QUANTUM CONTROL LANDSCAPE ANALYSIS
INCLUDING ITS APPLICATION IN NMR EXPERIMENTS

Qiuyang Sun

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
CHEMISTRY

Adviser: Herschel Rabitz

September 2016
Abstract

Scientists have attempted for decades to efficiently control the dynamics of quantum systems on the atomic and molecular scale with modulated electromagnetic fields, such as femtosecond laser pulses. The success of such experiments in achieving specified control objectives will rely on finding the optimal pulse shapes. The search for optimal solutions in such a high-dimensional control parameter space is usually accomplished by learning algorithms. The success of quantum optimal control in a wide range of applications motivated the analysis of the underlying control landscape, defined by the objective value as a functional of the control field, whose topology (e.g., local optima and saddle points) will dictate the ease of finding globally optimal control fields by iterative searches with gradient-based algorithms. Based on some assumptions theoretical analysis revealed many topological properties of the landscapes, which are both experimentally assessed and theoretically developed in this dissertation.

The experimental works focus on the landscapes in controlling nuclear spins by pulsed magnetic fields, with nuclear magnetic resonance (NMR) spectroscopy as a testbed. We choose this type of experimental settings because of its desirable physical properties and potential application in realizing quantum computation. A set of methodology for control landscape studies in NMR systems is constructed, leading to the first observation of landscape saddles in the laboratory. The dissertation also discusses a variety of theoretical topics about the control landscape in general, including (i) measuring “distances” from the landscape saddle points and driving to locate them; (ii) simultaneous optimization of multiple control objectives in the same quantum system; and (iii) generalizing the landscape
concept from the semiclassical setting to the full quantum setting. All these works aim to further enrich and develop the landscape theory in quantum optimal control, and apply it to physical systems with practical importance.
Contents

1 Introduction 1

1.1 Quantum optimal control and landscape analysis . . . . . . . . . . . . . . 4

1.1.1 Adaptive feedback control . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.1.2 Quantum control theory and landscape analysis: General formulation 5

1.1.3 Important theoretical conclusions about kinematic landscapes . . . . 8

1.1.4 Assumptions and limitations of current landscape analysis . . . . . 10

1.2 Quantum control of nuclear spins with NMR . . . . . . . . . . . . . . . . 13

2 Experimental exploration over a quantum control landscape through NMR 22

2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 23

2.2 Methods for experimental landscape roving . . . . . . . . . . . . . . . . . . 28

2.2.1 The laboratory landscape Rover algorithms . . . . . . . . . . . . . . . 28

2.2.2 Gradient and Hessian determination . . . . . . . . . . . . . . . . . . . 31

2.3 Experimental setup . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32

2.4 Results and discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34

2.4.1 Ascending and descending the landscape . . . . . . . . . . . . . . . . 35
2.4.2 Hessian observation over the landscape and exploration of the landscape top 36

2.4.3 Non-critical objective value level set exploration 40

2.5 Conclusion 45

3 Experimental observation of saddle points over the quantum control landscape of a two-spin system 49

3.1 Introduction 50

3.2 Theoretical analysis of the quantum control problem 52

3.3 Experimental setup 56

3.3.1 NMR experiments 56

3.3.2 Algorithms for laboratory landscape exploration 58

3.4 Results and discussion 60

3.4.1 Gradient ascent over the landscape 60

3.4.2 Finding the landscape saddles in the laboratory 61

3.4.3 Local landscape topology near the saddles: the Hessian signatures 65

3.4.4 Control on the $^{13}$C landscape 69

3.5 Conclusions 71

4 Identifying and avoiding singularity-induced local traps over control landscapes of spin chain systems 76

4.1 Introduction 77

4.2 Control landscape and singular controls 79
4.3 Singular traps in linear spin chain systems ........................................ 82
  4.3.1 Control system model and landscape gradient search ..................... 82
  4.3.2 Hessian analysis at the singular critical point .............................. 86
  4.3.3 The influence of more complex coupling in a three-spin case .......... 89
4.4 Illustrations ...................................................................................... 90
  4.4.1 Two-spin case: the likelihood of trapping in practice .................... 91
  4.4.2 Three-spin system: linear and triangular coupling models ............ 98
4.5 Discussion and conclusion ............................................................... 101

5 Measuring the distance from saddle points and driving to locate them over quantum control landscapes 108
  5.1 Introduction .................................................................................... 109
  5.2 The physical observable control landscape ...................................... 112
    5.2.1 Critical points over the landscape ........................................... 112
    5.2.2 Critical distance metric and saddle-seeking algorithm ............... 118
  5.3 The unitary transformation control landscape ................................ 123
  5.4 Illustrations .................................................................................... 127
    5.4.1 Numerical schemes ................................................................... 127
    5.4.2 The physical observable landscape .......................................... 129
    5.4.3 The unitary transformation landscape ...................................... 131
  5.5 Conclusion ...................................................................................... 133

6 The Pareto front shape in multiobservable quantum control 142
6.1 Introduction .................................................. 143
6.2 Pareto front shape with mutually commuting observables ............ 146
6.3 Numerical methods for identifying the Pareto front ...................... 153
6.4 Pareto front shape with non-commuting observables .................... 155
6.5 Experimental demonstration ...................................... 159
6.6 Conclusions .................................................. 162

7 Inherently trap-free convex landscapes for the full quantum optimal control 168
7.1 Introduction .................................................. 168
7.2 Full quantum landscape and its optimal solution ......................... 171
7.3 Illustrations .................................................. 175
7.4 Conclusions .................................................. 180

8 Future aspects of quantum control and landscape analysis .......... 186
8.1 Theoretical issues .............................................. 187
8.2 Quantum control in magnetic resonance ................................ 188
8.3 Quantum control in chemical physics ................................ 190
8.4 Control landscapes in alternative physical models ..................... 192
ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my advisor, Prof. Herschel Rabitz, for his continuous advice and support of my Ph.D study and related research, for his patience, motivation, and immense knowledge in a broad range of disciplines. His guidance helped me in all the time of research and writing of this dissertation.

Besides, I would like to thank the rest of my Thesis Advisory Committee, Prof. Annabella Selloni and Prof. Steven Bernasek, for their insightful comments and encouragement, and also for the hard questions which incented me to widen my research from various perspectives.

My sincere thanks also goes to Dr. István Pelczer, who provided me access to the NMR spectroscopic facilities and precious guidance on operating them, without which it would not be possible to conduct a major part of this research. Thanks to Dr. Re-Bing Wu from Tsinghua University, Dr. Tak-San Ho and Mr. Gregory Riviello for their collaboration on my work, and many other members of the Rabitz laboratory for fruitful discussions.

I acknowledge financial support of my research from the National Science Foundation, the Program in Plasma Science and Technology (Princeton Plasma Physics Laboratory), and the John Templeton Foundation.
Chapter 1

Introduction

Since the invention of laser in the 1960s, physicists and chemists have attempted to efficiently control the dynamics of quantum systems on the atomic and molecular scale with laser fields. For example, laser control has very attractive potential application in steering a complex chemical reaction to desirable pathways by inducing the cleavage or formation of particular bonds, thus significantly improving the reaction efficiency and selectivity (other applications will be discussed later). However, continuous wave (cw) laser fields generally fail to break a certain bond due to the internal vibrational relaxation, i.e., the energy being transferred to other vibrational modes and simply heating the whole molecule. The obstacle has been overcome since the 1990s by the technological development of femtosecond pulse shaping, which facilitates the control of reactions in an ultrafast timescale shorter than the relaxation [1]. The success of such experiments in achieving specified control objectives (e.g., maximizing the ratio of two products) will rely on finding the optimal pulse shape, i.e., the frequency, amplitude, phase, and/or polarization of the laser field. The search for optimal solutions in such a high-dimensional control parameter space is usually accomplished by stochastic or deterministic algorithms in both experiments and numerical simulations, by the scheme of adaptive feedback control (see Section 1.1.1).

The success of quantum optimal control in a wide range of applications motivated the
analysis of the underlying control landscape [2]. The landscape is defined by the objective value of a quantum control application as a functional of the control field as a time-domain function, or a function of multiple parameters that uniquely specify the field. The landscape topology will dictate the ease of finding globally optimal solutions by optimization searches. Of particular concern is the distribution and properties of the landscape critical points, especially the possible existence of traps (local optima) and saddle points with suboptimal objective values, which have a zero landscape gradient and thus may prevent gradient-based algorithms from efficiently approaching the global optimum. Based on some reasonable assumptions, theoretical analysis has revealed many important favorable properties of the landscapes for different types of control problems, including the feasibility of finding globally optimal controls with gradient-based algorithms (guaranteed by the absence of traps), as well as the robustness to noise at the landscape optimum. The conclusions shed light on the results of a large number of optimal control experiments and simulations. Violation of the assumptions in some special cases may bring about more complex landscape structures, which will be discussed later in this dissertation and is still open for further studies.

Despite its background in femtosecond laser control of chemical processes, quantum optimal control, including the landscape theory, also finds applications in a wide variety of other quantum phenomena controlled by modulated electromagnetic fields (see Ref. [3] for a review), such as high-harmonic generation, electron transport in nanoscale devices, and light control of neurons in living tissue to express light-sensitive ion channels, known as optogenetics [4]. The experimental works in this dissertation focus on control landscape studies of an important category of physical systems, i.e., nuclear spins controlled by pulsed magnetic fields with the common nuclear magnetic resonance (NMR) spectroscopy as a testbed. We choose this type of experimental settings for assessing the landscape principles because of its desirable physical properties, including simple and well-defined Hamiltonian involved, the mixed state at thermal equilibrium, slow relaxation, and high
signal-to-noise ratio, as well as its promising application in realizing quantum computation [5]. The implications of these experimental findings extend beyond for the control of other quantum phenomena with electromagnetic fields, as the underlying theory is generic.

The following chapters of this dissertation are organized as follows. Chapters 2-4 report control landscape studies with NMR experiments. Chapter 2 introduces the laboratory methodologies for exploring the control landscape, and demonstrates them in a simple single-spin system. Chapter 3 further applies the methods to a two-spin system of chloroform (\(^{13}\text{CHCl}_3\)), and observes the theoretically predicted landscape saddle points for the first time. Chapter 4 analyzes the impact of so-called singular controls on the optimal control of spin chains, and demonstrates the results in NMR experiments as well.

Chapters 5-7 discuss a variety of theoretical topics about the control landscape in general. Chapter 5 presents metrics for the “distances” from particular landscape critical points, which can be applied in locating the saddle controls with high precision or analyzing their roles in the optimization searches. Chapter 6 considers the problem of simultaneous optimization of multiple control objectives in the same quantum system, and discovers some general rules about the tradeoff relation between the competitive objectives. Finally, Chapter 7 generalizes the control landscape concept from the traditional semiclassical setting (a quantum system controlled by a classical field) to the more fundamental full quantum setting (controlled by another quantum system), and reveals the origins of the important features of semiclassical landscapes, including the traps and saddles. All these works aim to further enrich and develop the landscape theory in quantum optimal control, as well as apply it to physical systems with practical importance.
1.1 Quantum optimal control and landscape analysis

1.1.1 Adaptive feedback control

Experimental implementation of quantum optimal control can be a difficult task. First, numerical calculation of control solutions strongly depends on the employed model Hamiltonian. However, the Hamiltonians of real systems controlled in the laboratory are usually not well known, except for a few simple cases including the nuclear spins studied in this dissertation. The Hamiltonians for the system-environmental coupling could be even more mysterious. The computational complexity of accurately solving the quantum mechanical equations for realistic systems, like polyatomic molecules, is also a concern. Second, it is also difficult to reliably implement theoretical control designs in the laboratory, due to instrumental noise, the constraints on field parameters imposed by the apparatus, and other limitations.

A milestone in the development of optimal control experiments was the introduction of adaptive feedback control (AFC, also referred to as closed-loop laboratory control or learning control) in 1993 [6], which paved the way for optimal control of physical and chemical quantum phenomena with shaped light pulses. The AFC approach optimizes the applied control to meet a specified objective without relying on any exact priori knowledge about the system, but mainly by collecting and learning from the experimental data, making it especially suitable for highly complex systems which lack precise models, such as large molecules in a liquid solution [7]. A typical AFC procedure starts from some randomly picked initial controls, uses measurement results on the quantum system to evaluate the performance of the applied control and to refine it, thus forming an iterative closed loop until the control objective is reached as best as possible. Within each cycle of the loop, the external control (e.g. a shaped pulse) is applied to the system (e.g. an ensemble of molecules). The signal (e.g. population in a target state, or the yield of a desired reaction product) is detected and fed back to a learning algorithm. The algorithm evaluates each
control based on its measured outcome with respect to the control objective, and searches through the space of available controls to move towards an optimal solution. The procedure incorporates any laboratory constraints on the control fields.

The works in the dissertation realized purely experimental AFC of single and multiple nuclear spin systems on the testbed of pulsed nuclear magnetic resonance (NMR) spectroscopy, by gradient-based learning algorithms with experimentally measured gradients, cf. Eq. (1.6). The features of NMR control will be discussed in more details in Section 1.2 and the following chapters, from the perspectives of physical natures and experimental technologies.

1.1.2 Quantum control theory and landscape analysis: General formulation

In a broad range of quantum control experiments based on the method of AFC, people noticed that the optimal controls were generally easy to find in the iterative search procedure, considering the high-dimensional space of control parameters. Attempts to seek reasonable explanation of these phenomena from optimal control theory finally entailed the analysis of control landscapes. This section provides a brief framework of the physical model and typical problems of quantum control in the common semiclassical setting, i.e., the system and control being described by quantum and classical mechanics, respectively, and the main focuses of landscape analysis.

The time evolution of a closed quantum system (having no interaction with the environment) controlled by a classical field, expressed as a continuous time-domain function \( u(t) \), is governed by the Schrödinger equation

\[
\frac{i\hbar}{\hbar} \frac{d}{dt} U(t) = H[u(t)] U(t), \quad U(0) = I. \tag{1.1}
\]
Here, $U(t)$ is the unitary evolution operator of the system at time $t$, $\mathbb{I}$ is the identity operator, and $H$ is the time-dependent total Hamiltonian involving the control field $u(t)$. For example, the quantum dynamics of an atomic or molecular system interacting with a linearly polarized electric field in the dipole approximation can be adequately described by

$$H(t) = H_0 - u(t)H_1,$$  \hspace{1cm} (1.2)

where $H_0$ is the internal Hamiltonian and $H_1$ is an operator for the dipole moment along the polarization direction of $u(t)$. Based on the same dynamical rules, three commonly considered objective functions $J$ with practical meanings are given below.

(i) For observable control of a quantum ensemble, the goal is typically to maximize the expectation value of a target quantum observable $O$ at the final time $T$, given the initial density matrix $\rho_0$ at time $t = 0$. The corresponding cost function is

$$J = \text{Tr}[U(T)\rho_0U^\dagger(T)O]$$  \hspace{1cm} (1.3)

(ii) A simple but important special case of Eq. (1.3) is transition probability from an initial state $|i\rangle$ to a final state $|f\rangle$, both being pure states, i.e.,

$$J = |\langle f|U(T)|i\rangle|^2 = \text{Tr}[U(T)|i\rangle\langle i|U^\dagger(T)|f\rangle\langle f|],$$  \hspace{1cm} (1.4)

which is related to Eq. (1.3) by $\rho_0 = |i\rangle\langle i|$ and $O = |f\rangle\langle f|$.  

(iii) For evolution-operator control, the goal is to generate $U(T)$ such that its difference from a target unitary transformation $W$ is as small as possible. One commonly used form of cost function in this case can be expressed as

$$J = \|U(T) - W\|^2_{F} = 2N - 2\text{Re}\text{Tr}[W^\dagger U(T)]$$  \hspace{1cm} (1.5)
where $\| \cdot \|_F$ is the Frobenius norm of a matrix, and $\text{Re}$ takes the real part of a complex number. This type of objective (or its variations) can be used to characterize the fidelity of a quantum gate Pauli, the building blocks of quantum information sciences [8].

$J$ as a functional of the field $u(t)$ forms the aforementioned control landscape, and optimization of $J$ with the gradient ascent/descent algorithm is a process of ascending/descending the landscape along the steepest direction in the control space. We introduce a scalar parameter $s$ to describe the evolution of the field $u(t)$ from a given starting point $u_0(t)$, and the optimization process follows the differential equation

$$\frac{d}{ds} u(s, t) = \alpha(s) \frac{\delta J}{\delta u(s, t)}, \quad u(0, t) = u_0(t), \quad (1.6)$$

coupled to the quantum dynamical equation (1.1). $\alpha(s)$ is the step size controlling the optimization rate, which is positive/negative when the cost function $J$ is to be maximized/minimized. It is the discretized version of this equation that is applicable in simulations and experiments.

This simple algorithm can optimize the value of $J$ monotonically, and it terminates at the critical points of the landscape where the gradient of $J$ with respect to the control field is zero for all time, i.e.,

$$\frac{\delta J}{\delta u(t)} = 0, \quad \forall t \in [0, T]. \quad (1.7)$$

Should the landscape contain any suboptimal critical points, the gradient algorithm might terminate there prematurely, which is undesirable for the seek of globally optimal control solutions. A suboptimal critical control $u(t)$ can be further identified as a saddle point or a trap by the local landscape topology about it, assessed by the Hessian matrix which is the second derivative of $J$ with respect to $u(t)$:

$$\mathcal{H}(t, t') = \frac{\delta^2 J}{\delta u(t) \delta u(t')}, \quad t, t' \in [0, T]. \quad (1.8)$$
Roughly, a positive/negative semidefinite Hessian (having positive/negative eigenvalues only) indicates a local maximum/minimum, while an indefinite Hessian (having both positive and negative eigenvalues) indicates a saddle point on the landscape. The two types of critical points influence gradient ascent/descent trajectories passing by them in different ways: a trap can attract and terminate the nearby trajectories, while a saddle only temporarily slow them down in general (since there exist better solutions in $J$ value around a saddle). Analysis of control landscape topology, especially the distribution and properties of critical points, can help to explain and predict the behavior of optimization searches in typical quantum control problems like Eqs. (1.3)-(1.5).

1.1.3 Important theoretical conclusions about kinematic landscapes

For the control landscape problems (1.3)-(1.5) defined in a closed quantum system, the cost functions $J$ are all determined by the unitary evolution operator $U(T)$. Utilizing this fact can bring much simplification of the landscape analysis, since the theoretically infinite degrees of freedom in the control field $u(t)$ are compressed into only $N^2$ ones in a $N \times N$ unitary matrix. We assume that any element of the unitary group $U(N)$, or the special unitary group $SU(N)$ if the Hamiltonian $H(t)$ has a zero trace, can be produced as $U(T)$ by some admissible control $u(t)$ according to Eq. (1.1), a property referred to as evolution-operator controllability (see Section 1.1.4). The dependence of $J$ on $U(T)$ is called the *kinematic* landscape which is simpler to analyze, in contrast to the dependence on $u(t)$ known as the *dynamic* landscape. Some important conclusions on the properties of kinematic landscapes are summarized below, whose generalization to the dynamic landscapes will be addressed later in Section 1.1.4. We still focus on the distribution of kinematic critical points and their local topologies. The gradient and Hessian of $J$ are evaluated with respect to $U(T)$ on the unitary group, whose rigorous definitions can be found in [9].

For the landscape (1.3), a necessary and sufficient condition for a kinematic critical
point is that \( [U(T)\rho(0)U^\dagger(T),O] = 0 \), i.e., the system is driven to a final state \( \rho(T) = U(T)\rho_0U^\dagger(T) \) that commutes with the observable \( O \). Such critical points only exist at some discrete \( J \) values determined by the eigenspectra of \( \rho_0 \) and \( O \). Let \( \{\rho_i\} \) and \( \{O_i\} \) be the eigenvalues of \( \rho_0 \) and \( O \), respectively, both sorted in descending order, and the maximal and minimal values of \( J \) are given by [10]

\[
J_{\text{max}} = \sum_{i=1}^{N} \rho_i O_i, \quad J_{\text{min}} = \sum_{i=1}^{N} \rho_i O_{N-i+1}. \tag{1.9}
\]

The Hessian matrix of \( J \) evaluated at an intermediate critical point located between \( J_{\text{max}} \) and \( J_{\text{min}} \), if any, is proved to be indefinite [10]. Thus, the landscape has the local topology of a saddle point, instead of a local optimum, around such a critical point. The Hessian at any kinematic critical point has a rank (number of nonzero eigenvalues) of no more than \( N(N-1) \), and thus also possess a null space (the space of Hessian eigenvectors associated with zero eigenvalue). When a critical point is continuous varied within its Hessian null space, the \( J \) value will remain unchanged. This result implies that the optimal control solutions (and also the saddle controls) form critical manifolds, instead of existing as isolated points. An optimal control exhibits an inherent degree of robustness (tolerance to noise) when implemented in an application, since small noise components within the Hessian null space do not attenuate the optimal control performance. This can be an attractive behavior for attaining practical control, as noise is inevitably present in the laboratory, especially in laser pulse shaping. These conclusions also apply to the state transition landscape (1.4), a special case of (1.3), except that there are no additional critical points between its global maximum at \( J_{\text{max}} = 1 \) and global minimum at \( J_{\text{min}} = 0 \).

In parallel, the unitary transformation landscape (1.5) also has similar properties [11]. A necessary and sufficient condition for a critical point is that \( W^\dagger U = U^\dagger W \), where \( W \) is the targeted unitary operator. The critical points only exist at \((N+1)\) evenly space values of \( J \), i.e., \( J = 0, 4, \cdots, 4N \). Critical points other than the global maximum and minimum
are all saddle points.

These conclusions shed light on the general success of optimal control experiments and simulations. Most importantly, on a landscape containing saddle points only but not traps, a search trajectory starting from a random initial point with the gradient algorithm will almost always converge to a globally optimal solution.

1.1.4 Assumptions and limitations of current landscape analysis

Generalization of conclusions from kinematic landscape analysis (especially the absence of local traps) to dynamic landscapes, which are of more practical interest, is based on three underlying assumptions below. The dynamic landscape under conditions that violate any of the assumptions is an area for further research.

First, the closed quantum system must be controllable. Assessing the controllability of a system is one of the fundamental issues in the control theory, which has been well studied for closed [12] and open [13] quantum systems. Roughly, controllability describes the possibility of driving the system from an initial configuration to a final configuration within a finite time $T$. Controllability is determined by the equation of motion as well as properties of the Hamiltonian. For a closed system with unitary dynamics (1.1), which preserves the spectrum of the quantum state (i.e., the density matrix eigenvalues), evolution-operator controllability implies that for any unitary transformation $W$ there exists a finite time $T$ and a control $u(t)$ on the time interval $[0,T]$ such that $U(T) = e^{i\phi}W$, where $U(T)$ is the solution to (1.1) and $\phi$ is an arbitrary global phase. A criterion for such controllability is that the Lie algebra generated by the set of skew-Hermitian operators, $\{iH_0, iH_1\}$, has dimensionality $N^2$ (or $N^2 - 1$ for traceless Hamiltonians) [12]. Fortunately, this condition is almost always satisfied by random Hamiltonians $H_0$ and $H_1$, as shown by theoretical analysis [14]. Some special physical systems may require multiple independent control fields $u_k(t)$ acting on Hamiltonian terms $H_1^k$ to guarantee their controllability, including
the coupled nuclear spins studied in the following chapters.

Second, all control fields $u(t)$ are regular on the dynamic landscape. A subtle concept in landscape topology is the classification of a control as regular or singular [9]. Generally, a field $u(t)$ on the landscape $J[u(t)]$ is regular if the map $V_T: u(t) \mapsto U(T)$, which bridges the gap between kinematic and dynamic landscapes, is surjective in its vicinity, i.e., if for any local increment $\delta U(T)$ in the evolution operator there exists an increment $\delta u(t)$ in the control function such that $V_T[u(t)+\delta u(t)] = V_T[u(t)] + \delta U(T)$. This condition is equivalent to requiring that the Jacobian matrix $\frac{\delta U(T)}{\delta u(t)}$, which is a matrix-valued function in the time interval $[0, T]$, has full rank. A control $u(t)$ that violates this condition is referred to as singular. The phenomenon of singularity could bring about additional critical points on the dynamic landscape that have no counterparts on the kinematic one, or convert a kinematic saddle point to a trap-like critical point. Although the presence of singular controls is general, it requires much stronger conditions (depending on the specific control objective) for them to bring about noticeable landscape features like traps. Further discussion on singular critical points and traps with demonstration in practical systems will be given in Chapter 4.

Finally, theoretical landscape analysis originally deals with the ideal situation that the control field is unconstrained, i.e., it can be any time-domain function, but the presence of constraints is usually inevitable and significant in practice, e.g., in the shaping of ultrafast laser pulses. For both experiments and numerical simulations, a control has to be specified by a finite number of variables describing the intensity, phase, time duration, spectral frequencies, etc. of the field, whose values may be bounded for practical considerations. The control landscape structure with constrained fields in various forms is a focus of recent studies, many of which were performed by numerical simulation in few-level model systems for assessing the impact of constraints on seeking optimal control.

Moore et al. explored the control landscape for pure-state population transfer, Eq. (1.4), with a constrained class of control fields. The phases associated with spectral frequencies
were used as the control variables, in line with the laser pulse shaping technique. A “$2N - 2$ rule” was found that at least $2N - 2$ wisely-chosen independent controls ($N$ is the number of energy levels in the system) are required to achieve a high probability of reaching the landscape optimum with the gradient algorithm. When an insufficient number of controls and/or a weak control fluence are employed, trapping extrema and saddle points emerge on the landscape, which are not predicted by analysis of the unconstrained case [15, 16]. It is also showed on other types of landscapes that severe constraints on the number of control variables, the field strength, and the duration can prevent gradient searches from reaching the global optimum [17]. Zhdanov et al. addressed the observable landscape, Eq. (1.3), in a two-level system with upper bounds imposed on the field strength and control time, and found that although the traps always exist in this case, in practice they become trivially escapable if the control time is fixed and chosen long enough [18]. Rach et al. found that in the case of bandwidth-limited control pulses, the traps arising from the constraint can be eliminated [19]. In an experiment the strongly constrained landscape for laser-driven molecular fragmentation was observed, with only three phase variables as the control [20].

Besides the impact of constraints on landscape topology, Sugny et al. studied the constraint of time-integrated zero area on the control field, a physically fundamental requirement related to the fact that the dc component of an electromagnetic field is not a solution of Maxwells equation. This form of constraint can be crucial for laser pulses containing only a few optical cycles [21]. Shu et al. developed a theoretical method for optimizing the control fields subject to multiple constraints, while guaranteeing monotonic convergence towards the desired physical objective [22]. These works also aim to facilitate the search for optimal controls in practical problems.
1.2 Quantum control of nuclear spins with NMR

In the pulsed NMR spectroscopy, the system (an ensemble of atomic or nuclear spins) is typically steered by means of sequences of electromagnetic pulses, usually in the radiofrequency (RF) regime. These control methods have developed by the NMR community in the last 60 years, and attracted attention from the control perspective only recently \[14\].

A conventional one-dimensional NMR experiment, as those performed in the following chapters, consists of the sequential stages: (i) A relaxation delay of a few seconds, which allows the system to relax to its thermal equilibrium state; (ii) Implementation of pulses or pulse sequences that steer the nuclear magnetic moments, usually on the timescale of micro- to milliseconds; (iii) Real-time acquisition of the free induction decay (FID) signal that characterizes the free evolution of the system with the control pulse turned off. Fourier transform of FID gives the frequency-domain NMR spectrum, which characterizes the system state at the starting instant of data acquisition immediately following the control duration. Such experimental setting is very convenient for the fundamental study of the observable control problems in Eq. (1.3), and enhancement of signal intensity by optimal control methods is also meaningful for the application of NMR in chemical and biological analysis and medical imaging \[23\].

Like the general form in Eq. (1.2), a NMR system also has its internal Hamiltonian and control Hamiltonian. The former gives the energy levels of single and coupled nuclear spins in a static magnetic field, while the latter arise from application of RF pulses to the system at or near its resonant frequencies.

**Control of a single spin**

The dynamics of a single spin-$1/2$ particle in a magnetic field \( \mathbf{B} = (B_x, B_y, B_z)^T \) is governed by the Hamiltonian

\[
H = -\gamma \mathbf{B} \cdot \mathbf{I} = -\gamma (B_x I_x + B_y I_y + B_z I_z)
\]  

(1.10)
where $\gamma$ is the gyromagnetic ratio of the nucleus, and $I_a (a = x, y, z)$ is the spin angular momentum operator in the corresponding direction, related to the Pauli matrices by $I_a = \frac{\hbar}{2} \sigma_a$. In the basis of the two eigenstates of $I_z$, the Pauli matrices can be represented as

$$
\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad (1.11)
$$

and the general properties of quantum mechanical angular momentum operators are satisfied, i.e.,

$$
[I_x, I_y] = i\hbar I_z, \quad [I_y, I_z] = i\hbar I_x, \quad [I_z, I_x] = i\hbar I_y. \quad (1.12)
$$

In conventional liquid-state NMR experiments a strong static magnetic field $B_0$ is applied along the $z$ direction, providing the internal Hamiltonian as

$$
H_0 = -\gamma B_0 I_z := -\omega_0 I_z \quad (1.13)
$$

where $\omega_0/2\pi$ is the Larmor precession frequency. Eq. (1.13) shows that the energy splitting caused by the magnetic field, known as the Zeeman splitting, has an amount of $\hbar \omega_0$ proportional to the field strength $B_0$. Typical values of $B_0$ are 5-15 Tesla, resulting in precession frequencies $\omega_0$ of a few hundred MHz in the RF regime for $^1$H, the most useful isotope for chemical analysis. Spins of different isotopes can have very different $\gamma$ values and Larmor frequencies, and thus can be easily distinguished spectrally. Spins of the same isotope can also have distinct frequencies due to their chemical shifts, which arise from partial shielding of the external field $B_0$ by the electron cloud surrounding the nuclei. The magnitude of shielding depends on the chemical environment of each nucleus. The range of typical chemical shifts varies with isotopes, e.g., $\sim 10$ part per million (ppm) for $^1$H, $\sim 200$ ppm for $^{13}$C and $^{19}$F.

The state of a spin-1/2 can be manipulated by applying a shaped electromagnetic pulse
at the frequency $\omega_{RF}$ near resonance with the spin Larmor frequency $\omega_0$. Suppose that the pulse is circularly polarized in the $x$-$y$ plane with its magnetic fields given by

$$B_x(t) = A \cos(\omega_{RF}t + \phi), \quad B_y(t) = A \sin(\omega_{RF}t + \phi) \quad (1.14)$$

where $A$ is the amplitude of the RF field and $\phi$ its phase. Typical values of $|\gamma A|$ for $^1\text{H}$ are up to $\sim 10^2$ kHz, orders of magnitude weaker than the static field $B_0$. The time-dependent control Hamiltonian is thus

$$H_{\text{control}}(t) = -\gamma [B_x(t)I_x + B_y(t)I_y]. \quad (1.15)$$

The technology of arbitrary RF waveform generation [24] facilitates the quantum control of NMR systems.

**Spin Coupling**

The interaction between nuclei leads to additional Hamiltonian terms besides the internal and control Hamiltonians of each individual spin as given in Eqs. (1.13) and (1.15). Nuclear spins in molecules generally have two distinct interaction mechanisms, the direct dipole-dipole interaction, and the electron mediated Fermi contact interaction (known as $J$-coupling or scalar coupling) [5]. The dipole-dipole interaction is similar to the interaction between two bar magnets: It takes place purely through space (requiring no medium) and depends on the internuclear vector $\vec{r}_{ij}$ connecting nuclei $i$ and $j$. For molecules in liquid solution, dipole-dipole couplings are averaged away due to rapid tumbling, which is the case of all experiments in this dissertation. $J$-coupling is mediated by the electrons shared in the chemical bonds between the atoms, and due to the overlap of the shared electron wavefunction with the two coupled nuclei. The through-bond coupling strength $J$ depends on the associated nuclear species and decreases with the number of chemical bonds separating the nuclei. Typical values for $J$ can range from a few hundred Hz for one-bond
couplings to a few Hz for three- or four-bond couplings. The \( J \)-coupling Hamiltonian for two nuclei \( i \) and \( j \) is

\[
H_J = \frac{2\pi}{\hbar} J_{ij} \mathbf{I}_i \cdot \mathbf{I}_j = \frac{2\pi}{\hbar} J_{ij}(I_{ix}^i I_{ix}^j + I_{iy}^i I_{iy}^j + I_{iz}^i I_{iz}^j)
\]

where \( J_{ij} \) is the coupling strength in the unit of Hz, and operators like \( I_{ix}^i I_{iy}^j \) are the Kronecker product of single-spin operators on spins \( i \) and \( j \). When \( |\omega_{0,i} - \omega_{0,j}| \gg 2\pi |J_{ij}| \), a condition easily satisfied by heteronuclear spins or homonuclear spins with well separated chemical shifts, Eq. (1.16) simplifies to

\[
H_J = \frac{2\pi}{\hbar} J_{ij} I_{iz}^i I_{iz}^j
\]

which is known as the Ising-type coupling. A control landscape study based on this physical model will be provided in Chapter 4.

**Relaxation**

The interaction of a quantum ensemble with its environment leads to *decoherence*, or non-unitary evolution of its density matrix \( \rho(t) \). Generally, decoherence arises from the distinct evolution dynamics of individual particles within the ensemble. For nuclear spin systems the decoherence, also known as *relaxation*, is conventionally modeled by two parameters for each individual spin: the spin-lattice relaxation time \( T_1 \), and the spin-spin relaxation time \( T_2 \) (more complex relaxation mechanisms may appear in coupled spin systems). \( T_1 \) arises from couplings between the spins and the “lattice”, that is, excitation modes which can carry away energy quanta on the scale of the Larmor frequency. For example, these may be vibrational quanta, paramagnetic ions, chemical reactions such as ions exchanging with the solvent. \( T_2 \) relaxation originates from spin-spin couplings which are imperfectly averaged away, or unaccounted for in the system Hamiltonian. For example, in molecules in liquid solutions, spins on one molecule may have a long range, weak
interaction with spins on another molecule. Fluctuating magnetic fields, caused by spatial anisotropy of the chemical shift, local paramagnetic ions, or unstable laboratory fields, also contribute to $T_2$ [5].

In well prepared liquid-state NMR experiments $T_1$ and $T_2$ can easily be seconds or longer, unusually slow when compared with the lifetimes of vibrational, rotational, or electronic molecular energy levels. One advantage of slow relaxation is that it allows for manipulation and observation of quantum states on a relatively long timescale, which is especially important for the usage of nuclear spins as quantum bits in quantum computation [25, 8], although there is still controversy on the usefulness of NMR quantum computing [26].

Optimal control of spin systems and its realization with NMR spectroscopy has been well studied in the past two decades [27]. Important theoretical topics include the minimization of control time [29], the boundaries of objective functions in the presence of decoherence [30], etc. A gradient algorithm for designing optimal pulse shapes, known as GRAPE (gradient ascent pulse engineering) [28], has been applied to many NMR quantum control and computation experiments. In Chapters 2-4 we apply a similar algorithm in AFC by experimentally measuring the gradient, so that the optimal controls under the real laboratory conditions can be found. Moreover, second derivative analysis by estimation of the landscape Hessian will also be carried out.
Bibliography


Chapter 2

Experimental exploration over a quantum control landscape through NMR

The growing successes in performing quantum control experiments motivated the development of control landscape analysis as a basis to explain these findings. When a quantum system is controlled by an electromagnetic field, the observable as a functional of the control field forms a landscape. Theoretical analyses have revealed many properties of control landscapes, especially regarding their slopes, curvatures and topologies. A full experimental assessment of the landscape predictions is important for future consideration of controlling quantum phenomena. NMR is exploited here as an ideal laboratory setting for quantitative testing of the landscape principles. The experiments are performed on a simple two-level proton system in a H$_2$O-D$_2$O sample. We report a variety of NMR experiments roving over the control landscape based on estimation of the gradient and Hessian, including ascent or descent of the landscape, level set exploration, and an assessment of the theoretical predictions on the structure of the Hessian. The experimental results are fully consistent with the theoretical predictions. The procedures employed in this study provide
the basis for future multi-spin control landscape exploration where additional features are predicted to exist.

2.1 Introduction

The control of quantum phenomena is garnering increasing interest for fundamental reasons as well as potential applications. Quantum system optimal control seeks to meet a posed physical objective, such as selective breaking of chemical bonds, creation of particular molecular vibrational excitations, manipulation of electron transfer in biological complexes, etc. (see [1] for a review). The control objective is generally addressed through the introduction of tailored electromagnetic fields, whose behavior is honed for the particular application. A practical means for the discovery of optimally designed fields is through the use of closed-loop learning control in the laboratory [2], where the objective signal measurement is fed back to a learning algorithm to evaluate the success of a candidate control field for refinement until the objective is reached as best as possible. The generally successful outcomes of optimal control experiments, as well as the extensive simulations on model systems, indicate that it is relatively easy to find good solutions while searching through the vast space of possible control fields. Seeking a fundamental explanation of this good fortune motivated the development of quantum control landscape analysis [3, 4, 5] which, upon satisfaction of particular assumptions, provides quantitative predictions on the nature of landscape features that can be tested in the laboratory. Many of the theoretical predictions have been successfully tested in large numbers of simulations [6, 7], but experimental affirmation is important for fundamental and practical reasons. For laser field manipulation of atomic, molecular or condensed phase systems, various laboratory complexities can make quantitative testing of landscape features challenging in these settings. Nuclear magnetic resonance (NMR) is best known as a powerful tool for chemical analysis, but for our purposes, the advanced nature of NMR instrumentation producing high
signal-to-noise ratios (S/N) provides an ideal laboratory setting for testing the predictions from control landscape analysis. Control concepts have long been exploited experimentally in NMR [8, 9, 10, 11, 12, 13, 14, 15], and in this work we will utilize these capabilities for an initial study of control landscape principles in the most basic case of a two-level single-spin system. Although working with two or more coupled spins is needed to fully test the landscape analysis predictions [4, 5], the simple case of a single spin naturally must first be assessed.

In quantum control problems, like the case addressed here, the system evolution is governed by a time-dependent Hamiltonian, which is a function of a pulsed control field \( C(t) \). The amplitude, phase, and/or frequency of the field can be modulated to meet the control objective, i.e., maximizing or minimizing an experimental measurable \( J \) at the target time \( T \). The objective may be expressed as \( J = \text{Tr}[\rho(T)O] \), where \( O \) is a Hermitian observable operator and the density matrix \( \rho(T) \) depends on the control \( C(t) \) [4]. The functional dependence of \( J[C(t)] \) upon the control field forms a control landscape. The time interval \( T \) is chosen to be sufficiently long to permit unfettered control, while being short enough to consider the spin dynamics to form a closed system.

The topology of the landscape is important for determining the effectiveness of searches seeking an optimal control. A location on the landscape where the gradient satisfies \( \delta J/\delta C(t) = 0, \forall t \in [0, T] \), specifies a critical point. An analysis of the landscape Hessian \( \mathcal{H}(t, t') \equiv \delta^2 J/\delta C(t) \delta C(t') \) eigenvalues can reveal whether the critical point is a local extremum or saddle point. The number of positive and negative Hessian eigenvalues identify the intrinsic topological character at a critical point [5]. A critical point is locally maximal (or minimal) if the Hessian is negative (or positive) semidefinite, while a saddle point will have an indefinite Hessian spectrum.

The conclusions of control landscape analysis rest on the assumptions that (i) the system is controllable [16], (ii) the control to final state map, \( C(t) \mapsto \rho(T) \), is surjective [17] and (iii) the controls are unconstrained [18]. Upon satisfaction of these assumptions,
control landscape analysis predicts that the critical points only exist at particular values of the objective $J$ and that there are no local suboptimal maxima or minima (traps) over the landscape [4], thereby providing a basis to understand the observed relative ease of searching for optimal controls both in simulations [6, 7] and in the laboratory [1]. Although assumptions (i) and (ii) can be violated, the latter reported results suggest that they may be commonly satisfied. Assumption (iii) is always a concern in the laboratory where control resources are inevitably limited. The primary issue is whether the assumptions are satisfied to a practical degree in order to give good quality control performance.

The Hessian at a critical point has a specific maximum number of positive and negative eigenvalues dependent on the system’s Hilbert space dimension, the initial density matrix $\rho(0)$ and the nature of the observable operator $O$ [5]. The latter Hessian non-zero eigenvalues are accompanied by an infinite dimensional null space. The present chapter illustrates the most basic example of a control landscape for a proton spin-1/2 two-level system. Upon satisfaction of the three assumptions, this case is predicted to have no critical points except at the global maximum and minimum, and the Hessian at the landscape top (bottom) should possess at most two negative (positive) eigenvalues, respectively. This ideal control landscape picture from theoretical analysis might be altered by experimental imperfections such as relaxation through interaction with the environment and constraints on the control field [18], whose detailed impacts on the landscape structure remains an open challenge to assess. Extensive simulations have affirmed the predicted control landscape topology (see e.g., [6, 7]), and a recent laser experiment on pure state transitions in atomic rubidium was also consistent with expectations [19]. The present work demonstrates a set of experimental tools capable of roving over quantum control landscapes, including for future experiments in systems of interacting (coupled) spins and other circumstances.

The physical process underlying NMR can be viewed as the manipulation of magnetization vectors (proportional to the spin angular momentum) with pulsed magnetic fields in the radiofrequency (RF) regime as controls. NMR provides a desirable domain for study-
ing fundamental properties of quantum control landscapes for several reasons: (i) Most nuclear spin systems have relatively simple and well-defined Hamiltonians. An isolated single spin-1/2 nucleus only possesses two energy levels upon interaction with the magnetic field, so the Hilbert space of a \( n \)-spin system is \( 2^n \)-dimensional. Most molecules considered for laser control are too complex to model precisely, while spin systems under NMR control can often be modeled reliably with optimal control theory (OCT) as well as being readily amenable to the performance of optimal control experiments (OCEs) (see Figure 2.1). Thus, the option exists to either implement an OCT designed pulse shape in the laboratory to increase the efficiency of OCE, or utilize the laboratory OCE optimized pulse in a simulation to study the detailed controlled spin dynamics. (ii) Spin systems are well approximated as closed when the control interval \( T \) is much shorter than both the longitudinal (system-environment) relaxation time \( (T_1) \) and the transverse (spin-spin) relaxation time \( (T_2) \) [8]. (iii) NMR machines are at an advanced stage of engineering capable of producing high fidelity shaped control pulses, and the observed spectra generally have high S/N of \( 10^3 \sim 10^4 \), thus providing good accuracy for the measured observable. (iv) At conventional magnetic field strengths the Zeeman splitting is much smaller than \( k_B T \) at room temperature. Therefore, at thermal equilibrium a spin system is in a mixed state, which enables testing the general landscape topology predictions [4, 5]. The collective features (i)-(iv) above indicate that NMR provides a very attractive regime to test control landscape principles. Importantly, as these principles have a generic foundation, the implications of the findings extend beyond NMR to the control of other classes of systems using lasers or other sources.

Optimal control of nuclear spin systems has been treated theoretically and experimentally [9, 10, 11, 12, 13, 14, 15, 20]. Generally the experiments are executed with either of two approaches: (a) perform OCT to design an optimal control pulse and apply it in the laboratory [9, 10, 11]; (b) directly perform OCE by iterative enhancement of a spectral signal to determine the control pulse [12, 13, 14, 15]. Both approaches have achieved success in
the optimization of polarization transfer or dipolar decoupling in two-spin systems. However, these works have not addressed the analysis of the control landscape features, which is the focus of this chapter. With regard to Figure 2.1, in this work OCT was performed in parallel with the OCE studies in order to provide a full understanding of the observations.

The remainder of the chapter is organized as follows. Section 2.2 describes the methodology, including experimental determination of the gradient and Hessian and a series of algorithms as a basis to rove over the landscape. Section 2.3 gives the experimental setup for the two-level spin system in NMR (i.e., the proton in the HDO molecule in a mixture of 1% H$_2$O and 99% D$_2$O), the objective $J$ and control variables. Experimental results are provided in Section 2.4 and a brief conclusion is given in Section 2.5.

Figure 2.1: (Color online) The relationship between the dual loops of optimal control theory (OCT) and optimal control experiments (OCE) in NMR systems. Both control loops share the common features of implementing a field, calculating or measuring the cost function, assessing how well the objective is achieved, and varying the control iteratively guided by an algorithm. The high quality of spin system Hamiltonians permits balanced operation with OCT and OCE, as called for by the control objectives. In the present work on assessing control landscape analysis predictions, the experiments were exclusively performed with OCE and the findings subsequently affirmed in OCT simulations. In more complex scenarios with multiple spins the performance of OCT may provide reasonable control estimates to reduce the level of experimental effort.
2.2 Methods for experimental landscape roving

Experimental optimal control of quantum systems commonly use pulsed electromagnetic fields with adjustable amplitudes and phases over some spectral range as the input, and then the response signal from the physical system is recorded as the output. The various continuous control variables in the laboratory are inevitably discretized in some fashion, which we encapsulate here into a vector of \( D \) control variables, \( \vec{x} = (x_1, \cdots, x_D)^\top \).

The focus of this work is on control landscapes, and the gradient and Hessian characterize the local features around a particular point \( J(\vec{x}_0) \) on the landscape for control \( \vec{x}_0 \). The landscape gradient and Hessian evaluated at \( \vec{x}_0 \) respectively correspond to the vector \( \nabla J \),

\[
\nabla J(\vec{x}_0) = \left( \frac{\partial J}{\partial x_1}, \cdots, \frac{\partial J}{\partial x_D} \right)^\top \bigg|_{\vec{x}_0},
\]

(2.1)

and the matrix \( \mathcal{H} \) with elements

\[
\mathcal{H}_{ij}(\vec{x}_0) = \mathcal{H}_{ji}(\vec{x}_0) = \frac{\partial^2 J}{\partial x_i \partial x_j} \bigg|_{\vec{x}_0}, \quad i, j = 1, \cdots, D.
\]

(2.2)

We will construct “Rover” algorithms for systematically exploring the landscape utilizing the gradient and/or Hessian, whose experimental determination in the present two-level spin system will be described in Section 2.2.2.

2.2.1 The laboratory landscape Rover algorithms

In this work we introduce the concept of a control landscape Rover, which results from executing a suitable algorithm for taking an exploratory trajectory over the landscape in the laboratory. A trajectory can be characterized by a progress parameter \( s \geq 0 \), i.e., \( \vec{x}(s) \), and the corresponding objective value \( J[\vec{x}(s)] \). The trajectory roving over the landscape can be
described by an ordinary differential equation

\[
\frac{d\vec{x}(s)}{ds} = \vec{F}[\vec{x}(s)],
\]

(2.3)

where the form of \( \vec{F} \) is dictated by the particular landscape exploration goal. Four basic choices for \( \vec{F} \) are given below.

The landscape roving operations in this work are constructed from four elementary operations, which can be understood as (i) continued movement along a fixed specified direction in the control space, (ii) performance of steepest ascent or descent of the landscape, (iii) horizontal movement on the landscape that preserves a non-critical value of \( J \), and (iv) horizontal movement at a critical value of \( J \). Arbitrary roving of the landscape may be performed by interleaving these elementary procedures as desired, with various possible goals including categorization of local landscape features beyond those of critical points.

- **Movement along a fixed specified direction.**

  The movement in this case consists of “marching” in the same fixed direction \( \vec{c} \) in the space of controls,

  \[
  \frac{d\vec{x}(s)}{ds} = \vec{c},
  \]

  (4-i)

  such that \( \vec{x}(s) = \vec{x}(0) + s\vec{c} \). The choice of \( \vec{c} \) is application specific, but a simple circumstance arises in the three cases below when the function \( \vec{F} \) in Eq. (2.3) for vertical or horizontal movement (c.f., Eqs. (4-ii), (4-iii) or (4-iv)) is treated as a constant over a significant domain of the landscape.

- **Movement vertically.**

  Here the movement is guided by the gradient for ascent or descent of the landscape,

  \[
  \frac{d\vec{x}(s)}{ds} = \alpha \nabla J[\vec{x}(s)],
  \]

  (4-ii)

  where \( \alpha \) is a scale factor. By solving Eq. (4-ii) for \( \vec{x}(s) \), \( J \) will increase (\( \alpha > 0 \)) or decrease.
$(\alpha < 0)$ monotonically until a critical point is reached.

- **Movement horizontally over a non-critical level set.**

Although extremizing $J$ is a common physical goal, exploring all landscape features is important for assessing the basic theory. Thus, a third case is horizontal movement on the landscape, or level set exploration [21, 22] corresponding to continuous roving without changing the value of $J$. A quantum control level set is defined by a family of control solutions which all achieve the same objective value $J$, and the distinct members over the level set may show large variation of secondary characteristics. Simulations have explored non-critical level sets (i.e., away from where $\nabla J = 0$) [21] as well as the landscape top/bottom [22]. A trajectory on a non-critical level set is characterized by the local change in the control $d\vec{x}(s)$ being orthogonal to the gradient $\nabla J[\vec{x}(s)]$. In the $D$-dimensional control space, there are $(D - 1)$ linearly independent directions orthogonal to the gradient, so the level set trajectory from a specified starting point is not unique. The following differential equation satisfies these conditions:

$$
\frac{d\vec{x}(s)}{ds} = \vec{g} - \nabla J(\nabla J^\top \cdot \vec{g}) / (\nabla J^\top \cdot \nabla J),
$$

(4-iii)

where $\vec{g}$ is an arbitrary vector of length $D$ whose choice will guide the level set trajectory.

- **Movement horizontally over a critical level set.**

Eq. (4-iii) does not apply at landscape critical points, such as the top or bottom, where the gradient is zero. In this case the level set trajectory lying in the null space of the Hessian will keep the value of $J$ invariant at the critical point,

$$
\frac{d\vec{x}(s)}{ds} = \sum_i \vec{v}_i(s) \left[ \vec{v}_i^\top(s) \cdot \vec{h} \right],
$$

(4-iv)

where $\{\vec{v}_i\}$ are the eigenvectors of the Hessian with zero eigenvalue, and $\vec{h}$ is an arbitrary vector of length $D$ whose choice guides the critical level set exploration trajectory. In the
extreme case that the Hessian had no zero eigenvalues, the critical point would exist in isolation on the landscape forbidding movement as the r.h.s. of Eq. (4-iv) would be zero.

Eqs.(4(i-iv)) are special cases of Eq. (2.3) which will be used during the experiments reported in Section 2.4. Since the landscape gradient and Hessian at an arbitrary point can be measured in the laboratory, the differential equations can be solved in real time while performing the ongoing experiments. In this work the forward Euler method was found to be sufficient,

\[ \vec{x}(k + 1) = \vec{x}(k) + \beta \vec{F}[\vec{x}(k)], \quad k = 0, 1, \cdots \] (2.5)

where \( \vec{x}(k) \) is the control in the \( k \)-th iteration (i.e., the \( k \)-th step of \( s \)) and \( \beta \) is the step size.

In other applications, especially when the S/N is not high, statistical averaging of the data at each step and higher order integration methods may be needed.

### 2.2.2 Gradient and Hessian determination

Experimental determination of the gradient and Hessian of objective \( J \) in this work is based on making small increments about a current control \( \vec{x}_0 \) and then measuring the resultant changes in the associated \( J \) values. For the landscape gradient we found that a simple central difference method was stable,

\[ \frac{\partial J}{\partial x_i} \approx \frac{J(\cdots, x_i + d_i, \cdots) - J(\cdots, x_i - d_i, \cdots)}{2d_i}, \quad i = 1, \cdots, D, \] (2.6)

where \( d_i \) is a small increment of the variable \( x_i \) which should be reasonably chosen based on the nature of \( x_i \) and \( J \) in a particular experiment. Similarly, the second-order partial derivatives in the Hessian also can be standardly expressed with finite differences. However, this method has a high S/N requirement for measuring \( J \) because of noise sensitivity, and we found that estimation of the Hessian was problematic by direct application of finite differences. Statistical strategies can be employed to reliably extract quality gradients and
Hessians from experimental data [19, 23]. In this work we utilize least squares (LS) to determine the Hessian from the data $J(\vec{x}_0 + \Delta \vec{x})$ with a set of perturbations $\Delta \vec{x}$. For this purpose the landscape can be approximated about $\vec{x}_0$ by a second-order Taylor series

$$J(\vec{x}_0 + \Delta \vec{x}) \approx J(\vec{x}_0) + \nabla J(\vec{x}_0)^\top \cdot \Delta \vec{x} + \frac{1}{2} \Delta \vec{x}^\top \cdot \mathcal{H}(\vec{x}_0) \cdot \Delta \vec{x}. \quad (2.7)$$

Determining $\mathcal{H}(\vec{x}_0)$ also requires extracting the gradient $\nabla J(\vec{x}_0)$, and the linear system in Eq. (2.7) has $D(D + 3)/2$ unknowns: $D$ for the gradient and $D(D + 1)/2$ for the Hessian. With sufficient random samples of $\Delta \vec{x}$'s about $\vec{x}_0$ (i.e., typically $\sim$500 samples were used in this work) the overdetermined linear system can be solved with LS to obtain the Hessian. Although the gradient is also determined in solving Eq. (2.7), when only the gradient was called for in landscape ascent or descent and for non-critical level set exploration, the more efficient central difference method in Eq. (2.6) was used.

### 2.3 Experimental setup

The dynamics of a two-level system, i.e., a nuclear spin-1/2 with gyromagnetic ratio $\gamma$, is quantum-mechanically formulated as follows. A closed quantum system in a mixed state is described by its density matrix $\rho(t)$, whose time evolution is governed by the Hamiltonian $H(t)$ according to the Liouville-von Neumann equation,

$$\frac{d}{dt} \rho(t) = -\frac{i}{\hbar} [H(t), \rho(t)]. \quad (2.8)$$

In the case of a two-level spin system, a static and homogeneous magnetic field $B_0$ is implemented; its orientation defines the $z$-axis for the nuclear magnetization both in the laboratory and in the rotating frame. A shaped RF control pulse is applied orthogonal to $B_0$ (i.e., in the $x$-$y$ plane) with a carrier frequency of $\omega_{RF} > 0$. The pulse has adjustable amplitude $A(t)$ and phase $\phi(t)$, which are usually slowly varying functions in the time do-
main compared with the carrier frequency. In a rotating frame about the $z$-axis at frequency $\omega_{RF}$, whose orientations of $+x$ and $+y$ are defined with the phase angles of $0^\circ$ and $90^\circ$, respectively, we conveniently define two new functions in terms of $A(t)$ and $\phi(t)$ to specify a control pulse, i.e., $B_x(t) := A(t) \cos \phi(t)$ and $B_y(t) := A(t) \sin \phi(t)$. This formulation enables us to express the full Hamiltonian in the rotating frame as [8]

$$H(t) = -(\gamma B_0 - \omega_{RF}) I_z - \gamma [B_x(t) I_x + B_y(t) I_y],$$

(2.9)

where $I_x$, $I_y$ and $I_z$ are the spin angular momentum operators, related to the Pauli matrices as $I_j = \hbar \sigma_j / 2$, $j = x, y, z$. Although for special choices of $B_x(t)$ and $B_y(t)$ Eq. (2.8) may be solved analytically, here we seek to establish the capability of freely roving over the landscape and testing the landscape predictions which generally requires full freedom in the field forms.

The NMR experiments presented in this chapter were implemented on a Bruker Avance-III 500 MHz spectrometer, equipped with a TCI ($^{1}$H/$^{13}$C/$^{15}$N/$^{2}$H) cryoprobe and highly digitized and linear RF signal generator (SGU) (Bruker-Biospin, Billerica, MA). We used a sealed standard 1%H$_2$O/99%D$_2$O sample which contains a small amount of GdCl$_3$ to accelerate the $T_2$ relaxation processes. The measured values for $T_1$ and $T_2^*$ (which includes local inhomogeneity effects as well) are 187 ms and 70 ms, respectively. In this mixture the rapid exchange and the overwhelming excess of D$_2$O assures that the dominant species available for $^1$H-detection is HDO. The H/D coupling is effectively removed by the rapid exchange as well (i.e., exchange decoupling [24]). All these features permit treating the sample as a two-level system consisting of a single spin with a singlet resonance. All the experiments were performed at 295K after careful manual tuning and shimming of the magnet. The carrier frequency $\omega_{RF}$ was set exactly on-resonance, thus the $I_z$ term of the Hamiltonian in Eq. (2.9) can be dropped. In this work each shaped pulse with a fixed final time of $T = 500\mu$s is broken into four equal time intervals of constant field value,
so the control is given as a vector of length $D = 8$, i.e.,

$$\vec{x} = (B_x^1, \ldots, B_x^4, B_y^1, \ldots, B_y^4)^\top,$$

where $B_x^i$ ($B_y^i$) is the value of the corresponding control within the $i$-th time interval. The set of eight variables is an approximation to the true control as a freely varying continuous function of time, and in the present experiments this choice of discrete variables proved to be adequate for satisfying resource assumption (iii) of control landscape analysis. With the above experimental setup each measurement takes $\sim 3$ sec of laboratory time, including the pulse duration, data acquisition time of $\sim 0.5$ s, and $\sim 2.5$ s waiting for the system to relax back to its equilibrium state.

The control objective considered here is to steer the initial state $\rho(0) \propto I_z$ under thermal equilibrium to $-\langle I_x \rangle$ with a suitably shaped pulse. The objective $J = -\langle I_x \rangle$ is characterized by the integrated area of the corresponding singlet peak in the NMR spectrum. The zeroth-order phase correction [25] is calibrated manually so that the peak area represents the magnetization in the desired orientation $-x$, and then fixed throughout the experiment; the first-order phase correction [25] is always set to zero. If the maximal peak integral is $J_{\text{max}}$, then the full domain of $J$ will be $[-J_{\text{max}}, J_{\text{max}}]$ since the peak can be either positive or negative depending on the control pulse. To determine the level of noise we took 100 repeated measurements of several typical control pulses producing various $J$ values on the landscape; in each case the error approximately obeyed a Gaussian distribution with a standard deviation of $(10^{-4} \sim 10^{-3})J_{\text{max}}$.

### 2.4 Results and discussion

The laboratory experiments aimed to demonstrate the landscape roving capability while also assessing the particular predictions of landscape analysis. Starting from an arbitrarily chosen initial control field, we first performed optimization by ascent and descent of the landscape to reach the top and bottom in Section 2.4.1, resulting in a full landscape trajectory. The Hessian spectrum was evaluated at the top and bottom of the landscape in Section
2.4.2, along with performance of an extensive excursion through the space of controls while roving on the top of the landscape. Section 2.4.3 presents two non-critical level set exploration experiments optimizing a secondary characteristic to find the associated best control upon traversing a landscape level set.

### 2.4.1 Ascending and descending the landscape

Landscape roving starts with gradient ascent or descent, i.e., maximization or minimization of the objective $J$ by moving the control along the local gradient measured in each iteration, until either halting at a suboptimal critical point trap or reaching the full landscape top/bottom extrema. Arbitrary units are used for the objective and control fields below, and their respective scales will be consistent throughout Section 2.4 (e.g., a constant $90^\circ$ pulse with a length of $T = 500\mu s$ corresponds to a field strength of $\sim 37$). Figure 2.2(a) shows a typical optimization curve of $J$ containing $\sim 100$ iterations. The randomly chosen initial control corresponds to the 0th iteration, which was then optimized for both ascent and descent on the landscape to give the full curve. The Euclidean norm of the experimentally measured landscape gradient at each iteration is also displayed. As the search approached the landscape maximum or minimum, $J$ converged and continued to randomly vary in a narrow range with $J_{\text{min}} \simeq -J_{\text{max}}$ within the level of noise, while the gradient norm converged toward zero. The gradient norm curve is more noisy than that of $J$, especially in the near-optimal regions, due to noise amplification. The control landscape analysis predicts that upon satisfaction of the three underlying assumptions, there should be no suboptimal traps or saddle points on such a two-level system landscape [4], which is confirmed by the experiments.

The control field at the initial iteration as well as at the landscape maximum and minimum in the trajectory are shown in Figure 2.2(b). When a magnetic field interacts with a spin-1/2 proton it will rotate the magnetization vector clockwise about the magnetic field
direction. During maximization and minimization of the control objective $\langle I_z \rangle \rightarrow -\langle I_x \rangle$, the $B_x$ component of the control gradually diminishes while $B_y$ becomes dominant, providing the desired rotation of the initial magnetization along the $+z$ axis toward the desired orientation along $x$. Note that there are infinitely many optimal solutions for this control problem, the simplest among which is just a DC field in the $+y$ orientation of the rotating frame with a flip angle of $90^\circ$ to give the maximal value of $J$. However, in the trajectory we present here even at the landscape optimum the $B_x$ component is not zero. Once the initial control $\vec{x}(0)$ is specified, the gradient ascent trajectory $\vec{x}(s), s \geq 0$ is deterministic and only one optimal control is found. Thus, level set exploration is employed below to access other portions of the landscape.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{landscape.png}
\caption{(Color online) Landscape ascent/descent along the directions specified by the gradient of $J$ (the peak integral of the $^1$H NMR spectrum in HDO). (a) The evolution of the objective $J$ and the norm of its gradient $(\nabla J^\top \cdot \nabla J)^{1/2}$ during optimization. The initial condition corresponds to the 0-th iteration shown as a bold dot for $J$; the iterations proceed in the positive (negative) direction to ascend (descend) the landscape. (b) The control pulse shapes at the landscape maximum and minimum points in the above search trajectory, compared with the randomly picked initial control.}
\end{figure}

2.4.2 Hessian observation over the landscape and exploration of the landscape top

The Hessian matrix gives an assessment of the curvature at any location on the landscape. In particular, the landscape analysis for a two-level system predicts a specific rank
of at most 2 for the Hessian at the top and bottom critical points [5], which can be tested in the laboratory. To further assess the Hessian character we evaluated it using least squares (see Section 2.2.2) at five selected locations, $J/J_{\text{max}}=1.00, 0.71, 0.31, 0.03, \text{ and } -1.00$, along the gradient ascent/descent search trajectory in Figure 2.2. Approximately 500 small random control sampling variations were performed about each of the five points to determine the associated Hessian matrices with their eigenvalues shown in Figure 2.3. In the middle region of the landscape, the Hessian has positive and negative eigenvalues with comparable magnitudes. As $J$ approaches the landscape top, the Hessian spectrum moves toward being negative semidefinite. At the maximum point ($J/J_{\text{max}}=1.00$), two Hessian eigenvalues are significantly negative while the remaining six are zero within experimental error, in agreement with the theoretical prediction. Similarly, the Hessian spectrum at the landscape minimum $J/J_{\text{max}}=-1.00$ has two positive eigenvalues with the remaining six being zero within the level of noise, again agreeing with theory [5]. Simulations were performed considering the laboratory noise level, confirming the statistical quality of the results.

The Hessian eigenvectors associated with non-zero eigenvalues at a maximum (minimum) point of the landscape describe the independent paths for driving off the landscape top (bottom) region, and the magnitudes of the eigenvalues characterize the sensitivity of $J$ to variation along these paths in the control space. Similarly, the eigenvectors with zero eigenvalue specify directions for remaining on the top (bottom) of the landscape. The two eigenvectors $\vec{v}_1$ and $\vec{v}_2$ determined from the measured Hessian at the maximum point $J/J_{\text{max}}=1.00$ are presented in Figure 2.4, corresponding to the eigenvalues $\lambda_1 = -316$ and $\lambda_2 = -150$, respectively. The eigenvectors show that $J$ can be decreased by perturbing the optimal control in two independent coordinated ways: (i) adding an approximate constant field to the $B_y$ component expressed by $\vec{v}_1$, (ii) adding a specially shaped field to $B_x$ with a particular weaker variation of $B_y$, expressed by $\vec{v}_2$.

As an assessment of the landscape structure about an extremum point [19], straight
roving was performed with Eq. (4-i) along each of the Hessian eigenvectors in the vicinity of the initial critical point. Using the landscape maximum at \( \vec{x}_0 \) shown in Figure 2.2(b) as a starting point, we moved along each of the eight eigenvectors with the successive choice of \( \vec{c} = \vec{v}_i, i = 1, \cdots, 8 \) in Eq. (4-i). The process was characterized by the relative distance of a control \( \vec{x} \) from the starting point \( \vec{x}_0 \) in the control space, i.e., \( ||\vec{x} - \vec{x}_0||/||\vec{x}_0|| \), as shown in Figure 2.5. A parabolic drop of \( J \) was observed when moving along the directions specified by either of the two eigenvectors \( \vec{v}_1 \) and \( \vec{v}_2 \). In contrast, \( J \) remained almost constant while moving along the other six eigenvectors over appreciable roving distances of at least \( \pm 100\% \), indicating that the Hessian null space has broad extent. Importantly, through the choice of the vector \( \vec{h} \) in Eq. (4-iv) motion along the top of the landscape also can be expressed by a wandering trajectory specified as an arbitrary linear combination of the Hessian null space eigenvectors.

Building on the results in Figure 2.5 and the comment above, we implemented a trajectory to extensively “drive” over the top of the landscape by employing Eq. (4-iv), i.e., remeasuring the Hessian about the critical point in each iteration and moving in the null space of the local Hessian. Again we started from the maximum point found in Section 2.4.1 and chose a simple constant free vector \( \vec{h} = (1, 1, 1, 1, 1, 1, 1, 1) \) to obtain the control trajectory on the top of the landscape given in Figure 2.6. The Hessian in each iteration was estimated by the LS with 100 random samples, whose spectrum was always found to be dominated by two significantly negative eigenvalues, and the other six Hessian eigenvalues being essentially zero. The objective value stayed very close to the landscape maximum throughout the exploration, and the vector of control variables evolved from the starting point by a relative distance of more than \( 240\% \) at the final iteration. These results again suggest that the top of the landscape is quite extensive, although no attempt was made to fully explore its scope.
Figure 2.3: The Hessian spectrum evaluated at five selected values of $J/J_{\text{max}}$ along the gradient ascent (descent) search in the trajectory of Figure 2.2. The spectrum swings from being negative semidefinite at $J/J_{\text{max}} = 1.00$ to being positive semidefinite at $J/J_{\text{max}} = -1.00$. The existence of two (±) non-zero eigenvalues at the landscape bottom (top) is consistent with theoretical predictions.

Figure 2.4: (Color online) The two Hessian eigenvectors $\vec{v}_1$ and $\vec{v}_2$ respectively corresponding to the two dominant eigenvalues $\lambda_1$ and $\lambda_2$ at the landscape top ($J/J_{\text{max}}=1.00$).
Figure 2.5: (Color online) Driving on and off the landscape top by proceeding along directions specified by each of the eight Hessian eigenvectors at the starting point. Continued marching along the directions specified by either of the two eigenvectors with negative eigenvalues takes $J$ down the landscape along nearly parabolic paths. In contrast, marching along any of the six eigenvectors with approximately zero eigenvalue leaves $J \simeq J_{\text{max}}$ while the control accordingly changes form. In all cases, movement along any of the eigenvectors by relative distances even greater than $\pm 100\%$ still preserved the physical meaning of the eigenvectors reflected in their eigenvalues (i.e., being either negative or zero).

Figure 2.6: Shapes of the evolving control field in a 10-iteration exploration of the landscape top by roving iteratively in the Hessian null space utilizing controls given by solving Eq. (4-iv).

### 2.4.3 Non-critical objective value level set exploration

Section 2.4.2 examined the critical level set at the top of the landscape, while here we consider non-critical exploration at $J \neq J_{\text{max}}$ or $J_{\text{min}}$. As described by Eq. (4-iii), traversal of a non-critical level set requires movement along a path that is locally orthogonal to the
gradient $\nabla J$ on the landscape. In practice, this is performed by choosing any free vector $\vec{g}$ in Eq. (4-iii) with the projection operation on the r.h.s. removing the component of $\vec{g}$ along the gradient. A randomly varying $\vec{g}$ may lead to a random walk in the level set, but here we will take advantage of the freedom in choosing $\vec{g}$ by a specification that optimizes a secondary characteristic goal $K[\vec{x}(s)]$.

As a first example, we consider minimizing the energy $K_E$ of the control pulse:

$$K_E[\vec{x}(s)] \propto \int_0^T \left[ B_x^2(t) + B_y^2(t) \right] dt \propto \sum_{i=1}^4 \left[ (B_i^x)^2 + (B_i^y)^2 \right] = \|\vec{x}(s)\|^2. \quad (2.10)$$

Minimizing the pulse energy can be achieved by choosing the vector $\vec{g}(s)$ proportional to the derivative of the energy function, i.e.,

$$\vec{g}(s) = -\frac{\vec{x}(s)}{\|\vec{x}(s)\|} \quad (2.11)$$

The weight $\|\vec{x}(s)\|$ in the denominator scales $\vec{g}$ to the local magnitude of the control. The experimental results in Figure 2.7(a) start with an arbitrarily chosen initial control producing $J_0 = 0.585 J_{\text{max}}$ and then a level set trajectory is taken while minimizing $K_E(\vec{x})$. The pulse energy decreased by more than 50% over $\sim 20$ iterations, while the $J$ value varied in a narrow range of $\sim 0.003 J_{\text{max}}$. The error bars in Figure 2.7(a) for $J/J_{\text{max}}$ were determined from the standard deviations of five repeated measurements at each iteration. The fluctuation of the level set value was comparable with the random noise in measuring $J$, thereby showing good stability for the level set exploration. The evolution of the control field over the level set is given in Figure 2.7(b). The results show that the energy-optimal control pulse on the level set is converging towards a simple DC field in $y$ direction.
Figure 2.7: An experiment for energy-minimization on a non-critical point level set exploration: (a) The pulse energy drops significantly by over 50% at the end of the excursion while \( J \) remains nearly constant showing a variation of less than 0.3\% \( J_{\text{max}} \). (b) The controls at iterations 0 (initial), 5, 10, 15, and 22 (final) in (a), showing evolution of the control during the level set exploration.

As a second example of non-critical level set exploration, we desire to move as far as possible from the starting point in the control space. Therefore, we consider the secondary characteristic \( K_d \) of the Euclidean distance squared from the initial control \( \vec{x}(0) \):

\[
K_d[\vec{x}(s)] = \| \vec{x}(s) - \vec{x}(0) \|^2,
\]

and maximization of \( K_d \) can be achieved by defining the free vector proportional to the derivative of \( K_d \), i.e.,

\[
\vec{g}(s) = \vec{x}(s) - \vec{x}(0).
\]

The free vector is randomly chosen at the initial iteration to start the exploration, and after that it is determined by Eq. (2.13). In order to move a long distance on a landscape level set efficiently in the laboratory where measuring the gradient is time-consuming, we modestly sacrifice the precision of the level set. Thus, we adopt the procedure of continuing to exploit the current gradient direction as far as possible during the exploration in order to reduce the frequency of gradient measurement. We move along that constant direction combining Eqs. (4-i) and (4-iii) until the drop from the original level set value \( J_0 \) exceeds
a specified tolerance, and then we measure the gradient and correct the control by gradient ascent or descent to bring it back to the original level set value $J_0$ followed by further level set exploration, etc. The step size of the gradient correction is determined as follows.

According to the gradient ascent algorithm in Eq. (4-ii), we obtain

$$\frac{dJ}{ds} = \frac{\partial J}{\partial \bar{x}(s)} \cdot \frac{d\bar{x}(s)}{ds} = \alpha \left\| \frac{\partial J}{\partial \bar{x}(s)} \right\|^2 = \alpha \left\| \nabla J[\bar{x}(s)] \right\|^2. \quad (2.14)$$

A gradient correction of $J$ is performed with Eq. (2.5) (where $\beta = \alpha$ now) as follows at the $k$-th iteration $\bar{x}(k)$ in order to return $J$ back to $J_0$:

$$\bar{x}(k + 1) = \bar{x}(k) - \frac{J - J_0}{\parallel \nabla J \parallel^2} \nabla J \bigg|_{\bar{x}(k)} \quad (2.15)$$

with the coefficient $-(J - J_0)/\parallel \nabla J \parallel^2$ assuring ascent or descent as needed and conservative movement by the normalization in the denominator. Therefore, the full landscape level set algorithm is a combination of horizontal and vertical restoring elementary movements as follows:

1) Specify a non-critical initial control $\bar{x}(0)$, measure the objective value $J_0$ and the gradient at $\bar{x}(0)$, and randomly choose a free vector $\bar{g}$ for Eq. (4-iii);
2) Determine the roving direction by removing from $\bar{g}$ its projection on the gradient through the operation on the r.h.s. of Eq. (4-iii);
3) Move the control $\bar{x}$ along the level set roving direction by a constant distance (thus the coefficient $\beta$ in Eq. (2.5) is non-constant in this example) combining Eqs. (4-i) and (4-iii), and measure the new $J$;
4) If $|J - J_0| \leq \epsilon$, where $\epsilon$ is the tolerance for deviation from the original level set, go to step 3);

If $|J - J_0| > \epsilon$, measure the gradient and correct the control according to Eq. (2.15), then measure the new $J$ at the corrected control. When $|J - J_0| < \epsilon$ is obtained, reset the
free vector $\vec{g}$ by Eq. (2.13) and then go to step 2;

5) Iterate until $\vec{x}(s)$ has evolved a specified distance $\sqrt{K_d}$ in the control space from $\vec{x}(0)$.

Figure 2.8(a) shows a distance-maximization level set roving experiment using the algorithm above, where the tolerance for being acceptably close to $J_0$ in order to define a level set is chosen to be $|J - J_0| \leq 0.014J_{\text{max}}$. Over the set of ∼30 iterations the gradient correction was performed only three times, and after each correction $J$ returned to its initial value $J_0$ (dashed line in the figure) with good precision. By the end of the roving trajectory the control vector had changed by a relative distance of >250%, while taking only less than 20 min of laboratory time. The controls at several selected iterations, i.e., the initial and final ones as well as before and after each gradient correction, are shown in Figure 2.8(b). We see that the corrections only caused relatively small variations of the control while returning to the original $J_0$ value, and the overall procedure did not break the continuity of the level set exploration to the specified tolerance.

Figure 2.8: A landscape non-critical level set roving experiment aiming to move as far away as possible from an initial control: (a) Evolution of the objective $J$; the three sudden jumps correspond to gradient ascent corrections when the deviation of $J$ from the initial value $J_0 = 0.482J_{\text{max}}$ exceeded the tolerance of more than $0.014J_{\text{max}}$. The horizontal dashed line shows the position of the targeted level set. (b) The control fields at the initial and final iterations, as well as before (after) each gradient correction shown by the three groups of solid (dashed) lines. The experiment was terminated when the roving field evolved by a relative distance of ∼250%.
2.5 Conclusion

This work reported the first experimental study of an NMR control landscape including extensive roving. The high S/N in NMR experiments enabled the determination of the gradient and Hessian at any point over the landscape with good precision and efficiency. With knowledge of the local gradient and/or Hessian, we implemented various landscape roving algorithms serving specific purposes, including optimizing the objective $J$, finding controls on a non-critical level set while optimizing the value of a secondary characteristic and exploration of the top of the landscape. The theoretical predictions on the landscape topology were verified with good accuracy. The methodology developed in this chapter can be further exploited to assess the control landscape analysis in more complex coupled multi-spin systems, which are expected to have saddle points at particular values of $J$ with specified Hessian signatures. The findings of such landscape topology assessment studies are of interest in spin systems under control, but the implications extend beyond for the control of other quantum phenomena with electromagnetic fields.
Bibliography


Chapter 3

Experimental observation of saddle points over the quantum control landscape of a two-spin system

The growing successes in performing quantum control experiments motivated the development of control landscape analysis as a basis to explain these findings. When a quantum system is controlled by an electromagnetic field, the observable as a functional of the control field forms a landscape. Theoretical analyses have predicted the existence of critical points over the landscapes, including saddle points with indefinite Hessians. This chapter presents the first systematic experimental study of quantum control landscape saddle points. Nuclear magnetic resonance (NMR) control experiments are performed on a coupled two-spin system in a $^{13}$C-labeled chloroform ($^{13}$CHCl$_3$) sample. We address the saddles with a combined theoretical and experimental approach, measure the Hessian at each identified saddle point, and study how their presence can influence the search effort utilizing a gradient algorithm to seek an optimal control outcome. The results have significance beyond spin systems, as landscape saddles are expected to be present for the control of broad classes of quantum systems.
3.1 Introduction

The control of quantum phenomena is gathering increasing interest for fundamental reasons and potential applications. Quantum system optimal control is concerned with active manipulation of physical and chemical processes at the atomic and molecular scale, such as creation of particular molecular vibrational excitations, selective breaking of chemical bonds, and manipulation of electron transport in nanoscale devices (see [1] for a review). The control objective is generally addressed through the introduction of a semiclassical electromagnetic field whose shape is honed for the particular application. The generally successful outcomes of optimal control experiments and the extensive simulations on model systems suggest that while searching through the vast space of possible control fields, it is relatively easy to find good solutions. Seeking a fundamental explanation for this good fortune motivated the development of quantum control landscape analysis [2, 3, 4], which provides quantitative predictions on the landscape features that can be assessed in the laboratory. For laser field manipulation of atomic, molecular or condensed phase systems, various experimental complexities can make quantitative testing of landscape principles challenging. The advanced nature of nuclear magnetic resonance (NMR) instrumentation producing high signal-to-noise ratios (S/N) provides ideal laboratory circumstances for testing the predictions from control landscape analysis. In this work we will utilize these capabilities for the study of a control landscape in a two-spin system which possesses saddle points.

In quantum control applications, the system evolution is governed by a time-dependent Hamiltonian including an applied electromagnetic field. The amplitudes, phases, and/or frequencies of the field can be modulated to meet the control objective, i.e., maximizing or minimizing an experimental measurable objective $J$ at the target time $T$. The time interval $T$ is conventionally chosen to be sufficiently long to permit unfettered control, while being short enough to consider the dynamics to form a closed system. The functional depen-
dence of $J$ upon the control field forms a control landscape, whose topology, especially the distribution of critical points, may determine the effectiveness of a search for an optimal control. A critical point is defined as a location on the landscape where the gradient (the first functional derivative of $J$ with respect to the control) is zero for $t \in [0, T]$. An analysis of the Hessian matrix (second derivative of $J$) can identify the intrinsic topological character at a critical point [4]: a negative (or positive) semidefinite Hessian indicates that the critical point is locally maximal (or minimal) to second order, while an indefinite Hessian is associated with a saddle point.

Upon satisfaction of some underlying assumptions (see Section 3.2 for details), control landscape analysis predicts that the critical points only exist at particular values of the objective $J$ and that there are no local suboptimal maxima or minima (traps) over the landscape, i.e., the critical points located between the global maximum and minimum, if any, have to be saddle points [3]. Another important conclusion is the low rank of the Hessian at critical points, i.e., there exists a specific maximum number of positive and negative eigenvalues dependent on the nature of the quantum system and the control problem [4]. Many of the theoretical predictions have been successfully tested in large numbers of simulations [5, 6], but experimental affirmation in the laboratory is important for fundamental and practical reasons. A laser experiment on pure state transitions in atomic rubidium was found consistent with expectations in terms of the Hessian structure at the global minimum and maximum of the landscape [7], and as a foundation for the present chapter we explored the control landscape for a proton spin-1/2 two-level system [8]. The landscapes in both of the two latter works are devoid of saddle points. Thus, in order to fully test the predictions provided by landscape theoretical analysis we have to consider other more complex quantum systems with the simple example in this chapter being a coupled two-spin system.

NMR provides a desirable domain for studying fundamental properties of quantum control due to the simple and well-defined Hamiltonians involved, slow relaxation, high signal-to-noise ratio, etc. [8] The dynamical process underlying NMR can be viewed as
the manipulation of magnetization vectors (proportional to spin angular momenta) with pulsed magnetic fields in the radio frequency (rf) regime as controls. Quantum control of spin systems has been treated theoretically and experimentally [9, 10]. More specifically, for coupled two-spin systems the techniques of polarization or coherence transfer have attracted much interest [11, 12, 13] motivated by the desire of enhancing the signals of a low gyromagnetic ratio nucleus by transferring to it the larger magnetization of a higher gyromagnetic ratio nucleus [14]. Optimal control theory has been used to derive the physical limits of coherence or polarization transfer between two coupled spins in the presence of relaxation effects [15, 16, 17]. In the latter work the optimized pulse sequences were achieved experimentally and exhibited improved transfer efficiency compared with conventional schemes. The maximal amount of coherence that can be transferred in a given time in the absence of relaxation losses has also been studied [18, 19, 20]. The present chapter investigates the dynamics of a coupled two-spin system by addressing the underlying control landscape features.

The remainder of the chapter is organized as follows. In Section 3.2 we summarize the quantum-mechanical description of a two-spin system and the theoretical analysis of the associated topological nature of the control landscape. Section 3.3 includes the setup of our NMR experiments and some algorithmic and analysis tools that facilitate the landscape exploration in the laboratory. Experimental results are provided in Section 3.4 and a brief conclusion is given in Section 3.5.

3.2 Theoretical analysis of the quantum control problem

The dynamics of a heteronuclear system consisting of two weakly coupled spin-1/2 nuclei, labeled $I$ and $S$, is quantum-mechanically formulated as follows. The 4-dimensional Hilbert space is conventionally described by the basis of product operators [21] with $I_j := (\sigma_j/2) \otimes \mathbb{I}_2$, $S_j := \mathbb{I}_2 \otimes (\sigma_j/2)$, and $2I_jS_k := 2(\sigma_j/2) \otimes (\sigma_k/2)$, where $\sigma_j, \sigma_k \in \{\sigma_x, \sigma_y, \sigma_z\}$
are the Pauli matrices and $\mathbb{I}_2$ is the $2 \times 2$ identity matrix. Upon introduction of a static, homogeneous magnetic field $B_0$ whose orientation defines the $z$-axis, the internal Hamiltonian of the system is

$$H_0 = \hbar(\omega_0^I I_z + \omega_0^S S_z + 2\pi J_{IS} I_z S_z),$$

(3.1)

where $\omega_0^I = -\gamma^I B_0$ is the Larmor frequency of spin $I$ depending on its gyromagnetic ratio $\gamma^I$ (and similarly for spin $S$); $J_{IS}$ is the strength of the scalar coupling between spins $I$ and $S$ in units of frequency. Note that $H_0$ is given in the weak coupling regime, i.e., $2\pi |J_{IS}| \ll |\omega_0^I - \omega_0^S|$, which is commonly satisfied for two heteronuclear spins [22].

The system is simultaneously controlled by two pulsed oscillating magnetic fields along a particular orientation (defined as $x$) in the $x$-$y$ plane, whose carrier frequencies are resonant with the Larmor frequencies of spins $I$ and $S$ respectively [14], i.e.,

$$B^I(t) = 2A^I(t) \cos[\omega_0^I t + \phi^I(t)], \quad B^S(t) = 2A^S(t) \cos[\omega_0^S t + \phi^S(t)].$$

(3.2)

The envelope amplitudes $A^I, A^S$ and phases $\phi^I, \phi^S$ of the two control fields can be tailored as functions of time. Provided that the control fields are sufficiently weak such that each spin can be addressed individually, i.e., the interaction between $B^I$ and spin $S$ is negligible and vice versa, the full Hamiltonian in the laboratory frame is

$$H(t) = H_0 - \hbar[\gamma^I B^I(t) I_x + \gamma^S B^S(t) S_x].$$

(3.3)

Within the rotating wave approximation, the Hamiltonian $\tilde{H}$ in a doubly on-resonance rotating frame [22] reads

$$\tilde{H} \approx \hbar[2\pi J_{IS} I_z S_z - \gamma^I A^I(\cos \phi^I \cdot I_x + \sin \phi^I \cdot I_y) - \gamma^S A^S(\cos \phi^S \cdot S_x + \sin \phi^S \cdot S_y)]$$

:= $\hbar(2\pi J_{IS} I_z S_z - u_x^I \cdot I_x - u_y^I \cdot I_y - u_x^S \cdot S_x - u_y^S \cdot S_y),$$

(3.4)
where the controls $u_k^x(t) := \gamma_k A_k^x(t) \cos \phi_k(t)$ and $u_k^y(t) := \gamma_k A_k^y(t) \sin \phi_k(t)$, $k = I, S$, are introduced for each spin. This Hamiltonian form has been used in simulations [9], and provides full controllability for the system [23] to satisfy the first assumption of the control landscape analyses below. In the remainder of this chapter we will describe the controls in terms of $u^I_x, u^I_y, u^S_x$ and $u^S_y$ which have the units of Rabi frequency.

Some key results of control landscape analysis are briefly summarized below. For a generic closed quantum system under control, the Hamiltonian $H(t)$ including the control fields generates a unitary transformation $U(t)$ governed by the time-dependent Schrödinger equation,

$$i\hbar \frac{d}{dt} U(t) = H(t) U(t), \quad U(0) = \mathbb{I}. \quad (3.5)$$

The state of the system at time $t$ is described by its density matrix $\rho(t) = U(t) \rho(0) U^\dagger(t)$.

For a specified initial state $\rho(0)$ and target observable operator $O$, the control landscape

$$J = \text{Tr}[\rho(T)O] = \text{Tr}[U(T)\rho(0)U^\dagger(T)O] \quad (3.6)$$

can be treated as a function of the unitary transformation $U(T)$, which forms the kinematic landscape [4]. The topology of the kinematic landscape $J[U(T)]$ is simply determined by the eigenspectra of $\rho(0)$ and $O$. The necessary and sufficient condition for a critical control, defined at the landscape gradient being zero, is that

$$[\rho(T), O] = 0. \quad (3.7)$$

Thus, there are critical final states $\rho(T)$ located at some particular levels of the landscape [3], which are denoted by $J_{\text{crit}}$’s. The Hessian signatures at a critical point, i.e., the maximal allowed numbers of positive and negative eigenvalues of the Hessian, denoted by $D_+$ and $D_-$ respectively, are proved to be finite and can be determined from the degeneracies of $\rho(0)$ and $O$ by the contingency table method [4]. The set of critical points sharing common
values of $J_{\text{crit}}$ and $D_+, D_-$ form a critical submanifold (CM). The dynamical landscape (i.e., $J$ as a functional of the control fields) topology characterized by $J_{\text{crit}}$ and $D_+, D_-$ values is fully consistent with its kinematic counterpart if three assumptions are satisfied: (i) the system is controllable [23, 24], (ii) the control to final state map is surjective [25] and (iii) the controls are unconstrained [26]; a discussion of the assumptions is given in the cited references. Practical evidence suggests that assumptions (i) and (ii) may be commonly satisfied, while assumption (iii) is always a concern in the laboratory where control resources are inevitably limited. The primary issue is whether the assumptions are satisfied to a practical degree in order to give good quality control performance. The comparison of kinematic versus dynamic landscapes has been extensively discussed in a number of papers and we refer the reader to Ref. [25] for details.

In this chapter we use the thermal equilibrium state $\rho_{\text{eq}}$ of the two-spin system as the initial state, which is determined by the Boltzmann distribution and approximated under the condition that $\|H_0\| \ll k_B T$ as

$$\rho_{\text{eq}} = \frac{e^{-H_0/k_B T}}{\text{Tr}(e^{-H_0/k_B T})} \approx \frac{1}{4} (\mathbb{I}_4 - \frac{H_0}{k_B T}) = \frac{1}{4} - \frac{\hbar}{4k_B T} (\omega_0^I I_z + \omega_0^S S_z + 2\pi J_{IS} I_z S_z).$$

(3.8)

In conventional NMR experiments, $J_{IS} (10^0 \sim 10^2 \text{ Hz})$ is orders of magnitude smaller than $|\omega_0^I|$ and $|\omega_0^S|$ ($\sim 10^8 \text{ Hz}$), therefore the coupling term in $\rho_{\text{eq}}$ is negligible, and the landscape topology is invariant to whether this term is retained. We may ignore the term with $\mathbb{I}_4$ and simply specify $\rho_{\text{eq}} = I_z + r S_z$ in the following landscape analysis, where $r := \omega_0^S / \omega_0^I = \gamma^S / \gamma^I$ is defined on the range $0 < r < 1$ without loss of generality. For the target observable we choose $O = I_x$, the in-phase $x$-magnetization of spin $I$ [21], as an example, but the same landscape topology applies for any operator unitarily equivalent to $I_x$, such as $S_x$, $2I_z S_x$ and $2I_x S_z$. The topological nature of the landscape can be derived theoretically [4], and the outcome is summarized in Table 3.1. $\rho(0) = I_z + r S_z$ has four non-degenerate eigenvalues,
\[-(1+r)/2, -(1-r)/2, (1-r)/2, \text{ and } (1+r)/2,\]

while \(O = I_x\) has two doubly-degenerate eigenvalues, \(\pm 1/2\). Using these numbers, the landscape CMs are calculated to lie at five distinct values of \(J_{\text{crit}}\) [3], and we name them Top, Saddle B, Saddle A, Saddle B' and Bottom from the highest \(J_{\text{crit}}\) to the lowest (see Table 3.1). The Hessian signatures of each CM are also given in Table 3.1. We find that all three suboptimal CMs, lying between the global maximum \(J_{\text{crit}} = 1\) and minimum \(J_{\text{crit}} = -1\), have indefinite Hessians and therefore their topology is that of saddles. As \(J_{\text{crit}}\) transcends its values from 1 to -1, the index \(D_+\) increases from 0 to 8 while \(D_-\) drops from 8 to 0. The CMs at the Top and Bottom are symmetric in terms of \(D_+, D_-\) and \(J_{\text{crit}}\), and so are Saddle B and B'. Thus, we will focus on three representative CMs in the following study, i.e. Top, Saddle A and Saddle B. This analysis specifies a control landscape in a two-spin quantum ensemble system, which will be studied experimentally in the following sections to assess the theoretical topological predictions.

<table>
<thead>
<tr>
<th>CM</th>
<th>(J_{\text{crit}})</th>
<th>(D_+)</th>
<th>(D_-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top</td>
<td>1</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Saddle B</td>
<td>(r)</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Saddle A</td>
<td>0</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Saddle B'</td>
<td>(-r)</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Bottom</td>
<td>(-1)</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>

### 3.3 Experimental setup

#### 3.3.1 NMR experiments

The NMR experiments presented in this chapter were implemented on a Bruker Avance-III 500 MHz spectrometer, equipped with a dual \(^{13}\text{C}-^{1}\text{H}\) (DCH) cryoprobe and highly digitized and linear rf signal generators (SGU) (Bruker-Biospin, Billerica, MA). We used
a sealed $^{13}$C-labeled chloroform ($^{13}$CHCl$_3$) sample in DMSO-d$_6$ with a small quantity of added Gd(NO$_3$)$_3$ to reduce relaxation times. The scalar coupling between the two spins, $I = ^1$H and $S = ^{13}$C, is $J_{IS} = 215$ Hz as measured from the frequency difference of the two lines of the $^{13}$C doublet in the NMR spectrum. After careful manual tuning and shimming, all the experiments were performed at 295K. The sample was simultaneously irradiated with two independently shaped rf pulses, whose carrier frequencies were set exactly on-resonance with the $^1$H and $^{13}$C nuclear spins, respectively. Both pulses had a fixed length of $T = 5$ ms, which was picked to make $J_{IS} \cdot T \gtrsim 1$, allowing for full dynamical controllability [20] as well as relatively weak relaxation loss. Each shaped pulse was represented by its field values during four equal time intervals of the length $T/4 = 1.25$ ms, so the entire control could be expressed as a vector of length 16, i.e., $\vec{x} = [u^I_x(1), \ldots, u^I_x(4), u^I_y(1), \ldots, u^I_y(4), u^S_x(1), \ldots, u^S_x(4), u^S_y(1), \ldots, u^S_y(4)]^T$, where $u^I_x(1)$ is the value of the control $u^I_x$ at time interval 1, etc. Note that the control variables were converted back in terms of amplitudes and phases according to Eq. (3.4) when specifying the two pulses in the laboratory, with each interval having constant amplitude and phase. As a reference, constant 90° pulses for spins $^1$H and $^{13}$C over the period $T = 5$ms respectively have magnitudes of 20.2 and 18.5 by our setup (see the amplitudes of shaped pulses in Figures 3.1(a) and 3.7(a) for comparison). With the experimental setup a single measurement took 1.2s of laboratory time for data acquisition, and at least an additional $\sim 3$ s to allow for both spins to relax back to equilibrium, leading to overall $\sim 4.2$ s for recycle time.

The thermal equilibrium initial state of the sample is $\rho(0) = I_z + rS_z$, with $r = \gamma(^{13}$C)/$\gamma(^1$H) $\approx 0.25$ [22], and the objective $J$ can be read from the frequency-domain NMR spectrum. The $^1$H- or $^{13}$C-detected spectrum of the sample has a doublet peak with two lines separated by the coupling constant $J_{IS}$ (see Figure 3.1(b) for examples). The sum of the integrated areas of the two lines characterizes the in-phase magnetization of the corresponding spin along a particular orientation in the $x$-$y$ plane at the target time.
This orientation is specified by the experimental detector phase parameter, so for spin $I$ we can set the observable as $O(\theta) := \cos \theta I_x + \sin \theta I_y$ with any value of the phase angle $\theta$. By referring to simulation results we found a proper detector phase that gives $O = I_x$, although the topological nature of the landscape given in Table 3.1 is independent of $\theta$. To determine the noise level we took 100 repeated measurements of several typical control pulses producing $J$ values at different locations on the landscape. The error approximately obeyed a Gaussian distribution, and the standard deviation was about $10^{-4} \sim 10^{-3}$ of the maximum value of $J = 1$.

### 3.3.2 Algorithms for laboratory landscape exploration

A recent work introduced a flexible Rover algorithm [8] for guiding an exploratory trajectory over the landscape in the laboratory in order to reveal the underlying structure. A trajectory can be characterized by a progress parameter $s \geq 0$, i.e., the vector of controls $\vec{x}(s)$, and the corresponding objective value $J[\vec{x}(s)]$. The control trajectory corresponding to roving over the landscape can be generally described by the ordinary differential equation

$$\frac{d\vec{x}(s)}{ds} = \vec{F}[\vec{x}(s)],$$

(3.9)

where the form of the roving direction $\vec{F}$ is dictated by the particular landscape exploration goal, often aided by utilization of the gradient and/or Hessian at $\vec{x}(s)$. In this chapter we will employ various rover algorithms based on the form in Eq. (3.9) to optimize the objective value, approach saddle points, and explore the neighborhoods of saddles. Specific forms of $\vec{F}$ for the particular applications will be given in Section 3.4. As the landscape gradient and Hessian at an arbitrary point can be experimentally measured, Eq. (3.9) will be numerically solved in real time with the ongoing experiments. In this work the forward
Euler method was found to be sufficient,

\[ \vec{x}(k+1) = \vec{x}(k) + \alpha \cdot \vec{F}[\vec{x}(k)], \quad k = 0, 1, \cdots \] (3.10)

where \( \vec{x}(k) \) is the control in the \( k \)-th iteration (i.e., the \( k \)-th step of \( s \)) and \( \alpha > 0 \) is the step size. In other applications, especially when the S/N is not high, statistical averaging of the data at each step as well as higher order integration methods may be needed.

Over an experimental landscape \( J[\vec{x}] \), the gradient and Hessian about a current control \( \vec{x}_0 \) respectively correspond to a vector \( \vec{\nabla} J \) and matrix \( \mathcal{H} \). Experimental determination of the gradient and Hessian in this work is based on making small increments about \( \vec{x}_0 \) and then measuring the resultant changes in the associated \( J \) values. For the landscape gradient a simple central difference method was found to be stable for the present application (except when \( \vec{x}_0 \) is very close to a critical point),

\[ \frac{\partial J}{\partial x_i} \approx \frac{J(\cdots, x_i + d_i, \cdots) - J(\cdots, x_i - d_i, \cdots)}{2d_i}, \] (3.11)

where \( d_i \) is a small increment of the variable \( x_i \) which should be reasonably chosen based on the nature of \( x_i \) and \( J \) in a particular experiment. Estimation of the Hessian in a similar fashion can be problematic because of the higher sensitivity to noise involved in the second-order differencing. With this consideration, statistical strategies have been employed to reliably extract quality gradients, and especially Hessians, from experimental data \([7, 27]\).

In this work we utilize least squares (LS) to determine the Hessian from the data \( J[\vec{x}_0 + \Delta \vec{x}] \) over a random set of perturbation \( \Delta \vec{x} \). For this purpose the landscape can be approximated about \( \vec{x}_0 \) by a second-order Taylor series

\[ J[\vec{x}_0 + \Delta \vec{x}] \approx J[\vec{x}_0] + \vec{\nabla} J[\vec{x}_0] \cdot \Delta \vec{x} + \frac{1}{2} \Delta \vec{x}^\top \cdot \mathcal{H}[\vec{x}_0] \cdot \Delta \vec{x}. \] (3.12)

With sufficient random samplings of \( \Delta \vec{x} \)'s about \( \vec{x}_0 \) the overdetermined linear system
can be solved with LS to obtain the Hessian and gradient simultaneously; when only the gradient was required, Eq. (3.11) was implemented experimentally.

3.4 Results and discussion

3.4.1 Gradient ascent over the landscape

Sections 3.4.1 through 3.4.3 focus on the control landscape of the observable \( O = I_x \), with the experimental \( J \) value obtained from the integral of the \( ^1H \) doublet peak; Section 3.4.4 will examine like behavior of \( O = S_x \) for \(^{13}C \). Firstly, we started from the initial control with all field components set to zero and maximized the objective \( J \) with the gradient ascent algorithm [cf. Eq. (3.10)],

\[
\vec{x}(k + 1) = \vec{x}(k) + \alpha \vec{\nabla}J[\vec{x}(k)].
\]

(3.13)

With a proper step size \( \alpha > 0 \), \( J \) should increase monotonically until reaching a critical point. For this search trajectory it was found that \( J \) increased from zero and converged within \( \sim 20 \) iterations (see trajectory 1 in Figure 3.2); the control pulse at the optimum and the resultant NMR spectrum are shown in Figure 3.1. The same procedure was also performed in simulation, where the value of \( J \) could be calculated by numerically solving the Schrödinger equation (3.5) with the Hamiltonian in (3.4) and substituting \( U(T) \) into Eq. (3.6). We confirmed that the optimum obtained from this experimental search was actually the global maximum, i.e., it belonged to the CM Top lying at \( J_{\text{crit}} = 1 \) in Table 3.1. Thus, the experimental objective value at the maximum, originally in arbitrary units, is normalized to 1.0 and will be used as a scale for other experimental data. The pulse shapes at the landscape maximum determined experimentally and by simulation also exhibited good agreement (not shown here). Among the four components of the control, \( u_x^I \), \( u_y^I \), \( u_x^S \) and \( u_y^S \) in this illustration, only \( u_y^I \) was altered to gain a non-constant shape during
optimization of the target observable $O = I_x$, while the other three components stayed at zero. This field is a particular solution on the Top CM, where other solutions will generally have all four field components at coordinated nonzero values. From simulation we further discovered that the shape of $u_y^I$ at the landscape maximum point, obtained by optimizing from zero fields, is dependent on the final time $T$.

Figure 3.1: (Color online) (a) The controls at landscape critical points within three different submanifolds, Top, Saddle A and Saddle B, corresponding to the final iteration in trajectory 1 and the initial iterations in 2 and 3 in Figure 3.2, respectively. The initial iterations at two saddle points were found by simulation and then refined in the laboratory, as explained in Section 3.4.2. The components $u_x^I$ and $u_y^I$ address spin $I$ ($^1$H), while $u_x^S$ and $u_y^S$ address spin $S$ ($^{13}$C). For the critical points in Top and Saddle A shown here, the $u_x^S$ and $u_y^S$ components are approximately zero. (b) The $^1$H-detected NMR spectra in the chemical shift range of 7.79-8.79 ppm, acquired with control pulses at the three critical points. The small kink in the middle of the two lines is ascribed to the proton signal in unlabeled $^{12}$CHCl$_3$, which is eliminated when integrating the peaks as the objective $J$.

### 3.4.2 Finding the landscape saddles in the laboratory

Searching for landscape saddle points in the laboratory can be a challenging task, since a gradient ascent/descent trajectory may approach, but likely not arrive at a saddle. Here we first performed a simulation to find a control whose corresponding dynamics reach a particular desired saddle. This saddle-seeking procedure was performed by specifying a proper final state $\rho(T)$ consistent with the desired saddle. At a saddle point of the landscape $J = \text{Tr}[\rho(T)O]$, the state $\rho(T)$ must satisfy two conditions that $[\rho(T), O] = 0$ and $\text{Tr}[\rho(T)O]$ is a suboptimal value with $-1 < J < 1$. Suppose the system density matrix $\rho(0)$
is diagonalized to form $\Lambda^\rho$ and the observable operator is diagonalized as $O = U_O \Lambda^O U_O^\dagger$, then $\rho(T) = \rho_{\text{target}}$ within any CM will have the general form $\rho_{\text{target}} = U_O (\Pi \Lambda^\rho \Pi^\dagger) U_O^\dagger$, with $\Pi$ being a permutation matrix associated with a desired CM (including the saddles) [3]. It is easy to identify a particular permutation matrix $\Pi$ for any of the saddles in Table 3.1. Thus, having designed a target density matrix $\rho_{\text{target}}$ for the desired saddle CM in this way, we optimized the new cost function in a simulation

$$J := \text{Tr}[U(T)\rho(0)U^\dagger(T)\rho_{\text{target}}].$$

(3.14)

A control producing the maximum value of $J$ will satisfy $U(T)\rho(0)U^\dagger(T) = \rho_{\text{target}}$, and the control must correspond to a saddle point of the original landscape $J$. In the cases where $\rho(0)$ or $O$ have degeneracies, the choice of $\rho_{\text{target}}$ for a particular CM can be non-unique. We have also developed a more general theoretical method for optimizing an arbitrary initial control to move toward a given saddle submanifold, which does not require specifying $\rho(T)$ in advance [28].

Due to the generally high quality of the NMR control experiments and the good understanding of the Hamiltonian, we were able to transfer a simulation-determined control at a saddle point to the laboratory. Such trial fields often called for some further adjustment discussed below, thereby forming an efficient procedure for closely approaching a saddle. The experimentally measured gradient at a near-saddle control should be close to zero, but when this was not the case we employed the following experimental scheme to refine the control and draw it closer to creating a saddle point on the landscape. The procedure is based on minimizing the norm squared of the gradient, $\|\nabla J\|^2$, in order to approach the saddle. Thus, consider the derivative with respect to the progress variable $s$ given by the chain rule,

$$\frac{d}{ds} \|\nabla J[\vec{x}(s)]\|^2 = \frac{d\vec{x}^\tau}{ds} \cdot \frac{\partial}{\partial \vec{x}} \|\nabla J[\vec{x}]\|^2,$$

(3.15)
where we have
\[ \frac{\partial}{\partial \vec{x}} \| \vec{\nabla} J[\vec{x}] \|^2 = 2 \mathcal{H}[\vec{x}] \cdot \vec{\nabla} J[\vec{x}]. \] (3.16)

By combining Eqs. (3.15) and (3.16) the magnitude of \( \| \vec{\nabla} J \|^2 \) can be minimized by moving along the direction specified by \( d\vec{x}/ds = -\alpha \mathcal{H} \vec{\nabla} J \) in control space, i.e.,
\[ \vec{x}(k + 1) = \vec{x}(k) - \alpha \mathcal{H}[\vec{x}(k)] \cdot \vec{\nabla} J[\vec{x}(k)], \] (3.17)

where \( \alpha > 0 \). When many iterations are required, remeasurement of the Hessian could be slow in the laboratory. In our experiments we found that the initial field determined by simulation was close enough to the target saddle point such that the refinement could be accomplished within only a few rounds of iteration.

With the methods above we successfully found particular control fields within the CMs Saddle A and Saddle B, respectively, whose pulse shapes and corresponding NMR signals are given in Figure 3.1. In the example shown for Saddle A the control only excited the target spin \( I = ^1H \) (other solutions in the same saddle CM will generally also involve the field components addressing \(^{13}C\)). For Saddle B in Figure 3.1 both spins were significantly excited by their respective resonant pulses, and polarization or coherence transfer was thus induced between them. The lineshapes of the peaks in NMR spectra should generally be combination of absorption and dispersion modes if a phase correction is not performed [14]. However, for all the landscape critical points shown in Figure 3.1(b) both lines of the doublet have the absorption mode only and no dispersion feature. This circumstance arose because the dispersion mode results from the spin magnetization orthogonal to the detector phase, which does not commute with the observable \( O \) [e.g., if \( O = I_x \) then the dispersion lineshape corresponds to the components of \( I_y \) and \( 2I_yS_z \) in the final state \( \rho(T) \)], and must be absent at any critical point according to the criterion in Eq. (3.7).

After finding the two near-saddle controls they were used as starting points for landscape gradient ascents. In Figure 3.2 the trajectories starting from near the critical points
in Saddle A and B are labeled 2 and 3, which should be compared with trajectory 1 that started from zero. With a constant step size for each trajectory, the optimization rate of the objective $J$ from a near-saddle point is much lower than from a non-critical point with a relatively large gradient. The neighborhoods of the saddle points form low-gradient regions about some particular suboptimal values of the observable landscape, which increase the search effort of a gradient-based trajectory running close to them. In the laboratory the determination of the gradient to sufficient accuracy, especially when the gradient is relatively small, possibly can be problematic for ready escape from the neighborhood suboptimal saddle region. Thus, the gradient algorithm should be exploited with care in optimal control experiments, especially for complex quantum systems with little *a priori* knowledge of the landscape topology, where encountering a low gradient norm does not necessarily indicate that the global optimum of the landscape is being approached. A host of other algorithms (e.g., conjugate gradient, stochastic, etc.) can be exploited if ascent alone is the goal, but a gradient algorithm (more generally the Rover suite of procedures) is necessary for experiments seeking to identify topological and other landscape features.

![Figure 3.2: (Color online) Three gradient ascent trajectories starting from different initial controls. Trajectory 1: Zero control fields. Trajectory 2: Saddle A at $J \approx 0$. Trajectory 3: Saddle B at $J \approx 0.25$. The iteration step size is constant for all the three trajectories. The trajectory 1 starting at a non-critical point rapidly ascends on the landscape and climbs to the top, while the trajectories 2 and 3 linger near their respective saddles. Furthermore, trajectories 2 and 3 are gradient ascent (instead of saddle-seeking) trajectories, so the final iterations are not necessarily better saddle controls than the initial ones. In fact, the late iterations of trajectory 2 are clearly moving farther away from the saddle.](image-url)
Another gradient ascent trajectory is shown in Figure 3.3, where the initial control was obtained by perturbing the control at Saddle B above. During the optimization process, the rate of increase for $J$ slowed down when it approached the neighborhood of the saddle at $J \simeq 0.25$ again, but passed through that region smoothly and accelerated afterward. The procedure is also characterized by the evolution of the gradient norm $\|\vec{\nabla} J\|$, also shown in Figure 3.3. This behavior is typical of a gradient optimization coming near a saddle point [6]. Together with the trajectories in Figure 3.2, the results demonstrate that the saddle regions are distributed like “islands” at their corresponding $J$ values of the landscape, since other controls producing the same $J$ value as a saddle could have even large associated gradients with trajectories exhibiting no saddle-like features.

![Figure 3.3](image)

Figure 3.3: (Color online) A gradient ascent trajectory passing nearby Saddle B at $J \simeq 0.25$. The initial control at the 0th iteration is chosen by perturbing the control at Saddle B, which is shown in Figure 3.1(a). The trajectory slows down when again coming close to the saddle, but then continues up the landscape. The norm of the gradient is also shown.

### 3.4.3 Local landscape topology near the saddles: the Hessian signatures

In order to assess the theoretical predictions on Hessian signatures in Table 3.1, we measured the Hessian matrices at the three critical points found above, which belong to Top, Saddle A and Saddle B, by the LS method utilizing 1,000 random samples around each critical point. The eigenspectra of the respective Hessians are shown in Figure 3.4.
Although none of the eigenvalues, which are expected to be zero, exactly have that value under the imperfect experimental conditions, some eigenvalues have dominantly large magnitudes compared with the others. In addition, as confirmed by simulation, for particular dynamical controls some of the nonzero Hessian eigenvalues may be much smaller than others. This circumstance can result in the latter eigenvalues being practically indistinguishable from the zero eigenvalues due to experimental error influencing the quality of the Hessian. We may conclude that the experimental results are in agreement with theoretical predictions, as the numbers of significantly positive and negative eigenvalues do not exceed their upper bounds $D_+$ and $D_-$ upon comparing Figure 3.4 with Table 3.1.

Figure 3.4: (Color online) Hessian eigenspectra at the three critical points in Figure 3.1. Each Hessian was estimated by least squares with 1,000 random samples taken in the neighborhood of the fields at their respective critical points. Theoretical analysis summarized in Table 3.1 shows that Top, Saddle A and B should have at most 0/8, 4/4 and 2/6 positive/negative Hessian eigenvalues, respectively. The predictions are affirmed within experimental error.

Among the three cases, the Hessian for Saddle B has the smallest norm, resulting in the boundary between “zero” and “nonzero” eigenvalues being least clear in its eigenspectrum. For this Hessian a statistical bootstrapping strategy [7] was used for estimating the error associated with the eigenvalues. We took a total of $\sim 2,600$ random samples about the control in Saddle B, and randomly chose 10 subsets, each with sizes of $M = 600, 800, 1000, 1200,$ and 1400. The Hessian was then calculated with the data from each subset using LS, and for each sampling size $M$ the eigenvalues were sorted and averaged (see Figure 3.5).
error bars show the standard deviations of each single eigenvalue at a particular level of \( M \). In the Hessian spectrum there are a group of small eigenvalues which are roughly symmetrically distributed about zero, and the spacings between adjacent ones are comparable with the random error. These “zero” eigenvalues become smaller in magnitude as the sampling size \( M \) increases from 600 to 1400, i.e., more data averaging is utilized in determining the Hessian spectra. By extrapolating the eigenvalues with the expected scaling of \( \sim \frac{1}{\sqrt{M}} \) to the limit \( M \to \infty \), we find that the group of “zero” eigenvalues converge into a narrow interval around zero and clearly separate from the nonzero eigenvalues. The data analysis supports the theoretical conclusion of the presence of an extensive Hessian null space at the critical points [3, 4].

![Hessian eigenvalues extrapolation](image)

Figure 3.5: (Color online) Estimation of the Hessian eigenvalues of the control at Saddle B and extrapolation using the expected scaling as \( \sim \frac{1}{\sqrt{M}} \), where \( M \) is the number of random samples taken for determining the Hessian. The extrapolated spectrum shows several clear positive and negative eigenvalues along with a set of null eigenvalues consistent with theoretical predictions. The error bars at a given \( M \) are estimated from the eigenspectra determined from 10 data subsets of size \( M \), which are taken from the totally \( \sim 2,600 \) samples.

The Hessian eigenvectors associated with the nonzero eigenvalues at a landscape critical point describe the independent paths for driving off the CM, and the magnitudes of the corresponding eigenvalues characterize the sensitivity of \( J \) to variation along these paths in control space. For a saddle CM, moving along the Hessian eigenvectors with positive (or negative) eigenvalues should locally lead to a quadratic increase (or decrease) of the ob-
jective value. We explored the neighborhoods of the controls at Saddle A and B by roving along each of their respective dominant Hessian eigenvectors (for Saddle A, those associated with the four most negative and four most positive eigenvalues; for Saddle B, those associated with six most negative and two most positive ones), determined with the largest sampling size $M$ without extrapolation. The trajectories are shown in Figure 3.6, where the roving distance is calculated as the Euclidean distance of the vector of control variables, $\vec{x}$, from the saddle starting point $\vec{x}_0$. Parabola-like curves are obtained as expected, which can be precisely fitted by second-order polynomials, indicating that the higher-order terms in the Taylor expansion of the landscape are negligible in the regions explored. Some trajectories may look asymmetric about the starting point, which appears mostly due to the small, but nonzero, residual gradient at the experimentally determined near-saddle points.

![Figure 3.6](image.png)

**Figure 3.6:** (Color online) Driving off the critical points at Saddle A (a) and Saddle B (b), shown in Figure 3.1(a), along their respective Hessian eigenvectors associated with the dominant eigenvalues. The eigenvector $v_i$ corresponds to the $i$-th eigenvalue in the Hessian spectrum (see Figure 3.4). The presence of parabolas opening both upward and downward, resulting from positive and negative Hessian eigenvalues respectively, affirms the saddle topology of these two critical points. The roving distance along each eigenvector is in reference to the starting field at the respective saddle points.
3.4.4 Control on the $^{13}$C landscape

The control landscape with the observable being $O = S_x$ should also have the topological properties given in Table 3.1. The main distinction is that the spin $S = ^{13}$C is of a lower gyromagnetic ratio, or less sensitive, than the spin $I = ^1$H; therefore, we expect that utilizing spin coupling for transfer of spin polarization from $^1$H to $^{13}$C should provide a valuable assist for reaching the global maximum of the $^{13}$C landscape. For the two-spin system at its initial state $\rho(0) = I_z + rS_z$, if we only irradiate spin $S$ with its resonant rf pulse and leave spin $I$ unexcited, the component $I_z$ in the density matrix will be invariant, and thus the maximal reachable objective value for $O = S_x$ is $J = r$. This local maximum is caused by insufficient system controllability [23, 24], and corresponds to Saddle B in Table 3.1 kinematically upon removing the constraints on the control, as demonstrated by the experimental results in Figure 3.7 and explained below.

With the field components $u^I_x$ and $u^I_y$ fixed to zero we optimized $u^S_x$ and $u^S_y$ with the final fields shown in Figure 3.7(a). Then we measured the Hessian at the suboptimal critical point with all four components of the control released. To make up for the relatively lower S/N of the $^{13}$C nucleus, we took a greater number of random samples ($\sim 6,000$) for the LS Hessian estimation. In the Hessian spectrum six negative and two positive eigenvalues are clearly discernible in Figure 3.7(b) from the other approximately zero eigenvalues, consistent with the predicted Hessian signature of Saddle B. The eigenvectors of the two almost degenerate positive eigenvalues are found to be associated with $u^I_x$ and $u^I_y$ (see inset in Figure 3.7(b)), which implies that in order to overcome the critical value and further ascend the landscape by perturbing the control at this saddle point, the control field on spin $I$ must be turned on. We also note that the two positive Hessian eigenvalues at this type of saddle can vanish in some special cases due to singularity, i.e., violation of the assumption (ii) in Section 3.2, thus turning the saddle to a dynamical second-order trap. Nevertheless, this circumstance does not invalidate the landscape analysis derived in the kinematic picture.
(see [29, 30] for the discussion on similar examples). A more detailed assessment of this issue will be provided in an additional study [31], where the suboptimal maximum deduced from the Hessian is shown to actually remain a saddle when further analyzed considering higher-order variation of $J$ with respect to the controls.

We also addressed a global maximum point of the $^{13}$C landscape (Top in Table 3.1). The $J$ value at the maximum is enhanced from that at Saddle B by a factor of $\sim 4.1$, which is close to $\gamma^{(1)H}/\gamma^{(13)C}$ as in conventional polarization transfer experiments [12]. Figure 3.7(a) shows that at the landscape maximum both spins are manipulated by their resonant pulses, and the Hessian has eight significantly negative eigenvalues in Figure 3.7(b). The results presented here demonstrate the effectiveness, and even the necessity, of exploiting polarization transfer for signal enhancement of a low gyromagnetic ratio nucleus from the perspective of attaining optimal control performance.

![Figure 3.7: (Color online) Two critical points over the landscape of the observable $O = S_x$ belonging to Top and Saddle B in Table 3.1. The objective $J$ at Saddle B, determined from the peak integral in the $^{13}$C NMR spectrum, is $\sim 0.24$ when the $J$ value of Top is normalized to 1.0. (a) The control fields at the two critical points. Note that for Saddle B, the field addressing spin $I$ is zero. (b) The Hessian eigenspectra at the two critical points. The solid and dashed lines in the inset give the eigenvectors of the 15th and 16th (almost degenerate) eigenvalues at Saddle B, respectively, which turn on the field components for spin $I$ in order to enable climbing away from the saddle.](image-url)
3.5 Conclusions

This work reported systematic experimental observation of saddle points on a quantum control landscape with liquid-state NMR as the testbed by exploiting the Rover algorithms in a coupled heteronuclear two-spin system, i.e., $^1$H and $^{13}$C in a sample of $^{13}$CHCl$_3$. We first summarized the theoretical analysis of the landscape topology, and successfully found the predicted critical points of specified local topological nature in the experiments, especially for the two saddle points. Coming near a suboptimal saddle point manifold on gradient optimization of the control objective may slow down an ascent but should not stop the landscape climb, given that the gradient measurement is of sufficient accuracy. Similar behavior should be generally observed on any control landscape containing saddles. The Hessian eigenspectra at different critical points were measured, and they agreed with the theoretical predictions within the experimental noise level. The neighborhoods of the saddles were also explored to show the special local landscape structures in those regions. Numerical simulations assisted the experimental studies, especially in providing initial controls near the landscape saddle points. The findings of this particular landscape topological assessment concern two coupled spins under control, but the implications extend beyond for the control of other quantum mechanical phenomena with electromagnetic fields, as the landscape principles are generic.
Bibliography


[27] Gradient algorithm applied to laboratory quantum control, J. Roslund and H. Rabitz, 

[28] Measuring the distance from saddle points and driving to locate them over quantum 
   Theor. 48, 465305 (2015)

   Rev. Lett. 106, 120402 (2011); Comment on “are there traps in quantum control 

[30] Searching for quantum optimal control fields in the presence of singular critical 
   points, G. Riviello, C. Brif, R. Long, R.-B. Wu, K. Moore Tibbetts, T.-S. Ho, and H. 

[31] Identifying and avoiding singularity-induced local traps over control landscapes of 
   spin chain systems, Q. Sun, I. Pelczer, R.-B. Wu, G. Riviello, and H. Rabitz, 17, 17, 
   29714-29722 (2015)
Chapter 4

Identifying and avoiding

singularity-induced local traps over

control landscapes of spin chain systems

The wide success of quantum optimal control in experiments and simulations is attributed to the properties of the control landscape, defined by the objective value as a functional of the controls. Prior analysis has shown that on satisfaction of some underlying assumptions, the landscapes are free of suboptimal traps that could halt the search for a global optimum with gradient-based algorithms. However, violation of one particular assumption can give rise to a so-called singular control, possibly bringing about local traps on the corresponding landscapes in some particular situations. This chapter theoretically and experimentally demonstrates the existence of singular traps on the landscape in linear spin-1/2 chains with Ising couplings between nearest neighbors and with certain field components set to zero. The results in a two-spin example show how a trap influences the search trajectories passing by it, and how to avoid encountering such traps in practice by choosing sufficiently strong initial control fields. The findings are also discussed in the context of the generally observed success of quantum control.
4.1 Introduction

Quantum dynamics phenomena at the atomic and molecular scales can be controlled via the application of tailored electromagnetic fields [1]. The control objective, e.g., the expectation value of an observable, can be optimized by shaping the control fields. A control landscape is defined by the objective value $J$ as a functional of the controls, whose topology dictates the ease of finding the globally optimal solutions, especially when using gradient-based search algorithms. What lies at the core of control landscape analysis is the local topology around its suboptimal critical (or stationary) points, which is assessed primarily by a second derivative test, i.e., the definiteness of the Hessian matrix evaluated at a critical point. In this chapter, a suboptimal critical point with a nonzero and negative semidefinite Hessian is referred to as a trap for maximization of $J$, although it is not necessarily a local maximum when the Hessian has a nonempty null space (or kernel). More rigorously, such critical points are traps of at least second order, and behave similarly to local maxima for optimization algorithms utilizing first- and second-order derivatives only [5]. Based on the assumptions that (i) the closed quantum system is controllable [2, 3, 4], (ii) the control to final state map is surjective everywhere over the landscape [5, 6, 7], and (iii) the control fields are unconstrained, theoretical analyses reveal that the landscape is free of any traps [8, 9, 10], i.e., a gradient search will converge to a globally optimal solution from an arbitrary initial choice in the space of control fields. The trap-free landscape conclusion is a fortunate fiducial outcome of fully satisfying the three assumptions stated above, which provides a basis to understand the observed generally successful outcomes of quantum optimal control in both experiments [1] and simulations [11]. In particular circumstances the assumptions may be only partially satisfied, which can lead to additional landscape features appearing, including traps. In practice, satisfaction of assumption (iii) to an adequate degree is of concern and has been the focus of recent studies [12, 13, 14]. Notwithstanding the overall growing numbers of positive control experiments, each of the
assumptions deserve additional analysis to assess their ease of practical satisfaction.

This chapter focuses on so-called singular controls, which locally violate the assumption (ii) above and may produce suboptimal traps on the landscape. This prospect was considered earlier [15], and a few special examples of local traps based on theoretical model systems have been reported [?, ?], which can be ascribed to singularity. A general basis to assess the role of singular controls has been established in Ref. [6]. Carefully performed numerical simulations suggest that the singular controls are very rare, and most of them actually do not form traps for gradient searches [6, 7]. The singular traps known so far are associated with zero (or constant) fields [5, 16], and have very small attractive volumes, i.e., a search is likely to be halted in practice only if the initial control is chosen sufficiently close to the exact singular trap [7, 17]. This circumstance also applies to the findings here, but the possible presence of such traps and the means for their avoidance is a matter of considerable interest.

The control of dynamics in coupled spin-1/2 chains is of both theoretical and practical interest [18], because of its proposed utility in various applications including for efficient transfer of information in quantum computing devices [19], and in signal enhancement by polarization transfer in nuclear magnetic resonance (NMR) spectroscopy [20]. Analysis of the underlying control landscape structures could be valuable in these situations. Recently we developed experimental methodology for studying the control landscape of spin systems with NMR [21, 22]. This chapter focuses on the singular traps in control landscapes arising in linear spin chains and the associated influence of non-local spin couplings. Unlike many previous studies [5, 6, 7, 8, 9, 16] in which only a single control field was considered, the complete controllability of a multispin system generally calls for the introduction of multiple control fields when each spin is spectrally distinct from the others. In this situation, singularity may exist at special controls where some of the field components have zero value, which in turn may convert a regular landscape saddle point with an indefinite Hessian to a second-order local trap.
The remainder of the chapter is organized as follows. Section 4.2 provides the theoretical background for control landscapes and singular controls. Section 4.3 analyzes a particular landscape problem defined in linear spin chain systems and specifies the conditions for a singular control to become a local trap. Illustrative examples for a two-spin system are given in Section 4.4 with experimental and simulational results, confirming the existence of the singular traps and showing how to avoid them in practice by appropriate choices of initial controls. While the analysis here mainly focuses on spin chains with nearest neighbor coupling, treatment of a three-spin system with next nearest neighbor coupling demonstrates a reduction in singular behavior. Finally, Section 4.5 gives general conclusions and also discusses the overall special circumstances of zero (or constant) field singular controls, in the backdrop of the general success of quantum control experiments and the trap-free findings in large numbers of carefully performed simulations [11, 23].

4.2 Control landscape and singular controls

Consider a closed $N$-level quantum system under control, whose dynamics is described by the time-dependent Schrödinger equation (in units where $\hbar = 1$),

$$i \frac{d}{dt} U(t) = \left[H_0 + \sum_{m=1}^{M} u_m(t) H_m\right] U(t), \quad U(0) = \mathbb{1}_N, \quad (4.1)$$

where $H_0$ is the internal Hamiltonian, $H_1, \cdots, H_M$ are the linearly independent interaction Hamiltonian terms, and $u_1(t), \cdots, u_M(t)$ are their corresponding control fields, defined on the time interval $t \in [0, T]$. In the following we will encapsulate all of the $M$ control fields into an entire dynamic control $u(\cdot)$ when necessary, and each $u_m$ will be called a component of it. The circumstance in Eq. (4.1) arises in the present work with a set of weakly coupled spins where each spin is spectrally distinguishable from the others and thus may be individually addressed. A necessary and sufficient condition for complete controllability
of the system is that the Lie algebra generated by \(i H_0\) and \(i H_m, m = 1, \ldots, M\), is \(u(N)\) (or \(su(N)\) for a traceless Hamiltonian) [2, 3]. The density matrix \(\rho\), denoting the system state, evolves with time as \(\rho(t) = U(t)\rho(0)U^\dagger(t)\), or equivalently obeys the Liouville-von Neumann equation

\[
\frac{d}{dt} \rho(t) = -i \left[ H_0 + \sum_{m=1}^{M} u_m(t) H_m, \rho(t) \right].
\]  

(4.2)

A widely studied target of quantum control is to optimize the expectation value of a Hermitian observable operator \(O\) at the final time \(T\) [9, 10, 24]. The objective \(J\) specifies the landscape as a functional of the control \(u(\cdot)\) through \(U(T)\), expressed by

\[
J = \text{Tr}[\rho(T)O] = \text{Tr}[U(T)\rho(0)U^\dagger(T)O].
\]  

(4.3)

In order to simplify the analysis of the dynamic landscape (i.e., \(J\) as a functional of the control fields), the kinematic picture of the landscape is introduced, which is only concerned with the system state \(\rho(T)\) at the final time \(T\), but not the dynamics leading to the final state. Thus, in the kinematic picture the landscape \(J\) is simply treated as a function of \(\rho(T)\) defined on the space of matrices unitarily equivalent to \(\rho(0)\), or a function of \(U(T)\) defined on the unitary group \(U(N)\) [6, 10, 24]. Landscape analysis primarily aims at identifying the critical points where the first derivative of \(J\) is zero, which can slow down or even halt the gradient-based optimization searches. The first derivative of \(J\) in the dynamic and kinematic landscapes are related through the chain rule

\[
\frac{\delta J}{\delta u(\cdot)} = \left\langle \nabla J[\rho(T)], \frac{\delta \rho(T)}{\delta u(\cdot)} \right\rangle_{\rho(T)},
\]  

(4.4)

where \(\langle \cdot, \cdot \rangle\) represents the Riemannian metric, \(\nabla J[\rho(T)]\) is the kinematic gradient of \(J\), and \(\frac{\delta \rho(T)}{\delta u(\cdot)}\), which can be interpreted as a Jacobian matrix, is the state variation at \(T\) resulted from a variation \(\delta u(\cdot)\) in the control space [6]. When \(\frac{\delta \rho(T)}{\delta u(\cdot)}\) is full rank (i.e., the set \(\frac{\delta \rho(T)}{\delta u(\cdot)}\), \(\forall j \text{ and } k\), contains the maximal allowed number of linearly independent functions), we
have $\frac{\delta J}{\delta u(\cdot)} = 0$ if and only if $\nabla J[\rho(T)] = 0$. A criterion for a particular $\rho(T)$ to be such a kinematic critical point (KCP), is that [9]

$$[\rho(T), O] = 0. \quad (4.5)$$

Further analysis reveals that the topology of the kinematic landscape is determined solely by the initial state $\rho(0)$ and target observable $O$, and KCPs occur at specific discrete values of $J$, which can be computed from the eigenvalues of $\rho(0)$ and $O$ [9]. The local landscape topology about a critical point is assessed by the definiteness of its Hessian. By analyzing the kinematic Hessian form, i.e., the second derivative of $J$ with respect to $\rho(T)$, it can be shown that all the suboptimal KCPs located between the global maximum and minimum of the landscape, if any, must have both positive and negative Hessian eigenvalues, and thus be saddle points [10]. Therefore, the kinematic landscape does not contain any local traps for optimization searches.

The viability of lifting the landscape analysis results in the kinematic picture to the dynamic picture can be guaranteed if the Jacobian $\frac{\delta \rho(T)}{\delta u(\cdot)}$ is full rank [24], and such controls and critical points are called regular. In contrast, a dynamic control $u(\cdot)$ is referred to as singular if the Jacobian is rank deficient [6], in which case the final state $\rho(T)$ cannot be freely manipulated differentially by perturbing the control fields. The rank deficiency of $\frac{\delta \rho(T)}{\delta u(\cdot)}$ is called the corank of a singular control, characterizing the degree of its singularity [6]. Singularity can lead to changes of local landscape topology from at least two perspectives. First, it can induce nonkinematic critical points (NKCPs) on the dynamic landscape, which do not satisfy the condition in Eq. (4.5) and therefore have no counterparts in the kinematic picture. Although these singular NKCPs are dominantly saddles [6, 7], special examples are known in which the singular NKCPs have negative definite or semidefinite Hessians [16]. Second, singularity can qualitatively alter the Hessian feature at a KCP,
where the second derivative of $J$ in the two pictures is related by [6, 24]

$$\frac{\delta^2 J}{\delta u^2(\cdot)} = \left\langle \frac{\delta \rho(T)}{\delta u(\cdot)}, Q_{\rho(T)} \frac{\delta \rho(T)}{\delta u(\cdot)} \right\rangle_{\rho(T)},$$

(4.6)

where $Q_{\rho(T)}$ is the kinematic Hessian of $J$ at $\rho(T)$. The nonzero eigenvalues of $Q_{\rho(T)}$ are mapped one-to-one to the dynamic picture if $\frac{\delta \rho(T)}{\delta u(\cdot)}$ is full rank, but a dynamic critical point may lose some nonzero Hessian eigenvalues of its kinematic counterpart when it is singular. In some extreme cases, a kinematic saddle point can even be converted to a dynamic second-order maximum if all the positive Hessian eigenvalues vanish due to singularity [5, 17]. The situation analyzed in the following section will also fall into this category. To our knowledge, some of the singularity-induced traps discovered to date are true local maxima with strictly negative definite Hessians [16], while the others recover the topology of saddles when higher-order derivatives are further considered [5, 16, 17], including the case studied in this chapter. The local landscape topology at singular controls is related to the nature of the controlled quantum system in a complex and subtle manner.

### 4.3 Singular traps in linear spin chain systems

#### 4.3.1 Control system model and landscape gradient search

This chapter mainly addresses the model system of a linear chain consisting of $n$ spins-1/2 with Ising-type coupling, also known as weak coupling in NMR [18], only between nearest neighbors; the consequences of additional spin coupling complexity are discussed in Sections 4.3.3 and 4.5. We assume that the Larmor frequencies of the spins are well separated, so that each spin can be individually addressed by its on-resonance control field. A multiple rotating frame [25] simultaneously operating at the resonance frequencies of each spin is introduced to express the total Hamiltonian of the controlled system as (in
units where $\hbar = 1$)

$$H(t) = \sum_{j=1}^{n-1} J_{j,j+1} I_j^z I_{j+1}^z + \sum_{j=1}^{n} [u_x^j(t) I_x^j + u_y^j(t) I_y^j], \quad (4.7)$$

where $J_{j,j+1}$ is the coupling strength between the spins $j$ and $j+1$ in units of frequency. The multispin operator $I_a^j (a = x, y, z)$ denotes the tensor product of the single-spin operator $\sigma_a$ for the $j$th spin with identity operators for all the others [26, 27]:

$$I_a^j = \underbrace{\mathbb{1}_2 \otimes \cdots \otimes \mathbb{1}_2}_{j-1} \otimes \sigma_a \otimes \underbrace{\mathbb{1}_2 \otimes \cdots \otimes \mathbb{1}_2}_{n-j}, \quad (4.8)$$

where $\sigma_a (a = x, y, z)$ is a Pauli matrix and $\mathbb{1}_2$ is the $2 \times 2$ identity matrix. The products of the operators $I_a^j$ with different spin labels $j$, such as $I_x^1 I_y^2$, form an orthogonal basis for traceless $2^n$-dimensional Hermitian matrices [26]. The control portion of the Hamiltonian includes $2n$ linearly independent terms associated with the operators $I_x^j$ and $I_y^j (j = 1, \cdots, n)$, each addressed by a corresponding field $u_x^j(t)$ or $u_y^j(t)$ in units of frequency. The quantum system is completely controllable with the Hamiltonian form of Eq. (4.7) [2, 3].

Consider a control landscape problem in this system specified by

$$\rho(0) = \sum_{j=1}^{n} c_j I_z^j \quad (c_j > 0), \quad O = I_x^k, \; k \in \{1, \cdots, n\}. \quad (4.9)$$

Here, $\rho(0)$ is the traceless portion of the initial density matrix, which well approximates the thermal equilibrium state of $n$ spectrally distinguishable spins with gyromagnetic ratios proportional to the $c_j$’s, in conventional NMR experiments when the system is placed in a large static magnetic field along the $z$ direction [28]. The observable $O$ targets the in-phase coherence in the $x$ direction of the $k$th spin [26]. We will refer to the $k$th spin as the target spin, and all the others in the chain as spectator spins when necessary. A complete
list of KCPs over this landscape will not be given here; we only point out that there exist critical points at $J = \pm 2^{n-2}c_j, \forall j = 1, \cdots, n$ ($2^{n-2}$ is a coefficient arising from the number of spins $n$ in the system), most of which [29] have to be suboptimal saddle points in the kinematic picture [9, 10]. The landscape settings described here can be realized by heteronuclear NMR experiments in a straightforward manner; a two-spin example will be presented in Section 4.4.

The landscape analysis of the problem given by Eqs. (4.7) and (4.9) will start with gradient-based optimization of each single field component at some special dynamic controls. Invoking Eqs. (4.1) and (4.2), the gradient of the objective $J = \Tr[\rho(T)O]$ with respect to a single field $u_m(t)$ is given by [9, 30]

$$\frac{\delta J}{\delta u_m(t)} = -i \Tr \left\{ [\rho(t), U(t)O(T)U^\dagger(t)]H_m \right\}, \tag{4.10}$$

where $O(T) := U^\dagger(T)OU(T)$. For the observable $O = I_x^k$ and for a dynamic control with all the field components set to zero except $u_y^k(t)$, the gradient with respect to each of the $(2n - 1)$ zero fields will still be zero, as demonstrated below. In this special case, the total Hamiltonian reduces to

$$H(t) = \sum_{j=1}^{n-1} J_{j,j+1}I_z^jI_z^{j+1} + u_y^k(t)I_y^k. \tag{4.11}$$

Thus, the terms $\rho(t)$ and $U(t)O(T)U^\dagger(t)$ in the gradient formula of Eq. (4.10) can only evolve in some Hermitian subspaces. The time evolution of $\rho(t)$ from $\rho(0) = \sum_{j=1}^{n} c_j I_z^j$, obeying Eq. (4.2) with $H(t)$ given in Eq. (4.11), is constrained within the subspace spanned by the following product operators,

$$\rho(t) \in \text{span}\{I_z^j(j = 1, \cdots, n), I_x^k, I_y^kI_z^{k\pm 1}, I_z^{k-1}I_x^kI_z^{k+1}, I_z^{k-1}I_xI_z^kI_z^{k+1}\},$$
which is determined by calculating the commutator of $\rho(t)$ with $H(t)$, starting from $\rho(0)$, until no more linearly independent terms can be added in. Since $U(t)O(T)U^\dagger(t)$ is the unitary evolution of $O = I^k_x$ governed by the Hamiltonian of Eq. (4.11), we can also derive the corresponding subspace as

$$U(t)O(T)U^\dagger(t) \in \text{span}\{I^k_x, I^k_y, I^k_z, I^k_{x+1}, I^k_{z+1}, I^k_{x-1}, I^k_{z-1}\}.$$ 

Among the $2n$ distinct $H_m$ terms of Eq. (4.10), $H_m \in \{I^j_x, I^j_y\}_{j=1}^n$, only $I^k_y$ is present in the domain of $[\rho(t), U(t)O(T)U^\dagger(t)]$, so the landscape gradients with respect to control field components set a priori to zero [other than $u^k_y(t)$] all vanish. Thus, we have

**Remark 1.** When starting from an initial control with $u^k_y(t)$ being the only nonzero field, the other $(2n - 1)$ zero components will permanently stay at zero throughout an optimization process utilizing a simple gradient ascent algorithm [see Eq. (4.20) in Section 4.4]. Likewise, if $u^k_x(t)$ and $u^k_y(t)$ are the only nonzero components in the initial control, the remaining $(2n - 2)$ fields will stay at zero during a gradient ascent, which can be demonstrated in the same manner.

As such, although we have sufficient control resources, we will never get the chance to utilize all of them for some special initial control guesses. The consequence is that the special search trajectories may converge to suboptimal critical points, instead of the global optimum. When the control fields addressing the $(n - 1)$ spectator spins are all zero, the terms $I^j_z$ ($j \neq k$) in the initial state $\rho(0)$ will stay invariant throughout the control duration since they commute with the total Hamiltonian, and have no contribution to the observable $O = I^k_x$. In that case, the objective function $J = \text{Tr}[\rho(T)O]$ reduces to

$$J[U(T)] = c_k \text{Tr}[U(T)I^k_zU^\dagger(T)I^k_x].$$

The range of this function is $2^{n-2}[-c_k, c_k]$. The values $J = \pm 2^{n-2}c_k$ correspond to KCPs of
the original landscape, but not the global maximum or minimum in most cases. Therefore,

**Remark 2.** Gradient ascent from an initial control with field components on all spectator spins being zero will converge to a KCP at $J = 2^{n-2}c_k$, which is usually suboptimal.

In Section 4.3.2 we will determine whether such an end point is a saddle or a second-order trap in the dynamic picture by analyzing its Hessian matrix.

### 4.3.2 Hessian analysis at the singular critical point

For the landscape involving $M (= 2n$ in our spin chain system) independent control fields, the entire Hessian $\mathcal{H}(t, t')$ is specified as a real symmetric matrix composed of $M \times M$ submatrices $\mathcal{H}_{mm'}(t, t')$, with each of infinite dimension over the time domain $[0, T]$. An off-diagonal submatrix $\mathcal{H}_{mm'}$ with $m \neq m'$ is generally nonsymmetric by itself, but is the transpose of its opposite submatrix $\mathcal{H}_{m'm}$. In each submatrix, the triangular portion of the domain $0 \leq t' < t \leq T$ is given by [9, 31]

$$\mathcal{H}_{mm'}(t, t') := \frac{\delta^2 J}{\delta u_m(t) \delta u_{m'}(t')} = -\text{Tr}\{[[O(T), H_m(t)], H_{m'}(t')]\rho(0)\}$$

(4.13) for $m, m' = 1, \cdots, M$, and the opposite triangular portion can be determined by symmetry, $\mathcal{H}_{mm'}(t', t) = \mathcal{H}_{m'm}(t, t')$. The operator $H_m(t)$ is defined with the system dynamics in the Heisenberg picture, i.e., $H_m(t) := U^\dagger(t) H_m U(t)$. For our specific problem we conclude:

**Remark 3.** On the landscape $J = \text{Tr}[O(T)\rho(0)]$ with the Hamiltonian given in Eq. (4.7) and $\rho(0) = \sum_{j=1}^n c_j I_j^z$, $O = I_k^z$, a dynamic control must have a negative semidefinite Hessian if it satisfies two conditions: (i) all the field components except $u_k^y(t)$ are zero, and (ii) $O(T) = I_k^z$. Condition (ii) implies that such a control corresponds to a KCP lying at $J = 2^{n-2}c_k$.

**Proof.** The analysis below is carried out via a semi-explicit expression of each Hessian submatrix at a control satisfying the two conditions. With the Hamiltonian of Eq. (4.11) involving a nonzero control $u_k^x(t)$ only, the time evolution of the operator $I_k^x(t) = \cdots$
$U^\dagger(t)I_y^kU(t)$, for example, can be expressed as

$$I_y^k(t) = a_1 \cdot I_y^k + a_2 \cdot 2I_z^{k-1}I_y^k + a_3 \cdot 2I_z^kI_y^{k+1} + a_4 \cdot 4I_z^{k-1}I_y^kI_z^{k+1} + a_5 \cdot 2I_z^{k-1}I_y^k + a_6 \cdot 2I_z^kI_y^{k+1},$$  

(4.14)

where $a_1, \ldots, a_6$ are all real scalar functions of $t$. With the condition $O(T) = I_y^k$, we calculate the diagonal Hessian submatrix $H_{mm'}$ with $H_m = H_{m'} = I_y^k$ by Eq. (4.13):

$$[O(T), I_y^k(t)] = i[-a_1(t) \cdot I_x^k + a_2(t) \cdot 2I_z^{k-1}I_y^k + a_3(t) \cdot 2I_z^kI_y^{k+1} - a_4(t) \cdot 4I_z^{k-1}I_y^kI_z^{k+1}],$$  

(4.15)

$$\frac{\delta^2 J}{\delta u_y^k(t) \delta u_y^k(t')} = -2^{n-2}c_k[a_1(t)a_1(t') + a_2(t)a_2(t') + a_3(t)a_3(t') + a_4(t)a_4(t')], \quad t, t' \in [0, T].$$  

(4.16)

Hence, the subspace of the field component $u_y^k(t)$ gives four negative Hessian eigenvalues, whose corresponding eigenvectors are specified by linear combinations of the functions $a_1(t)$, $a_2(t)$, $a_3(t)$ and $a_4(t)$. If we perturb the field $u_y^k(t)$ and keep the other $(2n - 1)$ components invariant at zero value at the critical point, the $J$ value will drop quadratically if and only if the variation $\delta u_y^k(t)$ has a nonzero projection on the linear space spanned by the four functions above. Similarly, the field $u_z^k(t)$ also has four negative Hessian eigenvalues associated with it.

Regarding the two nearest neighbors of the target spin $k$, the time evolution of $I_x^k(t)$ under Eq. (4.11) can be formulated as well, e.g.,

$$I_x^{k-1}(t) = b_1 \cdot I_x^{k-1} + b_2 \cdot 2I_x^{k-2}I_y^{k-1} + b_3 \cdot 2I_z^{k-2}I_y^{k-1} + b_4 \cdot 4I_z^{k-2}I_x^{k-1}I_y^{k-1} + b_5 \cdot 2I_z^{k-1}I_y^{k-1}$$

$$+ b_6 \cdot 4I_z^{k-2}I_x^{k-1}I_y^{k-1} + b_7 \cdot 4I_z^{k-1}I_y^{k-1}I_z^{k-1} + b_8 \cdot 8I_z^{k-2}I_x^{k-1}I_y^{k-1}I_z^{k-1},$$  

(4.17)

where $b_1, \ldots, b_8$ are functions of $t$. Likewise,

$$\frac{\delta^2 J}{\delta u_x^{k-1}(t) \delta u_x^{k-1}(t')} = -2^{n-2}c_k[b_5(t)b_5(t') + b_6(t)b_6(t') + b_7(t)b_7(t') + b_8(t)b_8(t')], \quad t, t' \in [0, T].$$  

(4.18)
We find that the control subspace of the field \( u_{k-1}^k(t) \) also contains four Hessian eigenvectors with negative eigenvalues, given by linear combinations of \( b_5(t) \), \( b_6(t) \), \( b_7(t) \) and \( b_8(t) \). Similar situations occur for the field components \( u_{k-1}^k(t) \), \( u_{k+1}^k(t) \) and \( u_{y}^{k+1}(t) \), while the fields addressing spins other than \( k \) and \( k \pm 1 \) have zero contribution to the entire Hessian. It is also confirmed that the off-diagonal Hessian submatrices are all zero under conditions (i) and (ii), i.e., \( \delta^2 J / \delta u_m(t) \delta u_{m'}(t') \equiv 0 \) if \( u_m \) and \( u_{m'} \) are two different components. Therefore, the entire Hessian is semidefinite and has totally at most 24 negative eigenvalues, whose eigenvectors are associated with the control fields addressing the six operators \( I_{k,x/y}^k \) and \( I_{k,x/y}^{k+1} \). □

Unless \( J = 2^{n-2} c_k \) is the global maximum of the landscape, the results here demonstrate the existence of a suboptimal trap of at least second order. If we keep \( O(T) = I_z^k \) and alter the circumstances in either of the two ways, (i) making the component \( u_{k}^k(t) \) also nonzero, or (ii) adding a new coupling term \( J_{k-1,k+1} I_z^{k-1} I_z^{k+1} \) into the Hamiltonian, which couples the two nearest neighbors of the target spin, the simple Hessian structure as in Eqs. (4.16) and (4.18) will become more complex, and no longer be necessarily negative semidefinite. In those situations the existence of second-order traps cannot be guaranteed.

If the chain is not sufficiently long at least on one side of the target spin \( k \), when any of the four spins labeled \( (k \pm 1) \) or \( (k \pm 2) \) do not exist, the trap control may have less than 24 negative Hessian eigenvalues. For instance, if the chain contains a spin \( (k + 1) \) but not \( (k - 1) \), the terms of \( I_y^k(t) \) in Eq. (4.14) with coefficients \( a_2 \), \( a_4 \) and \( a_5 \) will disappear, and the subspace of \( u_{y}^k(t) \) will give only two (instead of four) negative eigenvalues whose eigenvectors are linear combinations of \( a_1(t) \) and \( a_3(t) \). The numbers of negative Hessian eigenvalues, \( D_- \), for different spin chain structures are counted with the method above and summarized in Table 4.1. Since only the five spins from \( (k - 2) \) to \( (k + 2) \) participate in the Hessian calculation, the presence of more spins in the chain will have no influence on the Hessian structure of our control problem targeting spin \( k \).

The emergence of the second-order trap on the landscape, as expected, is closely related
Table 4.1: Dependence of the number of negative Hessian eigenvalues at the local second-order trap, $D_-$, on the number of spins in the chain on both sides of the spin $k$ targeted by the observable $O = I_x^k$. The target spin is represented by • and the other spectator spins by ◦. Furthermore, · · · means that adding any more spins on that side does not change the Hessian index $D_-$.

<table>
<thead>
<tr>
<th>Spin chain structure</th>
<th>$D_-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>•</td>
<td>2</td>
</tr>
<tr>
<td>•-◦</td>
<td>6</td>
</tr>
<tr>
<td>•-◦-◦-◦ · · ·</td>
<td>8</td>
</tr>
<tr>
<td>◦-•-◦</td>
<td>16</td>
</tr>
<tr>
<td>◦-•-◦-◦ · · ·</td>
<td>20</td>
</tr>
<tr>
<td>· · ·◦-◦-◦-◦ · · ·</td>
<td>24</td>
</tr>
</tbody>
</table>

to singularity, or the rank deficiency of $\frac{\delta \rho(T)}{\delta u_m(t)}$, which can be calculated by

$$\frac{\delta \rho(T)}{\delta u_m(t)} = -iU(T)[H_m(t), \rho(0)]U^\dagger(T), \quad H_m \in \{I_x, I_y\}_{j=1}^n.$$  \hspace{1cm} (4.19)

Numerical simulations in two- and three-spin systems show that any dynamic control with only one of the $2n$ field components turned on should be singular (see Sections 4.3.3 and 4.4 for details). Although singular controls of this type can bring about new topological features, i.e., suboptimal second-order traps, to the dynamic landscape when overlapping with the kinematic saddle points, their significance in practice calls for special consideration, as starting with zero field components is not physically reasonable in a coupled spin system. The latter issue is examined in Section 4.4 and in ref. [7] for other like situations, as the domain of “attraction” around a trap is of primary importance.

4.3.3 The influence of more complex coupling in a three-spin case

The simple nearest neighbor coupling structure of the linear chain model considered here is crucial for the existence of the singular traps, as shown by the following example in the
three-spin case. Consider a control landscape with

\[ \rho(0) = I_1^1 + 0.8I_2^2 + 0.6I_3^3, \quad O = I_2^2, \]

defined in a linear chain model where spin 2 is coupled with its two (uncoupled) neighbors, spins 1 and 3. The condition (ii) above, \( O(T) = I_2^2 \), identifies a KCP lying at \( J = 1.6 \) compared to the global maximum \( J = 2.4 \) [9], whose kinematic Hessian form should have 6 positive and 26 negative eigenvalues [10]. On further satisfaction of condition (i), the corresponding singular dynamic critical point will have a semidefinite Hessian with 16 negative eigenvalues, and thus form a second-order trap. However, with the appearance of an additional next nearest neighbor coupling (i.e., triangular coupling) term \( J_{1,3}I_1^1I_3^3 \) in the Hamiltonian, a control satisfying the two conditions will still be critical but have an indefinite Hessian with 4 positive and 20 negative eigenvalues, being a saddle on the dynamic landscape instead of a second-order maximum. We see that the more complex coupling gives a lower degree of singularity, or Jacobian corank (cf. Section 4.2) at the same special dynamic control, which consequently influences the local landscape topology. In this example the Hessians are numerically computed with Eq. (4.13), with the coupling constants \( J_{1,2}, J_{2,3} \) and \( J_{1,3} \) set to be unequal. Section 4.5 contains further discussion on the role of coupling structure upon landscape topology.

### 4.4 Illustrations

In this section we demonstrate the existence of singular suboptimal traps in a two-spin systems as illustrations. We present experimental results and numerical simulations for a heteronuclear two-spin sample (\(^{13}\text{CHCl}_3\)), revealing how the choice of initial controls influences the likelihood of trapping in practice.

The gradient algorithm used for ascending the landscape and searching for the critical
points is described below. In optimal control experiments and simulations of a quantum system with dynamics having the form of Eq. (4.1), the objective function \( J \) can be optimized by updating the control field(s) iteratively along the gradient flow, until a critical point is encountered [7, 17, 30]. We introduce a scalar parameter \( s \geq 0 \) to describe the evolution of a control field \( u_m(s, t) \) in the landscape ascent process, which is guided by the differential equation

\[
\frac{\partial u_m(s, t)}{\partial s} = \alpha_m(s) \frac{\delta J}{\delta u_m(s, t)},
\]

(4.20)

where \( \alpha_m > 0 \) is the optimization rate associated with \( u_m \). We discretize the nominally smooth field \( u_m(t) \) to a piecewise constant function of \( L \) time intervals with identical length \( \Delta t = T/L \), and denote the field value within the \( l \)th interval \( (l = 1, \cdots, L) \) by \( u_m[l] \), where \( s \) is implicit. The differential equation (4.20) can be approximately solved by the forward Euler method with a constant step size, i.e.,

\[
u_m[k + 1, l] = u_m[k, l] + \alpha_m \left| \frac{\partial J}{\partial u_m[l]} \right|_k.
\]

(4.21)

where \( k = 0, 1, \cdots \) is the iteration index. In the laboratory, the partial derivatives \( \frac{\partial J}{\partial u_m[l]} \) can be measured by the central finite difference method [21] if the signal-to-noise ratio is sufficiently high, while in simulation they are approximated by the analytic gradient, given in Eq. (4.10), at the instants \( t = l\Delta t \), i.e., \( \frac{\partial J}{\partial u_m[l]} \approx \Delta t \frac{\delta J}{\delta u_m(l\Delta t)} \). Depending on application, higher-order (local) integration methods may also be utilized [17], although such techniques may be problematic to implement in the laboratory.

### 4.4.1 Two-spin case: the likelihood of trapping in practice

Here we will study the role of the singular local trap on a particular landscape in the simplest system of two coupled spins, illustrated experimentally and complemented with simulations. We consider neat \(^{13}\text{C}\)-labeled chloroform (\(^{13}\text{CHCl}_3\)) as the sample [22], and denote
the two nuclei $^1\text{H}$ and $^{13}\text{C}$ as spin 1 and 2, respectively. The coupling strength between the two spins is $J_{1,2}/2\pi = 215\text{Hz}$, measured from the frequency difference of the two lines of the doublet $^{13}\text{C}$ signal in the NMR spectrum. Two radiofrequency pulses as the control resources irradiate the sample simultaneously, with each pulse on resonance with $^1\text{H}$ or $^{13}\text{C}$, respectively, having a fixed length of $T = 5\text{ms}$ along with time-dependent envelope amplitudes $A_1, A_2$ and phases $\phi_1, \phi_2$. A doubly frame rotating at resonance with the Larmor frequencies of the two spins is constructed such that the nominal controls in Eq. (4.7) are realized by $u^1_x(t) := A^1(t) \cos \phi_1(t)$, $u^1_y(t) := A^1(t) \sin \phi_1(t)$, and similarly for $u^2_x$ and $u^2_y$ [21, 22]. The final time $T$ is chosen such that $J_{1,2}T \gtrsim 2\pi$ for the sake of system controllability [27, 32], and $T$ is broken into $L = 5$ time intervals, giving a total of 20 control variables. In the rotating frame, the full Hamiltonian of the controlled system is formulated as [30]

$$H(t) = J_{1,2}I^1_zI^2_z + u^1_x(t)I^1_x + u^1_y(t)I^1_y + u^2_x(t)I^2_x + u^2_y(t)I^2_y. \tag{4.22}$$

The traceless portion of the thermal equilibrium state at the experimental temperature (295 K) can be well approximated by the form $\rho_{eq} \approx c_1I^1_z + c_2I^2_z$, which is taken as the initial density matrix $\rho(0)$ of the landscape problem. For the sample molecule we have $c_1 : c_2 \approx 4 : 1$, determined by the intrinsic gyromagnetic ratios of $^1\text{H}$ and $^{13}\text{C}$ [25]. The observable $O = I^2_z$ is characterized by the integrated area of the doublet peak ascribed to the $^{13}\text{C}$ nucleus (the splitting is caused by its coupling with $^1\text{H}$) in the $^{13}\text{C}$-detected NMR spectrum. With these choices of $\rho(0)$ and $O$, the landscape $J = \text{Tr}[\rho(T)O]$ will possess KCPs at $J = \pm c_1, \pm c_2, 0$, with $\pm c_1$ being the global maximum and minimum [9].

From the landscape features at the special dynamic controls analyzed in Section 4.3, we can predict the general behavior of gradient optimization starting from various special initial guesses as explained below, which will be illustrated both experimentally and with simulations. (i) If we perform landscape gradient ascent from initial controls for which $u^1_x, u^1_y, u^2_x \equiv 0$, the search trajectory will converge to a suboptimal critical point located at
\( J = c_2 \), which is proved to be a second-order trap on the dynamic landscape. If the initial control only has \( u_x^1, u_y^1 \equiv 0 \), the search will also converge to \( J = c_2 \), but the critical point should have the topology of a saddle in this case. With such special initial controls, the zero field components in the initial control will permanently stay at zero during optimization with the gradient algorithm, and the global maximum of the landscape at \( J = c_1 \) cannot be approached. (ii) When the initial control is chosen to be close to the special zero field conditions in (i) leading to a trap or a saddle, the search is expected to slow down around \( J = c_2 \), and then speed up after escaping the neighborhood of the critical point. Under practical conditions when experimental or numerical error is present, the determination of an extremely small gradient could be inaccurate, and the search trajectory may wander randomly in a small region around the critical point. Other stochastic search methods (e.g., a genetic algorithm) may always be used, but they generally will not reveal subtle landscape features. (iii) When the initial control is far from the special conditions, e.g., each of the four components has a significant initial magnitude, the search trajectory will likely converge to the global maximum straightforwardly without halting at any suboptimal near-critical region. The suppositions in (ii) and (iii) are based on considering that the domain of attraction for the zero field induced singular trap is small on the scale of optimal field strengths, which will be affirmed below.

The distinct behavior described above is illustrated with three search trajectories starting from different initial controls in the laboratory utilizing the gradient algorithm (see Figure 4.1). We use the mean absolute value of \( u_m \), rescaled in units of \( T^{-1} = 1/(5\text{ms}) \) by referring to the experimentally determined amplitude of a \( \pi/2 \) pulse for \(^{13}\text{C} \), to characterize the magnitude of each control field component, i.e., \( \overline{|u_m|} = \frac{\sum_{l=1}^{L} |u_m[l]|}{L} \). The trajectories #1, #2 and #3 in Figure 4.1 are initialized as follows: constant fields are picked as the initial control, i.e., \( u_m[1] = \cdots = u_m[L] = u_{0,m}, \ m = 1, \cdots, 4 \); from trajectories #1, #2 to #3, each \( u_{0,m} \) increases accordingly with the ratio \( 0.25 : 0.58 : 1 \). Although exact zero fields was not chosen, trajectory #1 started at relatively small values. Despite the \( J \)
values at the initial guesses all being close to zero in Figure 4.1(a), the three trajectories evolve toward different destinations under the deterministic gradient algorithm. Trajectory #1 with the lowest initial field magnitude was manually terminated when its monotonicity broke, i.e., the $J$ value of the last iteration was slightly lower than that of its previous iteration, indicating that the system precision and noise levels were reached. Figure 4.1(b) for trajectory #1 shows that in the optimization process, the magnitudes of $u_x^1$, $u_y^1$ and $u_x^2$ stayed at relatively low levels compared with $u_y^2$. Thus, the movement of trajectory #1 in the control space was stopped somewhere in the neighborhood of the predicted local trap, for which the three components other than $u_y^2$ should be exactly zero. In trajectory #2 as shown in Figure 4.1(c), escaping the near-critical region in iterations 10-15 was accompanied by a dramatic increase of the magnitudes of $u_x^1$, $u_y^1$ and $u_x^2$, which can be viewed as a deviation from the trap condition. Trajectory #3, depicted in Figure 4.1(d), with highest initial control magnitudes more efficiently passed through the level of $J$ where the trap exists, and converged to the global maximum within fewer iterations than trajectory #2. The $J$ value where trajectory #1 is terminated and #2 slows down is approximately one-fourth of the global maximal value, in agreement with the relative positions of the two KCPs at $J = c_1$ and $c_2$ on the landscape.
Figure 4.1: (Color online) Three experimental gradient ascent trajectories on the control landscape of the two-spin sample of $^{13}$CHCl$_3$, starting from constant initial fields with different magnitudes. (a) The optimization of the objective value $J$. Trajectory #1 is terminated when $J$ drops in the last iteration, while the other two approach the global maximum. (b)-(d), respectively for trajectories #1, #2 and #3, showing the evolution of the mean field strength $|u_m|$ of each component, $u_1^x,u_1^y,u_2^x$ and $u_2^y$, in the optimization processes (given in the unit of $T^{-1}$).

Employing numerical simulation, we also assess the likelihood of an optimization search being “trapped” under convergence criteria based on the norm of landscape gradient, with the constraints on control resources (i.e., a set number $L$ of control variables) taken into consideration. In parallel with the experimental two-spin system, we set the $J$-coupling strength to $J_{1,2}/2\pi = 215$Hz and the final time to $T = 5$ms (digitized into either $L = 5$ or 10 equal time intervals). The initial state is chosen as $\rho(0) = I^1_z + 0.25I^2_z$, and the observable is $O = I^2_x$. The forward Euler method with a constant step size as in Eq. (4.21) is utilized for ascending the landscape, in which the gradient with respect to the piecewise
constant fields is approximated from the analytical expression in (4.10). The initial controls are generated by random numbers uniformly distributed around zero for each of the $4L$ control variables, and the mean absolute value $|u_m|$ of each of the four components, $u_x^1$, $u_y^1$, $u_x^2$ and $u_y^2$, is normalized to the same value $u_0$ (i.e., all the $L$ control variables of each field $u_m$ are proportionally rescaled such that $|u_m| = u_0$). An optimization process is terminated when the gradient norm $e$ becomes lower than a convergence threshold, chosen over the window $10^{-5} \sim 10^{-3}$ in Figure 4.2. The norm $e$ of the landscape gradient is defined as

$$e = \left\{ \frac{1}{L} \sum_m \sum_{l=1}^L \left( \frac{\delta J}{\delta u_m(l\Delta t)} \right)^2 \right\}^{1/2}.$$  

(4.23)

The proportion of optimization searches being practically trapped at various values of $u_0$ and $e$ is determined by 100 runs and given in Figure 4.2, with two different time resolutions of $L = 5$ and 10. Generally, the search is more likely to escape the trapping region when greater magnitude initial controls are chosen, and a more precise convergence threshold is employed. No fundamental difference is observed for the two time resolution values of $L = 5$ and 10 with the other parameters fixed, implying that the discretization constraints do not affect the landscape structure significantly even for $L = 5$. For this particular problem, the accumulation of larger numerical error in the case of $L = 5$ actually eases the deviation from the trapping region, according to the statistical data presented here. In Figure 4.3 we selected 20 initial controls for each distinct $u_0$, optimized them until $J > 0.99$, and then recorded the mean field strengths $u_{\text{opt}} = |u_m|$ of each component at the global optimum, which turned out to be much larger than the easily trapped initial field strengths in Figure 4.2. Therefore, if the field strengths of the initial control are sufficiently large, i.e., comparable to those of global optimal control fields for the landscape, the trapping region will be very unlikely to halt the search. This phenomenon was also observed elsewhere [7, 17] and is expected to widely apply, consistent with large numbers of reported successful simulations [11, 23].
Figure 4.2: (Color online) Simulation results from a model of the experimental two-spin system, showing the proportion of searches in 100 runs that are terminated prematurely at around $J = 0.25$ for various initial field strengths $u_0$ and convergence thresholds for the gradient norm $e$. (a) and (b) correspond to time resolution of $L = 5$ and 10, respectively.

Figure 4.3: (Color online) The mean field strengths of globally optimal controls ($J > 0.99$) of the two-spin landscape problem, obtained by optimizing from initial controls with different strengths $u_0$ in simulation.

Furthermore, we also numerically explored the higher-order landscape topology about the trap associated with trajectory #1 to check whether it is actually a local maximum. We randomly picked large numbers of perturbation directions $\delta u(\cdot)$ within the Hessian null space at a trap (including all four field components), and calculated the change of $J$ when the control $u_{\text{trap}}$ was varied along these directions. For sufficiently small variation, $J$ is dominated by its lowest-order nonzero derivative along $\delta u$. It is found that in a small neighborhood around $u_{\text{trap}}$ we have $J(u_{\text{trap}} + x\delta u) \approx J(u_{\text{trap}}) + k_4x^4$, where $x$ is the
distance from \(u_{\text{trap}}\) in the control space. The coefficient \(k_4\) can be positive or negative depending on the direction \(\delta u\), with two representative cases plotted in Figure 4.4. The numerical evidence implies that within the Hessian null space at the trap, the third derivative of \(J\) is zero while the fourth is indefinite. Thus, the trap identified in the two-spin example is actually only a trap to second order, but not a true local maximum.

![Figure 4.4](image)

Figure 4.4: (Color online) The second-order trap in the two-spin problem is shown not to be a local maximum. As we perturb a trap control \(u_{\text{trap}}\) along different directions \(\delta u\) in the Hessian null space, \(J\) may either increase or decrease from the value \(J_0\) of the trap as a quartic function of the distance from \(u_{\text{trap}}\), as the representative red and blue curves respectively show.

### 4.4.2 Three-spin system: linear and triangular coupling models

Here we show by numerical simulations in the three-spin case that the linear chain model is crucial for the existence of the singular traps. Consider a system of three spins-1/2 labeled 1, 2, and 3, with Ising-type coupling between each pair of them. The Hamiltonian can be expressed in the on-resonance rotating frame as

\[
H(t) = J_{1,2}I_1^z I_2^z + J_{1,3}I_1^z I_3^z + J_{2,3}I_2^z I_3^z + \sum_{j=1}^{3} [u_j^x(t)I_x^j + u_j^y(t)I_y^j].
\]

A triangular coupling model is formed when the three coupling constants \(J_{1,2}, J_{1,3}\) and \(J_{2,3}\) are all nonzero, while a linear chain model has only two nonzero couplings, the central spin of the chain coupled with the two terminal spins.
Consider the initial state \( \rho(0) = \sum_{j=1}^{3} c_j I^j_z \) \((c_1 > c_2 > c_3 > 0)\). We select several types of special dynamic controls containing particular zero components and determine by simulation the ranks of the Jacobian \( \frac{\delta \rho(T)}{\delta u(t)} \) with all of the six components involved. For each case, random non-constant fields for the nonzero components without any optimization are used in the calculation of Jacobian, and the results are summarized in Table 4.2. For a non-degenerate \( \rho(0) \) of the three-spin system with a Hilbert space dimensionality of \( N = 8 \), the full rank of the Jacobian should be \( N(N - 1) = 56 \) [6]. A dynamic control with all of its field components being nonzero is regular with a Jacobian rank of 56 for both linear and triangular coupling models, as the full controllability implies. Generally, turning off more field components or removing internuclear coupling channels might decrease the Jacobian rank and thus make the dynamic control “more singular”. In the following we will see how the singularity of such controls involving zero fields can influence the Hessian spectrum at landscape critical points.

Table 4.2: Degrees of singularity at special dynamic controls, with some of the six field components fixed to zero, under linear and triangular coupling models. \( u_y \) (terminal/central) indicates that the only nonzero component is \( u^k_y(t) \) irradiated on a spin \( k \), which is located at the terminal/central position of the linear chain, while “None” means that all the field components are set to zero. The different \( J \)-coupling constants are set to be unequal. The Jacobian rank is calculated by choosing random fields for the nonzero components in each case. A dynamic control with particular zero components and random nonzero components has a particular Jacobian rank, whether it is critical or not.

<table>
<thead>
<tr>
<th>Model</th>
<th>Nonzero component(s)</th>
<th>Rank of ( \frac{\delta \rho(T)}{\delta u(t)} )</th>
<th>Corank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>None</td>
<td>16</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>( u_y ) (terminal)</td>
<td>20</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>( u_y ) (central)</td>
<td>24</td>
<td>32</td>
</tr>
<tr>
<td>Triangular</td>
<td>None</td>
<td>24</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>( u_y ) (each single)</td>
<td>40</td>
<td>16</td>
</tr>
<tr>
<td>Either</td>
<td>All six</td>
<td>56</td>
<td>0</td>
</tr>
</tbody>
</table>

We define a collection of control landscape problems with the initial state \( \rho(0) \) above and observables \( O = I^k_x \), \( k = 1, 2, 3 \), which are all identical in the kinematic picture.
Table 4.3: Kinematic Hessian indices at some KCPs on a landscape of the three-spin system, defined by \( \rho(0) = \sum_{j=1}^{3} c_j I_j^z \) (\( c_1 > c_2 > c_3 > 0 \)) and \( O = I_x^k \) for \( k = 1, 2 \) or 3; the latter three landscapes are kinematically identical.

<table>
<thead>
<tr>
<th>Condition</th>
<th>( J_{\text{crit}} )</th>
<th>( D_+ )</th>
<th>( D_- )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c_1 &lt; c_2 + c_3 )</td>
<td>2c_1</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>2c_2</td>
<td>6</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>2c_3</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>( c_1 &gt; c_2 + c_3 )</td>
<td>2c_1</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>2c_2</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>2c_3</td>
<td>12</td>
<td>20</td>
</tr>
</tbody>
</table>

Utilizing the methods in [9, 10] we find that (i) there exist KCPs at (not limited to) \( J = \pm 2c_1, \pm 2c_2, \) and \( \pm 2c_3 \) for any \( k = 1, 2, \) or 3; (ii) the global maximum is \( J_{\text{max}} = \max\{c_1 + c_2 + c_3, 2c_1\} \); and (iii) the indices \( D_+ \) and \( D_- \), respectively denoting the numbers of positive and negative eigenvalues of the kinematic Hessian, have the relation that \( D_+ + D_- = 32 \) for any KCP of this landscape (see Table 4.3; note that the kinematic landscape topology can vary with the relative magnitudes of \( c_1, c_2 \) and \( c_3 \)).

An initial control will be referred to as a special initial control if its only nonzero field component is \( u_y^k(t) \), where \( k \) is the spin targeted by the observable \( O \). Regardless of whether the three \( J \)-coupling constants are zero or not, the landscape gradient ascent from such a special initial control will converge to a special critical point at \( J = 2c_k \) corresponding to \( O(T) = I_x^k \), and all the five zero components of the initial control will stay at zero during optimization. By Eq. (4.13) we calculate the Hessian matrices at the special critical points with linear and triangular coupling models, and check the dynamic Hessian indices \( D_+ \) and \( D_- \) (no greater than the corresponding \( D_+ \) and \( D_- \) in the kinematic picture), in order to discover under what conditions any dynamic traps can be formed (see Table 4.4). For generality the different nonzero coupling constants are set to be unequal.

If the target spin \( k \) is a terminal member of the linear chain (e.g., \( O = I_x^1, J_{1,2}, J_{2,3} \neq 0 \) while \( J_{1,3} = 0 \)), the Hessian at the special critical point will have \( D_+ = 0 \) and \( D_- = 8 \) for any of the three observables with \( k = 1, 2, 3 \). If spin \( k \) is located at the center and coupled
Table 4.4: Dynamic Hessian indices $D_+$ and $D_-$ at the special critical points for several observables, in linear and triangular coupling models. “Linear, terminal (central)” means that the spin $k$ targeted by the observable $O = I^k_x$ is located at the terminal (central) position of the three-spin chain. The different $J$-coupling constants are set to be unequal.

<table>
<thead>
<tr>
<th>Model</th>
<th>Observable</th>
<th>$D_+$</th>
<th>$D_-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear, terminal</td>
<td>$I^1_x$, $I^2_x$, $I^3_x$</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Linear, central</td>
<td>$I^1_x$, $I^2_x$, $I^3_x$</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Triangular</td>
<td>$I^1_x$</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>$I^2_x$</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>$I^3_x$</td>
<td>8</td>
<td>16</td>
</tr>
</tbody>
</table>

with the other two (e.g., $O = I^1_x$, $J_{1,2}, J_{1,3} \neq 0$ while $J_{2,3} = 0$), the Hessian indices will become $D_+ = 0$ and $D_- = 16$. Therefore, a special critical point at $J = 2c_k$ must be a local second-order maximum on the landscape of $O = I^k_x$ with the linear chain model, and form a trap unless it is the global maximum when $k = 1$ and $c_1 > c_2 + c_3$. When a triangular coupling model ($J_{1,2}, J_{1,3}, J_{2,3} \neq 0$) is employed, the special critical point at $J = 2c_k$ will still be encountered by the search trajectory from a special initial control, but not necessarily become a local maximum. For $O = I^2_x$ and $I^3_x$, the critical points recover their saddle topology qualitatively with both positive and negative Hessian eigenvalues; for $O = I^1_x$, the special critical point with a negative semidefinite Hessian ($D_+ = 0$, $D_- = 24$) is still a suboptimal trap when $c_1 < c_2 + c_3$, but the global maximum when $c_1 > c_2 + c_3$. These results imply that careful consideration is required before generalizing the second-order trap conclusion, derived in Section 4.3 with the linear spin chain model, to more complex systems including additional non-local couplings.

4.5 Discussion and conclusion

This chapter mainly considers the control landscape for Ising linear spin chain systems, and derives the sufficient condition for a type of suboptimal traps of at least second order over the landscape caused by singularity. The results extend the understanding of the
role of singular controls in landscape analysis. For quantum systems controlled by multiple fields, turning off some field components can produce singular controls of various degrees. The problem analyzed here provides an example in which the singular traps exist as a continuously varying manifold over the nonzero field components involved, and the distinct members of which can be identified by performing simple gradient optimization from some particularly chosen initial controls. To illustrate the theoretical conclusions, we performed NMR experiments with a heteronuclear two-spin molecule, and observed the influence of the predicted singular trap on practical gradient searches. Additional numerical simulations showed how the choice of initial controls influences the likelihood of being trapped in practice, and that the trap is actually not a local optimum when higher-order derivatives are taken into consideration. For the three-spin case we compared two different physical models of linear and triangular internuclear couplings, suggesting that the singular traps may disappear when more complex couplings are present in a set of spins. In optimal control experiments with spin systems the presence of landscape singularity could be a factor of significance. However, the influence of zero field traps appear easy to be avoided by starting with simple physically motivated initial choices for controls, based on the expectation that each coupled spin may have some role in the optimal dynamics.

The advances in this work are best understood in the overall context of control landscape principles. As stated in the Introduction, these principles rest on the satisfaction of three assumptions. The first assumption (i) on controllability and the second assumption (ii) on surjectivity concern inherent properties of the quantum system under control. The third assumption (iii) on complete access to all possible controls can never be fully satisfied, but the practical issue is having adequate control resources to meet the dynamic objective; while significant violation of this practical criterion will eventually lead to traps [12], reasonable considerations from spectroscopy appear to often provide guidance to identify suitable resources. Although the controllability assumption (i) is likely to be satisfied, violations can lead to traps on the landscape [4]. The surjectivity assumption (ii) is perhaps
the least understood, and the present work along with prior studies [5, 15, 16] have identified singular control induced traps existing only as zero (or constant) fields, which are unusual experimentally where propagating fields naturally arise as controls. Simple physically based trial fields can easily avoid constant field traps, as shown here and in other works [7]. A recent study introduced a special algorithm to specifically seek singular controls, and none of the discovered singular controls (all were non-constant) corresponded to traps [6]. Additionally extensive careful simulations for state preparation [11], optimization of general observables [?], and the creation of targeted unitary transformations [23] also did not encounter any fields corresponding to traps on the landscape while on the way to reaching high fidelity control performance. Interestingly, all of the known cases with singular control induced traps are simple systems with sparse coupling (e.g., Ising spin systems considered here). The greater body of simulation studies [11, 23], which do not encounter traps of any type, generally have Hamiltonians with more complex coupling structure. The landscape findings in the present work upon going from a three-spin linear chain to the triangular coupling arrangement can also be viewed as increasing coupling complexity. These collective results suggest that trap-inducing singular controls are a rarity, especially in systems of extensive coupling complexity. Most importantly, further study is needed to fully understand the circumstances when any of the three landscape assumptions break down and produce control outcomes that encroach on the nominal favorable landscape topology.
Bibliography


[29] Only the critical points of the greatest $c_j$ might be the top and bottom, but in some cases the actual top and bottom can still exceed them and do not belong to the category $J = \pm 2^{n-2}c_j$, as Section 4.3.3 shows in the three-spin case.


Chapter 5

Measuring the distance from saddle points and driving to locate them over quantum control landscapes

Optimal control of quantum phenomena involves the introduction of a cost functional $J$ to characterize the degree of achieving a physical objective by a chosen shaped electromagnetic field. The cost functional dependence upon the control forms a control landscape. Two theoretically important canonical cases are the landscapes associated with seeking to achieve either a physical observable or a unitary transformation. Upon satisfaction of particular assumptions, both landscapes are analytically known to be trap-free, yet possess saddle points at precise suboptimal $J$ values. The presence of saddles on the landscapes can influence the effort needed to find an optimal field. As a foundation to future algorithm development and analyses, we define metrics that identify the “distance” from a given saddle based on the sufficient and necessary conditions for the existence of the saddles. Algorithms are introduced utilizing the metrics to find a control such that the dynamics arrive at a targeted saddle. The saddle distance metric and saddle-seeking methodology is tested numerically in several model systems.
5.1 Introduction

The control of quantum phenomena is of fundamental interest and has various potential applications in many areas of physics, chemistry and biology [1]. In general, the goal of quantum optimal control is to achieve desired manipulation of a quantum system through the introduction of classical control fields, often shaped femtosecond laser pulses in the laboratory. The outcomes of growing numbers of optimal control experiments and simulations indicated that it was relatively easy to obtain good solutions for these problems while searching through the high-dimensional control spaces. Seeking a fundamental explanation of the very attractive behavior motivated analysis of the quantum control landscape [2]. A control landscape is formed by a cost functional $J$, which characterizes how well the objective is achieved at the target time $T$. The landscape is defined by the dependence of the cost functional $J$ upon the control field $\epsilon(t)$, $t \in [0, T]$. In the laboratory the control objective can be defined by various measurable quantities depending on the application [1]. Theoretical analysis has mostly focused on two types of control landscapes: the physical observable landscape [2, 3, 4, 5, 6, 7, 8, 9] and the unitary transformation landscape [10, 11, 12, 13]. The former aims at extremizing a target observable of the system, and the cost functional $J$ is defined as the expectation value of a Hermitian operator of the quantum ensemble at the target time. A well-studied special case of this type is the transition probability $P_{i \rightarrow f}$ from an initial state $|i\rangle$ to a final state $|f\rangle$, both being pure states [2, 3, 4, 5, 6]. The second type of control problem seeks to create a targeted unitary transformation $W$. In this circumstance $J$ characterizes the difference between the target $W$ and the unitary propagator $U$ at time $T$, produced by applying a certain control field to the quantum system. An important application in this case arises in quantum gate synthesis, i.e., generating particular unitary operations of high quality utilizing physical qubits (e.g., represented by photons, nuclear spins, etc.), which form the building blocks of quantum computation [14].
The topological details of the control landscapes, especially those of the slope and curvature (i.e., respectively characterized by the first derivative or gradient, and second derivative or Hessian matrix of $J$), are crucial for understanding the efficiency and overall behavior of algorithms searching for effective control fields [6, 15, 16]. Specifically, the distribution of controls with a zero gradient of $J$, called critical (or stationary) points, can especially affect the efficiency of gradient-based or other local control search algorithms. For both the observable and unitary transformation control landscapes, theoretical analysis based upon satisfaction of some underlying assumptions (see the discussion below Eq. (5.2)) has shown that the critical points only exist at some particular $J$ values; a set of continuously varying critical controls sharing the same $J$ value and local topology form a critical submanifold (CM). Besides the landscape top and bottom (global maximum and minimum), all the other critical points corresponding to intermediate $J$ values, if any, must be saddle points [8, 10], which do not form traps (suboptimal local maxima or minima) for a gradient search algorithm seeking to approach globally optimal solutions. The lack of traps over the landscape reveals part of the reason for the observed relative ease of searching for optimal controls, although coming near a saddle manifold could considerably slow down a search for an effective control.

In reference to the last comment, since the landscape has a low gradient near the saddle critical points, a search process for optimizing $J$ with gradient-based or other local algorithms will slow down when running in the neighborhood of the saddles. This behavior has been observed in simulations [16, 17, 18, 19] and experiments [20, 21], providing evidence for the existence of landscape saddle points and their impact. Whereas the global maximum (minimum) CM occupy the landscape top (bottom), the saddles are distributed at particular levels of $J$ value like “islands”, i.e., there are also controls that produce the same value of $J$ as the saddles, but are not critical points (i.e., they miss the saddle islands). For some initial controls the search may encounter a near-saddle region when passing through the corresponding $J$ value, while starting from some other initial controls may produce an
excursion to the landscape top without encountering saddles [16]. Moreover, theoretical analysis predicts that as the system dimensionality $N$ increases the number of saddles generally rises, but the “volume fractions” of the saddle near-critical sets appear to decrease to zero, implying that the saddles may become less likely to be encountered in searches [11, 22]. Thus, understanding the nature and influence of saddles upon seeking optimal controls is important, but knowledge of the extent of a particular saddle is unlikely to be discovered by simply trying different initial controls and optimizing them along the gradient flow of $J$, especially in quantum systems with high numbers of states $N$. Despite the fact that the topological properties of the CMs have been revealed [8, 11] theoretically, questions remain about (a) how to measure the “distance” of a given dynamic control from a particular CM on the landscape upon a climb to possibly provide guidance on finding better optimization algorithms, (b) how to precisely reach a saddle point, which is fundamentally important for the identification and systematic study of the extent of landscape saddles. These issues will be addressed here in a simulation context, but analogous questions arise when performing control experiments [20]. The ability set out in steps (a) and (b) opens up the concomitant prospect of seeking an algorithm to specifically avoid saddles, and thereby possibly accelerate the search for optimal controls.

In this work we develop algorithms for optimizing an arbitrary initial control such that the ensuing dynamics move toward any specified CMs for either a physical observable or a unitary transformation control landscape. The strategy is based on defining the distance of a control solution to a target CM (the distance being zero if and only if the control belongs to the target CM), and then seeking to minimize the distance to zero by a gradient-based method. For the observable and unitary transformation landscapes, the sufficient and necessary conditions for reaching a CM are provided, respectively, in sections 5.2 and 5.3. This analysis is based on the associated distance metrics, and analytic expressions of their derivatives are provided for constructing the gradient algorithms to home in on a targeted CM. As illustrations, in section 5.4 we will utilize these algorithms to search
for target saddle submanifolds in two numerical model systems. In summary, although basic theoretical landscape analysis can identify the $J$ values where a saddle point lies, the latter analysis cannot \textit{a priori} identify how close one is to a CM. The missing information is provided by the CM distance metrics in the chapter, followed by their use to directly approach a saddle.

5.2 The physical observable control landscape

5.2.1 Critical points over the landscape

In order to provide background for the methods introduced in the chapter, this section briefly summarizes some fundamental concepts and conclusions of quantum control landscape analysis, especially the characterization and topological properties of landscape critical points [7, 8].

The time evolution of a closed $N$-level quantum system under control can be described by its density matrix $\rho(t) = U(t, 0)\rho U^\dagger(t, 0)$, where $\rho := \rho(0)$ is the initial state and the unitary system propagator $U(t, 0)$ is generated by the Hamiltonian according to the time-dependent Schrödinger equation (in units where $\hbar = 1$)

$$i\frac{d}{dt} U(t, 0) = H[\epsilon(t)]U(t, 0), \quad U(0, 0) = \mathbb{I},$$

(5.1)

where $\mathbb{I}$ is the $N$-dimensional identity matrix. The governing Hamiltonian contains a control field $\epsilon(t)$ defined over the time interval $t \in [0, T]$ that guides the evolution of the system. The primary goal of the observable control problem is to find fields that extremize the expectation value of a desired Hermitian observable operator matrix $\theta$ at the target time $T$:

$$J_O[\epsilon(t)] = \text{Tr}[\rho(T)\theta] = \text{Tr}[U(T, 0)\rho U^\dagger(T, 0)\theta] := \text{Tr}(U \rho U^\dagger \theta),$$

(5.2)
where $U$ is the shorthand notation for $U(T, 0)$. Thus, the dynamic landscape $J_O[\epsilon(t)]$ on the space of control fields can be re-expressed on the unitary group $U(N)$ of dimension $N$, forming the kinematic landscape $J_O(U)$ on which some theoretical analysis [8] as well as simulations [11, 23, 24] were based. The kinematic landscape will preserve its topology when mapped to the dynamic picture if three assumptions are satisfied: (i) the quantum system is fully controllable, i.e., any unitary transformation $U$ can be produced at some time $T$ by a control field [25]; (ii) the map from the field $\epsilon(t)$ to $U$ is surjective, i.e., the Jacobian matrix $\delta U/\delta \epsilon(t)$ has full rank; and (iii) the control fields are unconstrained. Singular (respectively, regular) controls that violate the assumption (ii), i.e., having rank deficient Jacobians, can also become dynamic critical points which have no kinematic counterparts [26]. Despite the existence of singular critical points in certain cases, the collective reported results suggest that assumption (ii) may be commonly satisfied both in simulation [6] and in a variety of experiments [1]. Thus, without significantly limiting the scope of this work, we will exclude the singular controls from consideration and only deal with the regular controls in this chapter.

The necessary and sufficient condition for $U$ to be a critical point on the kinematic landscape is that the derivative of $J_O$ with respect to $U$ be zero. Under the parametrization $U(s, A) := e^{i s A} U = [\mathbb{1} + i s A + \mathcal{O}(s^2)] U$ in the neighborhood of $U$ on the unitary group, where $s$ is a real scalar and $A$ is a Hermitian matrix, $U$ is critical if and only if

$$\frac{d}{ds} J_O[U(s, A)] \bigg|_{s=0} = \frac{d}{ds} \text{Tr}(e^{i s A} U \rho U^\dagger e^{-i s A} \theta) \bigg|_{s=0} = i \text{Tr}(A[U \rho U^\dagger, \theta]) = 0, \quad \forall A^\dagger = A,$$

which is equivalent to

$$[U \rho U^\dagger, \theta] = 0.$$

This criterion, derived in different manners previously [7, 9], shows that at a critical point of the landscape in (5.2) the final state, $\rho(T) = U \rho U^\dagger$, must commute with the observable
\( \theta \). Therefore, they can always be simultaneously diagonalized by the same (generally non-unique) unitary matrix \( U \), i.e.,

\[
\rho(T) = U \Lambda_{\rho(T)} U^\dagger, \quad \theta = U \Lambda_{\theta} U^\dagger.
\] (5.5)

The diagonal matrices \( \Lambda_{\rho(T)} \) and \( \Lambda_{\theta} \) contain appropriately permuted eigenvalues of \( \rho \) and \( \theta \), respectively. Thus, the \( J_O \) value at any critical point can be written as [7]

\[
J_{O, \text{crit}} = \text{Tr}[U \Lambda_{\rho(T)} U^\dagger \Lambda_{\theta} U^\dagger] = \text{Tr}[\Lambda_{\rho(T)} \Lambda_{\theta}].
\] (5.6)

Eq. (5.6) shows that kinematic critical points only exist at some particular discrete \( J_O \) values over the landscape, which are determined merely by the eigenvalues of \( \rho \) and \( \theta \) and independent of the Hamiltonian. There can be infinitely many \( U \)'s corresponding to any particular critical point on the landscape, and in order to develop an algorithm searching for control fields reaching specified critical points we need to categorize and characterize them beforehand. This task was accomplished in Ref. [8], as briefly summarized below, by reducing the characterization of critical points to a simple finite combinatorial problem.

Above we only treated \( \rho \) and \( \theta \) as arbitrary Hermitian matrices, and in the following we will rewrite them in a more explicit form by specifying the multiplicities of each eigenvalue. Suppose that \( \rho \) and \( \theta \) have \( r \) and \( s \) distinct eigenvalues \((r, s \leq N)\), respectively, we can always express both \( \rho \) and \( \theta \) as diagonal matrices with their eigenvalues sorted in descending order (consistent with maintaining this ordering, they can be unitarily transformed without loss of generality; c.f. [8]), i.e.,

\[
\rho = \begin{pmatrix}
\rho_1 \mathbb{I}_{n_1} & & \\
& \ddots & \\
& & \rho_r \mathbb{I}_{n_r}
\end{pmatrix}, \quad \theta = \begin{pmatrix}
\theta_1 \mathbb{I}_{m_1} & & \\
& \ddots & \\
& & \theta_s \mathbb{I}_{m_s}
\end{pmatrix},
\] (5.7)
where \( \rho_1 > \cdots > \rho_r \) are distinct eigenvalues of \( \rho \) with \( n_1, \cdots, n_r \) multiplicities, and \( \theta_1 > \cdots > \theta_s \) are distinct eigenvalues of \( \theta \) with \( m_1, \cdots, m_s \) multiplicities. According to Theorem 1 in Ref. [8], with the formulation of (5.7), a unitary matrix \( U \) corresponds to a landscape critical point if and only if it lies in the double coset of some permutation matrix \( \pi \), i.e.,

\[
U = P \pi Q^\dagger: P \in U(m), Q \in U(n),
\]

where \( U(n) \) is the product group \( U(n_1) \times \cdots \times U(n_r) \) of unitary groups of dimensions \( n_1, \cdots, n_r \) associated with \( \rho \), and \( U(m) = U(m_1) \times \cdots \times U(m_s) \) is defined in the same manner relevant to \( \theta \). When \( \rho \) and \( \theta \) are both fully non-degenerate, a critical \( U \) can only differ from a permutation matrix by trivial independent phase factors \( e^{i\phi} \) for each corresponding entry, while the degeneracies of \( \rho \) and/or \( \theta \) will allow a critical \( U \) to have more complex structures.

Substituting the critical condition (5.8) into the landscape (5.2), the block-diagonal unitary matrices \( P \) and \( Q \) cancel out when acting on \( \theta \) and \( \rho \), and the \( J_O \) values at the critical points are

\[
J_{O, \text{crit}} = \text{Tr}[(P \pi Q^\dagger)\rho(P \pi Q^\dagger)\theta] = \text{Tr}[\pi(Q^\dagger \rho Q)\pi^\dagger(P^\dagger \theta P)] = \text{Tr}(\pi \rho \pi^\dagger \theta).
\]

Eq. (5.9), equivalent with (5.6) above, implies that the critical points at distinct landscape values \( J_O \) result from distinct permutations of the eigenvalues of \( \rho \) with respect to those of \( \theta \). However, the unitary matrix \( \pi \) itself does not serve as a desirable characterization of the critical points, since multiple \( \pi \)’s may give critical points located at the same level and sharing the same local landscape topology when \( \rho \) and/or \( \theta \) contain degenerate eigenvalues.

A general characterization of the critical points is specified by their topological information, which is provided by the contingency table method [8]. As formulated in Table 5.1, a contingency table \( K \) is an integer-valued matrix of size \( s \times r \), whose entry \( K_{ij} \) (\( i \leq s, j \leq r \)) is defined as the number of positions on the diagonal of \( \pi \rho \pi^\dagger \) and \( \theta \) where the
distinct eigenvalues $\rho_j$ and $\theta_i$ appear simultaneously. Note that the contingency table described here (a $s \times r$ matrix) is transposed from its original form (a $r \times s$ matrix) defined in [8], for the convenience of our following derivation. The contingency table consists of these non-negative overlap numbers that sum to $N$ by definition. Equivalently, the contingency table can also be interpreted as a “compressed” permutation matrix. Given any critical point associated with the permutation $\pi$, the contingency table can be calculated by summing up the rows of $\pi$ belonging to each distinct eigenvalue of $\theta$, and the columns belonging to each distinct eigenvalue of $\rho$, as the ordering of identical eigenvalues makes no difference. Hence, a $N \times N$ permutation matrix $\pi$ is compressed to a $s \times r$ contingency table $K$, and distinct $\pi$’s could be mapped to the same $K$. By Eq. (5.9) and the definition of a contingency table, the landscape objective value at a critical point corresponding to $K$ is

$$J_O(K) = \sum_{i,j}^{s,r} \rho_j K_{ij} \theta_i.$$  \hspace{1cm} (5.10)

<table>
<thead>
<tr>
<th>$n_1$</th>
<th>$\cdots$</th>
<th>$n_r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1$</td>
<td>$K_{11}$</td>
<td>$\cdots$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
</tr>
<tr>
<td>$m_s$</td>
<td>$K_{s1}$</td>
<td>$\cdots$</td>
</tr>
</tbody>
</table>

Table 5.1: The contingency table $K$ as a nonnegative integer-valued matrix, describing the association of distinct eigenvalues of $\rho$ and $\theta$, with multiplicities of $n_1, \cdots, n_r$ and $m_1, \cdots, m_s$, respectively, at particular critical points of the landscape in Eq. (5.2).

The set of critical points (either kinematic or dynamic) sharing a common contingency table form a critical submanifold (CM). For cases with non-degenerate $\rho$ and $\theta$, the landscape possesses $N!$ CMs, whose contingency tables are the permutation matrices $\pi$’s. The emergence of degeneracies in $\rho$ and/or $\theta$ can make multiple permutations equivalent with each other, and thus merge these CMs into a smaller number of higher-dimensional ones. Within each CM all the critical points share a common objective value $J_O$, although distinct CMs may also correspond to an identical $J$. Moreover, the kinematic Hessians evaluated
about each point in the same CM have identical numbers of positive and negative eigenvalues, which can be calculated simply from the contingency table associated with this CM. Hessian analysis reveals that besides the global maximum and minimum of the landscape, all the suboptimal CMs, if any, possess both positive and negative Hessian eigenvalues and thus have the topology of saddle points [8].

An example illustrating the utilization of a contingency table is provided below, which will be further exploited in section 5.4.2 with some variations as well. Consider a control problem in a four-level quantum system with degenerate initial state and target observable, i.e., $\rho = \theta = \text{diag}\{0.5, 0.5, 0, 0\}$, both having two distinct eigenvalues ($\rho_1 = \theta_1 = 0.5$ and $\rho_2 = \theta_2 = 0$). Therefore, the contingency table associated with each CM over this landscape should be a $2 \times 2$ matrix. To identify a critical point, as an example, we arbitrarily pick a permutation matrix as

$$
\pi = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix},
$$

which yields $\pi \rho \pi^\dagger = \text{diag}\{0.5, 0, 0, 0.5\}$. To determine the contingency table of this case, we compare $\pi \rho \pi^\dagger$ with $\theta$ and find that the overlap number, for $\rho_1 = 0.5$ and $\theta_1 = 0.5$, is $K_{11} = 1$, which appears at the first position on the diagonal of both $\pi \rho \pi^\dagger$ and $\theta$. Similarly, $K_{12}, K_{21},$ and $K_{22}$ also turn out to be 1, and an entire contingency table ($K^\beta$ below) is obtained, characterizing a particular CM on the landscape. After checking all the permutations of the eigenvalues of $\rho$ versus $\theta$ in the same manner we find three distinct contingency tables for this example,

$$
K^\alpha = \begin{pmatrix}
2 & 0 \\
0 & 2
\end{pmatrix}, \quad K^\beta = \begin{pmatrix}
1 & 1 \\
1 & 1
\end{pmatrix}, \quad K^\gamma = \begin{pmatrix}
0 & 2 \\
2 & 0
\end{pmatrix}.
$$
The $J_O$ values at the three CMs are calculated with (5.10) to be

$$J_O(K^\alpha) = 0.5, \quad J_O(K^\beta) = 0.25, \quad J_O(K^\gamma) = 0.$$  

Thus, $K^\alpha$ and $K^\gamma$ respectively correspond to the global maximum (top) and global minimum (bottom) of the landscape, while $K^\beta$ with an intermediate $J_O$ value represents the only saddle submanifold. In the following, the three contingency tables will be denoted respectively by $K_{\text{top}}$, $K_{\text{saddle}}$ and $K_{\text{bottom}}$. Note that the symmetry of the contingency tables in this example is not a general property applicable to other cases.

### 5.2.2 Critical distance metric and saddle-seeking algorithm

In this section we first derive an important property of each CM with regard to its contingency table, which can be applied to define a metric of the distance from the submanifold. Recalling (5.7) and (5.8), for the $\alpha$-th particular $\text{CM}^\alpha$ with the contingency table $K^\alpha$ we divide both $U$ and $\pi^\alpha$ (not necessarily unique) into $s \times r$ blocks according to the degeneracies of $\rho$ and $\theta$. Each block $U_{jk}$ or $\pi_{jk}^\alpha$ with the dimension $m_j \times n_k$ constitutes the components of $U$ and $\pi^\alpha$,

$$U = \begin{pmatrix} U_{11} & \cdots & U_{1r} \\ \vdots & \ddots & \vdots \\ U_{s1} & \cdots & U_{sr} \end{pmatrix}, \quad \pi^\alpha = \begin{pmatrix} \pi_{11}^\alpha & \cdots & \pi_{1r}^\alpha \\ \vdots & \ddots & \vdots \\ \pi_{s1}^\alpha & \cdots & \pi_{sr}^\alpha \end{pmatrix}, \quad (5.11)$$

and we note that generally the $U_{jk}$ and $\pi_{jk}^\alpha$ blocks are not unitary or square. Addressing each individual block in (5.8) we find that for any $U \in \text{CM}^\alpha$,

$$U_{jk} = P_j \pi_{jk}^\alpha Q_k^\dagger, \quad \forall j \leq s, k \leq r, \quad (5.12)$$

118
for some unitary matrices $P_j \in \mathcal{U}(m_j)$ and $Q_k \in \mathcal{U}(n_k)$, which implies that $U_{jk}$ and $\pi_{jk}^\alpha$ share the same singular values (should not be confused with singular controls described in section 5.2.1). Considering the nature of a permutation matrix it is easy to find that $\pi_{jk}^\alpha$ can only have singular value(s) of 1 or 0, with the number of 1’s being equal to the corresponding entry in the contingency table, $K^\alpha_{jk}$, and likewise for $U_{jk}$. Hence, we conclude with the following Theorem 1 (see Appendix A for the full proof):

**Theorem 1.** On the observable control landscape $J_O(U) = \text{Tr}(U \rho U^\dagger \theta)$, a unitary propagator $U$ belongs to a critical submanifold $CM^\alpha$, characterized by its contingency table $K^\alpha$, if and only if each $U_{jk}$ block only has singular value(s) 1 and/or 0 and the number of 1’s is exactly $K^\alpha_{jk}$. In particular, $U_{jk}$ must be zero when $K^\alpha_{jk} = 0$.

Addressing a target CM in simulation can be achieved by defining a distance metric from the CM, which should be a function of $U$, and minimizing the distance to zero by an appropriate algorithm, here chosen as the gradient method. Based on Theorem 1, we design a saddle distance metric with the singular values of all the $U_{jk}$ blocks. Suppose a block $U_{jk}$ is singular-value decomposed as

$$U_{jk} = X_{jk} \Sigma_{jk} Y_{jk}^\dagger,$$

(5.13)

where the columns of the unitary matrices $X_{jk}$ and $Y_{jk}$ are the left- and right-singular vectors of $U_{jk}$, respectively, and the diagonal entries of the rectangular diagonal matrix $\Sigma_{jk}$ are the singular values sorted in descending order, i.e, $\sigma_{jk,1} \geq \sigma_{jk,2} \geq \cdots$. By comparing the $K^\alpha_{jk}$ largest singular values with 1 and the remaining singular values with 0, we can determine how different each block $U_{jk}$ is from its desired form in $CM^\alpha$. Thus, we define the following scalar metric $D^\alpha_O(U)$ to characterize the distance of a kinematic control $U$
from CM$^\alpha$ on the landscape:

$$D^\alpha_O(U) = s,r \sum_{j,k} \left[ \sum_{l \leq K^\alpha_{j,k}} (1 - \sigma_{j,k,l})^2 + \sum_{l > K^\alpha_{j,k}} (\sigma_{j,k,l} - 0)^2 \right]$$

$$= N - 2 s,r \sum_{j,k} \sum_{l \leq K^\alpha_{j,k}} \sigma_{j,k,l} + \sum_{j,k} \sum_{l} \sigma^2_{j,k,l}$$

$$= 2N - 2 s,r \sum_{j,k} \sum_{l \leq K^\alpha_{j,k}} \sigma_{j,k,l}$$

$$= 2 s,r \sum_{j,k} \sum_{l \leq K^\alpha_{j,k}} (1 - \sigma_{j,k,l}). \quad (5.14)$$

In the above derivation we utilized the fact that the entries of any contingency table sum to $N$ in the second and fourth lines, and that the singular values squared of all the blocks of an $N$-dimensional unitary matrix sum to its Frobenius norm squared, $N$, in the third line. The final form of (5.14) shows that for the distance metric we only need to compare the $N$ selected singular values with 1, because once they are optimized to 1 all the others have to vanish automatically to preserve the norm of the entire unitary matrix. The range of this distance metric $D^\alpha_O$ is non-trivial and depends on the degeneracies of $\rho$ and $\theta$, and the maximal allowed distances from different CMs are not necessarily identical. By definition, the distance between two critical submanifolds CM$^\alpha$ and CM$^\beta$ can be calculated with the absolute difference of their contingency tables, i.e.,

$$D^\alpha_{O} \rightarrow^\beta = \sum_{jk} |K^\alpha - K^\beta|_{jk}. \quad (5.15)$$

Therefore, the maximal distance between two arbitrary CMs over the whole landscape cannot exceed $2N$, although we have verified that for some particular control problems with degenerate $\rho$ and $\theta$, the global maximal distance between two CMs could be lower than $2N$, and the distance between two distinct saddle CMs could be larger than that between
the landscape top and bottom.

In order to construct a gradient algorithm for approaching \( \text{CM}^\alpha \) by minimizing the distance metric \( D_\alpha^O \), we need the differential of \( D_\alpha^O \) with respect to \( U \). Starting from rearranging (5.13) as \( \Sigma_{jk} = X_{jk}^\dagger U_{jk} Y_{jk} \), the differential of the \( l \)-th singular value of \( U_{jk} \), treated as an arbitrary rectangular matrix, is given by

\[
\delta \sigma_{jk,l} = \Re(X_{jk}^\dagger \delta U_{jk} Y_{jk})_{ll},
\]

(5.16)

where \( \cdot \)_{ll} is the \((l, l)\)-th entry of a matrix, and \( \Re \) means the real part. Hence, the differential of \( D_\alpha^O \) reads

\[
\delta D_\alpha^O(U) = -2 \sum_{j,k} \sum_{l \leq K_{jk}} \Re \left( X_{jk}^\dagger [\delta U]_{jk} Y_{jk} \right)_{ll},
\]

(5.17)

where \([ \cdot ]_{jk}\) is the \((j, k)\)-th block of a matrix by the formulation in (5.11).

We note that in sorting the singular values of each \( U_{jk} \) block we actually defined \( D_\alpha^O \) as a piecewise function, which may have a non-unique differential at the edge of its subfunctions. Recall the definition of the distance metric in (5.14) and consider the situation that one \( U_{jk} \) block possesses two equal singular values \( \sigma_{jk, K_{jk}} = \sigma_{jk, K_{jk}+1} \), in which \( D_\alpha^O \) has two expressions by comparing either of the two singular values to 1. The two expressions will both satisfy the definition of \( D_\alpha^O \) and give the same value, but lead to different differentials at this point. However, in practice the piecewise differentiability should not halt the gradient-based search for a target CM, as exactly landing on the edge of the subfunctions is almost impossible in discrete gradient algorithms.

To search for the control fields belonging to \( \text{CM}^\alpha \) in the dynamic picture, we further need the derivative of the distance \( D_\alpha^O \) with respect to the field \( \epsilon(t) \). Within the electric
dipole approximation, the Hamiltonian has the form [7]

\[ H(t) = H_0 - \mu \cdot \epsilon(t), \quad t \in [0, T], \]  

(5.18)

where \( H_0 \) is the internal Hamiltonian of the quantum system and \( \mu \) is the dipole moment operator. Recalling Eq. (5.1), the derivative of \( U \) with respect to \( \epsilon(t) \) for this model Hamiltonian is

\[
\frac{\delta U}{\delta \epsilon(t)} = i U(T, t) \mu U(t, 0) = i U \mu(t),
\]  

(5.19)

where the shorthand notation \( \mu(t) := U^\dagger(t, 0) \mu U(t, 0) \) is used. Substituting (5.19) into (5.17) leads to the derivative of the distance metric \( D_O^{\alpha} \) in the dynamic picture (\( \Im \) means the imaginary part):

\[
\frac{\delta D_O^{\alpha}}{\delta \epsilon(t)} = 2 \sum_{j,k}^{K^{\alpha}} \sum_{l \leq K^{\alpha}_{jk}} \Im \left( X_{jk}^\dagger [U \mu(t)]_{jk} Y_{jk} \right)_{ll}.
\]  

(5.20)

With Eq. (5.20) a random initial control \( \epsilon_0(t) \) can be optimized such that \( U(T, 0) \) moves toward a target CM of the observable by a gradient algorithm. We introduce a progress parameter \( \kappa \geq 0 \) to characterize the evolution of the field \( \epsilon(\kappa, t) \) during the optimization, and a deterministic search trajectory is given by the ordinary differential equation

\[
\frac{d}{d\kappa} \epsilon(\kappa, t) = -\frac{\delta D_O^{\alpha}}{\delta \epsilon(\kappa, t)}, \quad \epsilon(0, t) = \epsilon_0(t).
\]  

(5.21)

As \( \kappa \) proceeds, the distance \( D_O^{\alpha} \) will decrease monotonically until a critical point in the desired CM\(^\alpha \) is addressed with arbitrarily high accuracy. No suboptimal traps were encountered in our numerical simulations where the quantum system is controllable, the control field is unconstrained, and surjectivity is assumed to be satisfied (see section 5.4.2 for details), although a proof is still open about the distance metric on the observable landscape being trap-free. However, as shown below in section 5.3 we have proved that the saddle
distance metric is trap-free for the unitary transformation control landscape.

Similar to the exploration of landscape top or bottom [5, 21], the exploration of an entire CM, especially a saddle submanifold, can be achieved by starting from a single critical point and continuously moving in the Hessian null space of the distance metric $D_O^\alpha$, so that the distance remains at zero (within some tolerance) in the process. This method is also applicable to the unitary transformation landscape described below.

5.3 The unitary transformation control landscape

In parallel with the material above, here we will define the distance metric to a critical submanifold on the unitary transformation landscape and construct the saddle-seeking algorithm. For the optimal creation of a targeted unitary transformation $W$ operating on a quantum system, a kinematic landscape $J_W(U)$ may be defined as [10]

$$J_W(U) = \|W - U\|_F^2 = \text{Tr} \left[ (W - U)\dag(W - U) \right] = 2N - \text{Tr}(W\dag U + U\dag W), \quad (5.22)$$

where $U := U(T, 0)$ is generated by a control field according to Eq. (5.1), and the cost function $J_W(U)$ characterizes the difference between two unitary matrices $U$ and $W$ by their Frobenius distance squared. The desired transformation $U = W$ is located at the bottom of the landscape where $J_W = 0$, while $U = -W$ is at the landscape top where $J = 4N$. To find the necessary condition for critical points over the unitary transformation landscape, we apply the parametrization $U(s, A) = e^{isA}U$ introduced in section 5.2.1 again,

$$\frac{d}{ds} J_W[U(s, A)] \bigg|_{s=0} = i\text{Tr} \left[ A(WU\dag - UW\dag) \right] = 0, \quad \forall A\dag = A, \quad (5.23)$$

which leads to the following criterion for critical $U$’s,

$$WU\dag = UW\dag, \quad \text{or} \quad W\dag U = U\dag W, \quad (5.24)$$
i.e., \( W^\dagger U \) is Hermitian and thus has a real-valued spectrum. Since \( W^\dagger U \) is also unitary, implying that its eigenvalues have a complex modulus of 1, the eigenvalues of \( W^\dagger U \) must be either 1 or -1 when \( U \) is a critical point (all values being 1 at \( J_W = 0 \) and correspondingly -1 at \( J_W = 4N \)). Thus the trace of \( W^\dagger U \) at landscape critical points can only take the discrete values \( N, N - 2, \ldots, -N \), and the corresponding \( J_W \) values calculated by (5.22) are

\[
J_{W,\text{crit}} = 0, 4, \ldots, 4N. \tag{5.25}
\]

The landscape possesses \((N + 1)\) CMs at distinct \( J_W \) values, and Hessian analysis reveals that besides the global maximum and minimum, the remaining \((N - 1)\) CMs are all saddles [10]. Note that on the unitary transformation landscape the CMs can be characterized simply by integers instead of contingency tables. By considering the eigenvalues of \( W^\dagger U \), we have the following sufficient and necessary condition for a \( U \) belonging to a particular CM:

**Theorem 2.** On the unitary transformation landscape \( J_W(U) = \|W - U\|^2_F \), a unitary propagator \( U \) belongs to the \( \alpha \)-th critical submanifold \((\alpha = 0, 1, \ldots, N)\), which lies at \( J_{W,\text{crit}} = 4\alpha \), if and only if the spectrum of \( W^\dagger U \) consists of \( \alpha \) (-1)'s and \((N - \alpha)\) 1's.

The sufficiency of this condition is clear, as a unitary matrix \( W^\dagger U \) with only \( \pm 1 \) eigenvalues must be Hermitian and thus specifies a critical point. Previously this property was used to define a metric \( S \) expressing the distance of an arbitrary \( U \) to the nearest CM over the unitary transformation landscape [16],

\[
S(U) = \sum_{j=1}^{N} (1 - |\Re \lambda_j|),
\]

(5.26)

where \( \lambda_j \)'s are the eigenvalues of \( W^\dagger U \). In the present chapter we will further provide the distance metric with respect to any specified CM, which is described as follows. Being a
normal matrix, $W^\dagger U$ can be diagonalized by unitary matrix $V$ as

$$W^\dagger U = V \Lambda V^\dagger, \quad (5.27)$$

where the diagonal entries of $\Lambda$ are the eigenvalues of $W^\dagger U$ sorted in ascending order by their real parts, i.e., $\Re \lambda_1 \leq \cdots \leq \Re \lambda_N$. To define a distance metric $D_\alpha^W$ from the $\alpha$-th CM, we can compare the $\alpha$ smaller $\Re \lambda_j$’s with -1, and the remaining $(N - \alpha)$ larger ones with 1. Therefore,

$$D_\alpha^W(U) = \sum_{j=1}^{\alpha} |\Re \lambda_j - (-1)| + \sum_{j=\alpha+1}^{N} |\Re \lambda_j - 1|$$

$$= \sum_{j=1}^{\alpha} (1 + \Re \lambda_j) + \sum_{j=\alpha+1}^{N} (1 - \Re \lambda_j), \quad \alpha = 0, 1, \cdots, N. \quad (5.28)$$

The globally maximal distance over the landscape is $2N$, the distance between the landscape top and bottom. From a particular CM$^\alpha$ the maximal allowed distance is its distance to either the top or bottom, i.e., $D_\alpha^W \leq \max \{2\alpha, 2(N - \alpha)\}$, so the CMs lying closer to the middle of the landscape ($J_W \sim 2N$) should have lower maximal distances from them.

We also find that the distance metric $D_\alpha^W$, as defined here, is closely related to the Frobenius distance from the given unitary operator to the targeted CM$^\alpha$ on the landscape (see Appendix B).

In parallel with section 5.2.2 we derive the differential of the distance $D_\alpha^W$ with respect to $U$. Rearrangement of (5.27) gives that $\Lambda = V^\dagger W^\dagger U V$, and

$$\delta \Lambda = \delta V^\dagger W^\dagger U V + V^\dagger W^\dagger \delta U V + V^\dagger W^\dagger U \delta V = (\delta V^\dagger V) \Lambda + V^\dagger W^\dagger \delta U V + \Lambda (V^\dagger \delta V). \quad (5.29)$$

Due to the unitarity of $V$ we have $\delta(V^\dagger V) = \delta V^\dagger V + V^\dagger \delta V = 0$, and thus the diagonal entries of the matrix $(\delta V^\dagger V) \Lambda + \Lambda (V^\dagger \delta V) = [\delta V^\dagger V, \Lambda]$ will be zero. Taking the diagonal entries on both sides of (5.29) we can find that the differentials of $W^\dagger U$’s eigenvalues are
related to $\delta U$ by
\[
\delta \lambda_j = (V^\dagger W^\dagger \delta UV)_{jj}, \quad j = 1, \cdots, N. \tag{5.30}
\]
Substituting (5.30) into the differential of (5.28) gives
\[
\delta D_\alpha^W(U) = \sum_{j=1}^\alpha \Re(V^\dagger W^\dagger \delta UV)_{jj} - \sum_{j=\alpha+1}^N \Re(V^\dagger W^\dagger \delta UV)_{jj}, \quad \alpha = 0, 1, \cdots, N. \tag{5.31}
\]
Similar to the case of the observable landscape, the differential of $D_\alpha^W$ is non-unique at points where $\Re \lambda_\alpha = \Re \lambda_{\alpha+1}$, but practically it should not hinder the search for the target CM (see the discussion below (5.17)). As the Hamiltonian $H(t)$ is a function of the control field $\epsilon(t)$, we can obtain the derivative of the distance $D_\alpha^W$ in the dynamic picture as well. Applying the electric dipole approximation again, we substitute (5.19) into (5.31) to reach
\[
\begin{align*}
\frac{\delta D_\alpha^W}{\delta \epsilon(t)} &= - \sum_{j=1}^\alpha \Im(V^\dagger W^\dagger U \mu(t) V)_{jj} + \sum_{j=\alpha+1}^N \Im(V^\dagger W^\dagger U \mu(t) V)_{jj}, \quad \alpha = 0, 1, \cdots, N. \tag{5.32}
\end{align*}
\]
Eq. (5.32) can be used in the gradient descent algorithm given in (5.21) to optimize an arbitrary control field toward the $\alpha$-th CM over the unitary transformation landscape.

Like the optimization of the landscape $J_W$ in (5.22), the dynamic gradient descent algorithm with (5.32) should not encounter hindering traps provided that the assumptions in section 5.2.1 underlying landscape analysis are satisfied, because the function $D_\alpha^W(U)$ is free of local minima in the kinematic picture. To demonstrate this, we only need to prove that for any $U$ with $D_\alpha^W(U) > 0$, i.e., $U \notin \text{CM}^\alpha$, there exists $\tilde{U}$ in the neighborhood of $U$ satisfying $D_\alpha^W(\tilde{U}) < D_\alpha^W(U)$. Such a $\tilde{U}$ can be constructed in the following way. Among the eigenvalues of $W^\dagger U$, $\{\lambda_j\}_{j=1}^N$, we arbitrarily pick a $\lambda_k$, which is away from its reference value (-1 or 1) and has positive contribution to the distance, and replace it with $\tilde{\lambda}_k$ with modulus 1 to generate a new eigenspectrum:
\[
\tilde{\Lambda} = \text{diag}\{\lambda_1, \cdots, \tilde{\lambda}_k, \cdots, \lambda_N\}
\]
Let \( \tilde{\lambda}_k \) be closer to its reference value than \( \lambda_k \), i.e., \( \Re \tilde{\lambda}_k < \Re \lambda_k \) for \( k \leq \alpha \) and \( \Re \tilde{\lambda}_k > \Re \lambda_k \) for \( k > \alpha \). Define a new unitary matrix \( \tilde{U} := WV\tilde{\Lambda}V^\dagger \), and it can be easily verified from the definition of the distance metric \( D^\alpha_W \) in (5.28) that \( D^\alpha_W(\tilde{U}) < D^\alpha_W(U) \). As \( \tilde{\lambda}_k \) approaches \( \lambda_k \) along the unit circle on the complex plane, \( \tilde{U} \) will converge to \( U \) in the unitary space. Therefore, none of the kinematic matrices \( U \) can be a local minimum of \( D^\alpha_W(U) \) except the global minimum where \( D^\alpha_W = 0 \). Upon satisfaction of the three assumptions in section 5.2.1, the distance function \( D^\alpha_W[\epsilon(t)] \) in the dynamic picture is also free of suboptimal traps for minimization.

### 5.4 Illustrations

Numerical simulations for testing the distance metrics and saddle-seeking algorithms on both types of landscapes are performed as follows. First, we optimize the cost functional \( J \) from an initial control field by following the gradient of \( J \), and simultaneously monitor the evolution of the distances from each CM along the search trajectory. Second, in an additional simulation we also select a target saddle submanifold, CM\(^\alpha\), and minimize the distance from it along the gradient flow, while monitoring the evolution of \( J \) and its gradient norm to verify that a desired saddle point was found.

#### 5.4.1 Numerical schemes

A dynamic model system is constructed by the Hamiltonian form in Eq. (5.18). To allow for full controllability of the quantum system the operators \( H_0 \) and \( \mu \) should be reasonably chosen [25], and the final time \( T \) should be sufficiently large, as it is also a control resource [28, 29, 30]. Throughout section 5.4 we consider a two-qubit system with isotropic coupling, described in the computational basis set \( \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} \):

\[
H_0 = \omega_1(\sigma_z \otimes I_2) + \omega_2(I_2 \otimes \sigma_z) + J_{12}(\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z),
\]
\[ \mu = \sigma_x \otimes I_2 + I_2 \otimes \sigma_x, \]

where \( \sigma_x, \sigma_y \) and \( \sigma_z \) are the Pauli matrices and \( I_2 \) is the \( 2 \times 2 \) identity matrix. The parameter values are set to \( \omega_1 = 1, \omega_2 = 1.5, \) and \( J_{12} = 0.25. \) Random numbers obeying the standard normal distribution \( \mathcal{N}(0, 1) \) for \( \xi_l \) at each time point \( l = 1 \cdots L \) were always used as the initial controls in order to assess the capability of our algorithms.

The control field \( \epsilon(t) \) is discretized into \( L \) time intervals on \( t \in [0, T] \) with an identical length \( \Delta t = T/L, \) and thus expressed by a \( L \)-dimensional vector \( \vec{\epsilon} = (\epsilon_1, \epsilon_2, \cdots, \epsilon_L)^\tau. \) For the minimization of \( D^\alpha, \) we utilize the forward Euler method (higher order integration methods may be applicable as well when necessary) with a constant step size \( \gamma > 0 \) for solving Eq. (5.21):

\[ \epsilon_l[k+1] = \epsilon_l[k] - \gamma \frac{\partial D^\alpha}{\partial \epsilon_l} \bigg|_{\vec{\epsilon}[k]}, \quad l = 1, 2, \cdots, L, \quad (5.33) \]

where \( k = 1, 2, \cdots \) is the iteration index. The partial derivatives can be well approximated by the analytical functional derivatives at the corresponding time points when the discretization has sufficient resolution [27], i.e.,

\[ \frac{\partial D^\alpha}{\partial \epsilon_l} \approx \frac{\delta D^\alpha}{\delta \epsilon(l \Delta t)} \cdot \Delta t. \quad (5.34) \]

The functional derivatives of \( D^\alpha \) for the two landscapes are given by equations (5.20) and (5.32). Likewise, for the optimization of \( J \) we can utilize the same numerical scheme and replace \( D^\alpha \) by \( J. \) The functional derivative of \( J \) for the observable landscape is [7]

\[ \frac{\delta J_O}{\delta \epsilon(t)} = i \text{Tr} \left\{ \mu(t)[\rho, U^\dagger \theta U] \right\}, \quad (5.35) \]
and for the unitary transformation landscape [16],

\[
\frac{\delta J_W}{\delta \epsilon(t)} = -i \text{Tr} \left[ (W^\dagger U - U^\dagger W) \mu(t) \right].
\] (5.36)

In Figures 5.1 and 5.2 we stopped at the first 100 iterations of the optimization procedures demonstrating that the objective being optimized, \( J \) or \( D^\alpha \), can reach a good value within relatively few iterations. Further optimization (not shown here) can finally lower the deviation from the theoretical limits to machine precision \( \sim 10^{-15} \), usually taking \( \sim 10^3 \) iterations with the iteration step sizes shown in the figures. The details of these two examples are described in the subsections below.

5.4.2 The physical observable landscape

To illustrate the methodology in a particular observable landscape \( J_O = \text{Tr}(U \rho U^\dagger \theta) \), we choose

\[
\rho = \text{diag}\{0.5, 0.5, 0, 0\}, \quad \theta = \text{diag}\{0.5, 0, 0, 0.5\},
\]

with \( \rho \) being a separable mixed state and \( \theta \) being an entangled state. Note that the \( \theta \) here does not follow the form in (5.7) since its eigenvalues are not sorted in descending order. To facilitate the calculation of the distance metrics, we introduce the transforms that \( \tilde{\theta} := S^\dagger \theta S \) and \( \tilde{U} := S^\dagger U \), where \( S \) a unitary transformation on \( \theta \) such that \( \tilde{\theta} = \text{diag}\{0.5, 0.5, 0, 0\} \), as required by (5.7). The landscape can be rewritten as

\[
J_O(\tilde{U}) = \text{Tr}(U \rho U^\dagger S \tilde{\theta} S^\dagger) = \text{Tr}(\tilde{U} \rho \tilde{U}^\dagger \tilde{\theta}).
\] (5.37)

All the distances \( D_O \) and their derivatives will be calculated from \( \tilde{U} \), the kinematic control variable of the mathematically equivalent landscape \( J_O(\tilde{U}) \) [8]. In accordance with the degeneracies of \( \rho \) and \( \theta \), i.e., \( n_1 = n_2 = m_1 = m_2 = 2 \), the formulation of (5.11) should
divide a $4 \times 4$ unitary transformation $\tilde{U}$ into four blocks:

$$
\begin{align*}
\tilde{U}_{11} &= \begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}, & \tilde{U}_{12} &= \begin{pmatrix} u_{13} & u_{14} \\ u_{23} & u_{24} \end{pmatrix}, & \tilde{U}_{21} &= \begin{pmatrix} u_{31} & u_{32} \\ u_{41} & u_{42} \end{pmatrix}, & \tilde{U}_{22} &= \begin{pmatrix} u_{33} & u_{34} \\ u_{43} & u_{44} \end{pmatrix},
\end{align*}
$$

where $u_{jk} := \langle j | \tilde{U} | k \rangle$. The distance of a given $\tilde{U}$ from one of the three CMs, i.e., landscape top, saddle or bottom, can be calculated with the singular values of these four blocks and the contingency table of the target CM according to Eq. (5.14).

Dynamically, we use the Hamiltonian of the two-qubit system in section 5.4.1, and set the target time as $T = 10$ and discretize it into $L = 200$ time steps. With these settings we first maximized the cost function $J_O$, starting from an initial control close to the landscape bottom, by following the gradient [i.e., $\gamma \rightarrow -\gamma$ in (5.33)], and $J_O$ was monotonically driven close to its maximal value of 0.5 [see figure 5.1(a)]. Most iterations were spent in the neighborhoods of the CMs. Within each iteration the distances from all the three CMs were calculated by (5.14) and recorded simultaneously. As $J_O$ passed through the saddle landscape “altitude” of 0.25, $D_{saddle}^O$ dramatically dropped in value, indicating a nearby pass of the saddle. As the search trajectory approached the landscape top, the distances converged to $D_{top}^O = 0$, while $D_{saddle}^O = 4$ and $D_{bottom}^O = 8$ approached their respective asymptotic values, agreeing with the distances between two CMs predicted by Eq. (5.15).

Next, we searched for a control in the saddle submanifold by starting with another random initial control and minimizing its distance from the saddle, $D_{saddle}^O$, toward zero. The evolution of the cost $J_O$ and the gradient norm of $J_O$, calculated with Eq. (5.35), was also recorded [see figure 5.1(b)]. The convergence of $J_O$ to 0.25 and its gradient norm to zero indicated that the search trajectory was closely approaching a critical point at $J_O = 0.25$, which belongs to the target saddle submanifold. The distances to the other two CMs also converged to their predicted values, $D_{top}^O = D_{bottom}^O = 4$. The accuracy of critical points found by the algorithm can be assessed by invoking the critical criterion
for this landscape [c.f. (5.4)], i.e., calculating the Frobenius norm of $[U \rho U^\dagger, \theta]$, which should be strictly zero at exact kinematic critical points. We verified that the accuracy can reach double precision ($\sim 10^{-15}$) after sufficient iterations. Therefore, fields creating target saddle points on the landscape of this example can be located with arbitrarily high accuracy by the method provided here.

![Figure 5.1](image)

Figure 5.1: Critical distance metric for an observable control landscape, $J_O = \text{Tr}(U \rho U^\dagger \theta)$. (a) Optimization of the cost functional $J_O$, together with evolution of the distances to the three critical submanifolds: top, saddle and bottom, whose landscape altitudes are shown by the horizontal dashed lines. (b) Searching for the target saddle: minimization of the distance to the saddle, together with evolution of $J_O$ and the Euclidean norm of its gradient, as well as the distances to the top and bottom.

### 5.4.3 The unitary transformation landscape

For the unitary transformation landscape $J_W(U) = \|W - U\|_F^2$, again we utilize the two-qubit model system and define the target transformation $W$ as the CNOT gate with a global phase factor chosen so that $W \in \text{SU}(4)$, since the Hamiltonian is traceless [30]:

$$W = \exp(-i\pi/4) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$
Here we set $T = 20$ and discretize it into $L = 200$ time intervals. We first minimized the cost function $J_W$ from a random initial control and monitored simultaneously the distances from all the CMs [figure 5.2(a)]. This landscape possesses five CMs at $J_W$=0, 4, 8, 12 and 16, identified by the eigenvalues of $W^\dagger U$ being $(1,1,1,1)$, $(-1,1,1,1)$, $(-1,-1,1,1)$, $(-1,-1,-1,1)$ and $(-1,-1,-1,-1)$, and the distances from them [calculated by (5.28)] are denoted $D_0^W$ through $D_4^W$, respectively. With $J_W$ decreasing to zero, the distances converged to $D_0^W = 0$, $D_1^W = 2$, $D_2^W = 4$, $D_3^W = 6$, and $D_4^W = 8$, as expected by comparing the eigenspectra of $W^\dagger U$ of different CMs. The optimization trajectory for $J_W$ made a close pass by the CM labeled 1 lying at $J_W = 4$, but did not come close to the other saddles.

In figure 5.2(b) we show a process of seeking a control field in the saddle submanifold at $J_W = 8$. As the distance from the target saddle, $D_2^W$, was minimized to zero with the gradient algorithm, the cost $J_W$ converged to the expected saddle value 8 and its gradient norm, calculated with Eq. (5.36), decayed toward zero, indicating that the search trajectory was moving toward the target CM. The distances to the other four CMs also converged to their predicted values, $D_0^W = D_4^W = 4$ and $D_1^W = D_3^W = 2$. We also checked the accuracy of the critical points found on this landscape by calculating the Frobenius norm of the criterion matrix $(W^\dagger U - U^\dagger W)$ [c.f. (5.24)], and the machine precision ($\sim 10^{-15}$) could be reached as well with a sufficient number of iterations, thus verifying the effectiveness of our algorithm for the unitary transformation landscape.
Figure 5.2: Critical distance metric in a unitary transformation control landscape, $J_W = \|W - U\|_F^2$. (a) Optimization of the cost functional $J_W$ and evolution of the distances, $D_W^0$, $D_W^1$, $D_W^2$, $D_W^3$ and $D_W^4$, to the five critical submanifolds at $J_W = 0$, 4, 8, 12 and 16 (the first three shown by horizontal dashed lines), respectively. (b) Minimization of the distance $D_W^2$, with evolution of $J_W$ and the Euclidean norm of the gradient of $J_W$, as well as the distances to the other four CMs.

## 5.5 Conclusion

In this chapter we developed methods for seeking critical saddle submanifolds, corresponding to suboptimal saddle objective values, over two types of quantum control landscapes: the observable landscape $J_O = \text{Tr}(U \rho U^\dagger \theta)$ and the unitary transformation landscape $J_W = \|W - U\|_F^2$. The algorithms provided here may not be the only or most efficient way of finding a single point in a specified CM; instead, our aim is to find the necessary and sufficient condition for every individual kinematic CM and define a reasonable metric to characterize the distance from it, which could be useful for understanding global landscape structure and its impact on seeking optimal control fields (see ref. [18] for an example). In the kinematic pictures of the two landscapes, the $U$'s belonging to particular CMs have some characteristic properties, i.e., the blocks of $U$ having singular values of 1 and/or 0 for the observable landscape, or $W^\dagger U$ having eigenvalues of $\pm 1$ for the unitary transformation landscape. Based on these properties we defined appropriate distance metrics, and derived their differentials with respect to the control field. By minimizing a particular distance
metric to zero with the gradient algorithm, an arbitrary initial control can be optimized to-ward the corresponding CM. The distance metric for the unitary transformation landscape is proved free of suboptimal traps for minimization in the kinematic picture, while a sim-ilar property for the observable landscape remains a challenge to prove, although no traps were encountered in seeking to approach a specified saddle CM. The numerical examples demonstrated that the algorithms can successfully address the targeted saddle submanifolds with arbitrarily high accuracy. The concepts and methodology provided here are general and can be applied in further studies of the landscape saddle points by simulation.

The tools developed here can aid in understanding landscape saddles, including how far a current field is from a saddle and how to approach or avoid the saddles. The latter prospect opens up consideration of new or modified algorithms accelerating the search for optimal fields. For example, we can add the distances to the saddles into an adaptive cost function as a penalty for getting too close to any of the saddles during the optimization of the landscape \( J \). We can then progressively turn off each of the saddle distance terms as we “pass” them, and finally we are left with just reaching the global optimum as all of the saddles are left behind. Such algorithmic developments are enabled by this work and left to future studies for exploration.

**Appendix A. Proof of Theorem 1**

In this appendix we provide a proof of Theorem 1 in section 5.2.2. We already know from (5.12) that for a kinematic critical point, the singular values of all the \( U_{jk} \) blocks are necessarily 1 or 0, like the singular values of permutation matrix blocks. The sufficiency of this condition, i.e., that each \( U_{jk} \) block can be written in the form of (5.12) implying that the entire \( U \) can be written in the form of (5.8), is proved as follows. Recall Eq. (5.13) that
a block $U_{jk}$ is singular-value decomposed as

$$U_{jk} = X_{jk} \Sigma_{jk} Y_{jk}^\dagger.$$  

If the diagonal of $\Sigma_{jk}$ contains $K^\alpha_{jk}$ 1’s with the other entries being 0, we can re-express $U_{jk}$ as

$$U_{jk} = \sum_{l=1}^{K^\alpha_{jk}} x_{jk,l} \cdot y_{jk,l}^\dagger, \quad (A.1)$$

where $x_{jk,l}$ or $y_{jk,l}$ is a column of $X_{jk}$ or $Y_{jk}$ corresponding to the singular value 1, whose length equals the degeneracy of the $j$-th distinct eigenvalue of $\theta (m_j)$ or the $k$-th of $\rho (n_k)$, respectively. Eq. (A.1) can be rearranged to give

$$U_{jk} y_{jk,l} = x_{jk,l}, \quad l = 1, \cdots, K^\alpha_{jk}. \quad (A.2)$$

Since all the $U_{jk}$ blocks totally possess $N$ singular values of 1, we can obtain $N$ such equations as (A.2) from an entire $U$ satisfying the singular-value conditions in Theorem 1. These equations can be extended to a dimension of $N$, i.e.,

$$\begin{pmatrix}
U_{11} & \cdots & U_{1r} \\
\vdots & \ddots & \vdots \\
U_{jk} & \vdots & \cdots \ y_{jk,l} & \cdots \\
U_{s1} & \cdots & U_{sr}
\end{pmatrix}
\begin{pmatrix}
0 \\
\vdots \\
0
\end{pmatrix}
= 
\begin{pmatrix}
0 \\
\vdots \\
0
\end{pmatrix} \quad (A.3)$$

where $x_{jk,l}$ and $y_{jk,l}$ are put in the $j$-th and $k$-th blocks of the extended column vectors. The complementary entries (also column vectors) are set to 0 in the $y$ vectors, and must be 0 in the $x$ vectors because the portion $x_{jk,l}$ already has a norm of unity. The block structures (number and distribution of non-zero entries) of the extended $x$ and $y$ vectors reflect the
degeneracies of $\theta$ and $\rho$. We observe from (A.3) that $U_{j'k} y_{jk,l} = 0$ for all $j' \neq j$, indicating that $y_{j'k,l}^\dagger \cdot y_{jk,l} = 0$ (if the block $U_{j'k}$ is non-zero), i.e., all the extended $y$ vectors with the same block structure have to be orthogonal, even if they originate from different blocks of $U$. Thus, the $y$ vectors can be permuted (leading to a permutation matrix $\pi^\alpha$) and grouped to give a block-diagonal unitary matrix $Q \in U(n) = U(n_1) \times \cdots \times U(n_r)$, and similarly for the $x$ vectors, $P \in U(m) = U(m_1) \times \cdots \times U(m_s)$. Therefore, an arbitrary $U$ with all the singular values of its blocks being 1 or 0 can be expressed in the form of (5.8), so it must be a kinematic critical point on the landscape (5.2).

**Appendix B. An additional property of the distance metric**

$D_W$

In this appendix we present an interesting property of the distance metric $D_W$, defined in (5.28), for the unitary transformation landscape (5.22): for a given unitary matrix $U_0$, $D_W^\alpha(U_0)$ corresponds to the shortest Frobenius distance from $U_0$ to any point in the target critical submanifold $CM^\alpha$, i.e.,

$$D_W^\alpha(U_0) = \frac{1}{2} \min \{ \|U^\alpha - U_0\|_F^2 \}, \quad U^\alpha \in CM^\alpha. \quad (B.1)$$

**Proof.** According to Theorem 2 we can express the critical points $U^\alpha \in CM^\alpha$ as [c.f. (5.27)]

$$U^\alpha = W V \Lambda^\alpha V^\dagger,$$

where $\Lambda^\alpha$ is a diagonal matrix with $\alpha$ (-1)'s and $(N - \alpha)$ 1's as its eigenvalues. Thus,

$$\|U^\alpha - U_0\|_F^2 = 2N - 2 \text{Tr}[V \Lambda^\alpha V^\dagger (U_0^\dagger W + W^\dagger U_0)].$$

136
Exploiting the logic for the analysis of $\text{Tr}(U\rho U^\dagger\theta)$ in (5.4) and taking $\Lambda^\alpha$ and $(U_0^\dagger W + W^\dagger U_0)$ as analogues of $\rho$ and $\theta$, respectively, we know that the critical points of $\|U^\alpha - U_0\|^2_F$ as a function of $V$ exist where

$$[V\Lambda^\alpha V^\dagger, (U_0^\dagger W + W^\dagger U_0)] = 0.$$ 

Suppose the spectrum of $W^\dagger U_0$ is $\{\lambda_j\}_{j=1}^N (\Re\lambda_1 \leq \cdots \leq \Re\lambda_N)$, then the eigenvalues of $(U_0^\dagger W + W^\dagger U_0)$ will be $\{2\Re\lambda_j\}_{j=1}^N$. By the commutativity of $V\Lambda^\alpha V^\dagger$ and $(U_0^\dagger W + W^\dagger U_0)$, we know that the maximal allowed value of $\text{Tr}[V\Lambda^\alpha V^\dagger(U_0^\dagger W + W^\dagger U_0)]$ should be $2(-\Re\lambda_1 - \cdots - \Re\lambda_\alpha + \Re\lambda_{\alpha+1} + \cdots + \Re\lambda_N)$, by sorting the eigenvalues of $\Lambda^\alpha$ and $(U_0^\dagger W + W^\dagger U_0)$ and summing up their pairwise products. Therefore,

$$\min \left\{ \|U^\alpha - U_0\|^2_F \right\} = 2N - 2(-\Re\lambda_1 - \cdots - \Re\lambda_\alpha + \Re\lambda_{\alpha+1} + \cdots + \Re\lambda_N)$$

$$= 2 \left[ \sum_{j=1}^\alpha (1 + \Re\lambda_j) + \sum_{j=\alpha+1}^N (1 - \Re\lambda_j) \right]$$

$$= 2D^\alpha_W(U_0).$$

Thus we reach the conclusion in (B.1).
Bibliography


Chapter 6

The Pareto front shape in multiobservable quantum control

Many scenarios in the sciences and engineering require simultaneous optimization of multiple objective functions, which are usually conflicting or competing. In such problems the Pareto front, where none of the individual objectives can be further improved without degrading some others, shows the tradeoff relations between the competing objectives. This chapter analyzes the Pareto front shape for the problem of quantum multiobservable control, i.e., optimizing the expectation values of multiple observables in the same quantum system. Analytic and numerical results demonstrate that with commuting observables the Pareto front is a convex polytope consisting of flat segments only, while with non-commuting observables the Pareto front includes convexly curved segments. We also discuss the situations where the weighted-sum method can or cannot continuously capture the points along the Pareto front. Illustrative examples with realistic physical conditions are presented, including NMR control experiments on a $^1$H-$^{13}$C two-spin system with two commuting or non-commuting observables.
6.1 Introduction

Quantum control aims at manipulating the dynamics of quantum systems across a wide range of physical, chemical and biological applications, usually via the implementation of shaped electromagnetic fields [1]. The objective value $J$, characterizing the degree that the goal of a specific application is achieved by a control, is a function of the control resources. Theoretical analysis of the objective as a function of the control, also known as the control landscape, has revealed many interesting properties of single-objective quantum optimal control in some generic physical problems such as pure state transitions, mixed state preparation and unitary transformation creation [2, 3]. However, in many circumstances we need to consider more than one criterion simultaneously for assessing the suitability of a control [5, 4, 6]. An important example is in the discrimination of similar quantum systems, known as optimal dynamic discrimination (ODD), with applications ranging from biological molecules [7] to homonuclear spins [8], by means of their different dynamics when interacting with a suitably shaped pulse. Multiple ODD experiments are typically performed to aid in the discrimination of the species involved to determine the concentration of each [9]. Another important case of multiobjective control is time minimization while driving quantum dynamics towards some physical objective [10, 11, 12], usually for the purpose of reducing decoherence losses in quantum computation. The tradeoff between the competitive objectives of maximizing a quantum gate fidelity and minimizing the control time has been studied in a number of contexts [13, 14, 15, 16, 17, 18]. Finally, in many circumstances competing objectives can arise due to limited control resources (e.g. the pulse structure) with an impact on the level of competitive control [19].

A wide variety of practical problems in many domains inherently involve the simultaneous optimization of multiple objectives depending on a set of variables [20]. In the situation of conflicting objectives, generally there does not exist a single solution that simultaneously reaches the global optima of all individual objectives. A solution is referred
to as *Pareto optimal*, or *nondominated* [21], if none of the objectives can be improved in value without degrading some of the other objective values. Originally introduced in economics, the concept of Pareto optimality has also been applied to wide ranging physical and engineering problems, such as the design of materials [22, 23] or sensor networks [24]. The set of Pareto optimal solutions defines the *Pareto front*, which separates the feasible and infeasible regions in the space of objective values [19].

Many efforts have been made to develop universal algorithms for determining the Pareto front [25, 27, 26, 28]. Regarding quantum control applications, experimental identification of the Pareto front has been achieved with stochastic multiobjective optimization algorithms, such as the nondominated sorted genetic algorithm (NSGA-II) [21, 26] and the multiobjective covariance matrix adaptation (MO-CMA) evolutionary algorithm [29, 30], through a process of random mutation, ranking, and selection. These strategies are generally less effective at locating the Pareto front than they are at locating an optimal solution to single objective control problems [31]; locating the Pareto front calls for identifying a family of controls specifying the front, thereby increasing the complexity of the task. In numerical simulations, a variety of deterministic algorithms have been designed for approaching and tracking the Pareto front in different applications [17, 32, 31, 19]. These methods are based on the *scalarization* of a multiobjective problem, i.e., converting the problem to one of single-objective optimization, which can be achieved by some common strategies including (i) assigning different weights to the individual objectives and combining them into a single objective [32, 31], and (ii) selecting one objective as primary while treating the others as constraints [33, 34, 19].

Besides the cases described above, another important multiobjective quantum control problem concerns optimizing the expectation values of multiple observable operators $O_m$’s in the *same* quantum ensemble, which is known as *multiobservable* control and can be viewed as a generalization of the state preparation control landscape [2] regarding a single observable only. This circumstance may have significance in various scenarios such as se-
lective excitation of multiple vibrational modes while suppressing others [35]. Theoretical analysis of this problem was reported in Refs. [32, 31], and various gradient-based deterministic algorithms for identifying and tracking the Pareto front were constructed therein. Although the range of each individual objective $J_m$ can be easily calculated from the initial density matrix $\rho_0$ and the single observable $O_m$ [2], the conflicting nature of the observables can make the feasible region smaller than the direct product of individual objective value ranges in most cases. The main goal of this chapter is to analyze the shape of the feasible region, whose boundary forms the Pareto front. The results provide insights into the tradeoff relations between the individual objectives associated with different observables, and can serve as guides for designing Pareto control experiments.

The remainder of the chapter is organized as follows. Section 6.2 proves that for a set of mutually commuting observables the Pareto front in the objective space must be a *convex polytope*, and a method is provided for determining its vertices. The analysis is illustrated by three examples with distinct physical circumstances. In Section 6.3 we introduce a numerical method for identifying the Pareto front by optimizing a weighted-sum objective function, and explain why the method is incapable of continuously capturing the points in the Pareto front in the case of commuting observables. Section 6.4 provides analytic and numerical results of the Pareto front for two non-commuting observables in specific examples, revealing the distinction between situations of commuting and non-commuting observables in terms of the Pareto front shape. Nuclear magnetic resonance (NMR) control experiments performed on a two-spin system are reported in Section 6.5 to demonstrate the Pareto concepts and methods described in the chapter, and a brief conclusion is given in Section 6.6.
6.2 Pareto front shape with mutually commuting observables

In a large body of quantum optimal control problems, a $N$-level closed quantum ensemble system is manipulated by externally implemented control fields to optimize the expectation value of an observable $O$ at the target time $T$. The objective function $J$ can be defined as [2]

$$J = \text{Tr}(\rho_T O) = \text{Tr}(U \rho_0 U^\dagger O). \quad (6.1)$$

The final density matrix $\rho_T$ of the system is related to the initial state $\rho_0$ by $\rho_T = U \rho_0 U^\dagger$, with $U$ being a unitary propagator generated by the Hamiltonian involving the controls. Given $\rho_0$ and $O$, the objective $J$ can be simply treated as a function of $U$ in what is referred to as a kinematic analysis. For simplicity we will assume in the following that the quantum system is fully controllable, i.e., any unitary transformation $U$ at time $T$ can be produced by some admissible control fields [36, 37]. Theoretical analysis shows that this assumption is almost surely satisfied by randomly chosen Hamiltonians [38].

As a natural generalization of (6.1) we seek to simultaneously optimize the expectation values of $M$ distinct observables, $\{O_m\}_{m=1}^M$, in the same quantum system. Define a multi-objective optimization problem with a vector-valued objective function $\mathbf{J}$ including the $M$ scalar functions $J_m$ as

$$\mathbf{J} = (J_1, \ldots, J_M)^T, \quad J_m = \text{Tr}(U \rho_0 U^\dagger O_m). \quad (6.2)$$

When the observables are conflicting, no unitary propagator $U$ can optimize all of the $J_m$’s simultaneously. The vector space for $\mathbf{J}$ is separated into the feasible and infeasible regions by the Pareto front. This chapter will focus on the feasible region shape in two situations: either the set of observables are mutually commuting or not. The main analytic result of
the former situation is presented as follows.

**Theorem.** In the multiobservable control problem of Eq. (6.2), if the observables \( \{O_m\}_{m=1}^M \) are mutually commuting, then the feasible region of the objective function \( J \) must have the shape of a **convex polytope** (including its interior), e.g., a two-dimensional polygon or a three-dimensional polyhedron.

**Proof.** Due to the mutual commutativity of the observables, they can be expressed as diagonal matrices in the same appropriate basis. We can further set the initial density matrix \( \rho_0 \) also to be diagonal in this basis without loss of generality, guaranteed by controllability. Let the diagonal entries of \( \rho_0 \) and \( O_m \) respectively be \( \rho_0^{(j)} \) and \( O_m^{(j)} \), \( j = 1, \cdots, N \), and then an individual objective \( J_m \) reads

\[
J_m = \sum_{j=1}^N \sum_{k=1}^N \rho_0^{(k)} |U_{jk}|^2 O_m^{(j)},
\]

(6.3)

where \( U_{jk} \) is the \((j, k)\)-th entry of the unitary matrix \( U \). Define a \( N \times N \) matrix \( P \) by

\[
P_{jk} := |U_{jk}|^2 \geq 0,
\]

which can be physically interpreted as the transition probability from the basis state \( |k\rangle \) to \( |j\rangle \) by the unitary transformation \( U \). Note that for given \( \rho_0 \) and \( O_m \), \( J_m \) is a **linear** function of the real scalar variables \( P_{jk} \)'s, which are subject to \( 2N \) **linear** equality constraints resulted from the unitarity of \( U \) that each row and column of \( P \) sum to one, i.e.,

\[
\sum_{j=1}^N P_{jk} = 1, \quad \forall k = 1, \cdots, N,
\]

\[
\sum_{k=1}^N P_{jk} = 1, \quad \forall j = 1, \cdots, N,
\]

(6.4)

as well as \( N^2 \) **linear** inequality constraints:

\[
P_{jk} \geq 0, \quad \forall j, k = 1, \cdots, N.
\]

(6.5)
Thus, the optimization of $M$ objectives $J_m$’s is converted to $M$ linear programming problems, sharing the same set of variables $P_{jk}$ subject to the same constraints. For convenience we encapsulate the $N^2$ variables $P_{jk}$ into a vector $p$.

\[
p_{N(j-1)+k} := P_{jk}, \quad (6.6)
\]

and define a $M \times N^2$ matrix $C$ of the coefficients (i.e., for all $m = 1, \cdots, M$) associated with the entries of $p$ in Eq. (6.3),

\[
c_{m,N(j-1)+k} := \rho_0^{(k)} O^{(j)}_m. \quad (6.7)
\]

Hence, the vector function in Eq. (6.2) is simply a linear mapping of $p$.

\[
J = Cp. \quad (6.8)
\]

According to the linear equality and inequality constraints, respectively in Eqs. (6.4) and (6.5), the admissible domain of the variable $p$ must be a convex polytope, and so will the range of $J$ under the linear mapping represented by $C$. Q.E.D.

By Eq. (6.8), the vertices of the polytope for $J$ are just images of those for $p$ under the mapping $C$. A vector $p$ located at a vertex of its domain must contain as many zeros as possible, so that it has the lowest degree of freedom under the inequality constraints in Eq. (6.5). Such a $p$ vector corresponds to a transition probability matrix $P = \Pi$, with $\Pi$ being a permutation matrix. Thus, the individual objective values $J_m$ at the vertices of the feasible region for $J$ are given by the general form

\[
J_m = \text{Tr}(\Pi \rho_0 \Pi^\dagger O_m), \quad m = 1, \cdots, M. \quad (6.9)
\]

At the vertices, the final state $\rho_T = \Pi \rho_0 \Pi^\dagger$, which is still diagonal, commutes with each
observable \( O_m \) and thus corresponds to critical points of each individual objective \( J_m \) [2]. Note that a point satisfying the form of Eq. (6.9) does not necessarily turn out to be a vertex; it could also lie inside the polytope region. The consequences of our analysis will be illustrated by the three examples below. The theoretically predicted feasible regions in all examples were confirmed by numerical tests: large numbers of random unitary matrices \( U \) were generated by \( U = e^{iA} \) with random Hermitian matrices \( A \)'s of sufficiently large magnitudes, and the objective vector \( J(U) \) gives a feasible sampling point. All the points were found to fall within the theoretical feasible regions.

**Example 1.** Consider a coupled two-spin system \( I-S \) with \( I \) and \( S \) denoting two discernible spins-1/2. We use the operators \( I_a \) and \( S_a \) \((a = x, y, z)\) to represent the angular momenta of each spin in the \( x, y, z \) directions, defined as \( I_a := (\sigma_a/2) \otimes I_2 \) and \( S_a := I_2 \otimes (\sigma_a/2) \), where \( \sigma_a \) is a Pauli matrix and \( I_2 \) is the \( 2 \times 2 \) identity matrix. In conventional NMR experiments, the thermal equilibrium state of the system can be well approximated by \( \rho_0 = I_z + rS_z \) (i.e., the traceless portion of the density matrix), with \( 0 < |r| < 1 \) corresponding to the different gyromagnetic ratio of spins \( I \) and \( S \) [39]. For the purpose of simultaneously optimizing the signal intensities of both spins in their corresponding NMR spectra in a single measurement, we can consider the two-observable optimization of \( J_m(U) = \text{Tr}(U\rho_0 U^d O_m) \), \( m = 1, 2 \), with \( O_1 = I_x \) and \( O_2 = S_x \). By finding all the \( 4! = 24 \) permutations of the nondegenerate \( \rho_0 \), we can calculate the dual critical points of \( J_1 \) and \( J_2 \) with Eq. (6.9) and thus construct the region, as plotted in Figure 6.1 with \( r = 0.25 \). The feasible region of this problem is an octagon with eight vertices at \( (J_1 = \pm 1, J_2 = \pm r) \) and \( (J_1 = \pm r, J_2 = \pm 1) \), and the other critical points fall inside or on the boundary of the octagon. When the individual objective \( J_1 \) reaches its global optima at \( \pm 1 \), the other objective \( J_2 \) can only take values within the range \([−|r|, |r|]\), and *vice versa*. Further improvement of \( J_2 \) can be achieved via polarization transfer [39], but the value of \( J_1 \) will inevitably deviate from its optimum at the same time; e.g., the maximization of \( J_1 \) and \( J_2 \) obeys a linear tradeoff relation that \( J_1 + J_2 \leq 1 + |r| \). Experimental verification of
the analysis in this example is given in Section 6.5.

Figure 6.1: (Color online) The Pareto front of Example 1, a two-observable problem defined in a two-spin system $I-S$, which has the shape of an octagon. The initial density matrix is chosen as $\rho_0 = I_z + 0.25 S_z$, and the two commuting observables are $O_1 = I_x$ and $O_2 = S_x$. The vertices are determined by permuting the diagonal entries of $\rho_0$ and calculating the corresponding $J_1$ and $J_2$ values.

**Example 2.** The orbitals of the electron in a hydrogen atom can be described by three quantum numbers (ignoring electron spin), the principal quantum number $n = 1, 2, \cdots$, the angular momentum quantum number $l = 0, 1, \cdots, n - 1$, and the magnetic quantum number $m_l = -l, \cdots, l$. The orthonormal basis set $|n, l, m_l\rangle$ contains eigenstates of two commuting observables: (i) the orbital angular momentum squared, $L^2$, and (ii) the angular momentum in $z$ direction, $L_z$. In units where $\hbar = 1$, we have

$$L^2|n, l, m_l\rangle = l(l+1)|n, l, m_l\rangle, \quad L_z|n, l, m_l\rangle = m_l|n, l, m_l\rangle. \quad (6.10)$$

Consider dual-observable optimization of $O_1 = L^2$ and $O_2 = L_z$ with a pure initial state $\rho_0 = |1, 0, 0\rangle\langle 1, 0, 0|$, where $|1, 0, 0\rangle$ is the atomic orbital with $n = 1$ and $l = m_l = 0$. In this infinite dimensional Hilbert space the expectation values of both observables are unbounded from above. We assume that the quantum bound states of the electron are fully controllable, consistent with dipolar selection rules of a linearly polarized field and
\( \rho_0 \) (i.e., there will be a reachable set of states). In the objective space, the feasible region has vertices at \( (J_1 = l(l + 1), J_2 = \pm l) \), i.e., \( (0, 0), (2, \pm 1), (6, \pm 2), (12, \pm 3), \ldots \), as depicted in Figure 6.2. By contrast, if we treat the electron as a classical particle, the boundary of the feasible region should simply be the parabola \( \langle L^2 \rangle = \langle L_z \rangle^2 \), which is also shown in the figure. As \( l \to \infty \) which requires the electron to have a high quantum number \( n \) (almost ionized), the slope of the quantum Pareto front will become closer to its classical counterpart. This example shows that quantization can break the smoothness of a Pareto front for multiobservable control, a phenomenon rooted in the fundamental distinction between quantum and classical descriptions of the world.

Figure 6.2: (Color online) The Pareto front of Example 2 on manipulating the orbital angular momentum of the electron in a hydrogen atom. A pure initial state \( \rho_0 = |1, 0, 0\rangle\langle 1, 0, 0| \) is chosen, while the two commuting observables are \( O_1 = L^2 \) and \( O_2 = L_z \). Both quantum and classical results of the Pareto front are provided to show the distinction caused by quantization.

**Example 3.** In a four-level system, we define a three-observable optimization problem by \( \rho_0 = |1\rangle\langle 1| \) and \( O_m = |m\rangle\langle m|, m = 1, 2, 3 \). The objectives \( J_1, J_2 \) and \( J_3 \) represent the populations of levels \( |1\rangle, |2\rangle \) and \( |3\rangle \) at the final state \( \rho_T \), respectively. The feasible region in the objective space \( (J_1, J_2, J_3) \) is a tetrahedron given by the inequalities

\[
J_1 + J_2 + J_3 \leq 1, \quad J_1, J_2, J_3 \geq 0,
\]  
(6.11)
since the populations of each level must be non-negative. However, if we decrease the system dimensionality by removing the level $|4\rangle$ while leaving $\rho_0$ and $O_m$’s unchanged, the feasible region will possess three vertices only and become a two-dimensional triangle given by

$$J_1 + J_2 + J_3 = 1, \quad J_1, J_2, J_3 \geq 0.$$  \hfill (6.12)

Since a linear combination of the observables, $O_1 + O_2 + O_3$, equals the identity matrix in this specific case, the individual objectives are related by an extra equality that $J_1 + J_2 + J_3 = \text{Tr}(\rho_T) = 1$, resulting in a two-dimensional feasible region in the three-dimensional objective space (see Figure 6.3). From this example we see that for the cases of mutually commuting observables, the dimensionality of the feasible region is upper bounded by a quantity we call “effective rank” of the observable set $\{O_m\}_{m=1}^M$, defined as $\text{Rank}\{(O_1, \ldots, O_M, \mathbb{I}_N)\}-1$ in a $N$-level system.

![Figure 6.3](image)

**Figure 6.3**: (Color online) The feasible region of *Example 3*, defined by $\rho_0 = |1\rangle\langle 1|$ and three mutually commuting observables $O_m = |m\rangle\langle m|$, $m = 1, 2, 3$. If the system only has the three levels $|1\rangle$, $|2\rangle$ and $|3\rangle$, the feasible region will be the colored triangle $ABC$. Once a fourth level (or more) is added in, the feasible region will be extended to the tetrahedron $OABC$. 

152
6.3 Numerical methods for identifying the Pareto front

Generally, the feasible region of a multi-objective problem can be generated numerically with various Pareto optimization algorithms [32, 31]. This section will start from the derivation of a classic weighted-sum algorithm, and then discuss its drawback when applied to the case of mutually commuting observables as described in Section 6.2. For the multi-observable optimization problem in Eq. (6.2) we introduce \( u \) to represent a generic control. \( u \) is a vector or function defined in a linear space, which generates the unitary propagator \( U \) transforming the initial state \( \rho_0 \) to \( \rho_T \) and gives the objective values \( J_m \). For example, \( u \) could be a function of time defined on the space \( L^2[0, T] \), serving as the control field. Denote the gradient of an individual objective \( J_m \) with respect to \( u \) by \( \nabla J_m \), then the variation of \( J_m \) caused by \( \delta u \) is

\[
\delta J_m = \nabla J_m \cdot \delta u, \quad m = 1, \ldots, M. \tag{6.13}
\]

Suppose \( u^* \) is a Pareto optimal control in a smooth segment of the Pareto front. Let the normal unit vector to the Pareto front at the point \( u^* \), pointing from the feasible to the infeasible region, be \( k = (k_1, \ldots, k_M)^T \). By the definition of a Pareto front, any allowed first-order variation of the objective value vector, \( \delta \mathbf{J} = (\delta J_1, \ldots, \delta J_M)^T \), at \( u^* \) must satisfy

\[
\mathbf{k} \cdot \delta \mathbf{J} = \sum_{m=1}^{M} k_m \delta J_m = \left( \sum_{m=1}^{M} k_m \nabla J_m \right) \cdot \delta u \leq 0, \tag{6.14}
\]

otherwise the infeasible region would be encroached. In order to guarantee Eq. (6.14) for any \( \delta u \), the gradients \( \nabla J_m \)'s have to satisfy

\[
\sum_{m=1}^{M} k_m \nabla J_m = 0, \tag{6.15}
\]
i.e., at the Pareto front the gradients of the individual objectives are linearly dependent. This first-order necessary condition for a Pareto optimum is well known, and it has been derived in different manners [40, 41].

The condition enables us to scalarize the vector-valued multiobjective optimization problem (6.2) for identifying its Pareto optima. Define a new objective function $J$ as a linear weighted sum of $J_m$'s with a set of weights $w = (w_1, \cdots, w_M)^T$,

$$J = w \cdot J = \sum_{m=1}^{M} w_m J_m,$$  

then a Pareto optimum satisfying Eq. (6.15) will also have a zero gradient on $J$ if $w$ is parallel to $k$. Therefore, a Pareto optimal point of the original problem may be located by finding the maximum of $J$, in which the $J_m$'s to be maximized (minimized) should be assigned positive (negative) weights. In the framework of multiobservable control (6.2), $J$ can also be viewed as the expectation value of a new single observable $O := \sum_{m=1}^{M} w_m O_m$,

$$J = \sum_{m=1}^{M} w_m \text{Tr}(\rho_T O_m) = \text{Tr} \left[ \rho_T \left( \sum_{m=1}^{M} w_m O_m \right) \right] = \text{Tr}(\rho_T O).$$  

Intuitively, by optimizing $J$ with continuously adjusted weights one should be capable of capturing different Pareto optimal points and thus reconstruct the entire Pareto front. However, it turns out that for mutually commuting observables, the Pareto front cannot be continuously sampled by this weighted-sum method. As shown in Section 6.2, the Pareto front consists of flat segments. The normal vector $k$ points at a constant direction within each segment, and changes abruptly across the edges of the polytope region. If the vector of weights $w$ for a chosen objective function $J$ is not parallel to the normal vector of any flat segment, then $J$ can only be optimized at the locations where multiple segments join together and the normal vector is undefined. As $w$ varies continuously, the gradient optimization trajectory of $J$ will converge to one of the polytope vertices for most choices.
of \( w \), as experimentally demonstrated in Section 6.5. This behavior agrees with the finding that the weighted-sum method succeeds in obtaining points from all parts of a Pareto front only when it is \textit{convex}, but fails to work when it is flat or concave \cite{42, 43}, as illustrated in Figure 6.4. For the latter situations more complicated numerical methods may be exploited to generate evenly distributed Pareto optimal points \cite{28}, which will not be covered by this chapter.

![Figure 6.4](image)

Figure 6.4: (Color online) Searching for the Pareto front by optimizing a weighted-sum objective \( J = w_1 J_1 + w_2 J_2 \). The dashed lines depict level sets with constant \( J \) values at different choices of \( w = (w_1, w_2) \), and in optimizing \( J \) along its own gradient we move across the level sets until stopping at a level set tangent (or nearly so, in practice) to the feasible region boundary. In (b) where the Pareto front (solid line) is smooth and convex, the search will stop where the normal vector (an arrow) is parallel to \( w \), while in (a) where the Pareto front consists of flat segments, the search will always stop at the vertex if \( w \) is aligned with any direction between the two dashed arrows.

### 6.4 Pareto front shape with non-commuting observables

This section will discuss the Pareto front shape in the more general and complex cases involving non-commuting observables. The theorem in Section 6.2 no longer applies because the observables cannot be simultaneously diagonalized in the same basis, and the feasible region does not have to be a polytope. No closed-form expression for the Pareto front is found except in some special problems. However, with some numerical methods, such as random sampling in the unitary space of \( U \) and the weighted-sum algorithm in Eq. (6.17), we can still sketch the Pareto front of a particular problem. For simplicity we deal
with two non-commuting observables here, which is sufficient for the purpose of revealing the fundamental distinction of commuting and non-commuting observables in Pareto optimization.

The Pareto front shape with non-commuting observables is first illustrated within a two-level system. A two-observable optimization problem is defined as:

\[ \rho_0 = (\mathbb{1}_2 + \sigma_z)/2, \quad O_1 = \sigma_x, \quad O_2(\theta) = \cos \theta \sigma_x + \sin \theta \sigma_y, \quad (6.18) \]

where \( \sigma_x, \sigma_y \) and \( \sigma_z \) are the Pauli matrices. The observables \( O_1 \) and \( O_2 \) are non-commuting except when \( \sin \theta = 0 \). In the circumstance of a single spin-1/2, this problem can be viewed as optimization of the angular momenta in two distinct orientations separated by an angle \( \theta \). Note that any two-observable problem in a two-level system can be converted to a form like Eq. (6.18) after some rescaling and frame rotation. The final state \( \rho_T \) after the control process corresponds to a point on the Bloch sphere, so

\[ \left[ \text{Tr}(\rho_T \sigma_x) \right]^2 + \left[ \text{Tr}(\rho_T \sigma_y) \right]^2 \leq 1. \quad (6.19) \]

Representing \( \text{Tr}(\rho_T \sigma_x) \) and \( \text{Tr}(\rho_T \sigma_y) \) in terms of \( J_1 \) and \( J_2 \) (i.e., corresponding to \( O_1 \) and \( O_2 \), respectively), we obtain

\[ J_1^2 - 2 \cos \theta J_1 J_2 + J_2^2 \leq \sin^2 \theta. \quad (6.20) \]

This inequality gives a feasible region whose boundary is an ellipse in the objective space \((J_1, J_2)\), the eccentricity of which depends on \( \theta \). In the extremal cases that \( \sin \theta = 0 \) or \( \cos \theta = 0 \) the ellipse converts to a line segment \((-1 \leq J_1 \leq 1, \ J_2 = \pm J_1)\) or a circle \((J_1^2 + J_2^2 = 1)\), respectively. The feasible region shapes at three representative values of \( \theta \), 0, \( \pi/4 \) and \( \pi/2 \), are plotted in Figure 6.5, which have been verified numerically with random samplings of \( U \). For such convexly curved segments as a Pareto front, the normal vector
direction changes continuously as we move on the Pareto front. This property enables us to systematically sample the Pareto front by optimizing weighted-sum objective functions in Eq. (6.17) with varying weights, as opposed to the case of flat Pareto front segments discussed in Section 6.3.

Figure 6.5: (Color online) The Pareto front of the two-observable problem in (6.18). The observables are non-commuting except when \( \sin \theta = 0 \). As the parameter \( \theta \) varies, the Pareto front may take on the extreme shapes of a circle, an ellipse or a line segment.

Generally in higher-dimensional quantum systems, the feasible region with two non-commuting observables cannot be simplified to a form like Eq. (6.20), and the Pareto front is not necessarily an ellipse. In all the examples we numerically tested with random \( \rho_0 \) and non-commuting observables, the Pareto fronts contained convex or flat segments only, but no concave ones. In some situations the Pareto front may even be a mixture of both flat and curved segments, and a three-level illustrative example is given as follows:

\[
\rho_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad O_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad O_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \quad (6.21)
\]

The Pareto front of this problem, as shown in Figure 6.6, was generated by random samplings followed by gradient optimization of weighted-sum functions in Eq. (6.17). It con-
sists of two line segments, $\overline{AB}$ and $\overline{BC}$, and an arc $\overarc{AC}$. Point $B$ located at the global maximum of both $J_1$ and $J_2$ forms the only nondifferentiable point in the Pareto front, where the normal vector changes abruptly from $(0, 1)^\top$ to $(1, 0)^\top$. To understand this behavior we recall that a Pareto optimum of the multiobservable problem must also be a critical point of the weighted-sum function in Eq. (6.17) with some choice of $w$. The condition is equivalent to that [2]

$$
\left[ \rho_T, \sum_{m=1}^{M} w_m O_m \right] = 0, \tag{6.22}
$$

i.e., $\rho_T$ commutes with a weighted sum of the single observables for some choice of $w_m$’s. In the two-observable case, if a Pareto optimal $\rho_T$ commutes with both $O_1$ and $O_2$, the coefficients $w_1$ and $w_2$ can be arbitrary. Thus, the Pareto front will have no unique normal vector at that point, possibly resulting in nondifferentiability. For commuting $O_1$ and $O_2$ there always exist such $\rho_T$’s that commute with both observables, which are responsible for the vertices of the feasible region polytope. For non-commuting observables the existence is not guaranteed; e.g., no Hermitian matrix commutes with both $\sigma_x$ and $\sigma_y$ except the identity or zero matrix. In the three-level example of Eq. (6.21) $O_1$ and $O_2$ are non-commuting; nevertheless, there still exists an accessible $\rho_T$ (i.e., $\rho_T = \rho_0$) that commutes with both $O_1$ and $O_2$, resulting in the nondifferentiable vertex $B$ on the Pareto front.
Figure 6.6: (Color online) The Pareto front of the three-level example in Eq. (6.21) is a mixture of both flat and convexly curved segments. The non-differentiable vertex $B$ corresponds to critical points of both individual objectives $J_1$ and $J_2$, and the Pareto front does not have a unique normal direction at point $B$.

6.5 Experimental demonstration

In this section we demonstrate the theoretical analysis on the shape of Pareto front by $^1$H and $^{13}$C NMR experiments performed on a 800 MHz spectrometer. We use $^{13}$C-labeled chloroform ($^{13}$CHCl$_3$) as the sample, and denote the two coupled nuclei $^1$H and $^{13}$C as spin $I$ and $S$, respectively. Two radio frequency pulses with an equal duration of $T = 5\text{ms}$ irradiate the sample simultaneously as the control resources, whose carrier frequencies are resonant with $^1$H or $^{13}$C. The amplitudes and phases of the control pulses are modulated in the time domain for optimally driving the dynamics of the two-spin system in the desired manner. The detailed experimental setup is similar to that used in our previous works [44, 39, 45].

The initial density matrix of the system is taken as $\rho_0 = I_z + 0.25S_z$, which is proportional to the traceless portion of the thermal equilibrium state. Two cases of multiobservable control are considered: (i) two commuting observables, $O_1 = I_x$ and $O_2 = S_x$ (c.f., Example 1 in Section 6.2); and (ii) two non-commuting observables, $O_1 = I_x$ and $O_2 = I_y$. 

159
The expectation values of these observables can be characterized by the integrated areas of the doublet peak ascribed to $^1\text{H}$ or $^{13}\text{C}$ in the corresponding NMR spectrum of that isotope. In case (i) the signals of the two isotopes are acquired simultaneously by the dual detectors and then processed individually, while in case (ii) the two observables, both associated with the spin angular momentum of $^1\text{H}$ but in two orthogonal orientations (denoted by $x$ and $y$), are successively measured by $^1\text{H}$ NMR with identical control pulses and detector phases differing by $90^\circ$.

For a control that steers the quantum system from $\rho_0$ to $\rho_T$, the two objectives $J_m = \text{Tr}(\rho_T O_m), m = 1, 2$, directly read from the NMR spectra, are given in arbitrary units. Utilizing a gradient algorithm and starting from the same initial control, we maximized the weighted-sum objective function

$$ J(w_1) = w_1 J_1 + (1 - w_1) J_2, \quad 0 \leq w_1 \leq 1, \quad (6.23) $$

along its own gradient to approach the Pareto front, and attempted to locate different points on the front by varying the weight $w_1$ from 0 to 1. The maximum values of $J_1$ and $J_2$ under the experimental condition were first determined by setting $w_1$ to 1 or 0 and then normalized to 1, as the theoretical maximum values of $J_1$ or $J_2$. The evolution of individual objective values ($J_1, J_2$) in the optimization processes is displayed in Figure 6.7.

The theoretical result for the Pareto front of case (i) is given in Figure 6.1, which has eight vertices. The trajectories in Figure 6.7(a) with a relatively small weight of $J_1$ ($w_1 = 0.3, 0.4$) converged toward the vertex $A(J_1 = 0.25, J_2 = 1)$, the global maximum of $J_2$ but not $J_1$, while the trajectory with a larger weight of $J_1$ ($w_1 = 0.6$) converged to another vertex $B(J_1 = 1, J_2 = 0.25)$ at the global maximum of $J_1$. We also verified from simulation that any trajectory with $0 < w_1 < 0.5$ or $0.5 < w_1 < 1$ should also converge to vertex $A$ or $B$, respectively, since the line segment $AB$ has a constant normal direction of $(1, 1)^T$. A trajectory with $w_1 = 0.5$ could stop somewhere between $A$ and $B$ in princi-
ple because all the points in the line segment give an identical $J$ value, which equals its global maximum. Therefore, the weighted-sum method will be incapable of continuously sampling the Pareto front with commuting observables $I_x$ and $S_x$. In case (ii), the feasible region is predicted to be the disc $J_1^2 + J_2^2 \leq 1$. The five search trajectories in Figure 6.7(b) converged to different points in the Pareto front, moving gradually along the arc from the maximum of $J_1$ to the maximum of $J_2$ as $w_1$ decreases from 1 to 0. The normal vector at each identified Pareto optimal point is roughly parallel to the vector of weights $w = (w_1, 1 - w_1)^T$ as expected. The collective experimental results are in agreement with our theoretical predictions on the Pareto front shape of multiobservable optimal control with commuting and non-commuting observables, and the effectiveness of the weighted-sum algorithm.

Figure 6.7: (Color online) Experimental NMR illustration of the theoretical Pareto principles. Two-observable quantum control experiment in the two-spin system of $^{13}$CHCl$_3$ ($I = ^1$H, $S = ^{13}$C): the evolution of $(J_1, J_2)$ during the optimization of a weighted-sum objective function $J(w_1) = w_1J_1 + (1 - w_1)J_2$. (a) $O_1 = I_x$ and $O_2 = S_x$ (commuting); (b) $O_1 = I_x$ and $O_2 = I_y$ (non-commuting). The dashed lines depict a portion of the theoretical Pareto fronts in the two cases. Red dashed and blue solid lines show optimization processes of an individual objective ($J_1$ or $J_2$) and combinations of both, respectively, each dot in which represents an experimental iteration.
6.6 Conclusions

This chapter deals with the generic problem of quantum multiobservable control, which seeks to simultaneously optimize the expectation values of $M$ distinct and usually conflicting observables in the same quantum ensemble with an initial state of $\rho_0$, steered by a single unitary transformation $U$. The individual objective functions are defined by $J_m(U) = \text{Tr}(U \rho_0 U^\dagger O_m)$, $m = 1, \cdots, M$. We particularly focus on the tradeoff relations between these competitive objectives described by the shape of Pareto front, given by the feasible region boundary in the space of objective values. We proved that when the set of observables $\{O_m\}$ are mutually commuting, each individual objective can be treated as a linear programming problem subject to some equality and inequality constraints, and the Pareto front in the objective space $(J_1, \cdots, J_M)$ has the shape of a convex polytope consisting of flat segments. When non-commuting observables are involved, convexly curved segments will emerge in the Pareto front while flat segments can also be present in certain cases. These conclusions on Pareto front shapes should apply to any quantum control circumstances with unitary system dynamics. When the full controllability condition is violated, the realized feasible region may shrink to a reachable subset. The results here provide a new perspective for understanding multiobjective optimization in quantum control, and reveal an interesting distinction between commuting and non-commuting observables as a feature of quantum mechanics.

We also discussed the capability of some numerical methods in Pareto front identification, especially the strategy of optimizing a scalar objective function defined as a linear weighted sum of the individual objectives. With the weighted-sum method one can continuously capture the points along a convex Pareto front segment by varying the relative weights assigned to the $J_m$’s, but for flat segments the optimization search will almost always converge to a vertex of the feasible region, as demonstrated in NMR control experiments performed on a $^1\text{H}$.$^{13}\text{C}$ two-spin system.
Bibliography


[34] Proper inequality constraints and maximization of index vectors, J. G. Lin, J. Optim. Theory Appl. 21, 505-521 (1977)


[42] Multicriteria truss optimization, J. Koski, in Ref. [20].


Chapter 7

Inherently trap-free convex landscapes for the full quantum optimal control

We present a fundamental analysis of quantum control landscapes as the principle to understand simulations and experiments. It is shown that the landscapes are unequivocally convex for the expectation value of an observable of one quantum system, which is controlled by another quantum system. The landscape globally optimal control can be efficiently determined, as shown in two illustrations including a comparison with a standard semiclassical solution (the quantum system controlled with classical fields). The findings show that the saddles and the rare local traps in semiclassical landscapes are attributed to the latter approximation.

7.1 Introduction

A large body of quantum optimal control simulations and experiments [1, 2] involve tailored electromagnetic fields acting on atoms or molecules to control a specified physical and/or chemical process, frequently accompanied by optimization of the performance of the control field [3, 4]. Aiming to explain the widely observed experimental successes
along with much larger numbers of almost perfect optimal control simulations, rigorous analyses were made in the past decade on the quantum control landscape, herein defined as the expectation value of a desired system observable as a function(al) of the control [5]. The landscape topology is important for establishing the feasibility of finding globally optimal controls, especially when myopic algorithms (e.g., gradient ascent/descent) are employed. A central concern is the properties of control landscape critical points, where the landscape gradient with respect to the control is zero. Of particular importance is the appearance of suboptimal local extrema as traps, which could halt a gradient search and prevent reaching the global optimum.

Most control studies, as well as the associated landscape analyses, are expressed in the semiclassical framework, i.e., a closed quantum system controlled by a classical field. It has been proved that with unconstrained classical control fields, the landscape for a controllable finite-level quantum system is in general free of any local suboptima, with all critical points being the global maximum or minimum, or possibly saddle points [5]. Inclusion of classical field constraints, for example, on pulse amplitude, frequency bandwidth and pulse duration, may result in additional topological features on the landscape, including possibly traps [6]. The origins of the rare traps without field constraints and the frequent presence of saddles have remained a mystery. This chapter shows unequivocally that these latter control landscape features are artifacts of adopting the semiclassical framework for the dynamics driven by a classical field. This fundamental conclusion will be drawn upon lifting any remnant of classical considerations.

Specifically, we analyze the control landscape by taking a full quantum (FQ) perspective wherein a quantum system $A$ (the target) is controlled by another quantum system $B$ (the control), which can be arbitrarily manipulated by suitable means [7, 8, 9, 10, 11, 12]. In principle, both quantum systems may take many forms including quantized radiation fields in an optical cavity [13, 14, 15, 16] as demonstrated by few-photon pulses in a recent experiment [17], bulk material systems, spins, etc., forming a diverse picture of FQ control.
scenarios. In this work the control is realized by the initial density matrix of the quantized system $B$ coupled with the target system $A$, in contrast to the semiclassical perspective of time-domain fields serving as the controls directly acting on quantum system $A$. We will demonstrate that the landscape expressed in the FQ context is convex, thus rigorously free of any local traps or saddles without requiring any auxiliary assumptions. Furthermore, we will show that the FQ optimal solution can be directly calculated from the eigenvectors of a specified operator matrix encapsulating all relevant physical information. The landscape study in this work subsumes so-called incoherent control [11], and more importantly expounds on semiclassical control.

In the semiclassical (sc) framework, the kinematic control landscape of the observable $O_A$ for a closed $N$-level quantum system $A$ controlled by an external classical field $u(t)$ over the duration $T$ has the cost function

$$J_{sc}[U_{sc}(T)] = \text{Tr}[U_{sc}(T)\rho_A(0)U_{sc}^\dagger(T)O_A],$$  \hspace{1cm} (7.1)$$

where the propagator $U_{sc}(T)$ belongs to the unitary group $U(N)$ designated as the kinematic control space, and $\rho_A(0)$ is the initial density matrix of $A$. It has been shown that the topology of $J_{sc}[U_{sc}(T)]$ is equivalent to that of the corresponding dynamical landscape $J_{sc}[u(t)]$, as a functional of the control field $u(t)$, upon satisfaction of three assumptions: (i) the system is controllable [18], (ii) the local mapping $\delta u(t) \mapsto \delta U_{sc}(T)$ bridging the kinematic and dynamical landscapes is surjective at any control [19, ?, ?], and (iii) the control fields $u(t)$ are unconstrained, or in practice have sufficient freedom to exploit conditions (i) and (ii) [6]. The landscape topology based on these sufficient assumptions have been the focus of many previous studies [5]. We remark that assumptions (i) and (ii) can be shown as being “almost always” satisfied [22, 23], consistent with the semiclassical landscape rarely exhibiting traps, but at most only non-trapping saddles.  

170
7.2 Full quantum landscape and its optimal solution

In the FQ setting, the total Hamiltonian of the composite system $A/B$ is time-independent and can be generally expressed as

$$H_{AB} = H_A^0 \otimes I_B + I_A \otimes H_B^0 + \sum_k H_A^k \otimes H_B^k,$$

(7.2)

where $I_A$ ($I_B$) is the identity operator in the Hilbert space of $A$ ($B$), and $H_A^0$ ($H_B^0$) is the respective uncoupled Hamiltonian of $A$ ($B$). The terms $H_A^k$ and $H_B^k$ are the interaction Hamiltonians associated with $A$ and $B$, respectively. Assuming the composite system $A/B$ to be closed, the total unitary propagator produced by the constant Hamiltonian is simply $U_{AB}(t) = \exp(-i/\hbar H_{AB}t)$. One way of making the semiclassical approximation from this formulation is to treat $H_A^k$ as the control Hamiltonians of the system $A$, and the expectation values of $H_B^k$ in the control $B$ as the corresponding classical fields [24].

The control objective, parallel to that of the semiclassical case, aims to optimize the expectation value of an observable $O_A$ at time $T$, initially given the state of $A$ as $\rho_A$. The composite system $A/B$ initially may be prepared (i) in a separable state $\rho = \rho_A \otimes \rho_B$ at $t = 0$, with $\rho_A$ being a counterpart of $\rho_A(0)$ in Eq. (7.1) and $\rho_B$ the initial state of the control $B$; or (ii) in an entangled initial state $\rho$ with the additional constraint that $\text{Tr}_B(\rho) = \rho_A$, i.e., the reduced density matrix of $A$ is still $\rho_A$. Once the initial state is created, $A$ and $B$ will freely evolve together under the total Hamiltonian $H_{AB}$ and generally become entangled.

The control scheme analyzed here will be feasible if the control $B$ can be prepared at a specified quantum state [25, 26], prior to its interaction with the target $A$.

In particular, for a separable initial state $\rho_A \otimes \rho_B$, the landscape in the FQ formulation can be defined as

$$J_Q[\rho_B] = \text{Tr}[U_{AB}(T)(\rho_A \otimes \rho_B)U_{AB}^\dagger(T)(O_A \otimes I_B)],$$

(7.3)
where the control $\rho_B$ is a positive semidefinite matrix of trace one. Optimization of the corresponding landscape can be posed as a semidefinite programming problem:

$$\begin{align*}
\text{max/min} & \quad J_Q[\rho_B] \\
\text{subject to:} & \quad \text{Tr}\rho_B = 1, \rho_B \succeq 0.
\end{align*} \tag{7.4}$$

From Eqs. (7.3) and (7.4), we note that (1) the cost function $J_Q$ is linear with respect to the control variable $\rho_B$, thus both convex and concave, and (2) the admissible set of $\rho_B$ is a closed convex set [27]. With these two conditions, and from the theory of convex optimization [28], it can be shown that the landscape $J_Q[\rho_B]$ is trap-free, i.e., a local maximum or minimum must also be a global one, if the admissible $\rho_B$ form a convex set. Specifically, for a local minimum $\rho_B^*$ of the function $J_Q[\rho_B]$ and an arbitrary $\rho_B' \neq \rho_B^*$, the convex combination $(1 - \lambda)\rho_B^* + \lambda\rho_B'$ with $\lambda \to 0_+$ is within the neighborhood of $\rho_B^*$, thus

$$J_Q[\rho_B^*] \leq J_Q[(1 - \lambda)\rho_B^* + \lambda\rho_B'] = (1 - \lambda)J_Q[\rho_B^*] + \lambda J_Q[\rho_B'], \tag{7.5}$$

which immediately leads to the relation $J_Q[\rho_B^*] \leq J_Q[\rho_B']$, so $\rho_B^*$ must also be a global minimum. Similarly it can be proved that a local maximum of $J_Q[\rho_B]$ is also a global maximum.

A level set of a control landscape is defined by the set of all controls with an identical cost function value, whose topology (especially the connectivity) was studied in the semiclassical formulation in Ref [29]. By the linearity of the function $J_Q$, we further find that all level sets of the FQ landscape are connected, as shown below. Given two initial states $\rho_{B,1}$ and $\rho_{B,2}$ of the control $B$ which are on the same level set of the FQ landscape $J_Q$, i.e., $J_Q[\rho_{B,1}] = J_Q[\rho_{B,2}] = J_0$, then any convex combination of $\rho_{B,1}$ and $\rho_{B,2}$ will also be on
the level set at $J_0$ since $J_Q$ is a linear function of $\rho_B$,

$$J_Q[\lambda \rho_B,1+(1-\lambda)\rho_B,2] = \lambda J_Q[\rho_B,1]+(1-\lambda)J_Q[\rho_B,2] = \lambda J_0+(1-\lambda)J_0 = J_0, \quad \lambda \in [0,1].$$

Therefore, the level set $\{\rho_B | J_Q[\rho_B] = J_0\}$ for any reachable $J_0$ value must be a convex set, and thus connected.

The simple convexity of the control landscape $J_Q$, Eq. (7.3), in the FQ formulation is in sharp contrast to the semiclassical counterpart $J_{sc}$, Eq. (7.1), which is a highly nonlinear functional of the control field $u(t)$. The control landscape features in these two different formulations are summarized in Table 7.1. Note that there is no controllability [11, 12] requirement for the FQ landscape to be trap-free, which is unnecessary for convexity of the admissible set of the control $\rho_B$, or the cost function $J_Q$. However, the forms of Hamiltonians $H_A^k$ and $H_B^k$ in Eq. (7.2), the flexibility in creating $\rho_B$, as well as the length of final time $T$ can influence the optimal value of $J_Q$ reachable in the FQ control scheme.

<table>
<thead>
<tr>
<th>Formulation</th>
<th>Semiclassical</th>
<th>Full quantum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nature of control</td>
<td>$u(t)$</td>
<td>$\rho_B$</td>
</tr>
<tr>
<td>Cost function</td>
<td>nonlinear</td>
<td>linear</td>
</tr>
<tr>
<td>Landscape topology</td>
<td>trap-free†</td>
<td>convex, trap-free*</td>
</tr>
</tbody>
</table>

Table 7.1: Summary of the control landscapes in two different formulations.
† Trap-free upon satisfaction of three key assumptions, and possibly with saddles. * No saddles or other suboptimal critical points present.

A complete optimal solution for the landscape (7.4) in the FQ formulation can be obtained by recasting Eq. (7.3) as

$$J_Q[\rho_B] = \text{Tr}(\rho_B \mathcal{O}_B)$$

(7.6)

where the partial trace

$$\mathcal{O}_B := \text{Tr}_A[U_{AB}^\dagger(T)(O_A \otimes I_B)U_{AB}(T)(\rho_A \otimes I_B)]$$

(7.7)
gives rise to the *landscape observable* associated with the control landscape \( J_Q[\rho_B] \) in the Hilbert space spanned by the density matrix \( \rho_B \). \( O_B \) plays the role of an “observer” enabling the information about \( A \) to be extracted [30]. All the necessary information, \( \rho_A, O_A \) and \( U_{AB}(T) \), for identifying the landscape optimum resides in the single operator \( O_B \). The upper and lower bounds of \( J_Q[\rho_B] \) in Eq. (7.6) can be given in terms of the eigenvalues of \( O_B \), i.e., \( O_B^{\min} \leq \text{Tr}(\rho_B O_B) \leq O_B^{\max} \), where \( O_B^{\max} \) and \( O_B^{\min} \) are the maximal and minimal eigenvalues of \( O_B \), respectively. To reach the global maximum (minimum) of \( J_Q, \rho_B \) must be composed of the eigenstate(s) of \( O_B \) corresponding to its maximal (minimal) eigenvalue. For a general degenerate eigenvalue \( O_B^* \) (* stands for max or min), the landscape optimal solutions

\[
\rho_B^* = \sum_i p_i |i\rangle\langle i|, \quad p_i \geq 0, \quad \sum_i p_i = 1, \tag{7.8}
\]

form a convex set of mixed states, with \{\(|i\rangle\langle i|\)\} being the subspace of degenerate eigenstates of \( O_B \) associated with \( O_B^* \), thus leading to the optimal cost function value \( J_Q[\rho_B^*] = O_B^* \). If \( O_B^* \) is nondegenerate, the optimal control \( \rho_B^* = |i\rangle\langle i| \) can only be a pure state, i.e., an extremal point on the boundary of the admissible set. We remark that in general there are infinitely many distinct control fields at the semiclassical dynamical landscape optimum, which require identification by deterministic or stochastic searching algorithms [31]. To the contrary, the optimal solution \( \rho_B^* \) of the FQ control landscape \( J_Q \), Eq. (7.3), can be explicitly determined directly from Eq. (7.8).

As an alternative of the main problem in Eqs. (7.3) and (7.4), if the initial state \( \rho \) of \( A/B \) is entangled and cannot be separated as \( \rho_A \otimes \rho_B \), the landscape in the FQ framework can instead be formulated as

\[
\begin{align*}
\max/\min \quad J_Q[\rho] &= \text{Tr}[U_{AB}(T)\rho U_{AB}^\dagger(T)(O_A \otimes I_B)] \\
\text{subject to:} \quad &\text{Tr}_B(\rho) = \rho_A, \quad \text{Tr} \rho = 1, \quad \rho \succeq 0. \tag{7.9}
\end{align*}
\]
It can be easily verified that this circumstance also entails a convex optimization problem, being a linear cost function in a convex admissible set of $\rho$, and thus the landscape $J_Q[\rho]$ is also free of local traps.

### 7.3 Illustrations

In the following we consider two illustrations of these solutions including a comparison with the semiclassical analog situation in one example. First, we consider the control of a two-level target system $A$ by another two-level control system $B$ (for example, two spins with one serving as $A$ and the other as $B$). The total Hamiltonian ($\hbar = 1$) is given by

\[
H_{AB} = \frac{\omega_A}{2} Z_A \otimes I_B + \frac{\omega_B}{2} I_A \otimes Z_B + H_{\text{int}},
\]

\[
H_{\text{int}} = X_A \otimes X_B + Y_A \otimes Y_B + Z_A \otimes Z_B \tag{7.10}
\]

where $X$, $Y$, and $Z$ denote the Pauli matrices on $A$ or $B$, $\omega_A$ and $\omega_B$ are transition frequencies of $A$ and $B$, respectively, set to $\omega_A = 2$ and $\omega_B = 3$ here. The model includes an isotropic coupling operator $H_{\text{int}}$ with three interaction Hamiltonian terms corresponding to the $x$, $y$, and $z$ directions. We choose a mixed initial state $\rho_A = (I_A/2 + Z_A/4)$ to specify a landscape problem regarding the target system $A$, and we consider the target observable $O_A = X_A$. Here $\rho_A$ could naturally occur, for example, in NMR with a thermalized spin in a static magnetic field along the $z$-axis.

The bounds of the landscape $J_Q[\rho_B]$ are calculated at different target times $T$ using Eq. (7.7) as depicted in Fig. 7.1(a). The extremal cost function values, $J_Q^{\text{max}}(T)$ and $J_Q^{\text{min}}(T)$, expand and contract as $T$ increases. The optimal solutions $\rho_B^*$ obtained from Eq. (7.8) at the landscape maximum or minimum vary smoothly, and their forms are exchanged at particular $T$ values where $J_Q^{\text{max}}$ and $J_Q^{\text{min}}$ meet at $J_Q = 0$. We also observed that at no time $T$ can $J_Q$ reach $\pm 1$ (the extremal eigenvalues of $O_A$) with the particular features specifying
the present composite system. This is consistent with the theoretical conclusion that with the initial density matrix as the control resource, full controllability can only be achieved by some special Hamiltonians [11].

As an example, we calculate the unique $\rho_B^*$ that maximizes $J_Q$ for $T = 0.78$ [a local maximum in the curve $J_Q^{\text{max}}(T)$], which is a pure state, and propagate the dynamics of $A/B$ from the initial state $\rho_A \otimes \rho_B^*$ under the Hamiltonian $H_{AB}$ during the time interval $[0, T = 0.78]$. A counterpart of Eq. (7.10) in the semiclassical picture involves three classical control fields $u_x, u_y, u_z$ such that the semiclassical Hamiltonian acting on $A$ reads

$$H_{sc}(t) = \frac{\omega_A}{2} Z_A + u_x(t) X_A + u_y(t) Y_A + u_z(t) Z_A.$$  (7.11)

A particular set of optimal classical fields (among many other possible solutions) for the same time interval $[0, T = 0.78]$ were obtained by optimizing arbitrary initial guesses with a gradient algorithm [5], corresponding to the semiclassical landscape maximum at $J_{sc}^{\text{max}} = 0.5$ [32]. Fig. 7.1(b) displays the time evolution of $\rho_A(t)$ in the FQ picture, given by $\rho_A(t) = \text{Tr}_B[U_{AB}(t)(\rho_A \otimes \rho_B^*)U_{AB}^\dagger(t)]$, and in the semiclassical picture, leading to their respective maximal cost function values $J_Q \simeq 0.66$ and $J_{sc} = 0.5$. The reference Bloch sphere in the figure has a radius of $1/2$, which is the Bloch vector length of the mixed state $\rho_A$ at $t = 0$. The semiclassical trajectory stays on the sphere since we have a closed system $A$ driven by classical fields. In contrast, the FQ trajectory can deviate beyond the sphere since we now have a quantum system $A$ driven by another quantum system $B$ [7].

Second, we consider the Jaynes-Cummings (JC) model [33], which describes a two-level atom (the target $A$) with a ground state $|g\rangle$ and an excited state $|e\rangle$, interacting with a quantized radiation field (the control $B$) containing a single bosonic mode with countably infinite number states $|n\rangle$, $n = 0, 1, \ldots$. In the rotating wave approximation, the total
Figure 7.1: (color online). (a) Upper and lower bounds $J_Q^{\text{max}}(T)$ (solid curve) and $J_Q^{\text{min}}(T)$ (dashed curve) of the landscape $J_Q$ at different target times $T$ in the FQ formulation, compared to the semiclassical bounds $J_{sc} = \pm 0.5$ shown by the dashed horizontal lines. The case at $T = 0.78$ marked by the red arrow is further analyzed below. (b) The target system’s density matrix evolution, $\rho_A(t)$, within the optimal control process in the FQ and semiclassical formulation for $T = 0.78$, leading to $J_Q^{\text{max}} \simeq 0.97$ and $J_{sc}^{\text{max}} = 0.5$. Both trajectories start from the mixed initial state $\rho_A = \frac{1}{2} I_A + \frac{1}{4} Z_A$, the north pole of the sphere with a radius of 1/2. We see that $\rho_{A,Q}(t)$ goes beyond the original sphere while $\rho_{A,sc}(t)$ stays on the sphere.
Hamiltonian is written as

$$H_{AB} = \frac{\omega}{2} \sigma_z + \nu a^\dagger a + \frac{\Omega}{2} (\sigma_+ a + \sigma_- a^\dagger)$$  \hspace{1cm} (7.12)$$

where $\omega$ and $\nu$ are the frequencies of the atom and the field, respectively, and $\Omega$ is the coupling strength. $a^\dagger$ and $a$ are the creation and annihilation operators of the field, while $\sigma_+ = |e\rangle\langle g|$, $\sigma_- = |g\rangle\langle e|$, and $\sigma_z = |e\rangle\langle e| - |g\rangle\langle g|$ are operators of the atom. Here we consider a control landscape with the form of Eq. (7.4), with the initial state of the quantized field as the control $\rho_B$ utilized to optimize the transition probability from the ground to the excited state in the atom, i.e., we specify that $\rho_A = |g\rangle\langle g|$ and $O_A = |e\rangle\langle e|$. Using Eq. (7.7), the resultant $O_B$ is a diagonal matrix with $\langle 0|O_B|0\rangle = 0$ and

$$\langle n|O_B|n\rangle = \sin^2 \alpha_n \sin^2 \left[\frac{T}{2} \sqrt{\Delta^2 + \Omega^2 n}\right]$$ \hspace{1cm} (7.13)$$

for $n = 1, 2, \cdots$, where $\Delta = \nu - \omega$ is the detuning and $\alpha_n := -\tan^{-1}(\frac{\Omega}{\Delta})$. The eigenvalues of $O_B$ are distributed within the interval $[0, 1]$, the maximum and minimum among which will determine the range of the landscape $J_Q$. In the on-resonance case of $\Delta = 0$ and thus $\sin^2 \alpha_n = 1$, for any $T > 0$ there exists some $n$ such that $\sin^2 \left[\frac{T}{2} \sqrt{\Delta^2 + \Omega^2 n}\right]$ approaches 1, and the matrix $O_B$ has an eigenvalue of 1, which means that full transition from $|g\rangle$ to $|e\rangle$ can be accomplished by the control $\rho_B^* = |n\rangle\langle n|$. In the off-resonance case that $\Delta \neq 0$, however, we observe that the upper bound for the eigenvalues of $O_B$ is always less than 1 for any finite $n$, since $|\alpha_n| < \pi/2$. Therefore, the full transition may only be asymptotically approached in the limit that $n \to \infty$, i.e., at infinite field strength. In numerical simulations we truncated the first $N_B$ levels of the quantized field, $|n\rangle$ with $n = 0, 1, \cdots, N_B - 1$, and calculated the bounds of the landscape $J_Q$ at different $T$ (see Fig. 7.2). The parameters are set to $\Omega = 0.2$ and $\Delta = 0$ or 0.1 to represent the on- or off-resonance cases. The restriction on the control space makes the full state transition
Figure 7.2: (color online). State transition $|g\rangle \rightarrow |e\rangle$ in a truncated JC model: upper bound curves of a FQ control landscape $J_Q$ at different $T$, with first $N_B$ levels of the quantized field selected as the control resource. The frequency detuning $\Delta = \nu - \omega$ is set to (a) $\Delta = 0$ (on-resonance) and (b) $\Delta = 0.1$ (off-resonance). Regardless of the value of $J_Q$ the landscape is trap free, and the optimal value of the control $\rho_B^*$ may be readily identified, as explained in the text. The lower bound for $J_Q$ is always zero, shown as the flat line.

unreachable for most values of the time $T$, but the impact becomes less significant as $N_B$ increases from 4, 8 to 16. A resonant quantized field exhibits better performance than a detuned one at the same level of truncation. The sharp changes of $J_Q$ evident in Fig. 7.2 arise from $\rho_B^*$ jumping from one solution $|n\rangle\langle n|$ to another $|n'\rangle\langle n'|$ as $T$ varies.

A semiclassical counterpart of Eq. (7.12) can be formulated as

$$H_{sc}(t) = \frac{\omega}{2} \sigma_z + \frac{1}{2}[u(t)\sigma_+ + u^*(t)\sigma_-]$$  \hspace{1cm} (7.14)$$

where $u(t)$ is a complex-valued classical field. The landscape maximum $J_{sc}^{\text{max}} = 1$ can be achieved as well with unconstrained $u(t)$. No saddles are present in this two-level single particle case for system $A$, but multi-level semiclassical cases will generally have landscape saddles; in contrast the FQ landscape as strictly convex is always free of hindering saddles.
7.4 Conclusions

In conclusion, this chapter provides a full quantum-quantum formulation for the control landscape aiming to optimize the expectation value of an observable in the target system. The target $A$ and the control $B$ are both treated quantum mechanically, which together undergo free evolution governed by a constant total Hamiltonian of the coupled bipartite system. The control consists of the initial density matrix $\rho_B$ of $B$, which would be prepared by any available means. Within this framework, optimization over the landscape $J_Q[\rho_B]$ with respect to the density matrix $\rho_B$ presents a convex problem with a convex admissible set of controls. Therefore, the full quantum control landscape is rigorously free of any suboptimal local extrema as either traps or saddles, if no additional constraints are imposed on $\rho_B$ to violate its convexity. The mathematical simplicity of the full quantum control problem permits finding the landscape optimal solutions $\rho_B^*$, and we show that the landscape optimum can always be achieved by some pure (or mixed, as appropriate) initial state $\rho_B^*$ of the control. The conclusions here imply that the rare events of local traps on the landscapes reported before [20, 21], as well as the common presence of saddle points [34], are technical artifacts in the semiclassical formulation, but not intrinsic phenomena in the broad subject of controlling quantum dynamical systems.

More generally, both the target system and the control can in principle be treated either classically or quantum mechanically, giving a tetrad of quantum-quantum, quantum-classical, classical-quantum, and classical-classical physical pictures. The characteristic control landscape in all these frameworks, rooted in the fundamental differences of quantum and classical mechanics, is an area of theoretical and practical importance, which has only been partially explored to date [35]. The present FQ landscape analysis provides a foundation for future research to draw together classical and quantum mechanical control in a seamless fashion [15].
Bibliography


The condition (2) can be understood as follows: a linear combination of arbitrary density matrices \( \{ \rho_B^{(k)} \} \), \( \varrho_B = \sum_k \lambda_k \rho_B^{(k)} \) with nonnegative coefficients \( \lambda_k \) summed to 1, must satisfy that \( \text{Tr} \varrho_B = 1 \) and \( \varrho_B \succeq 0 \), thus still being a physically allowed density matrix. The convexity of \( \rho_B \) may be violated by additional restrictions on the available initial states of the control \( B \), whose analysis is left for future studies.

Chapter 8

Future aspects of quantum control and landscape analysis

The technologies based on quantum control have the potential for truly revolutionary innovation. Quantum optimal control applications roughly fall into two categories: applications to chemistry-related areas, and applications to novel quantum technologies such as quantum computation and communication. Particularly for spin resonance, pulse sequences designed with optimal control methods have already been implemented in commercial NMR spectrometers, and the application in MRI is being pursued. Other chemical applications of coherent control include variants of spectroscopy as well as chemical analysis, where shaped laser pulses can improve the resolution and enhance the selectivity to a particular species. On the other hand, there has been an emerging industrial effort in the realization of quantum computation, e.g., by the development of superconducting qubits using optimal control techniques [1, 2]. Further industrial perspectives will be linked to development of quantum technologies [3]. In this final chapter of the dissertation, some future aspects of quantum control will be speculated, including its theory, methodology, and application, as well as the control landscape analysis as an effective tool for assessing its feasibility and complexity.
8.1 Theoretical issues

Controllability is always a primary concern in the theoretical analysis of quantum control. For closed quantum systems, although the controllability of Hamiltonians with a discrete spectrum has been solved by Lie algebra methods [4], the exact or approximate controllability of the Schrödinger equation with continuous or mixed spectrum is still an open question, which is important also in practice, since it covers dissociation and ionization processes being controlled by light fields in chemical physics. A bigger challenge is to better understand controllability in open quantum systems. For open systems with a discrete spectrum that undergo Markovian (i.e., memoryless) evolution, a rigorous understanding may be pursued by extending the standard Lie algebra techniques to non-unitary evolution, which does not apply to the controllability analysis for open quantum systems with non-Markovian dynamics, however. Beyond controllability is the issue of reachability, i.e., the reachable set of final states for a given control time, which is important for time optimal control [5]. Although the reachable set can be explored numerically in specific problems, a universal description is still lacking.

For the numerical simulation of controlled quantum dynamics, integration of the differential equations is a crucial issue that has only been partially addressed. In order to make the theory more useful for specific experiments, the gap between theory and experiments needs to be bridged. Despite recent progress, a systematic and efficient algorithm addressing this issue is still lacking. Improving the computational speed will be an important prerequisite to solve the complex control problems representing realistic experimental settings, toward the final goal of problem-specific design of controls in time-critical applications. For the optimization of control, the global optimization techniques that search for the global optimum are computationally expensive and generally do not perform very well. A rigorous understanding of the control landscape structure of specific applications will assist in the choice and development of algorithms.
Another novel area is the feedback control based on quantum measurement [6]. The quantum mechanical nature of systems makes them susceptible to measurement back-action. Thus, in the quantum domain, finding an experiment-class adapted balance between open-loop control techniques and closed-loop feedback counterparts is one of the major challenges. Most developments of quantum feedback control have been made in the area of quantum optics [7], and an emerging objective is to understand if and how the current quantum control techniques can be applied to hybrid systems involving quantum dots, superconducting qubits and opto-mechanical resonators. The existing quantum feedback theory has to be adapted to these new types of dynamics [3].

### 8.2 Quantum control in magnetic resonance

From the perspective of control sciences and engineering, further development of NMR control may proceed in the following directions [8]:

**Incoherent control of NMR.** Coherent control of NMR systems via radiofrequency pulses is a straightforward and effective approach. However, many quantum control tasks cannot be achieved solely by coherent control methods, for which incoherent control resources need to be introduced. The most important tools include quantum measurement, pulse gradient, phase cycling, and relaxation. The mathematical models for these decoherence effects are already available, which pave the way for developing more precise and effective incoherent (or mixed) control methods in the future.

**Robust control of NMR systems.** The controlled dynamics of quantum systems in the real world will inevitably involve some uncertainties, including perturbations and noises. For example, the suppression of the inhomogeneities in the static and radiofrequency magnetic fields has long been an important topic in NMR spectroscopy. The impact of uncertainties can be especially significant in large-scale spin systems, and robust control is necessary for realizing higher control precision.
**NMR feedback control.** The concept of feedback plays an important role in classical control theory, and systematic feedback control methods are also highly desirable for the quantum applications. Previous works on NMR feedback control usually focused on specific control tasks, while some fundamental questions are still unsolved, e.g., the impact of measurement on the quantum system, and the characterization of complexity in quantum feedback control.

**Manipulation of more complex sample systems.** Currently NMR quantum control is mostly restricted to liquid-state samples, and precise manipulation of more complex samples can be considered, including liquid-crystal [9], high spin, and solid-state samples.

Regarding optimal control, important fundamental topics in magnetic resonance that remain largely unsolved include the physical limits of quantum dynamics, such as quantum state reachability in open systems, and a more general understanding of the best possible performance in a given duration. Algorithmically, although gradient-based optimal control algorithms have been developed [10] and successfully applied to many magnetic resonance experiments in the past decade, an important goal is to make these algorithms easier to use, generally applicable, and more efficient. Depending on the applications, very different convergence rates could be encountered, due to the unknown structure of optimal control landscapes in practical conditions, e.g., relaxation, inhomogeneity of magnetic fields, or constraints on the field strength being present. With more efficient numerical optimization methods, optimal control theory will enable the design of problem-specific pulse sequences. The closed-loop AFC approach for NMR, as introduced in Chapter 2 of this dissertation, gives a promising alternative that automatically accounts for all the practical imperfections, and avoids the difficulties in building specific physical models and determining the parameters therein.

The magnetic resonance techniques have important applications in chemical and biological analysis via NMR, and in medical diagnostics via MRI. For identification and quantification of compounds in complex mixtures, methods derived from optimal control can
deliver more reliable and robust results. Optimal control methods are also expected to reduce the time required to determine structural and dynamical information of biomolecules like proteins. It is computationally expensive to cope with the large coupled spin networks in such molecules, especially in the presence of relaxation, and further improved numerical or analytical approaches are highly desirable. In medical imaging, optimal control methods are expected to lead to more sensitive and more efficient pulse sequences, so that a patient has to spend less time in a scanner for an examination.

A very useful outcome of the continued progress in optimal control technology could be that, for a given level of performance, the use of optimal control sequences could significantly reduce the instrument costs as well as costs of sample preparation and purification. With both being achieved by optimally tailored pulse sequences under non-ideal conditions, the complexity could be transferred from the instrument design to the mathematical optimization procedure. Another long-term goal of magnetic resonance techniques is the detection of the nuclear spins of individual molecules, such as a protein.

### 8.3 Quantum control in chemical physics

After about thirty years’ development, the dream of controlling a complete chemical reaction from its entrance channel to the products with light, especially in a coherent manner with shaped ultrafast laser pulses, seems to be within reach. Generally, the whole process may include the controlled cleavage of original chemical bonds, the controlled formation of new bonds, the controlled dynamics of the intermediate complex (most likely involving a conical intersection), as well as the stabilization of the reaction product [3]. Despite the success of coherent control for unimolecular dissociative processes, control of associative binary reactions still remains a great challenge. While the dependence of photodissociation on the pulse shape has been clearly demonstrated in large numbers of experiments (see Ref. [11] for an example) and relatively well understood, the role of laser pulses and their
shapes in the formation of new bonds is still ambiguous.

Recently, laser control of chemical reactions has been extended from unimolecular to bimolecular, and from bond cleavage to the forming of new bonds. One example investigated by different groups is the reaction of CO and H\textsubscript{2} yielding hydrocarbons, CO\textsubscript{2} and water, known as the Fischer-Tropsch reaction. G. Gerber et al. utilized shaped femtosecond laser pulses to initiate and control this reaction at room temperature and under high vacuum conditions on catalytic single crystal metal surfaces. With the AFC scheme aiming to maximize the yield of different product ions, they achieved selective control of reaction channels comprising the formation of new molecular bonds, especially the C-H bond, rather than the cleavage of existing bonds [12]. Interestingly, it was soon observed that the same laser driven reaction could also take place at atmospheric pressure by laser induction without any catalyst [13]. These experiments demonstrated the potential of shaped femtosecond laser pulses for steering a complex reaction in one or the other direction. Notice that the diffusion and collision of intermediates usually occur on a much longer timescale than photodissociation, especially in a gas-phase reaction, and the multiscale control process relies on repetitive implementation of laser pulses. The impact exerted by the control field (e.g., dissociation and ionization of reactants) is coupled with the intrinsic chemical kinetics of the system (collision of active species and formation of new bonds without participation of the laser field), making the detailed control mechanism ambiguous. Due to the long reaction time (up to hours in Ref. [13]), searching for the optimal field parameters with the closed-loop AFC approach would also be a challenge. In the long term, a vision of photochemistry is for synthesis by photoassociation via polarization shaped light, with the final product stabilized by laser cooling and trapped by light [14].

As of chemical analysis, the vision is for a light field shaped to a specific molecule or functional group generating a specific physical outcome, such as light emission or ionization, in order to lower the detection threshold of specific molecules. For example, an optimal dynamic detection method via optimal laser pulse shaping has been applied to
detect explosives [15].

8.4 Control landscapes in alternative physical models

This section discusses a few possibilities of applying landscape analysis to alternative physical models. Although they may not find direct applications in real life, the future works in this direction may enrich our understanding of various control and optimization problems in the nature.

Open or infinite-dimensional quantum systems. Currently, control landscape topology of open quantum systems has only been studied in a few specific and simple models, which exhibits remarkably different features from those of closed systems, e.g., the absence of any critical points [16]. General landscape analysis of open systems undergoing both Markovian and non-Markovian dynamics is left for future studies. Moreover, the generalization of landscape conclusions to infinite-dimensional closed quantum systems, especially for a target observable with unbounded and continuous spectrum, is non-trivial. All the analysis may require usage of new mathematical methods.

Quantum controller. In the semiclassical model commonly used in theoretical and simulational studies, a classical field represented by a time function is employed to control the dynamics of a quantum system. More fundamentally, the control can also be described quantum mechanically, especially in the subject of quantum optics. Beyond the preliminary work on full quantum control in Chapter 7 of this dissertation, the situation of a classical system with a quantum controller (e.g., a macromolecule in an optical cavity) is still an open area, which will complete the tetrad of quantum/classical control of dynamical systems.

Non-Hermitian Hamiltonian. A standard axiom of quantum mechanics requires that the Hamiltonian be Hermitian, since Hermiticity guarantees that the energy spectrum is real and that time evolution is unitary. In an alternative formulation of quantum mechanics,
the mathematical axiom of Hermiticity is replaced by the condition of the so-called $\mathcal{PT}$ symmetry. An unbroken $\mathcal{PT}$ symmetry of a Hamiltonian $H$, i.e., $H$ commuting with the $\mathcal{PT}$ operator, can also guarantee that the spectrum is real, as Hermiticity does. The condition of $\mathcal{PT}$ symmetry offers the possibility of studying new quantum theories that may even describe measurable physical phenomena (see [17, 18, ?] and references therein). As an analog of standard quantum control with Hermitian Hamiltonians, we may also consider the control of a system with non-Hermitian $\mathcal{PT}$-symmetric Hamiltonian as well, and analyze the underlying control landscape structure. An important feature of this model is that the evolution operator is no longer unitary by the conventional definition, and thus the kinematic landscape analysis making use of unitarity has to be reconstructed.

**Incoherent control with Einstein coefficients.** In 1916, before the formulation of Schrödinger equation, A. Einstein proposed three processes occurring in the light-atom interaction, referred to as *spontaneous emission*, *stimulated emission*, and *absorption*. Rate constants known as Einstein coefficients $A$ and $B$ are associated with these transition processes between two energy levels. As opposed to coherent control, the phase of the control field does not play role in the outcome, allowing us to use the much cheaper incoherent light as the control resources. To optimize a control objective, e.g., the population of a given level at the target time, we can still vary the radiation intensities of each spectral component in the time domain, and define the control landscape within the Einstein model.

In summary, the landscape analysis actually deals with a fundamental and general question: When will a high-dimensional optimization be “easy”, and why. We are still seeking to generalize and adapt it to various problems in quantum control as well as other areas in sciences and engineering, and find more applications for it in practice.
Bibliography


