7 Errors and Decoherence

7.1 Motivation

While the above chapters were mostly concerned with the potentials offered by quantum information, the present chapter introduces some of the obstacles that have prevented large-scale implementations of quantum information so far. A simple-minded summary of this is that quantum information is generally more fragile than classical information. It is therefore essential to protect quantum information and to eliminate errors that occur during the computational process. An additional complication was that the no-cloning theorem initially appeared to prevent error correction and therefore make large-scale reliable quantum computing impossible. However the introduction of quantum error correction in 1995 showed that active techniques can be employed to circumvent this problem and the threshold theorem finally showed that reliable quantum information processing is indeed possible - at least in theory. The combination of difficulty and promise has made this area a very active field that has generated a number of exciting results over the two decades of its existence.

7.1.1 Quantum vs. classical errors

Any physical implementation of a computational process is designed to transform an input information into the desired output by applying appropriate operations as prescribed by the algorithm. These algorithms break the computation into suitable elements that can be handled by the available hardware. The goal of the hardware design is therefore to build a device that implements the mathematical operations as precisely and efficiently as possible. Unfortunately, any real physical device deviates to some degree from the idealized mathematical operation; this holds true for classical as well as for quantum computers. We will refer to these differences between targeted and actual evolution as errors.

While one tries to approximate the mathematically ideal operations with a suitably engineered device, it is not always possible to avoid such deviations between the mathematically predicted result and the physically executed computation. Some of the most important goals of computer architectures and implementations is therefore to avoid, recognize and correct errors occuring during the computation.

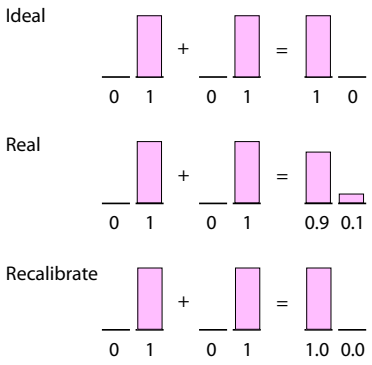

Figure 7.1: Calibration of digitally encoded data.

In classical computers, the most important design element for this purpose is the use of digital representation of information. As a result, every bit of information can be re-adjusted after every computational step to match the voltage corresponding to either the “0” or “1” state of the respective hardware.

This elementary error correction scheme can not be used in quantum computers, where the qubits can be in arbitrary superpositions of the relevant quantum mechanical states. As discussed in other parts of this

book, the input of a quantum computation is encoded in the exponentially many complex amplitudes of an initial state. They are continuous, rather than discrete. During the computation, they must be steered along a specific path in Hilbert space, whose dimension grows exponentially with the number of qubits. The final state contains the result of the computation, again in the continuous variables of the amplitudes and phases of the components. It is absolutely vital to maintain the phase coherence between the components of the state in order to perform a genuine quantum computation.

7.1.2 Sources of errors

We distinguish three effects that cause the results of a quantum computation to deviate from the ideal result:

- The gate operations are not perfect.
- The isolation between the quantum mechanical system (the quantum register) and the environment is not perfect. The spurious interactions with the environment cause unwanted transitions (=relaxation) and decay of the phase coherence (=dephasing or decoherence).
- The quantum system itself differs from the idealized model system considered in the design of the quantum computer. This includes, e.g. coupling constants that are slightly different from the ideal ones, and quantum states that are not included in the computational Hilbert space.

Section 7.2 summarizes the processes that lead to the loss of coherence in the system and therefore to the loss of quantum information.

7.1.3 Characterization of errors

Understanding the errors starts with differentiating between various categories. If only a single qubit is affected, the result is a single-qubit error. For obvious reasons, this is the most benign version of error, and the most thoroughly studied one. If the probability of single qubit errors is p_1, the probability that one error occurs in a system of N qubits is Np_1 and the probability that two errors occur is $N(N-1)p_1^2$. However, depending on the type of interactions between system and environments, it is also possible that two qubits are affected as a pair; this is the case of two-qubit errors. For schemes that try to avoid or correct errors, it is essential to consider the type of errors that the scheme is trying to address.

In addition to classifying the type of error, it is also important to measure the deviation of the state, to obtain quantitative information about the deviation between the targeted evolution and the actual evolution. Measures for the overlap of the actual state with the targeted state include the fidelity (→ ch. 7.3.1) and measures for the agreement between two evolutions, such as the process fidelity. Even more detailed information can be obtained by tomographic analysis of quantum states (→ ch. 7.3.3) and processes (→ ch. 7.3.4).

7.1.4 A counterstrategy

While one can (and should!) try to minimize these errors, it is important to realize that there are technical, financial as well as fundamental limits to the precision that can be achieved. It is, e.g., not possible to shield gravitational interactions between the system and the environment, or the quantum fluctuations in the apparatus that controls the gate operations and reads out the result.

To combat the detrimental effect of these imperfections on the results of computational processes, a number of options exist.

- Optimize the classical apparatus that controls the quantum system to make the gate operations as perfect as possible.

- Use robust gate operations, which are designed such that errors in experimental parameters tend to cancel rather than amplify. A typical example for this approach is the use of composite pulses in NMR [58].
- Use error correction schemes. (→ ch. 7.4)
- Store the information in areas of the Hilbert space that are least affected by the interaction between the system and its environment. (→ ch. 7.5)
- Use active schemes for decoupling the system from the environment, such as dynamical decoupling. (→ ch. 7.6)

It appears likely that any useful implementation of a quantum computer will require the implementation of all of these principles (and more) into its design. We discuss possible approaches to recognize and correct errors in quantum computers in Section 7.4. How information can be protected against decoherence will be discussed in Sections 7.5 and 7.6.

7.2 Errors and Decoherence

7.2.1 Phenomenology

Interference between two or more quantum states lies at the heart of the most striking quantum phenomena. As in classical wave optics, interference is possible only if the states keep a definite phase relationship, that is, if they are *coherent*. The destruction of coherence by uncontrollable interactions with environmental degrees of freedom is called *decoherence*. If decoherence occurs so fast that no interference phenomena can be observed, the resulting behavior can often be described in terms of classical physics.

If two states behave in the same way under the influence of the environment, they can stay coherent in spite of the coupling to the environment. If, on the other hand, they behave very differently, that is, if they can be easily distinguished from each other by the environment, they will lose coherence rapidly. This simple intuitive observation is important for quantum error correction and decoherence-free subspaces, to be discussed in later sections.

In this section we shall illustrate, by means of simple examples, how decoherence induced by interaction with the environment affects the state of a system, for example, a quantum information processing device. In the beginning the system is in a carefully prepared pure state; for a single qubit, this is

$$|\Psi(0)\rangle = a|0\rangle + b|1\rangle$$

or

$$|\Psi(0)\rangle = (a|0\rangle + b|1\rangle) \otimes |\Psi_e\rangle,$$

where $|\Psi_e\rangle$ summarises the complete state of the environment (which is always unknown). System and environment therefore form a product state.

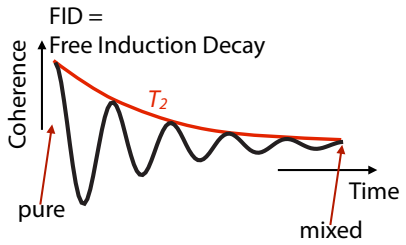

Figure 7.2: Decay of coherence by environmental perturbations.

The (complex) amplitudes of the initial state with respect to some basis in Hilbert space represent the quantum information to be processed. Classically, the uncontrollable interactions between system and environment cause the system evolution to deviate from the ideal evolution. Particularly fragile is the coherence a^*b, i.e. the relative phase of the two amplitudes.

If the environment is itself a quantum mechanical system, the interaction between system and environment builds up correlations between the system and environmental degrees of freedom (→ ch. 7.2.5). For the ideally prepared initial state, the environment also can be described as an (unknown) pure state $|\Psi_e\rangle$, which does not depend on the state of the system: system and environment factorize. The total quantum system, consisting of the quantum register and its environment, is then in a product state. Of course the preparation of the system's state requires interaction with other degrees of freedom; for the sake of simplicity we assume that those degrees of freedom can be separated sufficiently well from both system and environment once the preparation of the system's initial state is accomplished. Similarly, the gate operations require interactions with external degrees of freedom. We always treat these interactions as classical fields. This can be well motivated by noting that, e.g., the currents generating magnetic fields are generated by typically $> 10^{15}$ electrons, which are highly correlated, and therefore cannot possibly be described quantum mechanically.

The interaction between system and environment transforms this product state into a correlated state, which can be highly entangled. The state of the system alone (as represented by its density matrix) then in general is no longer pure, but mixed, as discussed in Chapter 4.

7.2.2 Semiclassical perturbation

The simplest description of the spurious interaction between system and environment uses a single spin-1/2 to describe the quantum register and a magnetic field for the environment. Since we discuss errors, we may restrict the analysis to the case when this field is weak compared to the static field that defines the energy of the basis states $|0\rangle$ and $|1\rangle$. In this limit, the most important effect of the error field is due to the component along the static field, which is usually chosen to be oriented along the z axis.

To illustrate its effect, we consider a system that is initially in a superposition state

$$|\Psi(0)\rangle = a|0\rangle + b|1\rangle.$$

If the two states $|0\rangle$ and $|1\rangle$ are eigenstates of the driving Hamiltonian $\mathscr{H}$ with eigenvalues E_0 and E_1, an ideal evolution transforms this state into

$$|\Psi(t)\rangle = a|0\rangle e^{-iE_0 t/\hbar} + b|1\rangle e^{-iE_1 t/\hbar}. \tag{7.1}$$

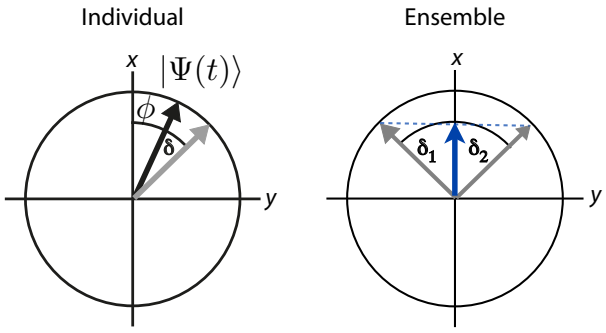

Figure 7.3: Coherent and incoherent contribution to the evolution.

Figure 7.3 (l.h.s.) represents this state as a spin-vector in the xy-plane. This corresponds to the case $|a| = |b| = \frac{1}{\sqrt{2}}$. Evolution corresponds to precession around the z-axis, and the resulting phase angle is $\phi = (E_1 - E_0)t/\hbar$.

Dephasing is due to additional (uncontrollable) interactions, which shift the energy of these eigenstates by a small (and in general time-dependent) amount δ_{E_i}. As a result, the average energy level shift difference changes the relative phase between the states by an angle

$$\delta(t) = \frac{1}{\hbar}\int_0^t (\delta_{E_1} - \delta_{E_0})dt'.$$

The state then becomes

$$|\psi(t)\rangle = a|0\rangle e^{-iE_0 t/\hbar} e^{i\delta/2} + b|1\rangle e^{-iE_1 t/\hbar} e^{-i\delta/2}.$$

In the example of Figure 7.3, this corresponds to a stochastic change of the orientation of the spin vector.

7.2.3 Ensemble average

Within the present picture of a single spin in a classical magnetic field, this additional phase increment arises from the fluctuating external field. The magnetic field has a well-defined value at all times, thereby causing a well-defined Larmor precession. However, the resulting precession angle differs between computational runs and it deviates from the mathematically correct representation by the random function $\delta(t)$. As shown in Figure 7.4, the resulting evolution of the spin corresponds to Brownian motion of the individual spin orientation.

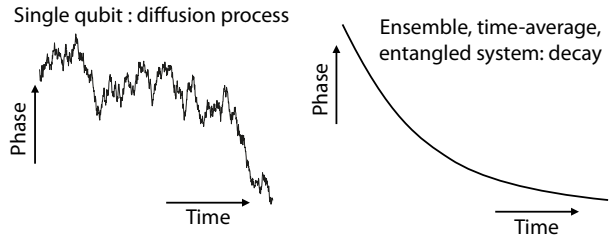

Figure 7.4: The left-hand part of the figure shows the evolution of a spin in a randomly varying magnetic field, which corresponds effectively to a diffusion process. The right-hand part shows how the average magnetization of an ensemble of spins decays when the individual spins evolve in random magnetic fields.

If we now consider an ensemble instead of a single quantum system, the random evolution of the individual members means that the average magnetization vector differs from that of the individual spins. Since the orientation of the individual spins (qubits) is progressively randomized as a function of time, the average magnetization vector becomes smaller, as shown in the right-hand part of Fig. 7.3. The decrease of the average magnetization can be obtained by taking the ensemble average

$$\langle S_x(t)\rangle = Tr\{e^{-i\delta(t)S_z}S_x e^{i\delta(t)S_z}S_x\} = \langle\cos\delta(t)\rangle,$$

where we have assumed that the system is initially oriented along the x-axis and we use an interaction representation where the evolution (7.1) due to the unperturbed Hamiltonian has been removed. This is generally known as the rotating frame. The $\langle\cos\delta(t)\rangle$ term can be evaluated by writing it as

$$\langle\cos\delta(t)\rangle = \frac{1}{2}\langle e^{i\delta} + e^{-i\delta}\rangle$$

and using the general property

$$\langle e^{iX}\rangle = e^{i\langle X\rangle - \langle X^2\rangle/2}$$

for a Gaussian random variable X to get

$$\langle\cos\delta(t)\rangle = e^{-\langle\delta^2(t)\rangle/2}. \tag{7.2}$$

For a random walk, $\langle\delta^2(t)\rangle$ is a linear function of time and the coherence decreases exponentially, as shown in the right-hand part of Figure 7.4. This simplified description becomes exact if the interaction that generates the random kicks does not have a memory (Markovian limit). It is then possible to describe their average effect by an exponential decay process. For the off-diagonal elements of a general density operator, one writes

$$\rho_{ij}(t) = \rho_{ij}(0)\, e^{-i(E_i-E_j)t/\hbar} e^{-t/T_2}.$$

Here, the first part describes the Hamiltonian evolution, in close analogy to eq. (7.1). The dephasing time T_2 is related to the RMS strength $\langle\delta^2(t)\rangle$ of the error field according to eq. (7.2). More detailed descriptions of these effects can be found in the NMR literature, where the effect is discussed as relaxation [59].

Different relaxation processes also cause the diagonal density operator elements to approach thermal equilibrium with a time constant T_1. These longitudinal relaxation processes also affect the quantum computation, causing a decay of the information. However, they are also needed, since they bring the system to the ground state, as required for initialization.

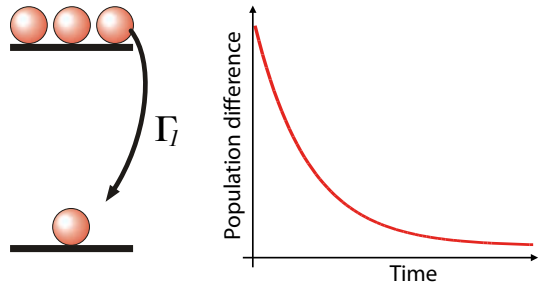

Figure 7.5: Relaxation of population difference.

The ensemble consideration is relevant not only for ensemble quantum computers, but also to quantum computers consisting of individual quantum systems. Even in these cases, a typical quantum computation will involve repeated runs of the computational process and the ensemble average corresponds then to the temporal average over the different runs.

7.2.4 Spin-boson model

We now move to models where the environment is quantum mechanical, rather than a classical field. In a quantum mechanical model, the phase-kicks are correlated to states of the external system, which is referred to as the bath. Typical examples for relevant degrees of freedom in the environment are phonons passing through the system or modes of the radiation field causing, e.g., spontaneous emission. For every state of this external system, the quantum register remains in a pure state, but the phase δ for this realization will be different from those for other states of the environment.

Since it is never possible to know exactly the state of the external system, one has to average over all accessible states of the external system. This averaging process changes the situation qualitatively: the vector representing the system is no longer only rotated by these additional phase kicks, it also becomes shorter. Technically, it is no longer in a pure state, but rather in a mixed state. In the simple picture given above, the vector no longer ends on the unit circle (or sphere), but remains inside it, in close analogy to the situation depicted in the right-hand part of figure 7.3.

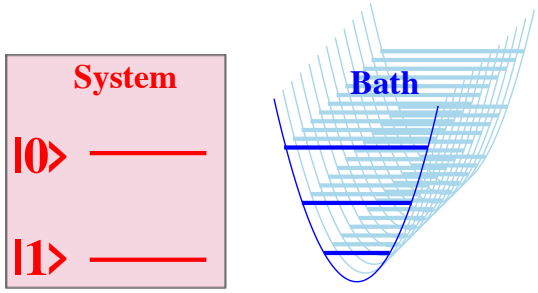

Figure 7.6: Spin-boson model: the system qubit is coupled to a bath of harmonic oscillators.

Such a situation can be represented, e.g., in the form of the so-called spin-boson model where the system is represented as a spin 1/2 (=qubit), and the environment as a system of bosonic modes, such as phonons or photons. In the general case, the interaction Hamiltonian between the two parts is

$$\mathscr{H}_I = \sum_k g_k \sigma_+ b_k + f_k \sigma_- b_k^\dagger + h_k \sigma_z b_k + h.c.$$

Here, g_k, f_k and h_k represent coupling constants, σ_α operators acting on the system qubit and b_k ladder operators acting on the k^{th} bath mode.

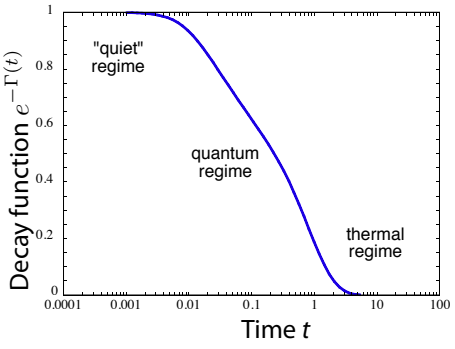

Figure 7.7: Decay of coherence in the spin-boson model. [60]

From this model, the system thermalizes [60] in

a process that can be split into a "quiet" regime, where the system remains almost coherent, a quantum regime where the interaction with the bath mode dephases the qubit and the thermal regime where it returns to equilibrium.

7.2.5 Spin-spin model

An even simpler quantum mechanical model is the spin-spin model, where the environment is reduced to a single spin 1/2. A more sophisticated versions of this is the central-spin model [61], where the environment is represented by several spins, which can be coupled amongst each other.

Here, we consider two interacting qubits: A (the system) and B (the environment). Each qubit is represented by a spin-$\frac{1}{2}$, and we assume that the two spins are coupled by an exchange interaction

$$\mathscr{H} = \frac{\omega}{\hbar}\vec{\mathbf{S}}_A \cdot \vec{\mathbf{S}}_B. \tag{7.3}$$

For $\omega > 0$ the ground state of this Hamiltonian is the singlet, with energy eigenvalue $-\frac{3}{4}\hbar\omega$, the triplet states have energy $+\frac{1}{4}\hbar\omega$ (see Appendix, ch. 14). The initial state is the most general product state (compare (4.26))

$$\begin{aligned} |\psi(0)\rangle &= \Big(a|\uparrow\rangle + b|\downarrow\rangle\Big)_A \otimes \Big(c|\uparrow\rangle + d|\downarrow\rangle\Big)_B \\ &= ac|\uparrow\uparrow\rangle + bc|\downarrow\uparrow\rangle + ad|\uparrow\downarrow\rangle + bd|\downarrow\downarrow\rangle. \end{aligned}$$

$|\psi(0)\rangle$ can be expressed in terms of the singlet and triplet states whose time evolution is simple.

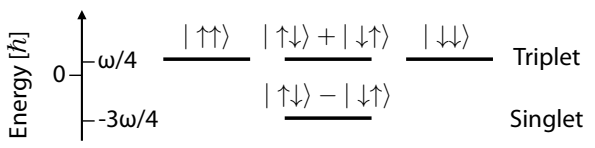

Figure 7.8: Eigenstates of the coupled 2-qubit system.

In particular, we can write the two states with antiparallel spin orientation as

$$\begin{aligned} |\uparrow\downarrow\rangle &= \frac{1}{\sqrt{2}}(|s\rangle + |t_0\rangle) \\ |\downarrow\uparrow\rangle &= \frac{1}{\sqrt{2}}(|s\rangle - |t_0\rangle), \end{aligned}$$

where

$$\begin{aligned} |t_0\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |s\rangle &= \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \end{aligned}$$

are the singlet and triplet states of this subspace. They are also eigenstates of the Hamiltonian and therefore evolve as

$$\begin{aligned} |t_0\rangle(t) &= |t_0\rangle(0)e^{-i\omega t/4} \\ |s\rangle(t) &= |s\rangle(0)e^{+3i\omega t/4}. \end{aligned}$$

The resulting time-dependent state $|\psi(t)\rangle$ is

$$\begin{aligned} & e^{i\omega t/4}|\psi(t)\rangle \\ = \; & ac|\uparrow\uparrow\rangle + bd|\downarrow\downarrow\rangle \\ & + \frac{1}{2}[ad(1+e^{i\omega t}) + bc(1-e^{i\omega t})]|\uparrow\downarrow\rangle \\ & + \frac{1}{2}[ad(1-e^{i\omega t}) + bc(1+e^{i\omega t})]|\downarrow\uparrow\rangle. \end{aligned} \tag{7.4}$$

This state is strictly periodic because the extremely simple model (7.3) contains only a single energy or frequency scale, ω. More complicated models of a system coupled to an environment of course will show more complex behavior. The general timescale on which decoherence phenomena happen, is inversely proportional to the coupling between system and environment (in our case, ω), as long as the different bath degrees of freedom interact independently with the system. If this is no longer the case, the system-bath interaction becomes effectively time-dependent [62]. This changes the effective strength as well as the characteristic behavior of the decoherence process.

7.2.6 Entanglement and mixing

The degree of entanglement between system A and environment B can be measured by the concurrence (4.30). A short calculation leads to the compact result

$$C = |ad - bc|^2 |\sin \omega t|. \tag{7.5}$$

The concurrence is a periodic function of time, as it should be for a periodically varying quantum state. The maximum value of C is determined by the initial state. If $|a| = |d| = 1$ or $|b| = |c| = 1$ the state can become maximally entangled; on the other hand, if $|a| = |c| = 1$ or $|b| = |d| = 1$ the state can never become entangled at all. In fact, in these two cases $|\psi(0)\rangle$ is a triplet state, $|\uparrow\uparrow\rangle$ or $|\downarrow\downarrow\rangle$, which is an eigenstate of $\mathscr{H}$ and thus remains unaffected by the coupling between the two qubits. All other cases where $C(t) \equiv 0$ are equivalent to this one, since $ad = bc$ only if A and B initially are in the same pure state, which can always be written as $|\uparrow\rangle$ in an appropriate spin-space coordinate system.

It appears tempting to exploit these states, which are not affected by the system-environment interaction, as long-lived quantum information. Unfortunately, this is not possible: preparing the system in such a state would require to prepare not only the system, but also the environment. Without such control, it is impossible to prepare system and environment into identical states. Not being able to control the corresponding degrees of freedom can be considered as a defining property of the environment. In addition, the environmental degrees of freedom are usually strongly coupled to additional degrees of freedom.

As an example for the time evolution of strongly entangled states, we consider the initial state

$$|\Psi(0)\rangle = |\uparrow\downarrow\rangle = \frac{1}{\sqrt{2}}(|t_0\rangle + |s\rangle),$$

corresponding to $a = d = 1$, $b = c = 0$. According to eq. (7.4) it evolves as

$$\begin{aligned} & e^{i\omega t/4}|\psi(t)\rangle \\ = \; & \frac{1}{2}[(1+e^{i\omega t})|\uparrow\downarrow\rangle + (1-e^{i\omega t})|\downarrow\uparrow\rangle]. \end{aligned}$$

At $t = \frac{\pi}{2\omega}$, this becomes

$$|\psi(\frac{\pi}{2\omega})\rangle = e^{-i\pi/8}[\frac{1+i}{2}|\uparrow\downarrow\rangle + \frac{1-i}{2}|\downarrow\uparrow\rangle]. \tag{7.6}$$

In this state, the system is maximally entangled with the environment. This can be seen, e.g., by calculating the system density operator by tracing over the environmental degree of freedom:

$$\begin{aligned} \rho_A(\frac{\pi}{2\omega}) &= Tr_B|\psi(\frac{\pi}{2\omega})\rangle\langle\psi(\frac{\pi}{2\omega})| \\ &= \frac{1}{2}(|\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow|) \\ &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned} \tag{7.7}$$

As usual, Tr_B denotes the trace over the Hilbert space of the environment B (see Chapter 4). Apparently this density operator is now diagonal. The spin has equal probabilities for being in the $\uparrow$ and $\downarrow$ states, but the phase information has been lost. The state is now a maximally mixed one, whereas the initial density operator $\rho(0)$ was pure.

7.2.7 Time dependence

The simple example discussed here is useful for demonstrating how the system becomes entangled with the environment and that this entanglement results in loss of coherence and therefore loss of information for the system.

On the other hand, this model is much simpler than any realistic situation. In particular, within this trivial model, the pure state could be recovered by simply letting the combined system-environment evolve for an identical period of time. This would recover

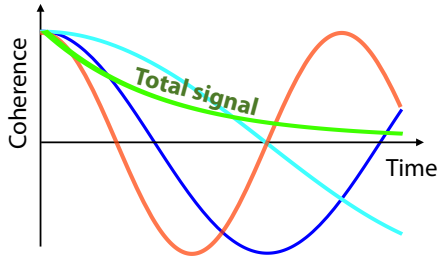

Figure 7.9: Oscillation of the coherence for different interaction strengths and total coherence for average over many couplings (green).

the full coherence and information content of the system qubit, apparently contradicting the second law of thermodynamics. This is corrected in more realistic models of the environment that have (infinitely) many degrees of freedom. The resulting evolution is then no longer periodic and it becomes impossible to recover the pure state from the mixed state.

This effect occurs also for other initial conditions, e.g., when the system is initially in a superposition state. As an example, we consider the case $a = b = c = -d = \frac{1}{\sqrt{2}}$, such that

$$|\psi(0)\rangle = \frac{1}{2}\left(|\uparrow\rangle + |\downarrow\rangle\right)_A \otimes \left(|\uparrow\rangle - |\downarrow\rangle\right)_B . \quad (7.8)$$

Note that the A part of this initial state is an eigenstate of $\mathbf{S}_x$ (4.13). A measurement of the x component of the system spin at $t = 0$ thus would clearly reveal the coherent nature of the state. At $t = \frac{\pi}{2\omega}$ this state evolves into the following maximally entangled state

$$\begin{aligned} & e^{i\pi/8}|\psi(\frac{\pi}{2\omega})\rangle \quad (7.9)\\ = \; & \frac{1}{2}\Big[|\uparrow\rangle_A \otimes \Big(|\uparrow\rangle - i|\downarrow\rangle\Big)_B \\ & -i|\downarrow\rangle_A \otimes \Big(|\uparrow\rangle + i|\downarrow\rangle\Big)_B\Big] \end{aligned}$$

Using the notation

$$\begin{aligned} |\rightarrow\rangle &= |\uparrow\rangle - i|\downarrow\rangle \\ |\leftarrow\rangle &= |\uparrow\rangle + i|\downarrow\rangle, \end{aligned}$$

the final state can be written as

$$\frac{1}{2}\left[|\uparrow\rangle_A \otimes |\rightarrow\rangle_B - i|\downarrow\rangle_A \otimes |\leftarrow\rangle_B\right].$$

The corresponding density matrix of A is again (7.7) and a measurement of $\mathbf{S}_x$ (of A) would yield zero. The initial information about the relative phase between $|\uparrow\rangle_A$ and $|\downarrow\rangle_A$ is lost.

The common feature of the two states $|\psi(\frac{\pi}{2\omega})\rangle$ (7.6) and (7.9) is the fact that the two basis states $|\uparrow\rangle_A$ and $|\downarrow\rangle_A$ of the system in both cases are strictly correlated to two mutually orthogonal states of the environment B. For (7.6) these are the eigenstates of $\mathbf{S}_z$ and for (7.9) the eigenstates of $\mathbf{S}_y$. This observation is an example of what was called "the fundamental theorem of decoherence" by Leggett [63]:

> If two mutually orthogonal states of the system of interest become correlated to two mutually orthogonal states of the environment, all effects of phase coherence between the two system states become lost.

In the situation just described, the final state of the system can be inferred from the final state of the environment; that is, the environment has "measured" the state of the system. This kind of reasoning can be applied to many instances of the quantum mechanical measurement problem, for example, the disappearance of the interference pattern in the standard two-slit experiment of quantum mechanics which occurs as soon as one measures through which of the two slits each single electron has passed.

7.2.8 Decoherence in large systems

The rate at which decoherence occurs in a given system is one of the most important parameters for as-

sessing the viability of a quantum computer implementation. However, it is important to realize that the rate at which quantum information is lost is not identical to the rate at which a single qubit undergoes decoherence. The difference is that during a typical computational process, information is spread over all qubits of the quantum register. It is therefore affected by decoherence processes acting on all qubits and decays correspondingly faster. The ultimate limit of this scaling process would be Schrödinger's cat [64]: in "classical" systems, the decoherence processes become so fast that it is no longer possible to observe quantum interference.

While it is generally assumed that the decay will be faster in larger quantum registers, there have been few experimental data to verify this prediction. While sufficiently large quantum computers are not available yet to test this, it is possible to measure the decoherence in model quantum registers consisting of correlated multi-qubit states. Nuclear spins in solids provide a particularly useful system for studying these processes, since correlated states of several thousand spins can be generated by suitable sequences of radio-frequency pulses.

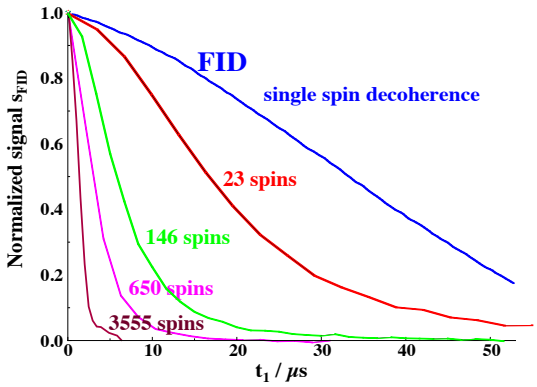

Figure 7.10: Decay of the coherence of quantum registers of different size.

Fig. 7.10 shows the decay of coherence in quantum registers of different sizes. Each model quantum register consists of a cluster of nuclear spins (^{1}H). Clearly, the larger quantum registers consisting of larger numbers of spins decay more rapidly, indicating that they are more fragile.

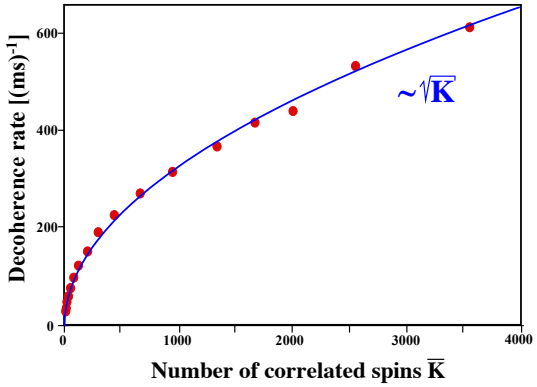

Figure 7.11: Decay rates of quantum registers of different size.

As shown in Fig. 7.11, the decoherence rate in this system grows approximately with the square root of the number of qubits [65, 66]. This is considerably less than what one would expect if each qubit interacted independently with the environment. If this behavior can be reproduced in other systems, the prospects for large-scale quantum computing may be brighter than one would expect from simple linear extrapolations.

How the decoherence rate increases with the number of qubits depends on the type of coupling to the environment that is responsible for the decoherence as well as on the encoding scheme used. In particular, large quantum systems contain many states that are relatively immune to environmental noise. These regions of Hilbert space are known as decoherence-free subspaces (→ ch. 7.5.4). While decoherence-free subspaces are a useful concept, we should not expect to find regions of Hilbert space that are completely immune to decoherence. Rather, these subspaces will be "sub-decoherent", i.e. the decoherence of states completely contained in them will be slower than for average quantum states.

In realistic systems, the external fields acting on the

different qubits are usually correlated to a finite degree. Depending on the degree of correlation, it should be possible to identify "clusters" of qubits for which the couplings are more strongly correlated than on average. The average rate at which information is lost from the quantum register can then be significantly reduced by suitable encoding schemes within such clusters of correlated spins [67].

7.3 Quality Measures

Quantum mechanical states encode information. Errors tend to degrade this information. The goal of error prevention is to reduce this degradation and the goal of error correction is to restore the information.

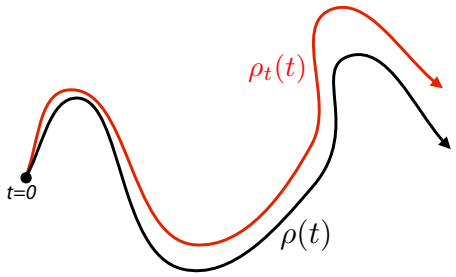

Figure 7.12: The distance between actual and targeted ideal evolution increases with time.

In order to assess the need for countermeasures, it is necessary to quantify deviations between the actual and the ideal information. Such distance measures correspond to the establishment of a metric and they can be used not only to follow the decay of information, but also to assess and optimize the success of the different countermeasures.

7.3.1 Distance and Fidelity

Measures of distance between different states exist also in classical information theory. A widely used measure is the Hamming distance between two bitstrings, which is defined by the number of bits that must be flipped to transform one into the other. As an example, the Hamming distance between the strings 00110 and 00101 is 2.

A distance metric for quantum states should specify how well a state $|\Psi_1\rangle$ agrees with the reference state $|\Psi_2\rangle$. In the case of pure states, it is possible to measure this by the scalar product $\langle\Psi_1|\Psi_2\rangle$. The scalar product has many of the properties that a useful distance measure should have: it is, e.g., independent of the coordinate system and therefore invariant under unitary transformations:

$$\langle U\Psi_1|U\Psi_2\rangle = \langle\Psi_1|\Psi_2\rangle.$$

It corresponds to an inverse distance in the sense that it is maximized if the two states are identical and it vanishes for orthogonal states.

Mixed states must be described by density operators, thus requiring different measures. One possible measure of the distance between two states (and thus of the error) is the trace-norm distance

$$D(\rho_1,\rho_2) = \frac{1}{2}\|\rho_1-\rho_2\|,$$

where

$$\|A\| = Tr\sqrt{A^\dagger A}.$$

Clearly, $D(\rho,\rho) = 0$ and for two pure orthogonal states ρ_1, ρ_2, the distance $D(\rho_1,\rho_2) = 1$ reaches the maximum possible value. If the two operators commute, the trace distance is equal to the sum over the differences between the eigenvalues.

Instead of measuring the distance, it is possible to measure how closely two states agree. The corresponding quantity is generally called the fidelity, and it can be considered as a generalization of overlap. Different definitions of the fidelity are used, including

$$F_1(\rho_1,\rho_2) = \|\sqrt{\rho_1}\sqrt{\rho_2}\|.$$

Other definitions are [68]

$$F_2(\rho_1,\rho_2) = Tr\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}$$

$$\begin{aligned} F_3(\rho_1,\rho_2) &= Tr\{\rho_1\rho_2\} \\ &\quad +\sqrt{1-Tr\{\rho_1^2\}}\sqrt{1-Tr\{\rho_2^2\}} \end{aligned}$$

$$F_4(\rho_1,\rho_2) = \frac{1-r}{2} + \frac{1+r}{2}F_3(\rho_1,\rho_2),$$

with $r = 1/(d-1)$ and d the dimension of the Hilbert space. In all these cases, the fidelity of a state with itself is unity, $F(\rho,\rho) = 1$, but the fidelity of general and orthogonal states may differ. Another fidelity measure that vanishes for orthogonal states is[68]

$$F_5(\rho_1,\rho_2) = \frac{|Tr\{\rho_1\rho_2\}|}{\sqrt{Tr\{\rho_1^2\}}\sqrt{Tr\{\rho_2^2\}}}.$$

This specific measure has the advantage that it does not require the evaluation of square roots of operators. Furthermore, it can be applied to deviation density operators (→ ch. 10.2.3), which is not possible for some other measures. It is quite similar to the expectation value of one density operator for the other state.

7.3.2 Process fidelity

In many cases, one is more interested in quantifying the agreement between two evolutions, instead of two states. This state-independent measure is called process fidelity. The evolutions may be described by two propagators U_1, U_2, where one might be a target operator, such as a quantum gate operation, and the other the actual propagator implemented in an experiment. The corresponding propagator fidelity can be defined in close analogy to F_5:

$$F(U_1,U_2) = \frac{|Tr\{U_1^\dagger U_2\}|}{\sqrt{Tr\{U_1^\dagger U_1\}}\sqrt{Tr\{U_2^\dagger U_2\}}}. \quad (7.10)$$

Again, this fidelity measure satisfies $F(U,U) = 1$.

As an example, we consider the Hadamard gate

$$\mathbf{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and compare it to the pseudo-Hadamard gate

$$\mathbf{h} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

If we apply these gates to the initial state

$$|\Psi_0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

which is oriented along the z-axis, it is rotated to the x-axis,

$$|\Psi_1\rangle = \mathbf{H}|\Psi_0\rangle = \mathbf{h}|\Psi_0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

for both operators. However, if we apply the operations again, now to $|\Psi_1\rangle$, we obtain

$$\begin{aligned} |\mathbf{H}|\Psi_1\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |\mathbf{h}|\Psi_1\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Clearly, the two operators do not represent the same process. This is verified if we calculate the process fidelity (7.10). Using

$$\mathbf{H}^\dagger\mathbf{h} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

we find the fidelity

$$F(U_1,U_2) = \frac{1}{2}\left|Tr\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\right| = 0.$$

7.3.3 Quantum state tomography

Determining the state of a quantum system is a very general problem. Clearly, this can not be done in a single measurement, since no procedure exists for distinguishing with certainty between non-orthogonal states in a single measurement. Repeated identical measurements can determine the amplitudes of the expansion coefficients in the chosen basis, but not the phases. The impossibility of doing a complete determination of a quantum state in a single measurement also follows directly from the no-cloning theorem: if it were possible, we could measure the state of the system and then prepare as many clones as required. It is therefore necessary to perform a series of measurements on identically prepared systems to extract the full quantum state. This is the scope of quantum state tomography.

If the system is not in a pure, but in a mixed state (to which it unavoidably evolves in the course of a computation process), a density operator is needed to fully describe the state. The density operator contains $(2^N)^2 = 2^{2N}$ elements. Determining all these numbers requires many measurements. The individual measurements must be designed to measure n^2 components of the density operator in an n^2-dimensional space of operators (the Liouville-space) acting on an n-dimensional Hilbert space.

For a single qubit, where $n = 2$ and $n^2 = 4$, four basis operators are needed. A convenient basis consists of the operators $\mathbf{1}, \mathbf{S}_x, \mathbf{S}_y, \mathbf{S}_z$. Measuring these observables allows us to reconstruct the density operator as

$$\begin{aligned}\rho &= \frac{1}{2}\left(Tr\{\rho\}\mathbf{1} + Tr\{\rho\mathbf{S}_x\}\mathbf{S}_x \right.\\ &\quad \left. + Tr\{\rho\mathbf{S}_y\}\mathbf{S}_y + Tr\{\rho\mathbf{S}_z\}\mathbf{S}_z\right).\end{aligned}$$

The individual results require in principle an average over an infinitely large sequence of measurements, but in practice, one will be satisfied with finite precision, which can be achieved by a finite number of measurements.

For a system of N qubits, the density operator can be expanded in an operator basis that consists of all possible (tensor) products of the operators

$$\mathbf{1}^i, \mathbf{S}_x^i, \mathbf{S}_y^i, \mathbf{S}_z^i,$$

where $i = 1..N$ runs over all qubits. This results in a total of $4^N = 2^{2N}$ operators that are orthogonal and form a complete basis for the expansion of the density operator.

To determine the expansion coefficients of the density operator in this basis, we have to measure 2^{2N} independent expectation values. This requires in principle the same number of independent observables. In general, nature does not provide a sufficient number of observables. As an example, products of spin operators like $\mathbf{S}_x^i\mathbf{S}_y^k$ are not directly observable. In the case of magnetic resonance, e.g., it is possible to measure the transverse components $\mathbf{S}_x^j$ and $\mathbf{S}_y^j$ of individual spins, which yields $2N \ll 2^{2N}$ observables. In the case of optical systems (e.g. ions, atoms, quantum dots), the same operators can be measured as electric dipole moments and populations can also be measured.

A significant extension if possible if one takes the time evolution and the couplings between the spins into account, which must be present in a quantum computer executing multi-qubit operations. For an Ising-type Hamiltonian, the evolution under a Hamiltonian term

$$\mathscr{H}^{ik} = d\mathbf{S}_z^i\mathbf{S}_z^k$$

drives an evolution

$$\mathbf{S}_x^i \leftrightarrow \mathbf{S}_y^i\mathbf{S}_z^k.$$

As a result, not only operators of the type

$$\mathbf{1}^1 \otimes \mathbf{1}^2 \otimes \cdots \mathbf{S}_x^j \otimes \cdots \mathbf{1}^N$$

are observable, but also

$$\mathbf{1}^1 \otimes \mathbf{1}^2 \otimes \cdots \mathbf{S}_y^j \otimes \cdots \otimes \mathbf{S}_z^k \otimes \cdots \otimes \mathbf{1}^N.$$

More precisely, the terms that are obtained in this way include all products that include exactly one transverse (x or y) term, while all other factors are either unity or $\mathbf{S}_z^k$ operators. This term represents so-called anti-phase transverse magnetization. It's expectation value $\langle S_y^j \rangle$ vanishes, i.e. it is also not directly observable. However, during free precession, the Hamiltonian term $S_z^j S_z^k$ turns it into observable S_x^j magnetization, which contributes to the observed FID signal. In the spectrum (i.e. after Fourier transformation), it is observed as two lines with opposite amplitudes. The number of such terms is $N2^N$ - more than the single spin terms, but still not enough for a full state tomography.

To measure the other components of the density operator, it is necessary to use unitary transformations that turn them into observable operators as listed above. This can be achieved by selective $\frac{\pi}{2}$ rotations applied to single qubits. Such a rotation of qubit k around the x-axis, e.g., turns the (unobservable) operator

$$\mathbf{1}^1 \otimes \mathbf{1}^2 \otimes \cdots \mathbf{S}_x^j \otimes \cdots \otimes \mathbf{S}_y^k \otimes \cdots \otimes \mathbf{1}^N.$$

into

$$\mathbf{1}^1 \otimes \mathbf{1}^2 \otimes \cdots \mathbf{S}_x^j \otimes \cdots \otimes \mathbf{S}_z^k \otimes \cdots \otimes \mathbf{1}^N,$$

which is observable, as discussed above. Since we have now four orthogonal operators ($\mathbf{1}$, $\mathbf{S}_x$, $\mathbf{S}_y$, $\mathbf{S}_z$) for every spin, we have a complete set of basis operators whose measurement allows the complete reconstruction of the density operator.

This procedure is called "quantum state tomography"[69, 70], in reference to X-ray tomography, where a sequence of two-dimensional pictures (or projections) is used to reconstruct the three-dimensional body being imaged. Figure 7.13 shows an example of such a tomographic analysis of the 2-qubit density operator that resulted from applying the Grover algorithm to a two-spin system [69]. The largest density operator element corresponds to the population of the $|11\rangle$ state.

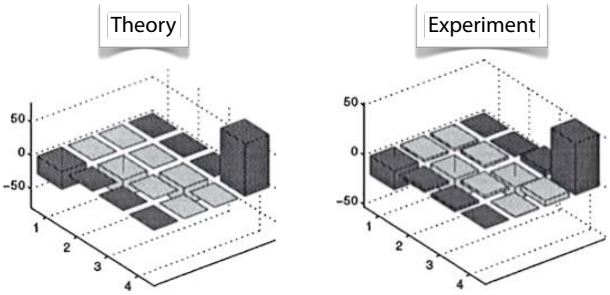

Figure 7.13: Theoretical and experimental density operator components during a Grover experiment [69].

7.3.4 Quantum process tomography

Quantum state tomography thus provides a complete characterization of a quantum state ρ. This is an essential prerequisite for the characterization of a quantum computer, but it is not sufficient. Since a general-purpose quantum computer must operate on arbitrary (possibly unknown) input states, its operation must be characterized with respect to every possible input. The algorithms require it to apply a (unitary) transformation to these input states. Accordingly, assessing its operation characteristics can only be done through an analysis of the corresponding (unitary) transformations. Furthermore, the effect of decoherence is to make the actual transformation non-unitary. The scope of quantum process tomography is a full determination of the actual (unitary or non-unitary) transformation for all possible input states.

The goal of quantum process tomography is the reconstruction of the evolution operator that transforms the unknown input state into the corresponding output state. In the case of unitary operations, this can be written in the form

$$\rho(T) = U(T)\,\rho(0)\,U^\dagger(T).$$

This can be achieved by preparing in subsequent experiments a complete set of basis states as input, ap-

plying the transformation and then performing density operator tomography on the final state $\rho(T)$.

In the context of quantum error correction, the relevant processes are often non-unitary. One possible approach to describe non-unitary evolutions uses a linear superposition of unitary processes:

$$\mathrm{E}(\rho) = \sum_i U_i \rho U_i^\dagger, \tag{7.11}$$

where $\mathrm{E}(\rho)$ is the general evolution operator acting on the state ρ. If we expand the unknown operators U_i in some convenient basis $\{E_m\}$, equation (7.11) can be rewritten as

$$\mathrm{E}(\rho) = \sum_{mn} E_m \rho E_n^\dagger \chi_{mn},$$

where the χ_{mn} are the expansion coefficients that need to be determined for a given process E. They form a Hermitian matrix and this representation is known as the 'chi matrix representation' [71]. The conventional choice of the operator basis $\{E_m\}$ is $\{\mathbf{1}, \mathbf{X}, -i\mathbf{Y}, \mathbf{Z}\}$ for each qubit. For a quantum register with N qubits, the total set of basis operators is then $\{\mathbf{1}, \mathbf{X}, -i\mathbf{Y}, \mathbf{Z}\}^{\otimes N}$, which consists of $4^N = 2^{2N}$ members and the χ-matrix has dimension $2^{2N} \times 2^{2N}$.

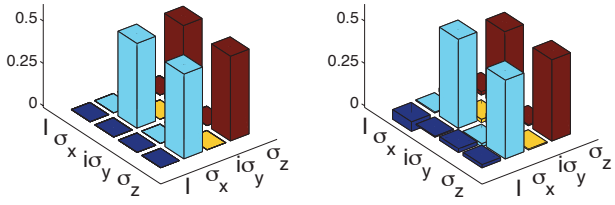

Figure 7.14: Theoretical and experimental process tomography for the Hadamard gate.

Figure 7.14 shows one example of such a chi-matrix, comparing the theoretical and experimental values.

One practical difficulty is that the measurement scheme itself is not error-free. If one tries to measure the process fidelity for a quantum gate that is close to an ideal one, the errors introduced by the measurement may be close to the errors of the gate operations. To obtain useful data, it is necessary to distinguish between the two types of errors. This can be done by measuring not the fidelity (or tomogram) of a single gate, but of a suitable combination of gate operations, provided their errors are independent of each other. Performing a series of such measurements that determine the fidelity of sequences of gate operations thus yields an average fidelity for the gate sequence. Provided the assumption of independent gate errors is valid, this allows one to calculate the fidelities of the individual gate operations.

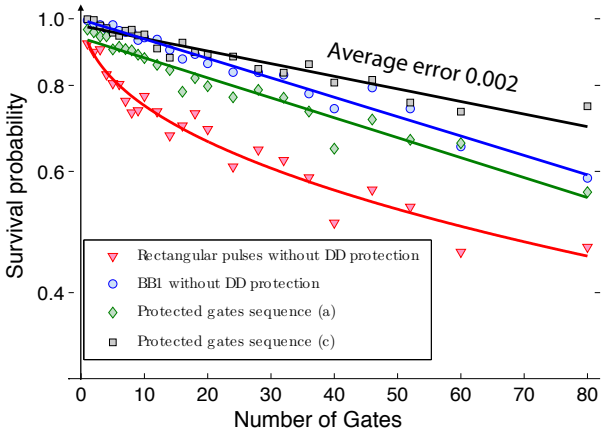

Figure 7.15: Fidelity decay during a randomized benchmarking experiment, comparing different types of gate operations [72].

Figure 7.15 compares the decay of fidelity during randomised benchmarking experiments for different types of gates. Comparison between these curves shows that the average fidelity for all of them is above 99%. In the uppermost curve, the fidelity per gate reaches 99.8 %.

Quantum process tomography is extremely useful for determining the actual dynamics of a quantum system. Applying it to large systems with many qubits, however, quickly becomes intractable, since the number of required operations grows exponentially with the size of the system. For N qubits, the Hilbert space has dimension 2^N and the number of density operator elements is 2^{2N}. For a unitary pro-

cess, the process can also be written as a $2^N \times 2^N$ matrix, but since decoherence results in deviations from unitarity, the total number of free parameters for a full characterization of the process grows as 2^{4N}. Determining these parameters requires a corresponding number of measurements:

N	1	2	3	4
2^{4N}	16	256	4096	65536

Clearly, such efforts quickly become impractical. It is then worth considering if the full information is really necessary. In many cases, it will be sufficient to measure not the full parameter set, but create 'summaries'. This is the purpose of reduced tomographic schemes, e.g. by symmetrizing the system [73]. The result consists, e.g., of the average probability of independent qubit flips.

7.4 Error correction

7.4.1 Basics

As errors are unavoidable in quantum as well as in classical computing, one must devise strategies for fighting them. *Error-correcting codes* do this by detecting erroneous qubits and correcting them. As in classical computation, redundancy is an indispensable ingredient here, and other than in classical computation, extreme care must be exerted not to garble the quantum information by the measurements involved in error detection.

Quantum information is not only potentially more "valuable" than classical information, but unfortunately also more vulnerable, because a qubit can be modified in more subtle ways than a classical bit, which can just be flipped from 0 to 1 or vice versa. Furthermore a classical bit can be protected against errors by basically copying it several times before transmission or processing and comparing the (different) results, an accidental simultaneous flip of many copies being extremely improbable. This is the basis of classical error correction.

No such procedure was in sight during the early years of quantum computing, and thus many scientists were very skeptical whether the attractive prospects of quantum computing could ever become a reality. Fortunately, methods for quantum error correction were soon discovered, based on coding schemes that permit detection of the presence and nature of an error (by converting it into a "syndrome" coded in ancillary qubits) *without* affecting the information stored in the encoded qubit. As we will discuss below these quantum error-correcting codes protect quantum information against large classes of errors. For simplicity we will restrict ourselves to errors that occur when information is transmitted through space (communication) or time (data storage) without being modified. The detection and correction of errors during the processing of data is the subject of *fault-tolerant computing* which we will only briefly mention at the end of the section.

The development of quantum error correction has culminated in the *threshold theorem* [74, 75, 76] stating that

> A quantum computation can be as long as required with any desired accuracy as long as the noise level is below a threshold value.

This result can be considered as a proof that reliable quantum computing is possible, at least in principle. How large this threshold is depends on the details of the error correction scheme and exact results are not available. Estimates for the admissible infidelity range from 10^{-2} to 10^{-4} for the different schemes. Achieving such levels of precision remains a challenging goal, but has been demonstrated in a few cases.

7.4.2 Classical error correction

To correct an error in a classical environment, one needs to detect it. The simplest way to do this is to generate copies of the input information to be protected from errors and to compare the outputs. More generally, the information must be encoded in some *redundant* way, which allows for reconstruction of the original data after partial destruction or loss. Of course, *completely* lost data cannot be recovered at all, but depending on the effort invested, the probability of complete loss can be made as small as desired.

The kind of error correction used and its probability of success depend on the kind of error expected. To keep things simple, suppose we want to transmit single classical bits 0 or 1, where each bit is transmitted successfully with probability $1-p$ and is flipped (once) with probability p, neglecting the possibility of multiple flips. In the simplest possible case, we encode the the *logical bit* 0_L in the *code word* 00 consisting of two *physical bits*, and likewise $1_L \mapsto 11$. If the receiver of the message detects that the two bits are identical, he may assume that the transmission is correct and accept it. If one of the two bits was flipped, the detected state is 01 or 10, which are outside of the set of legal codewords. The receiver will therefore detect that an error has occurred and may request re-transmission of the data. If the probability that one of the two bits flips is p, there is a probability of $2p(1-p) \approx 2p$ that a transmission error occurs and the transmission has to be repeated. In addition, there is a probability of p^2 that both bits have flipped. In this case, the error would go undetected.

If we do not want to only detect the presence of an error, but want (or must) also correct it, we can encode the logical bit in three physical bits. We choose for the logical state 0_L the code word 000 and $1_L \mapsto 111$. Thus 000 and 111 are the only two legal code words of the present coding scheme. If the error probabilities for the three bits are identical and independent of each other, the probability for error-free transmission of the logical bit is $(1-p)^3$, the probability that one of the three physical bits has flipped is is $3p(1-p)^2$, and so on. After transmission we check if all three bits of the code word are equal, and if they are not, we flip the one bit which does not conform to the other two. This leads to a wrong result if two or three bits were flipped during transmission, and the total probability for this to happen is $p^2(3-2p)$, which is much smaller than p for sufficiently small p.

Usually the bit-flip probability p grows with the distance (in space or time) of transmission, so that error correction must be repeated sufficiently frequently (but not too frequently, since copying and measuring operations may themselves introduce additional errors, which we have neglected here for simplicity). A larger number of physical bits per logical bit can be employed, increasing the probability of success, but also increasing the cost in terms of storage space or transmission time, as well as the complexity of the encoding and decoding schemes.

Of course in today's mature communication technology, far more sophisticated error correction schemes are in use than the one just presented, but they all rely on checking for damage and reconstructing the original information with the help of redundancy.

7.4.3 Quantum error correction

The classical error correction scheme discussed above is useless in the quantum regime, because it involves a measurement of every single bit transmitted. In the quantum case this entails a collapse of the qubit state to one of the measurement basis states, so that any information stored in the coefficients a and b of a superposition state $a|0\rangle + b|1\rangle$ is lost.

A quantum gate operation U, including the NOP operation is designed to drive the input state Ψ_{in} to the output state Ψ_{out} along a well-defined path in Hilbert space. In the case of a large quantum register, the dimension of Hilbert space is extremely large and de-

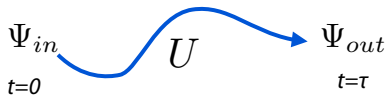

Figure 7.16: Evolution of the quantum state during a gate operation.

viations form this path can occur in all directions. Errors (from decoherence or from experimental imperfections) correspond to deviations from this path. One of the central ideas of quantum error correction is to detect the kind of error that has occurred (if any) *without* touching the information stored, and to subsequently reconstruct the original qubit state. Additional (or *ancillary*) qubits are needed in this process to store the kind of error (or *error syndrome*). Not every conceivable error is detectable or correctable; think of a multi-bit error converting one code word into a different legal code word in a classical redundant coding scheme. The more kinds of errors one wants to be able to correct, the more resources one needs. The code to be used must be chosen on the basis of a specific error model and the choice decides which errors can be detected and / or corrected.One of the specific problems related to the quantum nature of information was already addressed above: the fact that measurement may destroy the very information that was to be protected. This problem cannot be circumvented by just copying the information because of the no-cloning theorem (Section 4.2.11). Furthermore, in addition to the simple classical bit flip error, quantum mechanics allows for an entire continuum of possible errors, for example, continuous amplitude and phase changes. Fortunately the quantum error correction schemes developed during the past decade or so suffice to correct large classes of qubit errors.

One way to present the basic principle of quantum error correction is that the information is encoded in a Hilbert space whose dimension is larger than the minimum. Within this larger Hilbert space, it is then possible to choose two states as the basis states of the qubit in such a way that the interactions that cause the error do not transform one state directly into the other. Error detection then checks if the system contains contributions from other states and, if so, forces the system back into that part of Hilbert space that corresponds to the qubit.

7.4.4 Single spin-flip error

To begin with, let us discuss the transmission of qubits between a source A (Alice) and a receiver B (Bob). The transmission channel leaves each transmitted qubit either unchanged (with probability $1-p$) or flips it by applying an $\mathbf{X}$ operator (Section 4.2.1) (with probability p). The situation is completely analogous to the classical case discussed above. While quantum mechanics prevents Alice from copying quantum states for error protection, it provides her with entanglement as a new tool to achieve the same goal, as we will now see. In order to safely transmit the qubit state $a|0\rangle + b|1\rangle$ Alice initializes two further qubits in the state $|0\rangle$, so that the initial state of the three qubits is

$$|\psi_0\rangle = (a|0\rangle + b|1\rangle) \otimes |00\rangle = a|000\rangle + b|100\rangle.$$

These additional qubits are known as ancilla qubits or ancillas.

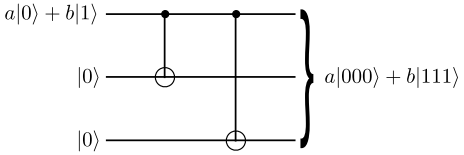

Figure 7.17: Circuit for encoding the input qubit in a logical qubit using three physical qubits.

Next she applies two CNOT gates, both with the first qubit as control and with the second and third qubits as targets, respectively. These two steps transform

the state to

$$\begin{aligned} |\psi_1\rangle &= \mathrm{CNOT}_{13}\mathrm{CNOT}_{12}|\psi_0\rangle \\ &= a|000\rangle + b|111\rangle. \end{aligned} \tag{7.12}$$

Alice thus encodes the information initially contained in the state of a single qubit in an entangled state of three qubits. This operation is *not* cloning: cloning (if it were possible) would lead to a product state of the three qubits with all of them in the same single-qubit state. Finally Alice sends the three qubits down the faulty channel, and relaxes.

Ideally, Bob receives the three-qubit state $|\psi_1\rangle$ without damage; this happens with probability $(1-p)^3$ since the three qubits have been transmitted independently. With probability $3p(1-p)^2$ one of the three qubits has been acted on by the "error operator" $\mathbf{X}$, and with probability $3p^2(1-p)$ one of the three possible pairs of two qubits have been flipped. Finally, with probability p^3 all three qubits have been flipped. Note that this is the only case where in spite of errors having occurred, Bob receives a combination of the legal "quantum code words" $|000\rangle$ and $|111\rangle$ and thus is unable to detect the error. In all other cases the entangled nature of Bob's state allows for error correction (which, however, is not always successful, as we will see). Note that the two components of Bob's state are always complements of each other; for example, if qubit 2 was flipped during transmission, Bob receives instead of $|\psi_1\rangle$ (7.12) the state

$$|\tilde{\psi}_1\rangle = a|010\rangle + b|101\rangle. \tag{7.13}$$

7.4.5 Error detection and correction

The goal of quantum error correction is to detect that an error has occurred and to correct it in such a way that the originally encoded quantum information is recovered. For this purpose, we need a measurement that detects the relevant errors and does not generate a measurement back-action that perturbs the correct states. For the present choice of encoding, suitable observables are the operators $\mathbf{Z}_1\mathbf{Z}_2$ and $\mathbf{Z}_1\mathbf{Z}_3$. Both legal code words, $|0\rangle_L = |000\rangle$ and $|1\rangle_L = |111\rangle$ are eigenstates of these operators with eigenvalue $+1$. Also, both components of Bob's state $|\tilde{\psi}_1\rangle$ are eigenstates of these operators with the same eigenvalue of ∓ 1. Since both components are eigenstates with the same eigenvalue, their linear combination is also an eigenstate with this eigenvalue,

$$\mathbf{Z}_1\mathbf{Z}_2|\tilde{\psi}_1\rangle = -|\tilde{\psi}_1\rangle = -a|010\rangle - b|101\rangle$$

and analogously for $\mathbf{Z}_1\mathbf{Z}_3$. Bob's state is thus always an eigenstate of $\mathbf{Z}_1\mathbf{Z}_2$ and $\mathbf{Z}_1\mathbf{Z}_3$, and the action of these two observables does not affect the state, apart from an unimportant global phase. By measuring $\mathbf{Z}_1\mathbf{Z}_2$ and $\mathbf{Z}_1\mathbf{Z}_3$ Bob can detect what kind of error has occurred (if any) and act accordingly. For the above example $\mathbf{Z}_1\mathbf{Z}_2 = -1$ and $\mathbf{Z}_1\mathbf{Z}_3 = 1$ from which Bob concludes that qubit 2 has been flipped. He applies $\mathbf{X}_2$ and thus restores the state $|\psi_1\rangle$, apart from a sign.

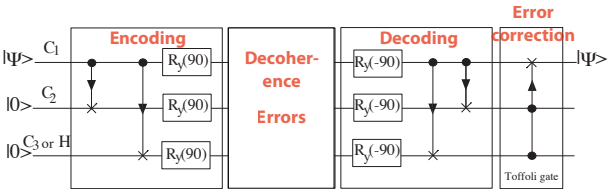

Figure 7.18: Error correction circuit for protection against a single-qubit-flip.

This procedure works for all cases where only one qubit was flipped, as one can verify easily. If two qubits are flipped, however, the error correction fails (as it does in the classical case). If, e.g., bits 1 and 3 are flipped, the transmitted state is $a|101\rangle + b|010\rangle$. Applying the two measurements yields the same values for $\mathbf{Z}_1\mathbf{Z}_2$ and $\mathbf{Z}_1\mathbf{Z}_3$ as the state $|\tilde{\psi}_1\rangle$ just discussed and is thus "corrected" to $a|111\rangle + b|000\rangle$.

There is a slightly different procedure for identifying the error which avoids any modification of Bob's state and which only employs CNOT gates. For that

procedure Bob needs two extra (ancilla) qubits prepared in the state $|00\rangle$. He then first carries out two CNOT operations with qubits 1 and 2 of the message as controls, respectively, and qubit 1 of the ancilla as target, and then two CNOTs with qubits 1 and 3 of the message as controls, respectively, and qubit 2 of the ancilla as target. The two ancilla qubits then contain the error syndrome: the first qubit is 0 if the first and second qubits of the message are equal, the second qubit of the ancilla compares the first and third qubits of the message. This procedure is an example for a more general strategy of storing the error syndrome in additional dimensions of the Hilbert space provided by ancillary qubits. This does not affect the information in the message, and the stored error syndrome can be used to correct the error, or to perform a *fault-tolerant* quantum computation which directly processes the encoded message and takes into account any errors which have been detected and stored as error syndromes. This eliminates (to some extent) the necessity to repeatedly decode and re-encode information, a procedure which is itself susceptible to errors.

7.4.6 Continuous errors

So far, we have assumed that a bit is either flipped or left invariant. A more realistic error model is a continuous error, which corresponds to a rotation around the corresponding axis. We start with the x-axis, which we have considered so far. A rotation around the x-axis by an angle θ corresponds to

$$R_x(\theta) = e^{-i\theta X/2} = \cos\frac{\theta}{2}\mathbf{1} - i\sin\frac{\theta}{2}\mathbf{X}.$$

As in our above example, we apply this rotation to the 2^{nd} qubit and obtain the state

$$\cos\frac{\theta}{2}\left(a|000\rangle + b|111\rangle\right) - i\sin\frac{\theta}{2}\left(a|010\rangle + b|101\rangle\right),$$

i.e. a superposition of the code word with no error and the code word with the error. Since these states have different eigenvalues for the syndrome extractor $\mathbf{Z}_1\mathbf{Z}_2$, performing a measurement with this operator will project it either onto the legal code word or on the one with the flipped bit. In the first case, we detect no error (and there is none), in the second case, we detect that the second bit has flipped and correct it. The code therefore does not only detect and let us correct discrete errors, but also works for continuous errors.

7.4.7 Decoding

The final step of the error correction protocol, is the decoding step: the logical qubit states are converted back to a single qubit. In our example, Bob has recovered the correct encoded state $|\psi_1\rangle a|000\rangle + b|111\rangle$. He can reconstruct Alice's original single-qubit state $a|0\rangle + b|1\rangle$ by repeating Alice's first two CNOT operations with qubit 1 as control and qubits 2 and 3 as targets, respectively:

$$\begin{aligned}\mathrm{CNOT}_{13}\mathrm{CNOT}_{12}|\psi_1\rangle &= a|000\rangle + b|100\rangle \\ &= (a|0\rangle + b|1\rangle)\otimes|00\rangle.\end{aligned}$$

The two ancilla qubits are no longer required and can be discarded. The result for the first qubit is $a|0\rangle + b|1\rangle$, as required.

The probability for this outcome is $1-3p^2-2p^3$, that is, in most cases, provided p is sufficiently small. In the case of an undetected double spin flip, the resulting state is $a|1\rangle + b|0\rangle$. The probability of failure is thus $\mathcal{O}(p^2)$, as compared to $\mathcal{O}(p)$ without error correction.

7.4.8 Phase errors

Next we consider another continuous type of error, which corresponds to a rotation around the z-axis. It turns out that this new error type can be corrected for by basically the same mechanism as for the rotation around the x-axis. The error is a random z axis

rotation given by

$$\begin{aligned}\mathbf{P}(\varepsilon) &= e^{i\varepsilon\phi\mathbf{Z}} = \begin{pmatrix} e^{i\varepsilon\phi} & 0 \\ 0 & e^{-i\varepsilon\phi} \end{pmatrix} \\ &= \cos(\varepsilon\phi)\mathbf{1} + i\sin(\varepsilon\phi)\mathbf{Z}.\end{aligned} \tag{7.14}$$

ϕ is a random angle between 0 and 2π, and ε is a "strength parameter" which controls the mean phase spread caused by $\mathbf{P}(\varepsilon)$ on average. The randomness in this operation is related to environmental degrees of freedom, for example, the random magnetic field discussed in Section 7.2.2. After the usual average over that randomness we have a combination of no error and a "phase flip" caused by the operator $\mathbf{Z}$:

$$\mathbf{Z}(a|0\rangle + b|1\rangle) = a|0\rangle - b|1\rangle. \tag{7.15}$$

Now, consider the action of $\mathbf{Z}$ in a different basis, given by the eigenstates $|+\rangle$ and $|-\rangle$ of $\mathbf{X}$:

$$|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}} \quad ; \quad \mathbf{X}|\pm\rangle = \pm|\pm\rangle : \tag{7.16}$$

obviously

$$\mathbf{Z}|\pm\rangle = |\mp\rangle, \tag{7.17}$$

that is, $\mathbf{Z}$ causes a bit flip in the basis given by the eigenstates of $\mathbf{X}$, and we have already seen how a bit flip can be corrected for. The basis change from $\mathbf{Z}$ eigenstates to $\mathbf{X}$ eigenstates and back is accomplished by a Hadamard gate $\mathbf{H}$ (4.16), formally

$$\mathbf{HZH} = \mathbf{X}. \tag{7.18}$$

In order to achieve error correction for a phase-flipping transmission channel, Alice prepares the state $|\psi_1\rangle$ (7.12) as before, and then applies $\mathbf{H}^{\otimes 3} = \mathbf{H}_1\mathbf{H}_2\mathbf{H}_3$ to $|\psi_1\rangle$:

$$\mathbf{H}^{\otimes 3}|\psi_1\rangle = a|{+}{+}{+}\rangle + b|{-}{-}{-}\rangle \tag{7.19}$$

before sending her 3-qubit message off. Bob can use almost the same procedure as before; however, he has to use $\mathbf{X}_1\mathbf{X}_2$ and $\mathbf{X}_1\mathbf{X}_3$ for error syndrome extraction and $\mathbf{Z}_1, \mathbf{Z}_2$, and $\mathbf{Z}_3$ for error correction, before applying $\mathbf{H}^{\otimes 3}$ to switch back to the computational basis.

7.4.9 Projection errors

Yet another kind of error that can happen to a single qubit is an "accidental measurement" resulting in a projection to $|0\rangle$ or $|1\rangle$. That kind of error can be related to a phase flip ($\mathbf{Z}$) by observing that the projectors to $|0\rangle$ and $|1\rangle$ (Section 4.2.1) can be written as

$$\begin{aligned}|0\rangle\langle 0| &= \mathbf{P}_\uparrow = \frac{1}{2}(\mathbf{1} + \mathbf{Z}); \\ |1\rangle\langle 1| &= \mathbf{P}_\downarrow = \frac{1}{2}(\mathbf{1} - \mathbf{Z}).\end{aligned} \tag{7.20}$$

Projectors onto more general Hilbert space vectors can be written as linear combinations of $\mathbf{1}, \mathbf{X}, \mathbf{Y}$, and $\mathbf{Z}$. This is clear from the fact that *any* 2×2 matrix can be written in terms of these operators.

→ **Problem 1**

The most general single-qubit error is given by a general unitary 2×2 matrix, combined with a projection to some axis, and can thus be written in terms of $\mathbf{1}, \mathbf{X}, \mathbf{Y}$, and $\mathbf{Z}$. We have seen that errors caused by $\mathbf{X}$ and $\mathbf{Z}$ can be corrected for by simple procedures, and given the fact that $\mathbf{ZX} = i\mathbf{Y}$, errors caused by $\mathbf{Y}$ should also be correctable.

7.4.10 General single qubit errors

The simple code that does the trick is a combination of the two procedures already discussed and was invented by Peter Shor [77]. Shor's code involves the idea of *concatenating* two redundant codes: the original logical qubit is redundantly encoded in three qubits in order to fight one kind of error, and then each of these three qubits is again encoded in three qubits to take care of the second type of error.

The encoding procedure consists of well-known steps. Alice first applies two CNOT gates with the original logical qubit as control and with the two additional qubits initialized to the state $|0\rangle$ as targets. Then she applies a Hadamard gate to each of

the three qubits. This maps the computational basis states as follows:

$$|0\rangle \to |+++\rangle \quad ; \quad |1\rangle \to |---\rangle. \tag{7.21}$$

Now Alice adds two fresh $|0\rangle$ qubits to each of the three code qubits in her possession, for a total of nine qubits, and again applies the two-CNOT encoding procedure to each of these qubit triplets. This yields one logical qubit encoded in entangled states of nine physical qubits:

$$\begin{aligned} |0\rangle &\to \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ &\quad \cdot(|000\rangle + |111\rangle) \\ |1\rangle &\to \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \\ &\quad \cdot(|000\rangle - |111\rangle). \end{aligned}$$

Assuming (as usual) that the encoding procedure is flawless, we discuss the correction of single-qubit errors. In order to detect a bit flip on the first qubit (or any qubit of the first triplet, in fact), Bob may again use the operators $\mathbf{Z}_1\mathbf{Z}_2$ and $\mathbf{Z}_1\mathbf{Z}_3$. Subsequent application of the appropriate $\mathbf{X}$ operator then corrects the error.

A phase flip on one of the first three qubits changes the sign within that block, that is, it changes $|000\rangle + |111\rangle$ to $|000\rangle - |111\rangle$ and vice versa. In order to detect such a sign change and its location Bob again only *compares* the signs of the three-qubit blocks one and two, and one and three. Since $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3$ is the operator for the simultaneous bit flip on qubits 1, 2, and 3, that is, it maps $|000\rangle \to |111\rangle$ and vice versa, the sign comparisons between blocks are performed by the somewhat clumsy operators $\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3\mathbf{X}_4\mathbf{X}_5\mathbf{X}_6$ and $\mathbf{X}_4\mathbf{X}_5\mathbf{X}_6\mathbf{X}_7\mathbf{X}_8\mathbf{X}_9$. A phase flip on any of the first three qubits can then be repaired by applying $\mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_3$.

If both a bit flip and a phase flip have occurred on, say, qubit 1, the two procedures outlined above will both detect and remove their respective "target errors", so that indeed all single-qubit errors caused by $\mathbf{X}, \mathbf{Z}$, or $\mathbf{ZX} = i\mathbf{Y}$ can be corrected. As argued above, this means that an entire continuum of arbitrary single qubit errors is kept at bay by really taking care of only a finite (and very small) set of errors. This remarkable fact is sometimes referred to as "discretization of errors", and it is instrumental to the whole concept of quantum error correction. Note that there is nothing similar for classical analog computing.

The Shor code is conceptually simple and easy to understand, but it needs nine physical qubits per logical qubit to provide protection against arbitrary single-qubit errors. There are codes providing the same degree of protection with 7 [78] and even 5 [79, 80] physical qubits per logical qubit.

7.4.11 Perfect 5-qubit code

A single qubit has four possible error conditions: no error, or a flip around the x-, y- or z-axis. We can represent them by the operators $\{\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$. Detecting the error condition therefore requires 2 bits of information. If we restrict ourselves to single-qubit errors, a system of 2 qubits has seven possible error conditions: no error or six independent single-qubit flips. In the general case of N qubits, the number of error conditions is thus $1 + 3N$.

Distinguishing between these error conditions requires

$$n = \log_2(1 + 3N)$$

bits of information, which can be obtained, e.g., through measurements on syndrome qubits.

N	1	2	3	4	5	6
$1+3N$	4	7	10	13	16	19
n	2	3	4	4	4	5

The above table shows, that $N = 5$ is the smallest number that allows to encode 1 bit with $N - 1$ ancilla bits and use them for differentiating between

all possible error conditions. $N = 5$ is therefore the smallest number of qubits required for a perfect error correction code, a code that can correct all possible single-qubit errors.

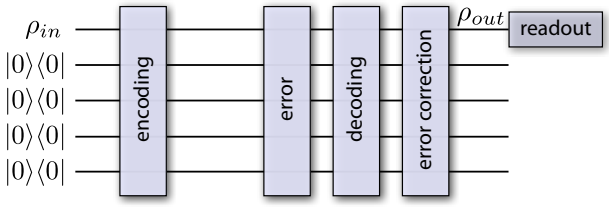

Figure 7.19: The simplest perfect error correction code uses 5 qubits.

A perfect 5-qubit code was proposed by Laflamme and coworkers [79]. Its logical basis states are

$$\begin{aligned} |0\rangle_L &= \frac{1}{\sqrt{8}}(|00000\rangle - |10111\rangle - |01011\rangle + |11100\rangle \\ &\quad + |10010\rangle + |00101\rangle + |11001\rangle + |01110\rangle) \\ |1\rangle_L &= \frac{1}{\sqrt{8}}(|11111\rangle - |01000\rangle + |10100\rangle - |00011\rangle \\ &\quad + |01101\rangle + |11010\rangle - |00110\rangle - |10001\rangle). \end{aligned} \tag{7.22}$$

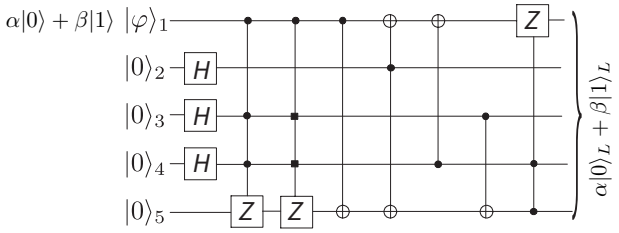

Figure 7.20: Encoding scheme for the 5-qubit QEC scheme. **H** are Hadamard gates and **Z** are π-phase gates.

Fig. 7.20 shows the gate operations required for encoding a single qubit in the five-qubit QEC scheme. It transforms the input state

$$\alpha|0\rangle + \beta|1\rangle \;\rightarrow\; \alpha|0\rangle_L + \beta|1\rangle_L.$$

The decoding operation is its inverse, $U_{de} = U_{en}^{\dagger}$.

7.4.12 Stabilizer codes

After the first error-correcting quantum codes were found, more general theoretical frameworks for the analysis and classification of codes were developed. One such framework is called *stabilizer* formalism, and the associated codes are stabilizer codes. We do not discuss the general formalism here, but concentrate on examples.

The approach is based on group theory, and the underlying group that we use here is the *Pauli group* for n qubits. In mathematics, a group G is defined as a set of elements that are combined with a binary operation $\cdot$, which is also called the group law. They must fulfill the following requirements

- Closure: For any pair of group elements $a, b \in G$, the result of the group operation must be in the group, $a \cdot b \in G$.
- Associativity: For all group elements $a, b, c \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
- Identity element: The group contains an element called identity and often written as $\mathbf{1}$, the the group operation yields $\mathbf{1} \cdot a = a \cdot \mathbf{1} = a$.
- Inverse element: For every group element a, there is an inverse element a^{-1}, such that $a \cdot a^{-1} = a^{-1} \cdot a = \mathbf{1}$.

For a single qubit the Pauli group consists of the unit matrix $\mathbf{1}$ and the three Pauli matrices $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$, all with prefactors $\pm 1, \pm i$. These matrices form a group under matrix multiplication: a product of two group elements is again a group element. For n qubits, direct products of matrices from the individual qubit Pauli groups form a group in a completely analogous way.

Suppose now that S is a subgroup of the n-qubit Pauli group and that a certain set V_S of n-qubit states is invariant under the action of all elements of S; then V_S is said to be the vector space *stabilized* by S, and S is called the *stabilizer*. The basis vectors of V_S can be used as code words for a *stabilizer code*. A simple example for $n = 3$ is provided by the set

$S = \{\mathbf{1}, \mathbf{Z}_1\mathbf{Z}_2, \mathbf{Z}_2\mathbf{Z}_3, \mathbf{Z}_1\mathbf{Z}_3\}$. Here, V_S is spanned by $|000\rangle$ and $|111\rangle$, which are both eigenstates of all four operators with eigenvalue $+1$.

The nontrivial elements of the stabilizer for this code work as error-syndrome extractors: they leave all states containing only legal code words intact and map all states affected by errors to other states. Different errors must be distinguishable by the syndrome extractors in order to be correctable. We have seen earlier that for the present simple three-qubit code only single-qubit flip errors can be corrected, while two-qubit flips lead to wrong transmission results and three-qubit flips are not detected at all.

The phase flip code discussed in section 7.4.8 has the stabilizer generators **XX1** and **1XX**, where we have omitted the indices. For the 9-qubit Shor code, the stabilizer set can be generated by the 8 operators **ZZ1111111**, **1ZZ111111**, **111ZZ1111**, **1111ZZ111**, **111111ZZ1**, **1111111ZZ**, **XXXXXX111**, and **111XXXXXX**.

For a code with n-qubit code words, one may classify errors by their weight, that is, by the number of nontrivial Pauli matrices applied to the code words. It is desirable to construct a code able to correct all errors up to a maximum weight w; such a code is called w-error-correcting. The achievable w depends on the similarity or distinguishability of the code words employed. If the minimum distance (as expressed by the number of differing qubits) between any two code words is d, then the maximum w is given by the integer part of $d/2$. Of course the minimum distance depends on the number k of logical (qu)bits encoded (as 2^k code words) in the n physical (qu)bits. In our example code for correcting single-qubit bit flips, the distance between the code words $|000\rangle$ and $|111\rangle$ was $d = 3$, which allowed us to correct $1 < d/2$ flipped bits.

Classical as well as quantum codes are often characterized by $[n,k,d]$. Our example of the simple bit-flip correcting code was a [3,1,3] code, which allowed to correct $w = 1$ bit flip errors. There is an elaborate theory of classical error-correcting codes, and in fact a class of quantum error-correcting codes may be derived from classical codes. These codes are called Calderbank–Shor–Steane (or CSS) codes [81, 82] after their inventors. They are a subclass of the stabilizer codes, as discussed in Chapter 10 of [33]. The codes with $n = 5$ [79, 80] and $n = 7$ [78] mentioned above both have $k = 1$ (that is, two code words, or one logical bit) and $d = 3$. It can be shown (see Chap 12 of [33]) that $n = 5$ is the minimum size for a 1-error-correcting quantum code. Nevertheless, the five-qubit code is of limited practical use because it involves complicated encoding and decoding procedures, and because fault-tolerant quantum logical operations are difficult to implement in this code.

7.4.13 Fault-tolerant computing

We have only discussed simple transmission (in space or time) of quantum information, without considering any logical operations (except those needed for quantum error correction). For quantum computing to become practical, it is necessary to perform logical operations in a fault-tolerant way.

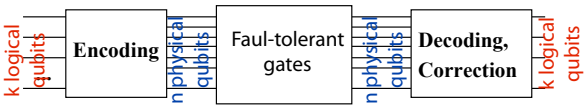

Figure 7.21: Basics of fault-tolerant quantum computing.

This means that all quantum gates (including those used in quantum error correction) should be implemented in such a way that they do not assume the input qubits or the gate operations to be perfectly free of errors. As a consequence gates should not operate on single logical qubits (which do not offer any possibility of detecting and correcting errors), but on the redundant code words of a quantum error-correcting code. During these operations care must be taken to keep errors from propagating too quickly through

the set of qubits employed. Of course the details of the implementations used in this field depend on the operations as well as the codes employed.

The logical basis states (7.22) of the 5-qubit code are chosen such that

$$\mathbf{NOT}_L = \mathbf{NOT}_i^{\otimes 5},$$

where $\mathbf{NOT}_L$ is the logical NOT operation and $\mathbf{NOT}_i$ are the NOT gates applied to the individual physical qubits. Since they correspond to rotations, the actual gate operations are

$$\mathbf{NOT}_i|0\rangle_i \to |1\rangle_i \qquad \mathbf{NOT}_i|1\rangle_i \to -|0\rangle_i.$$

Other gate operations must be implemented such that they generate a rotation within the subspace spanned by the two logical basis states $|0\rangle_L$ and $|1\rangle_L$ while leaving the remaining Hilbert space invariant. They can thus conveniently be written in a basis $\{|0\rangle_L, |1\rangle_L, |k\rangle\}$, where $|k\rangle$ refers to the remaining basis states. In this basis, every logical gate operation corresponds to the direct sum of a two-qubit gate and a unit operator:

$$\mathbf{H} = \mathbf{H}^{(2)} \oplus \mathbf{1}^{(30)} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & & & \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}.$$

This representation is also useful for visualizing some properties of computing with QEC: the first two basis states span the computational space. Any ideal gate operation therefore has the structure that it implements a unitary transformation $U^{(2)} \oplus \mathbf{1}^{(30)}$.

Figure 7.22 shows the experimental tomographic reconstruction of an experimental implementation of three single qubit gate operations encoded in a five-qubit system, with error correction applied to each gate[83].

More details, including the fault-tolerant implementation of a standard set of universal quantum gates

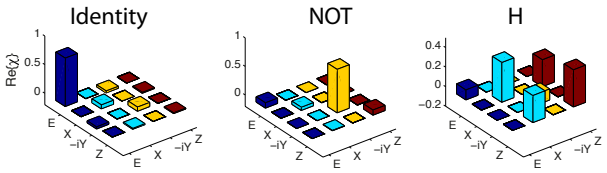

Figure 7.22: Process tomography for three different quantum gate operations on a single logical qubit encoded in five physical qubits. [83]

for the 7-qubit Steane code is discussed in Chapter 10 of [33]. This chapter also contains references to more technical treatments.

The techniques of quantum error-correction, employing concatenated multi-level encoding and fault-tolerant quantum logic, ensure that nontrivial quantum computations may become practical. Under physically reasonable assumptions about the noise present, it has been shown that

> Arbitrarily long quantum computations can be performed reliably and effectively, that is, with an affordable growth in resources such as storage, circuit size, or time, provided that the failure probability in individual quantum gates is below a certain constant threshold [74, 75, 76].

This important result is known as the threshold theorem; additional references to the original work may be found in [33]. A considerable amount of work is currently devoted to establishing the values of the threshold for different encoding and error correction schemes. Qualitatively, the tradeoff is clear: the lower the admissible error per gate (and thus the higher the challenge to implement it), the lower the necessary overhead for quantum error correction.

7.5 Avoiding errors

7.5.1 Robust operations

While error correction represents a necessary part of any quantum computer, the thresholds that have to be reached before error correction can be applied are very high. To make scalable quantum computing feasible, it is therefore necessary to implement strategies that reduce the error probability of each gate. Such efforts must encompass the complete hardware (and software) design.

Making gate operations robust means that they are designed such that deviations in the experimental parameters have only a small effect on the performance of the gate. As an example, a relative deviation $\delta \ll 1$ of the amplitude of a control field should have an effect on the fidelity that is at least of second order, $\propto \delta^2$. The principle of combining different rotations for eliminating imperfections was originally introduced into NMR in 1979 by Malcolm Levitt [84, 85].

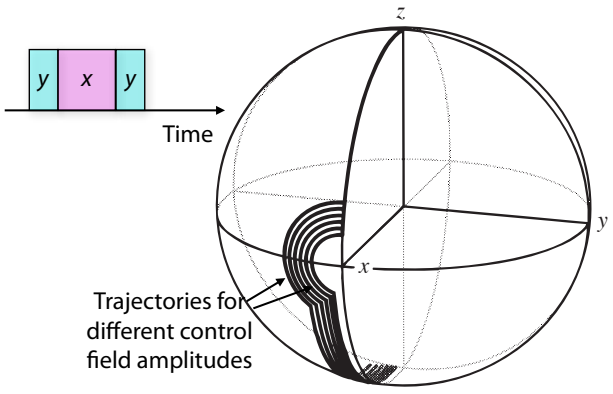

Figure 7.23: Error compensation in a composite pulse by combining rotations around orthogonal axes.

The principle of combining different rotations can be used to eliminate different types of imperfections, as shown in Fig. 7.23 for a simple composite pulse. Here three rotations are combined:

$$\left(\frac{\pi}{2}\right)_{\pi/2} (\pi)_0 \left(\frac{\pi}{2}\right)_{\pi/2},$$

where $(\alpha)_\beta$ is a rotation by an angle α around an axis in the xy-plane, which is oriented at an angle β from the x-axis. This combination achieves a nearly perfect inversion even if the amplitude of the gate differs by $\pm 10\%$. This is shown in Fig. 7.23 by the trajectories that correspond to different rf field strengths: They all end up close to the south pole of the Bloch sphere.

This effect can be analyzed, e.g., multiplying the rotation matrices for the individual rotations:

$$R_{\beta,x} = \begin{pmatrix} 1 & & \\ & \cos\beta & -\sin\beta \\ & \sin\beta & \cos\beta \end{pmatrix}$$

$$R_{\beta,y} = \begin{pmatrix} \cos\beta & & \sin\beta \\ & 1 & \\ -\sin\beta & & \cos\beta \end{pmatrix}$$

Multiplying the rotation matrices for

$$\left(\frac{\pi}{2}(1+\delta)\right)_{\pi/2} (\pi(1+\delta))_0 \left(\frac{\pi}{2}(1+\delta)\right)_{\pi/2},$$

we obtain, to first order in δ, the propagator

$$R_3 = \begin{pmatrix} -1 & & \\ & 1 & \\ & & -1 \end{pmatrix} - \delta\pi \begin{pmatrix} & & 1 \\ & & \\ 1 & & \end{pmatrix}.$$

Here, the first term is the ideal π-rotation around the y-axis, while the second term represents the first-order correction. In this case, it corresponds to a rotation around the z-axis. More complicated expansion schemes are required to eliminate also this term.

In many cases, nonideal behavior results not only from a single type of error, but from a combination of different errors. The most important errors are typically an error in the amplitude of the control field

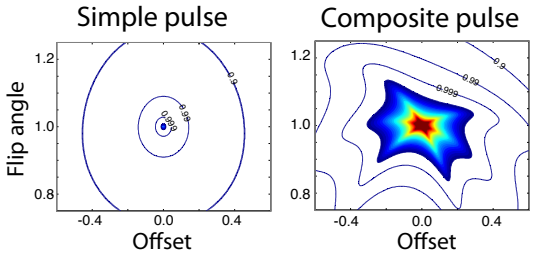

Figure 7.24: Comparison of the robustness of a simple rectangular pulse (left) with that of a compensated composite pulse (right).

and a deviation in the frequency, which is generally called “offset error”.

Fig. 7.24 illustrates how suitable composite pulses can compensate multiple error conditions for the example of a $\pi/2$ -rotation. The colored area indicates the range of parameters over which the actual rotation is sufficiently close to the target operation. For the simple pulse represented in the left-hand side, this area is vanishingly small; any significant deviation makes the pulse useless. The composite pulse whose performance is represented in the right hand panel generates rotations that are close to the target rotation even if the field strength, pulse duration or frequency offset deviate from their nominal values.

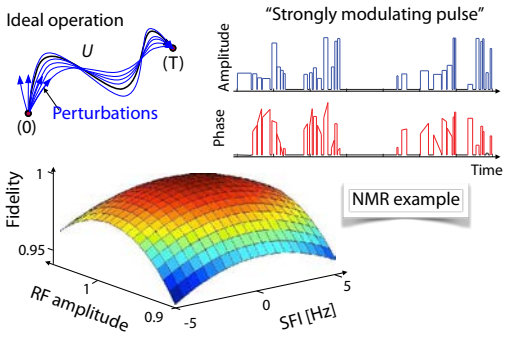

Figure 7.25: Example of a robust pulse designed by optimal control theory.[86]

The simple example presented above can be extened to more robust, more efficient gate operations by considering not only three, but hundreds to thousands of segments. The resulting large number of parameters (amplitudes, frequencies, phases) allows one to design almost arbitrary gate operations that are highly robust and fidelities close to unity. The discipline of optimal control theory offers the theoretical tools for designing such pulses. Fig. 7.25 shows, as an example, the amplitude and phase of such a pulse, which was designed for robust excitation of strongly coupled nuclear spins.

7.5.2 Robust sequences

Using robust gate operations goes a long way towards making quantum computing reliable. This goes at the price that the operations become longer and the amount of energy deposited in the system grows accordingly. This price can be reduced if on considers not only the fidelity of the individual gate operations, but the more relevant question of the fidelity of a sequence of gate operations.

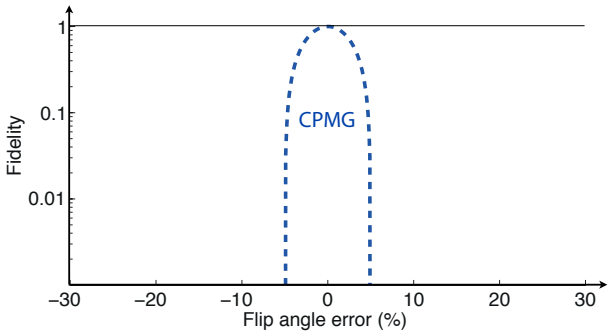

Figure 7.26: Fidelity a sequence of 20 π identical pulses as a function of the flip angle error of the individual pulses.

As an example of the effect of experimental imperfections, consider the dashed curve in Fig. 7.26. It shows the cumulative effect of 20 successive rotations by a nominal angle π around the same axis.

Under ideal conditions, this corresponds to the operation

$$\mathbf{NOT}^{20} = \left(e^{-i\pi \mathbf{S}_x}\right)^{20} = \mathbf{1} = \mathbf{NOOP}.$$

Accordingly, the fidelity was calculated with respect to this ideal case.

If the actual rotation angle differs by a few percent (e.g. 5 %), the error accumulates over the 20 pulses and the total propagator becomes

$$\left(e^{-i\pi 1.05 \mathbf{S}_x}\right)^{20} = e^{-i\pi \mathbf{S}_x} = \mathbf{NOT}.$$

The actual propagator thus has vanishing overlap with the target propagator, the fidelity of the operation is zero. This is the reason that the blue dashed curve in Fig. 7.26 tends to zero for flip angle errors of $\pm 5\%$.

This simple example is useful for illustrating some of the most useful schemes for avoiding errors. The loss of fidelity can be avoided if, instead of applying the 20 successive rotations around the same axis, one applies rotations around a series of different axes. Consider, e.g., the case that the rotations are applied alternating between the x- and $-x$-axes. In this case, the overall operation is

$$\mathbf{NOT}^{20} = \left(e^{-i(1+\delta)\pi \mathbf{S}_x} e^{i(1+\delta)\pi \mathbf{S}_x}\right)^{10} = \mathbf{1} = \mathbf{NOOP},$$

independent of the error δ. This simple 'trick' of alternating the rotation axis thus turns the highly error-prone sequence into a completely robust one, and this is achieved with zero overhead: the duration of the sequence and the amount of energy deposited remains the same!

This principle can be extended: switching not only between two possible orientations of the rotation axis, it is possible to find sequences that are much more robust against different types of experimental imperfections. This is illustrated in Fig. 7.27 by the two curves labelled XY-4 and KDD. In the case of XY-4, the rotation axis alternates between the x and

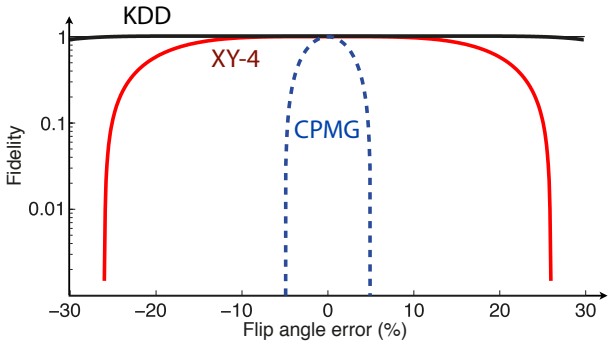

Figure 7.27: Fidelity of different pulse sequences after 20 π pulses as a function of the flip angle error of the individual pulses.

y axes, for the KDD sequence, a 10-step cycle is used for the rotation axes [87]. In all three cases, the error of the individual pulses is the same, but the compensated sequence performs almost flawlessly, even if the flip angle deviates by as much as 30 % from its nominal value.

7.5.3 Protection against decoherence

We now turn to the question how information can be protected from environmental noise by passive measures, without applying control fields or using error correction schemes.

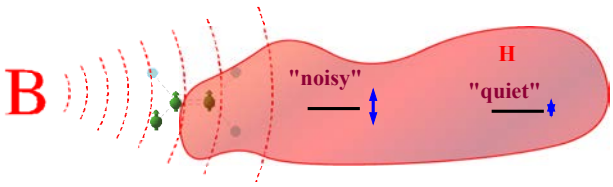

Figure 7.28: Storing and processing information in quiet regions of Hilbert space can reduce the error rate.

Fig. 7.28 illustrates the guiding idea: In most systems, noise does not affect the full Hilbert space in the same way. Instead, some areas are more noisy

than others. It is therefore essential to identify the quiet areas and preferentially use them for storing information.

For the discussion of decoherence processes, one has to distinguish between different types of coupling between the system and environment:

(i) Total decoherence. This is the most general case, essentially there are no restrictions on the operators that generate the decoherence.

(ii) Independent qubit decoherence. If the coupling operator contains only operators acting on individual qubits, errors of individual qubits are independent. This is the case typically considered in quantum error correction.

(iii) Collective decoherence. Here the coupling operators act in the same way on all qubits. They can thus be written in the form

$$\mathbf{F}_\alpha = \sum_i \mathbf{S}_\alpha^i, \tag{7.23}$$

where $\alpha = x,y,z$ marks the Cartesian component and i the index of the qubit. Clearly the perturbation has full permutation symmetry in this case. This symmetry can be exploited in a counterstrategy that we discuss in section 7.5.4. Only three independent perturbation operators exist in this case.

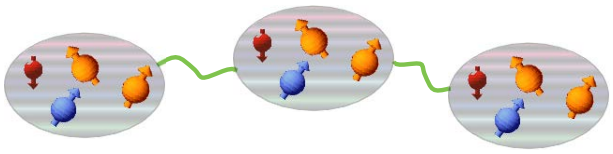

Figure 7.29: Schematic representation of cluster decoherence.

(iv) Cluster decoherence. This is an intermediate case, where clusters of qubits decohere collectively, while the different clusters decay independently.

The cases discussed above are idealized situations. Real systems may be close to one of them or intermediate between several limiting cases.

7.5.4 Decoherence-free subspaces

Decoherence-free subspaces represent a possibility for shielding quantum information from the decoherence processes caused by the environment by taking advantage of the symmetry properties of the coupling operators between the system and environment [60]. We follow the discussion of Lidar, Chuang and Whaley [88].

As discussed before, decoherence can be seen to arise from interactions with the bath. It is therefore useful to distinguish three contributions to the Hamiltonian of the full system (including the bath):

$$\mathscr{H} = \mathscr{H}_S \otimes \mathbf{1}_B + \mathbf{1}_S \otimes \mathscr{H}_B + \mathscr{H}_{\text{Int}}.$$

Here $\mathscr{H}_S$ is a pure system operator, $\mathscr{H}_B$ is a pure bath operator, and $\mathscr{H}_{\text{Int}}$ represents the coupling operator. It contains product operators

$$\mathscr{H}_{\text{Int}} = \sum_\alpha \mathbf{F}_\alpha \otimes \mathbf{B}_\alpha,$$

where $\mathbf{F}_\alpha$ are system operators and $\mathbf{B}_\alpha$ bath operators. If the system is a spin system, the $\mathbf{F}_\alpha$ are spin operators. Depending on the type of environment, the $\mathbf{B}_\alpha$ may be spatial coordinates, creation / annihilation operators, fields, spin components or some other degree of freedom.

Decoherence is the nonunitary part of the evolution of the system density matrix ρ_S, which, under appropriate conditions, can be written as [89]

$$\begin{aligned} &\frac{d}{dt}\rho_S + \frac{i}{\hbar}[\tilde{\mathscr{H}}_S, \rho_S] \\ = \;& \frac{1}{2}\sum_{\alpha,\beta} a_{\alpha\beta}\left([\mathbf{F}_\alpha, \rho_S \mathbf{F}_\beta^\dagger] + [\mathbf{F}_\alpha \rho_S, \mathbf{F}_\beta^\dagger]\right). \end{aligned} \tag{7.24}$$

Here $\tilde{\mathscr{H}_S}$ is the system Hamiltonian plus any possible unitary contributions arising from the system-bath interaction, and $a_{\alpha\beta}$ are elements of a positive semi-definite Hermitian matrix. The operators $\mathbf{F}_\alpha$ are the generators of the decoherence process. We may thus consider the possible decoherence processes in terms of these operators. In spin systems these are clearly the spin operators; for the typical case of spin-1/2 systems, they are multiples of the Pauli matrices.

Depending on the generators $\mathbf{F}_\alpha$, not all states are equally subject to decoherence. Decoherence-free subspaces exist if, for a certain set of states $|i\rangle$, the coupling to the environment does not generate a time evolution. For a formal analysis, we write the corresponding part of the density operator

$$\tilde{\rho} = \sum_{i,j} \tilde{\rho}_{i,j} |i\rangle\langle j|,$$

where the coefficients $\tilde{\rho}_{i,j}$ depend on the initial conditions. The condition for the existence of the decoherence-free subspace is then, that the right-hand side of (7.24) vanishes for this state:

$$\frac{1}{2}\sum_{\alpha,\beta} a_{\alpha\beta}\left([\mathbf{F}_\alpha, \tilde{\rho}\mathbf{F}_\beta^\dagger] + [\mathbf{F}_\alpha\tilde{\rho}, \mathbf{F}_\beta^\dagger]\right) = 0.$$

This condition can be fulfilled in a number of ways, depending on the initial conditions (via the $\tilde{\rho}_{i,j}$) and on the coupling to the bath (via the $a_{\alpha\beta}$). However, decoherence-free subspaces are only interesting if no additional constraints have to be imposed on the bath parameters (which are hard to control) or the initial conditions of the system (since we would like a general-purpose computer). Such additional constraints can be avoided if the states $|i\rangle$ satisfy the condition [88]

$$\mathbf{F}_\alpha|i\rangle = c_\alpha|i\rangle \tag{7.25}$$

for all operators $\mathbf{F}_\alpha$. This means that the states $|i\rangle$ of the decoherence-free subspace form a degenerate set of eigenstates for all error generators. Obviously this is a rather restrictive criterion. To motivate that relevant cases exist that approximately fulfill this condition, we discuss a few examples after we have finished the formal analysis.

7.5.5 Information capacity

Clearly, this concept is only useful if a significant amount of information can be encoded in a decoherence-free subspace. The answer depends on the type of decoherence, i.e. on the set of operators $\mathbf{F}_\alpha$. For collective decoherence, DFS turn out to be interesting, since the DFS asymptotically fill the Hilbert space completely. In this case there are only three independent perturbation operators, the total spin operators (7.23). As usual, we discuss the situation in terms of spins, with the understanding that they may also be representing pseudo-spins, i.e. general qubits.

While the condition (7.25) requires only that the states of the DFS have all the same eigenvalues c_α, we discuss here only the case $c_\alpha = 0$. This implies that the DFS is spanned by all singlet (total spin quantum number $S_T = 0$) states of, say, K spins (where K must be even). The number of these states can be determined by considering states with a given total spin z component S_T^z. The total number of $S_T^z = 0$ states is $\left(\begin{array}{c} K \\ K/2 \end{array}\right)$, the number of ways to pick $K/2$ down spins from a total of K spins. Some of these $S_T^z = 0$ states are the desired singlets, the others belong to subspaces with $S_T \neq 0$. Every such subspace contains exactly one $S_T^z = 1$ state. An $S_T^z = 1$ state corresponds to$K/2+1$ spins in the 'up' state and $K/2-1$ spins in the 'down' state, so the total number of $S_T^z = 1$ states is $\left(\begin{array}{c} K \\ K/2-1 \end{array}\right)$. Hence the number of $S_T = 0$ states (or subspaces, since each subspace is one-dimensional) is

$$\begin{aligned} \dim[\mathrm{DFS}(K)] &= \left(\begin{array}{c} K \\ K/2 \end{array}\right) - \left(\begin{array}{c} K \\ K/2-1 \end{array}\right) \\ &= \frac{K!}{\left(\frac{K}{2}\right)!\left(\frac{K}{2}+1\right)!}. \end{aligned} \tag{7.26}$$

The number N of logical qubits that can be stored in this DFS of K physical qubits then is

$$\begin{aligned} N &= \log_2 \dim[\mathrm{DFS}(K)] \\ &= K - \frac{3}{2}\log_2 K + \mathscr{O}(1), \end{aligned}$$

where we have used Stirling's formula (for large n)

$$\ln n! = (n+\frac{1}{2})\ln n - n + \mathscr{O}(1). \tag{7.27}$$

In the limit of large systems, $K \gg 1$, the information capacity of the DFS therefore asymptotically approaches that of the full Hilbert space. The result (7.26) for collective decoherence was first derived from group-theoretical considerations in [90].

In contrast to this case, where the decoherence-free subspaces asymptotically fill the whole Hilbert space, in the opposite limit of individual qubit decoherence or total decoherence, the amount of information that can be encoded in DFSs is negligibly small.

The last requirement that must be met is to implement gates in this DFS. This is easily achieved in the generic model, but actual implementations in physical systems are still rare and must be discussed for the specific examples. Here, we discuss the generic example of spins in magnetic fields.

7.5.6 Example: spin qubits

The simplest example of a decoherence-free subspace is provided by magnetic resonance if we consider the decoherence induced by randomly fluctuating homogeneous magnetic fields. They couple to the spin system through the sum of the z-components of the nuclear spin operators,

$$\mathscr{H}_z = b(t)\sum_i \mathbf{I}_z^i,$$

where $b(t)$ describes the fluctuating magnetic field and we consider a system of identical spins (a homonuclear spin system). This Hamiltonian generates a diffusion-like evolution of the spins, as discussed in section 7.2.2. In this case, $\mathbf{F} = \sum_i \mathbf{I}_z^i$ and according to (7.25) the states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ with $m_1 + m_2 = 0$ should form a decoherence-free subspace.

This can be seen by considering the effect of the randomly fluctuating field on the coherences $\rho_{ij} = \langle i|\rho|j\rangle$. All states $|i\rangle$ with the same z-component of the total spin, $m = \langle i|\sum_k \mathbf{I}_z^k|i\rangle$, have the same energy and are therefore shifted by the same amount if the external field fluctuates.[1] The effect of field fluctuations on off-diagonal density operator elements is then

$$i\hbar\frac{d}{dt}\rho_{ij} = b(t)\Delta m_{ij}\rho_{ij},$$

where

$$\Delta m_{ij} = \langle i|\sum_k \mathbf{I}_z^k|i\rangle - \langle j|\sum_k \mathbf{I}_z^k|j\rangle$$

and the sum runs over all spins. Δm_{ij} represents the change in the total magnetic spin quantum number, which is proportional to the difference in Zeeman energy between the two states $|i\rangle$ and $|j\rangle$. We can therefore eliminate the decoherence due to such a process if we encode a qubit not in a single spin but associate the logical states as

$$|0\rangle = |i\rangle \quad ; \quad |1\rangle = |j\rangle$$

with

$$\Delta m_{ij} = 0.$$

As shown in Fig. 7.30, the energy of the states in the $m = 0$ subspace does not depend on the strength of the magnetic field and therefore is not affected by fluctuations in the field.

[1] The energies are not exactly identical, since small energy differences (due to chemical-shift interactions) are used for addressing the individual qubits. However, these differences are small, of the order of 10^{-4} to 10^{-6} times the Zeeman energy.

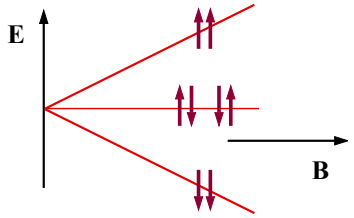

Figure 7.30: Energies of the 2-spin product states as a function of the strength of the external magnetic field.

In such an encoding scheme, the logical states are not associated with single physical qubits. As a result, one does not have immediate access to manipulate the system, i.e., to apply gate operations to these logical qubits. How this is done depends on the actual implementation and will not be discussed here.

From what has been said so far, it should be obvious that such an encoding scheme will only work for fluctuations of the field in the direction of the static field, i.e., along the z-axis. If more complex systems of coupling operators are present, it is still possible to design decoherence-free subspaces. While the general analysis is rather mathematical and mainly relies on existence proofs, without constructing an actual DFS [91], it is relatively easy to see that if a number of states are available that are immune to noise coupling to $\sum_i \mathbf{I}_z^i$, arbitrary linear combinations of these states are still immune to this type of noise. It is then possible to choose a suitable linear combination such that it is also immune to noise (e.g.) coupling to $\sum_i \mathbf{I}_x^i$.

The basic scheme has not only been discussed theoretically, but a number of experiemtal proofs of principle have also been performed. A single qubit of information was encoded in three spins in such a way that it was protected from global noise along all three axes [92]. The experimental results show that the information that is contained in the noiseless subspace decays significantly slower than the unprotected information. However, the encoding – decoding process is not error-free, so the fidelity with the encoding process is actually much lower than without the encoding for most of the range of experimental parameters.

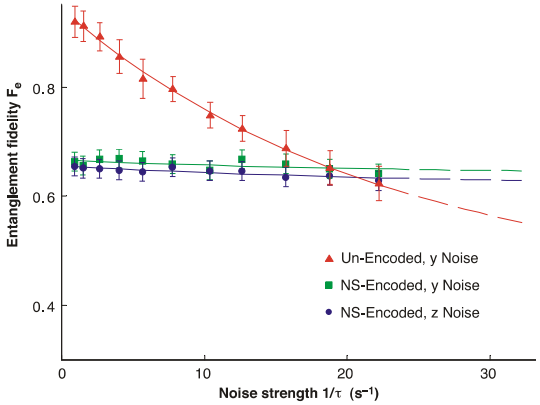

Figure 7.31: Experimental results for the decay of a DFS-encoded qubit.[92]

More recently, a complete quantum algorithm (Grover's algorithm on two qubits) was implemented in a decoherence-free subspace that was embedded in a four-spin system in such a way that it reliably reached the correct result in the presence of strong decoherence [93].

7.5.7 Clock transitions

The basic idea of using subspaces of the Hilbert space that are less sensitive to environmental perturbations than others has been exploited in different fields for a long time. A prominent example is that of atomic clocks where the evolution of the coherence in a chosen transition

$$|i\rangle\langle k|(t) = |i\rangle\langle k|(0)\, e^{-i\omega_{ik}t}$$

is used as a measure of time. Clearly, a variation of the level splitting $\hbar\omega_{ik}$ causes the clock to run too fast or too slow. The currently used time / frequency standard is defined as the duration of 9192631770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the cesium-133 atom.

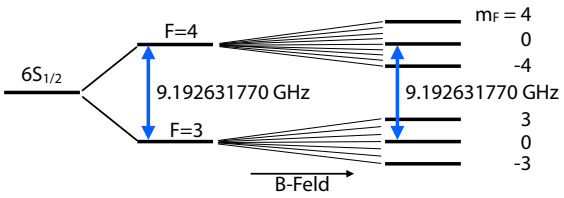

Figure 7.32: Ground state sublevels of cesium. The frequency of the clock transition ($m_F = 0 \leftrightarrow m_{F'} = 0$) transition is independent of the magnetic field (to first order).

A closer look at the level scheme of the cesium ground state (see Fig. 7.32) shows that the state splits not only into two hyperfine substates, but they again consist of a total of 16 Zeeman sublevel, which are shifted by the magnetic field by

$$\delta\mathscr{E} = m_F g_F \mu_B B_z,$$

where g_F is the Landé factor and μ_B the Bohr magneton and the $z-$axis is chosen along the magnetic field B. Accordingly, any perturbing magnetic field causes deviations of the atomic clock. The main exception from this is if the $m_F = 0 \leftrightarrow m_{F'} = 0$ transition is used, since these two states do not depend on magnetic fields acting on the atom.

The same principle can be applied to solid-state qubit systems. Fig. 7.33 shows, as an example, the energy level scheme of a Bi electron spin defect in silicon. The electron spin is $S = 1/2$, and therefore no field-independent state exists. However, the energy levels are also influenced by the nuclear spin $I = 9/2$ and the corresponding hyperfine interaction, which splits the system into an $F = 4$ and an $F = 5$ multiplet. In the presence of a magnetic field, these levels depend in a nonlinear way on the field strength, as shown in the upper part of Fig. 7.33. The middle panel shows, how most of the transition frequencies

$$\nu_{ik} = \left| \frac{\mathscr{E}_i - \mathscr{E}_k}{h} \right|$$

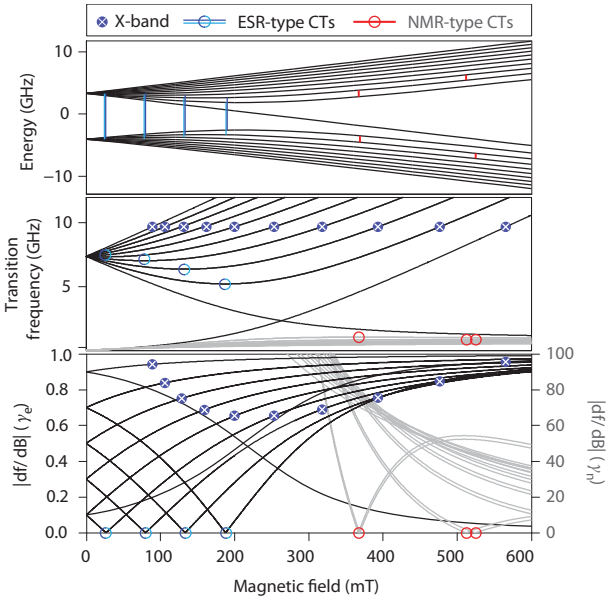

Figure 7.33: Energy levels, transition frequencies and field sensitivities for different transitions of a Bi qubit in Si.[94]

increase with the field strength, but some of them pass through a minimum, where

$$\frac{d\nu_{ik}}{dB} = 0.$$

These points are marked by blue circles. Since the frequencies are not affected by variations in the magnetic field strength at these points, the transitions are insensitive to magnetic field noise. In such a system, a phase relaxation time of $T_2 \approx 2.7$ s was measured recently [94], which is extremely long for a solid-state system.

This type of protection scheme has been applied primarily to magnetic fields, where the acronym ZEFOZ (=zero first-order Zeeman) is used. It can be used, e.g., to extend the dephasing time of spin qubits in rare earth ions. Fig. 7.34 shows that the application of a suitable magnetic field extends the dephasing time by up to three orders of magnitude.

The conditions for the existence of a suitable ZEFOZ

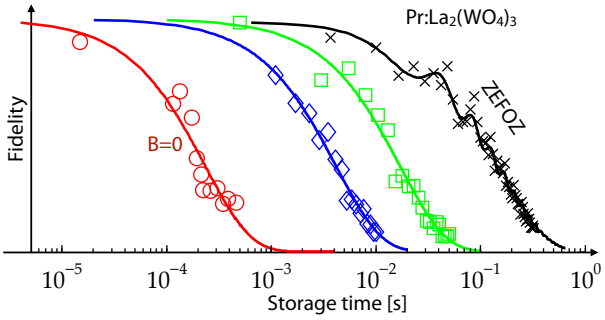

Figure 7.34: Storage of photons in the nuclear spin states of $Pr:La_2(WO_4)_3$.

point are sometimes difficult to find, since they depend on the details of the level strucure. A possible alternative is to apply near-resonant alternating fields that create dressed states forming decoherence-free subspaces. [95, 96]

7.5.8 The quantum Zeno effect

While the DFS-approach to protecting quantum information is purely passive, i.e. it requires no experimental actions, it is also possible to reduce decoherence by active means other than error correction. One such approach, called dynamical decoupling, will be discussed in section 7.6. Here, we briefly discuss a related approach, which is practically less relevant, but brings some interesting insight into the dynamics of quantum mechanical systems. It is based on the *quantum Zeno effect* [97]. The idea behind this radical simplification is to keep the quantum state error-free by projecting frequently (by a measurement) onto the subspace corresponding to the "no error" syndrome.

Zeno of Elea (ca. 490 – 430 b.C., southern Italy) was a student of Parmenides. He stated a number of paradoxa to defend the teachings of Parmenides, in particular the statement that motion is impossible and more than one thing cannot exist. One well known paradox is that of the race between Achilles and the tortoise. Achilles (the fastest man in antiquity) is ten times as fast as the tortoise. Nevertheless he cannot overtake her if she gets a head start of (e.g.) 10 m: Achilles first must cover these 10 m. During this time, the tortoise moves 1 m and is therefore still ahead. While he covers this meter, the tortoise moves another 0.1 m and so on, always staying ahead.

Another motion paradox "proves" that a body cannot move from A to B: for this, it would first have to move to the middle of the distance. For this it would first have to move to the middle of the first half, etc.

While these paradoxa are easily resolved, similar situations exist in quantum mechanics that are real. They have been discussed under the heading "quantum Zeno effect", although they cannot really be considered paradoxa.

We consider the evolution of a system that is initially (at $t = 0$) prepared in the state $|\psi_i\rangle$, which is an eigenstate of operator $\mathbf{A}$ with eigenvalue a_i. The state evolves under the influence of a Hamiltonian $\mathscr{H}$, which does not commute with $\mathbf{A}$. A possible example would be that the Hamiltonian is $\propto \mathbf{S}_z$ and the observable is $\mathbf{S}_x$. A measurement with $\mathbf{A}$ of the system after some time τ will then in general yield a result that is different from a_i.

For a qubit or a single spin 1/2, we can consider a spin in the $m_x = +1/2$ eigenstate of $\mathbf{S}_x$,

$$\Psi(0) = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

In a magnetic field $\vec{B}_0 || z$, the energy of the computational basis states are

$$\begin{aligned} \mathscr{E}_{|0\rangle} &= -\frac{\hbar}{2}\gamma B_0 = -\frac{\hbar\omega_L}{2} \\ \mathscr{E}_{|1\rangle} &= +\frac{\hbar}{2}\gamma B_0 = \frac{\hbar\omega_L}{2}, \end{aligned}$$

Therefore the state Ψ evolves to

$$\Psi(t) = \frac{1}{\sqrt{2}}(|0\rangle e^{i\omega_L t/2} + |1\rangle e^{-i\omega_L t/2}),$$

where $\omega_L = \gamma B_0$ is the Larmor frequency. The probability that a subsequent measurement of $\mathbf{S}_x$ at time t also finds the eigenvalue $+1/2$ is then

$$\begin{aligned} p_+ &= |\langle\Psi(0)|\Psi(t)\rangle|^2 = \\ &= \frac{1}{4}\left|e^{i\omega_L t/2} + e^{-i\omega_L t/2}\right|^2 \\ &= \left|\cos(\frac{\omega_L t}{2})\right|^2 = \frac{1}{2}(1+\cos(\omega_L t)). \end{aligned}$$

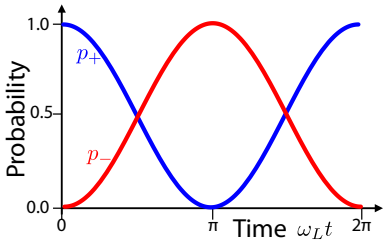

Figure 7.35: Probabilities p_+ for measuring the initial state and p_- for measuring the opposite state.

The probability of obtaining the opposite result is

$$p_- = \frac{1}{2}(1-\cos(\omega_L t)).$$

7.5.9 Repeated measurements

If such a measurement is performed, the projection postulate states that after the measurement the system is in an eigenstate of $\mathbf{A}$. If the measurement yielded the result +1/2, the system is again in the same initial state, and the evolution starts out again with the same time dependence. While the probability for this outcome is less than unity, the important point is that the first derivative of the time dependence,

$$\frac{d}{dt}p_+\Big|_{t=0} = -\frac{1}{2}\omega_L \sin(0) = 0,$$

vanishes after the projection. During short times after the measurement, the system therefore does not evolve significantly.

If a series of measurements is repeated with a separation (in time) of τ, the probability that n measurements in sequence will always find the system in state $\Psi(0)$, corresponding to $m_x = +1/2$, becomes

$$p_+ = \frac{1}{2^n}(1+\cos(\omega_L\tau))^n. \tag{7.28}$$

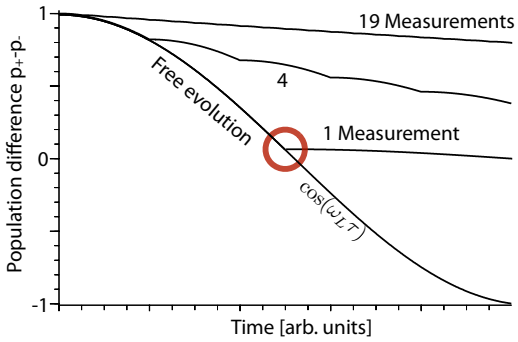

Figure 7.36: Quantum Zeno effect: the decay of a state becomes slower with increasing number of measurements.

Figure 7.36 shows how the evolution of the system changes as the measurement interval decreases. We now consider specifically the situation for short measurement intervals, $\omega_L\tau \ll 1$. In this limit, the cosine can be expanded as $\cos(x) \approx 1 - x^2/2$ and eq. (7.28) can be written as

$$\begin{aligned} p_+(n\tau) &\approx \frac{1}{2^n}\left(1+(1-\frac{(\omega_L\tau)^2}{2})\right)^n \\ &= \frac{1}{2^n}\left(2-\frac{(\omega_L\tau)^2}{2}\right)^n = \left(1-\frac{\omega_L^2\tau^2}{4}\right)^n. \end{aligned}$$

Writing $t = n\tau$ for the full duration, this becomes

$$p_+(t) \approx \left(1-\frac{\omega_L^2\tau t}{4n}\right)^n$$

Using the relation

$$\lim_{n\to\infty}\left(1-\frac{\varepsilon}{n}\right)^n = e^{-\varepsilon}$$

we can simplify the probability to

$$p_+(n\tau) = p_+(t) \approx \exp\left(-\frac{\omega_L^2 \tau}{4} t\right).$$

We therefore obtain an exponential decay with the decay rate $\omega_L^2\tau/4$. The system evolution is no longer periodic, but shows an exponential decay. The decay rate decreases with the interval between measurements. For quickly repeated measurements, it can therefore be made arbitrarily slow. This is referred to as the **quantum Zeno effect**:

> In the limit of frequent measurement, the system does not evolve.

7.5.10 Experimental example

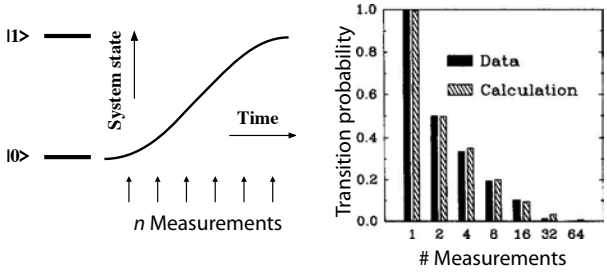

Figure 7.37: Experimental test of the quantum Zeno effect. Left-hand side: laser pulses measure the state of the ions while they are attempting to make a transition from state $|0\rangle$ to $|1\rangle$. Right-hand side: calculated and measured transition probability for increasing number of measurements [98].

These general quantum mechanical predictions can be verified experimentally, e.g., for trapped ions [98]. The left-hand part of figure 7.37 shows the principle of the experiment. The ions are initially in state $|0\rangle$, from where an RF field drives them into state $|1\rangle$. The amplitude of the RF field and its duration can be adjusted such that probability for the ion to make the transition from state $|0\rangle$ to $|1\rangle$ approaches unity at time τ.

To detect if the ions have arrived in state $|1\rangle$, one can use laser pulses that excite fluorescence from the ions if they are in state $|1\rangle$; with a suitable calibration, the fluorescence signal can be used to measure whether the ions are in this state. If such a laser pulse is applied first at time τ, it finds the ions in state $|1\rangle$ with almost unit probability. If, however, additional measurements are made at times $\tau_i = \tau\frac{i}{n}$ for $i = 1..n$, the probability of finding the system in state $|1\rangle$ at time τ is reduced to

$$p(n) = \frac{1}{2}[1 - \cos^n\left(\frac{\pi}{n}\right)].$$

For $n = 1,2,3,4$, we obtain $p(n) = 1, \frac{1}{2}, \frac{7}{16}, \frac{3}{8}$. In the limit of large n, the argument of the cosine tends to zero and

$$\cos^n\left(\frac{\pi}{n}\right) \to 1\,, \quad p(n) \to 0.$$

This prediction was verified experimentally by measurements on two hyperfine states of the $^9Be^+$ ground state [98]. A radio-frequency field was used to drive the transition. In the absence of measurements, this results in an oscillatory exchange of populations. The measurements were performed by selective laser irradiation, which generate a signal if the transition has occurred. The results, shown in the right-hand side of Figure 7.37, correspond to $1 - p(n)$ and are in good agreement with theory. For 64 measurements, the transition probability has dropped to a very small value, indicating that the evolution was quenched.

Clearly the slow-down of transition rates by measurement cannot be universal. As an example, consider an atom that is initially in the excited state. A possible measurement for the excited state population probability is a fluorescence measurement: as long as we do not observe a fluorescence photon from this atom, we know it is still in the excited

state. This would imply that, if we only "looked" at the atom often enough, it would therefore be impossible for the atom to decay. Similar arguments are used to explain why the decay of the proton has not yet been observed.

The main reason for this paradox is that the concept of a quantum mechanical measurement is not established with sufficient precision. A projection, i.e., a reduction of the wave packet, does not always occur in "standard" quantum mechanical measurements. If the interaction is weak (such as "looking" for a fluorescence photon), the reduction does not occur. One important point that must be considered is that a projective measurement can only occur during a finite time interval. If the interaction between the measurement apparatus and the system is weak, this time interval can become very long. In the experiment with the trapped ions, the interaction strength of the measurement is determined by the intensity of the laser used for the measurement. The projection postulate is well suited to the Stern–Gerlach type experiment, but completely unsuitable for experiments like NMR.

7.6 Fighting Decoherence

While passive methods for avoiding decoherence are useful, they are alone not sufficient. This section therefore discusses techniques that use active control operations for refocusing the environmental disturbance.

7.6.1 Refocusing

The basic experiment that refocuses environmental interactions uses a NOT gate to invert the unwanted time-evolution.

This approach to reducing decoherence was originally introduced in Nuclear Magnetic Resonance (NMR), in particular by Erwin Hahn [99], who

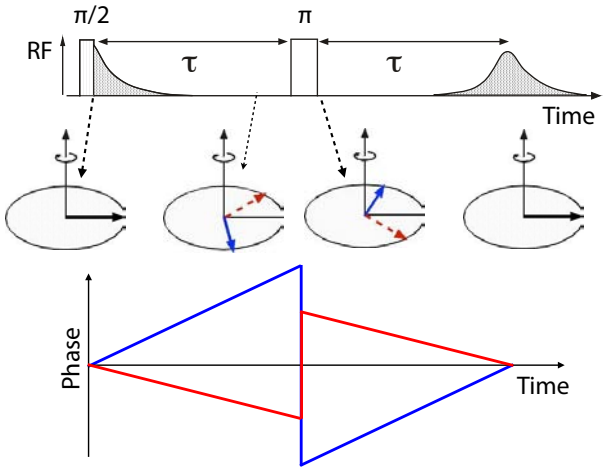

Figure 7.38: Phase reversal and echo formation by an inversion pulse applied to the qubit.

showed that a π-rotation (a NOT-gate) applied to a spin-1/2 system (a qubit) corresponds to an effective change of the sign of the perturbation Hamiltonian and therefore generates a time reversal of the corresponding evolution.

This principle can be understood by considering a qubit (=spin 1/2) in a superposition state. In an external field that splits the two states by a frequency ω, it evolve as

$$|\Psi\rangle(t) = \frac{1}{\sqrt{2}}\left(|0\rangle e^{i\omega t/2} + |1\rangle e^{-i\omega t/2}\right),$$

i.e. the relative phase ϕ of the coherence increases linearly with time, $\phi = \omega t$.

The NOT gate which is applied at time τ therefore changes the state to

$$|\Psi\rangle(t) = \frac{1}{\sqrt{2}}\left(|0\rangle e^{-i\omega t/2} + |1\rangle e^{i\omega t/2}\right).$$

The relative phase between the two components has thus been reversed, from $\phi_1 = \omega\tau$ to $\phi_1' = -\omega\tau$. Depending on the factor in $\omega\tau$ to which we associate this sign change, we can consider the effect as an inversion of the Hamiltonian for the period before the

pulse, from $\omega \to -\omega$ or to a reversal of the time evolution, from $\tau \to -\tau$.

As the evolution continues, the coherence continues to acquire phase,

$$\phi = -\omega\tau + \omega(t-\tau) = \omega + \omega(t-2\tau).$$

Over the next period of duration τ, the additional phase $\phi_2 = \omega\tau$ exactly cancels the phase $\phi_1^{'}$ and the sum of the two phases vanishes, $\phi_1^{'} + \phi_2 = 0$. It therefore appears as if the system had never undergone an evolution. Since this is true for all spins, independent of the interaction with the environment, the dephasing due to an inhomogeneous interaction is exactly cancelled by this refocusing pulse. All phases vanish and the qubits get back into phase, forming an echo at time τ after the refocusing pulse.

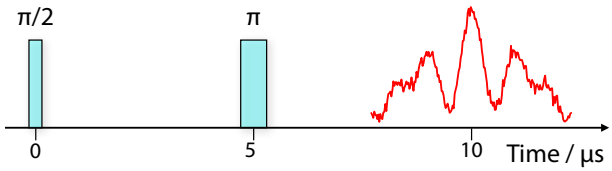

Figure 7.39: Experimental echo signal from a single electron spin of an NV center in diamond.

7.6.2 Fluctuations

In practice, refocusing never works perfectly. The most critical assumption is that the environment should be static, i.e. the interaction with the environment is time independent. In practice, there are always fluctuations. As a result of these fluctuations, a qubit may experience a different interaction with the environment after the refocusing pulse than before it. In this case, the phase acquired by the environmental interaction does not cancel and some destructive interference remains and the echo amplitude decays as a function of the refocusing time [99, 100]. This decay contains information about the time-dependence of the environment.

To explain how such a time-dependence arises, we introduce a simple model Hamiltonian. The free evolution Hamiltonian, without gate operations, is

$$\mathscr{H}_f = \mathscr{H}_{SE} + \mathscr{H}_E,$$

where $\mathscr{H}_E$ describes the environment and

$$\mathscr{H}_{SE} = \sum_\beta b^\beta E_z^\beta S_z$$

is the interaction between the system and the environment. E_z^β are operators of the environment and b^β the system-environment coupling constants. The index β runs over all modes of the environment. If the environmental Hamiltonian $\mathscr{H}_E$ does not commute with E_z^β, $\mathscr{H}_{SE}$ also undergoes a time evolution induced by $\mathscr{H}_E$.

A similar effect arises if the spins diffuse in an environment with an inhomogeneous magnetic field. The diffusion then changes the Larmor frequency of the spin and the refocusing becomes ineffective.

7.6.3 Dynamical Decoupling

A technique for reducing this effect was introduced by Carr and Purcell [100] and an improved version by Meiboom and Gill [101]: Instead of applying a single pulse in the middle of the period, they applied a sequence of pulses, with separations between them that were short compared to the timescale on which the environment changes.

As shown in Fig. 7.40, each pulse generates a new echo. Similar to the case of the Zeno-effect, the decay of the echo envelope slows down as the spacing between the pulses is decreased. If the pulse spacing becomes short compared to the environmental fluctuations, they become unimportant and refocusing is re-established.

The same idea was introduced in the context of quantum information processing under the name of dynamical decoupling (DD) [102].

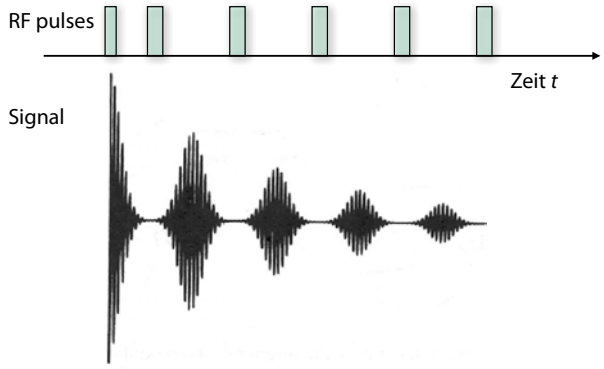

Figure 7.40: Echo train generated by a sequence of π-rotations.

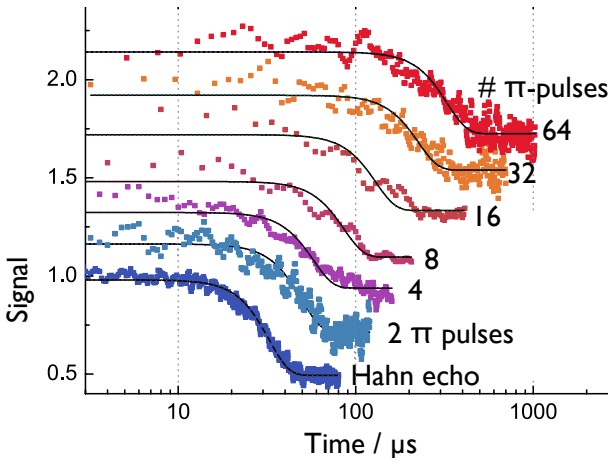

Figure 7.41: Decay of electron spin coherence for different numbers of refocusing pulses.

Fig. 7.41 shows that the application of refocusing pulses effectively decouples the qubit from the environment, increasing the survival time. The more pulses are applied (and thus the shorter the delay between the pulses), the longer the survival time of the electron spin coherence. For the conditions shown here (a single electron spin in a diamond NV-center), the coherence time increases by roughly one order of magnitude as the number of refocusing pulses is increased from 1 to 64.

7.6.4 Imperfections

If the refocusing pulses are ideal, i.e. perfect π-rotations of zero duration, it would be possible to keep increasing the number of refocusing pulses and thereby "freeze" the evolution of the system by completely isolating it from its environment. Unfortunately, experimental pulses are not perfect. They have finite durations, they may have a frequency offset, and most importantly, their flip angles differ from the target value, typically by as much as a few percent. The effect of such imperfections becomes most important when a large number of gate operations are used, such as in dynamical decoupling.

An ideal refocusing sequence works perfectly, independent of the initial condition of the system (which should be considered to be unknown). As Fig. 7.42 shows, this is not always the case. Here, the black symbols, representing experimental data points, show that the relaxation (=decoherence) time of the system increases by roughly two orders of magnitude as the delay between the pulses is reduced. However, a further reduction of the delay and therefore increased number of pulses does not lead to a further increase, but actually to a small reduction. Even more strikingly, the red symbols, which correspond to measurements where the initial condition is perpendicular to the rotation axis, indicate that an increasing number of pulses reduces the relaxation time of the system. In this case, the pulses

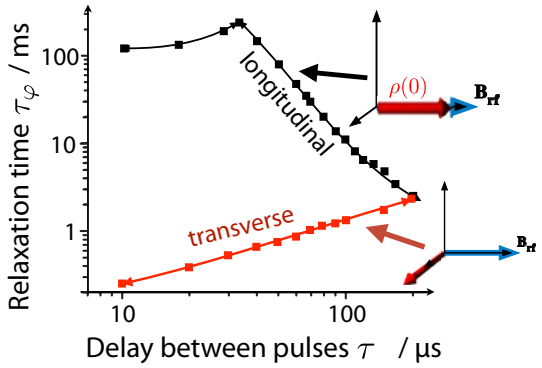

Figure 7.42: Decay times of nuclear spin coherence for different numbers of refocusing pulses. "Longitudinal" and "transverse" refer to the orientation of the spin with respect to the rotation axis of the pulses.

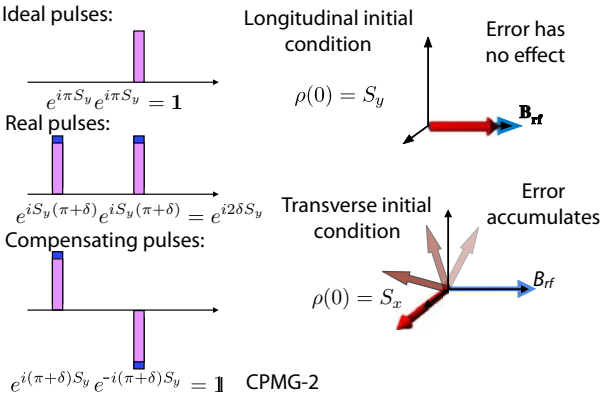

Figure 7.43: Effect of pulse imperfections: If the flip angle error is not set precisely, the error accumulates over a 2-pulse cycle. This does not affect the qubits if their polarization is aligned with the rotation axis, but if it is perpendicular to it, the errors accumulate. The problem can be solved by switching the rotation axis between opposite directions.

apparently do not help, but actually destroys spin coherence.

This effect can be understood by considering the effect of two π_Y pulses. The total propagator for two such pulses is

$$U_{id} = e^{i\pi S_y} e^{i\pi S_y} = \mathbf{1},$$

i.e. the system returns to its initial state.

If we consider now two pulses whose flip angle is $\pi + \delta$, where δ is the flip angle error, the total propagator becomes

$$U_\delta = e^{iS_y(\pi+\delta)} e^{iS_y(\pi+\delta)} = e^{i2\delta S_y}.$$

This is not a problem as long as the initial condition is aligned with the y-axis, $\rho(0) \propto S_y$: in this case, the density operator commutes with the error propagator, $[U_\delta, S_y] = 0$, indicating that the error does not affect the state. If, however, the initial state is $\rho(0) \propto S_x$, which is equally possible, the commutator does not vanish and the error causes a rotation of the qubit. This rotation accumulates over many cycles and results in a loss of coherence. This is an example of the general rule for quantum information processing:

> The quality of gate operations must be high for arbitrary initial conditions.

7.6.5 Error compensation

The basis idea of reducing the effect of pulse imperfections in sequences of gate operations, which was discussed in section 7.5.2, can be applied to the example of dynamical decoupling. The bottom left part of Fig. 7.43 shows how this problem can be solved: instead of applying all pulses with the same sense of rotation, one switches the rotation axes between the $\pm y$ direction. In this case, the propagator for a basic sequence element of two pulses with alternating sense of rotation is

$$U_{\pm\delta} = e^{iS_y(\pi+\delta)} e^{iS_y(\pi-\delta)} = \mathbf{1},$$

independent of the amplitude error δ.

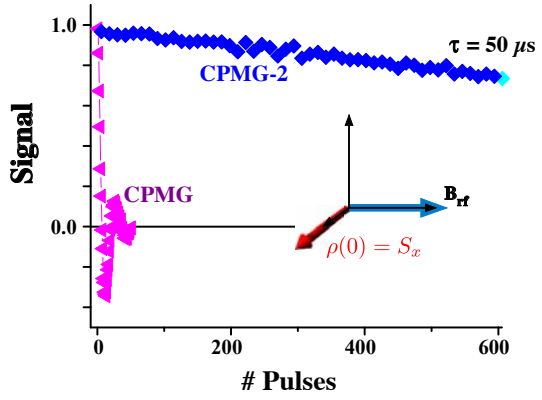

Figure 7.44: Comparison of the effect of different decoupling sequences.

Fig. 7.44 compares the performance of this compensated decoupling sequence (CPMG-2) with that of the standard uncompensated sequence (CPMG). Clearly, the compensated sequence reduces the decay rate by approx. two orders of magnitude.

For real-world operation, we have to look for gate operations that work reliably also if the precision of the experimental control fields is finite. Here, we discuss two possible approaches: first we show that it is possible to replace individual refocusing pulses by compensated pulses that implement very precise inversions, and then we discuss sequences that are inherently robust, i.e. insensitive to the imperfections of the individual pulses.

The simplest approach to make a sequence robust is by replacing every standard pulse by a robust composite pulse. The general approach to composite pulses was discussed in section 7.5.1. In the context of dynamical decoupling, we specifically need π-pulses that are robust against flip angle errors and frequency offset errors.

A composite π-pulse that is quite effective in compensating these errors simultaneously is the sequence

$$(\pi)_{\pi/6+\phi} - (\pi)_{\phi} - (\pi)_{\pi/2+\phi} - (\pi)_{\phi} - (\pi)_{\pi/6+\phi}. \tag{7.29}$$

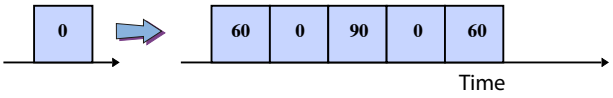

Figure 7.45: Robust inversion pulse.

If the 5 π-pulses are ideal, the sequence implements a π rotation around the ϕ-axis followed by a $-\pi/3$ rotation around the z axis. The phases are chosen such that errors cancel and do not change this overall rotation to first order. A comparison between this pulse and a normal rectangular pulse was shown in section 7.5.1. If the DD pulses are replaced by such pulses, the sequence becomes significantly more robust against pulse imperfections.

7.6.6 Robust DD

An alternative to the use of composite pulses consists in making the decoupling sequences fault-tolerant without compensating the error of each pulse, but by designing them in such a way that the error introduced by one pulse is compensated by the other pulses of the cycle. The first demonstration of this possibility is due to Maudsley [103], who noticed that sequences of identical pulses performed well for the longitudinal initial condition, but not for the transverse one. He suggested to alternate the phase of the π-pulses between x and y.

As shown in Fig. 7.46, the performance of this symmetrized sequence is independent of the initial condition. Various schemes are known for further improving the performance of this sequence.

A closely related approach is based on the robust π-pulse (7.29), which consists only of π-pulses. If the segments (i.e. individual $\pi-$pulses) are not applied directly after each other, but with equidistant delays

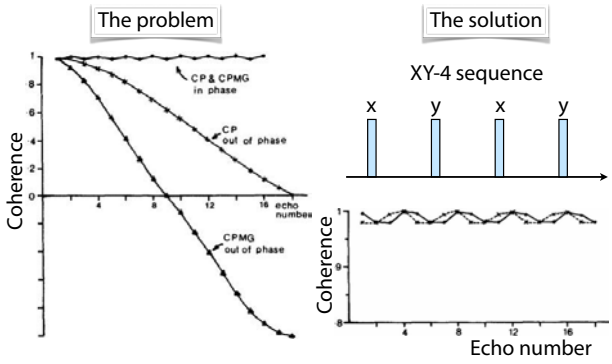

Figure 7.46: Comparison of the effect of different decoupling sequences.

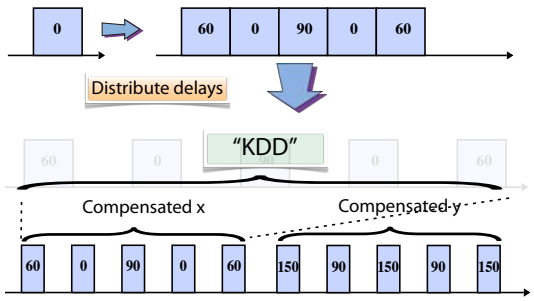

Figure 7.47: Robust DD sequence.

between them, one obtains a robust DD sequence

$$KDD_\phi = f_{\tau/2}(\pi)_{\pi/6+\phi} f_\tau(\pi)_\phi f_\tau(\pi)_{\pi/2+\phi} f_\tau(\pi)_\phi f_\tau(\pi)_{\pi/6+\phi} f_{\tau/2},$$

where f_τ describe delays of duration τ. The self-correcting sequence is created by combining 5-pulse blocks shifted in phase by $\pi/2$, such as [KDD$_\phi$ - KDD$_{\phi+\pi/2}$]^2, where the lower index gives the overall phase of the block. The cyclic repetition of these 20 π-pulses is referred to as the KDD sequence [87].

Figure 7.48 shows the overall error generated by a decoupling sequence where the individual pulses suffer from flip angle errors as well as offset errors. Without considering the effect of the environment, it shows the fidelity F after applying 100 pulses to

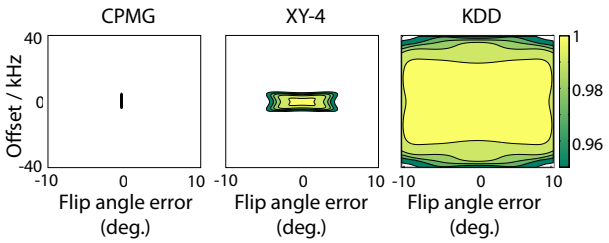

Figure 7.48: Error tolerance of different DD sequences: the uncompensate CPMG sequence and the compensated sequences XY-4 and KDD. The resulting fidelity for a sequence of 100 pulses is shown by color codes. The regions where the fidelity is lower than 0.95 are shown in white.

the system as a function of the two error parameters. Each panel contains the color-coded fidelity for one decoupling sequence. The best performance is achieved by the KDD sequence, whose cycle consists of 20 pulses.

Fig. 7.49 compares the experimental performance of different self-correcting sequences. The performance of the CDD sequences always saturates or decreases with increasing duty cycle under these experimental conditions. However, instead of saturating, the relaxation time for the KDD sequence continues to increase, as in the case of sequences with robust pulses. The KDD sequence combines the useful properties of robust sequences with those of sequences of robust pulses and can thus be used for both quantum computing and state preservation.

Dynamical decoupling is becoming a standard technique for preserving the coherence of quantum mechanical systems, which does not need control over

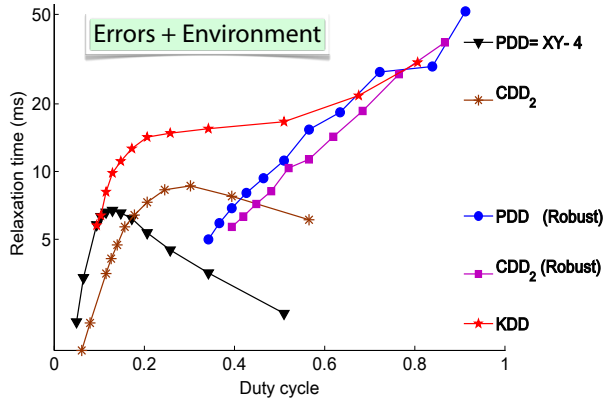

Figure 7.49: Experimental decoherence times for different compensated DD sequences as a function of the duty cycle. Experiments were done with nuclear spin qubits subjected to noise from an environment consisting of a nuclear spin bath.

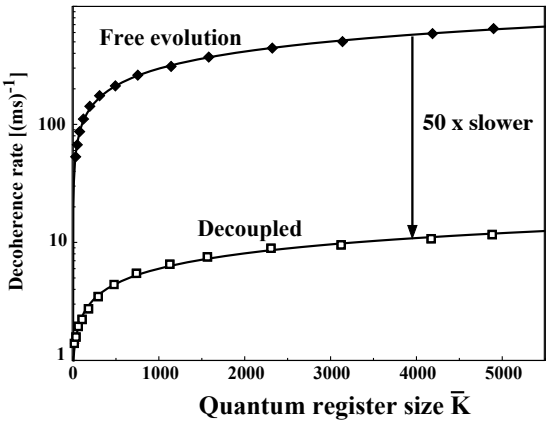

Figure 7.50: Scaling of decoherence rates with the number of qubits in the quantum register.

the environmental degrees of freedom.

7.6.7 DD for large systems

Of course, such measures for fighting decoherence become more important in large systems. We therefore have to test them also on systems with many qubits.

Fig. 7.50 shows an experimental test. The filled squares represent experimentally measured decoherence rates as a function of the number of correlated qubits, while the curve is a fit of the experimental data to a power law, $\propto K^{0.48}$. The upper part of the figure represents similar data as that in section 7.2.8.

As shown in Fig. 7.50 by the lower curve and data points, a suitable decoupling sequence allows one to reduce the decoherence rate by approximately a factor 50. The lower curve, labeled "Decoupled" has almost the same dependence on the number of qubits ($\propto K^{0.43}$), indicating that the decoupling works just as well for "large" quantum systems consisting of many correlated qubits, as for individual spins.

Further reading

Decoherence is discussed in many sources dealing with fundamental issues of quantum mechanics, such as the measurement problem and the quantum-classical boundary. In the present context Leggett's summer school lecture notes [63] are particularly useful. A compact and clear reference on quantum error correction is [104]; [33] discusses the topic in much more detail and from a more general perspective, with many references to original research articles. Preskill's lecture notes [30] also contain an in-depth discussion, pointing out relations to classical error-correcting codes. A review on decoherence-free subspaces and related topics is [105] and [106] reviews dynamical decoupling techniques, with an emphasis on robust sequences.

Problems and Exercises

1. Write the projector onto the general single-qubit state $\alpha|0\rangle + \beta|1\rangle$ as a linear combination of $\mathbf{1}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$.