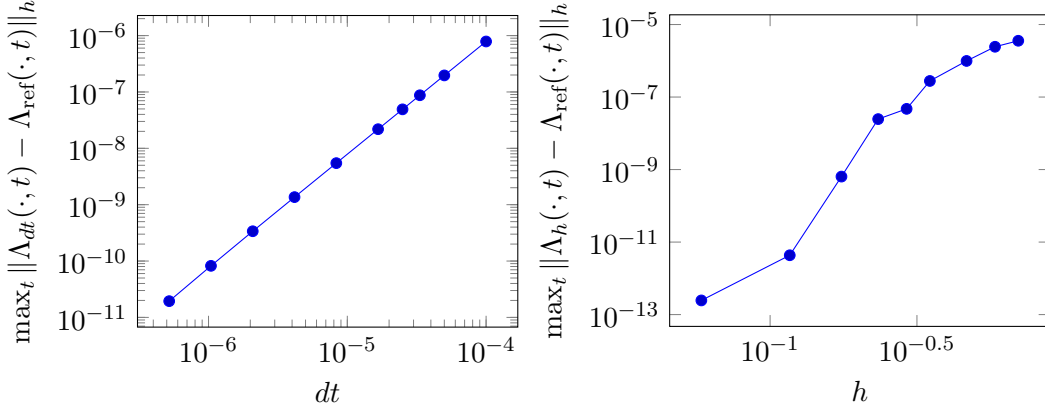
(a) The error for different time steps dt and fixed space steps $h = 9.1146 \cdot 10^{-4}$. The interpolation gives a slope factor of 2.01 confirming the theoretical convergence order of $p = 2$.

(b) The error for varying mesh size h and fixed time step $dt = 7.1429 \cdot 10^{-7}$. The analytical convergence can clearly be seen apart from the smallest values of h where the time error dominates.

Figure 5.2.: Accuracy of the numerical solution of the adjoint TDKS equation.

Algorithm 1 Strang splitting method (forward equation)

Input: Initial wavefunction $\Psi_0 \in \mathbb{C}^{N_x \times N_y \times N}$, potential $V(x, t, \Psi)$;
Output: Wavefunction $\Psi_j \in \mathbb{C}^{N_x \times N_y \times N}$, $j = 1, \ldots M$;

1: Set $k_x \leftarrow \frac{4\pi^2}{L_x^2}\left(0^2, 1^2 \ldots, \left(\frac{N_x}{2} - 1\right)^2, \left(\frac{N_x}{2}\right)^2, \left(-\frac{N_x}{2} + 1\right)^2, \ldots, (-2)^2, (-1)^2\right)$,

   $k_y \leftarrow \frac{4\pi^2}{L_y^2}\left(0^2, 1^2, \ldots, \left(\frac{N_y}{2} - 1\right)^2, \left(\frac{N_y}{2}\right)^2, \left(-\frac{N_y}{2} + 1\right)^2, \ldots, (-2)^2, (-1)^2\right)$,

   $K \leftarrow \underbrace{(1, \ldots, 1)^T}_{N_y \text{ times}} k_x + k_y^T \underbrace{(1, \ldots, 1)}_{N_x \text{ times}}$;

2: Assemble the matrix of the Laplace operator in Fourier space via element-wise exponentiation $(H_0)_{ij} \leftarrow \exp\left(-iK_{ij}\delta t\right)$;

3: **for** all timesteps $j = 0$ **to** $M - 1$ **do**

4:     Set potential $V \leftarrow V_{ext}(x, t_j, u_j) + V_{Hxc}(\rho(\Psi_j))$;

5:     **for** all particles $k = 1$ **to** N **do**

6:         $\psi_k^* \leftarrow \mathcal{FFT}\left(H_0 \mathcal{FFT}^{-1}\left[\exp\left(-iV\frac{\delta t}{2}\right)\psi_{k,j}\right]\right)$;

7:     **end for**

8:     Update V with intermediate Ψ^*, $V \leftarrow V_{ext}(x, t_{j+1}, u_{j+1}) + V_{Hxc}(\rho(\Psi^*))$;

9:     **for** all particles $k = 1$ **to** N **do**

10:        $\psi_{k,(j+1)} \leftarrow \exp\left(-iV\frac{\delta t}{2}\right)\psi_k^*$;

11:     **end for**

12: **end for**

In Algorithm 1, we begin with assembling the matrix representation of the Laplace operator in Fourier space (line 1) and store its exponential (line 2). The main task is done in line 6, where the wavefunction is first propagated in real space using the potential for half a time step. Then the FFT of Ψ is calculated, multiplied with the exponential of the Laplacian and transformed back to real space. In line 10, a second half-time step is performed using the potential that has been updated using the solution of the previous half-time step. Algorithm 2 is similar apart from the inclusion of the inhomogeneous right hand side g.

We remark that the potential V is diagonal in real space and the Laplace is diagonal in Fourier space. For these reasons, evaluation of the exponential matrices is computationally cheap because they are computed element-wise.

Algorithm 2 Strang splitting method (adjoint equation)

Input: Potential $V(x, t, \Psi)$, solution of forward equation $\Psi \in \mathbb{C}^{N_x \times N_y \times N \times (M+1)}$, terminal condition $\Lambda^T \in \mathbb{C}^{N_x \times N_y \times N}$;
Output: Adjoint variable $\Lambda_j \in \mathbb{C}^{N_x \times N_y \times N}$, $j = 0, \ldots, M-1$;

1: Define H_0 as in Algorithm 1, lines 1 and 2 replacing δt with $\frac{\delta t}{2}$;
2: **for** all timesteps $j = M - 1$ **to** 0 **do**
3:     Set potential $V \leftarrow V_{ext}(x, t_{j+1}, u_{j+1}) + V_{Hxc}(\rho(\Psi_{j+1}))$;
4:     **for** all particles $k = 1$ **to** N **do**
5:         $\lambda_k^* \leftarrow \mathcal{FFT}\left(H_0 \mathcal{FFT}^{-1}\left[\exp\left(-iV\frac{\delta t}{2}\right)\lambda_{k,(j+1)}\right]\right)$;
6:     **end for**
7:     Update V with new Ψ, $V \leftarrow V_{ext}(x, t_j, u_j) + V_{Hxc}(\rho(\Psi_j))$;
8:     **for** all particles $k = 1$ **to** N **do**
9:         Set $g_k \leftarrow 2\beta(\rho(\Psi') - \rho_d)\psi_k' + \sum_{l=1}^{N}\left(V_H\left(2\operatorname{Re}(\psi_l'\overline{\lambda_j^*})\right) + 2\frac{\partial V_{xc}}{\partial \rho}(\rho(\Psi'))\operatorname{Re}(\psi_l'\overline{\lambda_l^*})\right)\psi_k'$
        with $\Psi' = \frac{\Psi_j + \Psi_{j+1}}{2}$;
10:     $\lambda_{k,j} \leftarrow \exp\left(-iV\frac{\delta t}{2}\right)\mathcal{FFT}\left(H_0 \mathcal{FFT}^{-1}\left(\lambda_k^* + i\delta t\, g_k\right)\right)$;
11:     **end for**
12: **end for**

5.1.5. The Hartree potential

It is difficult to determine the Hartree potential V_H efficiently as it has a global dependence on ρ. For this reason, we use Algorithm 3, which is based on the idea that the convolution $\rho \star \frac{1}{|x|}$ can be obtained in Fourier space by multiplication. This method is described in [CRR09] and briefly illustrated below. An alternative approach in three dimensions is to solve the Poisson equation because in $3d$ (and only there) the Hartree potential is the Green's function of the Poisson equation with the density ρ as source term.

Algorithm 3 Hartree potential

Input: Grid parameters N_x, N_y, h, density ρ
Output: Hartree potential V_H;

1: Set $N_{max} = 2\max\{N_x, N_y\}$, extend ρ by zero to $\rho_{extended}$ on a square of size $N_{max} \times N_{max}$;

2: $R \leftarrow \frac{N_{max}h}{2}$;
3: **for** $k = \frac{2\sqrt{2}\pi}{RN_{max}}\frac{i}{I}$, $i = 1, \ldots I$ **do**
4:     $\bar{v}_H(k) \leftarrow 2\pi\int_0^R J_0(rk)\mathrm{d}r$, J_0 is Bessel function of the first kind;
5: **end for**
6: Create spline interpolant $\hat{v}_H$ from $\bar{v}_H$ and set $\hat{v}_H(0) \leftarrow 2\pi R$;
7: **for all** $(k_x, k_y) \in \frac{\pi}{R}(l_x, l_y)$, $l_x, l_y = 0, \ldots, N_{max}$ **do**
8:     $\hat{V}_H(k_x, k_y) \leftarrow \hat{v}_H\left(\sqrt{k_x^2 + k_y^2}\right)$;
9: **end for**
10: **return** $\mathcal{FFT}^{-1}(\hat{V}_H \mathcal{FFT}(\rho_{extended}))$ restricted to original domain;

The use of periodic boundary conditions results in the repeated interaction of the electron with its periodic copies, which is unphysical. To avoid this spurious effect, in Algorithm 3, the range of the Coulomb interaction is limited to a cut-off radius R and the domain of ρ is extended to a larger periodic domain including this cut-off region (line 1). To obtain the Fourier transform $\hat{V}_H(K) = \int_0^R J_0(Kr)\mathrm{d}r$ of $\frac{1}{|x|}$, we first integrate the Bessel function in line 4 for a number of radii k and then save the values for all grid points in $\hat{V}_H$ in line 8. Finally, the actual multiplication of $\hat{\rho}$ with $\hat{V}_H$ and the Fourier transforms are carried out in line 10 and the restriction to the original domain

is returned. We remark that the numerical approximation of $\hat{V}_H$ depends only on the number of grid points N_x, N_y and the cut-off radius R, so it can be stored and reused for the whole optimization procedure.

5.2. Numerical optimization schemes

In this section, we introduce the numerical optimization schemes used in the numerical experiments in Section 5.3 to find the solution of the reduced optimization problem

$$\min_{u \in V} \hat{J}(u).$$

We discuss the steepest descent, the NCG and the BFGS methods, see also [NW06, BCS17]. As the gradient gives the direction of steepest ascend, the idea of the steepest descent method is straight forward as given in Algorithm 4 and can be applied to all differentiable optimization problems in a Hilbert space V.

Algorithm 4 Steepest descent scheme

Input: Initial approx. u_0, maximum k_{max}, tolerance *tol*;
Output: Approximate optimal control u_k;
  1: Set $d_0 = -\nabla_V \hat{J}(u_0)$;
  2: Set $k = 0$;
  3: **while** $k < k_{max}$ **and** $\|d_k\|_V > tol$ **do**
  4:    Evaluate steepest descent $d_k = -\nabla_V \hat{J}(u_k)$ by Algorithm 6;
  5:    Compute steplength α_k along d_k by a given rule, e.g. Algorithm 7;
  6:    Set $u_{k+1} = u_k + \alpha_k d_k$;
  7:    Set $k \leftarrow k + 1$;
  8: **end while**
  9: **return** u_k;

5.2.1. Nonlinear conjugate gradient method

The steepest descent method is based on the plausible idea to take at each step the steepest direction downwards which is orthogonal to the level set of the functional to be minimised. However, this prevents the method to follow a curved valley, and renders the method sometimes very inefficient. This problem is overcome by the so-called conjugate directions that are orthogonal in a suitable scalar product that is adapted to the problem. This idea is the basis of the conjugate gradient method and its nonlinear extension, the nonlinear conjugate gradients method (NCG). An example for this behaviour is shown in Figure 5.3, where the steepest descent, the NCG method, and the BFGS scheme presented later are used to find the minimum of the Rosenbrock function

$$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2. \tag{5.29}$$

Indeed the steepest descent needs a lot of iterations because at each step a 90° turn to the level set is preformed. The NCG and the BFGS method are able to follow the curved valley and need only a few iterations.

Therefore, we focus mainly on the NCG method that is quite fast and robust. As opposed to Newton methods, see, e.g., [CB16, vWB10], the NCG method does not need the Hessian which is difficult to implement for the nonlinear TDKS approach. Later, we will give a short account on quasi-Newton methods that are also based on the gradient
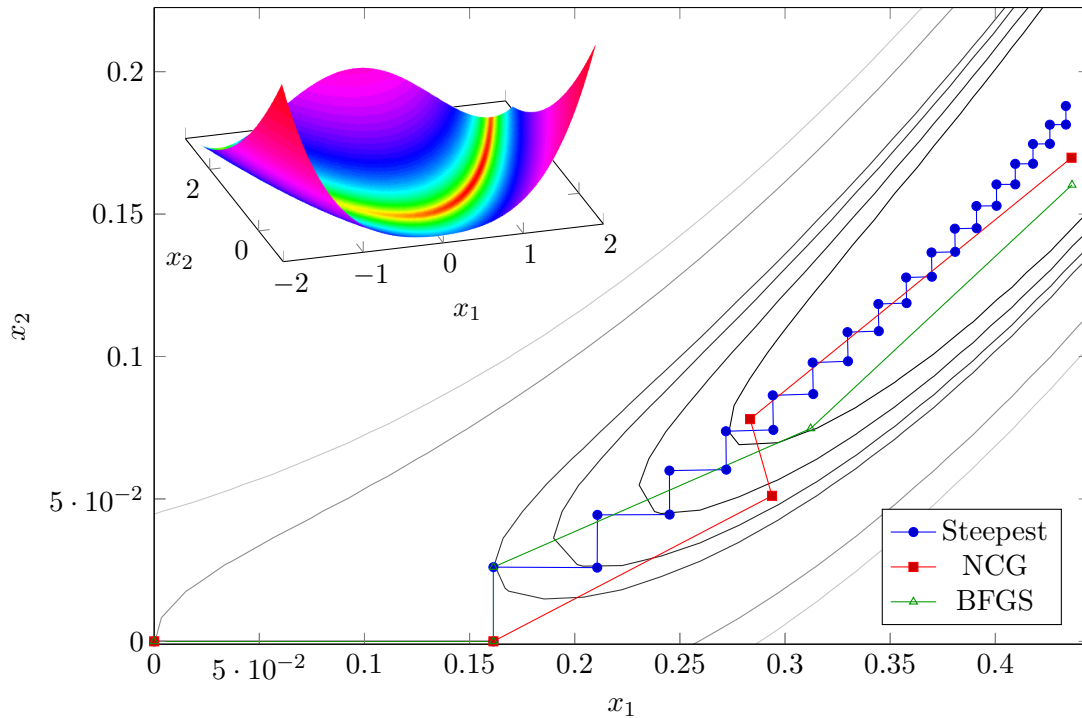
Figure 5.3.: Comparison of the classical optimization methods in the curved valley of the Rosenbrock function (5.29). The steepest descent always turns 90 degrees to the level set whereas the NCG and BFGS updates try to follow the bottom of the valley. The inset shows the Rosenbrock function (5.29).

and do not need the Hessian. To ensure convergence, the inclusion of a line search into the optimization method is necessary, see page 90 for a discussion on the line search.

The NCG schemes represent extensions of the linear conjugate gradient method to nonquadratic problems; see, e.g., [Sha78, GN92]. In the common variants, the basic idea is to avoid matrix operations and express the search directions recursively as follows:

$$d_{k+1} = -g_{k+1} + \beta_k \, d_k, \tag{5.30}$$

where $g_k = \nabla_V \hat{J}(u_k)$, $k = 0, 1, 2, \ldots$, with $d_0 = -g_0$. The iterates for a minimum point are given by

$$u_{k+1} = u_k + \alpha_k \, d_k, \tag{5.31}$$

where $\alpha_k > 0$ is a steplength obtained by a line search procedure. The parameter β_k is chosen such that (5.30)–(5.31) reduces to the linear CG scheme if $\hat{J}$ is a strictly convex quadratic function and α_k is the exact one-dimensional minimizer of $\hat{J}$ along d_k. In this case the NCG scheme terminates in at most n steps in exact arithmetic. This case provides a lower bound to the computational complexity of NCG schemes.

There are different formulas for β_k that satisfy the above condition of reduction to the CG scheme, which result in different performances depending on the (nonlinear) problem at hand. We employ the variant of Hager and Zhang [HZ05] based on the formula

$$\beta_k^{HZ} = \frac{\langle \sigma_k, \, g_{k+1} \rangle_V}{\langle d_k, \, y_k \rangle_V}, \qquad \sigma_k = y_k - 2d_k \frac{\langle y_k, \, y_k \rangle_V}{\langle y_k, \, d_k \rangle_V}, \tag{5.32}$$

where $y_k = g_{k+1} - g_k$ and V denotes the Hilbert space where the gradients are defined. This choice of β seems to be advantageous in solving quantum control problems, see, e.g.,

[BSV08, vWB08]. Notice that the definition of β_k is based on the V-space inner product. For optimization in complex Hilbert spaces, the inner products above are replaced with $\mathrm{Re}\,\langle\cdot,\,\cdot\rangle_V$; see [BSV08].

We summarize the NCG scheme in the Algorithm 5. For convergence of the NCG method we refer to [BCS17].

Algorithm 5 NCG scheme

Input: Initial approx. u_0, maximum k_{max}, tolerance *tol*;
Output: Approximate optimal control u_k;
  1: Set $k = 0$;
  2: Set $d_0 = -\nabla_V \hat{J}(u_0)$;
  3: **while** $k < k_{max}$ **and** $\|g_k\|_V > tol$ **do**
  4:    Determine step length $\alpha_k > 0$ along d_k by Algorithm 7 satisfying (5.33)–(5.34);
  5:    Set $u_{k+1} = u_k + \alpha_k\, d_k$;
  6:    Compute $g_{k+1} = \nabla_V \hat{J}(u_{k+1})$ by Algorithm 6;
  7:    Compute β_k by (5.32);
  8:    Set $d_{k+1} = -g_{k+1} + \beta_k\, d_k$;
  9:    Set $k \leftarrow k + 1$;
 10: **end while**
 11: **return** u_k;

At lines 2 and 6 of Algorithm 5, the gradient of the reduced cost functional $\nabla_V \hat{J}(u)$ is calculated by Algorithm 6. Notice that the descent direction d^k is computed at line 8 by means of the Hager-Zhang scheme [HZ05]. The step length α^k is obtained by the line search Algorithm 7.

Essential to the solution procedure is the computation of the reduced gradient corresponding to a given control u. This is the purpose of Algorithm 6, where the forward TDKS model (line 1) is solved first, followed by the solution of the backward TDKS equation (line 2). Line 3 is devoted to assembling the gradient.

Algorithm 6 Gradient of reduced cost functional

Input: External potential V_{ext} and control $u(t)$, Ψ^0, Λ^T;
Output: $\nabla_V \hat{J}(t)$;
  1: Solve the forward equation (4.8a) for the given control $u(t)$ and Ψ^0 to obtain $\Psi(x,t)$ using Algorithm 1;
  2: Solve the adjoint equation (4.8c) for the given control $u(t)$ and the solution of the forward equation $\Psi(x,t)$ to obtain $\Lambda(x,t)$ by Algorithm 2;
  3: Solve the ODE (4.9) to obtain μ and assemble the gradient according to (4.8e) with $u(t)$, $\Psi(x,t)$, and $\Lambda(x,t)$ from step 1 and 2, respectively;

To globalize the optimization method, it is necessary to use a line search procedure to determine the step length in Line 4 of Algorithm 5. Given two constants $0 < \delta < \sigma < 1$, two common criteria are given by (see also [NW06, BCS17] for further discussion)

- the Armijo condition to ensure sufficient decrease of the cost functional,

$$\hat{J}(u + \alpha d) \leq \hat{J}(u) + \delta\alpha \left\langle \nabla_V \hat{J}(u),\, d \right\rangle_V; \tag{5.33}$$

- and the Wolfe condition to ensure a sufficiently large step $\alpha_k d_k$,

$$\left\langle d_k,\, \nabla_V \hat{J}(u + \alpha_k d_k) \right\rangle_V \geq \sigma \left\langle d_k,\, g_k \right\rangle_V. \tag{5.34}$$

In Algorithm 7, we implement the bisection line search strategy presented in [vWB08] which is very robust for quantum control problems. The idea is to first increase the step size until the functional increases again after a local minimum. In the second part, the interval determined in the first step is successively refined via bisection using the finite difference approximation of the derivative of the functional defined in line 1.

Algorithm 7 Bisection line search

Input: Control $u(t)$, search direction d, cost functional $\hat{J}$;
Output: Step length α;
 1: define $\phi(\alpha) := \hat{J}(u + \alpha d)$, $d\phi(\alpha) := \frac{\phi(\alpha+\delta\alpha)-\phi(\alpha-\delta\alpha)}{2\delta\alpha}$, for fixed $\delta\alpha$;
 2: $\alpha^{old} \leftarrow 0$, choose initial α^{new};
 3: **while** $\phi(\alpha^{new}) < \phi(\alpha^{old})$ **do**
 4:     $\alpha^{old} \leftarrow \alpha^{new}$;
 5:     Increase α^{new}, e.g. $\alpha^{new} \leftarrow (\alpha^{new} + c_1)c_2$, $c_2 > 1$;
 6: **end while**
 7: $\alpha_l \leftarrow 0$, $\alpha_r \leftarrow \alpha^{new}$;
 8: **while** $|\alpha_r - \alpha_l| > tol$ **do**
 9:     **if** $d\phi\left(\frac{\alpha_l+\alpha_r}{2}\right) d\phi(\alpha_l) < 0$ **then**
10:         $\alpha_r \leftarrow \frac{\alpha_l+\alpha_r}{2}$;
11:     **else**
12:         $\alpha_l \leftarrow \frac{\alpha_l+\alpha_r}{2}$;
13:     **end if**
14: **end while**
15: **return** $\frac{\alpha_l+\alpha_r}{2}$;

5.2.2. Quasi-Newton methods

Quasi-Newton methods can be seen as extensions of the conjugate gradient method, in which additional storage is used to accelerate convergence. To achieve this, approximations of the Hessian matrix are constructed using low-rank updates based on gradient evaluations.

The BFGS method is a quasi-Newton method which makes successive rank-two updates to a matrix B such that it serves as an approximation to the true Hessian. Notice that the BFGS scheme may exhibit convergence rates superior to those of NCG schemes at the expense of additional computational effort.

Denote by B_k the k-th BFGS approximation to the Hessian. Then the BFGS search direction at the k-th step is given by $p_k = -B_k^{-1} g_k$, where $g_k = \nabla_V \hat{J}(u_k)$. Further, denote the difference between two successive updates of u as $u_{k+1} - u_k =: s_k = \alpha_k p_k$, where α_k is the steplength and denote by $y_k := g_{k+1} - g_k$ the difference of the gradients. The symmetric matrix B_k can be determined explicitly via the well-known recurrence formula, see, e.g., [NW06],

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^\top}{s_k^\top B_k s_k} + \frac{y_k y_k^\top}{y_k^\top s_k}. \tag{5.35}$$

To compute the search direction, it is necessary to invert the matrix B. We denote its inverse by $H = B^{-1}$. Using the Sherman–Morrison–Woodbury formula, we can also

establish a recurrence for H as follows:

$$H_{k+1} = H_k + \frac{s_k^\top y_k + y_k^\top H_k y_k}{(s_k^\top y_k)^2}(s_k s_k^\top) - \frac{H_k y_k s_k^\top + s_k y_k^\top H_k}{s_k^\top y_k} \tag{5.36}$$

$$= H_k + d_k c_k s_k s_k^\top - c_k z_k s_k^\top - c_k s_k z_k^\top, \tag{5.37}$$

where we introduced the vector $z_k := H_k y_k$ and the following abbreviations for the scalar values: $c_j := (s_j^\top y_j)^{-1}$ and $d_j := 1 + c_j\, y_j^\top z_j$.

In the case where the control u and the gradient of the objective function $\nabla_V \hat{J}$ are elements in a function space instead of vectors in $\mathbb{R}^n$, it is not immediately obvious how to directly use the formula (5.37) since it requires forming outer products. Moreover, to compute the search direction, we only need the action of H on a vector g and it is not necessary to construct any matrix. These facts are discussed in [vWB08], where a matrix-free BFGS is formulated; see also [MI99].

Suppose V is either $L^2(0, T; \mathbb{R})$ or $H^1(0, T; \mathbb{R})$ and that $x, y \in V$. Then, we can denote the function space analogue of the outer product as a dyadic operator $x \otimes y : V \to V$. The action of this operator on a third element $z \in V$ can be expressed in terms of the inner product $(x \otimes y)\, z = \langle y,\, z \rangle_V\, x$, see, e.g., [MF53, § 1.6]. In general function spaces, (5.37) is given by

$$H_{k+1} = H_k + d_k c_k s_k \otimes s_k - c_k z_k \otimes s_k - c_k s_k \otimes z_k. \tag{5.38}$$

Using this formula for H_k and $(x \otimes y)\, z = \langle y,\, z \rangle_V\, x$, we obtain the following result concerning the action of H_k to an arbitrary vector x. We have

$$H_k x = H_{k-1} x + d_{k-1} c_{k-1} \langle s_{k-1},\, x \rangle_V\, s_{k-1} - c_{k-1} \langle s_{k-1},\, x \rangle_V\, z_{k-1}$$
$$- c_{k-1} \langle z_{k-1},\, x \rangle_V\, s_{k-1}$$
$$= H_{k-1} x + c_{k-1} \langle s_{k-1},\, x \rangle_V\, r_{k-1} - c_{k-1} \langle z_{k-1},\, x \rangle_V\, s_{k-1},$$

where we defined $r_j := d_j$ and $s_j - z_j$. Using this formula recursively to eliminate all H_k, $k > 0$, on the right-hand side leads to the following:

$$H_k x = H_0 x + \sum_{j=0}^{k-1} c_j \left(\langle s_j,\, x \rangle_V\, r_j - \langle z_j,\, x \rangle_V\, s_j \right). \tag{5.39}$$

Now, we illustrate the computational steps to compute the BFGS solution. We start with the initial approximation u_0 and correspondingly determine $g_0 = \nabla_V \hat{J}(u_0)$. We set $p_0 = -g_0$ and minimize along p_0 with steplength α_0 given by a line search. We obtain u_1 and g_1, and therefore we can compute $y_0 = g_1 - g_0$ and $s_0 = \alpha_0 p_0$. The first step requires the initialization $H_0 = I$, and we have $z_0 = H_0 y_0$ and $p_1 = -H_0 g_1$. Correspondingly, minimizing along p_1 with line search with steplength α_1, we obtain $y_1 = g_2 - g_1$ and $s_1 = \alpha_1 p_1$.

Using (5.39) for $x = y_k$ and $x = -g_{k+1}$, we obtain the following:

$$z_k = H_k y_k = H_0 y_k + \sum_{j=0}^{k-1} c_j \left(\langle s_j,\, y_k \rangle_V\, r_j - \langle z_j,\, y_k \rangle_V\, s_j \right), \quad k \geq 1, \tag{5.40}$$

$$p_{k+1} = -H_k g_{k+1} = -H_0 g_{k+1} - \sum_{j=0}^{k} c_j \left(\langle s_j,\, g_{k+1} \rangle_V\, r_j - \langle z_j,\, g_{k+1} \rangle_V\, s_j \right). \tag{5.41}$$

Notice that with both the NCG and BFGS schemes, the new control u_{k+1} is composed of a linear combination of the original control and the gradients at every step. Further,

in the BFGS approach the current approximation to the inverse of the Hessian is stored, whereas in the matrix-free BFGS method the vectors $\{s_j, y_j, z_j\}$ are stored. These are the s_j vectors which are the search steps themselves, the y_j which are the differences between successive gradients, and the z_j vectors which are elements in the space spanned by $\{s_0, \ldots, s_{j-1}\}$. As a counterpart, the matrix-free BFGS formula requires progressively more computation for each optimization step, so it is important that the improved convergence properties at least compensate for the increased computational effort. The matrix-free BFGS scheme is implemented in Algorithm 8 below.

Algorithm 8 BFGS scheme

Input: Choose $H_0 = I$, initial approx. u_0, maximum k_{max}, tolerance tol;
Output: Approximate optimal control u_k;
  1: Set $g_0 = \nabla_V \hat{J}(u_0)$, $p_0 = -g_0$;
  2: Set $k \leftarrow 0$;
  3: **while** $k < k_{max}$ **and** $\|g_{k-1}\|_V > tol$ **do**
  4:    Compute $u_{k+1} = u_k + \alpha_k\, p_k$ with α_k satisfying (5.33)–(5.34);
  5:    Compute $g_{k+1} = \nabla_V \hat{J}(u_{k+1})$, $y_k = g_{k+1} - g_k$, $s_k = \alpha_k\, p_k$;
  6:    Compute z_k with (5.40);
  7:    Compute and save $c_k = \langle s_k, y_k \rangle_V^{-1}$, $d_k = 1 + c_k \langle y_k, z_k \rangle_V$, and $r_k = d_k\, s_k - z_k$;
  8:    Compute new search direction p_{k+1} with (5.41).
  9:    Set $k \leftarrow k + 1$;
10: **end while**
11: **return** u_k;

Notice that the superlinear convergence properties of quasi-Newton methods in Hilbert spaces are obtained only if the initial Hessian approximation is chosen as a compact perturbation of the Hessian at the optimal solution; see [Kup96, Sac86] for further discussion.

5.3. Numerical experiments

In this section, we present results of four different numerical experiments demonstrating the ability of our method to solve TDKS optimal control problems with different targets,

$$\min_{u \in H_0^1(0,T;\mathbb{R})} \hat{J}(u),$$

where $\hat{J}$ consists of different terms specified below. As the first example (defined in Section 5.3.1) is relatively fast to solve, we use this example to compare the NCG method with the BFGS method, to compare H^1 and L^2 optimization in Section 5.3.5 and to study the dependence of the runtime on the particle number in Section 5.3.6. The code for the experiments in this section is published in [SCB17a] and can be downloaded from `http://dx.doi.org/10.17632/p5g5hznkpy.1`.

5.3.1. Density tracking

In our first experiment, our objective is that the density of 2 electrons follows a prescribed trajectory ($\beta = 1$, $\eta = 0$),

$$\hat{J} = J_\beta + \frac{\nu}{2}\|u\|_{H^1(0,T)}^2, \qquad J_\beta = \frac{\beta}{2}\int_0^T \int_\Omega (\rho(x,t) - \rho_d(x))^2 \mathrm{d}x\mathrm{d}t. \qquad (5.42)$$
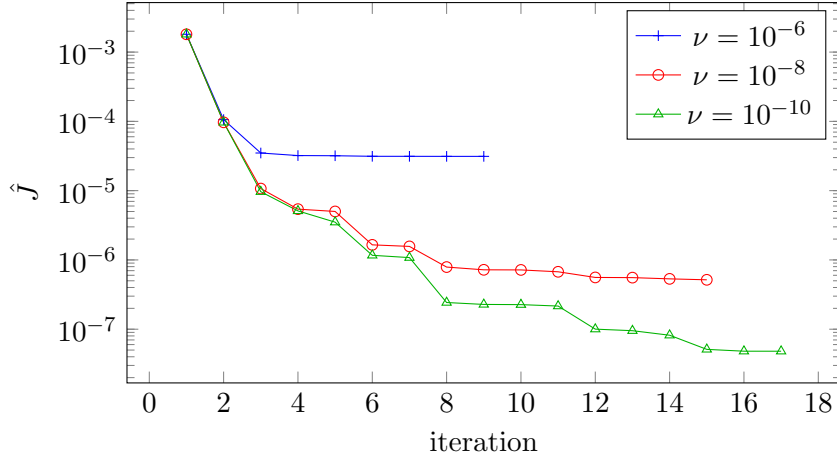
Figure 5.4.: The cost functional $\hat{J}$ as a function of the iteration number for different values of the regularization ν.

To this end, we chose a control function $u = u^{test}$, as an input to the TDKS model with $V_{ext}(x,t,u) = 50(x_1^2 + x_2^2) + u^{test}(t)(x_1^2 + x_2^2)$, and compute the corresponding solution to produce our target trajectory ρ_d, which determines the objective J_β. Starting with the initial guess for the control $u^0 = 0$, our purpose is then to track the density resulting from the prescribed forcing term u^{test} (dashed line in Figure 5.5). The stopping criterion for convergence is $\|\nabla \hat{J}\|_{H^1(0,T;\mathbb{R})} < 2 \cdot 10^{-7}$.

The results of this experiment are presented in Figures 5.4 and 5.5. In Figure 5.4, the monotonic decrease of the cost functional $\hat{J}$ can be seen. Figure 5.5 shows the initial guess u^0 for the optimization solver, the function u^{test}, which is used to generate the target density ρ_d, and the computed optimal control u^{opt} for different values of the Tikhonov regularization ν. As expected, the optimal control solution approaches u^{test} for decreasing ν; however an increasing computational effort is required for smaller ν as can be seen from the fact that the stopping criterion is only met after more iterations (Figure 5.4).

The detailed values for the different terms J_β and J_ν forming $\hat{J}$ and the number of iterations and the required CPU time are summarized in Table 5.1. It can be seen that decreasing the regularization ν is necessary to achieve small target values J_β because otherwise J_ν dominates the cost. Runtime is increasing slightly for smaller ν but only by a factor of 2 while decreasing ν from 10^{-6} to 10^{-10}.

ν	$\hat{J}$	J_β	J_ν	iteration	CPU
10^{-6}	$3.124639 \cdot 10^{-5}$	$6.7896 \cdot 10^{-6}$	$2.4457 \cdot 10^{-5}$	8	133
10^{-7}	$4.432995 \cdot 10^{-6}$	$4.9835 \cdot 10^{-7}$	$3.9346 \cdot 10^{-6}$	21	348
10^{-8}	$5.164362 \cdot 10^{-7}$	$5.8570 \cdot 10^{-8}$	$4.5787 \cdot 10^{-7}$	14	238
10^{-9}	$8.651593 \cdot 10^{-8}$	$3.9841 \cdot 10^{-8}$	$4.6675 \cdot 10^{-8}$	15	246
10^{-10}	$4.810496 \cdot 10^{-8}$	$4.3453 \cdot 10^{-8}$	$4.6524 \cdot 10^{-9}$	16	267

Table 5.1.: Convergence for the NCG method with fixed stopping criterion $\|\nabla \hat{J}\|_{H^1(0,T)} < 2 \cdot 10^{-7}$ for different values of ν. "CPU" is wall runtime in seconds, see also Section 5.3.6.
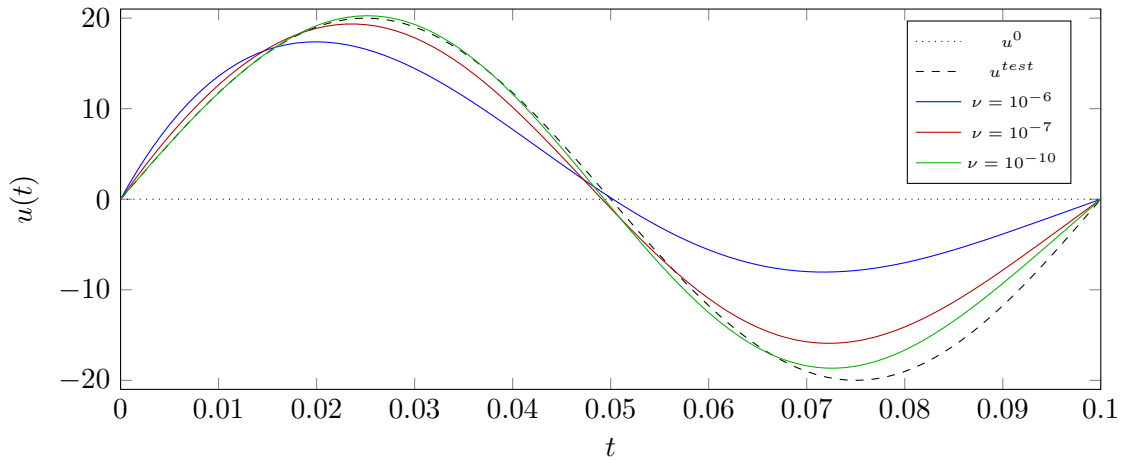
Figure 5.5.: Controls obtained for different Tikhonov regularization parameters ν.

Iteration	J_β	J_ν	$\|\nabla \hat{J}\|_{H^1(0,T)}$	step length
1	$1.8076 \cdot 10^{-3}$	0	$7.6244 \cdot 10^{-4}$	$5.6450 \cdot 10^3$
2	$9.6777 \cdot 10^{-5}$	$9.2621 \cdot 10^{-10}$	$2.8632 \cdot 10^{-5}$	$2.1488 \cdot 10^5$
3	$9.6494 \cdot 10^{-6}$	$2.8289 \cdot 10^{-9}$	$4.1384 \cdot 10^{-5}$	$5.3055 \cdot 10^3$
4	$5.0980 \cdot 10^{-6}$	$3.1314 \cdot 10^{-9}$	$1.8770 \cdot 10^{-6}$	$9.0253 \cdot 10^5$
5	$3.5048 \cdot 10^{-6}$	$3.2127 \cdot 10^{-9}$	$1.9393 \cdot 10^{-5}$	$1.2541 \cdot 10^4$
6	$1.1619 \cdot 10^{-6}$	$3.8234 \cdot 10^{-9}$	$3.4086 \cdot 10^{-6}$	$1.4793 \cdot 10^4$
7	$1.0769 \cdot 10^{-6}$	$3.8382 \cdot 10^{-9}$	$4.5825 \cdot 10^{-6}$	$7.8308 \cdot 10^4$
8	$2.3851 \cdot 10^{-7}$	$4.2691 \cdot 10^{-9}$	$2.2667 \cdot 10^{-6}$	$5.4753 \cdot 10^3$
9	$2.2356 \cdot 10^{-7}$	$4.2750 \cdot 10^{-9}$	$4.2153 \cdot 10^{-7}$	$2.0166 \cdot 10^4$
10	$2.2204 \cdot 10^{-7}$	$4.2757 \cdot 10^{-9}$	$6.8748 \cdot 10^{-7}$	$3.7822 \cdot 10^4$
11	$2.1161 \cdot 10^{-7}$	$4.2960 \cdot 10^{-9}$	$2.4825 \cdot 10^{-6}$	$3.5691 \cdot 10^4$
12	$9.5702 \cdot 10^{-8}$	$4.5303 \cdot 10^{-9}$	$8.8414 \cdot 10^{-7}$	$1.1481 \cdot 10^4$
13	$9.0635 \cdot 10^{-8}$	$4.5388 \cdot 10^{-9}$	$9.9239 \cdot 10^{-7}$	$3.0364 \cdot 10^4$
14	$7.7318 \cdot 10^{-8}$	$4.5580 \cdot 10^{-9}$	$1.3960 \cdot 10^{-6}$	$2.8538 \cdot 10^4$
15	$4.6520 \cdot 10^{-8}$	$4.6500 \cdot 10^{-9}$	$9.2407 \cdot 10^{-7}$	$6.0270 \cdot 10^3$
16	$4.3489 \cdot 10^{-8}$	$4.6526 \cdot 10^{-9}$	$2.2688 \cdot 10^{-7}$	$3.3448 \cdot 10^3$
17	$4.3453 \cdot 10^{-8}$	$4.6524 \cdot 10^{-9}$	$1.5028 \cdot 10^{-7}$	

Table 5.2.: Convergence of the NCG method for $\nu = 10^{-10}$.

In Table 5.4, we demonstrate that the BFGS quasi-Newton method can also be used to solve TDKS optimization problems. The BFGS method is slightly more expensive than the NCG method but provides a better guess for the step length as it collects some information about the curvature of the cost functional. This can be seen by the fact that the step length for the BFGS method (Table 5.3) is closer to 1 than for the NCG method (Table 5.2). As determining the step length is the computational most expensive part of the optimization, this is a serious advantage. However, the total computation times in terms of seconds of CPU runtime are comparable.

Iteration	J_β	J_ν	$\|\nabla \hat{J}\|_{H^1(0,T)}$	step length
1	0.0018076	0	$7.6243 \cdot 10^{-4}$	5645.0
2	$1.6052 \cdot 10^{-5}$	$2.5242 \cdot 10^{-9}$	$2.8632 \cdot 10^{-5}$	199981
3	$5.1226 \cdot 10^{-6}$	$3.1481 \cdot 10^{-9}$	$6.2073 \cdot 10^{-5}$	1.0075
4	$5.1132 \cdot 10^{-6}$	$3.1071 \cdot 10^{-9}$	$1.9644 \cdot 10^{-6}$	0.98288
5	$5.1132 \cdot 10^{-6}$	$3.1070 \cdot 10^{-9}$	$1.8585 \cdot 10^{-6}$	0.012353
6	$5.1132 \cdot 10^{-6}$	$3.1069 \cdot 10^{-9}$	$1.8585 \cdot 10^{-6}$	0.078476
7	$5.1132 \cdot 10^{-6}$	$3.1067 \cdot 10^{-9}$	$1.8584 \cdot 10^{-6}$	0.085039
8	$5.1132 \cdot 10^{-6}$	$3.1065 \cdot 10^{-9}$	$1.8583 \cdot 10^{-6}$	0.091601
9	$5.1132 \cdot 10^{-6}$	$3.1063 \cdot 10^{-9}$	$1.8582 \cdot 10^{-6}$	0.10035
10	$5.1132 \cdot 10^{-6}$	$3.1061 \cdot 10^{-9}$	$1.8581 \cdot 10^{-6}$	0.11101
11	$5.1132 \cdot 10^{-6}$	$3.1058 \cdot 10^{-9}$	$1.8579 \cdot 10^{-6}$	0.12195
12	$5.1132 \cdot 10^{-6}$	$3.1055 \cdot 10^{-9}$	$1.8578 \cdot 10^{-6}$	0.13835
13	$5.1132 \cdot 10^{-6}$	$3.1052 \cdot 10^{-9}$	$1.8576 \cdot 10^{-6}$	0.15695
14	$5.1132 \cdot 10^{-6}$	$3.1048 \cdot 10^{-9}$	$1.8575 \cdot 10^{-6}$	0.18101
15	$5.1132 \cdot 10^{-6}$	$3.1044 \cdot 10^{-9}$	$1.8573 \cdot 10^{-6}$	0.21492
16	$5.1132 \cdot 10^{-6}$	$3.1038 \cdot 10^{-9}$	$1.8571 \cdot 10^{-6}$	0.25921
17	$5.1132 \cdot 10^{-6}$	$3.1031 \cdot 10^{-9}$	$1.8569 \cdot 10^{-6}$	0.32921
18	$5.1132 \cdot 10^{-6}$	$3.1022 \cdot 10^{-9}$	$1.8566 \cdot 10^{-6}$	0.43621
19	$5.1132 \cdot 10^{-6}$	$3.1009 \cdot 10^{-9}$	$1.8563 \cdot 10^{-6}$	0.62786
20	$5.1132 \cdot 10^{-6}$	$3.0988 \cdot 10^{-9}$	$1.8560 \cdot 10^{-6}$	1.0240
21	$5.1131 \cdot 10^{-6}$	$3.0945 \cdot 10^{-9}$	$1.8557 \cdot 10^{-6}$	2.0720
22	$5.1122 \cdot 10^{-6}$	$3.0815 \cdot 10^{-9}$	$1.8562 \cdot 10^{-6}$	6.3308
23	$5.0624 \cdot 10^{-6}$	$2.9915 \cdot 10^{-9}$	$1.8659 \cdot 10^{-6}$	45.229
24	$2.3258 \cdot 10^{-7}$	$4.5983 \cdot 10^{-9}$	$2.2561 \cdot 10^{-6}$	95.796
25	$2.6512 \cdot 10^{-8}$	$4.8440 \cdot 10^{-9}$	$2.4853 \cdot 10^{-6}$	1.0569
26	$2.6512 \cdot 10^{-8}$	$4.8440 \cdot 10^{-9}$	$6.7373 \cdot 10^{-8}$	

Table 5.3.: Convergence for the BFGS method $\nu = 10^{-10}$.

ν	$\hat{J}$	J_β	J_ν	iteration	CPU
10^{-6}	$3.124073 \cdot 10^{-5}$	$6.6715 \cdot 10^{-6}$	$2.4569 \cdot 10^{-5}$	12	143
10^{-7}	$4.421684 \cdot 10^{-6}$	$3.9128 \cdot 10^{-7}$	$4.0304 \cdot 10^{-6}$	22	240
10^{-8}	$5.423582 \cdot 10^{-6}$	$5.1141 \cdot 10^{-6}$	$3.0953 \cdot 10^{-7}$	4	60
10^{-9}	$7.494184 \cdot 10^{-8}$	$2.6607 \cdot 10^{-8}$	$4.8335 \cdot 10^{-8}$	25	267
10^{-10}	$3.135697 \cdot 10^{-8}$	$2.6513 \cdot 10^{-8}$	$4.8440 \cdot 10^{-9}$	25	268

Table 5.4.: Convergence for the BFGS method with fixed stopping criterion $\|\nabla \hat{J}\|_{H^1(0,T)} < 2 \cdot 10^{-7}$ for different values of ν. "CPU" is runtime in seconds. For $\nu = 10^{-8}$ the algorithm stopped early due to an unsatisfactory search direction.

5.3.2. Asymmetric double-well

The second experiment is inspired by a more physical aim as in [CWG12]. In this case, we consider two electrons in the following asymmetric double well potential:

$$V_0(x) = \frac{1}{64}x_1^4 - \frac{1}{4}x_1^2 + \frac{1}{32}x_1^3 + \frac{1}{2}x_2^2, \tag{5.43}$$

as depicted in Figure 5.6. The initial condition is a Gaussian wave packet around the
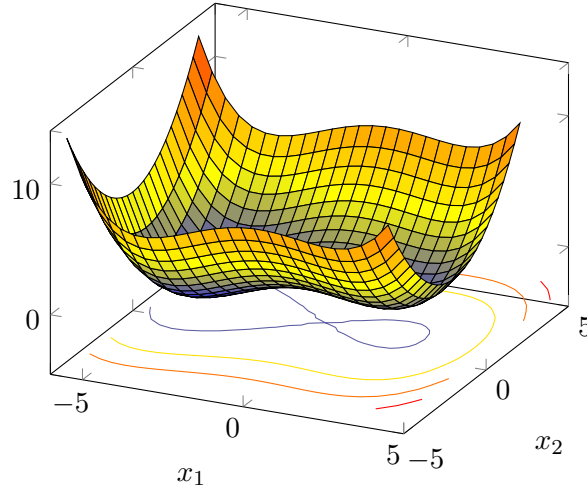
Figure 5.6.: The potential V_0 given in (5.43).

global minimum at $(x_1, x_2) = (-3.6, 0)$ which is close to the ground state of the system, $\psi^0(x) = 0.509803e^{-(x_1+4/3)^2/3}e^{-x_2^2/2}$, with $\|\psi^0\|_{L^2} = 1$. The aim is to drive the electrons to the right half-space $x_1 > 0$. This is modelled by the target

$$\hat{J} = J_\eta + \frac{\nu}{2}\|u\|_{H^1(0,T;\mathbb{R})}^2,$$

$$J_\eta = \frac{\eta}{2}\int_\Omega \chi_A(x)\rho(x,T)\mathrm{d}x, \ \text{with } A = \{(x_1, x_2) \in \mathbb{R}^2 | x_1 < 0\}.$$

A cut along the line $y = 0$ of the initial density, the target region, and the potential V_0 (scaled by a factor of 10) can be seen in Figure 5.7. The control mechanism considered

here is a laser pulse in dipole approximation in x_1-direction, i.e. $V_u(x) = x_1$. The stopping criterion for convergence is $\|\nabla \hat{J}\|_{H^1(0,T;\mathbb{R})} < 5 \cdot 10^{-5}$.

The results of the second experiment are presented in Figure 5.8. The monotonically decreasing cost functional $\hat{J}$ is shown in Figure 5.8a. The initial and the optimal control can be seen in Figure 5.8b. In Figure 5.8c, the initial density and the density $\rho^{opt}(T)$ at the final time T obtained with the optimal control u^{opt} are presented. A cut through Figure 5.8c at $x_2 = 0$ in Figure 5.8d shows the details of the densities together with the potential V_0.
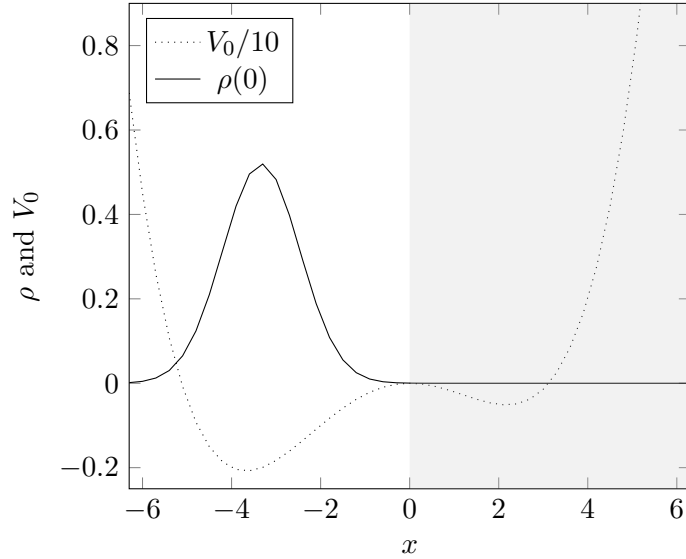
Figure 5.7.: The initial density $\rho(0)$, the potential V_0 and the target region $x_1 > 0$ (shaded in grey) along the line $x_2 = 0$.

As the results presented in Figure 5.8a show, the cost functional can be reduced by approximately 4 orders of magnitude and the density is almost completely localized in the desired set, i.e. $\int_{x_1<0} \rho(x,T)\mathrm{d}x = 2.6 \cdot 10^{-4}$. This result is obtained with $\nu = 10^{-7}$. For smaller values of ν, the stopping criterion is matched later and the objective can be further improved.

(a) The cost functional $\hat{J}$ as a function of the iteration number.

(b) The initial control u^0 and the optimal control u^{opt}.

(c) The initial density $\rho(0)$ and $\rho^{opt}(T)$ obtained with u^{opt}.

(d) The initial density $\rho(0)$ and $\rho^{opt}(T)$ obtained with u^{opt} along $x_2 = 0$.
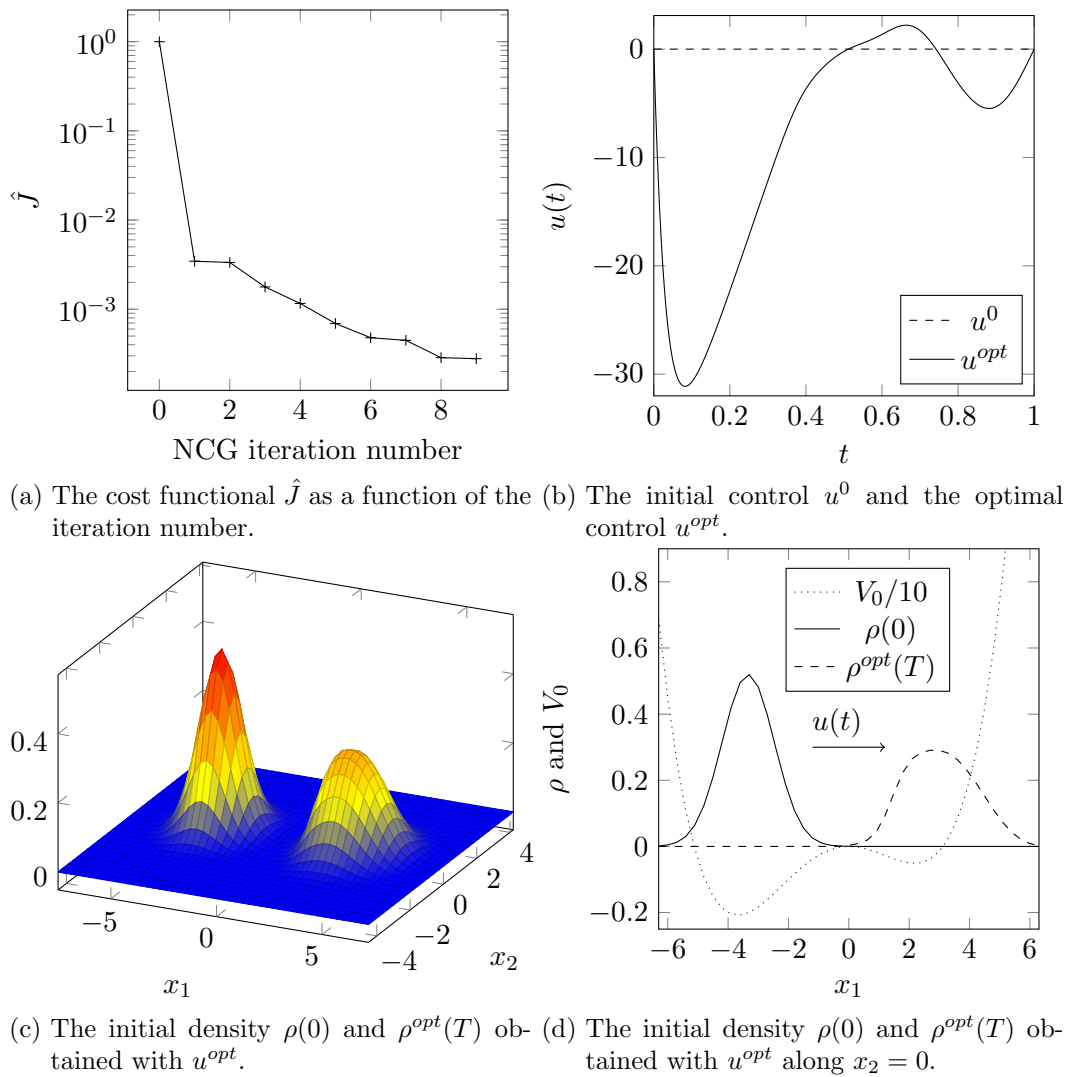
Figure 5.8.: Results obtained for the asymmetric double-well experiment.

5.3.3. Final time density

A common target in the literature, see, e.g., [KCM11], is to prescribe a terminal density ρ_T and minimize

$$\hat{J} = J_\gamma + \frac{\nu}{2}\|u\|^2_{H^1(0,T)}, \qquad\qquad J_\gamma = \frac{\gamma}{2}\int_\Omega (\rho(x,t) - \rho_T(x))^2 \mathrm{d}x. \qquad (5.44)$$

Even though the proof of the existence of a minimizer from Section 4.3 cannot applied here, we present a numerical experiment demonstrating the abilities of our code. We remark that sometimes it can be difficult to prescribe the exact desired shape. If that is the case, we recommend the usage of a target J_η as in the previous section 5.3.2, where only the region where the density is desired to be at the final time is prescribed but not its shape. However, in this and the next example it is essential to prescribe the exact shape of $\rho(T)$.

We consider the following problem. We have two electrons in coherent states of an harmonic oscillator such that without control they oscillate along the $x = y$ line. A summary on coherent states can be found in the Appendix A.3.1.

At $t = 0$ they are at the turning point $(x_1^{(0)}, x_2^{(0)})$. The target density has the same shape as the initial one, but is located at the mirrored point $(x_1^{(0)}, -x_2^{(0)})$. This means that the oscillation plane is rotated by $90°$ to the $x_1 = -x_2$ line.

For the time discretization, we use the final time $T = 0.31415$ and 2000 time steps. In space, we use a 41×41 grid with $h = 0.175$. The regularization parameter is $\nu = 10^{-8}$ and the stopping criterion is given by $\|\nabla \hat{J}\|_{H^1(0,T)} < 10^{-5}$. The initial guess is a sinusoidal control $u_0 = 100\sin(10t)$. The external potential is given by

$$V_{ext} = 50(x_1^2 + x_2^2) + u(t)x_1. \qquad (5.45)$$

The algorithm is successful in reducing the cost by three orders of magnitude as can be seen in Figure 5.9. The resulting optimal control is depicted in Figure 5.10 which has acquired several features compared to the simple initial half-sinus.
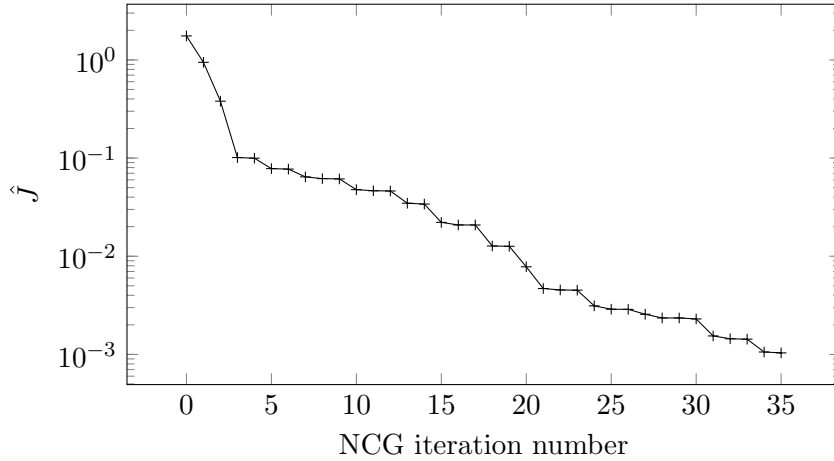
Figure 5.9.: The cost functional $\hat{J}$ as a function of the iteration number for the rotation of the oscillation plain.

In Figure 5.11 the initial density ρ_0 is depicted together with the density $\rho_{\text{free}}(T)$ that results from the evolution without control and the achieved optimal control $\rho(T)$. Figure 5.12 shows the path of the centre of the density

$$\frac{1}{N}\begin{pmatrix} \int_\Omega \rho(x,t)x_1 \, \mathrm{d}x \\ \int_\Omega \rho(x,t)x_2 \, \mathrm{d}x \end{pmatrix}.$$
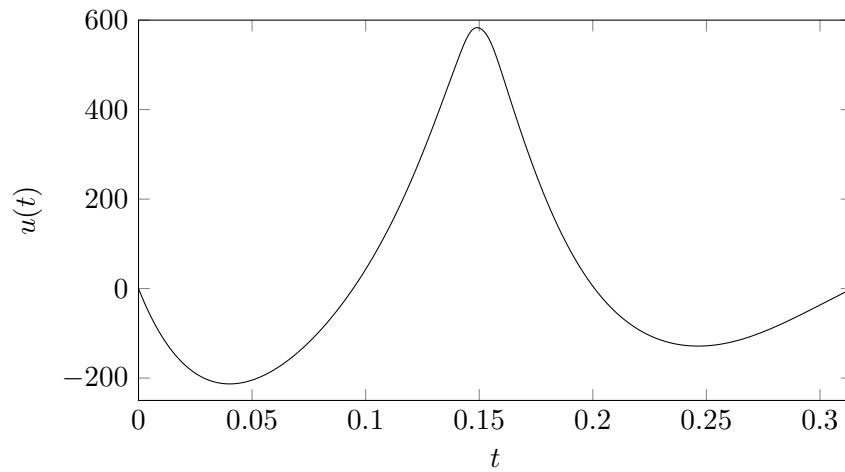
Figure 5.10.: The optimized control that rotates the oscillation plane of the electrons.

The blue path is the uncontrolled system whereas the red path shows the path with the optimal control applied. The dots are plotted at equal time; this means the closer the dots are the slower the density is moving. When the density is moving faster, a larger length is covered between two points.
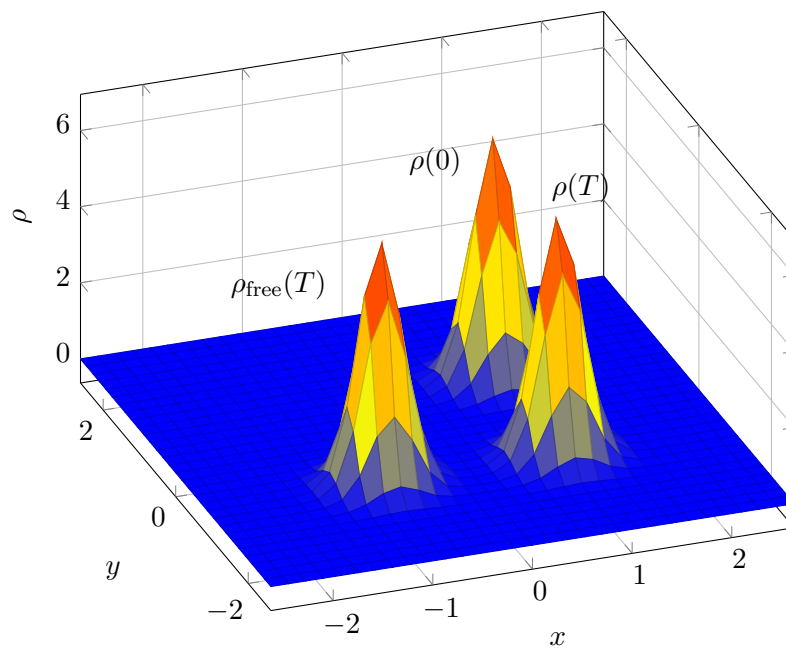

Figure 5.11.: The initial density $\rho(0)$ and the optimized final density $\rho(T)$ are shown together with the density $\rho_{\text{free}}(T)$ that the electrons attain if no control is applied.
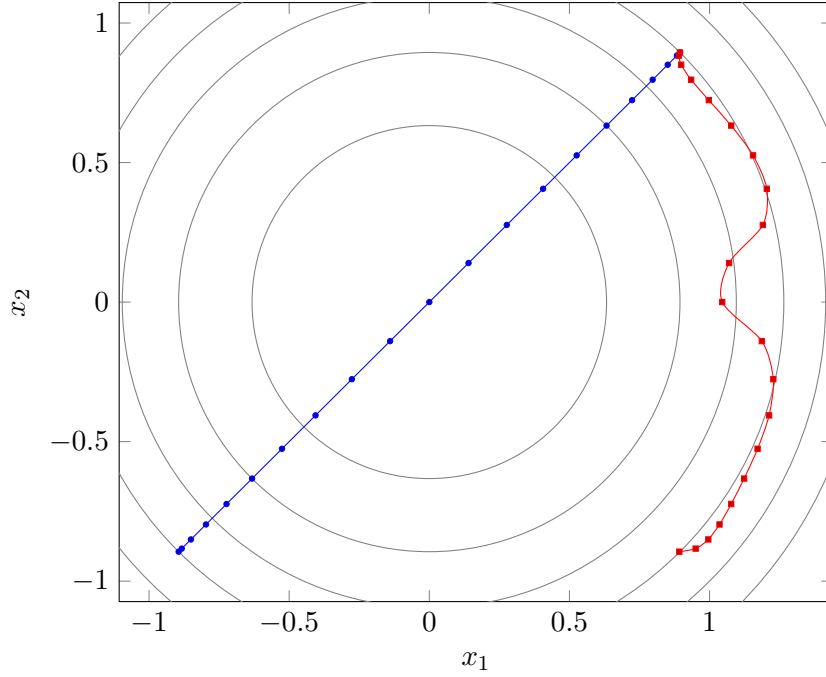
Figure 5.12.: The path of the centre of the density for the uncontrolled (blue) and the optimally controlled (red) system. Dots are at equal time steps and indicated the speed. The level sets of the confining potential (grey) are also equally spaced.

5.3.4. Helium excitation

In this section, we present a numerical experiment directly motivated from physics. We consider a neutral Helium atom in two dimensions with a so-called soft-Coulomb potential

$$V_0(x) = \frac{-2}{\sqrt{x_1^2 + x_2^2 + 1}},$$

see, e.g., [KCM11]. The density starts in the $1s$ ground state. The goal is to reach the $2p_x$ excited state at the final time T. The initial wavefunction Ψ^0 and the target density ρ_T are obtained with DFT calculations in OCTOPUS ([CAO$^+$06]). The initial and the target density are presented in Figure 5.13. The control mechanism used is a x-polarized laser pulse in dipole approximation. Therefore, the external potential is given by

$$V_{ext}(x,t) = \frac{-2}{\sqrt{x_1^2 + x_2^2 + 1}} + u(t)x_1. \tag{5.46}$$

For the numerical optimization, we choose the following parameters: $\gamma = 1$, $\nu = 10^{-5}$ in (5.44), and $tol = 2 \cdot 10^{-4}$. The initial guess for the control is a sinusoidal laser with a frequency given by the energy difference of the two states of $\Delta E = 0.2095$ (obtained via DFT calculations). The terminal time is then chosen to be six periods of the laser, i.e. $T = 7.898$.

Table 5.5 shows the value of the cost functional and the norm of the gradient. The NCG method converges in 22 iterations to a norm of the gradient of nearly 10^{-4}. In this procedure, the cost functional is reduced by nearly a factor of 40. The initial guess for the control and the resulting optimal control are depicted in Figure 5.14.
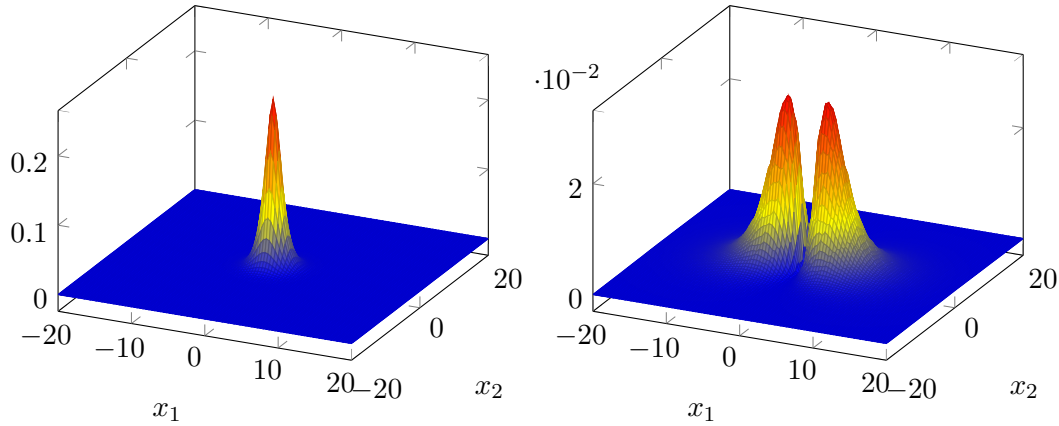
Figure 5.13.: The initial density of the $1s$ state (left) and the target density of the $2p_x$ state (right). The $2p_x$ state is much more extended and no longer spherical symmetric. Instead it is more widespread in the x_1-direction and has two separated peaks and a minimum along the x_1-axis.

Iteration	J_γ	J_ν	$\|\nabla \hat{J}\|_{H^1(0,T)}$
1	$8.1203 \cdot 10^{-2}$	$2.0195 \cdot 10^{-5}$	$1.1031 \cdot 10^{-1}$
2	$1.6869 \cdot 10^{-2}$	$2.6017 \cdot 10^{-5}$	$2.4866 \cdot 10^{-3}$
3	$8.6627 \cdot 10^{-3}$	$2.3876 \cdot 10^{-4}$	$1.0228 \cdot 10^{-2}$
4	$8.4563 \cdot 10^{-3}$	$2.4868 \cdot 10^{-4}$	$1.1157 \cdot 10^{-2}$
5	$5.5457 \cdot 10^{-3}$	$2.4197 \cdot 10^{-4}$	$1.1522 \cdot 10^{-3}$
6	$5.4486 \cdot 10^{-3}$	$2.4309 \cdot 10^{-4}$	$4.1367 \cdot 10^{-3}$
7	$4.7181 \cdot 10^{-3}$	$2.5769 \cdot 10^{-4}$	$6.7271 \cdot 10^{-3}$
8	$4.4747 \cdot 10^{-3}$	$2.5819 \cdot 10^{-4}$	$1.0153 \cdot 10^{-3}$
9	$2.5600 \cdot 10^{-3}$	$3.5726 \cdot 10^{-4}$	$3.3595 \cdot 10^{-3}$
10	$2.5206 \cdot 10^{-3}$	$3.6169 \cdot 10^{-4}$	$6.3656 \cdot 10^{-4}$
$\vdots$			
20	$2.3065 \cdot 10^{-3}$	$4.0875 \cdot 10^{-4}$	$2.5977 \cdot 10^{-4}$
21	$2.2775 \cdot 10^{-3}$	$4.1855 \cdot 10^{-4}$	$1.0667 \cdot 10^{-3}$
22	$2.2686 \cdot 10^{-3}$	$4.2155 \cdot 10^{-4}$	$1.1010 \cdot 10^{-4}$

Table 5.5.: Convergence of the NCG method.

Figure 5.15 shows the resulting optimized density $\rho(T)$ in the $x_1 - x_2$-plane. For additional understanding and a better comparison to the target density ρ_T of Figure 5.13, we provide comparisons along several lines.

First, in Figure 5.16a, we show $\rho((x_1, 0), T)$ and $\rho_T(x_1, 0)$ along the line $x_2 = 0$ as function of x_1. It can be seen that $\rho(T)$ exhibits the two peaks at the same position as ρ_T and a valley in between, even tough the valley does not fall off to zero. The decay for large values of x_1 is similar which means the states have similar extend.

In Figure 5.16b, we look at the densities as functions of x_2 for lines of fixed x_1. We choose the symmetric values of $x_1 = \pm 3.6$ which is close to the maximum of the peaks. As ρ_T is symmetric with respect to the line $x_1 = 0$, $\rho_T(-3.6, x_2) = \rho_T(+3.6, x_2)$ there is only one line for ρ_T. The optimized density $\rho(T)$ is not completely symmetric, as can also be seen in Figure 5.16a. However, for both $x_1 = -3.6$ and $x = +3.6$ both densities
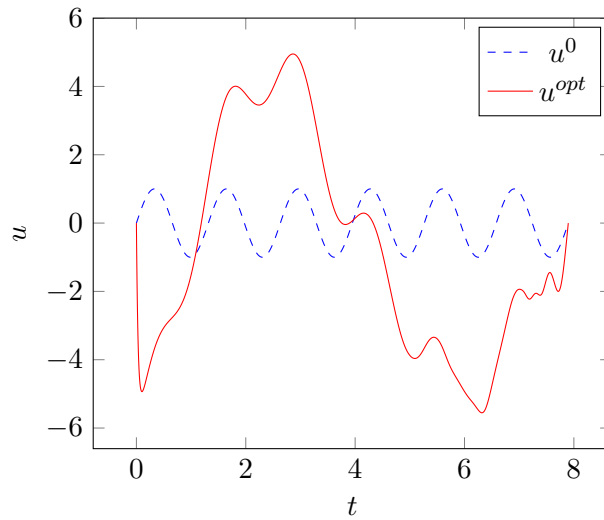
Figure 5.14.: The optimized control for the excitation $1s \to 2p_x$ (solid line).

have a similar shape, height and decay rate for large values of y.

To summarize, with the optimal control u^{opt} it is possible to bring the density from the $1s$ state to a state resembling the $2p_x$ very closely in shape and extend.
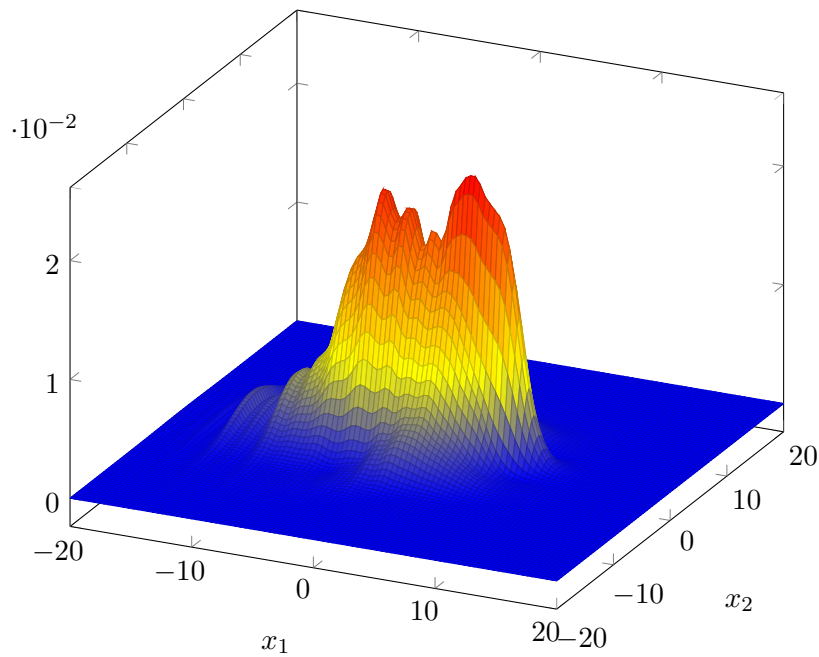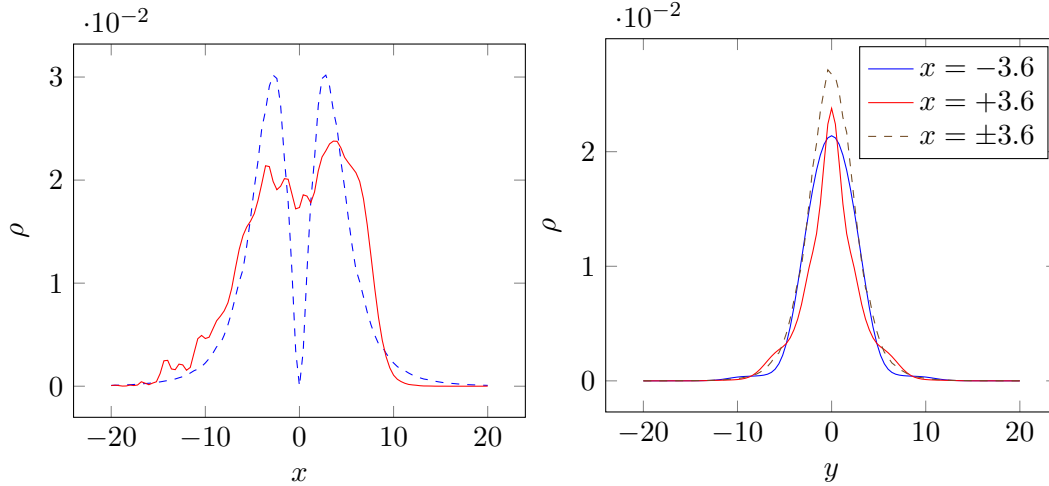

Figure 5.15.: The optimized density at the final time $\rho(T)$.

(a) This picture shows a cut along $x_2 = 0$ in Figure 5.15 together with the target density (dashed).

(b) Cuts along $x_1 = -3.6$ and $x_1 = +3.6$ in Figure 5.15 are shown together with $\rho_T(x_1 = \pm 3.6)$ (dashed).

Figure 5.16.: The optimized density at the final time $\rho(T)$ (solid line) and the target density ρ_T (dashed line) along the several lines.

5.3.5. Comparison of L^2 and H^1 optimization

In this section, we want to comment on our choice of an H^1 optimization. We recall that till today, all previous TDKS optimal control formulations consider $L^2(0,T)$ control spaces; see, e.g., [Cas13, CWG12, Wer06]. This choice is certainly reasonable but problematic for three reasons:

1. At the theoretical level, it appears difficult to prove existence of optimal solutions to the corresponding optimality system;

2. furthermore, it has not been shown that J_β is well-defined;

3. optimal controls in L^2 spaces may be very irregular and thus be difficult to implement in a laboratory.

On the other hand, we have that the forward and the adjoint equation (3.11) with $u \in H^1(0,T)$ can be proved to be uniquely solvable for any $u \in H^1(0,T)$; see Chapter 3. We also mention that with $u \in H^1(0,T)$ and the initial conditions $\psi_j^0 \in H_2(\Omega)$ the solution of the TDKS model results in a wavefunction $\psi_j \in L^4(0,T;L^4(\Omega))$ such that J_β is well-defined.

For these reasons, we strongly suggest to use an H^1 norm in J_ν to penalize the cost functional in $H^1(0,T)$. Now, given such an H^1 problem, using an L^2 gradient is problematic for the following reason. Finding the $L^2(0,T)$ gradient representation means to find $\nabla_u L \in L^2(0,T)$ such that $D_u L(\delta u) = \langle \nabla_u L, \, \delta u \rangle_{L^2(0,T)}$ for all $\delta u \in L^2(0,T)$. This is indeed possible if we assume that $u \in H^2(0,T)$ and require that $u(t) = 0$ (or $\dot{u}(t) = 0$) at $t = 0$ and $t = T$.

$$D_u L(u)(\delta u) = \nu \left\langle u, \, \delta u \right\rangle_{H^1(0,T)} - \left\langle \mathrm{Re} \left\langle \lambda, \, V_u \Psi \right\rangle_{L^2}, \, \delta u \right\rangle_{L^2(0,T)}$$

Then integration by parts gives

$$\nabla_u L = \nu(u - \ddot{u}) - \mathrm{Re} \left\langle \Lambda, \, V_u \Psi \right\rangle_{L^2}. \tag{5.47}$$

We remark that in general $u \in H^2(0,T)$ is an additional assumption. Next, notice that if a gradient-type optimization scheme is used with $\nabla_{L^2}\hat{J}(u) = \nabla_u L$, then we have the problem that the update $u^k = u^{k-1} - \alpha \nabla_{L^2}\hat{J}(u^{k-1})$ is inconsistent since the H^1 control obtained at the $(k-1)$th step is being updated with a L^2 function and the sequence $(u^k)_k$ cannot be guaranteed to be in H^1.

The most reasonable way to circumvent this problem is to work in the $H^1(0,T)$ space using the corresponding scalar product. Thus, we obtain the H^1 reduced gradient and the first-order optimality condition as given in (4.8e) where μ can be computed by solving the ODE problem (4.9). Notice that in this approach, while we do not require $u \in H^2(0,T)$, we obtain $u \in H_0^1(0,T)$. For more details concerning the derivation of the H^1 optimality system, see Chapter 4.

In the remaining part of this section, we want to illustrate the difference in the numerical behaviour between the L^2 and the H^1 optimization. To this end, we consider the tracking problem from Section 5.3.1 again using both L^2 and H^1 optimization. We remark that, due to the different norms determining the cost of the control, a comparison is not straightforward. To present a meaningful comparison, we consider parameter settings that achieve similar values of the tracking term, which is the main goal of the optimization.

		L^2				H^1			
tol	ν	$\frac{1}{2}\|\rho - \rho_d\|_Y^2$	$\|\nabla_{L^2}\hat{J}\|_{L^2}^2$	k	CPU	$\frac{1}{2}\|\rho - \rho_d\|_Y^2$	$\|\nabla_{H^1}\hat{J}\|_{H^1}^2$	k	CPU
10^{-6}	10^{-8}	$2.6 \cdot 10^{-7}$	$9.3 \cdot 10^{-7}$	15	202	$5.4 \cdot 10^{-5}$	$2.2 \cdot 10^{-8}$	3	57
	10^{-10}	$1.9 \cdot 10^{-7}$	$7.6 \cdot 10^{-7}$	22	288	$4.4 \cdot 10^{-6}$	$5.9 \cdot 10^{-8}$	3	59
	10^{-12}	$2.3 \cdot 10^{-7}$	$8.4 \cdot 10^{-7}$	17	228	$4.3 \cdot 10^{-6}$	$6.6 \cdot 10^{-8}$	3	59
	10^{-14}	$2.7 \cdot 10^{-7}$	$9.2 \cdot 10^{-7}$	15	202	$4.3 \cdot 10^{-6}$	$6.5 \cdot 10^{-8}$	3	59
	10^{-15}	$2.6 \cdot 10^{-7}$	$9.3 \cdot 10^{-7}$	15	201	$4.3 \cdot 10^{-6}$	$6.5 \cdot 10^{-8}$	3	59
10^{-8}	10^{-8}	$6.2 \cdot 10^{-8}$	$8.4 \cdot 10^{-7}$	40	523	$5.5 \cdot 10^{-5}$	$2.2 \cdot 10^{-9}$	7	133
	10^{-10}	$6.3 \cdot 10^{-8}$	$1.7 \cdot 10^{-6}$	40	525	$4.6 \cdot 10^{-7}$	$6.9 \cdot 10^{-9}$	9	175
	10^{-12}	$4.7 \cdot 10^{-8}$	$1.2 \cdot 10^{-6}$	40	521	$1.7 \cdot 10^{-8}$	$9.3 \cdot 10^{-9}$	14	271
	10^{-14}	$4.6 \cdot 10^{-8}$	$3.8 \cdot 10^{-7}$	40	519	$1.5 \cdot 10^{-8}$	$5.5 \cdot 10^{-9}$	17	329
	10^{-15}	$4.3 \cdot 10^{-8}$	$1.8 \cdot 10^{-6}$	40	524	$1.3 \cdot 10^{-7}$	$1.1 \cdot 10^{-8}$	10	210

Table 5.6.: Comparison between the L^2 and H^1 optimization problems. CPU is the wall time given in seconds (see Section 5.3.6 for discussion on time measurements) and k denotes the number of iterations.

As shown in Table 5.6, to achieve tracking values in the order of 10^{-7} our code needs a similar amount of CPU time for both problems. However, choosing a more strict tolerance for the norm of the gradient, only convergence for the H^1 problem is achieved; the L^2 solution process is stopped after the maximum number of iterations ($k = 40$), which will possibly converge in a much higher number of iterations. For small values of J_β of about 10^{-8}, the H^1 problem is solved significantly faster. This can be motivated by the fact that H^1 is a smaller function space to search the optimal control in comparison to L^2.

5.3.6. Dependence on particle number

DFT was introduced to overcome the exponential increase in complexity of the multi-particle Schrödinger equation. The Kohn-Sham problem (2.33) increases only linear with the number of particles N. The aim of this section is to illustrate this fact by looking at the computational performance of our code. To this end, we solve optimization problems similar to the one described in Section 5.3.1 with different numbers of electrons involved. Accordingly, the initial condition Ψ^0 and the density ρ_d are adapted such that the control

N	k	Wall/s	CPU/s	Wall per iter/s	CPU per iter/s
2	8	109	214	13.6	26.8
3	8	124	254	15.4	31.8
4	8	138	292	17.3	36.5
5	9	172	373	19.2	41.5
6	8	171	375	21.4	46.9
7	8	188	418	23.6	52.3
8	12	307	689	25.6	57.5
9	9	242	547	26.9	60.7
10	17	492	1119	28.9	65.8

Table 5.7.: Runtime of the optimization for different number of particles N. The number of NCG iterations is denoted by k. Both, the wall time per iteration and the CPU time per iteration increase linearly with the particle number N.

u^{test} is the same for all N. We use the following setting for these experiments: $T = 0.1$, $h = 0.175$ in (5.6); $M = 100$, $N_x = N_y = 41$ in Algorithm 1; $\nu = 10^{-10}$ and $tol = 10^{-6}$ in Algorithm 5.

For particle numbers $N = 2, \ldots, 10$, we record the number of NCG iterations k and measure the "wall time" and the "CPU time". The "wall time" is the real time elapsing on a clock between start and end of the optimization process. This time contains fluctuation due to other processes that run on the computer in the background. To minimize their influence the average of three different runs has been taken. The "CPU time" on the other hand is the time spend only by the MATLAB processes on the CPUs. In contrast to the wall time this is the sum of execution time spend on all CPUs.

The results in Table 5.7 show that the CPU time is relatively constant a factor of $2 - 2.2$ times larger than the Wall time; this factor is slightly increasing for larger N. This means that the built-in parallelization of MATLAB achieves to use on average 2 of the 4 available cores of the machine. The automatic built-in parallelization is mainly used for the FFT, where most of the computation time is spent. However, due to the spectral convergence of the method (see Section 5.1.4), the space grid can be relatively coarse and thus the matrices that are transformed via FFT are relatively small (in this example 41×41). It is well-known that the FFT cannot be efficiently parallelized for such small matrices, see [FFT] for a speedup benchmark of the FFTW library used by MATLAB.

We want to remark that in principle much better parallelization than this can be achieved. Instead of parallelizing the FFT in the innermost loop, it is possible to parallelize the outer loop over the particles (lines 5–7 and 9–11 in Algorithm 1 and lines 4–6 and 8–11 in Algorithm 2). Hence, it should be possible to arrive at nearly linear scaling of parallelization with the numbers of cores used, as long as the particle number N is a multiple of the number of cores. However, to implement this, detailed control over the execution path and the access of the wavefunction is needed, e.g., to avoid unnecessary synchronization barriers. This is not possible in MATLAB. Instead, a more low-level programming language like C needs to be utilised for this. However, as the implementation of a high-performance computing software is a complex separate subject that was not part of this project, we did not extend our work in this direction.

As can be seen in Table 5.7, the number of NCG iterations needed for convergence of the algorithm varies with the particle number as the optimization problem differs with N. To study the computational effort as a function of the particle number, we

determined therefore the time (wall and CPU) per iteration, shown in the fifth and sixth column of Table 5.7.

The values of wall time per iteration are depicted in Figure 5.17 together with the linear fit of the data,

$$t_{wall/iter,fit}(N) = (1.9446N + 9.6409)\text{s}, \tag{5.48}$$

which shows that the runtime is a nearly perfect affine linear function of the particle number.
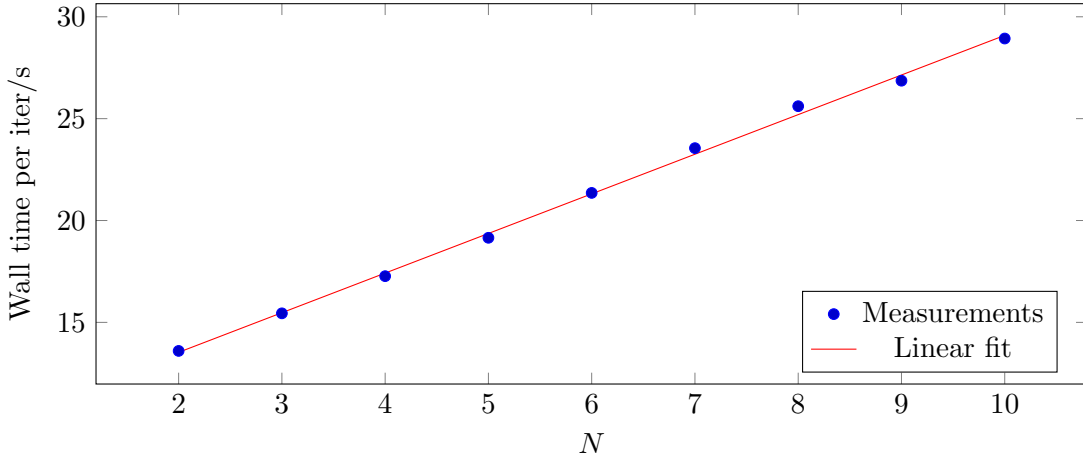

Figure 5.17.: The runtime of the optimization as wall time per NCG iteration is an affine linear function of the particle number N. Plotted are the values from Table 5.7 together with the linear fit (5.48).

To verify from the data values that the increase is indeed linear, we considered the model $t_{wall/iter} = aN^b + c$. To this end, we subtracted the fixed offset of $c = 9.6409$s from the data points and did a fit of $\log N$ to $\log(t_{wall/iter}/\text{s} - 9.6409)$. From this doubly logarithmic fit, we obtained and exponent of $b = 0.99845$. This result confirms that $t_{wall/iter}$ is indeed an affine linear function of N as expected and this demonstrates that TDDFT is capable of dealing with quantum systems consisting of a large number of particles.

5.4. Conclusion

This chapter consisted of three distinct parts. In the first part, we gave an introduction to operator splitting methods and discussed in particular the Strang splitting with spectral evaluation of the Laplacian. This method is advantageous because it provides second-order convergence in time and spectral accuracy in space thus leading to a very fast algorithm. The convergence was numerically verified for both the forward TDKS equation and the adjoint TDKS equation.

In the second part, we presented classical optimization method used to solve our optimization method. We introduced both the NCG and the BFGS method. A derivation of the matrix-free BFGS was shown.

Afterwards, we presented results of four different numerical experiments demonstrating the abilities of our method. Optimal control problems with different kinds of targets were solved. Furthermore, we investigated the difference between L^2 and H^1 optimization problems, showing the advantages of using the latter setting, and numerically con-

firmed the linear increase in complexity as a function of the particle number which was the main motivation for using TDDFT.

6. Summary

In this thesis, a comprehensive analysis of optimal control problems governed by the time-dependent Kohn-Sham (TDKS) equations was given. The TDKS equations are one of the most successful frameworks to calculate many-particle systems. The active control of chemical reactions on a molecular level is one of the ultimate goals of quantum control. Due to its ability to deal with a high number of particles, the TDKS equations seem to be the method of choice for the purpose of the control. There are already a large number of optimal control applications using the TDKS equations. However, both the theory of the TDKS equations and the theory of optimal control problems governed by them were still in its infancy. The present work has provided many contributions to close this gap.

First, the basic concepts of DFT and the Kohn-Sham approach were reviewed. Subsequently, this was extended to the time-dependent DFT framework of Runge and Gross. A mathematical rigorous statement of the van Leeuwen's theorem was given and a detailed proof for this theorem was presented. Furthermore, the TDKS equations in the LDA framework were introduced.

The TDKS equations were then analysed in detail. Using energy estimates and a Galerkin framework, the existence of a unique solution of this PDE was proved in a setting that is suitable for optimal control purposes. Furthermore, given stronger assumptions on the initial condition, higher regularity of the solution was shown. These results are essential for the TDKS optimal control problem to be well-defined.

With these preparations, it was possible to study optimal control problems governed by the TDKS equations with both tracking type targets and terminal targets. This was accomplished in the Lagrange framework. To this end, the differentiability properties of the TDKS equations were analysed. This analysis was very challenging for several reasons. One of the reasons was that the KS potential has a complicated nonlinear dependence on the wavefunction. Another reason was that the KS potential is a real-valued function of a complex variable, hence it is not complex differentiable. For the latter issue two different approaches were presented. Thereafter, the differentiability results were used to prove the existence of a minimizer of the optimal control problem and to characterize its solutions via a first-order optimality condition. Furthermore, the gradient of the reduced cost functional was derived which is essential for numerical implementations of optimization methods.

In the final chapter, results of several numerical experiments demonstrated the abilities of the method presented in this thesis. To this end, a very efficient discretization scheme for the TDKS equations was presented and the convergence order of the scheme was discussed. Moreover, gradient based numerical optimization methods were illustrated. These methods were used to perform various numerical experiments with three different kinds of optimization targets. Furthermore, L^2 and H^1 optimization problems were compared. Finally, it was verified that the numerical complexity of the optimization process scales indeed only linearly with the number of particles involved, thus confirming that the TDKS model is an excellent choice for multi-particle quantum control applications.

A. Appendix

A.1. Differentiability of the KS potential

In this section, we further discuss the issue of differentiability. In the discussion above, the reduced cost functional is a real valued function of the real control variable. Complex differentiability issues arise when trying to calculate the derivative using the chain rule with intermediate complex functions such as Ψ. This is a delicate step, and for clarity, in this section, we analyse the methodology used above. Thereafter, we introduce a different but equivalent approach rewriting complex equations as a system of real equations. A somewhat different problem is considered in [TV16] where a problem with a complex control is investigated.

The map defined by the linear SE is Fréchet differentiable in the standard sense. However, in the nonlinear TDKS model, the Kohn-Sham potential is a map from $\mathbb{C}$ to $\mathbb{R}$; and it is known from standard results in complex analysis that this map cannot be complex differentiable. This is because complex differentiability is stronger than real total differentiability by the requirement that the Cauchy-Riemann differential equations need to be satisfied.

Example A.1. *As an example, we show that the complex absolute value squared, $z \mapsto f(z) := |z|^2$, which also appears in the KS potential, is not complex differentiable but only real-differentiable. To this end, we consider the directional derivative of f at z in direction h,*

$$f'(z; h) = \lim_{t \searrow 0} \frac{(z + th)\overline{(z + th)} - z\overline{z}}{t} = z\overline{h} + \overline{z}h = 2\operatorname{Re}(\overline{z}h).$$

The map $h \mapsto f'(z; h) = 2\operatorname{Re}(\overline{z}h)$ is obviously not complex linear. Hence, f is not complex differentiable. However, $h \mapsto f'(z; h)$ is real-linear according to Definition 4.1 and therefore, f is real-differentiable.

In the following theorem, we want to address this point and show that the discrepancy between complex differentiability and real-differentiability is that the corresponding derivative is linear with respect to complex, instead of real, scalars.

Theorem A.2. *Let $f : \mathbb{C} \to \mathbb{C}$ and $\hat{f} : \mathbb{R}^2 \to \mathbb{R}^2$ such that $f(x+iy) = \hat{f}(x, y)_1 + i\hat{f}(x, y)_2$. Then the following statements are equivalent:*

    1. $\hat{f}$ is totally differentiable from $\mathbb{R}^2$ to $\mathbb{R}^2$;

    2. f is real-differentiable from $\mathbb{C}$ to $\mathbb{C}$.

Proof. First, we prove that 1. implies 2. We have

$$\hat{f}(x_1 + h_1, x_2 + h_2) = \hat{f}(x_1, x_2) + \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} + \hat{\phi}(h_1, h_2) \tag{A.1}$$

with

$$\lim_{(h_1, h_2) \to (0,0)} \frac{\hat{\phi}(h_1, h_2)}{\|(h_1, h_2)^T\|_2} = 0. \tag{A.2}$$

113

From (A.1), we find

$$f(x_1 + ix_2 + h_1 + ih_2) = f(x_1 + ix_2) + Df(h) + \phi(h_1 + ih_2) \qquad (A.3)$$

where $h = h_1 + ih_2$, $\phi(h_1 + ih_2) = \hat{\phi}(h_1, h_2)_1 + i\hat{\phi}(h_1, h_2)_2$, and

$$Df(h) = d_{11}h_1 + d_{12}h_2 + id_{21}h_1 + id_{22}h_2. \qquad (A.4)$$

Obviously, $Df(h + g) = Df(h) + Df(g)$ for $h, g \in \mathbb{C}$ and $Df(\alpha h) = \alpha Df(h)$ for $\alpha \in \mathbb{R}$, hence Df is real-linear. The fact that $\lim_{h \to 0} \frac{\phi(h_1 + ih_2)}{|h|} = 0$ follows directly from (A.2).

Next, we prove that 2. implies 1. We have for $x, h \in \mathbb{C}$

$$f(x + h) = f(x) + Df(h) + \phi(h) \qquad (A.5)$$

with

$$\lim_{h \to 0} \frac{\operatorname{Re} \phi(h) + i \operatorname{Im} \phi(h)}{|h|} = 0 \qquad (A.6)$$

and Df real-linear. Therefore, we have

$$\hat{f}(x_1 + h_1, x_2 + h_2) = \begin{pmatrix} \operatorname{Re} f(x) \\ \operatorname{Im} f(x) \end{pmatrix} + \begin{pmatrix} \operatorname{Re} Df(h) \\ \operatorname{Im} Df(h) \end{pmatrix} + \begin{pmatrix} \operatorname{Re} \phi(h) \\ \operatorname{Im} \phi(h) \end{pmatrix}. \qquad (A.7)$$

As $\operatorname{Re} Df(h)$ and $\operatorname{Im} Df(h)$ are linear with respect to real scalars, there exists a matrix $D\hat{f} \in \mathbb{R}^{2 \times 2}$ such that $\begin{pmatrix} \operatorname{Re} Df(h) \\ \operatorname{Im} Df(h) \end{pmatrix} = D\hat{f} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$.

From (A.6), it follows that $\lim_{(h_1, h_2) \to (0,0)} \frac{1}{\|(h_1, h_2)^T\|_2} \begin{pmatrix} \operatorname{Re} \phi(h) \\ \operatorname{Im} \phi(h) \end{pmatrix} = 0$. $\qquad \square$

Remark. Df defined in (A.4) is in general not linear with respect to multiplication with complex scalars. However, if the Cauchy-Riemann differential equations are satisfied, i.e. $d_{11} = d_{22}$ and $d_{12} = -d_{21}$, one can rewrite Df as

$$Df(h) = (d_{11} + id_{21})h$$

which is complex linear and hence f is complex differentiable.

Requiring only real-linearity of the derivative instead of complex linearity is therefore equivalent to requiring total differentiability of the function as a map from $\mathbb{R}^2$ to $\mathbb{R}^2$ without the requirement to satisfy the Cauchy-Riemann equations. All results that hold for total differentiable functions $\hat{f} : \mathbb{R}^2 \to \mathbb{R}^2$ do not make use of the special properties of complex differentiability and hold therefore also for real-differentiable functions from $\mathbb{C}$ to $\mathbb{C}$. In particular, the results known from real analysis as the mean value theorem, the chain rule and the implicit function theorem hold for real-differentiable complex functions. This is the approach used in Sections 4.3–4.4.

An alternative — and by Theorem A.2 fully equivalent — way of dealing with a complex valued optimization problem is to rewrite the N complex equations in $2N$ real equations for the real and imaginary part. As a supplement to the results already obtained, we consider this approach in the reminder of this section. To this end, we introduce the following vectors:

$$\hat{\Psi} = \begin{pmatrix} \psi_1^r \\ \psi_1^i \\ \psi_2^r \\ \psi_2^i \\ \vdots \\ \psi_N^i \end{pmatrix}, \quad \psi_j^r = \operatorname{Re} \psi_j, \quad \psi_j^i = \operatorname{Im} \psi_j, \qquad \hat{\Lambda} = \begin{pmatrix} \lambda_1^r \\ \lambda_1^i \\ \lambda_2^r \\ \lambda_2^i \\ \vdots \\ \lambda_N^i \end{pmatrix}.$$

With this definition, the TDKS equations (4.8a)–(4.8b) can be equivalently rewritten as

$$\frac{\partial}{\partial t}\begin{pmatrix}\psi_j^r\\\psi_j^i\end{pmatrix}=\begin{pmatrix}0&A\\-A&0\end{pmatrix}\begin{pmatrix}\psi_j^r\\\psi_j^i\end{pmatrix},\quad\begin{pmatrix}\psi_j^r\\\psi_j^i\end{pmatrix}(x,0)=\begin{pmatrix}\operatorname{Re}\psi_j^0\\\operatorname{Im}\psi_j^0\end{pmatrix},\quad j=1,\ldots,N,\qquad\text{(A.8)}$$

where $A:=-\nabla^2+V_{ext}(x,t;u)+V_{Hxc}(x,t;\Psi)$. In this picture, the density is given by

$$\rho=\sum_{j=1}^N(\psi_j^r)^2+(\psi_j^i)^2.\qquad\text{(A.9)}$$

Then, the optimization problem reads as follows:

$$\min_{(\hat\Psi,u)\in(\hat Z,H^1(0,T))}J(\hat\Psi,u)\qquad\text{subject to (A.8)},\qquad\text{(A.10)}$$

where $\hat Z=\left\{\hat\Psi\in\hat Z\,\middle|\,\hat\Psi'\in\hat X^*\right\}$, $\hat Z=L^\infty(0,T;H^2(\Omega;\mathbb{R}^{2N})\cap H_0^1(\Omega;\mathbb{R}^{2N}))$ and the cost functional is given by

$$J(\hat\Psi,u)=\frac{\beta}{2}\int_0^T\int_\Omega(\rho(x,t)-\rho_d(x,t))^2\mathrm{d}x\mathrm{d}t+\frac{\eta}{2}\int_\Omega\chi_A(x)\rho(x,T)\mathrm{d}x+\frac{\nu}{2}\|u\|_{H^1}^2.$$

To derive the optimality system, we introduce the Lagrange multipliers λ_j^r, λ_j^i, $j=1,\ldots,N$, and the following Lagrange function:

$$L(\hat\Psi,\hat\Lambda,u)=J(\hat\Psi,u)+\sum_{j=1}^N\int_0^T\int_\Omega\begin{pmatrix}\lambda_j^r&\lambda_j^i\end{pmatrix}\left(\frac{\partial}{\partial t}-\begin{pmatrix}0&A\\-A&0\end{pmatrix}\right)\begin{pmatrix}\psi_j^r\\\psi_j^i\end{pmatrix}\mathrm{d}x\mathrm{d}t$$

$$=J(\hat\Psi,u)+\sum_{j=1}^N\int_0^T\int_\Omega\lambda_j^r\left(\frac{\partial\psi_j^r}{\partial t}-A\psi_j^i\right)+\lambda_j^i\left(\frac{\partial\psi_j^i}{\partial t}+A\psi_j^r\right)\mathrm{d}x\mathrm{d}t.$$

Similar as in the previous section, we take the directional derivatives with respect to ψ_j^r, ψ_j^i, λ_j^r, λ_j^i, and u and require that these derivatives are zero for all admissible variations. The variation with respect to λ_j^r, λ_j^i reproduces the forward equation (A.8) and the variation with respect to ψ_j^r, ψ_j^i gives the following system of adjoint equations:

$$\frac{\partial}{\partial t}\begin{pmatrix}\lambda_j^r\\\lambda_j^i\end{pmatrix}=\begin{pmatrix}0&A\\-A&0\end{pmatrix}\begin{pmatrix}\lambda_j^r\\\lambda_j^i\end{pmatrix}+\begin{pmatrix}b_j&0\\0&b_j\end{pmatrix}\begin{pmatrix}\psi_j^r\\\psi_j^i\end{pmatrix},\qquad\text{(A.11a)}$$

$$b_j=\sum_{j=1}^N V_H\left(2c_j(\cdot,t)\right)(x,t)+2\frac{\partial V_{xc}}{\partial\rho}\sum_{j=1}^N c_j(x,t)-2\beta(\rho-\rho_d),\qquad\text{(A.11b)}$$

$$c_j(y,t)=\psi_j^i(y,t)\lambda_j^r(y,t)-\psi_j^r(y,t)\lambda_j^i(y,t),\qquad\text{(A.11c)}$$

$$\begin{pmatrix}\lambda_j^r(x,T)\\\lambda_j^i(x,T)\end{pmatrix}=-\eta\chi_A(x)\begin{pmatrix}\psi_j^r(x,T)\\\psi_j^i(x,T)\end{pmatrix}.\qquad\text{(A.11d)}$$

To derive the optimality condition, i.e. the zero gradient condition, the variation with respect to the control u is considered. We obtain $\nu u+\mu$ where μ is the $H^1(0,T)$-Riesz representative of the functional

$$g(\delta u):=-\sum_{j=1}^N\int_0^T\int_\Omega\left(-\lambda_j^r\delta u V_u\psi_j^i+\lambda_j^i\delta u V_u\psi_j^r\right)\mathrm{d}x\mathrm{d}t.\qquad\text{(A.12)}$$

Comparing this result with (4.8), we realize the correspondence

$$\lambda_j^i = -\operatorname{Re}\lambda_j, \qquad\qquad \lambda_j^r = \operatorname{Im}\lambda_j \qquad\qquad (A.13)$$

between the Lagrange multipliers of the two optimality systems. The differentiation of the KS potential in (A.8) exhibits no subtleties as the KS potential is totally differentiable over the field of the reals. Both the complex and the real approaches are equivalent and arrive at the same results. The calculations in Section 4.3–4.4 can be performed in the real setting without substantial changes by replacing, e.g., the complex absolute value with the real Euclidean norm.

A.2. Functional analysis

A.2.1. Embedding theorems

In this section, we collect embedding theorems and convergence results and that are used in this thesis.

Definition A.3. *We say X is embedded in Y, written as $X \hookrightarrow Y$, if $X \subset Y$ and there exists a constant c such that $\|v\|_Y \le c\|v\|_X$ for all $v \in X$.*

A map f is called compact, if $\overline{f(A)}$ is compact for all bounded sets A.

X is called compactly embedded in Y, written as $X \Subset Y$, if $X \hookrightarrow Y$ with a compact identity.

We recall the following statements from functional analysis, see, e.g. [Cia13, Dob06].

Theorem A.4. • *Let X be a reflexive Banach space. Then the closed unit ball is weakly sequentially compact.*

• *Let X, Y, Z be Banach spaces and $X \hookrightarrow Y \hookrightarrow Z$ with one of the embeddings being compact. Then $X \hookrightarrow Z$ is compact, i.e. $X \Subset Z$.*

• *Let X and Y be Banach spaces and $(f_n)_n$ be a sequence in X. If $X \Subset Y$ and $f^n \to f$ in X, then $f^n \to f$ in Y.*

Theorem A.5 (Embedding of L^p spaces). *Let $\Omega \subset \mathbb{R}^n$ with $\mu(\Omega) < \infty$ and $1 \le p \le q \le \infty$. Then*

$$L^q(\Omega) \hookrightarrow L^p(\Omega), \qquad \text{in particular } L^\infty(\Omega) \hookrightarrow L^p(\Omega) \hookrightarrow L^1(\Omega) \qquad (A.14)$$

Theorem A.6 (Morrey, embedding into continuous spaces). *If Ω satisfies the cone condition, then for $p > n$ we have the embedding $H^{1,p}(\Omega) \hookrightarrow C(\Omega) \cap L^\infty(\Omega)$. This means that any function $u \in H^{1,p}(\Omega)$ can be changed on a set of measure zero such that u is continuous on Ω.*

Theorem A.7 (Sobolev embeddings). *Let $\Omega \subset \mathbb{R}^n$ be a domain that satisfies the cone condition, $m \ge 1$, $j \ge 0$ be integers and $1 \le p < \infty$. Then the following embeddings hold; see [AF03, Theorem 4.12]:*

$H^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$	$\frac{1}{q} = \frac{1}{p} - \frac{m}{n}$ *if* $m < \frac{n}{p}$,
$H^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$	*for all* $1 \le q < \infty$ *if* $m = \frac{n}{p}$,
$H^{j+m,p}(\Omega) \hookrightarrow C^{j,\lambda}(\overline{\Omega})$, $0 < \lambda \le m - n/p$,	*if* $\frac{n}{p} < m < \frac{n}{p} + 1$,
$H^{j+m,p}(\Omega) \hookrightarrow C^{j,\lambda}(\overline{\Omega})$	$\forall 0 < \lambda < 1$ *if* $m = \frac{n}{p} + 1$,
$H^{j+m,p}(\Omega) \hookrightarrow C^{j,1}(\overline{\Omega})$	*if* $p = 1$ *and* $n + 1 < m$.

For the embeddings to continuous spaces the Lipschitz condition on Ω is needed that is stronger than the cone condition.

All of the above embeddings hold for arbitrary Ω if $H^{m,p}$ is replaced with $H_0^{m,p}$.

The third embedding gives $H^1((a,b)) \hookrightarrow C^0([a,b])$, but this is sharp in the sense that for $n \geq 2$ H^1-functions are not continuous ($-\log|x|^\alpha$ can be made to an example that is discontinuous), see also [Cia13, p. 332].

The following theorem improves the embeddings to be compact and can be found in [AF03, Theorem 6.3].

Theorem A.8 (Rellich-Kondrachov)**.** *Let* $\Omega \subset \mathbb{R}^n$ *be bounded and satisfy the cone condition,* $m \geq 1$, $j \geq 0$ *be integers and* $1 \leq p < \infty$. *Then the following embeddings are compact:*

$H^{m,p}(\Omega) \Subset L^q(\Omega)$ *for all* q *with* $1 \leq q < p^* = \frac{np}{n-mp}$ *if* $m < \frac{n}{p}$,

$H^{m,p}(\Omega) \Subset L^q(\Omega)$ *for all* q *with* $1 \leq q < \infty$ *if* $m = \frac{n}{p}$,

$H^{m,p}(\Omega) \Subset C(\overline{\Omega})$ *if* $\frac{n}{p} < m$.

All of the above embeddings hold for arbitrary Ω *if* $H^{m,p}$ *is replaced with* $H_0^{m,p}$.

This implies (using the first line for $n > p$, the second for $n = p$ and the chain $H^{1,p} \Subset C \hookrightarrow L^\infty \hookrightarrow L^p$) the following embedding independent of the dimension n:

$$H^{1,p}(\Omega) \Subset L^p(\Omega), \quad p \geq 1. \tag{A.15}$$

A.2.2. Evolution spaces

In this section, we collect some results concerning function spaces used to characterize evolution problems. These spaces are also called Bochner spaces. The reason for the introduction of Bochner spaces is that in evolution equations, like parabolic equations and the Schrödinger equation, the time and space variables appear with a different order of the derivative. This feature results in a different regularity of the solution in space and time and suggests a separate treatment of the space and time variables. More details on this topic and proofs to the statements presented here can be found in [Zei90, Chapter 23].

Definition A.9. *Let* X *be a Banach space and* $0 < T < \infty$. *Then* $C^m([0,T], X)$ *is the space of all continuous functions* $u : [0,T] \to X$, *which have continuous derivatives up to order* m, *and is endowed with the norm*

$$\|u\|_{C^m([0,T],X)} = \sum_{i=0}^m \max_{0 \leq t \leq T} \|u^{(i)}(t)\|_X. \tag{A.16}$$

Further, we define the Lebesgue space $L^p(0,T;X)$ *as the space of all measurable functions* $u : (0,T) \to X$ *with*

$$\|u\|_{L^p(0,T;X)} = \left(\int_0^T \|u(t)\|_X^p \mathrm{d}t \right)^{1/p} < \infty. \tag{A.17}$$

Next, we summarize some facts about the Lebesgue spaces $L^p(0,T;X)$.

Theorem A.10. *Let* X *be a Banach space and* $0 < T < \infty$, *then the following statements hold true:*

- $C^m([0,T], X)$ *and* $L^p(0,T;X)$ *are Banach spaces.*
- $C([0,T];X)$ *is dense in* $L^p(0,T;X)$ *and the embedding*

$$C([0,T];X) \hookrightarrow L^p(0,T;X), \tag{A.18}$$

is continuous.

- *If X is a Hilbert space, then $L^2(0, T; X)$ is a Hilbert space.*

- *If X is separable, then $L^p(0, T; X)$ is separable for $1 \leq p < \infty$.*

- *If $X \hookrightarrow Y$ is continuous, then the embedding*

$$L^r(0, T; X) \hookrightarrow L^p(0, T; Y), \quad 1 \leq p \leq r \leq \infty, \tag{A.19}$$

  is continuous.

- *The dual space $L^p(0, T; X)^*$ is norm-isomorphic to $L^q(0, T; X^*)$ with $\frac{1}{p} + \frac{1}{q} = 1$.*

The generalized derivative in Bochner spaces is defined by integration by parts just as the weak derivative in standard Sobolev spaces. In the following, we introduce the Sobolev space $W(0, T; V)$. For $V \subset H \subset V^*$ we define the space

$$W(0, T; V) := \{u \in X : u' \in X^*\},$$

where $X = L^p(0, T, V)$. This is a Banach space with the norm

$$\|u\|_W = \|u\|_{L^p(0,T;V)} + \|u'\|_{L^q(0,T;V^*)}. \tag{A.20}$$

For this space, we have the following embedding theorem.

Theorem A.11. *The embedding $W(0, T; V) \hookrightarrow C([0, T], H)$ is continuous. More precisely there exists a unique continuous function $\tilde{u} : [0, T] \to H$ that coincides almost everywhere with u. Furthermore,*

$$\max_{0 \leq t \leq T} \|\tilde{u}(t)\|_H \leq C\|u\|_W. \tag{A.21}$$

Moreover, the set of all polynomials $w : [0, T] \to V$ is dense in $W(0, T; V)$, $L^p(0, T; V)$ and $L^p(0, T; H)$.

The following compact embedding theorem is due to Aubin [Aub63] and Lions [Lio69, 1.5.2].

Theorem A.12. *Given three Banach spaces $B_0 \Subset B \hookrightarrow B_1$ with B_0, B_1 reflexive and the embedding $B_0 \hookrightarrow B$ being compact, then the space*

$$W := \left\{ v \,|\, v \in L^p(0, T; B_0), \, v' \in L^q(0, T; B_1) \right\}, \quad 1 < p, q < \infty,$$

with norm

$$\|v\|_W = \|v\|_{L^p(0,T;B_0)} + \|v'\|_{L^q(0,T;B_1)}, \quad 1 < p, q < \infty,$$

is compactly embedded in $L^p(0, T; B)$.

A.2.3. Functional derivatives

We summarize the different notions of derivatives in function spaces and recall the implicit function theorem for Banach spaces.

Definition A.13 (Gâteaux derivative)**.** *Let V_1, V_2 be locally convex spaces and $f : V_1 \to V_2$ be given. Take $x, h \in V_1$. We call f directional differentiable at x in direction h if the limit*

$$\delta f(x; h) := \lim_{t \searrow 0} \frac{f(x + th) - f(x)}{t} \tag{A.22}$$

exists in V_2.

The function f is called Gâteaux differentiable at x if f is directional differentiable in all directions and the mapping $A : h \mapsto \delta f(x; h)$ is linear and continuous. The mapping $A \in \mathcal{L}(V_1, V_2)$ is called the Gâteaux derivative of f, $\delta f(x; h) = Ah$. If and only if V_2 is the field belonging to V_1, the derivative is an element of the dual space V_1^*.

Definition A.14. *$f : V_1 \to V_2$ is called* Fréchet differentiable *at $a \in \Omega \subset V_1$ if there exists an $A \in \mathcal{L}(V_1; V_2)$ such that*

$$f(a + h) = f(a) + Ah + \|h\|_X \delta(h) \ \text{ with } \ \lim_{h \to 0} \delta(h) = 0 \in V_2. \tag{A.23}$$

Definition A.15. *Consider the setting of Definition A.13. If V_1 is a Hilbert space and V_2 is the field belonging to V_1, i.e. f is a functional, then the Riesz representative of Df using the inner product of V_1 is called the* gradient *of f and written as ∇f. Hence $\nabla f(x) \in V_1$,*

$$Df(x)(\delta x) = \langle \nabla f(x), \, \delta x \rangle_{V_1}.$$

We conclude this section with the implicit function theorem as found in [Cia13, Thm. 7.13-1]. We remark that this theorem also holds for real-Fréchet differentiable functions.

Theorem A.16 (Implicit function theorem)**.** *Let V, Y and Z be three Banach spaces and let O be an open subset of $V \times Y$. Consider a function $c : O \to Z$ such that $c \in C(O; Z)$ and assume that there exists a pair $(\hat{x}, \hat{y}) \in O$ such that*

(a) *$c(\hat{x}, \hat{y}) = 0$;*

(b) *c is Fréchet differentiable at $(\hat{x}, \hat{y})$;*

(c) *the derivative $D_x c(x, y) \in \mathcal{L}(V, Z)$ exists at all points $(x, y) \in O$ and $D_x c \in C(O; \mathcal{L}(V, Z))$;*

(d) *$D_x c(\hat{x}, \hat{y}) \in \mathcal{L}(V, Z)$ is a bijection, so that $(D_x c(\hat{x}, \hat{y}))^{-1} \in \mathcal{L}(Z, V)$,*

where $\mathcal{L}(Z, V)$ denotes the set of all bounded linear operators from Z to V.

Then there exists a function $x(y)$ of y in an open neighbourhood of $\hat{y}$ and it is differentiable at $\hat{y}$ with derivative given by

$$D_y x(\hat{y}) = (-D_x c(\hat{x}, \hat{y}))^{-1} D_y c(\hat{x}, \hat{y}). \tag{A.24}$$

A.3. Exact solutions of the harmonic oscillator

A.3.1. Coherent states of the harmonic oscillator

The coherent states of the harmonic oscillators — sometimes also called Glauber coherent states in the honour of Roy Glauber, Nobel price 2005 for quantum optics — are minimum uncertainty states in the harmonic oscillator; see also [Sch26] for their first derivation. They possess the characteristic property that their position expectation value performs a harmonic oscillation. Due to these two features they are in some sense the most classical quantum states realized in a harmonic potential and allow the connection between the quantum and classical harmonic oscillator via the Ehrenfest theorem.

The coherent states are important for application in quantum optics as they are used to describe laser modes due to their non-dissipative character. Furthermore, they are valuable test states in numerics. We make use of them as initial states for control applications that have known analytical solutions in the uncontrolled case.

To derive the formulas of the coherent states φ_α, we follow the discussion in [Sch07] and consider a harmonic oscillator in one dimension with the following Hamiltonian:

$$H = -\frac{1}{2}\nabla^2 + \frac{1}{2}\omega^2 x^2, \tag{A.25}$$

where we use natural units ($\hbar = m = 1$). Introducing the characteristic length $x_0 = \frac{1}{\sqrt{\omega}}$, the energy eigenstates of (A.25) are given by

$$\psi_n = \frac{1}{\sqrt{2^n n! \sqrt{\pi} x_0}} e^{-\frac{1}{2}\left(\frac{x}{x_0}\right)^2} \mathfrak{H}_n\left(\frac{x}{x_0}\right), \tag{A.26}$$

where $\mathfrak{H}_n$ is the nth Hermite polynomial defined by

$$\mathfrak{H}_n(x) = (-1)^n e^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2}.$$

From the minimum uncertainty property, one can derive that the coherent states are the eigenstates of the annihilation operator a which is defined as

$$a\psi_n = \sqrt{n}\psi_{n-1}, \quad n \geq 1, \qquad\qquad a\psi_0 = 0. \tag{A.27}$$

The eigenvalue relation

$$a\varphi_\alpha = \alpha\varphi_\alpha, \quad \alpha \in \mathbb{C}, \tag{A.28}$$

allows an expansion of the coherent states into the energy eigenstates as follows:

$$\varphi_\alpha(x,t) = e^{-(|\alpha|^2 + i\omega t)/2} \sum_{n=0}^{\infty} \frac{\left(\alpha e^{-i\omega t}\right)^2}{\sqrt{n!}} \psi_n(x). \tag{A.29}$$

The position expectation value can be calculated using the eigenstate property (A.28). One finds

$$\langle x \rangle = \sqrt{2}x_0|\alpha|\cos(\omega t - \delta),$$

which is a harmonic oscillation and $\langle x \rangle$ has the same time dependence as the classical oscillation.

From (A.29) and (A.26) an explicit formula for the coherent states can be obtained as follows:

$$\varphi_\alpha(x,t) = \frac{1}{\sqrt[4]{\pi}\sqrt{x_0}} \exp\left(-i\left(\frac{\omega t}{2} - \frac{|\alpha|^2}{2}\sin 2(\omega t - \delta) + \frac{\sqrt{2}|\alpha|x}{x_0}\sin(\omega t - \delta)\right)\right.$$
$$\left. -\frac{1}{2x_0^2}\left(x - \sqrt{2}x_0|\alpha|\cos(\omega t - \delta)\right)^2\right).$$

Taking the absolute value squared, we find the probability density

$$|\varphi_\alpha(x,t)|^2 = \frac{1}{\sqrt{\pi}x_0} \exp\left(-\frac{\left(x - \sqrt{2}x_0|\alpha|\cos(\omega t - \delta)\right)^2}{x_0^2}\right). \tag{A.30}$$

The density (A.30) describes a Gaussian that has constant width. This is achieved because all contributing states ψ_n stay in phase and justifies the name coherent states.

A.3.2. Two interacting electrons

To test the predictions of TDDFT, it is useful to have access to exact solution of multi-particle problems. Unfortunately, there are not many systems of interacting particles

that can be solved analytically. Here, we present some analytic solutions of the two-particle SE in the harmonic oscillator with the full Coulomb interaction between the two particles. The solutions in this section are based on the results of [Tau93].

We find the eigenstates for the stationary SE

$$H\Phi(x) = E\Phi(x),$$
$$H = \sum_{k=1}^{2} \left(\frac{1}{2}\nabla_k^2 + \frac{1}{2}\omega_0^2\|r_k\|^2\right) + \frac{1}{\|r_1 - r_2\|} \tag{A.31}$$

and then also provide analytic solutions for the corresponding time-dependent SE (A.36). In this section, we use natural units $\hbar = m = e = 1$ and denote the positions of the two electrons by $r_k = (x_k, y_k)^\top$, $k = 1, 2$.

To separate the problem, we introduce the relative coordinate $r = r_1 - r_2$ and the centre of mass (c.m.) coordinate $R = \frac{1}{2}(r_1 + r_2)$. Then the Hamiltonian H decouples exactly into $H = 2H_r + \frac{1}{2}H_R$ with

$$H_R = \frac{1}{2}\nabla_R^2 + \frac{1}{2}\omega_R^2 R^2, \tag{A.32}$$

$$H_r = \frac{1}{2}\nabla_r^2 + \frac{1}{2}\omega_r^2 r^2 + \frac{1}{2|r|}, \tag{A.33}$$

where $\omega_R = 2\omega_0$, $\omega_r = \frac{1}{2}\omega_0$. Due to the decoupling we can make the ansatz $\Phi(r, R) = \phi(R)\varphi(r)$ for the wavefunction.

We begin with solving the c.m. problem. The solutions of the non-interaction harmonic oscillator $-\frac{1}{2}\frac{d^2}{dx^2} + \frac{1}{2}\omega^2$ can be found in standard textbooks and is given in (A.26),

$$\tilde{\phi}_{n,\omega}(x) = \frac{\sqrt[4]{\omega}}{\sqrt[4]{\pi}\sqrt{2^n n!}}e^{-\frac{\omega x^2}{2}}\mathfrak{H}_n(\sqrt{\omega}x),$$

where $\mathfrak{H}_n$ is the n-th Hermite polynomial. Therefore, the eigenstates of the c.m. problem $H_R\phi(R) = E_R\phi(R)$ are given by

$$\phi_{nm,\omega_R}(R_x, R_y) = \tilde{\phi}_{n,\omega_R}(R_x)\tilde{\phi}_{m,\omega_R}(R_y), \tag{A.34}$$
$$R_x = \frac{x_1 + x_2}{2}, \quad R_y = \frac{y_1 + y_2}{2}.$$

For special frequencies $\omega_0 = \frac{1}{2l+1}$, $l = 1, 2, \ldots$ the relative equation $H_r\varphi = E_r\varphi$ can be solved. To this end, the relative equation is decomposed using spherical coordinates into a radial and an angular part. The angular part is given by the spherical harmonics as both $\frac{1}{2}\omega_r r^2$ and $\frac{1}{2|r|}$ are only functions of the radius. For a given angular quantum number l the solution of the radial part is given by

$$\frac{u_l(\rho)}{\sqrt{\rho}}, \quad u_l(\rho) = \rho^{l+1/2}e^{-\frac{\rho^2}{8(l+1)}}\left(1 + \frac{\rho}{2(l+1/2)}\right).$$

Combining this with the spherical harmonics, the solution to the radial equation in polar and cartesian coordinates is given by

$$\varphi_p(\rho, \theta) = \frac{1}{\sqrt{2\pi(2+4l)^l\left((3+4l)l! + \sqrt{2}(1+2l)^{3/2}\Gamma\left(l+\frac{1}{2}\right)\right)}}\frac{u_l(\rho)}{\sqrt{\rho}}e^{il\theta},$$

$$\varphi(x, y) = \varphi_p\left(\sqrt{x^2 + y^2}, \arctan\frac{y}{x}\right).$$

A solution to the full two-electron problem (A.31) is given by the product $\Phi_{nml} = \phi_{nm,\omega_R}\varphi_{l,\omega_r}$. For example, for $n = m = 0$, $l = 1$ we find

$$\Phi_{001}(r_1, r_2) = -\frac{(\|r_1 - r_2\| + 3)\,(x_1 - x_2 + i(y_1 - y_2))e^{-\frac{1}{6}\left(\|r_1\|^2 + \|r_2\|^2\right)}}{9\pi\sqrt{14 + 3\sqrt{6\pi}}} \tag{A.35}$$

with energy eigenvalue $E_{001} = \frac{4}{3}$.

Solutions of the time-dependent SE

$$i\frac{\partial\Psi}{\partial t}(r_1, r_2, t) = \left(\sum_{k=1}^{2}\left(\frac{1}{2}\nabla_k^2 + \frac{1}{2}\omega_0^2\|r_k\|^2\right) + \frac{1}{\|r_1 - r_2\|}\right)\Psi(r_1, r_2, t) \tag{A.36}$$

can be obtained by multiplying the stationary solutions with the phase $e^{iE_{nml}t}$ such that $\Psi_{nml}(r_1, r_2, t) = \Phi_{nml}(r_1, r_2)e^{iE_{nml}t}$,

$$\Psi_{001}(r_1, r_2, t) = -\frac{(\|r_1 - r_2\| + 5)\,(x_1 - x_2 + i(y_1 - y_2))^2 e^{-\frac{1}{10}\left(\|r_1\|^2 + \|r_2\|^2\right)}}{25\pi\sqrt{440 + 75\sqrt{10\pi}}}e^{i\frac{4}{3}t}. \tag{A.37}$$

This analytic solution for two electrons interacting via the Coulomb force and confined in a harmonic oscillator can be used as a test case for multi-particle problems. A more detailed study on these exact solutions and applications to non-V-representability can be found in [Tau93, TE10, TME09].

List of Figures

List of Tables

List of Algorithms

Bibliography

[AC09]   A. Anantharaman and E. Cancès, *Existence of minimizers for Kohn–Sham models in quantum chemistry*, Annales de l'Institut Henri Poincare (C) Non Linear Analysis **26** (2009), no. 6, 2425–2455.

[ADPM11]   L. Ambrosio, G. Da Prato, and A. Mennucci, *Introduction to Measure Theory and Integration*, Ed. della Normale, Pisa, 2011.

[AF03]   R. A. Adams and J. J.F. Fournier, *Sobolev spaces*, Pure and Applied Mathematics, vol. 140, Academic Press, Amsterdam, 2003.

[AH11]   B. Andrews and C. Hopper, *The Ricci Flow in Riemannian Geometry*, Lecture Notes in Mathematics, Springer Berlin Heidelberg, 2011.

[AMGGB02]   C. Attaccalite, S. Moroni, P. Gori-Giorgi, and G. B. Bachelet, *Correlation energy and spin polarization in the 2d electron gas*, Phys. Rev. Lett. **88** (2002), 256601.

[Aub63]   J.-P. Aubin, *Un théorème de compacité*, C.R. Acad. Sci. Paris **256** (1963), 5042–5044.

[BBD02]   C. Besse, B. Bidégaray, and S. Descombes, *Order estimates in time of splitting methods for the nonlinear Schrödinger equation*, SIAM Journal on Numerical Analysis **40** (2002), no. 1, 26–40.

[BCS17]   A. Borzì, G. Ciaramella, and M. Sprengel, *Formulation and Numerical Solution of Quantum Control Problems*, Society for Industrial and Applied Mathematics, Philadelphia, 2017, to appear.

[BG03]   T. Brixner and G. Gerber, *Quantum control of gas-phase and liquid-phase femtochemistry*, ChemPhysChem **4** (2003), 418–438.

[BH08]   A. Borzì and U. Hohenester, *Multigrid optimization schemes for solving bose–einstein condensate control problems*, SIAM Journal on Scientific Computing **30** (2008), no. 1, 441–462.

[BJM02]   W. Bao, S. Jin, and P. A. Markowich, *On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime*, Journal of Computational Physics **175** (2002), no. 2, 487–524.

[BJWM00]   M. H. Beck, A. Jäckle, G. A. Worth, and H.-D. Meyer, *The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets*, Physics Reports **324** (2000), no. 1, 1–105.

[BMC15]   J. A. Budagosky Marcilla and A. Castro, *Ultrafast single electron spin manipulation in 2d semiconductor quantum dots with optimally controlled time-dependent electric fields through spin-orbit coupling*, The European Physical Journal B **88** (2015), no. 1, 15.

Bibliography

[Boi]   S. Boixo, *Lecture notes TDDFT*, http://www.isi.edu/sites/default/files/users/serboixo/lecture11.pdf, last checked March 28, 2017.

[Bor12]   A. Borzì, *Quantum optimal control using the adjoint method*, Nanoscale Systems: Mathematical Modeling, Theory and Applications **1** (2012), 93–111.

[BS12]   A. Borzì and V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*, Society for Industrial and Applied Mathematics, Philadelphia, 2012.

[BSH02]   A. Borzì, G. Stadler, and U. Hohenester, *Optimal quantum control in nanostructures: Theory and application to a generic three-level system*, Phys. Rev. A **66** (2002), 053811.

[BSV08]   A. Borzì, J. Salomon, and S. Volkwein, *Formulation and numerical solution of finite-level quantum optimal control problems*, J. Comp. App. Math. **216** (2008), 170–197.

[CAO⁺06]   A. Castro, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, M. A. L. Marques, E. K. U. Gross, and A. Rubio, *octopus: a tool for the application of time-dependent density functional theory*, physica status solidi (b) **243** (2006), no. 11, 2465–2488.

[Cas13]   A. Castro, *Theoretical shaping of femtosecond laser pulses for ultrafast molecular photo-dissociation with control techniques based on time-dependent density functional theory*, ChemPhysChem **14** (2013), 1488–1495.

[Caz03]   T. Cazenave, *Semilinear Schrödinger Equations*, Courant Lecutre Notes, vol. 10, American Mathematical Society, 2003.

[CB16]   G. Ciaramella and A. Borzì, *A LONE code for the sparse control of quantum systems*, Computer Physics Communications **200** (2016), 312–323.

[CBDW15]   G. Ciaramella, A. Borzì, G. Dirr, and D. Wachsmuth, *Newton methods for the optimal control of closed quantum spin systems*, SIAM Journal on Scientific Computing **37** (2015), no. 1, A319–A346.

[Cia13]   P. G. Ciarlet, *Linear and nonlinear functional analysis with applications*, Society for Industrial and Applied Mathematics, Philadelphia, 2013.

[CL99]   E. Cancès and C. Le Bris, *On the time-dependent Hartree-Fock equations coupled with a classical nuclear dynamics*, Mathematical Models & Methods in Applied Sciences **9** (1999), no. 7, 963–990.

[CL10]   D. Conte and C. Lubich, *An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics*, ESAIM: Mathematical Modelling and Numerical Analysis **44** (2010), no. 4, 759–780.

[CM16]   E. Cancès and N. Mourad, *Existence of optimal norm–conserving pseudopotentials for Kohn–Sham models*, Comm. Math. Sci. **14** (2016), 1315–1352.

[Con08]   L. A. Constantin, *Dimensional crossover of the exchange-correlation energy at the semilocal level*, Phys. Rev. B **78** (2008), 155106.

[CRR09]   A. Castro, E. Räsänen, and C. A. Rozzi, *Exact Coulomb cutoff technique for supercell calculations in two dimensions*, Phys. Rev. B **80** (2009), 033102.

[CWG12]   A. Castro, J. Werschnik, and E. K. U. Gross, *Controlling the dynamics of many-electron systems from first principles: A combination of optimal control and time-dependent density-functional theory*, Phys. Rev. Lett. **109** (2012), 153603.

[Dob06]   M. Dobrowolski, *Angewandte Funktionalanalysis*, Springer, Berlin Heidelberg, 2006.

[DP10]   D. Dong and I. R. Petersen, *Quantum control theory and applications: a survey*, IET Control Theory & Applications **4** (2010), 2651–2671.

[ED11]   E. Engel and R. M. Dreizler, *Density Functional Theory, An Advanced Course*, Springer Heidelberg, 2011.

[Eva10]   L. C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010.

[FFT]   `http://www.fftw.org/parallel/xolas.html`, last checked May 21, 2017.

[FLLS16]   S. Fournais, J. Lampart, M. Lewin, and T. Ø. Sørensen, *Coulomb potentials and Taylor expansions in time-dependent density-functional theory*, Phys. Rev. A **93** (2016), 062510.

[FNRM16]   J. I. Fuks, S. E. B. Nielsen, M. Ruggenthaler, and N. T. Maitra, *Time-dependent density functional theory beyond Kohn-Sham Slater determinants*, Phys. Chem. Chem. Phys. **18** (2016), 20976–20985.

[FOS15]   E. Faou, A. Ostermann, and K. Schratz, *Analysis of exponential splitting methods for inhomogeneous parabolic equations*, IMA Journal of Numerical Analysis **35** (2015), no. 1, 161–178.

[FRAR15]   J. Flick, M. Ruggenthaler, H. Appel, and A. Rubio, *Kohn–Sham approach to quantum electrodynamical density-functional theory: Exact time-dependent effective potentials in real space*, Proceedings of the National Academy of Sciences of the United States of America **112** (2015), 15285–15290.

[FW16]   B. Feng and K. Wang, *Optimal bilinear control of nonlinear Hartree equations with singular potentials*, J. Optimization Theory and Applications **170** (2016), no. 3, 756–771.

[GBC$^+$15]   J. S. Glaser, U. Boscain, T. Calarco, P. C. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny, and K. F. Wilhelm, *Training schrödinger's cat: quantum optimal control*, The European Physical Journal D **69** (2015), no. 12, 1–24.

[GN92]   J. C. Gilbert and J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optim. **2** (1992), no. 1, 21–42. MR 1147881

Bibliography

[GU09]	V. Gol′dshtein and A. Ukhlov, *Weighted Sobolev spaces and embedding theorems*, Trans. Amer. Math. Soc. **361** (2009), no. 7, 3829–3850.
[HK64]	P. Hohenberg and W. Kohn, *Inhomogeneous electron gas*, Phys. Rev. **136** (1964), B864–B871.
[HMMS13]	M. Hintermüller, D. Marahrens, P. A. Markowich, and C. Sparber, *Optimal bilinear control of Gross-Pitaevskii equations*, SIAM Journal on Control and Optimization **51** (2013), no. 3, 2509–2543.
[HRBS07]	U. Hohenester, P.-K. Rekdal, A. Borzì, and J. Schmiedmayer, *Optimal quantum control of Bose-Einstein condensates in magnetic microtraps*, Phys. Rev. A **75** (2007), 023602.
[HZ05]	W. Hager and H. Zhang, *A new conjugate gradient method with guaranteed descent and an efficient line search*, SIAM Journal on Optimization **16** (2005), no. 1, 170–192.
[Jer15]	J. W. Jerome, *Time dependent closed quantum systems: nonlinear Kohn-Sham potential operators and weak solutions*, Journal of Mathematical Analysis and Applications **429** (2015), no. 2, 995–1006.
[JP14]	J. W. Jerome and E. Polizzi, *Discretization of time-dependent quantum systems: real-time propagation of the evolution operator*, Applicable Analysis **93** (2014), no. 12, 2574–2597.
[KCM11]	D. Kammerlander, A. Castro, and M. A. L. Marques, *Optimal control of the electronic current density: Application to one- and two-dimensional one-electron systems*, Phys. Rev. A **83** (2011), 043413.
[Koc06]	O. Koch, *Efficient computation of the MCTDHF approximation to the time-dependent Schrödinger equation*, Opuscula Mathematica **26** (2006), 473–487.
[Kre09]	K. Kreutz–Delgado, *The complex gradient operator and the $\mathbb{CR}$-calculus*, arXiv:0906.4835 (2009).
[KRG⁺89]	R. Kosloff, S. A. Rice, P. Gaspard, S. Tersigni, and D. J. Tannor, *Wavepacket dancing: Achieving chemical selectivity by shaping light pulses*, Chemical Physics **139** (1989), no. 1, 201–220.
[KS65]	W. Kohn and L. J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev. **140** (1965), A1133–A1138.
[Kup96]	F. S. Kupfer, *An infinite-dimensional convergence theory for reduced SQP methods in Hilbert space*, SIAM J. Optim. **6** (1996), 126–163.
[Lie83]	E. H. Lieb, *Density functionals for coulomb systems*, International Journal of Quantum Chemistry **24** (1983), no. 3, 243–277.
[Lie85]	_____, *Density Functionals for Coulomb Systems*, Density Functional Methods in Physics (Reiner M. Dreizler and Joao da Providencia, eds.), Nato ASI Series, 1985, p. 31–80.
[Lio69]	J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod; Gauthier-Villars, Paris, 1969.

[Lio71]   J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.

[Lub08]   C. Lubich, *On splitting methods for Schrödinger-Poisson and cubic non-linear Schrödinger equations*, Math. Comp. **77** (2008), 2141–2153.

[MB01]   N. T. Maitra and K. Burke, *Demonstration of initial-state dependence in time-dependent density-functional theory*, Phys. Rev. A **63** (2001), 042501.

[MF53]   P. M. Morse and H. Feshbach, *Methods of Theoretical Physics, Part I*, McGraw-Hill, New York, 1953.

[MI99]   S. McLoone and G. Irwin, *A variable memory quasi-newton training algorithm*, Neural Processing Letters **9** (1999), 77–89.

[MMN⁺12]   M. A. L. Marques, N. T. Maitra, F. M. S. Nogueira, E. K. U. Gross, and A. Rubio (eds.), *Fundamentals of Time-Dependent Density Functional Theory*, Lecture Notes in Physics, vol. 837, Springer Berlin Heidelberg, 2012.

[MOB12]   M. A. L. Marques, M. J. T. Oliveira, and T. Burnus, *Libxc: A library of exchange and correlation functionals for density functional theory* , Computer Physics Communications **183** (2012), no. 10, 2272–2281.

[MST06]   Y. Maday, J. Salomon, and G. Turinici, *Monotonic time-discretized schemes in quantum control*, Numer. Math. **103** (2006), no. 2, 323–338.

[MT03]   Y. Maday and G. Turinici, *New formulations of monotonically convergent quantum control algorithms*, J. Chem. Phys. **118** (2003), 8191–8196.

[MUN⁺06]   M. A. L. Marques, C. A. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. K. U. Gross, *Time-Dependent Density Functional Theory*, Lecture Notes in Physics, vol. 706, Springer, Berlin Heidelberg, 2006.

[NW06]   J. Nocedal and S. J. Wright, *Numerical Optimization*, 2 ed., Springer, New York, 2006.

[Pas80]   J. E. Pasciak, *Spectral and pseudospectral methods for advection equations*, Math. Comp. **35** (1980), 1081–1092.

[PDR88]   A. P. Peirce, M. A. Dahleh, and H. Rabitz, *Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications*, Phys. Rev. A **37** (1988), 4950–4964.

[Pen16]   M. Penz, *The Density-Potential Mapping in Quantum Dynamics*, Ph.D. thesis, Universität Insbruck, 2016.

[Pit61]   L. P. Pitaevsk, *Vortex lines in an imperfect bose gas*, Soviet Physics JETP-USSR **13** (1961), no. 2.

[PR11]   M. Penz and M. Ruggenthaler, *Domains of time-dependent density-potential mappings*, Journal of Physics A: Mathematical and Theoretical **44** (2011), no. 33, 335208.

Bibliography

[PY89]   R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, 1989.

[RCW⁺08a]   E. Räsänen, A. Castro, J. Werschnik, A. Rubio, and E. K. U. Gross, *Optimal laser control of double quantum dots*, Phys. Rev. B **77** (2008), 085324.

[RCW⁺08b]   E. Räsänen, A. Castro, J. Werschnik, A. Rubio, and E. K. U. Gross, *Coherent quantum switch driven by optimized laser pulses*, Physica E: Low-dimensional Systems and Nanostructures **40** (2008), no. 5, 1593–1595.

[Rem91]   E. Remmert, *Theory of Complex Functions*, Springer Heidelberg, 1991.

[RG84]   E. Runge and E. K. U. Gross, *Density-functional theory for time-dependent systems*, Phys. Rev. Lett. **52** (1984), 997–1000.

[RNvL13]   M. Ruggenthaler, S. E. B. Nielsen, and R. van Leeuwen, *Analytic density functionals with initial-state dependence and memory*, Phys. Rev. A **88** (2013), 022512.

[RPvL15]   M. Ruggenthaler, M. Penz, and R. van Leeuwen, *Existence, uniqueness, and construction of the density-potential mapping in time-dependent density-functional theory*, Journal of Physics: Condensed Matter **27** (2015), no. 20, 203202.

[RS78]   M. Reed and B. Simon, *Methods of modern mathematical physics*, vol. 4, Academic Press New York, 1978.

[Sac86]   E. W. Sachs, *Broyden's method in Hilbert space*, Math. Programming **35** (1986), no. 1, 71–82.

[Sal05]   J. Salomon, *Limit points of the monotonic schemes in quantum control*, Proceedings of the 44th IEEE Conference on Decision and Control, Sevilla (2005), 5854–5858.

[SCB17a]   M. Sprengel, G. Ciaramella, and A. Borzì, *A COKOSNUT code for the control of the time-dependent Kohn–Sham model*, Computer Physics Communications **214** (2017), 231–238.

[SCB17b]   _____, *Investigation of optimal control problems governed by a time-dependent Kohn-Sham model*, arXiv:1701.02679 (2017).

[SCB17c]   _____, *A theoretical investigation of time-dependent Kohn–Sham equations*, SIAM Journal on Mathematical Analysis **49** (2017), no. 3, 1681–1704.

[Sch26]   E. Schrödinger, *Der stetige Übergang von der Mikro- zur Makromechanik*, Naturwissenschaften **14** (1926), no. 28, 664–666.

[Sch05]   R. L. Schilling, *Measures, Integrals and Martingales*, Cambridge University Press, New York, 2005.

[Sch07]   F. Schwabl, *Quantenmechanik*, Springer, 2007.

[SGGS07]   M. Seidl, P. Gori-Giorgi, and A. Savin, *Strictly correlated electrons in density-functional theory: A general formulation with applications to spherical densities*, Phys. Rev. A **75** (2007), 042511.

[Sha78]   D. F. Shanno, *Conjugate gradient methods with inexact searches*, Math. Oper. Res. **3** (1978), 244–256.

[Sto32]   M. H. Stone, *On one-parameter unitary groups in Hilbert space*, Annals of Mathematics. Second Series **33** (1932), no. 3, 643–648.

[Tau93]   M. Taut, *Two electrons in an external oscillator potential: Particular analytic solutions of a Coulomb correlation problem*, Phys. Rev. A **48** (1993), 3561–3566.

[TCN09]   M. Thalhammer, M. Caliari, and C. Neuhauser, *High-order time-splitting Hermite and Fourier spectral methods*, Journal of Computational Physics **228** (2009), no. 3, 822–832.

[TE10]   M. Taut and H. Eschrig, *Exact solutions for a two-electron quantum dot model in a magnetic field and application to more complex systems*, Zeitschrift für Physikalische Chemie **224** (2010), 631.

[Tha12]   M. Thalhammer, *Convergence analysis of high-order time-splitting pseudospectral methods for nonlinear Schrödinger equations*, SIAM Journal on Numerical Analysis **50** (2012), no. 6, 3231–3258.

[TME09]   M. Taut, P. Machon, and H. Eschrig, *Violation of noninteracting V-representability of the exact solutions of the Schrödinger equation for a two-electron quantum dot in a homogeneous magnetic field*, Phys. Rev. A **80** (2009), 022517.

[Trö10]   F. Tröltzsch, *Optimal Control of Partial Differential Equations*, 1 ed., American Mathematical Society, Providence, Rhode Island, 2010.

[TV16]   F. Tröltzsch and A. Valli, *Optimal control of low-frequency electromagnetic fields in multiply connected conductors*, Optimization **65** (2016), no. 9, 1651–1673.

[Vig04]   G. Vignale, *Mapping from current densities to vector potentials in time-dependent current density functional theory*, Phys. Rev. B **70** (2004), 201102.

[vL99]   R. van Leeuwen, *Mapping from densities to potentials in time-dependent density-functional theory*, Phys. Rev. Lett. **82** (1999), 3863–3866.

[vWB08]   G. von Winckel and A. Borzì, *Computational techniques for a quantum control problem with H^1-cost*, Inverse Problems **24** (2008), no. 3, 034007, 23.

[vWB10]   G. von Winckel and A. Borzì, *QUCON: A fast Krylov–Newton code for dipole quantum control problems*, Computer Physics Communications **181** (2010), no. 12, 2158–2163.

[vWBV10]   G. von Winckel, A. Borzì, and S. Volkwein, *A globalized Newton method for the accurate solution of a dipole quantum control problem*, SIAM Journal on Scientific Computing **31** (2010), no. 6, 4176–4203.

[Wal98]   W. Walter, *Ordinary Differential Equations*, Graduate Texts in Mathematics, vol. 182, Springer, New York, 1998, Translated from the sixth German (1996) edition by Russell Thompson, Readings in Mathematics.

Bibliography

[Wal09]   R. M. Wald, *General Relativity*, Univ. of Chicago Press, Chicago, 2009.

[Wer06]   J. Werschnik, *Quantum Optimal Control Theory: Filter Techniques, Time-Dependent Targets, and Time-Dependent Density-Functional Theory*, Ph.D. thesis, Freie Universität Berlin, 2006.

[WLAWLL15] H.-Ch. Weissker, O. Lopez-Acevedo, R. L. Whetten, and X. López-Lozano, *Optical spectra of the special Au_{144} gold-cluster compounds: Sensitivity to structure and symmetry*, The Journal of Physical Chemistry C **119** (2015), no. 20, 11250–11259.

[Yse10]   H. Yserentant, *Regularity and Approximability of Electronic Wave Functions*, Lecture Notes in Mathematics, vol. 2000, Springer, Berlin, 2010.

[Zei90]   E. Zeidler, *Nonlinear Functional Analysis and its Applications II/A*, Springer, New York, 1990.

[ZK79]   J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Applied Mathematics and Optimization **5** (1979), no. 1, 49–62.