Exact and non-smooth control
of
quantum spin systems

Vento del Sud, portami via con te,
ho bisogno di riflettere,
tutto va troppo in fretta per me,
ho sempre l'impressione
di aver lasciato troppe cose indietro...

Vento del Sud, portami via con te,
per dimenticare quanto
l'abitudine è una giostra
da cui non ti conviene scendere
e anche per questo siamo fragili...

                                    Ottoohm

Acknowledgements

Tomorrow there's my PhD exam, and this evening I cannot really sleep. The last three years were very hard for me. I discovered a side of me that I didn't know before and to accept and admit it is probably the hardest part. Luckily many people helped and supported me. Without them I could not have succeed...thank you all!

Before any other person, I want to thank Mariagiovanna, the only person that really makes me feel free and happy and that cancels any fear from my soul. I love you.

I thank my Family, my Grandmother, my Aunt and the "Pallottas", that continuously support me. When I think to us all together outside behind our home and in the "sun of the south" during the "our old Sundays" I feel better and I would like to run back to you.

A special thank is for my "Boss", Prof. Alfio Borzì. Thank you for the chance that you gave me to study in Würzburg and discover new exciting topics. We both know that, even if you tried to brake my "uncontrolled side", I didn't change it (I won't change it!) and I will continue to explicitly hate "il vecchiaccio" and all the unfair people that behave like him. Grazie Boss!!

A huge thank is for Prof. Julien Salomon. You teach me how to think differently and outside any scheme. I will never forget the two days together in Stuttgart!

I want also to thank Prof. Gunther Dirr. Your patience, your calmness and your continuous happiness helped me a lot in the last three years. Thank you!

I would like to thank also Prof. Daniel Wachsmuth. Your lectures about infinite-dimensional optimization are the best lectures that I have studied. Working with you I learned a lot of mathematics and something more.

An enormous thank is for the WiReX group. First of all, I have to thank Mrs. Petra Markert-Autsch. With your continuous support you make everything possible in our group. An infinite-dimensional thank is for Suttida. I don't know how, but you managed to keep me calm in many bad moments. Because of you the atmosphere in our office was always friendly and full of happiness. Please never forget about "the duck" and keep always the music playing in our office!! Thank you!! A "quantum-thank" is for Martin. It was a great pleasure to enjoy with you our working travels, our hiking tours and the last days working together. I found in you a true friend and I hope that this friendship will continue in the future!! A particular thank is for Juri. Probably you do not remember, but your nodding during my talk in Munich was really important for me in that period! Thank you! A huge thank is for Mario, Tanvir, Masoumeh, Andreas, Beatrice, Duncan, Stephan and Christian for the nice atmosphere of our group. Thank you all!!

The last thank (and probably the most important) is for "the gladiator" Pablo, "the Roman commander" Alejandro, "il maestro" Giovanni and Gero (for which we still need a Roman's nickname). I don't have words to describe your friendship! Without you I could not have survived in Würzburg. As "the oracle" I see a great future for you all and that our friendships will never end!!! Thank you guys!!

29. 07. 2015
Gabriele

Abstract

An efficient and accurate computational framework for solving control problems governed by quantum spin systems is presented. Spin systems are extremely important in modern quantum technologies such as nuclear magnetic resonance spectroscopy, quantum imaging and quantum computing. In these applications, two classes of quantum control problems arise: optimal control problems and exact-controllability problems, with a bilinear control structure. These models correspond to the Schrödinger-Pauli equation, describing the time evolution of a spinor, and the Liouville-von Neumann master equation, describing the time evolution of a spinor and a density operator. This thesis focuses on quantum control problems governed by these models. An appropriate definition of the optimization objectives and of the admissible set of control functions allows to construct controls with specific properties. These properties are in general required by the physics and the technologies involved in quantum control applications.

A main purpose of this work is to address non-differentiable quantum control problems. For this reason, a computational framework is developed to address optimal-control problems, with possibly L^1-penalization term in the cost-functional, and exact-controllability problems. In both cases the set of admissible control functions is a subset of a Hilbert space. The bilinear control structure of the quantum model, the L^1-penalization term and the control constraints generate high non-linearities that make difficult to solve and analyse the corresponding control problems.

The first part of this thesis focuses on the physical description of the spin of particles and of the magnetic resonance phenomenon. Afterwards, the controlled Schrödinger-Pauli equation and the Liouville-von Neumann master equation are discussed. These equations, like many other controlled quantum models, can be represented by dynamical systems with a bilinear control structure.

In the second part of this thesis, theoretical investigations of optimal control problems, with a possible L^1-penalization term in the objective and control constraints, are considered. In particular, existence of solutions, optimality conditions, and regularity properties of the optimal controls are discussed. In order to solve these optimal control problems, semi-smooth Newton methods are developed and proved to be superlinear convergent. The main difficulty in the implementation of a Newton method for optimal control problems comes from the dimension of the Jacobian operator. In a discrete form, the Jacobian is a very large matrix, and this fact makes its construction infeasible from a practical point of view. For this reason, the focus of this work is on inexact Krylov-Newton methods, that combine the Newton method with Krylov iterative solvers for linear systems, and allows to avoid the construction of the discrete Jacobian.

In the third part of this thesis, two methodologies for the exact-controllability of quantum spin systems are presented. The first method consists of a continuation technique, while the second method is based on a particular reformulation of the exact-control problem. Both these methodologies address minimum L^2-norm exact-controllability problems.

In the fourth part, the thesis focuses on the numerical analysis of quantum control problems. In particular, the modified Crank-Nicolson scheme as an adequate time discretization of the Schrödinger equation is discussed, the first-discretize-then-optimize strategy is used to obtain a discrete reduced gradient formula for the differentiable part of the optimization objective, and implementation details and globalization strategies to guarantee an adequate numerical behaviour of semi-smooth Newton methods are treated.

In the last part of this work, several numerical experiments are performed to validate the theoretical results and demonstrate the ability of the proposed computational framework to solve quantum spin control problems.

Contents

Chapter 1

Introduction

The problem of steering a quantum dynamical system from an initial state to a target state is one of the main issues in many applications in nanotechnology. This class of problems appears in several disciplines, such as nuclear magnetic resonance (NMR) spectroscopy, magnetic resonance imaging (MRI), spin dynamics and quantum information processing, where quantum systems are controlled by the action of magnetic fields. These disciplines, motivated by medical and engineering applications, focus on the control of the spin orientation of particles and the magnetization of given nuclei; see, e.g., [9, 14, 20, 25, 46, 72, 73, 85, 125]. The issue of controlling the state of a quantum system appears also in applications involving chemical reactions; see, e.g., [14, 16, 17, 116, 128, 129]. In these applications, the control function represents an electric field, with the aim of steering the quantum system between eigenstates of its Hamiltonian; see, e.g., [14, 16, 116, 128, 129]. A quantum control problem is mathematically regarded as the control of a quantum model with a bilinear control structure. This means that, the product between control functions and quantum states appears in the governing model and generates highly non-linear dynamics.

The task of controlling a quantum system can be mathematically clarified as an optimal control problem, as an exact control problem and as a controllability problem. An optimal control problem has the purpose of computing control functions that minimize an appropriate cost functional, which contains a norm of the difference between the resulting terminal state and a given target state; see, e.g., [14, 25, 31, 46, 108, 109, 111]. An exact control problem aims at obtaining control functions that allow to reach exactly the given target state. This class of problem is more involved, and the corresponding literature is at its infancy [30, 31, 129, 130]. Further, a controllability problem aims at establishing the reachability of a given target state; see, e.g., [1, 2, 3, 4, 5, 39, 40, 102]. Another non-negligible aspect is related to the structure and the properties of the control functions. In fact, many applications motivated by physical and experimental reasons, require control functions with specific properties. These classes of quantum control problems pose several challenges from the scientific computing point of view: design and analyse quantum optimal control problems; develop numerical algorithms capable to solve quantum optimization problems in a fast and accurate way; prove that a quantum system is controllable and establish the conditions which are necessary and sufficient to guarantee controllability. In the present thesis, we contribute to all of these issues.

The main focus of this work is the control of quantum spin systems considered in NMR spectroscopy, MRI and other applications that involve the control of spin and magnetization state of particles and nuclei by means of magnetic fields. Despite this applicative purpose, most of the methodologies and the results presented in this thesis

are valid also for a more general class of quantum bilinear control problems.

In the field of control of spin systems, it is suitable to have criteria that allow to establish controllability of a given system. For this reason, controllability of quantum spin systems is an important topic in applied mathematics. The time evolution of spin systems can be regarded as the evolution of systems defined on Lie groups. General controllability results for bilinear systems evolving on Lie groups are given in [67, 68]. For controllability results regarding specifically spin systems, we refer to [2, 3, 4, 5, 39, 40, 42, 102]. Another important aspect is to guarantee controllability at a given time. The literature regarding these aspect is scarce, and pioneering works are given in, e.g., [1, 9, 39, 70, 72].

From the application point of view, the most important issue of quantum control problems is to determine control functions for specific transitions. This reason is boosting an intensive investigation in the optimal control theory for quantum systems, which has the purpose of designing fast control mechanisms that cannot be constructed simply based on perturbation theory [129]. This fact motivates the application of optimal control methods and related computational techniques to quantum systems [15, 59, 73, 88]; see [14] for a review. However, in most of the works on quantum optimal control problems, the focus is on numerical optimization techniques that allow to compute the required control functions. Pioneering results in the development of quantum optimal control algorithms can be found in [76, 115, 133]. Further progress in the development of efficient control schemes is documented in, e.g., [44, 59, 73, 88, 89, 91, 106]. In particular, advanced optimization methods are discussed in [15, 17, 18, 48, 62, 87, 128]. Investigation and analyses of discretization schemes for quantum control problems can be found in, e.g., [88, 129]. In these references, only first-order optimization methods were considered, while, to the best of our knowledge, we are not aware of works on second-order Newton methods for the control of quantum spin systems. On the other hand, a discussion of a Krylov-Newton (or Newton-Krylov) scheme for solving optimal spin-less quantum control problems governed by an infinite-dimensional Schrödinger equation can be found in [129, 130]. Moreover, in applications the need arises to constrain the control functions pointwise in time; see, e.g., [75, 105]. This requirement results in a lack of differentiability of the control problem so that a straightforward application of the Newton method is not possible. To overcome this limitation, in the field of PDE-constraint optimization problems, a semi-smooth Newton (SSN) method was developed to solve control-constrained problems; see, e.g., [56, 58, 78, 79, 80, 123]. The only works in the literature investigating semi-smooth Newton methods for the control of bilinear quantum systems are [29, 31, 32], that represent one of the main novelties of this research work.

In several applications, such as NMR and MRI, controls of practical interest are pointwise-bounded piecewise-constant time-dependent functions that can be exactly implemented in experiments [25, 37, 125]. Furthermore, in many applications in NMR spectroscopy, it is suitable to have control functions with a so-called sparse structure; see, e.g., [74, 85]. To obtain functions with this special structure, an L^1-penalization of the controls is considered in the cost functional. This penalization term increases the non-linearity of the problem, and makes the theoretical analysis and the numerical solution more involved. In the field of L^1-optimal control of linear PDEs, pioneering works are [7, 24, 35, 55, 110, 126, 127]. On the other hand, L^1-optimal control of quantum systems is a much less investigated subject, and we are aware only about the work [29], which represents one of the main novelties of this thesis.

In some spin control applications, the control obtained with the optimal control formulation does not satisfy the requirements of some physical experiments. Rather, it is essential to compute exact-control functions, that are controls capable to steer exactly

the trajectory of a quantum system to a given target state. However, research in the field of quantum exact-controllability problems focuses in large part on theoretical issues and much less is known concerning algorithms that allow to compute these functions and that are supported by a rigorous theoretical framework. In this area, the only works addressing computational schemes for exact-control problems with a rigorous theoretical background are [28, 30, 31], which are part of this thesis.

The main focus of our work is to investigate quantum control problems where the governing model has the following structure

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \tag{1.1}$$

where x is the state describing a quantum system, u_n are the control functions, and the operators A and B_n are in general skew-symmetric matrices. This general bilinear system can be a real representation of a Liouville-von Neumann equation, a (semi-discrete) Schrödinger equation, a Pauli equation and other quantum operator equations; see, e.g., [6, 25, 31, 40, 109].

One of the most challenging optimal control problems considered in this thesis is the following

$$
\begin{aligned}
\min_{x,u} \quad & J(x,u) := \frac{1}{2}\|x(T) - x_T\|_2^2 + \sum_{n=1}^{N_C} \left[\frac{\nu}{2}\|u_n\|_{L^2}^2 + \beta\|u_n\|_{L^1} \right] \\
\text{s.t.} \quad & \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \ \text{in} \ (0,T] \\
& x \in X \ , \ u \in U_{ad} \ ,
\end{aligned}
\tag{1.2}
$$

where X is the state space, x_T is a given target, U_{ad} is the set of admissible controls, and ν and β are non-negative parameters used to weight the L^2- and the L^1-norm of the controls. The term $\frac{1}{2}\|x(T) - x_T\|_2^2$ measures the distance between the terminal point of the trajectory $x(T)$ and the given target x_T. This term is referred to as tracking term.

The solutions to (1.2) can have different properties, depending on the values of ν and β, and on the structure of the admissible set U_{ad}. In particular, the coefficient ν weights the squared L^2-norm of the controls, that represents a regularization term; on the other hand, it penalizes high values of the controls. The coefficient β weights the L^1-norm of the controls, that, in contrast, penalizes small values of the controls and generates solutions that are said to be sparse; see, e.g., [29, 110, 126]. It is necessary to remark that, even if the L^1-norm increases the convexity of the problem, it cannot be regarded as a regularization term. This is due to the fact that it makes the problem non-smooth. Another source of non-smoothness and non-linearity is the control constraint set U_{ad}. This set is in general a closed and convex subset of a Hilbert space, and is used to guarantee particular features of the desired controls; for instance, in NMR spectroscopy, it is often required that the control functions are pointwise bounded. Both the L^1-penalization and the constraint U_{ad} increase the non-linearity of the optimization problem, which is mainly due to the bilinear structure of the quantum system (1.1). For these reasons, a theoretical characterization and, in particular, a numerical solution to (1.2) could be hard to perform. This is one of the main novelty of the thesis: we develop an efficient and robust SSN method, with proved theoretical convergence, for the solution to (1.2), and, in order to obtain a very fast and efficient scheme, we focus on a matrix-free implementation of the SSN method.

Furthermore, we provide a theoretical characterization of the optimal control solutions; see [29, 31, 32].

The second novelty of the present thesis is the development of two computational schemes, that are capable to address exact-controllability problems. In particular, both the methods aim to solve the following minimum L^2-norm problem [28, 29, 31]

$$
\begin{aligned}
\min_{x,u} \quad & J_u(x,u) := \frac{1}{2} \sum_{n=1}^{N_C} \|u_n\|_{L^2}^2 \\
\text{s.t.} \quad & \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \;,\; \text{in } (0,T) \;,\; x(0) = x_0 \;,\; x(T) = x_T \;.
\end{aligned}
\tag{1.3}
$$

The first method is based on a continuation technique that considers a sequence of positive weight parameters ν_k and the corresponding optimal control problems (1.2) (with $\beta = 0$). The corresponding sequence of optimal controls u^k is proved to converge to a solution to (1.3). This problem is addressed in [31], where a continuation method is developed for unconstrained exact-control problems. In this thesis, we extend these results in the case that the controls have to belong to a closed and convex set U_{ad}. The second computational method consists in a reformulation of the optimality system, that characterizes a solution to (1.3), in a new optimization problem. One of the unknowns of this new problem is the terminal condition of the adjoint equation, which is consequently referred as to a shooting-variable. The main feature of this reformulation, is that the new problem is proved to have, under reasonable conditions, regularity properties, that allow the use of standard optimization techniques. This framework is developed in [30].

Another novelty of this work is the analysis of the modified Crank-Nicolson (MCN) scheme as an appropriate discretization of (1.1), especially in the case that (1.1) is a Schrödinger equation. In fact, the time evolution of a Schrödinger equation is characterized by algebraic and geometrical properties, like norm-preservation and symplecticity, that are important in the optimization process. In the field of quantum control, the majority of the literature regarding the time-discretization of Schrödinger equations and quantum systems, focuses on schemes based on the use of exponential matrices; see, e.g., [37, 61, 73]. This allows to preserve the properties of the Schrödinger equation, but pays the price of a high computational effort. The MCN scheme provides an answer to this problem, since it combines the well-known computational efficiency of finite-differences with the preservation of all the properties of a Schrödinger equation. We remark that, this scheme was presented in [129] and derived by means of a Magnus expansion of the Hamiltonian following similar arguments as in [61]. However, there is no rigorous analysis regarding convergence and preservation of algebraic and geometrical properties. This analysis is presented in this thesis, and in particular, the analysis of algebraic and geometrical properties is performed with the help of the matrix Lie group theory.

Furthermore, since many scientific topics are involved in quantum control theory, the different scientific communities formulate quantum control problems with different formalisms that make a convenient transfer of knowledge difficult. For this reason, we provide a unifying formulation of quantum spin control problems. In particular, within this formulation, we discuss also existing theoretical results to provide a complete and consistent picture of quantum spin control problems.

For the sake of clarity, we summarize the main contributions achieved in this thesis:

1. The development of a semi-smooth Krylov-Newton scheme for the solution to the optimal control problem (1.2). This development consists of a theoretical investi-

gation, that provides the method with a theoretical proof of convergence, and in an implementation analysis, that discusses several implementation details for an adequate numerical behaviour.

2. The development of two computational methods for the exact control of (1.1). These schemes address the minimum L^2-norm problem (1.3).

3. The numerical, geometrical, and algebraic analysis of the modified Crank-Nicolson scheme for the numerical time-integration of a general Schrödinger equation.

4. A unifying formulation of quantum spin control problems. This formulation starts with a physical description of the main quantum spin models, and involves theoretical and numerical analysis, including results regarding controllability of spin systems. Some known results are revisited in this formulation.

The thesis is organized as follows. Chapter 2 provides a detailed introduction regarding the quantum systems addressed in this work. Starting with a physical description of the spin of particles and of the magnetic resonance phenomenon in Section 2.1, the Schrödinger-Pauli equation is discussed for spin systems controlled by means of magnetic fields in Section 2.2. A particular spectral discretization of the Pauli equation allows to obtain a finite-dimensional Schrödinger equation, that is used to derive the Liouville-von Neumann master (LvNM) equation. The LvNM equation is described in details in Section 2.3. In Section 2.4, it is shown that the Schrödinger-Pauli and the LvNM equations can be represented in the bilinear form (1.1).

Chapter 3 focuses on a theoretical investigation of problem (1.2). In particular, in Section 3.1 properties of the bilinear system (1.1) and its linearization are investigated. In Section 3.2, problem (1.2) with $\beta = 0$, $\nu \geq 0$ and different admissible sets U_{ad} is analysed. In particular, the existence of a minimizer is proved in Theorem 1 and first-order necessary conditions are addressed in Theorems 2, 3 and 4. In Section 3.3, problem (1.2) with $\beta = 0$ and piecewise-constant controls is studied. Section 3.4 focuses on the L^1-penalized problem (1.2) ($\beta > 0$). The existence of a minimizer and first-order necessary conditions are proved in Theorem 6 and Theorem 7, respectively. Characterizing results of the optimal controls are presented in Theorems 9, 10 and 11.

Chapter 4 aims to the development of Newton methods for the optimal control of quantum systems. In Section 4.1, an introduction of Newton and semi-smooth Newton methods is provided by discussing these methods for the solution of finite-dimensional different problems. This introduction has the purpose to explain in a simple way the main concepts of Newton and semi-smooth Newton methods, in the case of a minimization of finite-dimension functions, and simplify the reading of the following sections 4.2 and 4.3. In particular, Section 4.2 focuses on the description of a matrix-free Krylov-Newton method for the solution of (1.2) with $\beta = 0$ in the absence of control constraints. A proof of the local quadratic convergence is provided in Theorem 16. In Section 4.3, a semi-smooth Krylov-Newton method for the solution of the L^1-penalized problem (1.2) is presented, and proved to be locally superlinear convergent in Theorem 19.

In Chapter 5, the exact control of quantum systems is addressed. Section 5.1 provides a survey of some important controllability results for quantum spin systems. The two Sections 5.2 and 5.3 present two different computational methodologies, and corresponding theoretical analyses, for the exact control of quantum systems by solving problem (1.3). In particular, Section 5.2 focuses on a continuation procedure for the exact control of quantum systems, valid also in the case of constrained controls. Convergence of the

continuation method is proved in Theorems 25 and 27. Moreover, the rate of convergence is addressed in Theorems 26 and 28. Section 5.3 presents a methodology for the exact control of (1.1) based on a reformulation of (1.3). Regularity properties of this reformulation are proved in Theorems 29 and 30.

The focus of Chapter 6 is the numerical analysis of quantum control problems. Section 6.1 describes the modified Crank-Nicolson scheme as an adequate time discretization of the Schrödinger equation. This section provides a theoretical investigation of the MCN scheme regarding stability, convergence, and preservation of algebraic properties. These results are proved in Theorems 31, 33 and 32. In Section 6.2, the first-discretize-then-optimize strategy is used to obtain a discrete reduced gradient formula for the differentiable part of the cost in (1.2). Section 6.3 discusses implementation details and globalization strategies to guarantee an adequate numerical behaviour of Krylov-Newton and Krylov-SSN methods. In Sections 6.4 and 6.5, numerical schemes for performing the continuation procedure and the shooting-type method are discussed.

Chapter 7 presents numerical experiments that demonstrate the validity of the computational framework developed in this thesis. In particular, in Section 7.1 numerical experiments of L^1-penalized problems with possibly piecewise-constant controls are discussed. The experiments consider finite-dimensional spin systems and infinite-dimensional quantum systems and demonstrate the numerical convergence of the Krylov-Newton methods. Section 7.2 focuses on numerical experiments of exact-control problems. In particular, these experiments address the exact control of spin systems and the Pauli equation, and a case of a bang-bang exact-control problem. The two methodologies developed in Chapter 5 are used and compared. Section 7.3 presents the control of distributed inhomogeneous quantum spin systems.

Each chapter of this thesis is concluded with a "summary and remarks" section. In the appendix, we collect auxiliary results and additional informations, that make this thesis more self-contained.

We believe that the methodologies presented in this thesis will become the methods of choice in the applied sciences and boost further research in the field of quantum control problems.

Chapter 2

The physics of a spin and quantum control models

This chapter introduces the physics of a spin and the related mathematical models. In particular, physical behaviour of spin particles and the so-called magnetic resonance phenomenon are described. These are the main notions used in many applications in nanotechnology, like nuclear magnetic resonance (NMR) spectroscopy, magnetic resonance imaging (MRI), quantum computing, etc., where controlling magnetic fields allows to perform experiments with chemical, pharmaceutical, and medical purposes.

The main equation that models the spin dynamics is the famous Pauli equation. This is a Schrödinger equation whose Hamiltonian contains an additional term, known as Stern-Gerlach term, which expresses the interaction between the particle's spin and an external magnetic field. By means of a spectral discretization method and introducing the notion of density matrix, the Pauli equation can be transformed into the famous Liouville-von Neumann master (LvNM) equation, that is used to characterize the dynamics of nuclear magnetic resonance phenomena. The Pauli equation, the Liouville-von Neumann equation and their relationship are discussed in the present chapter. Furthermore, particular attention is posed in the description of the spin Hamiltonian, which is used in NMR and MRI applications. Both Pauli and LvNM equation admit a real representation, in which these assume the form of a dynamical system with bilinear control structure, that is usual in quantum control theory and applications. These representations are described in the last section of this chapter. Notice that the bilinear dynamical systems represent the governing equations of optimal control problems addressed in this thesis.

The chapter is organized as follows. In Section 2.1, the concepts of spin of particles and atomic nuclei is described, and its relationship with the magnetic resonance is discussed. Section 2.2 considers the Pauli equation, that is approximated by means of a spectral discretization method. Physical considerations show that for particular state transitions the solution of the obtained discrete system coincide with the exact solution of the Pauli equation. In Section 2.3, first we provide with a description of the density matrix, that is used to characterize the quantum state of spin systems. Then the LvNM equation is derived, and details regarding the control of spin systems involved in NMR are discussed. In Section 2.4, a general dynamical system with a bilinear control structure is presented as a real representation of quantum control systems.

2.1 Spin of Fermions and magnetic resonance

In quantum mechanics, there exist two types of angular momentum possessed by elementary particles [51, 83, 118]. The first one is the *orbital angular momentum*, which arises when a particle performs a motion in space, e.g., when an electron orbits a nucleus. This momentum represents the quantized counterpart of the classical notion of angular momentum. The second one is the *spin* or *intrinsic angular momentum*, that resembles the classical notion of particles spinning around their own axis. More precisely, the spin has no counterparts in classical mechanics and represents a quantum degree of freedom [103] that plays an important role in certain physical phenomena regarding, for instance, the interaction of elementary particles and atomic nuclei with magnetic fields; see, e.g., [20, 25, 46, 103]. Historically, the existence of the spin as a quantum degree of freedom was inferred from experiments, like the famous Stern–Gerlach experiment, in which particles are observed to possess angular momentum that cannot be accounted for by orbital angular momentum alone; see, e.g., [83, 103].

In this thesis, we focus on so-called Fermi particles, also called Fermions or spin-$\frac{1}{2}$ particles, which are particles with half-integer spin. To clarify this concept, consider a particle (for example an electron) irradiated by a spatially uniform and stationary magnetic field B_0, whose direction coincides with the z axis of the Cartesian coordinates. In this condition, the spin of the particle will be oriented in the same direction of the magnetic field, and can be pointing only in the two directions $+z$ and $-z$. These two conditions are referred to as spin-up (spin $+\frac{1}{2}$) and spin-down (spin $-\frac{1}{2}$). In particular, the spin-up and spin-down conditions correspond to two different energy levels, that are for convention assumed to be the lower and the higher energy level, respectively; see Figure 2.1(a). Moreover, the energetic distance between the two spin states is proportional to the amplitude of the magnetic field B_0 [20, 25], as shown in Figure 2.1(b).
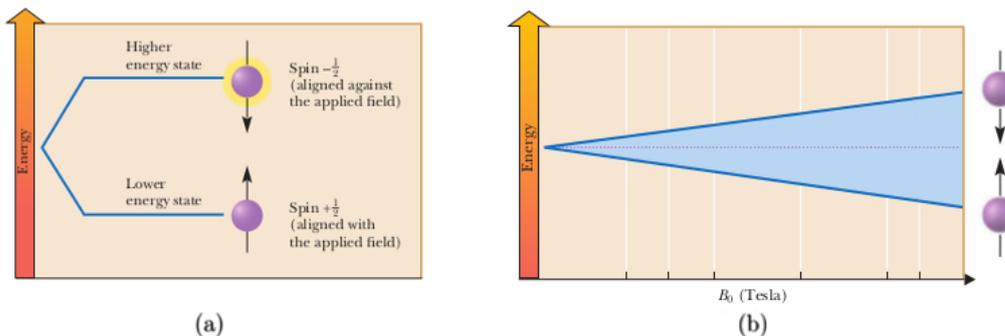
Figure 2.1: Spin states and energies: (a) The applied magnetic field B_0 create the two spin conditions corresponding to lower and higher energy states. (b) The energy of a spin state is proportional to the amplitude of the magnetic field B_0. In blue the energy difference between the allowed nuclear spin states is depicted. Source: Organic Chemistry - Brown et al. [20].

The energy corresponding to a spin state of a particle immersed in B_0 is given by the following scalar quantity

$$E = \gamma \hbar s B_0 \, , \tag{2.1}$$

where $\hbar$ is the reduced Planck constant, s is called spin quantum number, and γ is the gyromagnetic ratio, that is a constant related to the specific particle (or nuclei) given by

$$\gamma = \frac{g\mu_B}{\hbar} \, , \tag{2.2}$$

where g is the so-called g-factor, which is a dimensionless quantity that characterizes the magnetic moment of a particle, and μ_B is the nuclear magneton, that is a physical constant of magnetic moment. The nuclear magneton is given by $\mu_B = \frac{e\hbar}{\mu c}$, with e, μ the charge and the mass of the particle, respectively, and c the speed of light. The spin quantum number s is in general a half-integer and is used to identify the possible energy levels generated by the action of B_0. This phenomenon is known as Zeeman effect; see, e.g., [25] and references therein.

In quantum mechanics, scalar quantities are in general replaced by Hamilton operators [103, 83], and the Hamiltonian corresponding to the energy E in (2.1) is given by the following

$$H = g\mu_B I_z B_0 \ , \tag{2.3}$$

where I_z is a spin operator corresponding to the quantum number s, and we replace the gyromagnetic ratio with (2.2). Notice that the relationship between E and the corresponding Hamiltonian H is expressed by the following eigenvalue problem

$$H\varphi = E\varphi \ , \tag{2.4}$$

where φ is the wavefunction representing the quantum state of the system that corresponds to the energy E. Equation (2.4) is known as stationary Schrödinger equation. Notice that, in a more general situation, in which the particle (or the atomic nuclei) is irradiated by a magnetic field $\vec{B} = (B_x, B_y, B_z)$, the spin Hamiltonian, also known as Stern-Gerlach term, is given by the following

$$H = g\mu_B \vec{I} \cdot \vec{B} \ , \tag{2.5}$$

where $\vec{I}$ is the spin operator.

When immersed in a magnetic field B_0, a particle performs a particular rotational phenomenon, called precession, that is a rotation around the axis of the magnetic field B_0. This physical phenomenon is represented in the scheme shown in Figure 2.2. In particular, the precession frequency, that is in general called Larmor frequency, is obtained as

$$\nu_f = \gamma B_0 \ .$$

This quantity is used in the sequel of the present chapter for characterizing quantum spin control models.

The concept of spin and its relationship with external magnetic fields is particularly important in several nanotechnology applications, that are based on the so-called magnetic resonance, e.g., nuclear magnetic resonance (NMR) spectroscopy and magnetic resonance imaging (MRI). The concept of magnetic resonance can be explained as follows [20]. Consider an atomic nucleus immersed in a magnetic field B_0 and possessing a specific Larmor frequency. If irradiated by an electromagnetic radiation having frequency equal to the Larmor frequency, this nucleus absorbs energy and changes its spin state from the lower energy level (aligned with B_0) to the higher energy level (aligned against B_0). This phenomenon is shown in Figure 2.2. Resonance in this contest is the absorption of electromagnetic radiation by a precessing nucleus and its corresponding flip between energy levels [20]. This absorption can be detected by specific laboratory machines and used for applicative purposes, and the electromagnetic radiation is referred to as radiofrequency pulse or control.
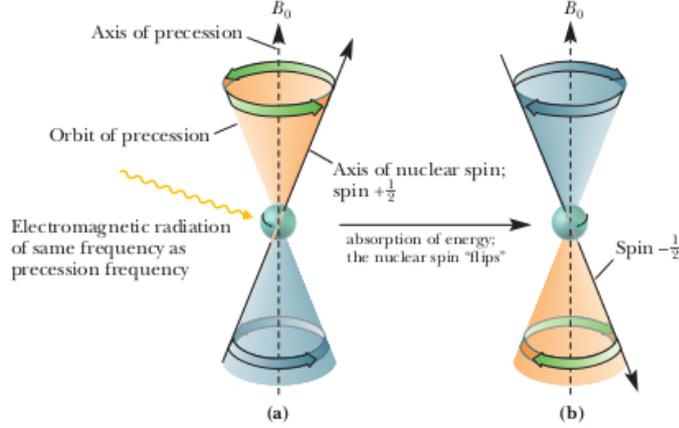
Figure 2.2: Magnetic resonance: (a) Precession of a spinning nucleus in an applied magnetic field. (b) Absorption of electromagnetic radiation occurs when the frequency of radiation is equal to the frequency of precession. Source: Organic Chemistry - Brown et al. [20].

2.2   The Pauli equation

The quantum state of a spin-$\frac{1}{2}$ particle is mathematically described by a so-called spinor, that is a vector-function $\vec{\psi}(t) \in L^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. Let $\vec{q} = (x, y, z) \in \mathbb{R}^3$ be the vector of space coordinates, a spinor is defined as follows; see, e.g., [83, 103, 118];

$$\vec{\psi}(t, \vec{q}) = \begin{pmatrix} \psi_+(t, \vec{q}) \\ \psi_-(t, \vec{q}) \end{pmatrix} , \tag{2.6}$$

where ψ_+ and ψ_- are wavefunctions corresponding to the spin-up and the spin-down conditions, respectively.

The time evolution of a spinor is governed by the Pauli equation. The Pauli equation, also known as the Schrödinger–Pauli equation, is the formulation of the Schrödinger equation for spin-$\frac{1}{2}$ particles, which takes into account the interaction of the particle's spin with an external, spatially uniform, and possibly time dependent magnetic field. This equation was formulated by Wolfgang Pauli in 1927 [95], and is given by the following; see, e.g., [83, 103, 118];

$$i\hbar \frac{\partial \vec{\psi}}{\partial t} = \left\{ \left[-\frac{1}{2\mu}\left(\vec{p} - \frac{e}{c}\vec{A}(\vec{q}, t)\right)^2 + V(\vec{q}) \right] \otimes I_2 + g\mu_B \vec{B}(t) \cdot \vec{I} \right\} \vec{\psi} , \tag{2.7}$$

where

- the physical constants $\hbar$, μ, e, c, g, μ_B represent the reduced Planck constant, the mass and the charge of the particle, the speed of light, the g-factor, and, the nuclear magneton, respectively;

- I_2 is the 2×2 identity matrix;

- $V(\vec{q}) \in \mathbb{R}$ is an external confining potential;

- $\vec{B}(t) \in \mathbb{R}^3$ is an external and time dependent magnetic field; it is given by

$$\vec{B}(t) = B_x(t)\hat{e}_x + B_y(t)\hat{e}_y + B_z\hat{e}_z , \tag{2.8}$$

with $(\hat{e}_x, \hat{e}_y, \hat{e}_z)$ the canonical Euclidean basis in $\mathbb{R}^3$; the component B_z is in general assumed to be constant, while B_x and B_y are time-dependent and are considered control functions;

- the operator $\vec{A}$ is the vector potential corresponding to the magnetic field $\vec{B}$ with $\vec{B} = \nabla \times \vec{A}$; it is given by

$$\vec{A} = \frac{1}{2}\vec{B} \times \vec{q}. \tag{2.9}$$

The vector potential is then

$$\vec{A} = \begin{vmatrix} \hat{e}_x & \hat{e}_y & \hat{e}_z \\ B_x & B_y & B_z \\ x & y & z \end{vmatrix} = \frac{1}{2}(zB_y - yB_z)\hat{e}_x + \frac{1}{2}(xB_z - zB_x)\hat{e}_y + \frac{1}{2}(yB_x - xB_y)\hat{e}_z \tag{2.10}$$

- $\vec{p}$ is the momentum operator, that is

$$\vec{p} = -i\hbar\big(\hat{e}_x\partial_x + \hat{e}_y\partial_y + \hat{e}_z\partial_z\big) \tag{2.11}$$

- $\vec{I}$ is the spin operator given by the following

$$\vec{I} = I_x\hat{e}_x + I_y\hat{e}_y + I_z\hat{e}_z , \tag{2.12}$$

where I_x, I_y, I_z are the three Pauli matrices, given explicitly by the following

$$I_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad I_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} , \quad I_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{2.13}$$

Notice that the term in (2.7) that is contained in the square brackets, and given by

$$H = \left[\frac{1}{2\mu}\Big(\vec{p} - \frac{e}{c}\vec{A}(\vec{q}, t)\Big)^2 + V(\vec{q})\right] , \tag{2.14}$$

represents the Hamiltonian of a Schrödinger equation describing the motion of a particle immersed in a magnetic field. The first term in H can be written as follows

$$\Big(\vec{p} - \frac{e}{c}\vec{A}\Big)^2 = |\vec{p}|^2 + \frac{e^2}{c^2}|A|^2 - \frac{e}{c}\vec{p}\cdot\vec{A} - \frac{e}{c}\vec{A}\cdot\vec{p}. \tag{2.15}$$

By recalling that $\vec{A}\cdot\vec{p} - \vec{p}\cdot\vec{A} = \mathrm{div}(\vec{A})$ [83], and noting that $\mathrm{div}(\vec{A}) = 0$, we obtain that $\vec{A}\cdot\vec{p}$ commutes with $\vec{p}\cdot\vec{A}$. For this reason and neglecting the term $\frac{e^2}{c^2}|A|^2$, the Hamiltonian H becomes as follows; see, e.g., [83, 103];

$$H = \left[\frac{1}{2\mu}|\vec{p}|^2 + \frac{e}{\mu c}\vec{A}(\vec{q}, t)\cdot\vec{p} + V(\vec{q})\right] . \tag{2.16}$$

Consequently, the Pauli equation assumes the following form

$$i\hbar\frac{\partial\vec{\psi}}{\partial t} = \left\{\left[\frac{1}{2\mu}|\vec{p}|^2 + \frac{e}{\mu c}\vec{A}(\vec{q}, t)\cdot\vec{p} + V(\vec{q})\right] \otimes I_2 + g\mu_B\vec{B}(t)\cdot\vec{I}\right\}\vec{\psi} . \tag{2.17}$$

Notice that the three terms appearing in H in (2.16), that are $|\vec{p}|^2$, $\vec{A}\cdot\vec{p}$, and V, represent the kinetic energy of the quantum system, the interaction between the momentum of the

particle and the external magnetic field, and the potential energy of the quantum system, respectively. The operator $g\mu_B \vec{B} \cdot \vec{I}$ in (2.17) is the Stern-Gerlach term and, as described in Section 2.1, represents the interaction between the external magnetic field and the particle's spin.

We remark that, the magnetic field $\vec{B}$ is used to control the quantum system. In particular, the component B_z is assumed to be constant in order to create the desired quantum states as described with B_0 in Section 2.1. The two components B_x and B_y are time dependent functions, and are used as controls with the purpose to steer the quantum system from an initial state to a desired target state. They can represent, for instance, the electromagnetic radiation discussed in Section 2.1 and shown in Figure 2.2.

2.2.1 The Pauli equation for a spin-$\frac{1}{2}$ particle in spherical coordinates

In this section, we describe the Pauli equation in spherical coordinates; see, e.g., [118]. This representation is usually used in quantum mechanics, since it allows a spectral discretization, that results to be useful for the analysis of the physical behaviour of quantum systems. Such a spectral discretization is discussed in the next section.

Let r be the radial distance, and θ, ϕ be the polar and azimuthal angles, respectively, and denote by $\vec{q} = (r, \theta, \phi)$. The Pauli equation (2.17) becomes as follows

$$i\hbar \frac{\partial \vec{\psi}}{\partial t} = \left\{ \left[-\frac{\hbar^2}{2\mu} \frac{1}{r} \frac{\partial^2}{\partial r^2} r + \frac{1}{2\mu r^2} L^2 - \frac{e}{4\mu c} \vec{B} \cdot \vec{L} + V(r) \right] I_2 + g\mu_B \vec{B} \cdot \vec{I} \right\} \vec{\psi} , \qquad (2.18)$$

where the spinor $\vec{\psi}$ is

$$\vec{\psi}(t, \vec{q}) = \begin{pmatrix} \psi_+(t, \vec{q}) \\ \psi_-(t, \vec{q}) \end{pmatrix} . \qquad (2.19)$$

The operator $\vec{L}$ is the following

$$\vec{L} = L_x \hat{e}_x + L_y \hat{e}_y + L_z \hat{e}_z , \qquad (2.20)$$

where

$$L_x = i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right) \qquad (2.21)$$

$$L_y = i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \qquad (2.22)$$

$$L_z = -i\hbar \frac{\partial}{\partial \phi} , \qquad (2.23)$$

and L^2 is given by

$$L^2 = L_x^2 + L_y^2 + L_z^2 = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} . \qquad (2.24)$$

Using (2.19), the Pauli equation (2.18) can be written explicitly as follows; see, e.g., [118];

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} H + g\mu_B B_z & g\mu_B(B_x - iB_y) \\ g\mu_B(B_x + iB_y) & H - g\mu_B B_z \end{pmatrix} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} , \qquad (2.25)$$

where the Hamiltonian H is given by

$$H = -\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}r + \frac{1}{2\mu r^2}L^2 - \frac{e}{4\mu c}\vec{B}\cdot\vec{L} + V(r) \ . \tag{2.26}$$

It is clear from (2.25), that the Pauli equation for a single particle is a system of two Schrödinger equations in ψ_+ and ψ_-:

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\psi_+(t,r,\theta,\phi) = &-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}\Big(r\psi_+(t,r,\theta,\phi)\Big) + \frac{1}{2\mu r^2}L^2\Big(\psi_+(t,r,\theta,\phi)\Big) \\
&-\frac{e}{4\mu c}\vec{B}(t)\cdot\vec{L}\Big(\psi_+(t,r,\theta,\phi)\Big) + (V(r) + g\mu_B B_z(t))\psi_+(t,r,\theta,\phi) \\
&+ g\mu_B(B_x(t) - iB_y(t))\psi_-(t,r,\theta,\phi) \ ,
\end{aligned}
\tag{2.27}
$$

and

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\psi_-(t,r,\theta,\phi) = &-\frac{\hbar^2}{2\mu}\frac{1}{r}\frac{\partial^2}{\partial r^2}\Big(r\psi_-(t,r,\theta,\phi)\Big) + \frac{1}{2\mu r^2}L^2\Big(\psi_-(t,r,\theta,\phi)\Big) \\
&-\frac{e}{4\mu c}\vec{B}(t)\cdot\vec{L}\Big(\psi_-(t,r,\theta,\phi)\Big) + (V(r) - g\mu_B B_z(t))\psi_-(t,r,\theta,\phi) \\
&+ g\mu_B(B_x(t) + iB_y(t))\psi_+(t,r,\theta,\phi) \ .
\end{aligned}
\tag{2.28}
$$

Notice that these equations are coupled by means of the two components of the external magnetic field B_x and B_y.

In the following section, we discuss a spectral discretization of the Pauli equation (2.25)-(2.28).

2.2.2 Spectral discretization of the Pauli equation

This section focuses on a discrete approximation of the Pauli-Schrödinger equation (2.25)-(2.28). For this purpose, we apply a spectral method and consider the following ansatz

$$\psi_\pm(t,r,\theta,\phi) = \sum_{n=0}^{\infty}\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} a_\pm^{n,\ell,m}(t)R_{n,\ell}(r)Y_\ell^m(\theta,\phi) \ , \tag{2.29}$$

where $Y_\ell^m(\theta,\phi)$ are spherical harmonics, $R_{n,\ell}(r)$ are eigenfunctions of the radial part of the Hamiltonian H, and the coefficients $a_+^{n,\ell,m}(t)$ and $a_-^{n,\ell,m}(t)$ are complex-valued time-dependent functions. Notice that the ansatz (2.29) is justified for specific choices of the potential $V(r)$, which guarantee that the radial functions $R_{n,\ell}(r)$ are properly defined.

The spherical harmonics are eigenfunctions of the operator L^2, that is $Y_\ell^m(\theta,\phi)$ solve the following eigenvalue problem

$$L^2 Y_\ell^m(\theta,\phi) = \hbar^2\ell(\ell+1)Y_\ell^m(\theta,\phi) \ . \tag{2.30}$$

The eigenfunctions $Y_\ell^m(\theta,\phi)$ satisfy the following orthogonality condition; see, e.g., [118];

$$\int_0^{2\pi}\int_0^{\pi} Y_\ell^m(\theta,\phi)Y_{\ell'}^{m'}(\theta,\phi)\sin\theta\,\mathrm{d}\theta d\phi = \delta_{\ell,\ell'}\delta_{m,m'} \ , \tag{2.31}$$

where $\delta_{mm'}$ is the Kronecker delta. Further, the following relations hold

$$L_z Y_\ell^m(\theta, \phi) = \hbar m Y_\ell^m(\theta, \phi) \tag{2.32}$$

$$L_\pm Y_\ell^m(\theta, \phi) = \hbar \sqrt{\ell(\ell+1) - m(m \pm 1)} Y_\ell^{m\pm 1}(\theta, \phi) \, , \tag{2.33}$$

where $L_\pm = L_x \pm i L_y$, and we have the following

$$L_x = \frac{1}{2}\left(L_+ + L_-\right) \tag{2.34}$$

$$L_y = \frac{1}{2i}\left(L_+ - L_-\right) \, . \tag{2.35}$$

More details regarding the spherical harmonics are given in the Appendix.

The radial functions $R_{n,\ell}(r)$ solve the following eigenvalue problem

$$O_\ell(V) R_{n,\ell}(r) = \lambda_{n,\ell} R_{n,\ell}(r) \, , \tag{2.36}$$

where

$$O_\ell(V) := \left\{ -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{\hbar^2}{2\mu} \frac{\ell(\ell+1)}{r^2} + V(r) \right\} \, . \tag{2.37}$$

The eigenfunctions $R_{n,\ell}(r)$ satisfy the following orthogonality condition

$$\int_0^\infty r^2 R_{n,\ell}(r) R_{n',\ell'}(r) dr = \delta_{\ell,\ell'} \delta_{n,n'} \, . \tag{2.38}$$

We remark that, since the operator $O_\ell(V)$ depends on the potential $V(r)$, the eigenfunctions $R_{n,\ell}(r)$ have to be computed separately for different cases of physical interest; see, e.g., the Appendix and [49, 51, 63, 83, 92, 103].

In the sequel, the following notation is used. The dependence on t, r, θ, ϕ is omitted for brevity, and we write (2.29) as follows

$$\begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} \sum\limits_{n,\ell,m} a_+^{n,\ell,m} R_{n,\ell} Y_\ell^m \\ \sum\limits_{n,\ell,m} a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m \end{pmatrix} \, . \tag{2.39}$$

where $\sum\limits_{n,\ell,m} = \sum\limits_{n=0}^\infty \sum\limits_{\ell=0}^\infty \sum\limits_{m=-\ell}^\ell$. Replacing (2.39) into (2.27), we get

$$\sum_{n,\ell,m} i\hbar \dot{a}_+^{n,\ell,m} R_{n,\ell} Y_\ell^m = \sum_{n,\ell,m} -a_+^{n,\ell,m} Y_\ell^m \frac{\hbar^2}{2\mu} \frac{1}{r} \frac{\partial^2}{\partial r^2}\left(r R_{n,\ell}\right) + a_+^{n,\ell,m} R_{n,\ell} \frac{1}{2\mu r^2} L^2\left(Y_\ell^m\right)$$

$$- a_+^{n,\ell,m} R_{n,\ell} \frac{e}{4\mu c} \vec{B} \cdot \vec{L}\left(Y_\ell^m\right) + (V(r) + g\mu_B B_z) a_+^{n,\ell,m} R_{n,\ell} Y_\ell^m$$

$$+ g\mu_B(B_x - i B_y) a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m \, , \tag{2.40}$$

and inserting (2.30) in (2.40), we write the following

$$
\begin{aligned}
\sum_{n,\ell,m} i\hbar \dot{a}_+^{n,\ell,m} R_{n,\ell} Y_\ell^m &= \sum_{n,\ell,m} -a_+^{n,\ell,m} Y_\ell^m \frac{\hbar^2}{2\mu} \frac{1}{r} \frac{\partial^2}{\partial r^2}\left(r R_{n,\ell}\right) + a_+^{n,\ell,m} R_{n,\ell} \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} Y_\ell^m \\
&\quad - a_+^{n,\ell,m} R_{n,\ell} \frac{e}{4\mu c} \vec{B} \cdot \vec{L}\left(Y_\ell^m\right) + (V(r) + g\mu_B B_z) a_+^{n,\ell,m} R_{n,\ell} Y_\ell^m \\
&\quad + g\mu_B (B_x - iB_y) a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m \\
&= \sum_{n,\ell,m} a_+^{n,\ell,m} Y_\ell^m \left\{ -\frac{\hbar^2}{2\mu} \frac{1}{r} \frac{\partial^2}{\partial r^2}\left(r R_{n,\ell}\right) + V(r) R_{n,\ell} + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} R_{n,\ell} \right\} \\
&\quad - a_+^{n,\ell,m} R_{n,\ell} \frac{e}{4\mu c} \vec{B} \cdot \vec{L}\left(Y_\ell^m\right) + g\mu_B B_z a_+^{n,\ell,m} R_{n,\ell} Y_\ell^m \\
&\quad + g\mu_B (B_x - iB_y) a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m \; .
\end{aligned}
\tag{2.41}
$$

Recalling that

$$
\frac{1}{r} \frac{\partial^2}{\partial r^2} r = \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) \; ,
\tag{2.42}
$$

and using (2.36) in (2.41), it holds that

$$
\begin{aligned}
\sum_{n,\ell,m} i\hbar \dot{a}_+^{n,\ell,m} R_{n,\ell} Y_\ell^m &= \sum_{n,\ell,m} a_+^{n,\ell,m} Y_\ell^m \lambda_{n,\ell} R_{n,\ell} - a_+^{n,\ell,m} R_{n,\ell} \frac{e}{4\mu c} \vec{B} \cdot \vec{L}\left(Y_\ell^m\right) \\
&\quad + g\mu_B B_z a_+^{n,\ell,m} R_{n,\ell} Y_\ell^m + g\mu_B (B_x - iB_y) a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m \; .
\end{aligned}
\tag{2.43}
$$

Next, consider the term $\vec{B} \cdot \vec{L}\left(Y_\ell^m\right)$ in (2.43). By means of (2.32)-(2.35), we have the following

$$
\begin{aligned}
\vec{B} \cdot \vec{L}\left(Y_\ell^m\right) &= B_x L_x\left(Y_\ell^m\right) + B_y L_y\left(Y_\ell^m\right) + B_z L_z\left(Y_\ell^m\right) \\
&= B_x \frac{1}{2}\left(\hbar\sqrt{\ell(\ell+1) - m(m+1)} Y_\ell^{m+1} + \hbar\sqrt{\ell(\ell+1) - m(m-1)} Y_\ell^{m-1}\right) \\
&\quad + B_y \frac{1}{2i}\left(\hbar\sqrt{\ell(\ell+1) - m(m+1)} Y_\ell^{m+1} - \hbar\sqrt{\ell(\ell+1) - m(m-1)} Y_\ell^{m-1}\right) \\
&\quad + B_z \hbar m Y_\ell^m \\
&= \frac{\hbar}{2}\left[B_x\left(K_{\ell,m}^+ Y_\ell^{m+1} + K_{\ell,m}^- Y_\ell^{m-1}\right) - iB_y\left(K_{\ell,m}^+ Y_\ell^{m+1} - K_{\ell,m}^- Y_\ell^{m-1}\right)\right] \\
&\quad + B_z \hbar m Y_\ell^m \\
&= \frac{\hbar}{2}\left[K_{\ell,m}^+ (B_x - iB_y) Y_\ell^{m+1} + K_{\ell,m}^- (B_x + iB_y) Y_\ell^{m-1}\right] + B_z \hbar m Y_\ell^m \; .
\end{aligned}
\tag{2.44}
$$

where $K_{\ell,m}^\pm := \sqrt{\ell(\ell+1) - m(m \pm 1)}$. Rearranging (2.43) and using (2.44), we obtain

the following

$$\sum_{n,\ell,m} i\hbar \dot{a}_+^{n,\ell,m} R_{n,\ell} Y_\ell^m = \sum_{n,\ell,m} a_+^{n,\ell,m} Y_\ell^m R_{n,\ell} \Big[\lambda_{n,\ell} + g\mu_B B_z\Big] + g\mu_B(B_x - iB_y)a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m$$

$$- a_+^{n,\ell,m} R_{n,\ell} \frac{e}{4\mu c} \vec{B} \cdot \vec{L}\Big(Y_\ell^m\Big)$$

$$= \sum_{n,\ell,m} a_+^{n,\ell,m} Y_\ell^m R_{n,\ell} \Big[\lambda_{n,\ell} + g\mu_B B_z\Big] + g\mu_B(B_x - iB_y)a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m$$

$$- a_+^{n,\ell,m} R_{n,\ell} \frac{e\hbar}{8\mu c} \Big[K_{\ell,m}^+(B_x - iB_y)Y_\ell^{m+1} + K_{\ell,m}^-(B_x + iB_y)Y_\ell^{m-1}\Big]$$

$$+ B_z \hbar m a_+^{n,\ell,m} R_{n,\ell} Y_\ell^m$$

$$= \sum_{n,\ell,m} a_+^{n,\ell,m} Y_\ell^m R_{n,\ell} \Big[\lambda_{n,\ell} + (g\mu_B + \hbar m)B_z\Big]$$

$$+ g\mu_B(B_x - iB_y)a_-^{n,\ell,m} R_{n,\ell} Y_\ell^m$$

$$- a_+^{n,\ell,m} R_{n,\ell} \frac{e\hbar}{8\mu c} \Big[K_{\ell,m}^+(B_x - iB_y)Y_\ell^{m+1} + K_{\ell,m}^-(B_x + iB_y)Y_\ell^{m-1}\Big].$$

$$(2.45)$$

By multiplying left- and right-hand sides of (2.45) with $\sin(\theta)r^2 R_{n',\ell'}Y_{\ell'}^{m'}$ and integrating, we get the following

$$\sum_{n,\ell,m} i\hbar \dot{a}_+^{n,\ell,m} \int_0^\infty \Big(r^2 R_{n,\ell} R_{n',\ell'}\Big) dr \int_0^{2\pi}\int_0^\pi Y_\ell^m Y_{\ell'}^{m'} \sin(\theta)d\theta d\phi$$

$$= \sum_{n,\ell,m} a_+^{n,\ell,m} \Big[\lambda_{n,\ell} + (g\mu_B + \hbar m)B_z\Big] \int_0^\infty \Big(r^2 R_{n,\ell} R_{n',\ell'}\Big) dr \int_0^{2\pi}\int_0^\pi Y_\ell^m Y_{\ell'}^{m'} \sin(\theta)d\theta d\phi$$

$$+ g\mu_B(B_x - iB_y)a_-^{n,\ell,m} \int_0^\infty \Big(r^2 R_{n,\ell} R_{n',\ell'}\Big) dr \int_0^{2\pi}\int_0^\pi Y_\ell^m Y_{\ell'}^{m'} \sin(\theta)d\theta d\phi \qquad (2.46)$$

$$- a_+^{n,\ell,m} \int_0^\infty \Big(r^2 R_{n,\ell} R_{n',\ell'}\Big) dr \frac{e\hbar}{8\mu c}\Big[K_{\ell,m}^+(B_x - iB_y) \int_0^{2\pi}\int_0^\pi Y_\ell^{m+1} Y_{\ell'}^{m'} \sin(\theta)d\theta d\phi$$

$$+ K_{\ell,m}^-(B_x + iB_y) \int_0^{2\pi}\int_0^\pi Y_\ell^{m-1} Y_{\ell'}^{m'} \sin(\theta)d\theta d\phi\Big].$$

Because of the orthogonality conditions (2.31) and (2.38), it holds that

$$\sum_{n,\ell,m} i\hbar \dot{a}_+^{n,\ell,m} \delta_{\ell,\ell'}\delta_{n,n'}\delta_{\ell,\ell'}\delta_{m,m'}$$

$$= \sum_{n,\ell,m} a_+^{n,\ell,m} \Big[\lambda_{n,\ell} + (g\mu_B + \hbar m)B_z\Big]\delta_{\ell,\ell'}\delta_{n,n'}\delta_{\ell,\ell'}\delta_{m,m'} + g\mu_B(B_x - iB_y)a_-^{n,\ell,m}\delta_{\ell,\ell'}\delta_{n,n'}\delta_{\ell,\ell'}\delta_{m,m'}$$

$$- a_+^{n,\ell,m}\delta_{\ell,\ell'}\delta_{n,n'}\frac{e\hbar}{8\mu c}\Big[K_{\ell,m}^+(B_x - iB_y)\delta_{\ell,\ell'}\delta_{m+1,m'} + K_{\ell,m}^-(B_x + iB_y)\delta_{\ell,\ell'}\delta_{m-1,m'}\Big].$$

$$(2.47)$$

Notice that $\delta_{\ell,\ell'}\delta_{\ell,\ell'} = \delta_{\ell,\ell'}$ and that $\delta_{m\pm1,m'} = \delta_{m,m'\mp1}$. Therefore (2.47) becomes

$$
\sum_{n,\ell,m} i\hbar\dot{a}_+^{n,\ell,m}\delta_{\ell,\ell'}\delta_{n,n'}\delta_{m,m'}
$$

$$
= \sum_{n,\ell,m} a_+^{n,\ell,m}\Big[\lambda_{n,\ell} + (g\mu_B + \hbar m)B_z\Big]\delta_{\ell,\ell'}\delta_{n,n'}\delta_{m,m'} + g\mu_B(B_x - iB_y)a_-^{n,\ell,m}\delta_{\ell,\ell'}\delta_{n,n'}\delta_{m,m'}
$$

$$
- a_+^{n,\ell,m}\delta_{\ell,\ell'}\delta_{n,n'}\frac{e\hbar}{8\mu c}\Big[K_{\ell,m}^+(B_x - iB_y)\delta_{m,m'-1} + K_{\ell,m}^-(B_x + iB_y)\delta_{m,m'+1}\Big] .
$$

$$(2.48)$$

By using the definition of the Kronecker-delta, and truncating the sums over n, ℓ, m, we obtain that

$$
i\hbar\dot{a}_+^{n,\ell,m} = a_+^{n,\ell,m}\Big[\lambda_{n,\ell} + (\hbar m + g\mu_B)B_z\Big] + g\mu_B(B_x - iB_y)a_-^{n,\ell,m}
$$
$$
- a_+^{n,\ell,m-1}\frac{e\hbar}{8\mu c}K_{\ell,m-1}^+(B_x - iB_y) - a_+^{n,\ell,m+1}\frac{e\hbar}{8\mu c}K_{\ell,m+1}^-(B_x + iB_y) ,
$$

$$(2.49)$$

for $n = 0,\ldots,\tilde{n} - 1$, $\ell = 0,\ldots,\tilde{\ell} - 1$ and $m = -\ell,\ldots,\ell$.

We can proceed in the same way for (2.28) to obtain the following

$$
i\hbar\dot{a}_-^{n,\ell,m} = a_-^{n,\ell,m}\Big[\lambda_{n,\ell} + (\hbar m - g\mu_B)B_z\Big] + g\mu_B(B_x + iB_y)a_+^{n,\ell,m}
$$
$$
- a_-^{n,\ell,m-1}\frac{e\hbar}{8\mu c}K_{\ell,m-1}^+(B_x - iB_y) - a_-^{n,\ell,m+1}\frac{e\hbar}{8\mu c}K_{\ell,m+1}^-(B_x + iB_y) ,
$$

$$(2.50)$$

for $n = 0,\ldots,\tilde{n} - 1$, $\ell = 0,\ldots,\tilde{\ell} - 1$ and $m = -\ell,\ldots,\ell$.

In order to write (2.49) and (2.50) in a compact form, we define the following

$$
a_\pm := \Big(a_\pm^{0,0,0}, a_\pm^{0,1,-1}, a_\pm^{0,1,0}, a_\pm^{0,1,1}, a_\pm^{0,2,-2}, a_\pm^{0,2,-1}, \ldots, a_\pm^{0,\tilde{\ell}-1,\tilde{\ell}-1}, a_\pm^{1,0,0}, \ldots, a_\pm^{\tilde{n}-1,\tilde{\ell}-1,\tilde{\ell}-1}\Big)^T .
$$

$$(2.51)$$

Notice that, the number of entries of $a_\pm$ due to ℓ and m is $\tilde{\ell}^2$. Therefore, we have that $a_\pm(t) \in \mathbb{C}^{\tilde{n}\tilde{\ell}^2}$. Consequently, (2.49) and (2.50) can be equivalently written as follows

$$
i\hbar\dot{a}_+ = \Big[\Lambda + I_{\tilde{n}} \otimes \Big(\hbar M + g\mu_B I_{\tilde{\ell}^2}\Big)B_z\Big]a_+ + (B_x - iB_y)\Big[g\mu_B I_{\tilde{n}\tilde{\ell}^2}\Big]a_-
$$
$$
+ (B_x - iB_y)\Big[I_{\tilde{n}} \otimes K_{low}\Big]a_+ + (B_x + iB_y)\Big[I_{\tilde{n}} \otimes K_{up}\Big]a_+ ,
$$

$$(2.52)$$

and

$$
i\hbar\dot{a}_- = \Big[\Lambda + I_{\tilde{n}} \otimes \Big(\hbar M - g\mu_B I_{\tilde{\ell}^2}\Big)B_z\Big]a_- + (B_x + iB_y)\Big[g\mu_B I_{\tilde{n}\tilde{\ell}^2}\Big]a_+
$$
$$
+ (B_x - iB_y)\Big[I_{\tilde{n}} \otimes K_{low}\Big]a_- + (B_x + iB_y)\Big[I_{\tilde{n}} \otimes K_{up}\Big]a_- ,
$$

$$(2.53)$$

where, for a given integer $k \in \mathbb{N}$, I_k is the $k \times k$ identity. The matrix Λ is a $\tilde{n}\tilde{\ell}^2 \times \tilde{n}\tilde{\ell}^2$ block-diagonal matrix. Every block is of the form $\lambda_{n,\ell}I_{2\ell+1}$, and Λ is given by the following

$$
\Lambda = \text{blk-diag}\Big(\lambda_{0,0}I_1 , \lambda_{0,1}I_3 , \lambda_{0,2}I_5 , \cdots , \lambda_{0,\tilde{\ell}-1}I_{2\tilde{\ell}-1} , \lambda_{1,0}I_1 , \cdots , \lambda_{\tilde{n}-1,\tilde{\ell}-1}I_{2\tilde{\ell}-1}\Big) .
$$

$$(2.54)$$

The diagonal matrix M has dimension $\tilde{\ell}^2 \times \tilde{\ell}^2$ and is given by

$$M = \text{diag}\big(0,\quad -1,0,1,\quad -2,-1,0,1,2,\quad -3,\cdots,\tilde{\ell}-1\big). \tag{2.55}$$

The matrices K_{low} and K_{up} have dimension $\tilde{\ell}^2 \times \tilde{\ell}^2$ and are defined as

$$K_{low} = \begin{pmatrix} 0 & & & & & & & & \\ K_{0,0}^+ & 0 & & & & & & & \\ & K_{1,-1}^+ & 0 & & & & & & \\ & & K_{1,0}^+ & 0 & & & & & \\ & & & K_{1,1}^+ & 0 & & & & \\ & & & & K_{2,-2}^+ & 0 & & & \\ & & & & & \ddots & \ddots & & \\ & & & & & & K_{\tilde{\ell}-1,\tilde{\ell}-3}^+ & 0 & \\ & & & & & & & K_{\tilde{\ell}-1,\tilde{\ell}-2}^+ & 0 \end{pmatrix}, \tag{2.56}$$

and

$$K_{up} = \begin{pmatrix} 0 & K_{1,-1}^- & & & & & \\ & 0 & K_{1,0}^- & & & & \\ & & 0 & K_{1,1}^- & & & \\ & & & 0 & K_{2,-2}^- & & \\ & & & & 0 & K_{2,-1}^- & \\ & & & & & \ddots & \ddots \\ & & & & & & 0 & K_{\tilde{\ell}-1,\tilde{\ell}-1}^- \\ & & & & & & & 0 \end{pmatrix}, \tag{2.57}$$

respectively.

Now, we define $a(t) \in \mathbb{C}^{2\tilde{n}\tilde{\ell}^2}$ as follows

$$a := \begin{pmatrix} a_+ \\ a_- \end{pmatrix}, \tag{2.58}$$

and consequently, the equations (2.52) and (2.53) can be written in the following equivalent system

$$\begin{aligned} i\hbar\dot{a} &= \Big[I_2 \otimes \Lambda + B_z\Big(\hbar I_{2\tilde{n}} \otimes M + g\mu_B I_z \otimes I_{\tilde{n}\tilde{\ell}^2}\Big)\Big]a \\ &+ g\mu_B\Big[(B_x I_x + B_y I_y) \otimes I_{\tilde{n}\tilde{\ell}^2}\Big]a \\ &+ \Big[(B_x - iB_y)\Big(I_{2\tilde{n}} \otimes K_{low}\Big) + (B_x + iB_y)\Big(I_{2\tilde{n}} \otimes K_{up}\Big)\Big]a. \end{aligned} \tag{2.59}$$

We write (2.59) in the following compact form

$$i\hbar\dot{a} = \Big[H_0 + B_x H_x + B_y H_y\Big]a, \tag{2.60}$$

where

$$H_0 = \Big[I_2 \otimes \Lambda + B_z\Big(\hbar I_{2\tilde{n}} \otimes M + g\mu_B I_z \otimes I_{\tilde{n}\tilde{\ell}^2}\Big)\Big], \tag{2.61}$$

$$H_x = \Big[g\mu_B I_x \otimes I_{\tilde{n}\tilde{\ell}^2} + I_{2\tilde{n}} \otimes K_+\Big], \tag{2.62}$$

and

$$H_y = \left[g\mu_B I_y \otimes I_{\tilde{n}\tilde{\ell}^2} + I_{2\tilde{n}} \otimes iK_- \right], \tag{2.63}$$

with K^+ and K^- defined as follows

$$K_+ := K_{up} + K_{low} \quad , \quad K_- := K_{up} - K_{low}. \tag{2.64}$$

Notice that the Hamiltonian in (2.60) is written as the sum of a free Hamiltonian H_0, and a controlled Hamiltonian composed by the sum of $B_x H_x$ and $B_y H_y$. The free Hamiltonian H_0 depends on the stationary component of the magnetic field B_z and represents the energy of the uncontrolled quantum system. The controlled Hamiltonian depends on the time-dependent components of the magnetic field B_x and B_y, that are referred to as control functions. The three matrices H_0, H_x and H_y are Hermitian. In fact, H_0 is a diagonal matrix with real entries. To see that H_x and H_y are Hermitian, first we recognize that the Pauli matrices I_x and I_y are Hermitian. Then, we notice that K_+ and K_- are symmetric and skew-symmetric with real entries, respectively. This is due to the following properties of the coefficients $K_{\ell,m}^{\pm}$

$$\begin{aligned} K_{\ell,m-1}^+ &= \sqrt{\ell(\ell+1) - (m-1)((m-1)+1)} \\ &= \sqrt{\ell(\ell+1) - (m-1)m} \\ &= K_{\ell,m}^-, \end{aligned} \tag{2.65}$$

$$K_{\ell,\ell}^+ = \sqrt{\ell(\ell+1) - \ell(\ell+1)} = 0, \tag{2.66}$$

and

$$\begin{aligned} K_{\ell,-\ell}^- &= \sqrt{\ell(\ell+1) - (-\ell)(-\ell-1)} \\ &= \sqrt{\ell(\ell+1) - \ell(\ell+1)} = 0. \end{aligned} \tag{2.67}$$

Consequently, by means of (2.65), (2.66) and (2.67) we have that K_+ and K_- are the following

$$K_+ = \begin{pmatrix} 0 & 0 & & & & & & \\ 0 & 0 & K_{1,-1}^+ & & & & & \\ & K_{1,-1}^+ & 0 & K_{1,0}^+ & & & & \\ & & K_{1,0}^+ & 0 & 0 & & & \\ & & & 0 & 0 & K_{2,-2}^+ & & \\ & & & & K_{2,-2}^+ & 0 & K_{2,-1}^+ & \\ & & & & & \ddots & \ddots & \ddots \\ & & & & & & K_{\tilde{\ell}-1,\tilde{\ell}-3}^+ & 0 & K_{\tilde{\ell}-1,\tilde{\ell}-2}^+ \\ & & & & & & & K_{\tilde{\ell}-1,\tilde{\ell}-2}^+ & 0 \end{pmatrix}, \tag{2.68}$$

and

$$K_- = \begin{pmatrix} 0 & 0 & & & & & & \\ 0 & 0 & K_{1,-1}^+ & & & & & \\ & -K_{1,-1}^+ & 0 & K_{1,0}^+ & & & & \\ & & -K_{1,0}^+ & 0 & 0 & & & \\ & & & 0 & 0 & K_{2,-2}^+ & & \\ & & & & -K_{2,-2}^+ & 0 & K_{2,-1}^+ & \\ & & & & & \ddots & \ddots & \ddots \\ & & & & & & -K_{\tilde{\ell}-1,\tilde{\ell}-3}^+ & 0 & K_{\tilde{\ell}-1,\tilde{\ell}-2}^+ \\ & & & & & & & -K_{\tilde{\ell}-1,\tilde{\ell}-2}^+ & 0 \end{pmatrix}. \tag{2.69}$$

2.2.3 Quantum numbers

In Section 2.2.2, a spectral discretization of the Pauli equation is applied by means of the ansatz (2.29), in which the three indexes n, ℓ, and m appear. These integers have a specific physical meaning in quantum mechanics. They are called quantum numbers, because every quantum state of a particle immersed in a uniform magnetic field is identified with a triple (n, ℓ, m). Further, as introduced in Section 2.1, the quantum number s identifies the spin of a particle. More specifically:

- n is the radial quantum number; it describes the radial behaviour of a quantum state; n is used together with the azimuthal quantum number ℓ to determine the principal quantum number n_p, that is given by

$$n_p = n + \ell + 1 \ .$$

  The principal quantum number n_p identifies the energy levels of the quantum system;

- ℓ is the azimuthal quantum number; it is a parameter that quantizes the momentum operator L^2, and describes the shape of the orbital; it can assume only the values $0, 1, 2, \ldots$;

- m is the magnetic quantum number; it determines the energy shift of an atomic orbital due to an external magnetic field, hence the name magnetic quantum number (Zeeman effect); m is an eigenvalue of the operator L_z and can assume only the values $-\ell, -\ell + 1, \ldots, \ell$;

- s is the spin quantum number; it parametrizes the intrinsic angular momentum of a given particle.

The quantum numbers are used in the following section to analyse the transitions that can be performed by means of a spatially uniform and time dependent magnetic field.

2.2.4 The Pauli equation as a finite-dimensional Schrödinger equation

In this section, we discuss the possible transitions allowed by a spatial uniform magnetic field $\vec{B}$. By means of this analysis, we notice that controlling the infinite-dimensional Pauli equation is equivalent to control a finite-dimensional Schrödinger-Pauli equation.

For this purpose, consider the system (2.60). It is composed of several subsystems that are connected only because of the action of the controls $B_x(t)$ and $B_y(t)$. If the corresponding quantum system is in a given eigenstate of the free Hamiltonian H_0, that is n' and ℓ' are fixed, the coefficients $a_{\pm}^{n,\ell,m}$ vanish for any $n \neq n'$ and $\ell \neq \ell'$. Due to the particular block-structure of the Hamiltonian in (2.60), the action of the time-dependent spatially uniform magnetic field cannot affect the quantum numbers n' and ℓ'. Mathematically, this means that $a_{\pm}^{n,\ell,m} = 0$, for any $n \neq n'$ and $\ell \neq \ell'$, are equilibrium points for the corresponding n, ℓ-blocks of equations. Hence, $B_x(t)$ and $B_y(t)$ can only affect the spin-orientation s and the transition between different quantum numbers m and m'. The possible transitions due to the controls B_x and B_y are represented in the following scheme

$$(n', \ell', m, s) \xrightarrow[\text{Pauli eq.}]{(B_x, B_y)} (n', \ell', m', s') \ ,$$

and we remark that, these transitions are not affected by the truncation of the sums in the ansatz (2.29).

The vectors of coefficients $a_\pm$ defined in (2.51) reduces to to the following

$$a_\pm^{n',\ell'} = \left(a_\pm^{n',\ell',-\ell'} \, , \, a_\pm^{n',\ell',-\ell'+1} \, , \, \ldots \, , \, a_\pm^{n',\ell',\ell'-1} \, , \, \ldots \, , \, a_\pm^{n',\ell',\ell'} \right)^T , \tag{2.70}$$

and we have that $a_\pm^{n',\ell'}(t) \in \mathbb{C}^{2\ell'+1}$. The vector a defined in (2.58) becomes as follows

$$a^{n',\ell'} := \begin{pmatrix} a_+^{n',\ell'} \\ a_-^{n',\ell'} \end{pmatrix} , \tag{2.71}$$

and we have that $a^{n',\ell'}(t) \in \mathbb{C}^{4\ell'+2}$.

Next, consider the overall system (2.60). First, we notice that the matrices Λ and M, defined in (2.54) and (2.55), respectively, can be written as follows

$$\Lambda = \text{blk-diag}_{n=0,\ldots,\tilde{n}-1} \left(\Lambda_n \right) , \tag{2.72}$$

and

$$M = \text{blk-diag}_{n=0,\ldots,\tilde{n}-1} \left(M_n \right) , \tag{2.73}$$

respectively. We have that

$$\Lambda_n = \text{blk-diag}_{\ell=0,\ldots,\tilde{\ell}-1} \left(\Lambda_{n,\ell} \right) , \tag{2.74}$$

where $\Lambda_{n,\ell} = \lambda_{n,\ell} I_{2\ell+1}$, and

$$M_n = \text{blk-diag}_{\ell=0,\ldots,\tilde{\ell}-1} \left(M_{n,\ell} \right) , \tag{2.75}$$

where $M_{n,\ell} = \text{diag}(-\ell,\ldots,\ell)$. Moreover, the matrices K_+ and K_- can be written as follows

$$\begin{aligned} K_+ &= \text{blk-diag}_{\ell=0,\ldots,\tilde{\ell}-1} \left(K_+^\ell \right) \\ K_- &= \text{blk-diag}_{\ell=0,\ldots,\tilde{\ell}-1} \left(K_-^\ell \right) \end{aligned} \tag{2.76}$$

where

$$K_+^\ell := \begin{pmatrix} 0 & K_{\ell,-\ell}^+ & & & & \\ K_{\ell,-\ell}^+ & 0 & K_{\ell,-\ell+1}^+ & & & \\ & K_{\ell,-\ell+1}^+ & 0 & K_{\ell,-\ell+2}^+ & & \\ & & \ddots & \ddots & \ddots & \\ & & & K_{\ell,\ell-2}^+ & 0 & K_{\ell,\ell-1}^+ \\ & & & & K_{\ell,\ell-1}^+ & 0 \end{pmatrix} , \tag{2.77}$$

and

$$K_-^\ell := \begin{pmatrix} 0 & K_{\ell,-\ell}^+ & & & & \\ -K_{\ell,-\ell}^+ & 0 & K_{\ell,-\ell+1}^+ & & & \\ & -K_{\ell,-\ell+1}^+ & 0 & K_{\ell,-\ell+2}^+ & & \\ & & \ddots & \ddots & \ddots & \\ & & & -K_{\ell,\ell-2}^+ & 0 & K_{\ell,\ell-1}^+ \\ & & & & -K_{\ell,\ell-1}^+ & 0 \end{pmatrix} . \tag{2.78}$$

Notice that $\Lambda_{n,\ell}, M_{n,\ell}, K_+^\ell, K_-^\ell \in \mathbb{R}^{2\ell+1 \times 2\ell+1}$. Consequently, for fixed $n = n'$ and $\ell = \ell'$, the Pauli equation (2.60) becomes as follows

$$i\hbar \dot{a}^{n',\ell'} = \left[H_0 + B_x H_x + B_y H_y\right] a^{n',\ell'} , \tag{2.79}$$

where

$$H_0 = \left[I_2 \otimes \Lambda_{n',\ell'} + B_z \left(\hbar I_2 \otimes M_{n',\ell'} + g\mu_B I_z \otimes I_{2\ell'+1}\right)\right] , \tag{2.80}$$

$$H_x = \left[g\mu_B I_x \otimes I_{2\ell'+1} + I_2 \otimes K_+^{\ell'}\right] , \tag{2.81}$$

and

$$H_y = \left[g\mu_B I_y \otimes I_{2\ell'+1} + I_2 \otimes iK_-^{\ell'}\right] . \tag{2.82}$$

For the sake of clarity, we provide the following simple example of system (2.79).

Example. Consider $n' = 1$ and $\ell' = 1$. We have that $2\ell' + 1 = 3$ and $m = -1, 0, 1$ and the finite-dimensional Pauli-Schrödinger equation (2.79) becomes as follows

$$a_\pm^{1,1} = \begin{pmatrix} a_\pm^{1,1,-1} & a_\pm^{1,1,0} & a_\pm^{1,1,1} \end{pmatrix}^T$$

$$a^{1,1} = \begin{pmatrix} a_+^{1,1,-1} & a_+^{1,1,0} & a_+^{1,1,1} & a_-^{1,1,-1} & a_-^{1,1,0} & a_-^{1,1,1} \end{pmatrix}^T$$

$$\Lambda_{1,1} = \begin{pmatrix} \lambda_{1,1} & & \\ & \lambda_{1,1} & \\ & & \lambda_{1,1} \end{pmatrix} \qquad M_{1,1} = \begin{pmatrix} -1 & & \\ & 0 & \\ & & 1 \end{pmatrix}$$

$$K_+^1 = \begin{pmatrix} 0 & K_{1,-1}^+ & 0 \\ K_{1,-1}^+ & 0 & K_{1,0}^+ \\ 0 & K_{1,0}^+ & 0 \end{pmatrix} \qquad K_-^1 = \begin{pmatrix} 0 & K_{1,-1}^+ & 0 \\ -K_{1,-1}^+ & 0 & K_{1,0}^+ \\ 0 & -K_{1,0}^+ & 0 \end{pmatrix}$$

where $K_{1,-1}^+ = \sqrt{2}$ and $K_{1,0}^+ = \sqrt{2}$. $\triangle$

Next, consider equation (2.79) for $n' = 0$ and $\ell' = 0$. In this specific case, we have that $K_+^0 = K_-^0 = 0$, $M_{0,0} = 0$ and $\Lambda_{0,0} = \lambda_{0,0}$. Hence, (2.79) reduces to the following

$$i\hbar \dot{a}^{0,0} = \left[H_0 + B_x H_x + B_y H_y\right] a^{0,0} , \tag{2.83}$$

where $H_0 = \lambda_{0,0} I_2 + g\mu_B B_z I_z$, $H_x = g\mu_B I_x$, and $H_y = g\mu_B I_y$. Moreover, since the addition of a constant term to the Hamiltonian generates only a translation of its eigenvalues, the term $\lambda_{0,0} I_2$ can be neglected, and the free Hamiltonian becomes as $H_0 = g\mu_B B_z I_z$. In our case, the finite-dimensional Schrödinger-Pauli equation (2.83) governs the time evolution of a spin-$\frac{1}{2}$ particle in magnetic field $\vec{B}$. The particle is assumed to be in the lowest energy level, that is $(n, \ell) = (0, 0)$, and (2.83) describes only transitions between spin states, that is

$$(0, 0, 0, s) \xrightarrow[\text{Pauli eq.}]{(B_x, B_y)} (0, 0, 0, s') .$$

We remark that, (2.83) is largely used in nanotechnology applications as NMR and MRI, where the purpose is to consider the time evolution of the spin and magnetic behaviour of specific nuclei.

Equation (2.83) can be generalized for systems of interacting particles. This generalization is described in the next section.

2.3 The Liouville-von Neumann master equation

In this section, a detailed discussion regarding the Liouville-von Neumann master (LvNM) equation is given. This equation characterizes the time-evolution of density matrices. A density matrix is a mathematical object used to characterize the quantum state of an ensemble of spin-particles, that are physically equal, but in different quantum states.

The LvNM equation is largely used in NMR, MRI, quantum computing, etc., where by means of specific magnetic fields, the aim is to control the dynamics of a quantum state characterized by a density matrix. This concept is summarized in the following Figure 2.3, where the tiny red arrows represent the spins, that form an ensemble (light blue disc). The big red arrow represents a magnetic field that is pulsed on the ensemble of spins with the purpose of performing specific spin-transitions.
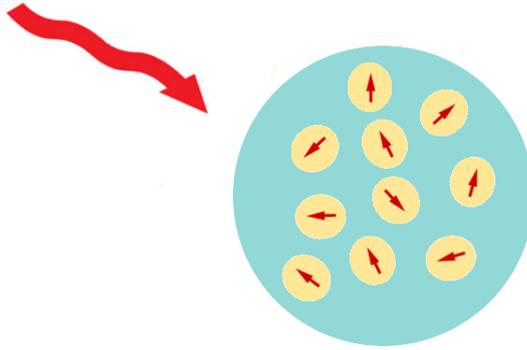

Figure 2.3: Control of an ensemble of spins: a magnetic field (big red arrow) is pulsed on an ensemble of spins (small red arrows in the light blue disc) in order to control the system and to steer the spin orientation from an initial state to a desired target state.

In particular, first we describe the density matrix as an alternative representation of the state of a quantum spin system. Second, we derive the LvNM equation that governs the time evolution of a density matrix. Then, we discuss a multi-spin controlled Hamiltonian corresponding to quantum spin systems especially considered in NMR.

2.3.1 The density matrix

A pure state of a quantum spin system can be represented by a normalized vector of a complex finite-dimensional Hilbert space $\mathcal{H}$ endowed with a scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. In this space, normalized vectors that differ only by a scalar phase factor are considered equivalent. Here and henceforth, an arbitrary vector of $\mathcal{H}$ is denoted by a and a^* represents the transpose complex-conjugate of a. We remark that, a can represent, for example, a finite-dimensional wavefunction and the coefficients $a_{\pm}^{n,\ell,m}$ described in Section 2.2.2.

Alternatively, a pure state can be described by a rank-one-projector of the form $a \otimes a^*$, where $\otimes$ denote the Kronecker product. However, in many applications, e.g., in NMR spectroscopy, the given quantum system can be regarded as a probabilistic mixture of many non-interactive subsystems. Such quantum ensembles are described by density matrices, that are Hermitian operators of the form

$$\rho = \sum_k p_k \left(a_k \otimes a_k^* \right) , \tag{2.84}$$

where $p_k \geq 0$ with $\sum_k p_k = 1$ can be interpreted as the fraction of subsystems being in state a_k. Notice that, since we assume that the vectors a_k are normalized, that is $\|a_k\|_2 = 1$, where $\|\cdot\|_2$ is the Euclidean norm, it holds that $\text{trace}(a_k \otimes a_k^*) = 1$. Moreover, in the case that $a_1, a_2, a_3, \ldots$, constitute an orthonormal basis for $\mathcal{H}$, the constants $p_1, p_2, p_3, \cdots \geq 0$ are the eigenvalues of ρ. In addition, the trace condition yields

$$\text{trace}(\rho) = \sum_k p_k = 1 \ .$$

This means that the eigenvalues of the density operator define a probability distribution.

These concepts lead to the following definitions, where we assume $\mathcal{H} = \mathbb{C}^N$; see, e.g., [25, 118].

Definition 1 (Density operator). *Let $\mathfrak{her}(N)$ denote the set of all Hermitian $N \times N$ matrices. Then $\rho \in \mathfrak{her}(N)$ is called density operator if it satisfies the following properties*

- *semi-positivity: $\rho \geq 0$,*
- *normalization: $\text{trace}(\rho) = 1$.*

Moreover, if a density operator ρ is a rank-1-projector (i.e., if ρ satisfies $\rho^2 = \rho$) then it is called a pure state, otherwise, it is called a mixed state.

In particular, if we consider subsystems of N_p coupled spin-$\frac{1}{2}$ particles, (common in NMR), the state space of a single spin-$\frac{1}{2}$ particle is given by $\mathbb{C}^2$ and the state space of N_p coupled spin-$\frac{1}{2}$ particles is obtained by the tensor product $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 \cong \mathbb{C}^{2^{N_p}}$. Thus the corresponding density operator belongs to the space $\mathfrak{her}(N)$, with $N = 2^{N_p}$.

For the sake of clarity, we discuss the following example of a density matrix.

Example. Consider the system depicted in Figure 2.4. It is an ensemble of four subsystems of two coupled spins, hence $N_p = 2$ and $N = 2^{N_p} = 4$.
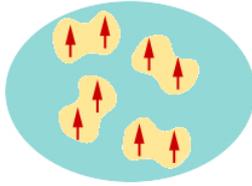
Figure 2.4: Ensemble of four spin subsystems. Each subsystem is formed by two coupled spins and its quantum state is represented by a vector $a_k \in \mathbb{C}^4$.

Assume that the states of the four subsystems are represented by the following four vectors

$$a_1 = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 & 0 & 0 \end{pmatrix}^T \qquad a_2 = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 & 0 & 0 \end{pmatrix}^T$$
$$a_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}^T \qquad a_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}^T \ ,$$

and notice that they form an orthonormal basis for $\mathcal{H} = \mathbb{C}^4$. The corresponding pure states are

$$a_1 \otimes a_1^* = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad a_2 \otimes a_2^* = \begin{pmatrix} 1/2 & -1/2 & 0 & 0 \\ -1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$a_3 \otimes a_3^* = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad a_4 \otimes a_4^* = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} .$$

Notice that $\|a_k\|_2 = 1$, where $\| \cdot \|_2$ is the Euclidean norm for $\mathbb{C}^4$. Consequently, it holds that $\text{trace}(a_k \otimes a_k^*) = 1$ for $k = 1, 2, 3, 4$. The density matrix is given by

$$\rho = 1/2 \begin{pmatrix} p_1 + p_2 & p_1 - p_2 & 0 & 0 \\ p_1 - p_2 & p_1 + p_2 & 0 & 0 \\ 0 & 0 & 2p_3 & 0 \\ 0 & 0 & 0 & 2p_4 \end{pmatrix} ,$$

and it is clear that p_1, p_2, p_3 and p_4 are its eigenvalues. $\triangle$

2.3.2 From the Pauli equation to the LvNM equation

The time evolution of a density matrix is governed by the Liouville-von Neumann master (LvNM) equation, that is [25, 70, 118]

$$\dot{\rho} = -i[H, \rho] , \tag{2.85}$$

where H is the Hamiltonian and $[\cdot, \cdot]$ represents the commutator operator, that is $[H, \rho] = H\rho - \rho H$.

This equation can be obtained by manipulating a finite-dimensional Schrödinger equation, modelling the evolution of a wavefunction a, that is

$$i\dot{a} = Ha , \tag{2.86}$$

where the reduced Plank constant $\hbar$ is assumed equal to 1 for simplicity and all energies are measured in frequency units. Here, at first H can be any Hermitian operator. Notice that (2.86) corresponds exactly to the Pauli equation (2.83), in the case that H is given by $[H_0 + B_x H_x + B_y H_y]$, as shown in the previous section.

Next, differentiating the density operator $\rho(t) = \sum_k p_k (a_k(t) \otimes a_k^*(t))$ and multiplying by i, we get the following; see, e.g., [25];

$$i\frac{d}{dt}\rho = i\frac{d}{dt}\sum_k p_k (a_k \otimes a_k^*) = \sum_k i(p_k \dot{a}_k \otimes a_k^* + p_k a_k \otimes \dot{a}_k^*)$$
$$= \sum_k H(p_k a_k \otimes a_k^*) - (p_k a_k \otimes a_k^*)H = H\rho - \rho H = [H, \rho] ,$$

that is the LvNM equation (2.85).

The LvNM equation (2.85) describes the time evolution of closed quantum spin systems, that are quantum systems which do not interact with the external environment. For the description of open systems, the LvNM equation (2.85) has to be modified by adding a double-commutator term as follows; see, e.g., [6, 25, 74, 82, 111];

$$\dot{\rho} = -i[H, \rho] - \sum_j [\tilde{V}_j, [\tilde{V}_j, \rho]] , \tag{2.87}$$

where $\tilde{V}_j$ are traceless matrices depending on the physics of the considered quantum systems. Equation (2.87) is also known as the Lindblad-Kossakowski master equation.

2.3.3 Controlled multi-spin Hamiltonian and rotating frame transformation

In the following, we consider finite-dimensional quantum systems composed by a set of non-interactive subsystems of N_p coupled spin-$\frac{1}{2}$ particles, for which the dimension of the corresponding density matrix is $N = 2^{N_p}$. We describe the Hamilton operator H for spin systems considered in NMR spectroscopy and discuss the so-called rotating frame transformation.

For these purposes, we construct a basis for the space $\mathfrak{her}(N)$ as described in the following. We consider the 2×2 identity matrix I_2 and the Pauli matrices I_x, I_y, I_z, given by (2.13), and we construct the matrices $I_{\alpha,k}$, for $\alpha = x, y, z, 2$ and $k = 1, \ldots, N_p$, as follows

$$I_{\alpha,k} := I_2 \otimes I_2 \otimes \cdots \otimes I_2 \otimes \underbrace{I_\alpha}_{k\text{-th position}} \otimes I_2 \otimes \cdots \otimes I_2 \ , \tag{2.88}$$

that is, $I_{\alpha,k}$ is obtained as the Kronecker product of $N_p - 1$ identity matrices I_2 and one of the matrices I_α. Particularly, the identity matrix appears everywhere except at the k-th position, which is occupied by I_α. By means of this construction, a basis for $\mathfrak{her}(N)$ is given by

$$\hat{\mathcal{B}}_{N_p} = \left\{ \hat{B}_{(\alpha_1, \ldots, \alpha_{N_p})} \mid \alpha_k \in \{x, y, z, 2\} \right\} \ , \tag{2.89}$$

where

$$\hat{B}_{(\alpha_1, \ldots, \alpha_{N_p})} = 2^{(q-1)} \prod_{k=1}^{N_p} I_{\alpha_k, k} \ , \tag{2.90}$$

where q is the number of indexes α_k different from 2; see, e.g., [70, 109]. Let us introduce the following inner product on $\mathfrak{her}(N)$

$$\langle A | B \rangle = \text{trace}(A^* B) \ , \text{ for } A, B \in \mathfrak{her}(N) \ . \tag{2.91}$$

This scalar product induces the so-called Hilbert-Schmidt norm

$$\|A\| = \sqrt{\langle A | A \rangle} = \sqrt{\text{trace}(A^* A)} \ , \text{ for } A \in \mathfrak{her}(N) \ . \tag{2.92}$$

Notice that $\hat{\mathcal{B}}_{N_p}$ is orthogonal with respect to the inner product (2.91). However,

$$\langle \hat{B}_k^* | \hat{B}_j \rangle = \delta_{k,j} 2^{N_p - 2} \ ,$$

where $\delta_{k,j}$ is the Kronecker delta, and the elements $\hat{B}_s$ have norm equal to $\sqrt{2^{N_p - 2}}$; see also [109]. We can normalize $\hat{\mathcal{B}}_{N_p}$ by multiplying all its elements by $\beta_{N_p} := \frac{1}{\sqrt{2^{N_p - 2}}}$, obtaining the following orthonormal basis

$$\tilde{\mathcal{B}}_{N_p} = \left\{ \tilde{B}_{(\alpha_1, \ldots, \alpha_{N_p})} := \beta_{N_p} \hat{B}_{(\alpha_1, \ldots, \alpha_{N_p})} \mid \alpha_k \in \{x, y, z, 2\} \right\} \ . \tag{2.93}$$

In the presence of external magnetic fields, a typical Hamiltonian has the following form [25, 46, 125]. We have

$$H = H_0 + H_{rf} \ , \tag{2.94}$$

where H_0 is the drift component related to a static magnetic field applied along the z-axis and H_{rf} represents radiofrequency electromagnetic radiations applied along the x and y axes and usually referred to as radiofrequency pulses. For a system of heteronuclear spins, that are spins having well-separated Larmor frequencies ν_k, see, e.g., [70],

$$|\nu_k - \nu_j| \gg |J_{k,j}| \ , \tag{2.95}$$

in a fixed laboratory coordinate frame, H_0 and H_{rf} take the following form [70, 125]

$$H_0 = \sum_{k=1}^{N_p} \nu_k I_{z,k} + \sum_{k<j} J_{k,j} I_{z,k} I_{z,j} \,, \tag{2.96}$$

and

$$H_{rf} = \sum_{k=1}^{N_p} V_k(t)\Big(\cos(\omega_k t + \phi_k(t))I_{x,k} + \sin(\omega_k t + \phi_k(t))I_{y,k}\Big) \,, \tag{2.97}$$

where $J_{k,j} \geq 0$ is the coupling constant of the Ising model between the k-th and j-th spin. Moreover, ω_k, ϕ_k, and V_k, represent frequencies, phases, and amplitudes of the external radiofrequency pulses, respectively. Notice that (2.97) is an approximation of the true radiofrequency Hamiltonian, and it is obtained neglecting the components rotating with frequencies $2\omega_k$, which are far from the resonance spectrum; see, e.g., [12, 108, 125].

Next, we discuss the so-called rotating-frame transformation. In NMR spectroscopy, because of the high values of frequencies of the static magnetic fields, it is convenient to represent the LvNM model in a coordinate system that "rotates" with frequencies similar or equal to the Larmor frequencies of the spin particles. In the literature, this rotating coordinate system is referred to as rotating frame coordinates. This representation allows a better description and understanding of the physical behaviour of the system. We also remark that, the rotating frame transformation represents a fundamental step for the numerical description and simulation of the NMR quantum spin systems: an adequate numerical description of a system rotating with a very high frequency could require a very fine mesh, and, consequently, a very high computational effort.

A general rotating frame transformation of the LvNM equation (2.85) is performed by means of a time-dependent unitary operator $\Upsilon(t)$, as follows [25, 108]

$$\tilde{\rho}(t) = \Upsilon(t)\rho(t)\Upsilon^*(t) \,. \tag{2.98}$$

Consequently, the LvNM equation assumes the following form

$$\dot{\tilde{\rho}} = -i[\tilde{H}, \tilde{\rho}] \,, \tag{2.99}$$

where $\tilde{H}$ is referred to as the effective Hamiltonian. The form of $\tilde{H}$ can be established as follows. Differentiating (2.98), we get

$$\begin{aligned}
\frac{d}{dt}\tilde{\rho} = \frac{d}{dt}\big(\Upsilon\rho\Upsilon^*\big) &= \Upsilon\frac{d}{dt}(\rho)\Upsilon^* + \frac{d}{dt}(\Upsilon)\rho\Upsilon^* + \Upsilon\rho\frac{d}{dt}(\Upsilon^*) \\
&= i\Upsilon[\rho, H]\Upsilon^* + \frac{d}{dt}(\Upsilon)\Upsilon^*\Upsilon\rho\Upsilon^* + \Upsilon\rho\Upsilon^*\Upsilon\frac{d}{dt}(\Upsilon^*) \\
&= i\Upsilon[\rho, H]\Upsilon^* + \frac{d}{dt}(\Upsilon)\Upsilon^*\tilde{\rho} + \tilde{\rho}\Upsilon\frac{d}{dt}(\Upsilon^*) \,.
\end{aligned} \tag{2.100}$$

Moreover, one has that

$$0 = \frac{d}{dt}\big(\Upsilon\Upsilon^*\big) = \frac{d\Upsilon}{dt}\Upsilon^* + \Upsilon\frac{d\Upsilon^*}{dt} \,, \tag{2.101}$$

which implies the following

$$\frac{d\Upsilon}{dt}\Upsilon^* = -\Upsilon\frac{d\Upsilon^*}{dt} \,. \tag{2.102}$$

Furthermore, one easily shows that

$$\Upsilon[\rho, H]\Upsilon^* = [\tilde{\rho}, \Upsilon H \Upsilon^*] . \tag{2.103}$$

Replacing (2.102) and (2.103) into (2.100), we obtain

$$
\begin{aligned}
\dot{\tilde{\rho}} &= -i\Upsilon[\rho, H]\Upsilon^* + \frac{d\Upsilon}{dt}\Upsilon^*\tilde{\rho} + \tilde{\rho}\Upsilon\frac{d\Upsilon^*}{dt} \\
&= i[\tilde{\rho}, \Upsilon H \Upsilon^*] - \Upsilon\frac{d\Upsilon^*}{dt}\tilde{\rho} + \tilde{\rho}\Upsilon\frac{d\Upsilon^*}{dt} \\
&= i[\tilde{\rho}, \Upsilon H \Upsilon^*] + \left[\tilde{\rho}, \Upsilon\frac{d\Upsilon^*}{dt}\right] \\
&= i\left[\tilde{\rho}, \Upsilon H \Upsilon^* - i\Upsilon\frac{d\Upsilon^*}{dt}\right] .
\end{aligned}
\tag{2.104}
$$

Hence, the effective Hamiltonian $\tilde{H}$ is given by the following

$$\tilde{H} = \Upsilon H \Upsilon^* - i\Upsilon\frac{d\Upsilon^*}{dt} . \tag{2.105}$$

In particular, if the transformation map is defined as $\Upsilon(t) := \prod_k \exp(-i\tilde{\omega}_k I_{z,k} t)$ [125], and H is given by (2.94), (2.96), and (2.97), in the rotating frame coordinates, the Hamiltonian becomes

$$\tilde{H} = \tilde{H}_0 + \tilde{H}_{rf} . \tag{2.106}$$

In this case, the drift Hamiltonian assumes the following form; see, e.g., [25, 46, 108];

$$\tilde{H}_0 = \sum_{k=1}^{N_p} (\nu_k - \tilde{\omega}_k) I_{z,k} + \sum_{k<j} J_{k,j} I_{z,k} I_{z,j} , \tag{2.107}$$

since the Ising Hamiltonian H_0 given by (2.96) commutes with $\Upsilon(t)$. Assuming that $\tilde{\omega}_k = \omega_k$, then the control term can be put in the following form [25, 72, 125]

$$\tilde{H}_{rf} = \sum_{k=1}^{N_p} (u_{x,k} I_{x,k} + u_{y,k} I_{y,k}) , \tag{2.108}$$

where $u_{x,k} : t \mapsto u_{x,k}(t) \in \mathbb{R}$ and $u_{y,k} : t \mapsto u_{y,k}(t) \in \mathbb{R}$ are regarded as control functions acting in the x- and y-direction of the k-th spin, respectively. We remark that, since the amplitudes V_k and the phases ϕ_k in (2.97) are time dependent and defined by the user [125], the control functions $u_{x,k}$ and $u_{y,k}$ are related to V_k and ϕ_k, by $u_{x,k}(t) = V_k(t)\cos(\phi_k(t))$ and $u_{y,k}(t) = V_k(t)\sin(\phi_k(t))$.

For the sake of clarity, we provide the following two examples regarding systems of 1 and 2 spin$-\frac{1}{2}$ particles, respectively.

Example.   *Case 1: one spin$-\frac{1}{2}$ system*
Consider the LvNM equation for one spin-$\frac{1}{2}$ given by $\dot{\rho} = -i[H, \rho]$, where the Hamilton operator is given by $H = H_0 + H_{rf}$ [6, 108, 111], with $H_0 = \nu_0 I_z$. Assuming that at high frequency fields the components rotating with frequency 2ω have negligible effects on the experiments [12, 108, 125], the control Hamiltonian is the following

$$H_{rf} = V(t)\big[\cos(\omega t + \phi(t))I_x + \sin(\omega t + \phi(t))I_y\big] ,$$

where ν_0 is the Larmor frequency of the spin, ω, ϕ and V are frequency, phase, and amplitude of the control radiofrequencies, respectively. We remark that the phase ϕ and the amplitude V are time dependent and chosen by the user [125].

To apply the rotating frame transformation, we consider the following unitary operator

$$\Upsilon = e^{i\tilde{\omega}tI_z} \ , \tag{2.109}$$

whose action defines a rotating frame reference which rotates around the z-direction at $\tilde{\omega}$ frequency. We first compute the transformed Pauli matrices as follows

$$\begin{aligned}
\Upsilon I_z \Upsilon^* &= I_z \ , \\
\Upsilon I_x \Upsilon^* &= \cos(\tilde{\omega}t)I_x - \sin(\tilde{\omega}t)I_y \ , \\
\Upsilon I_y \Upsilon^* &= \sin(\tilde{\omega}t)I_x + \cos(\tilde{\omega}t)I_y \ .
\end{aligned} \tag{2.110}$$

We also have that

$$\Upsilon \dot{\Upsilon}^* = -i\tilde{\omega}I_z \ . \tag{2.111}$$

The effective Hamiltonian is obtained as follows

$$\begin{aligned}
\tilde{H} &= \Upsilon H \Upsilon^* - i\Upsilon \dot{\Upsilon}^* \\
&= \nu_0 \Upsilon I_z \Upsilon^* + V(t)\cos(\omega t + \phi(t))\Upsilon I_x \Upsilon^* + V(t)\sin(\omega t + \phi(t))\Upsilon I_y \Upsilon^* - \tilde{\omega}I_z \\
&= (\nu_0 - \tilde{\omega})I_z + V(t)\cos(\omega t + \phi(t))(\cos(\tilde{\omega}t)I_x - \sin(\tilde{\omega}t)I_y) \\
&\quad + V(t)\sin(\omega t + \phi(t))(\sin(\tilde{\omega}t)I_x + \cos(\tilde{\omega}t)I_y) \\
&= (\nu_0 - \tilde{\omega})I_z + V(t)\big(\cos(\omega t)\cos(\phi(t)) - \sin(\omega t)\sin(\phi(t))\big)\big(\cos(\tilde{\omega}t)I_x - \sin(\tilde{\omega}t)I_y\big) \\
&\quad + V(t)\big(\sin(\omega t)\cos(\phi(t)) + \cos(\omega t)\sin(\phi(t))\big)\big(\sin(\tilde{\omega}t)I_x + \cos(\tilde{\omega}t)I_y\big) \ .
\end{aligned} \tag{2.112}$$

Choosing $\tilde{\omega} = \omega$, the previous expression simplifies as follows

$$\tilde{H} = (\nu_0 - \omega)I_z + V(t)\Big[\cos(\phi(t))I_x + \sin(\phi(t))I_y\Big] \ . \tag{2.113}$$

Now, define

$$u_1(t) := V(t)\cos(\phi(t)) \ ,$$

and

$$u_2(t) := V(t)\sin(\phi(t)) \ .$$

We obtain

$$\tilde{H} = (\nu_0 - \omega)I_z + u_1(t)I_x + u_2(t)I_y \ . \tag{2.114}$$

Notice that in NMR experiments, in order to obtain the resonance of the substance in analysis, ω is chosen equal to ν_0 and consequently, the drift term disappears. However, in practical situations, the nuclei in a molecule possess a range of chemical shifts. Because a radiofrequency can be applied at only one frequency ν_0, some nuclei will have resonant frequencies that are close to ν_0 while other nuclei will have resonant frequencies that are different from ν_0. Consequently, all nuclei cannot be expected to respond to the effect of a radiofrequency in an ideal way, hence, because of these characteristics of inhomogeneity, the drift sometimes does not vanish. $\triangle$

Example. *Case 2: 2 spin$-\frac{1}{2}$ systems*

In this case, we consider a system of two heteronuclear spins, I and S, such that their Larmor frequencies are well separated, that is $|\nu_I - \nu_S| \gg |J|$ [70], where ν_I and ν_S are the Larmor frequencies of the spin I and S, respectively, and J is the coupling constant. The Hamilton operator, in the laboratory frames and under the same assumption as in Case 1, has the following form

$$
\begin{aligned}
H = \nu_I \hat{I}_z + \nu_S \hat{S}_z + J \hat{I}_z \hat{S}_z \\
+ V_I(t) \Big[\cos(\omega_I t + \phi_I(t)) \hat{I}_x + \sin(\omega_I t + \phi_I(t)) \hat{I}_y \Big] \\
+ V_S(t) \Big[\cos(\omega_S t + \phi_S(t)) \hat{S}_x + \sin(\omega_S t + \phi_S(t)) \hat{S}_y \Big] \, ,
\end{aligned}
\tag{2.115}
$$

where $\hat{I}_\alpha = I_2 \otimes I_\alpha$ and $\hat{S}_\alpha = I_\alpha \otimes I_2$, for $\alpha = x, y, z$. To apply the rotating frame transformation, we consider the following unitary operator

$$
\Upsilon = e^{i\tilde{\omega}_I t \hat{I}_z} e^{i\tilde{\omega}_S t \hat{S}_z} \, ,
\tag{2.116}
$$

whose action defines two rotating frame references, one for each spin, rotating at frequencies $\tilde{\omega}_I$ and $\tilde{\omega}_S$, respectively. The following equalities hold

$$
\begin{aligned}
\Upsilon \hat{I}_z \hat{S}_z \Upsilon^* &= \hat{I}_z \hat{S}_z \\
\Upsilon \hat{I}_z \Upsilon^* &= \hat{I}_z \\
\Upsilon \hat{S}_z \Upsilon^* &= \hat{S}_z \\
\Upsilon \hat{I}_x \Upsilon^* &= \cos(\tilde{\omega}_I t) \hat{I}_x - \sin(\tilde{\omega}_I t) \hat{I}_y \\
\Upsilon \hat{I}_y \Upsilon^* &= \sin(\tilde{\omega}_I t) \hat{I}_x + \cos(\tilde{\omega}_I t) \hat{I}_y \\
\Upsilon \hat{S}_x \Upsilon^* &= \cos(\tilde{\omega}_S t) \hat{S}_x - \sin(\tilde{\omega}_S t) \hat{S}_y \\
\Upsilon \hat{S}_y \Upsilon^* &= \sin(\tilde{\omega}_S t) \hat{S}_x + \cos(\tilde{\omega}_S t) \hat{S}_y \, .
\end{aligned}
\tag{2.117}
$$

We also have that

$$
\Upsilon \dot{\Upsilon}^* = -i\tilde{\omega}_I \hat{I}_z - i\tilde{\omega}_S \hat{S}_z \, .
\tag{2.118}
$$

The effective Hamiltonian can be obtained as $\tilde{H} = \Upsilon H \Upsilon^* - i\Upsilon \dot{\Upsilon}^*$. Proceeding in the same way as in the previous case, and setting $\tilde{\omega}_I = \omega_I = \nu_I$ and $\tilde{\omega}_S = \omega_S = \nu_S$, we obtain that

$$
\tilde{H} = J \hat{I}_z \hat{S}_z + u_1(t) \hat{I}_x + u_2(t) \hat{I}_y + u_3(t) \hat{S}_x + u_4(t) \hat{S}_y \, ,
\tag{2.119}
$$

where

$$
u_1(t) := V_I(t) \cos(\phi_I(t)) \, ,
$$
$$
u_2(t) := V_I(t) \sin(\phi_I(t)) \, ,
$$
$$
u_3(t) := V_S(t) \cos(\phi_S(t)) \, ,
$$

and

$$
u_4(t) := V_S(t) \sin(\phi_S(t)) \, .
$$

$\triangle$

2.3.4 Control of inhomogeneous spin systems

In several applications in NMR, the considered ensemble of spins cannot be assumed to be physically homogeneous, in the sense that the radiofrequency pulses do not affect all the parts of the ensemble in the same way. This inhomogeneity can be modelled by dividing the ensemble in a finite number of groups, that are composed by the union of all the subsystems affected in the same way by the radiofrequency pulses. An example is shown in Figure 2.5, in which the ensemble of spins is divided into three groups (red, blue and green).
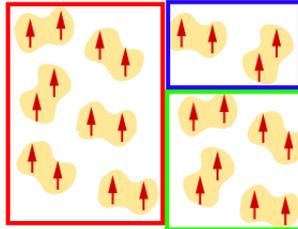

Figure 2.5: Inhomogeneous ensemble of spin divided into three groups.

Each of these groups is characterized by a density matrix and by a Hamilton operator. In particular, if we consider that the ensemble is divided into N_g number of groups, then the density operator corresponding to the entire ensemble is given by $\rho \in \mathfrak{her}(N_g N)$ with

$$\rho = \text{blk-diag}_{l=1,\dots,N_g} \left(\rho_l \right) . \tag{2.120}$$

The inhomogeneities are taken into account in the Hamiltonian by means of scaling parameters. The Hamiltonian corresponding to the l-th group is given as follows

$$\tilde{H}_l = \alpha_l \tilde{H}_0 + \hat{\alpha}_l \tilde{H}_{rf} , \tag{2.121}$$

where α_l and $\hat{\alpha}_l$ are real parameters, and $\tilde{H}_0$ and $\tilde{H}_{rf}$ are defined in (2.107) and (2.108), respectively. The scaling factors α_l and $\hat{\alpha}_l$ are used to perform the different action of the magnetic field on the ensemble. From the control point of view, this means that the same control functions are used for controlling spin systems characterized by different precession frequencies. Moreover, the control functions are scaled differently for each system.

2.4 Quantum systems with bilinear control structure

In quantum control theory, the controlled time evolution of a quantum system is in general described by means of a dynamical system having a bilinear control structure, as follows

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x , \tag{2.122}$$

where $x \in \mathbb{R}^{N_x}$ is a real representation of the quantum state, $u_n : [0,T] \to \mathbb{R}$ are the control functions, N_C is the number of controls, and A and B_n are skew-symmetric matrices in $\mathbb{R}^{N_x \times N_x}$. Equation (2.122) can be a real representation of a LvNM equation, semi-discrete Pauli and Schrödinger equations, quantum operator equation, etc.

In the next two sections, we show how to obtain (2.122) as a real representation of a LvNM equation and a semidiscrete Pauli or Schrödinger equation.

2.4.1 Real matrix representation of the LvNM equation

In this section, we discuss a real matrix representation of the LvNM equation (2.99), that allows to transform density operators into vectors in $\mathbb{R}^{N^2}$ with $N = 2^{N_p}$, that are called generalized Bloch vectors, and to represent the operators of the right-hand side of (2.99) by matrices in $\mathbb{R}^{N^2 \times N^2}$.

To this purpose, we consider the LvNM equation in the following form

$$\dot{\tilde{\rho}} = \mathcal{L}_{\tilde{H}}(\tilde{\rho}) \,, \tag{2.123}$$

where $\mathcal{L}_{\tilde{H}}(\tilde{\rho}) := -i[\tilde{H}, \tilde{\rho}]$, with

$$\tilde{H} = \tilde{H}_0 + \sum_{j=1}^{N_p} u_{x,j} \tilde{H}_{x,j} + u_{y,j} \tilde{H}_{y,j} \,. \tag{2.124}$$

Notice that (2.124) is a general form of the Hamiltonian given by (2.106)-(2.107)-(2.108).

We remark that (2.123) is in rotating-frame coordinates. To obtain the generalized Bloch vector, by means of the orthonormal basis $\tilde{\mathcal{B}}_{N_p}$, we fix an order of $\tilde{\mathcal{B}}_{N_p}$, i.e. $\tilde{\mathcal{B}}_{N_p} = \{\tilde{B}_1, \ldots, \tilde{B}_{N^2}\}$, and define the linear map $\mathcal{V} : \mathfrak{her}(N) \to \mathbb{R}^{N^2}$ as follows

$$\mathcal{V}(\tilde{\rho}) = \left(\langle \tilde{\rho} | \tilde{B}_1 \rangle, \langle \tilde{\rho} | \tilde{B}_2 \rangle, \ldots, \langle \tilde{\rho} | \tilde{B}_{N^2} \rangle \right)^T \,, \tag{2.125}$$

where $\langle \cdot | \cdot \rangle$ is the inner product defined in (2.91). By means of the above map, we get the following generalized Bloch vector

$$(x_1, x_2, \ldots, x_{N^2})^T = x := \mathcal{V}(\tilde{\rho}) \,, \tag{2.126}$$

and we remark that $\tilde{\rho} = \sum_{j=1}^{N^2} x_j \tilde{B}_j$.

Next, a matrix representation of $\mathcal{L}_{\tilde{H}}$, that is given explicitly by

$$\mathcal{L}_{\tilde{H}} = \mathcal{L}_{\tilde{H}_0} + \sum_{j=1}^{N_p} \left(u_{x,j} \mathcal{L}_{\tilde{H}_{x,j}} + u_{y,j} \mathcal{L}_{\tilde{H}_{y,j}} \right) \,,$$

with respect to $\tilde{\mathcal{B}}_{N_p}$, is given by $A_{\tilde{H}} \in \mathbb{R}^{N^2 \times N^2}$ written as follows

$$A_{\tilde{H}} = A + \sum_{j=1}^{N_p} \left(u_{x,j} A_{x,j} + u_{y,j} A_{y,j} \right) \,,$$

where A, $A_{x,j}$ and $A_{y,j}$ are the matrix representations of $\mathcal{L}_{\tilde{H}_0}$, $\mathcal{L}_{\tilde{H}_{x,j}}$, and $\mathcal{L}_{\tilde{H}_{y,j}}$, respectively. Their entries are obtained by the following scalar products

$$\langle \tilde{B}_k | \mathcal{L}_{\tilde{H}_0}(\tilde{B}_l) \rangle \,, \ \langle \tilde{B}_k | \mathcal{L}_{\tilde{H}_{x,j}}(\tilde{B}_l) \rangle \text{ and } \langle \tilde{B}_k | \mathcal{L}_{\tilde{H}_{y,j}}(\tilde{B}_l) \rangle \,. \tag{2.127}$$

For instance, denote $a_{k,l}$ the k, l entry of the matrix A. We can construct A by solving the following N^2 linear equations

$$\mathcal{L}_{\tilde{H}_0}(\tilde{B}_1) = a_{1,1} \tilde{B}_1 + a_{2,1} \tilde{B}_2 + a_{3,1} \tilde{B}_3 + \cdots + a_{N^2,1} \tilde{B}_{N^2}$$
$$\mathcal{L}_{\tilde{H}_0}(\tilde{B}_2) = a_{1,2} \tilde{B}_1 + a_{2,2} \tilde{B}_2 + a_{3,2} \tilde{B}_3 + \cdots + a_{N^2,2} \tilde{B}_{N^2}$$
$$\mathcal{L}_{\tilde{H}_0}(\tilde{B}_3) = a_{1,3} \tilde{B}_1 + a_{2,3} \tilde{B}_2 + a_{3,3} \tilde{B}_3 + \cdots + a_{N^2,3} \tilde{B}_{N^2}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\mathcal{L}_{\tilde{H}_0}(\tilde{B}_{N^2}) = a_{1,N^2} \tilde{B}_1 + a_{2,N^2} \tilde{B}_2 + a_{3,N^2} \tilde{B}_3 + \cdots + a_{N^2,N^2} \tilde{B}_{N^2} \,.$$

The solution to the previous linear system is given by $a_{k,l} = \langle \tilde{B}_k | \mathcal{L}_{\tilde{H}_0}(\tilde{B}_l) \rangle$.

This procedure results in a real matrix representation of the LvNM equation given by the bilinear dynamical system (2.122) with $N_C = 2N_p$ and $N_x = N^2 = 4^{N_p}$, and B_n and u_n are of the form $A_{\alpha,j}$ and $u_{\alpha,j}$, respectively, for some $\alpha = x, y$ and $j = 1, \ldots, N_p$. In particular, we order the control vector fields as $H_1 := \tilde{H}_{x,1}$, $H_2 := \tilde{H}_{y,1}$, $H_3 := \tilde{H}_{x,2}$, $H_4 := \tilde{H}_{y,2}$, $\ldots$, $H_{N_C-1} := \tilde{H}_{x,N_p}$, $H_{N_C} := \tilde{H}_{y,N_p}$. The drift A and the control terms B_n are skew-symmetric.

Notice that, formulas (2.127) are computationally very expensive, since all the entries of the matrices A and B_n are computed independently. For this reason, we describe the following alternative procedure. First, we compute a complex matrix representation of (2.99), defining the map $\mathcal{V}_{\mathbb{C}} : \mathfrak{her}(N) \to \mathbb{C}^{N^2}$ as follows

$$\mathcal{V}_{\mathbb{C}}(\tilde{\rho}) := vec(\tilde{\rho}),$$

that is, $\mathcal{V}_{\mathbb{C}}(\tilde{\rho})$ is a column vector containing sequentially the columns of $\tilde{\rho}$. Further, we define the following operators

$$A_{\mathbb{C}} := I_N \otimes H_0 - H_0^T \otimes I_N , \tag{2.128}$$

and

$$B_{\mathbb{C},n} := I_N \otimes H_n - H_n^T \otimes I_N , \tag{2.129}$$

where I_N is the $N \times N$ identity matrix. Consequently, the LvNM equation (2.99) becomes as follows

$$
\begin{aligned}
&\mathcal{V}_{\mathbb{C}}(\dot{\tilde{\rho}}) = \mathcal{V}_{\mathbb{C}}(-i[\tilde{H}, \tilde{\rho}]) \\
\Rightarrow & \frac{d}{dt} vec(\tilde{\rho}) = vec(-i[\tilde{H}, \tilde{\rho}]) \\
\Rightarrow & \frac{d}{dt} vec(\tilde{\rho}) = -i \, vec(\tilde{H}\tilde{\rho} - \tilde{\rho}\tilde{H}) \\
\Rightarrow & \frac{d}{dt} vec(\tilde{\rho}) = -i \, vec(\tilde{H}\tilde{\rho}I_N - I_N\tilde{\rho}\tilde{H}) \\
\Rightarrow & \frac{d}{dt} vec(\tilde{\rho}) = -i \left((I_N \otimes \tilde{H}) vec(\tilde{\rho}) - (\tilde{H}^T \otimes I_N) vec(\tilde{\rho}) \right) \\
\Rightarrow & \frac{d}{dt} vec(\tilde{\rho}) = -i \left(A_{\mathbb{C}} + \sum_{n=1}^{N_C} B_{\mathbb{C},n} \right) vec(\tilde{\rho}) ,
\end{aligned}
\tag{2.130}
$$

where we used the fact that $vec(DEF) = (F^T \otimes D) vec(E)$ for any three matrices D, E, F. Second, we construct the linear transformation operator T as follows

$$T = \left(T_1, \ldots, T_{N^2} \right) , \tag{2.131}$$

where $T_j = vec(\tilde{B}_j)$. Third, we get the real matrix representation as follows

$$
\begin{aligned}
\mathcal{V}(\rho) &= T^{-1} \mathcal{V}_{\mathbb{C}}(\rho) \\
A &= T^{-1} A_{\mathbb{C}} T \\
B_n &= T^{-1} B_{\mathbb{C},n} T .
\end{aligned}
\tag{2.132}
$$

Notice that the latter procedure, that first compute a complex representation and then transforms this into a real representation, is more efficient than the procedure given by (2.127), that computes directly a real representation. This is due to the fact that by

using (2.127) one has to compute the matrices A and B_n entry-by-entry, even if most of the entries are equal zero. This problem is avoided by passing through the complex representation, which allows a more compact computation, that can take into account the sparse structure of the matrices.

In the following, we show two examples of real matrix representation for two specific cases.

Example. *Case 1: one spin$-\frac{1}{2}$ system*
We want to obtain the real matrix representation of the following LvNM equation (in the rotating frame coordinates) corresponding to one spin$-\frac{1}{2}$ system, that is $\dot{\tilde{\rho}} = -i[\tilde{H}, \tilde{\rho}]$. According to the formulas (2.90) and (2.93), we have that $\tilde{B}_1 = \frac{\sqrt{2}}{2}I_2$, $\tilde{B}_2 = \sqrt{2}I_x$, $\tilde{B}_3 = \sqrt{2}I_y$, $\tilde{B}_4 = \sqrt{2}I_z$. By applying the procedure described in the previous section, we obtain the following real-matrix representation

$$\dot{x} = \left[A + u_1 B_1 + u_2 B_2\right]x \,,$$

where

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } B_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

$\triangle$

Example. *Case 2: 2 spin$-\frac{1}{2}$ system*
We want to obtain the real matrix representation of the LvNM equation $\dot{\tilde{\rho}} = -i[\tilde{H}, \tilde{\rho}]$ corresponding to two spin$-\frac{1}{2}$ system. In the examples in Section 2.3.3, we obtain that the effective Hamiltonian is given by

$$\tilde{H} = J\hat{I}_z\hat{S}_z + u_1\hat{I}_x + u_2\hat{I}_y + u_3\hat{S}_x + u_4\hat{S}_y \,.$$

According to the formulas (2.90) and (2.93), we have $\tilde{B}_1 = \frac{1}{2}I_2$, $\tilde{B}_2 = I_{x,2}$, $\tilde{B}_3 = I_{y,2}$, $\tilde{B}_4 = I_{z,2}$, $\tilde{B}_5 = I_{x,1}$, $\tilde{B}_6 = 2I_{x,1}I_{x,2}$, $\tilde{B}_7 = 2I_{x,1}I_{y,2}$, $\tilde{B}_7 = 2I_{x,1}I_{z,2}$, $\tilde{B}_9 = I_{y,1}$, $\tilde{B}_{10} = 2I_{y,1}I_{x,2}$, $\tilde{B}_{11} = 2I_{y,1}I_{y,2}$, $\tilde{B}_{12} = 2I_{y,1}I_{z,2}$, $\tilde{B}_{13} = I_{z,1}$, $\tilde{B}_{14} = 2I_{z,1}I_{x,2}$, $\tilde{B}_{15} = 2I_{z,1}I_{y,2}$, $\tilde{B}_{16} = 2I_{z,1}I_{z,2}$. By applying the procedure previously described, we obtain the following real matrix representation

$$\dot{x} = \left[A + u_1 B_1 + u_2 B_2 + u_3 B_3 + u_4 B_4\right]x \,.$$

where the obtained skew-symmetric matrices A and B_n are in $\mathbb{R}^{16 \times 16}$ and have the structures shown in the following figures, in which the small black dots denote the value 0, whereas the red dots and the blue dots denote positive and negative values, respectively.
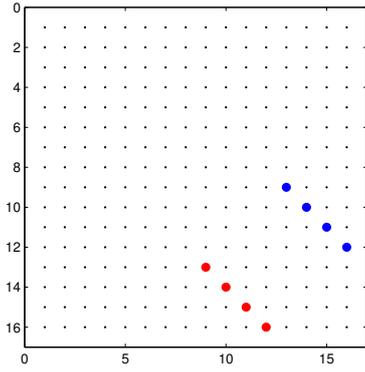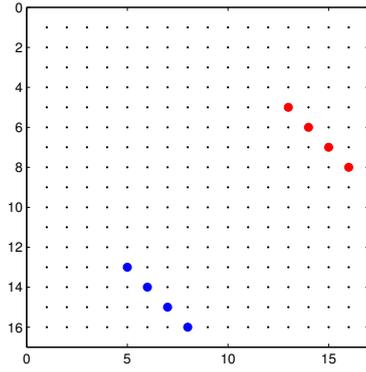
Figure 2.6: Structure of the matrix B_1.

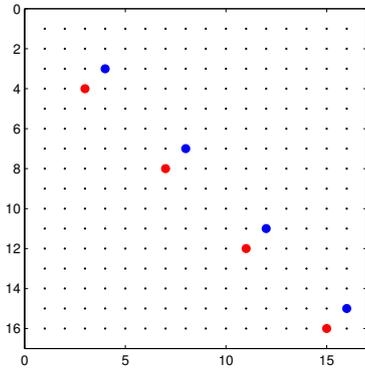

Figure 2.7: Structure of the matrix B_2.
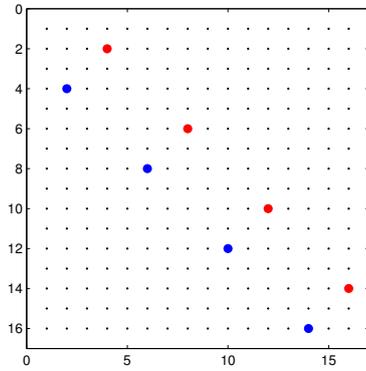

Figure 2.8: Structure of the matrix B_3.
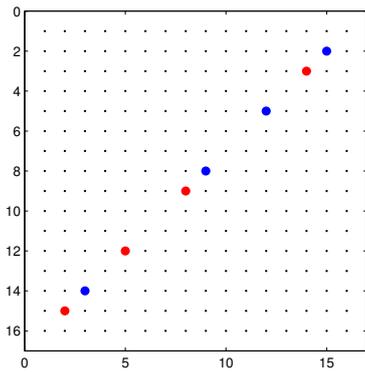

Figure 2.9: Structure of the matrix B_4.

Figure 2.10: Structure of the matrix A.

$\triangle$

2.4.2 Real representation of semidiscrete Pauli and Schrödinger equations

Consider a semi-discrete Schrödinger equation, that is a Schrödinger equation where the space operators are approximated by means of a discretization method, e.g., finite elements, spectral methods, etc. In a general form a semi-discrete Schrödinger equation can be written as follows

$$i\dot{\psi} = \big(H_0 + u_1 H_1 + i u_2 H_2\big)\psi \,, \tag{2.133}$$

where $\psi(t) \in \mathbb{C}^k$ is the discrete wavefunction, the integer k depends on the considered approximation, H_0 is a free Hamiltonian, $u_1 H_1$ and $u_2 H_2$ are control Hamiltonians with control functions $u_1, u_2 : [0, T] \to \mathbb{R}$. The matrices H_0 and H_1 are assumed to be symmetric and H_2 is assumed skew-symmetric. Notice that, for instance, equation (2.133) can represent

- the semi-discrete Pauli equation defined in (2.60), with $\psi = a$, H_0, $H_1 = H_x$, and $i H_2 = H_y$ defined by (2.61), (2.62) and (2.63), respectively, and $u_1 = B_x$ and $u_2 = B_y$;

- the semi-discrete Schrödinger equations used for the dipole control of a charged particle; see, e.g., [129, 130]. In this case H_0 is the discrete Laplace operator, $H_2 = 0$, H_1 is a diagonal matrix that represents the electric field, and u_1 is the control function, representing the time-dependent amplitude of the electric field.

By splitting real and imaginary parts of ψ, and defining $x \in \mathbb{R}^{N_x}$, with $N_x = 2k$, and $x := (\psi_{\mathbb{R}}, \psi_{\mathbb{I}})^T$, equation (2.133) becomes as follows

$$\dot{x} = \big(A + u_1 B_1 + u_2 B_2\big)x \,, \tag{2.134}$$

where

$$A = \begin{pmatrix} 0 & H_0 \\ -H_0 & 0 \end{pmatrix} \quad , \quad B_1 = \begin{pmatrix} 0 & H_1 \\ -H_1 & 0 \end{pmatrix} \quad \text{and} \quad B_2 = \begin{pmatrix} H_2 & 0 \\ 0 & H_2 \end{pmatrix} \,. \tag{2.135}$$

Since H_0 and H_1 are symmetric, and H_2 is skew-symmetric, A, B_1, and B_2 are skew-symmetric. Hence, the semi-discrete form (2.133) of an infinite-dimensional Schrödinger equation has a real representation with the same structure of (2.122).

2.5 Summary and remarks

The aim of this chapter was to describe the physics and the models that constitute the background that motivates optimal control and exact-control problems discussed in the thesis. The Pauli equation was discussed and used to introduce the Liouville-von Neumann master equation. These equations represent the main models for the description of the evolution of quantum spin systems. Both Pauli and LvNM equations admit real representations that belong to the class of dynamical systems with bilinear control structure, that are the focus of this thesis.

Chapter 3

Theoretical results about quantum optimal control problems

In this chapter, theoretical results concerning quantum optimal control problems are investigated. Existence and characterization properties of solutions and optimality conditions are proved for the following class of optimal control problems

$$
\begin{aligned}
\min_{x,u} \quad & J(x,u) := \frac{1}{2}\|x(T) - x_T\|_2^2 + \sum_{n=1}^{N_C}\left[\frac{\nu}{2}\|u_n\|_{L^2}^2 + \beta\|u_n\|_{L^1}\right] \\
\text{s.t.} \quad & \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]x \ , \ \text{in} \ (0,T) \\
& x \in X \ , \ u \in U_{ad} \subseteq L^2((0,T);\mathbb{R}^{N_C}) \ ,
\end{aligned} \tag{$\mathcal{P}$}
$$

where x is referred to as the state of the system, and belongs to the following set

$$
X := \left\{y \in H^1((0,T);\mathbb{R}^{N_x}) \ : \ y(0) = x_0\right\} , \tag{3.1}
$$

$x_0 \in \mathbb{R}^{N_x}$ is a given initial state, $u_n : [0,T] \to \mathbb{R}$ represent the control functions belonging to the admissible set U_{ad}, and N_C is the number of controls. The matrices A and B_n have real entries and their structure depends on the considered quantum system; in general, especially in NMR applications, B_n are skew-symmetric, and A is skew symmetric for closed systems and can have non-zero diagonal terms for open systems [6, 25, 111]; see also Section 2.4 and references therein. The vector $x_T \in \mathbb{R}^{N_x}$ is a given target state, and we assume that $\|x_T\|_2 = \|x_0\|_2$. This is a common situation in quantum control problem, in which in general the goal is to steer the quantum system from one eigenstate to another, and all the eigenstates are normalized.

In this work, we consider the following two admissible sets

$$
U_{ad,1} := \left\{v = (v_1,\ldots,v_{N_C}) \in L^2((0,T);\mathbb{R}^{N_C}) : |v_n(t)| \leq b \text{ a.e. in } (0,T) \text{ for } n = 1,\ldots,N_C\right\} , \tag{3.2}
$$

and

$$
U_{ad,2} := \left\{v \in L^2((0,T);\mathbb{R}^{N_C}) \ : \ \|v(t)\|_2 \leq b \text{ a.e. in } (0,T)\right\} , \tag{3.3}
$$

where b is a positive constant. We remark that these sets are considered in several quantum control applications, see, e.g., [9, 74, 75, 105]. In this chapter, we analyse the following optimal control cases

- Problem ($\mathcal{P}$) with $\beta = 0$, and both $U_{ad,1}$ and $U_{ad,2}$; this is an L^2-regularized optimal control problem which is considered in several quantum control applications; it is discussed in Section 3.2;

- Problem ($\mathcal{P}$) with $\beta = 0$, and piecewise constant controls belonging to $U_{ad,1}$; this optimal control problem is considered when the need of an exact implementation of the control functions in specific laboratory pulse shapers arises; see, e.g., [37]; this problem is discussed in Section 3.3;

- Problem ($\mathcal{P}$) with $\beta \neq 0$, and controls belonging to $U_{ad,1}$; this optimal control problem allows to generate control functions that are said to be "sparse", and that resembles the "pulsed shaped" controls often used in quantum control applications, like NMR and MRI; it is discussed in Section 3.4.

In our notation, given $k \in \mathbb{N}$, $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product in $\mathbb{R}^k$ and $\| \cdot \|_2$ the corresponding norm. We denote by $\langle \cdot, \cdot \rangle_{L^2}$ the L^2-inner product defined by

$$\langle y, z \rangle_{L^2} := \int_0^T \langle y(t), z(t) \rangle dt \ , \ \text{for every } y, z \in L^2((0,T); \mathbb{R}^k) \ .$$

The corresponding norm is denoted by $\| \cdot \|_{L^2}$. We denote by $\| \cdot \|_{L^q}$ the following norm

$$\|v\|_{L^q} := \left(\int_0^T \sum_{j=1}^k |v_j(t)|^q dt \right)^{\frac{1}{q}} \ , \ \text{for every } v \in L^q((0,T); \mathbb{R}^k) \ ,$$

for $1 \leq q < \infty$. Further, we denote by $\| \cdot \|_{\mathcal{L}} : \mathcal{L} \to \mathbb{R}$ the Hilbert-Schmidt norm, where $\mathcal{L}$ is the space of all $\mathbb{R}^{k \times k}$ matrices. Moreover, the notation $\sum_n = \sum_{n=1}^{N_C}$ is often used.

The chapter is organized as follows. In Section 3.1, main properties of the bilinear control system and its linearization are investigated. In Section 3.2, existence of solutions to L^2-quantum optimal control problems is proved, and optimality conditions are derived. Properties of the adjoint backward equation are also discussed. Section 3.3 investigates an L^2-quantum optimal control problem with piecewise-constant controls. In Section 3.4, L^1-quantum optimal control problems are investigated. Existence of and characterization properties of the optimal control solutions and optimality conditions are discussed. A summary section concludes this chapter.

3.1 Properties of bilinear quantum systems

In this section, we show some important properties of the following initial value problem

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \ \text{in } (0,T] \ , \ x(0) = x_0 \ , \tag{3.4}$$

and of its linearization, that is

$$\dot{\delta x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta x + \left[\sum_{n=1}^{N_C} \delta u_n B_n \right] x \ , \ \text{in } (0,T] \ , \ \delta x(0) = 0 \ . \tag{3.5}$$

We discuss existence, uniqueness and regularity properties of solutions to (3.4) and (3.5), and we show that the dynamics generated by (3.4) is norm preserving. In Proposition 1

we show existence and uniqueness of solutions to (3.4) and (3.5). In Proposition 2, we show that the dynamics of (3.4) is norm-preserving. Proposition 3 proves weak continuity of the map $u \mapsto x(u)$. In Proposition 4 and Proposition 5, we prove relationships between solutions to (3.4) and (3.5) evaluated pointwise in time and the L^1-norm of the controls. Proposition 6 shows a possible expansion of $u \mapsto x(u)$. In Proposition 7 we prove relationships between solutions to (3.4) and (3.5) evaluated pointwise in time and the L^2-norm of the controls. In Proposition 8, we investigate regularity properties of (3.4) and (3.5) written in operator forms.

First, we discuss existence and uniqueness of solutions to (3.4) and (3.5). Notice that, since the control functions u have to belong to a subset U_{ad} of the Lebesgue space $L^2((0,T); \mathbb{R}^{N_C})$, classical results, that requires that the right-hand side of the differential equation is continuous, are not valid. For this reason we seek existence and uniqueness of a weak-solution in the sense of Charathéodory, which requires the right-hand side to be measurable; see, e.g., [107, 131]. We have the following result.

Proposition 1. *Consider problems (3.4) and (3.5) with $x, \delta x \in H^1((0,T); \mathbb{R}^{N_x})$ and given $u, \delta u \in L^2((0,T); \mathbb{R}^{N_C})$. Then (3.4) and (3.5) admit unique solutions for any $T > 0$ and any initial condition.*

Proof. Let u be a given control function. Let us define $f_1 : \mathbb{R}^{N_x} \times \mathbb{R} \to \mathbb{R}^{N_x}$ as $f_1(x,t) := \left[A + \sum_n u_n B_n(t) \right] x$. Since $u \in L^2((0,T); \mathbb{R}^{N_C})$, then $f_1(x,t)$ is measurable for $t \in [0,T]$; furthermore it is linear and continuous with respect to x. Now, take $x, y \in \mathbb{R}^{N_x}$, we write that

$$\left\| f_1(x,t) - f_1(y,t) \right\|_2 = \left\| \left[A + \sum_n u_n(t) B_n \right](x-y) \right\|_2 \leq \left\| \left[A + \sum_n u_n(t) B_n \right] \right\|_{\mathcal{L}} \left\| x - y \right\|_2$$
$$\leq \left(\left\| A \right\|_{\mathcal{L}} + \sum_n |u_n(t)| \left\| B_n \right\|_{\mathcal{L}} \right) \left\| x - y \right\|_2 = \alpha(t) \left\| x - y \right\|_2 \, ,$$

where $\alpha(t) := \left\| A \right\|_{\mathcal{L}} + \sum_{n=1}^{N_C} |u_n(t)| \left\| B_n \right\|_{\mathcal{L}}$ is in $L^1((0,T); \mathbb{R})$. The previous inequality means that f_1 is Lipschitz in x. Moreover, for a given $x \in \mathbb{R}^{N_x}$ we have that

$$\left\| f_1(x,t) \right\|_2 = \left\| \left[A + \sum_n u_n(t) B_n \right] x \right\|_2 \leq \left\| \left[A + \sum_n u_n(t) B_n \right] \right\|_{\mathcal{L}} \left\| x \right\|_2$$
$$= \beta(t) \left\| x \right\|_2 \, , \ \forall t \in (0,T] \, ,$$

where $\beta(t)$ is equal to $\alpha(t)$. Hence, by [107] (Theorem 54), (3.4) admits a unique solution for any T and any initial condition; see, also [131].

Existence and uniqueness of solutions to (3.5) can be proved in the same way by considering $f_2(\delta x, t) := \left[A + \sum_n u_n(t) B_n \right] \delta x + \left[\sum_n \delta u_n(t) B_n \right] x$. $\qquad \square$

In the next proposition, we prove norm-preservation of the dynamical system (3.4), that is

$$\| x(t) \|_2 = \| x(0) \|_2 \, , \tag{3.6}$$

for $t \in [0,T]$. Notice that, to prove this property we need to assume that A and B_n are skew-symmetric matrices. This assumption is always satisfied in quantum control applications involving closed systems; see, e.g., Section 2.4 and references therein. Furthermore, the norm-preservation property is usual in quantum mechanics for closed systems, where the square of a wavefunction has a probabilistic meaning, and its L^2-norm is required to be constant, see, e.g., [83].

Proposition 2. *Assume that A, B_n, $n = 1, \ldots, N_C$ are skew-symmetric matrices. Then, the dynamics generated by (3.4) is norm-preserving.*

Proof. By multiplying the dynamical system in (3.4) from the left with x, we obtain that

$$\langle x, \dot{x} \rangle = \langle x, \big[A + \sum_n u_n B_n\big] x \rangle \ .$$

Now, notice that $\langle x, \dot{x} \rangle = \frac{1}{2} \frac{d}{dt} \|x\|_2^2$. Since A and B_n are assumed to be skew symmetric, we have that

$$\langle x, \big[A + \sum_n u_n B_n\big] x \rangle = 0 \ ,$$

which implies that

$$\frac{1}{2} \frac{d}{dt} \|x\|_2^2 = 0 \ .$$

This means that the dynamics generated by (3.4) is norm-preserving. $\qquad \square$

Next, the map $u \mapsto x(u)$ is analysed. Proposition 3 states the sequentially weak continuity of the map $u \mapsto x(u)$. Notice that a similar result can be found in [114], which is proved by means of the Ascoli-Arzelà theorem. Proposition 4 proves a relationship between the control u and the pointwise solution to (3.4).

Proposition 3. *Let $\{u^k\}_{k=1}^{\infty}$ be a sequence of controls such that*

$$u^k \rightharpoonup \hat{u}$$

in $L^2((0,T); \mathbb{R}^{N_C})$. Then the corresponding solutions of (3.4) $x^k := x(u^k)$ satisfy

$$x^k \to \hat{x} = x(\hat{u})$$

in $C([0,T]; \mathbb{R}^{N_x})$.

Proof. Consider a sequence of controls $\{u^k = (u_1^k, \ldots, u_{N_C}^k)\}_k$ in $L^2((0,T); \mathbb{R}^{N_C})$ such that $u_n^k \rightharpoonup \hat{u}_n$ in $L^2(0,T)$ as $k \to \infty$. The sequence $\{x^k\}_k$ in $H^1((0,T); \mathbb{R}^{N_x})$, defined as $x^k = x(u^k)$, is bounded by Proposition 2. Hence, since $H^1(0,T)$ is reflexive, we can extract a weakly convergent subsequence, i.e. $x^{k_j} \rightharpoonup \hat{x}$ in $H^1((0,T); \mathbb{R}^{N_x})$ as $j \to \infty$. Notice that the embedding $H^1(0,T) \to C[0,T]$ is compact, see, Theorem A 10 in the appendix, hence $x^{k_j} \to \hat{x}$ in $C([0,T]; \mathbb{R}^{N_x})$.

Now, consider the dynamical system corresponding to (u^{k_j}, x^{k_j}). By multiplying from the right with a test function $v \in H^1((0,T); \mathbb{R}^{N_x})$ and integrating over the interval $(0,T)$, we obtain

$$\int_0^T \langle \dot{x}^{k_j} - \big[A + \sum_n u_n^{k_j} B_n\big] x^{k_j}, v \rangle dt = 0 \ .$$

Since $u_n^k \rightharpoonup \hat{u}_n$ in $L^2(0,T)$, $x^{k_j} \to \hat{x}$ in $C([0,T]; \mathbb{R}^{N_x})$ and $\dot{x}^{k_j} \rightharpoonup \dot{\hat{x}}$ in $L^2((0,T); \mathbb{R}^{N_x})$ we have that

$$\int_0^T \langle \dot{x}^{k_j} - \big[A + \sum_n u_n^{k_j} B_n\big] x^{k_j}, v \rangle dt \to \int_0^T \langle \dot{\hat{x}} - \big[A + \sum_n \hat{u}_n B_n\big] \hat{x}, v \rangle dt \ ,$$

for all $v \in H^1((0,T); \mathbb{R}^{N_x})$. Since the above limit is true for all subsequences, and the limit $x(\hat{u})$ is unique, we have that $x(u^k) \to \hat{x} = x(\hat{u})$ in $C([0,T]; \mathbb{R}^{N_x})$. $\qquad \square$

Proposition 4. *Let v and w in $L^2((0,T);\mathbb{R}^{N_C})$ be two control functions, and consider the corresponding solutions to (3.4), $y = x(v)$ and $z = x(w)$. Assume that A and B_n are skew-symmetric. Then there exists a positive constant c_1 such that the following holds*

$$\|y(t) - z(t)\|_2^2 \leq c_1 \|v - w\|_{L^1} . \tag{3.7}$$

Moreover, if v and w are in $L^\infty((0,T);\mathbb{R}^{N_C})$, then there exists a positive constant $\tilde{c}_1$ such that the following holds

$$\|y(t) - z(t)\|_2 \leq \tilde{c}_1 \|v - w\|_{L^1} . \tag{3.8}$$

Proof. We define $\Delta x(t) := y(t) - z(t)$ a.e. in (0,T), which satisfies the following initial-value problem

$$\frac{d}{dt}\Delta x = A\Delta x + \sum_n \left(v_n B_n y - w_n B_n z \right) , \quad \Delta x(0) = 0 . \tag{3.9}$$

Taking the inner product of (3.9) with Δx, we obtain

$$
\begin{aligned}
\langle \Delta x, \frac{d}{dt}\Delta x \rangle &= \langle \Delta x, A\Delta x \rangle + \langle \Delta x, \sum_n \left(v_n B_n y - w_n B_n z \right) \rangle \\
&= \sum_n \langle \Delta x, v_n B_n y - w_n B_n z \rangle = \sum_n \langle y - z, v_n B_n y - w_n B_n z \rangle \\
&= \sum_n \left(-\langle y, v_n B_n z \rangle - \langle z, v_n B_n y \rangle \right) = \sum_n (v_n - w_n)\langle y, B_n z \rangle \\
&\leq \sum_n |v_n - w_n| \|B_n\|_{\mathcal{L}} \|x(0)\|_2^2 ,
\end{aligned}
\tag{3.10}
$$

where we used the skew-symmetry of A and B_n and the norm-preserving property of (3.4). Equation (3.10) is understood pointwise in time. Recalling that $\langle \Delta x(t), \frac{d}{dt}\Delta x(t) \rangle = \frac{1}{2}\frac{d}{dt}\|\Delta x(t)\|_2^2$, integrating (3.10) over $(0,t)$, and defining $K_0 = \max_n \|B_n\|_{\mathcal{L}}$, we obtain the following

$$
\begin{aligned}
\|y(t) - z(t)\|_2^2 &\leq 2K_0 \|x(0)\|_2^2 \sum_n \int_0^t |v_n(s) - w_n(s)| ds \\
&\leq 2K_0 \|x_0\|_2^2 \sum_n \|v_n - w_n\|_{L^1} = 2K_0 \|x_0\|_2^2 \|v - w\|_{L^1} ,
\end{aligned}
\tag{3.11}
$$

which is (3.7) with $c_1 = 2K_0 \|x_0\|_2^2$.

To prove estimate (3.8), we proceed in a similar way used in [100]. To this purpose, notice that $y = x(v)$ can be written as

$$y(t) = e^{At}x_0 + \int_0^t \sum_n v_n(s) e^{A(t-s)} B_n y(s) ds , \tag{3.12}$$

and the same for $z = x(w)$. Consequently, we write

$$
\begin{aligned}
y(t) - z(t) &= \sum_n \int_0^t v_n(s) e^{A(t-s)} B_n y(s) - w_n(s) e^{A(t-s)} B_n z(s) ds \\
&= \sum_n \int_0^t e^{A(t-s)} \left[v_n(s) B_n y(s) - w_n(s) B_n z(s) \right] ds \\
&= \sum_n \int_0^t e^{A(t-s)} \left[v_n(s) B_n y(s) - w_n(s) B_n z(s) + w_n(s) B_n y(s) - w_n(s) B_n y(s) \right] ds \\
&= \sum_n \int_0^t e^{A(t-s)} \left[(v_n(s) - w_n(s)) B_n y(s) + w_n(s) B_n (y(s) - z(s)) \right] ds .
\end{aligned}
\tag{3.13}
$$

Equation (3.13) implies the following

$$\|y(t) - z(t)\|_2 \le K_1\|v - w\|_{L^1} + K_2\int_0^t \|y(s) - z(s)\|_2 ds \,, \tag{3.14}$$

where $K_1 = \|e^A\|_{\mathcal{L}}K_0\|x_0\|_2$ and $K_2 = \|e^A\|_{\mathcal{L}}K_0 N_C W$, with $W = \max_n\big(\text{ess sup } |w_n|\big)$. Notice that W is bounded because w is in $L^\infty((0,T);\mathbb{R}^{N_C})$. Moreover, we use that $\|e^A\|_{\mathcal{L}} = \|e^{As}\|_{\mathcal{L}}$ for any $s \in \mathbb{R}$, which holds because A is skew-symmetric and e^{As} are orthogonal matrices, that are bounded [53]. Next, by applying to (3.14) the Gronwall's inequality in the integral form, we obtain

$$\begin{aligned}
\|y(t) - z(t)\|_2 &\le K_1\exp\Big(\int_0^t K_2 ds\Big)\|v - w\|_{L^1} \\
&\le K_1\exp\big(K_2 T\big)\|v - w\|_{L^1} \,,
\end{aligned} \tag{3.15}$$

that is (3.8) with $\tilde{c}_1 = K_1\exp\big(K_2 T\big)$. $\qquad\square$

In Proposition 4, we assume that A and B_n are skew symmetric. According to Proposition 2, this implies that the dynamics is norm preserving, that is, $\|x(t)\|_2 = \|x(0)\|_2$ for all t in $[0, T]$. This property holds for closed quantum systems. Regarding open quantum systems, we have that in general B_n are skew-symmetric but A is not. However, open systems can be characterized by the property $\|x(t)\|_2 \le \|x(0)\|_2$, which is due to the fact that the drift matrix A is given by the sum of a skew-symmetric matrix and a negative semi-definite matrix; see, e.g., [111]; this implies that $\langle x, Ax \rangle \le 0$. Consequently, the proof of Proposition 4 remains valid.

We remark that, by noticing that $\|v - w\|_{L^1} \le \sqrt{N_C T}\|v - w\|_{L^2}$, it follows from (3.7) and (3.8) that there exist two positive constants c_2 and $\tilde{c}_2$, such that

$$\|y(t) - z(t)\|_2^2 \le c_2\|v - w\|_{L^2} \,, \text{a.e. in } (0,T) \quad \text{for any } v, w \in L^2((0,T);\mathbb{R}^{N_C}) \tag{3.16a}$$

$$\|y(t) - z(t)\|_2 \le \tilde{c}_2\|v - w\|_{L^2} \,, \text{a.e. in } (0,T) \quad \text{for any } v, w \in L^\infty((0,T);\mathbb{R}^{N_C}) \,, \tag{3.16b}$$

where y and z are as in Proposition 4. However, we can obtain better estimates than (3.16a) and (3.16b), in the sense that we can prove (3.16b) for any $v, w \in L^2((0,T);\mathbb{R}^{N_C})$. This result is presented in Proposition 7, and to prove it we need the results given in the two following propositions.

The next proposition shows a result similar to (3.16a) and (3.16b) for the linearized problem (3.5). In particular, we remark that we can prove the estimate (3.18), that is similar to (3.16b), without assuming pointwise-boundedness of u and δu.

Proposition 5. *Let $\delta u \in L^2((0,T);\mathbb{R}^{N_C})$ and let $\delta x = \delta x(\delta u)$ be the corresponding unique solution to (3.5). Assume that A and B_n are skew-symmetric. Then the following estimate holds*

$$\|\delta x\|_{L^2} \le 2T\sqrt{N_C}K_{00}\|x_0\|_2\|\delta u\|_{L^2} \,, \tag{3.17}$$

where $K_{00} := \sum_{n=1}^{N_C}\|B_n\|_{\mathcal{L}}$. Notice that inequality (3.17) means that the map $\delta u \mapsto \delta x(\delta u)$ is Lipschitz continuous in $\delta u = 0$. Moreover, it holds that

$$\|\delta x(t)\|_2 \le 2\sqrt{TN_C}K_{00}\|x_0\|_2\|\delta u\|_{L^2} \,, \tag{3.18}$$

a.e. in $(0,T)$.

Proof. Consider the linearized problem (3.5). By multiplying (3.5) to the left with δx, we obtain

$$\langle \delta x, \dot{\delta x} \rangle = \langle \delta x, \Big[A + \sum_n u_n B_n \Big] \delta x \rangle + \langle \delta x, \Big[\sum_n \delta u_n B_n \Big] x \rangle \,. \tag{3.19}$$

Now, considering that $\langle \delta x(t), \dot{\delta x}(t) \rangle = \frac{1}{2} \frac{d}{dt} \| \delta x(t) \|_2^2$ and recalling the skew-symmetry of A and B_n, we get

$$\frac{1}{2} \frac{d}{dt} \| \delta x(t) \|_2^2 = \langle \delta x, \Big[\sum_n \delta u_n B_n \Big] x \rangle \,. \tag{3.20}$$

Integrating (3.20) over $(0, t)$, and using that $\delta x(0) = 0$, the dynamics of (3.4) is norm preserving and the Cauchy-Schwarz inequality, we obtain

$$
\begin{aligned}
\| \delta x(t) \|_2^2 &= 2 \int_0^t \langle \delta x, \Big[\sum_n \delta u_n B_n \Big] x \rangle ds = 2 \sum_n \int_0^t \delta u_n \langle \delta x, B_n x \rangle ds \\
&\le 2 \sum_n \int_0^t | \delta u_n | | \langle \delta x, B_n x \rangle | ds \le 2 \sum_n \int_0^T | \delta u_n | | \langle \delta x, B_n x \rangle | ds \\
&\le 2 \sum_n \int_0^T | \delta u_n | \| \delta x \|_2 \| B_n x \|_2 ds \le 2 \sum_n \int_0^T | \delta u_n | \| \delta x \|_2 \| B_n \|_{\mathcal{L}} \| x(0) \|_2 ds \\
&\le 2 \| x(0) \|_2 \sum_n \| B_n \|_{\mathcal{L}} \int_0^T | \delta u_n | \| \delta x \|_2 ds \\
&= 2 \| x(0) \|_2 \sum_n \| B_n \|_{\mathcal{L}} \langle | \delta u_n |, \| \delta x \|_2 \rangle_{L^2} \\
&\le 2 \| x_0 \|_2 K_{00} \sum_n \langle | \delta u_n |, \| \delta x \|_2 \rangle_{L^2} \\
&\le 2 \| x_0 \|_2 K_{00} \sum_n \Big(\int_0^T | \delta u_n |^2 ds \Big)^{1/2} \Big(\int_0^T \| \delta x \|_2^2 ds \Big)^{1/2} \\
&\le 2 \| x_0 \|_2 K_{00} \sqrt{N_C} \| \delta u \|_{L^2} \| \delta x \|_{L^2} \,,
\end{aligned}
\tag{3.21}
$$

where $K_{00} = \sum_n \| B_n \|_{\mathcal{L}}$. Now, integrating (3.21) over $(0, T)$, we obtain (3.17) as follows

$$
\begin{aligned}
\int_0^T \| \delta x(t) \|_2^2 dt &\le 2 \int_0^T \| x_0 \|_2 \sqrt{N_C} K_{00} \| \delta u \|_{L^2} \| \delta x \|_{L^2} dt \\
&\Rightarrow \| \delta x \|_{L^2}^2 \le 2T \| x_0 \|_2 \sqrt{N_C} K_{00} \| \delta u \|_{L^2} \| \delta x \|_{L^2} \\
&\Rightarrow \| \delta x \|_{L^2} \le 2T \| x_0 \|_2 \sqrt{N_C} K_{00} \| \delta u \|_{L^2} \,,
\end{aligned}
\tag{3.22}
$$

that is (3.17). Next, by replacing (3.22) in (3.21), we obtain (3.18). $\qquad\square$

Proposition 6. *Let $u \in L^2((0,T); \mathbb{R}^{N_C})$ and $h \in L^2((0,T); \mathbb{R}^{N_C})$. Consider the map $u \mapsto x(u)$, where $x(u) \in H^1((0,T); \mathbb{R}^{N_x})$ is the unique solution to (3.4) corresponding to u, and denote by $\delta x(u, h) \in H^1((0,T); \mathbb{R}^{N_x})$ the unique solution to (3.5) corresponding to h and u. Assume that A and B_n are skew-symmetric. The following expansion holds*

$$x(u + h) = x(u) + \delta x(u, h) + \vartheta(u, h) \,, \tag{3.23}$$

where $\vartheta(u, h) \in H^1((0,T); \mathbb{R}^{N_x})$ solves the following problem

$$\dot{\vartheta} = \Big[A + \sum_{n=1}^{N_C} u_n B_n \Big] \vartheta + \sum_{n=1}^{N_C} h_n B_n \big(\vartheta + \delta x \big) \;\; in \;\; (0, T] \,, \;\; and \;\; \vartheta(0) = 0 \,. \tag{3.24}$$

Moreover, there exist positive constants c_ϑ and $\tilde{c}_\vartheta$ such that

$$\|\vartheta(t)\|_2 \le c_\vartheta \|h\|_{L^2}^2 \,, \ a.e. \ \ in \ (0,T) \tag{3.25a}$$

$$\|\vartheta\|_{L^2} \le \tilde{c}_\vartheta \|h\|_{L^2}^2 \,. \tag{3.25b}$$

Proof. Denote by $y = x(u)$ and $z = x(u+h)$ the solutions to

$$\dot{y} = \Big[A + \sum_n u_n B_n\Big] y \quad y(0) = x_0 \tag{3.26a}$$

$$\dot{z} = \Big[A + \sum_n (u_n + h_n) B_n\Big] z \quad z(0) = x_0 \,, \tag{3.26b}$$

respectively. Now, we define $\vartheta + \delta x := z - y$. By subtracting term-by-term (3.26a) from (3.26b), we notice that the sum $\vartheta + \delta x$ solves the following problem

$$\dot{\vartheta} + \dot{\delta x} = \Big[A + \sum_n u_n B_n\Big](\vartheta + \delta x) + \sum_n h_n B_n z \,, \tag{3.27}$$

with $\vartheta(0) = 0$. Using that $z = y + \delta x + \vartheta$, and rearranging (3.27) we obtain

$$\begin{aligned} &\Big[\dot{\vartheta} - \Big[A + \sum_n u_n B_n\Big]\vartheta - \sum_n h_n B_n\big(\vartheta + \delta x\big)\Big] \\ &+ \Big[\dot{\delta x} - \Big[A + \sum_n u_n B_n\Big]\delta x - \sum_n h_n B_n y\Big] = 0 \,. \end{aligned} \tag{3.28}$$

Now, notice that (3.28) is satisfied if $\delta x = \delta x(u,h)$ solves (3.5), and $\vartheta = \vartheta(u,h)$ solves (3.24). Hence, expansion (3.23) is satisfied.

The estimates (3.25a) and (3.25b) can be proved similarly to (3.17) and (3.18). In fact, by multiplying (3.24) to the left with ϑ and recalling that A and B_n are assumed to be skew-symmetric, we obtain

$$\langle \vartheta, \dot{\vartheta} \rangle = \langle \vartheta, \sum_n h_n B_n \delta x \rangle \,. \tag{3.29}$$

Now, considering that $\langle \vartheta(t), \dot{\vartheta}(t) \rangle = \frac{1}{2}\frac{d}{dt}\|\vartheta(t)\|_2^2$, we get

$$\frac{1}{2}\frac{d}{dt}\|\vartheta(t)\|_2^2 = \langle \vartheta, \sum_n h_n B_n \delta x \rangle \,. \tag{3.30}$$

Integrating (3.30) over $(0,t)$, and using that $\vartheta(0) = 0$, the dynamics of (3.4) is norm preserving and using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \|\vartheta(t)\|_2^2 &= 2\int_0^t \langle \vartheta, \sum_n h_n B_n \delta x \rangle ds = 2\sum_n \int_0^t h_n \langle \vartheta, B_n \delta x \rangle ds \\ &\le 2\sum_n \int_0^T |h_n| \|\vartheta\|_2 \|B_n \delta x\|_2 ds \le 2\sum_n \int_0^T |h_n| \|\vartheta\|_2 \|B_n\|_{\mathcal{L}} \|\delta x\|_2 ds \\ &\le 2\sum_n \|B_n\|_{\mathcal{L}} \int_0^T |h_n| \|\vartheta\|_2 \|\delta x\|_2 ds \\ &\le 2K_{00} \sum_n \int_0^T |h_n| \|\vartheta\|_2 \|\delta x\|_2 ds \,. \end{aligned} \tag{3.31}$$

where $K_{00} = \sum_n \|B_n\|_{\mathcal{L}}$. Now, we use (3.18) to obtain the following

$$
\begin{aligned}
\|\vartheta(t)\|_2^2 &\leq 4\sqrt{TN_C}K_{00}^2\|x_0\|_2\|h\|_{L^2}\sum_n\int_0^T |h_n|\|\vartheta\|_2 ds \\
&\leq 4\sqrt{TN_C}K_{00}^2\|x_0\|_2\|h\|_{L^2}\sum_n \|h_n\|_{L^2}\|\vartheta\|_{L^2} \\
&\leq 4\sqrt{T}N_C K_{00}^2\|x_0\|_2\|h\|_{L^2}^2\|\vartheta\|_{L^2} \ .
\end{aligned}
\tag{3.32}
$$

Integrating (3.32) over $(0,T)$, we get

$$
\|\vartheta\|_{L^2} \leq 4\sqrt{T}TN_C K_{00}^2\|x_0\|_2\|h\|_{L^2}^2 \ ,
\tag{3.33}
$$

that is (3.25b) with $\tilde{c}_\vartheta = 4\sqrt{T}TN_C K_{00}^2\|x_0\|_2$. By replacing (3.33) into (3.32), we obtain (3.25a) with $c_\vartheta = 4T^{3/2}K_{00}^2\|x_0\|_2 N_C$, and the claim follows. $\square$

The next proposition discusses relationships between the state x and the control u, and in particular, local Lipschitz continuity of the map $u \mapsto x(u)$ is proved.

Proposition 7. *Let v and w in $L^2((0,T);\mathbb{R}^{N_C})$ be two control functions, and consider the corresponding solutions to (3.4), $y = x(v)$ and $z = x(w)$. Assume that A and B_n are skew-symmetric. Then there exist positive constants $\hat{c}_1$ and $\hat{c}_2$ such that for any $\|v - w\|_{L^2} \leq 1$ the following holds*

$$
\|y(t) - z(t)\|_2 \leq \hat{c}_1\|v - w\|_{L^2} \ , a.e. \ in \ (0,T)
\tag{3.34a}
$$

$$
\|y - z\|_{L^2} \leq \hat{c}_2\|v - w\|_{L^2} \ .
\tag{3.34b}
$$

Proof. The claim follows from Proposition 6. In fact, consider $v = u$ and $w = u + h$, by means of (3.23) we have that

$$
\|y(t) - z(t)\|_2 = \|\delta x(t) - \vartheta(t)\|_2 \leq \|\delta x(t)\|_2 + \|\vartheta(t)\|_2 \ .
\tag{3.35}
$$

By using of (3.18) and (3.25a) for $\|h\|_{L^2} \leq 1$ we have that

$$
\|y(t) - z(t)\|_2 \leq \left(2\sqrt{TN_C}K_{00}\|x_0\|_2 + c_\vartheta\right)\|h\|_{L^2} \ ,
\tag{3.36}
$$

which implies the following

$$
\|y(t) - z(t)\|_2 \leq \hat{c}_1\|v - w\|_{L^2} \ ,
\tag{3.37}
$$

with $\hat{c}_1 = \left(2\sqrt{TN_C}K_{00}\|x_0\|_2 + c_\vartheta\right)$.

To obtain (3.34b), we take the square of left- and right-hand side of (3.37) and integrate over $(0,T)$. Consequently, the following holds

$$
\|y - z\|_{L^2}^2 \leq \hat{c}_1^2 T\|v - w\|_{L^2}^2 \ ,
\tag{3.38}
$$

and the claim follows. $\square$

Next, consider the operator $c(\cdot,\cdot): X \times U \to P$ defined as follows

$$
c(x,u) := \frac{d}{dt}x - \left[A + \sum_{n=1}^{N_C} u_n B_n\right]x \ ,
\tag{3.39}
$$

with $X = \left\{ x \in H^1((0, T); \mathbb{R}^{N_x}) \mid x(0) = x_0 \right\}$, $U = L^2((0, T); \mathbb{R}^{N_C})$, and $P = L^2((0, T); \mathbb{R}^{N_x})$. In this way, (3.4) and (3.5) can be equivalently written respectively as

$$c(x, u) = 0 \quad \text{and} \quad Dc(x, u)(\delta x, \delta u) = 0 , \tag{3.40}$$

where $(\delta x, \delta u) \mapsto Dc(x, u)(\delta x, \delta u)$ is defined as

$$Dc(x, u)(\delta x, \delta u) := \frac{d}{dt} \delta x - \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta x - \left[\sum_{n=1}^{N_C} \delta u_n B_n \right] x . \tag{3.41}$$

The following proposition shows that c is invertible for a fixed u, and that it is Fréchet differentiable with Fréchet derivative given by $Dc(x, u)$. Moreover, surjectivity of the map $\delta x \mapsto Dc(x, u)(\delta x, \delta u)$ is proved.

Proposition 8. *The operator c is Fréchet differentiable, and invertible for fixed u. Moreover, its Fréchet derivative $Dc(x, u)$ is also invertible for fixed δu and hence surjective.*

Proof. We start proving Fréchet differentiability of $(x, u) \mapsto c(x, u)$. To this purpose, notice that

$$c(x + \delta x, u + \delta u) = c(x, u) + M(\delta x, \delta u) - \sum_{n=1}^{N_C} \delta u_n B_n \delta x ,$$

where $M(\delta x, \delta u) := \dot{\delta x} - \left[A + \sum_n B_n u_n \right] \delta x - \left[\sum_n B_n \delta u_n \right] x$. Hence, we write that

$$\| c(x + \delta x, u + \delta u) - c(x, u) - M(\delta x, \delta u) \|_{L^2} = \left\| \sum_n \delta u_n B_n \delta x \right\|_{L^2}$$
$$\leq K \left(\| \delta u \|_{L^2} + \| \delta x \|_{L^2} \right)^\alpha$$

for some positive constants K and $\alpha > 1$. Now, we obtain

$$\lim_{\| \delta u \|_{L^2} + \| \delta x \|_{L^2} \to 0} \frac{\| c(x + \delta x, u + \delta u) - c(x, u) - M(\delta x, \delta u) \|_{L^2}}{\| \delta u \|_{L^2} + \| \delta x \|_{L^2}}$$
$$\leq \lim_{\| \delta u \|_{L^2} + \| \delta x \|_{L^2} \to 0} \frac{K \left(\| \delta u \|_{L^2} + \| \delta x \|_{L^2} \right)^\alpha}{\| \delta u \|_{L^2} + \| \delta x \|_{L^2}} = 0 ,$$

which means that $(x, u) \mapsto c(x, u)$ is Fréchet differentiable and $Dc(x, u) = M$ is the Fréchet derivative of c.

To prove that c is invertible for fixed u, we recall that (3.4) admits a unique solution and the same holds for $c(x, u) = b$ for any $b \in P$. Hence, c is invertible. Using the same argument, we obtain that $Dc(x, u)$ is invertible for fixed x, u, and δu, hence surjective. $\quad \square$

3.2 L^2-regularized optimal control problems

In this section, we consider the following optimal control problem

$$\min_{x, u} \quad J(x, u) := \frac{1}{2} \| x(T) - x_T \|_2^2 + \sum_{n=1}^{N_C} \frac{\nu}{2} \| u_n \|_{L^2}^2$$

$$\text{s.t.} \quad \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x , \text{ in } (0, T] \tag{3.42}$$

$$x \in X , \ u \in U_{ad} \subseteq L^2((0, T); \mathbb{R}^{N_C}) ,$$

where X is defined in (3.1), and the admissible set U_{ad} is considered to be equal first to $U_{ad,1}$, and then to $U_{ad,2}$, defined in (3.2) and (3.3), respectively.

This section is organized as follows. In Lemma 1, we show that $U_{ad,1}$ and $U_{ad,2}$ are bounded, convex and closed. This result is used in Theorem 1 to prove existence of an optimal control solution to (3.42). In Lemma 2, we derive the gradient of (3.42), and in Theorem 2, the first-order optimality system for (3.42) with $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$ is obtained. In Theorem 3, we discuss the first-order optimality system for (3.42) with $U_{ad} = U_{ad,1}$. In Theorem 4 and Corollary 1, we derive the first-order optimality system for (3.42) with $U_{ad} = U_{ad,2}$. At the end of the section, we discuss some properties of the adjoint equations, that is the equation which characterizes the Lagrange multiplier corresponding to the bilinear constraint.

Lemma 1. *The admissible sets $U_{ad,1}$ and $U_{ad,2}$ as subsets of $L^2((0,T); \mathbb{R}^{N_C})$ are bounded, convex and closed, and hence weakly sequentially compact.*

Proof. Weak compactness of $U_{ad,1}$ is proved in [96]; see, also, [119]. Next, we consider $U_{ad,2}$. Let $u \in U_{ad,2}$, then

$$\|u\|_{L^2}^2 = \int_0^T \|u\|_2^2 dt \le \int_0^T |b|^2 dt = b^2 T .$$

Thus $U_{ad,2} \subset L^2((0,T); \mathbb{R}^{N_C})$ is bounded. Next, let $v, w \in U_{ad,2}$ and $\alpha \in [0,1]$, we have

$$\|\alpha v(t) + (1-\alpha)w(t)\|_2 \le \alpha \|v(t)\|_2 + (1-\alpha)\|w(t)\|_2 \le \alpha b + (1-\alpha)b = b .$$

Hence $U_{ad,2}$ is convex. Next, we prove that $U_{ad,2}$ is closed. Take a sequence $\{u^k\}_k$ in $U_{ad,2}$, such that $u^k \to \hat{u}$ in $L^2((0,T); \mathbb{R}^{N_C})$, that is

$$\lim_{k \to \infty} \|u^k - \hat{u}\|_{L^2} = 0 , \tag{3.43}$$

and seeking a contradiction we assume that $\hat{u} \notin U_{ad,2}$, that is there exists a positive ϵ and a measurable set I with measure $\mu(I) > 0$ such that $\|\hat{u}(t)\|_2 > b + \epsilon$ for $t \in I$. Hence, we write that

$$\|\hat{u}(t) - u^k(t)\|_2 \ge \left| \|\hat{u}(t)\|_2 - \|u^k(t)\|_2 \right| \ge \|\hat{u}(t)\|_2 - \|u^k(t)\|_2 > b + \epsilon - b = \epsilon , \tag{3.44}$$

for $t \in I$. Next, we write that

$$\|\hat{u} - u^k\|_{L^2}^2 = \int_0^T \|\hat{u}(t) - u^k(t)\|_2^2 dt \ge \int_I \|\hat{u}(t) - u^k(t)\|_2^2 dt > \int_I \epsilon^2 dt = \mu(I)\epsilon^2 > 0 ,$$

which contradicts (3.43). Hence $U_{ad,2}$ is closed.

Finally, since $U_{ad,2}$ is closed, convex and bounded, it follows from Theorem A 6, given in the Appendix, that it is weakly sequentially compact. $\qquad \square$

Theorem 1. *Assume that one of the following condition holds*

(a) $\nu > 0$, and $U_{ad} \subseteq L^2((0,T); \mathbb{R}^{N_C})$;
(b) $\nu = 0$, and $U_{ad} = U_{ad,1} \subsetneq L^2((0,T); \mathbb{R}^{N_C})$;
(c) $\nu = 0$, and $U_{ad} = U_{ad,2} \subsetneq L^2((0,T); \mathbb{R}^{N_C})$;

then Problem (3.42) admits a solution.

Proof. First, we prove (a). The cost functional $(x, u) \mapsto J(x, u)$ is continuous and convex, hence weakly lower semicontinuous; see Theorem A 8 in the appendix. Furthermore, it is weakly coercive with respect to u, that is $\lim_{\|u\|_{L^2} \to \infty} J(x, u) = \infty$. Now, take a minimizing sequence $\{u^k\}_k$ in $L^2((0, T); \mathbb{R}^{N_C})$, which implies the (minimizing) sequence $\{(u^k, x^k)\}_k$. Since J is weakly coercive with respect to u, then u^k is bounded and by reflexivity of L^2 we can extract a weak converging subsequence $u^{k_j} \rightharpoonup u$ in $L^2((0, T); \mathbb{R}^{N_C})$; see, e.g., Theorem A 5 in the appendix. By Proposition 3, we also have that $x^{k_j}(T) \to x(T)$. Lower semicontinuity of J allows us to write the following

$$
\begin{aligned}
J(x, u) &= \frac{1}{2}\|x(T) - x_T\|_2^2 + \frac{\nu}{2} \sum_{n=1}^{N_C} \|u_n\|_{L^2}^2 \\
&\leq \lim_{j \to \infty} \frac{1}{2}\|x^{k_j}(T) - x_T\|_2^2 + \liminf_{j \to \infty} \frac{\nu}{2} \sum_{n=1}^{N_C} \|u_n^{k_j}\|_{L^2}^2 \\
&= \liminf_{j \to \infty} J(x^{k_j}, u^{k_j}) = \inf J(x, u) \,,
\end{aligned}
$$

which proves (a).

The proofs of (b) and (c) are similar to (a), but since $\nu = 0$ it is not proved the coercivity of J, and the fact that $U_{ad,1}$ and $U_{ad,2}$ are weakly sequentially compact, proved in Lemma 1, has to be used to obtain the existence of a weak convergent subsequence. $\quad\square$

Using the operator defined in (3.40), problem (3.42) can be equivalently written in the following compact form

$$
\begin{aligned}
&\min_{x,u} \quad J(x, u) \\
&\text{s.t.} \quad c(x, u) = 0 \,, \quad x \in X \,, \ u \in U_{ad} \,,
\end{aligned}
\tag{3.45}
$$

where $c(x, u)$ is defined in (3.40). Since by Proposition 1, the state x is uniquely determined by the starting condition and the controls, the mapping $u \mapsto x = x(u)$ is well defined, and the problem (3.45) can be written in the following equivalent reduced form

$$
\min_{u \in U_{ad}} J_r(u) := J(x(u), u) \,.
\tag{3.46}
$$

A solution to (3.42) satisfies first-order necessary optimality conditions. Next, we fist derive the optimality system in the case $U_{ad} = L^2((0, T); \mathbb{R}^{N_C})$. Then we consider the cases $U_{ad} = U_{ad,1}$ and $U_{ad} = U_{ad,2}$.

Since by Proposition 8 the linearized constraint is surjective and recalling (3.40), then standard results guarantee the existence of a Lagrange multiplier $p \in L^2((0, T); \mathbb{R}^{N_x})$; see, e.g., Theorem A 14 given in the Appendix and the corresponding references. A way to obtain the first-order optimality system is to consider the Lagrange function given by

$$
L(x, u, p) = J(x, u) + \left\langle \dot{x} - \left[A + \sum_{n=1}^{N_C} u_n B_n\right] x, p \right\rangle_{L^2} \,,
\tag{3.47}
$$

and to consider its Fréchet derivatives; see, e.g., [16, 119]. Performing a formal computation of these derivatives, including integration-by-parts, yields equation (3.49). This suggest that p must belong to a function space more regular than $L^2((0, T); \mathbb{R}^{N_x})$. Therefore, for a rigorous proof one has to assume (3.49) and prove the optimality conditions [119]. This is expressed in the following results.

Lemma 2. *The gradient of the reduced problem* (3.46) *is given by*

$$\nabla_{u_n} J_r(u) = \nu u_n - \langle B_n x, p \rangle \ , \quad n = 1, \ldots, N_C \ , \tag{3.48}$$

where $x = x(u)$ *is the unique solution to* $c(x, u) = 0$ *and* $p \in H^1((0, T); \mathbb{R}^{N_x})$ *is the unique solution to the following problem*

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* p \ , \quad p(T) = -(x(T) - x_T) \ . \tag{3.49}$$

Proof. Notice that (3.49) is of the same form as $c(x, u) = 0$. Hence, existence and uniqueness of solution to (3.49) can be proved as in Proposition 1.

Consider that $x = x(u)$ is the unique solution to $c(x, u) = 0$. Let p be the unique solution of (3.49), hence $p \in H^1((0, T); \mathbb{R}^{N_x})$.

Now, consider the cost functional $(x, u) \mapsto J(x, u)$ and δx satisfying the linearized constraint (3.5). The equivalence of (3.45) and (3.46) allows to compute $\nabla_{u_n} J_r(u)$ as follows

$$
\begin{aligned}
\left\langle \nabla_u J_r(u), \delta u \right\rangle_{L^2} &= \left\langle \frac{\partial J(x, u)}{\partial x}, \delta x \right\rangle_{L^2} + \left\langle \frac{\partial J(x, u)}{\partial u}, \delta u \right\rangle_{L^2} \\
&= \left\langle x(T) - x_T, \delta x(T) \right\rangle + \left\langle \nu u, \delta u \right\rangle_{L^2} \\
&= \left\langle -p(T), \delta x(T) \right\rangle + \left\langle \nu u, \delta u \right\rangle_{L^2} \ ,
\end{aligned}
\tag{3.50}
$$

where we used that p satisfies (3.49). Now, by means of the integration-by-parts rule, we have the following

$$
\begin{aligned}
\left\langle -p(T), \delta x(T) \right\rangle &= \left\langle p(0), \delta x(0) \right\rangle + \left\langle -\dot{p}, \delta x \right\rangle_{L^2} - \left\langle p, \dot{\delta x} \right\rangle_{L^2} \\
&= \left\langle \left[A + \sum_n u_n B_n \right]^* p, \delta x \right\rangle_{L^2} - \left\langle p, \left[A + \sum_n u_n B_n \right] \delta x + \left[\sum_n \delta u_n B_n \right] x \right\rangle_{L^2} \\
&= \left\langle p, \left[A + \sum_n u_n B_n \right] \delta x \right\rangle_{L^2} - \left\langle p, \left[A + \sum_n u_n B_n \right] \delta x + \left[\sum_n \delta u_n B_n \right] x \right\rangle_{L^2} \\
&= \left\langle p, -\left[\sum_n \delta u_n B_n \right] x \right\rangle_{L^2} = \sum_n \left\langle \langle -B_n x, p \rangle, \delta u_n \right\rangle_{L^2} \ .
\end{aligned}
\tag{3.51}
$$

Hence, replacing (3.51) in (3.50), we obtain that

$$
\begin{aligned}
\left\langle \nabla_u J_r(u), \delta u \right\rangle_{L^2} &= \sum_n \left\langle \nabla_{u_n} J_r(u), \delta u_n \right\rangle_{L^2} \\
&= \sum_n \left\langle \nu u_n - \langle B_n x, p \rangle, \delta u_n \right\rangle_{L^2} \ ,
\end{aligned}
\tag{3.52}
$$

showing that $\nabla_{u_n} J_r(u) = \nu u_n - \langle B_n x, p \rangle$ for $n = 1, \ldots, N_C$. $\qquad \square$

Theorem 2. *Assume that the pair* $(x, u) \in H^1((0, T); \mathbb{R}^{N^2}) \times L^2((0, T); \mathbb{R}^{N_C})$ *is a minimizer for Problem* (3.42) *with* $U_{ad} = L^2((0, T); \mathbb{R}^{N_C})$. *Then there exists a unique Lagrange multiplier* $p \in H^1((0, T); \mathbb{R}^{N_x})$, *such that the triple* (x, u, p) *solves the following system*

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \quad x(0) = x_0 \tag{3.53a}$$

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* p \ , \quad p(T) = -(x(T) - x_T) \tag{3.53b}$$

$$\nu u_n - \langle B_n x, p \rangle = 0 \quad n = 1, \ldots, N_C \ . \tag{3.53c}$$

The system (3.53) *is called the first-order optimality system for Problem* (3.42).

Proof. By applying Theorem A 17, given in the appendix, the optimality condition for (3.42) can be written as a variational inequality as follows

$$\left\langle \frac{\partial J(x,u)}{\partial x}, \delta x \right\rangle_{L^2} + \left\langle \frac{\partial J(x,u)}{\partial u}, \delta u \right\rangle_{L^2} \geq 0 \ , \tag{3.54}$$

for all $\delta u \in L^2((0,T); \mathbb{R}^{N_C})$ and δx satisfying the linearized constraint (3.5). Since U_{ad} is assumed to be the entire space $L^2((0,T); \mathbb{R}^{N_C})$, condition (3.54) is equivalent to

$$\left\langle \nabla_u J_r(u), \delta u \right\rangle_{L^2} = 0 \ , \ \forall \delta u \in L^2((0,T); \mathbb{R}^{N_C}) \ . \tag{3.55}$$

Finally, Lemma 2 and (3.55) imply (3.53c). $\qquad\square$

Notice that the optimality system (3.53) is in agreement with Theorem A 15 in the appendix.

Next, we derive the optimality condition for (3.42) with $U_{ad} = U_{ad,1}$. Notice that the control constraint $u \in U_{ad}$, can be written in the following form $h(u) \in U_{cone}$, where U_{cone} is the non-negative cone in $L^2((0,T); \mathbb{R}^{N_C})$ [119], that is

$$U_{cone} = \left\{ v(\cdot) \in L^2((0,T); \mathbb{R}^{N_C}) \mid v(t) \geq 0 \, a.e. \, t \in (0,T) \right\} \ .$$

In our case both $U_{ad,1}$ and U_{cone} have empty interior with respect to L^2. For this reason, we cannot guarantee existence of Lagrange multipliers by using standard results, see [119] (Chapter 6). We then proceed as follows.

Theorem 3. *Assume that the pair $(x,u) \in H^1((0,T); \mathbb{R}^{N_x}) \times L^2((0,T); \mathbb{R}^{N_C})$ is a minimizer for (3.42) with $U_{ad} = U_{ad,1}$. Then there exist a unique $p \in H^1((0,T); \mathbb{R}^{N_x})$ and $\lambda_+ \in L^2((0,T); \mathbb{R}^{N_C})$ and $\lambda_- \in L^2((0,T); \mathbb{R}^{N_C})$ such that $(x,u,p,\lambda_+,\lambda_-)$ satisfy the following first-order optimality system*

$$\nu u_n - \langle B_n x, p \rangle + \lambda_{+,n} - \lambda_{-,n} = 0 \ , \ n = 1, \ldots, N_C \tag{3.56a}$$

$$\lambda_{+,n} \geq 0 \ , \ b - u_n \geq 0 \ , \ \lambda_{+,n}(b - u_n) = 0 \tag{3.56b}$$

$$\lambda_{-,n} \geq 0 \ , \ u_n + b \geq 0 \ , \ \lambda_{-,n}(u_n + b) = 0 \ , \tag{3.56c}$$

where x and p are solutions to (3.53a) and (3.53b), respectively.

Proof. Let $u \in U_{ad,1}$ be an optimal control and $x = x(u)$ denotes the corresponding optimal state, and $p = p(u,x)$ be the unique solution to the adjoint equation. Existence and uniqueness of p are obtained as in Theorem 2. Consider the optimality condition in the following variational inequality form

$$\left\langle \nabla_u J_r(u), v - u \right\rangle_{L^2} \geq 0 \ , \ \forall v \in U_{ad,1} \ . \tag{3.57}$$

This inequality is equivalent to the following condition

$$u_n = \mathcal{P}_{U_{ad,1}} \left(u_n - \theta \nabla_{u_n} J_r(u) \right) \ , \ n = 1, \ldots, N_C \ , \tag{3.58}$$

where $\mathcal{P}_{U_{ad,1}}$ is the projection operator from $L^2((0,T); \mathbb{R}^{N_C})$ to $U_{ad,1}$ and θ is an arbitrary positive constant. Assuming $\theta = \frac{1}{\nu}$, recalling the gradient $\nabla_u J_r(u)$ given in Lemma 2, and writing the projection operator in the pointwise form, we obtain the following

$$u_n(t) = \mathcal{P}_{U_{ad,1}} \left(u_n(t) - \frac{1}{\nu} \left(\nu u_n(t) - \langle B_n x(t), p(t) \rangle \right) \right) = \mathcal{P}_{U_{ad,1}} \left(\frac{1}{\nu} \langle B_n x(t), p(t) \rangle \right) , \tag{3.59}$$

By using the projection $\mathcal{P}_{U_{ad,1}}(v) = v - \max(0, v - b) - \min(0, v + b)$, we get the following

$$\nu u_n(t) - \langle B_n x(t), p(t) \rangle + \max\big(0, \langle B_n x(t), p(t) \rangle - \nu b\big) + \min\big(0, \langle B_n x(t), p(t) \rangle + \nu b\big) = 0 , \tag{3.60}$$

where we multiplied by ν. Now, we define $\lambda_{+,n}$ and $\lambda_{-,n}$ as follows

$$\lambda_{+,n}(t) := \max\big(0, \langle B_n x(t), p(t) \rangle - \nu b\big) , \tag{3.61a}$$

$$\lambda_{-,n}(t) := -\min\big(0, \langle B_n x(t), p(t) \rangle + \nu b\big) . \tag{3.61b}$$

Notice that $\lambda_{+,n} \geq 0$ and $\lambda_{-,n} \geq 0$. Furthermore, for $u_n(t) > 0$ it holds that

$$b - u_n(t) > 0 \;\Rightarrow\; \nu u_n(t) = \langle B_n x(t), p(t) \rangle \;\Rightarrow\; \lambda_{+,n} = \max\big(0, \nu u_n(t) - \nu b\big) = 0 , \tag{3.62}$$

and similarly, for $u_n(t) < 0$ we have

$$u_n(t) + b > 0 \;\Rightarrow\; \nu u_n(t) = \langle B_n x(t), p(t) \rangle \;\Rightarrow\; \lambda_{-,n} = -\min\big(0, \nu u_n(t) + \nu b\big) = 0 . \tag{3.63}$$

Consequently, the complementarity conditions (3.56b) and (3.56c) hold and the claim follows. $\qquad\square$

Next, we derive the optimality condition for (3.42) with $U_{ad} = U_{ad,2}$.

Theorem 4. *Assume that the pair $(x, u) \in H^1((0,T); \mathbb{R}^{N_x}) \times L^2((0,T); \mathbb{R}^{N_C})$ is a minimizer for (3.42) with $U_{ad} = U_{ad,2}$. Then there exist a unique $p \in H^1((0,T); \mathbb{R}^{N_x})$ and a $\lambda \in L^2((0,T); \mathbb{R}^{N_C})$ such that (x, u, p, λ) satisfy the following first-order optimality system*

$$\nu u_n - \langle B_n x, p \rangle + \lambda_n = 0 , \;\; n = 1, \ldots, N_C \tag{3.64a}$$

$$\lambda_n(t) = \langle B_n x(t), p(t) \rangle - \frac{\nu \langle B_n x(t), p(t) \rangle}{\max\big(\nu, \frac{1}{b} \|g(t)\|_2\big)} \tag{3.64b}$$

$$\text{with } g(t) := \Big(\langle B_1 x(t), p(t) \rangle, \ldots, \langle B_{N_C} x(t), p(t) \rangle \Big)^T , \tag{3.64c}$$

where $x(t)$ and $p(t)$ are solutions to (3.53a) and (3.53b), respectively.

Proof. Similarly as in Theorem 3, we have the optimality condition in the following variational inequality form

$$\big\langle \nabla_u J_r(u), v - u \big\rangle_{L^2} \geq 0 , \;\; \forall v \in U_{ad,2} . \tag{3.65}$$

This inequality is equivalent to the following condition

$$u_n = \mathcal{P}_{U_{ad,2}}\Big(u_n - \theta \nabla_{u_n} J_r(u) \Big) , \;\; n = 1, \ldots, N_C , \tag{3.66}$$

where $\mathcal{P}_{U_{ad,2}}$ is the projection operator from $L^2((0,T); \mathbb{R}^{N_C})$ to $U_{ad,2}$ and θ is an arbitrary positive constant. Assuming $\theta = \frac{1}{\nu}$, recalling the gradient $\nabla_u J_r(u)$ given in Lemma 2, and writing the projection operator in the pointwise form, we obtain the following

$$u_n(t) = \mathcal{P}_{U_{ad,2}}\Big(u_n(t) - \frac{1}{\nu}\big(\nu u_n(t) - \langle B_n x(t), p(t) \rangle\big) \Big) = \mathcal{P}_{U_{ad,2}}\Big(\frac{1}{\nu} \langle B_n x(t), p(t) \rangle \Big) . \tag{3.67}$$

By using the projection $\mathcal{P}_{U_{ad,2}}(v) = \frac{v}{\max\left(1, \frac{1}{b}\|v\|_2\right)}$, and considering $g : (0, T) \to \mathbb{R}^{N_C}$ as in (3.64c), we obtain the following

$$\nu u_n(t) - \frac{\nu \langle B_n x(t), p(t) \rangle}{\max\left(\nu, \frac{1}{b}\|g(t)\|_2\right)} = 0 \ . \tag{3.68}$$

Now, we define λ_n as follows

$$\lambda_n := \langle B_n x(t), p(t) \rangle - \frac{\nu \langle B_n x(t), p(t) \rangle}{\max\left(\nu, \frac{1}{b}\|g(t)\|_2\right)} \ . \tag{3.69}$$

Consequently, from condition (3.68) the claim follows. $\qquad\square$

The following corollary consists in a reformulation of the complementarity conditions (3.64). This reformulation allows to write the complementarity conditions (3.64) in a classical KKT-system form.

Corollary 1. *Consider the pointwise constraint $h(t) \leq 0$, where $h : (0, T) \to \mathbb{R}$ is defined as $h(t) := \|u(t)\|_2 - b$, a.e. in $(0, T)$. Then, there exists a function $\lambda_c \in L^\infty(0, T)$ such that the optimality system (3.64) can be equivalently written as follows*

$$\nu u_n(t) - \langle B_n x(t), p(t) \rangle + \frac{u_n(t)}{\|u(t)\|_2} \lambda_c(t) = 0 \ \text{ a.e. in } (0, T) \ , \ \ n = 1, \ldots, N_C$$

$$\lambda_c(t) \geq 0 \ , \ h(t) \leq 0 \ , \ h(t)\lambda_c(t) = 0 \ \text{ a.e. in } (0, T) \ , \tag{3.70}$$

where $x(t)$ and $p(t)$ are solutions to (3.53a) and (3.53b), respectively.

Proof. Notice that the optimality condition (3.68) implies the following equivalence

$$\left\| \frac{1}{\nu} g(t) \right\|_2 < b \Leftrightarrow \|u(t)\|_2 < b \ . \tag{3.71}$$

We want to construct λ_c such that the following equation is satisfied

$$\lambda_n(t) = \frac{u_n(t)}{\|u(t)\|_2} \lambda_c(t) \ , \ \text{a.e. in } (0, T) \ . \tag{3.72}$$

Now, from Theorem 4, we have that

$$\lambda_n(t) := \begin{cases} \text{if } \frac{1}{b}\|g(t)\|_2 < \nu \ , & 0 \\ \text{if } \frac{1}{b}\|g(t)\|_2 \geq \nu \ , & \langle B_n x(t), p(t) \rangle \left(1 - \frac{\nu b}{\|g(t)\|_2}\right) \end{cases} \ . \tag{3.73}$$

Considering (3.71), we write the following

$$\lambda_n(t) = \begin{cases} \text{if } \|u(t)\|_2 < b \ , & 0 \\ \text{if } \|u(t)\|_2 \geq b \ , & \langle B_n x(t), p(t) \rangle \left(1 - \frac{\nu b}{\|g(t)\|_2}\right) \end{cases} \ . \tag{3.74}$$

From (3.68), we obtain that

$$g_n(t) = u_n(t) \max\left(\nu, \frac{1}{b}\|g(t)\|_2\right) \tag{3.75}$$

Replacing (3.75) in (3.74), we get the following expression

$$\lambda_n(t) = \begin{cases} \text{if } \|u(t)\|_2 < b\,, & 0 \\ \text{if } \|u(t)\|_2 \geq b\,, & u_n(t)\frac{\|g(t)\|_2}{b}\left(1 - \frac{\nu b}{\|g(t)\|_2}\right) \end{cases}\,, \tag{3.76}$$

and equivalently

$$\lambda_n(t) = \begin{cases} \text{if } \|u(t)\|_2 < b\,, & 0 \\ \text{if } \|u(t)\|_2 \geq b\,, & \frac{u_n(t)}{\|u(t)\|_2}\|u(t)\|_2\frac{\|g(t)\|_2}{b}\left(1 - \frac{\nu b}{\|g(t)\|_2}\right) \end{cases}\,, \tag{3.77}$$

which suggests to define λ_c as follows

$$\lambda_c(t) := \begin{cases} \text{if } \|u(t)\|_2 < b\,, & 0 \\ \text{if } \|u(t)\|_2 \geq b\,, & \|u(t)\|_2\frac{\|g(t)\|_2}{b}\left(1 - \frac{\nu b}{\|g(t)\|_2}\right) \end{cases}\,. \tag{3.78}$$

Notice that λ_c satisfies (3.72) and the complementary conditions $\lambda_c(t) \geq 0$ and $\lambda_c(t)h(t) = 0$, a.e. in $(0, T)$.

Now, we show that $\lambda_c \in L^\infty(0, T)$. Since $\|u(t)\|_2 \leq b$, we have that

$$\lambda_c(t) \leq \|g(t)\|_2\left(1 - \frac{\nu b}{\|g(t)\|_2}\right) = \|g(t)\|_2 - \nu b \leq \|g(t)\|_2$$
$$= \left(\sum_{n=1}^{N_C} |\langle B_n x(t), p(t)\rangle|^2\right)^{1/2} \leq \left(\sum_{n=1}^{N_C} \|B_n x(t)\|_2^2\|p(t)\|_2^2\right)^{1/2}\,. \tag{3.79}$$

By Proposition 2, we know that the dynamics is norm-preserving, hence we obtain the following

$$\lambda_c(t) \leq \left(\sum_{n=1}^{N_C} \|B_n x(t)\|_2^2\|p(t)\|_2^2\right)^{1/2} \leq \left(\sum_{n=1}^{N_C} \|B_n\|_{\mathcal{L}}^2\|x(0)\|_2^2\|x(T) - x_T\|_2^2\right)^{1/2} \tag{3.80}$$
$$\leq \tilde{C}\|x(0)\|_2\|x(T) - x_T\|_2 \leq \tilde{C}\|x(0)\|_2\big(\|x(0)\|_2 + \|x_T\|_2\big)\,,$$

where $\tilde{C}$ is a positive constant depending on B_n and we used the Cauchy-Schwarz and triangular inequalities. The previous shows that $\lambda_c \in L^\infty(0, T)$. $\qquad\square$

We remark that the Filippov's Theorem and the Pontryagin Maximum Principle [13, 26, 84], could be applied to study existence of solutions and optimality conditions for problem (3.42). However, some of the hypotheses of these standard results could fail in the case of possibly non-smooth parts in the cost functional, whereas the framework that we considered can be applied with simple changes. Moreover, we prefer to proceed as above in order to obtain optimality systems in a form that is suitable for the application of Newton and semi-smooth Newton methods, as described in Chapter 4.2.

Properties of the adjoint equation and its linearization

Next, we discuss some properties of the solution to the adjoint equation (3.53b) and its linearization.

The following lemma states a similar result as in Proposition 4 for the backward equation (3.53b).

Proposition 9. *Let v and w in $L^2((0,T); \mathbb{R}^{N_C})$ be two control functions, and consider the corresponding solutions to (3.53b), $q = p(v,y)$ and $r = p(w,z)$, with $y = x(v)$ and $z = x(w)$ solutions to (3.53a). Assume that $\|x_T\|_2 = \|x_0\|_2$ and that A and B_n are skew-symmetric. Then there exists a positive constant c_3 such that the following holds*

$$\|q(t) - r(t)\|_2^2 \le c_3 \|v - w\|_{L^1} , \ a.e. \ in \ (0,T) . \tag{3.81}$$

Moreover, if v and w are in $L^\infty((0,T); \mathbb{R}^{N_C})$, then there exists a positive constant $\tilde{c}_3$ such that the following holds

$$\|q(t) - r(t)\|_2 \le \tilde{c}_3 \|v - w\|_{L^1} , \ a.e. \ in \ (0,T) . \tag{3.82}$$

Proof. Similarly to the proof of Proposition 4, we obtain

$$
\begin{aligned}
\|q(t) - r(t)\|_2^2 &\le \|q(T) - r(T)\|_2^2 + 2\sum_n \int_t^T |v_n(s) - w_n(s)| |\langle q(s), B_n^* r(s)\rangle| ds \\
&\le \|(y(T) - x_T) - (z(T) - x_T)\|_2^2 \\
&\quad + 2\sum_n \int_t^T |v_n(s) - w_n(s)| \|B_n\|_{\mathcal{L}} \|q(T)\|_2 \|r(T)\|_2 ds \\
&\le \|y(T) - z(T)\|_2^2 + 2K_0 \|y(T) - x_T\|_2 \|z(T) - x_T\|_2 \sum_n \int_0^T |v_n(s) - w_n(s)| ds \\
&\le c_1 \|v - w\|_{L^1} + 2K_0 (\|y(T)\|_2 + \|x_T\|_2)(\|z(T)\|_2 + \|x_T\|_2) \|v - w\|_{L^1} \\
&= \left(c_1 + 8\|x_0\|_2^2 K_0\right) \|v - w\|_{L^1} ,
\end{aligned}
\tag{3.83}
$$

which is (3.81) with $c_3 = c_1 + 8\|x_0\|_2^2 K_0$, with $K_0 = \max_n \|B_n\|_{\mathcal{L}}$.

The estimate (3.82) can be proved similarly as in Proposition 4. $\qquad\square$

The following proposition states boundedness of the solution to the linearized adjoint equation. For next usage, this result is derived for general terminal conditions for adjoint and linearized adjoint equations. However, in Proposition 11, we specify this result for the particular terminal condition $p(T) = -(x(T) - x_T)$ appearing in the optimality system (3.53).

Proposition 10. *Consider the following adjoint equation*

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]^* p , \ with \ p(T) = p_T , \tag{3.84}$$

with $\|p_T\|_2 \ne 0$, and and let δp be the unique solution to the corresponding linearized problem, that is

$$-\dot{\delta p} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]^* \delta p + \left[\sum_{n=1}^{N_C} B_n \delta u_n\right]^* p , \ with \ \delta p(T) = \delta p_T . \tag{3.85}$$

Then the following estimate holds

$$\|\delta p\|_{L^2} \le 2T\sqrt{N_C} K_{00} \|p_T\|_2 \|\delta u\|_{L^2} + \sqrt{T} \|\delta p_T\|_2 . \tag{3.86}$$

where $K_{00} = \sum_{n=1}^{N_C} \|B_n\|_{\mathcal{L}}$ as in Proposition 5.

Proof. By multiplying the differential equation in (3.85) from the left with δp and using the same arguments as in Proposition 5 for δx, we have

$$
\begin{aligned}
\|\delta p(t)\|_2^2 &= \|\delta p(T)\|_2^2 - 2\int_t^T \langle \delta p, \Big[\sum_n \delta u_n B_n\Big]p\rangle dt \\
&\leq \|\delta p_T\|_2^2 + 2\sum_n \int_0^T |\delta u_n||\langle \delta p, B_n p\rangle|dt \\
&\leq \|\delta p_T\|_2^2 + 2\sum_n \int_0^T |\delta u_n|\|\delta p\|_2\|B_n\|_{\mathcal{L}}\|p_T\|_2 dt \\
&\leq \|\delta p_T\|_2^2 + 2\|p_T\|_2 \sum_n \|B_n\|_{\mathcal{L}} \int_0^T |\delta u_n|\|\delta p\|_2 dt \\
&\leq \|\delta p_T\|_2^2 + 2\|p_T\|_2 \sqrt{N_C}K_{00}\|\delta u\|_{L^2}\|\delta p\|_{L^2} \ .
\end{aligned}
\tag{3.87}
$$

Now, integrating over $(0,T)$, we obtain

$$
\begin{aligned}
\int_0^T \|\delta p(t)\|_2^2 dt &\leq T\|\delta p_T\|_2^2 + 2\int_0^T \|p_T\|_2 \sqrt{N_C}K_{00}\|\delta u\|_{L^2}\|\delta p\|_{L^2} dt \\
&\Rightarrow \|\delta p\|_{L^2}^2 \leq T\|\delta p_T\|_2^2 + 2T\|p_T\|_2\sqrt{N_C}K_{00}\|\delta u\|_{L^2}\|\delta p\|_{L^2} \\
&\Rightarrow \|\delta p\|_{L^2}^2 - 2T\|p_T\|_2\sqrt{N_C}K_{00}\|\delta u\|_{L^2}\|\delta p\|_{L^2} - T\|\delta p_T\|_2^2 \leq 0 \ .
\end{aligned}
\tag{3.88}
$$

The discriminant of the previous quadratic inequality is

$$
\Delta = 4T^2 N_C K_{00}^2 \|p_T\|_2^2 \|\delta u\|_{L^2}^2 + 4T\|\delta p_T\|_2^2 > 0 \ \ \forall (\delta u, \delta p_T) \neq (0,0) \ ,
\tag{3.89}
$$

and we recall that $\|p_T\|_2 \neq 0$. Consequently, inequality (3.88) is satisfied for

$$
\|\delta p\|_{L^2} \leq T\sqrt{N_C}K_{00}\|p_T\|_2\|\delta u\|_{L^2} + \sqrt{T^2 N_C K_{00}^2\|p_T\|_2^2\|\delta u\|_{L^2}^2 + T\|\delta p_T\|_2^2} \ .
\tag{3.90}
$$

The previous inequality (3.90) allows us to write that

$$
\begin{aligned}
\|\delta p\|_{L^2} &\leq T\sqrt{N_C}K_{00}\|p_T\|_2\|\delta u\|_{L^2} \\
&+ \sqrt{T^2 N_C K_{00}^2\|p_T\|_2^2\|\delta u\|_{L^2}^2 + T\|\delta p_T\|_2^2 + 2(T\sqrt{N_C}K_{00}\|p_T\|_2\|\delta u\|_{L^2})\sqrt{T}\|\delta p_T\|_2} \\
&\leq 2T\sqrt{N_C}K_{00}\|p_T\|_2\|\delta u\|_{L^2} + \sqrt{T}\|\delta p_T\|_2 \ ,
\end{aligned}
\tag{3.91}
$$

which concludes the proof. $\qquad\square$

The following result specifies Proposition 10 when the adjoint and its linearization depend on the solutions to the constraint and linearized constraint equations.

Proposition 11. *Under the assumptions of Proposition 5 and Proposition 10, if $p(T) = -(x(T) - x_T)$ and $\delta p(T) = -\delta x(T)$, then the following holds*

$$
\|\delta p\|_{L^2} \leq 6T\sqrt{N_C}K_{00}\|x_0\|_2\|\delta u\|_{L^2}.
\tag{3.92}
$$

Proof. The claim follows from (3.86) by means of (3.18). $\qquad\square$

Proposition 12. *Let $u \in L^2((0,T); \mathbb{R}^{N_C})$ and $h \in L^2((0,T); \mathbb{R}^{N_C})$. Consider the map $u \mapsto p(u)$, where $p(u) \in H^1((0,T); \mathbb{R}^{N_x})$ is the unique solution to the adjoint equation (3.53b) corresponding to u, and denote by $\delta p(u,h) \in H^1((0,T); \mathbb{R}^{N_x})$ the unique solution to (3.85) corresponding to h, u and $\delta p(T) = -\delta x(T)$, where δx solves (3.5). Assume that A and B_n are skew-symmetric. The following expansion holds*

$$p(u+h) = p(u) + \delta p(u,h) + \vartheta_p(u,h) , \tag{3.93}$$

where $\vartheta_p(u,h) \in H^1((0,T); \mathbb{R}^{N_x})$ solves the following problem

$$-\dot{\vartheta}_p = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* \vartheta_p + \sum_{n=1}^{N_C} h_n B_n^* (\vartheta_p + \delta p) \quad in \;\; [0,T] , \;\; and \;\; \vartheta_p(T) = -\vartheta(T) , \tag{3.94}$$

where ϑ solves (3.24). Moreover, there exist positive constants c_{ϑ_p} and $\tilde{c}_{\vartheta_p}$ such that the following estimates hold

$$\|\vartheta_p(t)\|_2 \le c_{\vartheta_p} \|h\|_{L^2}^2 , \; a.e. \;\; in \, (0,T) \tag{3.95a}$$

$$\|\vartheta_p\|_{L^2} \le \tilde{c}_{\vartheta_p} \|h\|_{L^2}^2 . \tag{3.95b}$$

Furthermore, for every $v, w \in L^2((0,T); \mathbb{R}^{N_C})$ there exists a positive constant $\hat{c}_{\vartheta_p}$ such that

$$\|q(t) - r(t)\|_2 \le \hat{c}_{\vartheta_p} \|v - w\|_{L^2} , \tag{3.96}$$

where $q = p(v)$ and $r = p(w)$ are solutions to (3.53b) corresponding to v and w, respectively.

Proof. The proof of (3.23) is the same as in Proposition 6. Estimates (3.95a) and (3.95b) are obtained similarly as in Proposition 10 and using Proposition 6 and Proposition 11. Estimate (3.96) can be proved similarly as is Proposition 7. We omit these proofs for brevity. $\qquad \square$

3.3 Piecewise-constant L^2-regularized optimal control problems

In this section, a piecewise constant optimal control problem is discussed. This problems is considered when the need of an exact implementation of the control functions in specific laboratory pulse shapers arises; see, e.g., [37]. We consider problem ($\mathcal{P}$) with $\beta = 0$ and the admissible control space as a subset of a parametrized piecewise-constant vector space. In particular, we consider the following piecewise-constant optimal control problem, where U_{ad} resembles the set $U_{ad,1}$ defined in (3.2). We have

$$\min_{x,u} \;\; J(x,u) := \frac{1}{2}\|x(T) - x_T\|_2^2 + \sum_{n=1}^{N_C} \frac{\nu}{2}\|u_n\|_{L^2}^2$$

$$\text{s.t.} \;\; \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x , \; \text{in} \; (0,T] , \; x \in X \tag{3.97}$$

$$u \in U_{ad} := \big\{ v \in \mathcal{PC}((0,T); M; \mathbb{R}^{N_C}) \; : \; |v_n(t)| \le b$$

$$\text{a.e. in} \, (0,T) \; \text{for} \; n = 1, \ldots, N_C , \; b > 0 \big\} ,$$

where $\mathcal{PC}((0,T);M;\mathbb{R}^{N_C})$ is defined as follows. We decompose the time domain $(0,T]$ in M sub-intervals $\Delta_j := (t_{j-1}, t_j]$, such that

$$(0,T] = \bigcup_{j=1}^{M} \Delta_j , \qquad (3.98)$$

with $0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T$. Now, $\mathcal{PC}((0,T);M;\mathbb{R}^{N_C})$ is the set of all piecewise-constant functions in the following sense

$$\mathcal{PC}((0,T);M;\mathbb{R}^{N_C}) := \left\{ v \in L^2((0,T);\mathbb{R}^{N_C}) : v(t) = c_j \text{ for } t \in \Delta_j , \ j = 1, \ldots, M \right\} , \qquad (3.99)$$

where $c_j \in \mathbb{R}^{N_C}$, $j = 1, \ldots, M$. In particular, recall that $u_n(t) = c_j^n$ with $c_j^n \in \mathbb{R}$ for $t \in \Delta_j = (t_{j-1}, t_j]$, we consider parameters c_j^n given by the following equality

$$c_j^n = \frac{1}{\sqrt{\tilde{\Delta}_j}} u_{n,j} , \qquad (3.100)$$

for $t \in \Delta_j$ and $j = 1, \ldots, M$, where $\tilde{\Delta}_j = t_j - t_{j-1}$ is the measure of the interval Δ_j. We denote by $\tilde{u}_n = \left(u_{n,1}, \ldots, u_{n,M} \right) \in \mathbb{R}^M$ the parametrized control corresponding to u_n. This choice yields the equality between the L^2-norm of the control $u_n(t)$ and the Euclidean norm of its parametrization $\tilde{u}_n$, as follows

$$\|u_n\|_{L^2}^2 = \int_0^T |u_n(t)|^2 dt = \sum_{j=1}^{M} \int_{\Delta_j} \frac{1}{\tilde{\Delta}_j} |u_{n,j}|^2 dt = \sum_{j=1}^{M} |u_{n,j}|^2 = \|\tilde{u}_n\|_2^2 . \qquad (3.101)$$

Notice that, in this framework we have that $\mathcal{PC}((0,T);M;\mathbb{R})$ and $\mathbb{R}^M$ are isometrically isomorphic.

The existence of a solution to the optimization problem (3.97) is stated in the following theorem.

Theorem 5. *Assume that $\beta = 0$, $\nu \geq 0$, then problem* (3.97) *admits a solution.*

Proof. Notice that the admissible set $U_{ad} \subset \mathcal{PC}((0,T);M;\mathbb{R}^{N_C})$ is convex, bounded, and closed. To see it, we recall that $\mathcal{PC}((0,T);M;\mathbb{R}) \subset L^2(0,T)$ is isometrically isomorphic to $\mathbb{R}^M$. Since in finite-dimensional spaces all the norms are equivalent, we have that U_{ad} is a closed ball of radius b, and hence convex, closed and bounded.

With this argument, the proof of the existence of a solution to (3.97) is the same as in Theorem 1. $\qquad \square$

Next, we discuss the first-order optimality system for (3.97). First, we derive the reduced gradient of (3.97). For this purpose, consider the following Lagrange function

$$L(x,u,p) = J(x,u) + \langle p, \dot{x} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \rangle_{L^2}$$

$$= \frac{1}{2}\|x(T) - x_T\|_2^2 + \sum_{n=1}^{N_C} \frac{\nu}{2} \int_0^T |u_n|^2 dt + \int_0^T \langle p, \dot{x} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \rangle dt \qquad (3.102)$$

$$= \frac{1}{2}\|x(T) - x_T\|_2^2 + \int_0^T \langle p, \dot{x} - Ax \rangle dt + \sum_{n=1}^{N_C} \sum_{j=1}^{M} \frac{1}{\tilde{\Delta}_j} \int_{\Delta_j} \frac{\nu}{2} u_{n,j}^2 - u_{n,j} \langle p, B_n x \rangle dt .$$

Now, the gradient component $\nabla_{u_{n,j}} L(x, u, p)$ is obtained as the directional derivative of $L(x, u, p)$ with respect to $u_{n,j}$, and we write it in the following compact form

$$\nabla_{u_{n,j}} L(x, \tilde{u}, p) = \nu u_{n,j} - E_j(\langle p, B_n x \rangle) , \tag{3.103}$$

where $E_j : L^2((0, T); \mathbb{R}) \to \mathcal{PC}((0, T); M; \mathbb{R})$ is defined as follows

$$E_j(v)(t) := \frac{1}{\tilde{\Delta}_j} \int_{\Delta_j} v(s) ds , \ \forall t \in \Delta_j , \tag{3.104}$$

for $j = 1, \ldots, M$. Notice that E_j plays the role of a projector from L^2 onto $\mathcal{PC}$. This projector is needed since x and p are in $H^1((0, T); \mathbb{R}^N)$, and it is not guaranteed that $\langle B_n x, p \rangle$ is a piecewise-constant function. In general, we have $\langle B_n x, p \rangle \in H^1((0, T); \mathbb{R})$. On the other hand, piecewise-constant functions are not in the Sobolev space H^1; see, e.g., [47]. Consequently, to guarantee that the gradient belongs to the space of piecewise-constant functions, the projection of $\langle B_n x, p \rangle$ in $\mathcal{PC}((0, T); M; \mathbb{R})$ is necessary.

Now, similarly as in Theorem 3, the first-order necessary optimality system for (3.97) is given by

$$\nu u_{n,j} - E_j(\langle p, B_n x \rangle) + \lambda_{+,n,j} - \lambda_{-,n,j} = 0 \tag{3.105a}$$
$$\lambda_{+,n,j} \geq 0 , \ b - u_{n,j} \geq 0 , \ \lambda_{+,n,j}(b - u_{n,j}) = 0 \tag{3.105b}$$
$$\lambda_{-,n,j} \geq 0 , \ u_{n,j} + b \geq 0 , \ \lambda_{-,n,j}(u_{n,j} + b) = 0 , \tag{3.105c}$$

where x and p are solutions to (3.53a) and (3.53b), respectively, and $\lambda_{+,n,j}$ and $\lambda_{-,n,j}$ are the Lagrange multipliers corresponding to the control constraint. Furthermore, the complementarity conditions (3.105) can be written in the following projection form

$$C_{n,j}(\tilde{u}, \tilde{\mu}) := -\theta \tilde{\mu}_{n,j} + \max(0, u_{n,j} - b + \theta \tilde{\mu}_{n,j}) + \min(0, u_{n,j} + b + \theta \tilde{\mu}_{n,j}) = 0 \tag{3.106a}$$
$$\tilde{\mu}_{n,j} = E_j(\langle B_n x, p \rangle) - \nu u_{n,j} , \ n = 1, \ldots, N_C , \ j = 1, \ldots, M , \tag{3.106b}$$

where $\tilde{\mu}_{n,j} = -\nabla_{u_{n,j}} L(x, \tilde{u}, p)$ and θ is an arbitrary positive constant.

3.4 L^1-penalized optimal control problems

In this section, we investigate problem $(\mathcal{P})$ with $\beta > 0$, that is

$$\min_{x,u} \ J(x, u) := \frac{1}{2} \|x(T) - x_T\|_2^2 + \sum_{n=1}^{N_C} \left[\frac{\nu}{2} \|u_n\|_{L^2}^2 + \beta \|u_n\|_{L^1} \right]$$

$$\text{s.t.} \quad \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x , \ \text{in} \ (0, T] \tag{3.107}$$

$$x \in X , \ u \in U_{ad,1} ,$$

where X and $U_{ad,1}$ are defined in (3.1) and (3.2), respectively. This optimal control problem allows to generate control functions that are said to be "sparse", and that resembles the "pulsed shaped" controls often used in quantum control applications, like NMR and MRI.

This section aims at a theoretical characterization of the optimal control solutions to (3.107). In Theorem 6 we discuss existence of solutions to (3.107). In Theorem 7 we

derive first-order necessary optimality conditions. Lemma 3 contains preliminary results that are used in the sequel of the present thesis. Theorems 8, 9, 10 and 11 investigate the sparsity property of optimal controls and consider relationships between them and the weight parameters β and ν.

The existence of solutions to (3.107) is stated in the following theorem.

Theorem 6. *Assume that $\beta > 0$, $\nu \geq 0$, then problem* (3.107) *admits a solution.*

Proof. For $\nu \geq 0$ and $\beta > 0$ the map $(x, u) \mapsto J(x, u)$ is convex and continuous, hence weakly lower-semicontinuos, see Theorem A 8 in the appendix. Therefore the existence of a solution to the optimal control (3.107) can be proved by using Proposition 3 and Lemma 1 and following the same arguments as in Theorem 1. See also [33, 66, 119, 129] for more detailed discussions. □

Next, we study first-order necessary conditions. To this purpose, we notice that, in the space of solutions to the governing differential equation in (3.107), the reduced cost functional $\hat{J}(u) := J(x(u), u)$ can be written as follows

$$\hat{J}(u) = \varphi(u) + \phi(u) ,\tag{3.108}$$

where $u \mapsto \varphi(u) := \frac{1}{2}\|x(T) - x_T\|_2^2 + \frac{\nu}{2}\|u\|_{L^2}^2$ is Gâteaux differentiable, with derivative given by $(\varphi'(u))_n = \nu u_n - \langle B_n x, p \rangle$ (Lemma 2) where x and p solve

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right] x \; , \quad x(0) = x_0 ,\tag{3.109}$$

and

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]^* p \; , \quad p(T) = -(x(T) - x_T) ,\tag{3.110}$$

respectively (see, e.g., [31]). The map $u \mapsto \phi(u) := \beta\|u\|_{L^1}$ is non-differentiable in the classical sense, however it is convex and continuous and hence its subdifferential $\partial\phi(u)$ is non-empty, and it is given by the following; see, e.g., [34, 45];

$$\partial\phi(u) = \left\{ \hat{\lambda} \in L^2((0, T); \mathbb{R}^{N_C}) \; : \; \phi(v) - \phi(u) \geq \langle \hat{\lambda}, v - u \rangle_{L^2} , \; \forall v \in L^2((0, T); \mathbb{R}^{N_C}) \right\} .\tag{3.111}$$

Consequently, problem (3.107) in the reduced form reads as follows

$$\min_{u \in U_{ad,1}} \hat{J}(u) .\tag{3.112}$$

In the next theorem, we derive first-order necessary conditions for (3.107), and we derive a first-order optimality system. Notice that, similar optimality systems are discussed in [110] and in [126]. In particular, in [110] optimality conditions are discussed for linear systems by following non-smooth analysis, whereas in [126] the optimality system is obtained by means of minimizing a Hamiltonian function. In the next result, we follow the same arguments as in [110], and we prove a first-order necessary condition for the bilinear optimal control problem (3.107) by using Theorem A 18 [45], given in the appendix.

Theorem 7. *Assume that the pair $(x, u) \in H^1((0, T); \mathbb{R}^N) \times L^2((0, T); \mathbb{R}^{N_C})$ is a minimizer for* (3.107). *Then there exist a unique $p \in H^1((0, T); \mathbb{R}^N)$, which solves* (3.110), *and a $\hat{\lambda} \in \partial\phi(u)$ such that the following inequality condition is satisfied*

$$\langle \varphi'(u) + \hat{\lambda}, v - u \rangle_{L^2} \geq 0 \quad \forall v \in U_{ad,1} ,\tag{3.113}$$

where $\varphi(v)$, $\varphi'(v)$ and $\phi(v)$ are defined in (3.108). Moreover, there exist $\lambda_+ \in L^2((0,T);\mathbb{R}^{N_C})$ and $\lambda_- \in L^2((0,T);\mathbb{R}^{N_C})$ such that (3.113) is equivalent to the following system

$$\nu u_n - \langle B_n x, p \rangle + \lambda_{+,n} - \lambda_{-,n} + \hat{\lambda}_n = 0 \ , \ n = 1, ..., N_C \tag{3.114a}$$

$$\lambda_{+,n} \geq 0 \ , \ b - u_n \geq 0 \ , \ \lambda_{+,n}(b - u_n) = 0 \tag{3.114b}$$

$$\lambda_{-,n} \geq 0 \ , \ u_n + b \geq 0 \ , \ \lambda_{-,n}(u_n + b) = 0 \tag{3.114c}$$

$$\hat{\lambda}_n = \beta \ , \ a.e. \ in \ \{t \in (0,T) \ : \ u_n(t) > 0\} \tag{3.114d}$$

$$|\hat{\lambda}_n| \leq \beta \ , \ a.e. \ in \ \{t \in (0,T) \ : \ u_n(t) = 0\} \tag{3.114e}$$

$$\hat{\lambda}_n = -\beta \ , \ a.e. \ in \ \{t \in (0,T) \ : \ u_n(t) < 0\} \ , \tag{3.114f}$$

where $\hat{\lambda}_n \in \Lambda_{ad} := \{g \in L^2(0,T) \ : \ |g| \leq \beta \ a.e. \ on \ (0,T)\}$ and x and p are the unique solutions to (3.109) and (3.110), respectively.

Proof. Consider problem (3.112). Recalling that $U_{ad,1}$ is closed and convex, according to Proposition 2.2 in Chapter II of [45], a necessary condition for $u \in U_{ad,1}$ to be a minimizer for (3.112) is

$$\langle \varphi'(u), v - u \rangle_{L^2} + \phi(v) - \phi(u) \geq 0 \quad \forall v \in U_{ad,1} \ . \tag{3.115}$$

Notice that (3.111) implies that

$$\langle \varphi'(u), v - u \rangle_{L^2} + \phi(v) - \phi(u) \geq \langle \varphi'(u), v - u \rangle_{L^2} + \langle \hat{\lambda}, v - u \rangle_{L^2} \ , \tag{3.116}$$

for any $\hat{\lambda} \in \partial\phi(u)$. Hence, a sufficient condition for (3.115) is given by the following variational inequality

$$\langle \varphi'(u) + \hat{\lambda}, v - u \rangle_{L^2} \geq 0 \quad \forall v \in U_{ad,1} \ . \tag{3.117}$$

It follows from (3.115) that

$$\langle \varphi'(u), v \rangle_{L^2} + \phi(v) \geq \langle \varphi'(u), u \rangle_{L^2} + \phi(u) \quad \forall v \in U_{ad,1} \ ,$$

which means that

$$u \in \arg \min_{v \in U_{ad,1}} \left\{ \langle \varphi'(u), v \rangle_{L^2} + \phi(v) \right\} \ . \tag{3.118}$$

Now, we consider the map $f : v \mapsto f(v) = \langle \varphi'(u), v \rangle_{L^2}$ and introduce the indicator function $I_{U_{ad,1}}$ defined as follows

$$I_{U_{ad,1}}(v) := \begin{cases} 0 & \text{if} \quad v \in U_{ad,1} \\ \infty & \text{otherwise} \end{cases} \ .$$

Then (3.118) is equivalent to

$$u \in \arg \min_{v} \left\{ f(v) + \phi(v) + I_{U_{ad,1}}(v) \right\} \ , \tag{3.119}$$

and hence a minimizer u must satisfy the following necessary condition

$$0 \in \partial\left(f + \phi + I_{U_{ad,1}} \right)(u) \ . \tag{3.120}$$

Next, by using calculus for subdifferentials, we write

$$\partial\left(f + \phi + I_{U_{ad,1}} \right)(u) = \partial f(u) + \partial\phi(u) + \partial I_{U_{ad,1}}(u)$$
$$= \varphi'(u) + \partial\phi(u) + N_{U_{ad,1}}(u) \ ,$$

where we used the fact that the subdifferential of the indicator function $I_{U_{ad,1}}$ is equal to the normal cone $N_{U_{ad,1}}$ defined as

$$N_{U_{ad,1}}(u) := \left\{ z \in L^2((0,T); \mathbb{R}^{N_C}) \ : \ \langle z, v - u \rangle_{L^2} \leq 0 \right\} .$$

Consequently, from (3.120) we have

$$-\varphi'(u) \in \partial\phi(u) + N_{U_{ad,1}}(u) , \tag{3.121}$$

which means that there exists a $\hat{\lambda} \in \partial\phi(u)$ and a $z \in N_{U_{ad,1}}(u)$ such that

$$-\varphi'(u) - \hat{\lambda} = z \in N_{U_{ad,1}}(u) .$$

This equality together with the definition of $N_{U_{ad}}(u)$ imply that

$$\langle -\varphi'(u) - \hat{\lambda}, v - u \rangle_{L^2} \leq 0 ,$$

that is (3.113).

We remark that a similar proof of (3.113) can be found in [120] and further arguments regarding the equivalence between (3.115) and (3.113) can be found in [45, 64].

Next, we show the last statement of the Theorem. The previous variational inequality (3.117) can be equivalently written in a projection form; see, e.g., [119]. By recalling that $(\varphi'(u))_n = \nu u_n - \langle B_n x, p \rangle$ (Lemma 2), and following the proof of Theorem 3, (3.117) becomes as follows

$$\begin{aligned}
\left(\nu u_n(t) - \langle B_n x(t), p(t) \rangle + \hat{\lambda}_n \right) + \max\left(0, \langle B_n x(t), p(t) \rangle + \hat{\lambda}_n - \nu b\right) \\
+ \min\left(0, \langle B_n x(t), p(t) \rangle + \hat{\lambda}_n + \nu b\right) = 0 .
\end{aligned} \tag{3.122}$$

Now, the proof is the same as in Theorem 3 by defining

$$\lambda_{+,n} := \max\left(0, \langle B_n x(t), p(t) \rangle + \hat{\lambda}_n - \nu b\right) \tag{3.123}$$

and

$$\lambda_{-,n} := -\min\left(0, \langle B_n x(t), p(t) \rangle + \hat{\lambda}_n + \nu b\right) , \tag{3.124}$$

and the claim follows by noticing that $\hat{\lambda}_n \in \Lambda_{ad}$ together with (3.114d)-(3.114f) is an equivalent expression for $\hat{\lambda} \in \partial\phi(u)$. $\qquad\square$

Next, we prove continuity of the map $u \mapsto \langle B_n x(u), p(u) \rangle \in L^q(0,T)$, with $x(u)$ and $p(u)$ solutions to (3.109) and (3.110), respectively. Notice that, the Euclidean scalar product $\langle \cdot, \cdot \rangle$ is understood pointwise in time.

Lemma 3. *Consider the map $u \mapsto \Theta_n(u) := \langle B_n x(u), p(u) \rangle$. Under the assumptions of Proposition 4 and Proposition 9, we have*

(a) *for every $u \in L^2((0,T); \mathbb{R}^{N_C})$ it holds that $\Theta_n(u) \in L^\infty(0,T)$ and in particular, we have that $\Theta_n(u)(t)$ is pointwise in time bounded as follows*

$$|\Theta_n(u)(t)| \leq 2\|B_n\|_{\mathcal{L}} \|x_0\|_2^2 ; \tag{3.125}$$

(b) *the map $\Theta_n : L^2((0,T); \mathbb{R}^{N_C}) \to L^q(0,T)$, with $1 \leq q < \infty$, is locally Lipschitz continuous, in the sense that there exists a constant $K > 0$ such that*

$$\|\Theta_n(v) - \Theta_n(w)\|_{L^q} \leq K\|v - w\|_{L^2} ,$$

for any $v, w \in L^2((0,T); \mathbb{R}^{N_C})$ such that $\|v - w\|_{L^2} \leq 1$;

(c) the map $\Theta_n : L^2((0,T); \mathbb{R}^{N_C}) \to L^2(0,T)$ is Fréchet differentiable. In particular, by denoting with $\tilde{A}_n(u,h)$ the Fréchet derivative, it holds that

$$\|\Theta_n(u+h) - \Theta_n(u) - \tilde{A}_n(u,h)\|_{L^2} = O(\|h\|_{L^2}^2) , \qquad (3.126)$$

as $h \to 0$.

Proof. (a) Recalling Proposition 2 and that $\|x_T\|_2 = \|x_0\|_2$, we write that

$$\begin{aligned}
|\Theta_n(u)(t)| &= |\langle B_n x(t), p(t) \rangle| \\
&\leq \|B_n\|_{\mathcal{L}} \|x_0\|_2 \|x(T) - x_T\|_2 \\
&\leq \|B_n\|_{\mathcal{L}} \|x_0\|_2 (\|x_0\|_2 + \|x_T\|_2) \\
&= 2\|B_n\|_{\mathcal{L}} \|x_0\|_2^2 ,
\end{aligned}$$

which implies the claim.

(b) Let $y = x(v)$, $z = x(w)$, $q = p(v,y)$, $r = p(w,z)$. We have the following pointwise in time estimate

$$\begin{aligned}
|\Theta_n(v) - \Theta_n(w)| &= |\langle B_n y, q \rangle - \langle B_n z, r \rangle| \\
&= |\langle B_n y, q \rangle - \langle B_n z, r \rangle + \langle B_n z, q \rangle - \langle B_n z, q \rangle| \\
&= |\langle B_n(y-z), q \rangle + \langle B_n z, (q-r) \rangle| \\
&\leq \|B_n\|_{\mathcal{L}} \|q\|_2 \|y-z\|_2 + \|B_n\|_{\mathcal{L}} \|z\|_2 \|q-r\|_2 \\
&\leq \|B_n\|_{\mathcal{L}} \|y(T) - x_T\|_2 \|y-z\|_2 + \|B_n\|_{\mathcal{L}} \|x(0)\|_2 \|q-r\|_2 \\
&\leq \|B_n\|_{\mathcal{L}} 2\|x_0\|_2 \|y-z\|_2 + \|B_n\|_{\mathcal{L}} \|x_0\|_2 \|q-r\|_2 \\
&\leq \|B_n\|_{\mathcal{L}} \|x_0\|_2 (2\hat{c}_1 + \hat{c}_{\vartheta_p}) \|v-w\|_{L^2} ,
\end{aligned}$$

where $\hat{c}_1$ and $\hat{c}_{\vartheta_p}$ are given in Proposition 7 and Proposition 12. This implies that

$$\|\Theta_n(v) - \Theta_n(w)\|_{L^q} \leq T^{\frac{1}{q}} \|B_n\|_{\mathcal{L}} \|x_0\|_2 (2\hat{c}_1 + \hat{c}_{\vartheta_p}) \|v-w\|_{L^2} ,$$

which is the claim with $K = T^{\frac{1}{q}} \|B_n\|_{\mathcal{L}} \|x_0\|_2 (2\hat{c}_1 + \hat{c}_{\vartheta_p})$.

(c) First, let $h \in L^2((0,T); \mathbb{R}^{N_C})$ and denote by $\delta x(h)$ and $\delta p(h)$ the corresponding solutions to the linearized forward equation

$$\dot{\delta x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta x + \sum_{n=1}^{N_C} h_n B_n x , \quad \delta x(0) = 0 , \qquad (3.127)$$

and to the linearized backward equation

$$-\dot{\delta p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* \delta p + \sum_{n=1}^{N_C} h_n B_n^* p , \quad \delta p(T) = -\delta x(T) , \qquad (3.128)$$

respectively. According to Proposition 5 and Proposition 11 there exist positive constants K_3 and K_4 such that for any $h \in L^2((0,T); \mathbb{R}^{N_C})$ it holds that

$$\begin{aligned}
\|\delta x(h)(t)\|_2 &\leq K_3 \|h\|_{L^2} \\
\|\delta p(h)(t)\|_2 &\leq K_4 \|h\|_{L^2} ,
\end{aligned} \qquad (3.129)$$

a.e. in $(0, T)$. Next, we define the operator $\tilde{A}_n(u, \cdot) : L^2((0, T); \mathbb{R}^{N_C}) \to L^2(0, T)$ as follows

$$\tilde{A}_n(u, h) := \langle B_n \delta x(u, h), p(u) \rangle + \langle B_n x(u), \delta p(u, h) \rangle , \qquad (3.130)$$

where $\delta x(u, h)$ and $\delta p(u, h)$ are solutions to (3.127) and (3.128), respectively.

Now, we consider the following expansions with $h \in L^2((0, T); \mathbb{R}^{N_C})$ as in Proposition 6 and Proposition 12

$$x(u + h) = x(u) + \delta x(u, h) + \mathcal{O}(h^2) \qquad (3.131a)$$
$$p(u + h) = p(u) + \delta p(u, h) + \mathcal{O}(h^2) , \qquad (3.131b)$$

where we denote by $\mathcal{O}(h^2)$ the components in the expansions that are $O(\|h\|_{L^2}^2)$.

Now, by means of (3.130) and (3.131), we have

$$
\begin{aligned}
\Theta_n(u + h) - \Theta_n(u) - \tilde{A}_n(u, h) &= \langle B_n x(u + h), p(u + h) \rangle - \langle B_n x(u), p(u) \rangle \\
&\quad - \langle B_n \delta x(u, h), p(u) \rangle - \langle B_n x(u), \delta p(u, h) \rangle \\
&= \langle B_n \big(x(u) + \delta x(u, h) \big), p(u) + \delta p(u, h) \rangle \\
&\quad + \mathcal{O}(h^2) - \langle B_n x(u), p(u) \rangle \\
&\quad - \langle B_n \delta x(u, h), p(u) \rangle - \langle B_n x(u), \delta p(u, h) \rangle \\
&= \langle B_n \delta x(u, h), \delta p(u, h) \rangle + \mathcal{O}(h^2) .
\end{aligned}
\qquad (3.132)
$$

By using (3.129), we obtain

$$
\begin{aligned}
\|\Theta_n(u + h) - \Theta_n(u) - \tilde{A}_n(u, h)\|_{L^2} &= \|\langle B_n \delta x(u, h), \delta p(u, h) \rangle + \mathcal{O}(h^2)\|_{L^2} \\
&\leq \|B_n\|_{\mathcal{L}} \left(\int_0^T \|\delta x(u, h)(t)\|_2^2 \|\delta p(u, h)(t)\|_2^2 dt \right)^{1/2} + O(\|h\|_{L^2}^2) \\
&\leq \|B_n\|_{\mathcal{L}} \left(\int_0^T K_3^2 \|h\|_{L^2}^2 K_4^2 \|h\|_{L^2}^2 dt \right)^{1/2} + O(\|h\|_{L^2}^2) \\
&\leq \sqrt{T} \|B_n\|_{\mathcal{L}} K_3 K_4 \|h\|_{L^2}^2 + O(\|h\|_{L^2}^2) .
\end{aligned}
\qquad (3.133)
$$

This means that there exists a positive constant $\tilde{K}$ such that

$$\|\Theta_n(u + h) - \Theta_n(u) - \tilde{A}_n(u, h)\|_{L^2} \leq \tilde{K} \|h\|_{L^2}^2 . \qquad (3.134)$$

Consequently, the claim follows. $\qquad \square$

Notice that the hypothesis $\|x_T\|_2 = \|x_0\|_2$ that is used to prove Lemma 3, is in general satisfied for quantum control problems. Consequently, the inequality (3.125) in general holds.

Next, we investigate the behaviour of solutions to (3.107) with $\beta > 0$. Because of the L^1-norm in the cost functional J, the resulting solution u possesses a sparsity structure which is significantly different from a control given by a classical L^2-optimization as discussed in Theorem 3.1 in [126]. We state here this result for the sake of clarity.

Theorem 8. *If $u_n(t_j) u_n(t_k) < 0$ holds for two points $t_j < t_k$ in $[0, T]$, then there exist $\tilde{t}_j$ and $\tilde{t}_k$ in $[0, T]$ with $t_j < \tilde{t}_j < \tilde{t}_k < t_k$, such that $u_n = 0$ holds on $[\tilde{t}_j, \tilde{t}_k]$.*

Proof. This theorem is proved in [126] for an optimal control problem with $\nu = 0$. However, the proof holds also for our class of quantum control problems with $\nu > 0$. $\qquad \square$

This result can be interpreted as follows. If the optimal control u_n has different signs at two points, then there exists a non-empty open interval in-between where u_n vanishes. Moreover, in the particular case of bang-bang structure, u_n cannot jump directly between the lower and upper bound.

Next, we analyze the dependence of the optimal control solution u on the parameter β. In the following theorem, we show that there exists a $\hat{\beta}$ such that $u = 0$ solves (3.107) for all $\beta \geq \hat{\beta}$.

Theorem 9. *Assume that $\|x_T\|_2 = \|x_0\|_2$ and A and B_n are skew-symmetric. There exists a $\hat{\beta} > 0$ such that problem (3.107) is solved by $u = 0$ for all $\beta \geq \hat{\beta}$. In particular, an upper bound of $\hat{\beta}$ is given by*

$$\hat{\hat{\beta}} = 2\tilde{c}_1\|x_0\|_2 \, , \tag{3.135}$$

where $\tilde{c}_1$ is given in Proposition 4.

Proof. By denoting with $x^0(t)$ the evolution of the uncontrolled system ($u = 0$), we have

$$J(x, u) - J(x^0, 0) = \frac{\nu}{2}\|u\|_{L^2}^2 + \beta\|u\|_{L^1} + \frac{1}{2}\|x(T) - x_T\|_2^2 - \frac{1}{2}\|x^0(T) - x_T\|_2^2 \, . \tag{3.136}$$

We also have the following

$$
\begin{aligned}
\|x(T) - x_T\|_2^2 - \|x^0(T) - x_T\|_2^2 &= \|x(T) - x^0(T) + x^0(T) - x_T\|_2^2 - \|x^0(T) - x_T\|_2^2 \\
&= \|x(T) - x^0(T)\|_2^2 + \|x^0(T) - x_T\|_2^2 + 2\langle x(T) - x^0(T), x^0(T) - x_T\rangle - \|x^0(T) - x_T\|_2^2 \\
&= \|x(T) - x^0(T)\|_2^2 + 2\langle x(T) - x^0(T), x^0(T) - x_T\rangle \\
&\geq 2\langle x(T) - x^0(T), x^0(T) - x_T\rangle \\
&\geq -2\|x^0(T) - x_T\|_2\|x(T) - x^0(T)\|_2 \\
&\geq -2\left(\|x^0(T)\|_2 + \|x_T\|_2\right)\|x(T) - x^0(T)\|_2 \\
&\geq -4\|x_0\|_2\|x(T) - x^0(T)\|_2 \, .
\end{aligned}
\tag{3.137}
$$

Combining (3.136) with (3.137), we obtain

$$J(x, u) - J(x^0, 0) \geq \frac{\nu}{2}\|u\|_{L^2}^2 + \beta\|u\|_{L^1} - 2\|x_0\|_2\|x(T) - x^0(T)\|_2 \, . \tag{3.138}$$

By means of (3.8) in Proposition 4, the inequality (3.138) becomes

$$J(x, u) - J(x^0, 0) \geq \frac{\nu}{2}\|u\|_{L^2}^2 + \beta\|u\|_{L^1} - 2\tilde{c}_1\|x_0\|_2\|u\|_{L^1} \, . \tag{3.139}$$

Hence, if $\beta \geq 2\tilde{c}_1\|x_0\|_2$, then it holds that $J(x, u) - J(x_0, 0) \geq 0$ for any admissible pair $(u, x(u))$. $\square$

We remark that, Theorem 9 holds independently on the choice of the admissible set U_{ad}, that can be even the entire L^2 space. Next, a relationship between the L^1-norm of the optimal control and the weight parameters ν and β is investigated. For this reason, the following result is proved.

Theorem 10. *Assume that $\|x_T\|_2 = \|x_0\|_2$ and A and B_n are skew-symmetric. Then:*

(a) let x_0 be an equilibrium point, a necessary condition for $u = 0$ being a stationary point (and hence a minimum) is

$$\beta \geq \max_n \Big(\langle B_n x_0, x_T \rangle \Big) \; ; \tag{3.140}$$

(b) let $u \neq 0$ be a solution to (1.2) corresponding to $\nu > 0$ and β such that $\hat{\hat{\beta}} > \beta > 0$. Then the following holds

$$\|u\|_{L^1} \leq \frac{2 N_C T \big(\hat{\hat{\beta}} - \beta \big)}{\nu} \; , \tag{3.141}$$

where $\hat{\hat{\beta}}$ and $\hat{\beta}$ are as in Theorem 9.

Proof. (a) Assume that $u = 0$ is a stationary point. Then recalling that x_0 is an equilibrium point, (3.109) and (3.110) imply that $x(t) = x_0$ and $p(T) = -(x_0 - x_T)$. Hence, condition (3.114a) implies that for all n we have

$$-\langle B_n x_0, p(t) \rangle + \hat{\lambda}_n(t) = 0 \; , \tag{3.142}$$

for every $t \in [0, T]$. For $t = T$ we have

$$-\langle B_n x_0, -x_0 + x_T \rangle + \hat{\lambda}_n(T) = 0 \implies \hat{\lambda}_n(T) = \langle B_n x_0, x_T \rangle \implies \beta \geq \langle B_n x_0, x_T \rangle \; , \tag{3.143}$$

where we used the skew-symmetry of B_n and the conditions (3.114d)-(3.114f). The claim follows from (3.143).

(b) From the proof of Theorem 9, we have (3.139), that is,

$$J(x, u) - J(x^0, 0) \geq \frac{\nu}{2} \|u\|_{L^2}^2 + \beta \|u\|_{L^1} - 2\|x_0\|_2 \tilde{c}_1 \|u\|_{L^1} \; . \tag{3.144}$$

Using that $\|u\|_{L^1} \leq \sqrt{T N_C} \|u\|_{L^2}$ and recalling that $\hat{\hat{\beta}} = 2\|x_0\|_2 \tilde{c}_1$, for any stationary point $(u, x(u))$ we obtain that

$$\begin{aligned} J(x, u) - J(x^0, 0) &\geq \frac{\nu}{2 T N_C} \|u\|_{L^1}^2 + \beta \|u\|_{L^1} - \hat{\hat{\beta}} \|u\|_{L^1} \\ &= \Big(\frac{\nu}{2 T N_C} \|u\|_{L^1} + \beta - \hat{\hat{\beta}} \Big) \|u\|_{L^1} \; . \end{aligned} \tag{3.145}$$

Assume that $u \neq 0$ be a solution to (1.2) corresponding to $\beta < \hat{\beta}$ and $\nu > 0$, we write that

$$0 \geq J(x, u) - J(x^0, 0) \geq \Big(\frac{\nu}{2 T N_C} \|u\|_{L^1} + \beta - \hat{\hat{\beta}} \Big) \|u\|_{L^1} \; ,$$

which implies that

$$\|u\|_{L^1} \leq \frac{2 N_C T \big(\hat{\hat{\beta}} - \beta \big)}{\nu} \; .$$

$\square$

The interpretation of this result is the following. In the case that the estimate (3.8) does not hold, one can use the simple expression (3.140) in point (a) to check if $u = 0$ can be a minimum or not. Point (b) means that all the possible minimizers belong to a ball $B_{\tilde{y}}(0) \subset L^1((0, T); \mathbb{R}^{N_C})$ centered in $u = 0$ and radius $\tilde{y} = \frac{2 N_C T \big(\hat{\beta} - \beta \big)}{\nu}$. The estimate

(3.141) gives a relationship between the L^1-norm of the optimal control and the two parameters ν and β. According to (3.141) it is clear that $\|u_\beta\|_{L^1} \to 0$ linearly as $\beta \to \hat{\tilde{\beta}}$. Moreover, this implies that the measure of the region of the domain $(0,T)$ in which the control is non-zero converges to zero as $\beta \to \hat{\tilde{\beta}}$; this result is similar to a result obtained in [120] for an optimal control problem with linear control structure.

Now, we continue our analysis of the dependence of the optimal control u on the value of β, and we focus on the map $\Phi : (0,\tilde{\beta}) \to U_{ad,1} \subset L^2((0,T); \mathbb{R}^{N_C})$, with $\tilde{\beta} > 0$, defined as follows

$$\beta \mapsto \Phi(\beta) := u_\beta , \qquad (3.146)$$

where u_β solves (3.107) corresponding to β. Notice that, because of the bilinearity of the constraint (3.109), problem (3.107) is not, in general, uniquely solvable, and it can happen that for a given β there are many solutions to (3.107). Consequently, in order to have that Φ is well-defined, we need to assume that (3.107) is uniquely solvable, that is, there exists a unique $u_\beta \in U_{ad,1} \subset L^2((0,T); \mathbb{R}^{N_C})$ which solves (3.107) corresponding to β. We have the following result.

Theorem 11. *Assume that $\|x_T\|_2 = \|x_0\|_2$ and A and B_n are skew-symmetric. Consider $\tilde{\beta} > 0$, and assume that (3.107) with $\beta \in (0,\tilde{\beta})$ is uniquely solvable by $u_\beta \in U_{ad}$. Then the map $\Phi : (0,\tilde{\beta}) \to U_{ad,1}$ defined as $\beta \mapsto \Phi(\beta) := u_\beta$ is continuous. Moreover, if the weight parameter ν is large enough, then $\Phi : (0,\tilde{\beta}) \to U_{ad,1}$ is locally Lipschitz continuous.*

Proof. Consider a sequence of positive parameters $\{\beta_k\}_{k=1}^\infty$ in $(0,\tilde{\beta})$ such that $\beta_k \to \hat{\beta} \in (0,\tilde{\beta})$. Consider the corresponding sequence of optimal control functions $\{u^k\}_{k=1}^\infty$ solution to (3.107). Since $u^k \in U_{ad,1}$, with $U_{ad,1}$ closed, convex and bounded, then we can extract a weakly convergent subsequence $u^{k_j} \rightharpoonup \hat{u}$ in L^2, with $\hat{u} \in U_{ad,1}$. Denote by J_{β_j} and $J_{\hat{\beta}}$ the reduced cost functional corresponding to the elements of the sequence β_j and the limit of the sequence $\hat{\beta}$, respectively. Notice that $J_{\beta_{k_j}}(u^{k_j}) \leq J_{\beta_{k_j}}(u)$ for all $u \in U_{ad,1}$. By recalling Proposition 3, and using lower-semicontinuity of $J_{\beta_{k_j}}$, we obtain the following

$$J_{\hat{\beta}}(\hat{u}) \leq \liminf_{j\to\infty} J_{\beta_{k_j}}(u^{k_j}) \leq \liminf_{j\to\infty} J_{\beta_{k_j}}(u) , \ \forall u \in U_{ad,1}$$
$$\Rightarrow J_{\hat{\beta}}(\hat{u}) \leq J_{\hat{\beta}}(u) , \ \forall u \in L^2((0,T); \mathbb{R}^{N_C}) .$$

Hence, $\hat{u}$ is a solution to (3.107) corresponding to $\hat{\beta}$, that is $\hat{u} = \Phi(\hat{\beta})$. Since (3.107) corresponding to $\hat{\beta}$ is uniquely solvable, the limit of the sequence is unique and we obtain weakly sequentially continuity of Φ. By using Proposition 3, we obtain strong convergence of the sequences $x^k := x(u^k)$ and $p^k := p(u^k)$, that is $x^k \to \hat{x} = x(\hat{u})$ and $p^k \to \hat{p} = p(\hat{u})$. Consequently, it follows that

$$\langle B_n x^k, p^k \rangle \to \langle B_n \hat{x}, \hat{p} \rangle \ \ \text{in} \ \ C[0,T] , \qquad (3.147)$$

as $\beta^k \to \hat{\beta}$.

Now, let β_1 and β_2 be arbitrary positive parameters in $(0,\tilde{\beta})$. Let $v := \Phi(\beta_1)$ and $w := \Phi(\beta_2)$ be the corresponding solutions to (3.107), with $y = x(v)$, $z = x(w)$, $q = p(v,y)$ and $r = p(w,z)$. By means of condition (3.64c), we write that

$$\nu v_n - \langle B_n y, q \rangle + \hat{\mu}_n = 0 , \qquad (3.148)$$

and

$$\nu w_n - \langle B_n z, r \rangle + \tilde{\mu}_n = 0 , \qquad (3.149)$$

where $\hat{\mu}_n := \lambda_{+,n}(v_n) - \lambda_{-,n}(v_n) + \hat{\lambda}_n(v_n)$, and $\tilde{\mu}_n := \lambda_{+,n}(w_n) - \lambda_{-,n}(w_n) + \hat{\lambda}_n(w_n)$. Subtracting term-by-term (3.149) from (3.148) and taking the inner product with $v_n - w_n$, we obtain

$$\langle \nu(v_n - w_n) - (\langle B_n y, q \rangle - \langle B_n z, r \rangle) + (\hat{\mu}_n - \tilde{\mu}_n), v_n - w_n \rangle_{L^2} = 0 \,, \tag{3.150}$$

which implies

$$\nu \|v_n - w_n\|_{L^2}^2 = \langle (\langle B_n y, q \rangle - \langle B_n z, r \rangle), v_n - w_n \rangle_{L^2} - \langle \hat{\mu}_n - \tilde{\mu}_n, v_n - w_n \rangle_{L^2} \,. \tag{3.151}$$

From [110], we know that

$$\langle \hat{\mu}_n - \tilde{\mu}_n, v_n - w_n \rangle_{L^2} \leq \sqrt{T} |\beta_1 - \beta_2| \|v_n - w_n\|_{L^2} \,. \tag{3.152}$$

By using (3.152), we can write the following

$$\nu \|v_n - w_n\|_{L^2}^2 \leq \|\langle B_n y, q \rangle - \langle B_n z, r \rangle\|_{L^2} \|v_n - w_n\|_{L^2} + \sqrt{T} |\beta_1 - \beta_2| \|v_n - w_n\|_{L^2} \,, \tag{3.153}$$

which implies the following

$$\nu \|v_n - w_n\|_{L^2} \leq \|\langle B_n y, q \rangle - \langle B_n z, r \rangle\|_{L^2} + \sqrt{T} |\beta_1 - \beta_2| \,. \tag{3.154}$$

Continuity follows from (3.154) and (3.147).

To prove Lipschitz continuity, consider the following

$$\begin{aligned}
\|\langle B_n y, q \rangle - \langle B_n z, r \rangle\|_{L^2} &= \|\langle B_n(y - z), q \rangle - \langle B_n z, (r - q) \rangle\|_{L^2} \\
&\leq \|\langle B_n(y - z), q \rangle\|_{L^2} + \|\langle B_n z, (r - q) \rangle\|_{L^2} \\
&\leq (2\hat{c}_1 + \hat{c}_{\vartheta_p}) \sqrt{T} \|x_0\|_2 \|B_n\|_{\mathcal{L}} \|v - w\|_{L^2} \,,
\end{aligned} \tag{3.155}$$

where we used Proposition 7 and Proposition 12. By using (3.153) and (3.154), we write that

$$\begin{aligned}
\nu \|v - w\|_{L^2}^2 &\leq \sum_n \|\langle B_n y, q \rangle - \langle B_n z, r \rangle\|_{L^2} \|v_n - w_n\|_{L^2} + \sqrt{T N_C} |\beta_1 - \beta_2| \|v_n - w_n\|_{L^2} \\
&\leq (2\hat{c}_1 + \hat{c}_{\vartheta_p}) \sqrt{T} \|x_0\|_2 \|v - w\|_{L^2} \sum_n \|B_n\|_{\mathcal{L}} \|v_n - w_n\|_{L^2} \\
&\quad + \sqrt{T N_C} |\beta_1 - \beta_2| \|v_n - w_n\|_{L^2} \\
&\leq (2\hat{c}_1 + \hat{c}_{\vartheta_p}) \sqrt{T N_C} K_0 \|x_0\|_2 \|v - w\|_{L^2}^2 + \sqrt{T N_C} |\beta_1 - \beta_2| \|v_n - w_n\|_{L^2} \,,
\end{aligned} \tag{3.156}$$

where $K_0 = \max_n \|B_n\|_{\mathcal{L}}$. Equation (3.156) implies that

$$\|v - w\|_{L^2} \leq \frac{\sqrt{T N_C}}{\nu - (2\hat{c}_1 + \hat{c}_{\vartheta_p}) \sqrt{T N_C} K_0 \|x_0\|_2} |\beta_1 - \beta_2| \,, \tag{3.157}$$

that express the searched locally Lipschitz continuity for ν big enough, that is $\nu > (2\hat{c}_1 + \hat{c}_{\vartheta_p}) \sqrt{T N_C} K_0 \|x_0\|_2$. This completes our proof.

$\square$

For convenience, we conclude this section summarizing the properties of solutions to (3.107) discussed in this section. Moreover, we compare our results, which are derived for the optimal control of bilinear systems, with the ones presented by Stadler in [110], that

are proved for the optimal control of linear elliptic differential equations. Notice that in (3.107) there are three sources of non-linearity, that are, the bilinear control structure, the control constraints, and the L^1-penalization. We see that, the L^1-penalization in the cost does not increase significantly the non-linearity of the problem, that is mostly due to the other terms. This fact makes our problem significantly different from the linear problem considered in [110], where the only non-linearity arises from the L^1-control cost. Moreover, an additional difference between our problem and the one in [110] is due to the different cost functionals. In our case, the tracking term corresponds to terminal observation while in [110], a tracking of configuration is involved. This fact makes analysis and solution of our problem more difficult, because the map $u \mapsto x(u)$ is injective, but the end-point map $u \mapsto x(u)(T)$ is not.

- Theorem 8 states a sparsity property of the optimal control, that is, if the optimal control has a different sign at two points, then there exists a non-empty open interval in-between where the control is zero.

- In Theorem 9, we prove that for sufficiently large β, problem (3.107) has the trivial solution $u = 0$. This result is similar to Lemma 3.1 in [110].

- Theorem 10 shows that all the possible solutions to (3.107) are bounded according to (3.141) and contained in a ball centered in $u = 0$ with radius depending on ν and β.

- In Theorem 11, we prove that the map $\Phi : \beta \mapsto u_\beta$ is Lipschitz continuous. This result is weaker than Lemma 3.2 in [110]. In fact, because of the bilinearity, in order to guarantee that Φ is well-defined, we need to assume uniqueness of the solution and study the map $\beta \mapsto u_\beta$. Moreover, we need a strong regularity assumption on ν. Notice that, this assumption is used also in the sequel for obtaining coercivity of reduced Hessian operator in Lemma 4 and the reduced generalized Jacobian operator in Lemma 6.

3.5   Summary and remarks

In this chapter, optimal control problems governed by quantum spin dynamical systems were considered. These optimization problems were characterized by a possibly non-smooth cost functional and sets of admissible controls with different features, like point-wise bounded and piecewise-constant functions. In particular, the chapter focused on different aspects of quantum optimal control problems, as follows. First, several characterization properties of quantum spin dynamical systems were proved. These were used in this chapter as well as in the sequel of the thesis, and are important for proving some if the main results presented in this work. Second, existence and characterization properties of quantum optimal control solutions were discussed, especially in the case of L^1-penalized cost functional. Third, first-order optimality conditions were derived.

The novelties that characterize this chapter are:

- A characterization of optimal control solutions in the case of L^1-penalized cost functional. In the literature of L^1-optimization, optimal control problems governed by linear and semi-linear state constraints are investigated. On the other hand, much less is known in the case of bilinear control-state constraints. Moreover, the L^1-optimization approach is an absolute novelty in the field of quantum optimal

control. In fact, the corresponding mathematical literature considers usually more regular optimization problems.

- A rigorous mathematical description and characterization of quantum optimal control problems. This rigorous mathematical approach is not considered in the literature of quantum control problems, where physical purposes are dominant.

Chapter 4

Newton methods for the optimal control of quantum systems

Optimal control theory for quantum systems has the purpose to design fast control mechanisms that cannot be constructed simply based on perturbation theory [129]. This fact motivates the application of optimal control methods and related computational techniques to quantum systems [15, 59, 73, 88]; see [14] for a review. However, in most of the works on quantum optimal control problems, the focus is on numerical optimization techniques, that allow to compute the required control functions. Pioneering results in the development of quantum optimal control algorithms can be found in [76, 115, 133]. Further progress in the development of efficient control schemes is documented in, e.g., [17, 18, 44, 59, 73, 88, 89, 91, 106]. In these references, only first-order optimization methods were considered, while we are not aware of works on second-order Newton methods for the control of quantum spin systems. On the other hand, a discussion of a Krylov-Newton scheme for solving optimal spin-less quantum control problems governed by the infinite-dimensional Schrödinger equation can be found in [129, 130]. Moreover, in applications the need arises to constrain the control functions pointwise in time; see, e.g., [75, 105]. This requirement results in a lack of differentiability of the control problem so that a straightforward application of the Newton method is not possible. To overcome this limitation, in the field of PDE-constraint optimization problems, a semi-smooth Newton (SSN) method was developed to solve control-constrained problems; see, e.g., [56, 58, 78, 79, 80, 123].

In this chapter, we discuss Krylov-Newton methods for the solutions of possibly non-smooth optimal quantum control problems. The chapter is organized as follows. In Section 4.1 we provide a general introduction to Newton method and semi-smooth Newton method. This introductory discussion has the purpose to describe the main features of the Newton and SSN method in the case of finite-dimensional problems. These features are then addressed for infinite-dimensional problems in the other sections. In Section 4.2, first we describe Krylov-Newton methods for the solution to optimal control problems. Then, we show how to address our quantum optimal control problems by means of Krylov-Newton methods, and we prove convergence in Section 4.2.1. In Section 4.3, semi-smooth Krylov-Newton method is discussed in order to solve our non-smooth quantum optimal control problem characterized by a L^1-penalized cost functional and control constraint given by $U_{ad,1}$. Convergence of the SSN method for the solution to quantum control problem is proved in Section 4.3.1. In Section 4.3.2, we show how to use a SSN method in the case that the admissible set of controls is $U_{ad,2}$. A summary section concludes the chapter.

4.1 An introduction to Newton and semi-smooth Newton methods

The Newton method is an iterative procedure aimed at finding roots or minima of a given function. The peculiarity of this procedure is that it generates a sequence that can be proved to be quadratically convergent to the sought solution. For this reason, the use of the Newton method is suitable to solve root and optimization problems faster and more accurately than first-order methods, like the steepest-descent method, and other fixed point methods. On the other hand, particular conditions have to be satisfied to guarantee the fast convergence of the Newton method, and in many applications its implementation has to be carefully addressed for obtaining practical realizations.

In order to explain the basic ideas behind the Newton method and its convergence, consider the problem to find a root $\eta^* \in \mathbb{R}^N$ such that

$$\mathcal{F}(\eta^*) = 0 \,, \tag{4.1}$$

where $\mathcal{F} : \mathbb{R}^N \to \mathbb{R}^N$ is continuously differentiable, that is $\mathcal{F} \in C^1(\mathbb{R}^N)$. Denote by $\mathcal{J}(\eta)$ the Jacobian of $\mathcal{F}$ at η. The Newton method generates a sequence $\{\eta^k\}_{k=1}^\infty$ by means of the following two steps; see, e.g., [94, 112];

$$
\begin{aligned}
\text{S}_1: \quad & \delta\eta^k = -(\mathcal{J}(\eta^k))^{-1}\mathcal{F}(\eta^k) \\
\text{S}_2: \quad & \eta^{k+1} = \eta^k + \delta\eta^k \,.
\end{aligned}
\tag{4.2}
$$

Notice that the element η^{k+1} of the Newton sequence is well-defined if the Jacobian $\mathcal{J}(\eta^k)$ is invertible and the step S_1 is feasible. This condition is strict, and in practice, due to the non-linearity of $\mathcal{F}$, it can be difficult to show that $\mathcal{J}(\eta)$ is invertible for any $\eta \in \mathbb{R}^N$. More reasonable are the following assumptions [94, 112]: the Jacobian matrix is invertible in a neighbourhood $\mathcal{N}$ of the solution η^*, and there exists a positive constant K such that for any $\eta \in \mathcal{N}$ it holds that $\|(\mathcal{J}(\eta))^{-1}\|_\mathcal{L} \leq K$. Because of these assumptions, η^{k+1} is well defined if $\eta^k \in \mathcal{N}$ and we write the following

$$
\begin{aligned}
\|\eta^{k+1} - \eta^*\|_2 &= \|\eta^k - (\mathcal{J}(\eta^k))^{-1}\mathcal{F}(\eta^k) - \eta^*\|_2 \\
&= \|(\mathcal{J}(\eta^k))^{-1}\big(\mathcal{J}(\eta^k)(\eta^k - \eta^*) - \mathcal{F}(\eta^k)\big)\|_2 \\
&= \|(\mathcal{J}(\eta^k))^{-1}\big(\mathcal{J}(\eta^k)(\eta^k - \eta^*) + \mathcal{F}(\eta^*) - \mathcal{F}(\eta^k)\big)\|_2 \\
&\leq K\|\mathcal{F}(\eta^*) - \mathcal{F}(\eta^k) - \mathcal{J}(\eta^k)(\eta^* - \eta^k)\|_2 \,,
\end{aligned}
\tag{4.3}
$$

where we used that $\mathcal{F}(\eta^*) = 0$. Since it holds that

$$\mathcal{F}(\eta^*) - \mathcal{F}(\eta^k) = \int_0^1 \mathcal{J}(\eta^* + t(\eta^k - \eta^*))(\eta^* - \eta^k)dt \,, \tag{4.4}$$

we have the following

$$
\begin{aligned}
\|\eta^{k+1} - \eta^*\|_2 &\leq K\left\|\int_0^1 \mathcal{J}(\eta^* + t(\eta^k - \eta^*))(\eta^* - \eta^k)dt - \mathcal{J}(\eta^k)(\eta^* - \eta^k)\right\|_2 \\
&= \left\|\int_0^1 \big[\mathcal{J}(\eta^* + t(\eta^k - \eta^*)) - \mathcal{J}(\eta^k)\big](\eta^* - \eta^k)dt\right\|_2 \\
&\leq \int_0^1 \|\mathcal{J}(\eta^* + t(\eta^k - \eta^*)) - \mathcal{J}(\eta^k)\|_\mathcal{L}\|\eta^* - \eta^k\|_2 dt \,.
\end{aligned}
\tag{4.5}
$$

Now, by assuming that the Jacobian is Lipschitz continuous in $\mathcal{N}$, that is

$$\|\mathcal{J}(\hat{\eta}) - \mathcal{J}(\eta)\|_{\mathcal{L}} \leq L\|\hat{\eta} - \eta\|_2 \, , \ \forall \hat{\eta}, \eta \in \mathcal{N} \tag{4.6}$$

with $L > 0$, we obtain that

$$\|\eta^{k+1} - \eta^*\|_2 \leq KL\|\eta^k - \eta^*\|_2^2 \, . \tag{4.7}$$

Using this inequality inductively, we deduce that, if the starting point η^1 is sufficiently near η^*, the sequence $\{\eta^k\}_{k=1}^{\infty}$ generated by the Newton method converges quadratically to η^*. We remark that, one can consider the following assumption

$$\|\mathcal{F}(\eta + \delta\eta) - \mathcal{F}(\eta) - \mathcal{J}(\eta)(\delta\eta)\|_2 = O(\|\delta\eta\|_2^2) \quad \text{as} \quad \delta\eta \to 0 \, , \tag{4.8}$$

that, together with (4.3), still guarantees quadratic convergence of the Newton sequence. Furthermore, in the case that

$$\|\mathcal{F}(\eta + \delta\eta) - \mathcal{F}(\eta) - \mathcal{J}(\eta)(\delta\eta)\|_2 = o(\|\delta\eta\|_2) \quad \text{as} \quad \delta\eta \to 0 \, , \tag{4.9}$$

the Newton sequence converges superlinearly. Notice that (4.9) represents the Fréchet differentiability condition for $\mathcal{F}$ at η. We remark that, in finite-dimensional spaces it coincides with the classical concept of differentiability. Hence, Fréchet differentiability guarantees superlinear convergence of the Newton sequence. Notice that the convergence discussed in this chapter is known as Q-convergence; see, e.g., [57, 94, 123].

The Newton method is used also to solve minimization problems. Consider the following problem

$$\min_{\eta \in \mathbb{R}^N} \ f(\eta) \, , \tag{4.10}$$

where $f : \mathbb{R}^N \to \mathbb{R}$ is twice differentiable, and denote by η^* a stationary point, that is $\nabla f(\eta^*) = 0$. The Newton method described above can be applied to (4.10) by setting $\mathcal{F}(\eta) := \nabla f(\eta)$, which means that we aim to find a root η^* of the gradient of f. Obviously we have that $\mathcal{J}(\eta) = \nabla^2 f(\eta)$, that is the Jacobian of $\mathcal{F}$ coincides with the Hessian of f. In this optimization framework, the following theorem [94] can be proved with the same arguments as in the above discussion.

Theorem 12. *Suppose that f is twice differentiable. Assume that in a neighbourhood $\mathcal{N}$ of a stationary point η^* for (4.10) the Hessian $\nabla^2 f(\eta)$ is Lipschitz continuous and invertible with bounded inverse. Consider the Newton sequence $\{\eta^k\}_{k=1}^{\infty}$. Then*

1. *if the staring point η_1 is sufficiently close to η^*, the Newton sequence converges to η^*;*

2. *the rate of convergence of $\{\eta^k\}_{k=1}^{\infty}$ is quadratic;*

3. *the sequence of gradient norms $\{\|\nabla f(\eta^k)\|_2\}_{k=1}^{\infty}$ converges quadratically to zero.*

Next, we focus on the semi-smooth Newton method. In many applications, the hypothesis of differentiability of the map $\eta \mapsto \mathcal{F}(\eta)$ fails. Consequently, the Jacobian does not exist in a classical sense, and the Newton method described above cannot be applied. For this reason, the notion of differentiability has been generalized. The main assumption for this generalization is the Lipschitz continuity of the map $\eta \mapsto \mathcal{F}(\eta)$, which allows to use the Rademacher's theorem stated in the following; see, e.g., [34, 56, 57, 123].

Theorem 13 (Rademacher's Theorem). *If U is an open subset of $\mathbb{R}^N$ and $\mathcal{F} : U \to \mathbb{R}^M$ is Lipschitz continuous. Then $\mathcal{F}$ is almost everywhere (in the sense of Lebesgue measure) differentiable in U, that is the set $U_{nd} \subset U$ of all points at which $\mathcal{F}$ fails to be differentiable is a Lebesgue null set.*

Using this result, the notion of differentiability can be extended and generalized Jacobians can be constructed. We have the following definition [34, 57, 66, 123].

Definition 2. *Let $\mathcal{F}$ be locally Lipschitz. We define the following generalized Jacobians of $\mathcal{F}$ at η:*

(a) *the Bouligand subdifferential:*

$$\partial_B \mathcal{F}(\eta) := \left\{ S \in \mathbb{R}^{N \times M} \ : \ \exists \{\eta^k\}_{k=1}^{\infty} \subset \mathbb{R}^N \setminus U_{nd} \ : \ \eta^k \to \eta \,, \ \mathcal{J}(\eta^k) \to S \right\} ,$$

where U_{nd} is defined in Theorem 13 and $\mathcal{J}(\eta)$ denotes the Jacobian of $\mathcal{F}$ at η;

(b) *the Clarke's subdifferential is the convex hull of $\partial_B \mathcal{F}(\eta)$:*

$$\partial \mathcal{F}(\eta) := \mathrm{co}\ \partial_B \mathcal{F}(\eta) .$$

(c) *Qi's C-subdifferential: $\partial_C \mathcal{F}(\eta) := \partial \mathcal{F}_1(\eta) \times \cdots \times \partial \mathcal{F}_M(\eta)$.*

By means of these notions of differentiability, the Newton method given in (4.2) can be generalized as follows [97, 123]

$$
\begin{aligned}
\mathrm{S}_0 : \quad & \text{choose a generalized Jacobian } \mathcal{J}_g(\eta^k) \in \partial \mathcal{F}(\eta^k) \\
\mathrm{S}_1 : \quad & \delta\eta^k = -(\mathcal{J}_g(\eta^k))^{-1} \mathcal{F}(\eta^k) \\
\mathrm{S}_2 : \quad & \eta^{k+1} = \eta^k + \delta\eta^k ,
\end{aligned}
\tag{4.11}
$$

where we denote by $\mathcal{J}_g$ a generalized Jacobian. Now, we are interested in studying the convergence of the Newton sequence $\{\eta^k\}_{k=1}^{\infty}$ generated by (4.11). To this purpose, based on the Clarke's subdifferential, the following notion of semi-smoothness is introduced; see [97, 123].

Definition 3. *A function $\mathcal{F} : \mathbb{R}^N \to \mathbb{R}^M$ is said to be semi-smooth at η if and only if all the following conditions hold*

- *$\mathcal{F}$ is locally Lipschitz continuous;*
- *$\mathcal{F}$ is directionally differentiable at η;*
- *$\mathcal{F}$ satisfies the following condition*

$$\max_{S \in \partial \mathcal{F}(\eta + \delta\eta)} \ \|\mathcal{F}(\eta + \delta\eta) - \mathcal{F}(\eta) - S(\delta\eta)\|_2 = o(\|\delta\eta\|_2) \quad as \quad \delta\eta \to 0 . \tag{4.12}$$

Notice that condition (4.12) resembles equation (4.9), and semi-smoothness is used to guarantee convergence of the sequence $\{\eta^k\}_{k=1}^{\infty}$ given by (4.11). In particular, following arguments similar to the ones used for proving convergence of (4.2), the following result can be obtained [97, 98, 123].

Theorem 14. *Let η^* be a solution to $\mathcal{F}(\eta) = 0$. Assume that $\mathcal{F}$ is semi-smooth at η^* and every $S \in \partial \mathcal{F}(\eta^*)$ is invertible with $\|S^{-1}\|_{\mathcal{L}} \leq K$ for $K > 0$. Let the initial point η_1 be sufficiently close to η^*, then (4.11) generates a sequence $\{\eta^k\}_{k=1}^{\infty}$ that converges superlinearly to η^*.*

This result guarantees convergence of the semi-smooth Newton method, that can be successfully applied to solve nonlinear root and minimization problems in the case that the function $\mathcal{F}$ (the gradient ∇f) is non-differentiable in the classical sense. Moreover, we remark that, in the case that $\mathcal{F} \in C^1$ at η, then $\mathcal{F}$ is Lipschitz at η and $\partial \mathcal{F}(\eta) = \{\mathcal{J}(\eta)\}$ [34], and the semi-smooth Newton method coincide with the classical Newton method.

The discussion presented in this section concerns Newton and semi-smooth Newton methods for solving root and optimization problems defined in finite-dimensional vector spaces. This provides a general guideline that can be applied also to problems defined in infinite-dimensional spaces. However, in these cases other mathematical and computational problems arise. These problems are addressed in the following sections, where we discuss Newton and semi-smooth Newton methods for the optimal control of quantum spin systems.

4.2 Newton methods for quantum optimal control problems

In this section, we discuss the Newton method for the solution of quantum optimal control problems. The main difficulty in the implementation of the Newton method for optimal control problems comes from the dimension of the Hessian operator. In a discrete form the Hessian is a very large matrix, and this fact makes its construction infeasible from a practical point of view. This problem is avoided by using inexact Krylov-Newton methods, that combine the Newton method with Krylov iterative solvers for linear systems. These methods are also known in the literature as Newton-Krylov methods; see, e.g., [21, 43]. In this way, on one hand, the solution of the Newton linear system is not computed exactly, because the iteration procedure is truncated when a given tolerance is reached. On the other hand, Krylov iterative solvers allow to avoid the construction of the discrete Hessian by supplying a routine that returns as output a vector representing the action of the Hessian operator. In a continuous picture, this means that the optimal control problem is tackled in the space of solutions to constraint and adjoint equations, and corresponding linearized systems.

In the sequel, first we provide a general description of Krylov-Newton methods for general optimal control problems. Then we apply explicitly this class of Newton methods to the quantum control problem (3.42) with $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$. Furthermore, in Section 4.2.1, we study convergence of Krylov-Newton methods for the solution to the quantum optimal control problem (3.42).

Consider the following general optimal control problem

$$
\begin{aligned}
\min_{x,u} \quad & J(x,u) \\
\text{s.t.} \quad & c(x,u) = 0 \\
& x \in X \ , \ u \in U \ ,
\end{aligned}
\tag{4.13}
$$

where X and U are assumed to be Hilbert spaces, x and u are referred to as state and control, and $c(x,u) = 0$ represents the state constraint, that is in general a differential equation. The cost functional $J : X \times U \mapsto \mathbb{R}$ is assumed to be bounded from below, weakly lower semicontinuous and coercive with respect to u. In our discussion, the maps $(x,u) \mapsto J(x,u)$ and $(x,u) \mapsto c(x,u)$ are twice Fréchet differentiable. Moreover, we consider that the state constraint $c(x,u) = 0$ is uniquely solvable, and satisfies regularity

conditions in order to guarantee the existence of a Lagrange multiplier p in an appropriate Hilbert space P. From these assumptions it follows that (4.13) is solvable. In this framework, we consider the following Lagrange function

$$L(x, u, p) = J(x, u) + \langle p, c(x, u) \rangle_P \, , \tag{4.14}$$

where $\langle \cdot, \cdot \rangle_P$ denotes the inner product for P. A solution to (4.13), is characterized by the following first-order necessary optimality system

$$
\begin{aligned}
\nabla_x L(x, u, p) &= 0 \\
\nabla_u L(x, u, p) &= 0 \\
c(x, u) &= 0 \, ,
\end{aligned}
\tag{4.15}
$$

where $\nabla_x L$ and $\nabla_u L$ are the Fréchet derivatives of L with respect to x and u, respectively. We remark that (4.15) coincides with the result given in Theorem A 15 in the appendix. The first equation in (4.15), $\nabla_x L(x, u, p) = 0$, is the adjoint equation, that is assumed to be uniquely solvable for given x and u.

We aim to solve system (4.15) by means of the Newton method, by rewriting it as a root problem. To this purpose, let $\eta := (x, u, p)$ and define the following map [16, 90, 14]

$$\mathcal{F}(\eta) = \begin{pmatrix} \nabla_x L(x, u, p) \\ \nabla_u L(x, u, p) \\ c(x, u) \end{pmatrix} \, . \tag{4.16}$$

Notice that a root $\eta^* = (x^*, u^*, p^*)$ of $\mathcal{F}$ corresponds to a stationary point of (4.13). The Newton procedure is applied to the equation $\mathcal{F}(\eta) = 0$. Hence, similarly to (4.2), starting with some initial value η^1, we construct the Newton sequence $\{\eta^k\}_{k=1}^\infty$ as follows

$$
\begin{aligned}
S_1 : \quad & \mathcal{J}(\eta^k)(\delta\eta) = -\mathcal{F}(\eta^k) \\
S_2 : \quad & \eta^{k+1} = \eta^k + \delta\eta \, .
\end{aligned}
\tag{4.17}
$$

The Jacobian operator $\mathcal{J}(\eta)$ is given explicitly by

$$\mathcal{J} = \begin{pmatrix} \nabla_{xx}L & \nabla_{xu}L & \left(\nabla_x c\right)^* \\ \nabla_{ux}L & \nabla_{uu}L & \left(\nabla_u c\right)^* \\ \nabla_x c & \nabla_u c & 0 \end{pmatrix} \, ,$$

where we omit the dependence on $\eta = (x, u, p)$ for brevity. Notice that the Jacobian operator $\mathcal{J}$ is obtained by means of the second-Fréchet derivatives of L, and consequently coincides with the Hessian operator of the Lagrange function corresponding to (4.13). Moreover, since J and c are assumed to be twice Fréchet differentiable, the Lagrange function (4.14) is also twice Fréchet differentiable, and consequently, its Hessian has to be symmetric [23].

Next, we discuss Krylov-Newton methods. For this purpose, recall that first and third equations of (4.15) are uniquely solvable for a given control u. This means also, that the state and the adjoint equations are uniquely solvable for a given control u, that is $x = x(u)$ and $p = (u, x(u))$. For this reason, it is possible to introduce the so-called reduced cost functional, given by the following; see, e.g., [16, 119, 129];

$$J_r(u) := J(x(u), u) \, ,$$

and the corresponding reduced problem is given by

$$
\min_u \ J_r(u)
$$
$$
\text{s.t.} \ \ u \in U \ .
$$

(4.18)

Notice that (4.18) is, in contrast to (4.13), an unconstrained optimization problem. If u^* is a solution to (4.18), then the pair $(x(u^*), u^*)$ solves (4.13), see, e.g., [16, 119, 129]. The reduced gradient is defined as

$$
\mathcal{F}_r(u) := \nabla_u L\big(x(u), u, p(u, x(u))\big) \ .
$$

To obtain the reduced Hessian operator, we assume that the linearized constraint equation is solvable, that is, the operator $(\nabla_x c)$ is invertible. Hence, the following holds

$$
(\nabla_x c)(\delta x) + (\nabla_u c)(\delta u) = 0 \Rightarrow \delta x = -(\nabla_x c)^{-1}\big((\nabla_u c)(\delta u)\big) \ ,
$$

(4.19)

and the Hessian operator in the reduced form is given by the following; see, e.g., [14];

$$
\mathcal{J}_r(u) = \begin{pmatrix} -(\nabla_x c)^{-1}(\nabla_u c) \\ I \end{pmatrix}^* \begin{pmatrix} \nabla_{xx} L & \nabla_{xu} L \\ \nabla_{ux} L & \nabla_{uu} L \end{pmatrix} \begin{pmatrix} -(\nabla_x c)^{-1}(\nabla_u c) \\ I \end{pmatrix} \ .
$$

(4.20)

The action of $\mathcal{J}_r(u)$ on a function δu corresponds to solving the linearized constraint equation as in (4.19), then solving the linearized adjoint, that is

$$
(\nabla_{xx} L)(\delta x) + (\nabla_{xu} L)(\delta u) + (\nabla_x c)^*(\delta p) = 0 \ ,
$$

(4.21)

and considering the following definition

$$
\mathcal{J}_r(u)(\delta u) := (\nabla_{ux} L)(\delta x) + (\nabla_{uu} L)(\delta u) + (\nabla_u c)^*(\delta p) \ .
$$

(4.22)

To clarify formulas (4.20) and (4.22), let $\delta \eta = (\delta x, \delta u, \delta p)^T$ and consider the Newton linear system in (4.17), that is

$$
\begin{pmatrix} \nabla_{xx} L & \nabla_{xu} L & (\nabla_x c)^* \\ \nabla_{ux} L & \nabla_{uu} L & (\nabla_u c)^* \\ \nabla_x c & \nabla_u c & 0 \end{pmatrix} \begin{pmatrix} \delta x \\ \delta u \\ \delta p \end{pmatrix} = - \begin{pmatrix} \nabla_x L \\ \nabla_u L \\ c \end{pmatrix} \ .
$$

(4.23)

In the space of solutions to constraint $(c(x, u) = 0)$ and adjoint equation $(\nabla_x L(x, u, p) = 0)$, the Newton linear system (4.23) becomes as follows

$$
\begin{pmatrix} \nabla_{xx} L & \nabla_{xu} L & (\nabla_x c)^* \\ \nabla_{ux} L & \nabla_{uu} L & (\nabla_u c)^* \\ \nabla_x c & \nabla_u c & 0 \end{pmatrix} \begin{pmatrix} \delta x \\ \delta u \\ \delta p \end{pmatrix} = - \begin{pmatrix} 0 \\ \mathcal{F}_r \\ 0 \end{pmatrix} \ ,
$$

(4.24)

and we notice that solving the third equation corresponds to solve the linearized constraint equation as in (4.19). Then solving the first equation corresponds to solve the linearized adjoint equation (4.21). Now, in the space of linearized constraint and adjoint equations, it remains to solve the second equation, that is

$$
(\nabla_{ux} L)(\delta x) + (\nabla_{uu} L)(\delta u) + (\nabla_u c)^*(\delta p) = -\mathcal{F}_r \ ,
$$

(4.25)

and we recognize that the left-hand side of (4.25) is $\mathcal{J}_r(u)(\delta u)$, in agreement with (4.22). Consequently, the Newton method in a reduced form is given by

$$\mathcal{J}_r(u^k)(\delta u^k) = -\mathcal{F}_r(u^k)$$
$$u^{k+1} = u^k + \delta u^k \ ,$$

(4.26)

and the solution of the Newton linear system is obtained by means of a Krylov linear solver supplied with a routine that uses equations (4.19)-(4.21)-(4.22) for assembling the action of the reduced Hessian on a given function.

Next, we discuss explicitly the Krylov-Newton method for the solution of the quantum optimal control problem (3.42) with $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$. The function $\mathcal{F}(\eta)$ is given by the following

$$\mathcal{F}(\eta) := \begin{pmatrix} \tilde{c}(x,u,p) \\ \nabla_u L(x,u,p) \\ c(x,u) \end{pmatrix} \ ,$$

(4.27)

where $c(x,u) = 0$ is the state constraint, that is

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \quad x(0) = x_0 \ ,$$

(4.28)

and $\tilde{c}(x,u,p) = 0$ is the following corresponding adjoint equation

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* p \ , \quad p(T) = -(x(T) - x_T) \ .$$

(4.29)

The reduced form $\mathcal{F}_r(u)$ is obtained as follows. First, we solve the constraint equation $c(x,u) = 0$ and the adjoint equation $\tilde{c}(x,u,p) = 0$. Then, in the space of solutions to (4.28) and (4.29), we have that $\nabla_u L(x,u,p) = \nabla_u J_r(u)$, where the reduced gradient, derived in Lemma 2, is $\nabla_{u_n} J_r(u) = \nu u_n - \langle B_n x, p \rangle$. Consequently, we write that

$$\left(\mathcal{F}_r(u) \right)_n = \nu u_n - \langle B_n x, p \rangle \ .$$

(4.30)

The Hessian operator is obtained as follows. Differentiating the optimality system (3.53), or equivalently differentiating two times the Lagrange function (3.47), we obtain the following Hessian operator of problem (3.45) as follows

$$\mathcal{J}(x,u,p) = \begin{pmatrix} 0 & -B_1^* p & \cdots & \cdots & -B_{N_C}^* p & c(\cdot,u)^* \\ (-B_1^* p)^* & \nu I & 0 & \cdots & 0 & (-B_1 x)^* \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ (-B_{N_C}^* p)^* & 0 & \cdots & 0 & \nu I & (-B_{N_C} x)^* \\ c(\cdot,u) & -B_1 x & \cdots & \cdots & -B_{N_C} x & 0 \end{pmatrix} \ ,$$

(4.31)

where I is the identity operator. From $\mathcal{J}(x,u,p)$ we construct the reduced Hessian operator and its action on a vector function δu. For this purpose, let $\delta x = \delta x(\delta u)$ and $\delta p = \delta p(\delta u)$ be the unique solutions to linearized constraint and adjoint equations, that are given by

$$\dot{\delta x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta x + \sum_{n=1}^{N_C} (\delta u_n B_n) x \ , \quad \delta x(0) = 0$$

(4.32)

and

$$-\dot{\delta p} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]^* \delta p + \sum_{n=1}^{N_C} (\delta u_n B_n)^* p \ , \ \delta p(T) = -\delta x(T) \ . \qquad (4.33)$$

The action of $\mathcal{J}_r(u)$ on δu can be defined as follows

$$(\mathcal{J}_r(u)(\delta u))_n := \nu \delta u_n - \langle B_n x, \delta p(\delta u)\rangle - \langle B_n^* p, \delta x(\delta u)\rangle \ , \ \ n = 1, ..., N_C \ . \qquad (4.34)$$

The Krylov-Newton linear system, that is $\mathcal{J}_r(u)(\delta u) = -\mathcal{F}_r(u)$, is explicitly given by the following

$$\nu \delta u_n - \langle B_n x, \delta p(\delta u)\rangle - \langle B_n^* p, \delta x(\delta u)\rangle = -\left(\nu u_n - \langle B_n x, p\rangle\right) \ , \qquad (4.35)$$

with $n = 1, ..., N_C$. This linear system is solved by using a Krylov iterative solver, and a crucial advantage is the possibility to code this algorithm in a matrix-free way. This is essential to solve multiple spin systems with large N_p. The discrete Hessian operator is not stored and the overall algorithm is supplied by a procedure computing directly the action of the Hessian operator on a vector function [16]. This action corresponds to the solution to the linearized optimality system. Notice that the proposed procedure needs to be globalized by means of a line-search strategy. This and other implementation details are discussed in Section 6.

4.2.1 Convergence of the Krylov-Newton method for quantum optimal control

In this section, we discuss a convergence result of the Newton sequence in infinite-dimensional spaces, and study convergence of the Krylov-Newton method for solving the quantum control problem (3.42) with $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$.

The discussion presented in Section 4.1 can be extended to problems defined on general Banach spaces, and the following convergence result holds; see, e.g., [58, 90].

Theorem 15. *Consider the map* $\mathcal{F}_r : U \to Y$, *where* U *and* Y *are Banach spaces and let* $u^* \in U$ *such that* $\mathcal{F}_r(u^*) = 0$. *Assume that*

- $\mathcal{F}_r$ *is continuously Fréchet differentiable, that is*

$$\|\mathcal{F}_r(u + \delta u) - \mathcal{F}_r(u) - \mathcal{J}_r(u)(\delta u)\|_Y = o(\|\delta u\|_U) \quad as \quad \delta u \to 0 \ , \qquad (4.36)$$

  with $\delta u \mapsto \mathcal{J}_r(u)(\delta u)$ *continuous;*

- $\mathcal{J}_r(u)$ *is invertible in a neighbourhood of* u^*, *and the inverse is bounded in an appropriate operator norm.*

Consider the Newton sequence $\{u^k\}_{k=1}^{\infty}$ *generated by (4.26), that is*

$$\mathcal{J}_r(u^k)(\delta u^k) = -\mathcal{F}_r(u^k)$$
$$u^{k+1} = u^k + \delta u^k \ .$$

If u^1 *is sufficiently close to the solution* u^*, *then the Newton sequence* $\{u^k\}_{k=1}^{\infty}$ *converges locally superlinearly to* u^*. *Moreover, if in addition it holds that* $\mathcal{J}_r(u)$ *is* α-*Hölder continuous near* u^*, *or*

$$\|\mathcal{F}_r(u + \delta u) - \mathcal{F}_r(u) - \mathcal{J}_r(u)(\delta u)\|_Y = O(\|\delta u\|_U^{1+\alpha}) \quad as \quad \delta u \to 0 \ , \qquad (4.37)$$

then the order of convergence is $1 + \alpha$.

Next, we study convergence of the Newton method for the solution to the quantum spin control problem (3.42) with $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$. In particular, in order to apply Theorem 15, we set $U = Y = L^2((0,T); \mathbb{R}^{N_C})$, and we study regularity of the Jacobian $\mathcal{J}_r$. Due to the bilinearity of the quantum control problem (3.42), it is difficult to address directly invertibility of the Jacobian. For this reason, in Lemma 4, we investigate coercivity of $\mathcal{J}_r$ and derive a sufficient condition that guarantees this property. We have

Lemma 4. *The reduced Jacobian $\mathcal{J}_r$ is coercive if the weight parameter ν is large enough and the tracking term $\|x(T) - x_T\|_2$ is small enough. In particular, if*

$$\alpha := \left[\nu - 4K_{00}^2 T N_C \|x_0\|_2 \|x(T) - x_T\|_2 \right] > 0 , \tag{4.38}$$

with K_{00} as in Proposition 5, then $\mathcal{J}_r$ is coercive and it holds that

$$\langle \mathcal{J}_r(u)(h), h \rangle_{L^2} \geq \|\delta x(T; h)\|_2^2 + \alpha \|h\|_{L^2}^2 \quad \text{for any} \quad h \in L^2((0,T); \mathbb{R}^{N_C}) , \tag{4.39}$$

where δx is the solution to (4.32) corresponding to h.

Proof. Let $h \in L^2((0,T); \mathbb{R}^{N_C})$, and denote by $\delta x = \delta x(h)$ and $\delta p = \delta p(h)$ the corresponding solutions to linearized constraint (4.32) and adjoint (4.33). Recalling the action of the reduced Jacobian as in (4.34), we write the following

$$\begin{aligned}
\langle \mathcal{J}_r(u)(h), h \rangle_{L^2} &= \sum_n \langle \nu h_n - \langle B_n x, \delta p \rangle - \langle B_n \delta x, p \rangle, h_n \rangle_{L^2} \\
&= \nu \|h\|_{L^2}^2 - \sum_n \int_0^T h_n \big(\langle B_n x, \delta p \rangle + \langle B_n \delta x, p \rangle \big) dt \\
&= \nu \|h\|_{L^2}^2 - \sum_n \int_0^T \big(\langle h_n B_n x, \delta p \rangle + \langle h_n B_n \delta x, p \rangle \big) dt \\
&= \nu \|h\|_{L^2}^2 - \int_0^T \big(\langle \sum_n h_n B_n x, \delta p \rangle + \langle \sum_n h_n B_n \delta x, p \rangle \big) dt .
\end{aligned} \tag{4.40}$$

Now, according to (4.32) it holds that $\sum_n h_n B_n x = \dot{\delta x} - \big[A + \sum_n u_n B_n \big] \delta x$, and we write the following

$$\begin{aligned}
\int_0^T \langle \sum_n h_n B_n x, \delta p \rangle dt &= \int_0^T \langle \dot{\delta x} - \big[A + \sum_n u_n B_n \big] \delta x, \delta p \rangle dt \\
&= \int_0^T \langle \dot{\delta x}, \delta p \rangle dt + \int_0^T \langle -\big[A + \sum_n u_n B_n \big] \delta x, \delta p \rangle dt \\
&= \langle \delta x(T), \delta p(T) \rangle - \langle \delta x(0), \delta p(0) \rangle \\
&\quad + \int_0^T \langle \delta x, -\dot{\delta p} \rangle dt + \int_0^T \langle \delta x, -\big[A + \sum_n u_n B_n \big]^* \delta p \rangle dt \\
&= \langle \delta x(T), \delta p(T) \rangle + \int_0^T \langle \delta x, -\dot{\delta p} - \big[A + \sum_n u_n B_n \big]^* \delta p \rangle dt \\
&= \langle \delta x(T), \delta p(T) \rangle + \int_0^T \langle \delta x, \sum_n h_n B_n^* p \rangle dt \\
&= -\|\delta x(T)\|_2^2 + \int_0^T \langle \sum_n h_n B_n \delta x, p \rangle dt ,
\end{aligned} \tag{4.41}$$

where we used the integration-by-parts rule, $\delta x(0) = 0$, $\delta p(T) = -\delta x(T)$, and $\sum_n h_n B_n^* p = -\dot{\delta p} - \left[A + \sum_n u_n B_n\right]^* \delta p$. By replacing (4.41) into (4.40), we get

$$
\begin{aligned}
\langle \mathcal{J}_r(u)(h), h \rangle_{L^2} &= \nu \|h\|_{L^2}^2 + \|\delta x(T)\|_2^2 - 2\sum_n \int_0^T \langle h_n B_n \delta x, p \rangle dt \\
&\geq \nu \|h\|_{L^2}^2 + \|\delta x(T)\|_2^2 - 2\sum_n \int_0^T |h_n| \|B_n\|_{\mathcal{L}} \|\delta x\|_2 \|p\|_2 dt \\
&\geq \nu \|h\|_{L^2}^2 + \|\delta x(T)\|_2^2 - 2K_{00} \sum_n \int_0^T |h_n| \|\delta x\|_2 \|x(T) - x_T\|_2 dt \\
&= \nu \|h\|_{L^2}^2 + \|\delta x(T)\|_2^2 - 2K_{00} \|x(T) - x_T\|_2 \sum_n \int_0^T |h_n| \|\delta x\|_2 dt \\
&\geq \nu \|h\|_{L^2}^2 + \|\delta x(T)\|_2^2 - 2K_{00} \|x(T) - x_T\|_2 \sum_n \|h_n\|_{L^2} \|\delta x\|_{L^2} \\
&\geq \nu \|h\|_{L^2}^2 + \|\delta x(T)\|_2^2 - 2K_{00} \|x(T) - x_T\|_2 \sqrt{N_C} \|h\|_{L^2} \|\delta x\|_{L^2} ,
\end{aligned}
\tag{4.42}
$$

where we used the Cauchy-Schwarz inequality and K_{00} as in Proposition 5. By applying Proposition 5, it holds that

$$
\langle \mathcal{J}_r(u)(h), h \rangle_{L^2} \geq \|\delta x(T)\|_2^2 + \left[\nu - 4K_{00}^2 T N_C \|x_0\|_2 \|x(T) - x_T\|_2\right] \|h\|_{L^2}^2 ,
\tag{4.43}
$$

and the claim follows. $\qquad\square$

Since (4.38) is a sufficient condition for the coercivity of the reduced Hessian of (3.42), then it is also a sufficient optimality condition. Notice that, because of $x(T)$, α depends on the control u. We remark that, (4.38) can be a strict condition that is not always satisfied. On the other hand, this provides useful informations regarding regularity of problem (3.42). In particular:

- it is evident the benefit due to the regularization parameter ν;
- regularity increases as $\|x(T) - x_T\|_2$ decreases, and in the special case that u is an exact-controllability function, this term even vanishes;
- the same holds for the time T: the shorter is the time horizon, the more regular is the problem;
- the term $\|\delta x(T)\|_2$ in (4.39) provides a regularization effect, independently on condition (4.38).

It remains to discuss Fréchet differentiability required in Theorem 15. We recall that, the map $\mathcal{F}_r : L^2((0,T); \mathbb{R}^{N_C}) \to L^2((0,T); \mathbb{R}^{N_C})$ is given by $u \mapsto (\mathcal{F}_r(u))_n = \nu u_n - \langle B_n x, p \rangle$. Hence, Fréchet differentiability of $\mathcal{F}_r$ follows from Lemma 3, and in particular condition (4.37) is satisfied.

Consequently, all conditions in Theorem 15 are satisfied and we summarize our convergence result in the following theorem.

Theorem 16. *Under the assumptions of Lemma 3 and Lemma 4, the sequence $\{u^k\}_{k=1}^{\infty}$ generated by the Krylov-Newton scheme (4.26)-(4.35) for the optimal quantum control problem (3.42) with $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$, is locally superlinear convergent with order $1 + \alpha = 2$.*

4.3   Semi-smooth Newton method

In this section, Krylov-Newton methods are discussed for the solution to quantum optimal control problems characterized by control constraints and possibly non-smooth cost functionals. The resulting procedure is known as the semi-smooth Newton (SSN) method. Similarly to the previous section, first we provide a general description of the SSN method for general optimal control problems. Then, we discuss a Krylov-SSN method for the solution to (3.42), and present convergence results in Section 4.3.1.

In the same framework assumed for (4.13), consider the following optimal control problem

$$
\begin{aligned}
\min_{x,u} \quad & J(x, u) := J_1(x, u) + J_2(u) \\
\text{s.t.} \quad & c(x, u) = 0 \\
& x \in X \ , \ u \in U_{ad} \subset U \ ,
\end{aligned}
\tag{4.44}
$$

where the set of admissible controls U_{ad} is assumed to be convex and closed in U. In particular, the cost functional J is obtained as the sum of J_1 and J_2. The map $(x, u) \mapsto J_1(x, u)$ is assumed to be twice Fréchet differentiable, convex and lower semicontinuous, and $u \mapsto J_2(u)$ is considered convex, continuous and possibly non-smooth. Notice that, since J_2 is convex and continuous, the corresponding subdifferential ∂J_2 at a given u is

$$
\partial J_2(u) = \big\{ g \in U \ : \ J_2(v) - J_2(u) \ge \langle g, u - v \rangle_U \big\} \ ,
\tag{4.45}
$$

and is non-empty according to Theorem A 11; see, also [45, 66]. We remark that, a subdifferential of a convex function defined over a Banach space U, is rigorously defined as a subset of the dual space U^*, and in (4.45) the inner product is replaced by the duality-product $\langle \cdot, \cdot \rangle_{U^*, U}$. However, for the purposes of our discussion, U is a Hilbert space, and we can invoke the Riesz representation theorem. This allows to work with scalar product and avoid cumbersome notations that are not needed in our discussion.

Next, we discuss first-order optimality conditions for (4.44). To this end, we denote by p the Lagrange multiplier corresponding to the state constraint $c(x, u) = 0$, and $\tilde{c}(x, u, p) = 0$ the corresponding adjoint equation. By assuming that $c(x, u) = 0$ and $\tilde{c}(x, u, p) = 0$ are uniquely solvable for a given u, the maps $u \mapsto x(u)$ and $u \mapsto p(u)$ are well defined, and problem (4.44) can be studied in a reduced form, similarly to (4.18). A minimizer u for (4.44) is characterized by the following variational inequality condition; see Theorem A 18 in the appendix and, e.g., [45];

$$
\langle \nabla_u J_r(u), v - u \rangle_U + J_2(v) - J_2(u) \ge 0 \ , \ \forall v \in U_{ad} \ ,
\tag{4.46}
$$

where $\nabla_u J_r(u)$ denotes the derivative of J_1 in the reduced form (see, e.g. Lemma 2 and Theorem 7). By using (4.45), a sufficient condition for u to satisfy (4.46) is the following; see, e.g., [110] and references therein;

$$
\big\langle \nabla_u J_r(u) + \hat{\lambda}, v - u \big\rangle_U \ge 0 \ , \ \forall v \in U_{ad} \ ,
\tag{4.47}
$$

for any $\hat{\lambda} \in \partial J_2(u)$. Notice that equivalence between (4.46) and (4.47) is proved in Theorem 7; see also [45, 110, 120].

In order to apply a Newton method, the variational inequality (4.47) has to be reformulated as an equality condition, see, e.g., [110, 119, 123]. For example, in the case that $J_2 = 0$, because of convexity and closedness of U_{ad}, the variational inequality can be equivalently written in the following projection form; see, e.g., [33, 58, 80, 119, 123];

$$
u = \mathcal{P}_{U_{ad}} \Big(u - \theta \nabla_u J_r(u) \Big) \ ,
\tag{4.48}
$$

where θ is an arbitrary positive constant, $\mathcal{P}_{U_{ad}}$ is the projection operator from U to U_{ad}. In the case that $J_2 \neq 0$, $\hat{\lambda}$ in (4.46) is undetermined, and one has to reformulate the variational inequality in a different form in order to address the problem; see, e.g., [110]. Furthermore, we remark that the variational inequality (4.47) can be equivalently reformulated as a system of nonlinear complementarity conditions. To solve these non-linear complementarity problems so-called NCP-functions have been studied; see, e.g., [27, 38, 69] for finite-dimensional problems, and [50, 123] for infinite-dimensional problems.

In general, a reformulation of a variational inequality condition in an equality form is obtained by means of non-smooth functions. For our purposes, we assume that there exists a map $\Phi : U \to L^r((0,T); \mathbb{R})$, with $1 \leq r < \infty$, such that (4.47) is equivalently reformulated as follows

$$\Phi(u) = 0 \ . \tag{4.49}$$

For optimal control problems, this map is usually obtained as the composition $\Phi = \psi \circ F$ given by

$$\Phi(u)(t) = \psi(F(u)(t)) \ , \text{for } t \in (0,T) \ , \tag{4.50}$$

where $\psi : \mathbb{R}^m \to \mathbb{R}$ is a non-smooth and Lipschitz continuous function and $F : U \to L^{r'}((0,T); \mathbb{R}^m)$, with $1 \leq r \leq r' < \infty$ [123]. Consequently, similarly to Section 4.2, we aim to solve

$$\mathcal{F}_r(u) := \Phi(u) = 0 \ . \tag{4.51}$$

The use of the Newton method for the solution to (4.51) would require the evaluation of the Jacobian operator $\mathcal{J}_r$, that does not exist because of the non-smoothness of $\mathcal{F}_r$ due to ψ. For the development of a SSN, we have to choose an appropriate subdifferential for Φ. We refer to the generalized differential $\partial^\circ \Phi$ defined in [123], whose definition is motivated by a formal pointwise application of the chain rule. In fact, suppose for the moment that the map

$$F : u \in U \mapsto F(u) \in C([0,T]; \mathbb{R}^m) \tag{4.52}$$

is strictly differentiable, where $C([0,T]; \mathbb{R}^m)$ denotes the space of continuous function equipped with the max-norm. Then the function $f : U \to \mathbb{R}^m$ given, for a fixed $t \in (0,T)$, by

$$f : u \in U \mapsto F(u)(t) \ , \tag{4.53}$$

is strictly differentiable with derivative $f'(u) \in \mathcal{L}(U, \mathbb{R}^m)$, that is a linear operator obtained as

$$f'(u) : h \mapsto (F'(u)(h))(t) \ , \tag{4.54}$$

where $F'(u)$ denotes the derivative of F at u. If we assume that ψ possesses a subdifferential $\partial \psi$, then the chain rule for subdifferentials [34, 123] applied to the real-valued mapping $u \mapsto \Phi(u)(t) = \psi(f(u))$ yields

$$\partial\big(\Phi(u)(t)\big) \subset \partial\psi(f(u)) \circ f'(u) = \left\{ g \in U \ \middle| \ \begin{matrix} \langle g, h \rangle_U = \sum_{j=1}^m d_j(t)\big(F'_j(u)(h)\big)(t), \\ d(t) \in \partial\psi(F(u)(t)) \end{matrix} \right\} . \tag{4.55}$$

Furthermore, we can replace '$\subset$' with '$=$' if ψ is convex or if the operator $f'(u)$ is onto, see [34, 123]. Following the above motivation, the subdifferential $\partial^\circ \Phi$ is defined. To this purpose the following assumptions are required; see [123] Assumption 3.1;

Assumption 1. *There are $1 \leq r \leq r' < q \leq \infty$, such that*

(a) the operator $F : U \to L^{r'}((0,T); \mathbb{R}^m)$ is continuously Fréchet differentiable;

(b) the mapping $u \in U \mapsto F(u) \in L^q((0,T); \mathbb{R}^m)$ is locally Lipschitz continuous, that is for all $u \in U$ there exists an open neighbourhood $\mathcal{N} = \mathcal{N}(u)$ and a constant $L_F = L_F(\mathcal{N})$ such that

$$\sum_{j=1}^m \|F_j(u) - F_j(v)\|_{L^q} \leq L_F \|u - v\|_U \quad \forall u, v \in U ;$$

(c) the function $\psi : \mathbb{R}^m \to \mathbb{R}$ is Lipschitz, that is there exists L_ψ such that

$$|\psi(y) - \psi(z)| \leq L_\psi \|y - z\|_1 \quad \forall y, z \in \mathbb{R}^m ;$$

(d) ψ is semi-smooth (in the sense of Definition 3).

Notice that in Assumption 1 the only difference between the maps in (a) and (b) is the range space. Now, we can state the definition of $\partial^\circ \Phi$ as in [123].

Definition 4. *Let Assumption 1 hold. For $\Phi : U \to L^r((0,T); \mathbb{R})$ as in (4.50) the subdifferential $\partial^\circ \Phi$ at $u \in U$ is defined as follows*

$$\partial^\circ \Phi(u) := \left\{ S \in \mathcal{L}\big(U, L^r((0,T); \mathbb{R})\big) \;\middle|\; \begin{array}{l} S : h \mapsto \sum_{j=1}^m d_j(\cdot)\big(F'_j(u)(h)\big)(\cdot), \\ d \text{ measurable selection of } \partial\psi(F(u)) \end{array} \right\}, \quad (4.56)$$

where $\mathcal{L}\big(U, L^r((0,T); \mathbb{R})\big)$ is the space of all linear operators mapping from U to $L^r((0,T); \mathbb{R})$.

We remark that the superscript $\circ$ is chosen in [123] to indicate that $\partial^\circ \Psi$ is designed for superposition operators. By means of this subdifferential we can state the following definition of semi-smoothness.

Definition 5. *The operator Φ is semi-smooth at $u \in U$ if*

$$\sup_{S \in \partial^\circ \Phi(u+h)} \|\Phi(u+h) - \Phi(u) - S(h)\|_{L^r} = o(\|h\|_U) \quad as \quad h \to 0 \ in \ U . \qquad (4.57)$$

The following result motivates the choice of $\partial^\circ \Phi$ for constructing a semi-smooth Newton method; in fact, Theorem 17 states that $\partial^\circ \Phi$ is non-empty and Φ is semi-smooth; see [123] Proposition 4.8 and Theorem 5.2.

Theorem 17. *Under the Assumption 1, we have that*

(a) for all $u \in U$ the subdifferential $\partial^\circ \Phi(u)$ is non-empty and bounded in $\mathcal{L}\big(U, L^r((0,T); \mathbb{R})\big)$;

(b) the operator Φ is semi-smooth (in the sense of Definition 5).

Now, recalling that we aim to solve $\mathcal{F}_r(u) = 0$, with $\mathcal{F}_r := \Phi$, we consider the Newton sequence $\{u^k\}_{k=1}^\infty \subset U$ generated by the following three semi-smooth Newton steps

$$\begin{array}{lll} S_0 : & \text{choose a generalized Jacobian } \mathcal{J}_g(u^k) \in \partial^\circ \mathcal{F}_r(u^k) \\ S_1 : & \delta u^k = -(\mathcal{J}_g(u^k))^{-1} \mathcal{F}_r(u^k) & (4.58) \\ S_2 : & u^{k+1} = u^k + \delta u^k . \end{array}$$

We remark that in [123] the semi-smooth Newton procedure (4.58) presents an additional "smoothing" step, which is used in the case of a norm-discrepancy. This discrepancy

appears when $U = L^p$ with $p > r$ and, consequently, the L^p-norm is stronger than the L^r-norm. This framework arises, for instance, in the case that the complementarity condition is expressed by means of the Fischer-Burmeister functional; see [56, 123]. However, this situation does not appears in our problems, and for the sake of brevity, we omit a discussion regarding this discrepancy problem.

Convergence of the Newton sequence $\{u^k\}_{k=1}^\infty$ is discussed in the next theorem [123].

Theorem 18. *Let Assumption 1 hold, and assume that the generalized Jacobian $\mathcal{J}_g(u^k) \in \partial^\circ \mathcal{F}_r(u^k) \subset \mathcal{L}\big(U, L^r((0,T); \mathbb{R})\big)$ defined between appropriate spaces is invertible with a bounded inverse with respect to an appropriate norm. If u^1 sufficiently close to a solution u^* to $\mathcal{F}_r(u) = 0$, then the sequence $\{u^k\}_{k=1}^\infty \subset U$ generated by the semi-smooth Newton procedure (4.58) is locally superlinear convergent.*

We remark that, in Theorem 18 we used the term "appropriate spaces" to take into account the possible norm-discrepancy between U and L^r. However, this discrepancy disappear in the obvious situation that $U = L^p$ with $p = r$, which will be the case of our quantum control problems.

Next, we discuss explicitly the SSN method for the solution of the quantum control problem (3.107) with $U_{ad} = U_{ad,1}$. To this purpose, we reformulate the optimality system (3.114) for (3.107) following the approach presented in [110] for the optimal control of linear elliptic differential equations. In particular, conditions (3.114b)-(3.114f) can be equivalently written as follows

$$
\begin{aligned}
C_n(u,\mu) := u_n &- \max\big(0, u_n + \theta(\mu_n - \beta)\big) - \min\big(0, u_n + \theta(\mu_n + \beta)\big) \\
&+ \max\big(0, u_n - b + \theta(\mu_n - \beta)\big) + \min\big(0, u_n + b + \theta(\mu_n + \beta)\big) = 0 \,,
\end{aligned}
\tag{4.59}
$$

where θ is an arbitrary positive constant, the min- and max-functions are understood pointwise, $n = 1, ..., N_C$, and μ_n is given by

$$
\mu_n := \lambda_{+,n} - \lambda_{-,n} + \hat{\lambda}_n \,.
\tag{4.60}
$$

Hence, the optimality system (3.114) for (3.107) becomes as follows

$$
\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \,, \quad x(0) = x_0
\tag{4.61a}
$$

$$
-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* p \,, \quad p(T) = -(x(T) - x_T)
\tag{4.61b}
$$

$$
C_n(u,\mu) = 0
\tag{4.61c}
$$

$$
\mu_n = \langle B_n x, p \rangle - \nu u_n \,, \quad n = 1, ..., N_C \,.
\tag{4.61d}
$$

Now, we define the variable $\eta := (x, u, p)$ and write the problem of solving (4.61) as the following root problem. Find η such that

$$
\mathcal{F}(\eta) = 0 \,,
\tag{4.62}
$$

where $\mathcal{F}(\eta) := \left(\mathcal{F}_1(\eta), \mathcal{F}_2(\eta), \mathcal{F}_3(\eta) \right)^T$ and

$$\mathcal{F}_1(\eta) := -\dot{p} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* p \tag{4.63a}$$

$$\mathcal{F}_{2,n}(\eta) := C_n(x, u, p) , \ n = 1, ..., N_C \tag{4.63b}$$

$$\mathcal{F}_3(\eta) := \dot{x} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x . \tag{4.63c}$$

Notice that we replace in (4.61c) the function $\mu = \left(\mu_1, ..., \mu_{N_C} \right)$ obtained from (4.61d), and hence we write $C_n(u, \mu) = C_n(x, u, p)$.

It is clear from (4.59) and because of the min- and max-functions, that the function $C(\eta)$ is not differentiable in the classical sense. This fact prevents the use of a standard Newton method to solve (4.62). However, we can use non-smooth theory and the calculus for generalized gradients, see, e.g., [34], and apply the semi-smooth Newton scheme (4.58) for solving (4.62) with a guaranteed local superlinear convergence. Notice that convergence of the SSN method for the solution to (3.107) is discussed in Section 4.3.1.

To this purpose, we construct an adequate generalized Jacobian of $\mathcal{F}$, in the sense that we want to ensure semi-smoothness of C. First, using the calculus rules of sub-differentials [34, 123], we can obtain sub-gradients for $C_n(u, \mu)$. In order to guarantee that C_n is semi-smooth, we need to consider subgradients of C_n that are in the subdifferential $\partial^\circ C_n(u)$ defined in Definition 4. Semi-smoothness of C will be proved in Section 4.3.1. We consider particular choices of subgradients of the maps $y \mapsto \max(0, y)$ and $y \mapsto \min(0, y)$, given by the following characteristic functions [56, 110, 123]

$$\chi_{max}(v)(t) = \begin{cases} 1 & \text{if } v(t) \geq 0 \\ 0 & \text{if } v(t) < 0 \end{cases} \tag{4.64}$$

and

$$\chi_{min}(v)(t) = \begin{cases} 1 & \text{if } v(t) \leq 0 \\ 0 & \text{if } v(t) > 0 \end{cases} , \tag{4.65}$$

respectively. With this choice, subgradients of C_n are given by the following operators

$$\begin{aligned} \nabla_x C_n(x, u, p) &= \Psi_n(x, u, p) \langle B_n^* p, \cdot \rangle \\ \nabla_u C_n(x, u, p) &= I + (1 - \nu\theta) \Psi_n(x, u, p) I \\ \nabla_p C_n(x, u, p) &= \Psi_n(x, u, p) \langle B_n x, \cdot \rangle , \end{aligned} \tag{4.66}$$

where I is the identity operator and $\Psi_n(x, u, p)$ is defined as follows

$$\begin{aligned} \Psi_n(x, u, p) := & \left[- \chi_{max} \left(u_n + \theta(\mu_n(x, u, p) - \beta) \right) - \chi_{min} \left(u_n + \theta(\mu_n(x, u, p) + \beta) \right) \right. \\ & \left. + \chi_{max} \left(u_n - b + \theta\mu_n(x, u, p) \right) + \chi_{min} \left(u_n + b + \theta\mu_n(x, u, p) \right) \right] . \end{aligned} \tag{4.67}$$

Hence, the generalized Jacobian reads as follows

$$\mathcal{J}_g(\eta) = \begin{pmatrix} \nabla_x \mathcal{F}_1(\eta) & \nabla_u \mathcal{F}_1(\eta) & \nabla_p \mathcal{F}_1(\eta) \\ \nabla_x \mathcal{F}_2(\eta) & \nabla_u \mathcal{F}_2(\eta) & \nabla_p \mathcal{F}_2(\eta) \\ \nabla_x \mathcal{F}_3(\eta) & \nabla_u \mathcal{F}_3(\eta) & \nabla_p \mathcal{F}_3(\eta) \end{pmatrix} = \begin{pmatrix} \nabla_x \mathcal{F}_1(\eta) & \nabla_u \mathcal{F}_1(\eta) & \nabla_p \mathcal{F}_1(\eta) \\ \nabla_x C(\eta) & \nabla_u C(\eta) & \nabla_p C(\eta) \\ \nabla_x \mathcal{F}_3(\eta) & \nabla_u \mathcal{F}_3(\eta) & 0 \end{pmatrix} , \tag{4.68}$$

where

$$\nabla_u \mathcal{F}_1(\eta) = \Big((\nabla_u \mathcal{F}_1(\eta))_1 \cdots (\nabla_u \mathcal{F}_1(\eta))_{N_C} \Big) , \tag{4.69}$$

and

$$\nabla_u \mathcal{F}_3(\eta) = \Big((\nabla_u \mathcal{F}_3(\eta))_1 \cdots (\nabla_u \mathcal{F}_3(\eta))_{N_C} \Big) , \tag{4.70}$$

with

$$(\nabla_u \mathcal{F}_1(\eta))_n = -B_n^* p \quad \text{and} \quad (\nabla_u \mathcal{F}_3(\eta))_n = -B_n x . \tag{4.71}$$

We also have

$$\nabla_x \mathcal{F}_3(\eta) = \frac{d}{dt} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right] , \tag{4.72}$$

and

$$\nabla_p \mathcal{F}_1(\eta) = \nabla_x \mathcal{F}_3(\eta)^* . \tag{4.73}$$

Having obtained the generalized Jacobian $\mathcal{J}_g$, we construct the semi-smooth Newton sequence $\{\eta_k\}_{k=1}^{\infty}$ by using procedure (4.58). However, we remark that this procedure could be computationally infeasible because of the high dimension of the discrete generalized Jacobian. Consequently, similarly to Section 4.2, we want to construct a Krylov-SSN procedure. The SSN method requires at each iteration to solve the linear problem $\mathcal{J}_g(\eta)(\delta\eta) = -\mathcal{F}(\eta)$. This is given explicitly by the following

$$\begin{pmatrix} \nabla_x \mathcal{F}_1(x,u,p) & \nabla_u \mathcal{F}_1(x,u,p) & \nabla_p \mathcal{F}_1(x,u,p) \\ \nabla_x C(x,u,p) & \nabla_u C(x,u,p) & \nabla_p C(x,u,p) \\ \nabla_x \mathcal{F}_3(x,u,p) & \nabla_u \mathcal{F}_3(x,u,p) & 0 \end{pmatrix} \begin{pmatrix} \delta x \\ \delta u \\ \delta p \end{pmatrix} = - \begin{pmatrix} \mathcal{F}_1(x,u,p) \\ \mathcal{F}_2(x,u,p) \\ \mathcal{F}_3(x,u,p) \end{pmatrix} . \tag{4.74}$$

Now, since the state x and the adjoint p are uniquely determined by the control u, that is $x = x(u)$ and $p = p(u)$, we have $\mathcal{F}_3(x,u,p) = 0$ and $\mathcal{F}_1(x,u,p) = 0$, respectively. Therefore, the system (4.74) becomes as follows

$$\begin{pmatrix} \nabla_x \mathcal{F}_1(u) & \nabla_u \mathcal{F}_1(u) & \nabla_p \mathcal{F}_1(u) \\ \nabla_x C(u) & \nabla_u C(u) & \nabla_p C(u) \\ \nabla_x \mathcal{F}_3(u) & \nabla_u \mathcal{F}_3(u) & 0 \end{pmatrix} \begin{pmatrix} \delta x \\ \delta u \\ \delta p \end{pmatrix} = - \begin{pmatrix} 0 \\ \mathcal{F}_2(u) \\ 0 \end{pmatrix} . \tag{4.75}$$

The action of the third row on the vector of unknowns gives the following differential equation

$$\dot{\delta x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta x + \left[\sum_{n=1}^{N_C} \delta u_n B_n \right] x , \quad \delta x(0) = 0 , \tag{4.76}$$

that is, the linearization of the forward equation (3.53a). In the same fashion, the action of the first row gives the following

$$-\dot{\delta p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* \delta p + \left[\sum_{n=1}^{N_C} \delta u_n B_n \right]^* p , \quad \delta p(T) = -\delta x(T) , \tag{4.77}$$

that is, the linearization of the backward equation (3.53b). Equations (4.76) and (4.77) are uniquely solvable for a given δu, and their solutions are $\delta x = \delta x(\delta u)$ and $\delta p = \delta p(\delta u)$, respectively. Therefore, the linear system (4.75) becomes as follows

$$\nabla_x C(u)\big(\delta p(\delta u)\big) + \nabla_u C(u)\big(\delta u\big) + \nabla_p C(u)\big(\delta x(\delta u)\big) = -\mathcal{F}_2(u) , \tag{4.78}$$

where the only unknown is δu. Notice that the left hand side of (4.78) the action of the reduced generalized Jacobian $\mathcal{J}_{g,r}(u)$ on δu, and is given by

$$
\begin{aligned}
\left(\mathcal{J}_{g,r}(u)(\delta u)\right)_n &:= \left(\nabla_x C(u)\big(\delta p(\delta u)\big) + \nabla_u C(u)\big(\delta u\big) + \nabla_p C(u)\big(\delta x(\delta u)\big)\right)_n \\
&= \Big[1 + (1 - \nu\theta)\Psi_n(u)\Big]\delta u_n + \Psi_n(u)\Big[\langle B_n^* p, \delta x(\delta u)\rangle + \langle B_n x, \delta p(\delta u)\rangle\Big],
\end{aligned}
$$
(4.79)

for $n = 1, ..., N_C$ and where $\Psi_n(u)$ is defined in (4.67). Hence, the Krylov-Newton linear system is given by

$$
\begin{aligned}
\Big[1 + (1 - \nu\theta)\Psi_n(u)\Big]&\delta u_n + \Psi_n(u)\Big[\langle B_n^* p, \delta x(\delta u)\rangle + \langle B_n x, \delta p(\delta u)\rangle\Big] \\
&= u_n - \max\big(0, u_n + \theta(\mu_n - \beta)\big) - \min\big(0, u_n + \theta(\mu_n + \beta)\big) \\
&\quad + \max\big(0, u_n - b + \theta(\mu_n - \beta)\big) + \min\big(0, u_n + b + \theta(\mu_n + \beta)\big).
\end{aligned}
$$
(4.80)

In the next section, we discuss convergence of the presented SSN method for the solution to the quantum optimal control problem (3.107).

4.3.1 Convergence of the SSN method for quantum optimal control

In this section, the convergence of the SSN method presented in the previous section is discussed. In particular, we prove a convergence result of the SSN method, for the solution to the quantum optimal control problem (3.107), by means of Theorem 18.

It is well-known that the finite-dimensional maps $y \mapsto \max(0, y)$ and $y \mapsto \min(0, y)$ possess non-empty subdifferentials and are semi-smooth [56, 123]. Further, it is proved that for $1 \le r < q \le \infty$ the infinite-dimensional maps $\max, \min : L^q(0, T) \to L^r(0, T)$ are semi-smooth [56, 123]. Recalling that $x = x(u)$ and $p = p(x(u), u)$, we notice that the function $u \mapsto C(x(u), u, p(u))$ is obtained as the sum of compositions of max- and min-functions and the map $\mu : u \mapsto \mu(x(u), u, p(u))$. Consequently, to investigate semi-smoothness of C, we define the subdifferential $\partial^\circ C(u)$ as the following product space

$$
\partial^\circ C(u) := \partial^\circ C_1(u) \times \partial^\circ C_2(u) \times \cdots \times \partial^\circ C_{N_C}(u),
$$
(4.81)

where each of $\partial^\circ C_n(u)$ is defined as in Definition 4 [123]. We remark that $\partial^\circ C(u)$ resembles the Qi's subdifferential given in Definition 2. Then, we can prove semi-smoothness of C as follows.

Lemma 5. *Under the assumptions of Lemma 3, the map*

$$
C : L^2((0, T), \mathbb{R}^{N_C}) \to L^2((0, T), \mathbb{R}^{N_C}),
$$
(4.82)

defined according to (4.59) and (4.60) as $u \mapsto C(x(u), u, p(u))$, is semi-smooth in the sense of Definition 5 with $\partial^\circ C(u)$.

Proof. In this theorem, in order to avoid confusions, we use a notation that is slightly different to the one defined at the beginning of Chapter 3. In particular, we denote by $\|\cdot\|_{L_m^2}$ the following L^2-norm

$$
\|v\|_{L_m^2} := \left(\int_0^T \sum_{n=1}^m |v_n|^2 dt\right)^{1/2},
$$

for $v \in L^2((0,T); \mathbb{R}^m)$.

We prove the claim by means of Theorem 17 [123]. For this purpose, we show that Assumption 1 holds in our case. According to the notation used in Assumption 1, we have that $m = 2$ such that $\psi : \mathbb{R}^2 \to \mathbb{R}$ and $\Phi_n := \psi \circ (0, \mu_n)$ for $n = 1, \dots, N_C$. Notice that ψ is given by the sum of max and min functions (as in (4.59)), and we recall that max and min functions are semi-smooth and the sum of semi-smooth functions is still semi-smooth; see, e.g., [97] and references therein. Further, we have that $U = L^p((0,T); \mathbb{R}^{N_C})$ with $p = 2$, and $r = r' = 2$.

In this setting, we need that the map $\mu_n : L^2((0,T), \mathbb{R}^{N_C}) \to L^q((0,T), \mathbb{R})$ with $q > 2$ is locally Lipschitz continuous and that $\mu_n : L^2((0,T), \mathbb{R}^{N_C}) \to L^2((0,T), \mathbb{R})$ is Fréchet differentiable. For this purpose, we notice that the map μ_n is obtained as the sum of the identity with the map Θ_n defined in Lemma 3. Hence, by Lemma 3 and by the calculus rules for Lipschitz continuous and Fréchet differentiable maps, we obtain the desired properties of the map μ_n. Consequently, we can apply Theorem 17 to obtain semi-smoothness of $C_n = \Phi_n$ with respect to $\partial^\circ C_n$, that is

$$\sup_{S_n \in \partial^\circ C_n(u+h)} \|C_n(u+h) - C_n(u) - S_n(h)\|_{L_1^2} = o(\|h\|_{L_{N_C}^2}) \quad \text{as} \quad h \to 0 . \tag{4.83}$$

To prove semi-smoothness of C, we notice that

$$\|C(u+h) - C(u) - S(h)\|_{L_{N_C}^2} \leq \sum_n \|C_n(u+h) - C_n(u) - S_n(h)\|_{L_1^2} . \tag{4.84}$$

Then we write that

$$\sup_{S \in \partial^\circ C(u+h)} \|C(u+h) - C(u) - S(h)\|_{L_{N_C}^2}$$
$$\leq \sup_{S \in \partial^\circ C(u+h)} \sum_n \|C_n(u+h) - C_n(u) - S_n(h)\|_{L_1^2} \tag{4.85}$$
$$\leq \sum_n \sup_{S_n \in \partial^\circ C_n(u+h)} \|C_n(u+h) - C_n(u) - S_n(h)\|_{L_1^2} .$$

Recalling (4.83), we obtain the following

$$\sup_{S \in \partial^\circ C(u+h)} \|C(u+h) - C(u) - S(h)\|_{L_{N_C}^2} = o(\|h\|_{L_{N_C}^2}) \quad \text{as} \quad h \to 0 , \tag{4.86}$$

that is the required property. $\qquad\square$

In the next lemma we provide sufficient conditions for regularity of the reduced generalized Jacobian. We remark that Lemma 6 provides results similar as in Lemma 4. Consequently, all the remarks provided for Lemma 4 hold also for Lemma 6.

Lemma 6. *The reduced generalized Jacobian $\mathcal{J}_{g,r}$ is coercive if the parameters ν and θ are large enough and the tracking term $\|x(T) - x_T\|_2$ is small enough. In particular, if $\nu\theta \leq 1$ and*

$$\alpha := \left[\theta\nu - 4K_{00}^2 T N_C \|x_0\|_2 \|x(T) - x_T\|_2 \right] > 0 , \tag{4.87}$$

with K_{00} as in Proposition 5, then $\mathcal{J}_{g,r}$ is coercive and it holds that

$$\langle \mathcal{J}_{g,r}(u)(h), h \rangle_{L^2} \geq \|\tilde{\delta x}(T; h)\|_2^2 + \alpha\|h\|_{L^2}^2 \quad \text{for any} \quad h \in L^2((0,T); \mathbb{R}^{N_C}) , \tag{4.88}$$

where $\tilde{\delta x}$ is the solution to (4.76) corresponding to $\delta u_n = -\Psi_n(u)h_n$.

Moreover, the reduced generalized Jacobian $\mathcal{J}_{g,r}(u)$ is coercive if it holds that

$$8T N_C \|x_0\|_2^2 K_{00}^2 < \nu\theta \leq 1 . \tag{4.89}$$

Proof. We proceed similarly to Lemma 4. Let $h \in L^2((0,T); \mathbb{R}^{N_C})$, and denote by $\delta x = \delta x(h)$ and $\delta p = \delta p(h)$ the corresponding solutions to linearized constraint (4.76) and adjoint (4.77). Consider the inner product $\langle \mathcal{J}_{g,r}(u)(h), h \rangle_{L^2}$ with $h \in L^2((0,T); \mathbb{R}^{N_C})$. We have the following

$$
\begin{aligned}
\langle \mathcal{J}_{g,r}(u)(h), h \rangle_{L^2} = \|h\|_{L^2}^2 + (1 - \nu\theta) \int_0^T \sum_n \Psi_n(u) h_n^2 dt \\
+ \int_0^T \sum_n h_n \Psi_n(u) \Big[\langle B_n^* p, \delta x \rangle + \langle B_n x, \delta p \rangle \Big] dt .
\end{aligned}
\tag{4.90}
$$

Now, notice that $|\Psi_n(u)(t)| \leq 1$ a.e. in $(0,T)$ and recall that $\nu\theta \leq 1$, from (4.90) we get

$$
\begin{aligned}
\langle \mathcal{J}_{g,r}(u)(h), h \rangle_{L^2} \geq \|h\|_{L^2}^2 - (1 - \nu\theta)\|h\|_{L^2}^2 \\
- \int_0^T \sum_n (-\Psi_n(u) h_n) \Big[\langle B_n^* p, \delta x \rangle + \langle B_n x, \delta p \rangle \Big] dt .
\end{aligned}
\tag{4.91}
$$

Using (4.91) and proceeding similarly to (4.40), we obtain that

$$
\begin{aligned}
\langle \mathcal{J}_r(u)(h), h \rangle_{L^2} \geq \theta\nu \|h\|_{L^2}^2 \\
- \int_0^T \Big(\langle \sum_n (-\Psi_n(u) h_n) B_n x, \delta p \rangle + \langle \sum_n (-\Psi_n(u) h_n) B_n \delta x, p \rangle \Big) dt .
\end{aligned}
\tag{4.92}
$$

Now, we define $\tilde{h}_n := -\Psi_n(u) h_n$, and we denote by $\tilde{\delta x} = \delta x(\tilde{h})$ the solution to the linearized constraint (4.76) corresponding to $\delta u = \tilde{h}$, and according to (4.76) it holds that $\sum_n \tilde{h}_n B_n x = \dot{\tilde{\delta x}} - \big[A + \sum_n u_n B_n \big] \tilde{\delta x}$. Similarly to (4.41), we write

$$
\begin{aligned}
\int_0^T \langle \sum_n \tilde{h}_n B_n x, \delta p \rangle dt = \int_0^T \langle \dot{\tilde{\delta x}} - \big[A + \sum_n u_n B_n \big] \tilde{\delta x}, \delta p \rangle dt \\
= -\|\tilde{\delta x}(T)\|_2^2 + \int_0^T \langle \sum_n h_n B_n \tilde{\delta x}, p \rangle dt .
\end{aligned}
\tag{4.93}
$$

By replacing (4.93) into (4.92), and similarly to (4.42), we get

$$
\begin{aligned}
\langle \mathcal{J}_r(u)(h), h \rangle_{L^2} = \theta\nu \|h\|_{L^2}^2 + \|\tilde{\delta x}(T)\|_2^2 - \sum_n \int_0^T \langle h_n B_n (\delta x + \tilde{\delta x}), p \rangle dt \\
\geq \theta\nu \|h\|_{L^2}^2 + \|\tilde{\delta x}(T)\|_2^2 - K_{00} \|x(T) - x_T\|_2 \sum_n \int_0^T |h_n| \big(\|\delta x\|_2 + \|\tilde{\delta x}\|_2 \big) dt \\
\geq \theta\nu \|h\|_{L^2}^2 + \|\tilde{\delta x}(T)\|_2^2 \\
- \|x(T) - x_T\|_2 2\sqrt{T N_C} K_{00}^2 \|x_0\|_2 \big(\|h\|_{L^2} + \|\tilde{h}\|_{L^2} \big) \|h\|_{L^1} ,
\end{aligned}
\tag{4.94}
$$

where we used the Cauchy-Schwarz inequality, the estimate (3.18) and K_{00} as in Proposition 5. Next, by noticing that $|\Psi_n(u)(t)| \leq 1$, the following holds $\|\tilde{h}\|_{L^2} = \|\Psi(u) h\|_{L^2} \leq \|h\|_{L^2}$, and by recalling that $\|h\|_{L^1} \leq \sqrt{T N_C} \|h\|_{L^2}$, we obtain

$$
\langle \mathcal{J}_r(u)(h), h \rangle_{L^2} \geq \|\tilde{\delta x}(T)\|_2^2 + \Big[\theta\nu - 4 K_{00}^2 T N_C \|x_0\|_2 \|x(T) - x_T\|_2 \Big] \|h\|_{L^2}^2 ,
\tag{4.95}
$$

and coercivity of $\mathcal{J}_r$ follows.

To conclude the proof, notice that

$$
\begin{aligned}
\theta\nu - 4K_{00}^2 T N_C \|x_0\|_2 \|x(T) - x_T\|_2 &\geq \theta\nu - 4K_{00}^2 T N_C \|x_0\|_2 \Big(\|x(T)\|_2 + \|x_T\|_2 \Big) \\
&\geq \theta\nu - 4K_{00}^2 T N_C \|x_0\|_2 \Big(\|x_0\|_2 + \|x_0\|_2 \Big) \quad\quad (4.96) \\
&\geq \theta\nu - 8K_{00}^2 T N_C \|x_0\|_2^2 \,.
\end{aligned}
$$

Hence condition (4.89) implies (4.87) and the claim follows. $\qquad\square$

With the next Theorem 19, we discuss the locally superlinear convergence of the presented semi-smooth Newton scheme.

Theorem 19. *Under the assumptions of Lemma 5 and Lemma 6, the semi-smooth Newton method (4.58), for the solution to the quantum optimal control problem (3.107), is locally superlinear convergent.*

Proof. Notice that, Lemma 5 guarantees the semi-smoothness of the map C defined in (4.59) (and (3.106a)). Furthermore, Lemma 6 guarantees that the generalized Jacobian is invertible with bounded inverse. Consequently, the claim follows from Theorem 18. $\quad\square$

We remark that, to obtain local convergence of the semi-smooth Newton method, we need semi-smoothness of the map C and boundedness of the inverse of its generalized Jacobian. Regarding semi-smoothness, we remark that, since C is defined as $C : L^2((0,T); \mathbb{R}^{N_C}) \to L^2((0,T); \mathbb{R}^{N_C})$ there is no norm-discrepancies in the Newton procedure and the "smoothing" step discussed in [123] is not necessary in our case. Regarding invertibility of the generalized Jacobian, we provide in Lemma 6 sufficient conditions, that are given by (4.87) and (4.89), for the coercivity of the generalized Jacobian, and hence for the boundedness of its inverse. These conditions can be strict to be satisfied. However, as already discussed for Lemma 4, they provide useful informations regarding regularity of $\mathcal{J}_{g,r}$. In particular, similarly to Lemma 4:

- the benefit due to the product $\nu\theta$ is evident; however, due to the non-smoothness we need to assume that $\nu\theta \leq 1$ to obtain a result similar to Lemma 4;

- regularity increases as $\|x(T) - x_T\|_2$ decreases;

- the same holds for the time T: the shorter is the time horizon, the more regular is the problem;

- the term $\|\tilde{\delta x}(T)\|_2$ in (4.88) provides a regularization effect, independently on conditions (4.87) and (4.89).

4.3.2   SSN method in the case $U_{ad} = U_{ad,2}$

In this section, we show how to apply the SSN method to the spin optimization problem (3.42) with the admissible control set $U_{ad} = U_{ad,2}$, with

$$
U_{ad,2} := \Big\{ v \in L^2((0,T); \mathbb{R}^{N_C}) \ : \ \|v(t)\|_2 \leq b \,, \text{ a.e. in } (0,T) \,, \ b \in \mathbb{R}^+ \Big\} \,. \quad\quad (4.97)
$$

The control set $U_{ad,2}$ is often used in NMR applications with the purpose of bounding the amplitudes of control radio-frequencies used in the experiments; see, e.g., [75, 105]. This

restriction of the control space is usually due to the type of the considered spectrometer, or to specific experimental applications.

The first-order necessary optimality condition for u to be a minimizer for (3.42) is given by

$$\langle \nabla_u J_r(u), v - u \rangle_{L^2} \geq 0 \quad \forall v \in U_{ad,2} , \tag{4.98}$$

where the gradient $\nabla_u J_r(u)$ is given by (3.48) in Lemma 2. Since by Lemma 1 the set $U_{ad,2}$ is closed and convex, the variational inequality (4.98) can be equivalently written as follows; see, e.g., [58, 80];

$$u = \mathcal{P}_{U_{ad,2}} \Big(u - \theta \nabla_u J_r(u) \Big) , \tag{4.99}$$

where θ is an arbitrary positive constant and $\mathcal{P}_{U_{ad}}$ is the projection operator from $L^2((0,T); \mathbb{R}^{N_C})$ to $U_{ad,2}$. By virtue of the definition of the projection operator, it is possible to put (4.48) into the pointwise projection form [58, 80], as follows

$$u(t) = \mathcal{P}_{U_{ad,2}} \Big(u(t) - \theta \nabla_u J_r(u)(t) \Big) , \quad \text{a.e. in } (0, T) . \tag{4.100}$$

Consequently, similarly to (4.27), the optimality system is the following

$$\begin{aligned} \tilde{c}(x, u, p) &= 0 \\ u(t) &= \mathcal{P}_{U_{ad}} \Big(u(t) - \theta \nabla_u J_r(u)(t) \Big) , \quad \text{a.e. in } (0, T) \\ c(x, u) &= 0 . \end{aligned} \tag{4.101}$$

Now, consider explicitly the projector operator as follows [80]

$$\mathcal{P}_{U_{ad}}(v) = \frac{v}{\max(1, \frac{1}{b}\|v\|_2)} . \tag{4.102}$$

From (4.100) we get the following

$$u(t) = \frac{u(t) - \theta \nabla_u J_r(u)(t)}{\max(1, \frac{1}{b}\|u(t) - \theta \nabla_u J_r(u)(t)\|_2)} , \tag{4.103}$$

and by multiplying with $\max(1, \frac{1}{b}\|u(t) - \theta \nabla_u J_r(u)(t)\|_2)$, it holds that

$$u(t) \max\Big(1, \frac{1}{b}\|u(t) - \theta \nabla_u J_r(u)(t)\|_2\Big) - \Big(u(t) - \theta \nabla_u J_r(u)(t)\Big) = 0 . \tag{4.104}$$

By choosing $\theta = \frac{1}{\nu}$ and using $\nabla_{u_n} J_r(u) = \nu u_n - \langle B_n x, p \rangle$, the following holds

$$\max\Big(\nu, \frac{1}{b}\|g(x, p)(t)\|_2\Big) u_n(t) - g_n(x, p)(t) = 0 , \tag{4.105}$$

where $g_n(x, p) := \langle B_n x, p \rangle$ for $n = 1, \dots, N_C$. Consider $\eta := (x, u, p)$ and define the map $\mathcal{F}(\eta) = \big(\mathcal{F}_1(\eta), \mathcal{F}_2(\eta), \mathcal{F}_3(\eta)\big)^T$ as follows

$$\begin{pmatrix} \mathcal{F}_1(\eta) \\ \mathcal{F}_2(\eta) \\ \mathcal{F}_3(\eta) \end{pmatrix} := \begin{pmatrix} \tilde{c}(x, u, p) \\ \max\big(\nu, \frac{1}{b}\|g(x, p)\|_2\big) u - g(x, p) \\ c(x, u) \end{pmatrix} , \tag{4.106}$$

and notice that a root of $\mathcal{F}$ corresponds to a stationary point solving (4.101). By working in the space of solutions to $c(x, u) = 0$ and $\tilde{c}(x, u, p) = 0$, we write the function $\mathcal{F}$ in a reduced form:

$$\mathcal{F}_r(u) := \mathcal{F}_2\big(x(u), u, p(u)\big) = \max\Big(\nu, \frac{1}{b}\|g(x(u), p(u))\|_2\Big)u - g(x(u), p(u)) . \quad (4.107)$$

In order to construct the Newton procedure for solving $\mathcal{F}_r(u) = 0$, we need the Jacobian of $\mathcal{F}_r$. However, $\mathcal{F}_r$ is not differentiable with respect to u, because of the max function, and we need to construct a generalized Jacobian [56, 123]. For this purpose, we define the following set

$$\mathcal{S} := \Big\{t \in (0, T) \mid \frac{1}{b}\|g(x, p)(t)\|_2 > \nu\Big\} , \quad (4.108)$$

and the following characteristic function

$$\chi_{\mathcal{S}}(t) := \begin{cases} 1 & \text{if } t \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} . \quad (4.109)$$

In this setting, the action of the reduced generalized Jacobian on $\delta u \in L^2((0, T); \mathbb{R}^{N_C})$ is given by the following

$$
\begin{aligned}
\big(\mathcal{J}_{g,r}(u)(\delta u)\big)_n :=& \max\Big(\nu, \frac{1}{b}\|g(x(u), p(u))\|_2\Big)\delta u_n \\
&+ u_n \frac{1}{b}\chi_{\mathcal{S}} \frac{\sum_{n=1}^{N_C}\big(\langle B_n \delta x(\delta u), p(u)\rangle + \langle B_n x(u), \delta p(\delta u)\rangle\big)\langle B_n x(u), p(u)\rangle}{\|g(x, p)\|_2} \\
&- \langle B_n \delta x(\delta u), p(u)\rangle - \langle B_n x(u), \delta p(\delta u)\rangle , \quad n = 1, ..., N_C ,
\end{aligned}
$$
$$(4.110)$$

that has to be understood pointwise in $(0, T)$, and where $\delta x(\delta u)$ and $\delta p(\delta u)$ are the solutions to (4.32) and (4.33) corresponding to δu, respectively. Hence, the Newton linear system $\mathcal{J}_{g,r}(u)(\delta u) = -\mathcal{F}_r(u)$ is given by

$$
\begin{aligned}
&\max\Big(\nu, \frac{1}{b}\|g(x(u), p(u))\|_2\Big)\delta u_n \\
&+ u_n \frac{1}{b}\chi_{\mathcal{S}} \frac{\sum_{n=1}^{N_C}\big(\langle B_n \delta x(\delta u), p(u)\rangle + \langle B_n x(u), \delta p(\delta u)\rangle\big)\langle B_n x(u), p(u)\rangle}{\|g(x, p)\|_2} \\
&- \langle B_n \delta x(\delta u), p(u)\rangle - \langle B_n x(u), \delta p(\delta u)\rangle \\
&= -\Big[\max\Big(\nu, \frac{1}{b}\|g(x(u), p(u))\|_2\Big)u_n - g_n(x(u), p(u))\Big] ,
\end{aligned}
$$
$$(4.111)$$

for $n = 1, \ldots, N_C$.

Notice that convergence of the SSN method presented in this section can be discussed similarly as in Section 4.3.1. In particular, we remark that, since the Euclidean norm $\|\cdot\|_2$ is not differentiable in the origin, it could be thought to miss the Fréchet differentiability required for the convergence of the SSN method. However, this singular behaviour can be neglected because $\|\cdot\|_2$ is in composition with $\max(\nu, \cdot)$. Moreover, convergence can be also discussed by means of a pointwise evaluation of the non-differentiable map and its subdifferential as in [50].

4.4   Summary and remarks

In this chapter, Newton and semi-smooth Newton methods were described for the solution of possibly non-smooth optimal quantum control problems. Since the construction of Hessian/Jacobian operator would be infeasible for numerical implementations, we focused on the Krylov-Newton approach. This strategy allows to avoid the construction of the Hessian/Jacobian operator, making possible the use of Newton type methods for the solution of infinite dimensional problems. Further, we investigated convergence of the developed Newton methods.

The novelties of this chapter are:

- Development of Krylov-Newton and SSN methods for problems governed by bilinear systems, and analysis of their convergence. We remark that, in the literature quantum spin optimal control only first-order methods are considered, and we are not aware on scientific works involving SSN methods for the solution of general quantum optimal control problems. Consequently, investigations on Newton type methods represent an important novelty in this field.

- Development of SSN method for L^1-optimal control problems governed by bilinear systems and analysis of its convergence. In the literature of L^1-optimization, SSN methods are studied for the solution to optimal control problems governed by linear and semi-linear state constraints. On the other hand, much less in known in the case of bilinear control-state constraints. Furthermore, the use of this method is an absolute novelty in the field of quantum control, where "sparse" controls are usually constructed "by hand" following the intuition of experimentalists.

Chapter 5

Two methods for the exact-control of quantum spin systems

In this chapter, we present two computational methods to solve the following exact-control problem

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \ \text{in} \ (0, T) \ , \ \ x(0) = x_0 \ , \ x(T) = x_T \ , \quad (5.1)$$

that is we seek a control u such that the trajectory of (5.1) starting from the initial state x_0 reaches at time T the desired target state x_T. This problem is hard to solve, and we are not aware on methods, provided with a rigorous theoretical framework, that are capable to address it. To guarantee existence of an exact-control u for (5.1), one has to invoke controllability results. In particular, since our purpose is to address quantum spin systems, in Section 5.1, we discuss controllability results for quantum spin systems.

In the sequel of this section, we present two different approaches for solving (5.1). The first approach is based on a continuation technique, which allows to generate a sequence of optimal control functions that converges to a solution to (5.1). This framework is discussed in Section 5.2. The second approach consists in a reformulation of (5.1) as an optimization problem, that possesses suitable local regularity properties, and that can be numerically addressed by optimal control techniques. This second framework is discussed in Section 5.3. We remark that, even if the focus of the present work is to tackle spin systems, the presented computational methods remain valid for general exact-control problems governed by bilinear systems.

5.1 Controllability of quantum spin systems

In this section, we recall some fundamental controllability results of closed quantum spin systems governed by the LvNM equation

$$\dot{\tilde{\rho}} = -i \big[\tilde{H}, \tilde{\rho} \big] \ , \quad (5.2)$$

as well as the bilinear control system

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ . \quad (5.3)$$

In particular, we focus on Ising spin systems. From Section 2.3.3 we recall that $N = 2^{N_p}$ and $N_x = N^2$, with N_p the number of spins. Further, the Hamiltonian $\tilde{H}$ is given by

$\tilde{H} = \tilde{H}_0 + \tilde{H}_u$, where

$$\tilde{H}_0 = \sum_{k=1}^{N_p} (\nu_k - \tilde{\omega}_k) I_{z,k} + \sum_{k<j} J_{k,j} I_{z,k} I_{z,j} \tag{5.4}$$

is the free Hamiltonian, and

$$\tilde{H}_u = \sum_{j=1}^{N_p} u_{x,j} \tilde{H}_{x,j} + u_{y,j} \tilde{H}_{y,j} \tag{5.5}$$

is the control Hamiltonian. Moreover, from Section 2.4.1 we recall that (5.3) is a real matrix-representation of (5.2) with respect to the coordinate map $\mathcal{V} : \mathfrak{her}(N) \to \mathbb{R}^{N_x}$ defined in (2.125)-(2.126).

Systems (5.2) and (5.3) admit a lift to the special unitary group $SU(N)$ and to the special orthogonal group $SO(N_x)$, respectively. In particular, the lifted equations of (5.2) and (5.3) are given by

$$\dot{\Upsilon}(t) = -i\tilde{H}(t)\Upsilon(t) \, , \quad \Upsilon(0) = I_N \, , \tag{5.6}$$

and

$$\dot{\Theta}(t) = \left[A + \sum_{n=1}^{N_C} u_n(t) B_n \right] \Theta(t) \, , \quad \Theta(0) = I_{N_x} \, , \tag{5.7}$$

respectively. That is, solutions $\Upsilon(t)$ and $\Theta(t)$ of the lifted systems give rise to solutions of the original system via the group actions

$$(\Upsilon, \rho_0) \mapsto \Upsilon \rho_0 \Upsilon^* \quad \text{and} \quad (\Theta, x_0) \mapsto \Theta x_0 \, , \tag{5.8}$$

of $SU(N)$ and $SO(N_x)$, respectively, and all solutions of the original systems can be represented in the form

$$\tilde{\rho}(t) = \Upsilon(t) \rho_0 \Upsilon(t)^* \quad \text{and} \quad x(t) = \Theta(t) x_0 \, . \tag{5.9}$$

Therefore, controllability properties of the lifted systems are closely related to controllability properties of the original systems [40, 67].

For the lifted control systems, we denote by $\mathcal{R}_T(\Upsilon_0)$ and $\mathcal{R}_T(\Theta_0)$ the sets of all points reachable at time T from $\Upsilon_0 \in SU(N)$ and $\Theta_0 \in SO(N_x)$, respectively. Then, the entire reachable sets $\mathcal{R}(\Upsilon_0)$ and $\mathcal{R}(\Theta_0)$ are defined as

$$\mathcal{R}(\Upsilon_0) := \bigcup_{T \geq 0} \mathcal{R}_T(\Upsilon_0) \quad \text{and} \quad \mathcal{R}(\Theta_0) := \bigcup_{T \geq 0} \mathcal{R}_T(\Theta_0) \, , \tag{5.10}$$

respectively. In a similar way one defines the reachable sets $\mathcal{R}(\rho_0)$ of (5.2) and $\mathcal{R}(x_0)$ of (5.3). Notice that the right invariance of the corresponding vector fields of (5.6) and (5.7) implies that $\mathcal{R}(\Upsilon_0) = \mathcal{R}(I_N)\Upsilon_0$ and $\mathcal{R}(\Theta_0) = \mathcal{R}(I_{N_x})\Theta_0$, where $\mathcal{R}(I_N)$ and $\mathcal{R}(I_{N_x})$ are the reachable sets of the corresponding identity matrix.

With these preliminaries, we first discuss controllability of (5.2) and its lifted equation (5.6). Notice that due to (5.9) any solution of (5.2) is confined to the unitary orbit $\mathcal{O}(\rho_0) := \{\Upsilon \rho_0 \Upsilon^* \mid \Upsilon \in SU(N)\}$ of its initial value $\tilde{\rho}(0) = \rho_0$. This observation leads to several different notions of controllability in the literature [2, 40, 101].

Definition 6. *A closed spin system is called:*

- *density operator controllable, if the reachable set $\mathcal{R}(\rho_0)$ of any initial density operator ρ_0 coincides with its entire unitary orbit $\mathcal{O}(\rho_0)$, i.e. if system (5.2) is controllable on $\mathcal{O}(\rho_0)$ for all starting points ρ_0;*
- *operator controllable, if the reachable set $\mathcal{R}(I)$ coincides with the entire unitary group $SU(N)$, that is, if the system (5.6) is controllable.*

The following result shows that these two apparently different concepts of controllability coincide [2, 40, 101].

Theorem 20. *A closed spin system is operator controllable if and only if it is density operator controllable.*

The main result to establish operator controllability is the following theorem [67, 68], which provides a necessary and sufficient condition for controllability, that is based on the evaluation of the Lie algebra generated by the drift and control vector fields. For instance, for (5.6), the generated Lie algebra is the Lie algebra containing

$$-i\tilde{H}_0, -i\tilde{H}_{x,1}, -i\tilde{H}_{y,1}, \ldots$$

and iterated commutators

$$[-i\tilde{H}_0, -i\tilde{H}_{x,1}], [-i\tilde{H}_0, -i\tilde{H}_{y,1}], [-i\tilde{H}_{x,1}, -i\tilde{H}_{y,1}], [[-i\tilde{H}_0, -i\tilde{H}_{x,1}], -i\tilde{H}_{x,1}], \ldots \; ;$$

see, e.g., [40, 93].

Theorem 21. *A bilinear control system, evolving on a connected compact Lie group is controllable if and only if the Lie algebra generated by the drift and control vector fields coincides with the algebra of the Lie group.*

Notice that $SU(N)$ is a connected compact Lie group [53] and hence one can establish controllability by a direct evaluation of the generated Lie algebra. However, the computation of the generated Lie algebra can be cumbersome when the dimension N becomes large. To obtain a simple criterion for checking whether the Lie algebra generated by the drift and control terms of (5.6) coincides with the Lie algebra of $SU(N)$, that is the set $\mathfrak{su}(N)$ of all traceless skew-hermitian $N \times N$-matrices, we assign to any Ising Hamiltonian of the form (5.4) a so-called spin-spin interaction graph, defined as follows

Definition 7. *Consider the free component of the Ising Hamiltonian $\tilde{H}_0$ defined as in (5.4). The spin-spin interaction graph is defined as an undirected graph with vertex set $\{1, 2, ..., N_p\}$, representing the number of spin-$\frac{1}{2}$ particles, and edges drawn between vertex k and j if and only if the coupling constant $J_{k,j}$ in (5.4) does not vanish.*

We can associate to the coupling constants a matrix C by setting

$$C_{k,j} = \begin{cases} J_{k,j} & if \quad k < j \\ J_{k,j} & if \quad j < k \end{cases}.$$

Notice that, since the spin-spin interaction graph is undirected, the matrix C is symmetric with zero entries on the diagonal. For the sake of clarity, we provide the following example of spin-spin interaction graph.

Example. Consider a system of 5 coupled spins that corresponds to the following matrix of coupling constants

$$C = \begin{pmatrix} 0 & * & 0 & * & * \\ * & 0 & * & 0 & 0 \\ 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & 0 & * \\ * & 0 & 0 & * & 0 \end{pmatrix},$$

where we denote by $*$ the non-zero entries. The spin-spin interaction graph corresponding to $(J_{k,j})$ is the shown in Figure 5.1.
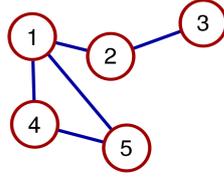

Figure 5.1: Connected spin-spin interaction graph.

$\triangle$

The following theorem provides a sufficient condition for controllability by using the notion of spin-spin interaction graph [19, 102].

Theorem 22. *If the spin-spin interaction graph of the drift term* (5.4) *is connected, then* (5.6) *is (density) operator controllable.*

Notice that, according to Definition 6, controllability implies that, for every element g in $SU(N)$, there exists a control such that the trajectory governed by (5.6) can be steered from I_N to g. However, Theorem 21 and Theorem 22 do not provide any information regarding the time T necessary to reach g. On the other hand, in many applications T has to be fixed a priori, and, consequently, informations on this aspects are required. To this purpose, we can use the semisimplicity property of a Lie group. A matrix Lie group is said to be semisimple if its corresponding Lie algebra is isomorphic to a direct sum of simple Lie algebras containing only trivial ideals. For a rigorous definition of a semisimple Lie group, see Appendix and [53]. This property is used to prove Theorem 7.2 in [68], that controllability at a given time T. This theorem is the following.

Theorem 23. *Consider a controllable right invariant control system evolving on a compact semisimple Lie group G. Then there exists $T > 0$ such that, for every $g, g' \in G$, there is a control that steers g into g' in exactly T units of time.*

Notice that $SU(N)$ is a semisimple matrix Lie group; see, e.g., [53]. Once controllability is established, by means of the semisimplicity of $SU(N)$ we can prove the following corollary to Theorem 23, which provides a strong controllability result for (5.6).

Corollary 2. *If* (5.6) *is (density) operator controllable then there exists a constant $T_* \geq 0$ such that for any $T \geq T_*$ one has $\mathcal{R}_T(I_N) = SU(N)$, i.e. for any sufficiently large $T \geq 0$ exact-time controllability of* (5.6) *holds.*

Proof. Assume that (5.6) is density operator controllable, then by Theorem 20, we have that $\mathcal{R}(I) = SU(N)$. Since $SU(N)$ is a compact and semisimple matrix Lie group, see, e.g., [53], then Theorem 23 [68] holds, hence there exists a constant time $T_* \geq 0$ such

that for every g and $g' \in SU(N)$ there is a control that steers g into g' in exactly T_* units of time. Now, let $T \geq T_*$ and let $g'' \in SU(N)$ be defined as $g'' := e^{-\tilde{H}_0(T-T_*)}g$. Consequently, the system will evolve from g into g'' in a time $\tilde{T} = T - T_* \geq 0$. Now, according to Theorem 23 [68], there exists a control that steers g'' into g' in exactly T_* units of time. Consequently, the evolution of the system from g to g' via g'' will be in $T = \tilde{T} + T_* \geq T_* > 0$ units of time, which completes our proof. $\qquad\square$

The result of Corollary 2 has one major drawback: it guarantees only the existence of $T_* \geq 0$, but it does not provide any estimate for $T_* \geq 0$. The literature regarding estimates of a minimum (infimum) controllability time is at its infancy, and in particular, there are only few works that establish an optimal time for specific spin transitions, see, [9, 39, 70, 72].

The previous controllability result on (5.6) can be carried over to (5.3) and (5.7), respectively. We have the following theorem.

Theorem 24.

(a) *System (5.7) is never controllable on the special orthogonal group $SO(N_x)$;*

(b) *If (5.6) is controllable, then the reachable set $\mathcal{R}(I)$ of (5.7) coincides with a closed subgroup of $SO(N_x)$ which is locally isomorphic to $SU(N)$;*

(c) *System (5.3) with $N > 2$ is neither controllable on $\mathbb{R}^{N_x}$ nor on S^{N_x-1}.*

(d) *If (5.6) is controllable, then the reachable set of (5.6) coincides with the image of $\mathcal{O}(\rho_0)$ under the map $\mathcal{V}$ defined in (2.125).*

Proof. (a): Recall that $N_x = N^2$. A simple argument concerning the dimensions of $SU(N)$ and $SO(N^2)$ shows that (5.7) is not controllable on $SO(N^2)$. In fact, for $N > 1$ we have that

$$dim(SU(N)) = N^2 - 1 < N^2(N^2 - 1)/2 = dim(SO(N^2)) .$$

(b): To see that the reachable set $\mathcal{R}(I)$ of (5.7) coincides with a closed subgroup of $SO(N^2)$, consider $\rho(t) = \Upsilon(t)\rho_0\Upsilon(t)^*$. Now, similarly as in Section 2.4.1 we write that $x = T^{-1}vec(\rho)$. Hence, it follows

$$x(t) = T^{-1}vec(\Upsilon(t)\rho_0\Upsilon(t)^*)$$
$$\Rightarrow \Theta(t)x_0 = T^{-1}\big((\Upsilon(t)^*)^T \otimes \Upsilon(t)\big)vec(\rho_0)$$
$$\Rightarrow \Theta(t)x_0 = T^{-1}\big((\Upsilon(t)^*)^T \otimes \Upsilon(t)\big)Tx_0$$
$$\Rightarrow \Theta(t) = T^{-1}\big((\Upsilon(t)^*)^T \otimes \Upsilon(t)\big)T .$$

Notice that $S_G := \{\Upsilon \otimes \Upsilon : \Upsilon \in SU(N)\}$ is a closed subgroup of $SU(N^2)$ and the same holds for $T^{-1}S_G T$. Moreover, since T is defined as in Section 2.4.1 and $T^*T = I$, we notice that each entry of $\Theta(t)$ is obtained as

$$\Theta_{k,j}(t) = T_k^*\big((\Upsilon(t)^*)^T \otimes \Upsilon(t)\big)T_j$$
$$= vec(\tilde{B}_k)^*\big((\Upsilon(t)^*)^T \otimes \Upsilon(t)\big)vec(\tilde{B}_j)$$
$$= \langle \tilde{B}_k | \big((\Upsilon(t)^*)^T \tilde{B}_j \Upsilon(t)\big) \rangle ,$$

where $\langle \cdot | \cdot \rangle$ is defined in (2.91), and consequently $\Theta_{k,j}(t)$ is real. It follows that $T^{-1}S_G T$ is a closed subgroup of $SO(N^2)$.

Furthermore, we have that

$$SO(N_x) \supset \mathcal{R}(I) = T^{-1}S_G T \leftrightharpoons S_G \leftrightharpoons SU(N) \,,$$

where the isomorphisms are understood locally, which means that the corresponding Lie algebras are isomorphic.

(c): System (5.3) is not controllable on $\mathbb{R}^{N^2}$ due to the fact that (5.3) is norm preserving. Further, arguments regarding the dimension of S^{N^2-1}, the unit sphere of $\mathbb{R}^{N^2}$, and the dimension of the orbit $\mathcal{O}(\rho_0)$ [54], show that even controllability on S^{N^2-1} fails. In particular, recalling that the dimension of the orbit $\mathcal{O}(\rho_0)$ is given by

$$dim(\mathcal{O}(\rho_0)) = dim(SU(N)) - dim(Stab(\rho_0)) \,,$$

where $Stab(\rho_0) = \{\Upsilon \in SU(N) \,:\, \Upsilon\rho_0\Upsilon^* = \rho_0\}$ is the stabilizer subgroup [39, 54], and following arguments as in [8], we have that

$$dim(\mathcal{O}(\rho_0)) \leq dim(SU(N)) - (N-1) = N^2 - N \,,$$

which means that $dim(\mathcal{O}(\rho_0)) < dim(S^{N^2-1})$ for $N > 2$.

(d): Just apply the map $\mathcal{V}$ defined in (2.125) to the orbit $\mathcal{O}(\rho_0) := \{\Upsilon\rho_0\Upsilon^* \mid \Upsilon \in SU(N)\}$. $\qquad\square$

Notice that in our discussion about controllability, we do not pay attention on the space of admissible controls U. This is due to the fact that, Theorem 21 and Theorem 23, proved in [68], hold for the class of piecewise-constant control functions. Consequently, if one enlarge this space to $U = L^2$, then the controllability property is preserved.

Next, we provide two examples of controllable spin systems.

Example. (one spin$-\frac{1}{2}$ system)
We consider the following LvNM equation in rotating frame coordinates

$$\dot{\tilde{\rho}} = -i[\tilde{H}, \tilde{\rho}] \,,$$

where the effective Hamiltonian is given by

$$\tilde{H} = (\nu_0 - \omega)I_z + u_1 I_x + u_2 I_y \,.$$

This systems admits the following lifted system evolving on $SU(N)$

$$\dot{\Upsilon} = -i\tilde{H}\Upsilon \,.$$

Since $\{iI_x, iI_y, iI_z\}$ is a basis for $\mathfrak{su}(2)$, then the Lie algebra generated by the drift and the input vector fields coincides with $\mathfrak{su}(2)$. Now, we notice that $SU(2)$ is a compact and connected matrix Lie group, see, e.g., [53]. Consequently, Theorem 21 holds and the lifted system is operator controllable, and Theorem 20 implies density operator controllability.
$$\triangle$$

Example. **(2 spin$-\frac{1}{2}$ system)**
We consider the following LvNM equation in rotating frame coordinates

$$\dot{\tilde{\rho}} = -i[\tilde{H}, \tilde{\rho}] \, ,$$

where the effective Hamiltonian is given by

$$\tilde{H} = J\hat{I}_z\hat{S}_z + u_1\hat{I}_x + u_2\hat{I}_y + u_3\hat{S}_x + u_4\hat{S}_y \, .$$

Notice that the spin-spin interaction graph of the drift term is connected, hence the (density) operator controllability of this system is obtained by means of Theorem 22.

We remark that, in principle, it would be possible to apply Theorem 21 directly. However, the computation of all the iterated commutators makes the application cumbersome, whereas Theorem 22 yields controllability straightforwardly.

$\triangle$

5.2 A continuation method for the exact-control of quantum systems

In this section, we present a continuation method for computing control functions that solve the exact-control problem (5.1). In order to explain the continuation method, consider the following two optimization problems

$$
\begin{aligned}
\min_{x,u} \quad & \tilde{J}(x,u) := \frac{1}{2}\|x(T) - x_T\|_2^2 \\
\text{s.t.} \quad & \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]x \, , \text{ in } (0,T] \, , \; x(0) = x_0 \, ,
\end{aligned}
\tag{$\tilde{P}$}
$$

and

$$
\begin{aligned}
\min_{x,u} \quad & J_\nu(x,u) := \frac{1}{2}\|x(T) - x_T\|_2^2 + \frac{\nu}{2}\sum_{n=1}^{N_C}\|u_n\|_{L^2}^2 \\
\text{s.t.} \quad & \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]x \, , \text{ in } (0,T] \, , \; x(0) = x_0 \, ,
\end{aligned}
\tag{P_ν}
$$

and we recall that, we consider closed quantum systems without relaxation phenomena.

The continuation strategy can be heuristically described as follows. Consider a sequence of weight parameters $\nu_k \to 0$, and denote by $\{J_{\nu_k}\}_{k=1}^{\infty}$ the corresponding sequence of cost functionals of (P_ν). This sequence converges to $J_0 = \frac{1}{2}\|x(T) - x_T\|_2^2$, that is the cost functional of $(\tilde{P})$, and a control $\tilde{u}$ such that $J_0 = 0$ is an exact-control, that is $\tilde{u}$ solve (5.1). The sequence $\{J_{\nu_k}\}_{k=1}^{\infty}$ generates a sequence of optimal control problems, and $\{u^k\}_{k=1}^{\infty}$ is the corresponding sequence of optimal controls solutions, that is u^k solves (P_ν) with $\nu = \nu_k$. Hence, the continuation procedure aims to reduce progressively the L^2-regularization term in (P_ν), in order to obtain at each iteration a better approximation to the solution to $(\tilde{P})$. Notice that each element u^k of the sequence $\{u^k\}_{k=1}^{\infty}$ is numerically generated by solving (P_ν) corresponding to $\nu = \nu_k$, and using as initial guess for the iterative optimal control solver the previous element u^{k-1} of the control sequence.

Next, we put the above heuristic procedure in a rigorous framework. In Theorem 25, we prove convergence of the continuation procedure. In particular, we prove that one can extract from the sequence $\{u^k\}_{k=1}^{\infty}$ a subsequence converging strongly in L^2 to a solution to $(\tilde{P})$. In Theorem 26, we prove a convergence ratio of the continuation method. Notice that, to prove Theorem 26 we need the intermediate results given in Lemma 7 and Lemma 8. These convergence results are proved in the case that the entire $L^2((0,T);\mathbb{R}^{N_C})$ represents the admissible control space U_{ad}. In Theorem 27, we extend the convergence results also for the case that U_{ad} is a closed, convex and bounded subset of $L^2((0,T);\mathbb{R}^{N_C})$. Theorem 28 proves a more general convergence result, that is useful to analyze the numerical experiments discussed in Section 7.

First, some remarks regarding the existence of solutions to (P_{ν}) and $(\tilde{P})$, are necessary for our purposes. Notice that the existence of solutions to problem (P_{ν}) is proved in Theorem 1. Regarding problem $(\tilde{P})$, if the target x_T belongs to the reachable set for the given starting point x_0, then the solution to $(\tilde{P})$ corresponds to $\tilde{J} = 0$. For this reason, to guarantee the existence of a solution to $(\tilde{P})$ we make the following assumption.

Assumption 2. *The LvNM equation* (5.3) *is controllable. Moreover, the target point x_T belongs to the set of the points reachable from x_0 in time T, such that a solution to $(\tilde{P})$ corresponds to $\tilde{J} = 0$.*

Theorem 25. *Consider problems $(\tilde{P})$ and (P_{ν}). Let $\{\nu_k\}_{k=1}^{\infty}$ be a positive sequence such that $\nu_k \to 0$ as $k \to \infty$. denote by $\tilde{u}$ a solution to $(\tilde{P})$ and u^k a solution to (P_{ν}) corresponding to ν_k. Then there exists a convergent subsequence, that is $u^{k_j} \to u^0$ in $L^2((0,T);\mathbb{R}^{N_C})$ as $j \to \infty$, and u^0 is a minimum-norm solution to $(\tilde{P})$, as follows*

$$\|u^0\|_{L^2} \leq \|\tilde{u}\|_{L^2} \,,$$

for all solutions $\tilde{u}$ of $(\tilde{P})$.

Proof. Using the monotonicity of $\tilde{J}$ and J_{ν_k}, we have

$$\tilde{J}(\tilde{u}) \leq \tilde{J}(u^k) = J_{\nu_k}(u^k) - \frac{\nu_k}{2}\|u^k\|_{L^2}^2 \leq J_{\nu_k}(\tilde{u}) - \frac{\nu_k}{2}\|u^k\|_{L^2}^2$$
$$= \tilde{J}(\tilde{u}) + \frac{\nu_k}{2}\|\tilde{u}\|_{L^2}^2 - \frac{\nu_k}{2}\|u^k\|_{L^2}^2 \,.$$

This implies that

$$\frac{\nu_k}{2}\|\tilde{u}\|_{L^2}^2 - \frac{\nu_k}{2}\|u^k\|_{L^2}^2 \geq 0 \Rightarrow \|\tilde{u}\|_{L^2} \geq \|u^k\|_{L^2} \,, \tag{5.11}$$

which means that u^k is bounded. Hence, by reflexivity of L^2, we can extract a weakly convergent subsequence, that is, $u^{k_j} \rightharpoonup u^0$ in $L^2((0,T);\mathbb{R}^{N_C})$ as $j \to \infty$. By Proposition 3, we have that $x^{k_j}(T) \to x^0(T)$ and consequently, we have the following

$$J_{\nu_{k_j}}(u^{k_j}) = \frac{1}{2}\|\underbrace{x^{k_j}(T)}_{\to x^0(T)} - x_T\|_2^2 \;+\; \underbrace{\frac{\nu_{k_j}}{2}\|u^{k_j}\|_{L^2}^2}_{\to 0} \,.$$

Notice that $J_{\nu_k}(u^k) \leq J_{\nu_k}(u)$ for all $u \in L^2((0,T);\mathbb{R}^{N_C})$. Hence, by lower-semicontinuity of J_{ν_k}, we obtain the following

$$\tilde{J}(u^0) = \frac{1}{2}\|x^0(T) - x_T\|_2^2 \leq \liminf_{j\to\infty} J_{\nu_{k_j}}(u^{k_j}) \leq \liminf_{j\to\infty} J_{\nu_{k_j}}(u) \,, \ \forall u \in L^2((0,T);\mathbb{R}^{N_C})$$
$$\Rightarrow \tilde{J}(u^0) \leq \tilde{J}(u) \,, \ \forall u \in L^2((0,T);\mathbb{R}^{N_C}) \,.$$

Hence, u^0 is a solution to $(\tilde{P})$.

Now, by (5.11) and lower-semicontinuity of the norm, we obtain the following

$$\|u^0\|_{L^2} \leq \liminf_{j \to \infty} \|u^{k_j}\|_{L^2} \leq \|\tilde{u}\|_{L^2} \Rightarrow \|u^0\|_{L^2} \leq \|\tilde{u}\|_{L^2} ,$$

for all solutions $\tilde{u}$ to $(\tilde{P})$. Hence, u^0 is a minimum norm solution to $(\tilde{P})$.

Next, recall that (5.11) holds for all the solutions $\tilde{u}$. Therefore, we have that $\|u^0\|_{L^2} \geq \|u^{k_j}\|_{L^2}$. Therefore, we obtain

$$\|u^0\|_{L^2} \leq \liminf_{j \to \infty} \|u^{k_j}\|_{L^2} \leq \|u^0\|_{L^2} .$$

Recalling that $u^{k_j} \rightharpoonup u^0$, we obtain the strong convergence $u^{k_j} \to u^0$. $\qquad\square$

In the sequel, we study the dependence of $\|u_\nu - \tilde{u}\|_{L^2}$ on ν and prove how fast this norm decays as $\nu \to 0$. We recall that the optimality system of problem (P_ν) is given in Theorem 2.

Lemma 7. *Let $(\tilde{x}, \tilde{u})$ be a solution to $(\tilde{P})$ and (x_ν, u_ν) be a solution to (P_ν). Assume that there exists a $\rho > 0$ and $\gamma > 0$ such that for all $u \in L^2((0,T); \mathbb{R}^{N_C})$, with $\|\tilde{u} - u\|_{L^2} < \rho$, the following holds*

$$\tilde{J}(x, u) \geq \gamma \|\delta x(T)\|_2^2 , \tag{5.12}$$

and δx solves

$$\dot{\delta x} = \left[A + \sum_{n=1}^{N_C} \tilde{u}_n B_n \right] \delta x + \sum_{n=1}^{N_C} (\tilde{u}_n - u_n) B_n \tilde{x} , \quad in \ (0, T] , \ \delta x(0) = 0 . \tag{5.13}$$

Then it holds that

$$\frac{\nu}{2} \|u_\nu - \tilde{u}\|_{L^2}^2 + \gamma \|\delta x(T)\|_2^2 \leq \nu \langle \tilde{u}, \tilde{u} - u_\nu \rangle_{L^2} .$$

Proof. For ν small enough, we have the following

$$\tilde{J}(x_\nu, u_\nu) \geq \gamma \|\delta x_\nu(T)\|_2^2 ,$$

where δx_ν denotes the solution to (5.13) corresponding to $(\tilde{u} - u_\nu)$. Since $\tilde{J}(\tilde{x}, \tilde{u}) = 0$, we have

$$\gamma \|\delta x_\nu(T)\|_2^2 \leq \tilde{J}(x_\nu, u_\nu) - \tilde{J}(\tilde{x}, \tilde{u})$$
$$= J_\nu(x_\nu, u_\nu) - J_\nu(\tilde{x}, \tilde{u}) - \nu \left(\frac{1}{2} \|u_\nu\|_{L^2}^2 - \frac{1}{2} \|\tilde{u}\|_{L^2}^2 \right)$$
$$\leq -\nu \left(\frac{1}{2} \|u_\nu\|_{L^2}^2 - \frac{1}{2} \|\tilde{u}\|_{L^2}^2 \right) ,$$

where we used the fact that $J_\nu(x_\nu, u_\nu) - J_\nu(\tilde{x}, \tilde{u}) \leq 0$. Now, since the following holds

$$\frac{1}{2} \|u_\nu\|_{L^2}^2 - \frac{1}{2} \|\tilde{u}\|_{L^2}^2 = \frac{1}{2} \|u_\nu\|_{L^2}^2 + \frac{1}{2} \|\tilde{u}\|_{L^2}^2 - \langle \tilde{u}, \tilde{u} \rangle_{L^2}$$
$$= \frac{1}{2} \|u_\nu\|_{L^2}^2 + \frac{1}{2} \|\tilde{u}\|_{L^2}^2 - \langle \tilde{u}, u_\nu \rangle_{L^2} - \langle \tilde{u}, \tilde{u} - u_\nu \rangle_{L^2}$$
$$= \frac{1}{2} \|u_\nu - \tilde{u}\|_{L^2}^2 - \langle \tilde{u}, \tilde{u} - u_\nu \rangle_{L^2} ,$$

we conclude that

$$\gamma \|\delta x_\nu(T)\|_2^2 \leq -\frac{\nu}{2} \|u_\nu - \tilde{u}\|_{L^2}^2 + \nu \langle \tilde{u}, \tilde{u} - u_\nu \rangle_{L^2} .$$

$\qquad\square$

We remark that, in Lemma 7 it is assumed (5.12), that is a local convexity assumption of problem $(\tilde{P})$. We are not able to prove (5.12), however in Section 7 we demonstrate by numerical experiments that this estimate holds in practice. Moreover, according to (3.23) and (3.25a) in Proposition 6, it holds that

$$\|\delta x(T) - (x(T) - x_T)\|_2 \leq c_\vartheta \|u - \tilde{u}\|_{L^2}^2 ,$$

and by using the reverse triangular inequality we get

$$\|\delta x(T)\|_2 \leq \|x(T) - x_T\|_2 + c_\vartheta \|u - \tilde{u}\|_{L^2}^2 ,$$

which resembles (5.12) (recalling that $\tilde{J}(x, u) = \frac{1}{2}\|x(T) - x_T\|_2^2$), and shows that the difference between $\|\delta x(T)\|_2$ and $\|x(T) - x_T\|_2$ decays, as $u \to \tilde{u}$, with an order of $\|u - \tilde{u}\|_{L^2}^2$.

Now, consider the following problem

$$\min_{x,u} \quad J_u(x, u) := \frac{1}{2} \sum_{n=1}^{N_C} \|u_n\|_{L^2}^2$$

$$\text{s.t.} \quad \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x , \text{ in } (0, T) , \ x(0) = x_0 , \ x(T) = x_T . \tag{P_u}$$

Since we assume controllability of the quantum system given by the LvNM equation (5.3) and Assumption 2, then (P_u) admits a solution. Notice also that a solution to (P_u) is also a solution to $(\tilde{P})$.

In the following result, we assume existence of a Lagrange multiplier q corresponding to the constraint of problem (P_u). The existence of such a Lagrange multiplier is not a trivial issue. In principle, in order to use standard existence result, one could prove controllability of the linearization of the end-point map constraint. Unfortunately, according to our experience, this controllability problem is not simple to solve, hence we make this existence assumption which is necessary for the purpose of our theoretical discussion.

Lemma 8. *Assume that there is a q_T such that $\tilde{u}$ satisfies*

$$\tilde{u}_n = \langle B_n \tilde{x}, q \rangle , \ n = 1, ..., N_C ,$$

where q is a Lagrange multiplier corresponding to the constraint of (P_u) and solves

$$-\dot{q} = \left[A + \sum_{n=1}^{N_C} \tilde{u}_n B_n \right]^* q , \ q(T) = q_T . \tag{5.14}$$

Then the following holds

$$\langle \tilde{u}, \tilde{u} - u_\nu \rangle_{L^2} = \langle q_T, \delta x_\nu(T) \rangle , \tag{5.15}$$

with δx_ν given by Lemma 7.

Proof. We have the following

$$\langle \tilde{u}, \tilde{u} - u_\nu \rangle_{L^2} = \sum_n \langle \langle B_n \tilde{x}, q \rangle, \tilde{u}_n - u_{\nu,n} \rangle_{L^2}$$

$$= \langle q, \sum_n (\tilde{u}_n - u_{\nu,n}) B_n \tilde{x} \rangle_{L^2} = \langle q, \dot{\delta x}_\nu - \left(A + \sum_n \tilde{u}_n B_n \right) \delta x_\nu \rangle_{L^2} \tag{5.16}$$

$$= -\langle \dot{q} + \left(A + \sum_n \tilde{u}_n B_n \right)^* q, \delta x_\nu \rangle_{L^2} + \langle q_T, \delta x_\nu \rangle \big|_0^T = \langle q_T, \delta x_\nu(T) \rangle ,$$

where $\dot{q} + \left(A + \sum_{n=1}^{N_C} \tilde{u}_n B_n \right)^* q = 0$ holds from (5.14). $\qquad\square$

In the following theorem, we prove an estimate which express the behaviour of $\|\tilde{u} - u_\nu\|_{L^2}$ as $\nu \to 0$.

Theorem 26. *Let $(\tilde{x}, \tilde{u})$ be a solution to (P_u) and let (x_ν, u_ν) be a solution to (P_ν). With the assumptions of Lemma 7 and Lemma 8, the following estimate holds*

$$\|\tilde{u} - u_\nu\|_{L^2} = O(\sqrt{\nu}) \ \ for \ \nu \to 0 \ . \tag{5.17}$$

Proof. From Lemma 7, we have the following

$$\frac{\nu}{2}\|u_\nu - \tilde{u}\|_{L^2}^2 + \gamma\|\delta x_\nu(T)\|_2^2 \leq \nu\langle\tilde{u}, \tilde{u} - u_\nu\rangle_{L^2} \ .$$

From Lemma 8, we write that

$$\nu\langle\tilde{u}, \tilde{u} - u_\nu\rangle_{L^2} = \nu\langle q_T, \delta x_\nu(T)\rangle \ .$$

Using Cauchy and Cauchy-Schwarz inequalities, we obtain

$$\nu\langle q_T, \delta x_\nu(T)\rangle = \langle\frac{\nu}{\sqrt{2\gamma}}q_T, \sqrt{2\gamma}\delta x_\nu(T)\rangle \leq \frac{\nu^2}{4\gamma}\|q_T\|_2^2 + \gamma\|\delta x_\nu(T)\|_2^2 \ .$$

Therefore we get the following inequalities

$$\frac{\nu}{2}\|u_\nu - \tilde{u}\|_{L^2}^2 + \gamma\|\delta x_\nu(T)\|_2^2 \leq \nu\langle q_T, \delta x_\nu(T)\rangle \leq \gamma\|\delta x_\nu(T)\|_2^2 + \frac{\nu^2}{4\gamma}\|q_T\|_2^2 \ ,$$

which implies the following

$$\|u_\nu - \tilde{u}\|_{L^2} \leq \sqrt{\nu}\Big(\|q_T\|_2\frac{1}{\sqrt{2\gamma}}\Big) \ .$$

$\square$

Notice that, the convergence results proved in Theorem 25 and Theorem 26 are valid in the case that the set of admissible controls U_{ad} coincides with the entire space $L^2((0,T); \mathbb{R}^{N_C})$. A natural extension of these results is to consider that U_{ad} is only a subset of $L^2((0,T); \mathbb{R}^{N_C})$. For this constrained case, we can prove the same convergence results in the following theorem.

Theorem 27. *Consider problems $(\tilde{P})$, (P_ν) and (P_u) with the additional constraint $u \in U_{ad}$, where U_{ad} is a closed, convex and bounded subset of $L^2((0,T); \mathbb{R}^{N_C})$. Let $\{\nu_k\}_{k=1}^\infty$ be a positive sequence such that $\nu_k \to 0$ as $k \to \infty$. denote by $\tilde{u}$ a solution to $(\tilde{P})$ and u^k a solution to (P_ν) corresponding to ν_k. Then there exists a convergent subsequence, that is $u^{k_j} \to u^0$ in $L^2((0,T); \mathbb{R}^{N_C})$ as $j \to \infty$, and $u^0 \in U_{ad}$ is a minimum-norm solution to $(\tilde{P})$, that is*

$$\|u^0\|_{L^2} \leq \|\tilde{u}\|_{L^2} \ ,$$

for all solutions $\tilde{u}$ of $(\tilde{P})$. Hence, u^0 is a solution to (P_u). Moreover, assume that

- *there exists a $\rho > 0$ and $\gamma > 0$ such, that for all u with $\|\tilde{u} - u\|_{L^2} < \rho$, the estimate (5.12) holds;*

- *there exists a q_T such that $\tilde{u}$ satisfies*

$$\sum_{n=1}^{N_C} \langle \tilde{u}_n - \langle B_n \tilde{x}, q\rangle, v_n - \tilde{u}_n\rangle_{L^2} \geq 0 \ , \ \forall v \in U_{ad} \ , \tag{5.18}$$

where q is a Lagrange multiplier corresponding to the state-constraint of (P_u) and solves (5.14).

Then the following estimate holds

$$\|\tilde{u} - u_\nu\|_{L^2} = O(\sqrt{\nu}) \ for \ \nu \to 0 \ . \tag{5.19}$$

Proof. The proof follows exactly the same arguments used in Theorem 25, Lemma 7, Lemma 8 and Theorem 26 with few small changes:

- in Theorem 25: the subsequence u^{k_j} converges weakly to u^0 in $L^2((0,T); \mathbb{R}^{N_C})$, and $u^0 \in U_{ad}$ because U_{ad} is closed, convex and bounded, hence weakly sequentially compact;

- in Lemma 8: from (5.18), by fixing $v = u_\nu$ we have

$$\langle \tilde{u}, u_\nu - \tilde{u}\rangle_{L^2} \geq \sum_n \langle \langle B_n \tilde{x}, q\rangle, u_{\nu,n} - \tilde{u}_n\rangle_{L^2} \ , \tag{5.20}$$

which implies that

$$\langle \tilde{u}, \tilde{u} - u_\nu\rangle_{L^2} \leq \sum_n \langle \langle B_n \tilde{x}, q\rangle, \tilde{u}_n - u_{\nu,n}\rangle_{L^2} \ . \tag{5.21}$$

Now, similarly as in (5.16), we obtain that

$$\sum_n \langle \langle B_n \tilde{x}, q\rangle, \tilde{u}_n - u_{\nu,n}\rangle_{L^2} = \langle q_T, \delta x_\nu(T)\rangle \ . \tag{5.22}$$

From (5.21) and (5.22) we get

$$\langle \tilde{u}, \tilde{u} - u_\nu\rangle_{L^2} \leq \langle q_T, \delta x_\nu(T)\rangle \ . \tag{5.23}$$

- in Theorem 26: use (5.20)-(5.23) instead of (5.15).

$\square$

We remark that, the estimate (5.17) is not sharp, in the sense that numerical evidence shows in some cases a faster rate of convergence. This fact will be clear from the numerical experiments discussed in Section 7. In order to provide this numerical evidence with a theoretical explanation, we consider the following theorem.

Theorem 28. *Let $(\tilde{x}, \tilde{u})$ be a solution to $(\tilde{P})$ and (x_ν, u_ν) be a solution to (P_ν). Assume that there exists a $\rho > 0$ and $\gamma > 0$ such that for all u, with $\|\tilde{u} - u\|_{L^2} < \rho$, the following holds*

$$\tilde{J}(x, u) \geq \gamma \|\tilde{u} - u\|_{L^2}^\alpha \ , \tag{5.24}$$

with $\alpha \geq 2$. Then it holds that

$$\|\tilde{u} - u_\nu\|_{L^2} = O(\nu^{1/(\alpha-1)}) \ for \ \nu \to 0 \ . \tag{5.25}$$

Proof. For ν small enough, as in Lemma 7, we obtain that

$$\gamma\|\tilde{u} - u_\nu\|_{L^2}^\alpha \leq -\frac{\nu}{2}\|u_\nu - \tilde{u}\|_{L^2}^2 + \nu\langle\tilde{u}, \tilde{u} - u_\nu\rangle_{L^2}$$
$$\leq \nu\|\tilde{u}\|_{L^2}\|\tilde{u} - u_\nu\|_{L^2} \, ,$$

which implies that

$$\|\tilde{u} - u_\nu\|_{L^2} \leq \nu^{1/(\alpha-1)}\left(\frac{1}{\gamma}\|\tilde{u}\|_{L^2}\right)^{1/(\alpha-1)} \, .$$

$\square$

Notice that in Theorem 28, we consider a relaxed version of the hypothesis (5.12). In particular, the exponent α is assumed to be greater or equal than 2. This choice is motivated by the fact that, recalling $\tilde{J}(x, u) = \frac{1}{2}\|x(T) - x_T\|_2^2$ and by using Proposition 7, it holds that

$$\gamma\|u - \tilde{u}\|_{L^2}^\alpha \leq \frac{1}{2}\|x(T) - x_T\|_2^2 \leq \frac{\hat{c}_1^2}{2}\|u - \tilde{u}\|_{L^2}^2 \, ,$$

where it is clear that α has to be greater or equal than 2. The estimate (5.24) is realistic. For instance, can be interpreted for $\alpha = 2$ as coercivity of the reduced Hessian operator: let $\tilde{u}$ be a solution with $\tilde{J}_r(\tilde{u}) = 0$, then we can write the following

$$\tilde{J}_r(u) \approx \tilde{J}_r(\tilde{u}) + \langle\nabla_u\tilde{J}_r(\tilde{u}), u - \tilde{u}\rangle_{L^2} + \langle\mathcal{J}_r(\tilde{u})(u - \tilde{u}), u - \tilde{u}\rangle_{L^2}$$
$$= \langle\mathcal{J}_r(\tilde{u})(u - \tilde{u}), u - \tilde{u}\rangle_{L^2}$$
$$\geq \tilde{\gamma}\|u - \tilde{u}\|_{L^2}^2 \, ,$$

where $\tilde{\gamma}$ is a positive constant.

The following result is a direct consequence of the above theorems.

Corollary 3. *Let $(\tilde{x}, \tilde{u})$ be a solution to (P_u) and consequently of $(\tilde{P})$. Consider the weight parameters sequence $\{\nu_k\}_{k=1}^\infty$ defined by the formula $\nu_{k+1} = \hat{\gamma}\nu_k$, where $\hat{\gamma} \in (0, 1)$. Then it holds that $\|u_{\nu_k} - u_{\nu_{k+1}}\|_{L^2} = O(\nu_k^\alpha)$.*

Proof. By Theorem 26 and using the triangular inequality, we have that

$$\|u_{\nu_k} - u_{\nu_{k+1}}\|_{L^2} \leq \|u_{\nu_k} - \tilde{u}\|_{L^2} + \|u_{\nu_{k+1}} - \tilde{u}\|_{L^2} \, ,$$

and consequently

$$\|u_{\nu_k} - u_{\nu_{k+1}}\|_{L^2} = O(\nu_k^\alpha) + O(\nu_{k+1}^\alpha) \, .$$

Now, since $\nu_{k+1} = \hat{\gamma}\nu_k$ we get that $\nu_{k+1}^\alpha = \gamma^\alpha\nu_k^\alpha = \gamma^\alpha\nu_k^\alpha = O(\nu_k^\alpha)$, and hence $\|u_{\nu_k} - u_{\nu_{k+1}}\|_{L^2} = O(\nu_k^\alpha)$. $\square$

5.3 A shooting-type method for the exact-control of quantum systems

In this section, we present a computational method to solve (5.1) with the additional requirement that the control functions have minimal norm. For this purpose, we consider the following steps.

(A) Embed (5.1) in a minimum-norm exact-controllability problem, that is (5.26).

(B) Consider the first-order optimality system of (5.26) given by (5.28).

(C) Embed (5.28) in the optimization problem (5.30).

(D) Derive the first-order optimality system of (5.30) in Proposition 2.

(E) Consider problem (5.30) in the reduced form, that is (5.36), having optimality system given by Proposition 16.

(F) Solve (5.36) by means of optimal control techniques. Notice that regularity properties of (5.36) are investigated in Section 5.3.1.

Notice that in the formulation of the optimality system (5.28) in step (B) it results that a boundary condition is undetermined, and hence has to be considered an unknown. The fact that a boundary condition has to be treated as an unknown motivates the fact that our computational scheme is addressed as a shooting-type method.

We start by discussing the step (A). Notice that, since (5.1) is a time-boundary-value problem, it is possible to solve it using the class shooting methods [112], although these methods have been less investigated in the case of bilinear control. However, problem (5.1) may admit many solutions, and it becomes necessary to complement the problem with a constraint on u.

A suitable way to constraint the controls is to consider (5.1) embedded in an optimization problem. For this reason, we focus on the following

$$
\begin{aligned}
\min_{x,u} \quad & J(x,u) := \frac{1}{2} \sum_{n=1}^{N_C} \|u_n\|_{L^2}^2 \\
\text{s.t.} \quad & \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \ \text{in } (0,T) \ , \ x(0) = x_0 \ , \ x(T) = x_T \\
& x \in H^1((0,T); \mathbb{R}^{N_x}) \text{ and } u \in L^2((0,T); \mathbb{R}^{N_C}) \ .
\end{aligned}
\tag{5.26}
$$

Notice that (5.26) admits a solution if the target x_T belongs to the set of all points reachable at time T from a given starting point x_0. Moreover, problems (5.1) and (5.26) are not equivalent. A solution of (5.26) is a minimum L^2-norm solution and solves also (5.1). On the other hand, a solution of (5.1) is not necessarily a solution to (5.26). We remark that, problem (5.26) is already considered in Section 5.2 to prove Lemma 8, and in this section we maintain Assumption 2.

Next, step (B) is addressed, and we discuss first-order optimality conditions for (5.26). To this purpose, the following assumption regarding the existence of Lagrange multipliers is necessary.

Assumption 3. *There exist Lagrange multipliers $p_T \in \mathbb{R}^{N_x}$ and $p \in H^1((0,T); \mathbb{R}^{N_x})$ corresponding to the constraint equation of the optimization problem (5.26). Moreover, p satisfies the following adjoint equation*

$$
-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* p \ , \ \text{in } [0,T) \ , \ p(T) = p_T \ .
\tag{5.27}
$$

In Assumption 3, we consider that there exists a vector p_T such that the corresponding solution p is the Lagrange multiplier associated with the state x. Notice that p_T is unknown and p is uniquely determined by p_T and the control u. Further, notice that, once the existence of $p \in H^1((0,T); \mathbb{R}^{N_x})$ is assumed, then (5.27) can be obtained by means of the standard Lagrange function approach.

A solution to (5.26) is characterized by the following first-order optimality system,

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \; , \;\; x(0) = x_0 \; , \;\; x(T) = x_T \tag{5.28a}$$

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^{*} p \; , \;\; p(T) = p_T \tag{5.28b}$$

$$u_n - \langle B_n x, p \rangle = 0 \; , \;\; n = 1, ..., N_C \; , \tag{5.28c}$$

where $\langle \cdot, \cdot \rangle$ represents the Euclidean scalar product.

Because of (5.28a), there is no clear approach of how to solve (5.28). For this reason, in the next step, we reformulate (5.28) in such a way that it can be solved by using appropriate optimization techniques.

Now, we address step (C). In order to solve (5.28), we consider the map

$$\mathcal{G} : H^1((0,T); \mathbb{R}^{N_x}) \times L^2((0,T); \mathbb{R}^{N_C}) \times H^1((0,T); \mathbb{R}^{N_x}) \to L^2((0,T); \mathbb{R}^{N_C}) \times \mathbb{R}^{N_x} \; ,$$

defined as follows

$$\mathcal{G}(x,u,p) := \begin{pmatrix} u_1 - \langle B_1 x, p \rangle \\ \vdots \\ u_{N_C} - \langle B_{N_C} x, p \rangle \\ x(T) - x_T \end{pmatrix} . \tag{5.29}$$

Since this map is obtained by using the gradient component (5.28c) and the terminal condition of (5.28a), a triple (x, u, p) is a solution of (5.28), and a stationary point for (P_u), if and only if it is a root of $\mathcal{G}$ with x and p solutions to (5.28a) and (5.28b), respectively.

We remark that, it could be possible to compute a root for $\mathcal{G}$ using a Newton method, however, according to our experience, the corresponding Jacobian operator is not sufficiently regular to be used successfully in computational algorithms. For this reason, in order to compute a root (x, u, p) of $\mathcal{G}$, we define our main optimization problem as follows

$$\min_{x,u,p} \;\; G(x,u,p) := \frac{1}{2} |||\mathcal{G}(x,u,p)|||^2 = \frac{1}{2} \sum_{n=1}^{N_C} \| u_n - \langle B_n x, p \rangle \|_{L^2}^2 + \frac{1}{2} \| x(T) - x_T \|_2^2$$

$$\text{s.t.} \;\; \dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \; , \;\; \text{in } (0,T] \; , \;\; x(0) = x_0 \tag{5.30}$$

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^{*} p \; , \;\; \text{in } [0,T) \; , \;\; p(T) = p_T$$

$$x, p \in H^1((0,T); \mathbb{R}^{N_x}) \;\; \text{and} \;\; u \in L^2((0,T); \mathbb{R}^{N_C}) \; .$$

In (5.30), the following notation is used. Consider any pair $a, b \in L^2((0,T); \mathbb{R}^{N_C}) \times \mathbb{R}^{N_x}$ given by $a = (a_1, a_2)$ and $b = (b_1, b_2)$, we define the inner product $(\cdot, \cdot)_G$ and the corresponding induced norm $||| \cdot |||$ is as follows

$$(a, b)_G := \sum_{n=1}^{N_C} \langle a_{1,n}, b_{1,n} \rangle_{L^2} + \langle a_2, b_2 \rangle \;\; , \;\; |||a||| := \sqrt{(a, a)_G} \; .$$

We remark that a solution $(\hat{x}, \hat{u}, \hat{p})$ of (5.30) with $G(\hat{x}, \hat{u}, \hat{p}) = 0$ is a root of $\mathcal{G}$, and hence a solution of the optimality system (5.28).

We address the forward equation in x and the backward equation in p as constraint equations in the minimization problem (5.30).

Existence and uniqueness of solutions $x, p \in H^1((0, T); \mathbb{R}^{N_x})$ of the constraint equations of (5.30) for any $T > 0$ and any initial and terminal condition, corresponding to a given $u \in L^2((0, T); \mathbb{R}^{N_C})$, can be proved as in Proposition 1. Hence, the solutions x and p are uniquely determined by the controls and the initial and terminal conditions, respectively. We have that $x = x(u, x_0)$ and $p = p(u, p_T)$. Consequently, we remark that the unknowns of (5.28) are the control $u \in L^2((0, T); \mathbb{R}^{N_C})$ and the terminal condition for the adjoint equation $p_T \in \mathbb{R}^{N_x}$.

In the following proposition, we state the existence of a solution of (5.30). Moreover, we analyze the relationship between the problems (5.26) and (5.30). In particular, the condition $G = 0$ is required to guarantee that a solution to (5.30) is a stationary point for (5.26).

Proposition 13. *A triple $(x, u, p) \in H^1((0, T); \mathbb{R}^{N_x}) \times L^2((0, T); \mathbb{R}^{N_C}) \times H^1((0, T); \mathbb{R}^{N_x})$, with $x = x(u, x_0)$ and $p = p(u, p_T)$, is a solution of (5.30) with $G(x, u, p) = 0$ if and only if it is a stationary point of (5.26).*

Proof. According to Assumptions 2 and 3, the control system (5.1) is solvable, hence problem (5.26) admits a solution that is a stationary point. Assume that (x, u, p) is a stationary point for (5.26), then it solves the optimality system (5.28). Hence, the constraint equations of (5.30) are satisfied, the norm of the gradient component is zero, that is $\sum_n \|u_n - \langle B_n x, p\rangle\|_{L^2}^2 = 0$, and the target point is reached. Consequently, (x, u, p) is a solution of (5.30) with $G = 0$.

On the other hand, assume that (x, u, p) solves (5.30) with $G = 0$. Then, the constraint and adjoint equations (as IVP) of (5.28) are satisfied. The fact that $G = 0$ implies that $\sum_n \|u_n - \langle B_n x, p\rangle\|_{L^2}^2 = 0$ and $\|x(T) - x_T\|_2^2 = 0$, which means that the gradient component of (5.28) is zero and the terminal condition $x(T) = x_T$ holds. Consequently, the triple (x, u, p) satisfies (5.28), hence it is a stationary point for (5.26). $\square$

We remark that a solution of (5.30) with $G = 0$ is only a stationary point for (5.26), hence it is not guaranteed that it is a minimum norm solution of (5.26).

Next, step (D) is discussed, and the optimality conditions used to characterize a solution to (5.30) are considered. To obtain the first-order optimality system, we follow the Lagrange multipliers approach. To this purpose, we denote by $y, q \in H^1((0, T); \mathbb{R}^{N_x})$ the Lagrange multipliers corresponding to x and p, respectively, and consider the Lagrange function corresponding to (5.30), that is given by

$$
L(x, u, p, y, q) = G(x, u, p) + \left\langle \dot{x} - \left[A + \sum_{n=1}^{N_C} u_n B_n\right] x, y \right\rangle_{L^2}
$$
$$
+ \left\langle -\dot{p} - \left[A + \sum_{n=1}^{N_C} u_n B_n\right]^* p, q \right\rangle_{L^2}.
$$

(5.31)

By means of (3.47), the optimality system for (5.30) is obtained in the following proposition.

Proposition 14. *The optimality system corresponding to* (5.30) *is given by*

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right] x \ , \quad x(0) = x_0 \ , \tag{5.32a}$$

$$-\dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]^* p \ , \quad p(T) = p_T \ , \tag{5.32b}$$

$$-\dot{y} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right]^* y + \sum_{n=1}^{N_C}\left[(u_n - \langle B_n x, p\rangle)B_n^* p\right] \ , \quad y(T) = -(x(T) - x_T) \ , \tag{5.32c}$$

$$\dot{q} = \left[A + \sum_{n=1}^{N_C} u_n B_n\right] q + \sum_{n=1}^{N_C}\left[(u_n - \langle B_n x, p\rangle)B_n x\right] \ , \quad q(0) = 0 \ , \tag{5.32d}$$

$$u_n - \langle B_n x, p\rangle - \langle B_n x, y\rangle - \langle B_n^* p, q\rangle = 0 \ , \quad n = 1, ..., N_C \ , \tag{5.32e}$$

where (5.32a) *and* (5.32b) *are the constraint equations,* (5.32c) *and* (5.32d) *are the corresponding adjoint equations, and* (5.32e) *gives the components of the gradient.*

Proof. Since $L(x, u, p, y, q)$ is linear with respect to the adjoint variables y and q, we obtain the constraint equations (5.32a) and (5.32b) as follows

$$\left\langle \nabla_y L(x, u, p, y, q), \delta y\right\rangle_{L^2} = \left\langle \dot{x} - \left[A + \sum_n u_n B_n\right]x, \delta y\right\rangle_{L^2} \ ,$$

and

$$\left\langle \nabla_q L(x, u, p, y, q), \delta q\right\rangle_{L^2} = \left\langle -\dot{p} - \left[A + \sum_n u_n B_n\right]^* p, \delta q\right\rangle_{L^2} \ .$$

For optimality, the two inner products $\left\langle \nabla_y L(x, u, p, y, q), \delta y\right\rangle_{L^2}$ and $\left\langle \nabla_q L(x, u, p, y, q), \delta q\right\rangle_{L^2}$ have to be equal to zero for all $\delta y \in L^2((0, T); \mathbb{R}^{N_x})$ and $\delta q \in L^2((0, T); \mathbb{R}^{N_x})$, respectively, thus (5.32a) and (5.32b) follow.

To obtain the adjoint equations (5.32c) and (5.32d), we consider the derivative with respect to x and p along the two directions δx and δp, respectively. We obtain (5.32c) as follows

$$\left\langle \nabla_x L(x, u, p, y, q), \delta x\right\rangle_{L^2} = \langle \delta x(T), x(T) - x_T\rangle$$
$$+ \int_0^T \langle \dot{\delta x} - \left[A + \sum_n u_n B_n\right]\delta x, y\rangle dt - \int_0^T \langle \sum_n (u_n - \langle B_n x, p\rangle)B_n \delta x, p\rangle dt$$
$$= \langle \delta x(T), x(T) - x_T\rangle + \left[\langle \delta x, y\rangle\right]_0^T$$
$$+ \int_0^T \langle -\dot{y} - \left[A + \sum_n u_n B_n\right]^* y - \sum_n (u_n - \langle B_n x, p\rangle)B_n^* p, \delta x\rangle dt$$
$$= \langle \delta x(T), x(T) - x_T\rangle + \left[\langle \delta x, y\rangle\right]_0^T$$
$$+ \left\langle -\dot{y} - \left[A + \sum_n u_n B_n\right]^* y - \sum_n (u_n - \langle B_n x, p\rangle)B_n^* p, \delta x\right\rangle_{L^2} \ .$$

Since the product $\left\langle \nabla_x L(x, u, p, y, q), \delta x\right\rangle_{L^2}$ has to be equal to zero for all $\delta x \in L^2((0, T); \mathbb{R}^{N_x})$, and we have that $\delta x(0) = 0$, we obtain the terminal condition $y(T) = -(x(T) - x_T)$ and the adjoint equation (5.32c).

To obtain the adjoint problem (5.32d), we proceed as follows

$$\left\langle \nabla_p L(x, u, p, y, q), \delta p \right\rangle_{L^2}$$

$$= \int_0^T \langle -\dot{\delta p} - \left[A + \sum_n u_n B_n \right]^* \delta p, q \rangle dt - \int_0^T \langle \sum_n (u_n - \langle B_n x, p \rangle) B_n x, \delta p \rangle dt$$

$$= -\left[\langle \delta p, q \rangle \right]_0^T + \int_0^T \langle \dot{q} - \left[A + \sum_n u_n B_n \right]^* q - \sum_n (u_n - \langle B_n x, p \rangle) B_n x, \delta p \rangle dt$$

$$= -\left[\langle \delta p, q \rangle \right]_0^T + \left\langle \dot{q} - \left[A + \sum_n u_n B_n \right] q - \sum_n (u_n - \langle B_n x, p \rangle) B_n x, \delta p \right\rangle_{L^2} .$$

The product $\left\langle \nabla_p L(x, u, p, y, q), \delta p \right\rangle_{L^2}$ has to be equal to zero for all $\delta p \in L^2((0, T); \mathbb{R}^{N_x})$ with $\delta p(T) = 0$. As a consequence, we have that $q(0) = 0$ and we obtain the adjoint equation (5.32d).

We derive the n-component of the gradient (5.32e) by means of the variation of the Lagrangian with respect to the control u_n as follows

$$\left\langle \nabla_{u_n} L(x, u, p, y, q), \delta u_n \right\rangle_{L^2} = \int_0^T (u_n - \langle B_n x, p \rangle) \delta u_n - \langle B_n x, y \rangle \delta u_n - \langle B_n^* p, q \rangle \delta u_n dt$$

$$= \left\langle u_n - \langle B_n x, p \rangle - \langle B_n x, y \rangle - \langle B_n^* p, q \rangle, \delta u_n \right\rangle_{L^2} .$$

Since this product has to be equal to zero for all $\delta u_n \in L^2(0, T)$, we obtain the optimality condition (5.32e). $\square$

In the following proposition, we discuss existence and uniqueness of solutions to the adjoint problems (5.32c) and (5.32d).

Proposition 15. *Given y_T and q_0, consider the following problems*

$$-\dot{y} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* y + \sum_{n=1}^{N_C} \left[(u_n - \langle B_n x, p \rangle) B_n^* p \right] \ , \quad y(T) = y_T \ , \tag{5.33}$$

and

$$\dot{q} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] q + \sum_{n=1}^{N_C} \left[(u_n - \langle B_n x, p \rangle) B_n x \right] \ , \quad q(0) = q_0 \ , \tag{5.34}$$

with $y, q, x, p \in H^1((0, T); \mathbb{R}^{N_x})$ and $u \in L^2((0, T); \mathbb{R}^{N_C})$. Then (5.33) and (5.34) admit unique solutions for any $T > 0$ and any y_T and q_0, respectively.

Moreover, assume that $\bigl(x(u, x_0), u, p(u, p_T)\bigr)$ is a stationary point for (5.26), then the problems (5.32c), which corresponds to problem (5.33) with $y_T = 0$, and (5.32d), which corresponds to (5.34) with $q_0 = 0$, admit the unique solutions $y(t) = 0$ and $q(t) = 0$, for all $t \in [0, T]$.

Proof. Existence and uniqueness of solution of (5.33) and (5.34) can be proved by means of known results; see, e.g., [107] and Proposition 1.

Next, consider problem (5.32d). Since $\bigl(x(u, x_0), u, p(u, p_T)\bigr)$ is a stationary point for (5.26), we have that $u_n - \langle B_n x, p \rangle = 0$, for $n = 1, ..., N_C$; hence, the forcing terms in the differential equations in (5.32c) and (5.32d) are zero. Consequently, since A and B_n are skew-symmetric, the dynamics are norm preserving, we have that (5.33) with $y_T = 0$ and (5.34) with $q_0 = 0$ admit the unique solutions $y = 0$ and $q = 0$. $\square$

Next, we address step (E), and discuss the reduced form of problem (5.30), which is suitable to be solved by means of appropriate numerical optimization methods. As mentioned above, the solutions of the constraint equations (5.32a) and (5.32b) are uniquely determined by the initial and terminal conditions, that are $x(0) = x_0$ and $p(T) = p_T$, respectively, and by the control vector function u. We have

$$x = x(u) \quad \text{and} \quad p = p(u, p_T) \,, \tag{5.35}$$

where the dependence of x from x_0 is omitted because it is an input of the problem. Consequently, problem (5.30) can be equivalently expressed in the following reduced form

$$\min_{u, p_T} \quad G_r(u, p_T) := G(x(u), u, p(u, p_T))$$

$$\text{s.t.} \quad (x(u), p(u, p_T)) \in \mathcal{S}_{ad} := \Big\{ (x, p) \mid x \text{ solves (3.53a) and } p \text{ solves (3.53b)} \Big\} \,. \tag{5.36}$$

We characterize a solution of (5.36) with the first-order optimality conditions given in the following result, which follows directly from Proposition 14.

Proposition 16. *The optimality system corresponding to problem* (5.36) *is given by*

$$\nabla_{u_n} G_r(u, p_T) := u_n - \langle B_n x, p \rangle - \langle B_n x, y \rangle - \langle B_n^* p, q \rangle = 0 \ , \ n = 1, ..., N_C \,, \tag{5.37a}$$

$$\nabla_{p_T} G_r(u, p_T) := -q(T) = 0 \,, \tag{5.37b}$$

where x*,* p*,* y *and* q *solve the following problems*

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \ , \quad x(0) = x_0 \,, \tag{5.37c}$$

$$- \dot{p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* p \ , \quad p(T) = p_T \,, \tag{5.37d}$$

$$- \dot{y} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* y + \sum_{n=1}^{N_C} \Big[(u_n - \langle B_n x, p \rangle) B_n^* p \Big] \ , \quad y(T) = -(x(T) - x_T) \,, \tag{5.37e}$$

$$\dot{q} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] q + \sum_{n=1}^{N_C} \Big[(u_n - \langle B_n x, p \rangle) B_n x \Big] \ , \quad q(0) = 0 \,. \tag{5.37f}$$

Proof. Consider Proposition 14 and its proof. We remark that, the gradient component of the reduced problem with respect to p_T is obtained from the fact that

$$\langle \nabla_{p_T} G_r(u, p_T), \delta q(T) \rangle = \langle -q(T), \delta q(T) \rangle = 0 \,, \tag{5.38}$$

for all $\delta q(T)$. Notice that, unlike in (5.30), in (5.36) p_T is not fixed, hence $\delta q(T)$ is not fixed to 0.

$\qquad\square$

Notice that in step (E), we write problem (5.30) in the reduced form, that is (5.36), and derive its first-order optimality system in Proposition 16. In this settings, it is possible to perform step (F) and address problem (5.30) by using optimal control techniques; see, e.g., [16].

In the next section, we discuss regularity properties of (5.30). These are necessary to guarantee a correct behaviour of optimal control solvers. In particular, we discuss regularity of the Hessian operator corresponding to (5.30), that allows us to obtain second order necessary and sufficient optimality conditions and to characterize the corresponding optimal control solutions.

5.3.1 Regularity properties

In this section, we investigate the reduced Hessian operator corresponding to (5.36) and its regularity properties. For this purpose, we first discuss the Hessian of problem (5.30), then we consider its reduced form corresponding to (5.36). In particular, we focus on its action on a given vector function. This aspect will be crucial in the development of a Krylov-Newton method as discussed in the next section.

By computing the second directional derivative of the Lagrange function (5.31), we write that

$$
\left\langle H(x,u,p) \begin{pmatrix} \delta x \\ \delta u \\ \delta p \\ \delta y \\ \delta q \end{pmatrix} , \begin{pmatrix} \delta x \\ \delta u \\ \delta p \\ \delta y \\ \delta q \end{pmatrix} \right\rangle_{L^2} = \left\langle \begin{pmatrix} H_x \\ H_u \\ H_p \\ H_y \\ H_q \end{pmatrix} , \begin{pmatrix} \delta x \\ \delta u \\ \delta p \\ \delta y \\ \delta q \end{pmatrix} \right\rangle_{L^2} , \tag{5.39}
$$

where H_x, H_u, H_p, H_y and H_q denote the following

$$
\begin{aligned}
H_x = {}& - \dot{\delta y} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* \delta y - \left[\sum_{n=1}^{N_C} \delta u_n B_n \right]^* y - \sum_{n=1}^{N_C}(u_n - \langle B_n x, p \rangle) B_n^* \delta p \\
& - \sum_{n=1}^{N_C}(\delta u_n - \langle B_n \delta x, p \rangle - \langle B_n x, \delta p \rangle) B_n^* p , \quad \text{with} \quad \delta y(T) = -\delta x(T) ,
\end{aligned} \tag{5.40}
$$

$$
\begin{aligned}
H_{u_n} = {}& \delta u_n - \langle B_n \delta x, p \rangle - \langle B_n x, \delta p \rangle - \langle B_n \delta x, y \rangle \\
& - \langle B_n x, \delta y \rangle - \langle B_n^* \delta p, q \rangle - \langle B_n^* p, \delta q \rangle ,
\end{aligned} \tag{5.41}
$$

$$
\begin{aligned}
H_p = {}& \dot{\delta q} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta q - \left[\sum_{n=1}^{N_C} \delta u_n B_n \right] q - \sum_{n=1}^{N_C}(u_n - \langle B_n x, p \rangle) B_n \delta x \\
& - \sum_{n=1}^{N_C}(\delta u_n - \langle B_n \delta x, p \rangle - \langle B_n x, \delta p \rangle) B_n x , \quad \text{with} \quad \delta q(0) = 0 ,
\end{aligned} \tag{5.42}
$$

$$
H_y = \dot{\delta x} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta x - \left[\sum_{n=1}^{N_C} \delta u_n B_n \right] x , \quad \text{with} \quad \delta x(0) = 0 , \tag{5.43}
$$

and

$$
H_q = -\dot{\delta p} - \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* \delta p - \left[\sum_{n=1}^{N_C} \delta u_n B_n \right]^* p , \quad \text{with} \quad \delta p(T) = \delta p_T . \tag{5.44}
$$

Notice that H_x, H_u, H_p, H_y and H_q represent the residuals of the linearized optimality system.

Now, we consider the reduced problem (5.36) and we denote by $\nabla^2 G_r(u, p_T)$ the corresponding reduced Hessian operator. We recall that the unknowns are the control u and the terminal condition p_T. Consequently, the action of $\nabla^2 G_r(u, p_T)$ on a vector $(\delta u, \delta p_T)^T \in L^2((0,T); \mathbb{R}^{N_C}) \times \mathbb{R}^N$ is given by the following

$$
\nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u_1 \\ \vdots \\ \delta u_{N_C} \\ \delta p_T \end{pmatrix} = \begin{pmatrix} H_{u_1}(x,u,p) \\ \vdots \\ H_{u_{N_C}}(x,u,p) \\ H_{p_T}(x,u,p) \end{pmatrix} , \tag{5.45}
$$

where δx, δp, δy and δq are solutions obtained by cancelling (5.43), (5.44), (5.40) and (5.42), respectively, and $H_{p_T}(x, u, p) = -\delta q(T)$. Hence, the action of the reduced Hessian operator can be obtained by solving the linearized equations (5.40) and (5.42)-(5.44) and the assembling (5.45).

With the following theorem, we prove regularity of the reduced Hessian operator.

Theorem 29. *Let (u, p_T) be a solution of (5.36) with $G_r(u, p_T) = 0$, then the reduced Hessian operator $\nabla^2 G_r(u, p_T)$ is positive semi-definite, and in particular we have that*

$$\left\langle \nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix}, \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix} \right\rangle_{L^2} = \|\delta x(T)\|_2^2 + \sum_{n=1}^{N_C} \|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2}^2 \, . \tag{5.46}$$

Proof. We denote by $x = x(u)$ and $p = p(u, p_T)$ the unique solutions of the constraint equations (5.37c) and (5.37d), respectively, and with $y = y(x, u, p)$ and $q = q(x, u, p)$, the unique solutions of the adjoint equations (5.37e) and (5.37f), respectively. We prove the claim in two steps.

Step 1: since (u, p_T) is a solution of (5.36) with $G_r(u, p_T) = 0$, then we have that $u_n - \langle B_n x, p \rangle = 0$ for $n = 1, ..., N_C$. Moreover, by Proposition 15, we know that $y = 0$ and $q = 0$. Consequently, the linearized adjoint equations $H_x = 0$ and $H_p = 0$ become as follows

$$-\dot{\delta y} = \left[A + \sum_n u_n B_n\right]^* \delta y + \sum_n \left(\delta u_n - \langle B_n \delta x, p \rangle - \langle B_n x, \delta p \rangle\right) B_n^* p \, , \tag{5.47}$$

with $\delta y(T) = -\delta x(T)$, and

$$\dot{\delta q} = \left[A + \sum_n u_n B_n\right] \delta q + \sum_n \left(\delta u_n - \langle B_n \delta x, p \rangle - \langle B_n x, \delta p \rangle\right) B_n x \, , \tag{5.48}$$

with $\delta q(0) = 0$. Now, similarly as in (3.39), we define $\mathcal{O}(u) : H^1((0, T); \mathbb{R}^{N_x}) \to L^2((0, T); \mathbb{R}^{N_x})$

$$\mathcal{O}(u) := \frac{d}{dt} - \left[A + \sum_n u_n B_n\right] , \tag{5.49}$$

whose adjoint (up to some boundary condition) is given by

$$\mathcal{O}(u)^* = -\frac{d}{dt} - \left[A + \sum_n u_n B_n\right]^* . \tag{5.50}$$

Notice that, by solving the equations $H_y = 0$ and $H_p = 0$, we have

$$\mathcal{O}(u)(\delta x + \delta q) = \dot{\delta x} + \dot{\delta q} - \left[A + \sum_n u_n B_n\right]\left(\delta x + \delta q\right)$$

$$= \sum_n \delta u_n B_n x + \sum_{n=1}^{N_C}\left(\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\right) B_n x \tag{5.51}$$

$$= \sum_n \left(2\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\right) B_n x \, ,$$

and analogously, solving $H_x = 0$ and $H_q = 0$, we have

$$\mathcal{O}(u)^*(\delta p + \delta y) = -\dot{\delta p} - \dot{\delta y} - \left[A + \sum_n u_n B_n\right]^*\left(\delta p + \delta y\right)$$

$$= \sum_n \left(2\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\right) B_n^* p \, . \tag{5.52}$$

Step 2: using (5.45) and (5.41) and the fact that $y = 0$ and $q = 0$, we have

$$
\left\langle \nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix}, \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix} \right\rangle_{L^2} =
$$

$$
= \left\langle \begin{pmatrix} \delta u_1 - \langle B_1 \delta x, p \rangle - \langle B_1 x, \delta p \rangle - \langle B_1 x, \delta y \rangle - \langle B_1^* p, \delta q \rangle \\ \vdots \\ \delta u_{N_C} - \langle B_{N_C} \delta x, p \rangle - \langle B_{N_C} x, \delta p \rangle - \langle B_{N_C} x, \delta y \rangle - \langle B_{N_C}^* p, \delta q \rangle \\ -\delta q(T) \end{pmatrix}, \begin{pmatrix} \delta u_1 \\ \vdots \\ \delta u_{N_C} \\ \delta p_T \end{pmatrix} \right\rangle_{L^2}
$$

$$
= \left\langle \begin{pmatrix} \delta u_1 - \langle B_1 x, \delta p + \delta y \rangle - \langle B_1^* p, \delta x + \delta q \rangle \\ \vdots \\ \delta u_{N_C} - \langle B_{N_C} x, \delta p + \delta y \rangle - \langle B_{N_C}^* p, \delta x + \delta q \rangle \\ -\delta q(T) \end{pmatrix}, \begin{pmatrix} \delta u_1 \\ \vdots \\ \delta u_{N_C} \\ \delta p_T \end{pmatrix} \right\rangle_{L^2}
$$

$$
= -\langle \delta q(T), \delta p_T \rangle + \sum_n \int_0^T \delta u_n^2 dt - \sum_n \int_0^T \left(\langle B_n x, \delta p + \delta y \rangle + \langle B_n^* p, \delta x + \delta q \rangle \right) \delta u_n dt
$$

$$
= -\langle \delta q(T), \delta p_T \rangle + \sum_n \int_0^T \delta u_n^2 dt - \int_0^T \left(\langle \sum_n \delta u_n B_n x, \delta p + \delta y \rangle + \langle \sum_n \delta u_n B_n^* p, \delta x + \delta q \rangle \right) dt
$$

$$
= -\langle \delta q(T), \delta p_T \rangle + \sum_n \int_0^T \delta u_n^2 dt - \int_0^T \left(\langle \mathcal{O}(u)(\delta x), \delta p + \delta y \rangle + \langle \mathcal{O}(u)^*(\delta p), \delta x + \delta q \rangle \right) dt ,
$$

$$(5.53)$$

the latter equation follows from solving $H_y = 0$, $H_q = 0$ and (5.49) and (5.50). Now, integrating by parts, we obtain

$$
\left\langle \nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix}, \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix} \right\rangle_{L^2} =
$$

$$
= -\langle \delta q(T), \delta p_T \rangle + \sum_n \int_0^T \delta u_n^2 dt - \left[\langle \delta x, \delta p + \delta y \rangle \right]_0^T - \int_0^T \langle \delta x, \mathcal{O}(u)^*(\delta p + \delta y) \rangle dt
$$

$$
+ \left[\langle \delta p, \delta x + \delta q \rangle \right]_0^T - \int_0^T \langle \delta p, \mathcal{O}(u)(\delta x + \delta q) \rangle dt
$$

$$
= -\langle \delta q(T), \delta p_T \rangle + \sum_n \int_0^T \delta u_n^2 dt - \langle \delta x(T), \delta p_T + \delta y(T) \rangle - \int_0^T \langle \delta x, \mathcal{O}(u)^*(\delta p + \delta y) \rangle dt
$$

$$
+ \langle \delta p_T, \delta x(T) + \delta q(T) \rangle - \int_0^T \langle \delta p, \mathcal{O}(u)(\delta x + \delta q) \rangle dt
$$

$$
= -\langle \delta x(T), \delta y(T) \rangle + \sum_n \int_0^T \delta u_n^2 dt - \int_0^T \langle \delta x, \mathcal{O}(u)^*(\delta p + \delta y) \rangle dt - \int_0^T \langle \delta p, \mathcal{O}(u)(\delta x + \delta q) \rangle dt
$$

$$
= \|\delta x(T)\|_2^2 + \sum_n \int_0^T \delta u_n^2 dt - \int_0^T \langle \delta x, \sum_n \left(2\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle \right) B_n^* p \rangle dt
$$

$$
- \int_0^T \langle \delta p, \sum_n \left(2\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle \right) B_n x \rangle dt
$$

$$(5.54)$$

where we use (5.51) and (5.52) and the fact that $\delta y(T) = -\delta x(T)$. We have the following

$$
\left\langle \nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix}, \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix} \right\rangle_{L^2} =
$$

$$
= \|\delta x(T)\|_2^2 + \sum_n \int_0^T \Big[\delta u_n^2 - 2\delta u_n \langle \delta x, B_n^* p \rangle + \Big(\langle \delta x, B_n^* p \rangle + \langle x, B_n^* \delta p \rangle \Big) \langle \delta x, B_n^* p \rangle
$$

$$
- 2\delta u_n \langle \delta p, B_n x \rangle + \Big(\langle \delta x, B_n^* p \rangle + \langle x, B_n^* \delta p \rangle \Big) \langle \delta p, B_n x \rangle \Big] dt
$$

$$
= \|\delta x(T)\|_2^2 + \sum_n \int_0^T \Big(\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle \Big)^2 dt \, ,
$$

$$(5.55)$$

which implies that

$$
\left\langle \nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix}, \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix} \right\rangle_{L^2} = \|\delta x(T)\|_2^2 + \sum_n \|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2}^2 \, .
$$

$$(5.56)$$

Consequently, we have

$$
\left\langle \nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix}, \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix} \right\rangle_{L^2} \geq 0 \, , \quad \forall (\delta u, \delta p_T) \, . \tag{5.57}
$$

$\square$

Next, coercivity of the reduced Hessian operator (5.45) is discussed. According to Theorem 29, the Hessian in (5.45) is positive-semidefinite for all the pair $(\delta u, \delta p_T)$. To improve this result, we characterize in Corollary 4 the set of all points in which (5.45) is indefinite and we discuss the relationship between (5.45) and the end-point map $\delta u \mapsto \delta x(T; \delta u)$. Next, we provide sufficient conditions for the coercivity of the reduced Hessian operator (5.45), see Theorem 30 and Corollary 5. Moreover, the coercivity of the reduced Hessian operator (5.45) allows us to characterize solutions to the minimum norm exact-control problem (5.26) as isolated points. This property is shown in Corollary 7.

Corollary 4. *Consider the assumptions of Theorem 29. Then, we have that*

$$
\left\langle \nabla^2 G_r(u, p_T) \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix}, \begin{pmatrix} \delta u \\ \delta p_T \end{pmatrix} \right\rangle_{L^2} = 0 \, , \tag{5.58}
$$

for all $(\delta u, \delta p_T)$ belonging to a convex neighbourhood of $(0,0)$. Moreover, if the map $\delta u \mapsto \delta x(T; \delta u)$ is injective in a neighbourhood $\mathcal{N}$ of $\delta u = 0$, then $\nabla^2 G_r(u, p_T)$ is positive definite in $\mathcal{N}$.

Proof. To prove the first claim, we recall (5.46) and consider the following optimization

problem

$$\min_{\delta u, \delta p_T} \quad F(\delta u, \delta p_T) := \|\delta x(T)\|_2^2 + \sum_{n=1}^{N_C} \|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2}^2$$

$$\text{s.t.} \quad \dot{\delta x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] \delta x + \left[\sum_{n=1}^{N_C} \delta u_n B_n \right] x \ , \ \delta x(0) = 0$$

$$- \dot{\delta p} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right]^* \delta p + \left[\sum_{n=1}^{N_C} \delta u_n B_n \right]^* p \ , \ \delta p(T) = \delta p_T \tag{5.59}$$

$$(\delta u, \delta p_T) \in S \subset L^2((0,T); \mathbb{R}^{N_x}) \times \mathbb{R}^{N_C} \ ,$$

where $(x(u), u, p(u, p_T))$ is a solution of (5.36) with $G_r(u, p_T) = 0$ and S is closed, convex and bounded subset of $L^2((0,T); \mathbb{R}^{N_x}) \times \mathbb{R}^{N_C}$. The existence of a solution of (5.59) follows from the fact that $F(\delta u, \delta p_T) \geq 0$ and $F(0, 0) = 0$. Hence $(\delta u, \delta p_T) = (0, 0)$ is a global minimum of (5.59).

Now, notice that the constraint differential equations in (5.59) are linear, and the maps $(\delta x, \delta u_n, \delta p) \mapsto \delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle$ and $\delta u \mapsto \delta x(T; \delta u)$ preserve convex combinations. Hence the convexity of the norms implies that F is convex. Since, S is a convex set and F a convex function, then the set of global minima of F is convex. Consequently, we obtain that $F(\delta u, \delta p_T) = 0$ for all $(\delta u, \delta p_T)$ belonging to a convex neighbourhood of $(0, 0)$ included in S; see, e.g., [66] (Theorem 2.14).

To prove the second argument, we consider the following. If $\delta u \mapsto \delta x(T; \delta u)$ is injective in a neighbourhood $\mathcal{N}$ of $\delta u = 0$, then in $\mathcal{N}$ we have that $\delta u = 0$ if and only if $\|\delta x(T; \delta u)\|_2 = 0$. Consequently, the positive definiteness of (5.45) follows. $\qquad \square$

Before to state and prove our main result regarding coercivity of the reduce Hessian operator, we need to consider the two following lemmas.

Lemma 9. *Let $(\tilde{u}, \tilde{p}_T)$ be a solution of (5.36) with $G_r(\tilde{u}, \tilde{p}_T) = 0$. If $\tilde{p}_T = 0$, then $\tilde{u} = 0$, that is $(\tilde{u}, \tilde{p}_T)$ is a trivial solution of (5.36).*

Proof. Assuming that $\tilde{p}_T = 0$ and recalling that (3.53b) is norm preserving, we get that $\tilde{p}(t; \tilde{p}_T) = 0$ a.e. on $(0, T)$. Since $(\tilde{u}, \tilde{p}_T)$ be a solution of (5.36) with $G_r(\tilde{u}, \tilde{p}_T) = 0$, we have that $\tilde{u}_n = \langle B_n x, \tilde{p} \rangle$ for $n = 1, ..., N_C$. Consequently, we obtain that $\tilde{u} = 0$. $\qquad \square$

Lemma 10. *Let (u, p_T) be a solution of (5.36) with $G_r(u, p_T) = 0$. Let δx and δp be the unique solutions of $H_y = 0$ and $H_q = 0$, respectively. Then the following estimates hold*

$$\|\delta x\|_{L^2} \leq 2T \sqrt{N_C} K_{00} \|x_0\|_2 \|\delta u\|_{L^2} \ , \tag{5.60}$$

and

$$\|\delta p\|_{L^2} \leq 2T \sqrt{N_C} K_{00} \|p_T\|_2 \|\delta u\|_{L^2} + \sqrt{T} \|\delta p_T\|_2 \ . \tag{5.61}$$

where $K_{00} = \sum_{n=1}^{N_C} \|B_n\|_{\mathcal{L}}$.

Proof. The claim follows straightforwardly from Proposition 5 and Proposition 10. Notice that we use Lemma 9 to ensure that $\|p_T\|_2 \neq 0$ in Proposition 10. $\qquad \square$

Theorem 30. *Let (u, p_T) be a solution of (5.36) with $G_r(u, p_T) = 0$. Let $K_{00} = \sum_{n=1}^{N_C} \|B_n\|_{\mathcal{L}}$ and*

$$\tilde{K}_n := \|B_n\|_{\mathcal{L}} \|x_0\|_2 \sqrt{T} \ , \tag{5.62}$$

and

$$K_n := 1 - 4\sqrt{N_C}K_{00}T\|B_n\|_{\mathcal{L}}\|p_T\|_2\|x_0\|_2 \,, \tag{5.63}$$

and assume that T is large enough, in the sense that there exists a positive constant ϵ of $O(1/T^{2m})$ with $m > 2$ such that the following holds

$$\|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2} \geq \sqrt{\epsilon}\big|K_n\|\delta u\|_{L^2} - \tilde{K}_n\|\delta p_T\|_2\big| \,. \tag{5.64}$$

Assume also that

$$C_{1,n} := 1 + \tilde{K}_n\Big(16TN_C K_{00}^2\|p_T\|_2^2\tilde{K}_n - 8\sqrt{TN_C}K_{00}\|p_T\|_2 - 1 + 4\sqrt{TN_C}K_{00}\|p_T\|_2\tilde{K}_n\Big) > 0 \,, \tag{5.65}$$

and

$$C_{2,n} := \tilde{K}_n + 4\sqrt{TN_C}K_{00}\tilde{K}_n\|p_T\|_2 - 1 > 0 \,, \tag{5.66}$$

for $n = 1, ..., N_C$. Then, the reduced Hessian operator $\nabla^2 G_r(u, p_T)$ is coercive as follows

$$\left\langle \nabla^2 G_r(u, p_T)\begin{pmatrix}\delta u \\ \delta p_T\end{pmatrix}, \begin{pmatrix}\delta u \\ \delta p_T\end{pmatrix}\right\rangle_{L^2} \geq \|\delta x(T)\|_2^2 + \epsilon\alpha\Big(\|\delta u\|_{L^2}^2 + \|\delta p_T\|_2^2\Big) \,, \quad \forall(\delta u, \delta p_T) \neq 0 \,, \tag{5.67}$$

where $\alpha > 0$ is given by

$$\alpha := \min_n\big\{(K_n^2 - K_n\tilde{K}_n) \,, \, (\tilde{K}_n^2 - K_n\tilde{K}_n)\big\} \,. \tag{5.68}$$

Moreover, $\nabla^2 G_r(u, p_T)$ is invertible in a neighbourhood of (u, p_T).

Proof. Consider the norm $\|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2}$ which appears in (5.56). We have that

$$\|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2} \geq \|\delta u_n\|_{L^2} - \|\langle \delta x, B_n^* p \rangle\|_{L^2} - \|\langle B_n x, \delta p \rangle\|_{L^2} \,. \tag{5.69}$$

Now, recalling that (5.37c) and (5.37d) are norm preserving and using the estimates (5.60) and (5.61), we obtain

$$\|\langle \delta x, B_n^* p \rangle\|_{L^2} \leq \|B_n\|_{\mathcal{L}}\|p_T\|_2 2T\sqrt{N_C}K_{00}\|x_0\|_2\|\delta u\|_{L^2} \,, \tag{5.70}$$

and

$$\|\langle B_n x, \delta p \rangle\|_{L^2} \leq \|B_n\|_{\mathcal{L}}\|p_T\|_2 2T\sqrt{N_C}K_{00}\|x_0\|_2\|\delta u\|_{L^2} + \|B_n\|_{\mathcal{L}}\sqrt{T}\|x_0\|_2\|\delta p_T\|_2 \,. \tag{5.71}$$

Replacing (5.70) and (5.71) in (5.69), we have

$$\|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2} \geq K_n\|\delta u\|_{L^2} - \tilde{K}_n\|\delta p_T\|_2 \,. \tag{5.72}$$

Moreover, for some positive ϵ of $O(1/T^{2m})$ with $m > 2$ it holds that

$$\|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2} \geq \sqrt{\epsilon}\big|K_n\|\delta u\|_{L^2} - \tilde{K}_n\|\delta p_T\|_2\big| \,. \tag{5.73}$$

Taking the square and using the Cauchy inequality, we obtain

$$\|\delta u_n - \langle \delta x, B_n^* p \rangle - \langle x, B_n^* \delta p \rangle\|_{L^2}^2 \geq \epsilon\Big(K_n^2\|\delta u\|_{L^2}^2 + \tilde{K}_n^2\|\delta p_T\|_2^2 - 2K_n\tilde{K}_n\|\delta u\|_{L^2}\|\delta p_T\|_2\Big)$$
$$\geq \epsilon\Big(K_n^2\|\delta u\|_{L^2}^2 + \tilde{K}_n^2\|\delta p_T\|_2^2 - K_n\tilde{K}_n\big(\|\delta u\|_{L^2}^2 + \|\delta p_T\|_2^2\big)\Big) \,. \tag{5.74}$$

Now, we take the sum over n and we look for a positive α such that the following holds

$$\sum_n \|\delta u_n - \langle \delta x, B_n^* p\rangle - \langle x, B_n^* \delta p\rangle\|_{L^2}^2$$
$$\geq \sum_n \epsilon\Big[K_n^2\|\delta u\|_{L^2}^2 + \tilde{K}_n^2\|\delta p_T\|_2^2 - K_n\tilde{K}_n\big(\|\delta u\|_{L^2}^2 + \|\delta p_T\|_2^2\big)\Big]$$
$$= \epsilon\Big[\sum_n\big(K_n^2 - K_n\tilde{K}_n\big)\Big]\|\delta u\|_{L^2}^2 + \Big[\sum_n\big(\tilde{K}_n^2 - K_n\tilde{K}_n\big)\Big]\|\delta p_T\|_2^2$$
$$\geq \epsilon\alpha\Big(\|\delta u\|_{L^2}^2 + \|\delta p_T\|_2^2\Big) . \tag{5.75}$$

We consider α defined in (5.68) and we notice that K_n, defined in (5.63), can be written as follows

$$K_n = 1 - 4T\sqrt{N_C}K_{00}\|B_n\|_{\mathcal{L}}\|x_0\|_2\|p_T\|_2 = 1 - 4\sqrt{TN_C}K_{00}\tilde{K}_n\|p_T\|_2 . \tag{5.76}$$

To guarantee the positivity of α, we have to require that $(K_n^2 - K_n\tilde{K}_n) > 0$ and $(\tilde{K}_n^2 - K_n\tilde{K}_n) > 0$. From these requirements, we derive the conditions (5.65) and (5.66) as follows

$$K_n^2 - K_n\tilde{K}_n = \big(1 - 4\sqrt{TN_C}K_{00}\tilde{K}_n\|p_T\|_2\big)^2 - \big(1 - 4\sqrt{TN_C}K_{00}\tilde{K}_n\|p_T\|_2\big)\tilde{K}_n$$
$$= 1 + 16TN_C K_{00}^2\|p_T\|_2^2\tilde{K}_n^2 - 8\sqrt{TN_C}K_{00}\|p_T\|_2\tilde{K}_n$$
$$- \tilde{K}_n + 4\sqrt{TN_C}K_{00}\|p_T\|_2\tilde{K}_n^2 > 0 , \tag{5.77}$$

and we obtain that

$$1 + \tilde{K}_n\Big(16TN_C K_{00}^2\|p_T\|_2^2\tilde{K}_n - 8\sqrt{T}\sqrt{N_C}K_{00}\|p_T\|_2$$
$$- 1 + 4\sqrt{T}\sqrt{N_C}K_{00}\|p_T\|_2\tilde{K}_n\Big) > 0 , \tag{5.78}$$

that is (5.65), and

$$\tilde{K}_n^2 - K_n\tilde{K}_n > 0 \Rightarrow \tilde{K}_n\big(\tilde{K}_n - K_n\big) > 0 \Rightarrow \big(\tilde{K}_n - K_n\big) > 0 , \tag{5.79}$$

which implies that

$$\tilde{K}_n + 4\sqrt{TN_C}K_{00}\tilde{K}_n\|p_T\|_2 - 1 > 0 , \tag{5.80}$$

that is (5.66). Finally, (5.56) and (5.75) imply that

$$\left\langle \nabla^2 G_r(u, p_T)\begin{pmatrix}\delta u \\ \delta p_T\end{pmatrix}, \begin{pmatrix}\delta u \\ \delta p_T\end{pmatrix}\right\rangle_{L^2} \geq \|\delta x(T)\|_2^2 + \sum_n \|\delta u_n - \langle \delta x, B_n^* p\rangle - \langle x, B_n^*\delta p\rangle\|_{L^2}^2$$
$$\geq \|\delta x(T)\|_2^2 + \epsilon\alpha\big(\|\delta u\|_{L^2}^2 + \|\delta p_T\|_2^2\big) , \ \forall(\delta u, \delta p_T) \neq 0 , \tag{5.81}$$

which implies that $\nabla^2 G_r$ is invertible in (u, p_T). Continuity of the map $(u, p_T) \mapsto \nabla^2 G_r(u, p_T)$ enables to conclude that $\nabla^2 G_r$ is invertible in a neighbourhood of (u, p_T). $\quad\square$

In the next corollary, we give a sufficient condition for the two assumptions (5.65) and (5.66) in Theorem 29 to hold.

Corollary 5. *Let (u, p_T) be a solution of (5.36) with $G_r(u, p_T) = 0$. Let K_{00}, K_n and $\tilde{K}_n$ be defined as in Theorem 29. Assume that*

$$C_{12,n} := 4\sqrt{TN_C}K_{00}\tilde{K}_n\|p_T\|_2 - 1 > 0 \tag{5.82}$$

for $n = 1, ..., N_C$. Then the conditions (5.65) and (5.66) are satisfied, hence the reduced Hessian operator $\nabla^2 G_r(u, p_T)$ is coercive with α given by (5.68).

Proof. Condition (5.66) follows immediately from (5.82) and the positivity of $\tilde{K}_n$.

Next, we show that (5.82) implies also (5.65). For this purpose, we write (5.65) as follows

$$16TM^2\tilde{K}_n^2\|p_T\|_2^2 + \left(4\sqrt{TN_C}K_{00}\tilde{K}_n^2 - 6\sqrt{TN_C}K_{00}\tilde{K}_n\right)\|p_T\|_2 + (1 - \tilde{K}_n) > 0 . \quad (5.83)$$

The discriminant of the previous quadratic inequality is

$$\begin{aligned}
\Delta &= \left(4\sqrt{T}\sqrt{N_C}K_{00}\tilde{K}_n^2 - 8\sqrt{T}\sqrt{N_C}K_{00}\tilde{K}_n\right)^2 \\
&\quad - 64TN_C K_{00}^2\tilde{K}_n^2(1 - \tilde{K}_n) = 16TN_C K_{00}^2\tilde{K}_n^4 > 0 .
\end{aligned} \quad (5.84)$$

Consequently, (5.83) is fulfilled if the following holds

$$\|p_T\|_2 > \frac{1}{2\sqrt{TN_C}K_{00}\tilde{K}_n} , \quad (5.85)$$

which is equivalent to (5.82). $\qquad\square$

We remark that, condition (5.82) is in agreement with Theorem 23 and Corollary 2: replacing $\tilde{K}_n$ in $C_{12,n}$ we obtain that

$$C_{12,n} = 4T\sqrt{N_C}K_{00}\|B_n\|_{\mathcal{L}}\|x_0\|_2\|p_T\|_2 - 1 ,$$

from which is clear that a "sufficiently large" T contributes to the fulfillment of (5.82).

The next corollary, which follows directly from Theorem 30, provides a relaxation on the conditions (5.65) and (5.66). The proof is similar to the one of Theorem 30, hence we omit it for brevity.

Corollary 6. *Let the assumptions of Theorem 29 hold, and assume the following*

$$C_3 := \sum_{n=1}^{N_C}\left(K_n^2 - K_n\tilde{K}_n\right) > 0 \quad and \quad C_4 := \sum_{n=1}^{N_C}\left(\tilde{K}_n^2 - K_n\tilde{K}_n\right) > 0 . \quad (5.86)$$

Then the reduced Hessian operator $\nabla^2 G_r(u, p_T)$ is coercive with

$$\alpha := \min\left\{\sum_{n=1}^{N_C}\left(K_n^2 - K_n\tilde{K}_n\right) , \sum_{n=1}^{N_C}\left(\tilde{K}_n^2 - K_n\tilde{K}_n\right)\right\} .$$

We remark that, if Theorem 30, Corollary 5 and Corollary 6 hold, then (u, p_T) such that $G_r(u, p_T) = 0$ is an isolated global minimum in a ball of finite radius centered in (u, p_T). This fact is expressed by the following corollary.

Corollary 7. *Let (u, p_T) be a solution of (5.36) with $G_r(u, p_T) = 0$. Let the assumptions of Theorem 30 hold, and assume that G_r is twice Fréchet differentiable in (u, p_T). Then, there exists a positive constant $\gamma > 0$ such that*

$$G_r(\hat{u}, \hat{p}_T) \geq G_r(u, p_T) + \gamma\left(\|\hat{u} - u\|_{L^2}^2 + \|\hat{p}_T - p_T\|_2^2\right) , \quad (5.87)$$

for all $(\hat{u}, \hat{p}_T)$ belonging to a ball centered in (u, p_T). In particular, inequality (5.87) ensure that global minima of (5.36) are isolated points.

Proof. By using that G is twice Fréchet differentiable, we have the following

$$
\begin{aligned}
G_r(\hat{u}, \hat{p}_T) - G_r(u, p_T) = {} & \big(\nabla G_r(u, p_T), (\delta u, \delta p_T)\big)_G \\
& + \frac{1}{2}\big(\nabla^2 G_r(u, p_T)(\delta u, \delta p_T), (\delta u, \delta p_T)\big)_G + r(\delta u, \delta p_T) ,
\end{aligned}
\tag{5.88}
$$

where $|r(\delta u, \delta p_T)| = o(|||(\delta u, \delta p_T)|||^2)$. Next, we notice that in a solution (u, p_T) it holds that $\nabla G_r(u, p_T) = 0$, and we write

$$
G_r(\hat{u}, \hat{p}_T) - G_r(u, p_T) \geq \frac{1}{2}\big(\nabla^2 G_r(u, p_T)(\delta u, \delta p_T), (\delta u, \delta p_T)\big)_G - |r(\delta u, \delta p_T)| .
\tag{5.89}
$$

Now, there exists a constant $\rho > 0$ such that $|r(\delta u, \delta p_T)| \leq \frac{\alpha}{4}$ for all $(\delta u, \delta p_T)$ such that $|||(\delta u, \delta p_T)||| \leq \rho$. This implies the following

$$
G_r(\hat{u}, \hat{p}_T) - G_r(u, p_T) \geq \frac{\alpha}{4}|||(\delta u, \delta p_T)|||^2 .
\tag{5.90}
$$

Now, inequality (5.87) follows with $\gamma = \frac{\alpha}{4}$ by setting $(\delta u, \delta p_T) = (\hat{u} - u, \hat{p}_T - p_T)$.

Finally, we observe that (5.87) guarantees that G_r is strictly convex in a neighbourhood of a global minimum (u, p_T). This implies that global minima of (5.36) are isolated points. $\square$

Corollary 7 has the important purpose of characterizing minimum points of (P_u). In fact, its meaning is the following. By Proposition 13 we know that global minima of (5.36) are stationary points of the minimum-norm problem (5.26). Under the assumptions of Corollary 7, global minima of (5.36) are isolated points, which implies that stationary points of (5.26) are isolated points, which means that minima of (5.26) are isolated points.

5.4 Summary and remarks

The main topic of the present chapter was the exact-control of bilinear systems. In Section 5.1, some controllability results for quantum spin systems were discussed. In Section 5.2 and Section 5.3, we presented two different methodologies for the solution of exact-control problems governed by bilinear systems. These are the main novelties of the present chapter. In fact, we are not aware of previous works in the literature that study methods for the exact-control of bilinear systems provided with a rigorous theoretical framework.

Chapter 6

Numerical and implementation aspects of quantum control problems

The aim of this chapter is to address numerical aspects of quantum control problems. The first problem that is discussed, is the discretization of the optimality system. This issue plays a fundamental role in the numerical solution of quantum optimal control problems. In particular, in Section 6.1, we focus on the modified Crank-Nicolson scheme as an adequate discretization method of dynamical quantum systems with a bilinear control structure. We show that this scheme is stable, convergent and preserves algebraical properties that are involved in the evolution of a wavefunction governed by a Schrödinger equation. In Section 6.2, the first-discretize-than-optimize strategy is discussed and applied in order to derive discrete optimality conditions, that are suitable for numerical implementations. The second part of this chapter focuses on implementation details of optimization methods for solving optimal control and exact-control problems presented in Chapter 3, Chapter 4 and Chapter 5. In particular, in Section 6.3.1, numerical and implementation details of globalized Krylov-Newton and Krylov-SSN methods are discussed. In Section 6.4, we present a scheme for the implementation of the continuation method, that is theoretically discussed in Section 5.2. Section 6.5 presents implementation details of a cascadic-NCG method and Krylov-Newton method for the solution to the exact-control problem introduced in Section 5.

6.1 The modified Crank-Nicolson method

The aim of this section is to discuss the so-called modified Crank-Nicolson (MCN) scheme for the discretization of the following bilinear system

$$\dot{x} = \left[A + \sum_{n=1}^{N_C} u_n B_n \right] x \,. \tag{6.1}$$

The MCN scheme was introduced by von Winckel et al. [129] as an appropriate time discretization of the Schrödinger equation. In particular, with the aim to obtain an unconditionally stable and L^2-norm preserving scheme, they considered a modification of the Crank-Nicolson method by combining numerical integrators based on the Magnus expansion, trapezoidal rule for the approximation of integrals, and the usual Crank-Nicolson scheme. The MCN scheme can be obtained as follows. Consider the following problem

$$\dot{\psi}(t) = H(t)\psi(t) \,, \quad t \in (t^k, t^k + 1] \,, \quad \text{and} \quad \psi(t^k) = \psi_0 \,, \tag{6.2}$$

where $t^{k+1} = t^k + h$. Numerical integrators based on the Magnus expansion aim at writing the solution as follows; see, e.g., [61];

$$\psi(t) = \exp(\Omega(t))\psi_0 \,, \tag{6.3}$$

for a suitable matrix $\Omega(t)$. In particular, it is shown in [61], that $\Omega(t)$ is given by the Magnus expansion:

$$\Omega(t) = \int_{t^k}^t H(s)ds - \frac{1}{2} \int_{t^k}^t \left[\int_{t^k}^s H(\tau)d\tau, H(s) \right] ds + \cdots \,, \tag{6.4}$$

where $[\cdot, \cdot]$ denotes the commutator. The ansatz (6.3) allows to obtain the following approximation

$$\tilde{\psi}(t^{j+1}) = \exp(\tilde{\Omega})\tilde{\psi}(t^j) \,, \tag{6.5}$$

where $\tilde{\Omega}$ is a suitable approximation of $\Omega(t)$ in (t^j, t^{j+1}). By truncating the Magnus expansion (6.4) at the first term, and approximating the following integral

$$\int_{t^j}^{t^{j+1}} H(s)ds \tag{6.6}$$

by means of the trapezoidal rule, we obtain

$$\tilde{\Omega} = \frac{h}{2}\left[H(t^{j+1}) + H(t^j)\right] \,. \tag{6.7}$$

Denoting by $\hat{H} = \frac{1}{2}\left[H(t^{j+1}) + H(t^j)\right]$ and replacing (6.7) into (6.5), we obtain

$$\tilde{\psi}(t^{j+1}) = \exp(h\hat{H})\tilde{\psi}(t^j) \,, \tag{6.8}$$

which is the solution to the following linear system

$$\dot{\tilde{\psi}}(t) = \hat{H}\tilde{\psi}(t) \quad \text{in} \quad (t^j, t^{j+1}] \,. \tag{6.9}$$

By applying a Crank-Nicolson scheme to approximate (6.9), we write

$$\psi^{j+1} - \psi^j = \frac{h}{4}\left[H(t^{j+1}) + H(t^j)\right]\left[\psi^{j+1} + \psi^j\right] \,, \tag{6.10}$$

where we used (6.7). The discretization scheme (6.10) is the MCN scheme introduced in [129], for the approximation of (6.2).

Now, we apply the MCN scheme for the approximation of (6.1). Consider the time interval $[0, T]$, and approximate it with a uniform mesh of size h and N_t points, such that $0 = t^1 < \cdots < t^{N_t} = T$, the MCN discretization of the bilinear system (6.1) is the following

$$x^{j+1} - x^j = \frac{h}{4}\left[2A + \sum_{n=1}^{N_C} (u_n^{j+1} + u_n^j)B_n\right]\left(x^{j+1} + x^j\right) \,, \tag{6.11}$$

where $j = 1, ..., N_t - 1$, $h = \frac{T}{N_t - 1}$ and the starting point $x^1 = x(0)$ is given.

Convergence analysis for the approximation scheme (6.8) is given in [61], and we remark that, since the Magnus expansion (6.4) allows to construct approximations of $\Omega(t)$ with the same properties of $H(t)$, the approximation scheme (6.8) preserves the same algebraic properties of the original system. On the other hand, the MCN scheme (6.10)-(6.11)

is obtained from (6.8) by applying a Crank-Nicolson method, and even if it is expected that (6.10)-(6.11) maintains the same properties of the Magnus discretization (6.8), there is no rigorous analysis that ensures convergence and preservation of algebraic properties. For this reason, in the following we consider the MCN method for the discretization of (6.1) and we present analysis of convergence and stability of the scheme in Section 6.1.1. Moreover, in Section 6.1.2, we prove that the MCN scheme preserves all the algebraic properties required for a correct integration of a Schrödinger equation.

6.1.1 Stability and accuracy of the MCN scheme

In this section, we focus on the analysis of the MCN scheme (6.11) by assuming that the matrices A and B_n are skew-symmetric. This situation is common for bilinear systems related to NMR spectroscopy. In particular, in Proposition 17 we study stability and convergence of the MCN scheme, and show that it preserves the Euclidean norm.

Proposition 17. *Assume that A and B_n are skew-symmetric. Then, the MCN scheme (6.11) is norm-preserving (independently on h). Moreover, consider $u \in C^2([0,T]; \mathbb{R}^{N_C})$ and the corresponding solution to the bilinear system (5.3) $x \in C^3([0,T]; \mathbb{R}^{N_x})$. Then the MCN scheme (6.11) is unconditionally stable, consistent and second-order accurate.*

Proof. We denote by $x(t^j)$, $u(t^j)$ the values of x and u on the j-grid point. Moreover, x^j represents the discrete approximation of x at the point j.

We start discussing the norm-preservation of the MCN scheme. By multiplying equation (6.11) from the left by $\left[x^{j+1} + x^j \right]$, we obtain what follows

$$\langle x^{j+1} + x^j, x^{j+1} - x^j \rangle = \langle x^{j+1} + x^j, \frac{h}{4} \Big[2A + \sum_n (u_n^{j+1} + u_n^j) B_n \Big] (x^{j+1} + x^j) \rangle .$$

Since A and B_n are skew-symmetric the right hand side of the previous equation is zero. Consequently, we get the following

$$0 = \langle x^{j+1} + x^j, x^{j+1} - x^j \rangle = \langle x^{j+1}, x^{j+1} \rangle - \langle x^j, x^j \rangle ,$$

which implies that $\|x^{j+1}\|_2^2 = \|x^j\|_2^2$ for any mesh size h. Notice that the norm-preservation implies the unconditional stability of the MCN scheme.

To show consistency of (6.11), we proceed similarly to [113] as follows. The truncation error is given by

$$R_j = \frac{x(t^{j+1}) - x(t^j)}{h} - \frac{1}{4} \Big[2A + \sum_n \Big(u_n(t^{j+1}) + u_n(t^j) \Big) B_n \Big] \Big(x(t^{j+1}) + x(t^j) \Big) , \quad (6.12)$$

where $t^{j+1} = t^j + h$. We have also that

$$\dot{x}(t^j) = \Big(A + \sum_n u_n(t^j) B_n \Big) x(t^j) . \quad (6.13)$$

Now, consider the following Taylor expansions.

$$x(t^{j+1}) = x(t^j) + h\dot{x}(t^j) + \frac{h^2}{2}\ddot{x}(t^j) + \frac{h^3}{6}\dddot{x}(t^j) + O(h^4) , \quad (6.14)$$

By expanding $\dot{x}(t^{j+1})$ on the right-hand side, we have

$$\left(A + \sum_n u_n(t^{j+1})B_n\right)x(t^{j+1}) = \dot{x}(t^{j+1})$$

$$= \dot{x}(t^j) + h\ddot{x}(t^j) + \frac{h^2}{2}\dddot{x}(t^j) + O(h^3)\,. \tag{6.15}$$

Next, we consider the Taylor expansion of $u_n(t^{j+1})$ and use (6.13) to obtain the following

$$\left(A + \sum_n u_n(t^{j+1})B_n\right)x(t^j) = \dot{x}(t^j) + h\left(\sum_n \dot{u}_n(t^j)B_n\right)x(t^j)$$

$$+ \frac{h^2}{2}\left(\sum_n \ddot{u}_n(t^j)B_n\right)x(t^j) + O(h^3)\,. \tag{6.16}$$

By expanding $x(t^{j+1})$ and using (6.13), we get

$$\left(A + \sum_n u_n(t^j)B_n\right)x(t^{j+1}) = \dot{x}(t^j) + h\left(A + \sum_n u_n(t^j)B_n\right)\dot{x}(t^j)$$

$$+ \frac{h^2}{2}\left(A + \sum_n u_n(t^j)B_n\right)\ddot{x}(t^j) + O(h^3)\,. \tag{6.17}$$

Replacing (6.14)-(6.17) in (6.12), we obtain

$$R_j = \frac{h^2}{24}\dddot{x}(t^j) + \frac{h}{4}\ddot{x}(t^j) - \frac{1}{4}\left\{h\left(\left[A + \sum_n u_n(t^j)B_n\right]\dot{x}(t^j) + \left[\sum_n \dot{u}_n(t^j)B_n\right]x(t^j)\right)\right.$$

$$\left. + \frac{h^2}{2}\left(\left[A + \sum_n u_n(t^j)B_n\right]\ddot{x}(t^j) + \left[\sum_n \ddot{u}_n(t^j)B_n\right]x(t^j)\right)\right\} + O(h^3)\,.$$

Now, we can use the following expression

$$\ddot{x} = \frac{d}{dt}\left(\left[A + \sum_n u_n B_n\right]x\right) = \left[A + \sum_n u_n B_n\right]\dot{x} + \left[\sum_n \dot{u}_n B_n\right]x\,,$$

to obtain that

$$R_j = h^2\left[\frac{1}{24}\dddot{x}(t^j) - \frac{1}{8}\left(\left[A + \sum_n u_n(t^j)B_n\right]\ddot{x}(t^j) + \left[\sum_n \ddot{u}_n(t^j)B_n\right]x(t^j)\right)\right] + O(h^3)\,.$$

Next, we define the following constant

$$K := \max_{t^j}\left\|\frac{1}{24}\dddot{x}(t^j) - \frac{1}{8}\left(\left[A + \sum_n u_n(t^j)B_n\right]\ddot{x}(t^j) + \left[\sum_n \ddot{u}_n(t^j)B_n\right]x(t^j)\right)\right\|_2\,,$$

and obtain that

$$\|R_j\|_2 \le Kh^2 + O(h^3)\,.$$

Hence, there exists a positive constant $\tilde{K}$ such that

$$R := \max_j \|R_j\|_2 \le \tilde{K}h^2\,, \tag{6.18}$$

which means that the MCN scheme is consistent.

We analyze convergence of the MCN scheme. We first rewrite (6.11) and (6.12) in the following forms

$$x(t^{j+1}) = x(t^j) + hC^j\big(x(t^{j+1}) + x(t^j)\big) + hR_j \, , \tag{6.19}$$

and

$$x^{j+1} = x^j + hC^j\big(x^{j+1} + x^j\big) \, . \tag{6.20}$$

Subtracting term-by-term (6.20) from (6.19) and defining the global error $e^j := x(t^j) - x^j$, we obtain the following

$$e^{j+1} = e^j + hC^j\big(e^{j+1} + e^j\big) + hR_j \, . \tag{6.21}$$

Denoting by $L_C = \max_j \|C^j\|_{\mathcal{L}}$, we write that

$$\begin{aligned}
\|e^{j+1}\|_2 &= \|e^j + hC^j\big(e^{j+1} + e^j\big) + hR_j\|_2 \\
&\le \|e^j\|_2 + hL_C\|e^{j+1} + e^j\|_2 + h\|R_j\|_2 \\
&\le \|e^j\|_2 + hL_C\|e^{j+1}\|_2 + hL_C\|e^j\|_2 + hR \, ,
\end{aligned} \tag{6.22}$$

Assuming that $h < 1/L_C$, the previous (6.22) implies the following

$$\|e^{j+1}\|_2 \le \frac{1 + hL_C}{1 - hL_C}\|e^j\|_2 + \frac{h}{1 - hL_C}R \, . \tag{6.23}$$

It follows by induction that

$$\|e^j\|_2 \le \left(\frac{1 + hL_C}{1 - hL_C}\right)^j\|e^1\|_2 + \left[\left(\frac{1 + hL_C}{1 - hL_C}\right)^j - 1\right]\frac{R}{2L_C} \, . \tag{6.24}$$

Now, notice that $\|e^1\|_2 = \|x^1 - x(0)\|_2 = 0$ and $\left(\frac{1+hL_C}{1-hL_C}\right) \le \exp\left(\frac{2hL_C}{1-hL_C}\right)$, and recalling (6.18), we obtain what follows

$$\begin{aligned}
\|e^j\|_2 &\le \frac{1}{2L_C}\left\{\exp\left[j\left(\frac{2hL_C}{1 - hL_C}\right)\right] - 1\right\}\tilde{K}h^2 \\
&\le \frac{1}{2L_C}\left\{\exp\left(\frac{2(t^j - t^1)L_C}{1 - hL_C}\right) - 1\right\}\tilde{K}h^2 \\
&\le \frac{1}{2L_C}\left\{\exp\left(\frac{2TL_C}{1 - hL_C}\right) - 1\right\}\tilde{K}h^2 \, ,
\end{aligned} \tag{6.25}$$

which shows that the scheme (6.11) is convergent and second-order accurate.

Notice that the condition $h < 1/L_C$ is not required for stability. However, it is a standard condition to guarantee accuracy of a one-step second order scheme [113]. $\quad\square$

6.1.2 An algebraic analysis of the MCN scheme

Consider a Schrödinger equation in the following form

$$i\dot{\psi}(t) = H(t)\psi(t) \, , \quad t \in (0, T] \, , \quad \psi(0) = \psi_0 \, . \tag{6.26}$$

where $\psi(t) \in \mathbb{C}^N$. The Hamiltonian $H(t)$ is Hermitian and is assumed to have the following structure

$$H(t) = H_A(t) + iH_B(t) \, , \tag{6.27}$$

where i is the imaginary unit, $H_A(t) \in \mathbb{R}^{N\times N}$ is a symmetric matrix and $H_B(t) \in \mathbb{R}^{N\times N}$ is a skew-symmetric matrix. We remark that, (6.26)-(6.27) can represent the following

- a finite-dimensional Schrödinger equation, see, e.g., [15, 116, 122]; in this case, $H_A(t) = H_0 + u(t)H_1$; H_0 is a diagonal matrix, with real entries, containing the values of energy corresponding to the considered quantum levels; H_1 is a symmetric matrix, with real entries, containing the possible transitions between the considered quantum energy levels and $H_B = 0$;

- the (space) discretization of an infinite-dimensional Schrödinger equation, see, e.g., [129]; in this case, $H_A(t) = H_0 + u(t)H_1$; H_0 is a symmetric matrix, with real entries; for instance, it could represent the discrete Laplace operator; H_1 is a symmetric matrix, with real entries; it could be, for instance, a diagonal matrix representing a potential energy, and $H_B = 0$;

- the (space) discretization of an infinite dimensional Schrödinger equation involving the action of magnetic fields and the effect of the spin of the particle; in this case, $H_A(t) = H_0 + u_1(t)H_1$, where H_0 and H_1 are symmetric matrices with real entries, and $H_B(t) = u_2(t)H_2$, with H_B a skew-symmetric matrix.

The main idea of our analysis, arises from the postulate of quantum mechanics which states that the time-evolution of a wave function $\psi(t)$, satisfying the Schrödinger equation (6.26), can be regarded as the time evolution of a unitary operator $\Upsilon(t)$, whose evolution is given by

$$\dot{\Upsilon}(t) = -iH(t)\Upsilon(t) , \ t \in (0, T] , \ \Upsilon(0) = I_N , \tag{6.28}$$

where I_N represents the identity matrix. Henceforth, we have that

$$\psi(t) = \Upsilon(t)\psi(0) , \ t \in [0, T] ; \tag{6.29}$$

see, e.g., [132]. We remark that, the (6.28) is referred to as the lifted-equation of (6.26). Since $H(t)$ is an Hermitian matrix, $iH(t)$ is skew-Hermitian. Consequently, it belongs to the matrix Lie algebra $\mathfrak{u}(N; \mathbb{C})$, and this implies that the operator $\Upsilon(t)$ belongs to the matrix Lie group $U(N; \mathbb{C})$ of unitary matrices; see, e.g., [53]. It is clear from equation (6.29), that the time evolution of the Schrödinger equation (6.26) is governed by a unitary propagator. Consequently, for an adequate numerical integration of (6.26), it is suitable to consider a numerical scheme that is capable to preserve the algebraic properties of a unitary propagator. The MCN scheme applied to (6.28) reads as follows

$$\Upsilon_{k+1} - \Upsilon_k = \tilde{\mathcal{H}}_k\left[\Upsilon_{k+1} + \Upsilon_k\right] , \tag{6.30}$$

with

$$\tilde{\mathcal{H}}_k = -i\frac{h}{4}\left[H(t^{k+1}) + H(t^k)\right] . \tag{6.31}$$

In what follows, we show that the MCN scheme (6.10) preserves the unitary properties.

Theorem 31. *The MCN scheme* (6.30)-(6.31) *preserves unitary propagation, that is*

$$\Upsilon_k \in U(N; \mathbb{C}) \Rightarrow \Upsilon_{k+1} \in U(N; \mathbb{C}) .$$

Proof. Assume that $\Upsilon^k \in U(N; \mathbb{C})$, which is equivalent to $\Upsilon_k^*\Upsilon_k = \Upsilon_k\Upsilon_k^* = I_N$. Multiply (6.30) from the left with Υ_k^* to get the following

$$\Upsilon_k^*\Upsilon_{k+1} - \Upsilon_k^*\Upsilon_k = \Upsilon_k^*\tilde{\mathcal{H}}_k\Upsilon_{k+1} + \Upsilon_k^*\tilde{\mathcal{H}}_k\Upsilon_k ,$$

which implies that
$$\Upsilon_k^* \Upsilon_{k+1} - I_N = \Upsilon_k^* \tilde{\mathcal{H}}_k \Upsilon_{k+1} + \Upsilon_k^* \tilde{\mathcal{H}}_k \Upsilon_k .$$

By taking the adjoint and using the fact that $\tilde{\mathcal{H}}_k$ is skew-hermitian, we have the following
$$\Upsilon_{k+1}^* \Upsilon_k - I_N = -\Upsilon_{k+1}^* \tilde{\mathcal{H}}_k \Upsilon_k - \Upsilon_k^* \tilde{\mathcal{H}}_k \Upsilon_k . \tag{6.32}$$

Next, by multiplying (6.30) from the left with Υ_{k+1}^*, we obtain
$$\Upsilon_{k+1}^* \Upsilon_{k+1} - \Upsilon_{k+1}^* \Upsilon_k = \Upsilon_{k+1}^* \tilde{\mathcal{H}}_k \Upsilon_{k+1} + \Upsilon_{k+1}^* \tilde{\mathcal{H}}_k \Upsilon_k . \tag{6.33}$$

Adding term-by-term (6.32) and (6.33), we obtain the following
$$\Upsilon_{k+1}^* \Upsilon_{k+1} - I_N = \Upsilon_{k+1}^* \tilde{\mathcal{H}}_k \Upsilon_{k+1} - \Upsilon_k^* \tilde{\mathcal{H}}_k \Upsilon_k . \tag{6.34}$$

By taking the adjoint of (6.34), we write
$$\Upsilon_{k+1}^* \Upsilon_{k+1} - I_N = -\left(\Upsilon_{k+1}^* \tilde{\mathcal{H}}_k \Upsilon_{k+1} - \Upsilon_k^* \tilde{\mathcal{H}}_k \Upsilon_k \right) . \tag{6.35}$$

Equations (6.34) and (6.35) imply that Υ_{k+1} is a unitary matrix. $\qquad\square$

An important geometric and algebraic property that characterizes a Schrödinger equation is the so-called symplecticity. In order to analyze this property, we consider the following. It is known that $U(N; \mathbb{C}) \cong U(2N; \mathbb{R})$, where $U(2N; \mathbb{R})$ represents the group of all $2N \times 2N$ unitary matrices with real entries. Moreover, the following equality holds [67]
$$U(2N; \mathbb{R}) = Sp(2N; \mathbb{R}) \cap O(2N; \mathbb{R}) , \tag{6.36}$$
where [53]

- $Sp(2N; \mathbb{R})$ is the Lie group of the symplectic matrices; a matrix M belongs to $Sp(2N; \mathbb{R})$ if and only if
$$M^T J M = J , \tag{6.37}$$
  where
$$J = \begin{pmatrix} 0 & I_N \\ -I_N & 0 \end{pmatrix} ; \tag{6.38}$$
  the corresponding Lie algebra is $\mathfrak{sp}(2N; \mathbb{R})$, that is the algebra generated by all the matrices C such that
$$C^T J + J C = 0 . \tag{6.39}$$
  A matrix $C \in \mathfrak{sp}(2N; \mathbb{R})$ has the following block structure
$$C = \begin{pmatrix} C_1 & C_2 \\ C_3 & -C_1^T \end{pmatrix} , \tag{6.40}$$
  where C_2 and C_3 are $N \times N$ symmetric matrices;

- $O(2N; \mathbb{R})$ is the Lie group of all $2N \times 2N$ orthogonal matrices; a matrix M belongs to $O(2N; \mathbb{R})$ if and only if
$$M M^T = M^T M = I_{2N} , \tag{6.41}$$
  where I_{2N} is the $2N \times 2N$ identity; the corresponding Lie algebra is $\mathfrak{o}(2N; \mathbb{R})$, that is the algebra of all skew-symmetric matrices.

To see the validity of the intersection (6.36) for the Schrödinger equation (6.26), we proceed as is Section 2.4.2: we split ψ in real and imaginary parts

$$\psi = \psi_\Re + i\psi_\Im \,, \tag{6.42}$$

and rewrite (6.26) as follows

$$\frac{d}{dt}\begin{pmatrix} \psi_\Re \\ \psi_\Im \end{pmatrix} = \begin{pmatrix} H_B & H_A \\ -H_A & H_B \end{pmatrix}\begin{pmatrix} \psi_\Re \\ \psi_\Im \end{pmatrix} \,. \tag{6.43}$$

Equation (6.43), in a more compact form, reads as follows

$$\dot{\Psi} = \mathcal{H}(t)\Psi \,, \tag{6.44}$$

where

$$\Psi := \begin{pmatrix} \psi_\Re \\ \psi_\Im \end{pmatrix} \,, \tag{6.45}$$

and

$$\mathcal{H}(t) := \begin{pmatrix} H_B(t) & H_A(t) \\ -H_A(t) & H_B(t) \end{pmatrix} \,. \tag{6.46}$$

We notice that, since H_B is skew-symmetric, $\mathcal{H}$ is a traceless skew-symmetric matrix having also the block-structure (6.40). This means that

$$\mathcal{H}(t) \in \mathfrak{sp}(2N;\mathbb{R}) \cap \mathfrak{o}(2N;\mathbb{R}) \,. \tag{6.47}$$

Consequently, equation (6.44) admits the following lifted equation

$$\dot{M}(t) = \mathcal{H}(t)M(t) \quad, \quad t \in (0,T] \quad, \quad M(0) = I_{2N} \,, \tag{6.48}$$

The time evolution of Ψ is obtained by the following linear action

$$\Psi(t) = M(t)\Psi(0) \quad, \quad t \in [0,T] \,. \tag{6.49}$$

The condition (6.47) and equation (6.48) imply that

$$M \in U(2N;\mathbb{R}) = Sp(2N;\mathbb{R}) \cap O(2N;\mathbb{R}) \,, \tag{6.50}$$

which means that the evolution of the Schrödinger equation (6.26) is characterized by the algebraic properties characterizing $Sp(2N;\mathbb{R})$ and $O(2N;\mathbb{R})$. For this reason, in the sequel we show that the MCN method preserves these algebraic properties.

The MCN scheme applied on (6.48) reads as follows

$$M_{k+1} - M_k = \frac{h}{4}\big[\mathcal{H}(t^{k+1}) + \mathcal{H}(t^k)\big]\big[M_{k+1} + M_k\big] \,. \tag{6.51}$$

Recalling equations (6.46) and (6.27), we define the following

$$\hat{\mathcal{H}}_k := \frac{h}{4}\big[\mathcal{H}(t^{k+1}) + \mathcal{H}(t^k)\big] = \begin{pmatrix} B_{k,k+1} & A_{k,k+1} \\ -A_{k,k+1} & B_{k,k+1} \end{pmatrix} \,. \tag{6.52}$$

where $A_{k,k+1} := H_A(t^{k+1}) + H_A(t^k)$ and $B_{k,k+1} := H_B(t^{k+1}) + H_B(t^k)$.

We remark that $A_{k,k+1}$ and $B_{k,k+1}$ are symmetric and skew-symmetric, respectively. Consequently $\hat{\mathcal{H}}_k \in \mathfrak{sp}(2N;\mathbb{R}) \cap \mathfrak{o}(2N;\mathbb{R})$. By means of (6.52), equation (6.51) becomes

$$M_{k+1} - M_k = \hat{\mathcal{H}}_k \big[M_{k+1} + M_k \big] . \tag{6.53}$$

We remark that equation (6.49) holds at any time. Therefore, by multiplying (6.53) to the right with $\Psi(0)$, we have

$$\Psi_k = M_k \Psi(0) . \tag{6.54}$$

Hence, we obtain

$$\Psi_{k+1} - \Psi_k = \hat{\mathcal{H}}_k \Big[\Psi_{k+1} + \Psi_k \Big] , \tag{6.55}$$

that is the MCN discretization of (6.44). This means that, the discrete approximated solution to (6.26) can be obtained by the discrete approximated solution to (6.53) combined with the linear action (6.54).

The following results show the algebraic properties of the MCN scheme. In particular, in Theorem 32 we prove that the MCN scheme (6.53) preserves symplecticity. Theorem 33 shows that the MCN scheme (6.53) preserves orthogonality. Corollaries 8 and 9 discuss norm-preserving properties of the scheme (6.53).

Theorem 32. *The MCN scheme* (6.53) *preserves symplecticity, in the following sense*

$$M_k \in Sp(2N;\mathbb{R}) \Rightarrow M_{k+1} \in Sp(2N;\mathbb{R}).$$

Proof. We assume that $M_k \in Sp(2N;\mathbb{R})$, which is equivalent to $M_k^T J M_k = J$, with J defined in (6.38).

Notice that $S := J\hat{\mathcal{H}}_k$ is a symmetric matrix. We multiply (6.53) from the left with $M_k^T J$, and we get the following

$$M_k^T J M_{k+1} - M_k^T J M_k = M_k^T J\hat{\mathcal{H}}_k M_{k+1} + M_k^T J\hat{\mathcal{H}}_k M_k ,$$

that implies the following

$$M_k^T J M_{k+1} - J = M_k^T S M_{k+1} + M_k^T S M_k .$$

By taking the transpose and using the symmetry of S and the skew-symmetry of J, we obtain

$$-M_{k+1}^T J M_k + J = M_{k+1}^T S M_k + M_k^T S M_k . \tag{6.56}$$

Now, we multiply (6.53) from the left with $M_{k+1}^T J$, and write that

$$M_{k+1}^T J M_{k+1} - M_{k+1}^T J M_k = M_{k+1}^T S M_{k+1} + M_{k+1}^T S M_k . \tag{6.57}$$

Subtracting term-by-term the last equation of (6.56) from (6.57), we obtain the following

$$M_{k+1}^T J M_{k+1} - J = M_{k+1}^T S M_{k+1} - M_k^T S M_k . \tag{6.58}$$

Taking the transpose of (6.58), we get

$$-(M_{k+1}^T J M_{k+1} - J) = M_{k+1}^T S M_{k+1} - M_k^T S M_k . \tag{6.59}$$

Equations (6.58) and (6.59) imply that $M_{k+1}^T J M_{k+1} - J = 0$, which is our claim.

$\square$

Theorem 33. *The MCN scheme* (6.53) *preserves orthogonality, in the following sense*

$$M_k \in O(2N;\mathbb{R}) \Rightarrow M_{k+1} \in O(2N;\mathbb{R}).$$

Proof. We assume that $M_k \in O(2N;\mathbb{R})$, which is equivalent to $M_k^T M_k = M_k M_k^T = I_{2N}$. By multiplying (6.53) from the left with M_k^T, we get the following

$$M_k^T M_{k+1} - M_k^T M_k = M_k^T \hat{\mathcal{H}}_k M_{k+1} + M_k^T \hat{\mathcal{H}}_k M_k .$$

that implies the following

$$M_k^T M_{k+1} - I_{2N} = M_k^T \hat{\mathcal{H}}_k M_{k+1} + M_k^T \hat{\mathcal{H}}_k M_k .$$

By taking the transpose and using the skew-symmetry of $\hat{\mathcal{H}}_k$, we obtain

$$M_{k+1}^T M_k - I_{2N} = -M_{k+1}^T \hat{\mathcal{H}}_k M_k - M_k^T \hat{\mathcal{H}}_k M_k . \tag{6.60}$$

By multiplying (6.53) from the left with M_{k+1}^T, we obtain that

$$M_{k+1}^T M_{k+1} - M_{k+1}^T M_k = M_{k+1}^T \hat{\mathcal{H}}_k M_{k+1} + M_{k+1}^T \hat{\mathcal{H}}_k M_k . \tag{6.61}$$

Adding term-by-term the last equation of (6.60) and (6.61), we obtain the following

$$M_{k+1}^T M_{k+1} - I_{2N} = M_{k+1}^T \hat{\mathcal{H}}_k M_{k+1} - M_k^T \hat{\mathcal{H}}_k M_k . \tag{6.62}$$

Taking the transpose of (6.62), we obtain

$$M_{k+1}^T M_{k+1} - I_{2N} = -(M_{k+1}^T \hat{\mathcal{H}}_k M_{k+1} - M_k^T \hat{\mathcal{H}}_k M_k) . \tag{6.63}$$

Equations (6.62) and (6.63) imply that $M_{k+1}^T M_{k+1} - I_{2N} = 0$, which is our claim. $\square$

Next, we discuss the norm-preservation of the MCN scheme. For this purpose, we recall the inner product $\langle \cdot | \cdot \rangle : \mathbb{R}^{2N \times 2N} \times \mathbb{R}^{2N \times 2N} \to \mathbb{R}$ defined in Section 2.3.3 as follows

$$\langle M_j | M_k \rangle = \text{trace}(M_j^T M_k) , \tag{6.64}$$

which induces the Hilbert-Schmidt norm $\| \cdot \|_{\mathcal{L}} := \sqrt{\langle \cdot | \cdot \rangle}$. Hence the Lie group $U(2N;\mathbb{R})$ endowed with (6.64) form a real normed Hilbert vector space. The norm-preserving property can be seen as a specific case of the property proved in the following result.

Corollary 8. *The MCN scheme* (6.53) *preserves the norm, that is*

$$\|M_k\|_{\mathcal{L}} = c \Rightarrow \|M_{k+1}\|_{\mathcal{L}} = c ,$$

for any constant c.

Proof. The norm-preserving property is induced by the orthogonality-preserving property. Recalling Theorem 33, we write that

$$\|M_k\|_{\mathcal{L}} = \sqrt{\text{trace}(M_k^T M_k)} = \sqrt{\text{trace}(I_{2N})} = \sqrt{\text{trace}(M_{k+1}^T M_{k+1})} = \|M_{k+1}\|_{\mathcal{L}} ,$$

and the claim follows. $\square$

Another consequence of the orthogonality-preserving properties of (6.53) is given in the following corollary. This shows that the orthogonality-preserving property of the scheme (6.53) induces the norm-preservation of the MCN scheme (6.55).

Corollary 9. *Assume that the MCN scheme* (6.53) *preserves orthogonality, then the MCN scheme* (6.55) *is norm-preserving.*

Proof. Recall Equation (6.54), that is $\Psi_k = M_k \Psi(0)$. Using Theorem 33, we write the following

$$\langle \Psi_k, \Psi_k \rangle = \langle M_k \Psi(0), M_k \Psi(0) \rangle = \langle \Psi(0), M_k^T M_k \Psi(0) \rangle = \langle \Psi(0), \Psi(0) \rangle , \tag{6.65}$$

that is our claim. $\square$

6.2 First-discretize-then-optimize strategy and discrete optimality conditions

An important aspect in the numerics of optimal control problems is the discretization of the reduced gradient. In fact, if the discrete gradient is not appropriately chosen, non-negligible inconsistencies could arise during the numerical solution of the optimal control problem. These inconsistencies can give rise to two main problems. On one hand, there is the possible emergence of spurious stationary point and minima, see, e.g., [22]. On the other hand, a gradient that is not consistent with the discrete cost functional can disturb the numerical optimization process, that can show a convergence rate slower than the theoretical expectation, or that can be even not converging in the worst situations. Moreover, another drawback of an inadequate discretization is that the discrete Hessian may not be symmetric, see, e.g., [129]. This is an important aspect in the application of second-order optimization methods.

For these reasons, in order to derive an adequate discrete reduced gradient, we consider the so-called first-discretize-then-optimize strategy; see, e.g., [16, 22, 129]. According to this discretization strategy, one has to proceed as follows. First, discretize the optimal control problem, which means to discretize the cost functional and the constraints. Second, construct the corresponding discrete Lagrangian function. Third, derive the first-order discrete optimality system and the corresponding linearized equations. With this procedure, a sufficiently fine mesh size is only required to obtain solutions to the differential equations that are sufficiently accurate to approximate differential models.

To apply the first-discretize-then-optimize strategy, we approximate the time interval $[0, T]$ with a uniform mesh of size h and N_t points, such that $0 = t^1 < \cdots < t^{N_t} = T$, and consider the following discrete $L^2((0, T); \mathbb{R}^m)$-scalar product

$$\langle a, b \rangle_{L_h^2} := h \sum_{j=2}^{N_t} \langle a^j, b^j \rangle \,, \tag{6.66}$$

where a and b are the discretizations of any two functions belonging to the $L^2((0, T); \mathbb{R}^m)$ space, and m is equal to N^2 for the state and to N_C for the control. We denote by $\| \cdot \|_{L_h^2}$ the norm induced by (6.66). Next, the discretization of problem (3.42) is the following

$$\min_{x,u} \quad J_h(x, u) := \frac{1}{2} \langle x^{N_t} - x_T, x^{N_t} - x_T \rangle + \frac{\nu}{2} h \sum_{n=1}^{N_C} \sum_{j=2}^{N_t} (u_n^j)^2$$

$$\text{s.t.} \quad \frac{1}{h} \big(x^j - x^{j-1} \big) = \frac{1}{4} \left[2A + \sum_{n=1}^{N_C} (u_n^j + u_n^{j-1}) B_n \right] \left[x^j + x^{j-1} \right] \tag{6.67}$$

$$\text{for } j = 2, \ldots, N_t \text{ and with } x^1 = x(0) \,.$$

Now, we define the constraint function $c(x, u)$ as follows

$$c_h(x, u)^j := \frac{1}{h} \big(x^j - x^{j-1} \big) - \frac{1}{4} \left[2A + \sum_{n=1}^{N_C} (u_n^j + u_n^{j-1}) B_n \right] \left[x^j + x^{j-1} \right] \,, \tag{6.68}$$

for $j = 2, \ldots, N_t$. The corresponding discrete Lagrangian function is given by

$$L_h(x, u, p) := \frac{1}{2} \langle x^{N_t} - x_T, x^{N_t} - x_T \rangle + \frac{\nu}{2} h \sum_{n=1}^{N_C} \sum_{j=2}^{N_t} (u_n^j)^2 + h \sum_{j=2}^{N_t} \langle p^j, c_h(x, u)^j \rangle \,. \tag{6.69}$$

The derivatives of the discrete Lagrangian with respect to p^j allow to obtain the following primal equation

$$x^j - x^{j-1} = \frac{h}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^j + u_n^{j-1})B_n\right]\left[x^j + x^{j-1}\right] , \tag{6.70}$$

for $j = 2, ..., N_t$ with $x^1 = x(0)$. The derivatives of L_h with respect to x^j yields the following adjoint dual equation

$$p^j - p^{j+1} = \frac{h}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^j + u_n^{j-1})B_n\right]^* p^j + \frac{h}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^{j+1} + u_n^j)B_n\right]^* p^{j+1} , \tag{6.71}$$

for $j = N_t - 1, \ldots, 2$, and with

$$(x^{N_t} - x_T) + h\left(\frac{1}{h}I - \frac{1}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^{N_t} + u_n^{N_t-1})B_n\right]^*\right)p^{N_t} = 0 . \tag{6.72}$$

By assuming that $u^1 = 0$ and deriving the discrete Lagrangian with respect to u^j, we obtain the discrete gradient, that is

$$\left(\nabla_u J_h\right)_n^j = \nu u_n^j - \frac{1}{4}\langle B_n(x^{j+1} + x^j), p^{j+1}\rangle - \frac{1}{4}\langle B_n(x^j + x^{j-1}), p^j\rangle , \tag{6.73}$$

for $j = 2, \ldots, N_t - 1$, and

$$\left(\nabla_u J_h\right)_n^{N_t} = \nu u_n^{N_t} - \frac{1}{4}\langle B_n(x^{N_t} + x^{N_t-1}), p^{N_t}\rangle , \tag{6.74}$$

for $n = 1, \ldots, N_C$. Notice that, in the case of an unconstrained problem ($U_{ad} = L^2$), the optimality condition is $\left(\nabla_u J_h\right)_n^j = 0$ for any j and any n. In a similar way, we can derive the linearized equations. In particular, the linearized constraint equation is given by

$$\begin{aligned} \delta x^j - \delta x^{j-1} &= \frac{h}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^j + u_n^{j-1})B_n\right]\left[\delta x^j + \delta x^{j-1}\right] \\ &+ \frac{h}{4}\left[\sum_{n=1}^{N_C}(\delta u_n^j + \delta u_n^{j-1})B_n\right]\left[x^j + x^{j-1}\right] , \end{aligned} \tag{6.75}$$

for $j = 2, \ldots, N_t$ with $\delta x^1 = 0$. The linearized adjoint equation is

$$\begin{aligned} \delta p^j - \delta p^{j+1} &= \frac{h}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^j + u_n^{j-1})B_n\right]^* \delta p^j + \frac{h}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^{j+1} + u_n^j)B_n\right]^* \delta p^{j+1} \\ &+ \frac{h}{4}\left[\sum_{n=1}^{N_C}(\delta u_n^j + \delta u_n^{j-1})B_n\right]^* p^j + \frac{h}{4}\left[\sum_{n=1}^{N_C}(\delta u_n^{j+1} + \delta u_n^j)B_n\right]^* p^{j+1} , \end{aligned}$$
$$\tag{6.76}$$

for $j = N_t - 1, \ldots, 2$, and

$$\delta x^{N_t} + h\left(\frac{1}{h}I - \frac{1}{4}\left[2A + \sum_{n=1}^{N_C}(u_n^{N_t} + u_n^{N_t-1})B_n\right]^*\right)\delta p^{N_t} - \frac{h}{4}\left[\sum_{n=1}^{N_C}(\delta u_n^{N_t} + \delta u_n^{N_t-1})B_n\right]^* p^{N_t} = 0 . \tag{6.77}$$

The action of the discrete reduced Hessian $\nabla_u^2 J_h$ on a discrete function δu is given by the following

$$
\begin{aligned}
\left(\nabla_u^2 J_h \delta u\right)_n^j &= \nu \delta u_n^j - \frac{1}{4}\left\langle B_n\left(x^{j+1} + x^j\right), \delta p^{j+1}\right\rangle - \frac{1}{4}\left\langle B_n\left(\delta x^{j+1} + \delta x^j\right), p^{j+1}\right\rangle \\
&\quad - \frac{1}{4}\left\langle B_n\left(x^j + x^{j-1}\right), \delta p^j\right\rangle - \frac{1}{4}\left\langle B_n\left(\delta x^j + \delta x^{j-1}\right), p^j\right\rangle ,
\end{aligned}
\tag{6.78}
$$

for $n = 1, \ldots, N_C$, $j = 2, \ldots, N_t - 1$, and

$$
\left(\nabla_u^2 J_h \delta u\right)_n^{N_t} = \nu \delta u_n^{N_t} - \frac{1}{4}\left\langle B_n\left(\delta x^{N_t} + \delta x^{N_t-1}\right), p^{N_t}\right\rangle - \frac{1}{4}\left\langle B_n\left(x^{N_t} + x^{N_t-1}\right), \delta p^{N_t}\right\rangle , \tag{6.79}
$$

for $n = 1, \ldots, N_C$.

Finally, we remark that, once the discrete reduced gradient is obtained, it is possible to use it to obtain discrete optimality conditions for optimal control problems with controls constraints, like the problems discussed in Section 3 and Section 4. The same arguments hold also for the construction of the discrete reduced Jacobian, in the case of a SSN method.

6.3 Globalized Newton and SSN algorithms

In this section, we present a globalized Krylov-Newton and semi-smooth Krylov Newton matrix-free procedures, and discuss implementation details. With matrix-free, we mean that the generalized Jacobian is never built during the iterations, and only its action on a given vector-function is computed in the Krylov-iterations for the solution of the Newton linear system. In particular, in Section 6.3.1 we discuss a globalized Krylov-Newton algorithm. In Section 6.3.2, we present a globalized Krylov-SSN algorithm.

6.3.1 A globalized Krylov-Newton scheme

In this section, a numerical scheme corresponding to the Krylov-Newton method is discussed. In particular, we describe the algorithm used to perform the procedure (4.17), that in a reduce form is given by

$$
\begin{aligned}
S_1 : & \quad \mathcal{J}_r(u^k)(\delta u) = -\mathcal{F}(u^k) \\
S_2 : & \quad u^{k+1} = u^k + \delta u ,
\end{aligned}
$$

which generates the Newton sequence $\{u^k\}_{k=1}^\infty$. The numerical computation of this sequence has two main drawbacks that have to be addressed. First, the dimension of the discrete Hessian can be very large and consequently the computational effort for its construction and inversion becomes too expensive. For this reason, the numerical procedure is based on the reduced formulation of the optimal control problem, and the Newton sequence $\{u^k\}_{k=1}^\infty$ is constructed in the space of solutions to constraint and adjoint equations. Second, the bilinearity of the quantum equation governing the optimal control problem makes, in general, the reduced cost functional non-convex. Consequently, a globalization strategy, specifically a line-search method, is required in the numerical procedure. We remark that it could be necessary in quantum control problems to initialize the Krylov-Newton procedure by using continuation techniques; see, e.g., [129, 130].

The Krylov-Newton procedure is given in the following algorithms 1-6. The following Algorithm 1 is the Krylov-Newton scheme, where one recognizes that in the while-loop the

Algorithm 2 is called to compute the reduced gradient for a given control u, the Algorithm 3 is called to solve the Newton linear system and compute the updating direction δu, and a line-search algorithm is called to globalize the procedure by computing a step-length α.

Algorithm 1 (Krylov-Newton scheme)

Input u, $k = 0$, k_{max}, tol;
Call Algorithm 2 to compute $\nabla_u J_r(u)$;
while $k < k_{max}$ **do**

- call Algorithm 3 to compute the search direction δu;

- perform a line-search to compute a step-length α;

- update: $u \leftarrow u + \alpha\, \delta u$;

- call Algorithm 2 to compute $\nabla_u J_r(u)$;

- if $\|\nabla_u J_r(u)\|_{L^2} \leq tol$, break;

- update: $k \leftarrow k + 1$;

end while

The reduced gradient is computed by means of Algorithm 2. This scheme consists in solving forward the bilinear constraint to obtain the state x, solving backward the adjoint problem to obtain the Lagrange multiplier p, and assembling the reduced gradient. It is clear from this routine that the iterative procedure given in Algorithm 1 is defined in the space of solutions to constraint and adjoint equations.

Algorithm 2 (Evaluation of the reduced gradient $\nabla_u J_r(u)$)

Input u;

- solve the forward problem (4.28) and get x;

- solve the backward problem (4.29) and get p;

- assemble $\nabla_u J_r(u)$ using (4.30);

The Newton linear system $\mathcal{J}_r(u)(\delta u) = -\nabla_u J_r(u)$ is solved by means of Algorithm 3, which consists in the application of a Krylov iterative solver. This aspect is crucial in the Krylov-Newton procedure, because a Krylov solver allows to avoid the construction of the discrete Hessian operator $\mathcal{J}$ (and its reduced form $\mathcal{J}_r$) by supplying a routine that computes the action of the reduced Hessian on a given vector function δu. This computation is performed by means of Algorithm 4. We remark that, depending on the eigenvalues of the Hessian and the specific Krylov solver, the computed direction δu can be an ascent direction, see, e.g., [16]. To address this situation, in the case where δu is an ascent direction, $-\delta u$ is used.

Algorithm 3 (Solve the Newton equation)

Input u, an initial guess for δu;

- compute δu by solving $\mathcal{J}_r(u)(\delta u) = -\nabla_u J_r(u)$ (use a linear system solver, e.g., GMRES or CG calling Algorithm 4 to apply $\mathcal{J}_r(u)$);

- if $\langle \mathcal{J}_r(u), \delta u \rangle_{L^2} > 0$, set $\delta u \leftarrow -\delta u$;

Algorithm 4 compute the action of the reduced Hessian $\mathcal{J}_r(u)$ on a given function δu. This computation is performed by solving the linearized constraint problem to obtain

---

δx, solving the linearized adjoint equation to obtain δp, and assembling the action of the reduced Hessian as described in Section 4.2 by means of (4.34).

Algorithm 4 (Apply the reduced Hessian to a vector δu)

Input x, p, u, δu;

- integrate the linearized forward problem (4.32) and get δx;

- integrate the linearized backward problem (4.33) and get δp;

- assemble $\mathcal{J}_r(u)(\delta u)$ by means of (4.34);

Since the constraint equation is bilinear, the objective is in general non-convex and therefore a line-search is required to globalize the Newton method. Once an updating direction δu is computed, then the line-search strategy is applied to compute a step-length α satisfying the strong Wolfe conditions (SWC), representing a sufficient reduction of the cost functional, as follows (Armijo's condition)

$$J_r(u + \alpha \delta u) \leq J_r(u) + c_1 \alpha \langle \delta u, \nabla J_r(u) \rangle_{L^2} , \tag{6.80}$$

such that a so called curvature condition

$$|\langle \delta u, \nabla J_r(u + \alpha \delta u) \rangle_{L^2}| \leq c_2 |\langle \delta u, \nabla J_r(u) \rangle_{L^2}| \tag{6.81}$$

is satisfied to avoid unacceptably small steps in the next iterations; see, e.g., [16, 52, 128, 129, 130]. The coefficients c_1 and c_2 are chosen such that $0 < c_1 \ll 1$ and $c_1 < c_2 < 1$. With a Newton-type method, it is recommended to begin with $\alpha = 1$ [94, 129, 130]. This is a reasonable choice when the cost functional is locally quadratic. However, in our non-convex optimization problem the desired step-length can be different and sometimes orders of magnitude smaller. In [16] and [130], an upper bound on the maximum feasible step-length is estimated using a quadratic approximation of the cost functional, that is $J_r(u + \alpha \delta u) \geq m_2 \alpha^2 + m_1 \alpha + m_0$, where the coefficients are $m_0 = \frac{\nu}{2} \langle u, u \rangle_{L^2} - J_r(u)$, $m_1 = \nu \langle u, \delta u \rangle_{L^2}$ and $m_2 = \frac{\nu}{2} \langle \delta u, \delta u \rangle_{L^2}$. Consequently, the upper bound on a feasible step-length is evaluated as follows

$$\alpha_{max} = \frac{\sqrt{m_1^2 - 4m_0 m_2} - m_1}{2m_2} . \tag{6.82}$$

We remark that, in [130] is proved that for c_1 small enough there exists an $\alpha \in (0, \alpha_{max}]$ that satisfies the SWC condition. Starting with the approximation α_{max}, a line-search method is developed in [130]. This scheme is based on a combination of two strategies: a cubic polynomial approximation and a bisection method. This line-search method is based on the following criteria which ensure that a twice differentiable function $\alpha \mapsto f(\alpha)$ must have at least one minimizer in a given interval (α_r, α_l). We have the two following propositions give the guideline for the formulation of Algorithm 6 and Algorithm 5 [130].

Proposition 18. *Suppose that f is continuously differentiable. If $f'(\alpha_l) < 0$ and $f'(\alpha_r) > 0$, then there must be at least one point $\alpha \in (\alpha_r, \alpha_l)$ such that α is a local minimum.*

Proposition 19. *Suppose that f is continuously differentiable. If $f'(\alpha_l) < 0$ and $f(\alpha_r) > f(\alpha_l)$ $\big(f'(\alpha_l) > 0$, $f(\alpha_r) < f(\alpha_l)\big)$, then there must be at least one point $\alpha \in (\alpha_r, \alpha_l)$ such that α is a local minimum.*

Notice that if the coefficient m_2 is very small, then the quadratic model could be not valid. This situation can occur if

- the norm of search direction δu is very small, which means that the optimization process reached a small neighbourhood of a stationary point;

- the regularization parameter ν is very small. This can happen when the goal is to maximize the tracking term of J.

Moreover, according to our experience, because of the bilinear structure of our problem, near a stationary point it can be difficult for the algorithm to satisfy the Wolfe's conditions. Motivated by these heuristics, we slightly modify the line-search strategy defined in [130], by adding another criterion based on the coefficient m_2 for the following choice:

- if m_2 is large enough, a "strong strategy" is chosen. To have a robust and efficient scheme we combine cubic approximations of the cost functional with a bisection method, as proposed in [16] and [130] and shown in Algorithms 6 and 5;

- if m_2 is small, then a "weak strategy" is chosen: the algorithm compute α such that it satisfies only the Armijo's condition.

The described line-search procedure is given in detail in the two following algorithms. In particular, Algorithm 5 is the bisection method based on Proposition 18 and Proposition 19. Algorithm 6 is the entire line-search strategy that is based on the quadratic/cubic models and makes use of Algorithm 5.

Algorithm 5 (Bisection scheme)

    Input tol, α_l and α_r which bracket a minimum point;
    Input two functions $f(\alpha) = J_r(u + \alpha\,\delta u)$ and $f'(\alpha) = \langle \delta u, \nabla J_r(u + \alpha\,\delta u)\rangle_{L^2}$;
    Compute $f_l = f(\alpha_l)$, $f_r = f(\alpha_r)$, $d_l = f'(\alpha_l)$, $d_r = f'(\alpha_r)$ and $L = \alpha_r - \alpha_l$;
    while $L > tol$ **do**
        Compute the midpoint $\alpha_m = \frac{1}{2}(\alpha_r + \alpha_l)$ and evaluate $f_a = f(\alpha_m)$ and $d_a = f'(\alpha_m)$;
        if ($\alpha = \alpha_m$ satisfies SWC) **then**
            $\alpha \leftarrow \alpha_m$
            return
        end if
        if ($d_l < 0$ **and** [$d_r > 0$ **or** $f_r > f_l$]) **then**
            $\alpha_r \leftarrow \alpha_m$ and $f_r \leftarrow f_a$ and $d_r \leftarrow d_a$;
        else
            if ($d_l > 0$ **and** [$d_r < 0$ **or** $f_r < f_l$]) **then**
                $\alpha_r \leftarrow \alpha_m$ and $f_r \leftarrow f_a$ and $d_r \leftarrow d_a$;
            else
                $\alpha_l \leftarrow \alpha_m$ and $f_l \leftarrow f_a$ and $d_l \leftarrow d_a$;
            end if
        end if
        $L \leftarrow (\alpha_r - \alpha_l)$;
    end while

Algorithm 6 (Line-search scheme based on SWC and quadratic/cubic model)

$\quad$ **Input** $J_r(u)$, $\nabla J_r(u)$, du, u, $\epsilon \ll 1$, $f(\alpha) = J_r(u + \alpha\,\delta u)$ and $f'(\alpha) = \langle \delta u, \nabla J_r(u + \alpha\,\delta u) \rangle_{L^2}$;

if $(\frac{\nu}{2}\|du\|_U^2 > \epsilon)$ **then**

$\quad$ Compute α_{max} with Equation (6.82);

$\quad$ **if** $(\alpha_{max} > 2)$ **then**

$\quad\quad$ Evaluate $f(1)$ and $f'(1)$;

$\quad\quad$ **if** $(\alpha = 1$ satisfies SWC$)$ **then**

$\quad\quad\quad$ $\alpha \leftarrow 1$;

$\quad\quad$ **else**

$\quad\quad\quad$ Construct cubic model on $[0, 1]$, compute its minimum α_m, and evaluate $f(\alpha_m)$ and $f'(\alpha_m)$;

$\quad\quad\quad$ **if** $(\alpha = \alpha_m$ satisfies SWC$)$ **then**

$\quad\quad\quad\quad$ $\alpha \leftarrow \alpha_m$;

$\quad\quad\quad$ **else**

$\quad\quad\quad\quad$ **if** $([0, \alpha_m]$ brackets a minimum$)$ **then**

$\quad\quad\quad\quad\quad$ $\alpha_r \leftarrow \alpha_m$;

$\quad\quad\quad\quad$ **else**

$\quad\quad\quad\quad\quad$ **if** $([0, 1]$ brackets a minimum$)$ **then**

$\quad\quad\quad\quad\quad\quad$ $\alpha_r \leftarrow 1$;

$\quad\quad\quad\quad\quad$ **else**

$\quad\quad\quad\quad\quad\quad$ $\alpha_r \leftarrow \alpha_{max}$;

$\quad\quad\quad\quad\quad$ **end if**

$\quad\quad\quad\quad$ **end if**

$\quad\quad\quad\quad$ $\alpha \leftarrow$ bisection on $(0, \alpha_r)$;

$\quad\quad\quad$ **end if**

$\quad\quad$ **end if**

$\quad$ **else**

$\quad\quad$ $\alpha \leftarrow$ bisection on $(0, \alpha_{max})$;

$\quad$ **end if**

else

$\quad$ $\alpha \leftarrow 1$;

$\quad$ Perform backtracking until the Armijo's condition (6.80) is satisfied;

end if

6.3.2 $\quad$ A globalized Krylov-SSN scheme

The numerical implementation of the theoretical SSN procedure (4.58), that is

$$\begin{aligned}
\text{S}_0: &\quad \text{choose a generalized Jacobian } \mathcal{J}_g(u^k) \in \partial^\circ \mathcal{F}_r(u^k) \\
\text{S}_1: &\quad \delta u^k = -(\mathcal{J}_{g,r}(u^k))^{-1}\mathcal{F}_r(u^k) \\
\text{S}_2: &\quad u^{k+1} = u^k + \delta u^k\ ,
\end{aligned}$$

is more involved than the Newton method described in the Section 6.3.1. This is due to the non-smoothness induced by the control-constraint U_{ad} and the L^1-norm in the cost in the case that $\beta > 0$:

- the first-order optimality system is addressed as a root problem in such a way that can generate a loss of informations regarding the reduced gradient; this is particularly evident in the case that $\beta > 0$, in which the reduced (generalized) gradient is undetermined;

- since the optimality condition is not $\nabla_u J_r(u) = 0$ in the case of an active control-constraint, an adequate stop-criterion has to be considered, and possibly it must take into account also the active set $\mathcal{A}$, that is the subset of the domain $[0, T]$ in which the control-constraint is active;

- because of the control-constraint and the loss of differentiability due to the L^1-norm in the cost, a standard line-search strategy could be not applicable; moreover, due to the active set, it could be convenient to have a globalization scheme that adapts the line-search criterion to the current situation during the iterative Newton procedure;

- the (generalized) Jacobian appearing in the Newton linear system does not coincide with the Hessian of the Lagrange function and could be not symmetric; this fact restricts the choice of the Krylov solver.

Because of these reasons, even if the global structure of the Krylov-SSN scheme is similar to the Krylov-Newton scheme described in Section 6.3.1, many implementation details have to be treated differently. We remark that it could be necessary to initialize the SSN method by using continuation techniques; see, e.g., [129, 130]. In the sequel, we describe the Krylov-SSN scheme for the case $U_{ad} = U_{ad,1}$ and $\beta > 0$. This scheme results to be the same also in the case $U_{ad} = U_{ad,2}$ and $\beta = 0$.

The Krylov-SSN procedure, corresponding to (4.58), is given in the following algorithms 7-12.

Algorithm 7 (Semi-smooth Krylov-Newton scheme)

Input u, $k = 0$, k_{max}, tol;
Call Algorithm 8 to compute $\mathcal{F}_r(u)$ and $\mathcal{A}$;
Set $\mathcal{A}_{old} \leftarrow \mathcal{A}$;
while $k < k_{max}$ **do**

- call Algorithm 9 to compute the search direction δu;

- depending on $\mathcal{A}$, compute α choosing an appropriate line-search;

- update: $u \leftarrow u + \alpha\, \delta u$;

- call Algorithm 8 to compute $\mathcal{F}_r(u)$ and $\mathcal{A}$;

- if a stop criterion (depending on $\mathcal{A}$, $\mathcal{A}_{old}$ and tol) is satisfied, break;

- update: $k \leftarrow k + 1$ and $\mathcal{A}_{old} \leftarrow \mathcal{A}$;

end while

It is proved in [65] that an adequate stopping criteria for a semi-smooth Newton method for the solution of optimal control problems governed by linear elliptic equations is related to the active set. For this purpose, define the following

$$\mathcal{A} := \mathcal{A}_1 \times \cdots \times \mathcal{A}_{N_C} \quad \text{with} \quad \mathcal{A}_n := \left\{ t \in (0, T) \mid |u_n(t)| \geq b \right\}, \qquad (6.83)$$

where $\mathcal{A} \neq \emptyset$ and $\mathcal{A}_{old}$ represent the active sets corresponding to the current and the previous iteration, respectively. If it holds that $\mathcal{A} = \mathcal{A}_{old}$, then a solution is found, and the semi-smooth Newton iterative procedure can be stopped.

This result is proved for optimal control problems where the control appears linearly in the constraint. Moreover, in [65] it is observed that the numerical iterations terminate after few steps with two consecutive active sets being equal. For these reasons, it is not guaranteed that this stopping criterion is always valid for our bilinear control problem. Consequently, recalling that we want to obtain a root u of $\mathcal{F}_r(u)$, we implement a stopping criterion that combines the residual of the L^2-norm of $\mathcal{F}_r(u)$ with the active-set criterion. In particular, we have the following criterion

- if ($\mathcal{A} = \mathcal{A}_{old}$ & $\|\mathcal{F}_r(u)\|_{L^2} < tol$) then break.

Notice that in the case $\mathcal{A} = \mathcal{A}_{old} = \emptyset$ with $\beta = 0$, this stop-criterion coincides with $\nabla_u J_r(u) = 0$.

The following Algorithm 8 solves the forward and backward problems and assembles $\mathcal{F}_r(u)$ and $\mathcal{A}$.

Algorithm 8 (Evaluation of $\mathcal{F}_r(u)$ and $\mathcal{A}$)

Input u;

- solve the forward problem (4.61a) and get x;

- solve the backward problem (4.61b) and get p;

- assemble $\mathcal{F}_r(u)$ using (4.63b);

- construct $\mathcal{A}$ using (6.83);

The following Algorithm 9 is used to solve the linear Newton system. This is obtained by means of a Krylov-iterative solver, e.g., GMRES and CG.

Algorithm 9 (Solve the Newton equation)

Input u, an initial guess for δu;
- compute δu by solving $\mathcal{J}_g(u)(\delta u) = -\mathcal{F}_r(u)$ (use a linear system solver, e.g., GMRES or CG calling Algorithm 10 to apply $\mathcal{J}_r(u)$);

- if $\mathcal{A} = \emptyset$, $\beta = 0$ and $\langle \mathcal{J}_{g,r}(u), \delta u \rangle_{L^2} > 0$, set $\delta u \leftarrow -\delta u$;

The following Algorithm 10 applies the reduced generalized Jacobian $\mathcal{J}_{g,r}(u)$ on the vector function δu.

Algorithm 10 (Apply the reduced generalized Jacobian to a vector δu)

Input x, p, u, δu;
- integrate the linearized forward problem (4.76) and get δx;

- integrate the linearized backward problem (4.77) and get δp;

- assemble $\mathcal{J}_{g,r}(u)(\delta u)$ by means of (4.79);

An important issue for the successful application of our Newton scheme, is the use of a robust line-search method for globalization purposes. If $\beta = 0$ and $\mathcal{A} = \emptyset$, the classical Armijo-Wolfe and Goldstein conditions can be used; see, e.g., [16, 52, 128, 129]. Otherwise, if $\beta \neq 0$ and/or $\mathcal{A} = \emptyset$, we do not have differentiability (in the classical sense) of the minimizing function and an alternative approach must be followed. In particular, one can apply a line-search procedure to the function $\vartheta : L^2((0,T); \mathbb{R}^{N_C}) \to \mathbb{R}$ defined as follows

$$\vartheta(u) := \frac{1}{2} \|\mathcal{F}_r(u)\|_{L^2}^2 \ .$$

For a detailed discussion on ϑ, see, e.g., [38, 98].

For this reasons, and according to our experience, an efficient globalization method is achieved with the following algorithm.

Algorithm 11 (Globalization: compute the step-length α along δu)

Input β, $\mathcal{A}_{old}$, δu;
- if $\beta = 0$ and $\mathcal{A}_{old} = \emptyset$ apply an Armijo-Wolfe based line-search to minimize J along δu (e.g., Algorithm 6) and break;

- if $\beta > 0$ and $\mathcal{A}_{old} = \emptyset$ apply the line-search Algorithm 12 to minimize J along δu and break;

- if $\mathcal{A}_{old} \neq \emptyset$ apply the line-search Algorithm 12 to minimize ϑ along δu and break;

Notice that, whenever classical differentiability is not guaranteed, we use the derivative-free line-search procedure given in Algorithm 12 [52], where the choice of the minimizing function f is performed by Algorithm 11. Furthermore, since Algorithm 12 does not make use of information about the derivative of f, it is not possible to establish a priori if δu is a descent direction [52]. Hence, it could be useful to use line-search criteria that consider also negative values of α. This explains Step 2. in Algorithm 12.

Algorithm 12 (Derivative-free line-search scheme)

Input δu, $\gamma > 0$, $\rho \in (0, 1)$ $\alpha_{min} > 0$;
- compute $f(u)$;

- set $\alpha \leftarrow 1$;

while $\alpha > \alpha_{min}$ **do**
1. if $f(u + \alpha \delta u) \leq f(u) - \gamma \alpha^2 \|\delta u\|_{L^2}^2$, break;

2. if $f(u - \alpha \delta u) \leq f(u) - \gamma \alpha^2 \|\delta u\|_{L^2}^2$, set $\alpha \leftarrow -\alpha$ and break;

3. set $\alpha \leftarrow \rho \alpha$;

end while

6.4   Continuation scheme for the exact-control of quantum spin systems

In this section, we implement our continuation method presented in Section 5.2 to compute exact-control functions.

Consider a weight parameter sequence given by $\nu_{k+1} = \gamma \nu_k$, with given ν_1 and $\gamma \in (0, 1)$. As proved above, corresponding to this sequence, a sequence of optimal solutions $\{(x_k, u_k)\}_k$ to (P_ν) is obtained that converges to the solution of $(\tilde{P})$.

A feature of the continuation method is that the solution at the step k is computed considering as a starting guess the solution obtained at the step $k - 1$. The detailed continuation scheme is described by the following algorithm.

Algorithm 13 (Continuation scheme)

Input u^0, $k = 1$, k_{max}, $\gamma \in (0, 1)$, an initial value of ν;
while $k < k_{max}$ **do**
- call Algorithm 1 (or Algorithm 7) to compute $u^k(\nu)$ (using u^{k-1} as starting condition);

- set $\nu \leftarrow \gamma \nu$;

- set $k \leftarrow k + 1$;

- if a stop criterion is satisfied, break;

end while

6.5 Optimization schemes for the shooting-type method

In this section, we present a numerical scheme which is specific for the formulation (5.36), that is

$$\min_{u,p_T} \quad G_r(u, p_T) := G(x(u), u, p(u, p_T))$$

$$\text{s.t.} \quad (x(u), p(u, p_T)) \in \left\{ (x, p) \mid x \text{ solves (3.53a) and } p \text{ solves (3.53b)} \right\}.$$

The numerical scheme makes use of a cascadic NCG method [16, 60] as an initialization procedure for a Krylov-Newton method. For completeness, we give all details regarding these procedures. See [15, 73, 116] for previous works on the use of NCG schemes to solve quantum control problems. We refer to [15, 36, 60] and references therein for details about the convergence of this method.

The iterative NCG procedure to solve (5.36) is given by the following algorithm.

Algorithm 14 (NCG scheme)

Require: u^0, p_T^0, $k = 0$, k_{max}, tol;

Call Algorithm 2 to compute $\nabla G_r(u^0, p_T^0)$;

Set $d^0 = -\nabla G_r(u^0, p_T^0)$;

while $k < k_{max}$ and $|||\nabla G_r(u^k, p_T^k)||| > tol$ **do**

- call Algorithm 16 to compute α along the direction d^k;

- set $(u^{k+1}, p_T^{k+1}) = (u^k, p_T^k) + \alpha\, d^k$;

- call Algorithm 15 to compute $\nabla G_r(u^{k+1}, p_T^{k+1})$;

- compute $y^k = \nabla G_r(u^{k+1}, p_T^{k+1}) - \nabla G_r(u^k, p_T^k)$;

- compute $\sigma^{k+1} = y^k - 2d^k \frac{(y^k, y^k)_G}{(d^k, y^k)_G}$;

- compute $\beta^{k+1} = \frac{(\nabla G_r(u^{k+1}, p_T^{k+1}), \sigma^{k+1})_G}{(d^k, y^k)_G}$;

- set $d^{k+1} = -\nabla G_r(u^{k+1}, p_T^{k+1}) + \beta^{k+1} d^k$;

- set $k \leftarrow k + 1$;

end while

In Algorithm 14, given (u, p_T), the gradient ∇G_r is obtained using the following algorithm.

Algorithm 15 (Evaluation of the gradient)

Require: u, p_T;

- integrate the constraint (5.37c) forward;

- integrate the constraint (5.37d) backward;

- integrate the adjoint (5.37e) backward;

- integrate the adjoint (5.37f) forward;

- assemble $\nabla_u G_r(u, p_T)$ using (5.37a);

- assemble $\nabla_{p_T} G_r(u, p_T)$ using (5.37b);

We implement a line-search strategy based on the Armijo's condition, see, e.g., [94,

---

52, 130], that is we use a step-length α that satisfies

$$G_r((u, p_T) + \alpha d) \leq G_r(u, p_T) + c_1 \alpha (d, \nabla G_r(u, p_T))_G . \tag{6.84}$$

More precisely, we implement a backtracking strategy, as shown in the next algorithm.

Algorithm 16 (Backtracking line-search scheme with Armijo's condition)

Input $G_r(u, p_T)$, $\nabla G_r(u, p_T)$, d, u, $k = 0$, k_{max}, $\gamma \in (0, 1)$, $c_1 \in (0, 1)$;
Set $\alpha = 1$;
while $k < k_{max}$ and $G_r((u, p_T) + \alpha d) > G_r(u, p_T) + c_1 \alpha (d, \nabla G_r(u, p_T))_G$ **do**

- evaluate $G_r((u, p_T) + \alpha d)$;

- if (6.84) is satisfied, then break;

- set $\alpha \leftarrow \gamma \alpha$;

- set $k \leftarrow k + 1$;

end while

According to our experience, Algorithm 14 shows a slow convergence in solving problem (5.36). In order to accelerate it, we embed it in the cascadic scheme. For a detailed discussion about this method see, e.g., [15, 16]. The NCG-cascadic procedure is given in the following algorithm.

Algorithm 17 (Cascadic scheme)

Require: u^0, p_T^0, $k = 1$, k_{max};
Require: Coarse space discretization grid;
Call Algorithm 14 to solve the problem and obtain u^1 and p_T^1;
while $k < k_{max}$ **do**

- refine the discretization grid;

- obtain a guess solution u^{k+1}, by interpolating u^k on the new grid;

- call Algorithm 14 to solve the problem and obtain u^{k+1} and p_T^{k+1};

- set $k \leftarrow k + 1$;

end while

We use the NCG-cascadic scheme to perform an adequate initialization for a Krylov-Newton method. In order to define a matrix-free procedure, we consider the reduced problem (5.36) with $x = x(u)$ and $p = p(u, p_T)$. Consequently, the Newton procedure consists, at a given step k, in solving

$$\begin{aligned}
\nabla^2 G_r(u^k, p_T^k)(d^k) &= -\nabla G_r(u^k, p_T^k) \\
(u^{k+1}, p_T^{k+1}) &= (u^k, p_T^k) + d^k .
\end{aligned} \tag{6.85}$$

A globalized implementation of this procedure is given by the following algorithm.

Algorithm 18 (Krylov-Newton scheme)

Require: u^0, p_T^0, $k = 0$, k_{max}, tol;
  while $k < k_{max}$ and $|||\nabla G_r(u^k, p_T^k)||| > tol$ **do**

- call Algorithm 2 to compute $\nabla G_r(u^{k+1}, p_T^{k+1})$;

- call Algorithm 19 to solve $\nabla^2 G_r(u^k, p_T^k)(d^k) = -\nabla G_r(u^k, p_T^k)$;

- call Algorithm 16 to compute α along the direction d^k;

- set $(u^{k+1}, p_T^{k+1}) = (u^k, p_T^k) + \alpha \, d^k$;

- set $k \leftarrow k + 1$;

  end while

The following algorithm is used to solve the Newton linear system.

Algorithm 19 (Solve the Newton linear system)

Input u, p_T, $\nabla G_r(u, p_T)$;
- Guess an initial value of d;

- Compute d by solving $\nabla^2 G_r(u, p_T)(d) = -\nabla G_r(u, p_T)$: use a Krylov-based linear system solver, e.g., GMRES or CG, calling Algorithm 20 to apply the reduced Hessian;

- If d is an ascending direction, then set $d \leftarrow -d$;

The action of the reduced Hessian can be evaluated by the following algorithm.

Algorithm 20 (Action of the reduced Hessian)

Require: $d = (\delta u, \delta p_T)$;
- integrate the linearized constraint (5.43) forward;

- integrate the linearized constraint (5.44) backward;

- integrate the linearized adjoint (5.40) backward;

- integrate the linearized adjoint (5.42) forward;

- assemble $\nabla^2 G_r(u, p_T)(d)$ by using (5.41);

6.6   Summary and remarks

This chapter provided Chapter 3, Chapter 4 and Chapter 5 with a numerical framework. Specifically, it focused on numerical and implementation details regarding optimal control and exact-control of bilinear quantum systems. In particular, the modified Crank-Nicolson scheme was analysed in several numerical and algebraic aspects, and the first-discretize-than-optimize strategy is used to obtain adequate discrete optimality conditions. Furthermore, implementation details, regarding the algorithms used to perform numerically the methodologies presented in Chapter 3, Chapter 4 and Chapter 5, were discussed. The main novelties of this chapter are:

- an algebraic analysis, based on the matrix Lie group theory, of the modified Crank-Nicolson scheme;

- implementation details of the semi-smooth Newton method for the optimal control of bilinear quantum system with an L^1-penalization in the cost functional.

Chapter 7

Numerical experiments

This chapter aims to present numerical experiments that demonstrate the validity of the computational framework developed in this thesis. In Section 7.1, L^1-penalized and piecewise-constant control problems are considered. These are used to show the ability of the computational methods developed in Chapter 3, Chapter 4 and Chapter 6. In particular, we solve L^1- and piecewise-constant optimal control problems governed by the LvNM equation and dipole-control problems, where the governing model is an infinite-dimensional Schrödinger equation. Convergence results of the SSN method are also discussed in this section. In Section 7.2, we consider experiments that perform the exact-control of the LvNM and Pauli equations. Several test cases are considered, and the two computational schemes developed in Chapter 5 are used and compared. In Section 7.3, the continuation procedure is used to solve control problems of inhomogeneous spin systems.

We remark that, for all the experiments, the quantum bilinear dynamical system is discretized by means of the modified Crank-Nicholson method, and the discrete approximation of the overall optimality system is obtained from the first-discretize-than-optimize strategy, which allows to obtain the discrete approximation of the corresponding backward adjoint equation and the gradient of the differentiable part of the cost functional J. The time interval $[0, T]$ is discretized with a uniform mesh of size $h = \frac{T}{N_t - 1}$ and N_t points, such that $0 = t^1 < \cdots < t^{N_t} = T$.

7.1 L^1-penalized problems and piecewise-constant controls

In this section, we present results of numerical experiments that demonstrate the efficiency of the methods developed in Chapter 5. For this purpose, we consider the following three test-cases:

- Case 1: Control of a system of two uncoupled spin-$\frac{1}{2}$ particles.
- Case 2: Control of a system of two coupled spin-$\frac{1}{2}$ particles.
- Case 3: Dipole quantum control of a charged particle.

In the analysis of these test cases, we make use of the following variables. First, we consider the so-called fidelity given by the following

$$C := \frac{\langle x(T), x_T \rangle}{\|x(T)\|_2 \|x_T\|_2} \ . \tag{7.1}$$

This quantity measures the projection of the terminal state $x(T)$ on the target state x_T, that is largely used in physics. Notice that $C = 1$ means that $x(T) = x_T$, that is, the trajectory reaches exactly the target. Second, in order to validate numerically the estimate (3.141) in Theorem 10, we define

$$\hat{K}(\beta, \nu) := \frac{2N_C T (\hat{\beta} - \beta)}{\nu} \ .$$

Consequently, (3.141) reads as follows

$$\|u\|_{L^1} \le \hat{K}(\beta, \nu) \ .$$

The three test-cases are discussed in the following sections.

7.1.1 Case 1: Control of two uncoupled spin-$\frac{1}{2}$ particles

Consider a system of two uncoupled spin-$\frac{1}{2}$ particles. The Hamiltonian operator of each particle is given by

$$H = \hat{\nu} I_z + u I_x \ ,$$

where $\hat{\nu}$ is the Larmor-frequency, u is the control and I_x and I_z are the Pauli matrices; see, e.g., Chapter 2.1 and [25] and references therein. The state of such a spin-system is represented by a density operator and the corresponding dynamics is governed by the Liouville-von Neumann master equation, that can be expressed, as in Section 2.3.4, in the following real matrix representation

$$\dot{x} = 2\pi \big(\tilde{A} + u\tilde{B} \big) x \ , \tag{7.2}$$

where

$$\tilde{A} = c \begin{pmatrix} A & 0 \\ 0 & -A \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} ,$$

with $c = 483$, and

$$\tilde{B} = \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix} \quad \text{with} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} .$$

We seek a control $u : [0, T] \to \mathbb{R}$, in $U_{ad,1}$, such that $|u(t)| \le 60$, that is capable to perform a transition from an initial state where both the spins are pointing in the z-direction, to a target state where both the spins are pointing in the y-direction. Consequently, starting and target vectors are

$$x_0 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} ,$$

and

$$x_T = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \end{pmatrix} .$$

For more details regarding spin-systems see [6, 25, 111] and references therein. We consider $T = 0.008$, which makes the system similar to the one considered in [9].

Within these settings, we solve (3.107) with different values of the parameters ν and β. In particular, we consider $\nu = 2^{-k}$ with $k = 0, ..., 9$ and $\beta \in \{0.1, 1, 2, 3, 4, 5, 6, 7\}$ and apply the semi-smooth Newton method described in Chapter 4. Moreover, the obtained solutions are compared with the piecewise constant optimal control solutions to (3.97),

corresponding to the same weight parameters ν and to different numbers of subintervals $M \in \{25, 50, 100\}$. We remark that, even if the number of subintervals varies, the overall time interval $[0, T]$ is discretized with the same number of mesh-points $N_t = 1001$.

Regarding the convergence criteria of our algorithms, we consider the following. The tolerances for the convergence of the semi-smooth Newton and of the Krylov linear solver are 10^{-7} and 10^{-8}, respectively. Moreover, we allow a maximum number of iterations equal to 100 for the semi-smooth Newton procedure and equal to NN_C for the Krylov solver. We remark that, in general, the desired tolerances are reached in a number of iterations that is lower than the allowed one.

In the following Figure 7.1, we show the value of the fidelity C resulting from the optimization procedure. It is evident that, the fidelity C increase as ν decreases. In particular, the solutions to (3.97) results in almost the same values of C, independently on the number constant steps M, and we are able to obtain very high values of the fidelity for $\nu \leq 2^{-7}$. On the other hand, we show the values of C resulting from the solutions to (3.107), and we observe the following. The fidelities corresponding to $\beta = 0.1$ are comparable to the ones obtained with piecewise constant solutions. Notice also that, C decreases when β increases, and its decay becomes faster for $\beta > 4$. Furthermore, we remark that the values of C corresponding to $\beta = 7$ are all zeros. This fact validates numerically the statement of Theorem 9.
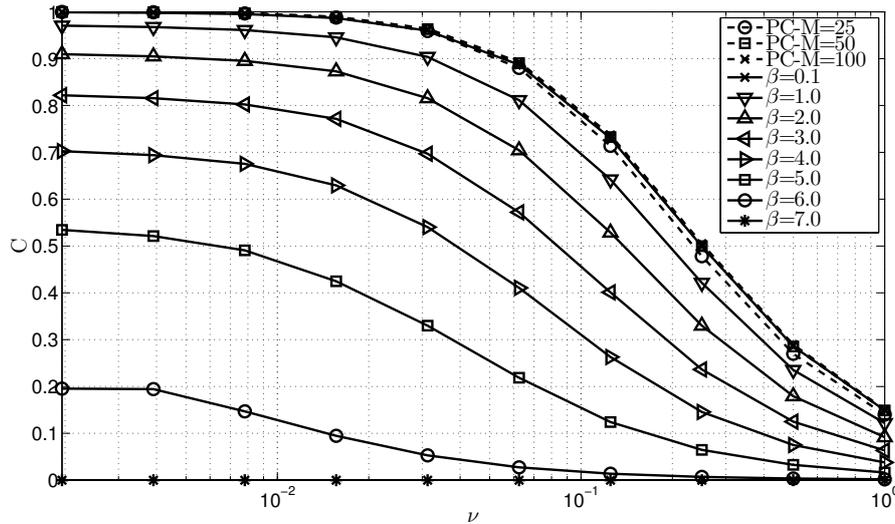

Figure 7.1: Values of the fidelity C obtained by solving problems (3.107) and (3.97) and corresponding to different values of penalization parameters ν and β.

Next, we want to demonstrate numerically the validity of Theorem 10. First, notice that for the considered system it holds that $\max_{n} \langle B_n x_0, x_T \rangle = 0$, which means that $u = 0$ is a candidate to be a stationary point of (3.107). By means of the Tables 7.1 and 7.2, we demonstrate the validity of the estimate (3.141). In particular, we show the values of the bound $\hat{K}(\beta, \nu)$ and the L^1-norm of the optimal control corresponding to $\beta = 0.1$ and $\beta = 3$. These results validate the estimate (3.141). We obtain similar results for the other considered values of β, and we omit them for brevity.

ν	$\|u\|_{L^1}$	$\hat{K}(\beta,\nu)$
2^0	0.0304	0.1104
2^{-1}	0.0589	0.2208
2^{-2}	0.1061	0.4416
2^{-3}	0.1660	0.8832
2^{-4}	0.2204	1.7664
2^{-5}	0.2581	3.5328
2^{-6}	0.2790	7.0656
2^{-7}	0.2878	14.1312
2^{-8}	0.2901	28.2624
2^{-9}	0.2896	56.5248

Table 7.1: Values of the constant $\hat{K}(\beta,\nu)$ and of the L^1-norm of the optimal controls corresponding to $\beta = 0.1$.

ν	$\|u\|_{L^1}$	$\hat{K}(\beta,\nu)$
2^0	0.0113	0.0640
2^{-1}	0.0224	0.1280
2^{-2}	0.0425	0.2560
2^{-3}	0.0729	0.5120
2^{-4}	0.1060	1.0240
2^{-5}	0.1317	2.0480
2^{-6}	0.1479	4.0960
2^{-7}	0.1557	8.1920
2^{-8}	0.1595	16.3840
2^{-9}	0.1615	32.7680

Table 7.2: Values of the constant $\hat{K}(\beta,\nu)$ and of the L^1-norm of the optimal controls corresponding to $\beta = 3$.

Next, we demonstrate that the semi-smooth Newton methods presented in Chapter 4 and Chapter 6 are locally superlinear convergent. For this purpose, we show in the next tables the decay of the residual norm $\|\mathcal{F}_r(u)\|_{L^2}$ with respect to the number of the iterations. Moreover, we show also the number of active points and the value of the step-length α computed at each iteration.

In Table 7.3, we show the convergence of the semi-smooth Newton method for solving problem (3.107) with $\beta = 3$ and $\nu = 2^{-5}$. It is evident that in the first 10 iterations the line-search algorithm performs a globalization process and generates values of α lower than 1. Then, the step-length $\alpha = 1$ satisfies line-search conditions and the method converges quadratically to the solution. We remark that, between iterations 5 and 11 the number of active points does not change, but the current approximation of the control is still far from the desired solution.

In Tables 7.4 and 7.5, we present convergence results of the semi-smooth Newton method for the solution to (3.97) with $\nu = 2^{-5}$ and with $\nu = 2^{-8}$, respectively. Corresponding to $\nu = 2^{-5}$, the active set $\mathcal{A}$ defined in (6.83) remains empty during the iterations, and the algorithm converges quadratically to the solution. On the other hand, corresponding to $\nu = 2^{-8}$ the active set $\mathcal{A}$ is non-empty, the line-search algorithm performs a globalization, and after that the rate of convergence is quadratic. We obtain similar results for the other considered values of penalization parameters, and we omit them for brevity.

iter.	$N°$ active p.	$\|\mathcal{F}_r(u)\|_{L^2}$	α
1	0	$5.20{\cdot}10^{-2}$	0.250
2	0	$6.46{\cdot}10^{-2}$	0.500
3	4	$1.60{\cdot}10^{-1}$	0.250
4	96	$6.21{\cdot}10^{-2}$	1.000
5	107	$3.45{\cdot}10^{-2}$	1.000
6	107	$7.24{\cdot}10^{-3}$	0.125
7	107	$3.51{\cdot}10^{-2}$	0.500
8	107	$2.71{\cdot}10^{-3}$	0.250
9	107	$2.15{\cdot}10^{-3}$	0.250
10	107	$2.14{\cdot}10^{-3}$	0.063
11	107	$3.83{\cdot}10^{-5}$	1.000
12	106	$7.65{\cdot}10^{-11}$	1.000
13	106	$4.72{\cdot}10^{-16}$	1.000

Table 7.3: Convergence of the semi-smooth Newton for the solution of (3.107) with $\beta = 3$ and $\nu = 2^{-5}$.

iter.	$N°$ active p.	$\|\mathcal{F}_r(u)\|_{L^2}$	α
1	0	$3.26{\cdot}10^{-3}$	1.000
2	0	$2.64{\cdot}10^{-6}$	1.000
3	0	$1.47{\cdot}10^{-10}$	1.000

Table 7.4: Convergence of the semi-smooth Newton for the solution of (3.97) with $\nu = 2^{-5}$ and $\mathcal{A} = \emptyset$.

iter.	$N°$ active p.	$\|\mathcal{F}_r(u)\|_{L^2}$	α
1	80	$7.65{\cdot}10^{-3}$	0.500
2	90	$4.78{\cdot}10^{-3}$	0.500
3	100	$7.18{\cdot}10^{-7}$	1.000
4	100	$1.13{\cdot}10^{-11}$	1.000

Table 7.5: Convergence of the semi-smooth Newton for the solution of (3.97) with $\nu = 2^{-8}$ and $\mathcal{A} \neq \emptyset$.

Next, we show some of the obtained optimal controls. In Figure 7.2, the optimal control solutions to (3.107) corresponding to $\beta = 0.1, 3, 6$ and different values of ν are depicted. We notice that, the amplitude of the optimal controls increases as ν decreases and the control constraint ($b = 60$) is active for $\nu \leq 2^{-7}$. From the shapes of the control, the behaviour described in Theorem 8 is evident. In particular, the sparsity of the control functions increases as ν decreases and the L^1-component of the cost starts to be dominating. Figure 7.3 shows the optimal controls obtained solving (3.107) with $\nu = 2^{-9}$ and different values of β. We observe that when β increases, the controls are more sparse, and the solution u vanishes for $\beta = 7$, in accordance with Theorem 9. Furthermore, the shape of the obtained controls varies (locally) continuously with β, demonstrating the validity of Theorem 11. In Figure 7.4, we present the piecewise constant optimal controls obtained solving (3.97). The amplitude of the controls increases as ν decreases and the control-constraint is active for $\nu \leq 2^{-7}$.
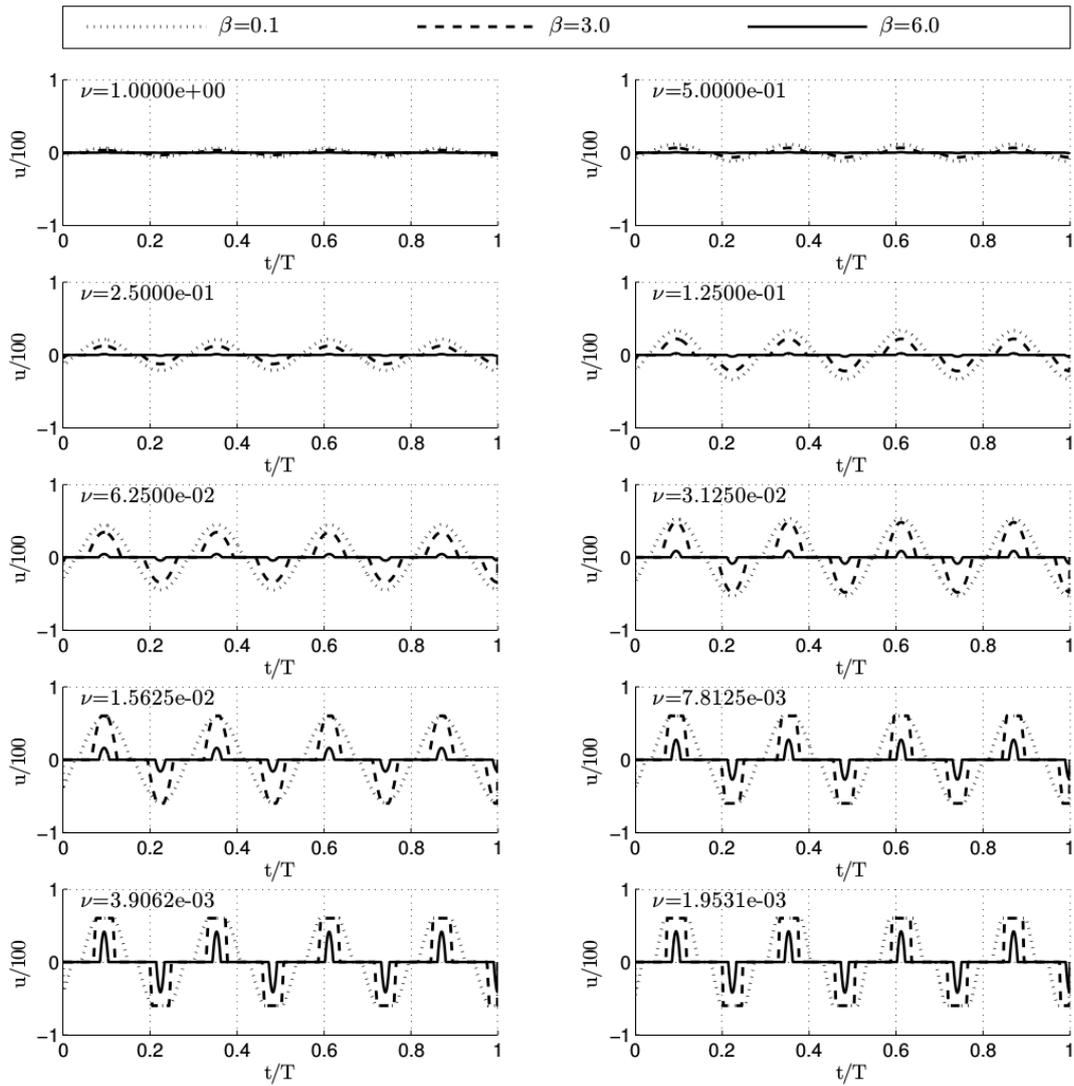
Figure 7.2: Optimal control solutions to (3.107) corresponding to different values of β and $\nu = 2^{-k}$ with $k = 0, 1, ..., 9$.
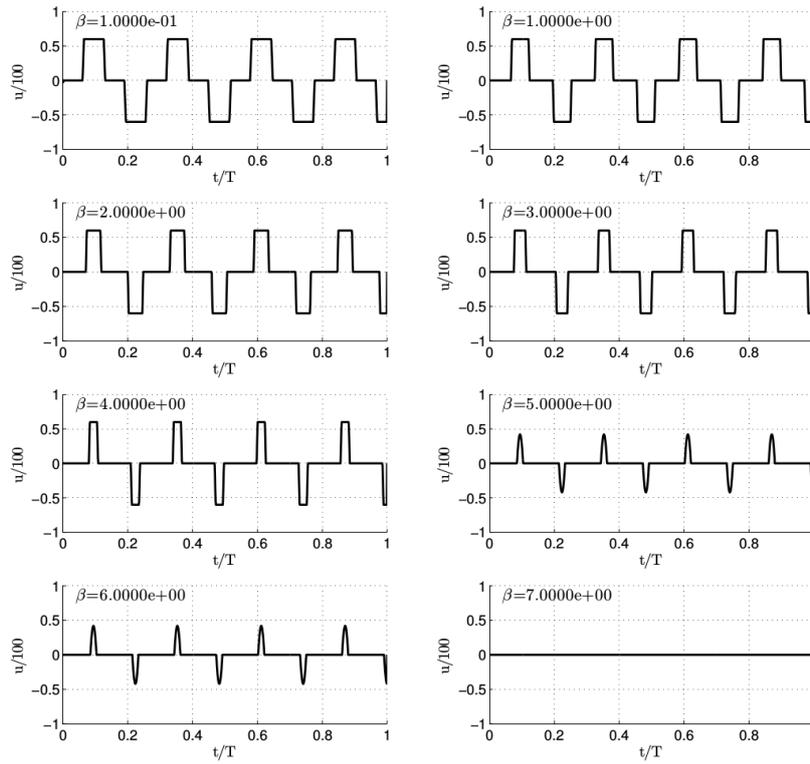
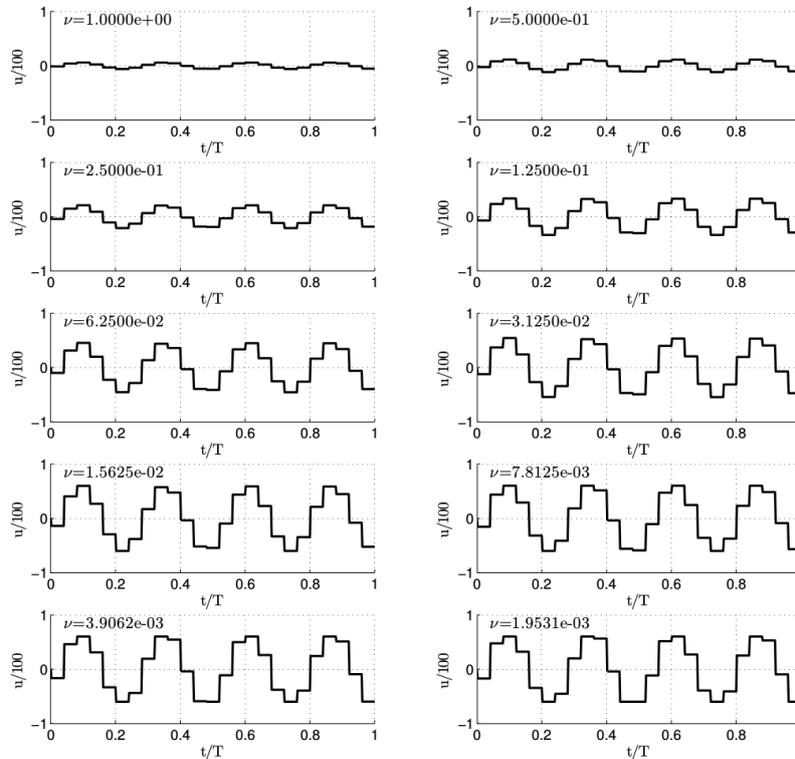Figure 7.3: Optimal control solutions to (3.107) corresponding to $\nu = 2^{-9}$ and different values of β.

Figure 7.4: Optimal control solutions to (3.97) with piecewise constant controls corresponding to $M = 25$ and $\nu = 2^{-k}$ with $k = 0, 1, ..., 9$.

7.1.2 Case 2: Control of two coupled spin-$\frac{1}{2}$ particles

Consider a system of two coupled spin-$\frac{1}{2}$ particles. As described in Section 2.3.3, the Hamiltonian of this system is given by

$$H = \hat{\nu}_1 \hat{I}_z + \hat{\nu}_2 \hat{S}_z + J_c \hat{I}_z \hat{S}_z + u_1 \hat{I}_x + u_2 \hat{I}_y + u_3 \hat{S}_x + u_4 \hat{S}_y \,,$$

where u_n with $n = 1, \dots, 4$ are the controls, $\hat{I}_\alpha = I_2 \otimes I_\alpha$, $\hat{S}_\alpha = I_\alpha \otimes I_2$, $\alpha = x, y, z$, I_α are the Pauli matrices, $\hat{\nu}_1$ and $\hat{\nu}_2$ are the Larmor-frequencies of the two particles, and J_c is the coupling constant. For details regarding this spin-model and its real matrix representation see [6, 25, 111] and references therein.

The state of this spin-system is represented by a density operator and the corresponding dynamics is governed by the Liouville-von Neumann master equation, that can be expressed in the following real matrix representation

$$\dot{x} = \left[A + \sum_{n=1}^{4} u_n B_n \right] x \,,$$

where A and B are skew-symmetric matrices in $\mathbb{R}^{16 \times 16}$. The starting and target vectors are given by

$$x(0) = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \end{pmatrix}^T$$
$$x_T = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T \,,$$

which physically correspond to a transition of the spin orientation of both the particles from the z direction to the x direction. We assume that $J_c = 1$, $\hat{\nu}_1 = 1.2$ and $\hat{\nu}_2 = 0.8$. Moreover, we consider $T = 2$ and a bound on the controls $b = 2$.

We consider the same settings as in the Case 1 for discretization and tolerances, and we perform several optimizations corresponding to different values of penalization parameters. The obtained results are shown in the next figure and tables and they lead to the same considerations as in Case 1.
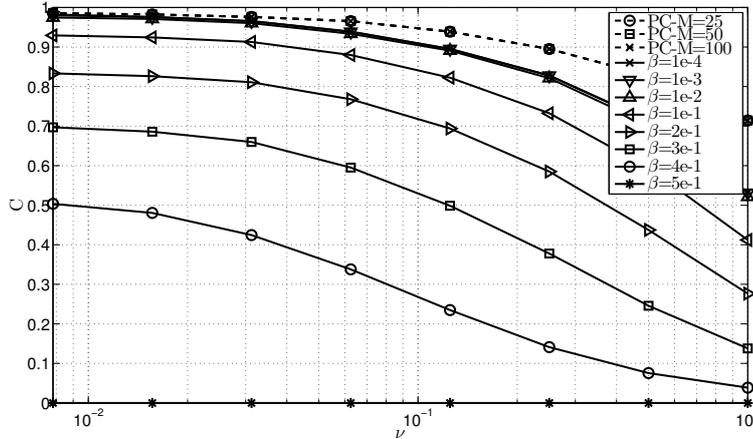

Figure 7.5: Values of the fidelity C obtained by solving problems (3.107) and (3.97) and corresponding to different values of penalization parameters ν and β.

ν	$\|u\|_{L^1}$	$\hat{K}(\beta,\nu)$
2^0	1.6806	7.9984
2^{-1}	2.4279	15.9968
2^{-2}	3.0223	31.9936
2^{-3}	3.4011	63.9872
2^{-4}	3.7877	127.9744
2^{-5}	4.1773	255.9488
2^{-6}	4.4715	511.8976
2^{-7}	5.1609	1023.7952

Table 7.6: Values of the constant $\hat{K}(\beta,\nu)$ and of the L^1-norm of the optimal controls corresponding to $\beta = 1\mathrm{e}{-4}$.

ν	$\|u\|_{L^1}$	$\hat{K}(\beta,\nu)$
2^0	1.1584	6.4000
2^{-1}	1.7193	12.8000
2^{-2}	2.1646	25.6000
2^{-3}	2.4469	51.2000
2^{-4}	2.6000	102.4000
2^{-5}	2.6801	204.8000
2^{-6}	2.7361	409.6000
2^{-7}	2.7698	819.2000

Table 7.7: Values of the constant $\hat{K}(\beta,\nu)$ and of the L^1-norm of the optimal controls corresponding to $\beta = 1\mathrm{e}{-1}$.

7.1.3 Case 3: Dipole quantum control

In this section, we consider the case of dipole-control of a charged particle in an infinite-potential well [129, 130]. This is an infinite-dimensional quantum control problem. However, as explained in Section 2.4.2, it is natural to apply our methodology to the discretized infinite-dimensional problem.

We consider the following optimal control problem

$$\min_{x,u} \quad J(\psi, u) := \frac{1}{2}\|\psi(\cdot, T) - \psi_T\|_{L^2(\Omega;\mathbb{C})}^2 + \frac{\nu}{2}\|u\|_{L^2}^2 + \beta\|u\|_{L^1}$$

$$\text{s.t.} \quad i\frac{\partial\psi}{\partial t} = \left[-\Delta + uX\right]\psi \quad \text{in} \quad \Omega = (0, L) \quad \text{for} \quad t \in (0, T]$$

$$\psi = 0 \quad \text{on} \quad \partial\Omega \quad , \quad \psi(\cdot, 0) = \psi_0$$

$$u \in U_{ad}$$

We consider $L = 2$, $T = 1$ and seek a control u in the admissible set $U_{ad} = U_{ad,1}$, defined in (3.2) with $N_C = 1$, such that transitions from the ground-state (first eigenstate of the free Hamiltonian) to the second (eigen-)state is performed; see [129].

In the Tables 7.8 and 7.9, in order to validate Theorem 10, we show the values of the bound $\hat{K}(\beta,\nu)$ and the L^1-norm of the optimal control corresponding to $\beta = 0.05$ and $\beta = 0.10$. Figure 7.6 shows the sparse-structure of the computed optimal controls. Notice that, the sparsity increases as β increases. In Figure 7.7, we show the piecewise-constant optimal control corresponding to $\beta = 0$ and $M = 25$, $M = 50$ and $M = 100$ subintervals. Figure 7.8 shows the fidelity corresponding to different values of the parameters ν and β.

ν	$\|u\|_{L^1}$	$\hat{K}(\beta,\nu)$
2^0	0.2093	0.9000
2^{-1}	0.3226	1.8000
2^{-2}	0.6315	3.6000
2^{-3}	1.1751	7.2000
2^{-4}	1.9605	14.4000
2^{-5}	2.8934	28.8000
2^{-6}	3.6427	57.6000
2^{-7}	4.0521	115.2000
2^{-8}	4.2543	230.4000
2^{-9}	4.3676	460.8000

Table 7.8: Values of the constant $\hat{K}(\beta,\nu)$ and of the L^1-norm of the optimal controls corresponding to $\beta = 0.05$.

ν	$\|u\|_{L^1}$	$\hat{K}(\beta,\nu)$
2^0	0.2093	0.8000
2^{-1}	0.2407	1.6000
2^{-2}	0.4732	3.2000
2^{-3}	0.8911	6.4000
2^{-4}	1.5106	12.8000
2^{-5}	2.1848	25.6000
2^{-6}	2.8959	51.2000
2^{-7}	3.2774	102.4000
2^{-8}	3.4483	204.8000
2^{-9}	3.5455	409.6000

Table 7.9: Values of the constant $\hat{K}(\beta,\nu)$ and of the L^1-norm of the optimal controls corresponding to $\beta = 0.10$.
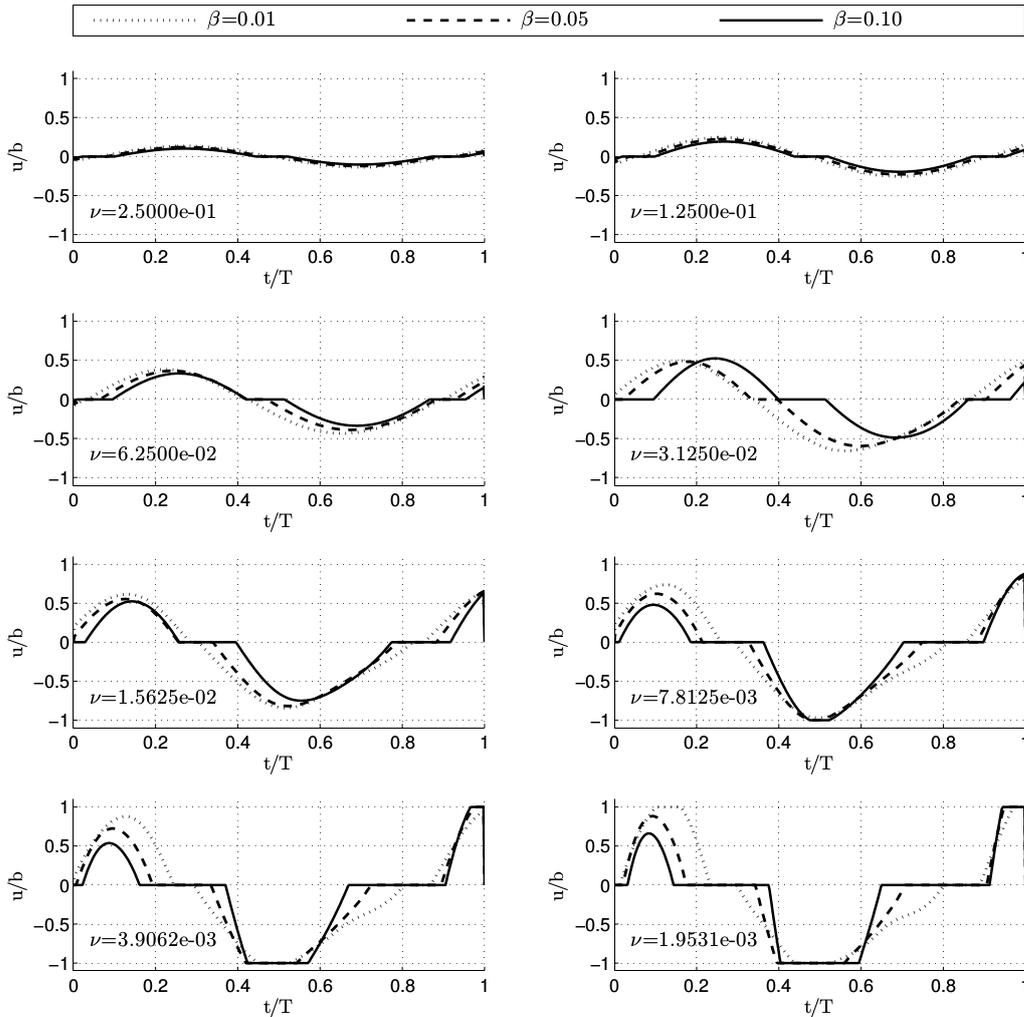

Figure 7.6: Optimal solutions to the dipole quantum control problem corresponding to different values of β and $\nu = 2^{-k}$ with $k = 1, ..., 9$.
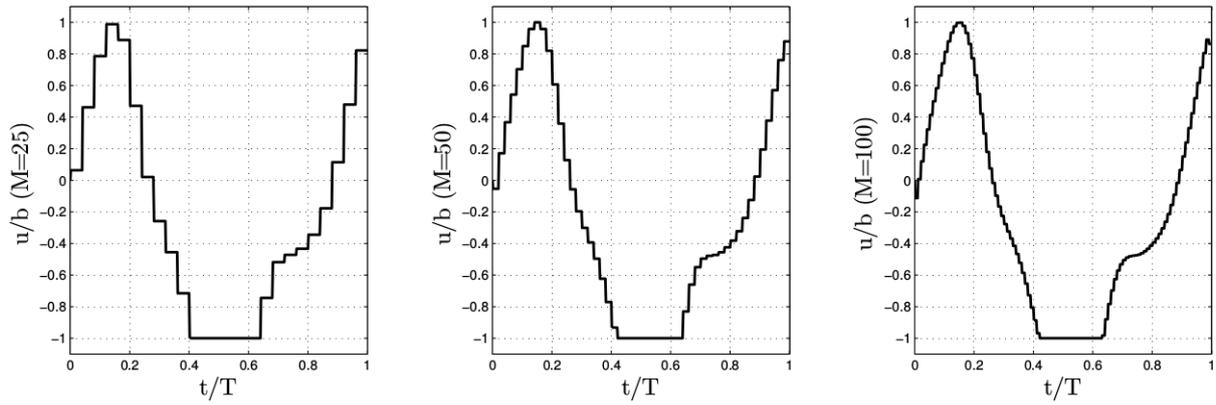
Figure 7.7: Piecewise constant optimal controls obtained for $\nu = 2^{-9}$ ($\beta = 0$) and different numbers of subintervals M.
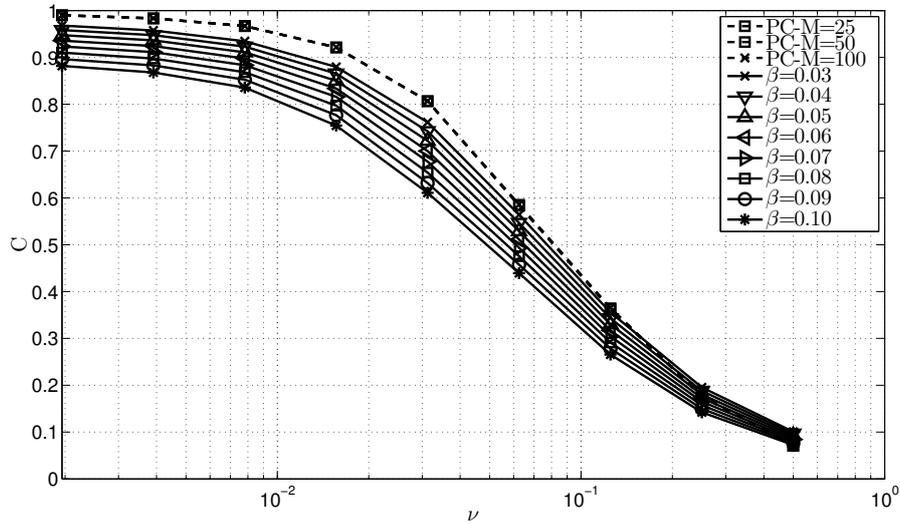

Figure 7.8: Values of the fidelity C obtained by solving problems (3.107) and (3.97) and corresponding to different values of penalization parameters ν and β.

Next, we consider two experiments where we look for piecewise-constant and sparse controls for the state transition from the ground state to the second eigenstate and the fourth eigenstate, respectively. We consider in both cases a number of subintervals $M = 100$. The obtained results are shown in the following figures. Figure 7.9 shows the optimal control for the transition from the first to the second eigenstate and corresponding to $\beta = 0.0120$ and $\nu = 0.0020$. The pointwise in time bound for the control is $b = 10$. From this picture, it is evident the sparse and piecewise-constant structure of the control function, which is also capable to obtain a value of the fidelity equal to $C = 0.9228$. Figure 7.10 shows the corresponding evolution of the absolute value of the wavefunction ψ.
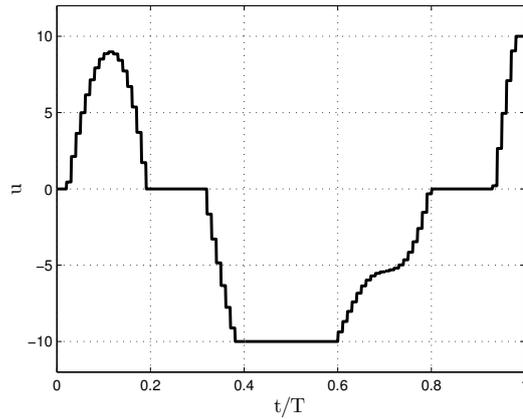

Figure 7.9: Sparse and piecewise-constant optimal control function obtained in the case that ψ_0 is the ground state and ψ_T is the second eigenstate. The corresponding weight parameters are $\beta = 0.0120$ and $\nu = 0.0020$, and the obtained fidelity is $C = 0.9228$.

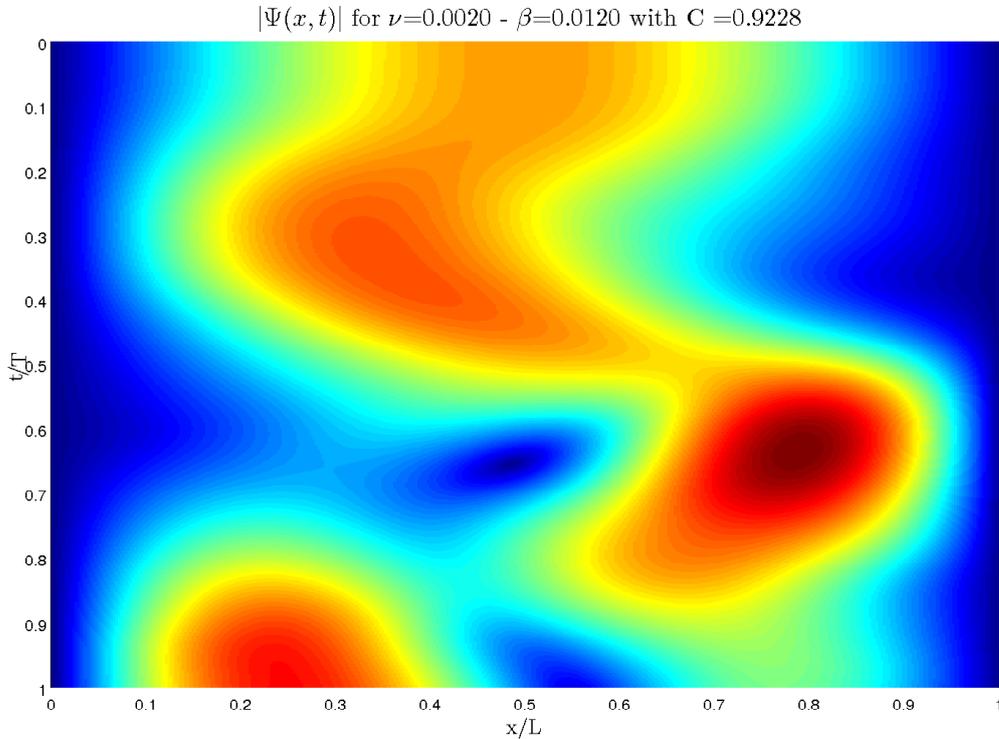Figure 7.10: This figure shows the evolution of the absolute value of the wavefunction $|\psi(x,t)|$ generated by the control function depicted in Figure 7.9.

The following two pictures show the results obtained in the case that the searched transition is from the ground state to the fourth eigenstate of the free-Hamiltonian. Figure 7.11 shows the optimal control solution corresponding to $\beta = 0.0120$ and $\nu = 0.0004$. This control function allows to obtain a fidelity equal to $C = 0.9646$. The pointwise in time bound for the control is $b = 30$. Also in this case, the optimal control has a sparse and piecewise-constant structure. Figure 7.12 shows the corresponding evolution of the absolute value of the wavefunction ψ.
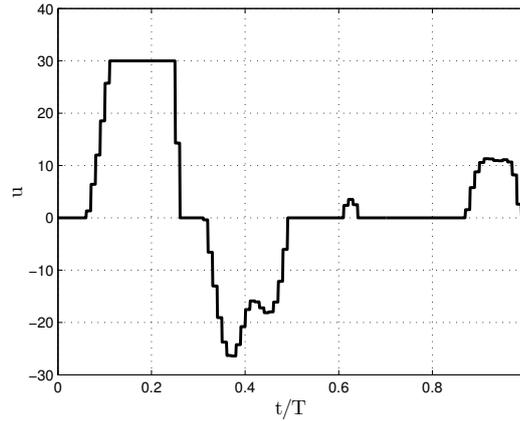
Figure 7.11: Sparse and piecewise-constant optimal control function obtained in the case that ψ_0 is the ground state and ψ_T is the fourth eigenstate. The corresponding weight parameters are $\beta = 0.0120$ and $\nu = 0.0004$, and the obtained fidelity is $C = 0.9646$.
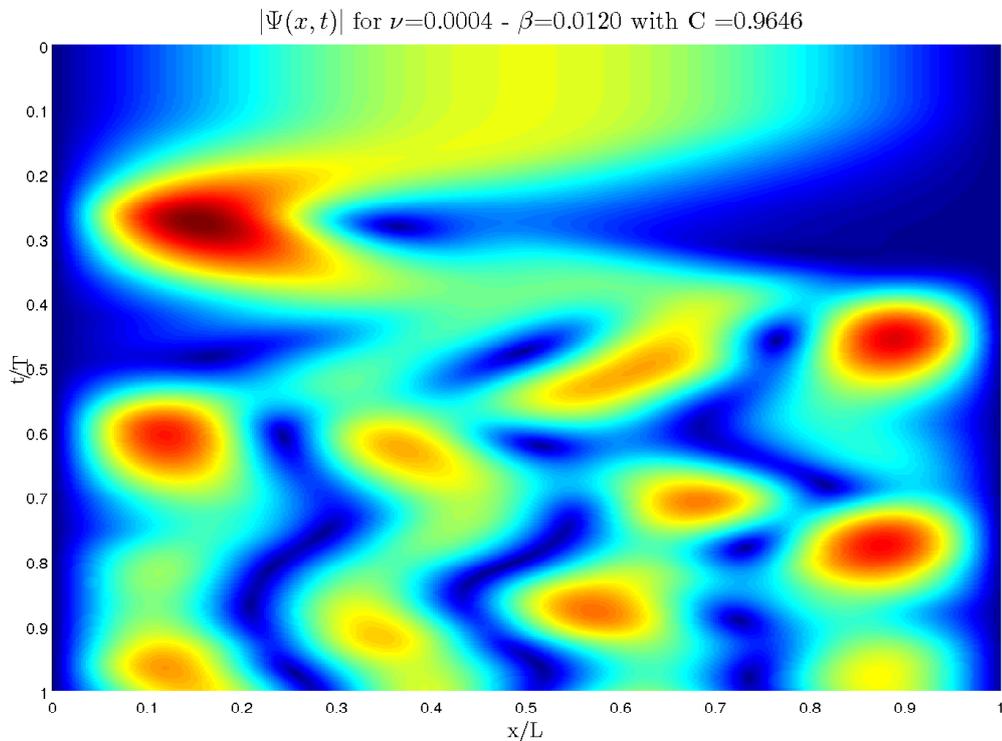

Figure 7.12: This figure shows the evolution of the absolute value of the wavefunction $|\psi(x,t)|$ generated by the control function depicted in Figure 7.11.

7.2 Exact-control problems

In this section, numerical experiments for the exact-control of spin systems are considered. These numerical experiments are performed by means of the computational methods developed in Chapter 5. In particular, in Section 7.2.1 the two computational methods are used to control exactly systems of spins coupled in a chain structure. The control functions are considered in $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$. A comparison between the two computational methods is also considered. In Section 7.2.2, we solve an exact-control problem of a system of two uncoupled spins having different Larmor frequencies. The control is pointwise in time bounded, that is $U_{ad} = U_{ad,1}$, and the resulting exact-control function shows a bang-bang structure. Section 7.2.3 concerns the exact-control of the Pauli equation (2.79)-(2.82), in the case of a zero-potential, that is $V(r) = 0$. In particular, these experiments aim to control the spin and the magnetic state of an electron.

7.2.1 Exact-control of systems of spin chains

We consider numerical experiments for the exact control of spin chains, that are systems of spins coupled in a chain structure. This model is important in several NMR and MRI applications. An example of a chain of spins is given in the following figure, which represents a chain of N_p spins coupled by constants $J_{k,j} > 0$.
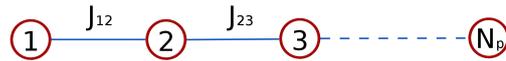

Figure 7.13: Chain of N_p spins.

As described in Section 2.3.3, the Hamiltonian corresponding to these systems is given by

$$\tilde{H} = \tilde{H}_0 + \tilde{H}_{rf} ,$$

where the free Hamiltonian is given by the following; see, e.g., [25, 46, 108];

$$\tilde{H}_0 = \sum_{k=1}^{N_p} \bar{\omega}_k I_{z,k} + \sum_{k<j} J_{k,j} I_{z,k} I_{z,j} .$$

In the case of a chain of spins, the matrix of coupling constants is a symmetric tridiagonal matrix with zeros diagonal entries. Notice that, because of the chain structure, these spin systems are controllable according to Theorem 22. The control Hamiltonian has the following form; see, e.g., [25, 72, 125];

$$\tilde{H}_{rf} = \sum_{k=1}^{N_p} \left(u_{x,k} I_{x,k} + u_{y,k} I_{y,k} \right) .$$

As explained in Section 2.3, the dynamics of a closed spin systems is governed by the Liouville-von Neumann equation, and its real representation corresponding to spin chains is given by the following dynamical system with bilinear control structure

$$\dot{x} = \left[A + \sum_{n=1}^{2N_p} u_n B_n \right] x , \qquad (7.3)$$

that is characterized as in Section 2.4.1. The aim of our experiments is to perform transition from an initial condition characterized by all the spins pointing in the z-direction, to a target state where all the spins are pointing in the x-direction. These two states are represented by the following vectors

$$x_0 = \mathcal{V}(I_{z,k}) \quad \text{and} \quad x_T = \mathcal{V}(I_{x,k}) \, ,$$

where the map $\mathcal{V}$ is defined in (2.125). We assume that, the admissible set of the controls is $U_{ad} = L^2((0,T); \mathbb{R}^{2N_p})$.

To solve this exact-control problem we use the two methods presented in Chapter 5. In particular, we consider this experiment corresponding to different numbers of spins belonging to the chain and different time intervals. Moreover, in order to analyze the results in a way which is of interest for NMR experiments, we compute the fidelity C defined in (7.1).

In the following paragraphs, we first perform numerical experiments by applying the continuation procedure developed in Section 5.2. Then, we consider the same experiments by applying the shooting-type computational scheme developed in Section 5.3. In the last paragraph, we compare the results obtained by means of the two computational methods.

Exact-control by means of the continuation method

In the following, we perform numerical experiments to demonstrate the efficiency of the continuation method developed in Section 5.2 and the validity the corresponding convergence results.

First, we consider two numerical experiments as test cases to analyze numerically the developed convergence results. These test cases are

- Case 1: control of system of one spin: $N_p = 1$, $\bar{\omega}_1 = 1$;

- Case 2: control of system of two coupled spins: $N_p = 2$, $\bar{\omega}_k = 0$, $J_{k,j} = 1$;

In particular, we verify numerically the assumption (5.12) and analyze the results given by (5.17). The experiments are performed by means of Algorithm 13 with $\gamma = 0.9$, $k_{max} = 300$ and starting with $\nu = 1$. The desired tolerances for the Newton method and for the Krylov-CG method are assumed equal to 10^{-14} and 10^{-12}, respectively. We remark that, since an analytical exact control solution $\tilde{u}$ is not available, we approximate it with a control u_{max}, which is the solution of the last step of the continuation method corresponding to a regularization parameter $\nu \approx 10^{-15}$. Furthermore, experiments with different numbers of discretization points are considered.

Consider Case 1. In the following Figure 7.14, we verify the assumption (5.12) of Theorem 7. For this purpose, we define $\Delta x(t) := x_\nu(t) - x_{max}(t)$, where $x_\nu = x(u_\nu)$ and $x_{max} = x(u_{max})$ are the solutions to (3.53a) corresponding to the controls u_ν and u_{max}, respectively. Notice that $\|\Delta x(T)\|_2^2$ and $\tilde{J}$ are equal up to a constant. We consider $\delta x(t)$ as solution to (5.13), that is $\delta x = \delta x(u_\nu - \tilde{u})$. With these settings, we compare $4\|\Delta x(T)\|_2$ and $\|\delta x(T)\|_2$ with a function $f(\nu) = 10\nu$. This comparison shows that the two norms decrease linearly with an order of ν.
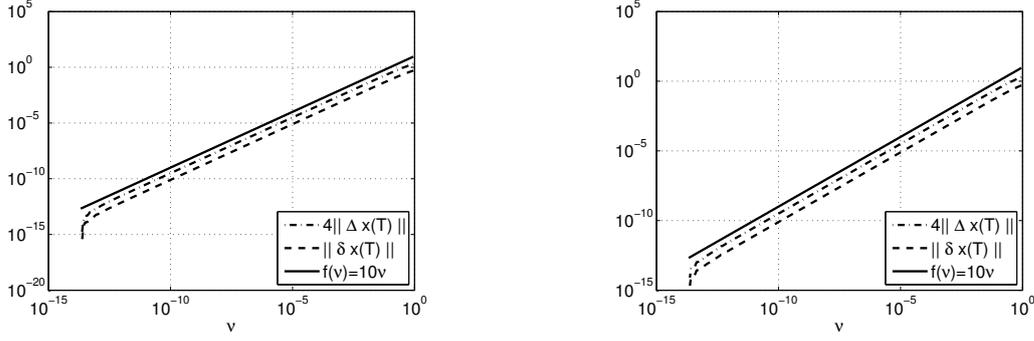
Figure 7.14: Case 1: $\|\Delta x(T)\|_2$ and $\|\delta x(T)\|_2$ with respect to ν in logarithmic scale with a mesh of $N_t = 201$ points (on the left) and $N_t = 2001$ points (on the right).

Figure 7.15 shows that the estimate (5.17) holds. Particularly, the obtained $\|\tilde{u} - u_\nu\|_{L^2}$ is compared with a function $f(\nu) = \nu$: the comparison shows that the norm decreases linearly with ν.
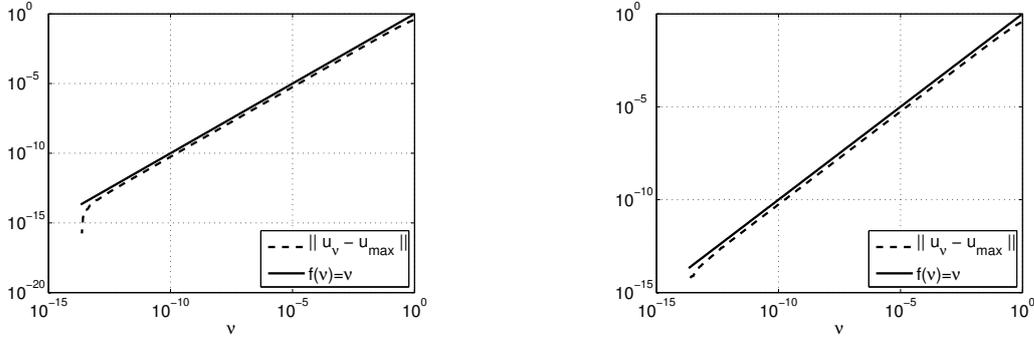

Figure 7.15: Case 1: $\|\tilde{u} - u_\nu\|_{L^2}$ with respect to ν in logarithmic scale with a mesh of $N_t = 201$ points (on the left) and $N_t = 2001$ points (on the right).

The results of numerical experiments shown in Figure 7.15 demonstrate that the estimate (5.17) is not optimal, in the sense that numerical evidence shows a linear rate of convergence, that is, $\|\tilde{u} - u_\nu\|_{L^2} = O(\nu)$. Therefore, our estimate provides an upper bound on the convergence rate. Moreover, the linear rate of convergence is explained by Theorem 28. In fact, from a comparison between Figure 7.14 and Figure 7.15, it results that $\|\Delta x(T)\|_2$ and $\|\tilde{u} - u_\nu\|_{L^2}$ decay with the same order. Consequently, the exponent α in Theorem 28 is equal to 2 and the linear convergence follows.

Consider Case 2. Similarly as for Case 1, with the help of Figure 7.16 we verify the assumption (5.12) and $4\|\Delta x(T)\|_2$ and $\|\delta x(T)\|_2$ are compared with a function $f(\nu) = 10\nu$. This comparison shows that the two norms decrease linearly with ν. Figure 7.17 shows that the estimate (5.17) holds. In this figure, the obtained $\|\tilde{u} - u_\nu\|_{L^2}$ shows a linear decreasing behaviour with an order of ν. In this experiment, when ν becomes very small, the effect of numerical errors due to discretization is evident. However, increasing the number of mesh points, with the help of Figure 7.17 we observe a decay of the numerical error.
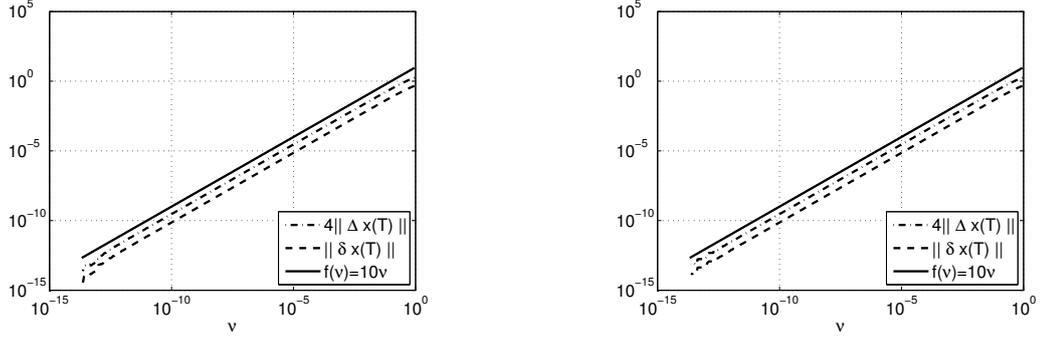
Figure 7.16: Case 2: $\|\Delta x(T)\|_2$ and $\|\delta x(T)\|_2$ with respect to ν in logarithmic scale with a mesh of $N_t = 2001$ points (on the left) and $N_t = 5001$ points (on the right).
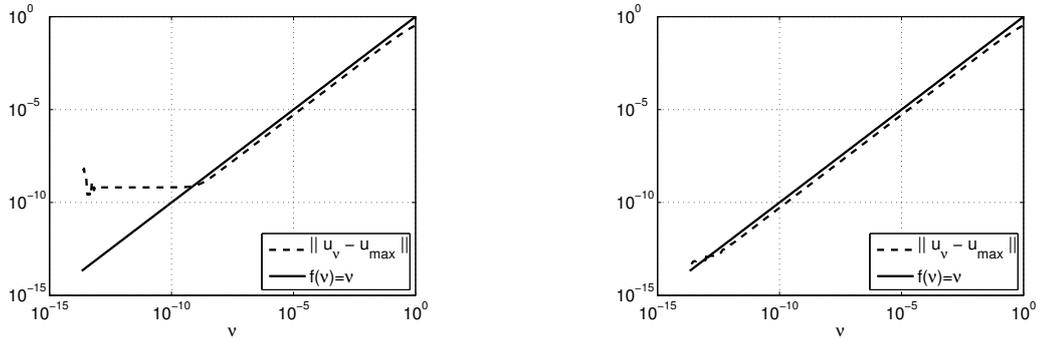
Figure 7.17: Case 2: $\|\tilde{u} - u_\nu\|_{L^2}$ with respect to ν in logarithmic scale with a mesh of $N_t = 2001$ points (on the left) and $N_t = 5001$ points (on the right).

Next, we are interested in studying the robustness of our computational scheme. For this purpose, we perform several experiments corresponding to different systems of coupled spins and different time intervals. The following settings are considered

- number of spins $N_p \in \{1, 2, 3\}$;
- Larmor frequencies $\bar{\omega}_k = 1$ for $N_p = 1$ and $\bar{\omega}_k = 0$ for $N_p = 2$ and $N_p = 3$;
- coupling constants $J_{k,j} = 1$;
- extremum of the time interval equal to $T \in \{1, 2, 5, 8, 10, 20\}$;
- number of discretization points of the interval $[0, T]$ equal to $N_t = 401$. We remark that in some cases we use a finer mesh; this situations are specified in the following tables;
- tolerance of the norm of the gradient for the Krylov-Newton method equal to 10^{-12};
- the sequence of weight parameter is generated by $\nu^{k+1} = \gamma \nu^k$ with $\gamma = 0.5$;
- starting value of the weight parameter $\nu^1 = 1$;
- maximum number of iterations for the continuation method equal to 20;
- the stop-criterion used is in the continuation procedure is related to the computed fidelity: if $C > 1 - 10^{-7}$ then the continuation procedure is stopped.

The results obtained by applying our computational method are shown in the following tables, that contain: ν the weight parameter corresponding to the last iteration of the continuation method; the number of performed iteration of the continuation method "cont

it"; the corresponding value of the cost functional J and the norm of the reduced gradient $\|\nabla_u J_r\|_{L^2}$; the obtained fidelity C and norm of the controls $\|u\|_{L^2}$.

Tables 7.10, 7.11 and 7.12 demonstrate that our continuation procedure is capable to compute exact-control function for almost all the considered cases.

T	ν	J	cont it	$\|\nabla_u J_r\|_{L^2}$	C	$\|u\|_{L^2}$
1	$2.44{\cdot}10^{-4}$	$3.01{\cdot}10^{-4}$	13	$3.75{\cdot}10^{-12}$	0.999999	1.57
2	$4.88{\cdot}10^{-4}$	$3.01{\cdot}10^{-4}$	12	$5.32{\cdot}10^{-12}$	0.999999	1.11
5	$9.76{\cdot}10^{-4}$	$2.41{\cdot}10^{-4}$	11	$1.60{\cdot}10^{-12}$	0.999999	0.70
8	$1.95{\cdot}10^{-3}$	$3.01{\cdot}10^{-4}$	10	$3.93{\cdot}10^{-12}$	0.999999	0.55
10	$1.95{\cdot}10^{-3}$	$2.41{\cdot}10^{-4}$	10	$2.27{\cdot}10^{-12}$	0.999999	0.49
20	$3.90{\cdot}10^{-3}$	$2.41{\cdot}10^{-4}$	9	$3.28{\cdot}10^{-12}$	0.999999	0.35

Table 7.10: Results of the optimizations performed for $N_p = 1$. The obtained values of J, ν and $\|u\|_{L^2}$ and the computed fidelity C show that for all values of T we are able to compute exact-control functions.

T	ν	J	cont it	$\|\nabla_u J_r\|_{L^2}$	C	$\|u\|_{L^2}$
1	$1.90{\cdot}10^{-6}$	$1.90{\cdot}10^{-4}$	20	$2.11{\cdot}10^{-10}$	0.999967	11.47
2	$1.90{\cdot}10^{-6}$	$1.94{\cdot}10^{-4}$	20	$3.03{\cdot}10^{-14}$	0.999965	11.49
5	$1.90{\cdot}10^{-6}$	$2.10{\cdot}10^{-4}$	20	$2.53{\cdot}10^{-12}$	0.999959	11.67
8	$9.76{\cdot}10^{-4}$	$5.54{\cdot}10^{-4}$	11	$2.97{\cdot}10^{-11}$	0.999999	1.06
10	$9.76{\cdot}10^{-4}$	$5.34{\cdot}10^{-4}$	11	$5.82{\cdot}10^{-11}$	0.999999	1.04
20	$1.95{\cdot}10^{-3}$	$4.65{\cdot}10^{-4}$	10	$1.29{\cdot}10^{-12}$	0.999999	0.69

Table 7.11: Results of the optimizations performed for $N_p = 2$. The obtained values of J, ν and $\|u\|_{L^2}$ and the computed fidelity C show that for almost all values of T we are able to compute exact-control functions. In the case that $T = 1$, $T = 2$ and $T = 5$ the numerical procedure is stopped because the number of allowed iterations is reached. However, the values of the corresponding fidelities C are very close to 1, and the results of the optimizations are physically relevant.

T	ν	J	cont it	$\|\nabla_u J_r\|_{L^2}$	C	$\|u\|_{L^2}$
1	$1.90{\cdot}10^{-6}$	$6.34{\cdot}10^{-5}$	20	$2.00{\cdot}10^{-11}$	0.999996	7.53
2	$1.90{\cdot}10^{-6}$	$3.56{\cdot}10^{-5}$	20	$1.75{\cdot}10^{-12}$	0.999999	6.10
5	$3.05{\cdot}10^{-5}$	$2.11{\cdot}10^{-4}$	16	$3.94{\cdot}10^{-13}$	0.999999	3.72
8	$1.22{\cdot}10^{-4}$	$1.39{\cdot}10^{-4}$	14	$7.53{\cdot}10^{-15}$	0.999999	1.51
10	$9.76{\cdot}10^{-4}$	$9.31{\cdot}10^{-4}$	11	$1.59{\cdot}10^{-10}$	0.999999	1.38
20	$3.81{\cdot}10^{-6}$	$1.88{\cdot}10^{-6}$	19	$1.66{\cdot}10^{-9}$	0.999999	0.91

Table 7.12: Results of the optimizations performed for $N_p = 3$. The obtained values of J, ν and $\|u\|_{L^2}$ and the computed fidelity C show that for almost all values of T we are able to compute exact-control functions. In the case that $T = 1$ and $T = 2$ the numerical procedure is stopped because the number of allowed iterations is reached. However, the values of the corresponding fidelities C are very close to 1, and the results of the optimizations are physically relevant.

Next, we consider other experiments similar to the ones previously discussed, but more difficult to solve. In particular, in order to break the physical symmetry of the considered spin systems, we consider non-zero values of the frequencies $(\bar{\omega}_k - \tilde{\omega}_k)$:

- $N_p = 2$: $\bar{\omega}_1 = 1.2$ and $\bar{\omega}_2 = 0.9$;

- $N_p = 3$: $\bar{\omega}_1 = 1.1$, $\bar{\omega}_2 = 1.0$ and $\bar{\omega}_3 = 0.9$.

The results obtained from these numerical experiments are shown in the following tables. In particular, tables 7.13 and 7.14 demonstrate the ability of the continuation procedure in computing exact-controls in almost all the considered cases.

T	ν	J	cont it	$\|\nabla_u J_r\|_{L^2}$	C	$\|u\|_{L^2}$
1	$1.90{\cdot}10^{-6}$	$1.90{\cdot}10^{-4}$	20	$4.33{\cdot}10^{-12}$	0.999967	11.47
2	$1.90{\cdot}10^{-6}$	$1.94{\cdot}10^{-4}$	20	$2.10{\cdot}10^{-12}$	0.999965	11.49
5	$1.90{\cdot}10^{-6}$	$1.88{\cdot}10^{-4}$	20	$4.17{\cdot}10^{-14}$	0.999967	11.35
8	$9.76{\cdot}10^{-4}$	$5.54{\cdot}10^{-4}$	11	$2.97{\cdot}10^{-11}$	0.999999	1.06
10	$9.76{\cdot}10^{-4}$	$5.34{\cdot}10^{-4}$	11	$5.84{\cdot}10^{-11}$	0.999999	1.04
20	$1.95{\cdot}10^{-3}$	$4.64{\cdot}10^{-4}$	10	$1.23{\cdot}10^{-12}$	0.999999	0.68

Table 7.13: Results of the optimizations performed for $N_p = 2$ and $\bar{\omega}_1 = 1.2$ and $\bar{\omega}_2 = 0.9$. The obtained values of J, ν and $\|u\|_{L^2}$ and the computed fidelity C show that for almost all values of T we are able to compute exact-control functions. In the case that $T = 1$, $T = 2$ and $T = 5$ the numerical procedure is stopped because the number of allowed iterations is reached. However, the values of the corresponding fidelities C are very close to 1, and the results of the optimizations are physically relevant. The optimizations for $T = 5$ and $T = 20$ are performed with a number of discretization points $N_t = 1601$.

T	ν	J	cont it	$\|\nabla_u J_r\|_{L^2}$	C	$\|u\|_{L^2}$
1	$1.90{\cdot}10^{-6}$	$6.34{\cdot}10^{-5}$	20	$1.46{\cdot}10^{-10}$	0.999996	7.53
2	$1.90{\cdot}10^{-6}$	$3.56{\cdot}10^{-5}$	20	$7.17{\cdot}10^{-11}$	0.999999	6.10
5	$3.05{\cdot}10^{-5}$	$2.11{\cdot}10^{-4}$	16	$5.60{\cdot}10^{-13}$	0.999999	3.72
8	$1.22{\cdot}10^{-4}$	$1.39{\cdot}10^{-4}$	14	$9.26{\cdot}10^{-15}$	0.999999	1.51
10	$4.88{\cdot}10^{-4}$	$4.66{\cdot}10^{-4}$	12	$1.89{\cdot}10^{-10}$	0.999999	1.38
20	$1.95{\cdot}10^{-3}$	$8.04{\cdot}10^{-4}$	10	$1.02{\cdot}10^{-14}$	0.999999	0.90

Table 7.14: Results of the optimizations performed for $N_p = 3$ and $\bar{\omega}_1 = 1.1$, $\bar{\omega}_2 = 1.0$ and $\bar{\omega}_3 = 0.9$. The obtained values of J, ν and $\|u\|_{L^2}$ and the computed fidelity C show that for almost all values of T we are able to compute exact-control functions. In the case that $T = 1$ and $T = 2$ the numerical procedure is stopped because the number of allowed iterations is reached. However, the values of the corresponding fidelities C are very close to 1, and the results of the optimizations are physically relevant. The optimizations for $T = 8$ and $T = 20$ are performed with a number of discretization points $N_t = 801$ and $N_t = 3201$, respectively.

Exact-control by means of the shooting-type method

We present numerical experiments in order to investigate efficiency and robustness of our shooting-type computational framework developed in Section 5.3. Problem (5.36) is solved by using Algorithm 17 to initialize the optimization procedure, and applying the Krylov-Newton Algorithm 18 to solve the problem. For this purpose, we consider the same problems as in the previous subsection and the following settings

- number of spins $N_p \in \{1, 2, 3\}$;
- Larmor frequencies $\bar{\omega}_k = 1$ for $N_p = 1$ and $\bar{\omega}_k = 0$ for $N_p = 2$ and $N_p = 3$;
- coupling constants $J_{k,j} = 1$;
- extremum of the time interval equal to $T \in \{1, 2, 5, 8, 10, 20\}$;
- number of mesh-refinements in the cascadic procedure equal to 3;
- initial number of discretization points of the interval $[0, T]$ equal to $N_t = 51$ (or greater if necessary);
- tolerance of the norm of the gradient for the NCG method equal to 10^{-3};
- maximum number of iteration for the NCG method equal to 100;
- tolerance of the norm of the gradient for the Krylov-Newton method equal to 10^{-8};
- maximum number of iteration for the Krylov-Newton method equal to 100;
- starting guess for the unknown $u = 0$ and $p_T = 0$.

In order to discuss the numerical behaviour of the computational method, we consider the following tables where G is the value of the cost functional of (5.36) evaluated in the computed solution; G_{init} is the value of the cost functional after the cascadic-NCG initialization; "Newton it" represents the number of iteration performed by the Krylov-Newton method; $\|\nabla G_r\|_{L^2}$ is the norm of the gradient of the reduced problem; C is the fidelity; $\|u\|_{L^2}$ is the norm of the control solution; $\|p_T\|_2$ is the norm of the shooting variable, that is the terminal condition of the adjoint equation of the minimum L^2-norm problem; the constants $C_{1,n}$, $C_{2,n}$, C_3 and C_4 are defined in Theorem 30 and Corollary 6.

In particular, tables 7.15, 7.17 and 7.19 show the results of the optimizations performed on the spin systems corresponding to $N_p = 1$, $N_p = 2$ and $N_p = 3$, respectively. The number of performed iterations by the Krylov-Newton method and the obtained values of the cost functional G show that the computational method is capable to solve efficiently almost all the considered exact-control problems. Moreover, the obtained values of G_{init} show that the NCG-cascadic approach is, in general, capable to provide an efficient initialization to the Newton solver. We remark that, even in the cases in which the convergence is not exactly obtained, the algorithm is still capable to provide high values of the fidelity C.

In tables 7.16, 7.18 and 7.20, we consider an a-posteriori analysis of the performed optimizations, and concerning the sufficient second-order optimality conditions given in Theorem 30 and Corollary 6. In particular, we computed $C_{1,n}$ and $C_{2,n}$ given by (5.65) and (5.66), respectively, and C_3 and C_4 given by (5.86). Notice that, all these coefficients are positive, hence, according to Theorem 30 and Corollary 5, the computed stationary points for the three cases are (global) minima of (5.36).

T	G	G_{init}	Newton it	$\|\nabla G\|_{L^2}$	C	$\|u\|_{L^2}$
1	$2.16 \cdot 10^{-22}$	$6.15 \cdot 10^{-7}$	3	$1.22 \cdot 10^{-11}$	0.999999	1.57
2	$2.95 \cdot 10^{-24}$	$1.36 \cdot 10^{-7}$	3	$2.34 \cdot 10^{-12}$	0.999999	1.11
5	$1.42 \cdot 10^{-24}$	$1.48 \cdot 10^{-7}$	3	$2.29 \cdot 10^{-12}$	0.999999	7.03
8	$2.01 \cdot 10^{-24}$	$2.36 \cdot 10^{-7}$	3	$3.20 \cdot 10^{-12}$	0.999999	5.56
10	$2.21 \cdot 10^{-24}$	$2.03 \cdot 10^{-8}$	3	$3.50 \cdot 10^{-12}$	0.999999	0.49
20	$3.50 \cdot 10^{-26}$	$5.67 \cdot 10^{-8}$	3	$3.92 \cdot 10^{-13}$	0.999999	0.35

Table 7.15: Results of the optimizations performed for $N_p = 1$. The obtained values of G and the computed fidelity C show that for all values of T we are able to compute exact-control functions.

T	C	$\|p_T\|_2$	$\min_n C_{1,n}$	$\min_n C_{2,n}$	C_3	C_4
1	0.999999	$1.72 \cdot 10^0$	$3.60 \cdot 10^{+2}$	$1.13 \cdot 10^{+1}$	$7.21 \cdot 10^{+2}$	$3.89 \cdot 10^{+1}$
2	0.999999	$8.40 \cdot 10^{-1}$	$3.49 \cdot 10^{+2}$	$2.30 \cdot 10^{+1}$	$6.99 \cdot 10^{+2}$	$5.49 \cdot 10^{+1}$
5	0.999999	$3.38 \cdot 10^{-1}$	$3.69 \cdot 10^{+2}$	$5.78 \cdot 10^{+1}$	$7.38 \cdot 10^{+2}$	$9.10 \cdot 10^{+1}$
8	0.999999	$2.11 \cdot 10^{-1}$	$3.81 \cdot 10^{+2}$	$9.23 \cdot 10^{+1}$	$7.63 \cdot 10^{+2}$	$1.18 \cdot 10^{+2}$
10	0.999999	$1.80 \cdot 10^{-1}$	$4.36 \cdot 10^{+2}$	$1.15 \cdot 10^{+2}$	$8.72 \cdot 10^{+2}$	$1.42 \cdot 10^{+2}$
20	0.999999	$8.75 \cdot 10^{-2}$	$4.37 \cdot 10^{+2}$	$2.29 \cdot 10^{+2}$	$8.75 \cdot 10^{+2}$	$2.08 \cdot 10^{+2}$

Table 7.16: In this table, conditions of Theorem 30 and Corollary 5 regarding the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 corresponding to the optimizations with $N_p = 1$. In particular, the norms of the terminal conditions p_T, the time T, the coefficients $C_{1,n}$, $C_{2,n}$, C_3 and C_4 are shown. According to Theorem 30 and Corollary 5, the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 guarantees that the computed stationary points of (5.36) in the different cases are isolated global minima.

T	G	G_{init}	Newton it	$\|\nabla G\|_{L^2}$	C	$\|u\|_{L^2}$
1	$4.06 \cdot 10^{-3}$	$3.94 \cdot 10^{-2}$	100	$2.96 \cdot 10^{-4}$	0.997976	4.25
2	$2.42 \cdot 10^{-14}$	$3.37 \cdot 10^{-2}$	44	$1.28 \cdot 10^{-11}$	0.999999	4.02
5	$1.26 \cdot 10^{-16}$	$6.85 \cdot 10^{-2}$	24	$1.63 \cdot 10^{-9}$	0.999999	2.66
8	$8.89 \cdot 10^{-21}$	$3.64 \cdot 10^{-7}$	3	$1.08 \cdot 10^{-10}$	0.999999	1.06
10	$7.71 \cdot 10^{-24}$	$8.42 \cdot 10^{-8}$	3	$5.16 \cdot 10^{-12}$	0.999999	1.04
20	$5.62 \cdot 10^{-26}$	$1.08 \cdot 10^{-7}$	3	$3.58 \cdot 10^{-13}$	0.999999	0.69

Table 7.17: Results of the optimizations performed for $N_p = 2$. The obtained values of G and the computed fidelity C show that for almost all values of T we are able to compute exact-control functions. In the case that $T = 1$, the numerical procedure is stopped because the number of allowed iterations is reached. However, the value of the corresponding fidelity C is very close to 1, and the result of the optimization is physically relevant.

T	C	$\|p_T\|_2$	$\min_n C_{1,n}$	$\min_n C_{2,n}$	C_3	C_4
1	0.997976	$3.16 \cdot 10^{+2}$	$2.04 \cdot 10^{+8}$	$4.56 \cdot 10^{+1}$	$8.18 \cdot 10^{+8}$	$8.09 \cdot 10^{+4}$
2	0.999999	$2.46 \cdot 10^{+1}$	$4.97 \cdot 10^{+6}$	$9.15 \cdot 10^{+1}$	$1.99 \cdot 10^{+7}$	$1.78 \cdot 10^{+4}$
5	0.999999	$1.03 \cdot 10^{+1}$	$5.52 \cdot 10^{+6}$	$2.28 \cdot 10^{+2}$	$2.21 \cdot 10^{+7}$	$2.97 \cdot 10^{+4}$
8	0.999999	$5.41 \cdot 10^{-1}$	$3.88 \cdot 10^{+4}$	$3.65 \cdot 10^{+2}$	$1.55 \cdot 10^{+5}$	$3.18 \cdot 10^{+3}$
10	0.999999	$5.33 \cdot 10^{-1}$	$5.88 \cdot 10^{+4}$	$4.56 \cdot 10^{+2}$	$2.35 \cdot 10^{+5}$	$4.38 \cdot 10^{+3}$
20	0.999999	$2.54 \cdot 10^{-1}$	$5.40 \cdot 10^{+4}$	$9.10 \cdot 10^{+2}$	$2.16 \cdot 10^{+5}$	$5.96 \cdot 10^{+3}$

Table 7.18: In this table, conditions of Theorem 30 and Corollary 5 regarding the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 corresponding to the optimizations with $N_p = 2$. In particular, the norms of the terminal conditions p_T, the time T, the coefficients $C_{1,n}$, $C_{2,n}$, C_3 and C_4 are shown. According to Theorem 30 and Corollary 5, the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 guarantees that the computed stationary points of (5.36) in the different cases are isolated global minima.

T	G	G_{init}	Newton it	$\|\nabla G\|_{L^2}$	C	$\|u\|_{L^2}$
1	$1.88{\cdot}10^{-2}$	$8.04{\cdot}10^{-2}$	100	$2.17{\cdot}10^{-4}$	0.993720	4.85
2	$4.18{\cdot}10^{-3}$	$7.21{\cdot}10^{-2}$	100	$1.37{\cdot}10^{-4}$	0.998607	4.36
5	$2.16{\cdot}10^{-3}$	$2.39{\cdot}10^{-2}$	96	$7.99{\cdot}10^{-5}$	0.999281	3.40
8	$1.49{\cdot}10^{-6}$	$1.75{\cdot}10^{-3}$	38	$4.10{\cdot}10^{-6}$	0.999999	1.51
10	$3.58{\cdot}10^{-9}$	$8.69{\cdot}10^{-6}$	5	$7.32{\cdot}10^{-11}$	0.999999	1.38
20	$3.53{\cdot}10^{-7}$	$1.13{\cdot}10^{-6}$	3	$2.53{\cdot}10^{-9}$	0.999999	0.91

Table 7.19: Results of the optimizations performed for $N_p = 3$. The obtained values of G and the computed fidelity C show that for almost all values of T we are able to compute exact-control functions. In the cases that $T = 1$ and $T = 2$, the numerical procedure is stopped because the number of allowed iterations is reached, while for $T = 5$ and $T = 8$, the iterative procedure is stopped because the linear Krylov solver was stagnant. However, the values of the corresponding fidelities C are very close to 1, and the results of the optimizations are physically relevant.

T	C	$\|p_T\|_2$	$\min_n C_{1,n}$	$\min_n C_{2,n}$	C_3	C_4
1	0.993720	$1.62{\cdot}10^{+3}$	$2.74{\cdot}10^{+10}$	$1.02{\cdot}10^{+2}$	$1.64{\cdot}10^{+11}$	$1.72{\cdot}10^{+6}$
2	0.998607	$7.79{\cdot}10^{+2}$	$2.51{\cdot}10^{+10}$	$2.05{\cdot}10^{+2}$	$1.51{\cdot}10^{+11}$	$2.33{\cdot}10^{+6}$
5	0.999281	$2.76{\cdot}10^{+2}$	$1.97{\cdot}10^{+10}$	$5.11{\cdot}10^{+2}$	$1.18{\cdot}10^{+11}$	$3.26{\cdot}10^{+6}$
8	0.999999	$2.14{\cdot}10^{+1}$	$3.06{\cdot}10^{+8}$	$8.18{\cdot}10^{+2}$	$1.83{\cdot}10^{+9}$	$5.14{\cdot}10^{+5}$
10	0.999999	$3.19{\cdot}10^{0}$	$1.05{\cdot}10^{+7}$	$1.02{\cdot}10^{+3}$	$6.34{\cdot}10^{+7}$	$1.06{\cdot}10^{+5}$
20	0.999999	$3.80{\cdot}10^{-1}$	$6.05{\cdot}10^{+5}$	$2.04{\cdot}10^{+3}$	$3.63{\cdot}10^{+6}$	$3.63{\cdot}10^{+4}$

Table 7.20: In this table, conditions of Theorem 30 and Corollary 5 regarding the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 corresponding to the optimizations with $N_p = 3$. In particular, the norms of the terminal conditions p_T, the time T, the coefficients $C_{1,n}$, $C_{2,n}$, C_3 and C_4 are shown. According to Theorem 30 and Corollary 5, the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 guarantees that the computed stationary points of (5.36) in the different cases are isolated global minima.

Next, similarly as in the previous subsection, we consider other experiments similar in which the physical symmetry of the considered spin systems is broken:

- $N_p = 2$: $\bar\omega_1 = 1.2$ and $\bar\omega_2 = 0.9$;
- $N_p = 3$: $\bar\omega_1 = 1.1$, $\bar\omega_2 = 1.0$ and $\bar\omega_3 = 0.9$.

The results obtained from these numerical experiments are shown in the following tables.

T	G	G_{init}	Newton it	$\|\nabla G\|_{L^2}$	C	$\|u\|_{L^2}$
1	$3.29{\cdot}10^{-3}$	$4.32{\cdot}10^{-2}$	100	$1.88{\cdot}10^{0}$	0.998402	4.51
2	$4.88{\cdot}10^{-11}$	$3.34{\cdot}10^{-2}$	55	$2.99{\cdot}10^{-9}$	0.999999	3.84
5	$5.08{\cdot}10^{-12}$	$7.36{\cdot}10^{-2}$	38	$3.12{\cdot}10^{-11}$	0.999999	2.65
8	$7.15{\cdot}10^{-9}$	$6.84{\cdot}10^{-7}$	4	$2.00{\cdot}10^{-10}$	0.999999	1.06
10	$1.77{\cdot}10^{-8}$	$7.49{\cdot}10^{-8}$	3	$2.04{\cdot}10^{-11}$	0.999999	1.04
20	$4.15{\cdot}10^{-9}$	$6.98{\cdot}10^{-8}$	3	$4.73{\cdot}10^{-11}$	0.999999	0.68

Table 7.21: Results of the optimizations performed for $N_p = 2$ and $\bar\omega_1 = 1.2$ and $\bar\omega_2 = 0.9$. The obtained values of G and the computed fidelity C show that for almost all values of T we are able to compute exact-control functions. In the case that $T = 1$, the numerical procedure is stopped because the number of allowed iterations is reached, while for $T = 5$, the iterative procedure is stopped because the linear Krylov solver was stagnant. However, the values of the corresponding fidelities C are very close to 1, and the results of the optimizations are physically relevant. The optimizations for $T = 5$ and $T = 20$ are performed with a number of discretization points $N_t = 1601$.

T	C	$\|p_T\|_2$	$\min_n C_{1,n}$	$\min_n C_{2,n}$	C_3	C_4
1	0.998402	$6.81 \cdot 10^{+2}$	$9.52 \cdot 10^{+8}$	$4.56 \cdot 10^{+1}$	$3.80 \cdot 10^{+9}$	$1.74 \cdot 10^{+5}$
2	0.999999	$1.98 \cdot 10^{+1}$	$3.22 \cdot 10^{+6}$	$9.15 \cdot 10^{+1}$	$1.28 \cdot 10^{+7}$	$1.43 \cdot 10^{+4}$
5	0.999999	$1.05 \cdot 10^{+1}$	$5.70 \cdot 10^{+6}$	$2.28 \cdot 10^{+2}$	$2.28 \cdot 10^{+7}$	$3.02 \cdot 10^{+4}$
8	0.999999	$6.06 \cdot 10^{-1}$	$4.86 \cdot 10^{+4}$	$3.65 \cdot 10^{+2}$	$1.94 \cdot 10^{+5}$	$3.56 \cdot 10^{+3}$
10	0.999999	$5.31 \cdot 10^{-1}$	$5.85 \cdot 10^{+4}$	$4.56 \cdot 10^{+2}$	$2.34 \cdot 10^{+5}$	$4.36 \cdot 10^{+3}$
20	0.999999	$2.55 \cdot 10^{-1}$	$5.45 \cdot 10^{+4}$	$9.10 \cdot 10^{+2}$	$2.18 \cdot 10^{+5}$	$5.98 \cdot 10^{+3}$

Table 7.22: In this table, conditions of Theorem 30 and Corollary 5 regarding the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 corresponding to the optimizations with $N_p = 2$ and $\bar{\omega}_1 = 1.2$ and $\bar{\omega}_2 = 0.9$. In particular, the norms of the terminal conditions p_T, the time T, the coefficients $C_{1,n}$, $C_{2,n}$, C_3 and C_4 are shown. According to Theorem 30 and Corollary 5, the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 guarantees that the computed stationary points of (5.36) in the different cases are isolated global minima.

T	G	G_{init}	Newton it	$\|\nabla G\|_{L^2}$	C	$\|u\|_{L^2}$
1	$2.01 \cdot 10^{-2}$	$8.26 \cdot 10^{-2}$	100	$2.49 \cdot 10^{-3}$	0.993274	5.41
2	$1.43 \cdot 10^{-3}$	$3.48 \cdot 10^{-2}$	100	$1.27 \cdot 10^{-1}$	0.999521	4.68
5	$2.99 \cdot 10^{-3}$	$3.20 \cdot 10^{-2}$	100	$2.96 \cdot 10^{-1}$	0.999017	3.29
8	$4.17 \cdot 10^{-6}$	$1.72 \cdot 10^{-3}$	55	$8.29 \cdot 10^{-6}$	0.999998	1.51
10	$9.13 \cdot 10^{-8}$	$2.23 \cdot 10^{-5}$	10	$2.57 \cdot 10^{-9}$	0.999999	1.38
20	$1.05 \cdot 10^{-9}$	$1.07 \cdot 10^{-5}$	6	$5.78 \cdot 10^{-9}$	0.999999	0.90

Table 7.23: Results of the optimizations performed for $N_p = 3$ and $\bar{\omega}_1 = 1.1$, $\bar{\omega}_2 = 1.0$ and $\bar{\omega}_3 = 0.9$. The obtained values of G and the computed fidelity C show that for all values of T allows to compute good control functions. In the cases that $T = 1$, $T = 2$ and $T = 5$, the numerical procedure is stopped because the number of allowed iterations is reached, while for $T = 8$, the iterative procedure is stopped because the linear Krylov solver was stagnant. However, the values of the corresponding fidelities C are very close to 1, and the results of the optimizations are physically relevant. The optimizations for $T = 8$ and $T = 20$ are performed with a number of discretization points $N_t = 801$ and $N_t = 3201$, respectively.

T	C	$\|p_T\|_2$	$\min_n C_{1,n}$	$\min_n C_{2,n}$	C_3	C_4
1	0.993274	$2.95 \cdot 10^{+2}$	$9.03 \cdot 10^{+8}$	$1.02 \cdot 10^{+2}$	$5.42 \cdot 10^{+9}$	$3.12 \cdot 10^{+5}$
2	0.999521	$3.72 \cdot 10^{+2}$	$5.74 \cdot 10^{+9}$	$2.05 \cdot 10^{+2}$	$3.44 \cdot 10^{+10}$	$1.11 \cdot 10^{+6}$
5	0.999017	$1.16 \cdot 10^{+2}$	$3.51 \cdot 10^{+9}$	$5.11 \cdot 10^{+2}$	$2.10 \cdot 10^{+10}$	$1.37 \cdot 10^{+6}$
8	0.999998	$2.20 \cdot 10^{+1}$	$3.22 \cdot 10^{+8}$	$8.18 \cdot 10^{+2}$	$1.93 \cdot 10^{+9}$	$5.27 \cdot 10^{+5}$
10	0.999999	$4.24 \cdot 10^{0}$	$1.86 \cdot 10^{+7}$	$1.02 \cdot 10^{+3}$	$1.12 \cdot 10^{+8}$	$1.42 \cdot 10^{+5}$
20	0.999999	$1.37 \cdot 10^{0}$	$7.86 \cdot 10^{+6}$	$2.04 \cdot 10^{+3}$	$4.72 \cdot 10^{+7}$	$1.30 \cdot 10^{+5}$

Table 7.24: In this table, conditions of Theorem 30 and Corollary 5 regarding the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 corresponding to the optimizations with $N_p = 3$ and $\bar{\omega}_1 = 1.1$, $\bar{\omega}_2 = 1.0$ and $\bar{\omega}_3 = 0.9$. In particular, the norms of the terminal conditions p_T, the time T, the coefficients $C_{1,n}$, $C_{2,n}$, C_3 and C_4 are shown. According to Theorem 30 and Corollary 5, the positivity of $C_{1,n}$, $C_{2,n}$, C_3 and C_4 guarantees that the computed stationary points of (5.36) in the different cases are isolated global minima.

Comparison of the results

In this section, the results presented in the two previous subsections which are obtained by means of the two methods presented in Chapter 5, are compared. In particular, in the following tables, we show the values of fidelities obtained from the continuation method C_{cont} and the shooting-type method C_{shoot}, and the corresponding norms of the optimal controls $\|u_{cont}\|_{L^2}$, $\|u_{shoot}\|_{L^2}$ and $\|u_{cont} - u_{shoot}\|_{L^2}$. These results demonstrate that, for large values of T the two computational methods allow to obtain the same exact-control functions. This is demonstrated also by means of the plots of the obtained exact-controls corresponding to $T = 10$.

The following table shows the results obtained in the case $N_p = 1$. The Table 7.25 compare results presented in tables 7.10 and 7.20. In particular, a comparison of the fidelities and the norms of the optimal controls demonstrate that for all the vales of T the two computational methods result in the same exact-control function. This is also validated by Figure 7.18, which shows the control functions obtained in the case that $T = 10$.

T	C_{cont}	C_{shoot}	$\|u_{cont}\|_{L^2}$	$\|u_{shoot}\|_{L^2}$	$\|u_{cont} - u_{shoot}\|_{L^2}$
1	0.999999	0.999999	1.57	1.57	$5.11{\cdot}10^{-4}$
2	0.999999	0.999999	1.11	1.11	$2.69{\cdot}10^{-4}$
5	0.999999	0.999999	0.70	0.70	$2.07{\cdot}10^{-4}$
8	0.999999	0.999999	0.55	0.55	$1.48{\cdot}10^{-4}$
10	0.999999	0.999999	0.49	0.49	$1.54{\cdot}10^{-4}$
20	0.999999	0.999999	0.35	0.35	$1.07{\cdot}10^{-4}$

Table 7.25: Comparison of the results obtained for $N_p = 1$. The values of the fidelities C_{cont} and C_{shoot} show that for all the values T the two computational methods are capable to obtain exact controls. Furthermore, the values $\|u_{cont}\|_{L^2}$, $\|u_{shoot}\|_{L^2}$ and $\|u_{cont} - u_{shoot}\|_{L^2}$ show that the two frameworks allow to compute the same control functions.
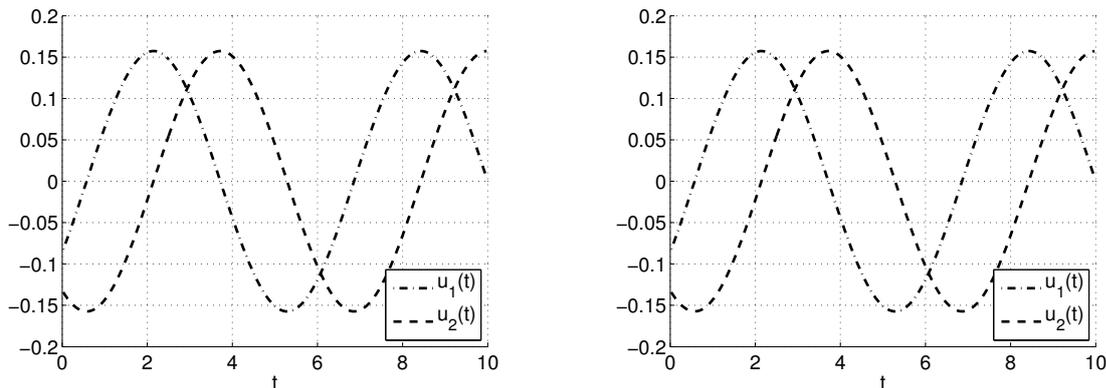

Figure 7.18: Control functions obtained by means of the continuation method (left) and the shooting-type method (right) in the case $N_p = 1$ and $T = 10$. A comparison between the two pictures shows that the two frameworks allow to compute the same control functions.

Next, we consider the results obtained in the case $N_p = 2$. In this case, the Table 7.26 compare results presented in tables 7.11 and 7.21. A comparison of the fidelities and the norms of the optimal controls demonstrate that for almost all the vales of T the two computational methods result in the same exact-control function. This fact is also validated by Figure 7.19, which shows the control functions corresponding to $T = 10$.

T	C_{cont}	C_{shoot}	$\|u_{cont}\|_{L^2}$	$\|u_{shoot}\|_{L^2}$	$\|u_{cont} - u_{shoot}\|_{L^2}$
1	0.999967	0.997976	11.47	4.25	-
2	0.999965	0.999999	11.49	1.11	-
5	0.999959	0.999999	11.67	0.70	-
8	0.999999	0.999999	1.06	1.06	$7.92 \cdot 10^{-4}$
10	0.999999	0.999999	1.04	1.04	$9.10 \cdot 10^{-4}$
20	0.999999	0.999999	0.69	0.69	$5.02 \cdot 10^{-4}$

Table 7.26: Comparison of the results obtained for $N_p = 2$. The values of the fidelities C_{cont} and C_{shoot} show that for almost all the values T the two computational methods are capable to obtain exact controls. Furthermore, the values $\|u_{cont}\|_{L^2}$, $\|u_{shoot}\|_{L^2}$ and $\|u_{cont} - u_{shoot}\|_{L^2}$ show that the two frameworks allow to compute the same control functions for $T = 8$, $T = 10$ and $T = 20$.
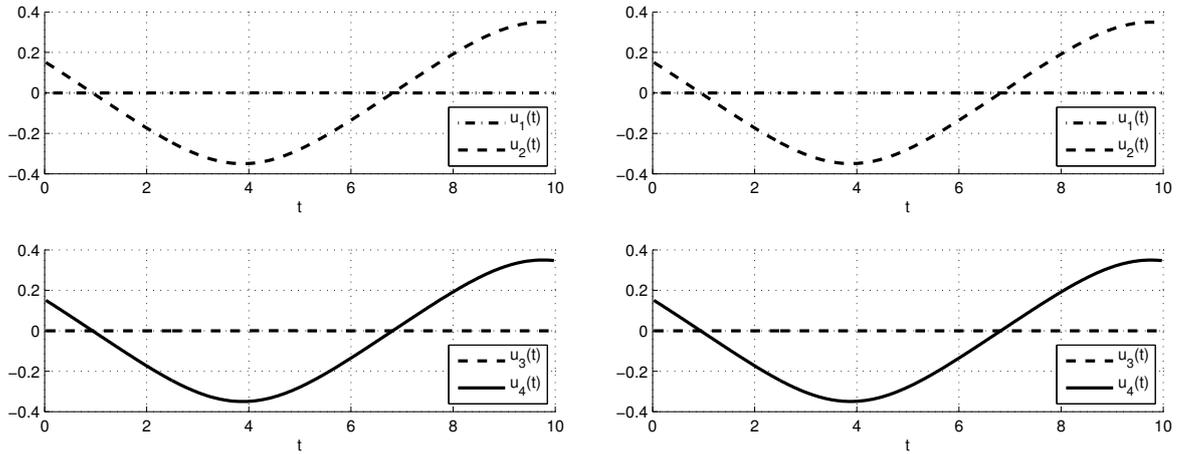

Figure 7.19: Control functions obtained by means of the continuation method (left) and the shooting-type method (right) in the case $N_p = 2$ and $T = 10$. A comparison between the two pictures shows that the two frameworks allow to compute the same control functions.

The following table shows the results obtained in the case $N_p = 3$. The Table 7.27 compare results presented in tables 7.12 and 7.22. In particular, a comparison of the fidelities and the norms of the optimal controls demonstrate that for almost all the vales of T the two methods result in the same exact-control function. This is also validated by Figure 7.20, which shows the control functions obtained in the case that $T = 10$.

T	C_{cont}	C_{shoot}	$\|u_{cont}\|_{L^2}$	$\|u_{shoot}\|_{L^2}$	$\|u_{cont} - u_{shoot}\|_{L^2}$
1	0.999996	0.993720	7.53	4.85	-
2	0.999999	0.998627	6.10	4.36	-
5	0.999999	0.999281	3.72	3.40	-
8	0.999999	0.999999	1.51	1.51	$7.50 \cdot 10^{-3}$
10	0.999999	0.999999	1.38	1.38	$8.93 \cdot 10^{-4}$
20	0.999999	0.999999	0.91	0.91	$1.02 \cdot 10^{-2}$

Table 7.27: Comparison of the results obtained for $N_p = 3$. The values of the fidelities C_{cont} and C_{shoot} show that for almost all the values T the two computational schemes are capable to obtain exact controls. Furthermore, the values $\|u_{cont}\|_{L^2}$, $\|u_{shoot}\|_{L^2}$ and $\|u_{cont} - u_{shoot}\|_{L^2}$ show that the two frameworks allow to compute the same control functions for $T = 8$, $T = 10$ and $T = 20$.
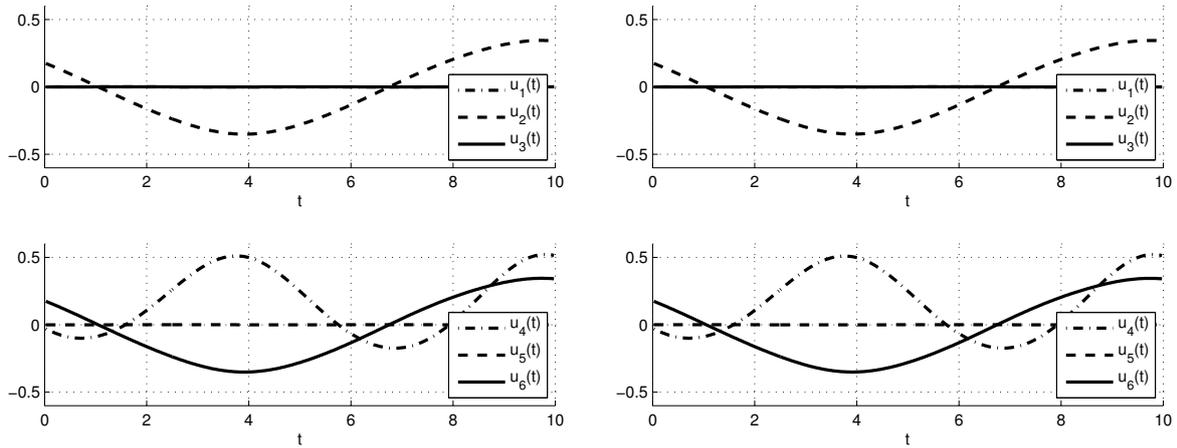

Figure 7.20: Control functions obtained by means of the continuation method (left) and the shooting-type method (right) in the case $N_p = 3$ and $T = 10$. A comparison between the two pictures shows that the two frameworks allow to compute the same control functions.

The following table shows the results obtained in the case $N_p = 2$ with $\bar{\omega}_1 = 1.2$ and $\bar{\omega}_2 = 0.9$. The Table 7.28 compare results presented in tables 7.13 and 7.23. In particular, a comparison of the fidelities and the norms of the optimal controls demonstrate that for almost all the vales of T the two methods result in the same exact-control function. This is also validated by Figure 7.21, which shows the control functions obtained in the case that $T = 10$.

T	C_{cont}	C_{shoot}	$\|u_{cont}\|_{L^2}$	$\|u_{shoot}\|_{L^2}$	$\|u_{cont} - u_{shoot}\|_{L^2}$
1	0.999967	0.998402	11.47	4.51	-
2	0.999965	0.999999	11.49	3.84	-
5	0.999967	0.999999	11.35	2.65	-
8	0.999999	0.999999	1.06	1.06	$6.54 \cdot 10^{-4}$
10	0.999999	0.999999	1.04	1.04	$6.62 \cdot 10^{-4}$
20	0.999999	0.999999	0.68	0.68	$3.92 \cdot 10^{-5}$

Table 7.28: Comparison of the results obtained for $N_p = 2$ with $\bar{\omega}_1 = 1.2$ and $\bar{\omega}_2 = 0.9$. The values of the fidelities C_{cont} and C_{shoot} show that for almost all the values T the two computational methods are capable to obtain exact controls. Furthermore, the values $\|u_{cont}\|_{L^2}$, $\|u_{shoot}\|_{L^2}$ and $\|u_{cont} - u_{shoot}\|_{L^2}$ show that the two frameworks allow to compute the same control functions for $T = 8$, $T = 10$ and $T = 20$.
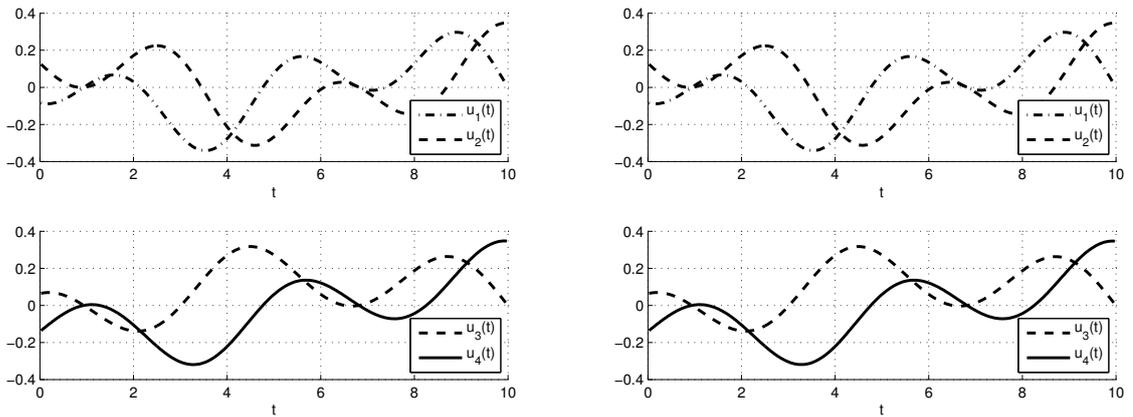

Figure 7.21: Control functions obtained by means of the continuation method (left) and the shooting-type method (right) in the case $N_p = 2$, with $\bar{\omega}_1 = 1.2$ and $\bar{\omega}_2 = 0.9$, and $T = 10$. A comparison between the two pictures shows that the two frameworks allow to compute the same control functions.

Next, we consider the results obtained in the case $N_p = 3$ with $\bar{\omega}_1 = 1.1$, $\bar{\omega}_2 = 1.0$ and $\bar{\omega}_2 = 0.9$. In this case, the Table 7.29 compare results presented in tables 7.14 and 7.24. A comparison of the fidelities and the norms of the optimal controls demonstrate that for almost all the vales of T the two computational schemes result in the same exact-control function. This fact is also validated by Figure 7.22, which shows the control functions corresponding to $T = 10$.

T	C_{cont}	C_{shoot}	$\|u_{cont}\|_{L^2}$	$\|u_{shoot}\|_{L^2}$	$\|u_{cont} - u_{shoot}\|_{L^2}$
1	0.999996	0.993274	7.53	5.41	-
2	0.999999	0.999521	6.10	4.68	-
5	0.999999	0.999017	3.72	3.29	-
8	0.999999	0.999998	1.51	1.51	$3.43 \cdot 10^{-2}$
10	0.999999	0.999999	1.38	1.38	$7.49 \cdot 10^{-4}$
20	0.999999	0.999999	0.90	0.90	$8.25 \cdot 10^{-5}$

Table 7.29: Comparison of the results obtained for $N_p = 3$ with $\bar{\omega}_1 = 1.1$, $\bar{\omega}_2 = 1.0$ and $\bar{\omega}_2 = 0.9$. The values of the fidelities C_{cont} and C_{shoot} show that for almost all the values T the two computational methods are capable to obtain exact controls. Furthermore, the values $\|u_{cont}\|_{L^2}$, $\|u_{shoot}\|_{L^2}$ and $\|u_{cont} - u_{shoot}\|_{L^2}$ show that the two frameworks allow to compute the same control functions for $T = 8$, $T = 10$ and $T = 20$.
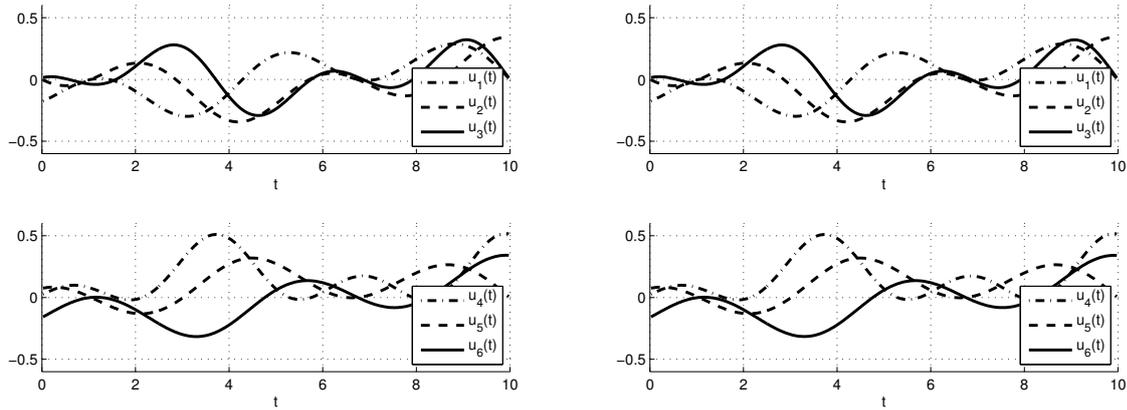

Figure 7.22: Control functions obtained by means of the continuation method (left) and the shooting-type method (right) in the case $N_p = 2$, with $\bar{\omega}_1 = 1.1$, $\bar{\omega}_2 = 1.0$ and $\bar{\omega}_2 = 0.9$, and $T = 10$. A comparison between the two pictures shows that the two frameworks allow to compute the same control functions.

7.2.2 A case of bang-bang exact-control of two uncoupled spins

In this section, we consider an exact-control problem in the case that the control is pointwise in time bounded, that is $u \in U_{ad,1}$. We remark that, the literature regarding controllability results in the case of bounded controls is at its infancy. For this reason, we refer to the work done by Assémat et al. in [9], where the exact-control of a system of two uncoupled spins is considered. In particular, similarly as in Section 2.3.4, we have the following system

$$\dot{x} = 2\pi \big(\tilde{A} + u\tilde{B}\big)x \, ,$$

where

$$\tilde{A} = c \begin{pmatrix} A & 0 \\ 0 & -A \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \, ,$$

with $c = 483$, and

$$\tilde{B} = \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix} \quad \text{with} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \, .$$

We seek a control $u : [0,T] \to \mathbb{R}$, such that $|u(t)| \leq 120.77$, that is capable to perform a transition from an initial state where both the spins are pointing in the z-direction, to a target state where both the spins are pointing in the $-z$-direction. Consequently, starting and target vectors are

$$x_0 = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \, ,$$

and

$$x_T = \begin{pmatrix} 0 & 0 & -1 & 0 & 0 & -1 \end{pmatrix} \, .$$

In [9] is proved that the optimal time to perform such a transition is $T = 0.006409$, which is the value that we assume.

In order to solve this exact-control problem, we apply our continuation method. The time interval is discretized with $N_t = 20001$ points and corresponds to a value of the regularization parameter of the order of 10^{-8}. The corresponding fidelity $C = \frac{\langle x(T), x_T \rangle}{\|x(T)\|_2 \|x_T\|_2}$ is equal to 0.999999. The obtained exact-control is shown in the Figure 7.23. This is an exact-control with a bang-bang structure, and is exactly the same as presented in [9].
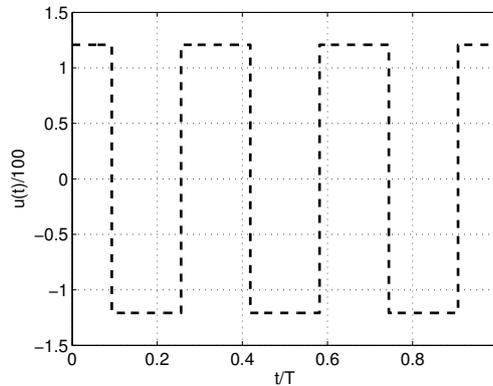

Figure 7.23: Bang-bang control function computed for the inversion of a system of two uncoupled spins having different Larmor frequency.

In order to demonstrate the validity Theorem 27 and estimate (5.19), we show results corresponding to different number of discretization points. In particular, similarly as in Figure 7.4 and Figure 7.17, the two pictures in Figure 7.24 describe the decay of the norm $\|u_\nu - u_{max}\|_{L^2}$ with respect to the regularization parameter ν. We compare this decay with the two functions $f(\nu) = C_1\nu$ and $f(\nu) = C_2\nu^{1/2}$. In particular, in the left picture, we show the decay of $\|u_\nu - u_{max}\|_{L^2}$ for different number of discretization points. It is clear that, as the number of discretization point increases, the discretization error induces effects for smaller values of ν. Furthermore, in order to demonstrate the numerical validity of (5.19), we perform a curve-fitting, in order to compute two parameters α and β such that the curve $\alpha\nu^\beta$ fits the black curve corresponding to $N_t = 20001$. In particular, we exclude from this analysis the points corresponding to large values of ν and the points corresponding to the part of the curve affected by numerical errors. The considered part of the fitting analysis is shown in Figure 7.23 (right). The computed parameters are $\alpha = 666.758$ and $\beta = 0.528$, which show that the decay $\|u_\nu - u_{max}\|_{L^2}$ is slightly faster than the estimated $O(\sqrt{\nu})$.
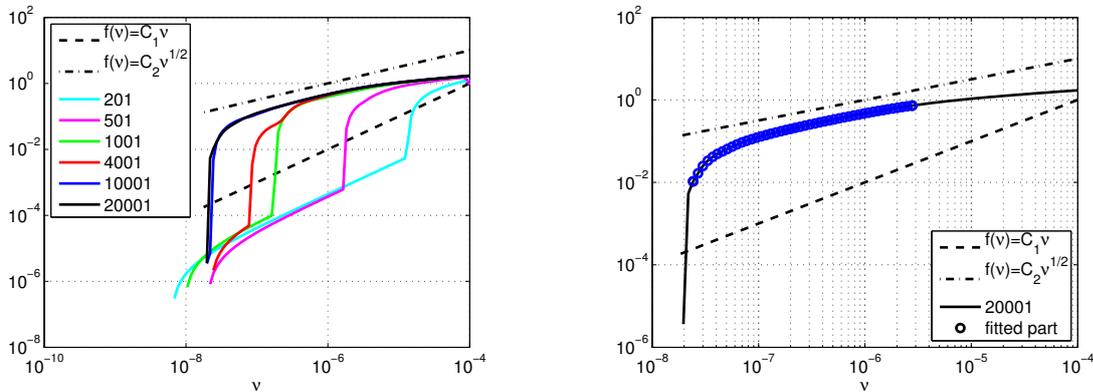

Figure 7.24: The picture on the left shows the decay of $\|u_\nu - u_{max}\|_{L^2}$ with respect to ν for different number of discretization points. The picture on the right shows the part of the black curve used in the fitting analysis that demonstrate the validity of the estimate (5.19).

7.2.3 Exact-control of the Pauli equation

In this section, we consider numerical experiments that demonstrates the validity of the presented computational methods for the exact-control of the Pauli equation (2.79)-(2.82) for $V(r) = 0$ and with $U_{ad} = L^2((0,T); \mathbb{R}^{N_C})$. In particular, we initialize the numerical solution by means of the continuation method described in Section 5.2, and then we address problem (5.36) by means of a Krylov-Newton method, where the terminal state of the adjoint p_T is used as a shooting-variable. This strategy is summarized in the following algorithm.

Algorithm 21 (Exact control method)

- call Algorithm 13 to obtain an initialization u_{init};

  while $C < tol$ **do**

  - Call a Krylov-Newton solver (Algorithm 18) to solve (5.36) and minimize G_r;

  - Refine the discretization mesh and interpolate the current solution to the new mesh;

  end while

We consider numerical experiments in which we want to perform in time $T = 1$ a spin inversion and a transition from $m_0 = -\ell$ to $m_d = \ell$, with $\ell = 1$, 2, 4, 6, 8 and 10.

Some of the controls obtained by the presented Algorithm 21 are shown in the next Figure 7.25. In particular, the left and the right picture show the controls obtained for $\ell = 1$ and $\ell = 10$, respectively. Notice that the obtained control solutions are capable to obtain values of G and C vary close to 0 and 1, respectively, that correspond to the exact control. This is expressed in the next Table 7.30, in which the value of the fidelity C, the functional G, the norm of the reduced gradient of G, and the number of discretization points for the time interval $[0, T]$ are shown. Moreover, in order to show that the continuation procedure is capable to provide an adequate initialization, the values C_{init} is considered. This correspond to the fidelity obtained by the initialization, performed by means of the continuation procedure (Algorithm 13).

ℓ	C_{init}	C	G	$\|\nabla G_r\|_{L^2}$	N_t
1	0.9997635	0.9999999	$2.38 \cdot 10^{-11}$	$1.98 \cdot 10^{-11}$	1601
2	0.9999518	0.9999999	$7.15 \cdot 10^{-11}$	$8.22 \cdot 10^{-11}$	1601
4	0.9998752	0.9999999	$1.45 \cdot 10^{-9}$	$9.44 \cdot 10^{-12}$	1601
6	0.9999286	0.9999999	$5.13 \cdot 10^{-10}$	$6.48 \cdot 10^{-11}$	3201
8	0.9997721	0.9999999	$3.47 \cdot 10^{-10}$	$2.91 \cdot 10^{-9}$	4801
10	0.9998180	0.9999999	$8.63 \cdot 10^{-10}$	$2.47 \cdot 10^{-9}$	4801

Table 7.30: Results obtained from the numerical experiments. In particular, C_{init} is the fidelity corresponding to the control obtained by the initialization procedure (continuation method in Algorithm 13); C is the fidelity corresponding to the control solution obtained by Algorithm 21; G_r is the value of the cost functional of the problem (5.36) evaluated in the computed controls; $\|\nabla G_r\|_{L^2}$ is the norm of the reduced gradient of (5.36); N_t is the number of the points used for the discretization of the time interval $[0, T]$.
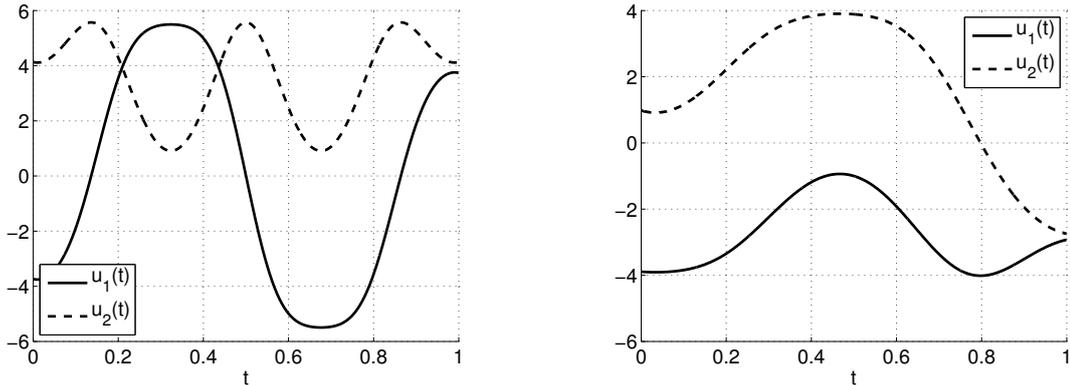

Figure 7.25: Exact control functions obtained by means of the presented computational scheme in the cases $\ell = 1$ and $\ell = 10$. In particular, the controls in the left picture correspond to $\ell = 1$, and the controls in the right picture correspond to $\ell = 10$. These control functions perform the spin inversion and the transition from $m = -\ell$ to $m = \ell$.

7.3 Problems of controlling distributed spin systems

In this section, we present numerical results obtained for the control of distributed spin systems. As discussed in Section 2.3.4, these systems are characterized by a density operator $\rho \in \mathfrak{her}(N_g N)$ with having the following block-diagonal form

$$\rho = \text{blk-diag}_{l=1,\dots,N_g} (\rho_l) .$$

The evolution of such an operator is governed by the LvNM equation

$$\dot{\rho} = -i[H, \rho] .$$

The Hamiltonian H is the following block-diagonal matrix

$$H = \begin{pmatrix} \tilde{H}_1 & & & & \\ & \ddots & & & \\ & & \tilde{H}_l & & \\ & & & \ddots & \\ & & & & \tilde{H}_g \end{pmatrix}$$

where the l-th block is given by

$$\tilde{H}_l = \alpha_l \tilde{H}_0 + \hat{\alpha}_l \tilde{H}_{rf} ,$$

with α_l and $\hat{\alpha}_l$ are real parameters, and $\tilde{H}_0$ and $\tilde{H}_{rf}$ are defined in (2.107) and (2.108), respectively. We refer to α_l as detuning parameters and to $\hat{\alpha}_l$ as inhomogeneity parameters. The parameters α_l and $\hat{\alpha}_l$ represents inhomogeneities of the stationary magnetic-field along the z-axis and of the control magnetic fields, respectively.

Consider inhomogeneous spin systems with $N = 2$ that are controlled by means of two control functions. A real matrix representation is obtained similarly as in Section 2.4.1, and the LvNM equation becomes as follows

$$\dot{x} = 2\pi \big(\tilde{A} + u_1 \tilde{B}_1 + u_2 \tilde{B}_2\big) x , \tag{7.4}$$

where

$$\tilde{A} = \begin{pmatrix} \alpha_1 A & & \\ & \ddots & \\ & & \alpha_g A \end{pmatrix} \quad \text{with} \quad A = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} ,$$

$$\tilde{B}_1 = \begin{pmatrix} \hat{\alpha}_1 B_1 & & \\ & \ddots & \\ & & \hat{\alpha}_g B_1 \end{pmatrix} \quad \text{with} \quad B_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} .$$

and

$$\tilde{B}_2 = \begin{pmatrix} \hat{\alpha}_1 B_2 & & \\ & \ddots & \\ & & \hat{\alpha}_g B_2 \end{pmatrix} \quad \text{with} \quad B_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} .$$

We consider three numerical experiments that are of interest in NMR spectroscopy [75, 105]. In these experiments, we want to perform specific spin transitions by means of control functions $u \in U_{ad,2}$, with $U_{ad,2}$ defined in (3.3). The characteristics of these experiments are given in the following Table 7.31. In this table, we have

- "Detuning": it represents the interval of detuning parameters; in particular, the detuning parameters α_l are obtained by discretizing the detuning interval with a uniform mesh, and the number of discretization points is given by "Det. points";

- "Inhomog.": it represents the interval of inhomogeneity parameters; in particular, the inhomogeneity parameters α_l are obtained by discretizing the inhomogeneity interval with a uniform mesh, and the number of discretization points is given by "Inh. points";

- T: time of the experiment;

- b: value of the bound on the controls; it is used to define $U_{ad,2}$ in (3.3);

- "transition": it represents the desired spin transition; in particular, $z \to y$ means that we aim to change the spin orientation from the z direction to the y direction, and $z \to -z$ means that we aim to change the spin orientation from the z direction to the $-z$ direction (spin inversion).

Case	Detuning	Det. points	Inhomog.	Inh. points	T	b	transition
1	$[-15\cdot10^3,15\cdot10^3]$	201	-	-	$82.5\cdot10^{-6}$	10^4	$z \to y$
2	$[-10\cdot10^3,10\cdot10^3]$	201	-	-	$92.5\cdot10^{-6}$	10^4	$z \to -z$
3	$[-10\cdot10^3,10\cdot10^3]$	201	$[0.8,1.20]$	5	$192.5\cdot10^{-6}$	10^4	$z \to y$

Table 7.31: Characteristics of the considered experiments.

To solve these spin control problems, we use the continuation method discussed in Chapter 5 and 6. We remark that, it is not proved that the considered inhomogeneous spin systems are controllable. Hence, we do not expect exact-control results, however, the experiments aim to obtain values of the fidelity C, defined in (7.1), as close as possible to 1. The results of the optimizations are given in the following. In particular, Table 7.32 summarizes the obtained results, and shows the number of discretization point N_t of the time interval $[0,T]$, the reached value of weight parameter ν and the obtained fidelity C. The corresponding control functions are depicted in Figure 7.26, Figure 7.27 and Figure 7.28, respectively. Notice that, even if the considered inhomogeneous spin systems are not proved to be controllable, our computational scheme allows to obtain high values of the fidelity. Moreover, we remark that, reducing the value of ν does not significantly increase the obtained fidelities.

Case	N_t	ν	C
1	101	$10^{-3} \cdot 0.5^{10}$	0.916918
2	101	$10^{-3} \cdot 0.5^{10}$	0.791679
3	101	$10^{-3} \cdot 0.5^{10}$	0.974128

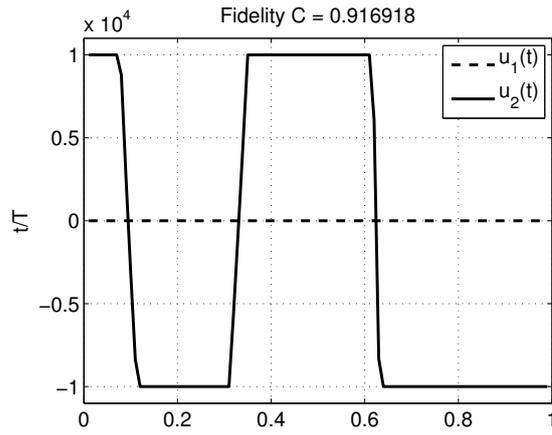Table 7.32: Characteristics of the considered experiments.

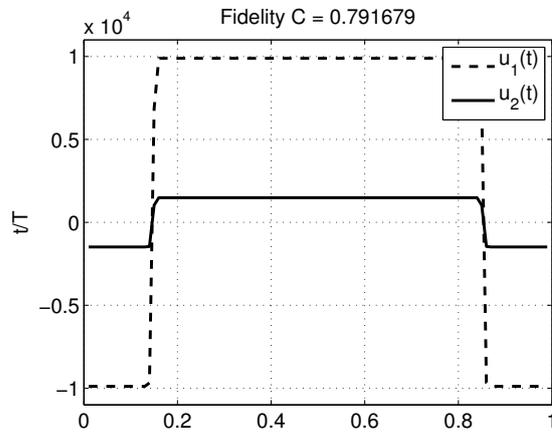Figure 7.26: Case 1: Optimal controls functions.
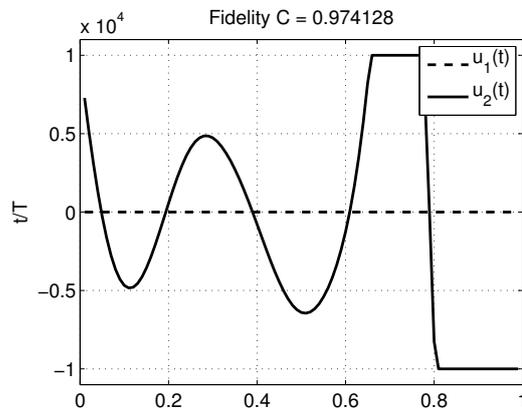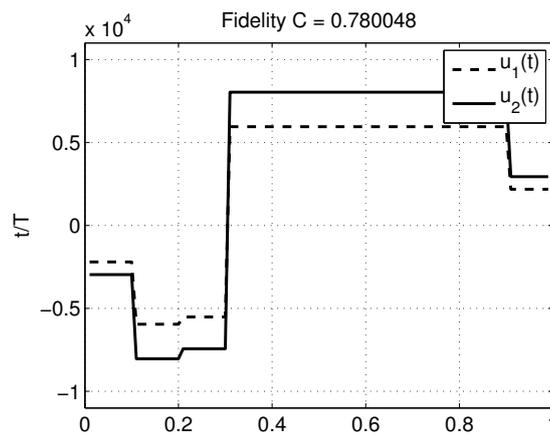

Figure 7.27: Case 2: Optimal controls functions.

Figure 7.28: Case 3: Optimal controls functions.

Next, we repeat the same experiments in the case of piecewise-constant controls. The considered number of piecewise-constant subintervals is $M = 10$ (see Section 3.3). The obtained results are shown in the following Table 7.33 and figures 7.29, 7.30 and 7.31. We remark that, the piecewise-constant optimal control are capable to obtain almost the same values of fidelities shown in Table 7.32.

Case	N_t	ν	C
1	101	$10^{-3} \cdot 0.5^{10}$	0.908541
2	101	$10^{-3} \cdot 0.5^{10}$	0.780048
3	101	$10^{-3} \cdot 0.5^{10}$	0.970660

Table 7.33: Characteristics of the considered experiments.
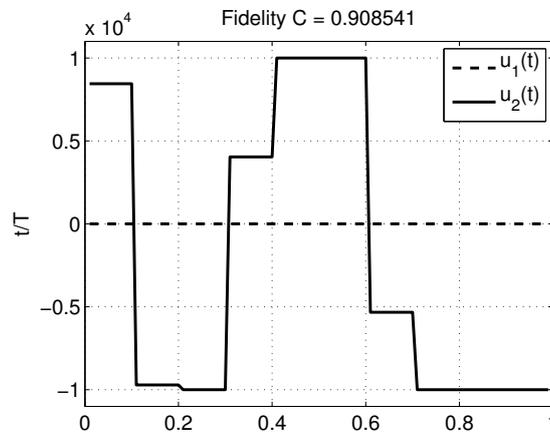
Figure 7.29: Case 1: Piecewise-constant optimal controls functions.

Figure 7.30: Case 2: Piecewise-constant optimal controls functions.
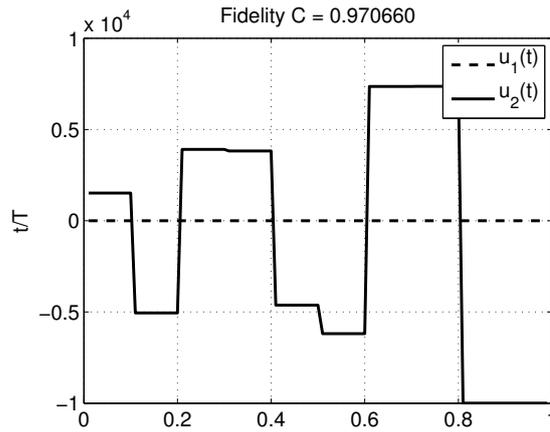
Figure 7.31: Case 3: Piecewise-constant optimal controls functions.

7.4   Summary and remarks

In this section, numerical experiments regarding the control of quantum systems were shown. The experiments, involving the control of spin systems that are of interest in NMR spectroscopy, were used to demonstrate the validity of the computational framework developed in this thesis. We showed that, the developed semi-smooth Newton method can be successfully used to address quantum optimal control problems with L^1-penalization term in the cost functional and with piecewise-constant control functions. An infinite-dimensional dipole-control problem was also considered, and the SSN method was capable to obtain sparse and piecewise-constant controls. Moreover, several exact-control problems governed by the LvNM equation and by the Pauli equation were addressed by means of the two methodologies developed in Chapter 5. Finally, we considered the control of inhomogeneous spin systems, which are in general considered in NMR experiments. Also in these last cases, the methodologies presented in this thesis were able to address successfully the considered quantum spin control problems.

Appendix A

Appendix

A.1 The angular equations and the spherical harmonics functions

The spherical harmonics, denoted with $Y_\ell^m(\theta, \phi)$, are eigen-functions of the operator

$$L^2 = L_x^2 + L_y^2 + L_z^2 \, ,$$

i.e., $Y_\ell^m(\theta, \phi)$ solve the following eigen-problem

$$L^2 Y_\ell^m(\theta, \phi) = \hbar^2 \ell(\ell + 1) Y_\ell^m(\theta, \phi) \, . \tag{A.1}$$

Since L^2 is an Hermitian operator, its eigen-functions $Y_\ell^m(\theta, \phi)$ satisfy the following orthogonality condition

$$\int_0^{2\pi} \int_0^\pi Y_\ell^m(\theta, \phi) Y_{\ell'}^{m'}(\theta, \phi) \sin\theta \, \mathrm{d}\theta d\phi = \delta_{\ell, \ell'} \delta_{m, m'} \, ,$$

where $\delta_{mm'}$ is the Kronecker delta, and have the following parity property

$$Y_\ell^m(\pi - \theta, \phi + \pi) = (-1)^\ell Y_\ell^m(\theta, \phi) \, .$$

Further, we have the following relations

$$L_z Y_\ell^m(\theta, \phi) = \hbar m Y_\ell^m(\theta, \phi)$$
$$L_\pm Y_\ell^m(\theta, \phi) = \hbar \sqrt{\ell(\ell + 1) - m(m \pm 1)} Y_\ell^{m\pm 1}(\theta, \phi) \, ,$$

where $L_\pm = L_x \pm i L_y$.

Spherical harmonics are computed as a product of associated Legendre functions and complex exponentials

$$Y_\ell^m(\theta, \phi) = c_{\ell, m} P_\ell^m(\cos\theta) e^{im\phi} \, , \tag{A.2}$$

where the normalization constant is

$$c_{\ell, m} = (-1)^m \sqrt{\frac{(2\ell + 1)(\ell - |m|)!}{4\pi(\ell + |m|)!}} \, .$$

The separation of variable which allows to write $Y_\ell^m(\theta, \phi)$ in the form (A.2) is explained in the next subsection.

A.1.1 The polar and azimuthal angular equations

We consider the angular equation given by (A.1) and (2.24). It is explicitly given by the following

$$-\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)Y(\theta,\phi) - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}Y(\theta,\phi) = \ell(\ell+1)Y(\theta,\phi) ,\qquad (A.3)$$

and we analyze its solutions. To begin with, we divide (A.3) by $Y(\theta,\phi)$ and we get

$$-\frac{1}{Y(\theta,\phi)\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)Y(\theta,\phi) - \frac{1}{Y(\theta,\phi)\sin^2\theta}\frac{\partial^2}{\partial\phi^2}Y(\theta,\phi) = \ell(\ell+1) .\qquad (A.4)$$

Using spherical harmonics, we write $Y(\theta,\phi) = Y_1(\theta)Y_2(\phi)$ and replacing it into (A.4) we obtain

$$-\frac{1}{Y_1(\theta)\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)Y_1(\theta) - \frac{1}{Y_2(\phi)\sin^2\theta}\frac{\partial^2}{\partial\phi^2}Y_2(\phi) = \ell(\ell+1) .\qquad (A.5)$$

Multiplying (A.5) with $\sin^2\theta$, we have

$$\frac{\sin\theta}{Y_1(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)Y_1(\theta) + \ell(\ell+1)\sin^2\theta = -\frac{1}{Y_2(\phi)}\frac{\partial^2}{\partial\phi^2}Y_2(\phi) .\qquad (A.6)$$

The only non-trivial solution of (A.6) has to satisfy the following polar equation

$$\frac{\sin\theta}{Y_1(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)Y_1(\theta) + \ell(\ell+1)\sin^2\theta = m^2 ,\qquad (A.7)$$

and azimuthal equation

$$-\frac{1}{Y_2(\phi)}\frac{\partial^2}{\partial\phi^2}Y_2(\phi) = m^2 ,\qquad (A.8)$$

where we consider m^2 as the separation constant. The solution to the azimuthal equation (A.8) is given by

$$Y_{2,m}(\phi) = e^{im\phi} ,$$

where the subscript m means that the solution depends on the constant m.

Now, we consider the polar equation (A.7) and we rearrange it as follows

$$\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)Y_1(\theta) + \ell(\ell+1)\sin^2\theta\, Y_1(\theta) - m^2 Y_1(\theta) = 0 ,\qquad (A.9)$$

The first term can be written as

$$\begin{aligned}
\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)Y_1(\theta) &= \sin\theta\left(\cos\theta\frac{\partial}{\partial\theta} + \sin\theta\frac{\partial^2}{\partial\theta^2}\right)Y_1(\theta) \\
&= \left(\sin\theta\cos\theta\frac{\partial}{\partial\theta} + \sin^2\theta\frac{\partial^2}{\partial\theta^2}\right)Y_1(\theta) .
\end{aligned}\qquad (A.10)$$

Consequently, (A.9) becomes

$$\left(\sin\theta\cos\theta\frac{\partial}{\partial\theta} + \sin^2\theta\frac{\partial^2}{\partial\theta^2}\right)Y_1(\theta) + \ell(\ell+1)\sin^2\theta\, Y_1(\theta) - m^2 Y_1(\theta) = 0 .\qquad (A.11)$$

Now, we apply a change of variable using $x(\theta) = \cos\theta$. We have that

$$\frac{d\,Y_1(x(\theta))}{d\theta} = \frac{d\,Y_1(x)}{dx}\frac{d\,x(\theta)}{d\theta} = -\sin\theta\frac{d\,Y_1(x)}{dx}\ , \tag{A.12}$$

and

$$\frac{d^2\,Y_1(x(\theta))}{d\theta^2} = -\cos\theta\frac{d\,Y_1(x)}{dx} + \sin^2\theta\frac{d^2\,Y_1(x)}{dx^2}\ . \tag{A.13}$$

Substituting the above equations in (A.11) we obtain

$$\left(-\sin^2\theta\cos\theta\frac{\partial}{\partial x} + \sin^2\theta\left(-\cos\theta\frac{\partial}{\partial x} + \sin^2\theta\frac{\partial^2}{\partial x^2}\right)\right)Y_1(x)$$
$$+ \ell(\ell+1)\sin^2\theta\,Y_1(x) - m^2 Y_1(x) = 0\ . \tag{A.14}$$

Dividing (A.14) by $\sin^2\theta$ we get

$$\left(-2\cos\theta\frac{\partial}{\partial x} + \sin^2\theta\frac{\partial^2}{\partial x^2}\right)Y_1(x) + \ell(\ell+1)\,Y_1(x) - \frac{m^2}{\sin^2\theta}Y_1(x) = 0\ . \tag{A.15}$$

Now, using the fact that $\cos\theta = x$, and $\sin^2\theta = 1 - \cos^2\theta = 1 - x^2$ we obtain

$$\left(-2x\frac{\partial}{\partial x} + (1-x)^2\frac{\partial^2}{\partial x^2}\right)Y_1(x) + \ell(\ell+1)\,Y_1(x) - \frac{m^2}{(1-x)^2}Y_1(x) = 0\ . \tag{A.16}$$

This equation is the associated Legendre equation, whose solutions are the associated Legendre polynomials which are defined in terms of derivatives of ordinary Legendre polynomials

$$P_\ell^m(x) = (-1)^m(1-x^2)^{m/2}\frac{d^m}{dx^m}P_\ell(x)\ , \tag{A.17}$$

where $P_\ell(x)$ are the Legendre polynomials. The associated Legendre polynomials are described in the following subsection.

A.1.2 Associated Legendre polynomials

The associated Legendre polynomials solves equation (A.16) and are given by

$$P_\ell^m(x) = (-1)^m(1-x^2)^{m/2}\frac{d^m}{dx^m}P_\ell(x)\ , \tag{A.18}$$

where $P_\ell(x)$ is the ℓ-th Legendre polynomial, which is conveniently defined by the following Rodrigues formula

$$P_\ell(x) = \frac{1}{2^\ell\ell!}\frac{d^\ell}{dx^\ell}\left[(x^2-1)^\ell\right]\ . \tag{A.19}$$

It can be shown that $P_\ell^m(x)$ can be computed recursively by the following formulas

$$P_m^m(x) = a_m^m(1-x^2)^{|m|/2} \tag{A.20}$$
$$P_{m+1}^m(x) = a_{m+1}^m x P_m^m(x) \tag{A.21}$$
$$P_\ell^m(x) = a_\ell^m x P_{\ell-1}^m(x) + b_\ell^m P_{\ell-2}^m(x)\ , \tag{A.22}$$

with

$$a_m^m = \sqrt{\frac{1}{4\pi} \prod_{k=1}^{|m|} \frac{2k+1}{2k}} \tag{A.23}$$

$$a_\ell^m = \sqrt{\frac{4\ell^2 - 1}{\ell^2 - m^2}} \tag{A.24}$$

$$b_\ell^m = -\sqrt{\frac{2\ell + 1}{2\ell - 3} \frac{(\ell - 1)^2 - m^2}{\ell^2 - m^2}} \ . \tag{A.25}$$

A.2   The radial equation and its wavefunctions

Consider the following operator

$$O_\ell(V) := \left\{ -\frac{\hbar^2}{2\mu} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{\hbar^2}{2\mu} \frac{\ell(\ell+1)}{r^2} + V(r) \right\} \tag{A.26}$$

then, the radial equation is defined as the following eigen-problem

$$O_\ell(V) R_{n,\ell}(r) = \lambda_{n,\ell} R_{n,\ell}(r) \ . \tag{A.27}$$

The eigenfunctions $R_{n,\ell}(r)$ satisfy the following orthogonality condition

$$\int_0^\infty r^2 R_{n,\ell}(r) R_{n',\ell'}(r) dr = \delta_{\ell,\ell'} \delta_{n,n'} \ .$$

Since the operator $O_\ell(V)$ depends on the potential $V(r)$ its eigen-functions $R_{n,\ell}(r)$ have to be computed in different cases of physical interest. In particular, we recall here the solutions to (2.36) in the following main cases

- $V(r) := -\frac{e^2}{r}$, which is the Coulomb potential, used for describing the behaviour of an Hydrogen atom;

- $V(r) := \frac{1}{2}\mu\omega r^2$, which is used for defining the three dimensional harmonic oscillator;

- $V(r) := \begin{cases} \infty & \text{if } r > a \\ 0 & \text{if } r \le a \end{cases}$, with $a > 0$ which is the so-called infinite spherical well.

We remark that in the literature it is possible to find solutions to (A.27) for other potentials of physical interest, see, e.g., [63].

A.2.1   The Coulomb potential

We consider the Coulomb potential, which is used to describe an Hydrogen atom. It is given by the following

$$V(r) := -\frac{e^2}{r} \tag{A.28}$$

The radial equation becomes as follows

$$\left\{ -\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{\hbar^2}{2\mu}\frac{\ell(\ell+1)}{r^2} - \frac{e^2}{r} \right\} R_{n,\ell}(r) = \lambda_{n,\ell}R_{n,\ell}(r) . \tag{A.29}$$

Its solutions are given by the following

$$R_{n,\ell}(r) = c_{n,\ell}e^{-\frac{r}{na_0}}\left(\frac{2r}{na_0}\right)^{\ell} L_{n-\ell-1}^{2\ell+1}\left(\frac{2r}{na_0}\right) , \tag{A.30}$$

where

$$a_0 = \frac{\hbar^2}{\mu e^2} , \tag{A.31}$$

is the so-called Bohr radius. The coefficients $c_{n,\ell}$ are given by the following

$$c_{n,\ell} = \sqrt{\left(\frac{2}{na_0}\frac{(n-\ell-1)!}{2n[(n+\ell)!]^3}\right)^3} . \tag{A.32}$$

The functions $L_n^\alpha(x)$ are the generalized Laguerre polynomials, which are given by the following

$$L_n^\alpha(x) = \frac{x^{-\alpha}e^x}{n!}\frac{d^n}{dx^n}\left(e^{-x}x^{n+\alpha}\right) , \tag{A.33}$$

and they can be computed by means of the following recursive formulas

$$L_0^\alpha(x) = 1 \tag{A.34}$$

$$L_1^\alpha(x) = 1 + \alpha - x \tag{A.35}$$

$$L_{n+1}^\alpha(x) = \frac{(2n+1+\alpha-x)L_n^\alpha(x) - (n+\alpha)L_{n-1}^\alpha(x)}{n+1} . \tag{A.36}$$

The corresponding energies are given by the following

$$\lambda_{n,\ell} = -\frac{\hbar^2}{2\mu a_0^2 n^2} . \tag{A.37}$$

A.2.2   The harmonic oscillator

We consider the harmonic oscillator, which is obtained by means of the following potential

$$V(r) := \frac{1}{2}\mu\omega^2 r^2 . \tag{A.38}$$

Hence, the radial equation becomes as follows

$$\left\{ -\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{\hbar^2}{2\mu}\frac{\ell(\ell+1)}{r^2} + \frac{1}{2}\mu\omega^2 r^2 \right\} R_{n,\ell}(r) = \lambda_{n,\ell}R_{n,\ell}(r) . \tag{A.39}$$

Its solutions are given by the following

$$R_{n,\ell}(r) = r^\ell e^{-\frac{\beta^2 r^2}{2}}\sum_{q=0}^{\infty} c_q r^q , \tag{A.40}$$

where

$$\beta = \sqrt{\frac{\mu\omega}{\hbar}} \ . \tag{A.41}$$

The coefficients c_q are given by the following formulas

$$c_0 \neq 0 \tag{A.42}$$

$$c_1 = 0 \tag{A.43}$$

$$c_{q+2} = \frac{\left[(2q + 2\ell + 3)\beta^2 - \epsilon_{n,\ell}\right]}{(q+2)(q+2\ell+3)} c_q \ , \tag{A.44}$$

where

$$\epsilon_{n,\ell} = \left(2n + 2\ell + 3\right)\beta^2 \ . \tag{A.45}$$

Notice that according to (A.42)-(A.44) the coefficients a_q with q odd are zero.

The corresponding energies are given by the following

$$\lambda_{n,\ell} = \hbar\omega\left(n + \ell + \frac{3}{2}\right) \ . \tag{A.46}$$

A.2.3   The infinite spherical well

We consider the infinite spherical well potential, which is defined as follows

$$V(r) := \begin{cases} \infty & \text{if } r > a \\ 0 & \text{if } r \leq a \end{cases} \ . \tag{A.47}$$

with $a > 0$. Outside the well the wave function is zero. Inside the well, the radial equation (2.36) become as follows

$$\left\{-\frac{\hbar^2}{2\mu}\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r} + \frac{\hbar^2}{2\mu}\frac{\ell(\ell+1)}{r^2}\right\} R_{n,\ell}(r) = \lambda_{n,\ell}R_{n,\ell}(r) \ . \tag{A.48}$$

Its solutions are given by the following

$$R_{n,\ell}(r) = \sqrt{\frac{2}{a}}\, j_\ell\left(\frac{\beta_{n,\ell}}{a}r\right) \ , \tag{A.49}$$

where $j_\ell(x)$ are the spherical Bessel functions that are given by

$$j_\ell(x) = (-x)^\ell\left(\frac{1}{x}\frac{d}{dx}\right)^\ell\frac{\sin x}{x} \ , \tag{A.50}$$

and $\beta_{n,\ell}$ is the n-th zero of the ℓ-th spherical Bessel function. The corresponding energies are given by

$$\lambda_{n,\ell} = \frac{\hbar}{2\mu a^2}\beta_{n,\ell}^2 \ . \tag{A.51}$$

The Bessel functions are oscillatory and each one has an infinite number of zeros. They have to be computed numerically [51].

A.3 Lie algebra and Lie groups

In this section, we state some main definitions and results regarding matrix Lie groups and Lie algebras. Our main references are [53, 67, 93].

A matrix Lie group is defined as follows.

Definition A 1. *Let $GL(n; \mathbb{C})$ denote the general linear group of all $n \times n$ matrices with complex entries. A matrix Lie group is any subgroup G of $GL(n; \mathbb{C})$ with the following property: if $\{A_m\}_{m=1}^{\infty}$ is any sequence of matrices in G, and $A_m \to A$, then either $A \in G$, or A is not invertible.*

Two important properties of a matrix Lie group are compactness and connectedness. They are defined in the following.

Definition A 2. *A matrix Lie group G is said to be compact if the following two conditions are satisfied:*

- *if $\{A_m\}_{m=1}^{\infty}$ is any sequence of matrices in G, and A_m converges to a matrix A, then A is in G;*
- *there exists a constant C such that for all $A \in G$, $|A_{j,k}| \leq C$ for all $1 \leq j, k \leq n$.*

Notice that, this definition says that G is compact if it is a closed and bounded subset of $\mathbb{C}^{n^2}$.

Definition A 3. *A matrix Lie group G is said to be (path-)connected if given two matrices A and B in G, there exists a continuous path $A(t)$, $a \leq t \leq b$, lying in G with $A(a) = A$ and $A(b) = B$.*

A Lie algebra corresponding to a matrix Lie group is defined as follows.

Definition A 4. *Let G be a matrix Lie group. The Lie algebra of G, denoted by $\mathfrak{g}$, is the set of all matrices X such that $\exp(tX)$ is in G for all $t \in \mathbb{R}$.*

A Lie algebra can be defined in a general way (independently on a matrix Lie group) as a vector space endowed with a particular bilinear operation. This is done in the following definition.

Definition A 5. *A finite-dimensional real (complex) Lie algebra is a finite-dimensional real (complex) vector space $\mathfrak{g}$, together with a map $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, with the following properties:*

- *$[\cdot, \cdot]$ is bilinear;*
- *$[X, Y] = -[X, Y]$ for all $X, Y \in \mathfrak{g}$;*
- *$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ for all $X, Y, Z \in \mathfrak{g}$.*

The following theorem states that a Lie algebra corresponding to a matrix Lie group (Definition 4) is a real vector space and a Lie algebra in the sense of Definition 5.

Theorem A 1. *Let G be a matrix Lie group, $\mathfrak{g}$ its Lie algebra, and X and Y elements of $\mathfrak{g}$. Then it holds that*

- *$sX \in \mathfrak{g}$ for all real numbers s;*
- *$X + Y \in \mathfrak{g}$;*
- *$XY - YX \in \mathfrak{g}$.*

Moreover, the Lie algebra $\mathfrak{g}$ of G is a real Lie algebra (according to Definition A 5).

An important property of a matrix Lie group is the so-called semisimplicity. In particular, a matrix Lie group G is said to be semisimple if the corresponding Lie algebra $\mathfrak{g}$ is semisimple. In order to define semisimplicity of a Lie algebra $\mathfrak{g}$, we need to introduce the definitions of ideal and simple Lie algebra, as follows.

Definition A 6. *Let $\mathfrak{g}$ be a (complex) Lie algebra, then an ideal in $\mathfrak{g}$ is a complex subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ with the property that for all X in $\mathfrak{g}$ and H in $\mathfrak{h}$, we have that $[X, H] \in \mathfrak{h}$.*

Definition A 7. *A (complex) Lie algebra $\mathfrak{g}$ is said to be simple if the only ideals in $\mathfrak{g}$ are $\mathfrak{g}$ and $\{0\}$ and $\dim \mathfrak{g} \geq 2$.*

A semisimple Lie algebra is defined as follows.

Definition A 8. *A (complex) Lie algebra $\mathfrak{g}$ is said to be semisimple if it is isomorphic to a direct sum of simple Lie algebras.*

We remark that, it could be difficult to establish semisimplicity of a Lie algebra by using Definition 8. In order to obtain a simpler and more practical condition to establish semisimplicity, consider the so-called Killing form on $\mathfrak{g}$, that is a bilinear form $\mathcal{K} : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{K}$, where $\mathbb{K}$ stays for $\mathbb{R}$ or $\mathbb{C}$, defined as follows; see, e.g., [40, 53, 67];

$$\mathcal{K} : (A, B) \mapsto \operatorname{trace}(\operatorname{ad}_A \circ \operatorname{ad}_B) \, , \text{ for any } A, B \in \mathfrak{g} \, ,$$

where the linear operator $\operatorname{ad}_A : \mathfrak{g} \rightarrow \mathfrak{g}$ is defined as $\operatorname{ad}_A(\cdot) := [A, \cdot]$. The Killing form allows to characterize a semisimple Lie group by means of the following theorem; see, e.g., [40, 53, 67].

Theorem A 2. *A matrix Lie group is called semisimple if its corresponding Lie algebra $\mathfrak{g}$ is semisimple, that is the so-called Killing form is nondegenerate on $\mathfrak{g}$:*

$$\operatorname{trace}(\operatorname{ad}_A \circ \operatorname{ad}_B) = 0 \quad \forall A \in \mathfrak{g} \quad \Rightarrow \quad B = 0 \, ,$$

and

$$\operatorname{trace}(\operatorname{ad}_A \circ \operatorname{ad}_B) = 0 \quad \forall B \in \mathfrak{g} \quad \Rightarrow \quad A = 0 \, .$$

A.4 Results of functional analysis

In this section we recall some important concept and results of functional analysis that are used in the present thesis. Our main references are [33, 47, 66, 96, 119]. In particular, we recall definition and results regarding weak convergence, semicontinuity and embeddings.

Definition A 9. *Let X be a normed vector space and let X^* denote its dual. A sequence $\{x_n\}_{n=1}^\infty$ of elements $x_n \in X$ is said to converge weakly in X if there exists $x \in X$ such that*

$$\text{for each } x^* \in X^* \, , \quad x^*(x_n) \rightarrow x^*(x) \quad \text{as } n \rightarrow \infty \, ,$$

and such an x is called the weak limit of the sequence $\{x_n\}_{n=1}^\infty$. Weak convergence is usually denoted by $x_n \rightharpoonup x$.

Theorem A 3. *Let X and Y be normed vector spaces over the same field $\mathbb{K}$. It holds:*

- *let $A \in \mathcal{L}(X, Y)$, then*

$$x_n \rightharpoonup x \ in \ X \quad implies \quad Ax_n \rightharpoonup Ax \ in \ Y \ ;$$

- *let $B \in \mathcal{L}(X \times Y, \mathbb{K})$, then*

$$x_n \rightharpoonup x \ in \ X \ \ and \ \ y_n \rightarrow y \ in \ Y \quad implies \quad B(x_n, y_n) \rightarrow B(x, y) \ in \ \mathbb{K} \ .$$

Theorem A 4 (Banach-Saks-Mazur)**.** *Let X be a real normed vector space. Let C be a non-empty, convex, and closed subset of X, and let $\{x_k\}_{k=1}^{\infty}$ be a sequence of points $x_k \in C$ that weakly converges to $x \in X$ as $k \rightarrow \infty$. Then the weak limit x belongs to C.*

Theorem A 5 (Banach-Eberlein-Šmulian)**.**

(a) *Any bounded sequence in a reflexive Banach space contains a weakly convergent subsequence.*

(b) *Conversely, a Banach space in which every bounded sequence contains a weakly convergent subsequence is reflexive.*

A direct consequence of the previous theorems is the following important result.

Theorem A 6. *Let X be a reflexive Banach space. Let C be a non-empty, convex, closed and bounded subset of X. Then C is weakly sequentially compact, that is every sequence contains a subsequence that weakly converges to some $x \in C$.*

Theorem A 7. *The following Banach spaces are reflexive:*

(a) *Any finite-dimensional normed vector space;*

(b) *Any Hilbert space;*

(c) *Any closed subspace of reflexive Banach space;*

(d) *The dual space of any reflexive Banach space;*

(e) *The spaces ℓ^p, $1 < p < \infty$, and the Lebesgue spaces $L^p(\Omega)$, $1 < p < \infty$, with Ω any subset of $\mathbb{R}^N$.*

Definition A 10. *Let X be a normed vector space and let X^* denote its dual. A function $J : X \rightarrow \mathbb{R} \cup \{\infty\}$ is said sequentially lower semicontinuous if*

$$\lim_{k \rightarrow \infty} x_k = x \ in \ X \quad implies \quad J(x) \leq \liminf_{k \rightarrow \infty} J(x_k) \ .$$

Furthermore, let $U \subset X$ be non-empty, then $J : U \rightarrow \mathbb{R} \cup \{\infty\}$ is said sequentially weakly lower semicontinuous if

$$x_k \in U \rightharpoonup x \in U \ as \ k \rightarrow \infty \quad implies \quad J(x) \leq \liminf_{k \rightarrow \infty} J(x_k) \ .$$

Theorem A 8. *Let X be a normed space. Then a convex and continuous function $J : X \rightarrow \mathbb{R} \cup \{\infty\}$ is sequentially weakly lower semicontinuous on X.*

Definition A 11. *Let X and Y be two normed vector spaces. The space X is said continuously embedded in Y, $X \hookrightarrow Y$, if $X \subset Y$ and there exists a constant c such that $\|x\|_Y \leq c\|x\|_X$ for all $x \in X$. Furthermore, X is said compactly embedded in Y, $X \subset\subset Y$, if $X \hookrightarrow Y$ and every bounded sequence $\{x_k\}_{k=1}^{\infty}$ in X contains a subsequence converging in Y.*

Theorem A 9 (Sobolev imbedding theorems)**.** *Let Ω be a domain in $\mathbb{R}^N$, let $m \geq 1$ be an integer, and let $1 \leq p < \infty$. Then the following continuous embeddings hold:*

    (a) $W^{m,p}(\Omega) \hookrightarrow L^{p^*}(\Omega)$ with $\frac{1}{p^*} = \frac{1}{p} - \frac{m}{N}$ if $m < \frac{N}{p}$,

    (b) $W^{m,p}(\Omega) \hookrightarrow L^q(\Omega)$ for all q with $1 \leq q < \infty$ if $m = \frac{N}{p}$,

    (c) $W^{m,p}(\Omega) \hookrightarrow C^{0,m-N/p}(\bar{\Omega})$ if $\frac{N}{p} < m < \frac{N}{p}+1$,

    (d) $W^{m,p}(\Omega) \hookrightarrow C^{0,\lambda}(\bar{\Omega})$ for all λ with $0 < \lambda < 1$ if $m = \frac{N}{p}+1$,

    (e) $W^{m,p}(\Omega) \hookrightarrow C^{0,1}(\bar{\Omega})$ if $m > \frac{N}{p}+1$.

Theorem A 10 (Rellich-Kondrachov compact imbedding theorems)**.** *Let Ω be a domain in $\mathbb{R}^N$, let $m \geq 1$ be an integer, and let $1 \leq p < \infty$. Then the following compact embeddings hold:*

    (a) $W^{m,p}(\Omega) \subset\subset L^q(\Omega)$ for all q with $1 \leq q < p^*$ if $m < \frac{N}{p}$,

    (b) $W^{m,p}(\Omega) \subset\subset L^q(\Omega)$ for all q with $1 \leq q < \infty$ if $m = \frac{N}{p}$,

    (c) $W^{m,p}(\Omega) \subset\subset C(\bar{\Omega})$ if $m > \frac{N}{p}$.

A.5   Non-smooth calculus

In this section, we recall some main results of non-smooth calculus. Our main references are [34, 45, 66].

Definition A 12. *Let X be a locally convex topological vector space, and denote by X^* its dual space. Let $F : X \to \mathbb{R} \cup \{\infty\}$ a convex function. The function F is said subdifferentiable at $x \in X$ if $F(x)$ is finite and there exists $g \in X^*$ such that*

$$F(y) - F(x) \geq g(y - x) \ \ \forall y \in X \ .$$

The element g is called subgradient of F at x. The set of all these linear functionals is called subdifferential:

$$\partial F(x) := \big\{ g \in X^* \ : \ F(y) - F(x) \geq g(y - x) \ \forall y \in X \big\} \ .$$

Theorem A 11. *Let X be a locally convex vector space, and $F : X \to \mathbb{R} \cup \{\infty\}$ be convex, finite and continuous at x. Then $\partial F(x) \neq \emptyset$.*

Theorem A 12. *Let X be a locally convex vector space, and $F_1, F_2 : X \to \mathbb{R} \cup \{\infty\}$ be proper and continuous functions. Let $x \in \mathrm{dom}F_1 \cap \mathrm{dom}F_2$, such that F_1 is continuous at x. Then it holds that*

$$\partial(F_1 + F_2)(x) = \partial F_1(x) + \partial F_2(x) \ ,$$

for all $x \in X$.

Theorem A 13. *Let X, Y be two Banach spaces. Let $F : Y \to X$ and $G : X \to \mathbb{R}$. Suppose that F is strictly differentiable at y and that G is Lipschitz near $F(y)$. Then $f = G \circ F$ is Lipschitz near y and it holds that*

$$\partial f(y) \subset \partial G \circ F'(y) \ ,$$

where F' is the derivative of F.

A.6 Standard results in optimization and optimal control theory

In this section, we briefly summarize well known results regarding necessary optimality conditions for the following general optimal control problem

$$
\begin{aligned}
\min_{x,u} \quad & J(x,u) \\
\text{s.t.} \quad & c(x,u) = 0 \\
& x \in X , \ u \in U ,
\end{aligned}
\tag{P_{OC}}
$$

where $c(x,u) = 0$ denotes the differential constraint. The associated Lagrangian function is defined as

$$
L(x,u,p) := J(x,u) + \big(c(x,u), p\big)_{P,P^*} ,
$$

where P is the space of the residuals of the constraint equation, P^* is its dual space and $(\cdot,\cdot)_{P,P^*}$ represents the duality product between $c(x,u)$ and the corresponding Lagrange multiplier $p \in P^*$. To guarantee the existence of Lagrange multipliers in P^* it is possible to consider the following result, see, e.g., [66, 119, 134, 135].

Theorem A 14. *Let Y and P be Banach spaces, C a closed and convex subset of Y and K a closed convex cone in P. Let $J : Y \to \mathbb{R}$ and $c : Y \to P$ be Fréchet differentiable at $\tilde{y} \in Y$ solution to*

$$
min \quad J(y) \ \ s.t. \ \ y \in C \ \ and \ \ c(y) \in K .
$$

Assume that one of the following conditions is satisfied

- *the Zowe-Kurcyusz constraint qualification (ZKCQ) holds, i.e.*

$$
P = c'(\tilde{y}) Z_C(\tilde{y}) - Z_K(c(\tilde{y})) ,
$$

  where $Z_C(\tilde{y})$ and $Z_K(c(\tilde{y}))$ are the cones of feasible directions of C at $\tilde{y}$ and of K at $c(\tilde{y})$, respectively.

- *the slater condition holds, that is*

$$
\exists y \in C \ \ s.t. \ \ c(\tilde{y}) + c'(\tilde{y})(y - \tilde{y}) \in \operatorname{int} K ;
$$

- *$c'(\tilde{y})$ is surjective.*

Then there exists a Lagrange multiplier $p \in P^*$.

The first-order necessary optimality condition of (P_{OC}) is given by the following result see, e.g., [16, 86, 119].

Theorem A 15. *Assume that at $(\tilde{x}, \tilde{u}) \in X \times U$ it holds that*

- *J and c are Fréchet differentiable maps;*

- *$c(\tilde{x}, \tilde{u}) = 0$ admits a solution;*

- *c satisfies some regularity conditions, e.g. (ZKCQ), to guarantee the existence of the associated Lagrange multiplier $\tilde{p} \in P^*$.*

If the pair $(\tilde{x}, \tilde{u}) \in (X, U)$ is a local minimizer for the optimal control problem (P_{OC}). Then, the triple $(\tilde{x}, \tilde{u}, \tilde{p})$ is a KKT point and satisfies the following system

$$\nabla_x J(x, u) + \left(\nabla_x c(x, u)\right)^* p = 0$$

$$\nabla_u J(x, u) + \left(\nabla_u c(x, u)\right)^* p = 0$$

$$c(x, u) = 0 .$$

The first equation is termed as the adjoint equation, the second is the gradient component of the optimality conditions and the last is the constraint of the problem, sometimes termed as primal equation.

Notice that the structure of the considered optimal control problem allows to consider the state as a function of the control, i.e. $x = x(u)$. For this reason, it is possible to introduce the so-called reduced cost functional, given by the following; see, e.g., [16, 119, 129];

$$J_r(u) := J(x(u), u) ,$$

and the corresponding reduced problem is given by

$$\min_u \ J_r(u)$$
$$\text{s.t.} \ u \in U . \qquad\qquad (P_{OCr})$$

Notice that (P_{OCr}) is, in contrast to (P_{OC}), an unconstrained optimization problem. If $\tilde{u}$ is a solution to (P_{OCr}), then the pair $(x(\tilde{u}), \tilde{u})$ solves (P_{OC}), see, e.g., [16, 119, 129]. The reduced gradient is defined as

$$\nabla_u J_r(u) := \nabla_u J(x, u) + \left(\nabla_u c(x, u)\right)^* p .$$

Now, assume that the cost functional J and the constraint c are twice Fréchet differentiable. To obtain the reduced Hessian operator, we first consider the Hessian of the Lagrangian function, that is

$$H(L(x, u, p)) = \begin{pmatrix} \nabla_{xx} L & \nabla_{xu} L & \left(\nabla_x c\right)^* \\ \nabla_{ux} L & \nabla_{uu} L & \left(\nabla_u c\right)^* \\ \nabla_x c & \nabla_u c & 0 \end{pmatrix} .$$

This operator act on a vector $(\delta x, \delta u, \delta p)^T$ where δx, δu and δp are any variations of x, u and p in the space of the solutions of constraint and adjoint equations. Then, assuming that the linearized constraint equation is solvable, i.e. the operator $\left(\nabla_x c\right)$ is invertible, we write that

$$\left(\nabla_x c\right)\delta x + \left(\nabla_u c\right)\delta u = 0 \Rightarrow \delta x = -\left(\nabla_x c\right)^{-1}\left(\nabla_u c\right)\delta u .$$

We obtain the Hessian operator in the reduced form, that is

$$H_r(x, u) = \begin{pmatrix} -\left(\nabla_x c\right)^{-1}\left(\nabla_u c\right) \\ I \end{pmatrix}^* \begin{pmatrix} \nabla_{xx} L & \nabla_{xu} L \\ \nabla_{ux} L & \nabla_{uu} L \end{pmatrix} \begin{pmatrix} -\left(\nabla_x c\right)^{-1}\left(\nabla_u c\right) \\ I \end{pmatrix} .$$

Furthermore, the action of H_r on δu can be obtained solving the linearized constraint equation, then solving the linearized adjoint, that is

$$\left(\nabla_{xx} L\right)\delta x + \left(\nabla_{xu} L\right)\delta u + \left(\nabla_x c\right)^* \delta p = 0 ,$$

and assembling as follows

$$H_r \delta u := \big(\nabla_{ux} L\big)\delta x + \big(\nabla_{uu} L\big)\delta u + \big(\nabla_u c\big)^* \delta p \, .$$

A second-order necessary condition for a local minimum of the problem (P_{OCr}) is given by the following theorem; see, e.g., [16, 66, 119].

Theorem A 16. *Under the assumption of Theorem 15, consider that $J(x, u)$ and $c(x, u)$ are twice Fréchet differentiable maps, and that the linearized constraint equation is solvable. If the triple $(\tilde{x}, \tilde{u}, \tilde{p})$ is an optimal solution to the problem (P_{OCr}), then the reduced Hessian operator $H_r(\tilde{x}, \tilde{u})$ is positive semidefinite on the kernel of the linearized constraints, in the sense that*

$$\left(\begin{pmatrix} \nabla_{xx} L & \nabla_{xu} L \\ \nabla_{ux} L & \nabla_{uu} L \end{pmatrix}_{(\tilde{x}, \tilde{u}, \tilde{p})} \begin{pmatrix} \delta x \\ \delta u \end{pmatrix}, \begin{pmatrix} \delta x \\ \delta u \end{pmatrix} \right) \geq 0 \, ,$$

for all $(\delta x, \delta u) \in X \times U$ satisfying the linearized constraint equation

$$\big(\nabla_x c\big)\delta x + \big(\nabla_u c\big)\delta u = 0 \, .$$

Next, we consider a general optimal control problem with inequality constraints on the controls, that is

$$\begin{aligned} \min_{x,u} \quad & J(x, u) \\ \text{s.t.} \quad & c(x, u) = 0 \\ & x \in X \, , \ u \in U_{ad} \subset U \, . \end{aligned} \qquad (P_{OC2})$$

A necessary optimality condition to this problem is the following; see, e.g., [16, 66, 119].

Theorem A 17. *If $U_{ad} \subset U$ is a convex and closed subset of the Hilbert space U and $J_r : U_{ad} \to \mathbb{R}$ is differentiable, then the solution $\tilde{u}$ of the optimization problem (P_{OC2}) satisfies the variational inequality*

$$\langle \nabla_u J_r(\tilde{u}), u - \tilde{u} \rangle_U \geq 0 \, , \ \forall u \in U_{ad} \, .$$

Another useful result that establish optimality conditions is given by the following theorem [45].

Theorem A 18. *Let X be a reflexive Banach space. Consider a map $F : U \to \mathbb{R}$, where $U \subseteq X$. Consider the following problem*

$$\inf_{u \in U} F(u) \, . \qquad (A.52)$$

It holds that

(a) *if $U = X$ and $\tilde{u}$ is a solution to (A.52), then it holds that $0 \in \partial F(\tilde{u})$, where $\partial F(\tilde{u})$ is a subdifferential of F at $\tilde{u}$;*

(b) *if $U \subset X$ is a non-empty, closed and convex, and $F = F_1 + F_2$, with F_1 and F_2 being lower semicontinuous functions of U into $\mathbb{R}$, F_1 being Gâteaux differentiable with derivative F_1', F_2 being convex, then a necessary condition for $\tilde{u}$ to be a minimizer for (A.52) is the following*

$$\langle F_1'(\tilde{u}), v - \tilde{u} \rangle + F_2(v) - F_2(\tilde{u}) \geq 0 \, , \ \forall v \in U \, . \qquad (A.53)$$

If in addition F_1 is convex, then the previous condition is also sufficient.

Bibliography

[1] A. Agrachev, T. Chambrion, *An estimation of the controllability time for single-input systems on compact Lie groups*, ESAIM: Control, Optimization and Calculus of Variations, Vol. 12 (2006), 409-441;

[2] F. Albertini, D. D'Alessandro, *The Lie algebra structure and controllability of spin systems*, Linear algebra and its applications, 350 (2002), 213-235;

[3] F. Albertini, D. D'Alessandro, *Notions of controllability for bilinear multilevel quantum system*, IEEE Trans. on Automatic Control, 48(8) (2003), 1399-1403;

[4] D. D'Alessandro, *Controllability of one and two interacting spins*, Math. Control Signals Systems, 16 (2003), 1-25;

[5] C. Altafini, *Controllability of quantum mechanical systems by root space decomposition of $\mathfrak{su}(N)$*, Journal of Mathematical Physics, 43(5) (2002), 2051-2062;

[6] P. Allard, M. Helgstrand, T. Härd, *The complete homogeneous master equation for a Heteronuclear two-spin system in the basis of cartesian product operators*, J. Magn. Reson., 134, 7-4 (1998), No. MN981509;

[7] W. Alt, C. Schneider, *Linear-quadratic control problems with L^1-control cost*, Optimal control applications and methods, (2014) DOI: 10.1002/oca.2126;

[8] A. Arvanitoyeorgos, *An Introduction to Lie Groups and the Geometry of Homogeneous Spaces*, Student Mathematical Library (AMS) Volume 22, 2003;

[9] E. Assémat, M. Lapert, Y. Zhang, M. Braun, S.J. Glaser, D. Sugny, *Simultaneous time-optimal control of the inversion of two spin-$\frac{1}{2}$ particles*, Physical Review A, 82 (2010), pp. (013415)1-6;

[10] K. Beauchard, J.-M. Coron, P. Rouchon, *Controllability issues for continuous-spectrum systems and ensemble controllability of Bloch equations*, Commun. Math. Phys., 296(2) (2010), 525-557;

[11] M. Belhadj, J. Salomon, G. Turinici, *A stable toolkit method in quantum control*, J. Phys. A, 41 (36) (2008), 362001-362011;

[12] F. Bloch, A. J. Siegert, *Magnetic resonance for nonrotating fields*, Physical Review 57 (1940), 522-527;

[13] B. Bonnard, M. Chyba, *Singular Trajectories and their Role in Control Theory*, Mathématiques et Applications, Springer (2003);

[14] A. Borzì, *Quantum optimal control using the adjoint method*, Nanoscale Systems MMTA, 1 (2012), pp. 93-111;

[15] A. Borzì, J. Salomon, S. Volkwein, *Formulation and numerical solution of finite-level quantum optimal control problems*, J. Comp. App. Math., 216 (1) (2008), 170-197;

[16] A. Borzì, V. Schulz, *Computational Optimization of Systems Governed by Partial Differential Equations*, Computational Science and Engineering, SIAM, Philadelphia, 2012;

[17] A. Borzì, G. Stadler, U. Hohenester, *Optimal quantum control in nanostructures: Theory and application to a generic three-level system*, Phys. Rev. A 66, 053811 (2002);

[18] A. Borzì, U. Hohenester, *Multigrid optimization schemes for solving Bose-Einstein condensates control problems*, SIAM J. Sci. Comp., 30 (2008), 441-462;

[19] R. W. Brockett, N. Khaneja, *On the stochastic control of quantum ensembles*, System theory, 518 (2000), 75-96;

[20] W. H. Brown, C. F. Foote, D. L. Iverson, E. Anslyn, *Organic Chemistry - Seventh Edition*, Brooks/Cole, Cengage Learning USA (2012-2014);

[21] P. N. Brown, V. Saad, *Convergence theory of nonlinear Newton-Krylov algorithms*, SIAM J. Optimization, 4 (1994), pp. 297-330;

[22] J. Burkardt, M. Gunzburger, J. Peterson, *Insensitive functionals, inconsistent gradients, spurious minima, and regularized functionals in flow optimization problems*, International Journal of Computational Fluid Dynamics, 16 (2002), pp. 171-185;

[23] H. Cartan, *Calcul Différentiel*, Hermann, Paris (1967);

[24] E. Casas, R. Herzog, G. Wachsmuth, *Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional*, SIAM J. Optim., 22(3) (2012), 795-820;

[25] J. Cavanagh, W J. Fairbrother, A. G. Palmer III, M. Rance, N. J. Skelton, *Protein NMR Spectroscopy - Principles and Practice - Second Edition*, Elsevier, Academic Press New York (2007);

[26] L. Cesari, *Optimization - Theory and Applications*, Springer-Verlag (1983);

[27] B. Chen, X. Chen. C. Kanzow, *A Penalized Fischer-Burmeister Ncp-Function: Theoretical Investigation And Numerical Results*, Math. Programm. 88, 211-216 (2000);

[28] G. Ciaramella, *Towards a minimum L^2-norm exact control of the Pauli equation*, proceeding in SIAM Conference on Control and its Applications (July 8-10 - Maison de la Mutualité, Paris - France), accepted;

[29] G. Ciaramella, A. Borzì, *On L^1 and piecewise-constant optimal control of bilinear quantum systems*, (2014) - submitted;

[30] G. Ciaramella, J. Salomon, A. Borzì, *A method for solving exact-controllability problems governed by closed quantum spin systems*, (2015) International Journal of Control (IJC), 88(4), pp. 682-702;

[31] G. Ciaramella, A. Borzì, G. Dirr, D. Wachsmuth, *Newton methods for the optimal control of closed quantum spin systems*, (2015) SIAM Journal on Scientific Computing (SISC), 37(1), pp. A319-A346;

[32] G. Ciaramella, A. Borzì, *SKRYN: A fast semismooth-Krylov-Newton method for controlling Ising spin systems*, (2015) CPC (Computer Physics Communication), 190, pp. 213-223;

[33] P. G. Ciarlet, *Linear and Nonlinear Functional Analysis with Applications*, SIAM, Philadelphia (2013);

[34] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society Series of Monographs and Advanced Texts, John Wiley and Sons Inc. (1983), New York;

[35] C. Clason, K. Kunisch, *A duality-based approach to elliptic control problems in non-reflexive Banach spaces*, ESAIM Control Optim. Calc. Var., 17 (2011), 243-266;

[36] Y. H. Dai, Y. Yuan, *A nonlinear conjugate gradient with a strong global convergence property*, SIAM J. Optim., 10, 177-182 (1999);

[37] I. Degani, A. Zanna, L. Saelen, R. Nepstad, *Quantum control with piecewise constant control functions*, SIAM J. Sci. Comput., 31(5) (2008), pp. 3566-3594;

[38] T. De Luca, F. Facchinei, C. Kanzow, *A semismooth equation approach to the solution of nonlinear complementarity problems*, Mathematical Programming, 75 (1996), pp. 407-439;

[39] G. Dirr, U. Helmke, K. Hüper, M. Kleisteuber, Y. Liu, *Spin dynamics: a paradigm for time optimal control on compact Lie groups*, Journal of Global Optimization, 35 (2006), 443-474;

[40] G. Dirr, U. Helmke, *Lie theory for quantum control*, GAMM-Mitteilungen, 31 (2008), 59-93;

[41] P. Ditz and A. Borzì, *A cascadic monotonic time-discretized algorithm for finite-level quantum control computation*, Computer Physics Communications, 178 (2008), 393-399.

[42] D. Dong, I. R. Petersen, *Quantum control theory and application: A survey*, IET Control Theory and Applications, 4(2), 2651-2671;

[43] S. Eisenstat, H. Walker, *Choosing the forcing term in an inexact Newton method*, SIAM J. Sci. Comput., 17-1 (1996), pp. 16-32;

[44] R. Eitan, M. Mundt, D. J. Tannor, *Optimal Control with Accelerated Convergence: Combining the Krotov and Quasi-Newton Methods*, Phys. Rev. 83 (2011), 053426;

[45] I. Ekeland, R. Témam, *Convex Analysis and Variational Problems*, SIAM Classics In Applied Mathematics vol.28, SIAM Philadelphia (1999);

[46] R. R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, International Series of Monographs on Chemistry, Vol. 14, Oxford Science Publications, (1989 - Reprinted 2004);

[47] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, Vol. 19, American Society, Providence, RI, (2002);

[48] P. de Fouquieres, S. G. Schirmer, S. J. Glaser, I. Kuprov, *Second order gradient ascent pulse engineering*, arXiv:1102.4096v2, 2011;

[49] S. Gasiorowicz, *Quantum Physics - Third Edition*, Wiley and Sons (2003);

[50] M. Gerdts, *A nonsmooth Newton's method for control-state constrained optimal control problems*, Mathematics and Computers in Simulation, 12, 79:925-936 (2008);

[51] D. J. Griffiths (1995) *Introduction to Quantum Mechanics*, Prentice Hall;

[52] L. Grippo, M. Sciandrone, *Metodi di Ottimizzazione non Vincolata*, Springer-Verlag Milano Italy (2011);

[53] B. C. Hall, *Lie Groups, Lie Algebras and Representations: an Elementary Introduction*, Springer (2003);

[54] U. Helmke, J.B. Moore, *Optimization and Dynamical Systems*, Springer, (1994);

[55] R. Herzog, G. Stadler, G. Wachsmuth, *Directional sparsity in optimal control of partial differential equations*, SIAM J. Control Optim., 50 (2012), N.02, pp. 943-963;

[56] M. Hintermüller, K. Ito, K. Kunisch, *The primal-dual active set method as a semismooth Newton method*, SIAM J. Optimization, 13 (2003), pp. 865-888;

[57] M. Hintermüller, *Semismooth Newton Methods and Applications*, Oberwolfach-Seminar on "Mathematics of PDE-Constrained Optimization" at Mathematisches Forschunginstitut in Oberwolfach, November 2010;

[58] M. Hinze, R. Pinnau, M. Ulbrich, S. Ulbrich, *Optimization with PDE Constraints*, Mathematical modelling: Theory and applications, V. 23, Springer;

[59] TS Ho, H. Rabitz, *Accelerated monotonic convergence of optimal control over quantum dynamics*, J. Math. Chem., 48 (2010), 1010;

[60] W. W. Hager, H. Zhang, *A new conjugate gradient method with guaranteed descent and an efficient line search*, SIAM Journal on Optimization 16:1 (2005), 170-192;

[61] M. Hochbruck, C. Lubich, *On Magnus integrators for time-dependent Schrödinger equations*, SIAM J. Numer. Anal., 41 (2003), pp. 945-963;

[62] H. J. Hogben, M. Krzystyniak, G.T.P. Charnock, P.J. Hore, I. Kuprov, *Spinach - a software library for simulation of spin dynamics in large spin systems*, J. Magn. Reson., 208 (2011) 179-194;

[63] S. M. Ikhadir, R. Sever, *Exact solutions of the radial Schrödinger equation for some physical potentials*, Central Eur.J.Phys. 5(2007), 516;

[64] K. Ito, K. Kunisch, *Augmented Lagrangian methods for nonsmooth convex optimization in Hilbert spaces*, Nonlinear Analysis, TMA, 41 (2000), pp. 591-616;

[65] K. Ito, K. Kunisch, *Semi-smooth Newton methods for state-constrained optimal control problems*, Systems and Control Letters, 50 (2003), pp. 221-228;

[66] J. Jahn, *Introduction to the Theory of Nonlinear Optimization - Third Edition*, Springer-Verlag Berlin Heidelberg (2007);

[67] V. Jurdjevic, *Geometric Control Theory*, Cambridge University Press, New York, (1997);

[68] V. Jurdjevic, H. Sussmann, *Control systems on Lie groups*, Journal Diff. Eqns., 12 (1972), 313-329;

[69] C. Kanzow, H. Kleinmichel, *A new class of semismooth Newton-type methods for nonlinear complementarity problems*, Computational Optimization and Applications 11 (1998), pp. 227-251;

[70] N. Khaneja, S. J. Glaser, R. Brockett, *Sub-Riemannian geometry and time optimal control of three spin systems: quantum gates and coherence transfer*, Physical Review A, 65(3) (2002) 032301;

[71] N. Khaneja, S. Glaser, *Cartan decomposition of $SU(2^n)$ and control of spin systems*, Chem. Phys. 267 (2001), 11:23;

[72] N. Khaneja, R. Brockett, S. Glaser, *Time optimal control in spin systems*, Physical Review A, 63(3) (2001), 032308;

[73] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, S. J. Glaser, R. Brockett, *Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms*, J. Magn. Reson. 172 (2005), 296-305;

[74] N. Khaneja, Jr-S. Li, C. Khelet, B. Luy, S.J. Glaser, *Broadband relaxation-optimized polarization transfer in magnetic resonance*, Proc. Natl. Acad. Sci. USA 101 (2004), pp. 14742-14747;

[75] K. Kobzar, T. E. Skinner, N. Khaneja, S. J. Glaser, B. Luy, *Exploring the limits of broadband excitation and inversion pulses*, J. Magn. Reson. 170(2004), 236-243;

[76] A. I. Konnov, V. F. Krotov, *On global methods for the successive improvement of control processes*, Avtomat. i Telemekh., 10 (1999), 77-88;

[77] K. Kreutz-Delgado, *The complex gradient operator and the $\mathbb{CR}$-calculus*, arXiv:0906.4835 (2009), pp. 380-382;

[78] A. Kröner, *Semi-smooth Newton methods for optimal control of the dynamical Lamé system with control constraints*, Numerical Functional Analysis and Optimization, 34 (7) (2013), 741-769;

[79] A. Kröner, K. Kunisch, and B. Vexler, *Semi-smooth Newton methods for optimal control of the wave equation with control constraints*, SIAM J. Control Optim., 49(2) (2011), 830-858;

[80] K. Kunisch, X. Lu, *Optimal control for elliptic systems with pointwise euclidean norm constraints on the controls*, Mathematical Programming, Volume 142 (1-2) (2013), 461-483;

[81] I. Kurniawan, G. Dirr, U. Helmke, *The dynamics of open quantum systems: Accessibility results*, Proc. Appl. Math. Mech., GAMM Annual Meeting, Zurich (2007);

[82] I. Kurniawan, *Controllability aspects of the Lindblad-Kossakowski master equation: a Lie-theoretical approach*, PhD Thesis, (2009) University of Würzburg;

[83] L. D. Landau, E. M. Lifshitz, *Quantum Mechanics (Non-relativistic theory) - Third Edition*, Butterworth Heinemann (1958);

[84] E. B. Lee, L. Markus, *Foundations of Optimal Control Theory*, Robert E. Krieger Publishing Company, Malabar, Florida (1986);

[85] M. H. Levitt, *Symmetrical composite pulse sequences for NMR population inversion. I. Compensation of radiofrequency field inhomogeneity*, Journal of Magnetic Resonance, 48 (1982), pp. 234-264;

[86] D. G. Luenberger, *Optimization by Vextor Space Methods*, Wiley, New York, (1969);

[87] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Guslys, T. Schulte-Herbrüggen, *Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework*, Phys. Rev. A 84 (2011), 022305;

[88] Y. Maday, J. Salomon, G. Turinici, *Monotonic time-discretized schemes in quantum control*, Num. Math., 103 (2) (2006), 323-338;

[89] Y. Maday, J. Salomon, G. Turinici, *Monotonic Parareal Control for Quantum Systems*, SIAM J. Num. Anal., 45 (6) (2007), 2468-2482;

[90] K. D. Malanowski, *Convergence of the Lagrange-Newton method for optimal control problems*, Int. J. Appl. Math. Comput. Sci., 14(4) (2004), 531-540;

[91] I. I. Maximov, J. Salomon, G. Turinici, N. C. Nielsen, *A smoothing monotonic convergent optimal control algorithm for NMR pulse sequence design*, J. Chem. Phys., 132 (2010), 084107-1-084107-9;

[92] J. Mawhin, A Ronveaux, *Schrödinger and Dirac equations for the Hydrogen atom, and Laguerre polynomials*, Archive for History of Exact Sciences, 64(4) (2010), pp. 429-460;

[93] H. Nijmeijer, A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, (2010) Springer-Verlag New York;

[94] J. Nocedal, S. Wright, *Numerical Optimization*, (2006) Springer;

[95] W. Pauli, *Zur Quantenmechanik des magnetischen Elektrons*, Zeitschrift für Physik, 43 (1927), pp. 601-623;

[96] P. Philip, *Optimal Control of Partial Differential Equations*, Lecture Notes, LMU Munich (2013);

[97] L. Qi, D. Sun, *A nonsmooth version of Newton's methods*, Math. Programming, 58 (1993), pp. 353-367;

[98] L. Qi, D. Sun, *A survey of some nonsmooth equations and smoothing Newton methods*, Progress in Optimization - Applied Optimization, 30 (1999), pp. 121-146;

[99] D. Reich, M. Ndong, C. Koch, *Monotonically convergent optimization in quantum control using Krotov's method*, J. Chem. Phys. 136, 104103 (2012);

[100] J. Salomon, *Limit points of the monotonic schemes in quantum control*, Proceedings of the 44th IEEE Conference on Decision and Control, Sevilla, (2005);

[101] S. Schirmer, I. Pullen, A. Solomon, *Controllability of quantum systems*, Lagrangian and Hamiltonian methods in nonlinear control 2013: Workshop, IFAC Proceedings Volumes, Springer Verlag, New York (2003); e-print: http://arxiv.org/pdf/quant-ph/0302121;

[102] T. Schulte-Herbrüggen, *Aspects and Prospects of High-Resolution NMR*, PhD thesis, ETH Zürich (1998), Diss. ETH No. 12752;

[103] R. Shankar, *Principles of Quantum Mechanics - Second Edition*, Kluwer Academic - Plenum Publishers (1994);

[104] N. Schaeffer, *Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations*, Geochemistry, Geophysics, Geosystems, 14(3) (2013), pp. 751-758;

[105] T. E. Skinner, T. O. Reiss, B. Luy, N. Khaneja, S. J. Glaser, *Tailoring the optimal control cost function to a desired output: application to minimizing phase errors in shorts broadband excitation pulses*, J. Magn. Reson., 172 (2005), 17-23;

[106] S.E. Sklarz, D. J. Tannor, *Loading a Bose-Einstein condensate onto an optical lattice: An application of optimal control theory to the nonlinear Schrödinger equation*, Phys. Rev A 66 (2002), 053619;

[107] E. D. Sontag, *Mathematical Control Theory - Deterministic Finite Dimensional Systems - Second Edition*, Text in Applied Mathematics 6, (1998) Springer;

[108] O. W. Søresen, *Polarization transfer experiments in high-resolution NMR spectroscopy*, Progress in NMR Spectroscopy, 21 (1989), 503-569;

[109] O. W. Søresen, G. H. Eich, M. H. Levitt, G. Bodenhausen, R. R. Ernst, *Product operator formalism for the description of NMR pulse experiments*, Progress in NMR Spectroscopy, 16 (1983), 163-192;

[110] G. Stadler, *Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices*, Computational Optimization and Applications, 44(2) (2009), pp. 159-181;

[111] D. Stefanatos, S. J. Glaser, N. Khaneja, *Relaxation optimized transfer of spin order in Ising spin chains*, Phys. Rev. A 72 (2005), pp. (062320)1-6;

[112] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis - Second Edition*, Springer-Verlag (1993);

[113] E. Süli, D. Mayers, *An Introduction to Numerical Analysis*, Cambridge University Press (2003);

[114] H. J. Sussmann, *The Bang-Bang problem for certain control systems in $GL(n, \mathbb{R})$*, SIAM J. Control 10 N.3 (1972), 470-476;

[115] D. J. Tannor, V. Kazakov and V. Orlov, *Control of photochemical branching: novel procedures for finding optimal pulses and global upper bounds*, Time Dependent Quantum Molecular Dynamics, J. Broeckhove and L. Lathouwers, eds., (Plenum, 1992), 347-360;

[116] S. H. Tersigni, P. Gaspard, S.A. Rice, *On using shaped light pulses to control the selectivity of product formation in a chemical reaction: an application to a multiple level system*, J. Chem. Phys. 93 (1990), 1670-1680;

[117] V. Thomée, *Galerkin finite element methods for parabolic problems*, Springer-Verlag, Berlin, 2006;

[118] B. Thaller, *Advanced Visual Quantum Mechanics*, Springer, New York, 2005;

[119] F. Tröltzsch, *Optimal Control of Partial Differential Equations: Theory, Methods and Applications*, Grad. Stud. Math. 112, American Mathematical Society, Providence, RI, 2010;

[120] F. Tröltzsch, A. Valli, *Optimal control of low-frequency electromagnetic fields in multiply connected conductors*, DFG-Research Center Matheon, Matheon preprint, December 2014;

[121] G. Turinici, H. Rabitz, *Quantum wavefunction controllability*, Chemical Physics 267 (1), 1-9;

[122] G. Turinici, *On the controllability of bilinear quantum systems*, Mathematical models and methods for ab initio Quantum Chemistry, 2000;

[123] M. Ulbrich, *Semismooth Newton methods for operator equations in function spaces*, SIAM J. Optimization, 13 (2003), pp. 805-842;

[124] A. van den Bos, *Complex gradient and Hessian*, IEE Proc.-Vis. Image Signal Processing (British), 141(6) (1994), pp. 380-382;

[125] L. M. K. Vandersypen, I. L. Chuang, *NMR Techniques for quantum control and computation*, Rev. Mod. Phys., 76 (2005), pp. 1037-1069;

[126] G. Vossen, H. Maurer, *On L^1-minimization in optimal control and applications to robotics*, Optimal control applications and methods, 27 (2006), pp. 301-321, DOI: 10.1002/oca.781;

[127] G. Wachsmuth, D. Wachsmuth, *Convergence and regularization results for optimal control problems with sparsity functional*, ESAIM: Control, Optimisation and Calculus of Variations, 17(3) (2011), pp. 858-886;

[128] G. von Winckel, A. Borzì, *Computational techniques for a quantum control problem with H^1-cost*, Inverse Problems, 24 (2008), pp. (034007)1-23;

[129] G. von Winckel, A. Borzì, S. Volkwein, *A globalized Newton method for the accurate solution of a dipole quantum control problem*, SIAM J. Sci. Comput., 31 (2009), pp. 4176-4203;

[130] G. von Winckel, A. Borzì, *QUCON: A fast Krylov-Newton code for dipole quantum control problems*, Computer Physics Communications, 181 (2010), pp. 2158-2163;

[131] W. Walter, *Ordinary Differential Equations*, Springer-Verlag New York (1998);

[132] H. Yserentant, *Regularity and Approximability of Electronic Wave Functions*, Springer-Verlag Berlin Heidelberg (2010);

[133] W. Zhu, H. Rabitz, *A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator*, J. Chem. Phys. 109 (1998), 385;

[134] E. Zeidler, *Applied Functional Analysis - Main Principles and their Applications*, Applied Mathematical Sciences 109, Springer-Verlag;

[135] J. Zowe, S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5(1) (1979), 49-62;