Smooth Optimal Control of Coherent Quantum Dynamics

Björn Bartels

School of Soft Matter

Freiburg Institute for Advanced Studies

Albert-Ludwigs-Universität Freiburg

Dissertation
zur Erlangung des Doktorgrades
der Fakultät für Mathematik und Physik
der Albert-Ludwigs-Universität
Freiburg im Breisgau

März 2015

Dekan: Prof. Dr. Dietmar Kröner
Betreuer der Arbeit: PD Dr. Florian Mintert
Referent: J.-Prof. Dr. David Groß
Koreferent: Prof. Dr. Stefan Dittmaier
Datum der mündlichen Prüfung: 15. Juni 2015
Prüfer:
apl. Prof. Dr. Thomas Filk
Prof. Dr. Tobias Schätz
PD Dr. Florian Mintert
J.-Prof. Dr. David Groß

Nous voulons, tant ce feu nous brûle le cerveau [...]
Plonger [...]au fond de l'Inconnu pour trouver du nouveau.

Charles Baudelaire, *Les fleurs du mal*

To my wife Michelle

Danksagung

Zu allererst möchte ich Florian Mintert für die interessante Themenstellung danken, und dafür dass er stets für Fragen bereit stand. Außerdem danke ich meiner Arbeitsgruppe, das sind Albert, Björn, Farhang, Federico, Jian, Łukasz, Mario, Mark und Sol; sie haben zu einem guten Arbeitsklima beigetragen. Weiterhin gilt mein Dank der Verwaltung des FRIAS, speziell Olga (Reisekostenabrechnung) und Helen (alle übrigen Fragen) - sie sind verantwortlich dafür, dass ich mich am FRIAS stets wohl und inspiriert gefühlt habe. Zum Schluss danke ich von ganzem Herzen meiner Familie, vor allem meiner Frau Michelle, für die moralische Unterstützung.

Abstract

The present thesis describes an optimal control approach using trigonometric functions. The result are very simple and smooth pulses, which motivates the notion of *Smooth Optimal Control*. Variational calculus with Floquet theory can be exploited in order to derive many expressions analytically.

Narrow-band pulses are constructed to enable AC magnetic field sensing with ensembles of nitrogen-vacancy centers. These pulses are robust against inhomogeneous broadening and inhomogeneities of the control field.

The versatility of smooth optimal control is shown by applying it to selected examples of composite quantum systems. In particular, we demonstrate how to manipulate the dynamics of two spins such that the presence of entanglement can be guaranteed during a prescribed time window. Furthermore, in order to combat decoherence, time-optimal entangling gates are constructed. Another method to mitigate decoherence effects is to decouple a spin from its spin bath. Finally, in a spin chain, the end spins are entangled in the presence of static noise. Moreover, the single spin dynamics is generalized to the coupling of an ensemble of spins to a quantum-mechanical cavity, and the effective control of the emission of this cavity is shown by driving it with suitably shaped pulses.

Zusammenfassung

Die vorliegende Arbeit beschreibt eine Optimale-Kontroll-Methode, die trigonometrische Funktionen verwendet. Das Ergebnis sind sehr einfache, glatte Pulse, was den Terminus *Glatte Optimale Kontrolle* rechtfertigt. Variationsrechnung mit Floquet-Theorie kann dazu benutzt werden, um viele Ausdrücke analytisch herzuleiten.

Schmalbandige Pulse werden konstruiert, um kleine zeitlich veränderliche Magnetfelder mit Hilfe von Stickstoff-Fehlstellen-Zentren-Ensembles zu messen. Diese Pulse sind sowohl robust gegenüber inhomogener Verbreiterung als auch gegenüber Inhomogenitäten im Kontrollfeld.

Die ganze Vielseitigkeit von glatter optimaler Kontrolle wird gezeigt durch die Anwendung auf ausgewählte Beispiele aus dem Bereich der Mehr-Teilchen-Quantendynamik, genauer wird die Dynamik zweier Spins so manipuliert, dass Verschränkung über ein endliches Zeitfenster garantiert wird. Darüber hinaus wird Dekohärenz durch Erzeugung zeitoptimaler Gatter entgegengewirkt. Eine andere Methode, Dekohärenzeffekte zu vermeiden, besteht in der Abkopplung eines Spins von seinem Spinbad. Schließlich werden in einer Spinkette die Endspins in Anwesenheit von statischem Rauschen miteinander verschränkt. Desweiteren wird die Einzel-Spin-Dynamik dahingehend verallgemeinert, dass die Kopplung eines Spinensembles an eine quantenmechanische Kavität betrachtet wird und die effektive Kontrolle der Strahlungsemission dieser Kavität durch Verwendung speziell designter Pulse gezeigt wird.

Contents

Introduction

It is astonishing how succesful time-independent quantum mechanics is. An exciting phenomenon like superconductivity [1] can be described with the concept of quasi-particles; other interesting effects like Bose-Einstein condensation [2] can be explained in terms of statistical quantum mechanics. They both have in common that they can be described by static, equilibrium quantum mechanics. Yet, there are other phenomena that are intrinsically time-dependent. For example, if an atom is irradiated with a laser beam, it is of great interest to study the population of different atomic levels as a function of time. In this context, *Fermi's Golden Rule* [3] makes a statement about the final populations of the levels, and its derivation is based on time-dependent quantum mechanics. In the dynamics of large molecules, on the other hand, excitation energies can be calculated by using time-dependent density-functional theory [4, 5].

In all the above cases, one starts from the time-dependence and uses it in order to explain a certain effect. In Optimal Control (OC), on the other hand, one starts from a desired effect and uses the time dependence in order to reach it. More precisely, OC strives for a time-dependent control Hamiltonian $\mathcal{H}_c(t)$ with the help of which one can manipulate a system described by the time-independent Hamiltonian $\mathcal{H}_0$. The goal of OC is to drive the system to desired properties. These desired properties can be the overlap of an initial state $|\psi_i\rangle$ with a final state $|\psi_f\rangle$ (population transfer), as it is encountered in physical chemistry [6, 7]. Indeed, the first area of application of OC was nuclear magnetic resonance (NMR) [8, 9]. In this field, besides population transfer, one is particularly interested in the synthesis of unitary gates, this means not only transfer one state to another but to map a whole family of states to another one [10]. At the same time, technological advances appear, enabling the control of quantum systems, almost perfectly isolated and therefore protected against decoherence, like nitrogen-vacancy centers [11], trapped ions [12] or atoms

trapped in optical lattices [13]. In this way, quantum information processing got into the scope of OC. Here, OC is used to optimize even more interesting properties of quantum systems, like the value of an entanglement measure [14] (see Sec. 1.1.5 for a self-contained review) or the time interval over which entanglement exceeds a certain value [15]. Although one- and two-qubit gates have already been implemented in NMR, its application to systems with long coherence times is still lacking. This application would pave the way for a quantum computer.

Whereas very powerful OC algorithms exist, many of them generate pulses which contain high-frequency components. However, in some experiments [16, 17], the spectral width of the pulses to be generated is limited by nature. As an example, a cavity can generate the control signal, which has a well-defined resonance frequency and a small resonance width, *i.e.* only a small frequency band would be filtered out of a broadband pulse, leading to a completely different result than if all frequencies were present. That is why, in our approach we search for a description of the pulses which excludes high frequency components from the very beginning. This narrow band control has the advantage that it uses only a few control parameters, corresponding to the number of frequency components present in the pulse. By using very few control parameters, one obtains pulses of a quite simple shape, which gives hope to answer the question *why* a certain control goal can be obtained, thus going beyond the mere question of existence *if* such a control goal can be reached. This advantage has to be seen in contrast to many other algorithms [18], where the pulses seem - at least to an unexperienced eye - almost random [9] and it is impossible to extract any information about the functionality of the pulses.

By doing pulse shaping in frequency space, we can calculate many quantities analytically. Namely, we can compute the time evolution operator to arbitrary numerical precision, from which we can derive the quality of a pulse. However, if we want to improve the quality of a pulse, we also need information about how the system evolves if one slightly modifies the control Hamiltonian. By exploiting this information, we can improve the quality of a pulse in an iterative manner. In our framework, we can use variational calculus in order to find this information.

One important aspect that can be obtained by OC is the robustness against parameter fluctuations. In the above experiments, these parameters can be the

coupling of a spin to a cavity or the resonance frequency of this spin. Whereas a strong rectangular pulse, *i.e.* a pulse whose envelope is not modulated in time, is always optimized for a certain set of internal parameters, the modulation of a time-dependent pulse can be designed such that it gains robustness against these parameters. It is the goal of this thesis to construct pulses which work not only for one spin with a specific resonance frequency and coupling but for a whole ensemble of spins with different parameters.

These robust pulses can be used in order to measure small magnetic fields. Small magnetic fields are of a broad interest ranging from biological effects [19, 20] to applications on the human body [21]. Magnetic field sensing can be realized by driving solid state spins with a sequence of control pulses and measuring the angle by which these spins have dephased. Up to now, sensing has mostly been done with single spins and rectangular pulses [22, 23]. Examples where ensembles of spins are used can be found in [24, 25]. If the advantage of a higher signal attainable by more than one spin has not yet been envisioned, this is because spin ensembles typically suffer from parameter fluctuations that deteriorate the signal. In the present thesis, we want to overcome this restriction and tailor smooth pulses for AC magnetic field sensing with ensembles of nitrogen-vacancy centers.

Since magnetic field sensing can be modeled by an ensemble of non-interacting spins, the whole dynamics can be studied by considering the dynamics of a single spin. Indeed, only few works [26, 14, 27] investigate OC in the presence of many-body quantum dynamics. In the present work, we want to progress in this realm of OC. Since entanglement is the key element of quantum information processing, it will be interesting how entanglement dynamics can be influenced by OC. Moreover, we will drop the assumption that an ensemble can be described as a collection of non-interacting spins and will see how their dynamics can be manipulated.

The current work is structured as follows: After presenting in the first chapter the underlying theoretical and experimental concepts appearing in this thesis, we will explain in detail our OC approach - Smooth Optimal Control - in chapter 2. Chapter 3 is dedicated to the control of ensembles of independent spins. The major focus of this chapter will be to build a pulse sequence for magnetometry, *i.e.* how to measure small AC magnetic fields with the help of nitrogen-vacancy centers. In the last and most important chapter, we will

apply Smooth Optimal Control to the dynamics of more than one body. Most interestingly, we will attack the problem of an ensemble whose members interact with each other via a common coupling to a cavity. The work ends with some concluding remarks.

1. State of the art

In this chapter the state of the art of methods used in this thesis will be presented. It will be distinguished between theoretical and experimental concepts.

1.1. Theoretical Concepts

In the present section, the theoretical concepts underlying this work will be summarized. As explained in the introduction, optimal control will be used to manipulate a quantum system such that it exhibits certain desired properties. It is therefore of a great importance to understand the main ideas of OC theory, which will be presented in the following. As the function space in which one searches for the solution of the OC problem is very large, analytical results are difficult to obtain. However, there exists a number of OC algorithms with the help of which the system can be driven to the desired properties with high accuracy. As two particularly successful examples, we will describe the GRAPE and Krotov's algorithms.

In order to assess the quality of the pulses generated by optimal control, we have to calculate the dynamics induced by these pulses, *i.e.* to solve the time-dependent Schrödinger equation. Floquet's theorem gives a statement about the solutions of the Schrödinger equation for Hamiltonians periodic in time. As we will develop in chapter 2 an optimal control approach which is based on time-periodic Hamiltonians, an overview of Floquet theory will be provided.

A common task in quantum control is to implement a quantum gate, *i.e.* an operation which maps a set $\mathcal{A}$ of states onto another set $\mathcal{B}$ of states. If one wants to verify in the experiment if this quantum gate is correctly implemented, one measures to which output states a certain class of input states is transferred. This procedure is called a *quantum process tomography*, which will be explained in detail in the following.

Entanglement is a feature concerning the combination of two systems and is intrinsically quantum mechanical. As we will control many-body quantum dynamics, entanglement theory is an important tool, which will be described at the end of this chapter.

1.1.1. Optimal Control Theory

In this section, we will introduce the basic tools of optimal control theory. Fig. 1.1 can be considered as the summary of this part. As two important examples, the GRAPE and Krotov's algorithms will be described.

1.1.1.1. Basics of Optimal Control

In nature, many processes can be described by differential equations. The Schrödinger equation

$$i\frac{\partial}{\partial t}|\Psi(t)\rangle = \mathcal{H}(t)|\Psi(t)\rangle \tag{1.1}$$

is the equation of motion of quantum mechanics. Here, since we will work with a finite number of two-level systems (called 'qubits'), the Hamiltonian $\mathcal{H}(t)$ is a finite-dimensional Hermitian matrix, which can depend on time. The solution of the Schrödinger equation describes how the state vector $|\Psi(t)\rangle$ evolves in time. This state vector is not an observable in itself, but with its help experimentally accessible expectation values $\langle A\rangle = \mathrm{Tr}(|\Psi\rangle\langle\Psi|A)$ of an observable A can be computed. In the following, we will focus on the Schrödinger equation as equation of motion, but we keep in mind that it is only a special case of a linear differential equation $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), t)$, where the function $\mathbf{f}(\mathbf{x}(t), t)$ is linear in $\mathbf{x}$. This is why what follows can easily be generalized to different linear differential equations.

As we argued in the introduction, the above equation of motion merely *describes* reality. However, in the following, we want to be in the position to *design* reality according to our wishes. This is the interest of optimal control. In order to have a coherent picture of OC, we start from a time-independent *system* or *drift* Hamiltonian $\mathcal{H}_0$. As an example, one can give a spin chain whose inter-spin couplings are determined by their distance. Since one cannot change this distance, one cannot change the system Hamiltonian. This is why one adds a time-dependent control Hamiltonian $\mathcal{H}_c(t)$. In the example of the

spin chain, this could be the control of each spin by a laser beam, whose shape is modulated in time.

The goal of optimal control is formulated in terms of a target functional $\mathcal{F}(\mathcal{H}_c(t), |\Psi(t)\rangle)$ which has to be maximized and which can depend on both the control Hamiltonian $\mathcal{H}_c(t)$ and the state $|\Psi(t)\rangle$, fulfilling the Schrödinger equation Eq. (1.1) with the Hamiltonian $\mathcal{H}(t) = \mathcal{H}_0 + \mathcal{H}_c(t)$. A typical OC goal would be the transfer from an initial state $|\Psi_i\rangle$ to a final state $|\Psi_f\rangle$ at the moment t_f in time. The corresponding target functional would be the overlap $\mathcal{F}_0(|\Psi(t_f)\rangle) = |\langle\Psi_f|\Psi(t_f)\rangle|^2$ between the evolved state $|\Psi(t_f)\rangle = \mathcal{U}(t_f)|\Psi_i\rangle$ and the target state $|\Psi_f\rangle$. Here, $\mathcal{U}$ is the time evolution operator, which again fulfills the Schrödinger equation:

$$i\dot{\mathcal{U}}(t) = \mathcal{H}(t)\mathcal{U}(t) \tag{1.2}$$

with the initial condition $\mathcal{U}(0) = \mathbb{1}$. As one sees, if the overlap is equal to one, the evolved state is identical (up to an irrelevant global phase) to the target state: $|\Psi(t_f)\rangle = |\Psi_f\rangle$. If one wishes, in addition, to limit the power of the applied pulse, one can do this by adding a penalty functional $\mathcal{F}_p = -p\int_0^{t_f} \mathrm{d}t\|\mathcal{H}_c(t)\|^2$, meaning that the full target functional is $\mathcal{F} = \mathcal{F}_0 + \mathcal{F}_p$. One can see in the expression for the penalty functional that it is maximized by $\|\mathcal{H}_c(t)\| \to 0$ (be aware of the fact that the penalty functional is strictly non-positive). With the help of the parameter p one can control the importance of low power pulses: $p \to 0$ means that the intensity of the pulse is very unimportant, whereas $p \to \infty$ favours low power pulses.

If we summarize the task of OC, it is to find a control Hamiltonian $H_c(t)$ such that a certain target functional $\mathcal{F}(\mathcal{H}_c(t), |\Psi(t)\rangle)$ becomes maximal. Unfortunately, the space of all possible control Hamiltonians is extremely high-dimensional, and it is therefore a highly complex task to maximize the target functional. However, the physical reality is simpler: in the real world, only certain Hamiltonians are realized. This restriction is obvious in an experimental situation: In the case of a spin chain *e.g.* single spin properties can be easily manipulated by the experimentator, while one has no access to many-body interactions. This is why one can write

$$\mathcal{H}_c(t) = \sum_{k=1}^{N} f_k(t)\mathbf{h}_k, \tag{1.3}$$

where $f_k(t)$ are scalar functions depending on time and $\mathbf{h}_k$ are time-independent single spin operators. By performing the decomposition Eq. (1.3), one replaces the search in a high-dimensional space of Hamiltonians by a search in the space of scalar functions. However, this space is still very large, which suggests to decompose the pulses $f_k(t)$ in terms of a certain basis $g_{jk}(t)$:

$$f_k(t) = \sum_{j=1}^{n} a_{jk} g_{jk}(t) \tag{1.4}$$

Now, the OC task is a lot simpler, it consists in searching scalar tuples in contrast to scalar functions. One of these decompositions could be to express the pulses in terms of periodic functions, as we will see later.

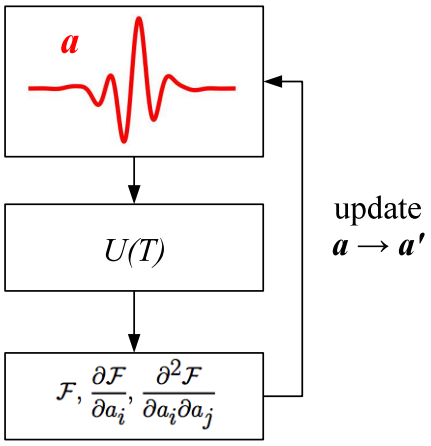

Figure 1.1.: Typical structure of a gradient-based optimal control algorithm: After computing the time evolution $\mathcal{U}$ at the end time T, one evaluates the target functional $\mathcal{F}$ and its derivatives. With the help of the derivatives a better pulse is created.

Mathematically, in order to maximize the functional $\mathcal{F}$, one has to solve the equation $\frac{\partial \mathcal{F}}{\partial a_{ij}} = 0$ for every a_{ij}. These equations are typically impossible to be solved analytically. The common way is therefore to solve the maximization problem by numerical algorithms. The basic steps of these algorithms are the following (see Fig. 1.1): In the first step, one has to determine how the control changes the system, *i.e.* one has to solve the equation of motion Eq. (1.2) for the time evolution operator $\mathcal{U}(t)$. In the second step, one assesses the quality of the pulse by evaluating the target functional. Furthermore, one has to find out how the quality of the pulse, represented by the target functional $\mathcal{F}$, changes if one slightly modifies the control amplitude a_{jk} by a small amount δa_{jk}, *i.e.*

one searches for the derivatives $\frac{\partial \mathcal{F}}{\partial a_{jk}}$ of the target functional with respect to the control parameters a_{jk}. By using the information about the derivatives, one can in the third step update the control parameters to yield a better pulse, *e.g.* one can use a method of steepest ascent, following the direction of the gradient $\nabla_{\mathbf{a}}\mathcal{F} = \left\{\frac{\partial \mathcal{F}}{\partial a_{jk}}\right\}_{jk}$,

$$\mathbf{a} \to \mathbf{a}' = \mathbf{a} + \alpha \nabla_{\mathbf{a}} \mathcal{F}. \tag{1.5}$$

Here, the step size α is still to be determined. The simplest choice is a constant value of α, a more sophisticated solution is provided by a line search, which means that one increases α if the target functional is increased, otherwise one deminishes it. A more complicated update rule using also the second derivatives with respect to the control parameters is discussed in Sec. 2.2. If after the update, the target functional has not yet obtained the desired value, one can iterate the three steps of the optimal control algorithm. Ideally, the target functional is increased in each iteration step, before obtaining its optimal value.

1.1.1.2. GRAPE

Here, we will give an example for an optimal control algorithm. GRAPE (gradient ascent pulse engineering) is a very powerful optimal control algorithm with applications mainly in NMR [10, 28]. If one follows Fig. 1.1 about the construction of a general optimal control algorithm, the first step is to calculate the time evolution operator. Therefore, it is advantageous to choose a decomposition Eq. (1.4) which is as simple as possible such that the propagator can be readily computed. In the framework of GRAPE, one chooses piecewise constant functions

$$\theta_{t_1,t_2}(t) = \begin{cases} 1 & \text{if } t_1 \leq t \leq t_2 \\ 0 & \text{otherwise} \end{cases} \tag{1.6}$$

as the basis $g_{jk}(t)$ of the control pulses, $g_{jk}(t) = \theta_{t_j,t_{j+1}}(t)$. Such a decomposition is given in Fig. 1.2. Therefore, the pulse is fully characterized by the step heights a_{jk} of the piecewise constant functions. Let us from now on assume equal time intervals, $\Delta t = t_{j+1} - t_j$.

The decomposition into piecewise constant functions has advantages and disadvantages: An advantage is that if the step size Δt is sufficiently small, one can be sure that almost any dynamics is implementable. Another advantage is

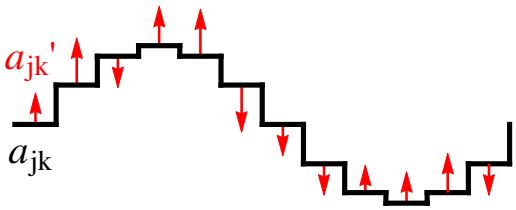

Figure 1.2.: Starting point of the GRAPE algorithm: the pulse is decomposed into piecewise constant functions with step heights (= control parameters). The update of the a_{jk}'s with a higher fidelity is provided by another set of control parameters a'_{jk}.

that the computation of the time evolution operator is particularly easy (see below). The drawback of choosing piecewise constant functions is that high frequency components can be generated rather easily, because the difference $|a_{j+1,k} - a_{jk}|$ between two adjacent steps is not limited. As one can see in [9] or other applications of OC related to NMR, the resulting pulses look nearly random. One possibility to avoid the high frequency problem is to use filter functions that smoothen the high frequency signal [29]. Furthermore, because of the high number of control parameters it is very difficult, if not impossible, to extract any information about the functionality of these pulses.

Following Fig. 1.1, one first has to calculate the time evolution operator. Here, the advantage of having chosen piecewise constant functions as the basis of the pulses can be understood: Since the Hamiltonian is constant in the time interval between t_j and t_{j+1}, the time evolution operator reads

$$\mathcal{U}_j(t) = \exp\left(-i\left(\mathcal{H}_0 + \sum_{k=1}^{N} a_{jk}\mathbf{h}_k\right)t\right), \qquad t \in [t_j, t_{j+1}]. \tag{1.7}$$

If one searches for the full time evolution $\mathcal{U}(t_f)$ at the end time $t_f = n\Delta t$, then one has to apply the operators $\mathcal{U}_j \equiv \mathcal{U}_j(\Delta t)$ one after the other:

$$\mathcal{U}(t_f) = \mathcal{U}_n \cdot \ldots \cdot \mathcal{U}_1 \tag{1.8}$$

The second step of Fig. 1.1 is to calculate the fidelity from the time evolution operator: A common fidelity measure for the transfer from the state ρ_0 to the state ρ_f at the moment t_f in time is given by

$$\mathcal{F}(\rho) = \langle \rho_f, \rho(t_f) \rangle. \tag{1.9}$$

Here, one defines the scalar product $\langle \rho_1, \rho_2 \rangle := \mathrm{Tr}(\rho_1^\dagger \rho_2)$. The fidelity is therefore nothing else but an overlap between the target state ρ_f and the final state $\rho(t_f)$ of the system after the application of control. Exploiting the fact that the trace is invariant under cyclic permutations, one can write

$$\begin{aligned} \mathcal{F} &= \langle \rho_f, \mathcal{U}_n \cdots \mathcal{U}_1 \rho(0) \mathcal{U}_1^\dagger \cdots \mathcal{U}_n^\dagger \rangle \\ &= \langle \underbrace{\mathcal{U}_{j+1}^\dagger \cdots \mathcal{U}_n^\dagger \rho_f \mathcal{U}_n \cdots \mathcal{U}_{j+1}}_{\lambda_j}, \underbrace{\mathcal{U}_j \cdots \mathcal{U}_1 \rho_0 \mathcal{U}_1^\dagger \cdots \mathcal{U}_j^\dagger}_{\rho_j} \rangle. \end{aligned} \tag{1.10}$$

In this notation, one understands that the fidelity can be computed as the scalar product between the states ρ_j and λ_j. The interpretation of the two states is that $\rho_j \equiv \rho(t_j)$ is obtained by a propagation with the initial condition $\rho(0) = \rho_0$ (forward propagation), whereas λ_j corresponds to the fictitious case where the initial condition is $\rho(0) = \rho_f$ and the final state $\mathcal{U}_{j+1}^\dagger \cdots \mathcal{U}_n^\dagger \rho_f \cdot \mathcal{U}_n \cdots \mathcal{U}_{j+1}$ at the moment t_j in time (backward propagation). This interpretation is illustrated in Fig. 1.3: One state is propagated forward, the other backwards in time, both are evaluated at the same moment t_j in time.

The full power of the decomposition with respect to the states ρ_j and λ_j will become obvious after one has calculated the gradient of the fidelity with respect to the control parameters, which is the third step of Fig. 1.1. Here, one has

$$\frac{\partial \mathcal{F}}{\partial a_{jk}} = \langle \lambda_j, \frac{\partial \mathcal{U}_j}{\partial a_{jk}} \cdots \mathcal{U}_1 \rho(0) \mathcal{U}_1^\dagger \cdots \mathcal{U}_j^\dagger \rangle + \langle \lambda_j, \mathcal{U}_j \cdots \mathcal{U}_1 \rho(0) \mathcal{U}_1^\dagger \cdots \frac{\partial \mathcal{U}_j^\dagger}{\partial a_{jk}} \rangle \tag{1.11}$$

The quantity to calculate is therefore

$$\frac{\partial \mathcal{U}_j}{\partial a_{jk}} = \frac{\partial}{\partial a_{jk}} \exp\left(-i \left(\mathcal{H}_0 + \sum_{k=1}^{N} a_{jk} \mathbf{h}_k \right) \Delta t \right). \tag{1.12}$$

This expression can be put into the form $\frac{\mathrm{d}}{\mathrm{d}x} e^{A+xB}\Big|_{x=0}$ with $A = -i\left(\mathcal{H}_0 + \sum_{k=1}^{N} a_{jk}\mathbf{h}_k\right)\Delta t$, $B = -i\mathbf{h}_k \Delta t$ and $x = \delta a_{jk}$. If the operators $\mathbf{h}_k$ commuted with each other and with $\mathcal{H}_0$, then the result would simply be

$$\frac{\mathrm{d}}{\mathrm{d}x} e^{A+xB}\Big|_{x=0} = e^A B = B e^A. \tag{1.13}$$

However, in general, these operators do not commute. This is why we have to resort to the formula

$$\frac{\mathrm{d}}{\mathrm{d}x} e^{A+xB}\Big|_{x=0} = \int_0^1 e^{A\tau} B e^{(1-\tau)A} \mathrm{d}\tau. \tag{1.14}$$

Therefore, one can write

$$\frac{\partial \mathcal{U}_j}{\partial a_{jk}} = -i \int_0^{\Delta t} \mathrm{d}\tau \, \mathcal{U}_j(\tau) \mathbf{h}_k \mathcal{U}_j(-\tau) \mathcal{U}_j. \tag{1.15}$$

One can calculate this expression exactly, as shown in [18], or use the approximation $\Delta t \| \mathcal{H}_0 + \sum_k a_{jk} \mathbf{h}_k \| \ll 1$ [10]. For small time intervals, the operator $\mathcal{U}(\tau)$ ($\tau \in [0, \Delta t]$) reduces to the identity and one obtains

$$\frac{\partial U_j}{\partial a_{jk}} \approx -i \Delta t \mathbf{h}_k U_j. \tag{1.16}$$

Reinserting this into Eq. (1.11), one obtains

$$\frac{\partial \mathcal{F}}{\partial a_{jk}} \approx -i \Delta t \langle \lambda_j | [\mathbf{h}_k, \rho_j] \rangle \tag{1.17}$$

This computation of the gradient is in contrast to conventional methods that use a rule of finite differences

$$\frac{\partial \mathcal{F}}{\partial a_{jk}} \approx \frac{\mathcal{F}(a_{jk} + \delta a_{jk}) - \mathcal{F}(a_{jk})}{\delta a_{jk}} \tag{1.18}$$

with a small δa_{jk} in order to approximate the gradient.

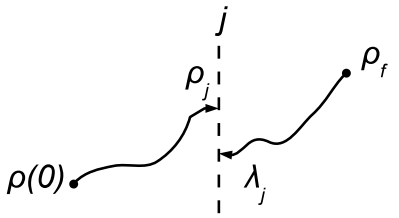

Figure 1.3.: The calculation of the target functional is illustrated by propagating forward the initial state $\rho(0)$ and backward the final state ρ_f. In order to compute the gradient, only these two states are necessary, whereas conventional methods need n propagations.

Finally, in the last step in Fig. 1.1, one updates the control parameters by a method of steepest ascent (see Eq. (1.5)). We will now see why GRAPE is faster than conventional optimal control methods: In the framework of these conventional methods a control parameter is updated and then the time evolution is recalculated, as it had changed due to the update of the control parameter. This is what is called a sequential update ('one after the other'). In the framework of GRAPE, however, to calculate $\frac{\partial \mathcal{F}}{\partial a_{jk}}$ only requires the two forward and backward propagations ρ_j and λ_j and all control parameters can be updated at once (concurrent update).

1.1.1.3. Krotov's algorithm

An algorithm that uses a sequential update is Krotov's method [30]. The whole mathematical understanding of this algorithm is beyond the scope of this thesis but we want nevertheless describe its idea. The idea is to maximize the target functional $\mathcal{F}$ by extremalizing the variation $\delta\mathcal{F}$, *i.e.* $\delta\mathcal{F} = 0$. In the following, we assume a target functional

$$\mathcal{F} = \langle\psi(t_f)|O|\psi(t_f)\rangle - \alpha \int_0^{t_f} \sum_j f_j(t)^2 \mathrm{d}t, \tag{1.19}$$

where O is an operator ($O = |\psi(0)\rangle\langle\psi(0)|$ in the case of an overlap fidelity) and the last term penalizes strong fields by the factor α. By using variational calculus, in the same manner as Euler-Lagrange equations are obtained, one is able to derive the update rule [31]

$$f_j(t) = -\frac{1}{\alpha} \mathrm{Im}\langle\chi(t)|\mathbf{h}_j|\psi(t)\rangle. \tag{1.20}$$

Here, in analogy with Eq. (1.10), $|\psi(t)\rangle$ is the forward propagated solution of the Schrödinger equation (initial condition $|\psi(0)\rangle = |\psi_0\rangle$) and $|\chi(t)\rangle$ is the backward propagated solution with initial condition $|\chi(t_f)\rangle = O|\psi(t_f)\rangle$, where t_f is the total duration of the pulse. A characteristic feature of Krotov's method is that it is continuous in time. However, a computer has a limited memory. This is why $f_j(t)$ can only be stored at finitely many moments of time $t_1, \ldots, t_n$. Because of the discreteness of time on a computer, the pulse has to be updated successively (sequential update) for all time slices $t = t_1, \ldots, t_n$. An advantage of Krotov's method is that, as the time slices can be made arbitrarily small, it is expected that it also remains continuous in time, *i.e.* the generation of high-frequency components is less probable.

Krotov's method has become a very popular OC algorithm. It is used in different fields such as quantum information theory [32, 33] or quantum chemistry [31, 34].

1.1.2. Floquet Theory

So far, only piecewise constant Hamiltonians were considered, from now on the focus will be on periodic Hamiltonians, as they appear in smooth optimal control (SOC) described in chapter 2. As mentioned above, Floquet theory is

important in order to assess the quality of the SOC pulses. It is a powerful tool that permits us to calculate the solutions of the time-dependent Schrödinger equation Eq. (1.1) with a periodic Hamiltonian $\mathcal{H}(t+T) = \mathcal{H}(t)$ and finally to calculate the time evolution operator $\mathcal{U}$. Floquet's theorem [35] states that for Hamiltonians periodic in time the fundamental solutions $|\Psi_k\rangle$ of Eq. (1.1) can be written in the form

$$|\Psi_k(t)\rangle = e^{-i\epsilon_k t}|\Phi_k(t)\rangle, \tag{1.21}$$

where the $|\Phi_k\rangle$ are again periodic in time, $|\Phi_k(t+T)\rangle = |\Phi_k(t)\rangle$. If one casts Eq. (1.21) back into Eq. (1.1), one obtains the eigenvalue problem

$$\mathcal{K}|\Phi_k\rangle = \epsilon_k|\Phi_k\rangle \tag{1.22}$$

with the Floquet operator $\mathcal{K} = \mathcal{H} - i\partial_t$. Because of this equation describing an eigenvalue problem, one calls the ϵ_k (quasi-)energies and the corresponding $|\Phi_k\rangle$ eigenstates. One uses the term *quasi*-energies in order to distinguish them from the energies appearing in the stationary Schrödinger equation. Whereas the energies, as the eigenvalues of a Hamiltonian, have a physical meaning, this is not true in the case of the Floquet quasi-energies. It is noteworthy that the Floquet picture has many points in common with the Schrödinger picture: The Schrödinger equation $i\frac{\mathrm{d}}{\mathrm{d}t}|\Psi\rangle = \mathcal{H}|\Psi\rangle$ with a time-independent Hamilton operator $\mathcal{H}$ can be transformed into an eigenvalue problem $\mathcal{H}|\Phi_k\rangle = E_k|\Phi_k\rangle$ by means of the ansatz $|\Psi_k(t)\rangle = e^{-iE_k t}|\Phi_k\rangle$. One can see by this ansatz that a Floquet eigenstate is the formal analogon of a stationary state. The only difference between Floquet and Schrödinger picture is that the eigenstates in the Floquet picture still depend periodically on time, whereas the energy eigenstates in the Schrödinger picture are constant in time. This additional time-dependence can already be seen in the expression of the Floquet operator, which is not stationary but depends on time. Indeed, $\mathcal{K}$ not only acts via $\mathcal{H}_0$ on the original finite-dimensional Hilbert space $\mathbb{C}^d$ (where d is the dimension of the system) but also acts via the derivative with respect to time and $\mathcal{H}_c(t)$ on the space $L^2([0,T))$ of T-periodic square integrable functions. The latter space is infinite-dimensional and so is the eigenvalue problem, in contrast to the finite dimensional Schrödinger equation.

Besides the similarity of the Schrödinger and Floquet picture, it is important to mention that the Bloch theorem [36] from solid state physics is very similar

to Floquet's theorem. Instead of the time domain, the Hamilton operator is periodic in space, *i.e.* the potential fulfills $V(\mathbf{r}+\mathbf{R}) = V(\mathbf{r})$, where the vector $\mathbf{R}$ describes the periodicity in space, and the solutions of the Schrödinger equation are $\psi(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{\mathbf{k}}(\mathbf{r})$ with periodic functions $u_{\mathbf{k}}(\mathbf{r}+\mathbf{R}) = u_{\mathbf{k}}(\mathbf{r})$.

The periodic structure of Floquet's theorem suggests to solve the time-dependent Schrödinger equation in Fourier space. To this end one performs the elementary steps of a Fourier transformation, namely one writes $|\Phi_k\rangle$ as a Fourier series $|\Phi_k\rangle = \sum_{\nu=-\infty}^{\infty} e^{i\nu\Omega t}|\chi_{k\nu}\rangle$ with $\Omega = 2\pi/T$, as well as the Fourier-transformed Hamilton operator $\mathcal{H}_\nu = 1/T\int_0^T \mathrm{d}t\, e^{-i\nu\Omega t}\mathcal{H}(t)$. Then one reinserts this into Eq. (1.22), multiplies both sides by $e^{-i\mu\Omega t}$, $\mu \in \mathbb{Z}$, and integrates them with respect to t from 0 to T. This yields

$$1/T\int_0^T \mathrm{d}t\, e^{-i\mu\Omega t}(\mathcal{H}(t)-i\partial_t)\sum_{n=-\infty}^{\infty} e^{i\nu\Omega t}|\chi_{k\nu}\rangle = \epsilon_k/T\int_0^T \mathrm{d}t\, e^{-i\mu\Omega t}\sum_{n=-\infty}^{\infty} e^{i\nu\Omega t}|\chi_{k\nu}\rangle, \tag{1.23}$$

which gives

$$\sum_{\nu=-\infty}^{\infty} 1/T\int_0^T \mathrm{d}t\, \left(e^{-i(\mu-\nu)\Omega t}\mathcal{H}(t) + \nu\Omega e^{i(\nu-\mu)\Omega t}\right)|\chi_{k\nu}\rangle = \varepsilon_k|\chi_{k\mu}\rangle \tag{1.24}$$

and finally the set of coupled equations

$$\mu\Omega|\chi_{k\mu}\rangle + \sum_{\nu=-\infty}^{\infty}\mathcal{H}_{\mu-\nu}|\chi_{k\nu}\rangle = \varepsilon_k|\chi_{k\mu}\rangle,\ \forall\mu\in\mathbb{Z}. \tag{1.25}$$

Solving Eq. (1.25) is equivalent to finding the eigensystem of the eigenvalue problem [37]

$$\tilde{\mathcal{K}}|\chi_k\rangle = \epsilon_k|\chi_k\rangle, \tag{1.26}$$

where $|\chi_k\rangle = (\dots, |\chi_{k,-1}\rangle, |\chi_{k,0}\rangle, |\chi_{k,1}\rangle, \dots)^T$ and

$$\tilde{\mathcal{K}} = \begin{pmatrix} \ddots & & & & & \\ \ddots & \mathcal{H}_0 - \mathbb{1}\Omega & \ddots & & & \\ \dots & \mathcal{H}_1 & \mathcal{H}_0 & \mathcal{H}_{-1} & \dots \\ & & \ddots & \mathcal{H}_0 + \mathbb{1}\Omega & \ddots \\ & & & & \ddots \end{pmatrix} \tag{1.27}$$

is the (infinite-dimensional) Fourier-transformed Floquet operator. We have therefore shifted the problem of solving a differential equation (the Schrödinger equation) to the problem of diagonalizing the Floquet matrix $\tilde{\mathcal{K}}$.

As in the experiment one can only implement a finite number of frequency components, there is a highest non-vanishing Fourier component $\mathcal{H}_n$. The consequence of this fact is that $|\tilde{\mathcal{K}}_{ii}-\tilde{\mathcal{K}}_{jj}| \gg |\tilde{\mathcal{K}}_{ij}|$, *i.e.* for large $i-j$, the mixing between the off-diagonal elements becomes negligeable and the eigenvalues and eigenvectors of the truncated version $\tilde{\mathcal{K}}_{\mathrm{trunc}}$ of the Floquet matrix represent a good approximation of those of the full matrix $\tilde{\mathcal{K}}$.

By the following symmetry one can construct all eigenvalues and eigenvectors from a very limited set of eigenvalues and eigenvectors: Since the Floquet matrix $\tilde{\mathcal{K}}$ satisfies $S\tilde{\mathcal{K}}S^\dagger = \tilde{\mathcal{K}} + \Omega\mathbb{1}$, where S shifts the entries of $|\chi_k\rangle$ by one, *i.e.* $(S|\chi_k\rangle)_\nu = (|\chi_k\rangle)_{\nu+1}$, the vector $S^m|\chi_k\rangle$ is an eigenvector for any integer m if $|\chi_k\rangle$ is an eigenvector. In other words, if $|\chi_k\rangle$ is an eigenvector to the eigenvalue ϵ_k, then $S^m|\chi_k\rangle$ is an eigenvector to the eigenvalue $\epsilon_k - m\Omega$. In this way, one obtains a set of eigenvalues, *i.e.* the knowledge of the right d eigenvalues of $\tilde{\mathcal{K}}$, namely those for which $(\epsilon_k - \epsilon_j)/\Omega \notin \mathbb{Z}$, $k \neq j$ is valid, where $d = \dim(\mathcal{H})$, permits us to know all of them by translation. The interval in which such a set of eigenvalues lies is also called a Brillouin zone. This terminology is the same as in Bloch's theorem: If $\mathbf{k}$ is a wave vector, then $\mathbf{k} + \mathbf{K}$ is also a wave vector. Here, $\mathbf{K}$ is a reciprocal lattice vector with the property $e^{i\mathbf{K}\cdot\mathbf{R}} = 1$, and $\mathbf{R}$ determines the periodicity. The first Brillouin zone is built up by the wave vectors $\mathbf{k}$ which cannot be connected by a reciprocal lattice vector $\mathbf{K}$.

In order to compute the time evolution operator $\mathcal{U}$, which has been the ultimate goal of the above calculations, one has to know only the spectrum of one Brillouin zone, *i.e.* one can limit the summation of basis states $|\Phi_k\rangle$ in $\mathcal{U}$ to the summation of d terms:

$$\mathcal{U}(t) = \sum_{k=1}^{d} |\Psi_k(t)\rangle\langle\Psi_k(0)| = \sum_{k=1}^{d}\sum_{\nu=-\infty}^{\infty} e^{i(\nu\Omega-\epsilon_k)t}|\chi_{k\nu}\rangle\langle\Phi_k(0)| \ . \tag{1.28}$$

Numerically, one is able to compute the time evolution operator exactly, as the resulting error can be reduced to a value which is equal to the machine precision. More information about the numerics of Floquet theory can be found in the appendix A.

1.1.3. Quantum states and tomography

For what follows, one needs to generalize the concept of a state. Up to now, the concept of a state was limited to vectors $|\psi\rangle$ in a Hilbert space $\mathscr{H}$. If one

wishes to unify statistical and quantum mechanics, one has to consider the case where the system is in the state $|\psi_k\rangle$ only with a certain probability p_k. The physics is therefore characerized by an ensemble of states. The central object describing this situation is the *density matrix*

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|. \tag{1.29}$$

This matrix is Hermitian because $|\psi_k\rangle\langle\psi_k|$ is Hermitian and p_k real-valued. The density matrix is also positive-semidefinite, *i.e.* only possesses non-negative eigenvalues, because the probabilities p_k are non-negative. Furthermore, it is $\mathrm{Tr}\,\rho = 1$ because $\sum_k p_k = 1$. If $\rho = |\psi\rangle\langle\psi|$ (equivalent to $\mathrm{Tr}(\rho^2) = 1$), the density matrix represents a *pure* state, otherwise it is called a *mixed* state (characterized by $\mathrm{Tr}(\rho^2) < 1$).

In the experiment, the exact density matrix is typically not known. In quantum state tomography, which is a generalization of quantum process tomography (to be discussed in Sec. 1.1.4), the quantum state is reconstructed by doing sufficiently many measurements. In order to measure observables more than one time, this implies that there are sufficiently many copies of the system, which can be realized by doing the same experiment over and over again. As an example of a quantum state tomography, one can take a qubit: the operators $\mathbb{1}/\sqrt{2}$, $\sigma_x/\sqrt{2}$, $\sigma_y/\sqrt{2}$, $\sigma_z/\sqrt{2}$, where σ_k are the Pauli matrices, form a basis of the four-dimensonal Hilbert space. A state ρ can then be expanded in terms of this basis [38]:

$$\rho = \frac{1}{2}\left(\mathrm{Tr}(\rho)\mathbb{1} + \mathrm{Tr}(\sigma_x\rho)\sigma_x + \mathrm{Tr}(\sigma_y\rho)\sigma_y + \mathrm{Tr}(\sigma_z\rho)\sigma_z\right) \tag{1.30}$$

The expansion coefficients are given by the expectation values of the corresponding operators. However, in the experiment, one can only perform a finite number of measurements. This means that the density matrix can only be reconstructed approximately. The standard deviation of this matrix decreases as $1/\sqrt{N}$ [38], where N is the number of measurements. When measuring *e.g.* σ_z, one gets the outcomes $z_1, \ldots, z_N \in \{-1, 1\}$, *i.e.* the expectation value of σ_z can be approximated by $\frac{1}{N}\sum_{k=1}^{N} z_k$. Therefore, by measuring the spin in x-, y- and z-direction and obtaining the outcomes x_k, y_k and z_k ($k = 1, \ldots, N$), the density matrix can be reconstructed as

$$\rho \approx \frac{1}{2}\left(\mathbb{1} + \frac{1}{N}\sum_{k=1}^{N}\left(x_k\sigma_x + y_k\sigma_y + z_k\sigma_z\right)\right) \tag{1.31}$$

1.1.4. Quantum Process Tomography

Floquet theory as discussed in Sec. 1.1.2 permits us to theoretically construct the dynamics induced by a periodic Hamiltonian. If one aims at an experimental implementation of such a Hamiltonian, one typically needs to verify that indeed the desired propagator has been realized. Such a task is typically referred to as quantum process tomography. As a quantum process can be considered the dynamics induced by the pulses that are constructed with the methods presented in 1.1.1.

In the present section, quantum process tomography will be described. In order to do that, one first needs the notion of a *quantum process*: For a pure state $|\psi\rangle$, a quantum process $\mathcal{E}$ is represented by a unitary operator U, *i.e.* which preserves the scalar product:

$$|\psi\rangle \to |\psi'\rangle = \mathcal{E}(|\psi\rangle) = U|\psi\rangle \tag{1.32}$$

For a mixed state, a quantum process $\mathcal{E}$ is represented by a set of *Kraus operators* $\{E_k\}$, each of which is applied to a state $|\psi_k\rangle$ of the statistical mixture, *i.e.*

$$\rho \to \rho' = \mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger. \tag{1.33}$$

Preservation of trace, $\operatorname{Tr}\rho' = \operatorname{Tr}\rho = 1$, results in the constraint $\sum_k E_k^\dagger E_k = \mathbb{1}$ on the operators E_k.

Let us now come to quantum process tomography: Whereas the concept of Kraus operators $\{E_k\}$ is only a formal tool, in the experiment the quantum process can often be modelled by using another basis $\{A_k\}$, which is connected to E_k via $E_k = \sum_j a_{kj} A_j$. The quantum process then reads

$$\mathcal{E}(\rho) = \sum_{m,n} \chi_{mn} A_m \rho A_n^\dagger, \tag{1.34}$$

where $\chi_{mn} = \sum_k a_{km} a_{kn}^*$ are the entries of a matrix which is positive Hermitian by construction. A common choice [38] for a single qubit is $A_1 = \mathbb{1}$, $A_2 = \sigma_x$, $A_3 = -i\sigma_y$, $A_4 = \sigma_z$. In order to do quantum process tomography, one prepares the system successively in one of the basis states ρ_j, which were selected beforehand. For a single qubit, a common choice [39] of the basis are the pure states $\rho_j = |\psi_j\rangle\langle\psi_j|$ with $|\psi_1\rangle = |0\rangle$, $|\psi_2\rangle = |1\rangle$, $|\psi_3\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$

and $|\psi_4\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$, which form a complete set of states. The result is the following decomposition

$$\mathcal{E}(\rho_j) = \sum_k \lambda_{jk}\rho_k, \tag{1.35}$$

where the coefficients λ_{jk} can be determined by a quantum state tomography of the $\mathcal{E}(\rho_j)$. The task of quantum process tomography is to determine the process matrix χ. In order to do that, we write $A_m\rho_j A_n^\dagger$ in the basis ρ_j:

$$A_m\rho_j A_n^\dagger = \sum_k \beta_{jkmn}\rho_k, \tag{1.36}$$

The coefficients β_{jkmn} can be found by standard linear algebra [38]. Combining Eqs. (1.35) and (1.36), one obtains:

$$\sum_{k,m,n} \chi_{mn}\beta_{jkmn}\rho_k = \sum_k \lambda_{jk}\rho_k \tag{1.37}$$

Since the ρ_j form a basis, they are linear independent, which results in

$$\sum_{m,n} \beta_{jkmn}\chi_{mn} = \lambda_{jk}. \tag{1.38}$$

If one introduces the multi-indices $m' \equiv (j,k)$ and $n' \equiv (m,n)$, the tensor β can be represented by a matrix and the matrices λ and χ by vectors. Then, by inverting β, one arrives at the process matrix χ.

For the case of a qubit, if one defines the output states

$$\begin{aligned}
\rho_1' &= \mathcal{E}(|\psi_1\rangle\langle\psi_1|) \\
\rho_4' &= \mathcal{E}(|\psi_2\rangle\langle\psi_2|) \\
\rho_2' &= \mathcal{E}(|\psi_3\rangle\langle\psi_3|) - i\mathcal{E}(|\psi_4\rangle\langle\psi_4|) - (1-i)\frac{\rho_1' + \rho_4'}{2} \\
\rho_3' &= \mathcal{E}(|\psi_3\rangle\langle\psi_3|) + i\mathcal{E}(|\psi_4\rangle\langle\psi_4|) - (1+i)\frac{\rho_1' + \rho_4'}{2},
\end{aligned} \tag{1.39}$$

the process matrix χ can be calculated according to [40] as

$$\chi = \frac{1}{2}\begin{pmatrix} \mathbb{1} & \sigma_x \\ \sigma_x & -\mathbb{1} \end{pmatrix}\begin{pmatrix} \rho_1' & \rho_2' \\ \rho_3' & \rho_4' \end{pmatrix}\begin{pmatrix} \mathbb{1} & \sigma_x \\ \sigma_x & -\mathbb{1} \end{pmatrix}. \tag{1.40}$$

The problem of quantum process tomography is that the number of required measurements grows exponentially with the number of involved particles [38].

Fortunately, in the present work, we will only consider the quantum tomography of one spin.

Unfortunately, in the experiment every measurement is afflicted by errors which cannot be described by positive semidefinite process matrices. In order to remove these processes, one tries to find the nearest possible process matrix to the experimental one. For this, it can be exploited that according to the Cholesky decomposition, a positive semidefinite matrix can be written in the form [39]

$$\tilde{\chi} = T^\dagger T \tag{1.41}$$

with the lower triangular matrix

$$T = \begin{pmatrix} t_1 & 0 & 0 & 0 \\ t_5 + it_6 & t_2 & 0 & 0 \\ t_{11} + it_{12} & t_7 + it_8 & t_3 & 0 \\ t_{15} + it_{16} & t_{13} + it_{14} & t_9 + it_{10} & t_4 \end{pmatrix}. \tag{1.42}$$

The task is then to find the parameters $t_1, \ldots, t_{16}$ such that the distance between the unphysical and the correct process

$$\Delta\chi = \|\chi - \tilde{\chi}\|_{HS} \tag{1.43}$$

becomes minimal. Here, $\|\chi\|_{HS} := \sqrt{\mathrm{Tr}(\chi\chi^\dagger)}$ is the Hilbert-Schmidt norm. The search of the coefficients t_k can be done by standard optimization algorithms.

1.1.5. Entanglement Theory

Entanglement has no classical counterpart and is the central element thanks to which quantum algorithms outperform their classical counterparts [38]. In chapter 4, control will be used in order to generate entanglement. In this section we will present the well established theory of bipartite entanglement as well as the latest advances in multipartite entanglement.

1.1.5.1. Bipartite Entanglement

As entanglement is an effect that concerns (at least) two particles, one first has to know what structure a Hilbert space formed by two particles has. The natural way to combine two Hilbert spaces is the tensor product '$\otimes$': The tensor product between two states $|u\rangle = (u_1, \ldots, u_m)$ and $|v\rangle = (v_1, \ldots, v_n)$

(from spaces with dimensions m and n, respectively) consists in multiplying every entry of $|u\rangle$ with every entry of $|v\rangle$:

$$|u\rangle \otimes |v\rangle = (u_1 v_1, \ldots, u_1 v_n, \ldots, u_m v_1, \ldots, u_m v_n) \tag{1.44}$$

The tensor product has many of the useful properties that one knows from the product in $\mathbb{R}$, namely additivity in both arguments. However, the tensor product is not commutative. If the tensor product $\mathscr{H}_A \otimes \mathscr{H}_B$ of two Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$ is constructed, its dimension is $\dim(\mathscr{H}_A \otimes \mathscr{H}_B) = mn$, where $\dim(\mathscr{H}_A) = m$ and $\dim(\mathscr{H}_B) = n$.

In order to see what entanglement is, let us consider two parties A and B with Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$, respectively. For a pure state $|\psi\rangle \in \mathscr{H}_A \otimes \mathscr{H}_B$ from the combined Hilbert space we say that it is entangled if it is not separable, *i.e.* if it cannot be written in a product form $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ with $|\psi_A\rangle \in \mathscr{H}_A$, $|\psi_B\rangle \in \mathscr{H}_B$. A mixed state $\rho \in \mathscr{H}_A \otimes \mathscr{H}_B$ is entangled if it cannot be written as a sum

$$\rho = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)} \tag{1.45}$$

with $p_i \geq 0$ and the density matrices $\rho_A^{(i)} \in \mathscr{H}_A$, $\rho_B^{(i)} \in \mathscr{H}_B$. In particular, even if a state cannot be written in the form $\rho = \rho_A \otimes \rho_B$, it does not yet mean that the state is entangled, as it can contain statistical mixtures of separable states.

In order to see what entanglement actually means, we will consider a concrete example: Let us assume that two qubits are in the entangled state $|\psi\rangle = (|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle)/\sqrt{2}$. This can be achieved, *e.g.* in subatomic physics if a spin zero particle decays into two particles. In order to conserve the total spin quantum number, the spin quantum numbers of the particles after the decay must have opposite sign. If, now, party A measures the outcome '0', the state collapses into the state

$$|0\rangle\langle 0| \otimes \mathbb{1}|\psi\rangle \propto |0\rangle \otimes |1\rangle \tag{1.46}$$

This means that if A measures '0', then the outcome '1' of B is determined before performing any measurement. This also holds for other measurements. The dependence of the measurements on each other is even more spectacular if the parties A and B are spatially separated because then it seems that

information is transmitted instantaneously. However, it can be shown [38] that information cannot be exchanged via a quantum channel with a speed higher than the speed of light.

There are different possibilities to judge if a state is entangled or not, one of those are separability criterions. The following remarks concern pure states. They can be brought from their general form

$$|\psi\rangle = \sum_{i,j} w_{ij} |\psi_A^{(i)}\rangle \otimes |\psi_B^{(j)}\rangle \tag{1.47}$$

to the Schmidt decomposition

$$|\psi\rangle = \sum_{i} \sqrt{\lambda_i} |\phi_A^{(i)}\rangle \otimes |\phi_B^{(i)}\rangle. \tag{1.48}$$

This follows from the singular value decomposition [38]. Now, a bipartite state is separable if and only if all but one Schmidt coefficient λ_i vanish. Otherwise, the state is entangled. Here, one can draw an analogy between entanglement and mixedness $\mathrm{Tr}(\rho_A^2)$ of the reduced density matrix $\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|)$. If the reduced state ρ_A is pure, then the initial pure state $|\psi\rangle$ is separable. If, on the other hand, the reduced density matrix is maximally mixed, then $|\psi\rangle$ is maximally entangled.

Since for mixed states, the definition of entanglement is more complex, it can be expected that separability criteria are also harder to find. These criteria are beyond the scope of this thesis but can be found in detail in [41].

If up to know, we have only looked at criteria that permit us to know whether a state is entangled or not, however, it is more interesting to establish a quantitative concept of entanglement. This quantitative approach is provided by entanglement measures. It is beyond the scope of this work to give a deeper insight into entanglement measures. This is why we restrict ourselves to the elementary information which is needed in chapter 4. In this sense, for pure states, the *von Neumann entropy* $E(\rho_A) = -\mathrm{Tr}(\rho_A \log_2 \rho_A)$ of the reduced state $\rho_A = \mathrm{Tr}_B |\psi\rangle\langle\psi|$ represents an entanglement measure. The Einstein-Podolski-Rosen (EPR)-state $|\mathrm{EPR}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ maximizes $E(\rho_A)$.

For mixed states, to find an entanglement measure is not so easy anymore because it is difficult to distinguish between quantum and classical correlations. One way out of this dilemma is the introduction of the *convex roof*: It

means that starting from a pure state entanglement measure one calculates the infimum of the convex sum over all decompositions $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, *i.e.*

$$E(\rho) = \inf_{\{p_i, |\psi_i\rangle\}} \sum_i p_i E(|\psi_i\rangle). \tag{1.49}$$

To calculate the convex roof can be very complicated, if not impossible, in practice. For the case of two qubits, *i.e.* $\dim \mathscr{H}_A = \dim \mathscr{H}_B = 2$ and E being the entanglement of formation (see [42]), the problem has been solved exactly. To arrive at the final expression, one first has to define the spin-flipped state

$$\tilde{\rho} = \sigma_y \otimes \sigma_y \rho^* \sigma_y \otimes \sigma_y \tag{1.50}$$

Then, one labels the eigenvalues λ_i of the matrix $\rho\tilde{\rho}$ in decreasing order: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. The so-called *concurrence* can be calculated as

$$C(\rho) = \max\{0, \sqrt{\lambda_1} - \sqrt{\lambda_2} - \sqrt{\lambda_3} - \sqrt{\lambda_4}\} \tag{1.51}$$

In the case of pure states it is

$$C(|\psi\rangle) = |\langle\psi|\sigma_y \otimes \sigma_y |\psi^*\rangle|. \tag{1.52}$$

The entanglement measure, called *entanglement of formation*, is finally

$$E(\rho) = \mathcal{E}(C(\rho)) \tag{1.53}$$

with

$$\mathcal{E}(C) = -\frac{1+\sqrt{1-C^2}}{2} \log_2 \frac{1+\sqrt{1-C^2}}{2} - \frac{1-\sqrt{1-C^2}}{2} \log_2 \frac{1-\sqrt{1-C^2}}{2}. \tag{1.54}$$

It has to be mentioned that there exist also other entanglement measures than entanglement of formation. A representative collection can be found in [43]. It is worth noting that these entanglement measures can in general not be compared, as they all define a different notion of entanglement.

1.1.5.2. One- and Two-Qubit operations

In quantum information processing, single- and two-qubit operations are performed one after the other, in order to implement a quantum algorithm. Single-qubit operations can be characterized by three angels ϕ_x, ϕ_y, ϕ_z, such that

$$\mathcal{U}_{(\phi_x,\phi_y,\phi_z)} = \exp\left(-i \sum_{k=x,y,z} \phi_k \sigma_k / 2\right) \tag{1.55}$$

An even more interesting question is the parametrization of two-qubit operations because they generate entanglement. This can be done up to single qubit operations as a function of only three parameters α_x, α_y and α_z [44]:

$$\mathcal{U}_{(\alpha_x,\alpha_y,\alpha_z)} = \exp\left(-i \sum_{k=x,y,z} \alpha_k \sigma_k^{(1)} \otimes \sigma_k^{(2)}\right) \tag{1.56}$$

The gates $\mathcal{U}_{(\alpha_x,\alpha_y,\alpha_z)}$ which generate maximal entanglement are characterized by the two inequalities [44]

$$\alpha_x + \alpha_y \geq \pi/4 \tag{1.57}$$

$$\alpha_y + \alpha_z \leq \pi/4 \tag{1.58}$$

The separable initial state $|\Psi\rangle$ with which $\mathcal{U}_{(\alpha_x,\alpha_y,\alpha_z)}|\Psi\rangle$ is a maximal entangled state reads [44]

$$|\Psi\rangle = \sum_{k=1}^{4} |\mu_k| e^{i\lambda_k} |\Phi_k\rangle \tag{1.59}$$

with

$$\begin{aligned} \lambda_1 &= \alpha_x - \alpha_y + \alpha_z \\ \lambda_2 &= -\alpha_x + \alpha_y + \alpha_z \\ \lambda_3 &= -\alpha_x - \alpha_y - \alpha_z \\ \lambda_4 &= \alpha_x + \alpha_y - \alpha_z \end{aligned} \tag{1.60}$$

and

$$\begin{aligned} |\Phi_1\rangle &= \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \\ |\Phi_2\rangle &= \frac{-i}{\sqrt{2}} (|00\rangle - |11\rangle) \\ |\Phi_3\rangle &= \frac{1}{\sqrt{2}} (|01\rangle - |10\rangle) \\ |\Phi_4\rangle &= \frac{-i}{\sqrt{2}} (|01\rangle + |10\rangle) . \end{aligned} \tag{1.61}$$

The coefficients μ_k obey the following equations (we can choose $\mu_1 = 0$ without loss of generality):

$$\sin(\alpha_3)|\mu_2|^2 + \sin(\alpha_1)|\mu_4|^2 = 0 \tag{1.62}$$

$$|\mu_3|^2 + \cos(\alpha_3)|\mu_2|^2 + \cos(\alpha_1)|\mu_4|^2 = 0 \tag{1.63}$$

$$|\mu_2|^2 + |\mu_3|^2 + |\mu_4|^2 = 1, \tag{1.64}$$

where we have defined

$$\begin{aligned} \alpha_1 &= 4(\alpha_x + \alpha_y) \\ \alpha_2 &= 4(\alpha_x + \alpha_z) \\ \alpha_3 &= 4(\alpha_y + \alpha_z). \end{aligned} \tag{1.65}$$

1.1.5.3. Multipartite Entanglement

Whereas a good understanding of bipartite entanglement exists, there are still many open problems concerning multipartite entanglement. A pure state is called *genuinely multipartite entangled* if it is not separable with respect to any bipartition [45]. For three parties A, B and C these bipartitions would be $AB|C$, $A|BC$ and $AC|B$. A mixed state ρ is called genuinely multipartite entangled if *any* decomposition into pure states ($p_i > 0$)

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \tag{1.66}$$

contains at least one genuinely multipartite entangled state. Besides genuinely multipartite entanglement, there exists a whole variety of weaker notions of multipartite entanglement.

If we want to generalize the above notions from bipartite to multipartite entanglement, we first need a notion of the tangle. For a pure state $\rho = |\psi\rangle\langle\psi|$ it is defined in the following way [45]:

$$\mathcal{F} = \mathrm{Tr}(\rho \otimes \rho V) \tag{1.67}$$

with

$$V = 4\left(\mathbf{P}_+ - \bigotimes_{n=1}^{N} P_+^{(n)} - (1 - 2^{1-N})\mathbf{P}_-\right). \tag{1.68}$$

Here, $P_\pm^{(n)} = (1 \pm \Pi_n)/2$ are the projection operators on the (anti-)symmetric subspace and Π_n the permutation operator on the n-th subsystem, *i.e.* $\Pi|ij\rangle = |ji\rangle$. The operator $\mathbf{P}_+$ ($\mathbf{P}_-$) is the symmetric (antisymmetric) global operator. They are built up by an even (odd) number of antisymmetric operators $P_-^{(n)}$. The form of Eq. (1.68) can be explained by the fact that each summand detects a different kind of multipartite entanglement (with respect to a certain bipartition). For this, one has to keep in mind that separable states

are invariant under permutation of two factors, $\Pi_1(|\phi\rangle \otimes |\psi\rangle) \otimes (|\phi\rangle \otimes |\psi\rangle) = (|\phi\rangle \otimes |\psi\rangle) \otimes (|\phi\rangle \otimes |\psi\rangle)$, whereas entangled states are not invariant, *e.g.*:

$$\begin{aligned} \Pi_1|\mathrm{EPR}\rangle \otimes |\mathrm{EPR}\rangle &= \Pi_1(|0000\rangle + |0011\rangle + |1100\rangle + |1111\rangle)/2 \\ &= (|0000\rangle + |1001\rangle + |0110\rangle + |1111\rangle)/2 \\ &\neq |\mathrm{EPR}\rangle \otimes |\mathrm{EPR}\rangle \end{aligned} \tag{1.69}$$

Therefore, $P_-^{(i)} = 1 - \Pi_i$ vanishs for separable states. In Eq. (1.68) *e.g.* the summand $P_-^{(1)} \otimes \cdots \otimes P_-^{(N)}$ detects all states that are not fully separable, *i.e.* cannot be written in a form $\bigotimes_{i=1}^N |\psi_i\rangle$. Indeed, the operators $P_\pm^{(n)}$ are invariant under local unitary transformations, making of $\mathcal{F}$ an entanglement measure.

As an example for Eq. (1.67), let us calculate the case $N = 2$. By inserting all the definitions, one has

$$\begin{aligned} V &= 4\left(P_-^{(1)}P_-^{(2)} + P_+^{(1)}P_+^{(2)} - P_+^{(1)}P_+^{(2)} - \frac{1}{2}\left(P_-^{(1)}P_+^{(2)} + P_+^{(1)}P_-^{(2)}\right)\right) \\ &= (1-\Pi_1)(1-\Pi_2) - \frac{1}{2}\left((1-\Pi_1)(1+\Pi_2) + (1+\Pi_1)(1-\Pi_2)\right) \quad (1.70) \\ &= 2\Pi_1\Pi_2 - \Pi_1 - \Pi_2. \quad (1.71) \end{aligned}$$

Now one has to evaluate expressions like the following:

$$\begin{aligned} \mathrm{Tr}(\rho \otimes \rho\Pi_1) &= \sum_{\substack{i_1,i_2 \\ j_1,j_2}} \langle i_1 i_2|\langle j_1 j_2|\rho \otimes \rho\Pi_1|i_1 i_2\rangle|j_1 j_2\rangle \\ &= \sum_{\substack{i_1,i_2 \\ j_1,j_2}} \langle i_1 i_2|\rho|j_1 i_2\rangle\langle j_1 j_2|\rho|i_1 j_2\rangle \\ &= \sum_{i_1,j_1} \langle i_1|\,\mathrm{Tr}_2(\rho)|j_1\rangle\langle j_1|\,\mathrm{Tr}_2(\rho)|i_1\rangle = \mathrm{Tr}(\rho_1^2) \end{aligned} \tag{1.72}$$

Here, $\rho_1 = \mathrm{Tr}_2\,\rho$ is the reduced density matrix with respect to the first system. If one collects all terms, one obtains (with $\mathrm{Tr}(\rho^2) = 1$ for pure states)

$$\mathcal{F} = \mathrm{Tr}(\rho \otimes \rho V) = 2\,\mathrm{Tr}(\rho^2) - \mathrm{Tr}(\rho_1^2 + \rho_2^2) \tag{1.73}$$

This result is in accordance with intuition: The more the reduced density matrices are mixed, the more information they retain about the other part of the system, *i.e.* the more they are entangled.

It is interesting that generalized to mixed states by a convex roof expression, $\inf_{p_k,|\psi_k\rangle} \sum_k p_k \mathcal{F}(|\psi_k\rangle\langle\psi_k|)$, the expression Eq. (1.67) is still a lower bound of

the tangle. This means that if one finds a state ρ with $\mathcal{F}(\rho) > 0$, then ρ is multipartite entangled.

As a concluding remark, we have to mention that there can be several classes of maximally entangled states which cannot be transformed into each other by local transformations. In a three qubit system *e.g.* two inequivalent classes of maximally entangled states are formed by the *Greenberger-Horn-Zeilinger (GHZ) state* $|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$ and the *W state* $|\mathrm{W}\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle)$ [41].

1.2. Experimental Concepts

In the following chapter, we will describe the experimental setup which has been used to obtain the experimental results presented in chapter 3. The central element of these experiments are nitrogen-vacancy (NV) centers, which can show very long coherence times, on the order of seconds [46]. As NV centers with their excellent coherence properties are used in the experiments, we will first discuss the physics of NV centers. After that, we will present how NV centers can be manipulated coherently. Then, we will describe how NV centers can be used in order to measure small magnetic fields.

1.2.1. Nitrogen-Vacancy Centers

An NV center is a defect occuring in diamond, and it consists of a nitrogen atom, that has replaced a carbon atom from the diamond lattice, and a missing carbon atom, *i.e.* a vacancy. The axis on which the nitrogen atom and the vacancy lie is also referred to as the *NV axis*. As can be seen in Fig. 1.4, the three carbon atoms that are grouped around the vacancy contribute three free electrons to the NV complex. Another two electrons are delivered by the lone pair of the nitrogen atom. The complex described above is called the NV^0. What we will deal with, however, is the negatively charged NV^-. There, an additional electron is taken from the diamond lattice, such that six electrons contribute to the NV^-. In the following, it will always be referred to the NV^- as *the* NV center.

As the NV complex is built by an even number of spins, their total spin must be an integer number. Indeed, as derived from a combination of molecular orbitals [11, 47] and confirmed experimentally [48], the NV center has a spin-1

triplet both in the ground and the excited state (energy difference 1.945 eV, see Fig. 1.4) and a spin-0-singlet, lying between the ground and the excited states. The spin-singlet is a meta-stable dark state, which means that the system does not radiate when undergoing the transition from this state to the ground state.

The Hamiltonian describing the ground state of the NV center in the presence of a magnetic field $\mathbf{B}$ is the following [49]:

$$\mathcal{H} = \Delta S_z^2 - \gamma_e \mathbf{B}\cdot\mathbf{S} - \sum_n \gamma_n \mathbf{B}\cdot\mathbf{I}_n - \sum_n \mathbf{S}\cdot\mathbf{A}_n\cdot\mathbf{I}_n - \sum_{n>m} \mathbf{I}_n\cdot\mathbf{C}_{nm}\cdot\mathbf{I}_m \quad (1.74)$$

Here, $\mathbf{S}$ is the spin-1-operator of the six electrons participating in the NV complex, while $\mathbf{I}_n$ labels the nuclear spins, *i.e.* the carbon and nitrogen nuclei. The first term is the zero-field splitting Δ between the states characterized by the magnetic quantum numbers $m_S = 0$ and $m_S = \pm 1$, where the latter states are degenerate for zero magnetic field. The second and the third term describe the Zeeman splitting of the electronic and nuclear levels, respectively, due to the magnetic field $\mathbf{B}$; the gyromagnetic ratios with which the nuclei and electrons couple to the magnetic field are γ_e and γ_n, respectively. The fourth term results from the hyperfine interaction between the electronic spin and the nuclei, the last term stands for the dipole-dipole interaction between the nuclei. The tensors $\mathbf{A}_n$ and $\mathbf{C}_{nm}$ account for an anisotropy, they are $\mathbf{A}_n, \mathbf{C}_{nm} \propto \mathbb{1}$ in the case of a totally isotropic system.

We will now see by which approximations the general Hamiltonian Eq. (1.74) can be simplified: If the density of NV centers is sufficiently small - which will be the case in the present experiments - one can neglect the spin-spin interaction $\sum_{n>m} \mathbf{I}_n\cdot\mathbf{C}_{nm}\cdot\mathbf{I}_m$. If one applies a strong magnetic field along the NV axis, one can lift the degeneracy between the states $m_S = \pm 1$ and split up the three levels of the ground state. However, there are still sufficiently many hyperfine levels that broaden the m_S-levels. By using a hole burning scheme described in [40], one can move all but one hyperfine level out of resonance. One can now drive the transition between the ground state $m_S = 0$, named $|0\rangle$, and a hyperfine level of one of the $m_S = \pm 1$, let us say $m_S = 1$ which will be called $|1\rangle$. The whole dynamics is now constrained to two levels, which form the qubit that we will work with. If ω_0 is the transition frequency between the two levels $|0\rangle$ and $|1\rangle$, then the original Hamiltonian Eq. (1.74) is reduced to

the Hamiltonian

$$\mathcal{H}_0 = \frac{\omega_0}{2}\sigma_z. \tag{1.75}$$

In chapter 3, we will always start from this Hamiltonian.

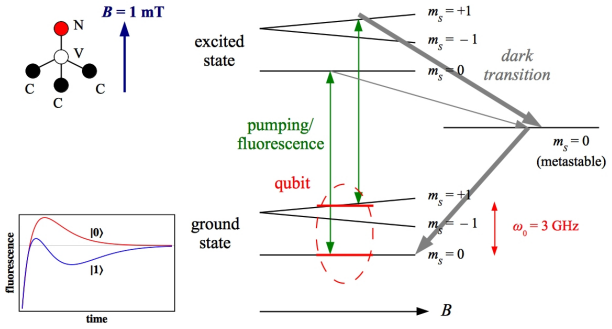

Figure 1.4.: Setup to measure population differences in the two qubit states $|0\rangle$ and $|1\rangle$, respectively: By using a magnetic field B (left), one selects a qubit transition. As the transition to a dark singlet state depends on the magnetic quantum number m_S (right), the fluorescence of the $m_S = +1$-level (left) presents a dip as compared with the one of $m_S = 0$.

Another issue one has to tackle is the problem of reading out population differences of the two qubit levels. In order to measure the population in the ground state and in the $m_S = +1$-level, one uses a sophisticated method developed in [50]: With the help of a laser, the transition to the first excited level is driven (see Fig. 1.4); this transition conserves the magnetic quantum number m_S. From the excited state a certain fraction follows a dark transition to a metastable singlet state, this fraction is higher in the $m_S = +1$ than in the $m_S = 0$ state. The remaining possibility for the system is to decay to the ground state by emitting fluorescence photons. If one observes the fluorescence as a function of time, the curve for $m_S = 1$ will present a more pronounced dip than the $m_S = 0$-curve. This dip is caused by the fraction of spins which have decayed non-radiatively from the metastable singlet state to the $m_S = 0$ ground state and therefore do not contribute to the fluorescence signal. For times $t \to \infty$ the curves for $m_S = 0$ and $m_S = 1$ saturate to a common equilibrium. This also represents a method how to initialize the system in the $|0\rangle$ state: One applies a very long laser pulse such that all populations have decayed to the $|0\rangle$ level. Now, in the experiment one can only observe a superposition of the $m_S = 0$ and $m_S = 1$ fluorescence curves. As the

absolute fluorescence varies from time to time, one measures instead a relative fluorescence I_{rel}, namely one devides the integrated fluorescence during the first 200 ns by the fluorescence during 200 ns of the steady state (approximately after 3 μs). If one applies a microwave pulse that drives the qubit transition $|0\rangle \rightarrow |1\rangle$, one can observe how the fluorescence curve changes its shape and oscillates between the curves for $m_S = 0$ and $m_S = 1$. By setting the relative fluorescence for $m_S = 0$ equal to one, one normalizes the fluorescence such that it is equivalent to the population $P(|0\rangle)$ in the level $|0\rangle$. The population is averaged by repeating the experiment about 10^7 times.

1.2.2. Optimal Control with Nitrogen-Vacancy Centers

In this section, we will explain how to do optimal control with an ensemble of NV centers. This leads to the two immediate questions: Why do we use ensembles and why do we need optimal control? After answering those, we will conclude this section by stating the optimal control problem that one has to solve.

1.2.2.1. Why to use ensembles?

In this work, we will deal with an ensemble of non-interacting NV centers. The reason of choosing an NV ensemble rather than a single NV center is two-fold:

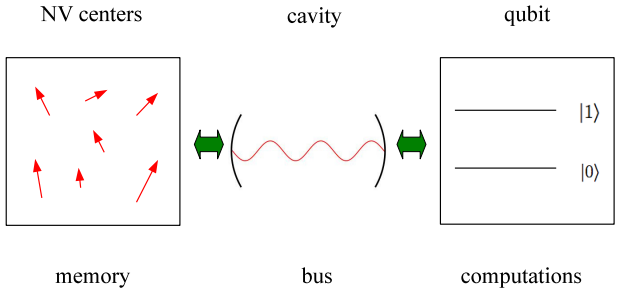

Figure 1.5.: Proposal by Kubo *et al.* [16] to build a hybrid quantum device: NV centers serve as a quantum memory, via the bus (cavity) a state is transported to a superconducting qubit, where quantum computations can be done. After that, the state is swapped back into the memory.

First, NV centers show very long coherence times, *i.e.* they are ideally suited to serve as a quantum memory. However, the coupling to an NV center (or in

general to a spin) is often too weak in order to write or read out a state. If, on the other hand, one uses ensembles of spins, the interaction is enhanced by a factor $\sqrt{N}$, where N is the number of spins [17, 51]. This enhancement can be explained by the coupling to collective excitations [52]. As there are systems in which qubits are easier to manipulate than NV centers, it is only logical to combine different quantum systems in order to build a hybrid quantum circuit [16] (see Fig. 1.5). The hybrid quantum system in [16] consists of a memory, a bus and a processing unit (CPU). The bus serves as a transmitter of the quantum state from the memory to the CPU. As soon as the state is stored in the CPU, it can be manipulated coherently, *i.e.* quantum gate operations can be performed. After these computations, the state is again transferred back via the bus into the memory. In the proposal of Kubo *et al.* [16] as the memory one uses an ensemble of NV centers, as the bus a cavity and as the CPU a superconducting qubit. The levels in this sort of qubit are the two directions (clock- or counterclock-wise) of a superconducting current.

The second reason why one uses ensembles is that spins can serve as sensors to measure *e.g.* small magnetic fields [22]. The signal coming from the spin can be amplified by using an ensemble of spins, since the signal grows with the number of participating spins. On the other hand, the coherence time T_2^* connected to different resonance frequencies of the spins becomes shorter with the number of spins. This is the case because due to stress and impurities the resonance frequencies ω_0 vary from spin to spin. This phenomenon is called *inhomogeneous broadening*. If the control Hamiltonian $\mathcal{H}_c(t)$ is optimized for one specific spin, it is not anymore suited to drive another spin with a different resonance frequency. As the signal increases with the number of spins but the coherence time decreases, there is an optimal number of spins with which one can perform sensing.

1.2.2.2. Why to use optimal control?

In the experiment that Tobias Nöbauer and Andreas Angerer performed [40], an NV ensemble is controlled by an antenna. Each of the non-interacting NV centers can be modeled by a two-level Hamiltonian Eq. (1.75). Due to inhomogeneous broadening, the distribution of the resonance frequencies ω_0 of the NV centers obey a Gaussian law.

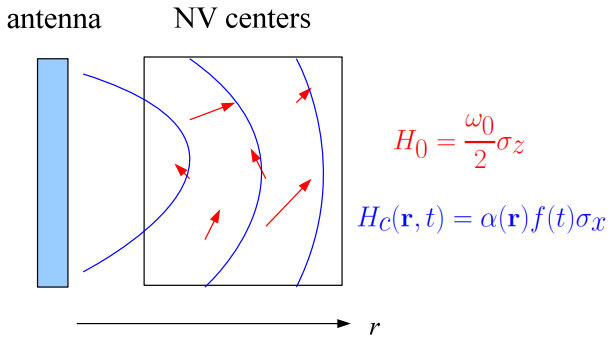

Figure 1.6.: An antenna emits an inhomogeneous control field (different $\alpha(\mathbf{r})$), which can be used to manipulate an inhomogeneously broadened ensemble of NV centers (different ω_0).

The antenna, on the other hand, emits a time-dependent control field

$$\mathcal{H}_c(t) = \cos(\omega_0 t)\left(f_x(t)\sigma_x + f_y(t)\sigma_y\right). \tag{1.76}$$

It is important to note that the control field is classical and not quantum, *i.e.* the Hamiltonian is not expressed in terms of annihilation and creation operators a and $a^\dagger$ of a quantum field. As the NV centers are spatially distributed in a large probe volume V, they experience the inhomogeneous, spatially dependent field of the antenna. It follows that the spins experience a percentage $\alpha(\mathbf{r})\mathcal{H}_c(t)$ of the control field, depending on the spatial position $\mathbf{r}$ of the NV center. The factor α is the larger, the nearer the NV center is to the antenna.

If the NV centers were manipulated by time-independent pulses (hard pulses), the outcome of these pulses would suffer from the fact that they are always designed to work for a precise value of ω_0 and α and fail for other values. One can see this explicitly in an example: At resonance $\omega_0 = 0$ and only control of the x-component ($f_y \equiv 0$) by a rectangular pulse of duration T and strength $f_x \equiv 1$, the target functional of transferring the spin from the state $|0\rangle$ to the state $|1\rangle$ is given by

$$|\langle 1|\exp(-i\alpha T\sigma_x)|0\rangle|^2 = \sin^2(\alpha T). \tag{1.77}$$

Now, if $\alpha = 1$, the rectangular pulse of duration $T = \pi/2$ transfers the spin from $|0\rangle$ to $|1\rangle$. However, the other spins which have another distance from the antenna, *i.e.* another value of α, they only show a fidelity of $\sin^2(\alpha\pi/2)$. What is important in the case of rectangular pulses, is that the motion of the

spin on the Bloch sphere follows a curve of the same longitude. On the Bloch sphere, the state of a qubit is visualized according to

$$|\theta, \phi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \tag{1.78}$$

with the latitude θ (measured from the z-axis) and longitude ϕ (measured from the x-axis). If, however, the pulses were time-dependent, one could design them in a way that each spin follows a different trajectory on the Bloch sphere, but ends in the same point (here $|1\rangle$) at the end of the control. Therefore, a time-dependent pulse can gain robustness with respect to different values of α and ω_0.

1.2.2.3. Rotating Wave Approximation

Up to now, three important approximations have been made: All spins are independent of each other, one spin can be regarded as a two level system, and the field of the antenna is classical. In this section, we will make an additional approximation, the *rotating wave approximation*: If one considers the spin dynamics from a rotating frame

$$\mathcal{U}_0 = e^{-i\mathcal{H}_0 t} = e^{-i\omega_0 t\sigma_z/2} = \mathbb{1}\cos\left(\frac{\omega_0}{2}t\right) - i\sigma_z\sin\left(\frac{\omega_0}{2}t\right), \tag{1.79}$$

one obtains

$$\mathcal{U}_0^\dagger\mathcal{H}\mathcal{U}_0 = \frac{\omega_0}{2}\sigma_z + \cos(\omega_0 t)\sum_{k=1}^{n}\mathbf{A}_k\sin(k\Omega t) \tag{1.80}$$

with the SOC-Hamiltonian $\mathcal{H} = \frac{\omega_0}{2}\sigma_z + \cos(\omega_0 t)\sum_{k=1}^{n}(a_k\sigma_x + b_k\sigma_y)\sin(k\Omega t)$ (see Eqs. (1.75) and (1.76)) and

$$\begin{aligned}\mathbf{A}_k = \quad & (a_k\sigma_x + b_k\sigma_y)\left(\cos^2\left(\frac{\omega_0}{2}t\right) - \sin^2\left(\frac{\omega_0}{2}\right)t\right) \\ & +2\cos\left(\frac{\omega_0}{2}t\right)\sin\left(\frac{\omega_0}{2}t\right)(b_k\sigma_x - a_k\sigma_y)\end{aligned} \tag{1.81}$$

Expanding the trigonometric functions into exponentials yields

$$\mathbf{A}_k = \begin{pmatrix} 0 & (a_k - ib_k)e^{i\omega_0 t} \\ (a_k + ib_k)e^{-i\omega_0 t} & 0 \end{pmatrix}. \tag{1.82}$$

Multiplying with $\cos(\omega_0 t)$ gives

$$\cos(\omega_0 t)\mathbf{A}_k = \begin{pmatrix} 0 & \frac{1}{2}(a_k - ib_k)\left(e^{2i\omega_0 t} + 1\right) \\ \frac{1}{2}(a_k + ib_k)\left(e^{-2i\omega_0 t} + 1\right) & 0 \end{pmatrix} \tag{1.83}$$

Here, one can make the rotating wave approximation: Since one drives the spins near the resonance, one can eliminate the rapidly oscillating terms $e^{\pm 2i\omega_0 t} \approx 0$ that are averaged out in the end. One finally obtains

$$\mathcal{U}_0^\dagger \mathcal{H} \mathcal{U}_0 = \frac{\omega_0}{2}\sigma_z + \frac{1}{2}\sum_{k=1}^{n}(a_k\sigma_x + b_k\sigma_y)\sin(k\Omega t) \tag{1.84}$$

In the following, we will absorb the factor 1/2 in the coefficients a_k and b_k, also we will regard ω_0 as the detuning from the resonance frequency.

1.2.2.4. The Optimal Control Problem

Coming back to the formulation of our optimal control problem, we have now identified the Hamiltonian Eq. (1.84) as underlying the spin dynamics. As we have seen in Sec. 1.2.2.2, the pulses that we have to design are to work for all experimentally relevant values of detunings ω_0 and control amplitudes α. Because of this, one has to average the target functional over ω_0 and α such that a high average target functional means that the pulse works for all parameters α and ω_0 one has averaged over. Concerning the target functional, we have to distinguish between state transfer and gate implementation.

If one aims at a state transfer from $|\psi_i\rangle$ to $|\psi_f\rangle$ in the time t_f, one can use the averaged overlap (also referred to as 'fidelity')

$$\langle \mathcal{F}_{\text{fid}} \rangle = \left\langle |\langle \psi_f | \mathcal{U}(t_f) | \psi_i \rangle|^2 \right\rangle_{\omega_0,\alpha} \tag{1.85}$$

as the target functional, where $\mathcal{U}(t_f)$ is the propagator at time t_f.

If one wishes to implement a unitary transformation $\mathcal{U}_f$, one has to minimize the distance (Hilbert-Schmidt norm $\|\cdot\|_{HS}$) between the actual $\mathcal{U} \equiv \mathcal{U}(t_f)$ and the desired time evolution $\mathcal{U}_f$:

$$\begin{aligned}
\|\mathcal{U} - \mathcal{U}_f\|_{HS}^2 &= \operatorname{Tr}(\mathcal{U} - \mathcal{U}_f)(\mathcal{U}^\dagger - \mathcal{U}_f^\dagger) && (1.86)\\
&= \operatorname{Tr}\left(\mathcal{U}\mathcal{U}^\dagger - \mathcal{U}\mathcal{U}_f^\dagger - \mathcal{U}_f\mathcal{U}^\dagger + \mathcal{U}_f\mathcal{U}_f^\dagger\right) && (1.87)\\
&= 2\operatorname{Tr}\mathbb{1} - 2\operatorname{Re}\operatorname{Tr}\mathcal{U}\mathcal{U}_f^\dagger \equiv 2(d - \operatorname{Re}\langle \mathcal{U}, \mathcal{U}_f\rangle_{HS}), && (1.88)
\end{aligned}$$

where d is the dimension of the system and $\langle\cdot,\cdot\rangle_{HS}$ the Hilbert-Schmidt scalar product. Therefore, one has to maximize the averaged operator fidelity

$$\langle \mathcal{F}_{\text{op}} \rangle = \langle \operatorname{Re}\langle \mathcal{U}, \mathcal{U}_f\rangle_{HS}/d \rangle_{\omega_0,\alpha}\,. \tag{1.89}$$

Before going to a detailed analysis, let us first observe some general symmetry principles of OC pulses:

If the pulse $f_{x,y}(t)$ implements a state transfer from $|\psi_i\rangle$ to $|\psi_f\rangle$, then $f_{x,y}(-t)$ causes the state transfer from $|\psi_f\rangle$ to $|\psi_i\rangle$, since $\mathcal{U}(-t)|\psi_f\rangle = \mathcal{U}(-t)\mathcal{U}(t)|\psi_i\rangle = |\psi_i\rangle$

If a pulse $f_{x,y}(t)$ of length t_f works for an inhomogeneous broadening $\Delta\omega$, *i.e.* if a high fidelity is obtained for all spin frequencies from an inhomogeneously broadened ensemble, then the squeezed pulse $f'_{x,y}(t) = \kappa f_{x,y}(\kappa t)$ copes with an inhomogeneous broadening of $\kappa\Delta\omega$, since the transformation $t = \kappa t'$ entails

$$\begin{aligned} i\frac{\partial}{\kappa\partial t'}|\psi(\kappa t')\rangle &= \left(\frac{\omega_0}{2}\sigma_z + f_x(\kappa t') + f_y(\kappa t')\sigma_y\right)|\psi(\kappa t')\rangle \\ \Leftrightarrow i\frac{\partial}{\partial t'}|\psi(\kappa t')\rangle &= \left(\frac{\kappa\omega_0}{2}\sigma_z + \kappa f_x(\kappa t')\sigma_x + \kappa f_y(\kappa t')\sigma_y\right)|\psi(\kappa t')\rangle. \end{aligned} \quad (1.90)$$

The new fundamental frequency is then $\Omega' = \kappa\Omega$, *i.e.* the pulse ends at $t = t_f/\kappa$. By squeezing the pulse, one can therefore obtain higher fidelities but at the cost of a higher intensity and a broader bandwidth.

If the control Hamiltonians $f_x(t)\sigma_x$ and $f_y(t)\sigma_y$ implement $\mathcal{U}_f$ and the Hamiltonian is transformed by a transformation M ($\mathcal{H}' = M\mathcal{H}M^\dagger$, $MM^\dagger = \mathbb{1}$) and the inhomogeneous broadening is invariant under M, then the transformation $\mathcal{U}'_f = M\mathcal{U}_f M^\dagger$ is implemented by the controls $f_x(t)M\sigma_x M^\dagger$ and $f_y(t)M\sigma_y M^\dagger$. This is because

$$\mathrm{Tr}(\mathcal{U}_f^\dagger\mathcal{U}) = \mathrm{Tr}(\mathcal{U}_f^\dagger M^\dagger M\mathcal{U}M^\dagger M) = \mathrm{Tr}((M\mathcal{U}_f M^\dagger)^\dagger M\mathcal{U}M^\dagger). \quad (1.91)$$

The simplest example is the transformation

$$M = \begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix} \quad (1.92)$$

that exchanges σ_x and σ_y:

$$M\sigma_x M^\dagger = \sigma_y \qquad M\sigma_y M^\dagger = \sigma_x \qquad M\sigma_z M^\dagger = -\sigma_z, \quad (1.93)$$

i.e. if one exchanges $f_x(t)$ and $f_y(t)$, the x- and y-axis are exchanged. In other words, a rotation around the x-axis becomes a rotation around the y-axis and vice versa. It is therefore not necessary to calculate the pulse that implements a rotation around the y-axis if one already has calculated the corresponding rotation around the x-axis. The last equation $M\sigma_z M^\dagger = -\sigma_z$ implies that the transformed pulse works now for an ensemble whose resonance frequencies are

reflected around the point $\omega_0 = 0$. For a symmetric distribution, which is the case in most applications, the transformed and the original ensembles are the same.

1.2.3. Experimental Setup

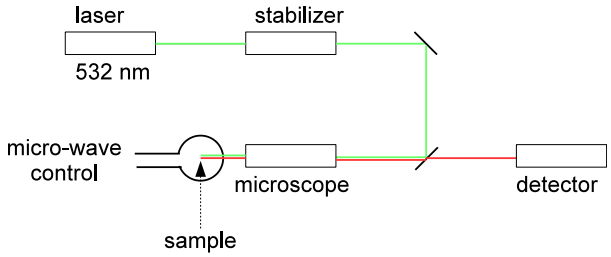

Figure 1.7.: The schematic setup of the NV experiment s depicted [53]: The sample is excited by a 532 nm laser, the fluorescence light is directed onto a detector, while the sample is manipulated by a microwave pulse.

Let us now come to the concrete experimental setup (see Fig. 1.7) we will refer to in chapter 3. The beam of a 532 nm laser is stabilized by an arrangement of acousto-optical modulators, lenses and beam splitters (see [40]). Then, through a microscope, the beam is guided onto the sample. Optimal control can be done with the sample by controlling it with the help of a micro-wave signal. The antenna that emits the control field is a gold wire which has the shape of a loop. Inside this loop, there are the NV centers, which experience the inhomogeneous field of the loop (see Fig. 1.8). The fluorescence light is collected by the microscope and directed onto a detector.

1.2.4. Magnetic Field Sensing

Magnetometry is the ability of measuring small magnetic fields. If one deals with magnetometry, one has to distinguish the cases of DC (static) and AC (time-dependent) fields. Both have in common that they measure a phase shift that results from the application of a magnetic field. In the following, a magnetic field B is assumed to be applied along the z-axis. In the case of a DC field (see Fig. 1.9), one first transfers the spin from the state $|0\rangle$ to the state $(|0\rangle + |1\rangle)/\sqrt{2}$ (called a $\pi/2$-pulse) [53]. On the Bloch sphere, a $\pi/2$

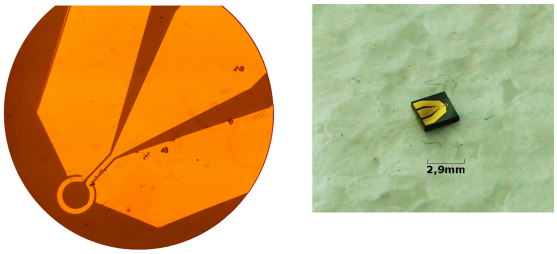

Figure 1.8.: On the left: a gold wire forms the antenna. Inside the loop, NV centers are to be manipulated. On the right: the black diamond containing NV centers with the antenna on the top (photos taken by T. Nöbauer [53]).

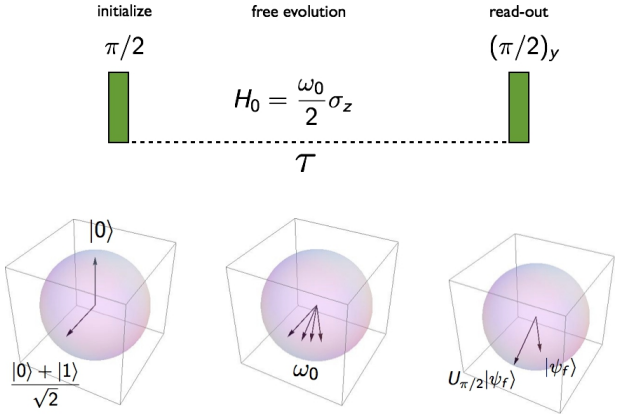

Figure 1.9.: Scheme for measuring DC magnetic fields: After a $\pi/2$-pulse, the spins precess in the equatorial plane of the Bloch sphere (see below) due to their resonance frequency ω_0. The read-out is done by applying a $(\pi/2)_y$-pulse.

pulse corresponds to the mapping of the north pole to a point in the equatorial plane of the Bloch sphere (left sphere in Fig. 1.9). Then, one lets the spins precess in the equatorial plane for a time τ. This evolution follows the internal dynamics $e^{-i\omega_0\tau\sigma_z/2}$ (second sphere in Fig. 1.9) with the resonance frequency ω_0 of the spin. The read-out consists of a $(\pi/2)_y$-pulse and a projection on the state $|1\rangle$ (right sphere in Fig. 1.9). The above sequence is also called a Ramsey sequence [54]. The result of this sequence is the fidelity

$$\left|\langle 1|e^{-i\pi\sigma_y/4}e^{-i\omega_0\tau\sigma_z/2}\left(|0\rangle + |1\rangle\right)/\sqrt{2}\right|^2 = \cos^2\left(\frac{\omega_0\tau}{2}\right) = \frac{1}{2}\left(1+\cos(\omega_0\tau)\right). \tag{1.94}$$

An additional static B field in form of a Hamiltonian $H_B = B\sigma_z$ will cause

an additional shift $e^{-iB\tau\sigma_z}$. If one considers the output $P(|1\rangle)$ as a function of τ, one can see that the magnetic field causes a higher frequency. From this frequency shift $\Delta\phi = B\tau$ one can calculate the value of B as $B = \Delta\phi/\tau$.

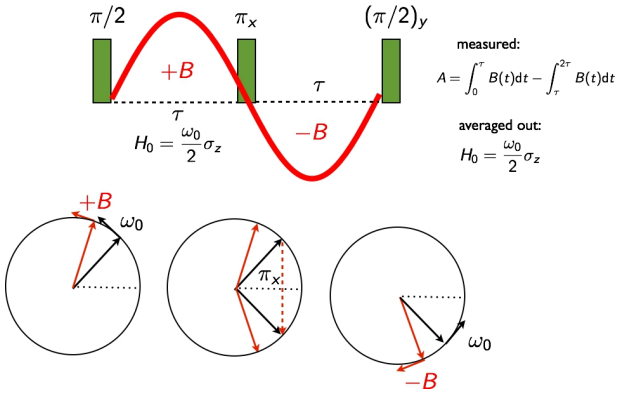

Figure 1.10.: Scheme for measuring AC magnetic fields: A π_x-pulse would lead to a refocusing of the spins after the second free evolution. However, an AC magnetic field, as triggered above, induces an additional phase shift A.

In the case of an AC magnetic field, one cannot use the above scheme because fields would be averaged out over time. Instead, one adds another π_x-pulse (on the Bloch sphere the rotation by π around the x-axis) in between the free evolution (see Fig. 1.10) [55]. This is what is called a (Hahn) spin echo sequence. Let us first discuss the pulse scheme without magnetic field. The idea of spin echo is that during the first evolution the spins precess with their frequency ω_0 (see black arrow in the left sphere of Fig. 1.10), then they are rotated around the x-axis by an angle of π (second sphere in Fig. 1.10), and during the second free evolution they exactly get back the distance they had covered during the first evolution (black arrow in the right sphere of Fig. 1.10), so that they are finally refocused after a time τ. The mechanism of spin echo can be illustrated by runners in a stadium. They start to run as fast as possible, everyone at his own speed (free evolution). Then, on a whistle (the π_x-pulse), they run back to the point where they came from. Now, a fast runner has covered a long distance, but has also a long distance to get back. On the other hand, a slower runner has covered a smaller distance and also has to get back a smaller distance. So, in the end, all runners will come back at the same time.

Now, the magnetic field comes into play: One assumes that this field is periodic in time and that its periodicity is adapted to the control sequence in

a way that it is positive during the first free evolution (see left sphere in Fig. 1.10) and negative during the second (right sphere in Fig. 1.10). This can be achieved experimentally by suitably triggering the control sequence [53]. Therefore, the spins will obtain a positive dephasing $+B\tau$ during the first interval and a negative dephasing $-B\tau$ during the second evolution. In this way one will have a phase shift of

$$\Delta\phi = \int_0^\tau B(t)\mathrm{d}t - \int_\tau^{2\tau} B(t)\mathrm{d}t. \tag{1.95}$$

By this method, one thus measures the lowest Fourier component of the time-dependent field $B(t)$.

In order to assess the quality of the above pulse scheme in the presence of experimental imperfections, let us next calculate the sensitivity $\delta B_{\min}$. This is the minimal magnetic field δB_{min} that permits us to distinguish a field B_1 from the field $B_1 + \delta B_{min}$ by performing asymptotically many measurements. These measurements are limited by the standard deviation σ_{sn} of the shot-noise, which is always present when counting the number of photons. Therefore, the signal dS that permits us to distinguish between two magnetic fields has to be larger than the experimental shot-noise. The sensitivity $\delta B_{\min}$ is thus defined as the limiting case where the signal dS can still be distinguished from shot-noise (signal-to-noise ratio = 1):

$$\delta B_{\min} = \frac{\sigma_{sn}}{\mathrm{d}S}. \tag{1.96}$$

Here, $\mathrm{d}S = \left|\frac{\partial S}{\partial B_0}\right|$ is the slope of the signal as a function of the magnetic field [40]. According to Eq. (1.94), the fidelity after a free evolution with a phase shift $\Delta\phi$ can be written as $\mathcal{F}_0 + S(\Delta\phi)$. One therefore defines the *spin echo signal* as $S = CN_{\mathrm{tot}}\cos(\Delta\phi)$, where C is the spin-echo contrast and N_{tot} the total number of collected photons. In order to get a rough estimate of the sensitivity, one takes as $B(t)$ not a sine function but a rectangular one. Since the surface under a rectangular pulse is greater than the surface under any other pulse of equal duration, a rectangular pulse provides the maximal attainable sensitivity. As the signal has the value B_0 during the first free evolution τ and $-B_0$ during the second free evolution, this gives $\Delta\phi = 2B_0\tau$. It follows that

$$\left|\frac{\partial S}{\partial B_0}\right| = 2CN_{\mathrm{tot}}\tau\sin(2B_0\tau). \tag{1.97}$$

Again, we are interested in the highest signal, *i.e.* $\sin(2B_0\tau) = 1$. Since the shot-noise represents a high number N_{tot} of independent events, it can be modeled as a Poissonian process, *i.e.* $\sigma_{sn} = \sqrt{N_{\text{tot}}}$. As the total number of collected photons depends on the measurement time T and volume V, it seems reasonable to divide by these quantities such that the normalized sensitivity reads

$$\delta B_{\text{min}} = \frac{1}{C\tau\sqrt{N_{\text{tot}}TV}} \tag{1.98}$$

One can see that the longer the free precession time τ is, the more photons are collected and the better is the sensitivity δB_{min}. Unfortunately, the spin echo contrast C deminishes with τ on the time scale of the decoherence time T_2 [55]. Therefore, an optimal precession time τ_{opt} exists. The spin echo contrast diminishes approximately as [56]

$$C(\tau) = C_0 e^{-\left(\frac{\tau}{T_2}\right)^n}, \tag{1.99}$$

where $n \in [1, 2]$ is a fitting parameter. In order to yield the optimal precession time, we calculate the derivative of δB_{min} with respect to τ, which is equivalent to solving the following equation:

$$\frac{e^{\left(\frac{\tau_{\text{opt}}}{T_2}\right)^n}\left(n\left(n\left(\frac{\tau_{\text{opt}}}{T_2}\right)^n - 1\right)\right)}{\tau^2} = 0 \tag{1.100}$$

The solution of this equation is

$$\tau_{\text{opt}} = \frac{T_2}{n^{1/n}} \tag{1.101}$$

A common value of this is $\tau \approx 0.5T_2$, as we will see soon.

2. Smooth Optimal Control with Floquet Theory

In this chapter, we will describe the method of Smooth Optimal Control (SOC) that we have developed. In Sec. 1.1.2 we already explained how Floquet theory can be used to compute the time evolution operator, which is the first step of our algorithm (see Fig. 1.1). In this chapter, we will present a new technique with which one can compute the derivatives with respect to the control parameters with the help of perturbation theory. The chapter will be concluded with the description of the concurrent update of the control amplitudes. In addition, the reader can find in the appendix A some remarks about the numerical calculation of the Floquet eigenvalues.

In the framework of SOC, the pulse is expanded into a Fourier series with a fundamental frequency Ω, *i.e.* the functions in Eq. (1.4) are

$$g_{jk}(t) = \sin(j\Omega t). \tag{2.1}$$

From this follows that the Hamiltonian is periodic in time, $\mathcal{H}(t+T) = \mathcal{H}(t)$, with the period $T = 2\pi/\Omega$. We therefore can invoke Floquet's theorem in order to calculate the time evolution operator. In contrast to GRAPE, where the above functions were discontinuous step functions, in our framework these functions are smooth - which motivates the notion of *smooth* optimal control. Whereas GRAPE is discrete in time, SOC is discrete in frequency space. This discreteness bears some important advantages, as we will see later.

One may argue that working with periodic Hamiltonians is a big restriction. In practice, however, the duration t_f of the pulse can always be chosen such that $\mathcal{H}(0) = \mathcal{H}(t_f) = 0$, which can be achieved by switching the pulse on at $t = 0$ and switching it off at $t = t_f$. In this way, the period T of the Hamiltonian, *i.e.* the time T for which $\mathcal{H}(t) = \mathcal{H}(t+T)$ is fulfilled, naturally coincides with the duration t_f of the pulse. It has to be mentioned that, since

the Hamiltonian for times $t > t_f$ is of no interest, one can periodically continue $\mathcal{H}(t)$. The condition of dealing with a time-periodic Hamiltonian therefore has been achieved without any restriction in practice.

2.1. Derivatives with Perturbation Theory

Whereas with GRAPE it is still manageable to calculate the first derivatives of the propagator with respect to the control parameters, the calculation of the second derivatives entails a lot of effort and is only possible numerically [28]. With SOC, on the contrary, one is capable with little effort to compute any order of the derivatives analytically. This will be shown in the following.

In order to compute the derivatives of the time evolution operator $\mathcal{U}$ with respect to the control parameters a_{jk}, which is the second step of our algorithm from Fig. 1.1, one has to know how the Floquet eigenvalues and eigenvectors change under a modification of the control amplitudes. As the Hamiltonian $\mathcal{H}$ depends linearly on the control parameters, so does the Floquet operator $\tilde{\mathcal{K}}$:

$$\tilde{\mathcal{K}}(\mathbf{a} + \delta\mathbf{a}) = \tilde{\mathcal{K}}(\mathbf{a}) + \delta\mathbf{a} \cdot \mathbf{K}_{\text{int}}, \tag{2.2}$$

where $\mathbf{K}_{\text{int}} = \nabla_{\mathbf{a}} \tilde{\mathcal{K}}|_{\mathbf{a}=0}$. One can now do perturbation theory in Floquet space with the unperturbed matrix $\tilde{\mathcal{K}}(\mathbf{a})$ and the small perturbation $\delta\mathbf{a} \cdot \mathbf{K}_{\text{int}}$. It is important to note that perturbation theory in Floquet space does not differ conceptually from perturbation theory in the Schrödinger picture, the only difference is that in Floquet space one has to deal with infinitely large matrices.

The first step of perturbation theory is to expand the perturbed eigenvalues and eigenvectors in powers of the perturbation parameter $\delta\mathbf{a}$,

$$\epsilon_k(\mathbf{a} + \delta\mathbf{a}) \approx \epsilon_k(\mathbf{a}) + \delta\mathbf{a} \cdot \boldsymbol{\epsilon}_k^{(1)} = \epsilon_k(\mathbf{a}) + \delta\mathbf{a} \cdot \nabla_{\mathbf{a}} \epsilon_k, \tag{2.3}$$

with $\boldsymbol{\epsilon}_k^{(1)}$ being the first perturbative correction. It results from this expansion that the m-th derivatives with respect to the control parameters are nothing else but the m-th perturbative corrections in Floquet space. In the following, for the ease of notation, we will replace the two indices i and j in a_{ij} by one index j. If one expands eigenvalues and eigenvectors in a series in δa_j, one obtains

$$\left(\tilde{\mathcal{K}} + \sum_j \delta a_j K_{\text{int}}^{(j)}\right) \sum_{\alpha \in \mathbb{N}_0^n} (\delta\mathbf{a})^\alpha |\chi_k^{(\alpha)}\rangle = \sum_{\alpha \in \mathbb{N}_0^n} (\delta\mathbf{a})^\alpha \epsilon_k^{(\alpha)} \sum_{\beta \in \mathbb{N}_0^n} (\delta\mathbf{a})^\beta |\chi_k^{(\beta)}\rangle. \tag{2.4}$$

Collecting all first order terms in δa_j yields:

$$\tilde{\mathcal{K}}|\chi_k^{(1_j)}\rangle + K_{\text{int}}^{(j)}|\chi_k^{(0)}\rangle = \epsilon_k^{(1_j)}|\chi_k^{(0)}\rangle + \epsilon_k^{(0)}|\chi_k^{(1_j)}\rangle \tag{2.5}$$

Here, $\epsilon_k^{(0)} \equiv \epsilon_k$ and $|\chi_k^{(0)}\rangle \equiv |\chi_k\rangle$ are the unperturbed eigenvalues and eigenvectors, and $\epsilon_k^{(1_j)}$ and $|\chi_k^{(1_j)}\rangle$ are the first perturbative expressions with perturbation in the j-th coordinate. By multiplying with $\langle\chi_k^{(0)}|$ and taking $\langle\chi_k^{(0)}|\chi_k^{(\alpha)}\rangle = \delta_{0\alpha}$, one obtains:

$$\epsilon_k^{(1_j)} = \frac{\partial \epsilon_k}{\partial a_j} = \langle\chi_k|K_{\text{int}}^{(j)}|\chi_k\rangle \tag{2.6}$$

Expressing

$$|\chi_k^{(1_j)}\rangle = \sum_m \langle\chi_m^{(0)}|\chi_k^{(1_j)}\rangle|\chi_m^{(0)}\rangle \tag{2.7}$$

in terms of $|\chi_m^{(0)}\rangle$ and inserting in Eq. (2.5) gives

$$\langle\chi_m^{(0)}|\chi_k^{(1_j)}\rangle = \begin{cases} \frac{\langle\chi_m|K_{\text{int}}^{(j)}|\chi_k\rangle}{\epsilon_k - \epsilon_m} & k \neq m \\ 0 & k = m \end{cases} \tag{2.8}$$

If one introduces the (pseudo-)inverse

$$\mathcal{I}_k := \sum_{n\neq k} \frac{1}{\epsilon_n - \epsilon_k}|\chi_n\rangle\langle\chi_n| \tag{2.9}$$

of $\tilde{\mathcal{K}} - \epsilon_k$ in the space orthogonal to $|\chi_k\rangle$ and the operator $\mathcal{T}_{kj} := \mathcal{K}_{\text{int}}^{(j)} - \frac{\partial \epsilon_k}{\partial a_j}\mathbb{1}$, it results that $|\chi_k^{(1_j)}\rangle$ can be written in the form

$$\frac{\partial|\chi_k\rangle}{\partial a_j} = -\mathcal{I}_k\mathcal{T}_{kj}|\chi_k\rangle. \tag{2.10}$$

The form Eq. (2.9) implies that the eigenvalues ϵ_n are non-degenerate. It is worth noting that the term $\mathcal{I}_k\frac{\partial\epsilon_k}{\partial a_j}|\chi_k\rangle$ vanishes because $|\chi_k\rangle$ is by construction orthogonal to $|\chi_n\rangle$ for all $n \neq k$. However, we keep the current definition because $\mathcal{T}_{kj}$ will later be applied to other vectors.

Let us now proceed to the second order perturbation theory. If one collects in Eq. (2.4) all second order terms proportional to $\delta a_j \delta a_l$, one obtains

$$\begin{aligned} & \tilde{\mathcal{K}}|\chi_k^{(1_j,1_l)}\rangle + K_{\text{int}}^{(j)}|\chi_k^{(1_l)}\rangle + K_{\text{int}}^{(l)}|\chi_k^{(1_j)}\rangle \\ = \;& \epsilon_k^{(1_j,1_l)}|\chi_k^{(0)}\rangle + \epsilon_k^{(1_j)}|\chi_k^{(1_l)}\rangle + \epsilon_k^{(1_l)}|\chi_k^{(1j)}\rangle + \epsilon_k^{(0)}|\chi_k^{(1_j,1_l)}\rangle \end{aligned} \tag{2.11}$$

By projecting on $\langle\chi_k|$ it follows that

$$\begin{aligned} \frac{\partial^2\epsilon_k}{\partial a_j \partial a_l} &= \langle\chi_k|K_{\text{int}}^{(j)}|\chi_k^{(1_l)}\rangle + \langle\chi_k|K_{\text{int}}^{(l)}|\chi_k^{(1_j)}\rangle \\ &= -\langle\chi_k|\left(\mathcal{T}_{kj}\mathcal{I}_k\mathcal{T}_{kl} + \mathcal{T}_{kl}\mathcal{I}_k\mathcal{T}_{kj}\right)|\chi_k\rangle. \end{aligned} \tag{2.12}$$

Expanding again $|\chi_k^{(1_j,1_l)}\rangle$ in terms of the basis $|\chi_m^{(0)}\rangle$ yields

$$|\chi_k^{(1_j,1_l)}\rangle = \sum_m \langle\chi_m^{(0)}|\chi_k^{(1_j,1_l)}\rangle|\chi_m^{(0)}\rangle. \tag{2.13}$$

By replacing $|\chi_k^{(1_j,1_l)}\rangle$ from the latter equation in Eq. (2.11), one obtains

$$\langle\chi_m^{(0)}|\chi_k^{(1_j,1_l)}\rangle = \begin{cases} \frac{\langle\chi_m^{(0)}|K_{\text{int}}^{(j)}|\chi_k^{(1_l)}\rangle + \langle\chi_m^{(0)}|K_{\text{int}}^{(l)}|\chi_k^{(1_j)}\rangle - \epsilon_k^{(1_j)}\langle\chi_m^{(0)}|\chi_k^{(1_l)}\rangle - \epsilon_k^{(1_l)}\langle\chi_m^{(0)}|\chi_k^{(1_j)}\rangle}{\epsilon_k^{(0)} - \epsilon_m^{(0)}} & k \neq m \\ 0 & k = m \end{cases}. \tag{2.14}$$

At the second order, an issue occurs that is related to normalization. While the vector $|\chi_m\rangle$ is still normalized at first order,

$$\begin{aligned} & \left(\langle\chi_k^{(0)}| + \delta a\langle\chi_k^{(1)}|\right)\left(|\chi_k^{(0)}\rangle + \delta a|\chi_k^{(1)}\rangle\right) && (2.15)\\ = \;& \langle\chi_k^{(0)}|\chi_k^{(0)}\rangle + 2\delta a\,\mathrm{Re}\langle\chi_k^{(1)}|\chi_k^{(0)}\rangle + O((\delta a)^2) && (2.16)\\ = \;& 1 + O((\delta a)^2) && (2.17)\end{aligned}$$

it is not anymore at second order, where we have

$$\left\|\left(|\chi_k^{(0)}\rangle + \delta a|\chi_k^{(1)}\rangle + (\delta a)^2|\chi_k^{(2)}\rangle\right)\right\|^2 = 1 + (\delta a)^2\underbrace{\langle\chi_k^{(1)}|\chi_k^{(1)}\rangle}_{=1} + O((\delta a)^3) \neq 1. \tag{2.18}$$

In order to normalize $|\chi_k^{(2)}\rangle$, we start from the general expression $|\chi_k^{(2)}\rangle = \sum_j c_j|\chi_j^{(0)}\rangle$: For $j \neq k$, the coefficients c_j are determined by Eq. (2.14), whereas c_k is fixed by normalization. Since $|\delta\chi_k^{(2)}\rangle = c_k|\chi_k^{(0)}\rangle$ is orthogonal to $|\chi_j^{(0)}\rangle$ for all $j \neq k$, it does not alter the result of perturbation theory and can be seen as a correction term ensuring normalization. By using Eq. (2.18), the calculation leads to the condition

$$\langle\delta\chi_k^{(1_j,1_l)}|\chi_k^{(0)}\rangle + \langle\chi_k^{(0)}|\delta\chi_k^{(1_j,1_l)}\rangle + \langle\chi_k^{(1_j)}|\chi_k^{(1_l)}\rangle + \langle\chi_k^{(1_l)}|\chi_k^{(1_j)}\rangle = 0. \tag{2.19}$$

It follows that the correction term can be written as

$$|\delta\chi_k^{(1_j,1_l)}\rangle = -\langle\chi_k^{(1_j)}|\chi_k^{(1_l)}\rangle|\chi_k^{(0)}\rangle. \tag{2.20}$$

Using Eq. (2.10) one can finally write

$$\begin{aligned}\frac{\partial^2|\chi_k\rangle}{\partial a_{ij}\partial a_l} \;=\; & \left(\mathcal{I}_k\mathcal{T}_{kj}\mathcal{I}_k\mathcal{T}_{kl} + \mathcal{I}_k\mathcal{T}_{kl}\mathcal{I}_k\mathcal{T}_{kj}\right)|\chi_k\rangle \\ & -\frac{1}{2}\langle\chi_k|\left(\mathcal{T}_{kj}\mathcal{I}_k^2\mathcal{T}_{kl} + \mathcal{T}_{kl}\mathcal{I}_k^2\mathcal{T}_{kj}\right)|\chi_k\rangle|\chi_k\rangle. \end{aligned} \tag{2.21}$$

If one desires to calculate the derivative or curvature along a specific (normalized) direction $\mathbf{e}$, one has to replace the direction $\delta\mathbf{a}$ in Eq. (2.2) by the direction $\delta a\,\mathbf{e}$, such that the small perturbation of the Floquet matrix becomes $\delta a\,\mathbf{e}\cdot\mathbf{K}_{\text{int}}$. As we will see in detail in the next section, the speed of convergence with which the OC algorithm finds an optimal value of the target functional is the higher the more information is known about the derivatives. In this sense, the curvature along the direction of the gradient is an important quantity which can be evaluated in the SOC framework with nearly no additional overhead, whereas it is not clear how this can be done in the framework of GRAPE.

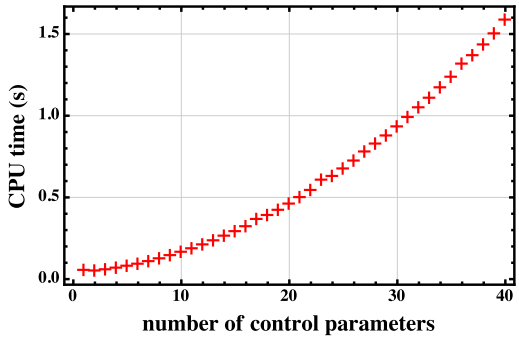

Figure 2.1.: The time needed to calculate the derivatives with respect to the control parameters is depicted as a function of the number of control parameters. The more control parameters one uses, the more time is needed. The increase is a cross-over between polynomial and exponential behavior.

To conclude this section, we will measure the time consumed in determining the derivatives. It is clear that using second order perturbation theory is more time-consuming than calculating only the first derivatives. In general, as the time needed to compute the derivatives depends strongly on the number of control parameters, it is recommendable to determine beforehand how many frequency components one really requires. As the dimension of the Hilbert space increases exponentially with the number of involved particles, the question of the required number of frequency components is even more delicate in the case of Hamiltonians of more than one spin. In Fig. 2.1, we plotted the CPU time needed to calculate the derivatives with respect to the control parameters as a function of the number of control parameters. We used the

test Hamiltonian

$$\mathcal{H} = \underbrace{\frac{\omega_0}{2}\sigma_z}_{\mathcal{H}_0} + \underbrace{\sum_{k=1}^{n_{\max}} a_k \sin(kt)\sigma_x}_{\mathcal{H}_c(t)} . \tag{2.22}$$

We have set $\omega_0 = 0.3592$, and the control amplitudes a_k were chosen randomly for each $n_{\max}$ from an interval of $[-1, 1]$. As one can see, the CPU time increases rapidly with increasing number of control parameters. If one compares the computational effort with existing optimal control routines like GRAPE, one has to state that the calculation of the derivative with respect to a *single* control parameter can be performed with very low effort, but on the other hand there are hundreds or thousands of such parameters, whereas SOC only necessitates the computation of around ten (first) derivatives but with more effort for each derivative. From this fact one can expect that the computational effort of GRAPE and SOC is approximately comparable.

From a technical point of view, in MATLAB, the Penrose inverse *pinv* provides an efficient algorithm to compute the pseudo-inverse Eq. (2.9), which is about ten times faster than a direct computation of the operator $\mathcal{I}_k$ by hand. It has to be said that from all steps of the algorithm, the computation of the derivatives consumes by far the most time compared to the other steps. It is therefore of great interest to implement the computation of the derivatives efficiently. This efficient implementation is obtained by using the operators $\mathcal{I}_k$ and $\mathcal{T}_{kj}$ - which only have to be calculated once and reappear at the second order perturbation theory. Furthermore, what is also time consuming is the evaluation of $K_{\text{int}}^{(j)}$, but fortunately this has only to be calculated when the size of the truncated Floquet matrix is changed, as it only depends on the structure of the Hamiltonian and not on the explicit values of the control amplitudes a_j.

2.2. Concurrent Update

We will now proceed to the last step of our algorithm from Fig. 1.1, namely the update of the pulses. There are numerous ways to update the pulse in order to get a better fidelity in each iteration step. The most intuitive method is the method of steepest ascent, where the new pulse $\mathbf{a}'$ lies on the line $\mathbf{a} + \alpha\nabla\mathcal{F}$. The parameter α, *i.e.* the distance covered in the direction of the gradient $\nabla\mathcal{F}$, can be determined *e.g.* by a direct line search, where α is enlarged if the fidelity is improved and reduced if the fidelity is deminished.

The speed of convergence can be enhanced by exploiting information about the second derivatives with respect to the control parameters. For example, the step size of a gradient-based optimization can be approximated by exploiting the curvature along the direction $\mathbf{e}_{\nabla\mathcal{F}} = \nabla\mathcal{F}/\left\|\nabla\mathcal{F}\right\|$ of the gradient $\nabla\mathcal{F}$. In this case, as explained in the previous section, the expression $\delta\mathbf{a}\cdot\mathbf{K}_{\text{int}}$ in Eq. (2.2) has to be replaced by $\delta a\,\mathbf{e}_{\nabla\mathcal{F}}\cdot\mathbf{K}_{\text{int}}$. This computation is by far easier than in the framework of GRAPE, where all second derivatives have to be calculated in order to determine a directional derivative at second order. It has to be mentioned that second-order GRAPE has only been developped recently [28]. Concerning the update of the pulse, one has to distinguish between a convex and concave landscape. If the curvature $d^2\mathcal{F}_{gr}$ along the gradient is negative, *i.e.* one approaches a local maximum, then the landscape of the fidelity in the neighborhood of $\mathbf{x}_0 = \mathbf{a}$, where $\mathbf{a}$ are the control parameters, can be approximated by a (convex) quadratic function $f(\mathbf{x}) = d^2\mathcal{F}_{gr}(\mathbf{x}-\mathbf{x}_0)^2 + \nabla\mathcal{F}\cdot(\mathbf{x}-\mathbf{x}_0)+c$. If the landscape was really quadratic, the maximum could be found by setting the first derivative of $f(\mathbf{x})$ equal to zero. This results in the condition $\nabla f(\mathbf{x}) = 2d^2\mathcal{F}_{gr}(\mathbf{x}-\mathbf{x}_0)+\nabla\mathcal{F} = 0$. The maximum would therefore lie at $\mathbf{x} = \mathbf{x}_0 - \frac{\nabla\mathcal{F}}{2d^2\mathcal{F}_{gr}}$, *i.e.* the update of the pulse for a negative curvature is

$$\mathbf{a}' = \mathbf{a} + \frac{\nabla\mathcal{F}}{2\left|d^2\mathcal{F}_{gr}\right|}. \tag{2.23}$$

If, on the other hand, the curvature is positive, *i.e.* the landscape is concave, one still has to update the pulse along the direction of the gradient. If one again models the landscape by a quadratic function $f(\mathbf{x}) = d^2\mathcal{F}_{gr}(\mathbf{x}-\mathbf{x}_0)^2 + \nabla\mathcal{F}\cdot(\mathbf{x}-\mathbf{x}_0)+c$, one obtains the condition $d^2\mathcal{F}_{gr}(\mathbf{x}-\mathbf{x}_0)^2 \ll \nabla\mathcal{F}\cdot(\mathbf{x}-\mathbf{x}_0)$ such that the approximation to second order is valid. This condition results in a constraint on $\mathbf{x}-\mathbf{x}_0$:

$$\mathbf{x}-\mathbf{x}_0 \ll \frac{\nabla F}{d^2\mathcal{F}_{gr}} \tag{2.24}$$

For a concave landscape, we therefore update the pulse in the following way:

$$\mathbf{a}' = \mathbf{a} + \delta\frac{\nabla\mathcal{F}}{d^2\mathcal{F}_{gr}} \tag{2.25}$$

The parameter δ specifies how strongly the update deviates from the original landscape. For our purposes, we have chosen $\delta = 0.05$, meaning a deviation of 5 %.

For the case in which the full curvature $\frac{\partial^2 \mathcal{F}}{\partial a_j \partial a_l}$ is accessible, we developed a more sophisticated method: First of all, the step size should depend on the curvature. Again, one has to distinguish between regions of positive and negative curvature. In order to know in which directions the landscape is convex and in which concave, one has to diagonalize the matrix of curvature $K_{jl} = \frac{\partial^2 \mathcal{F}}{\partial a_j \partial a_l}$ and consider their eigenvectors $|k_j\rangle$ and eigenvalues κ_j. Next, it is helpful to construct the operators

$$P = \sum_j \theta(\epsilon - \kappa_j)|k_j\rangle\langle k_j| \quad \text{and} \quad Q = \sum_j \theta(\kappa_j - \epsilon)|k_j\rangle\langle k_j|. \tag{2.26}$$

They project on the directions where the curvature is lower or higher than a certain threshold ϵ. This threshold should be adapted to the fidelity reached, because the approximation to second order is the more accurate the nearer one gets to a local maximum, and these local maxima are characetrized by high fidelities. In this work, we used the following values for ϵ:

$\mathcal{F}$	ϵ
<0.6	-10
$[0.6, 0.8)$	-1
$[0.8, 0.95)$	-0.1
≥ 0.95	-0.01

We first consider the update for a convex landscape: Similar to Eq. (2.23), but where the matrix of curvature K replaces the curvature along the gradient and the operator P projects onto the directions where the landscape is convex, one updates the control amplitudes $\mathbf{a}$ according to

$$\mathbf{a}' = \mathbf{a} + \frac{1}{2} P K^{-1} \nabla \mathcal{F} \tag{2.27}$$

if $\operatorname{Tr} P > \delta$. Here, $\delta = 0.1$ is a small parameter ensuring that the curvature is substantially greater than zero. This parameter has been introduced heuristically, as we have discovered that convergence is the worse the flatter the landscape is. Being in the concave part of the landscape ($\operatorname{Tr} Q > \delta$) entails the following update rule

$$\mathbf{a}' = \mathbf{a} + \alpha \frac{\nabla \mathcal{F} \cdot Q \cdot \nabla \mathcal{F}}{|\nabla \mathcal{F} \cdot QK \cdot \nabla \mathcal{F}|} Q \cdot \nabla \mathcal{F} \tag{2.28}$$

Here, we used $\alpha = 0.1$, specifying the deviation from the original landscape. This update rule can be explained by the same reasoning as in Eq. (2.24), but

with $(\nabla\mathcal{F}\cdot Q\cdot\nabla\mathcal{F})Q\cdot\nabla\mathcal{F}$ replacing the gradient and $|\nabla\mathcal{F}\cdot QK\cdot\nabla\mathcal{F}|$ replacing the curvature along the gradient.

In the following, we want to compare the three update methods explained above. As a test model, we optimized a system of 20 non-interactive spins governed by the Hamiltonian Eq. (1.84) (with $\Omega = 1$ and $n = 5$ Fourier modes) and showing a Gaussian inhomogeneous broadening of 1 MHz full width at half maximum and a control amplitude ranging from $\alpha = 0.75$ to $\alpha = 1.25$. As the target functional we used the average fidelity Eq. (1.85). As one can see in Fig. 2.2, in the beginning, *i.e.* where the control landscape is rather flat, the method using the first derivative only is the fastest one (here we used a constant step size $\alpha = 1$). However, after 30 iteration steps the convergence slows down significantly and after 40 iteration steps it is outperformed by the method using the first and second derivatives. The latter method is obviously the best choice if the landscape is curved. Here, one can fully exploit the additional information about the curvature provided by the second derivatives. Concerning the method using the first derivative and the curvature along the direction of the gradient, it is the slowest one in the beginning. However, after 60 iteration steps it speeds up considerably and attains a convergence speed comparable to the method that uses the full information about the curvature. If one takes the best of each method, one would begin with a search only using the first derivative, and as the control landscape becomes curved one would proceed by using in addition the curvature along the direction of the gradient, as this is computationally less expansive than calculating the full curvature. There is certainly more than one criterion when switching from the optimization using the first derivative only to the method that uses the curvature along the direction of the gradient. Here, we switched when the fidelity did not substantially increase over five iteration steps (here: when it is less than 0.002). One can see in Fig. 2.2 that the hybrid method attains a fidelity comparable to the method that uses the curvature along the gradient after 100 iteration steps. However, the hybrid method has consumed less computational time during the first 40 iteration steps, where the first derivative only has been used.

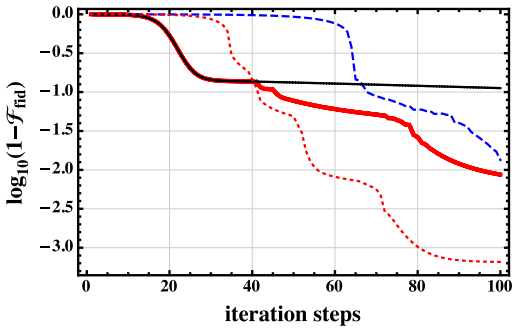

Figure 2.2.: Convergence using the first derivative only (black continuous line), using the first derivative and the curvature along the direction of the gradient (blue dashed line), using the first and second derivatives (red dotted line) and using a combination of first derivative and curvature along the gradient (red, bold line).

Concluding remarks

After having introduced SOC, we would like to describe the advantages that characterize SOC in comparison with existing state of the art techniques: First of all, one has the full control of the spectral width, as it is set by the choice of the Fourier components used. This distinguishes the SOC approach from methods where the spectral width is artificially limited a posteriori by a filter function. As we have already seen in Fig. 2.2 and as we will see in more detail in the next chapter, comparatively few (typically: 10) optimization parameters are required in order to yield a very good result of the optimization. The very restricted number of parameters gives hope that one can explain why certain pulses work better than others. This question is very difficult to be answered in state-of-the-art techniques like GRAPE, where hundreds or thousands of control parameters are needed. Another advantage of SOC is due to Floquet theory, that is many calculations can be performed analytically, *e.g.* non-linear target functionals can be implemented without additional effort (see Sec. 4.3), time-optimal pulses are easily implemented (see Sec. 4.1) and also averaging over time is performed fairly easily (see Sec. 4.4). In the framework of GRAPE, these quantities are only numerically accessible.

3. Control of Single Spin Dynamics

The following chapter is dedicated to the coherent control of a single spin (or an ensemble of non-interacting spins). Most of the material presented in this chapter results from a collaboration with the experimentalists Tobias Nöbauer, Andreas Angerer and Johannes Majer from the Atominstitut Vienna. Experiment and theory are then compared with each other. The main findings are published in [40]. The main goal of this chapter is to measure small magnetic fields with an ensemble of NV centers. The unavoidable inhomogeneous broadening of such ensembles will be compensated with suitably designed pulses, and we will show that the pulses designed with SOC clearly outperform typically employed hard pulse sequences and that they permit us to compensate inhomogeneous broadening essentially perfectly despite their narrow spectral structure. In order to measure small magnetic fields, both state transfer and operator pulses are designed. It has to be said from the very beginning that the robustness we achieve by SOC can also - at least in principle - be attained by using GRAPE. However, SOC has several advantages mentioned at the end of the previous chapter, above all the inherent smoothness of the pulses. Furthermore, GRAPE has its main applications in NMR, only recently it has been applied to NV centers [57, 58].

3.1. Designing Ensemble Pulses

If one wants to design a pulse for a specific experiment, one has to know the following data given by the experiment: The form and size of the inhomogeneous broadening of the resonance frequencies ω_0 (see Eq. (1.75)) depend on the NV sample. Unless it is not specified otherwise, we will assume an inhomogeneous broadening with a Gaussian form and 10 MHz full width at half

maximum (FWHM). The range of the control field experienced by the spins is determined by the geometry of the gold wire that generates the control field (see Sec. 1.7) and by the volume V the microscope focuses on. If not specified otherwise, the control field ranges from weak fields of 75% in the middle of the probe volume to strong fields of 125% in the vicinity of the gold wire. As an ensemble, we take one with 100 spins and test its representativeness at the end by applying the resulting pulse on a different ensemble. By doing this, we ensure that the average fidelity is independent on the test ensemble that one uses in order to calculate the average. As another important quantity characterizing the pulse, the maximal spectral width of the pulse is given by the number of frequency components present in the optimal control pulse. In addition to the above quantities one also needs to know the maximal pulse length. This pulse length is limited by the single spin decoherence time T_2, which is determined by the interaction with surrounding carbon and nitrogen nuclei [59]. Another quantity that one needs in order to design an optimal control pulse is the maximal Rabi frequency, *i.e.* the maximal value A_{thr} of the control amplitudes $f_x(t)$ and $f_y(t)$ in Eq. (1.76). To limit this value, one introduces a penalty functional $\mathcal{F}_p = -p\frac{t_f}{2}\mathbf{a}^2$, where $\mathbf{a}$ are the control amplitudes, t_f is the duration of the pulse and $p \geq 0$ is a parameter which specifies how strongly high amplitudes are penalized. In practice, one starts the algorithm with a strong penalty of $p = 0.1$ and calculates the maximal Rabi frequency A_{max} in each iteration step. If $A_{\text{max}} < A_{\text{thr}}$, p is decreased by a small amount Δp (here we used $\Delta p = 1/600$), otherwise it is increased by Δp. The penalty for high amplitudes not only limits the maximal power of the pulse but also prevents the algorithm from being trapped in local maxima.

3.1.1. State Transfer Pulses

In the following, we will deal with the state transfer problem $|\psi_i\rangle \rightarrow |\psi_f\rangle$ described in Sec. 1.2.2.4 and the corresponding target functional Eq. (1.85). In this context, we recall that a π-pulse is a pulse which causes a transfer from the state $|0\rangle$ to the state $|1\rangle$, a $\pi/2$-pulse a pulse which realizes the transfer $|0\rangle \rightarrow (|0\rangle + |1\rangle)/\sqrt{2}$. In the following example, we optimized a pulse for 8 MHz Gaussian FWHM and a control amplitude as stated above. The experiment is done with an NV ensemble that contains at least 99.999% of

^{12}C. The decoherence time is $T_2 = 140\,\mu$s, the dephasing time connected to the inhomogeneous broadening is $T_2^* = 10\,\mu$s.

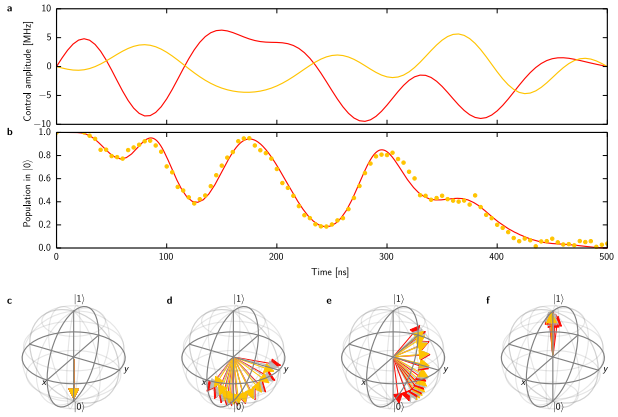

Figure 3.1.: SOC pulse (a: red $f_x(t)$, orange $f_y(t)$) with $n = 10$ frequency components, which leads to the ensemble trajectory b. The theoretical prediction is shown by the solid line, the dots correspond to the measurements that have been done. In c-f are depicted the spins on the Bloch sphere at the moments $t = 0$, 150, 300 and 500 ns in time. The red spins only experience 90% of the control amplitude, the grey ones 100% and the orange ones 110%. Within these three groups the spins show eleven detunings equally spaced in the interval $[-7, +7]$ MHz. [40]

Our first objective was to verify that the pulses that we design induce the desired dynamics in the experiment. For this purpose we constructed a pulse with ten frequency components that induces a propagator $\mathcal{U}_{\alpha,\omega_0}(t_f)$, where t_f is the duration of the pulse, such that $|\langle 1|\mathcal{U}_{\alpha,\omega_0}(t_f)|0\rangle|^2 \approx 1$ (π-pulse) for any spin resonance frequency ω_0 and relative control amplitude α in the respective interval (see Sec. 1.2.2.2). The time-dependent, ensemble-averaged spin population $P(|0\rangle) = \langle\langle 0|\mathcal{U}_{\alpha,\omega_0}|0\rangle\rangle_{\alpha,\omega_0}$ can be measured experimentally and comparison between the theoretical prediction and experimental data permits us to verify the agreement between theory and experiment. Fig. 3.1 depicts the averaged trajectory $P(|0\rangle)$, *i.e.* the population in the state $|0\rangle$, for both the experimental observation (dots) and the theoretical prediction (continuous line). One sees that the trajectory reaches zero at the end of the pulse (here after 500 ns), meaning that the state $|1\rangle$ is reached. Therefore, we really deal with a π-pulse. Furthermore, the theoretical prediction is reproduced with

high accuracy in the experiment, meaning that our theoretical description is right. In addition, we plotted the SOC pulses that led to this trajectory. As one can see, they are very smooth and do not contain high frequency components by definition - as one would expect for a pulse with ten frequency components. We limited the maximal Rabi frequency with the above algorithm to $A_{\max} = 10$ MHz. This goal is achieved, as one can see in Fig. 3.1 a), where the amplitudes vary between -10 and 7 MHz. In order to illustrate the theoretical dynamics of the single spins, four snapshots of a spin ensemble on the Bloch sphere are taken at different moments in time. As one can see, the dynamics of the spins is rather complicated but in the end all spins are refocused in the state $|1\rangle$. As mentioned in Sec. 1.2.2.2, the 'detour' made by some of the spins is crucial for the robustness of the pulse. Moreover, it is important to note that the pulse is the more robust the more of the Bloch sphere the spins explore.

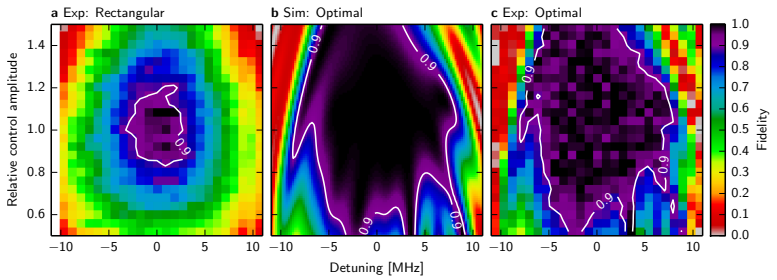

Figure 3.2.: Fidelity maps are depicted for a rectangular pulse (a) with the same maximal amplitude as the smooth pulse, (b) shows the theoretical prediction of a SOC pulse, (c) the experimental result. The white lines indicate a fidelity of 90%. [40]

In order to assess the quality of a pulse, the best way is to generate a fidelity map, where the fidelity, encoded in a color scheme, is shown as a function of the detuning ω_0 and the relative control amplitude α, as introduced in Sec. 1.2.2.2. The fidelity is probed with a single NV, which is detuned or whose relative control amplitude is varied. The result is shown in Fig. 3.2: One can see that both theoretical and experimental plot show a region with a fidelity of more than 90% ranging from -8 to +8 MHz detuning and a relative control amplitude from 0.5 to 1.5. The experimental plot resembles the theoretical plot taken at a lower resolution. This lower resolution is due to the unavoidable presence of shot noise in the experiment. In order to compare SOC with the state of the

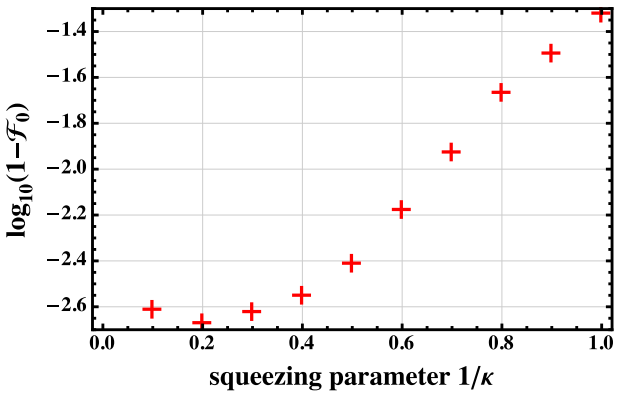

Figure 3.3.: Logarithmic infidelity of a π-pulse with $n = 10$ frequency components as a function of the inverse squeezing, measured by the parameter κ. The initial pulse length was $t_f = 500$ ns (corresponding to $\kappa = 1$).

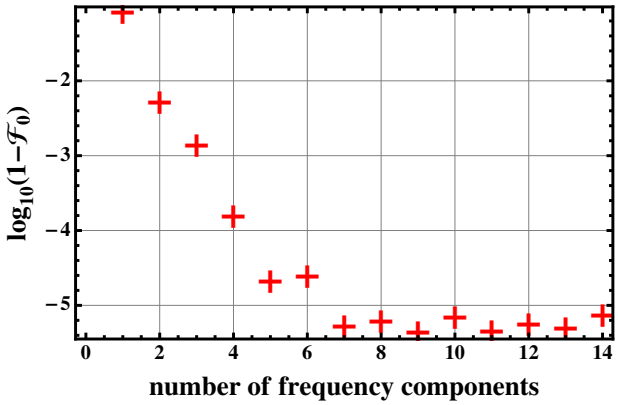

Figure 3.4.: Logarithmic infidelity of a $\pi/2$-pulse as a function of the number of frequency components n. The inhomogeneous broadening had a rectangular shape with width $\Delta\omega = 0.5$ MHz, the control amplitude had a range from 80 to 120%, the duration of the pulse was $t_f = 1\,\mu$s.

art techniques one has to draw a comparison with rectangular pulses. Here, different choices are possible: One could compare with the shortest possible rectangular pulse. However, this choice is not appropriate because given a highest implementable frequency, retangular pulses can always be made shorter than any other pulse, resulting in an artificially higher fidelity. The second option would be to compare with the rectangular pulse of the same duration, this however would disadvantage the rectangular pulse, which could be realized

with much higher intensity. The third and fairest solution is to compare pulses with the same maximal Rabi frequency because this is the natural limitation of the pulses.

For a rectangular pulse, one can identify a region from -4 to 4 MHz detuning and 0.8 to 1.2 relative control amplitude, where the fidelity is higher than 90%. Beyond this region, fidelity rapidly decreases. Concerning the SOC pulse, the fidelity is higher than 90% in a much larger region, more precisely from -8 to 8 MHz detuning - the double of the rectangular pulse - and for a relative control amplitude ranging (nearly) from 0.5 to 1.5 - being the double of the rectangular pulse, too. We have therefore shown that SOC is better than conventional control techniques working with rectangular pulses. One also has to say that the robustness against different control amplitudes - from $\alpha = 0.5$ to 1.5 - is finally much better than the targeted control amplitudes (from $\alpha = 0.75$ to 1.25).

It is impressive how far one can push the robustness against different control amplitudes: We were able to generate a π-pulse with $n = 10$ frequency components that is robust against an inhomogeneous broadening of 1 MHz FWHM and control amplitudes ranging from $\alpha = 0.04$ to 1.96, *i.e.* the nearest spin to the antenna experiences a control field that is about 50 times stronger than this of the farest spin. By limiting the maximal Rabi frequency to $A_{\max} = 20$ MHz, this pulse yields an average fidelity of more than 95%. Therefore, SOC is ideally suited to control an ensemble of NV centers in nanodiamonds. As the orientations of these nanodiamonds are randomly distributed over the sample, they experience a broad range of control amplitudes [60].

As we have seen in Sec. 1.2.2.4, a given pulse can be improved by squeezing it, *i.e.* by decreasing the pulse duration by a factor κ and enlarging the control amplitude by the same factor κ. In Fig. 3.3, the logarithmic infidelity is shown as a function of the inverse squeezing parameter $1/\kappa$. One can see that the fidelity gets exponentially better with increasing squeezing until it saturates for $1/\kappa < 0.4$ to $\log_{10}(1 - \mathcal{F}_0) = -2.6$. This can be understood in the following way: Squeezing means nothing else but magnifying the central region in the above fidelity map - region over which one has to average. As the fidelity is higher in the central region, the fidelity of a squeezed pulse must increase. However, in the limit of $\kappa \to \infty$, the fidelity cannot exceed the fidelity $\mathcal{F}_{\Delta\omega=0,\,\alpha=1}$ of no detuning and the exact relative control amplitude.

It is important to note that one squeezes a pulse which is *already* optimized. In particular, there is *no* additional optimization for a given squeezing; the improvement in fidelity only comes from the squeezing.

In Sec. 2.1 we have seen that the computational cost for computing the derivatives with respect to the control parameters increases fast with the number of frequency components. This means that many frequency components should be avoided. Here, the logarithmic infidelity of a $\pi/2$-pulse is shown in Fig. 3.4 as a function of the number of frequency components. In order to compensate for local maxima, every data point is the result of the best of five iterations. As more frequency components mean more freedom in designing a pulse, one may expect the fidelity to increase with increasing number of frequency components. Indeed, one can see in Fig. 3.4 that the fidelity becomes exponentially better until it saturates to $1-10^{-5}$ for $n \geq 7$ frequency components. This means that one does not have to go to a large number of frequency components. In order to be sure that the fidelity does not increase substantially with the number of frequency components, one takes more than 7 frequency components, typically about $n \approx 10$ components are enough. In this way, the computational cost of many frequency components can be avoided. In addition, smooth pulses typically look very simple and almost intuitive. It is expected that the fidelity with respect to the number of frequency components behaves in a monotonic way, as a pulse with more frequency components can always be generated from less components by adding sufficiently many zeros. The fact that the fidelity for $n = 6$ frequency components is less than for $n = 5$ has to do with the existence of local maxima in the control landscape [61]. These local maxima are also present in the region where the fidelity saturates.

Another important aspect is the behavior of the fidelity as a function of the duration t_f of the pulse. It is found that the fidelity does not depend significantly on the pulse duration. A reason for this is the interplay of two effects: on the one hand, a longer pulse provides more possibilities to control the system, *i.e.* results in a higher fidelity; on the other hand, the detrimental effect of inhomogeneous broadening is the stronger, the longer the pulse is. In the end, these two tendencies cancel each other out.

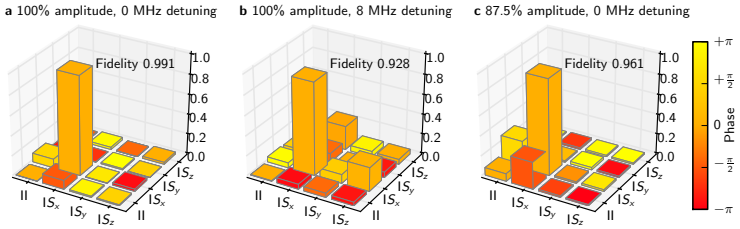

Figure 3.5.: Quantum process tomography for a π_x-pulse: The modulus and the phase of the process matrix χ are displayed for the correct amplitude (a), for the correct amplitude and 8 MHz detuning (b) and for 87.5% of the optimal amplitude and no detuning (c). As one can see, the loss in fidelity is little. [40]

3.1.2. Operator Pulses

In the following, we will deal with operator pulses, *e.g.* pulses that implement a rotation $\mathcal{U}_f = e^{-i\pi\sigma_x/2}$ by π around the x-axis, abbreviated as π_x. The corresponding target functional is given by the operator fidelity Eq. 1.89. In theory, one could also generate a fidelity map for operator pulses. In the experiment, however, this operator fidelity is not directly measurable but one needs to reconstruct the quantum channel induced by the pulse by means of quantum process tomography described in Sec. 1.1.4. With the reconstructed channel one can then compare experiment and theory.

The operations which are necessary for quantum state tomography are done with the shortest possible rectangular pulses, which have a high fidelity. If one had used SOC pulses, the result of the quantum process tomography would depend heavily on the operations with which one realizes the input states of the quantum state tomography.

When dealing with operator pulses, in the case of NV centers, there is an additional difficulty to overcome: Since the signal for measuring populations (see Sec. 1.2.1) only depends on the magnetic quantum number m_S, it is only possible to measure the z-component of the spins. However, if one likes to measure the x- or y-component after the application of a certain pulse - as it is necessary for quantum state tomography -, one can do this by the following measurement scheme: One rotates the spins first by a $(\pi/2)_x$- or $(\pi/2)_y$-pulse ($\mathcal{U}_f = e^{-i\pi\sigma_k/4}$, $k = x, y$), then performs the pulse, next rotates the state back by a $(-\pi/2)_x$- or $(-\pi/2)_y$-pulse ($\mathcal{U}_f = e^{i\pi\sigma_k/4}$, $k = x, y$) and then measures the z-component in order to determine the x- or y-component.

In the experiment, the process matrix for a π_x-pulse for a single spin is to be found. A π_x-pulse can be described by the unitary transformation $\mathcal{U} = e^{-i\pi\sigma_x/2} = \sigma_x$. From this, it follows that the density matrix ρ' after the pulse can be written in the form $\rho' = \mathcal{U}\rho\mathcal{U}^\dagger = \sigma_x\rho\sigma_x$. This means that the process matrix χ in Eq. (1.34) can be expressed in terms of the basis $\{\mathbb{1}, \sigma_x, -i\sigma_y, \sigma_z\}$ by the matrix

$$\chi = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{3.1}$$

Compared to this ideal matrix, Figure 3.5 shows the experimental process matrix χ in the case of the ideal amplitude (a), the ideal amplitude and 8 MHz detuning (b) and of an amplitude which is 12.5% smaller than the ideal one (c). The results are represented in a bar chart, where the height of the bars represents the modulus and the color the phase of the matrix elements χ_{mn}. The matrix with the ideal amplitude and no detuning agrees with high accuracy with the theoretical prediction. The dominant entry of the process matrix is χ_{22} and has a modulus of 0.991 and a phase of zero. The next largest entries are χ_{12} and χ_{21} and have a modulus of less than 0.1. The phases of the other entries are distributed over the whole range of $[0, 2\pi)$, which is not surprising, as a fluctuation on the top of a complex number near zero can easily alter its phase. The process matrix for the ideal amplitude and a detuning of 8 MHz is still very close to the theoretical prediction, as the main entry χ_{22} has a modulus of 0.928 and zero phase. The next largest entries are χ_{24} and χ_{42} with a modulus of 0.2. Also the process matrix for 87.5% of the original amplitude and no detuning is in agreement with the theory: The leading entry χ_{22} has a modulus of 0.961 and zero phase. The next largest entries are χ_{12} and χ_{21} with a modulus of 0.25, the other entries being close to zero.

As the above analysis shows, the constructed pulse induces a π_x-gate not only for one spin with resonance frequency $\omega_0 = 0$ and control amplitude $\alpha = 1$ but also for spins which are detuned from the resonance frequency or do not experience the full control amplitude. Therefore, the operator pulse π_x will also work with *ensembles* of NV centers, and this is what is needed in magnetometry.

In Sec. 1.1.4, we considered the possibility that the process matrix may be not positive semidefinite. In the three examples of quantum process tomog-

raphy, the distance (calculated with the Hilbert-Schmidt norm) between the experimental process matrix χ and the ideal one $\tilde{\chi}$ is only 2 to 3% of $\|\tilde{\chi}\|_{HS}$ [40]. In Ref. [39], the authors present a quantum process tomography with NV centers. The deviation of the experimental process matrix from the ideal one, calculated therein, lies slightly higher than ours. This entails that our experimentally found χ represents to a very good approximation a physical process.

3.2. Magnetometry

Now that we have the necessary pulses $\pi/2$, π_x and $(\pi/2)_y$ at hand, we are able to do magnetometry, as it is described in Sec. 1.2.4. The experiment described here was performed by T. Nöbauer and A. Angerer with a sample that had a decoherence time of $T_2 \approx T_2^* \approx 2.2\,\mu$s. The time of free precession was $\tau = 1.2\,\mu$s, in accordance with Eq. (1.101). In contrast to the measurements quoted in Figs. 3.2 and 3.5, an ensemble of NV centers is now used in the experiment. Scanning the detuning and the relative control amplitude over a broad range for both rectangular and SOC pulses with the same maximal amplitude permits us to experimentally verify the robustness of our method. The resulting sensitivity encoded in a color scale is shown in Fig. 3.6 a). As one can see, the SOC pulses lead to a uniform high sensitivity of at least 10^{-6} T Hz$^{-1/2}$ over the detuning range of 0 to 14 MHz and a relative control amplitude ranging from 0.65 to 1. On the contrary, the rectangular pulse has a comparable sensitivity for a detuning of less than 5 MHz and then rapidly looses one to two orders of magnitude in sensitivity as detunings increase. This behavior is illustrated by Figs. 3.6 b) and c), where one has averaged along constant control amplitude (1.0) and along constant detuning (0), respectively. Whereas the SOC pulses are only slightly better than the rectangular ones for varying relative control amplitude, SOC shows its whole power for strongly broadened ensembles. Whereas the averaged sensitivities obtained with SOC and rectangular pulses nearly coincide for detunings smaller than 4 MHz, SOC almost gains two orders of magnitude in sensitivity for a detuning of 14 MHz. Here, one can see a well-known trend which one has already seen in the case of the fidelity maps of state transfer pulses (see Sec. 3.1.1): the robustness against different detunings is very limited for rectangular pulses. In summary,

we have shown that the sensitivity in the case of SOC magnetometry is not only better than methods using state-of-the-art rectangular pulses, but is also by far more robust against different detunings and relative control amplitudes. As in the case of other OC algorithms, this robustness is an inherent property of our OC framework.

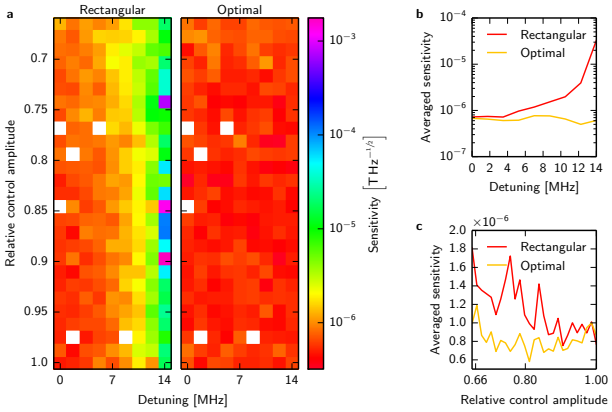

Figure 3.6.: Sensitivity map (a) for SOC and rectangular pulse (white points stand for data points that were discarded because of experimental instabilities). In addition we integrated along a relative control amplitude 1.0 (b) and no detuning (c). The SOC pulse sequence performs by far better than the rectangular one. [40]

4. Control of Composite Quantum Systems

In this chapter, the use of Smooth Optimal Control will be extended to control the dynamics of more than one body. The goal of this chapter is to show the versatility of SOC when applied to composite quantum systems. The first three sections are also discused in [15]. More precisely, it is demonstrated how time-optimal pulses can be designed in order to entangle two interacting spins. Next, it is shown how entanglement dynamics can be influenced such that the time interval for which entanglement - both bipartite and multipartite - is high can be enhanced. After that, the focus will be on spin chains and mediated interactions. Before describing in detail the interaction of a cavity and a spin ensemble, the decoupling of a spin from its environment will be studied.

4.1. Time-Optimal Gates

A crucial task in quantum information processing is to implement a unitary gate $\mathcal{U}_d \in U(4)$ on two qubits. Unfortunately, every system is subject to decoherence as it is coupled to an environment. This will limit the precision with which one can implement a quantum operation. The simplest way to avoid decoherence is to make the duration of the pulse that implements the desired gate as short as possible. If the duration of a gate is substantially shorter than the decoherence time, the destructive effects of decoherence can be circumvented. In this context, the notion of time optimality has to be introduced: A pulse is said to be *time-optimal*, if a desired fidelity cannot be reached in a shorter time. The common way [62] to obtain time-optimal pulses is to start the optimization algorithm with an initial time t_0 and then decrease the pulse duration, until a certain fidelity threshold cannot be reached

anymore. Depending on the desired precision, this procedure can be very time-consuming.

On the other hand, time-optimality can be implemented in the current OC framework by introducing an appropriate penalty functional $\mathcal{F}_p$. In this framework, the duration of a gate becomes a control parameter in itself. The simplest choice for the penalty functional is $\mathcal{F}_p = -pt_f$, where t_f is the duration of the pulse and $p \geq 0$ is a parameter specifying how strongly long times are penalized. In this way, long pulse durations make the penalty functional highly negative, enforcing the algorithm to favor smaller pulse durations, corresponding to a higher (the penalty functional is negative) value of the penalty functional. The reason why time-optimality can be obtained in such an easy way within the SOC framework will become clear in the following: Since the pulse should be switched on at $t = 0$ and switched off at $t = t_f$, the zeros of the fundamental frequency component Ω have to coincide with the beginning and the end of the control, $g_{1k}(0) = g_{1k}(t_f) = 0$ in Eq. (1.4) for all $k \in \mathbb{N}$. This constraint implies that Ωt_f is a constant, while Ω is varied, namely, when using trigonometric functions, it is $\Omega t_f = \pi$ because $\sin(0) = \sin(k\Omega t_f) = \sin(k\pi) = 0$ with k being an integer. One can now replace t_f in the time evolution operator $\mathcal{U}(t_f)$ by $t_f = \pi/\Omega$. Since the Floquet operator $\tilde{\mathcal{K}}$ (Eq. (2.2)) is linear in Ω in the same way as in the control amplitudes $\mathbf{a}$, one can use the same perturbation ansatz in order to calculate the derivatives with respect to Ω. In this way, the fundamental frequency Ω becomes a control parameter itself.

For the problem of implementing a gate $\mathcal{U}_d$, the target functional reads $\mathcal{F} = \mathcal{F}_0 + \mathcal{F}_p$, with the gate fidelity $\mathcal{F}_0 = \mathrm{Re}\left[\mathrm{Tr}\left[\mathcal{U}(t_f)^\dagger \mathcal{U}_d\right]\right]/4$ and the penalty functional $\mathcal{F}_p$ from above. Having chosen a value for p, the algorithm will end in a state where $\mathcal{F}_0$ and t_f will have certain values. In practice, however, one is more interested in finding the shortest possible time in which a gate can be constructed with a certain initially specified fidelity $\mathcal{F}_{\mathrm{thr}}$. One problem that one faces when running our algorithm is the trapping into local minima of slow gates (see below in detail). To circumvent this difficulty, we first optimized the SOC pulse for a fixed pulse duration $t = t_0$ ($p = 0$), which is smaller than the characteristic time scale t_{char}, on which entanglement can be created by the system itself. After the optimization for $t = t_0$, the time-optimality algorithm is started with a high penalty of $p = 1$. However, a too high parameter p would result in a fast gate but with very bad fidelity. Because of this, p is

varied during the SOC algorithm, namely p is decreased by an amount of $\Delta p = 0.01$ in each iteration step. As soon as the target fidelity $\mathcal{F}_{\text{thr}}$ is reached, p is again increased by Δp in order to get a faster gate with the same fidelity.

In the present work, the implementation of entangling gates is studied for two interacting spins whose system Hamiltonian has the following form:

$$\mathcal{H}_0 = \frac{\omega_1}{2}\sigma_z^{(1)} + \frac{\omega_2}{2}\sigma_z^{(2)} + g_x\sigma_x^{(1)}\sigma_x^{(2)} + g_y\sigma_y^{(1)}\sigma_y^{(2)} \tag{4.1}$$

Here, $\sigma_k^{(j)}$ is the k-th Pauli matrix of the j-th spin. The coupling constants g_x and g_y and the (possibly small) splittings ω_1 and ω_2 are chosen at random and can be found in the caption of Fig. 4.1. The two spins can be manipulated separately by the control Hamiltonian

$$\mathcal{H}_c(t) = \sum_{\substack{k=x,y \\ j=1,2}} f_k^{(j)}(t)\sigma_k^{(j)}, \tag{4.2}$$

where the control pulse

$$f_k^{(j)}(t) = \sum_{n=1}^{n_{\max}} a_n^{(k,j)} \sin(n\Omega t) \tag{4.3}$$

has $n_{\max}$ frequency components of a fundamental frequency Ω.

As quoted in Eq. (1.56), a unitary gate $\mathcal{U}_d$ can be characterized up to local transformations by only three parameters α_x, α_y and α_z. A common task in quantum information processing is to drive the system from an initial product state to a maximally entangled state. The unitary gates $\mathcal{U}_{(\alpha_x,\alpha_y,\alpha_z)}$ that can do this obey the inequalities Eqs. (1.57) and (1.58).

In Fig. 4.1, the maximally entangling gate $U_{(0.5,0.4,0.3)}$ with a fidelity threshold of $\mathcal{F}_{\text{thr}} = 1 - 10^{-4}$ has been implemented. The starting point of our algorithm is determined by the characteristic time scale t_{char} on which entanglement can be created. This time scale can be estimated by $t_{\text{char}} = \pi/(4g_{\max})$, where $g_{\max}$ is the largest coupling constant in Eq. (4.1). This estimate is derived from the fact that $\exp(-i\pi/4\sigma_x \otimes \sigma_x)$ provides a maximally entangling gate. In Fig. 4.1, one has $t_{\text{char}} \approx 0.08\,\mu$s. The starting point for the search of the time-optimal solution was a pulse optimized for $t_0 = 0.06\,\mu$s, while using $n_{\max} = 6$ Fourier components. In Fig. 4.1 a), one can see the evolution of the gate fidelity $\mathcal{F}_0$ and of the pulse duration t_f as a function of the iteration steps. The fidelity is at first increased by using longer pulse durations. As

soon as the algorithm passes the fidelity threshold $\mathcal{F}_{\text{thr}}$ (here, after about 350 iteration steps), the weighting parameter p is increased and the pulse duration decreased. The pulse duration then saturates to the final value of $t_f = 0.077\,\mu$s. Fig. 4.1 b) is the result of the above optimization; it shows the time evolution of the logarithmic infidelity. As one can see, the intrinsic dynamics alone is not able to entangle the spins with high fidelity, whereas control can implement a maximally entangling gate with fidelity $\mathcal{F}_0 = 1 - 10^{-4}$ by using only $n_{\max} = 6$ Fourier components. More frequency components improve the result only modestly.

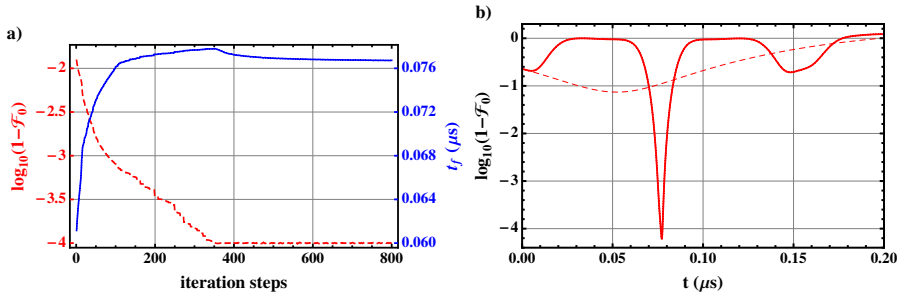

Figure 4.1.: a) Typical run of the optimization routine to construct the entangling gate $\mathcal{U}_{(0.5,0.4,0.3)}$ for $g_x = 5.40$ MHz, $g_y = 9.95$ MHz, $\omega_1 = 0.13$ MHz, $\omega_2 = 0.26$ MHz in minimal time. The time t_f (solid line) is minimized, such that a gate fidelity (dashed line) of more than 99.99% is reached. b) Result of the above optimization: Logarithmic infidelity as a function of time with (solid line) and without control (dashed line). Without control the gate cannot be realized with high fidelity. [15]

To be sure that the solution for $t_f = 0.077\,\mu$s is really time-optimal, the SOC algorithm has been run for fixed times $t < t_{\text{char}}$ ($p = 0$). In this case, one should see that the fidelity threshold $\mathcal{F}_{\text{thr}} = 1 - 10^{-4}$ is reached in the time $t_f = 0.077\,\mu$s. When optimizing for fixed times, one particularity (see Fig. 4.2) is observed: there are two branches one can follow; the first (green stars) is obtained without any restriction, for the second branch (red crosses) the algorithm was started with the optimal solution for $t_f = 0.08\,\mu$s and then one optimizes $\mathcal{F}_0$ for successively smaller times. It turns out that the green branch represents a set of local maxima. Indeed, if one lets run the time-optimality algorithm as it is described above, one always lands on the curve with the red crosses, strongly suggesting that one has found the time-optimal solution.

If, on the other hand, one starts the optimization from an arbitrary point, without pre-optimization, the solution always lies on the curve with the green stars. Furthermore, it is easy to achieve a higher fidelity, as the curve $\mathcal{F}_0(t_f)$ drops down fast when approaching t_{char}, *e.g.* going from $\mathcal{F}_{\text{thr}} = 1 - 10^{-4}$ to $\mathcal{F}_{\text{thr}} = 1 - 10^{-6}$ only requires 2% more time.

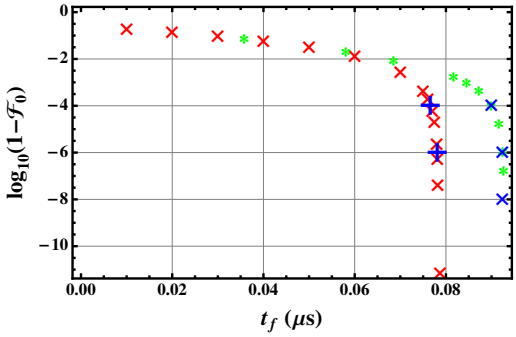

Figure 4.2.: Logarithmic gate infidelity $\log_{10}(1 - \mathcal{F}_0)$ as a function of the pulse duration t_f for $n_{\max} = 6$ frequency components, with pre-optimization of $t = 0.08\,\mu$s (red '$\times$') and without initial optimization (green '$*$'). The symbol '$+$' depicts the time-optimal pulses for a required gate fidelity of $\mathcal{F}_0 = 1 - 10^{-4}$ and $\mathcal{F}_0 = 1 - 10^{-6}$, respectively, and with pre-optimization; they always lie on the red curve. The blue '$\times$' correspond to runs of the time-optimality algorithm without pre-optimization and lie on the green curve of local maxima.

4.2. Creating and Maintaining Entanglement

Entanglement is the essential part when constructing algorithms that are more powerful than their classical analogues [38]. In the following, we will discuss control pulses that help to create entanglement and maintain it over a finite time window. The maintenance of entanglement is important for experimental situations where time can only be triggered up to a certain accuracy Δt or where entanglement is needed for a longer time in order to perform other operations onto the system. In the following, one will distinguish between bipartite and multipartite entanglement.

4.2.1. Bipartite Entanglement

The easiest way to entangle two particles is to transfer them from an initially unentangled state $|\Psi_i\rangle = |\psi_1\rangle|\psi_2\rangle$ to a maximally entangled state $|\Psi_f\rangle = |\Psi_E\rangle$ by using the fidelity $\mathcal{F}_{\text{fid}} = |\langle\Psi_f|\mathcal{U}(t_f)|\Psi_i\rangle|^2$ as the target functional. This is what is displayed in Fig. 4.3. We assumed the same system Hamiltonian Eq. (4.1) with controls Eqs. (4.2) and (4.3) and the separable state

$$|\Psi(0)\rangle = \bigotimes_{k=1}^{N} \left(\cos(\theta_k/2)\,|0\rangle + e^{i\phi_k}\sin(\theta_k/2)\,|1\rangle\right) \tag{4.4}$$

with $\theta_1 = 1.59$, $\theta_2 = 2.10$, $\phi_1 = 5.23$, and $\phi_2 = 0.57$ as initial condition. The maximally entangled target state was

$$|\Psi_f\rangle = \alpha_1|00\rangle + \alpha_2|01\rangle + \alpha_3|10\rangle + \alpha_4|11\rangle \tag{4.5}$$

with $\alpha_1 = -0.46 + 0.32i$, $\alpha_2 = -0.02 - 0.43i$, $\alpha_3 = -0.33 - 0.26$, $\alpha_4 = -0.53 + 0.18i$. As one can see, the fidelity quickly drops down after having reached its maximal value. However, in experimental situations [63], where one desires robustness against small variations Δt in time, one is rather interested in a broader time interval of high entanglement. Unfortunately, if one optimizes the entanglement for two different initial control amplitudes $\mathbf{a}_0$, one obtains two different control pulses but which follow nearly the same evolution of entanglement, as can be seen in Fig. 4.3. This means that it is very difficult to design the time evolution in a desired fashion. It seems that demanding a precise final state restricts the dynamics in Hilbert space to a very tiny region, which is incompatible with other dynamical constraints on the evolution of entanglement.

In order to enlarge the region in Hilbert space to which the dynamics are constrained, one can use an entanglement measure $\mathcal{E}$ as the target functional. As entanglement is preserved by local unitary transformations, this provides more freedom to design the time evolution of entanglement. In this work, the tangle $C^2 = |\langle\Psi(t_f)|\sigma_y \otimes \sigma_y|\Psi^*(t_f)\rangle|^2$ [42] is used as an entanglement measure. Similarly to Sec. 4.1, the target functional reads $\mathcal{F} = C^2 + \mathcal{F}_p$ with a penalty functional $\mathcal{F}_p = -p\left(\partial^2 C^2/\partial t^2|_{t=t_f}\right)^2$, which penalizes high curvatures with respect to time at $t = t_f$, therefore enforces a plateau in the time evolution of entanglement. The derivatives with respect to time can be obtained

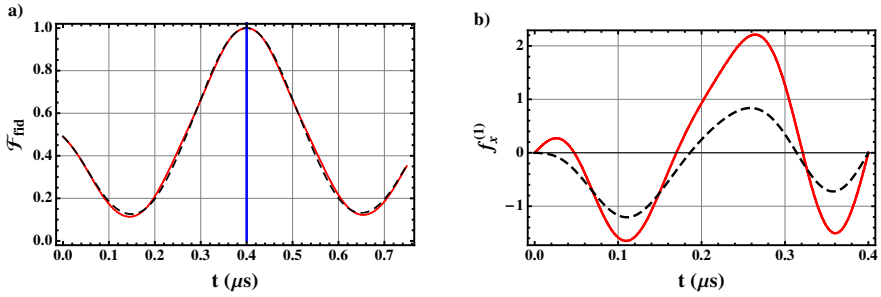

Figure 4.3.: Driving two coupled spins with $g_x = 2.7\,\mathrm{MHz}$, $g_y = 6.2\,\mathrm{MHz}$, $\omega_1 = 0.3\,\mathrm{MHz}$, and $\omega_2 = 0.2\,\mathrm{MHz}$ to a maximally entangled state. Although the control pulses are rather different (see b), as an example, only the component $f_x^{(1)}(t)$ is depicted), the time evolutions of the fidelity $\mathcal{F}_{\rm fid}$, represented in a) by the continuous and dashed lines, are very similar.

analytically without resorting to a method of finite differences as

$$\frac{\partial^n \mathcal{U}}{\partial t^n} = i^n \sum_{k,\nu} (\nu\Omega - \epsilon_k)^n |\chi_{k\nu}\rangle\langle\Phi_k(0)| e^{i(\nu\Omega-\epsilon_k)t}. \tag{4.6}$$

In practice, a fixed small value of p is sufficient in order to decrease the curvature by several orders of magnitude (practically to zero, below $p = 10^{-4}$ is used) while maintaining maximal the tangle at $t = t_f$. In Fig. 4.4, the tangle C^2 with and without minimization of the curvature and without any optimal control is displayed as a function of time for two durations $t_f = 0.2$ and 0.4 μs of control. Not surprisingly, without control no substantial entanglement can be created. With control, maximal entanglement can be easily generated by using only six frequency components. By penalizing high curvatures, the curvature of the tangle is decreased from $10^2(\mu\mathrm{s})^{-2}$ (no penalty, $p = 0$) to $10^{-7}(\mu\mathrm{s})^{-2}$ (penalty $p = 10^{-4}$). In the case of $t_f = 0.4\,\mu$s, without minimization the control keeps entanglement above 99.9% for 4.6 ns, while the minimization of the curvature results in 98 ns of high entanglement, *i.e.* more than twenty times longer. This longer time requires only a moderate increase in the maximal control amplitude from 3.0 to 5.4 MHz.

As one can see in Fig. 4.4, the time interval of high entanglement is the longer, the longer the control time is. This is not surprising, since a longer duration of the pulse corresponds to more possibilities to control the system. However, a limitation of the time interval T of high entanglement is given by

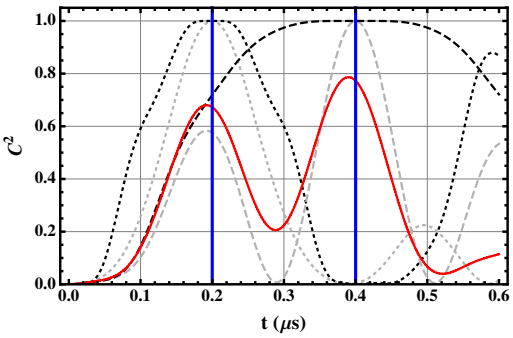

Figure 4.4.: Time evolution of the tangle C^2 for two interacting spins (same parameters as in Fig. 4.3). Without control (solid line) maximal entanglement can never be reached. The grey curves are obtained by maximizing the tangle at $t_f = 0.2\,\mu$s (dotted curve) and $t_f = 0.4\,\mu$s (dashed curve), respectively (these moments in time are emphasized by fat vertical lines), by using a control pulse with $n_{\max} = 6$ frequency components and fundamental frequency $\Omega = \pi/t_f$. The curves obtained with an additional minimization of the curvature are depicted in black. The time interval for which the spins are highly entangled is considerably longer in these cases. [15]

the minimal time $t_{\min}$ which is needed in order to provide maximal entanglement. An upper limit of T is therefore given by $T = 2(t_f - t_{\min})$, since the plateau cannot be extended below $t_{\min}$, otherwise one could have obtained entanglement in a time smaller than $t_{\min}$. The region for which the evolution of entanglement can be approximated by the second derivative only is already quite impressive. Nevertheless, the plateau cannot be increased indefinitely because at some point the validity of the second derivative reaches its limit. This case is given for $t_f \gtrsim 0.6\,\mu$s. If one wishes to extend the plateau even further, one has to maximize the tangle and minimize its curvature at several moments $t_1, \ldots, t_N$ in time. This corresponds to using the new fidelity $\tilde{\mathcal{F}} = \frac{1}{N}\sum_{n=1}^{N} \mathcal{F}(t_n)$, *i.e.* an average over several intervals for each of which the approximation by the second derivative is valid. In the case of Fig. 4.5, one could enlarge the width of the plateau of high entanglement ($\mathcal{F} > 99.9\%$) for $t_f = 0.6\,\mu$s by almost a factor of three by only using a second node at $t_2 = 0.5\,\mu$s.

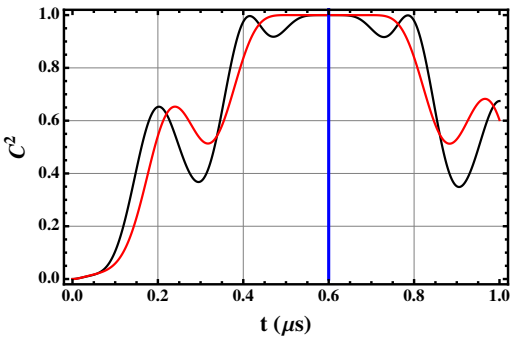

Figure 4.5.: It is shown how one can create a plateau of high entanglement at $t = 0.6\,\mu$s (black curve) and by using the combined fidelity $\frac{1}{2}\left(\mathcal{F}(t_1) + \mathcal{F}(t_2)\right)$ with $t_1 = 0.5\,\mu$s and $t_2 = 0.6\,\mu$s (red curve). In the latter case, the time interval of high entanglement is considerably increased.

4.2.2. Multipartite Entanglement

If it is possible to prolong a time interval of high bipartite entanglement, the question if this is still manageable in the case of multipartite entanglement (see Sec. 1.1.5.3) is only natural. As a target functional, the lower bound Eq. (1.67) of the tangle is used.

In the following, it is assumed that the total density matrix represents a pure state. In this case the above lower bound for mixed states becomes an entanglement measure of pure states in its own. For $N = 3$, one then obtains

$$\mathcal{F}_{N=3} = \frac{1}{2}\left(6 - \mathrm{Tr}(\rho_{23}^2 + \rho_{13}^2 + \rho_{12}^2) - \mathrm{Tr}(\rho_1^2 + \rho_2^2 + \rho_3^2)\right), \tag{4.7}$$

The maximal value of $\mathcal{F}_{N=3} = 3/2$ is obtained by the Greenberger-Horn-Zeilinger state $|\mathrm{GHZ}_3\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$, since it results in maximally mixed single-spin reduced density matrices. The W state $|\mathrm{W}_3\rangle = (|100\rangle + |010\rangle + |001\rangle)/\sqrt{3}$, on the other hand, only achieves $\mathcal{F}_{N=3} = 4/3$. For $N = 4$, one has:

$$\begin{aligned}\mathcal{F}_{N=4} = \frac{1}{4}\,(14 \quad &- \quad \mathrm{Tr}(\rho_{234}^2 + \rho_{134}^2 + \rho_{124}^2 + \rho_{123}^2) \\ &- \quad \mathrm{Tr}(\rho_{12}^2 + \rho_{13}^2 + \rho_{14}^2 + \rho_{23}^2 + \rho_{24}^2 + \rho_{34}^2) \\ &- \quad \mathrm{Tr}(\rho_1^2 + \rho_2^2 + \rho_3^2 + \rho_4^2)\big)\end{aligned} \tag{4.8}$$

For this case, neither the generalized GHZ state $|\mathrm{GHZ}_4\rangle = (|0000\rangle + |1111\rangle)/\sqrt{2}$ ($\mathcal{F}_{N=4} = 7/4$) nor the W state $|\mathrm{W}_3\rangle = (|1000\rangle + |0100\rangle + |0010\rangle + |0001\rangle)/2$

($\mathcal{F}_{N=4} = 3/2$) obtain the maximum of $\mathcal{F}_{N=4} = 2$. The maximum is instead obtained by a state which maximizes the mixedness of the single-spin reduced density matrices ($\mathrm{Tr}\,\rho_i^2 = \frac{1}{2}$, $i = 1, 2, 3, 4$) and which yields $\mathrm{Tr}(\rho_{1,2}^2 + \rho_{1,3}^2 + \rho_{1,4}^2) = \mathrm{Tr}(\rho_{3,4}^2 + \rho_{2,4}^2 + \rho_{2,3}^2) = 1$.

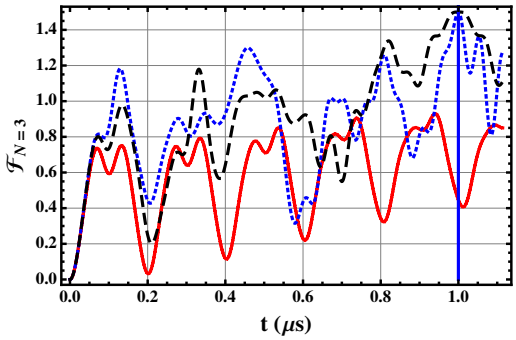

Figure 4.6.: Time evolution of the genuine three-body entanglement ($\omega_1 = 0.16$ MHz, $\omega_2 = 0.12$ MHz, $\omega_3 = 0.70$ MHz, $g_x^{(1)} = 4.93$ MHz, $g_x^{(2)} = 2.52$ MHz, $g_y^{(1)} = 4.11$ MHz, $g_y^{(2)} = 0.65$ MHz) without control (red continuous line), with optimization of the fidelity (blue dotted line) and with an additional minimization of the curvature (black dashed line).

In order to test the SOC algorithm for multipartite entanglement, we use a spin chain

$$\mathcal{H}_0 = \sum_{k=1}^{N} \frac{\omega_k}{2} \sigma_z^{(k)} + \sum_{k=1}^{N-1} \left(g_x^{(k)} \sigma_x^{(k)} \sigma_x^{(k+1)} + g_y^{(k)} \sigma_y^{(k)} \sigma_y^{(k+1)} \right) \tag{4.9}$$

with nearest neighbor couplings $g_j^{(k)}$ and a small splitting ω_k in z-direction, whose values can be found in the caption of Fig. 4.6. We only control the end spins by the same control Hamiltonian as in Eqs. (4.1) and (4.3). As initial state, one chose the state with $\theta_1 = 0.66$, $\theta_2 = 2.63$, $\theta_3 = 0$, $\phi_1 = \phi_2 = \phi_3 = 0$, according to Eq. (4.4). Fig. 4.6 displays the time evolution of the controlled system (with and without minimization of the curvature) as well as of the uncontrolled one. As expected, without control there is no chance to attain a maximally three-body entangled state. With a control pulse that contains $n_{\max} = 6$ Fourier components, one obtains a fidelity of $1.5 - 10^{-11}$. The minimization of the curvature enlarges the duration of high entanglement ($\mathcal{F}_{N=3} > 1.499$) from 7 to 60 ns, while the power needed to reach this control goal is even a few percent lower than without targeting

minimal curvature. This can be explained by a cooperative effect of the control and the internal dynamics of the spin chain which itself creates many-body entanglement. Control becomes the more difficult, the larger the system is: for $N = 4$, one is able to enlarge the duration of high entanglement ($\mathcal{F}_{N=4} > 1.999$) from 7 to 25 ns. This is what one would intuitively expect: the longer the spin chain is, the less control can be exerted on the whole chain by only controlling the end spins.

4.3. Mediated Interaction in Spin Chains

As demonstrated in Sec. 4.2, entanglement cannot be generated on a time scale substantially shorter than $t_{\text{char}} = \pi/(4g_{\max})$. If, however, the direct interaction between two spins is weak (*e.g.* if dipole-dipole interaction becomes negligible due to a large separation between the to-be-entangled spins [64]), then the spins cannot be entangled on a reasonable time scale. If, on the other hand, one uses spins that lie in between the to-be-entangled spins, *i.e.* if each spin of such a chain is coupled to its nearest neighbor, then entanglement can be generated between the end spins by exploiting the mediated interaction between them. In the experiment, the situation of a spin chain could be achieved by implanting spins in between the to-be-entangled spins. Recent proposals are based on nitrogen-vacancy centers [65, 66]. As will be shown, Smooth Optimal Control is capable to entangle the end spins of a chain even if the couplings of the spins are known only up to a certain accuracy. This corresponds to an experimentally relevant situation where the couplings can only be measured up to a certain accuracy. But let us first start from the assumption that all couplings are known exactly:

One uses the Hamiltonian Eq. (4.9) and only controls the end spins with a control Hamiltonian of the form of Eqs. (4.2) and (4.3). As maximal entanglement is the goal of the optimization, an entanglement measure seems to be the right choice for the target functional. In contrast to Sec. 4.2.1, one deals with mixed states, as one is only interested in the reduced density matrix of the end spins. Unfortunately, most entanglement measures are non-analytical for mixed states, meaning that the target functional and its derivatives have to be evaluated numerically. This computation is not only time-consuming but

also unreliable. Instead, one uses the lower bound Eqs. (1.67) and (1.68)

$$\mathcal{F} = 2\,\mathrm{Tr}(\rho_{1N}^2) - \mathrm{Tr}(\rho_1^2) - \mathrm{Tr}(\rho_N^2) \tag{4.10}$$

of the tangle of the end spins. Here, ρ_{1N} is the reduced density matrix of the two end spins, whereas ρ_1 and ρ_N are the reduced single-spin density matrices of the end spins. Because of the analytic (here: quadratic) dependence on the density matrices, the target functional and its derivatives with respect to the control parameters can be calculated analytically.

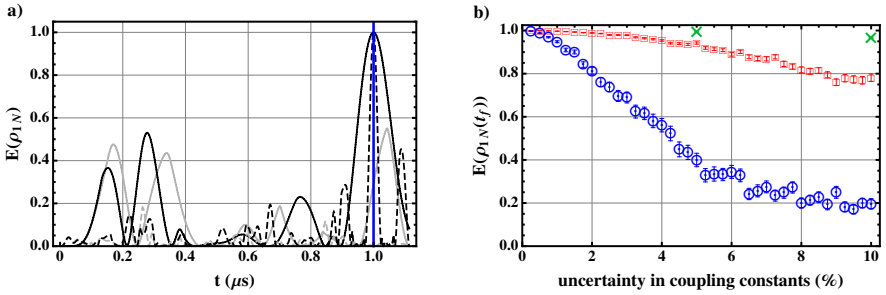

Figure 4.7.: a) Time evolution of the entanglement of formation E of the end spins in two random configurations of a chain with $N = 3$ (solid line, $g_x^{(1)} = 0.78$ MHz, $g_x^{(2)} = 1.48$ MHz, $g_y^{(1)} = 1.27$ MHz, $g_y^{(2)} = 2.65$ MHz, $\omega_1 = 0.91$ MHz, $\omega_2 = 0.97$ MHz, and $\omega_3 = 0.40$ MHz) and $N = 4$ spins (dotted line, $g_x^{(1)} = 4.36$ MHz, $g_x^{(2)} = 1.61$ MHz, $g_x^{(3)} = 5.33$ MHz, $g_y^{(1)} = 1.02$ MHz, $g_y^{(2)} = 8.82$ MHz, $g_y^{(3)} = 1.29$ MHz, $\omega_1 = 0.57$ MHz, $\omega_2 = 0.55$ MHz, $\omega_3 = 0.81$ MHz, and $\omega_4 = 0.42$ MHz). In both cases, the system dynamics of the chain (grey) are not able to entangle the end spins, whereas a control pulse with $n_{\max} = 6$ frequency components (black) can generate maximal entanglement at $t_f = 1\,\mu$s (see fat vertical line). [15] b) Entanglement of formation of the controlled system of $N = 3$ (rectangles) and 4 (circles) if the pulse optimized for the above coupling configuration is applied to a system where the coupling constants are only known up to a certain error. The error bars result from a test ensemble of 100 coupling configurations. In the case of $N = 3$ spins, the maximization of the averaged target functional for a test ensemble of ten coupling configurations (crosses) significantly increases the amount of entanglement (5%: 94.2%→ 99.3%; 10%: 78.1%→ 96.7%).

We tested SOC on chains with $N = 3$ and 4 spins and random configurations of couplings (see Fig. 4.7 a)). The initial states are random, too, since the

result should be independent of the initial state: it is $\theta_1 = 1.39$, $\theta_2 = 1.28$, $\theta_3 = 0.71$, $\phi_1 = 6.03$, $\phi_2 = 0.95$, and $\phi_3 = 5.30$ for $N = 3$ and $\theta_1 = 1.60$, $\theta_2 = 1.31$, $\theta_3 = 0.94$, $\theta_4 = 0.44$, $\phi_1 = 5.24$, $\phi_2 = 6.21$, $\phi_3 = 5.85$, and $\phi_4 = 6.07$ for $N = 4$, according to the notation Eq. (4.4). As illustrated in Fig. 4.7 a), where the time evolution of entanglement of formation is depicted, one is able to create entanglement between the end spins with an infidelity of at most 10^{-5} at $t_f = 1\,\mu$s by using a control pulse with $n_{\max} = 6$ frequency components. This has to be compared with the case of no control, where only little entanglement ($E < 0.6$) can be generated. Nevertheless, if one wants to entangle the end spins in an experiment, one knows the coupling constants only up to a certain extent [67]. In order to estimate how much entanglement is lost in the case of limited knowledge of the couplings, one can apply the pulse optimized for certain couplings to an ensemble of spins whose couplings deviate by an uncertainty ϵ from the orginal ones and measure the value of the target functional. This is what is depicted in Fig. 4.7 b): the value of the target functional is the more affected, the more spins are involved. For $N = 4$ spins only 20% entanglement can be generated in the presence of 10% uncertainty, compared to 80% entanglement for $N = 3$. This coincides with the intuitive picture according to which the dynamics is the more perturbed the more spins can participate in such a perturbation.

Similarly to the inhomogeneous broadening in chapter 3, one can reduce the decrease of entanglement by using as a target functional the averaged target functional $\langle \mathcal{F} \rangle_{\mathbf{g}}$, where $\mathbf{g} = \{g_{x,y}^{(k)}\}_{k=1,\ldots,N-1}$ corresponds to a coupling configuration. Surprisingly, one can construct a robust pulse based on a small ensemble of only ten coupling configurations. More precisely, one optimized the averaged target functional for the case of $N = 3$ spins and an uncertainty of 5 and 10%, respectively. In order to guarantee that the number of coupling configurations with which one has calculated the average is sufficient, the pulses were tested on an ensemble of 100 coupling configurations. As one can see in Fig. 4.7 b), the optimization of the averaged target functional with 5% uncertainty results in an improvement from 94.2 to 99.3%, whereas for 10% uncertainty the entanglement is improved from 78.1 to 96.7%. In terms of control amplitude $|\mathbf{a}|^2$, the improvement of 10% uncertainty requires the double amplitude compared to the case of no uncertainty, whereas the amplitude

needed to achieve robustness for 5% uncertainty is roughly the same as the amplitude needed to maximize entanglement without uncertainty.

Concerning the computational power required by SOC, spin chains clearly show the limitations of SOC in the case of many-body systems: Since the dimension of the Hilbert space increases exponentially according to 2^N, where N is the number of qubits, the dimension of the Floquet matrix increases as $2^N \times 2^N$. As one needs to take approximately a hundred Brillouin zones, the matrices with which one has to work quickly become very large. The consequence of these large matrices is that also the time necessary to perform all matrix manipulations exceeds any reasonable limit. Whereas the pulse that entangles the end spins of a chain of $N = 3$ spins is found after a few hours of iterations, the time needed for $N = 4$ spins lies on the time scale of days. In this way, chains of $N = 5$ or more spins become numerically hardly tractable. Concerning the robustness of the pulses with respect to different coupling configurations, the required time of the algorithm for reaching the optimum scales linearly with the size of the ensemble. Furthermore, the size of the ensemble has to be enlarged with the loss of fidelity due to different coupling configurations. Only by using a larger ensemble, one can guarantee that the average is independent of the size of the ensemble. This is why concerning the robust pulses we restricted us to the case of $N = 3$ spins. One may argue that one can enlarge the number of qubits with which one can deal by truncating the Hilbert space, *e.g.* by permitting for just one excitation in the whole system. In the case of one excitation one could reduce the dimension of the Hilbert space from 2^N to $N + 1$ (ground state + one excitation in each spin). However, our calculations have shown that the end spins cannot be entangled in a system whose Hilbert space is truncated. One infers from this fact that the whole Hilbert space is necessary in order to distribute entanglement.

Since the peak in the time evolution of entanglement in Fig. 4.7 a) is rather sharp, one may consider using the same method as in Sec. 4.2 to reduce the curvature. Unfortunately, this attempt results in very poor values for both entanglement and curvature. The reason why it is not possible to generate maximal entanglement and minimal curvature, is the result of a fundamental limitation. While in Sec. 4.2 the interaction between the two spins was direct, it is here indirectly induced by nearest neighbor couplings. In order to understand the influence of the indirect coupling on the minimization of curvature,

one can consider the following Hamiltonian:

$$\mathcal{H}_\alpha = (1-\alpha)\left(g_x^{(N)}\sigma_x^{(N)}\sigma_x^{(1)} + g_y^{(N)}\sigma_y^{(N)}\sigma_y^{(1)}\right) + \sum_{k=1}^{N}\frac{\omega_k}{2}\sigma_z^{(k)} + \alpha\sum_{k=1}^{N-1}\left(g_x^{(k)}\sigma_x^{(k)}\sigma_x^{(k+1)} + g_y^{(k)}\sigma_y^{(k)}\sigma_y^{(k+1)}\right) \tag{4.11}$$

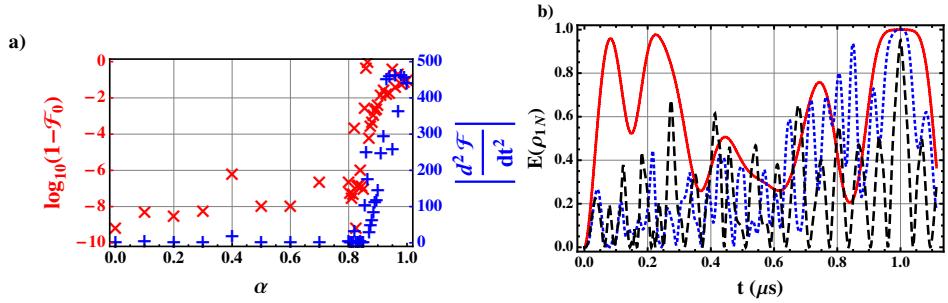

Figure 4.8.: a) Lower bound of entanglement $\mathcal{F}_0$ between the end spins (red '$\times$') and their curvature at the end of the control (blue '$+$') as a function of the indirect interaction α taking place in the dynamics ($g_x^{(3)} = 2.32$ MHz, $g_y^{(3)} = 1.74$ MHz). For $\alpha > 0.85$ entanglement and curvature are not independent anymore. b) Time evolution of entanglement of formation of the end spins for three different values of indirect coupling: $\alpha = 0$ (only direct interaction, red solid line), $\alpha = 0.85$ (blue dotted line), $\alpha = 1$ (only indirect coupling, dashed black line).

Here $\alpha \in [0, 1]$ is a parameter which specifies the strength of the direct interaction between the end spins: $\alpha = 0$ means only direct coupling, whereas $\alpha = 1$ represents only indirect coupling. We optimized entanglement and curvature for different values of α for a system with $N = 3$ spins. The result of this optimization is shown in Fig. 4.8 a) where the peak value of entanglement (characterized in terms of the lower bound Eq. (4.10)) and its curvature is depicted as a function of the parameter α. One can see in Fig. 4.8 a) that for $\alpha < 0.85$ the infidelity and the curvature are arbitrarily low, while they both increase for $\alpha > 0.85$. Therefore, there exists a threshold $\alpha_{\text{thr}} = 0.85$, beyond which entanglement and curvature cannot be optimized independently from each other, *i.e.* 15% of direct interaction between the end spins are necessary in order to vary the curvature independently of the entanglement. Fig. 4.8 b) is an illustration of Fig. 4.8 a): it depicts the time evolution of entanglement for

three different values of α. For $\alpha = 0$ (only direct coupling) the plateau of high entanglement is maximal; one also obtains a fairly broad interval of maximal entanglement for $\alpha = 0.85$, whereas $\alpha = 1$ (only indirect coupling) results in a sharp peak (high curvature) which does not reach maximal entanglement.

4.4. Decoupling of system and environment

Decoherence is the main source of errors in quantum information processing. In the case of NV centers, decoherence is mainly caused by the surrounding carbon and nitrogen atoms [67]. One possibility to include decoherence effects into our model, is to describe the time evolution by a Lindblad master equation [68]. However, since spin dynamics are often dominated by non-Markovianity [69], a Markovian Lindblad equation seems to be hardly reliable. Another solution consists in exlicitly taking N spins (system + environment) and calculating their unitary evolution $\mathcal{U}(t)$. However, as elucidated in the previous section, this is only possible up to $N = 4$ spins. If one takes in the following a system $\mathcal{S}$ of dimension $d_\mathcal{S}$ with an initial state $|\psi_\mathcal{S}\rangle$ and an environment $\mathcal{E}$ of dimension $d_\mathcal{E}$ with the initial state $|\phi_\mathcal{E}\rangle$, then the evolution of the system $\mathcal{S}$ can be obtained by tracing out all environmental degrees of freedom $\mathcal{E}$:

$$\rho_\mathcal{S}(t) = \mathrm{Tr}_\mathcal{E}\left(\mathcal{U}(t)|\psi_\mathcal{S}\rangle|\phi_\mathcal{E}\rangle\langle\psi_\mathcal{S}|\langle\phi_\mathcal{E}|\mathcal{U}^\dagger(t)\right) \tag{4.12}$$

This state has to be compared to the evolution of the system without environment, which would be

$$\rho_f(t) = e^{-i\mathcal{H}_\mathcal{S}t}|\psi_\mathcal{S}\rangle\langle\psi_\mathcal{S}|e^{i\mathcal{H}_\mathcal{S}t}, \tag{4.13}$$

where $\mathcal{H}_\mathcal{S}$ is the Hamiltonian of the system. If the system should be decoupled from the environment during a time interval $[0, t_f]$, this would lead to a time-averaged target functional

$$\mathcal{F}_{\mathrm{av}} = \frac{1}{t_f}\int_0^{t_f} \mathrm{d}t \|\rho_\mathcal{S}(t) - \rho_f(t)\|_{HS}^2 = \frac{1}{t_f}\int_0^{t_f} \mathrm{d}t\, \mathrm{Tr}\left(\rho_\mathcal{S}(t) - \rho_f(t)\right)\left(\rho_\mathcal{S}^\dagger(t) - \rho_f^\dagger(t)\right) \tag{4.14}$$

Time-averaged target functionals are particularly easy to evaluate in the framework of SOC, as they mean to merely integrate an exponential, *e.g.* with

$|\psi_0\rangle = |\psi_\mathcal{S}\rangle|\phi_\mathcal{E}\rangle$ and the basis $|\phi_l\rangle$ of $\mathcal{E}$ as well as the basis $|\psi_l\rangle$ of $\mathcal{S}$

$$\begin{aligned}
\int_0^{t_f} \mathrm{d}t\, \mathrm{Tr}\, \rho_\mathcal{S}(t)\rho_\mathcal{S}^\dagger(t) &= \int_0^{t_f} \mathrm{d}t\, \mathrm{Tr} \sum_{\substack{l_1,l_2\\ k_1,k_2,k_3,k_4\\ \nu_1,\nu_2,\nu_3,\nu_4}} \langle\phi_{l_1}|\chi_{k_1\nu_1}\rangle\langle\Phi_{k_1}(0)|\psi_0\rangle e^{i(\nu_1\Omega t-\epsilon_{k_1})t} \\
&\quad \times\langle\psi_0|\Phi_{k_2}(0)\rangle\langle\chi_{k_2\nu_2}|\phi_{l_1}\rangle e^{-i(\nu_2\Omega t-\epsilon_{k_2})t} \\
&\quad \times\langle\phi_{l_2}|\chi_{k_3\nu_3}\rangle\langle\Phi_{k_3}(0)|\psi_0\rangle e^{i(\nu_3\Omega t-\epsilon_{k_3})t} \\
&\quad \times\langle\psi_0|\Phi_{k_4}(0)\rangle\langle\chi_{k_4\nu_4}|\phi_{l_2}\rangle e^{-i(\nu_4\Omega t-\epsilon_{k_4})t} \\
&= \sum_{\substack{l_1,l_2,l_3\\ k_1,k_2,k_3,k_4\\ \nu_1,\nu_2,\nu_3,\nu_4}} \frac{e^{i((\nu_1-\nu_2+\nu_3-\nu_4)\Omega-(\epsilon_{k_1}-\epsilon_{k_2}+\epsilon_{k_3}-\epsilon_{k_4}))t_f}-1}{i((\nu_1-\nu_2+\nu_3-\nu_4)\Omega-(\epsilon_{k_1}-\epsilon_{k_2}+\epsilon_{k_3}-\epsilon_{k_4}))} \\
&\quad \times\langle\psi_{l_3}|\langle\phi_{l_1}|\chi_{k_1\nu_1}\rangle\langle\Phi_{k_1}(0)|\psi_0\rangle\langle\psi_0|\Phi_{k_2}(0)\rangle\langle\chi_{k_2\nu_2}|\phi_{l_1}\rangle \\
&\quad \times\langle\phi_{l_2}|\chi_{k_3\nu_3}\rangle\langle\Phi_{k_3}(0)|\psi_0\rangle\langle\psi_0|\Phi_{k_4}(0)\rangle\langle\chi_{k_4\nu_4}|\phi_{l_2}\rangle|\psi_{l_3}\rangle \quad (4.15)
\end{aligned}$$

Here, we inserted the time evolution operator from Eq. (1.28) and performed the integral over the exponentials. As one can see, the evaluation is cumbersome: indeed, if one takes 100 Brillouin zones, in order to evaluate the sum from above, one has to add $d_\mathcal{S}d_\mathcal{E}^2(d_\mathcal{S}d_\mathcal{E})^4\cdot 10^8$ summands, which is extremely time-consuming from a numerical point of view. The high number of summands can be explained by the fact that in the time-average Eq. (4.15) the propagator $\mathcal{U}$ appears four times. One therefore searches for a target functional where the propagator appears less often.

In order to find a new target functional, one first notices that the influence Eq. (4.12) of the environment on the system can be described by a quantum channel with Kraus operators F_k, *i.e.*

$$\rho_\mathcal{S} = \sum_{k=1}^{d_\mathcal{E}} F_k|\psi_\mathcal{S}\rangle\langle\psi_\mathcal{S}|F_k^\dagger. \tag{4.16}$$

The Kraus operators read $F_k = (\mathbb{1}_{d_\mathcal{S}}\otimes\langle\phi_k|)\mathcal{U}(\mathbb{1}_{d_\mathcal{S}}\otimes|\phi_\mathcal{E}\rangle) \equiv \langle\phi_k|\mathcal{U}|\phi_\mathcal{E}\rangle$, where $|\phi_k\rangle$ is a basis state of $\mathcal{E}$. Now, we want the system to evolve as if no environment was present. This requires that

$$\sum_{k=1}^{d_\mathcal{E}} F_k|\psi_\mathcal{S}\rangle\langle\psi_\mathcal{S}|F_k^\dagger \stackrel{!}{=} \mathcal{U}_\mathcal{S}|\psi_\mathcal{S}\rangle\langle\psi_\mathcal{S}|\mathcal{U}_\mathcal{S}, \tag{4.17}$$

where $\mathcal{U}_\mathcal{S}(t_f) = e^{-i\mathcal{H}_\mathcal{S}t_f}$ is the time evolution without environment at the moment t_f in time, induced by the system Hamiltonian $\mathcal{H}_\mathcal{S}$. If Eq. (4.17) is

valid, it follows that $F_k = c_k \mathcal{U}_\mathcal{S}$ with an arbitrary complex number c_k. If one multiplies F_k with $\mathcal{U}_\mathcal{S}^\dagger$ and takes the trace, one obtains

$$\mathrm{Tr}(F_k \mathcal{U}_\mathcal{S}^\dagger) = d_\mathcal{S} c_k. \tag{4.18}$$

Next, one will make use of the following identity:

$$\begin{aligned}
\sum_{k=1}^{d_\mathcal{E}} \mathrm{Tr}_\mathcal{S}(F_k F_k^\dagger) &= \sum_{k=1}^{d_\mathcal{E}} \mathrm{Tr}_\mathcal{S}\left(\langle\phi_k|\mathcal{U}|\phi_\mathcal{E}\rangle\langle\phi_\mathcal{E}|\mathcal{U}^\dagger|\phi_k\rangle\right) && (4.19)\\
&= \sum_{k=1}^{d_\mathcal{E}} \mathrm{Tr}\left(\mathcal{U}\, \mathbb{1}_{d_\mathcal{S}} \otimes |\phi_\mathcal{E}\rangle\langle\phi_\mathcal{E}|\, \mathcal{U}^\dagger\right) && (4.20)\\
&= \sum_{k=1}^{d_\mathcal{E}} \mathrm{Tr}\left(\mathcal{U}^\dagger\mathcal{U}\, \mathbb{1}_{d_\mathcal{S}} \otimes |\phi_\mathcal{E}\rangle\langle\phi_\mathcal{E}|\right) && (4.21)\\
&= \langle\phi_\mathcal{E}|\phi_\mathcal{E}\rangle\, \mathrm{Tr}_\mathcal{S}(\mathbb{1}_{d_\mathcal{S}}) = d_\mathcal{S} && (4.22)
\end{aligned}$$

One can finally calculate the difference between actual and desired dynamics:

$$\begin{aligned}
\sum_{k=1}^{d_\mathcal{E}} \|F_k - c_k\mathcal{U}_\mathcal{S}\|_{HS}^2 &= \sum_{k=1}^{d_\mathcal{E}} \mathrm{Tr}\Bigg(F_k F_k^\dagger - \frac{1}{d_\mathcal{S}} F_k \mathcal{U}_\mathcal{S}^\dagger\, \mathrm{Tr}(\mathcal{U}_\mathcal{S} F_k^\dagger)\\
&\quad - \frac{1}{d_\mathcal{S}}\, \mathrm{Tr}(F_k\mathcal{U}_\mathcal{S}^\dagger)\mathcal{U}_\mathcal{S} F_k^\dagger + \frac{1}{d_\mathcal{S}^2}\, \mathrm{Tr}(F_k\mathcal{U}_\mathcal{S}^\dagger)\mathcal{U}_\mathcal{S}\mathcal{U}_\mathcal{S}^\dagger\, \mathrm{Tr}(\mathcal{U}_\mathcal{S} F_k^\dagger)\Bigg)\\
&= d_\mathcal{S} - \frac{1}{d_\mathcal{S}} \sum_{k=1}^{d_\mathcal{E}} \left|\mathrm{Tr}(F_k\mathcal{U}_\mathcal{S}^\dagger)\right|^2 && (4.23)
\end{aligned}$$

Since $\sum_{k=1}^{d_\mathcal{E}} \|F_k - c_k\mathcal{U}_\mathcal{S}\|^2 \geq 0$, a good target functional is

$$\mathcal{F} = \frac{1}{d_\mathcal{S}^2} \sum_{k=1}^{d_\mathcal{E}} \left|\langle F_k, \mathcal{U}_\mathcal{S}\rangle_{HS}\right|^2. \tag{4.24}$$

$\mathcal{F}$ is non-negative and adopts its maximal value $\mathcal{F} \equiv 1$ if and only if $F_k \propto \mathcal{U}_\mathcal{S}$ for all k. In the new target functional, the propagator $\mathcal{U}$ only appears twice. However, this still means to add 10^4 summands, when considering 100 Brillouin zones. Indeed, a numerical evaluation of the time-average is faster than computing the analytical expression. We did calculations for the paradigmatic model of a central spin $\mathcal{S}$ coupled to $N_\mathcal{E}$ environmental spins, which is described by the following Hamiltonian

$$\mathcal{H} = \frac{\omega_0}{2}\sigma_z^{(\mathcal{S})} + \sum_{k=1}^{N_\mathcal{E}} \frac{\omega_k}{2}\sigma_z^{(k)} + \sum_{k=1}^{N_\mathcal{E}} \left(g_x^{(k)}\sigma_x^{(\mathcal{S})}\sigma_x^{(k)} + g_y^{(k)}\sigma_y^{(\mathcal{S})}\sigma_y^{(k)}\right), \tag{4.25}$$

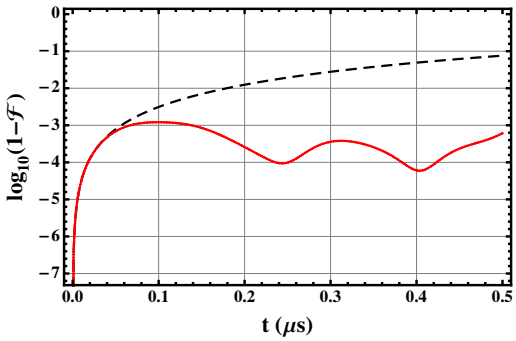

Figure 4.9.: The logarithmic infidelity is depicted for two interacting spins as a function of time for the uncontrolled case (black, dashed line) and the decoupling pulse controlling both spins (red, continuous line).

where we only controlled the system spin via a control Hamiltonian

$$\mathcal{H}_c = f_x(t)\sigma_x^{(\mathcal{S})} + f_y(t)\sigma_y^{(\mathcal{S})}. \tag{4.26}$$

Unfortunately, even if the target functional Eq. (4.24) obtained high values at single points in time, it was never possible to attain a substantial improvement of the time-averaged fidelity compared with the case of no control. For two spins ($\omega_{\mathcal{S}} = \omega_1 = 10$ Hz, $g_x = 0.25$ MHz, $g_y = 0.5$ MHz), for example, we obtained only an improvement of less than 0.1%. If, however, we control both spins by the the same type of control Hamiltonian as above, we obtain an improvement of 99.9% compared to 97.4% without control. In Fig. 4.9, the time evolution of the target functional Eq. (4.24) is depicted for the case of no control and control of both spins. One sees that in the case of control the fidelity is always better than 99.9%, whereas it becomes monotonically worse in the case of no control. We learn from the above study that in order to implement a certain dynamics, one has to control each spin individually. Unfortunately, in the experiment, it is very rare that one can address each environmental spin individually. This is why, in the following, we will concentrate on the problem of decoupling the system spin from the environmental ones at a specific *moment* in time by only controlling the system spin.

We performed calculations for $N_{\mathcal{E}} = 1, 2$ and 3 environmental spins. In all cases, we were able to decouple the dynamics of the central spin from the environmental ones, independently of the choice of the final time t_f. With $n_{\max} = 6$ Fourier components, minimal fidelities of 95% ($N_{\mathcal{E}} = 3$), 99% ($N_{\mathcal{E}} = 2$) and

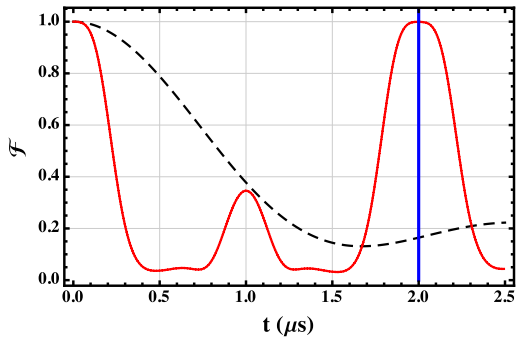

Figure 4.10.: Fidelity $\mathcal{F}$ of three environmental spins being decoupled, for the uncontrolled case (black, dashed line) and the decoupling pulse (red, continuous line). The system is characterized by $\omega_{\mathcal{S}} = \omega_k =$ 10 Hz, $g_x^{(k)} = 0.25$ MHz, $g_y^{(k)} = 0.50$ MHz ($k = 1, 2, 3$).

99.9% ($N_{\mathcal{E}} = 1$) are achieved. Most importantly, the obtained pulses work independently of the initial state $|\phi_{\mathcal{E}}\rangle$ of the environment. In Fig. 4.10, one can see the comparison of the time evolution between the controlled and the uncontrolled case. The uncontrolled system does not return to its initial (decoupled) state, whereas with control the system is decoupled from the environment at $t = t_f$. One might suspect that like in the previous chapter the fidelity presents a sharp peak in time. However, one even obtains a plateau of high decoupling. As the target functional does not depend on the exact state of the system and environment, one obtains a relative freedom for the control in Hilbert space. The plateau of the fidelity is a consequence of this freedom.

To conclude this section, we would like to compare our approach with state of the art decoupling scenarios. The most commonly used decoupling method is dynamical decoupling [70, 71]. Dynamical decoupling can be described as a generalization of a spin echo sequence. Whereas spin echo compensates for inhomogeneous broadening, dynamical decoupling can cope with a very general class of system-environment couplings. In spin echo, the π_x-pulse effectively behaves like a propagator backwards in time; dynamical decoupling uses a series of strong, approximately instantaneous pulses such that the spins evolve towards their initial state. However, in experimental situations where hard pulses cannot be realized because of a limited bandwidth, dynamical decoupling will not work. This is the regime in which the smooth continuous driving of SOC is advantageous. A disadvantage of our method is that one has to know the couplings to the environment exactly, whereas dynamical decoupling only

needs to know the type of interaction, *e.g.* $\sigma_z \otimes E$, where E is an operator acting on the Hilbert space of the environment. On the other hand, with SOC one can gain a certain robustness against different couplings by averaging over a few coupling configurations. As a test, we consider the decoupling from one spin ($N_{\mathcal{E}} = 1$): by using a test ensemble of 20 coupling configurations with an uncertainty of 20%, we were able to achieve an extremely robust pulse with an average fidelity of 99.99% compared to $(94.1 \pm 0.7)\%$ for a pulse which has not been optimized for different coupling configurations (the error results from an application to a test ensemble of 100 coupling configurations). The robustness, however, is paid at the price of a pulse intensity $|\mathbf{a}|^2$ which is ten times larger than without averaging.

4.5. A Cavity interacting with a Spin Ensemble

In chapter 3, the coupling of a spin ensemble to the classical field of an antenna was studied. There, one neglected both the back action of a spin to another spin via the antenna and the quantum character of the control field. In the present section, we will treat a spin ensemble coupled to a cavity which is controlled by SOC pulses. The Hamiltonian of such a system is [72]

$$\mathcal{H}(t) = \omega_c a^\dagger a + \frac{1}{2}\sum_{k=1}^{N} \omega_k \sigma_k^z + i \sum_{k=1}^{N} \left(g_k \sigma_k^- a^\dagger - g_k^* \sigma_k^+ a \right) \tag{4.27}$$

$$- \; i \left(\eta(t) a^\dagger e^{-i\omega_p t} - \eta^*(t) a e^{i\omega_p t} \right), \tag{4.28}$$

where N is the number of spins, ω_c the central frequency of the cavity, ω_k are the frequencies of the spins, g_k the coupling constants between cavity and spins, and $\eta(t)$ is a time-dependent external field. In the rotating frame defined by

$$\mathcal{H}_{\text{rot}} = e^{i\mathcal{H}_0 t} \mathcal{H}\, e^{-i\mathcal{H}_0 t} - \mathcal{H}_0 \tag{4.29}$$

with

$$\mathcal{H}_0 = \omega_p \left(a^\dagger a + \frac{1}{2}\sum_{k=1}^{N} \sigma_k^z \right), \tag{4.30}$$

one obtains

$$\mathcal{H}_{\text{rot}} = (\omega_c - \omega_p)\, a^\dagger a + \frac{1}{2}\sum_{k=1}^{N} (\omega_k - \omega_p)\, \sigma_k^z + i \sum_{k=1}^{N} \left(g_k \sigma_k^- a^\dagger - g_k^* \sigma_k^+ a \right) - i \left(\eta a^\dagger - \eta^* a \right). \tag{4.31}$$

If the external field is on resonance with the cavity ($\omega_c = \omega_p$), the Hamiltonian takes the form

$$\mathcal{H}_{\text{res}} = \frac{1}{2}\sum_{k=1}^{N} \delta\omega_k \sigma_k^z + i\sum_{k=1}^{N}\left(g_k\sigma_k^- a^\dagger - g_k^*\sigma_k^+ a\right) - i\left(\eta a^\dagger - \eta^* a\right), \tag{4.32}$$

where $\delta\omega_k = \omega_k - \omega_p$ is the detuning of the spins from the resonance frequency.

In the following, one wants to have an equation of motion for the cavity operator a. This equation is of the Volterra type and is derived in the appendix B. There, we followed the derivation in [72]. Starting from the Heisenberg equation of motion, taking the quantum mechanical average $\langle\cdot\rangle$ and making the Holstein-Primakoff approximation $\langle\sigma_k^z\rangle \approx -1$ [73], one obtains

$$\dot{A}(t) = -[\kappa + i(\omega_c - \omega_p)]A(t) + \sum_{k=1}^{N} g_k B_k(t) - \eta(t) \tag{4.33}$$

$$\dot{B}_k(t) = -(\gamma + i\delta\omega_k)B_k(t) - g_k^* A(t), \tag{4.34}$$

where $A(t) = \langle a(t)\rangle$ and $B_k(t) = \langle\sigma_k^-\rangle$, and the cavity losses κ and the losses of the spin ensemble γ are introduced by hand. The above equations result in the Volterra equation

$$A(t) = F(t) + \int_0^t \mathrm{d}\tau\, K(t-\tau)A(\tau) \tag{4.35}$$

with the driving force

$$F(t) = -\int_0^t \mathrm{d}\tau\, \eta(\tau) e^{-(i(\omega_c-\omega_p)+\kappa)(t-\tau)} \tag{4.36}$$

and the memory kernel

$$K(t-\tau) = e^{-(i(\omega_c-\omega_p)+\kappa)(t-\tau)}\Omega^2 \int_{-\infty}^{\infty} \mathrm{d}\omega\, \rho(\omega) \frac{e^{-(i(\omega-\omega_c)+\gamma-\kappa)(t-\tau)} - 1}{i(\omega-\omega_c)+\gamma-\kappa} \tag{4.37}$$

as well as the spectral density

$$\rho(\omega) = \sum_k \frac{g_k^2 \delta(\omega - \omega_k)}{\Omega^2}. \tag{4.38}$$

4.5.1. Discretization

Since the operators of the cavity are bosonic operators, the space they act on is infinitely large. In order to be solved on a computer, one has to discretize

the problem. In particular, the infinite ladder of states, typical for a harmonic oscillator, has to be cut off at a finite size. Here, one just considers one excitation in the cavity and the spins. Since the average excitation of a single spin is little - fact used by making the Holstein-Primakoff approximation -, the limitation to one excitation in the cavity and the spins is justified. The states are labelled as follows:

$$\begin{aligned}
|1\rangle &= |0\downarrow\ldots\downarrow\rangle \\
|2\rangle &= |0\uparrow\downarrow\ldots\downarrow\rangle \\
&\vdots \\
|N+1\rangle &= |0\downarrow\ldots\downarrow\uparrow\rangle \\
|N+2\rangle &= |1\downarrow\ldots\downarrow\rangle \\
|N+3\rangle &= |1\uparrow\downarrow\ldots\downarrow\rangle \\
&\vdots \\
|2N+2\rangle &= |1\downarrow\ldots\downarrow\uparrow\rangle
\end{aligned} \tag{4.39}$$

If the Hamiltonian Eq. (4.31) acts on these states, one obtains for $k = 1$

$$\mathcal{H}|1\rangle = -\sum_k \frac{\delta\omega_k}{2}|1\rangle - i\eta|N+2\rangle, \tag{4.40}$$

for $k = 2, \ldots, N+1$

$$\mathcal{H}|k\rangle = -\left(\sum_{j\neq k-1}\frac{\delta\omega_j}{2} - \frac{\delta\omega_{k-1}}{2}\right)|k\rangle + ig_{k-1}|N+2\rangle - i\eta|k+N+1\rangle \tag{4.41}$$

for k=$N+2$

$$\mathcal{H}|N+2\rangle = -\sum_k \frac{\delta\omega_k}{2}|N+2\rangle - i\sum_k g_k^*|k+1\rangle + i\eta^*|1\rangle \tag{4.42}$$

and finally for $k = N+3, \ldots, 2N+2$:

$$\mathcal{H}|k\rangle = -\left(\sum_{j\neq k-N-2}\frac{\delta\omega_j}{2} - \frac{\delta\omega_{k-N-2}}{2}\right)|k\rangle + i\eta^*|k-N-1\rangle \tag{4.43}$$

If one projects onto the state $\langle m|$, one obtains the matrix representation of the truncated Hamiltonian, with which one can calculate the Floquet eigenvalues and eigenvectors and finally the time evolution as described by Eq. (1.28). In contrast to the exponential dependence in the previous sections, the dimension

of the (truncated) Hilbert space depends now linearly on the number of spins. This permits us to consider spin ensemble of a bigger size ($N \approx 30$ with optimization, $N \approx 100$ without optimization, only in order to calculate the time evolution). In contrast to the previous sections, to consider only the one-excitation subspace is a good approximation because the average excitation of an ensemble spin is very low. In the above basis, the reduced density matrix of the cavity can be expressed as

$$\begin{aligned}\rho_c = \sum_{k=1}^{N+1} (&\langle k|\rho|k\rangle|0\rangle\langle 0| + \langle k|\rho|k+N+1\rangle|0\rangle\langle 1| \\ &+\langle k+N+1|\rho|k\rangle|1\rangle\langle 0| + \langle k+N+1|\rho|k+N+1\rangle|1\rangle\langle 1|)\end{aligned} \tag{4.44}$$

One can now calculate the squared expectation value $|\langle a(t)\rangle|^2$:

$$|\langle a\rangle|^2 = |\mathrm{Tr}\,(\rho_c a)|^2 = \left|\sum_{n=0,1}\langle n|\rho_c a|n\rangle\right|^2 = |\langle 1|\rho_c|0\rangle|^2 = \left|\sum_{k=1}^{N+1}\langle k+N+1|\rho|k\rangle\right|^2 \tag{4.45}$$

Next, one has to discretize the spectral density, *i.e.* to draw important representatives from it. To this end, one defines an interval $[\omega_s - \epsilon\omega_s, \omega_s + \epsilon\omega_s]$ and devides it into M subintervals with a width $\Delta\omega$. Each subinterval is centered around a frequency ω_k (see Fig. 4.11). If the subinterval is sufficiently small, then the approximation

$$\int_{\omega_k-\Delta\omega/2}^{\omega_k+\Delta\omega/2} \mathrm{d}\omega\, \rho(\omega) \approx \rho(\omega_k)\Delta\omega \tag{4.46}$$

holds. On the other hand, one has

$$\int_{\omega_k-\Delta\omega/2}^{\omega_k+\Delta\omega/2} \mathrm{d}\omega\, \rho(\omega) = \frac{1}{\Omega^2}\sum_{j=1}^{M} g_j^2 \int_{\omega_k-\Delta\omega/2}^{\omega_k+\Delta\omega/2} \mathrm{d}\omega\, \delta(\omega-\omega_j) = \frac{g_k^2}{\Omega^2}. \tag{4.47}$$

It follows that $g_k = \Omega\sqrt{\Delta\omega\rho(\omega_k)}$. In order to calculate $\Delta\omega$, one exploits the normalization condition

$$\int_{-\infty}^{\infty} \mathrm{d}\omega\, \rho(\omega) = \sum_{k=1}^{M} \rho(\omega_k)\Delta\omega = 1. \tag{4.48}$$

In total, one arrives at

$$g_k = \Omega\sqrt{\frac{\rho(\omega_k)}{\sum_{j=1}^{M}\rho(\omega_j)}}. \tag{4.49}$$

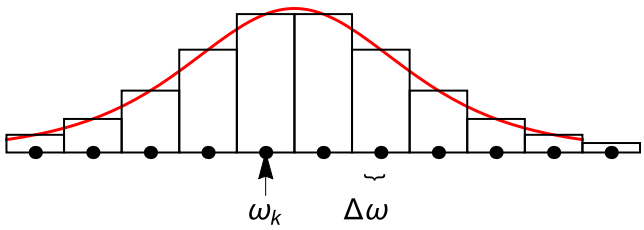

Figure 4.11.: Discretization of the spectral density. The frequency axis is divided into M subintervals with values ω_k and a width $\Delta\omega$ of the intervals.

In order to test our discretization scheme, we take a Lorentzian spectral density

$$\rho(\omega) = \frac{1}{\pi\Delta}\frac{1}{1+\left(\frac{\omega-\omega_s}{\Delta}\right)^2}. \tag{4.50}$$

As an experimentally relevant example one takes $\Delta/(2\pi) = 4$ MHz, $\Omega/(2\pi) = 9.55$ MHz, $\omega_s/(2\pi) = 2.6915$ GHz, it is $\omega_p = \omega_c = \omega_s$ [72]. As the Hamiltonian Eqs. (4.40)-(4.43) cannot account for losses, one sets $\gamma = \kappa = 0$. For the case of a Lorentzian spectral density and constant $\eta(t) \equiv \eta = 55.8$ kHz (rectangular pulse), the solution of the Volterra equation can be found analytically, it states $|\langle a(t)\rangle|^2 = y(t)^2$ with

$$y(t) = 4\frac{\eta\Delta}{\Delta^2+4\Omega_R^2} + \frac{\eta e^{-\Delta t/2}}{\Omega_R(\Delta^2+\Omega_R^2)}x(t) \tag{4.51}$$

$$x(t) = -4\Omega_R\Delta\cos(\Omega_R t) + (4\Omega_R^2-\Delta^2)\sin(\Omega_R t) \tag{4.52}$$

and the Rabi frequency $\Omega_R = \sqrt{4\Omega^2-\Delta^2}$. After the pulse is turned off, say after the time t_{off}, one has the following evolution:

$$y_{\text{off}}(t) = \frac{\eta e^{-\Delta(t-t_{\text{off}})/2}}{\Omega_R(\Delta^2+\Omega_R^2)}x_{\text{off}}(t) \tag{4.53}$$

$$x_{\text{off}}(t) = x(t-t_{\text{off}}) \tag{4.54}$$

The evolution can be described as follows (see Fig. 4.12, where the time evolution of $|\langle a(t)\rangle|^2$ is depicted): during the pulse is switched on, oscillations saturate to a constant level of $|\langle a\rangle|_{\text{sat}} = \frac{4\eta\Delta}{\Delta^2+4\Omega_R^2}$, *i.e.* the cavity is pumped by the external field η; after the pulse is switched off, the cavity excitation decays to zero on a time scale set by the inhomogeneous broadening Δ, and energy is transferred into the spin ensemble.

As a test, one performs calculations with different widths ϵ and different ensemble sizes M. One started with an initially empty cavity and no excitations in the spins. It is found out that an ensemble of $M \approx 100$ spins and a width of $\epsilon = 0.02$ is enough to model the analytical solution. As one can see in Fig. 4.12, an ensemble of $M = 70$ spins, on the other hand, results in important finite-size effects after 500 ns.

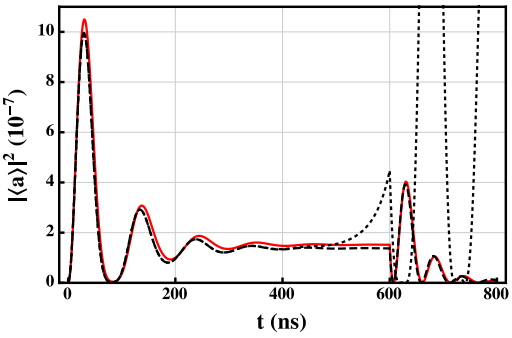

Figure 4.12.: Exact solution of the Volterra equation for a Lorentzian spectral density (red continuous curve) and approximation by ensembles of $N = 70$ (dotted line) and 100 spins (dashed line, both have the same width $\epsilon = 0.02$). The rectangular pulse is switched off at $t = 600$ ns. 100 spins are enough in order to simulate the Lorentzian spectral density, whereas 70 spins do not suffice.

An even more relevant case for the experiment is that of a *q-Gaussian* distribution [72]

$$\rho_q(\omega) = \left(1 - (1-q)\frac{(\omega - \omega_s)^2}{\Delta^2}\right)^{\frac{1}{1-q}}. \tag{4.55}$$

Here ρ_q is of a Lorentzian shape for $q = 2$, while $q \to 1$ results in a Gaussian distribution. For Fig. 4.13, where the time evolution of $|\langle a \rangle|^2$ is depicted, the following parameters were used: $\gamma_q/(2\pi) = 9.44$ MHz, $\Omega/(2\pi) = 8.56$ MHz, $q = 1.389$. Here, γ_q is defined via $\gamma_q = 2\Delta\sqrt{\frac{2q-2}{2q-2}}$. Because the q-Gaussian with $q = 1.389$ is narrower than the Lorentzian distribution ($q = 2$), one needs fewer spins distributed within a smaller interval. Indeed, a width of $\epsilon = 0.005$ and 30 spins yield a very good result without any finite size effect.

4.5.2. Optimization

Now, one can optimize the pulse $\eta(t) = \sum_{k=1}^{n} a_k \sin(k\Omega_{\text{fund}}t)$ for a target functional $\mathcal{F}(\mathbf{a})$ to be specified later. We can do this in two ways: by using SOC

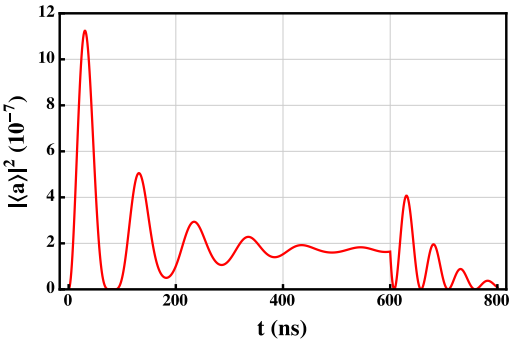

Figure 4.13.: Evolution of $|\langle a\rangle|^2$ for a q-Gaussian spectral density; approximation by 30 spins and a width of $\epsilon = 0.005$.

with a finite number of spins out of an interval of width ϵ or, since the Volterra equation (4.35) is linear in both $\eta(\tau)$ and $A(t)$, one can directly optimize the Volterra equation. If $A_k(t)$ is the solution of

$$A_k(t) = -\int_0^t \mathrm{d}\tau\, \sin(k\Omega_{\mathrm{fund}}\tau)e^{-(i(\omega_c-\omega_p)+\kappa)(t-\tau)} + \int_0^t \mathrm{d}\tau\, K(t-\tau)A_k(\tau), \quad (4.56)$$

then $A(t) = \sum_{k=1}^n a_k A_k(t)$ is the solution of

$$A(t) = -\int_0^t \mathrm{d}\tau \sum_{k=1}^n a_k \sin(k\Omega_{\mathrm{fund}}\tau)e^{-(i(\omega_c-\omega_p)+\kappa)(t-\tau)} + \int_0^t \mathrm{d}\tau\, K(t-\tau)A(\tau), \quad (4.57)$$

One can now optimize a target functional $\mathcal{F}(\mathbf{a})$ in the same way as above.

In the following, we want to manipulate the emission of light from the cavity. Since the operator a annihilates an excitation in the cavity this means that we have to use a target functional which contains the averaged cavity operator $|\langle a\rangle|^2$. In order to control the emission of light, one can imagine two elementary processes: the suppression and the enhancement of the emission of light. The suppression from a moment t_1 to a moment t_2 in time is achieved by minimizing the integral

$$\mathcal{F}_{\mathrm{sup}} = \int_{t_1}^{t_2} \mathrm{d}t |\langle a(t)\rangle|^2, \quad (4.58)$$

whereas the enhancement at a moment T in time can be obtained by using the target functional

$$\mathcal{F}_{\mathrm{enh}} = |\langle a(T)\rangle|^2. \quad (4.59)$$

We first use the method with the discretization of Sec. 4.5.1. The result of the minimization of Eq. (4.58) can be found in Fig. 4.14 (the same parameters

as in Fig. 4.13 are used): With the help of SOC - $n = 10$ frequency components were used - the integral is minimized between $t_1 = 100$ and $t_2 = 280$ ns. As one can see, this results in a suppression of light during approximately 200 ns. In order to understand that this time is long and has not been occurred 'by accident', one compares it with the characteristic oscillation time without control. This Rabi period can be estimated by $T_R \approx 100$ ns, *i.e.* the suppression of light has indeed been enforced during a time which is the double of T_R.

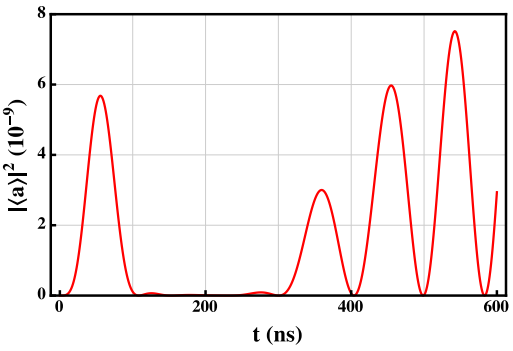

Figure 4.14.: Minimization of the integral $\int_{t_1}^{t_2} \mathrm{d}t\, |\langle a \rangle|^2$ with $t_1 = 100$ and $t_2 = 280$ ns. Emission of light is supressed during $t_1 < t < t_2$.

We will now try the second method, which consists in directly optimizing the Volterra equation. In order to describe an experimentally relevant situation, the cavity/spin losses $\gamma = 50$ Hz and $\kappa = 0.4$ MHz are assumed [72]. We used $n = 10$ frequency components and fixed the total amplitude to $\sum_{k=1}^{n} a_k^2 = P_0$. In this way, one amplitude $a_m = \pm\sqrt{P_0 - \sum_{k \neq m} a_k^2}$ can be eliminated. Here, we used $P_0 = 100$ MHz. As a first test, we minimized the integral of $|\langle a \rangle|^2$ from zero to a moment T in time. As one can see in Fig. 4.15, which displays the time evolution of $|\langle a(t) \rangle|^2$, there is no problem of supressing the emission of light during 50, 100 or 170 ns. Here, one has to mention that the total duration of control is only 200 ns, *i.e.* the time of suppression of light is indeed long.

In a second example, one maximizes the amplitude $|\langle a(t) \rangle|^2$ at a moment $t = T$ in time. As one can see in Fig. 4.16, where the time evolution of $|\langle a(t) \rangle|^2$ is depicted, a peak can easily be generated at the moments $T = 50$, 100 and 170 ns. These peaks are also the highest peaks in the time evolution of $|\langle a(t) \rangle|^2$.

It is important to mention that the dynamics can only be influenced as long as the control pulse is applied. After control is turned off, the dynamics entirely follow the internal time evolution induced by the Hamiltonian Eq. (4.31) with $\eta \equiv 0$. It is therefore impossible to maximize $|\langle a \rangle|$ or to minimize it over a finite time interval *outside* the control window.

If one compares the both methods, optimization by discretized SOC and direct optimization of the Volterra equation, one sees that the direct optimization is substantially faster than SOC because it circumvents the time-consuming operations in Floquet space. On the other hand, the variation of parameters is possible without additional overhead in the framework of SOC, whereas the direct optimization of the Volterra equation makes it necessary to recalculate the response functions A_k Eq. (4.56). Another advantage of the direct optimization of the Volterra equation is that one can consider spin and cavity losses, which is not possible in the Hamiltonian approach of SOC. Furthermore, if the width of the spectral density is broad, one needs a large number of spins to model it. In this case, SOC might be very slow because its computational time depends on the number of involved spins.

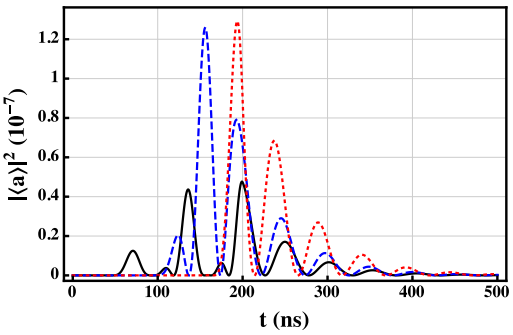

Figure 4.15.: Minimization of the integral $\int_0^T \mathrm{d}t\, |\langle a \rangle|^2$ by using a direct optimization of the Volterra equation. The time of suppression of light emission can be prolonged to $T = 50$ (black, continuous line), 100 (blue, dashed line) and 170 ns (red, dotted line). The duration of the pulses is 200 ns.

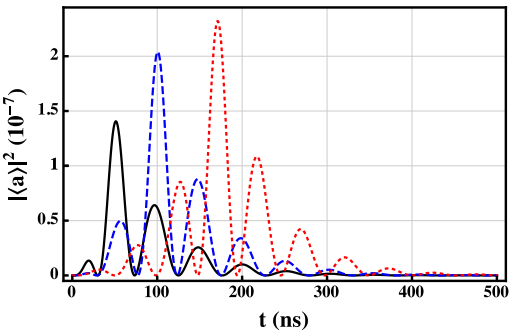

Figure 4.16.: Maximization of $|\langle a(t)\rangle|^2$ at the moment $t = T$ in time. A peak is enforced at the following moments in time: $T = 50$ (black, continuous line), 100 (blue, dashed line) and 170 ns (red, dotted line). The duration of the pulses is 200 ns.

Conclusions and Outlook

In the present thesis, an optimal control framework has been developed which is based on control pulses being expressed in terms of Fourier modes. Variational calculus with Floquet theory made it possible to calculate many quantities analytically. In particular, the dependence on the variable 'time' permits us to create time-optimal pulses and to extend control targets from single points in time to finite time windows. Furthermore, it could be shown that entanglement can be distributed within a spin chain by using only mediated interactions. This was even possible when these interactions are not known exactly, making out of the robustness against parameter fluctuations an important property of our approach.

Concerning the control of single spin dynamics, high accuracy AC magnetic field sensing has been demonstrated with ensembles of nitrogen-vacancy centers. In particular, the supremacy over state-of-the-art techniques with rectangular pulses has been shown. Since the quality of the sensing scheme is limited by the decoherence time T_2, as induced by the nuclear spin bath, one goal remains to mitigate the effect of decoherence by using an appropriate target functional. More precisely, the decoupling of a system spin from several environmental spins, as presented in Sec. 4.4, seems to be attractive in order to circumvent the detrimental effects of the environment.

In Sec. 4.5, we generalized the coupling of an ensemble of non-interacting spins to a classical antenna to an ensemble which interacts via the common coupling to a quantum-mechanical cavity. There, it could be seen that a continuity of spins can be accurately modelled by using just a few dozens of spins. If the spin-cavity coupling makes it necessary to consider a big number of spins, *e.g.* for q-Gaussian spectral densities with a high value of q, we identified the direct optimization of the Volterra equation as an attractive alternative. More precisely, it has been shown that the dynamics can be manipulated in a way that the emission of light from the cavity is suppressed during a finite

time interval or enhanced at a moment T in time. By exploiting these two elementary processes, a long term goal would be to write an excitation from the spin ensemble into the cavity and read it out again. In this sense, it would be of a great benifit to combine cavity and spins with a superconducting qubit as described in Sec. 1.2.2.1 and therefore to pave the way for a quantum computer. Moreover, as entanglement is the central element of a quantum computer, one may use the methods developed in Chapter 4, as they constitute the main capacities of quantum computers: long-lasting, time-optimal entanglement which can be created between two spins, even if they are only coupled by intermediate spins.

Last but not least, we would like to emphasize the inherent simplicity of our approach. The relatively low number of optimization parameters, when compared to other optimal control methods, gives us great hope to learn more about the functionality of our pulses.

Appendix

A. Floquet Theory with MATLAB

In order to calculate the Floquet spectrum, we diagonalize with the program MATLAB a truncated version $\tilde{\mathcal{K}}_{\text{trunc}}$ of the Floquet matrix $\tilde{\mathcal{K}}$. Typically, for $n \approx 10$ Fourier components and moderate amplitudes $|a| \approx 1/t_f$ (where t_f is the duration of the pulse) one does not need more than 100 Brillouin zones. In order to speed up calculations, it is important to dynamically adapt the size of the truncated Floquet matrix. For this purpose, we need a quantity that measures the error due to the truncation. As this measure, we chose the non-unitarity $\mathscr{E} = |\text{Tr}(\mathcal{U}\mathcal{U}^\dagger - \mathbb{1})|$ of the time evolution operator $\mathcal{U}$. As soon as $\mathscr{E} > \epsilon$ (here we chose $\epsilon = 10^{-10}$), the Floquet matrix will be enlarged by one Brillouin zone. By this procedure we calculate the required number of Brillouin zones. In Fig. A.1 we computed the size of the Floquet matrix using the Hamiltonian Eq. (1.84). As one can see in Fig. A.1 a), the number of required Brillouin zones increases approximately linearly with the number of frequency components. Here, for each number of frequency component one takes random amplitudes $a_k \in [-1, 1]$. However, one does not want a number of required Brillouin zones that results from a higher amplitude. That is why one normalizes each control amplitudes $\mathbf{a}$ by the largest amplitude $a_{\max} = \max_k |a_k|$: $\mathbf{a} \to \mathbf{a}/a_{\max}$. Whereas the number of required Brillouin zones depends linearly on the number of frequency components, one has to keep in mind that the computational effort scales quadratically with the number of Brillouin zones (as one deals with *matrices*). On the other hand, the number of required Brillouin zones only scales moderately with the amplitude $|\mathbf{a}|^2$ of the control pulse, as one can extract from Fig. A.1 b): for $|\mathbf{a}|^2 = 1000$ one needs about twice as many Brillouin zones as for $|\mathbf{a}|^2 = 1$. Here, one started with a random pulse with $n = 10$ frequency components and multiplied it by a scaling factor in order to obtain different values of $|\mathbf{a}|^2$.

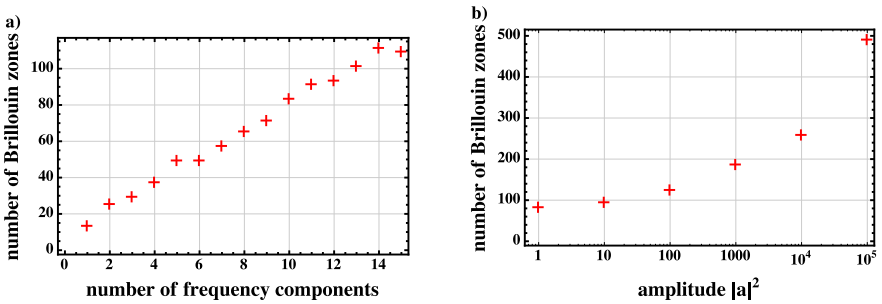

Figure A.1.: a) Number of required Brillouin zones as a function of the number of frequency components. b) Number of Brillouin zones for $n = 10$ frequency components as a function of the consumed power $|\mathbf{a}|^2$.

Another important aspect when dealing with numerical Floquet theory concerns the question which eigenvalues of the truncated matrix one has to take. The answer is that one has to take always the eigenvalues from the central Brillouin zone, *i.e.* containing the eigenvalues nearest to zero. Eigenvalues at the end of the spectrum, on the contrary, suffer from finite-size effects: typically, the difference between eigenvalues from outer Brillouin zones will deviate from $m\Omega$ ($m \in \mathbb{Z}$, see Sec. 1.1.2).

B. Derivation of the Volterra equation

Here, we want to derive an equation of motion for the cavity operator a, which will be of the Volterra type with a certain memory kernel. If one sets up the Heisenberg equations of motion, one obtains

$$\dot{a}(t) = i[H, a(t)] = -i(\omega_c - \omega_p)a(t) + \sum_{k=1}^{N} g_k \sigma_k^- - \eta(t) \tag{B.1}$$

$$\dot{\sigma_-^k}(t) = i[H, \sigma_-^k(t)] = -i\delta\omega_k \sigma_k^-(t) + g_k^* \sigma_k^z(t) a \tag{B.2}$$

Now, the number of spins N is very large, *i.e.* the excitation of a single spin is very low. Therefore, one can use the Holstein-Primakoff approximation $\langle \sigma_k^z \rangle \approx -1$ [73]. If one considers cavity losses κ and losses γ of the spin ensemble and takes the average $\langle \cdot \rangle$, one obtains

$$\dot{A}(t) = -[\kappa + i(\omega_c - \omega_p)]A(t) + \sum_{k=1}^{N} g_k B_k(t) - \eta(t) \tag{B.3}$$

$$\dot{B}_k(t) = -(\gamma + i\delta\omega_k)B_k(t) - g_k^* A(t), \tag{B.4}$$

where $A(t) = \langle a(t) \rangle$ and $B_k(t) = \langle \sigma_k^- \rangle$. For simplicity, one assumes g_k to be real. Making the transformation $B_k(t) = \tilde{B}_k(t) e^{-\Delta_k t}$, $\Delta_k = i\delta\omega_k + \gamma$ yields

$$\dot{\tilde{B}}_k(t) = -g_k A(t) e^{\Delta_k t}. \tag{B.5}$$

Integrating this equation gives

$$\tilde{B}_k(t) = \tilde{B}_k(0) - g_k \int_0^t \mathrm{d}\tau A(\tau) e^{\Delta_k \tau} \tag{B.6}$$

Inserting into (B.3) yields ($\Delta_c = \kappa + i(\omega_c - \omega_p)$)

$$\dot{A}(t) = -\Delta_c A(t) - \eta(t) + \sum_{k=1}^{N} g_k \tilde{B}_k(0) e^{-\Delta_k t} - \sum_{k=1}^{N} g_k^2 \int_0^t \mathrm{d}\tau e^{-\Delta_k (t-\tau)} A(\tau) \tag{B.7}$$

One can again eliminate the term $-\Delta_c A(t)$ by transforming $A(t) = \tilde{A}(t)e^{-\Delta_c t}$:

$$\dot{\tilde{A}}(t) = -\eta(t)e^{\Delta_k t} + \sum_{k=1}^{N} g_k B_k(0) e^{-(\Delta_k - \Delta_c)t} - \int_0^t \sum_{k=1}^{N} g_k^2 e^{-(\Delta_k - \Delta_c)(t-\tau)} \tilde{A}(\tau) \mathrm{d}\tau \tag{B.8}$$

By exploiting $\Delta_k - \Delta_c = i(\omega_k - \omega_c) + \gamma - \kappa$, setting $\tilde{A}(0) = \tilde{B}_k(0) = 0$ and introducing the spectral density

$$\rho(\omega) = \sum_k \frac{g_k^2 \delta(\omega - \omega_k)}{\Omega^2} \tag{B.9}$$

one arrives at

$$\tilde{A}(t) = -\int_0^t \mathrm{d}\tau\, \eta(\tau) e^{\Delta_c \tau} - \Omega^2 \int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}\tau \int_{-\infty}^{\infty} \mathrm{d}\omega \rho(\omega) e^{-(i(\omega-\omega_c)+\gamma-\kappa)(t'-\tau)} \tilde{A}(\tau) \tag{B.10}$$

Partial integration with respect to t' gives

$$\begin{aligned}
&\int_0^t \mathrm{d}t'\, e^{-(i(\omega-\omega_c)+\gamma-\kappa)t'} \int_0^{t'} \mathrm{d}\tau\, e^{(i(\omega-\omega_c)+\gamma-\kappa)\tau} \tilde{A}(\tau) \\
= &-\left[\frac{e^{-(i(\omega-\omega_c)+\gamma-\kappa)t'}}{i(\omega-\omega_c)+\gamma-\kappa} \int_0^{t'} \mathrm{d}\tau\, e^{(i(\omega-\omega_c)+\gamma-\kappa)\tau} \tilde{A}(\tau) \right]_{t'=0}^{t} \\
&+ \int_0^t \mathrm{d}\tau \frac{e^{-(i(\omega-\omega_c)+\gamma-\kappa)\tau}}{i(\omega-\omega_c)+\gamma-\kappa} e^{(i(\omega-\omega_c)+\gamma-\kappa)\tau} \tilde{A}(\tau) \\
= &\int_0^t \mathrm{d}\tau\, \frac{1 - e^{-(i(\omega-\omega_c)+\gamma-\kappa)(t-\tau)}}{i(\omega-\omega_c)+\gamma-\kappa} \tilde{A}(\tau)
\end{aligned} \tag{B.11}$$

By recasting everything into (B.10) one obtains the following Volterra equation

$$A(t) = F(t) + \int_0^t \mathrm{d}\tau\, K(t-\tau) A(\tau) \tag{B.12}$$

with the driving force

$$F(t) = -\int_0^t \mathrm{d}\tau\, \eta(\tau) e^{-(i(\omega_c - \omega_p)+\kappa)(t-\tau)} \tag{B.13}$$

and the memory kernel

$$K(t-\tau) = e^{-(i(\omega_c-\omega_p)+\kappa)(t-\tau)} \Omega^2 \int_{-\infty}^{\infty} \mathrm{d}\omega\, \rho(\omega) \frac{e^{-(i(\omega-\omega_c)+\gamma-\kappa)(t-\tau)} - 1}{i(\omega-\omega_c)+\gamma-\kappa}. \tag{B.14}$$

Bibliography

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. Theory of superconductivity. *Phys. Rev.*, 108:1175–1204, Dec 1957.

[2] A Einstein. Quantentheorie des einatomigen idealen Gases - Zweite Abhandlung. *Sitzungsberichte der Preußischen Akademie der Wissenschaften*, page 3, 1925.

[3] J. Sakurai. *Advanced Quantum Mechanics*. Addison Wesley, 1967.

[4] Erich Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52:997–1000, Mar 1984.

[5] R. Eric Stratmann, Gustavo E. Scuseria, and Michael J. Frisch. An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules. *The Journal of Chemical Physics*, 109(19):8218–8224, 1998.

[6] József Somlói, Vladimir A. Kazakov, and David J. Tannor. Controlled dissociation of I_2 via optical transitions between the x and b electronic states. *Chemical Physics*, 172(1):85 – 98, 1993.

[7] Wusheng Zhu, Jair Botina, and Herschel Rabitz. Rapidly convergent iteration methods for quantum optimal control of population. *The Journal of Chemical Physics*, 108(5):1953–1963, 1998.

[8] Anthony P. Peirce, Mohammed A. Dahleh, and Herschel Rabitz. Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications. *Phys. Rev. A*, 37:4950–4964, Jun 1988.

[9] Kyryl Kobzar, Burkhard Luy, Navin Khaneja, and Steffen J. Glaser. Pattern pulses: design of arbitrary excitation profiles as a function of pulse amplitude and offset. *Journal of Magnetic Resonance*, 173(2):229 – 235, 2005.

[10] Navin Khaneja, Timo Reiss, Cindie Kehlet, Thomas Schulte-Herbrüggen, and Steffen J. Glaser. Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms. *J. Magn. Reson.*, 172:296, 2005.

[11] N. B. Manson, J. P. Harrison, and M. J. Sellars. Nitrogen-vacancy center in diamond: Model of the electronic structure and associated dynamics. *Phys. Rev. B*, 74:104303, Sep 2006.

[12] H. Häffner, C.F. Roos, and R. Blatt. Quantum computing with trapped ions. *Phys. Rep.*, 469:155, 2008.

[13] Immanuel Bloch, Jean Dalibard, and Wilhelm Zwerger. Many-body physics with ultracold gases. *Rev. Mod. Phys.*, 80:885, 2008.

[14] Felix Platzer, Florian Mintert, and Andreas Buchleitner. Optimal dynamical control of many-body entanglement. *Phys. Rev. Lett.*, 105(2):020501, Jul 2010.

[15] Björn Bartels and Florian Mintert. Smooth optimal control with Floquet theory. *Phys. Rev. A*, 88:052315, Nov 2013.

[16] Y. Kubo, I. Diniz, A. Dewes, V. Jacques, A. Dréau, J.-F. Roch, A. Auffeves, D. Vion, D. Esteve, and P. Bertet. Storage and retrieval of a microwave field in a spin ensemble. *Phys. Rev. A*, 85:012333, Jan 2012.

[17] K. Sandner, H. Ritsch, R. Amsüss, Ch. Koller, T. Nöbauer, S. Putz, J. Schmiedmayer, and J. Majer. Strong magnetic coupling of an inhomogeneous nitrogen-vacancy ensemble to a cavity. *Phys. Rev. A*, 85:053806, May 2012.

[18] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. G. Schirmer, and T. Schulte-Herbrüggen. Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework. *Phys. Rev. A*, 84:022305, Aug 2011.

[19] V. V. Lednev. Possible mechanism for the influence of weak magnetic fields on biological systems. *Bioelectromagnetics*, 12(2):71–75, 1991.

[20] Richard L. Moore. Biological effects of magnetic fields: studies with microorganisms. *Canadian Journal of Microbiology*, 25(10):1145–1151, 1979.

[21] Joseph L. Kirschvink, Atsuko Kobayashi-Kirschvink, Juan C. Diaz-Ricci, and Steven J. Kirschvink. Magnetite in human tissues: A mechanism for the biological effects of weak ELF magnetic fields. *Bioelectromagnetics*, 13(S1):101–113, 1992.

[22] J. R. Maze, P. L. Stanwix, J. S. Hodges, S. Hong, J. M. Taylor, P. Cappellaro, L. Jiang, M. V. Gurudev Dutt, E. Togan, A. S. Zibrov, A. Yacoby, R. L. Walsworth, and M. D. Lukin. Nanoscale magnetic sensing with an individual electronic spin in diamond. *Nature*, 455:644–647, 2008.

[23] L. T. Hall, J. H. Cole, C. D. Hill, and L. C. L. Hollenberg. Sensing of fluctuating nanoscale magnetic fields using nitrogen-vacancy centers in diamond. *Phys. Rev. Lett.*, 103:220802, Nov 2009.

[24] L M Pham, D Le Sage, P L Stanwix, T K Yeung, D Glenn, A Trifonov, P Cappellaro, P R Hemmer, M D Lukin, H Park, A Yacoby, and R L Walsworth. Magnetic field imaging with nitrogen-vacancy ensembles. *New Journal of Physics*, 13(4):045021, 2011.

[25] Sungkun Hong, Michael S. Grinolds, Linh M. Pham, David Le Sage, Lan Luan, Ronald L. Walsworth, and Amir Yacoby. Nanoscale magnetometry with nv centers in diamond. *MRS Bulletin*, 38:155–161, 2 2013.

[26] Patrick Doria, Tommaso Calarco, and Simone Montangero. Optimal control technique for many-body quantum dynamics. *Phys. Rev. Lett.*, 106:190501, May 2011.

[27] Armin Rahmani and Claudio Chamon. Optimal control for unitary preparation of many-body states: Application to Luttinger liquids. *Phys. Rev. Lett.*, 107:016402, Jul 2011.

[28] P. de Fouquières, S. G. Schirmer, S. J. Glaser, and Ilya Kuprov. Second order gradient ascent pulse engineering. *J. Magn. Reson.*, 212(2):412 – 417, 2011.

[29] Georg Jäger and Ulrich Hohenester. Optimal quantum control of Bose-Einstein condensates in magnetic microtraps: Consideration of filter effects. *Phys. Rev. A*, 88:035601, Sep 2013.

[30] V. F. Krotov. *Global Methods in Optimal Control Theory.* Marcel Dekker, New York, 1996.

[31] Wusheng Zhu and Herschel Rabitz. A rapid monotonically convergent iteration algorithm for quantum optimal control over the expectation value of a positive definite operator. *The Journal of Chemical Physics*, 109(2):385–391, 1998.

[32] Oriol Romero-Isart and Juan José García-Ripoll. Quantum ratchets for quantum communication with optical superlattices. *Phys. Rev. A*, 76:052304, Nov 2007.

[33] R. Schmidt, A. Negretti, J. Ankerhold, T. Calarco, and J. T. Stockburger. Optimal control of open quantum systems: Cooperative effects of driving and dissipation. *Phys. Rev. Lett.*, 107:130404, Sep 2011.

[34] Daniel M. Reich, Mamadou Ndong, and Christiane P. Koch. Monotonically convergent optimization in quantum control using Krotov's method. *The Journal of Chemical Physics*, 136(10):–, 2012.

[35] Gaston Floquet. Sur les équations différentielles linéaires à coefficients périodiques. *Ann. École Norm. Sup.*, 12:47, 1883.

[36] Felix Bloch. Über die Quantenmechanik der Elektronen in Kristallgittern. *Zeitschrift für Physik*, 52(7-8):555–600, 1929.

[37] G. Goelman, S. Vega, and D. B. Zax. Design of broadband propagators in two-level systems. *Phys. Rev. A*, 39:5725–5743, Jun 1989.

[38] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2000.

[39] M. Howard, J. Twamley, C. Wittmann, T. Gaebel, F. Jelezko, and J. Wrachtrup. Quantum process tomography and Lindblad estimation of a solid-state qubit. *New Journal of Physics*, 8(3):33, 2006.

[40] T. Nöbauer, A. Angerer, B. Bartels, M. Trupke, S. Rotter, J. Schmiedmayer, F. Mintert, and J. Majer. Smooth optimal quantum control for robust solid state spin magnetometry. arXiv:1412.5051 (2014).

[41] Florian Mintert, André R. R. Carvalho, Marek Kuś, and Andreas Buchleitner. Measures and dynamics of entangled states. *Phys. Rep.*, 415:207, 2005.

[42] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245, 1998.

[43] Martin B. Plenio and Shashank Virmani. An Introduction to entanglement measures. *Quant.Inf.Comput.*, 7:1–51, 2007.

[44] B. Kraus and J. I. Cirac. Optimal creation of entanglement using a two-qubit gate. *Phys. Rev. A*, 63:062309, May 2001.

[45] Leandro Aolita, Andreas Buchleitner, and Florian Mintert. Scalable method to estimate experimentally the entanglement of multipartite systems. *Phys. Rev. A*, 78(2):022308, 2008.

[46] P. C. Maurer, G. Kucsko, C. Latta, L. Jiang, N. Y. Yao, S. D. Bennett, F. Pastawski, D. Hunger, N. Chisholm, M. Markham, D. J. Twitchen, J. I. Cirac, and M. D. Lukin. Room-temperature quantum bit memory exceeding one second. *Science*, 336(6086):1283–1286, 2012.

[47] M. W. Doherty, F. Dolde, H. Fedder, F. Jelezko, J. Wrachtrup, N. B. Manson, and L. C. L. Hollenberg. Theory of the ground-state spin of the NV^- center in diamond. *Phys. Rev. B*, 85:205203, May 2012.

[48] J. P. Goss, R. Jones, S. J. Breuer, P. R. Briddon, and S. Öberg. The twelve-line 1.682 eV luminescence center in diamond and the vacancy-silicon complex. *Phys. Rev. Lett.*, 77:3041–3044, Sep 1996.

[49] Jeronimo Maze Rios. *Quantum manipulation of nitrogen-vacancy centers in diamond: from basic properties to applications.* PhD thesis, Harvard University, 2010.

[50] M. Steiner, P. Neumann, J. Beck, F. Jelezko, and J. Wrachtrup. Universal enhancement of the optical readout fidelity of single electron spins at nitrogen-vacancy centers in diamond. *Phys. Rev. B*, 81:035205, Jan 2010.

[51] Y. Kubo, F. R. Ong, P. Bertet, D. Vion, V. Jacques, D. Zheng, A. Dréau, J.-F. Roch, A. Auffeves, F. Jelezko, J. Wrachtrup, M. F. Barthe, P. Bergonzo, and D. Esteve. Strong coupling of a spin ensemble to a superconducting resonator. *Phys. Rev. Lett.*, 105:140502, Sep 2010.

[52] R. H. Dicke. Coherence in spontaneous radiation processes. *Phys. Rev.*, 93:99–110, Jan 1954.

[53] Tobias Nöbauer. *Sensing, coherent coupling and optimal control with nitrogen-vacancy colour centres in diamond*. PhD thesis, Technical University of Vienna, 2013.

[54] Norman F. Ramsey. A molecular beam resonance method with separated oscillating fields. *Phys. Rev.*, 78:695–699, Jun 1950.

[55] F. Dolde, H. Fedder, M. W. Doherty, T. Nöbauer, F. Rempp, G. Balasubramanian, T. Wolf, F. Reinhard, L. C. L. Hollenberg, F. Jelezko, and J. Wrachtrup. Electric-field sensing using single diamond spins. *Nat. Phys.*, 7(6):459–463, 2011.

[56] P. L. Stanwix, L. M. Pham, J. R. Maze, D. Le Sage, T. K. Yeung, P. Cappellaro, P. R. Hemmer, A. Yacoby, M. D. Lukin, and R. L. Walsworth. Coherence of nitrogen-vacancy electronic spin ensembles in diamond. *Phys. Rev. B*, 82:201201, Nov 2010.

[57] G. Waldherr, Y. Wang, S. Zaiser, M. Jamali, T. Schulte-Herbrüggen, H. Abe, T. Ohshima, J. Isoya, J. F. Du, P. Neumann, and J. Wrachtrup. Quantum error correction in a solid-state hybrid spin register. *Nature*, 506(7487):204–207, 2014.

[58] Florian Dolde, Ville Bergholm, Ya Wang, Ingmar Jakobi, Boris Naydenov, Sébastien Pezzagna, Jan Meijer, Fedor Jelezko, Philipp Neumann, Thomas Schulte-Herbrüggen, Jacob Biamonte, and Jörg Wrachtrup. High-fidelity spin entanglement using optimal control. *Nat. Commun.*, 5:3371, 2014.

[59] Nan Zhao, Sai-Wah Ho, and Ren-Bao Liu. Decoherence and dynamical decoupling control of nitrogen vacancy center electron spins in nuclear spin baths. *Phys. Rev. B*, 85:115303, Mar 2012.

[60] Chi-Cheng Fu, Hsu-Yang Lee, Kowa Chen, Tsong-Shin Lim, Hsiao-Yun Wu, Po-Keng Lin, Pei-Kuen Wei, Pei-Hsi Tsao, Huan-Cheng Chang, and Wunshain Fann. Characterization and application of single fluorescent nanodiamonds as cellular biomarkers. *Proceedings of the National Academy of Sciences*, 104(3):727–732, 2007.

[61] Katharine W. Moore and Herschel Rabitz. Exploring constrained quantum control landscapes. *J. Chem. Phys.*, 137(13):134113, 2012.

[62] Michael H Goerz, Tommaso Calarco, and Christiane P Koch. The quantum speed limit of optimal controlled phasegates for trapped neutral atoms. *J. Phys. B: At. Mol. Opt. Phys.*, 44(15):154011, 2011.

[63] Robert Raussendorf and Hans-Jörg Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188, 2001.

[64] P. C. Maurer, J. R. Maze, P. L. Stanwix, L. Jiang, A. V. Gorshkov, A. A. Zibrov, B. Harke, J. S. Hodges, A. S. Zibrov, A. Yacoby, D. Twitchen, S. W. Hell, R. L. Walsworth, and M. D. Lukin. Far-field optical imaging and manipulation of individual spins with nanoscale resolution. *Nat. Phys.*, 6:912, 2010.

[65] N. Y. Yao, L. Jiang, A. V. Gorshkov, Z.-X. Gong, A. Zhai, L.-M. Duan, and M. D. Lukin. Robust quantum state transfer in random unpolarized spin chains. *Phys. Rev. Lett.*, 106:040505, Jan 2011.

[66] Yuting Ping, Brendon W. Lovett, Simon C. Benjamin, and Erik M. Gauger. Practicality of spin chain wiring in diamond quantum technologies. *Phys. Rev. Lett.*, 110:100503, Mar 2013.

[67] G. de Lange, Z. H. Wang, D. Ristè, V. V. Dobrovitski, and R. Hanson. Universal dynamical decoupling of a single solid-state spin from a spin bath. *Science*, 330(6000):60–63, 2010.

[68] G. Lindblad. On the generators of quantum dynamical semigroups. *Comm. Math. Phys.*, 48:119, 1976.

[69] N. V. Prokof'ev and P. C. E. Stamp. Theory of the spin bath. *Reports on Progress in Physics*, 63(4):669, 2000.

[70] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. *Phys. Rev. A*, 58:2733–2744, Oct 1998.

[71] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.*, 82:2417–2421, Mar 1999.

[72] S. Putz, D. O. Krimer, R. Amsüss, A Valookaran, T. Nöbauer, J. Schmiedmayer, S. Rotter, and J. Majer. Protecting a spin ensemble against decoherence in the strong-coupling regime of cavity qed. *Nat. Phys.*, 10(10):720–724, 2014.

[73] H. Primakoff and T. Holstein. Many-body interactions in atomic and nuclear systems. *Phys. Rev.*, 55:1218–1234, Jun 1939.