Randomized Dynamical Decoupling Strategies and Improved One-Way Key Rates for Quantum Cryptography

Vom Fachbereich Physik
der Technischen Universität Darmstadt
zur Erlangung des Grades
eines Doktors der Naturwissenschaften
(Dr. rer. nat.)

genehmigte Dissertation von
Dipl.-Phys. Oliver Kern
aus Mainz

Darmstadt 2009
D17

Referent:	Prof. Dr. G. Alber
Korreferent:	Prof. Dr. J. Berges
Tag der Einreichung:	29.01.09
Tag der Prüfung:	25.05.09

Randomized Dynamical Decoupling Strategies and Improved One-Way Key Rates for Quantum Cryptography

Abstract

The present thesis deals with various methods of quantum error correction. It is divided into two parts. In the first part, dynamical decoupling methods are considered which have the task of suppressing the influence of residual imperfections in a quantum memory. Such imperfections might be given by couplings between the finite dimensional quantum systems (qudits) constituting the quantum memory, for instance. The suppression is achieved by altering the dynamics of an imperfect quantum memory with the help of a sequence of local unitary operations applied to the qudits. Whereas up to now the operations of such decoupling sequences have been constructed in a deterministic fashion, strategies are developed in this thesis which construct the operations by random selection from a suitable set. Formulas are derived which estimate the average performance of such strategies. As it turns out, randomized decoupling strategies offer advantages and disadvantages over deterministic ones. It is possible to benefit from the advantages of both kind of strategies by designing combined strategies. Furthermore, it is investigated if and how the discussed decoupling strategies can be employed to protect a quantum computation running on the quantum memory. It is shown that a purely randomized decoupling strategy may be used by applying the decoupling operations and adjusted gates of the quantum algorithm in an alternating fashion. Again this method can be enhanced by the means of deterministic methods in order to obtain a combined decoupling method for quantum computations analogously to the combining strategies for quantum memories.

The second part of the thesis deals with quantum error-correcting codes and protocols for quantum key distribution. The focus is on the BB84 and the 6-state protocol making use of only one-way communication during the error correction and privacy amplification steps. It is shown that by adding additional errors to the preliminary key (a process called noisy preprocessing) followed by the use of a structured block code, higher secure key rates may be obtained. For the BB84 protocol it is shown that iterating the combined preprocessing leads to an even higher gain. In order to speed up the numerical evaluation of the key rates, results of representation theory come into play. If a coherent version of the protocol is considered, the block code used in the preprocessing stage becomes a concatenated stabilizer code which is obtained by concatenating an outer random code with an inner deterministic one. This concatenated stabilizer code is used to compute an improved lower bound on the quantum capacity of a certain quantum channel (the so-called qubit depolarizing channel).

Zufallsbasierte dynamische Entkopplungsmethoden und verbesserte Schlüsselraten für die Quantenkryptographie mit Einwegkommunikation

Zusammenfassung

Die vorliegende Arbeit befaßt sich mit verschiedenen Methoden der Quantenfehlerkorrektur. Sie ist in zwei Teile gegliedert. Im ersten Teil werden dynamische Entkopplungsmethoden betrachtet, welche die Aufgabe haben, den Einfluß verbleibender Unvollkommenheiten in einem Quantenspeicher zu unterdrücken. Solche Unvollkommenheiten sind z. B. gegeben durch Kopplungen zwischen den einzelnen endlichdimensionalen Quantensystemen (Qudits), welche zusammen den Quantenspeicher bilden. Um die Unterdrückung zu realisieren, wird die Dynamik eines fehlerbehafteten Quantenspeichers mit Hilfe einer Sequenz von lokalen unitären Operationen, die auf die einzelnen Qudits angewandt werden, modifiziert. Während die Operationen einer solchen Entkopplungssequenz bislang deterministisch ausgewählt wurden, werden in dieser Arbeit Strategien entwickelt, welche die Operationen durch zufällige Auswahl aus einer geeigneten Menge bestimmen. Es werden Formeln hergeleitet, welche die mittlere Leistung solcher Strategien abschätzen. Dabei zeigt sich, daß die zufallsbasierten dynamische Entkopplungsstrategien gegenüber den deterministischen Vor- und Nachteile bieten. Es ist möglich von den Vorteilen beider Arten von Strategien zu profitieren, indem man geeignete kombinierte Strategien entwickelt. Weiterhin wird untersucht, inwiefern sich die diskutierten Entkopplungsstrategien einsetzen lassen, um eine auf dem Quantenspeicher laufende Quantenrechnung zu schützen. Es wird gezeigt, daß sich eine rein zufallsbasierte Entkopplungsmethode verwenden läßt, indem speziell angepaßte Gatter des zu rechnenden Quantenalgorithmus und Entkopplungsoperationen abwechselnd angewandt werden. Diese Methode läßt sich wiederum mittels deterministischer Verfahren erweitern um analog zu den kombinierten Enkopplungsstrategien für Quantenspeicher kombinierte Entkopplungsmethoden für Quantenrechner zu erhalten.

Im zweiten Teil der Arbeit geht es um quantenfehlerkorrigierende Codes und quantenkryptographische Protokolle. Es wird das BB84- und das 6-State-Protokoll zur sicheren Schlüsselverteilung unter Verwendung von Einwegkommunikation während der Fehlerkorrektur und Privatspärhenverstärkung betrachtet. Es wird gezeigt, daß durch das nachträgliche Hinzufügen von Fehlern im vorläufigen Schlüssel („noisy preprocessing“) in Verbindung mit der Nutzung eines bestimmten Blockcodes höhere Schlüsselraten erzielt werden können. Für das BB84-Protokoll wird weiter gezeigt, daß sich die erzielten Vorteile verstärken lassen, falls das kombinierte „preprocessing“ iterativ verwendet wird. Die numerische Berechnung der jeweiligen Schlüsselraten wird dabei durch das Verwenden von Resultaten der Darstellungstheorie beschleunigt. Bei einer kohärenten Betrachtung der Protokolle entspricht der verwendete Blockcode einem verketteten Stabilizer-Code, bei dem ein äußerer zufälliger Code mit einem deterministischen inneren Code verkettet wird. Mittels dieses verketteten quantenfehlerkorrigierenden Codes wird eine verbesserte untere Schranke für die Quantenkapazität eines bestimmten Quantenkanals (genannt „qubit depolarizing channel“) berechnet.

Contents

1 Introduction and Preliminaries

1.1 Introduction and Outline

Quantum mechanics is a theory which appears rather counterintuitive: For instance, certain variables of a quantum mechanical system like position and momentum cannot both be determined with arbitrary accuracy, particles might penetrate a barrier (tunnel effect), and cats might be dead and alive at the same time [Sch35]. Quantum information theory tries to generalize classical information theory to the quantum world by considering a quantum mechanical two-level system (a qubit) as basic information carrier. It turns out that the properties of quantum information, i. e. the information encoded in a quantum system, are in strong contrast to the properties we now about classical information: While classical information can be copied perfectly (resulting in the enormous success of file sharing networks), quantum information, in general, cannot be copied (no cloning theorem [Die82; WZ82]). Although it cannot be duplicated, quantum information may be teleported [BBC$^+$93] by using distributed entangled states as a resource. During the last two and a half decades the idea emerged to use quantum mechanics to implement technical applications which might not exist in a purely classical world. Two particularly important concepts are quantum computing (promising faster computation) and quantum cryptography (promising unconditionally secure communication).

The first example of a quantum algorithm that is more efficient than any possible classical algorithm is the Deutsch-Jozsa algorithm [DJ92]. Given a black box quantum computer known as an oracle that implements a binary function which is either constant or balanced, the algorithm is able to determine if the function is constant or balanced by using the oracle only once. By contrast, a classical computer might have to use a corresponding classical oracle on more than half the input values in the worst case. The reason behind this speedup is the quantum parallelism which arises from the ability of a quantum memory to exist in a coherent superposition of states. While the Deutsch-Jozsa algorithm is merely of academic interest, the potential of quantum computing drew lots of attention in 1994 when Shor presented polynomial-time algorithms for prime factorization and discrete logarithms [Sho94] due to their potential to break current cryptosystems: Nearly all current cryptosystems can be divided into two families. One family is based on the assumption that an efficient algorithm for prime factorization does not exist (an example is the RSA public key encryption protocol [RSA78]), while the other one is based on the assumption that computation of the discrete logarithm is hard (examples are the Diffie-Hellman key distribution protocol [DH76] and the Elgamal public key encryption protocol [Elg85]). The security of such cryptosystems is then provided by the fact that an eavesdropper with limited computational power is unable to solve these hard problems. With Shor's algorithms the only reason that current cryptosystems can still be considered save is the tremendous difficulty to build a working quantum computer. While the factorization of the number $15 = 3 \times 5$ has been demonstrated on an nuclear magnetic resonance (NMR) quantum computer consisting of 7 qubits [VSB$^+$01], factorization of a 1024 bit number requires about 2000 qubits [PZ03] and is completely out of range of current technology.

The main obstacle in the realization of a quantum computer is a process called decoherence which arises from an interaction of the quantum information carriers with the environment and which destroys any coherent superpositions on a very short time scale. But even if the quantum computer could be perfectly isolated, it still has to be accessible to perform manipulations (quantum gates) with very high accuracy. In addition, any imperfections and interactions between the finite-dimensional quantum information carriers (qudits) constituting the quantum memory of the quantum computer tend also

to shorten the time scale of reliable quantum computation. A way out might be the use of quantum error-correcting techniques: One technique is to employ quantum error-correcting codes, introduced by Shor in 1995 [Sho95]. By encoding the quantum information in a subspace of the total available space, they allow a recovery step involving a syndrome measurement to reverse certain decoherence processes. Another technique is called dynamical decoupling [Zan99; VKL99]. It alters the dynamics of a quantum memory by applying a series of local unitary operations to the qudits. If the couplings to the environment effectively cancel out in the resulting dynamics, decoherence is suppressed. This technique is inspired by refocusing techniques in NMR spectroscopy [EBW87].

While on the one hand quantum mechanics seems to question the security of established classical cryptosystems, on the other hand it provides key distribution protocols whose security does not rely on any unproven assumption but is guaranteed by the validity of quantum mechanics itself. The idea of quantum key distribution (QKD) is due to Bennett and Brassard who, inspired by a paper of Wiesner written in the late 60's and not accepted for publication until 1983 [Wie83], invented the first QKD protocol (now called BB84 protocol) in 1984 [BB84]. To establish a secret key between two distant parties connected via a quantum channel and an authenticated classical channel, a QKD protocol demands one party (usually called Alice) to send non-orthogonal quantum states to the other party (usually called Bob). The no cloning theorem [Die82; WZ82] prevents an eavesdropper with full access to the quantum channel from copying these states in a perfect manner. Hence any action of an eavesdropper leaves traces in the transmitted states which can be recognized by Alice and Bob by comparing measurement and preparation data of a subset of randomly chosen check states. Thereby, one takes the conservative point of view that any noise in the channel is caused by an eavesdropper. As long as the action of an eavesdropper seems harmless enough (i. e. as long as the detected error rate is low enough), Alice and Bob should be able to generate a random, correct and secure key from their raw data. A security proof for a QKD protocol typically gives a lower bound on the length of the secret key that can be obtained from the raw data for a given error rate. While it is still in question whether it will ever be possible to build large-scale quantum computers, the first generation of QKD systems is already available commercially [QKD]. The reason that a QKD system is so much easier to build than a quantum computer is that it shares only the need to prepare and measure quantum systems but does not need to store and manipulate them. Since the key rates of such devices are typically low, it is important to find security proofs which allow the secure key generation rate for a given error rate to be as high as possible.

The goal of this thesis is to develop improved dynamical decoupling techniques in order to contribute to the field of quantum computing and to provide improved secret key generation rates for quantum key distribution protocols in order to contribute to the field of quantum cryptography. To achieve this goal, the thesis deals with different kinds of quantum error correction: Part I of the thesis studies the potential of randomized dynamical decoupling strategies which are able to stabilize a quantum memory and even a running quantum computation against residual imperfections and interactions. Part II of the thesis considers quantum error-correcting codes which are used to compute improved lower bounds on the capacity of the qubit depolarizing channel. Furthermore, these codes are used to obtain improved secret key rates for variants of QKD protocols like the BB84 protocol [BB84] and the 6-state protocol [Bru98]. The later protocol is a natural extension of BB84, which makes use of four different quantum states, and makes use of two additional quantum states. A more detailed introduction and outline is given in the following subsections.

1.1.1 Part I: Random Decoupling

In order to use a quantum system as a quantum memory, or even more demanding, to use it for quantum computations, we must be able to apply some kind of control. Let us assume that the quantum system is a quantum register formed by a set of qudits. Usually the experimentally easiest kind of control

is to apply single qudit gates realized by a local control Hamiltonian. Dynamical control of a local Hamiltonian allows the time evolution of a quantum system to be modified. A well known example is given by the refocusing techniques used to manipulate nuclear spin Hamiltonians [EBW87]. There are various possible controls tasks: For example, for a closed quantum system, we might want to simulate a time evolution according to a Hamiltonian which is different from the system Hamiltonian [WRJB02a; BDNB04]. In particular, we might want to simulate a vanishing Hamiltonian, a task we call decoupling from now on. For an open quantum system, we might try to suppress decoherence by simulating vanishing couplings with the environment [Zan99; VKL99], or we might try to generate at least a noiseless subsystem [Zan00; VKL00]. In the simplest control scenario, the so-called bang-bang control scenario, the local control Hamiltonian generates a set of pulses belonging to a suitable control scheme instantaneously. Then, for all control tasks, the fundamental deterministic control strategy is to apply the pulses belonging to the control scheme in a cyclic manner over and over again. The control scheme is designed in such a way that, in lowest order average Hamiltonian theory (AHT) [EBW87], the resulting dynamics corresponds to the Hamiltonian to be simulated. Assuming the pulses to be ideal, the finite time interval Δt in between subsequent pulses is the only obstacle preventing a control task to be achieved in a perfect manner.

One of the goals of part I of this thesis is to devise and analyze improved control strategies which lead to a better performance for a fixed time interval Δt, or in other words, which lead to a suppression of the residual higher order terms in AHT. Let us focus for now on the control task of decoupling a closed quantum system. In the context of decoupling, a control scheme is said to be a decoupling scheme. As we will see, the fundamental deterministic control strategy leads to an average fidelity decay which is quadratic in time. Thereby, the strength of the decay is determined by the strength of the system Hamiltonian, by the size of the decoupling scheme, and by the time interval Δt in between subsequent pulses. For example, an improved strategy commonly used by the NMR community is a symmetrized version of the fundamental strategy: In spite of doubling the size of the decoupling scheme, it leads to a decrease of the strength of the fidelity decay but keeps its quadratic-in-time nature. An interesting result, first observed in the author's diploma thesis [Ker04] (see also [KAS05]), is that a control strategy based on random selection of the elements of a decoupling scheme leads to a fidelity decay which is only linear in time. Subsequently, randomized decoupling was proposed for open quantum systems by Viola and Knill [VK05], who confirmed the linear-in-time decay by constructing a strict lower bound for the worst case fidelity [VK05; Vio05]. Meanwhile control strategies combining the advantages of purely deterministic and randomized strategies have been devised by the author [KA05] and by Santos and Viola [SV06; VS06].

Since, in general, the pulses of a decoupling sequence interfere with the application of an additional Hamiltonian implementing a quantum gate, the protection of a running quantum computation against imperfections of the quantum memory is not straightforward [VLK99]. Another goal of this thesis is to study how the devised decoupling strategies might be used in order to protect quantum computations. In the bang-bang control scenario, one option for deterministic strategies is to apply quantum gates instantaneously in between completed decoupling cycles. Under the more realistic assumption that quantum gates (especially two-qudit gates) are generated within a finite time interval by the means of bounded controls, more advanced techniques are required in order to combine decoupling and computation. For instance, the dynamically corrected gate (DCG) of Khodjasteh and Viola [KV09] combines a single decoupling cycle with the generation of a quantum gate. It turns out that the decoupling pulses of a randomized decoupling strategy can be alternated with especially adjusted quantum gates, a method which was called Pauli random error correction (PAREC) by the author and collaborators [KAS05]. In order to benefit from the advantages of both methods, DCGs might be combined with the PAREC method. Another scenario arises if the two-qudit gates of a quantum computer are generated by the couplings between adjacent qudits. In this case a selective decoupling method is used which switches

off all but the desired coupling. The fundamental selective decoupling strategy can be improved by combining it with a randomized decoupling strategy.

Outline

Chapter 2: Dynamical Decoupling. Chapter 2 deals with dynamical decoupling strategies for quantum memories in the bang-bang control scenario. In order to improve the fundamental deterministic decoupling strategy, new randomized strategies are considered. The performance of these strategies is analyzed by (i) deriving formulas expressing the average fidelity and (ii) by considering the variance of the fidelity. The chapter closes with a numerical simulation of any strategy on a quantum memory perturbed by Heisenberg interactions. The idea of the embedded decoupling strategy was published in [KA05]:

> O. Kern and G. Alber.
> Controlling Quantum Systems by Embedded Dynamical Decoupling Schemes.
> *Phys. Rev. Lett.*, **95**(25), 250501 (2005). arXiv:quant-ph/0506038v1.

Chapter 3: Decoupling and Computation. This chapter focuses on the fundamental problem of combining dynamical decoupling and quantum computation. Here, we allow the quantum gates as well as the decoupling pulses to be generated within a finite time interval. After presenting an overview of known results, the PAREC method is proposed, which is based on alternating the decoupling pulses of a randomized decoupling strategy with specially adjusted quantum gates forming the quantum algorithm. We derive a formula for the fidelity decay of a quantum computation perturbed by static imperfections with and without the PAREC method. It is shown that the PAREC method is a realization of an idea of Prosen and Žnidarič [PŽ01], who proposed to stabilize a quantum computation against static imperfections by increasing the decay of the correlation function measuring the fidelity decay. Eventually, we consider the dynamically corrected gates (Euler-DCGs) of Khodjasteh and Viola [KV09] which correspond to an implementation of a deterministic decoupling strategy for the purpose of computation. We propose to implement the PAREC method by using only Euler-DCGs in order to benefit from the advantages of both methods. Some of the results of this chapter have already been published. The PAREC method together with numerical evidence was already devised in the author's diploma thesis [Ker04] and has been published in [KAS05]:

> O. Kern, G. Alber, and D. L. Shepelyansky.
> Quantum error correction of coherent errors by randomization.
> *Eur. Phys. J. D*, **32**(1), 153–156 (2005). arXiv:quant-ph/0407262v1.

The comparison of the PAREC method with the idea of Prosen and Žnidarič together with a formula for the average fidelity for the special case of instantaneous gates and decoupling pulses was given in [GKAJ08]:

> D. Geberth, O. Kern, G. Alber, and I. Jex.
> Stabilization of quantum information by combined dynamical decoupling and detected-jump error correction.
> *Eur. Phys. J. D*, **46**(2), 381–394 (2008). arXiv:0712.1480v1.

Chapter 4: Selective Recoupling and Randomized Decoupling. Instead of implementing a two-qudit quantum gate with the help of an external gate Hamiltonian, a quantum computer might use existing inter-qudit couplings. Now, a non-operation is implemented by using a decoupling scheme which effectively switches off all couplings. To implement a certain two-qudit gate, a selective decoupling (or selective recoupling) scheme is employed which removes all but the desired coupling. By drawing on a

particular example, this chapter shows how a selective decoupling strategy can be improved by devising a combined selective decoupling strategy involving randomized decoupling. While a corresponding combined decoupling strategy can be devised quite easily, the non-vanishing lowest order AHT term of the selective decoupling strategy makes things a bit more difficult. This chapter is a slightly enhanced version of [KA06]:

> O. Kern and G. Alber.
> Stabilizing selective recoupling schemes by randomization.
> *Phys. Rev. A*, **73**(6), 062302 (2006). arXiv:quant-ph/0602167v1.

Appendix A and B: Chapter A of the appendix contains some examples of difference schemes and orthogonal arrays. This data can be used to obtain decoupling schemes as explained in section 2.2. Chapter B explains how certain quantum maps can be implemented as quantum algorithms. Such quantum algorithms are used in chapters 3 and 4 as test algorithms for the numerical simulations of the PAREC method and the improved selective decoupling method, respectively. More detailed information on quantum maps and their implementation on a quantum computer can be found in the author's diploma thesis [Ker04].

1.1.2 Part II: Codes and Cryptography

One of the fundamental theorems in classical information theory is Shannon's noisy channel coding theorem [Sha48]. If classical information is to be transmitted over a classical noisy channel, Shannon's theorem assures that the transmission can be performed error-free as long as the transmission rate is below a maximum rate. This maximum rate is called the capacity of the channel. To achieve an error-free transmission, error-correcting codes have to be employed. It turns out that the full capacity of a channel can be achieved by using randomly constructed block codes. In quantum information theory, the analogous theorem is the quantum noisy channel coding theorem which states that quantum information can be sent reliably over a noisy quantum channel as long as the transmission rate is below the quantum capacity of the channel. Quantum information which is to be sent over a noisy quantum channel has to be encoded using quantum error-correcting codes. Surprisingly, it turns out that in contrast to the classical case, randomly constructed quantum codes do not achieve the full capacity of a quantum channel: By considering a concatenated quantum code obtained by encoding the information encoded by a random code one more time with a so-called cat code, Shor and Smolin showed that error-free transmission over the so-called qubit depolarizing channel becomes possible at a higher rate than achievable by the random code alone [SS96; DSS98]. In this thesis we extend these calculations to cat codes of larger size and obtain improved lower bounds on the capacity of the qubit depolarizing channel.

The quantum capacity of a noisy quantum channel has a close connection with the security of a quantum key distribution protocol. If the parties Alice and Bob are able to determine how the quantum channel (i. e. the eavesdropper) acts on the quantum states sent from Alice to Bob, they might use a quantum error-correcting code to transmit these states error-free, i. e. in such a way that the eavesdropper does not learn anything about them. A security proof of the BB84 protocol following this idea was given by Lo and Chau [LC99]. Unfortunately, in order to implement such a protocol, Alice and Bob have to be able to manipulate quantum states during the encoding and the recovery step. By making use of the special structure of a certain class of quantum codes, the so-called CSS codes [CS96; Ste96], Shor and Preskill showed that a protocol based on encoding the states is equivalent to the original prepare and measure protocol [SP00]. Hence, results on the achievable transmission rate over a certain type of quantum channels (so-called memoryless Pauli channels) with the help of CSS codes can be used to prove the security of certain QKD protocols up to a certain error rate. As discussed in

the previous paragraph, randomly constructed CSS codes give a lower bound on the obtainable secure key rate, but concatenation of such codes with deterministic ones leads to even better bounds.

Another way to improve the secret key rates is to add noise to the raw key bits before they are processed into the final key. Such a procedure is known as local randomization or noisy preprocessing and was discovered by Renner et al. [KGR05; RGK05]. A security proof of a QKD protocol involving noisy preprocessing is not so straightforward as the Lo-Chau or Shor-Preskill proof. The difficulty lies in the fact that a security proof based on perfect quantum error correction assures that Alice and Bob could in principle share perfectly entangled states, which are sufficient but not necessary for the generation of a secure key [HHHO05]. A more sophisticated security proof involving CSS codes and noisy preprocessing was given by Renes and Smith [RS07]. Recently it was shown by the same authors for BB84 that by combining both methods — noisy preprocessing and the use of the concatenated cat code — even higher secure key rates can be obtained [SRS08]. In this thesis it will be shown that these results can also be applied to the 6-state protocol. Furthermore, an iterated version of the combined preprocessing protocol is considered. In order to evaluate the formulas expressing the secure key rates efficiently, results from representation theory have to be used. In this context a matlab program was developed which calculates the Schur basis of the Hilbert space of n qudits of dimension q.

Outline

Chapter 5: Classical Error Correction This chapter provides an introduction to the theory of classical error-correcting codes. The main focus is on linear codes. A linear code with q^k codewords of length n is a k dimensional subspace of the space $\mathbb{F}_q^n$ containing all strings of length n with entries from the field $\mathbb{F}_q$. Shannon's noisy coding theorem is proven for the binary symmetric channel by using random linear codes and typical set decoding.

Chapter 6: Quantum Error-Correcting Codes We introduce the theory of quantum error-correcting codes. An important class of quantum codes are the so-called stabilizer codes which might be viewed as the linear codes of quantum error correction. A stabilizer code encoding k qudits of dimension q into n is characterized by a k dimensional self-orthogonal subspace (called stabilizer) in the space $\mathbb{F}_q^{2n}$ with respect to a symplectic inner product. We explain how a unitary encoding of such a code corresponds to an extension of a basis of the stabilizer to a hyperbolic basis of $\mathbb{F}_q^{2n}$. Then we specialize in the class of CSS codes, which form a subclass of stabilizer codes with a direct connection to classical linear codes, and show how the description of an encoding can be simplified. Finally, we discuss the concatenation of two stabilizer codes.

Chapter 7: Quantum Channel Capacity While Shannon's noisy coding theorem is one of the fundamental theorems of classical information theory, this chapter deals with the quantum analog of Shannon's noisy coding theorem. For a certain class of channels — so-called memoryless Pauli channels — coding theorems are proven which provide lower bounds on the capacity. These theorems use (i) random stabilizer codes, (ii) random CSS codes, and (iii) random stabilizer codes concatenated with deterministic inner ones for encoding and joint-typical set decoding to implement the recovery operation. The last theorem is used in connection with a specific deterministic inner code — a so-called cat code — to obtain better lower bounds on the capacity of the qubit depolarizing channel.

Chapter 8: Quantum Cryptography This chapter shows how the results of the combined preprocessing step for BB84 [SRS08] can be applied to the 6-state protocol. We make use of the detailed analysis of the concatenated cat code provided by chapters 6 and 7, and employ the security proof of Renner [Ren05, corollary 6.5.2]. In addition, for the BB84 protocol, an iterative version of this preprocessing

scheme is considered. It is explained how the secret key rates can be efficiently evaluated by using insights from representation theory. The chapter is an enhanced version of the following article [KR08]:

> O. Kern and J. M. Renes.
> Improved one-way rates for BB84 and 6-state protocols.
> *Quant. Inf. & Comp.*, **8**(8/9), 0756–0772 (2008). arXiv:0712.1494v2.

Appendix C and D: Chapter C of the appendix contains some technical results mainly concerning error-correcting codes. Chapter D explains the eigenfunction method [CPW02] which can be used to obtain a computer program calculating the Schur transform. The Schur transform is a unitary transformation relating the standard computational basis of n qudits of dimension q with the Schur basis associated with the representation theory of the symmetric group S_n and the general linear group GL_q.

1.2 Preliminaries

The understanding of this thesis requires the knowledge of basic quantum mechanics and basic representation theory. In addition, the theory of error-correcting codes comes into play in part II. This section provides a brief overview of the necessary fundamentals of classical information theory, quantum mechanics and representation theory. An introduction to error-correcting codes will be given in chapter 5. An overview over classical information theory can be found in the book of MacKay [Mac03]. For an introduction to quantum mechanics we refer to the two books of Cohen-Tannoudji et al. [CTDL77]. A comprehensive introduction to quantum computation and quantum information can be found in the book of Nielsen and Chuang [NC00], which contains also a brief introduction to quantum mechanics and classical information theory. In addition we refer to the lecture notes of Preskill [Pre98]. An introduction to group representation theory can be found in the book of Tung [Tun85].

1.2.1 Probabilities and Entropy

A discrete random variable X is characterized by a set of outcomes $A = (a_1, \ldots, a_s)$ ($s = |A|$) together with an associated probability distribution $P = (p_1, \ldots, p_s)$ such that X takes the values $a_i \in A$ with probability $\Pr(X = a_i) = p_i$. The probabilities p_i are non-negative numbers which sum up to one. The uncertainty of the outcome a random variable is characterized by the Shannon entropy of its probability distribution.

Definition 1.2.1 (Shannon entropy)**.** The q-ary Shannon entropy of a discrete probability distribution $P = (p_1, \ldots, p_s)$ is defined as

$$H_{s[\log_q]}(P) = -\sum_{i=1}^{s} p_i \log_q p_i, \tag{1.1}$$

where the value of $0 \log_q 0$ is taken to be 0, which is consistent with the limit $\lim_{p\to 0} p \log_q p = 0$.

Alternatively, we might say that the entropy of the random variable X is given by

$$H_{s[\log_q]}(X) = -\sum_{a\in A} \Pr(a) \log_q \Pr(a). \tag{1.2}$$

If the logarithm is taken to the base 2 the entropy is expressed in bits. Otherwise we stress such a fact by denoting the base b as $H_{[\log_b]}$. If a discrete probability distribution P consists of s elements, we write $H_s(P)$ to indicate the number of summands. For $s = 2$ it is sufficient to denote the first element of a probability distribution $P = (p, 1-p)$, i. e. we write $H_2(p) \equiv H_2(P) = H_2(p, 1-p)$.

Let us consider an additional random variable Y which is characterized by the set of outcomes $B = (b_1, \ldots, b_r)$ ($r = |B|$) and the probability distribution $Q = (q_1, \ldots, q_r)$. Then the conditional entropy of X given Y is defined as

$$\begin{aligned} H_{[\log_q]}(X|Y) &= \sum_{b\in B} \underbrace{H_{[\log_q]}(X|Y=b)}_{-\sum_{a\in A} \Pr(a|b)\, \log_q\, \Pr(a|b)} \cdot \Pr(b) \\ &= -\sum_{b\in B, a\in A} \Pr(a,b) \log_q \Pr(a|b). \end{aligned} \tag{1.3}$$

If X and Y are independent random variables, i. e. if $\Pr(a_i, b_j) = \Pr(a_i)\Pr(b_j) = p_i q_j$, it follows that $H_{[\log_q]}(X|Y) = H_{[\log_q]}(X)$. For general X and Y the relation $H_{[\log_q]}(X|Y) = H_{[\log_q]}(X,Y) - H_{[\log_q]}(Y)$ can be shown to hold, where $H_{[\log_q]}(X,Y)$ denotes the joint entropy of X and Y:

$$H_{[\log_q]}(X,Y) = -\sum_{b\in B, a\in A} \Pr(a,b) \log_q \Pr(a,b). \tag{1.4}$$

Definition 1.2.2 (Mutual information)**.** The mutual information of two discrete random variables X and Y is defined as

$$I_{[\log_q]}(X:Y) = H_{[\log_q]}(X) + H_{[\log_q]}(Y) - H_{[\log_q]}(X,Y). \tag{1.5}$$

It is easy to verify the relations $I_{[\log_q]}(X:Y) = H_{[\log_q]}(X) - H_{[\log_q]}(X|Y) = H_{[\log_q]}(Y) - H_{[\log_q]}(Y|X) = I_{[\log_q]}(Y:X)$. Hence the mutual information measures how much the uncertainty of X is reduced when Y is known (and vice versa). The mutual information is always non-negative and 0 if and only if X and Y are independent variables.

For prime q the Galois field $\mathbb{F}_q$ contains the numbers $0, 1, \ldots, q-1$ and addition and multiplication are performed modulo q. The vector space $\mathbb{F}_q^n$ contains the q^n vectors $(0,0,\ldots,0), (0,0,\ldots,1), \ldots, (q-1, q-1, \ldots, q-1)$ of length n with entries from $\mathbb{F}_q$. Note that $\mathbb{F}_q^n$ forms a group with respect to addition modulo q.

Definition 1.2.3. The Hamming distance $\mathrm{dist}(\vec{x}, \vec{y})$ between two vectors $\vec{x}, \vec{y} \in \mathbb{F}_q^n$ is defined as the number of places in which the two vectors differ. The Hamming weight $\mathrm{wt}(\vec{x})$ of a vector $\vec{x} \in \mathbb{F}_q^n$ is defined as the Hamming distance between $\vec{x}$ and the null vector $\vec{0} = (0, \ldots, 0)$.

We close this subsection proving the Chernoff bound for binomial distributions which will be used frequently in part II of the thesis to obtain asymptotic bounds. The proof is taken from the book [Rom92].

Lemma 1.2.1 (Chernoff bound)**.** *Let Y be a random variable which follows a (n,p) binomial distribution, i. e. $\Pr(Y=k) = \binom{n}{k} p^k (1-p)^{n-k}$. Then, for any $\lambda < p$ such that $n\lambda \in \mathbb{N}_0$,*

$$\Pr(Y \le n\lambda) \le \left(\frac{p}{\lambda}\right)^{\lambda n} \left(\frac{1-p}{1-\lambda}\right)^{n(1-\lambda)}. \tag{1.6}$$

Proof. Let us define the random variable $X = e^{tY}$ with $t < 0$. Since X takes only positive values, the Markov bound applies:

$$\Pr(X \ge a) \le \langle X \rangle / a. \tag{1.7}$$

It follows that the probability of Y taking on a value less than b is upper bounded by

$$\Pr(Y \le b) = \Pr(X \ge e^{tb}) \le \langle X \rangle / e^{tb}. \tag{1.8}$$

Plugging the expectation value of X,

$$\langle X \rangle = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot e^{tk} = (pe^t + 1 - p)^n, \tag{1.9}$$

into the upper bound for $\Pr(Y \leq b)$ leads to

$$\Pr(Y \leq b) = \sum_{k=0}^{b} \binom{n}{k} p^k (1-p)^{n-k} \leq (pe^t + 1 - p)^n \cdot e^{-tb}. \tag{1.10}$$

Let us set $\lambda = b/n$ now. The right hand side is minimized for $e^t = \frac{1-p}{p}\frac{\lambda}{1-\lambda}$ which lies in $[0,1]$ if $\lambda < p$. □

For $p = 1/2$ the Chernoff bound leads to the tail inequality (see e. g. [Wel88, section 3.5]):

Corollary 1.2.2 (Tail inequality)**.** *For any λ, with $0 \leq \lambda < 1/2$ and $n\lambda \in \mathbb{N}_0$,*

$$\sum_{k=0}^{\lambda n} \binom{n}{k} \leq \lambda^{-\lambda n}(1-\lambda)^{-n(1-\lambda)} = 2^{nH_2(\lambda)}. \tag{1.11}$$

1.2.2 Quantum Mechanics

The state of a quantum mechanical system S is represented by a density operator ρ which is a non-negative operator of trace one acting on the associated Hilbert space $\mathcal{H}_S$ of the system. We denote the set of operators as $\mathcal{L}(\mathcal{H}_S)$ and the subset of density operators as $\mathcal{S}(\mathcal{H}_S)$. A state ρ is said to be a pure state if $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle \in \mathcal{H}_S$ such that $\langle\psi|\psi\rangle = 1$.

Time Evolution and Measurements

The time evolution of a pure quantum state is specified by the Schrödinger equation,

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle, \tag{1.12}$$

where $H(t) \in \mathcal{L}(\mathcal{H}_S)$ denotes the self-adjoint Hamiltonian of the system. Correspondingly, the time evolution of a general quantum state ρ is described by the von Neumann equation,

$$i\hbar \frac{d}{dt}\rho(t) = [H(t), \rho(t)], \tag{1.13}$$

where the brackets denote a commutator, i. e. $[A, B] = AB - BA$. As a consequence, the time evolution operator of a closed quantum system is unitary,

$$\rho(t) = U(t,0)\rho(0)U^\dagger(t,0), \tag{1.14}$$

with

$$U(t,0) = \mathcal{T}\exp\Big(-i\int_0^t H(t')dt'\Big), \tag{1.15}$$

where $\mathcal{T}$ denotes the Dyson time-ordering operator.

If the quantum system S forms a part of a larger quantum system, S is said to be an open quantum system. Then the resulting time evolution of the open system alone is not necessarily unitary anymore, but is given by a trace-preserving completely positive map (tpcp-map) $\mathcal{E} : \mathcal{S}(\mathcal{H}_S) \to \mathcal{S}(\mathcal{H}_S)$. Any tpcp-map $\mathcal{E}$ can be represented in terms of an operator sum decomposition $\{E_\mu\}$ such that $\sum_\mu E_\mu^\dagger E_\mu = \mathcal{I}$ and

$$\mathcal{E} : \rho \mapsto \mathcal{E}(\rho) = \sum_\mu E_\mu \rho E_\mu^\dagger, \tag{1.16}$$

where $\mathcal{I}$ denotes the identity operator.

A von Neumann measurement is characterized by a self-adjoint measurement operator M with spectral decomposition $M = \sum_\mu m_\mu P_\mu$, where the m_μ denote distinct measurement values and the P_μ denote

orthogonal projections ($\sum_\mu P_\mu = \mathcal{I}$). If we perform a measurement of M on the state ρ, we obtain the result μ with probability $p_\mu = \mathrm{tr}(P_\mu \rho)$. Conditioned on the measurement result the state changes from ρ to $P_\mu \rho P_\mu / \mathrm{tr}(P_\mu \rho)$. A more general measurement is specified by a positive operator valued measure (POVM), which consists of a set $\{F_\mu\}$ of positive operators such that $\sum_\mu F_\mu = \mathcal{I}$. In this case the probability of getting the result μ is given by $p_\mu = \mathrm{tr}(F_\mu \rho)$.

Entropy and Quantum Mutual Information

Definition 1.2.4 (von Neumann entropy)**.** The von Neumann entropy of a quantum state $\rho \in \mathcal{S}(\mathcal{H}_S)$ is defined by

$$S_{[\log_q]}(\rho) = -\mathrm{tr}\big(\rho \log_q \rho\big). \tag{1.17}$$

For a bipartite quantum system AB the joint von Neumann entropy of the state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined by

$$S_{[\log_q]}(A,B) \equiv S_{[\log_q]}(\rho_{AB}) = -\mathrm{tr}\big(\rho_{AB} \log_q \rho_{AB}\big). \tag{1.18}$$

By analogy with the Shannon entropies the conditional entropy of system A given system B is defined by

$$S_{[\log_q]}(A|B) = S_{[\log_q]}(A,B) - S_{[\log_q]}(B), \tag{1.19}$$

where $S_{[\log_q]}(B) \equiv S_{[\log_q]}(\rho_B)$ denotes the entropy of the reduced state $\rho_B = \mathrm{tr}_B(\rho_{AB})$ (tr_B denotes the partial trace with respect to system B). In contrast to the conditional Shannon entropy, the conditional von Neumann entropy might become negative.

Definition 1.2.5 (Quantum mutual information)**.** The quantum mutual information of a bipartite quantum system AB in the state $\rho_{AB} \in \mathcal{S}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is defined by

$$I_{[\log_q]}(A:B) = S_{[\log_q]}(A) + S_{[\log_q]}(B) - S_{[\log_q]}(A,B). \tag{1.20}$$

As it is the case for the classical mutual information, the relation $I_{[\log_q]}(A:B) = S_{[\log_q]}(A) - S_{[\log_q]}(A|B) = S_{[\log_q]}(B) - S_{[\log_q]}(B|A) = I_{[\log_q]}(B:A)$ holds. The quantum mutual information is always non-negative.

Quantum Registers

A two-dimensional quantum mechanical system is called a qubit. Finite-dimensional quantum mechanical systems of higher dimension are called qudits. Let $\mathcal{H}_q = \mathbb{C}^q$ denote the Hilbert space of a qudit of dimension q, and fix an orthonormal basis $\{|0\rangle, \ldots, |q-1\rangle\}$ of $\mathcal{H}_q$. A quantum register consisting of n qudits of dimension q is defined on the Hilbert space $\mathcal{H}_q^{\otimes n}$. An orthonormal basis of $\mathcal{H}_q^{\otimes n}$ is given by the set of n-fold product states of the one-qudit basis states,

$$\mathcal{H}_q^{\otimes n} = \mathrm{span}\big\{|i_1, i_2, \ldots, i_n\rangle\big\}, \tag{1.21}$$

with $0 \le i_j < q$ for $j \in \{1, 2, \ldots, n\}$ and $|i_1, i_2, \ldots, i_n\rangle = |i_1\rangle \otimes |i_2\rangle \otimes \cdots \otimes |i_n\rangle$. A short hand notation for the basis states is given by $|i_1, i_2, \ldots, i_n\rangle = |\vec{i}\rangle$ with $\vec{i} \in \mathbb{F}_q^n$.

Pauli Operators

We consider qudits of of prime dimensions. The Pauli X and Z operators acting on $\mathcal{H}_q$ are defined by*

$$X|i\rangle = |i+1 \pmod q\rangle \tag{1.22a}$$

$$Z|i\rangle = \omega^i |i\rangle, \tag{1.22b}$$

where $\omega = \exp(2\pi i/q)$ is a complex primitive q-th root of unity. It follows that $ZX = \omega XZ$.

*Some authors use the definition $X|i\rangle = |i-1 \pmod q\rangle$. See, for instance, [Ham03].

Definition 1.2.6. For any vector $\vec{a} = (\vec{a}^x, \vec{a}^z) = (a_1^x, \ldots, a_n^x, a_1^z, \ldots, a_n^z) \in \mathbb{F}_q^{2n}$, let the Pauli operator $XZ(\vec{a})$ acting on $\mathcal{H}_q^{\otimes n}$ be defined by

$$XZ(\vec{a}) = \begin{cases} i^{a_1^x a_1^z} X^{a_1^x} Z^{a_1^z} \otimes \cdots \otimes i^{a_n^x a_n^z} X^{a_n^x} Z^{a_n^z} & \text{for } q = 2 \\ X^{a_1^x} Z^{a_1^z} \otimes \cdots \otimes X^{a_n^x} Z^{a_n^z} & \text{for } q \geq 3 \end{cases}, \tag{1.23}$$

so that the eigenvalues of $XZ(\vec{a})$ are powers of ω.

Remark. If we write the operator $XZ(\vec{a})$ as $XZ((\vec{a}^x, \vec{a}^z))$ for some $\vec{a} = (\vec{a}^x, \vec{a}^z) \in \mathbb{F}_q^{2n}$, we will use the shorthand notation $XZ(\vec{a}^x, \vec{a}^z)$ omitting the braces of $\vec{a} = (\vec{a}^x, \vec{a}^z)$. For instance, the identity operator $\mathcal{I}$ is given by the operator $XZ(\vec{0}, \vec{0})$ with $\vec{0} = (0, \ldots, 0) \in \mathbb{F}_q^n$.
If we represent the qubit Pauli operators in the $\{|0\rangle, |1\rangle\}$-basis, we obtain the well known Pauli matrices,

$$XZ(0,0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad XZ(1,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad XZ(1,1) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad XZ(0,1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.24}$$

which are also denoted as $\mathcal{I}, X, Y$ and Z. Hence, the qubit Pauli operators are Hermitian. For $q \geq 3$ we obtain

$$XZ(\vec{a}) \cdot XZ(\vec{b}) = \omega^{\sum_i a_i^z b_i^x} XZ(\vec{a} + \vec{b}), \tag{1.25}$$

while for $q = 2$ this expression holds up to some powers of i. As a consequence, $XZ(\cdot)$ gives rise to a unitary projective representation of $\mathbb{F}_q^{2n}$, which by itself forms a group under addition modulo q:

$$XZ(\cdot): \ \mathbb{F}_q^{2n} \ni \vec{a} \mapsto XZ(\vec{a}) \in \mathfrak{P}_q^n. \tag{1.26}$$

The full Pauli group is given by

$$\mathfrak{P}_q^n = \begin{cases} \{\mu XZ(\vec{a}) \mid \mu \in \{\pm 1, \pm i\}, \vec{a} \in \mathbb{F}_q^{2n}\} & , q = 2 \\ \{\omega^j XZ(\vec{a}) \mid j \in \mathbb{F}_q, \vec{a} \in \mathbb{F}_q^{2n}\} & , q \geq 3 \end{cases}. \tag{1.27}$$

Its order is $4 \cdot 4^n$ for qubits and $q \cdot q^{2n}$ in general ($q \geq 3$). If two elements of the Pauli group are identical up to a phase ω^p, $p \in \mathbb{F}_q$, (or some power of i for $q = 2$ respectively), we write $XZ(\vec{a}) \sim \omega^p XZ(\vec{a})$. We denote the set containing all n-fold tensor products of Pauli operators as

$$\mathcal{P}_q^n = \{XZ(\vec{a}) \mid \vec{a} \in \mathbb{F}_q^{2n}\}. \tag{1.28}$$

Note that $|\mathcal{P}_q^n| = q^{2n}$ while $|\mathfrak{P}_q^n| = q \cdot q^{2n}$ (for $q \geq 3$).

Definition 1.2.7. The symplectic inner product between elements $\vec{a}$ and $\vec{b}$ of $\mathbb{F}_q^{2n}$ is defined as

$$(\vec{a}, \vec{b})_{sp} = \sum_{i=1}^{n} a_i^z b_i^x - a_i^x b_i^z \pmod{q}. \tag{1.29}$$

Remark. With the help of the inner product defined above, the order of a product of two Pauli operators $XZ(\vec{a})$ and $XZ(\vec{b})$ can be inverted,

$$XZ(\vec{a}) \cdot XZ(\vec{b}) = \omega^{(\vec{a}, \vec{b})_{sp}} XZ(\vec{b}) \cdot XZ(\vec{a}). \tag{1.30}$$

Two operators commute if and only if the symplectic inner product between $\vec{a}$ and $\vec{b}$ vanishes.

Bell States

Definition 1.2.8 (Bell states)**.** Let $\mathcal{H}_q$ denote the Hilbert space of a qudit of dimension q and let $\mathcal{H}_A = \mathcal{H}_q^{\otimes n}$, $\mathcal{H}_B = \mathcal{H}_q^{\otimes n}$. Then the states

$$|\Phi_{\vec{x}}\rangle_{AB} = \frac{1}{\sqrt{q^n}} \sum_{\vec{j} \in \mathbb{F}_q^n} |\vec{j}\rangle_A \otimes XZ(\vec{x})_B |\vec{j}\rangle_B, \quad \vec{x} \in \mathbb{F}_q^{2n}, \tag{1.31}$$

are called Bell states. They are maximally entangled and form an orthonormal basis of $\mathcal{H}_A \otimes \mathcal{H}_B$.

1.2.3 Representation Theory

This subsection provides a brief overview of the basics of the representation theory of finite groups. Representation theory will be relevant for decoupling in part I (if the elements of a decoupling scheme form a projective representation of an underlying group) and as a tool for the evaluation of the secure key rates of the quantum key distribution protocols in part II.

We consider a finite group G of order n_G, i. e. G contains $n_G = |G|$ elements. If the elements of G commute with one another, the group is called an abelian group.

Definition 1.2.9. An element $b \in G$ is said to be conjugate to an element $a \in G$ if there exists $u \in G$ such that $b = uau^{-1}$. Elements conjugate to one another form a conjugacy class.

Since conjugacy is an equivalence relation, each element of G belongs to one and only one of the classes. If we denote the number of classes by n_ζ and the number of elements in class i by n_i, we have $\sum_{i=1}^{n_\zeta} n_i = n_G$. A class containing the inverse of all elements in the class is called ambivalent. If a group is abelian, each element forms a class by itself.

Definition 1.2.10. A representation (rep) R of G is a group homomorphism from G to a group $R(G)$ of operators on a vector space $\mathcal{V}$,

$$R : G \ni a \mapsto R(a) = R_a \in \mathcal{L}(\mathcal{V}). \tag{1.32}$$

From the definition of a group homomorphism we have $R_{ab} = R_a \cdot R_b$ for all $a, b \in G$. The dimension $d = \dim(\mathcal{V})$ of $\mathcal{V}$ is called the dimension of the rep.

Remark. If $\mathcal{V}$ is the vector space over the field $\mathbb{C}$, a map from G to a set $R(G)$ of operators on $\mathcal{V}$ satisfying

$$R_{ab} = r(a, b) \cdot R_a \cdot R_b, \tag{1.33}$$

with $r(a, b) \in \mathbb{C}$ for all $a, b \in G$, is called a projective representation.

We will always assume that the vector space $\mathcal{V}$ is an inner product space over the field $\mathbb{C}$. Let us fix an orthonormal basis $\{|j\rangle\}_{j=0\ldots d-1}$ of $\mathcal{V}$. Then,

$$R_a|j\rangle = \sum_{i=0}^{d-1} D_{ij}(a)|i\rangle, \tag{1.34}$$

with $D_{ij}(a) = \langle i|R_a|j\rangle$, and we obtain

$$R_a R_b|j\rangle = R_a \sum_{i=0}^{d-1} D_{ij}(b)|i\rangle = \sum_{k,i=0}^{d-1} D_{ki}(a)D_{ij}(b)|k\rangle = R_{ab}|j\rangle = \sum_{k=0}^{d-1} D_{kj}(ab)|k\rangle.$$

Since the $\{|j\rangle\}$ form a basis, it follows that $D_{kj}(ab) = \sum_i D_{ki}(a)D_{ij}(b)$ or $D(ab) = D(a) \cdot D(b)$. Hence, the group of matrices $D(G) = \{D(a) \mid a \in G\}$ forms a matrix representation of G. If $R(G)$ is a representation of G on a vector space $\mathcal{V}$, and A is a non-singular operator on $\mathcal{V}$, then it is obvious that $R'(G) = AR(G)A^{-1}$ also forms a representation of G on $\mathcal{V}$. In this case $R(G)$ and $R'(G)$ are related by a similarity transformation.

Definition 1.2.11. Two representations of a group G on a vector space $\mathcal{V}$ which are related by a similarity transformation are said to be equivalent representations.

Definition 1.2.12. If the group representation space is an inner product space and if the operators R_g are unitary for all $g \in G$, then the representation $R(G)$ is called a unitary representation.

Remark. It can be shown that every representation of a finite group on an inner product space is equivalent to a unitary representation (see e. g. [Tun85, theorem 3.3]). In the following we consider only unitary representations.

Definition 1.2.13. Let $R(G)$ be a representation of G on a vector space $\mathcal{V}$. A subspace $\mathcal{V}_1$ of $\mathcal{V}$ is called invariant subspace of $\mathcal{V}$ with respect to $R(G)$ if $R_g|\varphi\rangle \in \mathcal{V}_1$ for all $g \in G$ and for all $|\varphi\rangle \in \mathcal{V}_1$.

Remark. If a space $\mathcal{V}_1$ is an invariant subspace of a representation $R(G)$ on $\mathcal{V}$, then $\mathcal{V}_1$ itself is a representation space.

Theorem 1.2.3. *If an operator A commutes with all operators R_g of a rep $R(G)$, then the eigenspace $\mathcal{V}_\lambda$ of A is a representation space of G.*

Proof. We show that $\mathcal{V}_\lambda$ is an invariant subspace of $R(G)$ on $\mathcal{V}$. Let $|\phi_\lambda\rangle \in \mathcal{V}_\lambda$ so that $A|\phi_\lambda\rangle = \lambda|\phi_\lambda\rangle$. Then, $AR_g|\phi_\lambda\rangle = R_gA|\phi_\lambda\rangle = \lambda R_g|\phi_\lambda\rangle$ and it follows that $R_g|\phi_\lambda\rangle \in \mathcal{V}_\lambda$ for all $|\phi_\lambda\rangle \in \mathcal{V}_\lambda$ and all $R_g \in R(G)$. □

Definition 1.2.14. A representation $R(G)$ on $\mathcal{V}$ is irreducible if there is no non-trivial invariant subspace in $\mathcal{V}$ with respect to $R(G)$ (we may also say that the representation space is irreducible). Otherwise the representation is reducible.

Since we consider only unitary representations, reducible always means fully reducible: Let $\mathcal{V}_1$ be an invariant subspace of the representation space $\mathcal{V}$, and let $\mathcal{V}_2$ be the space orthogonal to $\mathcal{V}_1$, i. e. $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$. Then, since $\langle R_g v_2|v_1\rangle = \langle v_2|R_g^\dagger v_1\rangle = \langle v_2|R_{g^{-1}}v_1\rangle = 0$ for all $|v_1\rangle \in \mathcal{V}_1$, all $|v_2\rangle \in \mathcal{V}_2$ and all $g \in G$, it follows that $\mathcal{V}_2$ remains invariant, too. In other words, the operators R_g of a reducible representation $R(G)$ become block-diagonal for a proper choice of basis. For instance, if the representation space $\mathcal{V}$ decomposes into two irreducible invariant subspaces $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2$ of dimension d_1 and $d_2 = d - d_1$, we write $R(G) = D^{(1)}(G) \oplus D^{(2)}(G)$ and

$$R_g \mapsto D(g) = \begin{pmatrix} D^{(1)}(g) & 0 \\ 0 & D^{(2)}(g) \end{pmatrix}, \tag{1.35}$$

where $D^{(1)}(g)$ is a $d_1 \times d_1$ matrix and $D^{(2)}(g)$ is a $d_2 \times d_2$ matrix. In general we obtain the relation

$$R(G) = \bigoplus_{\nu \in \mathcal{J}} \tau_\nu \cdot D^{(\nu)}(G), \tag{1.36}$$

where ν labels inequivalent irreducible representations and τ_ν denotes the number of times a certain irreducible representation ν occurs. The dimension of the irrep $D^{(\nu)}(G)$ is denoted by d_ν. Hence there exists an orthonormal basis

$$\big\{|\nu\ l_\nu\ m_\nu\rangle \,|\, \nu \in \mathcal{J},\ l_\nu = 1 \dots \tau_\nu,\ m_\nu = 1 \dots d_\nu\big\}, \tag{1.37}$$

in which the operators R_g are block-diagonal, i. e.

$$R_g|\nu\ l_\nu\ m_\nu\rangle = D^{(\nu)}(g)|\nu\ l_\nu\ m_\nu\rangle = \sum_{m'_\nu=1}^{d_\nu} D^{(\nu)}_{m'_\nu m_\nu}(g)|\nu\ l_\nu\ m'_\nu\rangle. \tag{1.38}$$

We label the subspace of the representation space $\mathcal{V}$ which is spanned by the set of basis vectors with fixed ν by $\mathcal{V}_\nu$,

$$\mathcal{V}_\nu = \operatorname{span}\big\{|\nu\ l_\nu\ m_\nu\rangle \,|\, l_\nu = 1 \dots \tau_\nu, m_\nu = 1 \dots d_\nu\big\}. \tag{1.39}$$

Since $\mathcal{V}_\nu$ has the form of a tensor space ($|\nu\ l_\nu\ m_\nu\rangle = |l_\nu\rangle \otimes |m_\nu\rangle$), we write $\mathcal{V}_\nu = \mathcal{C}_\nu \otimes \mathcal{D}_\nu$, where the dimension of $\mathcal{C}_\nu$ is given by τ_ν and the dimension of $\mathcal{D}_\nu$ is given by d_ν. The representation space $\mathcal{V}$ decomposes into a direct sum of orthogonal subspaces,

$$\mathcal{V} = \bigoplus_{\nu \in \mathcal{J}} \mathcal{V}_\nu = \bigoplus_{\nu \in \mathcal{J}} \mathcal{C}_\nu \otimes \mathcal{D}_\nu. \tag{1.40}$$

If we restrict an irreducible representation (irrep) $D^{(\nu)}(G)$ of a group G to elements of a subgroup $G_s \subset G$, we obtain a subduced representation denoted as $D^{(\nu)}(G) \downarrow G_s$. A subduced rep is in general reducible and can be decomposed into a direct sum of irreps of G_s,

$$D^{(\nu)}(G) \downarrow G_s = \bigoplus_{\mu} \tau_{\mu}^{(\nu)} \cdot D^{(\mu)}(G_s), \tag{1.41}$$

where $\tau_{\mu}^{(\nu)}$ denotes the number of times the irrep $D^{(\mu)}(G_s)$ occurs in $D^{(\nu)}(G) \downarrow G_s$. If $\tau_{\mu}^{(\nu)} \leq 1$ for all possible ν and μ, then G_s is called a canonical subgroup of G. A canonical subgroup chain is a group chain $G \supset G_1 \supset G_2 \cdots \supset G_n$ such that G_{i+1} is a canonical subgroup of G_i ($i = 0, \ldots, n-1$ with $G \equiv G_0$) and G_n is abelian.

Theorem 1.2.4 (Schur's lemma i)**.** *Let A be an operator commuting with all operators of a rep $R(G)$ of G on $\mathcal{V}$, and let $\mathcal{V}_\nu \subseteq \mathcal{V}$ be an irreducible rep space of G and an invariant subspace of A. Then $\mathcal{V}_\nu$ is necessarily an eigenspace of A.*

Proof. Let us assume that the invariant subspace $\mathcal{V}_\nu$ of A decomposes into two eigenspaces of A, $\mathcal{V}_\nu = \mathcal{V}_{\nu,1} \oplus \mathcal{V}_{\nu,2}$. According to theorem 1.2.3, each of these spaces would be a representation space, which is in contradiction to $\mathcal{V}_\nu$ being an irreducible rep space. Hence, the only possibility is that $A\mathcal{V}_\nu = \nu\mathcal{V}_\nu$. □

Remark (i). The representative of an operator A in $\mathcal{V}_\nu$ is a multiple of the identity: Let a basis of $\mathcal{V}_\nu$ be given by $\{|i\rangle\}_{i=0,\ldots,d_\nu-1}$. Then the matrix representative of A in $\mathcal{V}_\nu$ is given by $D_{ij}^{(\nu)}(A) = \langle i|A|j\rangle = \nu\delta_{ij}$. If $\mathcal{V}_\nu = \mathcal{V}$ we obtain the result that the only operator commuting with all operators of an irrep $R(G)$ is a multiple of the identity.

Remark (ii). A direct consequence of Schur's lemma is that an irrep of an abelian group must be of dimension one.

Theorem 1.2.5 (Schur's lemma ii)**.** *Let $D^{(\mu)}(G)$ and $D^{(\nu)}(G)$ be two irreps of G on the spaces $\mathcal{V}_\mu$ and $\mathcal{V}_\nu$ respectively, and let A be a linear transformation from $\mathcal{V}_\nu$ to $\mathcal{V}_\mu$ which satisfies $AD^{(\nu)}(g) = D^{(\mu)}(g)A$ for all $g \in G$. Then, either $A = 0$, or $\mathcal{V}_\mu$ and $\mathcal{V}_\nu$ are isomorphic and $D^{(\mu)}(G) = AD^{(\nu)}(G)A^{-1}$, i. e. the irreps μ and ν are equivalent.*

Proof. It is easy to verify that the range of A is an invariant subspace of $\mathcal{V}_\mu$ with respect to $D^{(\mu)}(G)$. Since $D^{(\mu)}(G)$ is irreducible it follows that either the range is 0 (which implies $A = 0$) or the range is $\mathcal{V}_\mu$. Similarly, the null space of A in $\mathcal{V}_\nu$ is an invariant subspace of $\mathcal{V}_\nu$ with respect to $D^{(\nu)}(G)$. Since $D^{(\nu)}(G)$ is irreducible it follows that either the null space is equal to $\mathcal{V}_\nu$ (implying $A = 0$) or the null space is 0 (implying that A is a one-to-one mapping). Hence A is either an isomorphism between $\mathcal{V}_\mu$ and $\mathcal{V}_\nu$ or it vanishes. □

The second part of Schur's lemma can be used to prove the orthonormality of irreducible representation matrices.

Theorem 1.2.6 (Orthonormality of irreducible representation matrices)**.** *Let $D^{(\nu)}(G)$ and $D^{(\mu)}(G)$ denote two inequivalent irreducible representations of G, and let the dimension of the μ representation be given by d_μ. Then the following orthonormality condition holds,*

$$\frac{d_\mu}{n_G} \sum_{g \in G} D_{ki}^{\dagger(\mu)}(g)\ D_{jl}^{(\nu)}(g) = \delta_{\mu\nu}\delta_{ij}\delta_{kl}, \tag{1.42}$$

with $D_{ki}^{\dagger(\mu)}(g)$ denoting the complex conjugate of the matrix element $D_{ik}^{(\mu)}(g)$.

Definition 1.2.15 (Group algebra)**.** The group algebra $\mathbb{C}G$ is defined as the complex vector space spanned by the group elements, i. e. any element a in $\mathbb{C}G$ can be written as $a = \sum_{g\in G} a_g g$ with $a_g \in \mathbb{C}$. For two elements a, b in $\mathbb{C}G$ the product

$$ab = \Big(\sum_{g\in G} a_g g\Big)\Big(\sum_{h\in G} b_h h\Big) = \sum_{g,h\in G} a_g b_h\, gh = \sum_{g'\in G}\Big(\sum_{h\in G} a_{g'h^{-1}} b_h\Big) g' \tag{1.43}$$

turns $\mathbb{C}G$ into an algebra.

Any representation R of G extends by linearity to a representation of the elements in $\mathbb{C}G$. Let $\mathcal{A} = R(\mathbb{C}G)$ denote the algebra generated by R, and let its commutant $\mathcal{A}'$ be defined as the set of elements that commutes with all the elements in $\mathcal{A}$, $\mathcal{A}' = \{V \in \mathcal{L}(\mathcal{V}) \mid VA = AV \text{ for all } A \in \mathcal{A}\}$. The following theorem follows from the orthonormality of irreducible representation matrices and the first part of Schur's lemma.

Theorem 1.2.7. *In the $\{|\nu\ l_\nu\ m_\nu\rangle\}$-basis (as defined in equation (1.37)) corresponding to the representation R, $\mathcal{A} = R(\mathbb{C}G)$ and $\mathcal{A}'$ take the form*

$$\mathcal{A} \cong \bigoplus_{\nu\in\mathcal{J}} \mathbb{1}_{\tau_\nu} \otimes \mathrm{Mat}(d_\nu \times d_\nu, \mathbb{C}) \tag{1.44}$$

$$\mathcal{A}' \cong \bigoplus_{\nu\in\mathcal{J}} \mathrm{Mat}(\tau_\nu \times \tau_\nu, \mathbb{C}) \otimes \mathbb{1}_{d_\nu}, \tag{1.45}$$

where $\mathbb{1}_n$ denotes an $n \times n$ dimensional identity matrix and $\mathrm{Mat}(n \times n, \mathbb{C})$ denotes the set of $n \times n$ matrices with entries in $\mathbb{C}$.

Remark. In part I of this thesis we are sometimes going to deal with projective representations $R(G)$ of G on $\mathcal{V}$. In this case we assume that the set of unitary matrices $\{R_g = R(g) \mid g \in G\}$ generates a finite group $\hat{G}$ larger than G and consider the ordinary irreducible representations of $\hat{G}$. If we define the center of $\hat{G}$ by $Z(\hat{G}) = \{z \in \hat{G} \mid gz = zg \text{ for all } g \in \hat{G}\}$ then the quotient group $\hat{G}/Z(\hat{G})$ is isomorphic to the original group G.

Let us close this subsection revisiting the set $\mathcal{P}_q^n$ of n-fold tensor products of Pauli operators. This set is an example of a so-called nice error basis. Such a basis was defined by Knill in [Kni96] as follows:

Definition 1.2.16. Let G be a group of order $|G| = d^2$ and let its identity element be denoted by e. A nice error basis is a set $\mathcal{E} = \{D(g) \in \mathsf{U}_d \mid g \in G\}$ of unitary $d \times d$ matrices such that (i) $D(e)$ is given by the identity matrix, (ii) $\mathrm{tr}(D(g))/d = \delta_{g,e}$ for all $g \in G$, and (iii) $D(g)D(h) = \alpha(g,h)D(gh)$ for all $g, h \in G$, where $\alpha(g,h)$ is a function from $G \times G$ to $\mathbb{C} \setminus \{0\}$.

A consequence of conditions (i) and (iii) is that the map $G \ni g \mapsto D(g) \in \mathsf{U}_d$ defines a projective representation of G on a d-dimensional Hilbert space $\mathcal{H}$. It follows from condition (ii) that the matrices $D(g)$ are pairwise orthogonal with respect to the trace inner product $\langle A, B\rangle = \mathrm{tr}(A^\dagger B)/d$. Hence they form a basis for the operators acting on $\mathcal{H}$ and the projective representation of G on $\mathcal{H}$ must be irreducible. Since the matrices are unitary we have $|\det D(g)| = 1$ for all $g \in G$ and it follows from (iii) that $|\alpha(g,h)| = 1$. The group G is also called the index group. It is easy to verify that the set $\mathcal{P}_q^n$ of Pauli operators with the index group given by $\mathbb{F}_q^{2n}$ fulfills the definition of a nice error basis with $\alpha(\vec{g}, \vec{h}) = \omega^{\sum_i g_i^z h_i^x}$ for $\vec{g} = (\vec{g}^x, \vec{g}^z), \vec{h} = (\vec{h}^x, \vec{h}^z) \in \mathbb{F}_q^{2n}$ (compare with (1.25)).

Finally, let us define the notation of several groups we are going to encounter. The symmetric group S_n on the finite set $\{1, 2, \ldots, n\}$ consists of all permutations of the set and has order $n!$. The general linear group of degree q over the field $\mathbb{C}$ is the group of $q \times q$ invertible matrices with entries from $\mathbb{C}$. It is denoted by $\mathsf{GL}_q = \mathsf{GL}(q, \mathbb{C})$. Subgroups of GL_q are the unitary group U_q containing unitary matrices, the special unitary group SU_q containing unitary matrices with unit determinant, and the 3-dimensional rotation group which is the special orthogonal group of degree 3 over the field $\mathbb{R}$ and is denoted by $\mathsf{SO}_3 = \mathsf{SO}(3, \mathbb{R})$.

Part I

Random Decoupling

2 Dynamical Decoupling

This chapter deals with dynamical decoupling strategies in the bang-bang control scenario. After giving an introduction to dynamical control theory and average Hamiltonian theory (AHT), we present an overview over known construction methods for dynamical decoupling schemes. The main focus is then on improved decoupling strategies which are based on a fixed decoupling scheme. The performance of these strategies is analyzed by deriving formulas for the average fidelity decay. For any randomized strategy, in addition, the variance of the fidelity is studied. With the help of a numerical simulation of a quantum memory perturbed by Heisenberg interactions, these formulas are validated and conclusions concerning a general guideline for optimal decoupling are drawn.

We start by presenting the necessary framework in section 2.1. The overview over known construction methods for efficient decoupling schemes will then be given in section 2.2. Improved control strategies based on a given decoupling scheme are explored in section 2.3. Finally, we present the results of the numerical simulation in section 2.4.

2.1 Dynamical Control of Quantum Systems

Let S be a quantum system defined on a finite d-dimensional Hilbert space $\mathcal{H}_S$ and let its dynamics be generated by the system Hamiltonian $H_0 \in \mathcal{L}(\mathcal{H}_S)$. Typically the quantum system S under consideration will be a quantum register consisting of n qudits of dimension q so that $\mathcal{H}_S = \mathcal{H}_q^{\otimes n}$ and the system Hamiltonian describes some static imperfections. We assume that we are able to apply a certain set of local control operations which are realized by the time-dependent control Hamiltonian $H_c(t) \in \mathcal{L}(\mathcal{H}_S)$. Local means that H_c is a sum over one qudit Hamiltonians, i. e. $H_c(t) = \sum_{i=1}^{n} h_i^{(i)}(t) \otimes \mathcal{I}_{\{1,\ldots,n\}\setminus\{i\}}$ with some time-dependent $h^{(i)}(t) \in \mathcal{L}(\mathcal{H}_q)$. In turn the total Hamiltonian is given by

$$H(t) = H_0 + H_c(t) \tag{2.1}$$

and according to the Schrödinger equation our system evolves in time as

$$U(t) = \mathcal{T} \exp\Big(-i \int_0^t H(t')dt'/\hbar\Big), \tag{2.2}$$

where $\mathcal{T}$ denotes the Dyson time-ordering operator. Analogous to (2.2) let us denote the time evolution due to $H_c(t)$ alone by $U_c(t)$, i. e.

$$U_c(t) = \mathcal{T} \exp\Big(-i \int_0^t H_c(t')dt'/\hbar\Big). \tag{2.3}$$

We now define the toggled frame as the frame that continuously follows the applied control, $\tilde{U}(t) = U_c^\dagger(t)U(t)$. The time evolution in the toggled frame is determined by the Schrödinger equation

$$i\hbar \frac{d\tilde{U}(t)}{dt} = \tilde{H}(t)\tilde{U}(t), \tag{2.4}$$

where the toggled frame Hamiltonian is given by

$$\tilde{H}(t) = U_c^\dagger(t) H_0 U_c(t). \tag{2.5}$$

Dynamical control of $H_c(t)$ and in turn of $U_c(t)$ allows us to modify the time evolution in the toggled frame. There are different possible control tasks. If we deal with a quantum memory for example, we

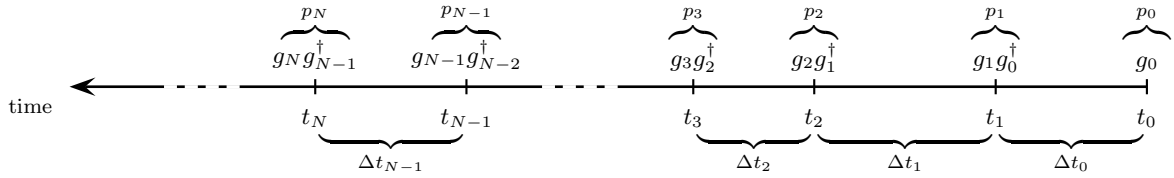


Figure 2.1: Schematic representation of bang-bang control. At time t_i the pulse p_i is applied instantaneously.

may want to freeze the evolution by demanding $\tilde{U}(t) \approx \mathcal{I}$ in order to preserve the stored data. Another goal is the simulation of other Hamiltonians (see e. g. [WRJB02a; BDNB04]), i. e. we would like the system to evolve as $\tilde{U}(t) \approx \exp(-iH_0' t/\hbar)$ with $H_0' \neq H_0$. The former of these tasks is called decoupling.

2.1.1 Bang-Bang Control

In the quantum bang-bang control scenario [VL98] it is assumed that we are able to apply a strong control $H_c(t)$ over a very short time interval. In this case the resulting control action can be described as a quasi-instantaneous application of unitary pulses p_i at times $t_i = \sum_{k=0}^{i-1} \Delta t_k$, $i \in \mathbb{N}_0$. Since the control is assumed to be local, these pulses are of the form $p_i = u_1^{(1,i)} \otimes u_2^{(2,i)} \otimes \cdots \otimes u_n^{(n,i)}$, where $u_c^{(a,i)}$ denotes the unitary $u^{(a,i)} \in \mathsf{U}_q$ being applied to the c-th qudit. After a time t_N we obtain the total time evolution

$$U(t_N) = p_N \, f_{\Delta t_{N-1}} \, \cdots \, p_2 \, f_{\Delta t_1} \, p_1 \, f_{\Delta t_0} \, p_0, \tag{2.6}$$

as depicted in figure 2.1. Here, $f_{\Delta t_j} = \exp(-iH_0 \Delta t_j/\hbar)$ denotes free evolution due to H_0 over the time interval Δt_j. Defining $g_i = p_i \ldots p_1 p_0$ we note that this evolution can be written as

$$U(t_N) = g_N (g_{N-1}^\dagger f_{\Delta t_{N-1}} g_{N-1}) \ldots (g_1^\dagger f_{\Delta t_1} g_1)(g_0^\dagger f_{\Delta t_0} g_0). \tag{2.7}$$

The time evolution operator U_c at time $t_i + s$ with $s \in [0, \Delta t_i)$ is given by $U_c(t_i + s) = g_i$, i. e. U_c jumps from g_{i-1} to $g_i = (g_i g_{i-1}^\dagger) g_{i-1} \equiv p_i g_{i-1}$ at time t_i. Since $g_j^\dagger f_{\Delta t_j} g_j = \exp(-i g_j^\dagger H_0 g_j \Delta t_j/\hbar)$, let us define the toggled frame Hamiltonians $\tilde{H}_i = g_i^\dagger H_0 g_i$. After switching to the toggled frame $\tilde{U}(t_N) = U_c^\dagger(t_N) U(t_N)$, the time evolution of equation (2.7) becomes

$$\tilde{U}(t_N) = \exp(-i\tilde{H}_{N-1} \Delta t_{N-1}/\hbar) \ldots \exp(-i\tilde{H}_1 \Delta t_1/\hbar) \exp(-i\tilde{H}_0 \Delta t_0/\hbar). \tag{2.8}$$

To keep the notation as simple as possible, we set $\hbar = 1$ for the remaining chapters.

2.1.2 Average Hamiltonian Theory

A convenient tool which is commonly used to analyze the resulting dynamics of a dynamical control scheme in the toggled frame is the average Hamiltonian theory (AHT) [EBW87]. Let the time evolution in the toggled frame be generated by the time-dependent toggling frame Hamiltonian $\tilde{H}(t)$ of equation (2.5). After a time t this results in the time evolution operator

$$\tilde{U}(t) = \mathcal{T} \exp\Big(-i \int_0^t \tilde{H}(t') dt'\Big), \tag{2.9}$$

which can be written in terms of an average Hamiltonian $\overline{H}$ (which depends on t) as

$$\tilde{U}(t) = \exp\Big(-i\overline{H} t\Big). \tag{2.10}$$

AHT expresses this average Hamiltonian as an infinite series of self-adjoint operators called Magnus expansion,

$$\overline{H} = \overline{H}^{(0)} + \overline{H}^{(1)} + \overline{H}^{(2)} + \dots, \tag{2.11}$$

the first three terms of which are given by

$$\overline{H}^{(0)} = \frac{1}{t}\int_0^t dt_1 \tilde{H}(t_1) \tag{2.12a}$$

$$\overline{H}^{(1)} = -\frac{i}{2t}\int_0^t dt_2 \int_0^{t_2} dt_1 [\tilde{H}(t_2), \tilde{H}(t_1)] \tag{2.12b}$$

$$\overline{H}^{(2)} = -\frac{1}{6t}\int_0^t dt_3 \int_0^{t_3} dt_2 \int_0^{t_2} dt_1 \Big([\tilde{H}(t_3), [\tilde{H}(t_2), \tilde{H}(t_1)]] + [[\tilde{H}(t_3), \tilde{H}(t_2)], \tilde{H}(t_1)] \Big). \tag{2.12c}$$

To obtain these expressions, we write (2.9) as an infinite series,

$$\begin{aligned} \tilde{U}(t) &= \mathcal{I} - it \sum_{n=0}^{\infty} \frac{(-i)^n}{t} \int_0^t dt_{n+1} \int_0^{t_{n+1}} dt_n \dots \int_0^{t_2} dt_1 \tilde{H}(t_{n+1}) \tilde{H}(t_n) \dots \tilde{H}(t_1) \\ &\equiv \mathcal{I} - it \sum_{n=0}^{\infty} h_n, \end{aligned} \tag{2.13}$$

and expand (2.10) as

$$\tilde{U}(t) = \mathcal{I} + \sum_{n=1}^{\infty} \frac{(-it)^n}{n!} \big(\overline{H}^{(0)} + \overline{H}^{(1)} + \overline{H}^{(2)} + \dots \big)^n. \tag{2.14}$$

By noting that both h_j and $\overline{H}^{(j)}$ are of order $j+1$ in $\tilde{H}$ and by comparing expressions of the same order in the last two equations, we obtain the expressions $h_0 = \overline{H}^{(0)}$, $h_1 = \overline{H}^{(1)} - it(\overline{H}^{(0)})^2/2$, et cetera, which eventually lead to (2.12). In the bang-bang scenario at the time $t = t_N$ the Hamiltonians (2.12a)–(2.12c) become

$$\overline{H}^{(0)} = \frac{1}{t_N} \sum_{j=0}^{N-1} \tilde{H}_j \Delta t_j \tag{2.15a}$$

$$\overline{H}^{(1)} = -\frac{i}{2t_N} \sum_{i>j=0}^{N-1} [\tilde{H}_i, \tilde{H}_j] \Delta t_i \Delta t_j \tag{2.15b}$$

$$\overline{H}^{(2)} = -\frac{1}{6t_N} \sum_{i\geq j\geq k=0}^{N-1} \Big([\tilde{H}_i, [\tilde{H}_j, \tilde{H}_k]] + [[\tilde{H}_i, \tilde{H}_j], \tilde{H}_k] \Big) \Delta t_i \Delta t_j \Delta t_k \times \begin{cases} 1/2 & \text{if } i = j \text{ or } j = k \\ 1 & \text{else} \end{cases}. \tag{2.15c}$$

Finally, we state a theorem that will be used later on in this chapter to improve the performance of dynamical control schemes. A proof of this theorem can be found in [Bur81].

Theorem 2.1.1. *If the toggled frame Hamiltonian is symmetric in time, i. e. if* $\tilde{H}(t - t') = \tilde{H}(t')$ *for* $t' \in [0, t]$*, all odd orders in the Magnus expansion* (2.11) *of* $\tilde{U}(t) = \exp(-i\overline{H}t)$ *vanish, i. e.* $\overline{H}^{(k)} = 0$ *for* $k = 1, 3, 5, \dots$.

2.1.3 The Fundamental Control Strategy

To achieve a certain control task like the simulation of a Hamiltonian H_0', we make use of the simple structure of the zeroth order term $\overline{H}^{(0)}$ in the bang-bang setting.

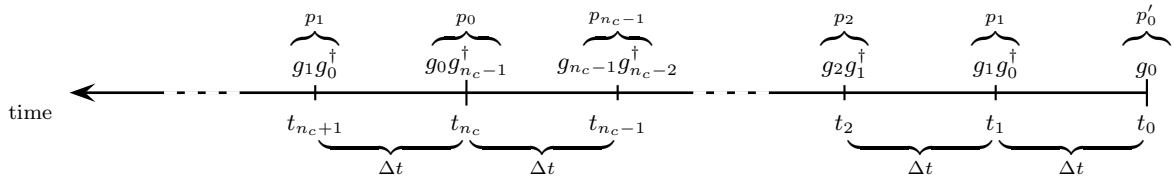


Figure 2.2: Schematic representation of the cyclic (or periodic) control strategy (PDD).

Definition 2.1.1. A set of unitaries $\{g_j\}_{j=0}^{n_c-1}$ and relative times $\{\Delta t_j\}_{j=0}^{n_c-1}$ such that

$$\overline{H}^{(0)} = \frac{1}{t_c}\sum_{j=0}^{n_c-1} \tilde{H}_j \Delta t_j \equiv \frac{1}{t_c}\sum_{j=0}^{n_c-1} g_j^\dagger H_0 g_j \Delta t_j = H_0' + \mathfrak{c}\cdot\frac{1}{d}\mathcal{I}, \tag{2.16}$$

where $t_c = \sum_{j=0}^{n_c-1}\Delta t_j$, $\mathfrak{c} = \mathrm{tr}(H_0) - \mathrm{tr}(H_0')$, and $d = \dim(\mathcal{H}_S)$, is called a control scheme of length n_c for the simulation of the Hamiltonian H_0' with the system Hamiltonian H_0.

Remark. Without loss of generality, we usually assume all of the involved Hamiltonians to be traceless. In this case we have a vanishing constant $\mathfrak{c} = 0$.

If we would like to achieve decoupling we set $H_0' \equiv 0$. In this case a control scheme $\{g_j, \Delta t_j\}_{j=0}^{n_c-1}$ is called a decoupling scheme. An overview over various decoupling schemes for different types of H_0 is given in section 2.2. It turns out that most of the times it is sufficient to consider control schemes with constant relative time intervals, i. e. $\Delta t_j = \Delta t$ for all $j \in \{0, \ldots, n_c - 1\}$. In the following we will always be dealing with such schemes.

The most basic control strategy is called cyclic (or periodic) dynamical decoupling* (PDD). It consists of repeating the pulse sequence $p_0, \ldots, p_{n_c-1}$, with $p_j = g_j g_{j-1}^\dagger$ for $j = 1, \ldots, n_c - 1$ and $p_0 = g_0 g_{n_c-1}^\dagger$ (with the exception that the first p_0 is simply given by g_0) constructed using the elements of a control scheme $\{g_j\}_{j=0}^{n_c-1}$ satisfying

$$\frac{1}{n_c}\sum_{j=0}^{n_c-1} g_j^\dagger H_0 g_j = H_0' + \mathfrak{c}\cdot\frac{1}{d}\mathcal{I}, \tag{2.17}$$

over and over again (compare with figure 2.2): At the time $t_j = j\cdot\Delta t$, $j \in \mathbb{N}_0$, the pulse $p_{j \bmod n_c}$ is applied. As a result, the time evolution in the toggled frame after a time $T = m\cdot t_c$, $m \in \mathbb{N}$, $t_c = n_c\Delta t$, is given by

$$\tilde{U}(T = m\cdot t_c) = \Big(\exp(-i\tilde{H}_{n_c-1}\Delta t)\ldots\exp(-i\tilde{H}_1\Delta t)\exp(-i\tilde{H}_0\Delta t)\Big)^m = \exp(-i\overline{H}t_c\cdot m), \tag{2.18}$$

where the zeroth order term in the Magnus expansion of $\overline{H}$ is given by (2.17). In the limit of $m\to\infty$ and $\Delta t \to 0$ with $T = m\cdot n_c\Delta t$ held constant, the influence of the higher order terms in the Magnus expansion decreases and PDD achieves its task perfectly: $\lim_{\Delta t\to 0}\tilde{U}(T) = \exp\big(-i\overline{H}^{(0)}T\big) = \exp(-iH_0'T)\cdot e^{-iT\mathfrak{c}/d}$. In a realistic experiment we do not achieve this limit. Therefore it is important to (i) quantify the error caused by the higher order terms and (ii) devise control strategies which keep the error for finite Δt as small as possible. In fact the main focus of the first part of this thesis is on (ii) and is dealt with in section 2.3. We proceed with (i) in the next subsection.

2.1.4 Performance Measure

If the control task is the simulation of a Hamiltonian H_0', the goal of a dynamical control strategy is to achieve a time evolution $\tilde{U}(T)$ in the toggled frame which is (up to a global phase) as close to

*We call it a decoupling strategy even so it might be used for the purpose of simulating some Hamiltonian.

$\tilde{U}_{\mathrm{id}}(T) = \exp(-iH_0'T)$ as possible. To quantify this closeness we define the pure state fidelity

$$F_{|\psi\rangle}(T) = \left|\langle\psi|\tilde{U}_{\mathrm{id}}^{\dagger}(T)\,\tilde{U}(T)|\psi\rangle\right|^2. \tag{2.19}$$

As long as $F_{|\psi\rangle}(T)$ stays close to one, we know that our control strategy was successful (at least if the quantum system was in the initial state $|\psi\rangle$). To drop the dependence on $|\psi\rangle$, we might consider the worst case fidelity,

$$F_w(T) = \min_{|\psi\rangle\in\mathcal{H}_S} \left|\langle\psi|\tilde{U}_{\mathrm{id}}^{\dagger}(T)\,\tilde{U}(T)|\psi\rangle\right|^2, \tag{2.20}$$

as it was done in [VK05] for the purpose of finding a lower bound, or we might consider the average fidelity

$$F_a(T) = \int \left|\langle\psi|\tilde{U}_{\mathrm{id}}^{\dagger}(T)\,\tilde{U}(T)|\psi\rangle\right|^2 d\psi. \tag{2.21}$$

Here, the integration involved in the definition of the average fidelity has to be performed over the uniform (Haar) measure on the relevant quantum state space with the normalization $\int d\psi = 1$.

More generally, for a trace-preserving quantum operation $\mathcal{E}$ (i. e. a trace-preserving completely positive map), the average fidelity is defined as

$$F_a(\mathcal{E}) = \int \langle\psi|\mathcal{E}(|\psi\rangle\langle\psi|)|\psi\rangle d\psi. \tag{2.22}$$

Let $|\Phi\rangle$ be a maximally entangled state (e. g. a Bell state) between the quantum system under consideration and an ancilla system of the same dimension $d = \dim(\mathcal{H}_S)$. Then the entanglement fidelity is defined as

$$F_e(\mathcal{E}) = \langle\Phi|(\mathcal{I}\otimes\mathcal{E})(|\Phi\rangle\langle\Phi|)|\Phi\rangle, \tag{2.23}$$

where $\mathcal{I}$ denotes the identity operation acting on the ancilla system. The entanglement fidelity measures the degree to which the entanglement of quantum state is preserved by a quantum operation $\mathcal{E}$. Apparently, it is independent of the choice of the maximally entangled state since any two maximally entangled states are related by a unitary acting only on the ancilla. Both fidelity measures are not independent but are related by [HHH99; Nie02]

$$F_a(\mathcal{E}) = \frac{dF_e(\mathcal{E})+1}{d+1} = F_e(\mathcal{E}) + \mathcal{O}\left(\frac{1-F_e(\mathcal{E})}{d}\right). \tag{2.24}$$

Thus, in the case of a quantum system which consists of a large number of qudits, i. e. $d = q^n \gg 1$, the difference between both measures tends to zero. If we set $\mathcal{E}(\rho) = \tilde{U}_{\mathrm{id}}^{\dagger}(T)\tilde{U}(T)\,\rho\,\tilde{U}^{\dagger}(T)\tilde{U}_{\mathrm{id}}(T)$, we obtain

$$F_e(\mathcal{E}) = F_e(T) = \left|\frac{1}{d}\,\mathrm{tr}\big(\tilde{U}_{\mathrm{id}}^{\dagger}(T)\tilde{U}(T)\big)\right|^2. \tag{2.25}$$

Typically, the evaluation of the entanglement fidelity is much simpler than the direct evaluation of the average fidelity (2.21). Therefore, in view of its close relationship to the average fidelity our subsequent discussion will mainly concentrate on the behavior of the entanglement fidelity.

Let us now consider the control task of decoupling, i. e. $\tilde{U}_{\mathrm{id}}(T) = \mathcal{I}$, and let us estimate the entanglement fidelity given by (2.25) for the PDD control strategy. The resulting fidelity has to be compared with the fidelity which is obtained in the absence of any decoupling. Without loss in generality, we assume that $\mathrm{tr}(H_0) = 0$.

No Decoupling (none)

Let us start examining the decay of the entanglement fidelity (2.25) in the absence of any decoupling. In this case the time evolution due to the control alone is trivial, $U_c(t) = \mathcal{I}$, and the time evolution in the toggled frame coincides with the time evolution in the Schrödinger picture, i. e. $\tilde{U}(T) = U(T) =$

$\exp(-iH_0T)$. In order to derive a series expansion of the fidelity, we write the system Hamiltonian H_0 as λH_0 and expand in λ (setting $\lambda = 1$ in the end). Such a series expansion up to fourth order in λ leads to

$$\begin{aligned} F_e^{\mathsf{none}}(T) &= \Big|\frac{1}{d}\operatorname{tr}\big(\tilde{U}(T)\big)\Big|^2 \\ &= 1 - \frac{1}{d}\operatorname{tr}(H_0^2)T^2 + \Big(\frac{1}{2}\big(\frac{1}{d}\operatorname{tr}(H_0^2)\big)^2 + \frac{1}{6d}\operatorname{tr}(H_0^4)\Big)\frac{1}{2}T^4 + \mathcal{O}(\lambda^6 T^6). \end{aligned} \tag{2.26}$$

Hence, for sufficiently small times, the fidelity decay is quadratic in time and its strength is determined by the trace of the square of the system Hamiltonian H_0. By comparison with numerical simulations for various H_0, we found that a good approximation of $F_e^{\mathsf{none}}(T)$ valid for $0 \le T \lesssim \sqrt{2}/\sqrt{\operatorname{tr}(H_0^2)/d}$, or in other words as long as $F_e^{\mathsf{none}}(T) \gtrsim 0.1$, is given by the simple expression

$$F_{e\ \mathrm{app}}^{\mathsf{none}}(T) = \exp\Big(-\frac{1}{d}\operatorname{tr}(H_0^2)T^2\Big). \tag{2.27}$$

Viola and Knill [VK05, theorem 3] gave a strict lower bound on the worst case fidelity (2.20) for PDD by using the matrix norm $\|A\|_2 = \max|\operatorname{eig}(\sqrt{A^\dagger A})|$ and setting $\kappa = \|H_0\|_2$. Analogous to this bound, a corresponding lower bound in the absence of decoupling is given by

$$F_w^{\mathsf{none}}(T) = \min_{|\psi\rangle\in\mathcal{H}_S} \big|\langle\psi|\tilde{U}(T)|\psi\rangle\big|^2 > 1 - \kappa^2T^2 + \mathcal{O}\big(\kappa^3T^3\big). \tag{2.28}$$

The PDD Fidelity

By using a suitable control scheme $\{g_j\}_{j=0}^{n_c-1}$, we have $\overline{H}^{(0)} = 0$ and $\tilde{U}(T) = \exp\big(-i\sum_{j=1}^{\infty}\overline{H}^{(j)}T\big)$ for $T = m \cdot t_c$ with $m \in \mathbb{N}$ and $t_c = n_c\Delta t$ (compare with (2.18)). Writing H_0 as λH_0, we obtain

$$\begin{aligned} F_e^{\mathsf{PDD}}(T) = \Big|\frac{1}{d}\operatorname{tr}\big(\tilde{U}(T)\big)\Big|^2 &= 1 - \frac{1}{d}\operatorname{tr}\Big(\big(\sum_{j=1}^{\infty}\overline{H}^{(j)}\big)^2\Big)T^2 + \ldots \\ &= 1 - \frac{1}{d}\operatorname{tr}\big(\big(\overline{H}^{(1)}\big)^2\big)T^2 + \mathcal{O}(\lambda^5 t_c^3 T^2). \end{aligned} \tag{2.29}$$

To evaluate this short time estimation, we have to calculate $\overline{H}^{(1)}$. A rough estimate based on the fact that $\overline{H}^{(1)}$ is a sum over $\mathcal{O}(n_c^2)$ terms of the form $\tilde{H}_i\tilde{H}_j$ leads to $\overline{H}^{(1)} = \mathcal{O}(\lambda^2 t_c)$. As before, we argue that a good approximation of $F_e^{\mathsf{PDD}}(T)$ is given by

$$F_{e\ \mathrm{app}}^{\mathsf{PDD}}(T) = \exp\Big(-\frac{1}{d}\operatorname{tr}\big(\big(\overline{H}^{(1)}\big)^2\big)T^2\Big), \tag{2.30}$$

as long as the fidelity has not become too small, i. e. for times T such that $F_e^{\mathsf{PDD}}(T) \gtrsim 0.1$. A strict lower bound on the worst case fidelity was given by Viola and Knill [VK05, theorem 3]:

$$F_w^{\mathsf{PDD}}(T) > 1 - \kappa^4 t_c^2 T^2 + \mathcal{O}\big(\kappa^5 t_c^3 T^2\big). \tag{2.31}$$

2.1.5 Open Quantum Systems

Up to this point we considered a closed quantum system S and the task of dynamical decoupling was the removal of inter-qudit couplings. In a real-world scenario, there will always be an interaction of the system with its surrounding environment E. As a result, entanglement between the system and the environment may arise causing the quantum system to evolve in a non-unitary way and to undergo a process called decoherence. Zanardi [Zan99] and Viola et al. [VKL99] proposed that dynamical decoupling techniques may be applied to decouple such systems from their environment. This subsection summarizes the main idea.

In this subsection we consider S to be an open system, i. e. to be part of a larger closed system formed by S and E together. Then the total system is defined on the Hilbert space $\mathcal{H}_{SE} = \mathcal{H}_S \otimes \mathcal{H}_E$, where $\mathcal{H}_S$ and $\mathcal{H}_E$ denote the system and environment Hilbert space. The Hamiltonian of the total system is given by the sum of the Hamiltonian H_0 of the system S and the Hamiltonian H_E of the environment, plus additional terms describing the couplings of the system with the environment,

$$H_{0,SE} = H_0 \otimes \mathcal{I}_E + \mathcal{I}_S \otimes H_E + \sum_\alpha S_\alpha \otimes E_\alpha. \tag{2.32}$$

Here, the E_α's are supposed to be linearly independent and, without loss of generality, the coupling operators S_α are assumed to be traceless. We proceed as in the case of a closed system: By applying a time-dependent local control $H_c(t)$ on the system S, the total Hamiltonian becomes time dependent,

$$H_{SE}(t) = H_{0,SE} + H_c(t) \otimes \mathcal{I}_E, \tag{2.33}$$

and we switch to the toggled frame defined by $\tilde{U}(t) = U_c^\dagger(t) \otimes \mathcal{I}_E \cdot U_{SE}(t)$, where $U_{SE}(t)$ denotes the time evolution operator of the combined system evolving according to (2.33), and $U_c(t)$ is defined as in (2.3) as the time evolution operator of the system evolving according to $H_c(t)$ alone. The time evolution in the toggled frame is determined by the toggled frame Hamiltonian

$$\tilde{H}_{SE}(t) = U_c^\dagger(t) H_0 U_c(t) \otimes \mathcal{I}_E + \mathcal{I}_S \otimes H_E + \sum_\alpha U_c^\dagger(t) S_\alpha U_c(t) \otimes E_\alpha. \tag{2.34}$$

A decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ that applies to all the coupling operators S_α satisfies

$$\overline{S}_\alpha^{(0)} = \frac{1}{n_c} \sum_{j=0}^{n_c} g_j^\dagger S_\alpha g_j = \mathfrak{c}_\alpha \cdot \frac{1}{d} \mathcal{I}, \text{ with } \mathfrak{c}_\alpha = \operatorname{tr}(S_\alpha), \tag{2.35}$$

for all α. If we use such a scheme in connection with the periodic dynamical decoupling (PDD) control strategy, we achieve the desired decoupling from the environment in lowest order AHT:

$$\overline{H}_{SE}^{(0)} = \overline{H}_0^{(0)} \otimes \mathcal{I}_E + \mathcal{I}_S \otimes \big(H_E + \sum_\alpha \frac{\mathfrak{c}_\alpha}{d} E_\alpha\big) \equiv \overline{H}_0^{(0)} \otimes \mathcal{I}_E + \mathcal{I}_S \otimes H'_E. \tag{2.36}$$

As it was discussed before, in the fast control limit, i. e. for $\Delta t \to 0$ and $m \to \infty$ with the total time $T = m \cdot n_c \Delta t$ held constant, lowest order AHT becomes exact and we obtain

$$\tilde{U}(T) = \exp\big(-i\overline{H}_{SE}^{(0)} T\big) = \exp\big(-i\overline{H}_0^{(0)} T\big) \otimes \exp\big(-iH'_E T\big). \tag{2.37}$$

For quantum memories the decoupling scheme should also satisfy $\overline{H}_0^{(0)} = \mathfrak{c} \cdot \frac{1}{d}\mathcal{I}$, so that (up to a global phase determined by $\mathfrak{c}$) $\tilde{U}(T) = \mathcal{I}_S \otimes \exp\big(-iH'_E T\big)$.

2.1.6 Noiseless Subsystems

Dynamical decoupling was defined as a dynamical control setting in which the time evolution of a quantum system is made to freeze. This is achieved by applying a decoupling scheme for the system Hamiltonian H_0 in a way specified by a certain control strategy (as for example PDD). As a result, the average Hamiltonian in the toggled frame vanishes. As discussed in the preceding subsection, for open quantum systems in principle the same method can be applied, provided that the decoupling scheme also applies to the coupling operators which are responsible for the interaction with the environment. A less demanding goal is the dynamical generation of a noiseless subsystem [Zan00; VKL00]. Instead of trying to protect the whole quantum system, control schemes are applied in order to preserve parts of the system. Information can then safely be stored by encoding it into such a part.

Let $G = \{\mathfrak{g}_j\}_{j=0}^{n_G-1}$ be a finite group of order n_G, and let $R : \mathfrak{g}_j \mapsto R(\mathfrak{g}_j) = g_j \in \mathsf{U}_d$ be a unitary representation of G on the d-dimensional system Hilbert space $\mathcal{H}_S$ (for our qudit quantum register $\mathcal{H}_S = \mathcal{H}_q^{\otimes n}$ and $d = q^n$). As explained in the introduction in subsection 1.2.3, the representation R decomposes into a sum of irreps of G,

$$R(G) = \bigoplus_{\nu\in\mathcal{J}} \tau_\nu \cdot D^{(\nu)}(G), \tag{2.38}$$

where the multiplicity of the irrep labeled by ν is denoted as τ_ν and the dimension of the irrep $D^{(\nu)}(G)$ is denoted by d_ν. Since the representation space of R is the system Hilbert space $\mathcal{H}_S$, any of the results of subsection 1.2.3 concerning the representation space apply to $\mathcal{H}_S$: There exists an orthonormal basis

$$\big\{|\nu\ l_\nu\ m_\nu\rangle \,|\, \nu \in \mathcal{J},\ l_\nu = 1\ldots\tau_\nu,\ m_\nu = 1\ldots d_\nu\big\}, \tag{2.39}$$

in which the operators g_j are block-diagonal, i. e.

$$g_j|\nu\ l_\nu\ m_\nu\rangle = D^{(\nu)}(g_j)|\nu\ l_\nu\ m_\nu\rangle = \sum_{m'_\nu=1}^{d_\nu} D^{(\nu)}_{m'_\nu m_\nu}(g_j)|\nu\ l_\nu\ m'_\nu\rangle. \tag{2.40}$$

The subspace of $\mathcal{H}_S$ which is spanned by the set of basis vectors with fixed ν is labeled by $\mathcal{H}_\nu$,

$$\mathcal{H}_\nu = \mathrm{span}\big\{|\nu\ l_\nu\ m_\nu\rangle \,|\, l_\nu = 1\ldots\tau_\nu, m_\nu = 1\ldots d_\nu\big\}, \tag{2.41}$$

and has the form of a tensor space ($|\nu\ l_\nu\ m_\nu\rangle = |l_\nu\rangle \otimes |m_\nu\rangle$), i. e. we write $\mathcal{H}_\nu = \mathcal{C}_\nu \otimes \mathcal{D}_\nu$, where the dimension of $\mathcal{C}_\nu$ is given by τ_ν and the dimension of $\mathcal{D}_\nu$ is given by d_ν. The Hilbert space decomposes as

$$\mathcal{H}_S = \bigoplus_{\nu\in\mathcal{J}} \mathcal{H}_\nu = \bigoplus_{\nu\in\mathcal{J}} \mathcal{C}_\nu \otimes \mathcal{D}_\nu. \tag{2.42}$$

Let $\mathcal{A} = R(\mathbb{C}G)$ denote the group algebra generated by R, and let its commutant $\mathcal{A}'$ be defined as the set of elements that commute with all the elements in $\mathcal{A}$, $\mathcal{A}' = \{V \in \mathcal{L}(\mathcal{H}_S) \,|\, VA = AV \text{ for all } A \in \mathcal{A}\}$. According to theorem 1.2.7 the elements of $\mathcal{A}$ and $\mathcal{A}'$ become block-diagonal in the $\{|\nu\ l_\nu\ m_\nu\rangle\}$-basis,

$$\mathcal{A} \cong \bigoplus_{\nu\in\mathcal{J}} \mathbb{1}_{\tau_\nu} \otimes \mathrm{Mat}(d_\nu \times d_\nu, \mathbb{C}) \tag{2.43}$$

$$\mathcal{A}' \cong \bigoplus_{\nu\in\mathcal{J}} \mathrm{Mat}(\tau_\nu \times \tau_\nu, \mathbb{C}) \otimes \mathbb{1}_{d_\nu}, \tag{2.44}$$

where $\mathbb{1}_n$ denotes an $n \times n$ dimensional identity matrix and $\mathrm{Mat}(n \times n, \mathbb{C})$ denotes the set of $n \times n$ matrices with entries in $\mathbb{C}$.

We start by describing the idea of a noiseless subsystem [ZR97a; ZR97b; LCW98]. Let us imagine that the quantum system S under consideration is open and its Hamiltonian is given by equation (2.32). If H_0 and the coupling operators S_α are elements of $\mathcal{A}$, we have

$$H_0 = \bigoplus_{\nu\in\mathcal{J}} \mathcal{I}_{\mathcal{C}_\nu} \otimes D^{(\nu)}(H_0), \tag{2.45}$$

and corresponding expressions for the S_α. It follows that information encoded in the $\mathcal{C}_\nu$-part of the subspace $\mathcal{H}_\nu$ remains unchanged over time: Let the information be described by $\rho = \sum_{i,j=1}^{\tau_\nu} \rho_{ij}|i\rangle\langle j|$ and let it be encoded in $\mathcal{H}_\nu$ as

$$\rho_{\mathcal{C}_\nu} \otimes \sigma_{\mathcal{D}_\nu} = \sum_{i,j=1}^{\tau_\nu} \sum_{k,l=1}^{d_\nu} \rho_{ij}\sigma_{kl}|\nu i_\nu k_\nu\rangle\langle\nu j_\nu l_\nu| \tag{2.46}$$

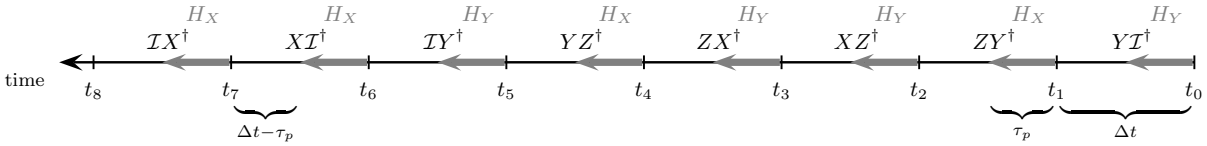


Figure 2.3: Schematic representation of an Euler decoupling cycle based on the decoupling set $\mathcal{G} = \{\mathcal{I}, X, Y, Z\}$ and the generators $\Gamma = \{X, Y\}$. The above cycle of length $t_c = |\mathcal{G}| \cdot |\Gamma| \cdot \Delta t = 8\Delta t$ is based on the Eulerian cycle on the Cayley graph of $\mathcal{G}$ with respect to Γ shown in figure 2.4. It is repeated over and over again. H_X denotes a potentially time-dependent control Hamiltonian which generates the generator X, i. e. up to a phase we have $X = \mathcal{T}\exp\big(-i\int_0^{\tau_p} H_X(t')dt'\big)$. H_Y is defined analogously. As a result, the applied control generates the gates denoted in the second line.

for some arbitrary $\sigma = \sum_{k,l=1}^{d_\nu} \sigma_{kl}|k\rangle\langle l|$. Denoting the time evolution operator of the total system as $U_{SE}(t)$ and assuming that the environment is initially not entangled with the system, we obtain

$$\begin{aligned}\big((\rho_{\mathcal{C}_\nu} \otimes \sigma_{\mathcal{D}_\nu})_S \otimes \tau_E\big)(t) &= U_{SE}(t)\ \big((\rho_{\mathcal{C}_\nu} \otimes \sigma_{\mathcal{D}_\nu})_S \otimes \tau_E\big)\ U_{SE}^\dagger(t) \\ &= \exp\Big(-it\,\mathcal{I}_{\mathcal{C}_\nu} \otimes \big(D^{(\nu)}(H_0) \otimes \mathcal{I}_E + \mathcal{I}_{\mathcal{D}_\nu} \otimes H_E + \sum_\alpha D^{(\nu)}(S_\alpha) \otimes E_\alpha\big)\Big)\times \\ &\qquad\qquad \rho_{\mathcal{C}_\nu} \otimes \sigma_{\mathcal{D}_\nu} \otimes \tau_E\ \exp\big(+it\ldots\big) \\ &= \rho_{\mathcal{C}_\nu} \otimes U_{\mathcal{D}_\nu E}(t)\,(\sigma_{\mathcal{D}_\nu} \otimes \tau_E)\,U_{\mathcal{D}_\nu E}^\dagger(t). \end{aligned} \tag{2.47}$$

Hence the $\{\mathcal{C}_\nu\}_{\nu\in\mathcal{J}}$ are indeed noiseless (or decoherence-free) subsystems. In the special case that $d_\nu = 1$, $\mathcal{C}_\nu$ is a noiseless subspace.

Unfortunately, the interactions of a typical quantum system hardly allow the existence of large noiseless subsystems. Hence Zanardi and Viola et al. [Zan00; VKL00] came up with the idea to modify the interactions in terms of dynamical control, such that the resulting symmetrized dynamics allows for larger noiseless subsystems. Let a control scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ of length $n_c = n_G$ be defined by a unitary projective representation R of a group $G = \{\mathfrak{g}_j\}_{j=0}^{n_G-1}$ acting on the system Hilbert space $\mathcal{H}_S$, i. e. $g_j = R(\mathfrak{g}_j)$. As a result of the applied control scheme (let us assume here for simplicity that we use the PDD control strategy in the fast control limit), the operators H_0 and S_α become

$$\Pi_{\mathcal{G}}(X) = \frac{1}{n_c}\sum_{j=0}^{n_c-1} g_j^\dagger X g_j, \tag{2.48}$$

with $X \in \{H_0, S_\alpha\}$. Since $\Pi_{\mathcal{G}}(X)$ commutes with any element g_j of the control scheme, it follows that $\Pi_{\mathcal{G}}(X)$ is in $\mathcal{A}'$. As a result, the subsystems $\{\mathcal{D}_\nu\}_{\nu\in\mathcal{J}}$ are noiseless. The standard decoupling scenario ($\Pi_{\mathcal{G}}(X) = \mathfrak{c}_X \cdot \mathcal{I}$ for all $X \in \{H_0, S_\alpha\}$) is included as a special case: If the representation is irreducible, the set $\mathcal{J}$ consist of only one element ν and we have $\tau_\nu = 1$ and $d_\nu = \dim(\mathcal{H}_S)$.

2.1.7 Bounded Controls

The current chapter of this thesis deals with dynamical decoupling in the bang-bang control scenario, i. e. we assume a strong control Hamiltonian such that any applied control pulse may be considered as being applied instantaneously. Of course such a scenario is an idealization. This subsection discusses the effects of bounded controls.

In order to analyze the effects of bounded controls, let us assume we apply the control scheme $\{g_j\}_{j=0}^{n_c-1}$ of length n_c using the fundamental control strategy (also called periodic dynamical decoupling), i. e. we repeat the pulse sequence $p_0, \ldots, p_{n_c-1}$, with $p_j = g_{j+1}g_j^\dagger$ for $j = 0, \ldots, n_c - 1$ and $g_{n_c} = g_0$,

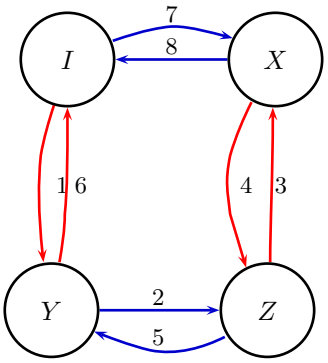


Figure 2.4: Eulerian cycle on the Cayley graph of $\mathcal{G} = \{\mathcal{I}, X, Y, Z\}$ with respect to the generators $\Gamma = \{X, Y\}$. The edges colored by X are depicted in blue, those colored by Y are shown in red.

over and over again[†]. But instead of applying the pulses $p_{j \bmod n_c}$ instantaneously at times $j \cdot \Delta t$, $j \in \mathbb{N}_0$, we now assume that each pulse is generated by switching on a possibly time-dependent control Hamiltonian $H_j(t')$ for a time $\tau_p < \Delta t$ during the time interval $[j \cdot \Delta t, j \cdot \Delta t + \tau_p]$ such that $p_j = p_j(\tau_p) = \mathcal{T} \exp\bigl(-i \int_0^{\tau_p} H_j(t') dt'\bigr)$. As a result, after $m \in \mathbb{N}$ such cycles of length $t_c = n_c \Delta t$, the time evolution operator in the toggled frame is given by $\tilde{U}(T = m t_c) = \exp\bigl(-i \overline{H} T\bigr)$, where in lowest order AHT $\overline{H}$ is given by equation (2.12a):

$$\begin{aligned} \overline{H}^{(0)} &= \frac{1}{t_c} \int_0^{t_c} dt_1 \tilde{H}(t_1) \\ &= \frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger \left(\frac{1}{\Delta t} \int_0^{\tau_p} p_j^\dagger(t') \; H_0 \; p_j(t') dt' + H_0 \cdot (1 - \tau_p / \Delta t) \right) g_j . \end{aligned} \tag{2.49}$$

For $\tau_p \to 0$ this expression reduces to the corresponding expression (2.17) of the bang-bang scenario. If the control scheme $\{g_j\}_{j=0}^{n_c-1}$ is for the simulation of the Hamiltonian H_0' with the system Hamiltonian H_0, this means that for $\tau_p = 0$ we would get

$$\overline{H}^{(0)} = \frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger H_0 g_j = H_0' + \mathfrak{c} \cdot \frac{1}{d} \mathcal{I}, \tag{2.50}$$

with $\mathfrak{c} = \mathrm{tr}(H_0) - \mathrm{tr}(H_0')$. For finite τ_p the first term within the braces in equation (2.49) depends on j and prevents the bang-bang control condition from above to be fulfilled.

If the elements of the control scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ are defined by a unitary projective representation R of a group $G = \{\mathfrak{g}_j\}_{j=0}^{n_c-1}$ acting on the system Hilbert space $\mathcal{H}_S$, i. e. if we have $g_j = R(\mathfrak{g}_j)$, this problem may be circumvented by using the so-called Eulerian decoupling proposed by Viola and Knill [VK03]. Before we describe their idea, we have to make some definitions. First, let $\mathcal{A} = R(\mathbb{C}G)$ denote the corresponding group algebra, and let its commutant $\mathcal{A}'$ be defined as the set of elements that commutes with all the elements in $\mathcal{A}$. Second, the Cayley graph of $\mathcal{G}$ with respect of to a set of generators is defined as follows:

Definition 2.1.2 (Cayley graph)**.** Let $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ be a finite group of order n_c, and let $\Gamma = \{p_i\}_{i=1}^{|\Gamma|}$ be a generating set. Then the Cayley graph of $\mathcal{G}$ with respect to Γ is defined as the directed multigraph whose edges are colored by the generators $p_i \in \Gamma$, such that vertex g_j is joined to vertex g_k by an edge of color p_i if and only if $g_k = p_i g_j$ (or $p_i = g_k g_j^\dagger$).

Last, an Eulerian path in the Cayley graph is defined as a path which uses each edge exactly once. The proposal of Viola and Knill is now to replace the basic PDD cycle of length $n_c = |\mathcal{G}|$ by a cycle

[†]In subsection 2.1.3 the original definition of p_j was $p_j = g_j g_{j-1}^\dagger$. Here it is changed it to $p_j = g_{j+1} g_j^\dagger$ in order to close the basic cycle with g_0 instead of g_{n_c-1}.

corresponding to an Eulerian path of length $n_c \cdot |\Gamma|$. As a consequence, instead of (2.49), we obtain in lowest order AHT

$$\overline{H}^{(0)} = \frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger \Big(\frac{1}{|\Gamma|} \sum_{i=1}^{|\Gamma|} \frac{1}{\Delta t} \Big(\int_0^{\tau_p} p_i^\dagger(t') \; H_0 \; p_i(t') dt' + H_0 \cdot (\Delta t - \tau_p) \Big) \Big) g_j, \tag{2.51}$$

where, as before, $p_i(t) = \mathcal{T} \exp\big(-i \int_0^t H_i(t')dt'\big)$ is generated using a possibly time-dependent control Hamiltonian H_i $(i = 1, \ldots, |\Gamma|)$. By using the definitions

$$\Pi_{\mathcal{G}}(X) = \frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger X g_j \tag{2.52}$$

$$F_\Gamma(X) = \frac{1}{|\Gamma|} \sum_{i=1}^{|\Gamma|} \frac{1}{\tau_p} \int_0^{\tau_p} p_i^\dagger(t') \; X \; p_i(t') dt', \tag{2.53}$$

this expression can be written as

$$\overline{H}^{(0)} = \Pi_{\mathcal{G}}\big(F_\Gamma(H_0)\big) \cdot \frac{\tau_p}{\Delta t} + \Pi_{\mathcal{G}}\big(H_0\big) \cdot (\Delta t - \tau_p)/\Delta t. \tag{2.54}$$

Due to the following theorem this is equal to $\Pi_{\mathcal{G}}\big(H_0\big)$ and we arrive at the standard control condition (2.50) of the bang-bang scenario.

Theorem 2.1.2 ([VK03])**.** *Let X be any time-independent operator acting on the system Hilbert space $\mathcal{H}_S$. If the control Hamiltonians $H_i(t)$ are in the group algebra $\mathcal{A} = R(\mathbb{C}G)$ for all $t \in [0, \tau_p]$ and all $i \in \{1, \ldots, |\Gamma|\}$, then $\Pi_{\mathcal{G}}\big(F_\Gamma(X)\big) = \Pi_{\mathcal{G}}\big(X\big)$.*

Proof. If $H_i(t) \in \mathcal{A}$ it follows that $p_i(t) \in \mathcal{A}$ for all $t \in [0, \tau_p]$ and all $i \in \{1, \ldots |\Gamma|\}$. Hence, $F_\Gamma(Y) = Y$ for any time-independent operator $Y \in \mathcal{A}'$. We are now going to show that $Q(X) = \Pi_{\mathcal{G}}\big(F_\Gamma(X)\big)$ is a projector. First, we note that $Q^2(X) = \Pi_{\mathcal{G}}\big(F_\Gamma\big(\Pi_{\mathcal{G}}\big(F_\Gamma(X)\big)\big)\big) = \Pi_{\mathcal{G}}\big(\Pi_{\mathcal{G}}\big(F_\Gamma(X)\big)\big)$, which follows from $F_\Gamma(Y) = Y$ for $Y \in \mathcal{A}'$. By using the fact that $\Pi_{\mathcal{G}}$ is a projector, we find that $Q^2(X) = Q(X)$. Since the range of Q is in $\mathcal{A}'$, we have $Q = \Pi_{\mathcal{G}}$ if and only if Q acts on $\mathcal{A}'$ as the identity. Let $Y \in \mathcal{A}'$, then $Q(Y) = \Pi_{\mathcal{G}}(Y) = Y$. □

As an example we consider the decoupling scheme for one qubit given by the Pauli group $\mathcal{G} = \{\mathcal{I}, X, Y, Z\}$. As a set of generators we choose $\Gamma = \{X, Y\}$. The Cayley graph of $\mathcal{G}$ with respect to Γ is shown in figure 2.4. An Eulerian path is obtained by following the numbers $1, \ldots, 8$. The decoupling cycle corresponding to this path is depicted in figure 2.3.

The above method increases the length of a basic PDD cycle by a factor $|\Gamma|$. For local system Hamiltonians shorter decoupling schemes may be devised using Eulerian orthogonal arrays [Woc06]. For a geometric perspective on the theory of decoupling with bounded controls we refer to [Che06].

2.2 Decoupling Schemes

A decoupling scheme for the system Hamiltonian H_0 acting on the system Hilbert space $\mathcal{H}_S$ was defined in definition 2.1.1 as a set of unitaries $\{g_j\}_{j=0}^{n_c-1}$ and relative times $\{\Delta t_j\}_{j=0}^{n_c-1}$ such that

$$\frac{1}{t_c} \sum_{j=0}^{n_c-1} g_j^\dagger H_0 g_j \Delta t_j = \mathrm{tr}(H_0) \cdot \frac{1}{d} \mathcal{I}, \tag{2.55}$$

where $t_c = \sum_{j=0}^{n_c-1} \Delta t_j$ and $d = \dim(\mathcal{H}_S)$. In this section we give an overview over known decoupling schemes for different types of system Hamiltonians. All these schemes work with constant relative time

intervals, i. e. $\Delta t_j = \Delta t$ for all $j \in \{0, \ldots, n_c - 1\}$. Since the quantum system under consideration forms a quantum register consisting of n qudits of dimension q we have $\mathcal{H}_S = \mathcal{H}_q^{\otimes n}$ and the local control assumption requires the unitaries g_j to be of the form $g_j = g_1^{(1,j)} \otimes g_2^{(2,j)} \otimes \cdots \otimes g_n^{(n,j)}$, where $g_k^{(i,j)}$ denotes the unitary $g^{(i,j)} \in \mathsf{U}_q$ being applied to the k-th qudit.

2.2.1 General Hamiltonians

We start with decoupling schemes which apply to all traceless Hamiltonians H_0 acting on $\mathcal{H}_S$.

Definition 2.2.1. An annihilator is a decoupling scheme $\{g_j, \Delta t_j\}_{j=0}^{n_c^2-1}$ satisfying

$$\frac{1}{t_c} \sum_{j=0}^{n_c^2-1} g_j^\dagger H_0 g_j \Delta t_j = 0, \tag{2.56}$$

for all traceless system Hamiltonians H_0.

It was shown in [WRJB02a] that an annihilator has to contain at least $n_c = \dim(\mathcal{H}_S)^2$ elements g_j and that the relative times for such a minimal annihilator have to be equal, i. e. $\Delta t_j = \Delta t$ for all $j \in \{0, \ldots, n_c - 1\}$. Annihilators can be found using the following group-theoretic averaging procedure [Zan99; VKL99].

Theorem 2.2.1. *Let $G = \{\mathfrak{g}_j\}_{j=0}^{n_c^2-1}$ be a finite group of order n_c, and let $R : \mathfrak{g}_j \mapsto g_j \in \mathsf{U}_d$ be an irreducible representation of G on a d-dimensional Hilbert space $\mathcal{H}_S$. Then, for any $H_0 \in \mathcal{L}(\mathcal{H}_S)$,*

$$\Pi_{R(G)}(H_0) \equiv \frac{1}{n_c} \sum_{j=0}^{n_c^2-1} g_j^\dagger H_0 g_j = \mathrm{tr}(H_0) \cdot \frac{1}{d} \mathcal{I}. \tag{2.57}$$

Proof. First we note that the left hand side of the above equation commutes with all the unitaries g_j. Since the g_j form an irreducible representation, Schur's lemma (theorem 1.2.4) tells us that the only operator commuting with all the g_j is a multiple of the identity. The correct factor is obtained by taking the trace on both sides of the equation. □

This theorem was shown in [WRJB02a] to hold for irreducible projective representations as well. Since, by definition, any nice error basis (see definition 1.2.16) forms an irreducible projective representation, it can be used as an annihilator. A particular example for a nice error basis — and hence for an annihilator — for $\mathcal{H}_S = \mathcal{H}_q^{\otimes n}$ is the set of Pauli operators,

$$\mathcal{P}_q^n = \{XZ(\vec{a}) \mid \vec{a} \in \mathbb{F}_q^{2n}\}, \tag{2.58}$$

as defined in section 1.2.

Decoupling according to theorem 2.2.1 corresponds to the special case of a dynamical generated noiseless subsystem (subsection 2.1.6) which is identical with the whole system.

2.2.2 Local Hamiltonians

Let us first define a map mapping an operator of the form $A = A_1^{(1)} \otimes \cdots \otimes A_s^{(s)}$ acting on $\mathcal{H}_q^{\otimes s}$ to an operator acting on $\mathcal{H}_q^{\otimes n}$ with $n \geq s$ via

$$A \mapsto \big[A\big]_{(k_1,k_2,\ldots,k_s)} = A_{k_1}^{(1)} \otimes \cdots \otimes A_{k_s}^{(s)} \otimes \mathcal{I}_{\{1,2,\ldots,n\}\setminus\{k_1,\ldots,k_s\}}, \tag{2.59}$$

for any $1 \leq k_1 < k_2 < \cdots < k_s \leq n$. Here, the index i in $A_i^{(j)}$ indicates that the operator $A^{(j)} \in \mathcal{L}(\mathcal{H}_q)$ acts on the i-th qudit. Using this kind of notation, a t-local Hamiltonian is defined as follows:

$$H_0 = \sum_{s=1}^{t} \sum_{k_1=1}^{n-s+1} \sum_{k_2=k_1+1}^{n-s+2} \cdots \sum_{k_s=k_{s-1}+1}^{n} \sum_{\vec{a} \in \mathbb{F}_q^{2s} \setminus \{\vec{0}\}} J_{\vec{a}}^{k_1 \ldots k_s} \big[XZ(\vec{a})\big]_{(k_1 \ldots k_s)}. \tag{2.60}$$

Since the Pauli operators form an operator basis, any Hamiltonian that couples no more than t of the qudits can be written as in (2.60). Decoupling schemes for t-local qubit Hamiltonians ($q = 2$) have been devised by Leung [Leu02] in terms of Hadamard matrices and by Stollsteimer and Mahler [SM01] using orthogonal arrays [HSS99]. The orthogonal array approach was generalized to qudits by Wocjan et al. in [WRJB02b]. Eventually it was shown by Rötteler and Wocjan [RW06] that both methods are equivalent. We proceed explaining the generalized orthogonal array approach.

Definition 2.2.2 (Orthogonal arrays)**.** Let $\mathcal{A}$ be an alphabet containing a symbols. An orthogonal array $OA_\lambda(n_c, n, t, a)$ with a levels, strength t and index λ is an $n \times n_c$ matrix $M = (m_{ij})$ with entries from $\mathcal{A}$ if any $s \times n_c$ sub-matrix (obtained from M by selecting s rows) contains any possible s-tuple of elements from $\mathcal{A}$ exactly λ times as a column.

Let $\{u_i\}_{i=0}^{q^2-1}$ denote an annihilator for the one qudit Hilbert space $\mathcal{H}_q$ (for example we could choose the set of Pauli operators, i. e. $\{u_i\}_{i=0}^{q^2-1} = \mathcal{P}_q$). Given an $OA_\lambda(n_c, n, t, q^2)$ with q^2 levels, a control scheme $\{g_j\}_{j=0}^{n_c-1}$ can be obtained as follows: The j-th unitary g_j is constructed using the $(j+1)$-th column of the orthogonal array as $g_j = u_{m_{1,j+1}} \otimes u_{m_{2,j+1}} \otimes \cdots \otimes u_{m_{n,j+1}}$. The following theorem due to Wocjan and Rötteler [WRJB02b; RW06] shows that such a control scheme is in fact a decoupling scheme for any t-local Hamiltonian.

Theorem 2.2.2. *A control scheme $\{g_j\}_{j=0}^{n_c-1}$ constructed from an $OA_\lambda(n_c, n, t, q^2)$ with q^2 levels and strength t as described above, is a decoupling scheme for all t-local Hamiltonians acting on $\mathcal{H}_S = \mathcal{H}_q^{\otimes n}$.*

Proof. The annihilator $\{u_i\}_{i=0}^{q^2-1}$ for the one-qudit Hilbert space $\mathcal{H}_q$ consists of the elements of a nice error basis for operators acting on $\mathcal{H}_q$. Hence the collection of all s-fold tensor products of the u_i's forms a nice error basis for $\mathcal{H}_q^{\otimes s}$ and we obtain

$$\frac{1}{q^{2s}} \sum_{i_1,\ldots,i_s=0}^{q^2-1} \big(u_{i_1}^\dagger \ldots u_{i_s}^\dagger\big) H\big(u_{i_1} \ldots u_{i_s}\big) = 0 \tag{2.61}$$

for all traceless Hamiltonians H acting on $\mathcal{H}_q^{\otimes s}$. Let us pick now the term characterized by $(k_1 \ldots k_s)$ and $\vec{a}$ from the t-local H_0 given by (2.60). For the control scheme $\{g_j\}_{j=0}^{n_c-1}$ constructed from the OA we obtain

$$\begin{aligned}
&\frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger \; J_{\vec{a}}^{k_1\ldots k_s} \big[XZ(\vec{a})\big]_{(k_1\ldots k_s)} \; g_j \\
&= \frac{1}{n_c} \sum_{j=1}^{n_c} \big(u_{m_{1,j}}^\dagger \otimes \cdots \otimes u_{m_{n,j}}^\dagger\big) \; J_{\vec{a}}^{k_1\ldots k_s} \big[XZ(\vec{a})\big]_{(k_1\ldots k_s)} \; \big(u_{m_{1,j}} \otimes \cdots \otimes u_{m_{n,j}}\big) \\
&= \left[J_{\vec{a}}^{k_1\ldots k_s} \frac{1}{n_c} \sum_{j=1}^{n_c} \big(u_{m_{k_1,j}}^\dagger \otimes \cdots \otimes u_{m_{k_s,j}}^\dagger\big) \; XZ(\vec{a}) \; \big(u_{m_{k_1,j}} \otimes \cdots \otimes u_{m_{k_s,j}}\big) \right]_{(k_1\ldots k_s)} \\
&= \left[J_{\vec{a}}^{k_1\ldots k_s} \frac{1}{q^{2s}} \sum_{i_1,\ldots,i_s=0}^{q^2-1} \big(u_{i_1}^\dagger \otimes \cdots \otimes u_{i_s}^\dagger\big) \; XZ(\vec{a}) \; \big(u_{i_1} \otimes \cdots \otimes u_{i_s}\big) \right]_{(k_1\ldots k_s)} = 0.
\end{aligned} \tag{2.62}$$

The last line is obtained by noting that the OA contains each possible s-tuple (with $s \leq t$) with entries in $\mathbb{F}_q^2$ equally often. □

Remark. A decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ for a t-local Hamiltonian acting on $\mathcal{H}_q^{\otimes n}$ based on an orthogonal array $OA(n_c, n, t, q^2)$ can be extended to a decoupling scheme $\{g'_j\}_{j=0}^{n_c-1}$ for a t-local Hamiltonian acting on $\mathcal{H}_q^{\otimes n+1}$ by setting $g'_j = (g_j)_{\{1\ldots n\}} \otimes \mathcal{I}_{n+1}$, as long as there are no local terms in the Hamiltonian which act only the $(n+1)$-th qudit.

Physical interactions are typically described by 2-local Hamiltonians. Hence orthogonal arrays of strength two are of special importance. Using a construction method based on Hamming codes [HSS99, chapter 5.3], orthogonal arrays $OA\big(s^i, (s^i-1)/(s-1), 2, s\big)$, with s being a prime power (here $s = q^2$) and $i \geq 2$, can be obtained. It follows that any 2-local Hamiltonian acting on up to n qudits of dimension q can be decoupled using a decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ of length n_c, where an upper bound on n_c is given by $n_c \leq n(s-1)s + 2s - s^2$. Even though this bound is far from optimal (orthogonal arrays exist which cannot be obtained by the Hamming code method), it shows that the length of a decoupling scheme scales linearly with the number of qudits. In the appendix A.2 we list the orthogonal arrays $OA(16, 5, 2, 4)$, $OA(32, 9, 2, 4)$ and $OA(48, 13, 2, 4)$, which can be used to decouple up to 5, 9 and 13 qubits, respectively.

Diagonal Couplings

Let us consider now the special case of an n-qubit Hamiltonian H_0 involving only bipartite couplings,

$$H_0 = \sum_{k_1=1}^{n-1} \sum_{k_2=k_1+1}^{n} \sum_{\vec{a}\in\mathbb{F}_2^2\setminus\{\vec{0}\}} J_{\vec{a}}^{k_1,k_2} \big[XZ(\vec{a}) \otimes XZ(\vec{a})\big]_{(k_1,k_2)}. \tag{2.63}$$

These kind of couplings are called diagonal couplings, since the coefficient matrix $J_{\vec{a}}$ is diagonal when compared with the one for the general case $\big(\sum_{\vec{a},\vec{c}} J_{\vec{a},\vec{c}} XZ(\vec{a}) \otimes XZ(\vec{c})\big)$. It was shown by Stollsteimer and Mahler [SM01] that such Hamiltonians can be decoupled using decoupling schemes constructed from difference schemes [HSS99, chapter 6]. The advantage over corresponding decoupling schemes using orthogonal arrays is the shorter length of such schemes. We generalize this approach to the qudit case. For qudits of dimension $q \geq 3$ let us consider the Hamiltonian

$$H_0 = \sum_{k_1=1}^{n-1} \sum_{k_2=k_1+1}^{n} \sum_{\vec{a}\in\mathbb{F}_q^2\setminus\{\vec{0}\}} J_{\vec{a}}^{k_1,k_2} \big[XZ(\vec{a}) \otimes XZ^\dagger(\vec{a})\big]_{(k_1,k_2)}. \tag{2.64}$$

For H_0 to be Hermitian, the coefficients must satisfy $J_{\vec{a}}^{k_1,k_2} = J_{-\vec{a}}^{k_1,k_2}$ since $XZ(-\vec{a}) \otimes XZ^\dagger(-\vec{a}) = XZ(\vec{a})^\dagger \otimes XZ(\vec{a})$. It follows that the Hamiltonian is symmetric with respect to k_1 and k_2.

Remark. Note that interactions of the form (2.64) might be of interest for quantum computation, since the swap gate, $U_{\mathrm{SWAP}}|\phi\rangle \otimes |\psi\rangle = |\psi\rangle \otimes |\phi\rangle$, which can be written as

$$U_{\mathrm{SWAP}} = \frac{1}{q^2} \sum_{\vec{a}\in\mathbb{F}_q^2} XZ(\vec{a}) \otimes XZ^\dagger(\vec{a}), \tag{2.65}$$

can be generated (up to a global phase) as $U_{\mathrm{SWAP}} = \exp\big(-iH_{\mathrm{SWAP}}\pi/2\big)$ by the interaction

$$H_{\mathrm{SWAP}} = \frac{1}{q} \sum_{\vec{a}\in\mathbb{F}_q^2\setminus\{\vec{0}\}} XZ(\vec{a}) \otimes XZ^\dagger(\vec{a}) \tag{2.66}$$

which is of the form (2.64). In the qubit case the square root swap gate — $\exp\big(-iH_{\mathrm{SWAP}}\pi/4\big)$ — in connection with all single qubit gates forms a universal set of gates.

Definition 2.2.3 (Difference schemes)**.** A difference scheme $D(n_c, n, s)$ based on a finite abelian group $(\mathcal{A}, +)$ of order s is an $n \times n_c$ matrix $M = (m_{ij})$ such that for all $1 \leq i < j \leq n$, the vector difference between the i-th and the j-th row contains each element of $\mathcal{A}$ equally often.

Necessarily n_c is a multiple of s. It can be shown that if a difference scheme $D(n_c, n, s)$ exists, then $n \leq n_c$ [HSS99, chapter 6].

Let the set of Pauli operators be given by $\mathcal{P}_q = \{XZ(\vec{a}) \mid \vec{a} \in \mathbb{F}_q^2\}$. Given a $D(n_c, n, q^2)$ based on $\mathbb{F}_q^2$, a control scheme $\{g_j\}_{j=0}^{n_c-1}$ can be constructed as follows: The j-th unitary g_j is constructed using the $(j+1)$-th column of the difference scheme as $g_j = XZ(m_{1,j+1}) \otimes XZ(m_{2,j+1}) \otimes \cdots \otimes XZ(m_{n,j+1})$.

Theorem 2.2.3. *A control scheme $\{g_j\}_{j=0}^{n_c-1}$ constructed from a $D(n_c, n, q^2)$ as described above, is a decoupling scheme for all n-qudit Hamiltonians involving diagonal qudit-qudit couplings as in* (2.64).

Proof. Let us pick a single term characterized by (k_1, k_2) and $\vec{a}$ from H_0 in (2.64). For the control scheme $\{g_j\}_{j=0}^{n_c-1}$ constructed from a $D(n_c, n, q^2) = (m_{ij})$, $i = 1 \dots n$, $j = 1 \dots n_c$, $m_{ij} \in \mathbb{F}_q^2$, we obtain

$$
\begin{aligned}
&\frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger \; J_{\vec{a}}^{k_1,k_2} \big[XZ(\vec{a}) \otimes XZ^\dagger(\vec{a})\big]_{(k_1,k_2)} \; g_j \\
&= \frac{1}{n_c} \sum_{j=1}^{n_c} XZ^\dagger(m_{1,j}) \otimes \cdots \otimes XZ^\dagger(m_{n,j}) \; J_{\vec{a}}^{k_1,k_2} \big[XZ(\vec{a}) \otimes XZ^\dagger(\vec{a})\big]_{(k_1,k_2)} \; XZ(m_{1,j}) \otimes \cdots \otimes XZ(m_{n,j}) \\
&= \left[J_{\vec{a}}^{k_1,k_2} \frac{1}{n_c} \sum_{j=1}^{n_c} \big(XZ^\dagger(m_{k_1,j}) \otimes XZ^\dagger(m_{k_2,j})\big) \big(XZ(\vec{a}) \otimes XZ(\vec{a})^\dagger\big) \big(XZ(m_{k_1,j}) \otimes XZ(m_{k_2,j})\big) \right]_{(k_1,k_2)}.
\end{aligned}
$$

Using the symplectic inner product as in (1.30), the order of the Pauli operators can be inverted leading to

$$
\begin{aligned}
&= \left[J_{\vec{a}}^{k_1,k_2} XZ(\vec{a}) \otimes XZ^\dagger(\vec{a}) \right]_{(k_1,k_2)} \frac{1}{n_c} \sum_{j=1}^{n_c} \omega^{(\vec{a}, m_{k_1,j} - m_{k_2,j})_{sp}} \\
&= \left[J_{\vec{a}}^{k_1,k_2} XZ(\vec{a}) \otimes XZ^\dagger(\vec{a}) \right]_{(k_1,k_2)} \frac{1}{q^2} \sum_{\vec{d} \in \mathbb{F}_q^2} \omega^{(\vec{a}, \vec{d})_{sp}} = 0.
\end{aligned} \tag{2.67}
$$

The last line is obtained by noting that in the difference scheme (m_{ij}), with $m_{ij} \in \mathbb{F}_q^2$, the vector difference between row k_1 and k_2 contains each element in $\mathbb{F}_q^2$ exactly n_c/s times. As it can be seen from the last two lines, the position of the dagger operator is not important: The decoupling scheme also eliminates couplings of the form $XZ(\vec{a})^\dagger \otimes XZ(\vec{a})$. □

Construction methods for difference schemes $D(q^m, q^m, q^2)$ with q prime and $m \geq 2$ are known (see for example [HSS99, chapter 6.1]). It follows that any Hamiltonian with diagonal couplings between up to n qudits of dimension q can be decoupled using a decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ of length n_c, where an upper bound on n_c is given by $n_c \leq nq - q$. This bound is of the order $\mathcal{O}(nq)$ and has to be compared with the bound for orthogonal arrays which was $\mathcal{O}(nq^4)$. In the appendix A.1 we list difference schemes $D(4\lambda, 4\lambda, 4)$ for $\lambda \in \{1, 2, 3, 4\}$, which can be used in order to decouple up to 4λ qubits, respectively.

There exist diagonal couplings for which even shorter decoupling schemes can be devised. A famous example are dipolar inter-qubit couplings,

$$
H_0 = \sum_{k_1=1}^{n-1} \sum_{k_2=k_1+1}^{n} J^{k_1,k_2} \big[2Z \otimes Z - X \otimes X - Y \otimes Y\big]_{(k_1,k_2)}, \tag{2.68}
$$

for which a decoupling scheme of constant length $n_c = 3$ is given by the set $\{g_j\}_{j=1}^3$ [WHH68] of non-selective $\pi/2$ pulses,

$$
g_j = \exp\Big(-\frac{i}{2} \alpha_j \frac{\pi}{2}\Big)^{\otimes n}, \quad \text{with } \alpha_j = X, Y, Z \text{ for } j = 1, 2, 3. \tag{2.69}
$$

The $\pi/2$ pulses convert the diagonal terms in the Hamiltonian H_0 in a cyclic manner, thereby achieving the decoupling condition $\tilde{H}_1 + \tilde{H}_2 + \tilde{H}_3 = 0$ with $\tilde{H}_j = g_j^\dagger H_0 g_j$.

2.2.3 Selective Decoupling

In the preceding subsection, among others, decoupling schemes for general and diagonal Hamiltonians involving only bipartite inter-qudit couplings have been presented. These schemes turn off all qudit-qudit couplings in

$$H_0 = \sum_{k_1=1}^{n-1} \sum_{k_2=k_1+1}^{n} \sum_{\vec{a},\vec{b}\in\mathbb{F}_q^2\setminus\{\vec{0}\}} J_{\vec{a},\vec{b}}^{k_1,k_s} \big[XZ(\vec{a}) \otimes XZ(\vec{b})\big]_{(k_1,k_2)}, \tag{2.70}$$

or its diagonal counterpart (2.64). Under certain circumstances we might want to keep one (or more than one) particular coupling alive, i. e. we want to simulate the Hamiltonian

$$H_0' = \sum_{\vec{a},\vec{b}\in\mathbb{F}_q^2\setminus\{\vec{0}\}} J_{\vec{a},\vec{b}}^{k_1,k_s} \big[XZ(\vec{a}) \otimes XZ(\vec{b})\big]_{(k_1,k_2)}, \tag{2.71}$$

for some fixed pair (k_1, k_2) with $1 \leq k_1 < k_2 \leq n$. This control task is called selective decoupling. An example for such a scenario is a quantum computer in which the two qudit gates are generated by the qudit-qudit couplings. A control scheme $\{g_j'\}_{j=0}^{n_c-1}$ for the simulation of H_0' can easily be obtained from the corresponding decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ as follows [SM01]: Let g_j be given by $g_1^{(1,j)} \otimes g_2^{(2,j)} \otimes \cdots \otimes g_n^{(n,j)}$, where $g_k^{(i,j)}$ denotes the unitary $g^{(i,j)}$ being applied to the k-th qudit. We set $g_j' = g_j$ and apply the following modifications:

- For general couplings, the decoupling scheme was constructed with the help of an orthogonal array. To keep the (k_1, k_2)-coupling, we replace the unitaries $g_{k_1}^{(k_1,j)}$ and $g_{k_2}^{(k_2,j)}$ by $\mathcal{I}_{k_1}$ and $\mathcal{I}_{k_2}$.
- For diagonal couplings, the decoupling scheme was constructed with the help of a difference scheme. To keep the (k_1, k_2)-coupling, we replace $g_{k_2}^{(k_2,j)}$ by $g_{k_2}^{(k_1,j)}$ (or vice versa $g_{k_1}^{(k_1,j)}$ by $g_{k_1}^{(k_2,j)}$).

2.2.4 Nearest-Neighbor Couplings

A general 2-local n-qudit Hamiltonian H_0 involves couplings between up to $n(n-1)/2$ pairs. If the only inter-qudit couplings involved in H_0 are nearest-neighbor couplings and the qudits are arranged on a linear chain, i. e. if

$$H_0 = \sum_{k=1}^{n-1} \sum_{\vec{a},\vec{c}\in\mathbb{F}_q^2\setminus\{\vec{0}\}} J_{\vec{a},\vec{c}}^{k,k+1} \big[XZ(\vec{a}) \otimes XZ(\vec{c})\big]_{(k,k+1)}, \tag{2.72}$$

far shorter decoupling schemes can be devised as the ones discussed in the preceding subsection. Let $\{u(j)\}_{j=0}^{q^2-1}$ denote an annihilator for the one qudit Hilbert space $\mathcal{H}_q$. A decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ of constant length $n_c = q^2$ can be constructed by letting the elements of the annihilator act on the even numbered qudits, i. e. by setting $g_j = \mathcal{I}_1 \otimes u(j)_2 \otimes \mathcal{I}_3 \otimes u(j)_4 \otimes \ldots$ for all $j \in \{0, \ldots, q^2-1\}$.

Theorem 2.2.4. *An n-qudit Hamiltonian H_0 involving only nearest-neighbor couplings as in (2.72) can be decoupled using a decoupling scheme of constant length $n_c = q^2$ as it is described above.*

Proof. Let us pick a term in (2.72) with odd k (for even k the proof goes analogously). Then,

$$\begin{aligned}
&\frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger \; J_{\vec{a},\vec{c}}^{k,k+1} \big[XZ(\vec{a}) \otimes XZ(\vec{c})\big]_{(k,k+1)} \; g_j \\
&= \frac{1}{n_c} \sum_{j=0}^{n_c-1} \big(\mathcal{I}_1 \otimes u(j)_2^\dagger \otimes \mathcal{I}_3 \otimes u(j)_4^\dagger \ldots\big) \; J_{\vec{a},\vec{c}}^{k,k+1} \big[XZ(\vec{a}) \otimes XZ(\vec{c})\big]_{(k,k+1)} \; \big(\mathcal{I}_1 \otimes u(j)_2 \otimes \mathcal{I}_3 \otimes u(j)_4 \ldots\big) \\
&= \left[J_{\vec{a},\vec{c}}^{k,k+1} \; XZ(\vec{a}) \otimes \Big(\frac{1}{q^2} \sum_{j=0}^{q^2-1} u(j)^\dagger XZ(\vec{c}) u(j)\Big)\right]_{(k,k+1)} = 0.
\end{aligned} \tag{2.73}$$

The last step is due to the fact that the set $\{u(j)\}_{j=0}^{q^2-1}$ forms an annihilator and $XZ(\vec{c})$ is traceless for $\vec{c} \in \mathbb{F}_q^2 \setminus \{\vec{0}\}$. □

2.3 Control Strategies

Dynamical control over a local Hamiltonian allows the time evolution of a quantum system to be modified. In the bang-bang scenario, a control scheme consisting of a set of unitaries generated by the local Hamiltonian, is used to achieve a certain control task. For example, for a closed quantum system, we might want to simulate a time evolution according to a Hamiltonian which is different from the system Hamiltonian. In particular, the simulation of a vanishing Hamiltonian is called decoupling. For an open system, we might try to generate a noiseless subsystem (see subsection 2.1.6). For all these tasks, the fundamental control strategy (as discussed in subsection 2.1.3) is to apply the pulses determined by the control scheme with the help of the local control Hamiltonian over and over again. Assuming the pulses to be ideal, the finite time interval in between subsequent pulses is the only obstacle preventing a control task to be achieved in a perfect manner. For the task of decoupling, it was shown in subsection 2.1.4, that the fundamental control strategy (PDD) leads to an average fidelity decay which is quadratic in time. The strength of the decay is determined by (i) the strength of the system Hamiltonian, (ii) by the length of the decoupling scheme, and (iii) by the time interval Δt in between subsequent pulses.

In this section, we consider control strategies which improve the average fidelity decay of a given decoupling scheme for a fixed time interval Δt. The standard technique used by the nuclear magnetic resonance (NMR) community is a symmetrized version of the PDD strategy, which leads to a decrease of the strength of the decay, but keeps its quadratic-in-time nature. In the author's diploma thesis [Ker04] it was observed that a control strategy based on a random selection of the elements of a decoupling scheme leads to a fidelity decay which is only linear in time. Subsequently, randomized decoupling was proposed for open quantum systems by Viola and Knill [VK05]. The linear-in-time decay was confirmed by constructing a lower bound on the worst case fidelity [VK05; Vio05]. Control strategies combining the advantages of purely deterministic and randomized strategies have been devised by the author [KA05] and by Santos and Viola [SV06; VS06], and have been explored numerically for open [SV05] and closed systems [SV08]. We start presenting the deterministic strategies in subsection 2.3.1, and proceed with the randomized strategies in subsection 2.3.2. For most of the strategies we calculate a short time expansion of the average fidelity decay, which allows us to discuss the advantages and disadvantages of a certain strategy. Even though we focus on decoupling, the control strategies discussed in this section are applicable to other control tasks as well. We label the strategies using the abbreviations introduced by Santos and Viola in [SV06; VS06; SV08].

As in the preceding chapters, let S be a closed quantum system defined on a finite-dimensional Hilbert space $\mathcal{H}_S$ of dimension $d = \dim(\mathcal{H}_S)$, and let its Hamiltonian be given by H_0 acting on $\mathcal{H}_S$. Without loss of generality H_0 is assumed to be traceless, i. e. $\mathrm{tr}(H_0) = 0$. Occasionally, we write H_0 as λH_0 and use powers of λ to indicate the dependence on H_0. We assume that a certain decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ of length n_c for H_0 is given.

2.3.1 Deterministic Strategies

Periodic Dynamical Decoupling (PDD)

The fundamental decoupling strategy, as described in subsection 2.1.3, is called periodic dynamical decoupling. At the time $t_i = i \cdot \Delta t$, $i \in \mathbb{N}_0$, the local control Hamiltonian is used to generate the pulse $p_{i \bmod n_c}$, where $p_j = g_j g_{j-1}^\dagger$ (for $j = 0 \ldots n_c - 1$) is defined in terms of the elements g_j of the decoupling scheme by setting $g_{-1} = g_{n_c-1}$ with the exception that the first p_0 is simply given by $p_0' = g_0$ (compare

with figure 2.2). As a result, the time evolution in the toggled frame after the time $T = m \cdot t_c$ with $m \in \mathbb{N}$ and $t_c = n_c \Delta t$ is given by

$$\tilde{U}(T = m \cdot t_c) = \Big(\exp(-i\tilde{H}_{n_c-1}\Delta t) \dots \exp(-i\tilde{H}_1 \Delta t) \exp(-i\tilde{H}_0 \Delta t)\Big)^m = \exp(-i\overline{H} t_c \cdot m), \tag{2.74}$$

with $\tilde{H}_j = g_j^\dagger H_0 g_j$. The zeroth order term in the Magnus expansion of $\overline{H}$ vanishes by definition of the decoupling scheme,

$$\overline{H}^{(0)} = \frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger H_0 g_j = 0, \tag{2.75}$$

and, as it was shown in subsection 2.1.4, the decay of the entanglement fidelity,

$$F_e^{\mathsf{PDD}}(T) = 1 - \frac{1}{d} \operatorname{tr}\big(\big(\overline{H}^{(1)}\big)^2\big) T^2 + \mathcal{O}(\lambda^5 t_c^3 T^2), \tag{2.76}$$

$$F_{e\ \mathrm{app}}^{\mathsf{PDD}}(T) = \exp\Big(-\frac{1}{d} \operatorname{tr}\big(\big(\overline{H}^{(1)}\big)^2\big) T^2\Big), \tag{2.77}$$

is in lowest order only due to the first order term in $\overline{H}$, which is given by (2.15b):

$$\overline{H}^{(1)} = -\frac{i}{2n_c} \sum_{i>j=0}^{n_c-1} [\tilde{H}_i, \tilde{H}_j] \Delta t = \mathcal{O}(\lambda^2 t_c). \tag{2.78}$$

A strict lower bound on the worst case fidelity (2.20) was given in [VK05] by using the matrix norm $\|A\|_2 = \max|\operatorname{eig}(\sqrt{A^\dagger A})|$ and setting $\kappa = \|H_0\|_2$,

$$F_w^{\mathsf{PDD}}(T) > 1 - \kappa^4 t_c^2 T^2 + \mathcal{O}\big(\kappa^5 t_c^3 T^2\big). \tag{2.79}$$

In summary, the fidelity decay using PDD is of the order $\mathcal{O}(\lambda^4 t_c^2 T^2)$ and is caused mainly by the first order term (2.78) in the Magnus expansion of a single PDD cycle of length $t_c = n_c \Delta t$. Suppose we cannot decrease the time interval in between pulses below a certain value Δt. Then, to optimize the fidelity decay of the PDD strategy, we have to find a decoupling scheme as small as possible (i. e. we minimize n_c). The performance of a minimal decoupling scheme may be optimized further by noting that the first order term (2.78) depends on the order of the elements g_j in the decoupling scheme: There are $n_c!$ possibilities and the term $\operatorname{tr}\big(\big(\overline{H}^{(1)}\big)^2\big)/d$ becomes minimal for the new decoupling scheme $\{g'_j\}$ specified by $g'_j = g_{\pi(j)}$, where $\pi \in \mathsf{S}_{n_c}$ denotes a particular permutation of $0, 1, \dots, n_c - 1$. We might also say that π denotes a particular path which traverses the elements of the decoupling scheme. Unfortunately, such an optimal path is hard to find, depends on H_0, and the improvement might be relatively small.

Symmetric Dynamical Decoupling (SDD)

The decoupling technique commonly used by the NMR community is a symmetrized version of the PDD strategy. We call it symmetric dynamical decoupling. Let us construct a symmetrized decoupling scheme $\{g'_j\}_{j=0}^{n'_c-1}$ of length $n'_c = 2n_c$ from the given decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ of length n_c as follows:

$$g'_j = \begin{cases} g_j & \text{for } j = 0, \dots, n_c - 1 \\ g_{2n_c-1-j} & \text{for } j = n_c, \dots, 2n_c - 1 \end{cases}. \tag{2.80}$$

The SDD strategy is to apply the new scheme using the PDD strategy. As a consequence, the time evolution of a single SDD cycle of length $t'_c = n'_c \Delta t$ in the toggled frame is given by

$$\begin{aligned} \tilde{U}(t'_c) = \exp(-i\tilde{H}_0\Delta t) \exp(-i\tilde{H}_1\Delta t) \dots \exp(-i\tilde{H}_{n_c-1}\Delta t) \times \\ \exp(-i\tilde{H}_{n_c-1}\Delta t) \dots \exp(-i\tilde{H}_1\Delta t) \exp(-i\tilde{H}_0\Delta t) = \exp(-i\overline{H} t'_c). \end{aligned} \tag{2.81}$$

Each cycle is symmetric in time, and according to theorem 2.1.1, all odd orders in the Magnus expansion of $\overline{H}$ vanish. Hence, any resulting error is generated mainly by the second order term (2.15c),

$$\overline{H}^{(2)} = -\frac{1}{6n_c'} \sum_{i\geq j\geq k=0}^{n_c'-1} \Big([\tilde{H}_{f(i)}, [\tilde{H}_{f(j)}, \tilde{H}_{f(i)}]] + [[\tilde{H}_{f(i)}, \tilde{H}_{f(j)}], \tilde{H}_{f(k)}] \Big) \Delta t^2 \times \begin{cases} 1/2 & \text{if } i=j \text{ or } j=k \\ 1 & \text{else} \end{cases}, \quad (2.82)$$

where $f(i) = i$ for $i \in \{0, \ldots, n_c - 1\}$ and $f(i) = 2n_c - 1 - i$ for $i \in \{n_c, \ldots, 2n_c - 1\}$. The above expression can be simplified as explained by the following lemma.

Lemma 2.3.1. *The second-order term in the Magnus expansion of a single SDD cycle as given by the above equation is equal to the second-order term in the Magnus expansion of the corresponding PDD cycle, i. e.*

$$\overline{H}^{(2)} = -\frac{1}{6n_c} \sum_{i\geq j\geq k=0}^{n_c-1} \Big([\tilde{H}_i, [\tilde{H}_j, \tilde{H}_k]] + [[\tilde{H}_i, \tilde{H}_j], \tilde{H}_k] \Big) \Delta t^2 \times \begin{cases} 1/2 & \textit{if } i=j \textit{ or } j=k \\ 1 & \textit{else} \end{cases}. \quad (2.83)$$

Proof. Let us divide the interval $[0, t_c']$ into the two subintervals $[0, t_c]$ and $[t_c, t_c']$. If we calculate the average Hamiltonian for each of these subintervals, we obtain vanishing zeroth-order terms of the form of equation (2.75). The results presented in [Bur81, section IV.D] state that in such a case the second-order term of the entire interval is given by the sum of the second-order terms of the subintervals, divided by two. The proof is finished by noting that the second-order term of each of the subintervals is given by (2.83). □

Analogously to equations (2.76), (2.77) and (2.79), we obtain the expressions

$$F_e^{\mathsf{SDD}}(T) = 1 - \frac{1}{d} \operatorname{tr}\big((\overline{H}^{(2)})^2 \big) T^2 + \mathcal{O}(\lambda^8 t_c'^6 T^2), \quad (2.84)$$

$$F_{e\ \text{app}}^{\mathsf{SDD}}(T) = \exp\Big(-\frac{1}{d} \operatorname{tr}\big((\overline{H}^{(2)})^2 \big) T^2 \Big), \quad (2.85)$$

$$F_w^{\mathsf{SDD}}(T) > 1 - \kappa^6 t_c^4 T^2 + \mathcal{O}\big(\kappa^8 t_c'^6 T^2\big), \quad (2.86)$$

and the estimate $\overline{H}^{(2)} = \mathcal{O}(\lambda^3 t_c^2)$.

In summary, the fidelity decay using SDD is of the order $\mathcal{O}(\lambda^6 t_c^4 T^2)$ and is caused mainly by the second order term (2.83) in the Magnus expansion of a single SDD cycle of length $t_c' = 2n_c \Delta t$. For $\kappa t_c < 1$ this is an improvement over PDD in the sense that the strength of the SDD decay ($\mathcal{O}(\lambda^6 t_c^4)$) is smaller than the corresponding PDD strength ($\mathcal{O}(\lambda^4 t_c^2)$). As it was the case for PDD, the performance of SDD might be optimized further by choosing an optimal path $\pi \in \mathsf{S}_{n_c}$ traversing the elements g_j of the underlying decoupling scheme, i. e. an order of the elements such that $\operatorname{tr}\big((\overline{H}^{(2)})^2 \big)/d$ is minimal.

Higher Order Decoupling

A natural question is whether the SDD approach can be generalized to suppress even higher order terms in the Magnus expansion. For a given decoupling scheme of length n_c, we have the set $\{\tilde{H}_j\}_{j=0}^{n_c-1}$ of toggled frame Hamiltonians. Is there a set of indices $\{j(i)\}_{i=1}^N$, $j(i) \in \{0, \ldots, n_c - 1\}$, and relative times $\{\Delta t_i\}_{i=1}^N$ of length N such that the sequence

$$\tilde{U}\big(T = \sum_{i=1}^{N} \Delta t_i\big) = \exp(-i\tilde{H}_{j(N)} \Delta t_N) \ldots \exp(-i\tilde{H}_{j(1)} \Delta t_1) = \exp(-i\overline{H}T) \quad (2.87)$$

has vanishing zeroth, first and second order terms in the Magnus expansion ? (SDD is obtained for $N = 2n_c$, $\Delta t_i = \Delta t$, $T = 2n_c\Delta t$ and $j(i) = \{i \text{ for } i = 0 \dots n_c - 1 \text{ and } 2n_c - 1 - i \text{ for } i = n_c \dots 2n_c - 1\}$. It leads to a vanishing zeroth and first order term.) According to (2.15c), the second order Magnus term is of third order in H_0. Sets $\{j(i)\}_{i=1}^N$ and $\{\Delta t_i\}_{i=1}^N$ of length N satisfying $\overline{H} \sim \mathcal{O}(H_0)^m$ can be found using a Trotter-Suzuki decomposition [Suz91], but according to the non-existence theorem of positive decompositions (ibd.), they always involve negative times Δt_i for $m \geq 4$. This fact forbids general higher order decoupling according to some simple rule (see also the comment in [KL07, section V]). (Nevertheless, there exist specific examples for which second order decoupling is achievable by repetition of a decoupling scheme traversing a series of different paths, see for example the 'H2' scheme in [SV08].)

Concatenated Dynamical Decoupling (CDD)

When using the PDD strategy, the time evolution of a single cycle in the toggled frame is given by (2.74),

$$\tilde{U}(t_c) = \exp(-i\tilde{H}_{n_c-1}\Delta t) \dots \exp(-i\tilde{H}_1\Delta t)\exp(-i\tilde{H}_0\Delta t) = \exp(-i\overline{H}t_c). \tag{2.88}$$

Khodjasteh and Lidar proposed a concatenated dynamical decoupling strategy [KL05], which tries to fight the remaining higher order terms in the Magnus expansion of $\overline{H}$ as follows: As a first step, the basic PDD cycle is embedded into an additional one,

$$\tilde{U}(n_c \cdot t_c) = g_{n_c-1}^\dagger \tilde{U}(t_c) g_{n_c-1} \cdot \dots \cdot g_1^\dagger \tilde{U}(t_c) g_1 \cdot g_0^\dagger \tilde{U}(t_c) g_0, \tag{2.89}$$

leading to a cycle of length n_c^2. We may now either repeat this cycle over and over again (called periodic concatenated level 2 decoupling (PCDD2)), or iterate the embedding process one more time to obtain a cycle of length n_c^3. After k recursive embeddings, one obtains a cycle of length n_c^k. Periodic decoupling with such a cycle is called periodic concatenated level k decoupling (PCDDk) [SV06; VS06; SV08]. The CDD strategy is to repeat the embedding process ad infinitum.

In order to achieve a good performance with CDD, the underlying decoupling scheme should be able to suppress the correlations in the remaining effective Hamiltonian of the k-th embedded cycle for increasing k. Since these correlations increase with k, we expect CDD to work best when the decoupling scheme is an annihilator of short length n_c. Due to the fact that the length of a minimal annihilator is equal to the dimension of the system Hilbert space, it will be hard to meet this criterion. In fact, CDD was proposed to decouple a single qubit from its environment [KL05], in which case an annihilator of length four is given by the Pauli operators $\mathcal{I}, X, Y$, and Z.

2.3.2 Randomized Strategies

Naive Random Decoupling (NRD)

The simplest randomized control strategy is to apply the pulses p_i at times $t_i = i\Delta t$, $i \in \mathbb{N}_0$, where $p_i = g_{r(i)} g_{r(i-1)}^\dagger$ is constructed by picking the elements of the decoupling scheme at random: The indices $r(i) \in \{0, \dots, n_c - 1\}$ are chosen independently according to a uniform distribution. As a result, after a time $T = t_N$ the time evolution operator in the toggled frame is given by

$$\tilde{U}(T = t_N) = \exp(-i\tilde{H}_{r(N-1)}\Delta t) \dots \exp(-i\tilde{H}_{r(1)}\Delta t)\exp(-i\tilde{H}_{r(0)}\Delta t). \tag{2.90}$$

The resulting decay of the entanglement fidelity (2.25) (corresponding to the average state fidelity) depends on the particular choice of indices. To obtain a general statement, we take the average over all random realizations (denoted by $\mathbb{E}$), i. e. we define

$$F_e^{\mathsf{NRD}}(T) = \mathbb{E}\Big|\frac{1}{d}\operatorname{tr}\big(\tilde{U}(T)\big)\Big|^2 \tag{2.91}$$

as the relevant performance measure.

Theorem 2.3.2. *In lowest order, the average NRD fidelity* (2.91) *is given by*

$$F_e^{NRD}(T) = 1 - \frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t T + \mathcal{O}(\lambda^4 \Delta t^2 T). \tag{2.92}$$

Proof. Writing H_0 as λH_0, we calculate the fidelity (2.91) up to fourth order in λ. This allows any result to be used later on to obtain the variance of the fidelity. We start by expanding each of the products in (2.90) as

$$\exp(-i\tilde{H}_{r(s)}\Delta t) = \mathcal{I} - i\tilde{H}_{r(s)}\Delta t - \frac{1}{2}\tilde{H}_{r(s)}^2\Delta t^2 + \frac{i}{6}\tilde{H}_{r(s)}^3\Delta t^3 + \frac{1}{24}\tilde{H}_{r(s)}^4\Delta t^4 + \mathcal{O}(\lambda^5), \tag{2.93}$$

with $0 \le s \le N-1$. Taking the trace leads to

$$\begin{aligned}\frac{1}{d}\operatorname{tr}\big(\tilde{U}(T)\big) &= 1 - \frac{1}{2}\sum_s \frac{1}{d}\operatorname{tr}\big(\tilde{H}_{r(s)}^2\big)\Delta t^2 - \sum_{s>u}\frac{1}{d}\operatorname{tr}\big(\tilde{H}_{r(s)}\tilde{H}_{r(u)}\big)\Delta t^2 + \frac{i}{6}\sum_s \frac{1}{d}\operatorname{tr}\big(\tilde{H}_{r(s)}^3\big)\Delta t^3 \\ &+ \frac{i}{2}\sum_{s>u}\frac{1}{d}\operatorname{tr}\big(\tilde{H}_{r(s)}\tilde{H}_{r(u)}^2 + \tilde{H}_{r(s)}^2\tilde{H}_{r(u)}\big)\Delta t^3 + i\sum_{s>u>v}\frac{1}{d}\operatorname{tr}\big(\tilde{H}_{r(s)}\tilde{H}_{r(u)}\tilde{H}_{r(v)}\big)\Delta t^3 + \cdots + \mathcal{O}(\lambda^5). \end{aligned} \tag{2.94}$$

The fidelity is obtained by averaging the absolute square of the above expression over all random realizations. With the help of the decoupling condition (2.16) for traceless H_0,

$$\frac{1}{n_c}\sum_{j=0}^{n_c-1}\tilde{H}_j = 0, \tag{2.95}$$

and due to the independence of the random selections, we obtain

$$\begin{aligned}\mathbb{E}\Big|\frac{1}{d}\operatorname{tr}\big(\tilde{U}(T)\big)\Big|^2 &= 1 - \frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t T + \frac{1}{4}\Big(\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t T\Big)^2 + \frac{1}{12}\frac{1}{d}\operatorname{tr}\big(H_0^4\big)\Delta t^3 T \\ &+ \frac{1}{2}\frac{1}{n_c}\sum_{j=0}^{n_c-1}\frac{1}{n_c}\sum_{j'=0}^{n_c-1}\Big(\frac{1}{d}\operatorname{tr}\big(\tilde{H}_j\tilde{H}_{j'}\big)\Big)^2\Delta t^2 T(T-\Delta t) \\ &+ \frac{1}{4}\frac{1}{d}\operatorname{tr}\Big(\Big(\frac{1}{n_c}\sum_{j=0}^{n_c-1}\tilde{H}_j^2\Big)^2\Big)\Delta t^2 T(T-\Delta t) + \mathcal{O}(\lambda^5). \quad \square \end{aligned} \tag{2.96}$$

Remark. As it turns out by looking at various numeric examples, a good approximation of (2.91), valid for all times $T \ge 0$ and in lowest order identical to (2.92), is given by

$$F_{e\ \mathrm{app}}^{\mathsf{NRD}}(T) = \exp\Big(-\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t T\Big). \tag{2.97}$$

A strict lower bound on the average worst case fidelity was given in [VK05],

$$F_w^{\mathsf{NRD}}(T) = \mathbb{E}\min_{|\psi\rangle\in\mathcal{H}_S}\big|\langle\psi|\tilde{U}(T)|\psi\rangle\big|^2 > 1 - 4\kappa^2\Delta t T + \mathcal{O}\big(\kappa^3\Delta t^2 T\big), \tag{2.98}$$

with $\kappa = \|H_0\|_2$. The bound remains valid for time-dependent system Hamiltonians $H_0(t)$ if $\|H_0(t)\|_2 < \kappa$ for $0 \le t \le T$ and the decoupling condition (2.55) is satisfied for $0 \le t \le T$. For an appropriate redefinition of κ, the bound applies to open quantum systems as well [VK05; Vio05].

In summary, NRD offers some interesting advantages over deterministic strategies like PDD and SDD: The fidelity decay ($\mathcal{O}(\lambda^2\Delta t T)$) is only linear in time, while it is quadratic in time for PDD and SDD. The strength of the decay does not depend on the length n_c of the underlying decoupling scheme. As a consequence, it is always possible to choose an annihilator as decoupling scheme. Since the lower bound guarantees a linear decay also for time dependent Hamiltonians, it is possible to apply NRD

even if the system Hamiltonian is completely unknown. An additional advantage over PDD is that the NRD strategy remains applicable if we use bounded control instead of bang-bang control (a fact that turns out in subsection 3.2.4), while the PDD cycles have to be replaced by the longer Euler cycles of subsection 2.1.7. A disadvantage is the higher strength of the decay ($\mathcal{O}(\lambda^2\Delta t)$) compared to PDD and SDD ($\mathcal{O}(\lambda^4 t_c^2)$ and $\mathcal{O}(\lambda^6 t_c^4)$, respectively). As pointed out in [VK05], NRD outperforms PDD if $\kappa^2\Delta t T \cdot n_c^2 \gg 1$, i. e. for long times and/or long decoupling schemes.

The linear-in-time fidelity decay of NRD was first observed in the author's diploma thesis [Ker04, chapter 4.2] where a quantum memory consisting of $n = 10$ qubits was protected against inter-qubit couplings by using a decoupling scheme of length $n_c = 4^n$ given by the set of Pauli operators $\mathcal{P}_2^n$. As it will be shown in section 3.2, in contrast with any periodic strategy, NRD allows the protection of a quantum computation in which the quantum gates are applied in between subsequent decoupling pulses [KAS05; GKAJ08]. In this context, NRD using a decoupling scheme given by the set of Pauli operators was called Pauli random error correction (PAREC).

For any decoupling strategy which involves some kind of randomization, in addition to the average fidelity, an important quantity is its variance. It is a measure of how close the fidelity of a single run comes to the average fidelity: The smaller the variance, the smaller the expected difference.

Theorem 2.3.3. *In lowest non-vanishing order, the variance of the NRD fidelity* (2.91) *is given by*

$$\sigma^2_{NRD}(T) = 2T(T-\Delta t)\frac{1}{n_c}\sum_{j=0}^{n_c-1}\frac{1}{n_c}\sum_{j'=0}^{n_c-1}\Big(\frac{1}{d}\operatorname{tr}\big(\tilde{H}_j\tilde{H}_{j'}\big)\Big)^2\Delta t^2 + \mathcal{O}(\lambda^6). \tag{2.99}$$

Proof. We calculate the quantity

$$\sigma^2_{\mathsf{NRD}}(T) = \mathbb{E}\Big(\Big|\frac{1}{d}\operatorname{tr}\big(\tilde{U}(T)\big)\Big|^2\Big)^2 - \Big(\mathbb{E}\Big|\frac{1}{d}\operatorname{tr}\big(\tilde{U}(T)\big)\Big|^2\Big)^2 \tag{2.100}$$

up to fourth order in λ as it was done in the proof of theorem 2.3.2. The term whose square is subtracted on the right hand side is given by (2.96). □

Remark (i). Equation (2.99) can be upper and lower bounded as follows: Using the fact that $\langle A, B\rangle = \operatorname{tr}(A^\dagger B)$ denotes the Hilbert-Schmidt inner product, the Cauchy-Schwarz inequality, $|\langle A, B\rangle|^2 \leq \langle A, A\rangle \cdot \langle B, B\rangle$, in connection with $T(T-\Delta t) < T^2$ leads to an upper bound. Since the averaging is performed over a non-negative expression, we obtain a lower bound by picking the elements where $j = j'$. Altogether,

$$\frac{2T(T-\Delta t)}{n_c}\Big(\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t\Big)^2 \leq \sigma^2_{\mathsf{NRD}}(T) \leq 2\Big(\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t T\Big)^2. \tag{2.101}$$

Remark (ii). If the elements g_j of the decoupling set $\{g_j\}_{j=0}^{n_c-1}$ form a group, equation (2.99) simplifies to

$$\sigma^2_{\mathsf{NRD}}(T) = 2T(T-\Delta t)\frac{1}{n_c}\sum_{j=0}^{n_c-1}\Big(\frac{1}{d}\operatorname{tr}\big(H_0\tilde{H}_j\big)\Big)^2\Delta t^2 + \mathcal{O}(\lambda^6). \tag{2.102}$$

While the average NRD fidelity does not depend on the length of the decoupling scheme, equation (2.99) in connection with the lower bound in (2.101) leads to the conclusion that its variance actually becomes smaller, the greater the length of the decoupling scheme. We expect the variance to become minimal if the underlying decoupling scheme is an annihilator. This feature is in strong contrast to PDD and SDD where smaller decoupling schemes increase the performance.

Embedded Decoupling (EMD)

In order to combine the advantages of the PDD and the NRD strategy, the following embedded dynamical decoupling strategy has been devised by the author in [KA05]. Let $\tilde{U}(t_c)$ denote the time evolution

operator of a single PDD cycle in the toggled frame (compare with (2.74)),

$$\tilde{U}(t_c) = \exp(-i\tilde{H}_{n_c-1}\Delta t)\ldots\exp(-i\tilde{H}_1\Delta t)\exp(-i\tilde{H}_0\Delta t) = \exp(-i\overline{H}t_c). \tag{2.103}$$

By definition of the decoupling scheme, the zeroth order term in $\overline{H}$ vanishes and we have the residual Hamiltonian $\overline{H} = \overline{H}^{(1)} + \overline{H}^{(2)} + \ldots$, with $\overline{H}^{(1)}$ given by (2.78). Let us now take a second decoupling scheme $\{\gamma_j\}_{j=0}^{\nu_c-1}$ eliminating the residual Hamiltonian. The embedded decoupling strategy is to apply the NRD strategy at times $i \cdot t_c$, $i \in \mathbb{N}_0$, using the second decoupling set to suppress the residual Hamiltonian of the PDD cycles. As a result, after a time $T = N \cdot t_c$, $N \in \mathbb{N}_0$, we obtain the following time evolution,

$$\begin{aligned}\tilde{U}(T = N\cdot t_c) &= \gamma_{r(N-1)}^\dagger \tilde{U}(t_c)\gamma_{r(N-1)} \ \cdots\ \gamma_{r(1)}^\dagger \tilde{U}(t_c)\gamma_{r(1)}\ \gamma_{r(0)}^\dagger \tilde{U}(t_c)\gamma_{r(0)} \\ &= \exp(-i\tilde{\overline{H}}_{r(N-1)}\Delta t)\ldots\exp(-i\tilde{\overline{H}}_{r(1)}\Delta t)\exp(-i\tilde{\overline{H}}_{r(0)}\Delta t),\end{aligned} \tag{2.104}$$

where $\tilde{\overline{H}}_{r(i)} = \gamma_{r(i)}^\dagger \overline{H}\,\gamma_{r(i)}$ for $i = \{0, 1, \ldots, N-1\}$, and $r(i) \in \{0, 1, \ldots, \nu_c - 1\}$. Typically, we choose the second decoupling set to be an annihilator given by the set of Pauli operators, i. e. $\{\gamma_j\}_{j=0}^{\nu_c-1} = \mathcal{P}_q^n$. To analyze the performance of EMD, we can simply adopt the results obtained for NRD if we apply the substitutions $H_0 \mapsto \overline{H}$ and $\Delta t \mapsto t_c$. In particular, to obtain the lowest order results, it suffices to replace H_0 with $\overline{H}^{(1)}$. Hence, we obtain the entanglement fidelity

$$F_e^{\mathsf{EMD}}(T) = 1 - \frac{1}{d}\operatorname{tr}\big((\overline{H}^{(1)})^2\big)t_c T + \mathcal{O}((\lambda^2 t_c)^4 t_c^2 T), \tag{2.105}$$

$$F_{e\ \mathrm{app}}^{\mathsf{EMD}}(T) = \exp\Big(-\frac{1}{d}\operatorname{tr}\big((\overline{H}^{(1)})^2\big)t_c T\Big), \tag{2.106}$$

the worst case fidelity

$$F_w^{\mathsf{EMD}}(T) > 1 - 4\kappa^4 t_c^3 T + \mathcal{O}\big(\kappa^6 t_c^5 T\big), \tag{2.107}$$

and the variance

$$\sigma_{\mathsf{EMD}}^2(T) = 2T(T - t_c)\frac{1}{\nu_c}\sum_{j=0}^{\nu_c-1}\frac{1}{\nu_c}\sum_{j'=0}^{\nu_c-1}\Big(\frac{1}{d}\operatorname{tr}\big(\tilde{\overline{H}}_j^{(1)}\tilde{\overline{H}}_{j'}^{(1)}\big)\Big)^2 t_c^2 + \mathcal{O}(\lambda^{12}), \tag{2.108}$$

with $\tilde{\overline{H}}_j^{(1)} = \gamma_j^\dagger \overline{H}^{(1)}\gamma_j$. The fidelity decay is of order $\mathcal{O}(\lambda^4 t_c^3 T)$ and does indeed combine the advantage of the linear-in-time decay of NRD with the stronger suppression of PDD. As it was discussed in the PDD paragraph, the performance of PDD depends slightly on the order of the elements in the decoupling scheme, or in other words, on the path which traverses the elements during a cycle. To eliminate this dependence and to achieve an average performance, we might choose a random path for each basic cycle (compare with the RPD strategy). We label such an embedded strategy involving the additional path randomization EMDr. An overview over the dependencies of the average fidelity decay for different control strategies can be found in table 2.1.

Embedded Symmetric Decoupling (ESDD)

The embedded decoupling strategy described in the preceeding paragraph can naturally be extended to an underlying SDD scheme, as it was done implicitly in [KA06] (see chapter 4). We call the resulting decoupling strategy embedded symmetric dynamical decoupling. For a single SDD cycle, equation (2.103) becomes

$$\begin{aligned}\tilde{U}(t_c') = \exp(-i\tilde{H}_0\Delta t)\exp(-i\tilde{H}_1\Delta t)\ldots\exp(-i\tilde{H}_{n_c-1}\Delta t)\times \\ \exp(-i\tilde{H}_{n_c-1}\Delta t)\ldots\exp(-i\tilde{H}_1\Delta t)\exp(-i\tilde{H}_0\Delta t) = \exp(-i\overline{H}t_c'),\end{aligned} \tag{2.109}$$

strategy	decay
none	$\mathcal{O}(\lambda^2 T^2)$
NRD	$\mathcal{O}(\lambda^2 \Delta t T)$
PDD	$\mathcal{O}(\lambda^4 t_c^2 T^2)$
EMD,EMDr,RPD	$\mathcal{O}(\lambda^4 t_c^3 T)$
SDD	$\mathcal{O}(\lambda^6 t_c^4 T^2)$
ESDD,ESDDr,SRPD	$\mathcal{O}(\lambda^6 t_c^5 T)$

Table 2.1: The average fidelity decay of various control strategies using an underlying decoupling scheme of length n_c and a pulse distance in time Δt ($t_c = n_c \Delta t$) to suppress the system Hamiltonian λH_0. Note that the strength of the decay of NRD does not depend on the length n_c.

with $t_c' = n_c' \Delta t$ and $n_c' = 2n_c$, and the Magnus expansion of the residual Hamiltonian $\overline{H}$ contains only terms of second and higher order, i. e. $\overline{H} = \overline{H}^{(2)} + \overline{H}^{(4)} + \ldots$, with $\overline{H}^{(2)}$ given by (2.83). As it was done in the analysis of the performance of EMD, we can simply adopt the results obtained for NRD if we apply the substitutions $H_0 \mapsto \overline{H}$ and $\Delta t \mapsto t_c'$ in the corresponding expressions. To obtain the lowest order results, it suffices to replace H_0 with $\overline{H}^{(2)}$, and we obtain the average fidelity

$$F_e^{\mathsf{ESDD}}(T) = 1 - \frac{1}{d} \operatorname{tr}\big(\big(\overline{H}^{(2)}\big)^2\big) t_c' T + \mathcal{O}((\lambda^3 t_c'^2)^4 t_c'^2 T), \tag{2.110}$$

$$F_{e\ \mathrm{app}}^{\mathsf{ESDD}}(T) = \exp\Big(-\frac{1}{d} \operatorname{tr}\big(\big(\overline{H}^{(2)}\big)^2\big) t_c' T\Big), \tag{2.111}$$

the worst case fidelity

$$F_w^{\mathsf{ESDD}}(T) > 1 - 4\kappa^6 t_c'^5 T + \mathcal{O}\big(\kappa^9 t_c'^8 T\big), \tag{2.112}$$

and the variance

$$\sigma_{\mathsf{ESDD}}^2(T) = 2T(T - t_c') \frac{1}{\nu_c} \sum_{j=0}^{\nu_c - 1} \frac{1}{\nu_c} \sum_{j'=0}^{\nu_c - 1} \Big(\frac{1}{d} \operatorname{tr}\big(\tilde{\overline{H}}_j^{(2)} \tilde{\overline{H}}_{j'}^{(2)}\big)\Big)^2 t_c'^2 + \mathcal{O}(\lambda^{18}). \tag{2.113}$$

As in the EMD case, we might bring the decoupling elements after each cycle into a new random order. We label such a strategy involving this additional randomization by ESDDr (to be compared with SRPD).

Random Path Decoupling (RPD)

Another approach to combine the advantages of the deterministic and randomized strategies is called random path decoupling. It was proposed by Viola and Knill [VK05] and explored by Santos and Viola in [SV06; VS06]. While the performance of RPD was conjectured to be comparable with EMD [VS06], we are going to prove this conjecture. The RPD strategy is basically to apply PDD, but now each PDD cycle is constructed from a randomly reordered decoupling scheme. In other words, each PDD cycle traverses the elements of the decoupling scheme according to a random path. The time evolution operator of such a PDD cycle is given by

$$\tilde{U}_\pi(t_c) = \exp(-i\tilde{H}_{\pi(n_c-1)}\Delta t) \ldots \exp(-i\tilde{H}_{\pi(1)}\Delta t) \exp(-i\tilde{H}_{\pi(0)}\Delta t) = \exp(-i\overline{H}_\pi t_c), \tag{2.114}$$

where $\pi \in \mathsf{S}_{n_c}$ denotes a randomly chosen permutation of the elements of the decoupling scheme. The reordering obviously does not affect the zeroth order term in the Magnus expansion of $\overline{H}_\pi = \overline{H}_\pi^{(0)} + \overline{H}_\pi^{(1)} + \ldots$, which is still given by (2.75) (i. e. $\overline{H}_\pi^{(0)} = 0$ for all π), but the first order term,

$$\overline{H}_\pi^{(1)} = -\frac{i}{2n_c} \sum_{i>j=0}^{n_c-1} [\tilde{H}_{\pi(i)}, \tilde{H}_{\pi(j)}] \Delta t \tag{2.115}$$

depends on π.

Lemma 2.3.4. *The average of $\overline{H}_\pi^{(1)}$ taken over all $\pi \in \mathsf{S}_{n_c}$ vanishes, i. e. we have*

$$\left\langle \overline{H}_\pi^{(1)} \right\rangle_{\pi \in \mathsf{S}_{n_c}} = \frac{1}{n_c!} \sum_{\pi \in \mathsf{S}_{n_c}} \overline{H}_\pi^{(1)} = 0. \tag{2.116}$$

Proof. This result is a simple consequence of the fact that $[\tilde{H}_i, \tilde{H}_j] = -[\tilde{H}_j, \tilde{H}_i]$. □

According to the above lemma, we are in a situation similar to EMD, where the residual Hamiltonian $\overline{H}^{(1)} + \overline{H}^{(2)} + \dots$ of a fixed PDD cycle is eliminated on average by the additional pulses generated by random selection from the second decoupling scheme. While EMD achieves the suppression perfectly in the sense that the average taken over all toggled residual Hamiltonians vanishes, it is unclear whether RPD achieves annihilation of the second- and higher-order terms in the residual Hamiltonian as well. (It will be shown in the next paragraph that annihilation is still achieved for the second-order term.) Therefore, we expect RPD to perform slightly worse than EMD (or EMDr if we eliminate the influence of the order of the decoupling elements). In fact RPD is equivalent to EMDr, if we replace each element of the second decoupling scheme by the identity. Nevertheless, RPD offers the advantage that no second decoupling scheme is involved. Hence, all the applied pulses are of the form $g_j g_i^\dagger$ for some $i, j \in \{0, \dots, n_c - 1\}$.

Symmetric Random Path Decoupling (SRPD)

The RPD strategy of the preceding paragraph can be improved by symmetrizing the randomly traversed PDD cycles as it was done by the SDD strategy. The resulting strategy is called symmetric random path decoupling [SV06; VS06]. Using SRPD, a basic random cycle of length $n_c' = 2n_c$ is given by

$$\begin{aligned}\tilde{U}_\pi(t_c') = \exp(-i\tilde{H}_{\pi(0)}\Delta t)\exp(-i\tilde{H}_{\pi(1)}\Delta t)\dots\exp(-i\tilde{H}_{\pi(n_c-1)}\Delta t)\times \\ \exp(-i\tilde{H}_{\pi(n_c-1)}\Delta t)\dots\exp(-i\tilde{H}_{\pi(1)}\Delta t)\exp(-i\tilde{H}_{\pi(0)}\Delta t) = \exp(-i\overline{H}_\pi t_c'),\end{aligned} \tag{2.117}$$

with $\pi \in \mathsf{S}_{n_c}$, and by using lemma 2.3.1, the lowest non-vanishing term in the Magnus expansion of $\overline{H}_\pi$ is given by

$$\overline{H}_\pi^{(2)} = -\frac{1}{6n_c} \sum_{i \geq j \geq k = 0}^{n_c - 1} \Big([\tilde{H}_{\pi(i)}, [\tilde{H}_{\pi(j)}, \tilde{H}_{\pi(k)}]] + [[\tilde{H}_{\pi(i)}, \tilde{H}_{\pi(j)}], \tilde{H}_{\pi(k)}] \Big) \Delta t^2 \times \begin{cases} 1/2 & \text{if } i = j \text{ or } j = k \\ 1 & \text{else} \end{cases}. \tag{2.118}$$

Lemma 2.3.5. *The average of the above expression taken over all permutations $\pi \in \mathsf{S}_{n_c}$ vanishes, i. e. we have*

$$\left\langle \overline{H}_\pi^{(2)} \right\rangle_{\pi \in \mathsf{S}_{n_c}} = 0. \tag{2.119}$$

Proof. We start with the observation that all the terms in the sum forming $\overline{H}_\pi^{(2)}$ with $i = j$ or $j = k$ add up to zero:

$$\begin{aligned}\sum_{i > j = 0}^{n_c - 1} \Big([\tilde{H}_{\pi(i)}, [\tilde{H}_{\pi(i)}, \tilde{H}_{\pi(j)}]] + [[\tilde{H}_{\pi(i)}, \tilde{H}_{\pi(j)}], \tilde{H}_{\pi(j)}] \Big) = \\ \sum_{i \neq j = 0}^{n_c - 1} \Big(\tilde{H}_{\pi(i)}\tilde{H}_{\pi(i)}\tilde{H}_{\pi(j)} - 2\tilde{H}_{\pi(i)}\tilde{H}_{\pi(j)}\tilde{H}_{\pi(i)} + \tilde{H}_{\pi(j)}\tilde{H}_{\pi(i)}\tilde{H}_{\pi(i)} \Big) = 0.\end{aligned} \tag{2.120}$$

The last identity follows if we extend the sum by the terms $i = j$ and use the fact that $\overline{H}^{(0)} = 0$. Hence, the average over all permutations can be taken over the simpler expression

$$\overline{H}_{\pi}^{(2)} = -\frac{1}{6n_c} \sum_{i>j>k=0}^{n_c-1} \Big([\tilde{H}_{\pi(i)}, [\tilde{H}_{\pi(j)}, \tilde{H}_{\pi(k)}]] + [[\tilde{H}_{\pi(i)}, \tilde{H}_{\pi(j)}], \tilde{H}_{\pi(k)}] \Big) \Delta t^2, \tag{2.121}$$

and as in the proof of lemma 2.3.4, the property $[\tilde{H}_{\pi(i)}, \tilde{H}_{\pi(j)}] = -[\tilde{H}_{\pi(j)}, \tilde{H}_{\pi(i)}]$ leads to the vanishing mean. □

Since it remains unclear whether SRPD eliminates the remaining higher order terms in the Magnus expansion as well, we expect it to perform slightly worse than an average ESDD or ESDDr, respectively.

2.4 Example

In the preceding section various decoupling strategies and their advantages have been discussed. We are now going to examine the performance of these strategies by means of numerical simulations. Results on the entanglement fidelity obtained numerically are compared with the corresponding formulas which have been derived in the preceding section. We start by presenting the model, a quantum register perturbed by Heisenberg couplings, in subsection 2.4.1. Then, in subsection 2.4.2, we focus on the variance of the naive random decoupling NRD strategy using different decoupling sets. In subsection 2.4.3, we compare different strategies in order to identify the best one. Finally, we conclude in subsection 2.4.4 with a general guideline for a good decoupling strategy.

2.4.1 The Model

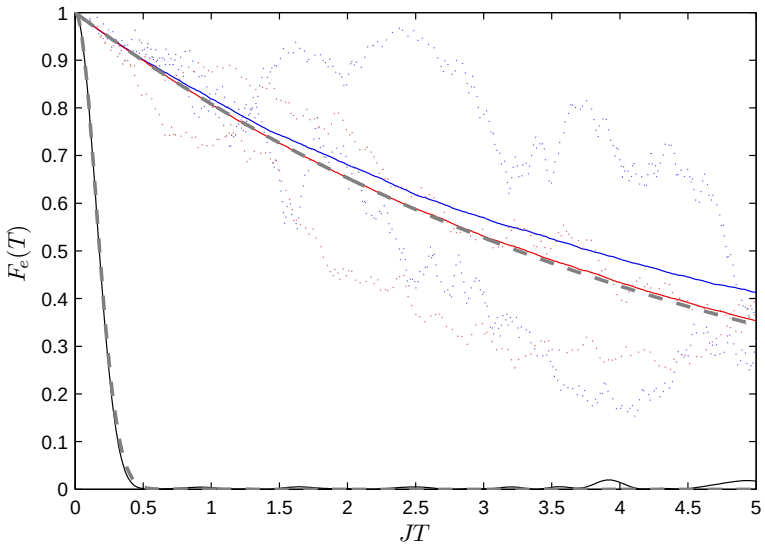
We choose the same model Hamiltonian as in [SV06], i. e. we consider $n = 8$ qubits with Heisenberg couplings arranged on a linear chain,

$$H_0 = \sum_{k_1=1}^{n-1} \sum_{k_2=k_1+1}^{n} J^{k_1,k_2} \big[X \otimes X + Y \otimes Y + Z \otimes Z \big]_{(k_1,k_2)}, \tag{2.122}$$

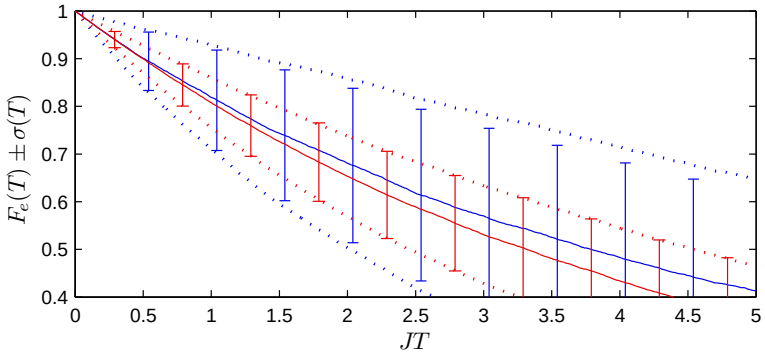
where the coupling strength between qubits k_1 and k_2 decays cubically with their separation distance, i. e. $J^{k_1,k_2} = J \cdot |k_1 - k_2|^{-3}$. We construct a decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ of length $n_c = 8$ for H_0 by using the difference scheme $D(8, 8, 4)$ listed in table A.2 in a way explained in theorem 2.2.3. Another decoupling scheme for H_0 is given by the annihilator $\{\gamma_j\}_{j=0}^{\nu_c-1}$ of length $\nu_c = 4^8$ consisting of Pauli operators, i. e. $\gamma_j = XZ(j)$ with $j \in \mathbb{F}_2^{2\cdot 8}$.

2.4.2 The Naive Random Strategy

We performed a numerical simulation of model (2.122) over the time $0 \le T \le 5J^{-1}$. The resulting entanglement fidelity without decoupling, $F_e^{\mathsf{none}}(T)$, drops down to zero after the time $\approx 0.5J^{-1}$ and is shown in figure 2.5a (*black, solid line*). It is in excellent agreement with our estimation $F_{e\ \mathrm{app}}^{\mathsf{none}}(T)$ given by (2.27) (*dashed line*). In addition, figure 2.5a shows the numerically obtained NRD fidelity $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T)$ when using the small decoupling set $\{g_j\}_{j=0}^{n_c-1}$ of length $n_c = 8$ with a pulse distance in time of $\Delta t = 0.01J^{-1}$ (*blue, solid line*). The index num in $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T)$ indicates that the quantity differs from the definition of $F_e^{\mathsf{NRD}}(T)$ in equation (2.91) with respect to the average over the random pulse realizations: The latter quantity was defined by averaging over all realizations, while $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T)$ is averaged over a random subset of simulated runs. The NRD fidelity based on the small decoupling set (*blue, solid line*) is compared with the corresponding NRD fidelity based on the annihilator $\{\gamma_j\}_{j=0}^{\nu_c-1}$ of length $\nu_c = 4^8$ (*red, solid line*). Both fidelities have been obtained by averaging over 1500 single runs



(a) Entanglement fidelity.



(b) Entanglement fidelity and its root mean square.

Figure 2.5: The entanglement fidelity of a quantum register with $n = 8$ qubits, perturbed by the Hamiltonian given in (2.122). The time interval between adjacent decoupling pulses is $\Delta t = 0.01 J^{-1}$. The NRD fidelities are averaged over 1500 runs.
(a) Without decoupling (*solid line, black*), with NRD using the set $\{g_j\}_{j=0}^{7}$ (*solid line, blue*), and NRD using the set $\{\gamma_j\}_{j=0}^{4^8-1}$ (*solid line, red*). For both of the NRD strategies two individual runs are shown (*dotted lines*). The *dashed lines* indicate the estimations given by (2.27) and (2.97), respectively.
(b) In addition to the two NRD fidelities $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T)$ (*solid lines*), we indicate the intervals $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T) \pm \sigma_{\mathsf{NRD}}^{\mathrm{num}}(T)$ (*error bars*) and $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T) \pm \sigma_{\mathsf{NRD}}^{\mathrm{app}}(T)$ (*dotted lines*), where $\sigma_{\mathsf{NRD}}^{\mathrm{num}}(T)$ denotes the standard deviation of the numerical fidelity and $\sigma_{\mathsf{NRD}}^{\mathrm{app}}(T)$ the corresponding estimation given by (2.124).

with independent random pulse realizations. As predicted by our short time expansion (2.92), both fidelities are identical for short times. In the region where higher orders become relevant, NRD based on the small decoupling set performs slightly better. Our estimation $F_{e\ \mathrm{app}}^{\mathsf{NRD}}(T)$ (2.97) (*dashed line*) is in excellent agreement with the NRD fidelity using the annihilator (*red, solid line*). To evaluate the estimations $F_{e\ \mathrm{app}}^{\mathsf{none}}(T)$ (2.27) and $F_{e\ \mathrm{app}}^{\mathsf{NRD}}(T)$ (2.97), we need the quantity $\mathrm{tr}(H_0^2)/d \approx 21.30J^2$.

We are now going to study the variance of the NRD fidelities. In figure 2.5b we indicate the value of the quantity $\sigma_{\mathsf{NRD}}^{2\ \mathrm{num}}(T)$, which is defined as in (2.100) with the average over all random realizations (denoted by $\mathbb{E}$) being replaced by the average over the subset of simulated random realizations, by plotting $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T) \pm \sigma_{\mathsf{NRD}}^{\mathrm{num}}(T)$ (*error bars*) in addition to $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T)$ (*solid line*). As in figure 2.5a, the plots corresponding to the NRD strategy based on the small decoupling set are depicted in *blue*, while plots corresponding to the NRD strategy based on the annihilator are depicted in *red*. It can be seen that the variance is smaller with the annihilator as the underlying decoupling set. A short time estimation for the variance $\sigma_{\mathsf{NRD}}^2(T)$ was calculated in equation (2.99). Evaluating this expression for the two different decoupling sets leads to

$$\sigma_{\mathsf{NRD}}^2(T) \approx 2T^2\Delta t^2 \times \begin{cases} 92.47J^4 & ,\text{for } \{g_j\}_{j=0}^{7} \\ 21.00J^4 & ,\text{for } \{\gamma_j\}_{j=0}^{4^8-1} \end{cases} + \mathcal{O}(J^6). \tag{2.123}$$

As it turns out, this expression overestimates the variance for longer times. Hence, we propose the following estimation,

$$\sigma_{\mathsf{NRD}}^{2\ \mathrm{app}}(T) = 2T(T-\Delta t)\mathbb{E}_j\mathbb{E}_{j'}\Big(\frac{1}{d}\,\mathrm{tr}\big(\tilde{H}_j\tilde{H}_{j'}\big)\Big)^2\Delta t^2 \times \exp\Big(-2\frac{1}{d}\,\mathrm{tr}\big(H_0^2\big)\Delta t T\Big), \tag{2.124}$$

which we expect to deliver a good approximation for all relevant times. Here, $\mathbb{E}_j$ ($\mathbb{E}_{j'}$) denotes the average taken over all elements of the underlying decoupling set, i. e. $\tilde{H}_j = g_j^\dagger H_0 g_j$ for the small set $\{g_j\}_{j=0}^{n_c-1}$ of length $n_c = 8$ and $\tilde{H}_j = \gamma_j^\dagger H_0 \gamma_j$ for the annihilator $\{\gamma_j\}_{j=0}^{\nu_c-1}$ of length $\nu_c = 4^8$. Note that for short times the exponential can be neglected and this estimation is identical to the (exact) short time expression (2.99). We put (2.124) to the test by plotting the quantities $F_{e\ \mathrm{num}}^{\mathsf{NRD}}(T) \pm \sigma_{\mathsf{NRD}}^{\mathrm{app}}(T)$ for both NRD cases (*dotted lines* in figure 2.5b). As expected, the estimation (2.124) is excellent for short times. For longer times it remains excellent when using the annihilator, but slightly overestimates the variance when the small decoupling set is involved.

A decoupling strategy like NRD will be of interest only as long as the resulting fidelity is reasonably high. In this range, it doesn't make any difference from what kind of decoupling set the elements for NRD are chosen: all choices lead essentially to the same performance. But since one is interested in a reliable result, one might prefer an annihilator like the set of Pauli operators to constitute the underlying decoupling set, because of the smaller variance.

2.4.3 Comparison of Strategies

We are now going to compare the long-time performance of different decoupling strategies. For this purpose we simulated the time evolution of model (2.122) up to the time $T = 100J^{-1}$. All decoupling strategies apply their pulses at times $t_i = i \cdot \Delta t$, $i \in \{0, 1, \ldots, 2000\}$, with $\Delta t = 0.05J^{-1}$. Each of the entanglement fidelities of the randomized strategies is averaged over 100 individual runs with independent random selections.

For this setting, a simulation of various decoupling strategies up to the time $T = 50J^{-1}$ has already been done by Santos and Viola in [SV06] (even though for a slightly different underlying decoupling scheme). We extend this research by taking a closer look on the influence of the traversing path of the decoupling elements and by comparing the obtained fidelities with their corresponding estimations, which have been obtained in section 2.3. In addition, we study the variance of the randomized schemes and analyze the performance of the EMDr and ESDDr strategies, which have not been considered in [SV06].

PDD_i	traversing path	$\mathrm{tr}\big(\big(\overline{H}_i^{(1)}\big)^2\big)/d$	$\mathrm{tr}\big(\big(\overline{H}_i^{(2)}\big)^2\big)/d$
PDD_1	$g_0, g_2, g_4, g_7, g_1, g_3, g_5, g_6$	$0.09252 J^4 \Delta t^2$	$16.2032 J^6 \Delta t^4$
PDD_2	$g_0, g_1, g_2, g_3, g_4, g_5, g_6, g_7$	$5.5994 J^4 \Delta t^2$	$389.5980 J^6 \Delta t^4$
PDD_3	$g_0, g_1, g_6, g_5, g_2, g_3, g_4, g_7$	$36.963 J^4 \Delta t^2$	$1971.425 J^6 \Delta t^4$

Table 2.2: The trace of the squared first- and second-order terms of the residual Hamiltonian of a PDD cycle as a function of the order of the decoupling elements. From top to bottom: optimal order (i. e. the order which minimizes $\mathrm{tr}\big(\big(\overline{H}_i^{(1)}\big)^2\big)/d$), standard order (close to average performance), and worst order.

Influence of the Traversing Path

Let us start with an examination of the performance of the fundamental decoupling strategy (PDD) based on the decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ constructed using the difference scheme $D(8,8,4)$ listed in table A.2. As it was discussed in the PDD paragraph in subsection 2.3.1, the resulting fidelity decay is mainly due to the first order term in the Magnus expansion of a single decoupling cycle of length $t_c = n_c \Delta t$, and we proposed the estimate (2.77)

$$F_{e\ \mathrm{app}}^{\mathsf{PDD}}(T) = \exp\Big(-\frac{1}{d}\,\mathrm{tr}\big(\big(\overline{H}^{(1)}\big)^2\big) T^2\Big). \tag{2.125}$$

Since, with exception of the vanishing zeroth-order term, all orders in the Magnus expansion depend on the order of the elements in the decoupling scheme, the performance of PDD may be optimized by finding the permutation $\pi \in \mathsf{S}_{n_c}$ which minimizes $\mathrm{tr}\big(\big(\overline{H}_\pi^{(1)}\big)^2\big)$, or in other words by finding an optimal traversing path for the elements of the decoupling scheme. We calculated the latter quantity for all permutations and found that it lies in the range $0.09252 J^4 \Delta t^2 \leq \mathrm{tr}\big(\big(\overline{H}_\pi^{(1)}\big)^2\big)/d \leq 36.963 J^4 \Delta t^2$. Permutations corresponding to these extremal values are shown in table 2.2. We label the PDD strategy based on the optimal path as PDD_1, the one corresponding to the standard path as PDD_2, and the worst one as PDD_3. The resulting fidelities $F_e^{\mathsf{PDD}_i}(T)$, $i \in \{1,2,3\}$, are compared in figure 2.6a (*blue, solid lines*). As it is to be expected from the estimation $F_{e\ \mathrm{app}}^{\mathsf{PDD}}(T)$, we have $F_e^{\mathsf{PDD}_1}(T) > F_e^{\mathsf{PDD}_2}(T) > F_e^{\mathsf{PDD}_3}(T)$. In figure 2.6a we also depicted the improved estimations

$$\begin{aligned} F_{e\ \mathrm{app}}^{\mathsf{PDD}_i}(T) &= \exp\Big(-\frac{1}{d}\Big(\mathrm{tr}\big(\big(\overline{H}_i^{(1)} + \overline{H}_i^{(2)}\big)^2\big)\Big) T^2\Big) \\ &= \exp\Big(-\frac{1}{d}\Big(\mathrm{tr}\big(\big(\overline{H}_i^{(1)}\big)^2\big) + \mathrm{tr}\big(\big(\overline{H}_i^{(2)}\big)^2\big)\Big) T^2\Big), \end{aligned} \tag{2.126}$$

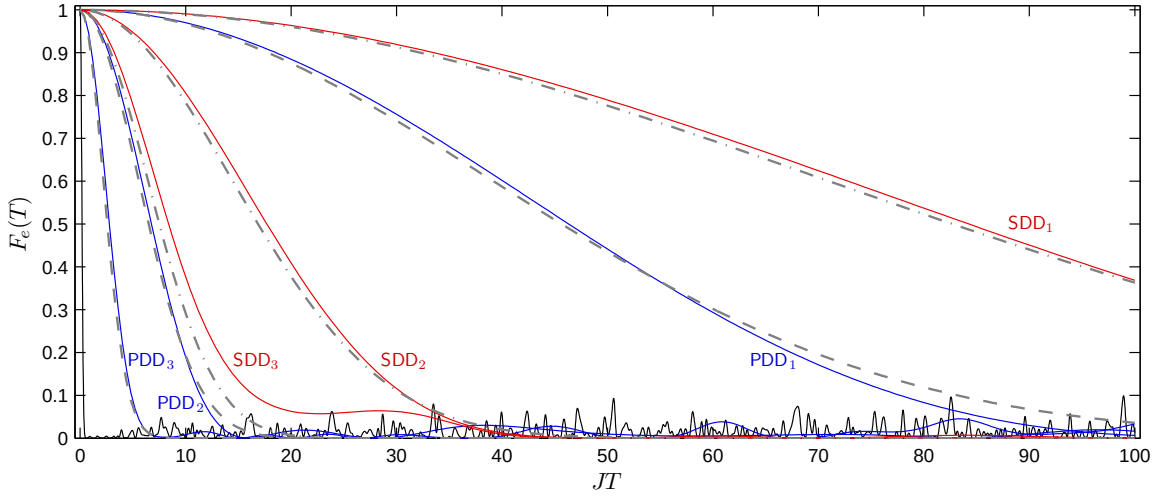
with the first and second-order Magnus term of the i-th path given in table 2.2, as *dashed* lines. It can be seen that they are quite close to the actual curves $F_e^{\mathsf{PDD}_i}(T)$.

The better deterministic control strategy is SDD which achieves a vanishing first-order Magnus term by doubling the length of a single decoupling cycle. Hence, the expected fidelities of the three traversing paths are given by (2.85),

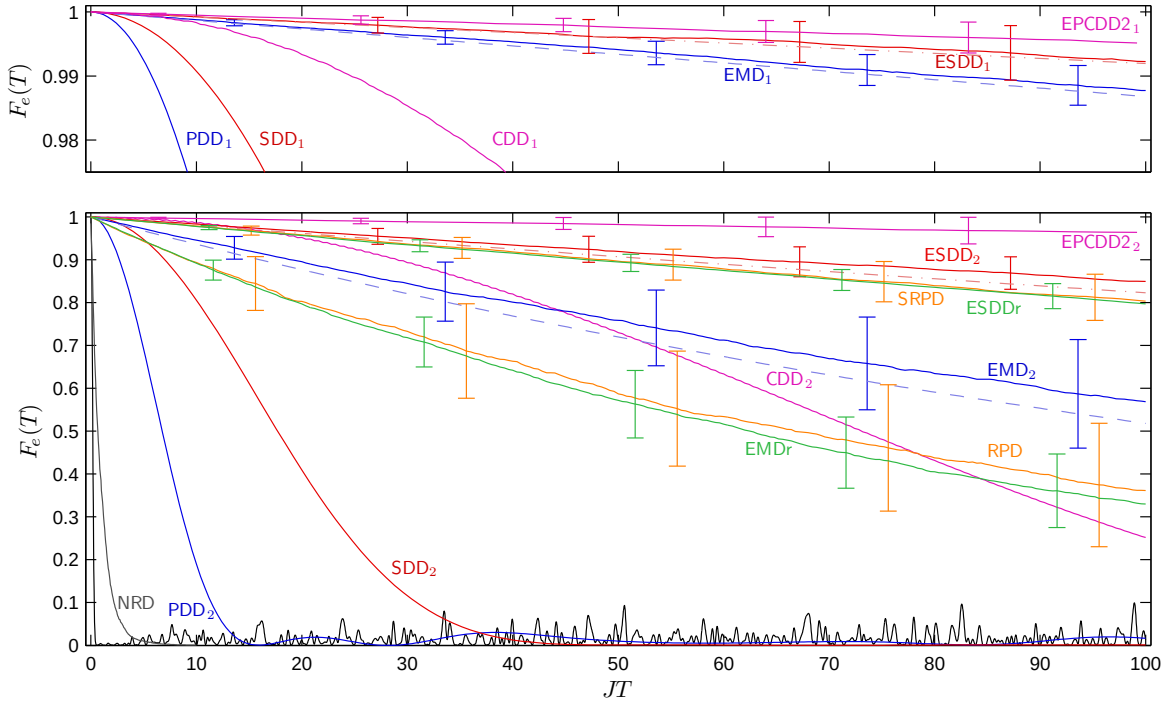
$$F_{e\ \mathrm{app}}^{\mathsf{SDD}_i}(T) = \exp\Big(-\frac{1}{d}\,\mathrm{tr}\big(\big(\overline{H}_i^{(2)}\big)^2\big) T^2\Big). \tag{2.127}$$

They are shown in figure 2.6a as *dashed-dotted* lines, and are in good agreement with the actual SDD fidelities $F_e^{\mathsf{SDD}_i}(T)$ (*red, solid lines*). In principle $F_{e\ \mathrm{app}}^{\mathsf{SDD}_1}(T)$ is not necessarily the best SDD fidelity since we minimized the quantity $\mathrm{tr}\big(\big(\overline{H}_\pi^{(1)}\big)^2\big)$ which is now vanishing. Hence, in order to obtain the optimal SDD fidelity we should search for the permutation π which minimizes $\mathrm{tr}\big(\big(\overline{H}_\pi^{(2)}\big)^2\big)$. Although we did not perform this search (due to computational limitations), we expect the optimal SDD fidelity to be quite close to $F_e^{\mathsf{SDD}_1}(T)$.

The last remaining deterministic strategy we are going to consider is CDD. It turns out that for the model and decoupling scheme under consideration, CDD leads to the same fidelity as PCDD2 repeating



(a) Entanglement fidelity of PDD and SDD for three different permutations of the decoupling scheme.



(b) Entanglement fidelity of various deterministic and randomized strategies. The upper part shows an enlarged representation of the high fidelity range $[0.975, 1]$.

Figure 2.6: The entanglement fidelity of a quantum register with $n = 8$ qubits, perturbed by the Hamiltonian given in (2.122). The time interval between adjacent decoupling pulses is $\Delta t = 0.05J^{-1}$. All randomized fidelities are averaged over 100 individual runs.
(a) Without decoupling (*solid line, black*), with PDD using the decoupling set $\{g_j\}_{j=0}^{7}$ for three different traversing paths (*blue, solid lines*), the corresponding estimations (*dashed lines*), the corresponding SDD fidelities (*red, solid lines*) and their estimations (*dashed-dotted lines*).
(b) Strategies using the standard path (labeled as 2): PDD$_2$ (*blue*), SDD$_2$ (*red*), CDD$_2$ (*purple*) EMD$_2$ (*blue*), ESDD$_2$ (*red*), and EPCDD2$_2$ (*purple*). Fully randomized strategies: NRD (*gray*), RPD (*orange*), EMDr (*green*), SRPD (*orange*), and ESDDr (*green*). The upper part shows strategies using the optimal path (labeled as 1): PDD$_1$ (*blue*), SDD$_1$ (*red*), CDD$_1$ (*purple*) EMD$_1$ (*blue*), ESDD$_1$ (*red*), and EPCDD2$_1$ (*purple*). In addition the estimations (2.128) and (2.129) for EMD$_i$ and ESDD$_i$ are shown (*dashed lines*). The standard deviation of the randomized strategies is indicated by error bars.

a PCDD2 cycle of length $n_c^2\Delta t$. This is a result of the fact that the residual Hamiltonian of such a cycle cannot be eliminated by the decoupling scheme which was designed to eliminate the system Hamiltonian H_0. Again, the fidelity depends on the traversing path of the underlying PDD cycle. We show CDD_i for the optimal PDD path ($i = 1$) and the standard path ($i = 2$) in figure 2.6b (*purple*). It can be seen that the CDD_i fidelity surpasses the SDD_i fidelity. Since, for the model under consideration, the performance of CDD_i is equal to the performance of $\mathsf{PCDD2}_i$, this means that periodic dynamical decoupling using a single $\mathsf{PCDD2}_i$ cycle of length $n_c^2\Delta t$ is superior than periodic dynamical decoupling based on a SDD_i cycle of length $2n_c\Delta t$. Hence, according to the estimating formulas for periodic decoupling strategies, the trace of the square of the residual Hamiltonian of a $\mathsf{PCDD2}_i$ cycle has to be smaller than the one of a SDD_i cycle.

The randomized strategies which depend on the traversing path are EMD and ESDD, for which the estimations (2.106) and (2.111) have been proposed:

$$F_{e\ \mathrm{app}}^{\mathsf{EMD}_i}(T) = \exp\Big(-\frac{1}{d}\Big(\mathrm{tr}\big(\big(\overline{H}_i^{(1)}\big)^2\big) + \mathrm{tr}\big(\big(\overline{H}_i^{(2)}\big)^2\big)\Big)T \cdot n_c\Delta t\Big) \tag{2.128}$$

$$F_{e\ \mathrm{app}}^{\mathsf{ESDD}_i}(T) = \exp\Big(-\frac{1}{d}\,\mathrm{tr}\big(\big(\overline{H}_i^{(2)}\big)^2\big)T \cdot 2n_c\Delta t\Big). \tag{2.129}$$

The improvement over PDD and SDD is the conversion of the quadratic decay into a linear-in-time one. We show the fidelities $F_{e\ \mathrm{num}}^{\mathsf{EMD}_i}(T)$ (*blue*) and $F_{e\ \mathrm{num}}^{\mathsf{ESDD}_i}(T)$ (*red*) for $i = 1, 2$ in the lower and upper part of 2.6b, respectively. The corresponding approximations $F_{e\ \mathrm{app}}^{\mathsf{EMD}_i}(T)$ and $F_{e\ \mathrm{app}}^{\mathsf{ESDD}_i}(T)$ are also shown (*dashed lines*). Analogous to ESDD, we might as well embed the $\mathsf{PCDD2}_i$ cycles into a naive random decoupling scheme based on an annihilator. We label the resulting strategy EPCDD2 for embedded periodic concatenated second level dynamical decoupling. In figure 2.6b, $F_{e\ \mathrm{num}}^{\mathsf{EPCDD2}_i}(T)$ is depicted for $i = 1, 2$ (*purple*). As to be expected from the result that the $\mathsf{PCDD2}_i$ fidelity surpasses the SDD_i fidelity, $\mathsf{EPCDD2}_i$ is superior to ESDD_i. In fact, the best decoupling strategy we found for our model is $\mathsf{EPCDD2}_1$ for the optimized traversing path. It has to be compared with the best previously known strategy in [SV06], which was SRPD (SRPD will be discussed in the next paragraph) and which achieves a fidelity of ≈ 0.8 at $T = 100J^{-1}$, while $\mathsf{EPCDD2}_1$ manages to sustain the fidelity nearly perfectly. The standard deviation of the fidelity of each randomized decoupling strategy is indicated in figure 2.6b by error bars.

Fully Randomized Strategies

Randomized decoupling strategies which do not involve a fixed traversing path through the elements of the decoupling set are NRD, RPD, and EMDr as well as their symmetrized counterparts SRPD and ESDDr. We refer to these strategies as being fully randomized. The NRD fidelity based on the set of Pauli operators performs quite poor, as it can be seen from the *gray* curve in figure 2.6b. This fact can be understood by looking at the estimation given by (2.97),

$$F_{e\ \mathrm{app}}^{\mathsf{NRD}}(T) = \exp\Big(-\frac{1}{d}\,\mathrm{tr}\big(H_0^2\big)T\Delta t\Big). \tag{2.130}$$

Even though the fidelity decay is linear in time, the value of $\mathrm{tr}\big(H_0^2\big)/d \approx 21.30J^2$ is huge compared to the worst (i. e. largest) first-order term $\mathrm{tr}\big(\big(\overline{H}_3^{(1)}\big)^2\big) \approx 36.963J^4\Delta t^2 = 0.0924J^2$ relevant for PDD. A higher suppression of H_0 is obtained by using the random path decoupling (RPD) strategy, which chooses the traversing path through $\{g_j\}_{j=0}^7$ for each successively applied PDD cycle of length $n_c\Delta t = 8\Delta t$ at random. While the EMD fidelity depends on the particular choice of a fixed path, RPD delivers an average EMD fidelity, i. e. we propose that a good approximation is given by

$$F_{e\ \mathrm{app}}^{\mathsf{RPD}}(T) = \exp\Big(-\frac{1}{d}\mathbb{E}_\pi\Big(\mathrm{tr}\big(\big(\overline{H}_\pi^{(1)}\big)^2\big) + \mathrm{tr}\big(\big(\overline{H}_\pi^{(2)}\big)^2\big)\Big)T \cdot n_c\Delta t\Big), \tag{2.131}$$

where $\mathbb{E}_\pi$ denotes the average over all permutations $\pi \in \mathsf{S}_{n_c}$. The numerically obtained fidelity $F_{e\ \mathrm{num}}^{\mathsf{RPD}}(T)$ is depicted in figure 2.6b in *orange*. The symmetrized counterpart of RPD is SRPD and makes use of random SDD cycles of length $2n_c\Delta t$. As a result, SRPD removes the first-order Magnus terms and leads to the improved fidelity

$$F_{e\ \mathrm{app}}^{\mathsf{SRPD}}(T) = \exp\Big(-\frac{1}{d}\mathbb{E}_\pi \operatorname{tr}\big(\big(\overline{H}_\pi^{(2)}\big)^2\big)T \cdot 2n_c\Delta t\Big). \tag{2.132}$$

$F_{e\ \mathrm{num}}^{\mathsf{SRPD}}(T)$ is also shown in figure 2.6b in *orange*. From RPD and SRPD we obtain the strategies EMDr and ESDDr by plugging in additional pulses in between subsequent PDD or SDD cycles, where these additional pulses are constructed by random selection from a second decoupling set (typically an annihilator given by the set of Pauli operators). Since the average over the residual Hamiltonian of the underlying cycles vanishes for the random path strategies even if we do not apply this additional embedding[‡], we expect the resulting fidelity to be effectively identical with the one of RPD and SRPD. This fact is confirmed by the data shown in figure 2.6b, although a bit surprisingly the EMDr and ESDDr fidelities appear to be slightly worse. Nevertheless, the EMDr and ESDDr fidelities shown in figure 2.6b (*green*) indicate an advantage: The square root of the variance indicated by the length of the error bars is approximately only half the size as the corresponding quantity for RPD and SRPD. This feature might be important in practice, since it is a priori unknown whether a particular single run of a randomized strategy delivers a fidelity above or below average.

2.4.4 Conclusions

The general guideline for the construction of a good decoupling strategy for a system Hamiltonian H_0 turned out to be the following:

- We start by looking for a deterministic strategy, for which the average Hamiltonian $\overline{H} = \overline{H}^{(0)} + \overline{H}^{(1)} + \overline{H}^{(2)} + \dots$ of a basic decoupling cycle gets as small as possible. Such a strategy is usually based on a decoupling scheme of length n_c for H_0, which satisfies the decoupling condition $\overline{H}^{(0)} = 0$. In order to minimize the residual Hamiltonian, the length n_c should be as small as possible (since we have $\overline{H}^{(i)} = \mathcal{O}((H_0)^{i+1}(n_c\Delta t)^i)$). The standard trick to improve a given decoupling scheme is to make it symmetric in time. Even though the length of such a symmetrized scheme is twice the length of the basic decoupling scheme, this leads to a vanishing first-order term $\overline{H}^{(1)}$. In addition we saw that the residual Hamiltonian depends on the order of the elements of the decoupling scheme. By finding an optimal order, the remaining quantity $\overline{H}^{(1)}$ (or for $\overline{H}^{(1)} = 0$ the quantity $\overline{H}^{(2)}$) can be minimized. For our example, the basic decoupling scheme was based on a difference scheme of length $n_c = 8$ and the best deterministic decoupling strategy we found was the PCDD2 cycle of length n_c^2 for an order of the decoupling elements which minimized the quantity $\operatorname{tr}\big(\big(\overline{H}^{(1)}\big)^2\big)$.
- The second step is to suppress the residual Hamiltonian. In principle we could use the same guideline that was used in the first step for the suppression of H_0, but because of the complicated structure of the typically highly correlated residual Hamiltonian, a small decoupling scheme usually does not exist. Instead we have to use an annihilator like the set of Pauli operators. Because of the large length of this second decoupling scheme (which is equal to the square of the dimension of the system Hilbert space), now the method of choice is naive random decoupling. Hence, we end up with an embedded decoupling scheme. For our example, the best result was obtained for EPCDD2, while the second best result was obtained for ESDD (in both cases for an optimal order of the decoupling elements).

[‡]This might not be true for terms of third and higher order in the Magnus expansion of a basic cycle.

While it might be hard to find a deterministic strategy which surpasses SDD for a given decoupling scheme, the SDD strategy can always be applied. If we are not able to determine a good order of the decoupling elements, we might ensure at least an average performance by using the symmetric random path strategy (SRPD) instead of embedding the SDD strategy. The variance of SRPD can then be minimized by an additional embedding of the basic SRPD cycles in a naive random decoupling strategy based on an annihilator (leading to ESDDr). In addition, SRPD is the method of choice if we cannot afford the second decoupling scheme, i. e. if we are restricted to apply only pulses of the form $g_i g_j^\dagger$, with g_i being an element of the basic decoupling scheme $\{g_j\}_{j=0}^{n_c-1}$ for H_0.

Let us close this chapter by giving a small outlook. According to the results presented in the last subsection, NRD alone seems to be a rather poor choice for decoupling. Nevertheless it holds many useful features: For example, it can be applied even if the system Hamiltonian is time dependent. Even more important, in chapter 3 NRD turns out to be applicable even if the decoupling pulses have to be implemented using bounded controls, and in addition, it turns out to be able to stabilize quantum computations.

So far only the control task of decoupling has been considered. We expect similar results for the task of simulating a non-vanishing Hamiltonian. For example, the potential of ESDD for the simulation of a two qubit gate Hamiltonian in the context of a selective decoupling scheme will be explored in chapter 4.

The assumption that the decoupling pulses can be applied in a perfect manner is a strong idealization. In practice, each pulse will be non-ideal and we have to distinguish between systematic and random pulse errors. An important question is how such errors affect the performance of a given decoupling strategy. First results concerning this question have been obtained by Santos in Viola with the help of numerical simulations [SV08]. In addition, the question arises whether decoupling sequences might be designed that are stable against pulse imperfections. For instance, an Eulerian decoupling cycle (as discussed in subsection 2.1.7) projects any systematic errors of the decoupling pulses (which are elements of the group algebra $\mathcal{A} = R(\mathbb{C}G)$) into the commutant $\mathcal{A}'$ and an additional subsystem encoding might protect against these residual errors [VK03].

Remark. The latter fact can be seen by looking at equation (2.51) in which the effect of systematic pulse errors is reflected by replacing the left H_0 by $H_0 + H_i^{\mathrm{err}}(t')$ where $H_j^{\mathrm{err}}(t')$ specifies the error of the pulse

$$p_j^{\mathrm{err}} = p_j^{\mathrm{err}}(\tau_p) = \mathcal{T} \exp\Big(-i \int_0^{\tau_p} \big(H_j(t') + H_j^{\mathrm{err}}(t')\big) dt'\Big),$$

while the corresponding ideal pulse is given by

$$p_j = p_j(\tau_p) = \mathcal{T} \exp\Big(-i \int_0^{\tau_p} H_j(t') dt'\Big).$$

3 Decoupling and Computation

In chapter 2 we studied dynamical decoupling methods which were designed to suppress the influence of imperfections in a quantum memory. A more demanding goal is to use these methods to protect a running quantum computation, which consists of a sequence of one- and two-qudit quantum gates. While we assumed in chapter 2 that the decoupling pulses are applied quasi-instantaneously using a *strong* local control Hamiltonian (with the exception of subsection 2.1.7), we are going to assume that the experimentally more demanding quantum gates (especially the two-qudit quantum gates) are realized by applying a *weak* gate Hamiltonian over a finite time interval τ_g larger than the time interval Δt in between subsequent decoupling pulses. As a consequence, in general, the applied decoupling scheme also alters the gate Hamiltonians. Solutions for this fundamental problem have been discussed by Viola et al. in [VLK99]. In particular, by using a subsystem encoding it becomes possible to achieve universal control via a set of gate Hamiltonians which commute with the decoupling pulses, and hence remain unaffected. For example, the hybrid decoupling and computing scheme analyzed in [KL08] by Khodjasteh and Lidar is based on the above approach. Even more general, we might assume that the decoupling pulses are realized over a finite time interval as well. In this case the dynamically corrected gates based on an Eulerian decoupling cycle (Euler-DCGs) proposed recently by Khodjasteh and Viola [KV09] are able to achieve simultaneous computation and decoupling: An Euler-DCG is generated by extending an Eulerian path in the Cayley graph of the Eulerian decoupling strategy [VK03] described in subsection 2.1.7, by applying a corresponding gate Hamiltonian after completing the path. In addition, in order to get a vanishing lowest order average Hamiltonian, a gate leading to the same error as the gate Hamiltonian, but implementing the identity, is applied after visiting each of the non-identity vertices in the Cayley graph for the last time.

In this chapter we consider the most general setting, i. e. we consider decoupling pulses which are generated by applying a local control Hamiltonian for a time τ_p and quantum gates which are generated by applying a two-qudit gate Hamiltonian for a time τ_g. We are going to show that a quantum computation can be stabilized against static imperfections by executing the quantum gates in between subsequent decoupling pulses. This is in contrast with the Euler-DCGs of Khodjasteh and Viola [KV09], where a quantum gate is effectively implemented only in between completed cycles. Thereby, our decoupling pulses are constructed by random selection from an annihilator as the set of Pauli operators, or in other words by using the naive random decoupling (NRD) strategy presented in the preceding chapter. Our method has been published in [KAS05], where we devised the acronym Pauli random error correction (PAREC), and provided numerical evidence of its error suppressing properties. We derive a formula for the fidelity decay of a stabilized quantum computation (for the special case of instantaneous gates and pulses we derived such a formula in [GKAJ08]). A numerical simulation of the PAREC method is performed for the quantum computation of a quantum map running on a quantum computer perturbed by Heisenberg couplings. The PAREC method is compared with an idea of Prosen and Žnidarič [PŽ01], who proposed to stabilize a quantum computation against static imperfections by increasing the decay of the correlation function measuring the fidelity decay. It turns out that our approach does exactly that, i. e. it leads to an ultimate decay of correlations. Eventually, we consider the Euler-DCGs of Khodjasteh and Viola [KV09]. By implementing each quantum gate as an Euler-DCG, a deterministic decoupling method for quantum computations is obtained. We propose to implement the PAREC method by using only Euler-DCGs in order to benefit from the advantages of both methods.

Another scenario in which the decoupling strategies of the preceding chapter may be used to improve the performance of a quantum computation is given if the quantum gates are implemented using a selective decoupling scheme. It will be dealt with in chapter 4.

We start by presenting an overview of known results on the fundamental problem of combining quantum computation and dynamical decoupling in section 3.1. The PAREC method based on the randomized decoupling strategy is presented, analyzed and simulated in section 3.2. In section 3.3, we compare the PAREC method with the idea of Prosen and Žnidarič [PŽ01], who proposed to increase the correlation decay. Eventually, we present the Euler-DCGs of Khodjasteh and Viola [KV09] in section 3.4 and show how they might be combined with the PAREC method.

3.1 Decoupling and Quantum Logic

Let us consider a quantum register S defined on a d-dimensional Hilbert space $\mathcal{H}_S$. Typically the register consists of n qudits of dimension q such that $d = q^n$. For the sake of simplicity, we assume S to be a closed system perturbed by static imperfections modeled by the system Hamiltonian H_0 acting on $\mathcal{H}_S$. (It is straightforward to extend any of the forthcoming results to the case where S is an open system coupled to an environment E via a set of coupling operators as in subsections 2.1.5 and 2.1.6). In this section we assume that the decoupling pulses are applied quasi-instantaneously by using a strong local control Hamiltonian, or in other words, by using bang-bang control, but all results are also applicable if the Euler decoupling method ([VK03], subsection 2.1.7) for bounded strength control is applied. The fundamental control strategy, called periodic dynamic decoupling (PDD, subsection 2.3.1), repeats a basic control cycle traversing all the elements of a control scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ over and over again. The length $t_c = n_c \Delta t$ of such a basic cycle is determined by the number n_c of elements in the control scheme and by the time Δt in between subsequent pulses. Let us assume now, that we would like to generate a certain two-qudit quantum gate by applying a possibly time-dependent gate Hamiltonian $H_g(t)$ for a time $\tau_g = m \cdot t_c$, $m \in \mathbb{N}$. Then, the total Hamiltonian is given by the sum of the Hamiltonians describing the static imperfections (H_0), the quantum gate ($H_g(t)$), and the decoupling pulses ($H_c(t)$),

$$H(t) = H_0 + H_g(t) + H_c(t), \tag{3.1}$$

for $t \in [0, \tau_g]$. As in section 2.1, we switch to the toggled frame $\tilde{U}(t) = U_c^\dagger(t) \cdot U(t)$. As a result of the control, we obtain (in lowest order AHT) the effective total Hamiltonian

$$\overline{H}^{(0)} = \Pi_{\mathcal{G}}(H_0) + \Pi_{\mathcal{G}}(H_g), \tag{3.2}$$

where we assumed for simplicity that the gate Hamiltonian remains constant over the time interval τ_g, and where we used the definition

$$\Pi_{\mathcal{G}}(X) = \frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger X g_j, \tag{3.3}$$

for any operator X acting on $\mathcal{H}_S$. Hence, any gate Hamiltonian gets altered by the applied decoupling scheme. In particular, a time-independent gate Hamiltonian H_g becomes $\Pi_{\mathcal{G}}(H_g)$. We are now going to discuss solutions to this problem. For the remaining section, let us assume that the elements of the control scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ are defined by a unitary projective representation R of a group $G = \{\mathfrak{g}_j\}_{j=0}^{n_c-1}$ acting on the system Hilbert space $\mathcal{H}_S$, i. e. we assume that $g_j = R(\mathfrak{g}_j)$. We will call G the underlying index group. Assuming that the elements in $\mathcal{G}$ generate a larger but finite group $\hat{G}$, we consider the ordinary irreducible representations of $\hat{G}$. As in subsections 2.1.6 and 2.1.7 we denote the corresponding group algebra $R(\mathbb{C}G)$ by $\mathcal{A}$ and its commutant by $\mathcal{A}'$.

3.1.1 Universal Computation on a Subsystem

As discussed in subsection 2.1.6, the Hilbert space of the quantum register decomposes with respect to the irreps $\mathcal{J}$ of $\mathcal{G}$,

$$\mathcal{H}_S = \bigoplus_{\nu\in\mathcal{J}} \mathcal{H}_\nu = \bigoplus_{\nu\in\mathcal{J}} \mathcal{C}_\nu \otimes \mathcal{D}_\nu, \tag{3.4}$$

where $\tau_\nu = \dim(\mathcal{C}_\nu)$ denotes the degeneracy and $d_\nu = \dim(\mathcal{D}_\nu)$ denotes the dimension of the irrep $\nu \in \mathcal{J}$. Since, for any operator X acting on $\mathcal{H}_S$, $\Pi_\mathcal{G}(X)$ commutes with all the group elements, it follows that $\Pi_\mathcal{G}(X)$ is in $\mathcal{A}'$. Hence, the subsystems $\{\mathcal{D}_\nu\}_{\nu\in\mathcal{J}}$ are dynamically generated noiseless subsystems ([Zan00; VKL00], subsection 2.1.6). In order to generate a universal set of gates acting on subsystem $\mathcal{D}_\nu$, we have to apply gate Hamiltonians which belong to the group algebra $\mathcal{A}$. Unfortunately, according to equation (3.2), this is impracticable since such a Hamiltonian gets projected onto $\mathcal{A}'$. A very elegant solution appears for the case that $\Pi_\mathcal{G}(H_0) \in \mathcal{A}' \cap \mathcal{A} = \bigoplus_{\nu\in\mathcal{J}} \lambda_\nu \mathcal{I}_\nu$, with $\lambda_\nu \in \mathbb{C}$ and $\mathcal{I}_\nu$ denoting the identity acting on $\mathcal{H}_\nu$: In this case we might use one of the subsystems $\{\mathcal{C}_\nu\}_{\nu\in\mathcal{J}}$ as a noiseless subsystem and generate the corresponding quantum gates using a gate Hamiltonian belonging to $\mathcal{A}'$. Any Hamiltonian belonging to $\mathcal{A}'$ remains unaffected by the action of $\Pi_\mathcal{G}$ [Zan00; VKL00]. The method becomes infeasible if $\mathcal{G}$ acts irreducible on $\mathcal{H}_S$. Then, the set $\mathcal{J}$ contains only one element ν with $d_\nu = \dim(\mathcal{H}_S)$ and $\tau_\nu = 1$.

In the above scenario, universal control is achieved via a set of gate Hamiltonians which commute with the decoupling pulses, and hence remain unaffected. For instance, the hybrid decoupling and computing scheme analyzed in [KL08] by Khodjasteh and Lidar is based on the assumption that the computational operations commute with the decoupling pulses.

3.1.2 Universal Computation using Multiple Decoupling Schemes

By using a decoupling scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ defined by a unitary projective representation R, any time-independent gate Hamiltonian H_g gets projected onto the commutant $\mathcal{A}'$ of the group algebra $\mathcal{A}$ via $\Pi_\mathcal{G}(H_g)$ (compare with (3.2)). Hence, the only applicable gate Hamiltonians are those which belong to $\mathcal{A}'$. If an additional decoupling group $\tilde{\mathcal{G}} = \{\tilde{g}_j\}_{j=0}^{\tilde{n}_c-1}$, with group algebra $\tilde{\mathcal{A}}$ and commutant $\tilde{\mathcal{A}}'$, is available, it becomes also possible to apply any gate Hamiltonian belonging to $\tilde{\mathcal{A}}'$. Let $A \in \mathcal{A}'$ and let $B \in \tilde{\mathcal{A}}'$. It was recognized by Viola et al. in [VLK99], that by applying A and B interchangeably, any gate $U_g = e^L$ could be created, where L belongs to the Lie algebra generated by iA and iB under commutation. Additional decoupling groups $\tilde{\mathcal{G}}$ might be generated by employing the following trick: We apply the additional bang-bang pulses P and $P^\dagger$ at the beginning and the end of a single $\mathcal{G}$-decoupling cycle, respectively. As a result, the time evolution of a single PDD cycle is changed from

$$\tilde{U}(t_c) = \exp\big(-ig_{n_c-1}^\dagger(H_0+H_g)g_{n_c-1}\Delta t\big)\ldots\exp\big(-ig_1^\dagger(H_0+H_g)g_1\Delta t\big)\exp\big(-ig_0^\dagger(H_0+H_g)g_0\Delta t\big) \tag{3.5}$$

(compare with (2.74)) to $P^\dagger\tilde{U}(t_c)P$, and lowest order AHT leads to $H_g \mapsto \Pi_{\tilde{\mathcal{G}}}(P^\dagger H_g P) \in \tilde{\mathcal{A}}'$ with $\tilde{\mathcal{G}} = P^\dagger\mathcal{G}P$. The decoupling of $H_0 \mapsto \Pi_\mathcal{G}(H_0) = \lambda\cdot\mathcal{I}$ (with $\lambda\in\mathbb{R}$) remains unaffected since $\Pi_{\tilde{\mathcal{G}}}(P^\dagger H_0 P) = P^\dagger\Pi_\mathcal{G}(H_0)P = P^\dagger\lambda\mathcal{I}P = \lambda\cdot\mathcal{I}$. Note that for $\tilde{\mathcal{A}}' \neq \mathcal{A}'$, P must not be in $\mathcal{A}$. If, in addition to $\mathcal{G}$, a large enough set of bang-bang pulses $P \notin \mathcal{A}$ is available, it might become feasible to construct a universal set of gates [VLK99]. Again, the method becomes infeasible if $\mathcal{G}$ acts irreducible on $\mathcal{H}_S$: Then, $\mathcal{A}' = \lambda\mathcal{I}$, with $\lambda\in\mathbb{C}$, generates only a trivial action.

3.1.3 Gates via Fast Switching

In the previous two subsections we assumed that a gate Hamiltonian H_g was switched on over a period corresponding to an integer number of decoupling cycles, each of which is of length $t_c = n_c\cdot\Delta t$. As a consequence, in lowest order AHT, H_g became projected onto $\Pi_\mathcal{G}(H_g)$. Let us now assume that we are able to switch H_g on and off for shorter periods Δt, a scenario which is called 'weak strength/fast

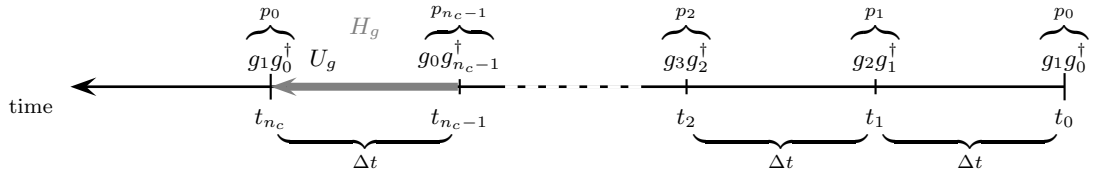


Figure 3.1: Schematic representation of a PDD cycle, which tries to implement a quantum gate $U_g = \mathcal{T}\exp\big(-i\int_0^{\tau_g} H_g(t')dt'\big)$ with $\tau_g = \Delta t$ by switching on the gate Hamiltonian H_g during the period where the control visits the identity element g_0 of the control scheme $\{g_j\}_{j=0}^{n_c-1}$.

switching' in [VLK99]. If H_g is switched on only during the interval Δt corresponding to the identity element $g_0 \in \mathcal{G}$, lowest order AHT leads to

$$\overline{H}^{(0)} = \Pi_{\mathcal{G}}(H_0) + \frac{1}{|\mathcal{G}|} H_g. \tag{3.6}$$

Now any quantum gate $U_g = \exp(-iH_g \cdot m t_c)$ with $m \in \mathbb{N}$ could be generated by repeating such a cycle an integer number of times. If we are also able to switch on the Hamiltonians $g_j H_g g_j^\dagger$ during the j-th part of the cycle (for $j = 1, \ldots, n_c - 1$), the factor $1/|\mathcal{G}|$ in the above equation vanishes [VLK99]. Note that this method works even if the control scheme $\mathcal{G}$ is not related to an underlying index group.

3.1.4 Dynamically Corrected Gates

In this subsection we present an idea due to Khodjasteh and Viola [KV09], who proposed to combine decoupling and computation by constructing dynamically corrected gates (DCGs)*. We consider a decoupling scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ of length n_c, where $g_0 = \mathcal{I}$ denotes the identity element. The basic PDD cycle of length $t_c = n_c \cdot \Delta t$ is constructed by traversing the elements of the decoupling scheme in the order $g_1, g_2, \ldots, g_{n_c-1}, g_0$, i. e. we close the cycle by visiting the identity element. If the gates implementing a quantum computation could be generated instantaneously, they could simply be executed in between subsequent cycles without introducing any errors. Instead, we assume that a quantum gate U_g has to be generated by switching on a time-dependent gate Hamiltonian $H_g(t)$ for a time $\tau_g = \Delta t$: $U_g \equiv U_g(\tau_g)$ with $U_g(t) = \mathcal{T}\exp\big(-i\int_0^t H_g(t')dt'\big)$ for $t \in [0, \tau_g]$. In order to combine a decoupling cycle with the generation of a quantum gate U_g, we apply the corresponding gate Hamiltonian during the last part of the cycle, in which the control visits the identity element. A schematic representation is given in figure 3.1. As a consequence, the time evolution of such a cycle is given by

$$\tilde{U}(t_c) = U_g \cdot \mathcal{T}\exp\Big(-i\int_0^{\Delta t} U_g^\dagger(t') H_0 U_g(t') dt'\Big) \cdot \exp\big(-i g_{n_c-1}^\dagger H_0 g_{n_c-1} \Delta t\big) \ldots \exp\big(-i g_1^\dagger H_0 g_1 \Delta t\big), \tag{3.7}$$

and in lowest order AHT the average Hamiltonian of such a cycle is given by

$$\overline{H}^{(0)} = \frac{1}{n_c \Delta t}\Bigg(g_0^\dagger \underbrace{\int_0^{\Delta t} U_g^\dagger(t') H_0 U_g(t') dt'}_{\Phi_g} g_0 + \sum_{j=1}^{n_c-1} g_j^\dagger H_0 g_j \Delta t\Bigg). \tag{3.8}$$

Because of the lowest order gate error Φ_g, we do not obtain the usual result $\overline{H}^{(0)} = \Pi_{\mathcal{G}}(H_0)$. The idea of Khodjasteh and Viola [KV09] is now to produce the same error during all the non-identity steps of

*In [KV09] the idea of dynamically corrected gates was presented in the context of Eulerian decoupling using bounded controls; here we consider the simpler case of instantaneous decoupling pulses.

the decoupling cycle. As a result, the lowest order average Hamiltonian of such a cycle would be given by

$$\overline{H}^{(0)} = \frac{1}{n_c \Delta t} \sum_{j=0}^{n_c-1} g_j^\dagger \Phi_g g_j = \Pi_{\mathcal{G}}\Big(\frac{1}{\Delta t}\int_0^{\Delta t} U_g^\dagger(t') H_0 U_g(t') dt'\Big). \tag{3.9}$$

The above expression leads to a trivial time evolution, if we demand a decoupling scheme which satisfies $\Pi_{\mathcal{G}}\big(\Phi_g\big) = \lambda \cdot \mathcal{I}$, with $\lambda \in \mathbb{C}$ (this point will be further discussed in subsection 3.4.1 dealing with Euler-DCGs).

We close this subsection by showing how these additional errors could be generated. Khodjasteh and Viola [KV09] proposed the following trick: Let us assume that the quantum gate $U_g \equiv U_g(\tau_g) = \exp(-iH_g\tau_g)$ is generated using a fixed gate Hamiltonian H_g whose strength is modulated by a time-dependent pulse shape $f(t)$ such that $\int_0^1 f(t')dt' = 1$:

$$U_g(t) = \exp\Big(-iH_g\tau_g \cdot \frac{1}{\tau_g}\int_0^t f(t'/\tau_g)dt'\Big). \tag{3.10}$$

Assuming that $f(t) = 0$ for $t \notin [0,1]$ we could generate an identity gate $\mathcal{I} \equiv U_I(\tau_g)$ by using the following pulse shape:

$$U_I(t) = \exp\Big(-iH_g\tau_g \cdot \frac{2}{\tau_g}\int_0^t \big(f(2t'/\tau_g) - f(2-2t'/\tau_g)\big)dt'\Big). \tag{3.11}$$

Calculating the lowest order error Φ_I of such an identity gate,

$$\Phi_I = \int_0^{\Delta t} U_I^\dagger(t') H_0 U_I(t') dt', \tag{3.12}$$

is straightforward and shows that indeed $\Phi_I = \Phi_g$. Hence, in order to generate the additional errors, we have to implement these identity gates by switching on the Hamiltonian in the exponent of (3.11) during the first $n_c - 1$ steps of the decoupling cycle.

3.2 Pauli Random Error Correction

The methods for quantum computation in the presence of decoupling, which have been discussed in the preceding section, all have some drawbacks: The first two proposals, subsystem-encoding and multiple decoupling schemes, become infeasible if the decoupling group acts irreducible on the system Hilbert space. The fast-switching method demands the ability to switch a gate Hamiltonian on and off quickly, and in addition, weakens the interaction strength of any applied gate Hamiltonian by a factor in inverse proportion to the size of the decoupling set. Eventually, dynamically corrected quantum gates demand a decoupling set which satisfies the decoupling condition for perturbations which have been twisted by the gate errors (3.9), and in addition, demands the generation of additional identity-gates mirroring the gate errors.

We are now going to present a method which uses naive random decoupling (NRD, subsection 2.3.2) to stabilize arbitrary quantum algorithms against static imperfections (like inter-qudit couplings, for instance) in a rather simple way. While any method based on deterministic decoupling strategies, like for instance the method of dynamically corrected gates ([KV09], subsection 3.1.4), is only allowed to implement quantum gates in between completed decoupling cycles, random decoupling allows the quantum gates to be implemented in between subsequent decoupling pulses. It will be shown that, as it is the case for NRD in the absence of any computation, the fidelity decay caused by static imperfections will be slowed down to a linear-in-time one. Our method was proposed for the first time in the author's diploma thesis [Ker04] and subsequently in [KAS05], where the acronym Pauli random error correction

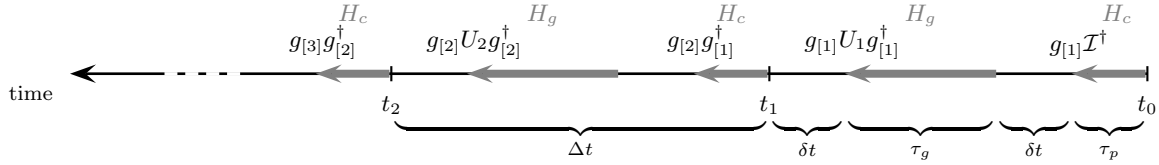


Figure 3.2: Schematic representation of the PAREC method. The gate sequence of the original quantum algorithm $U_{QA} = \cdots \cdot U_3 \cdot U_2 \cdot U_1$ is replaced by an alternating sequence of randomly chosen decoupling pulses $g_{[i+1]}g^\dagger_{[i]}$ of duration τ_p generated by the local control Hamiltonian $H_c(t)$, and twisted quantum gates $g_{[i]}U_ig^\dagger_{[i]}$ of duration τ_g generated by a gate Hamiltonian $H_g(t)$.

(PAREC) was devised. In these publications, all pulses and gates were assumed to be of the bang-bang kind, and only numerical evidence of the resulting linear-in-time decay was provided. We derived a formula for the resulting fidelity decay in [GKAJ08]. In this section, we consider the more general case of bounded controls generating finite decoupling pulses of duration τ_p and finite quantum gates of duration τ_g.

We start with a detailed description of the PAREC method in subsection 3.2.1. To evaluate the stabilizing properties of PAREC, we have to compare a stabilized computation with an unprotected one. Before we proceed with an analysis of the fidelity decay of an unprotected quantum computation in subsection 3.2.3, we derive a general second order expansion of the entanglement fidelity of a perturbed quantum algorithm in subsection 3.2.2. The fidelity decay of a stabilized computation is analyzed in subsection 3.2.4. Eventually, in subsection 3.2.5, we present the results of a numerical simulation of a protected and an unprotected quantum algorithm, which allow us to put the derived fidelity formulas to the test.

3.2.1 Implementation

Let us consider $n_a \in \mathbb{N}$ iterations of a quantum algorithm given by the ideal unitary transformation $U_{QA} = U_{n_g} \cdots U_3 \cdot U_2 \cdot U_1$, where U_i, $i = 1, \ldots, n_g$, denotes an elementary one- or two-qudit quantum gate. In the PAREC method before each quantum gate U_i of the τ-th iteration ($\tau = 1, ..., n_a$) of the unitary transformation U_{QA}, a unitary of the form $g_{[\tau,i]}g^\dagger_{[\tau,i-1]}$ is applied. Here, the unitaries $g_{[\tau,i]}$ (with $g_{[\tau,0]} = g_{[\tau-1,n_g]}$ and $g_{[1,0]} = \mathcal{I}$) are drawn at random from a decoupling set $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$, i. e. the index $[\tau, i]$ is in $\{0, 1, \ldots, n_c - 1\}$ for all $\tau = 1, ..., n_a$ and all $i = 1, \ldots, n_g$. Simultaneously the changes on the quantum algorithm due to these random unitary gates have to be compensated by replacing each elementary quantum gate U_i of the τ-th iteration of the original algorithm by $U_i^{(\tau)} = g_{[\tau,i]}U_ig^\dagger_{[\tau,i]}$. The locality of the control assures that any quantum gate acting on m qudits remains an m-qudit gate: With $g_{[\tau,i]} = g_{[\tau,i],1} \otimes g_{[\tau,i],2} \otimes \cdots \otimes g_{[\tau,i],n} \in \mathsf{U}_q^{\otimes n}$ it follows that

$$\begin{aligned} U_i^{(\tau)} &= g_{[\tau,i]} \cdot U_i \cdot g^\dagger_{[\tau,i]} \\ &= \left(g_{[\tau,i],k_1} \otimes g_{[\tau,i],k_2} \cdot U_i \cdot g^\dagger_{[\tau,i],k_1} \otimes g^\dagger_{[\tau,i],k_2}\right) \otimes \mathcal{I}_{\{1,\ldots,n\}\setminus\{k_1,k_2\}}, \end{aligned} \tag{3.13}$$

for any $m = 2$ qudit gate U_i acting on qudits k_1 and k_2, for instance. Furthermore, after the last quantum gate $U_{n_g}^{(n_a)}$ a final unitary gate $g^\dagger_{[n_a,n_g]}$ is applied. As a result each iteration of a unitary transformation U_{QA} is replaced by $2n_g$ unitary quantum gates so that after n_a iterations one obtains

the result

$$\begin{aligned} U_{QA}^{n_a} &= U_{QA}\ldots U_{QA}\cdot U_{QA} \\ &= g^\dagger_{[n_a,n_g]}\big(U^{(n_a)}_{n_g}\cdot\ldots\cdot g_{[n_a,3]}g^\dagger_{[n_a,2]}\cdot U_2^{(n_a)}\cdot g_{[n_a,2]}g^\dagger_{[n_a,1]}\cdot U_1^{(n_a)}\cdot g_{[n_a,1]}g^\dagger_{[n_a-1,n_g]}\big)\cdot \\ &\qquad \ldots\big(U^{(2)}_{n_g}\cdot\ldots\cdot g_{[2,3]}g^\dagger_{[2,2]}\cdot U_2^{(2)}\cdot g_{[2,2]}g^\dagger_{[2,1]}\cdot U_1^{(2)}\cdot g_{[2,1]}g^\dagger_{[1,n_g]}\big)\cdot \\ &\qquad\qquad \big(U^{(1)}_{n_g}\cdot\ldots\cdot g_{[1,3]}g^\dagger_{[1,2]}\cdot U_2^{(1)}\cdot g_{[1,2]}g^\dagger_{[1,1]}\cdot U_1^{(1)}\cdot g_{[1,1]}\mathcal{I}^\dagger\big) \\ &\equiv g^\dagger_{[n_a,n_g]}\big(V^{(n_a)}_{2n_g}\ldots V_2^{(n_a)}V_1^{(n_a)}\big)\ldots\big(V^{(2)}_{2n_g}\ldots V_2^{(2)}V_1^{(2)}\big)\big(V^{(1)}_{2n_g}\ldots V_2^{(1)}V_1^{(1)}\big) \end{aligned} \tag{3.14}$$

with $V_{2k}^{(\tau)} = U_k^{(\tau)}$ and $V_{2k-1}^{(\tau)} = g_{[\tau,k]}g^\dagger_{[\tau,k-1]}$ for $k = 1,...,n_g$. A particular PAREC implementation of the quantum Fourier transform (QFT) is schematically represented in figure 3.4 for the special case of $n = 4$ qubits and $\mathcal{G} = \mathcal{P}_2^n$ given by the set of Pauli operators (2.58). Definitely, this random application of decoupling elements together with the associated change of elementary quantum gates does not affect any quantum algorithm.

In this section, we consider the general case of bounded controls, i. e. we assume that the decoupling pulses $V_{2k-1}^{(\tau)} \equiv V_{2k-1}^{(\tau)}(\tau_p)$ are generated by switching on a local control Hamiltonian H_c for a time τ_p,

$$V_{2k-1}^{(\tau)}(t) = \mathcal{T}\exp\Big(-i\int_0^t H_c(t')dt'\Big),\ \text{for } t\in[0,\tau_p], \tag{3.15}$$

and the quantum gates $V_{2k}^{(\tau)} \equiv V_{2k}^{(\tau)}(\tau_g)$ are generated by switching on a gate Hamiltonian H_g for a time τ_g,

$$V_{2k}^{(\tau)}(t) = \mathcal{T}\exp\Big(-i\int_0^t H_g(t')dt'\Big),\ \text{for } t\in[0,\tau_g]. \tag{3.16}$$

The situation is depicted in figure 3.2, where we consider a single iteration of U_{QA}. The time in between subsequent decoupling pulses is denoted as usual as Δt. As a consequence, the time δt of free evolution in between gates and pulses is given by $\delta t = (\Delta t - \tau_g - \tau_p)/2$.

While equation (3.14) denotes the ideal time evolution of an iterated quantum algorithm $U_{QA}^{n_a}$ employing the PAREC method, the total time evolution in the presence of static imperfections described by a Hamiltonian H_0 is given by

$$\begin{aligned} U^{n_a}_{QA\ \text{perturbed}} = \big(G^{(n_a)}_{n_g}\ldots G_2^{(n_a)}G_1^{(n_a)}\big)\cdot\ldots\cdot\big(G^{(2)}_{n_g}\ldots G_2^{(2)}G_1^{(2)}\big)\cdot \\ \big(G^{(1)}_{n_g}\ldots G_2^{(1)}G_1^{(1)}\big)\cdot\mathcal{T}\exp\Big(-i\int_0^{\tau_p} V_1^{(1)\dagger}(t')H_0V_1^{(1)}(t')dt'\Big), \end{aligned} \tag{3.17}$$

where we used the abbreviations

$$\begin{aligned} G_k^{(\tau)} = g^\dagger_{[\tau,k]}\cdot\mathcal{T}\exp\Big(-i\int_0^{\tau_p} V_{2k+1}^{(\tau)\dagger}(t')H_0V_{2k+1}^{(\tau)}(t')dt'\Big)\cdot\exp(-iH_0\delta t)\cdot \\ \underbrace{g_{[\tau,k]}\cdot U_k\cdot g^\dagger_{[\tau,k]}}_{V_{2k}^{(\tau)}(\tau_g)}\cdot\mathcal{T}\exp\Big(-i\int_0^{\tau_g} V_{2k}^{(\tau)\dagger}(t')H_0V_{2k}^{(\tau)}(t')dt'\Big)\cdot\exp(-iH_0\delta t)\cdot g_{[\tau,k]}. \end{aligned} \tag{3.18}$$

In other words, to obtain the total time evolution, the quantum gate U_k, $k = 1,\ldots,n_g$, in the τ-th iteration of the ideal quantum algorithm U_{QA} is replaced by the gate

$$G_k^{(\tau)} = g^\dagger_{[\tau,k]}\exp(-iH^\tau_{kl})g_{[\tau,k]}\cdot U_k\cdot g^\dagger_{[\tau,k]}\exp(-iH^\tau_{kr})g_{[\tau,k]}, \tag{3.19}$$

where in lowest order AHT the average Hamiltonians H^τ_{kl} and H^τ_{kr} are given by

$$H^\tau_{kl} = \int_0^{\tau_p} V_{2k+1}^{(\tau)\dagger}(t')H_0V_{2k+1}^{(\tau)}(t')dt' + H_0\delta t \tag{3.20a}$$

$$\text{and } H^\tau_{kr} = \int_0^{\tau_g} V_{2k}^{(\tau)\dagger}(t')H_0V_{2k}^{(\tau)}(t')dt' + H_0\delta t, \tag{3.20b}$$

respectively. If the decoupling pulses and the quantum gates are applied in the bang-bang limit ($\tau_p \to 0$, $\tau_g \to 0$), we obtain the simpler and exact expressions $H_{kl}^{\tau} = H_{kr}^{\tau} = H_0 \Delta t/2$.

3.2.2 Expansion of the Entanglement Fidelity

In the following we are mainly interested in the entanglement fidelity comparing a unitary operation U and its slightly perturbed version U_δ. Thus the relevant quantum operation $\mathcal{E}$ involves a single unitary Kraus operator K which is given by $K = U^\dagger \cdot U_\delta$. On the basis of (2.24) in the case of high dimensional quantum systems the average fidelity is approximately given by the entanglement fidelity (2.25)

$$F_e(\mathcal{E}) = \left| \frac{1}{d} \operatorname{tr}\big(U^\dagger U_\delta\big) \right|^2 \tag{3.21}$$

which is determined by the absolute square of a fidelity amplitude

$$A_e = \frac{1}{d} \operatorname{tr}\big(U^\dagger U_\delta\big). \tag{3.22}$$

In this subsection a perturbative short-time approximation of the fidelity amplitude is derived, which will be used at several occasions in the current and the following section. Let us consider n_a iterations of a quantum algorithm given by the ideal unitary transformation $U_{QA} = U_{n_g} \cdots U_3 \cdot U_2 \cdot U_1$, i. e. we set $U^\dagger = U_{QA}^{-n_a}$ in (3.22). We make the general assumption that the ideal time evolution is perturbed, where the j-th quantum gate of the τ-th iteration of U_{QA} is replaced by the perturbed unitary quantum gate

$$U_j \mapsto \exp(-i\delta H_{jl}^{\tau}) U_j \exp(-i\delta H_{jr}^{\tau}). \tag{3.23}$$

The index τ in (3.23) takes into account that perturbations may be different in successive iterations of the unitary transformation U_{QA}.

Lemma 3.2.1. *A second order expansion of the fidelity amplitude A_e (3.22) after n_a iterations of the perturbed quantum algorithm with respect to δH_{jl}^{τ} and δH_{jr}^{τ} is given by*

$$\begin{aligned}
A_e(n_a) = 1 &- \sum_{p=l,r} \sum_{\tau=1}^{n_a} \sum_{j=1}^{n_g} \frac{1}{d} \Big[i \operatorname{tr}\big(\delta H_{jp}^{\tau}\big) + \frac{1}{2} \operatorname{tr}\big((\delta H_{jp}^{\tau})^2\big) \Big] \\
&- \sum_{\tau=1}^{n_a} \sum_{j=2}^{n_g} \sum_{k=1}^{j-1} \frac{1}{d} \Big[\operatorname{tr}\big(\delta H_{jl}^{\tau}(j) \delta H_{kl}^{\tau}(k)\big) + \operatorname{tr}\big(\delta H_{jl}^{\tau}(j) \delta H_{kr}^{\tau}(k-1)\big) \\
&\qquad + \operatorname{tr}\big(\delta H_{jr}^{\tau}(j-1) \delta H_{kl}^{\tau}(k)\big) + \operatorname{tr}\big(\delta H_{jr}^{\tau}(j-1) \delta H_{kr}^{\tau}(k-1)\big) \Big] \\
&- \sum_{\tau=1}^{n_a} \sum_{j=1}^{n_g} \frac{1}{d} \operatorname{tr}\big(U_j^\dagger \delta H_{jl}^{\tau} U_j \delta H_{jr}^{\tau}\big) \\
&- \sum_{\tau_1=2}^{n_a} \sum_{\tau_2=1}^{\tau_1-1} \sum_{j,k=1}^{n_g} \frac{1}{d} \Big[\operatorname{tr}\big(U^{\tau_2-\tau_1} \delta H_{jl}^{\tau_1}(j) U^{\tau_1-\tau_2} \delta H_{kl}^{\tau_2}(k)\big) \\
&\qquad + \operatorname{tr}\big(U^{\tau_2-\tau_1} \delta H_{jl}^{\tau_1}(j) U^{\tau_1-\tau_2} \delta H_{kr}^{\tau_2}(k-1)\big) + \operatorname{tr}\big(U^{\tau_2-\tau_1} \delta H_{jr}^{\tau_1}(j-1) U^{\tau_1-\tau_2} \delta H_{kl}^{\tau_2}(k)\big) \\
&\qquad + \operatorname{tr}\big(U^{\tau_2-\tau_1} \delta H_{jr}^{\tau_1}(j-1) U^{\tau_1-\tau_2} \delta H_{kr}^{\tau_2}(k-1)\big) \Big] + \mathcal{O}\big((\delta H)^3\big),
\end{aligned} \tag{3.24}$$

with the abbreviation

$$\delta H_{jp}^{\tau}(i) = U_1^\dagger U_2^\dagger \ldots U_i^\dagger \cdot \delta H_{jp}^{\tau} \cdot U_i \ldots U_2 U_1 \equiv U_{1\ldots i}^\dagger \, \delta H_{jp}^{\tau} \, U_{i\ldots 1}. \tag{3.25}$$

The terms linear in the perturbing Hamiltonians δH_{jp}^{τ} vanish if all Hamiltonians involved are traceless.

Proof. To obtain the expansion, all terms of the form $\exp(-i\delta H_{jl}^{\tau})$ and $\exp(-i\delta H_{jr}^{\tau})$ are expanded as $\exp(-i\delta H_{jr}^{\tau}) = \mathcal{I} - i\delta H_{jr}^{\tau} - \frac{1}{2}\big(\delta H_{jr}^{\tau}\big)^2 + \ldots$. □

Remark. Note that all the terms of (3.24) involving $\mathrm{tr}\big(\cdot\big)$ terms are real valued so that up to second order the fidelity $F_e(n_a) = |A_e(n_a)|^2$ is simply obtained by multiplying all these terms of $A_e(n_a)$ with a factor of magnitude two.

3.2.3 Fidelity Decay of Unprotected Computations

Before we are going to derive a formula for the entanglement fidelity of a quantum computation in the presence of static imperfections which is protected by the PAREC method, we have to examine the corresponding fidelity decay of an unprotected computation. Typically, the fundamental unitary transformation U_{QA} constituting a quantum algorithm can be decomposed into a sequence of n_g elementary one- and two-qudit quantum gates, i.e.

$$U_{QA} = U_{n_g} \cdot \cdots \cdot U_3 \cdot U_2 \cdot U_1. \tag{3.26}$$

Let us assume in our subsequent discussion that the quantum algorithm under consideration involves n_a iterations of such a fundamental unitary transformation U_{QA}. Such quantum algorithms appear in the context of search algorithms, for example [Gro97]. Furthermore, let us focus our attention on the case of static imperfection in which the perturbing influence on such a quantum algorithm arises from a fixed and time-independent Hamiltonian coupling H_0 between the qudits constituting the quantum information processor. Without loss in generality, H_0 is taken to be traceless throughout the remaining section. We assume that an elementary quantum gate U_g is generated by switching on a possibly time-dependent gate Hamiltonian H_g for a time τ_g, i. e. we have $U_g \equiv U_g(\tau_g)$ with

$$U_g(t) = \mathcal{T}\exp\big(-i\int_0^t H_g(t')dt'\big) \tag{3.27}$$

for $t \in [0, \tau_g]$. Instead, because of the imperfections, after the time τ_g we obtain the perturbed evolution

$$U_g' = \mathcal{T}\exp\big(-i\int_0^{\tau_g}(H_g(t') + H_0)dt'\big) = U_g \cdot \mathcal{T}\exp\big(-i\int_0^{\tau_g} U_g^{\dagger}(t')H_0U_g(t')dt'\big) \tag{3.28}$$

Let us assume in addition, that subsequent quantum gates are performed after time intervals of duration $\Delta t \geq \tau_g$, i. e. in between subsequent gates there is also a period $\Delta t - \tau_g$ of free evolution during which the inter-qudit couplings perturb the quantum algorithm. Hence, in order to describe the perturbed quantum algorithm, we replace each elementary quantum gate U_j in (3.26) by

$$U_j \mapsto U_j \cdot \mathcal{T}\exp\big(-i\int_0^{\tau_g} U_j^{\dagger}(t')H_0U_j(t')dt'\big) \cdot \exp\big(-iH_0(\Delta t - \tau_g)\big), \tag{3.29}$$

$$\equiv U_j \cdot \exp\big(-i\delta H_j\big), \tag{3.30}$$

where (in lowest order AHT) the Hamiltonian δH_j is given by

$$\delta H_j = \overline{H}^{(0)}\Delta t = \int_0^{\tau_g} U_j^{\dagger}(t')H_0U_j(t')dt' + H_0 \cdot (\Delta t - \tau_g). \tag{3.31}$$

The total time T taken by the n_a iterations of the quantum algorithm U_{QA} is $T = n_a \cdot n_g\Delta t$. Equation (3.30) allows us to use the second order expansion of the fidelity amplitude which was derived in the

preceding subsection: By setting $\delta H_{jl}^{\tau} = 0$ and $\delta H_{jr}^{\tau} = \delta H_j$, equation (3.24) reduces to

$$F_e(n_a) = |A(n_a)|^2 = 1 - n_a \sum_{j,k=1}^{n_g} \frac{1}{d} \operatorname{tr}\big(U_{1\ldots j-1}^{\dagger} \delta H_j U_{j-1\ldots 1} \cdot U_{1\ldots k-1}^{\dagger} \delta H_k U_{k-1\ldots 1}\big)$$
$$- 2 \sum_{\tau=1}^{n_a-1} (n_a - \tau) \sum_{j,k=1}^{n_g} \frac{1}{d} \operatorname{tr}\big(U_{QA}^{-\tau} \cdot U_{1\ldots j-1}^{\dagger} \delta H_j U_{j-1\ldots 1} \cdot U_{QA}^{\tau} \cdot U_{1\ldots k-1}^{\dagger} \delta H_k U_{k-1\ldots 1}\big) + \mathcal{O}\big(H_0^3\big). \quad (3.32)$$

Here, the first term in the sum of (3.32) describes the influence of perturbations occurring in the same iteration $\tau \in \{1, \ldots, n_a\}$ and the second double sum describes their influence in different iterations.

Let us switch now to the simpler scenario of instantaneously applied gates. By letting $\tau_g \to 0$, we find that the effective perturbation $\delta H_j = H_0 \Delta t$ becomes the same for all quantum gates. In this case, the short-time behavior of the entanglement fidelity $F_e(n_a)$ has been studied in detail by Frahm et al. [FFS04]. In particular, these authors demonstrated that whenever an ideal unitary transformation of a quantum map U_{QA} can be modeled by a random matrix after n_a iterations the corresponding decay of the entanglement fidelity is given by

$$F_e^{\mathsf{QMap}}(n_a) = 1 - \frac{n_a}{t_a} - \frac{2}{d\sigma} \frac{n_a^2}{t_a} + \mathcal{O}\big(H_0^3\big), \quad (3.33)$$

where σ denotes the relative fraction of the chaotic component of the phase space of this map and t_a is defined by

$$\frac{1}{t_a} = \sum_{j,k=1}^{n_g} \frac{1}{d} \operatorname{tr}\big(U_{1\ldots j-1}^{\dagger} H_0 U_{j-1\ldots 1} \cdot U_{1\ldots k-1}^{\dagger} H_0 U_{k-1\ldots 1}\big) \Delta t^2 = \alpha \cdot n_g^2 \frac{1}{d} \operatorname{tr}(H_0^2) \Delta t^2, \quad (3.34)$$

with $\alpha \leq 1$. Furthermore, numerical studies indicate that the behavior of higher order terms is such that the fidelity decay becomes approximately exponential, i. e.

$$F_{e\ \mathrm{app}}^{\mathsf{QMap}}(n_a) = \exp\Big(-\frac{n_a}{t_a} - \frac{2}{d\sigma} \frac{n_a^2}{t_a}\Big). \quad (3.35)$$

While this formula was derived considering instantaneously applied quantum gates ($\tau_g = 0$), it should remain valid for finite $\tau_g \in [0, \Delta t]$ as well. The fidelity in the above expression has to be compared with the fidelity of a quantum memory after the time $T = n_a \cdot n_g \Delta t$, which was derived in subsection 2.1.4:

$$F_{e\ \mathrm{app}}^{\mathsf{none}}(T = n_a \cdot n_g \Delta t) = \exp\Big(-\frac{1}{d} \operatorname{tr}(H_0^2) T^2\Big) = \exp\Big(-\frac{n_a^2}{t_a}\Big) \text{ with } \alpha = 1. \quad (3.36)$$

It can be seen that the application of a quantum map slows down the quadratic fidelity decay by a factor $2/(d\sigma)$. Hence, the more chaotic the quantum map ($\sigma \to 1$), the slower is the fidelity decay. This is essentially the observation of Prosen and Žnidarič [PŽ01], who proposed to stabilize a quantum algorithm U_{QA} against static imperfections by devising more chaotic gate decompositions (see section 3.3).

3.2.4 Fidelity Decay of Protected Computations

The goal of this subsection is to derive a formula for the entanglement fidelity of a quantum computation which is perturbed by static imperfections, and protected using the PAREC method. It will be shown that the quadratic time dependence of the resulting fidelity decay (3.33) of an unprotected computation will be converted into a linear one.

As we showed in subsection 3.2.1, the total time evolution in the presence of static imperfections of n_a iterations of a quantum algorithm $U_{QA} = U_{n_g} \dots U_2 U_1$ which is stabilized using the PAREC method, is obtained by replacing the k-th quantum gate ($k = 1, \dots, n_g$) of the τ-th iteration by the gate (3.19)

$$G_k^{(\tau)} = g_{[\tau,k]}^\dagger \exp(-iH_{kl}^\tau) g_{[\tau,k]} \cdot U_k \cdot g_{[\tau,k]}^\dagger \exp(-iH_{kr}^\tau) g_{[\tau,k]}. \tag{3.37}$$

Hence, by setting

$$\delta H_{kl}^\tau = g_{[\tau,k]}^\dagger \cdot H_{kl}^\tau \cdot g_{[\tau,k]} \tag{3.38a}$$

$$\text{and } \delta H_{kr}^\tau = g_{[\tau,k]}^\dagger \cdot H_{kr}^\tau \cdot g_{[\tau,k]}, \tag{3.38b}$$

equation (3.24) yields the second order expansion of the entanglement fidelity between the total time evolution (3.17) and the ideal time evolution $U_{QA}^{n_a}$. (We neglect the first term in (3.17) which describes the time evolution of the first decoupling pulse.) We proceed by calculating the quantities $\mathbb{E}\,\delta H_{kl}^\tau$ and $\mathbb{E}\,\delta H_{kr}^\tau$, where $\mathbb{E}$ denotes the average taken over all random selections $g_{[\tau,k]}$ from the decoupling set $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$.

According to (3.20a), H_{kl}^τ depends on the random index $[\tau, k]$, because the time integral of the integrand $V_{2k+1}^{(\tau)\dagger}(t') H_0 V_{2k+1}^{(\tau)}(t')$ involves the unitary $V_{2k+1}^{(\tau)}(t')$ generating the pulse $V_{2k+1}^{(\tau)}(\tau_p) = g_{[\tau,k+1]} g_{[\tau,k]}^\dagger$. If the elements of the decoupling set $\mathcal{G}$ form a projective representation R of a group $G = \{\mathfrak{g}_j\}_{j=0}^{n_c-1}$ (i. e. if $g_j = R(\mathfrak{g}_j)$), the pulse $g_{[\tau,k+1]} g_{[\tau,k]}^\dagger$ corresponds to a random member $g_{j'}$ of the group. In the following we make this assumption and are going to use the notation $\Pi_{\mathcal{G}}(X)$, which was introduced in (3.3) as the projection of the operator X onto the commutant $\mathcal{A}'$ of the group algebra $\mathcal{A} = R(\mathbb{C}G)$. Hence, the average becomes

$$\mathbb{E}\,\delta H_{kl}^\tau = \Pi_{\mathcal{G}}\left(\frac{1}{n_c} \sum_{j'=0}^{n_c-1} \int_0^{\tau_p} g_{j'}^\dagger(t') H_0 g_{j'}(t') dt' \right) + \Pi_{\mathcal{G}}(H_0) \cdot \delta t \tag{3.39}$$

$$= \Pi_{\mathcal{G}}(H_0) \cdot (\tau_p + \delta t), \tag{3.40}$$

where $g_{j'}(t) = \mathcal{T} \exp\big(\int_0^t H_c(t') dt'\big)$ for $t \in [0, \tau_p]$ denotes the unitary generating the pulse $g_{j'} \equiv g_{j'}(\tau_p)$. The last identity is obtained analogously to the proof of theorem 2.1.2 by demanding that the control Hamiltonian $H_c(t')$ generating $g_{j'}(t)$ is within the group algebra $\mathcal{A}$ for all $t' \in [0, \tau_p]$ and for all $j \in \{0, 1, \dots, n_c - 1\}$.

In order to calculate $\mathbb{E}\,\delta H_{kr}^\tau$, we note that according to (3.20b), H_{kr}^τ depends on the random index $[\tau, k]$ because the time integral of $V_{2k}^{(\tau)\dagger}(t') H_0 V_{2k}^{(\tau)}(t')$ involves the unitary $V_{2k}^{(\tau)}(t')$ generating the twisted quantum gate $V_{2k}^{(\tau)}(\tau_g) = g_{[\tau,k]} \cdot U_k \cdot g_{[\tau,k]}^\dagger$. Let us assume now that the quantum gate $U_k = \exp\big(-iK \int_0^{\tau_g} f(t') dt'\big)$ is generated by a gate Hamiltonian K, shaped by a pulse form $f(t)$ such that $\int_0^{\tau_g} f(t) dt = 1$. The corresponding twisted gate could now be generated by the altered gate Hamiltonian $K'_{[\tau,k]} = g_{[\tau,k]} \cdot K \cdot g_{[\tau,k]}^\dagger$, i. e. $g_{[\tau,k]} \cdot U_k \cdot g_{[\tau,k]}^\dagger = \exp\big(-iK'_{[\tau,k]} \int_0^{\tau_g} f(t') dt'\big)$. Then,

$$\begin{aligned}
\mathbb{E}\,\delta H_{kr}^\tau &= \frac{1}{n_c} \sum_{j=0}^{n_c-1} g_j^\dagger \left(\int_0^{\tau_g} \exp\Big(+iK'_j \int_0^{t'} f(t'') dt''\Big) H_0 \exp\Big(-iK'_j \int_0^{t'} f(t'') dt''\Big) dt' + H_0 \delta t \right) g_j \\
&= \int_0^{\tau_g} \exp\Big(+iK \int_0^{t'} f(t'') dt''\Big) \Pi_{\mathcal{G}}(H_0) \exp\Big(-iK \int_0^{t'} f(t'') dt''\Big) dt' + \Pi_{\mathcal{G}}(H_0) \cdot \delta t \qquad (3.41) \\
&= \Pi_{\mathcal{G}}(H_0) \cdot (\tau_g + \delta t), \qquad (3.42)
\end{aligned}$$

where the last step is obtained provided that the action of $\Pi_{\mathcal{G}}(H_0)$ is trivial.

We are now going to use the results of the preceding two paragraphs on $\mathbb{E}\,\delta H_{lr}^\tau$ and $\mathbb{E}\,\delta H_{kr}^\tau$ to calculate the average $\mathbb{E}$ of the second order expansion of the entanglement fidelity given by equation (3.24). For

a traceless Hamiltonian H_0 a suitable decoupling scheme $\mathcal{G}$ leads to $\Pi_{\mathcal{G}}(H_0) = 0$ and we obtain the expectation value of the amplitude

$$\begin{aligned}\mathbb{E}A_e(n_a) = 1 - \frac{1}{2}\sum_{\tau=1}^{n_a}\sum_{j=1}^{n_g}\Big(\frac{1}{d}\operatorname{tr}\big((H_{jl}^{\tau})^2\big) + \frac{1}{d}\operatorname{tr}\big((H_{jr}^{\tau})^2\big)\Big) \\ - \sum_{\tau=1}^{n_a}\sum_{j=1}^{n_g}\mathbb{E}\frac{1}{d}\operatorname{tr}\big(U_j^{\dagger}\delta H_{jl}^{\tau}U_j\delta H_{jr}^{\tau}\big) + \mathcal{O}\big((\delta H)^3\big).\end{aligned} \tag{3.43}$$

In order to derive a simple expression for the fidelity, we are now going to consider the limit in which the pulses and gates are generated instantaneously ($\tau_p, \tau_g \to 0$), but we stress that the crucial step in the derivation of our fidelity formula was performed for the general case of finite pulses. In the bang-bang limit, we have $H_{kl}^{\tau} = H_{kr}^{\tau} = H_0\Delta t/2$ and (3.43) simplifies to

$$\mathbb{E}A_e(n_a) = 1 - \frac{n_a}{4}\frac{1}{d}\Big(n_g\operatorname{tr}\big(H_0^2\big) + \sum_{j=1}^{n_g}\frac{1}{n_c}\sum_{i=0}^{n_c-1}\operatorname{tr}\big(U_j^{\dagger}\,g_i^{\dagger}H_0g_i\,U_j\,g_i^{\dagger}H_0g_i\big)\Big)\Delta t^2 + \mathcal{O}\big(H_0^3\big), \tag{3.44}$$

$$\geq 1 - \frac{n_an_g}{2}\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t^2 + \mathcal{O}\big(H_0^3\big). \tag{3.45}$$

The last inequality can be obtained by recalling that $\operatorname{tr}(A^{\dagger}B)$ constitutes a Hermitian inner product for which the Cauchy-Schwarz inequality applies. We proved the following theorem:

Theorem 3.2.2. *Let a quantum computation consist of n_a iterations of a quantum algorithm U_{QA} consisting of n_g quantum gates. The entanglement fidelity between an ideal computation and a non-ideal computation protected by the* PAREC *method is (on average) given by*

$$F_e^{\mathsf{PAREC}}(n_a) \geq 1 - n_an_g\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t^2 + \mathcal{O}\big(H_0^3\big), \tag{3.46}$$

where the Hamiltonian H_0 describes the imperfections of the quantum computer, and Δt denotes the time in between subsequent quantum gates (compare with figure 3.2).

Remark. In subsection 2.3.2 we derived a short time expansion of the entanglement fidelity $F_e^{\mathsf{NRD}}(T)$ (2.92) of a quantum memory protected by NRD and argued that a good approximation (valid for all times T) is given by the exponential $F_{e\ \mathrm{app}}^{\mathsf{NRD}}(T)$ given by (2.97). Analogously, we propose that for all numbers of iterations n_a a good approximation of the PAREC fidelity is given by

$$F_{e\ \mathrm{app}}^{\mathsf{PAREC}}(n_a) = \exp\Big(-n_an_g\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t^2\Big). \tag{3.47}$$

As it turned out in this subsection, the decoupling scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ employed by PAREC has to satisfy the decoupling condition $\Pi_{\mathcal{G}}(H_0) = \lambda\cdot\mathcal{I}$, with $\lambda \in \mathbb{R}$. In addition, for a finite pulse width τ_p, the elements of the decoupling scheme should also form a group (or at least a projective representation of a group) in order to arrive at (3.40). Since the order n_c of the decoupling group does not enter in the formula for the resulting fidelity decay, it is always possible to choose $\mathcal{G}$ to be an annihilator, such as the set $\mathcal{P}_q^n$ of Pauli operators. Equation (3.46) explicitly exhibits the dependence of the entanglement fidelity decay on the number n_g of elementary quantum gates and the strictly linear dependence on the numbers of iterations of the unitary transformation U_{QA}.

Several straightforward improvements of the basic relation (3.46) are possible. For example, it is also possible to apply the random decoupling pulses not before each elementary quantum gate but less often. One random decoupling pulse between each iteration of a quantum algorithm, for example, is already enough to get rid of the terms of (3.32) quadratic in n_a. In this case (3.46) is replaced by the inequality

$$F_e^{\mathsf{PAREC}}(n_a) \geq 1 - n_an_g^2\frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t^2 + \mathcal{O}\big(H_0^3\big), \tag{3.48}$$

at the expense that the term linear in n_a has a coefficient quadratic in the number of elementary quantum gates per iteration n_g.

In order to determine the decay of the average entanglement fidelity of a quantum memory stabilized by NRD we use (3.46) and specialize to the case of n_a iterations of a quantum algorithm consisting of n_g identity gates. Denoting the total interaction time between the qudits of the quantum memory by $T = n_a n_g \Delta t$ one obtains the result

$$F_e^{\mathsf{PAREC}}(n_a) \geq 1 - n_a n_g \frac{1}{d} \operatorname{tr}\big(H_0^2\big)\Delta t^2 + \mathcal{O}\big(H_0^3\big) = 1 - \frac{1}{d} \operatorname{tr}\big(H_0^2\big)\Delta t T + \mathcal{O}\big(H_0^3\big), \tag{3.49}$$

which is identical to the average NRD fidelity (2.92),

$$F_e^{\mathsf{NRD}}(T) = 1 - \frac{1}{d} \operatorname{tr}\big(H_0^2\big)\Delta t T + \ldots, \tag{3.50}$$

of theorem 2.3.2 derived in subsection 2.3.2 by considering bang-bang control. Since we derived the PAREC fidelity by considering bounded controls (generating the decoupling pulses within a finite time interval τ_p), this fact indicates that the NRD strategy remains applicable even if only bounded controls are available.

3.2.5 Numerical Example

We close the discussion of the PAREC method with a numerical simulation. Let us consider a quantum computer with $n = 8$ qubits arranged on a linear chain, which are perturbed by Heisenberg couplings,

$$H_0 = \sum_{k_1=1}^{n-1} \sum_{k_2=k_1+1}^{n} J^{k_1,k_2} \big[X \otimes X + Y \otimes Y + Z \otimes Z\big]_{(k_1,k_2)}, \tag{3.51}$$

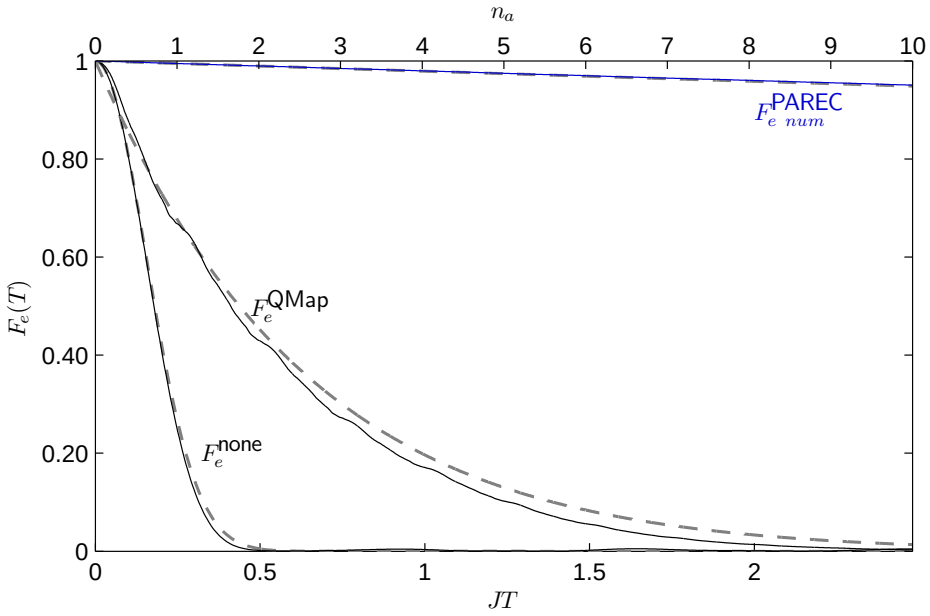
where the coupling strength between qubits k_1 and k_2 decays cubically with their separation distance, i. e. $J^{k_1,k_2} = J \cdot |k_1 - k_2|^{-3}$. Note that these are the same imperfections as assumed for the numerical simulations of the decoupling strategies in section 2.4. As a quantum algorithm we consider multiple iterations of the quantum tent map,

$$U_{QA} = \exp\Big(-\frac{i}{2} m^2 T\Big) \exp\Big(-ikV(q)\Big), \tag{3.52}$$

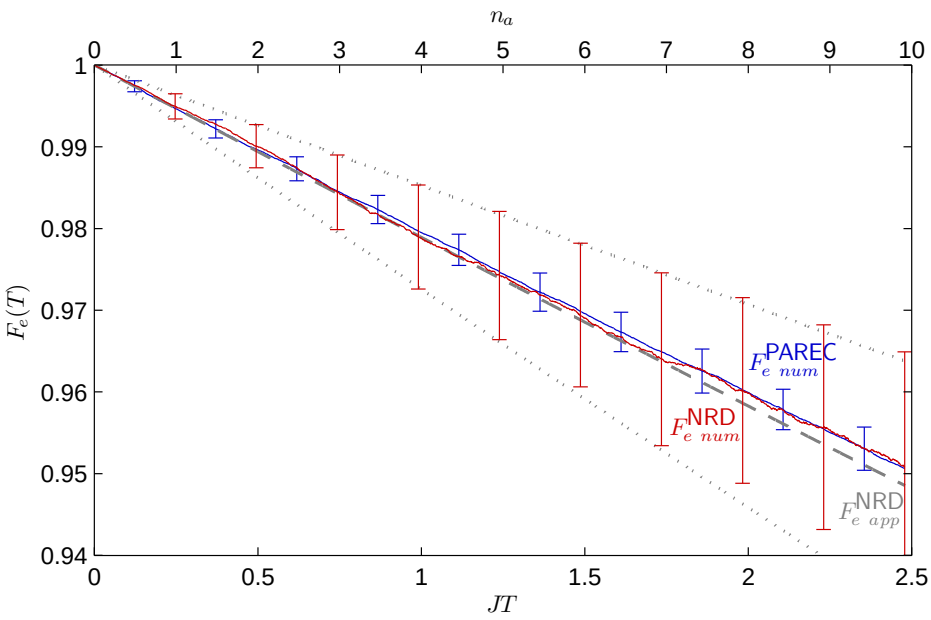
with parameters $T = 2\pi/2^n$ and $kT = 1.7$. A definition of the operators m and q and the tent-map potential V can be found in appendix B.2. It is also explained in the appendix that each iteration of the tent map can be decomposed into $n_g = \frac{9}{2}n^2 - \frac{11}{2}n + 4$ elementary one- and two-qubit quantum gates, which for $n = 8$ leads to $n_g = 248$. We assume that the gates and pulses are performed instantaneously, and that the time interval Δt in between subsequent quantum gates is given by $\Delta t = 0.001 J^{-1}$. The simulations cover $n_a = 10$ iterations. Hence, the total run time of the quantum computation is given by $T = 10 n_g \Delta t = 2.48 J^{-1}$. The results of our simulations are presented in figure 3.3.

In figure 3.3a, we compare the fidelity F_e^{QMap} of the unprotected quantum computation with the corresponding fidelity F_e^{none} of an unprotected quantum memory. The corresponding estimations $F_{e\ \mathrm{app}}^{\mathsf{QMap}}$ given by (3.35) with $\alpha = 0.294$ and $F_{e\ \mathrm{app}}^{\mathsf{none}}$ given by (2.27) are also shown (*dashed lines*). It can be seen that the quantum computation itself leads to a slow down of the fidelity decay. If the computation is stabilized using the PAREC method, the resulting fidelity $F_{e\ \mathrm{num}}^{\mathsf{PAREC}}$ (*blue*) is significantly improved and in good agreement with the predicted fidelity $F_{e\ \mathrm{app}}^{\mathsf{PAREC}}$ of equation (3.47) (*dashed line*). (The index num indicates the fact that the fidelity is obtained numerically by averaging over a subset of 70 random pulse realizations.)

Figure 3.3b shows an enlarged part of the high fidelity region. In addition to $F_{e\ \mathrm{num}}^{\mathsf{PAREC}}$ (*blue*), the fidelity $F_{e\ \mathrm{num}}^{\mathsf{NRD}}$ (*red*) of a quantum memory protected by the naive random decoupling strategy (NRD)



(a) Entanglement fidelity of a $n = 8$ qubit quantum computation.



(b) Enlarged part of figure 3.3a.

Figure 3.3: Entanglement fidelity of a $n = 8$ qubit quantum computation perturbed by the imperfections given in (3.51). Each of the $n_a = 10$ iterations of the quantum algorithm consists of $n_g = 248$ quantum gates. The time in between subsequent gates is given by $\Delta t = 0.001 J^{-1}$.
(a) The fidelity F_e^{QMap} of an unprotected computation, the fidelity F_e^{none} of a quantum memory, the fidelity $F_{e\ \mathrm{num}}^{\mathsf{PAREC}}$ (*blue*) of the stabilized computation, and the corresponding estimations $F_{e\ \mathrm{app}}^{\mathsf{QMap}}$ (3.35), $F_{e\ \mathrm{app}}^{\mathsf{none}}$ (2.27), and $F_{e\ \mathrm{app}}^{\mathsf{PAREC}}$ (3.47) (*dashed lines*).
(b) The fidelity $F_{e\ \mathrm{num}}^{\mathsf{PAREC}}$ (*blue*) of the stabilized quantum computation, the fidelity $F_{e\ \mathrm{num}}^{\mathsf{NRD}}$ (*red*) of a quantum memory stabilized by NRD, and its estimation $F_{e\ \mathrm{app}}^{\mathsf{NRD}}$ (2.97) (*dashed line*). In addition, the standard deviation of the PAREC and the NRD fidelity is indicated by error bars. The estimate $\sigma_{\mathsf{NRD}}^{\mathrm{app}}$ (2.124) of the NRD standard deviation is indicated by $F_{e\ \mathrm{app}}^{\mathsf{NRD}} \pm \sigma_{\mathsf{NRD}}^{\mathrm{app}}$ (*dotted lines*).

is shown together with its corresponding estimation $F_{e\ \mathrm{app}}^{\mathsf{NRD}}$ (*dashed*) given by equation (2.97). The memory protected via NRD corresponds to a trivial quantum computation (all the quantum gates are identity gates) which is protected by the PAREC method. As predicted by equations (3.49) and (3.50), all three fidelities are quite close to each other. Let us focus now on the variance of F_e^{PAREC} and F_e^{NRD}. An estimation of the latter quantity σ_{NRD}^2 was proposed in subsection 2.4.2 to be given by (2.124). This estimation is indicated by the two *dotted lines* representing $F_{e\ \mathrm{app}}^{\mathsf{NRD}} \pm \sigma_{\mathsf{NRD}}^{\mathrm{app}}$. It is in good agreement with the actual standard deviation $\sigma_{\mathsf{NRD}}^{\mathrm{num}}$ indicated by the error bars (*red*). An interesting observation is that the standard deviation $\sigma_{\mathsf{PAREC}}^{\mathrm{num}}$ of the fidelity $F_{e\ \mathrm{num}}^{\mathsf{PAREC}}$ of the stabilized computation (indicated by the *blue* error bars) is considerably smaller.

3.3 Stabilizing Computations by Increasing the Correlation Decay

The fidelity decay of an unprotected quantum computation in the presence of static imperfections depends on the decomposition of the quantum algorithm into elementary one- and two-qudit gates (subsection 3.2.3). The first non-trivial term in a short-time expansion of the fidelity is called correlation function. The larger the value of this correlation function, the faster the decay of the fidelity. Based on this observation, Prosen and Žnidarič [Pro02; PŽ01] proposed to stabilize quantum algorithms by rewriting them in such a way, that the new gate decomposition leads to an increased decay of the correlation function. For a particular type of imperfections, they demonstrated their idea by designing an alternative gate decomposition for the quantum Fourier transform [PŽ01]. An open question is how to find good gate decompositions for general algorithms and general imperfections. In this section, we are going to demonstrate that the PAREC method of the preceding section provides a solution to this question: By viewing the random decoupling pulses as additional quantum gates, PAREC translates an arbitrary quantum algorithm consisting of n_g quantum gates into one containing twice as much gates. This new gate decomposition leads (on average) to an ultimate decay of the correlation function.

We start in the first subsection with a summary of the main results of [PŽ01]. The second subsection explains how the PAREC method wipes out the correlations. As in subsection 3.2.3 we consider n_a iterations of quantum algorithm $U_{QA} = U_{n_g} \dots U_2 U_1$. To keep things as simple as possible, we assume that the gates (quantum gates and decoupling pulses) are applied instantaneously ($\tau_g, \tau_p \to 0$).

3.3.1 Fidelity and Correlation Decay

In special cases in which an ideal unitary transformation U_{QA} is not decomposed into elementary gates we may simplify (3.32) by taking $n_g = 1$ thus obtaining the fidelity decay

$$F_e(n_a) = 1 - \sum_{\tau=-(n_a-1)}^{n_a-1} (n_a - |\tau|) \frac{1}{d} \operatorname{tr}\big(U_{QA}^{-\tau} H_0 U_{QA}^{\tau} H_0\big) \Delta t^2 + \mathcal{O}\big(H_0^3\big). \tag{3.53}$$

This expression has been studied previously by Prosen [Pro02]. It indicates that the faster the decay of the correlation function $\operatorname{tr}\big(U_{QA}^{-\tau} H_0 U_{QA}^{\tau} H_0\big)$ the slower the decay of the fidelity. According to an original proposal by Prosen and Žnidarič [PŽ01] this characteristic feature of the fidelity decay can be exploited for stabilizing a quantum algorithm against static imperfections. This aspect was investigated in detail by these authors for the special case of $n_a = 1$. In this case (3.32) reduces to the simpler form

$$F_e = 1 - \sum_{j,k=1}^{n_g} \underbrace{\frac{1}{d} \operatorname{tr}\big(U_{1\dots j-1}^{\dagger} H_0 U_{j-1\dots 1} \cdot U_{1\dots k-1}^{\dagger} H_0 U_{k-1\dots 1}\big) \Delta t^2}_{C(j,k)} + \mathcal{O}\big(H_0^3\big). \tag{3.54}$$

Prosen and Žnidarič based their error suppression method on the idea to rewrite a quantum algorithm U_{QA} in such a way that for the new gate decomposition the sum over the off-diagonal elements of the

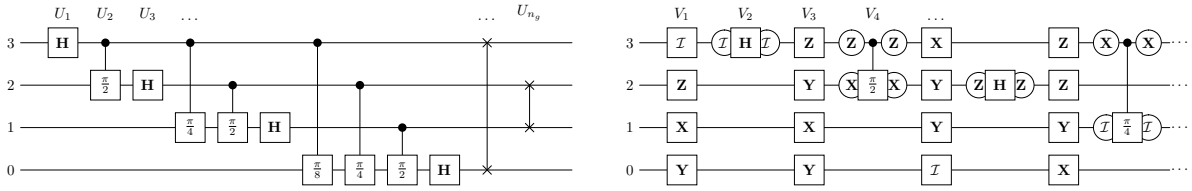


Figure 3.4: Quantum circuit of the quantum Fourier transform for $n = 4$ qubits *(left)*. The first four gates of the same circuit involving the PAREC method *(right)*.

correlation matrix $C(j,k)$ becomes smaller than for the original gate sequence (thereby using possibly even a larger number of quantum gates). They considered as an example perturbations of the form $H_0\Delta t = V\delta$ with V being represented by a d-dimensional matrix randomly chosen from the Gaussian unitary ensemble (GUE). Thus, on average the matrix elements of V fulfill the condition $\langle V_{jk}V_{lm}\rangle = \delta_{jm}\delta_{kl}/d$. With this kind of imperfections on average the correlation function becomes

$$\langle C(j,k)\rangle = \left(\left|\frac{1}{d}\,\mathrm{tr}\big(U_{j-1}\ldots U_2U_1\cdot U_1^\dagger U_2^\dagger\ldots U_{k-1}^\dagger\big)\right|^2 - \frac{1}{d^2}\right)\delta^2. \tag{3.55}$$

The $1/d^2$-term comes from the fact that according to our assumption of traceless perturbing Hamiltonians also our matrices V have to be chosen traceless. (In the case of a non-traceless perturbation V this restriction can be achieved by the replacement $V \mapsto V - \mathcal{I}\cdot\mathrm{tr}(V)/d$). It should be mentioned that this latter $1/d^2$-term was not taken into account in reference [PŽ01] so that these authors investigated the quantity $\left|\frac{1}{d}\,\mathrm{tr}\big(U_{j-1\ldots1}\cdot U_{1\ldots k-1}^\dagger\big)\right|^2\delta^2$.

In order to demonstrate their idea, Prosen and Žnidarič considered the quantum Fourier transformation (QFT) as an example. Typically, this unitary transformation U_{QA} is decomposed into $n_g = \lfloor n(n+2)/2\rfloor$ quantum gates which involve Hadamard operations, controlled-phase gates, and swap gates. (compare with the left-hand side of figure 3.4, see also subsection B.2.1 of the appendix). Instead, Prosen and Žnidarič used a different decomposition involving $n_g' = \lfloor n(2n+1)/2\rfloor$ quantum gates. In figure 3.5 the correlation matrix $\langle C(j,k)\rangle$ is depicted for both gate decompositions. Compared to the conventional gate decomposition (left) the off-diagonal elements of this correlation matrix are suppressed significantly by this new gate decomposition (middle). Diagonal values are always constant, i.e. $\langle C(j,j)\rangle/\delta^2 + 1/d^2 = 1$.

Though of interest this proposal of Prosen and Žnidarič leaves important questions unanswered. How can such an improved gate sequence be found for an arbitrary quantum algorithm ? How can this be achieved for repeated iterations of a unitary quantum map ? Is it possible to suppress all off-diagonal elements of the correlation function perfectly ? All these questions can be addressed and solved in a rather straightforward way utilizing NRD decoupling as described in the preceding section.

3.3.2 Destroying Correlations with the PAREC Method

In this subsection it is explicitly shown that the PAREC method is capable of canceling the off-diagonal terms of the correlation function $\langle C(j,k)\rangle$ (3.55) perfectly. According to equation (3.14), the PAREC method translates a quantum algorithm U_{QA} consisting of n_g quantum gates, into one containing $n_g' = 2n_g + 1$ quantum gates. Let us consider the stabilizing properties of the PAREC method with respect to static imperfections which can be characterized by traceless perturbing Hamiltonians of the form $H_0\Delta t \equiv \frac{1}{2}\cdot\big(V - \mathcal{I}\cdot\mathrm{tr}(V)/d\big)\cdot\delta$ with V chosen randomly from the Gaussian unitary ensemble (GUE). The strength of the interaction is reduced by the factor $1/2$ so that the situation is equivalent to the one depicted in figure 3.2, where Δt denotes the time interval in between 'real' subsequent quantum gates (not counting the decoupling pulses as gates). These perturbations describe physical situations

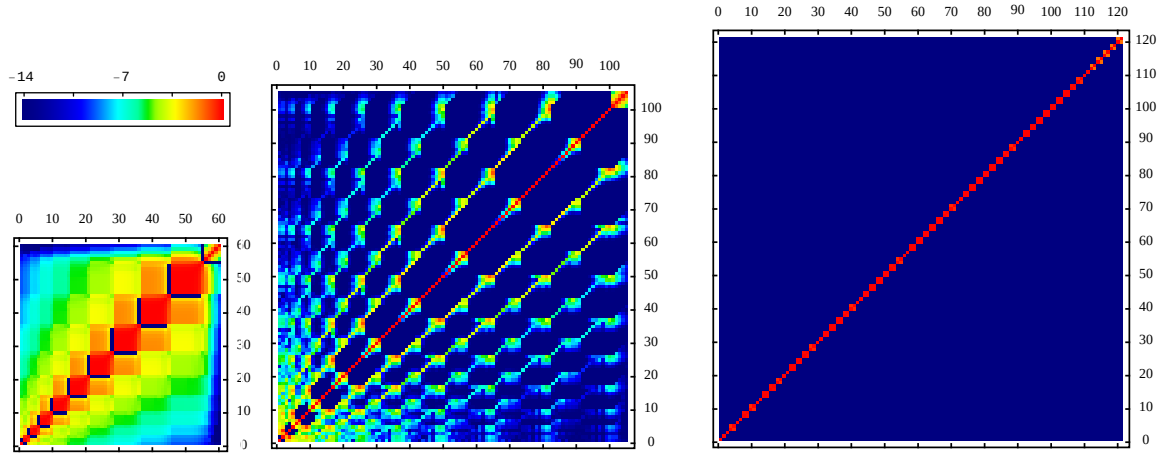


Figure 3.5: $\ln\left[\langle C(j,k)\rangle/\delta^2 + 1/d^2\right]$ for the QFT with $n = 10$ qubits using the usual gate decomposition with $n_g = 60$ gates *(left)*, the decomposition by Prosen using $n'_g = 105$ gates *(middle)* and $\ln\left[\mathbb{E}\langle C(j,k)\rangle/\delta^2 + 1/d^2\right]$ using the PAREC method with $n'_g = 2n_g + 1 = 121$ gates *(right)*.

in which in each individual realization of a quantum algorithm the inter-qudit Hamiltonian perturbing the dynamics of the qudits of the quantum information processor is time independent but random. To eliminate such GUE-governed static imperfections we have to choose an annihilator, such as the set of Pauli operators $\mathcal{P}_q^n$, as a decoupling set $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$. As a result the fidelity averaged over all possible random gates reduces to the expression

$$\mathbb{E}\langle F_e\rangle = 1 - \frac{1}{4}\sum_{j,k=1}^{n'_g} \mathbb{E}\langle C(j,k)\rangle + \mathcal{O}\big(H_0^3\big), \tag{3.56}$$

where the factor $1/4$ is a consequence of the reduced interaction strength $\delta/2$. In view of the statistical independence of subsequent Pauli operations almost all off-diagonal terms of the correlation function vanish, i.e.

$$\mathbb{E}\langle C(j,k)\rangle = \delta^2 \cdot \begin{cases} 1 - \frac{1}{d^2} & ,\text{if } j = k \\ \left|\frac{1}{d}\operatorname{tr} U_{(j-1)/2}\right|^2 - \frac{1}{d^2} & ,\text{if } j \text{ odd and } j = k+1 \\ \left|\frac{1}{d}\operatorname{tr} U_{(k-1)/2}\right|^2 - \frac{1}{d^2} & ,\text{if } k \text{ odd and } k = j+1 \\ 0 & ,\text{else.} \end{cases} \tag{3.57}$$

Here, it has been taken into account that for all unitary matrices U the relation

$$\mathbb{E}\left|\frac{1}{d}\operatorname{tr}\big(gU\big)\right|^2 \equiv \frac{1}{d^2}\sum_{j=0}^{n_c-1}\left|\frac{1}{d}\operatorname{tr}\big(g_jU\big)\right|^2 = \frac{1}{d^2} \tag{3.58}$$

holds since the average is performed over all unitary random Pauli gates $g_j \in \mathcal{G} = \mathcal{P}_q^n$ which are elements of an orthonormal unitary error basis. As a result the expectation value of the entanglement fidelity becomes

$$\begin{aligned}\mathbb{E}\langle F_e\rangle &= 1 - (2n_g+1)\frac{\delta^2}{4}(1-d^{-2}) - 2\frac{\delta^2}{4}\sum_{j=1}^{n_g}\Big(\left|\frac{1}{d}\operatorname{tr} U_j\right|^2 - d^{-2}\Big) + \mathcal{O}(\delta^3) \\ &\geq 1 - n_g\delta^2(1-d^{-2}) + \mathcal{O}(\delta^3).\end{aligned} \tag{3.59}$$

Alternatively this expression can also be derived by averaging (3.46) over all elements of the GUE after substituting the relevant perturbing Hamiltonian $H_0\Delta t \equiv \big(V - \mathcal{I}\cdot\mathrm{tr}(V)/d\big)\cdot\delta$ and setting $n_a = 1$. For the special case of a quantum Fourier transform (QFT) the resulting values of $\mathbb{E}\langle C(j,k)\rangle$ are shown on the right-hand side of figure 3.5. In this figure they are also compared to the corresponding values resulting from the improved QFT proposed by Prosen.

The PAREC method works not only for general algorithms, but also for general imperfections: In the general case of a traceless Hamiltonian H_0, the quantity $C(j,k)$ defined in (3.54) becomes on average (compare with (3.44))

$$\mathbb{E}C(j,k) = \Delta t^2 \cdot \begin{cases} \frac{1}{d}\,\mathrm{tr}\big(H_0^2\big) & ,\text{if } j = k \\ \mathbb{E}_i \frac{1}{d}\,\mathrm{tr}\big(U_{(j-1)/2}^\dagger\, g_i^\dagger H_0 g_i\, U_{(j-1)/2}\, g_i^\dagger H_0 g_i\big) & ,\text{if } j \text{ odd and } j = k+1 \\ \mathbb{E}_i \frac{1}{d}\,\mathrm{tr}\big(U_{(k-1)/2}^\dagger\, g_i^\dagger H_0 g_i\, U_{(k-1)/2}\, g_i^\dagger H_0 g_i\big) & ,\text{if } k \text{ odd and } k = j+1 \\ 0 & ,\text{else,} \end{cases} \tag{3.60}$$

for any decoupling scheme $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$ satisfying the standard decoupling condition

$$\frac{1}{n_c}\sum_{j=0}^{n_c-1} g_j^\dagger H_0 g_j = 0. \tag{3.61}$$

3.4 Stabilizing Computations using Dynamically Corrected Gates

The PAREC method of section 3.2 combines quantum computation with the naive random decoupling (NRD) strategy of subsection 2.3.2. Unfortunately, the suppression potential of NRD is rather low. If the imperfections of a quantum computer are described by a Hamiltonian λH_0, the decay of the entanglement fidelity after the time T is of the order $\mathcal{O}\big(\lambda^2\Delta t T\big)$, where Δt denotes the time interval in between the application of subsequent decoupling pulses. On the other hand, periodic dynamical decoupling (PDD, subsection 2.3.1) is able to achieve a decay of the order $\mathcal{O}\big(\lambda^4(\Delta t n_c T)^2\big)$. Even though the quadratic time dependence of PDD is inferior to the linear one of NRD, the fact that the imperfection strength λ enters in the fourth power is a serious advantage. In subsection 3.1.4, we discussed the dynamically corrected gate (DCG) of Khodjasteh and Viola [KV09], which combines a single PDD cycle with the generation of a quantum gate. By implementing each gate constituting a quantum algorithm as a DCG, a complete quantum computation might be stabilized against imperfections. In contrast to the PAREC method, that way the resulting fidelity decay of the stabilized algorithm would benefit from the PDD characteristics. This chapter considers the general case of decoupling pulses being generated by turning on a bounded control Hamiltonian for a time $\tau_p > 0$. Hence, in place of PDD, the Eulerian decoupling strategy ([VK03], subsection 2.1.7) has to be applied. In subsection 3.4.1, we consider a generalization of the DCG approach of subsection 3.1.4 from PDD to Eulerian decoupling. (In fact the original DCG proposal of Khodjasteh and Viola [KV09] was for Eulerian decoupling.) We compare the error suppression potential of the PAREC method and the Euler-DCG method for quantum algorithms. By embedding PDD cycles within NRD, the embedded decoupling (EMD) strategy was devised in subsection 2.3.2, which combines the advantages of both strategies in order to protect a quantum memory. Motivated by this idea, we propose to combine Euler-DCGs with the PAREC method in order to protect quantum computations in subsection 3.4.2.

3.4.1 Dynamically Corrected Gates (Euler-DCGs)

Subsection 3.1.4 dealt with a dynamically corrected gate (DCG) combining a single PDD cycle with the generation of a quantum gate U_g. Thereby, the decoupling pulses constituting the PDD cycle were assumed to be implemented instantaneously (i. e. in the bang-bang fashion), while the quantum gate

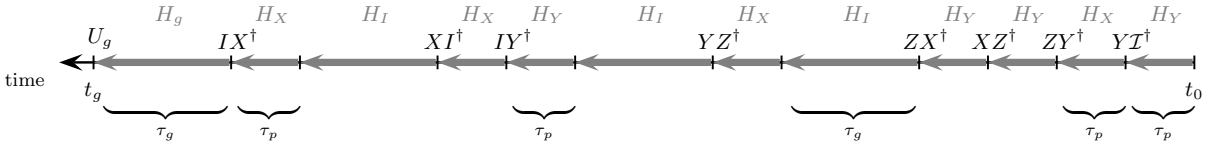


Figure 3.6: Schematic representation of an Euler-DCG based on the decoupling set $\mathcal{G} = \{\mathcal{I}, X, Y, Z\}$ and the generators $\Gamma = \{X, Y\}$. The above cycle of length $t_g = |\mathcal{G}| \cdot |\Gamma| \cdot \tau_p + |\mathcal{G}| \cdot \tau_g$ is based on the Eulerian path in the Cayley graph of $\mathcal{G}$ with respect to Γ shown in figure 3.7. H_X denotes a potentially time-dependent control Hamiltonian which generates the generator X, i. e. up to a phase we have $X = \mathcal{T} \exp\left(-i \int_0^{\tau_p} H_X(t')dt'\right)$. H_Y is defined analogously. Furthermore, H_g denotes the Hamiltonian generating the quantum gate $U_g = \mathcal{T} \exp\left(-i \int_0^{\tau_g} H_g(t')dt'\right)$ and H_I denotes the Hamiltonian mirroring the error of H_g, but implementing the identity. The gates generated by the applied Hamiltonians are denoted in the second line.

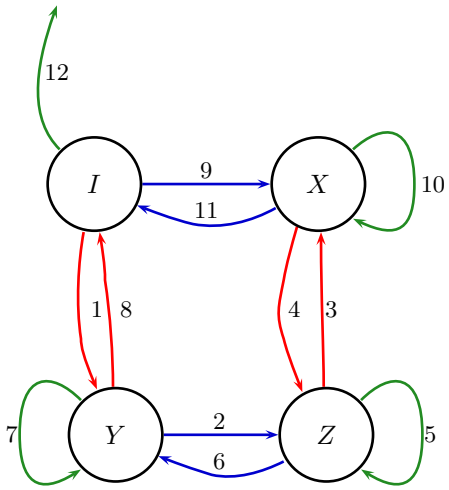


Figure 3.7: Eulerian path in the Cayley graph of $\mathcal{G} = \{\mathcal{I}, X, Y, Z\}$ with respect to the generators $\Gamma = \{X, Y\}$. The edges colored by X are depicted in blue, those colored by Y are shown in red. After a vertex is visited for the last time, a loop (depicted in green) is applied (with the exception that the final loop of the vertex assigned to the identity element is not closed).

U_g was assumed to be generated within the finite time τ_g using bounded controls: $U_g \equiv U_g(\tau_g)$ with $U_g(t) = \mathcal{T}\exp\big(-i\int_0^t H_g(t')dt'\big)$ for $t \in [0, \tau_g]$. Let us assume now that the decoupling pulses have to be generated using bounded controls as well. The standard decoupling condition demands that the action of the lowest-order average Hamiltonian of a basic decoupling cycle is trivial. If the decoupling scheme is given by the set $\mathcal{G} = \{g_j\}_{j=0}^{n_c-1}$, the decoupling condition for a PDD cycle becomes (compare with equation (2.75))

$$\Pi_{\mathcal{G}}(H_0) \equiv \frac{1}{n_c}\sum_{j=0}^{n_c-1} g_j^\dagger H_0 g_j = \mathrm{tr}(H_0)\cdot\frac{1}{d}\mathcal{I}. \tag{3.62}$$

As we know from subsection 2.1.7, in order to maintain the decoupling condition from above for finite pulses of duration τ_p, the PDD cycle of length $t_c = n_c\Delta t$ has to be replaced by an Eulerian cycle. To construct an Eulerian cycle, the elements of the decoupling scheme $\mathcal{G}$ have to form a group (strictly speaking a projective representation R of a group is sufficient). After choosing a subset of generators Γ, an Eulerian cycle is obtained by choosing an Eulerian path in the Cayley graph of $\mathcal{G}$ with respect to Γ [VK03].

We are now going to show how an Eulerian decoupling cycle has to be modified in order to generate a dynamically corrected gate. As discussed in subsection 3.1.4, a DCG generates the quantum gate U_g within the last step of a PDD cycle visiting the identity element. The error produced by generating the gate has to be mirrored during all the remaining steps of the PDD cycle. Since an Eulerian cycle visits each element $g_j \in \mathcal{G}$ exactly Γ times, this means that we have to implement the identity-gates mirroring the gate error only once, say after an element is visited for the last time. As a result, the duration of an Euler-DCG is given by $t_g = n_c|\Gamma|\tau_p + n_c\tau_g$ (compared with the scenario depicted in figure 2.3, we set $\Delta t = \tau_p$ for simplicity). Hence the zeroth-order average Hamiltonian of an Euler-DCG implementing U_g is given by

$$\begin{aligned}\overline{H}^{(0)} &= \frac{1}{t_g}\sum_{j=0}^{n_c-1} g_j^\dagger\Big(F_\Gamma(H_0)\cdot|\Gamma|\tau_p + \int_0^{\tau_g} U_g^\dagger(t')H_0U_g(t')dt'\Big)g_j \\ &= \frac{1}{|\Gamma|\tau_p+\tau_g}\Big(\Pi_{\mathcal{G}}(H_0)\cdot|\Gamma|\tau_p + \Pi_{\mathcal{G}}\big(\int_0^{\tau_g} U_g^\dagger(t')H_0U_g(t')dt'\big)\Big),\end{aligned} \tag{3.63}$$

where we used definition (2.53) and theorem 2.1.2 from subsection 2.1.7. If, in addition, the gate Hamiltonian generating $U_g(t')$ is an element of the group algebra $\mathcal{A} = R(\mathbb{C}G)$, analogous to theorem 2.1.2 we finally arrive at $\overline{H}^{(0)} = \Pi_{\mathcal{G}}(H_0)$, i. e we recover the standard decoupling condition that the action of $\Pi_{\mathcal{G}}(H_0)$ has to be trivial. (Note that in order to achieve universal quantum computation, not all the gate Hamiltonians are allowed to be in $\mathcal{A}$. Hence, finding a decoupling scheme satisfying (3.63) is not trivial. The problem might be solved by a suitable subsystem encoding, for example. Another possibility would be to employ multiple decoupling groups with different group algebras.) To illustrate the method, figure 3.6 shows an Euler-DCG corresponding to the Eulerian path in the Cayley graph of $\mathcal{G} = \{\mathcal{I}, X, Y, Z\}$ with respect to $\Gamma = \{X, Y\}$ which is depicted in figure 2.3.

Fidelity of Protected Computations

We are now going to analyze the entanglement fidelity of a quantum computation whose gates are all realized by Euler-DCGs. The analysis is performed as in subsection 3.2.3. The computation consists of n_a iterations of a quantum algorithm $U_{QA} = U_{n_g}\ldots U_2\cdot U_1$ which is decomposed into n_g elementary quantum gates. In order to describe the time evolution of the perturbed algorithm, in subsection 3.2.3 the j-th gate U_j of the ideal algorithm was replaced by the perturbed gate $U_j\cdot\exp\big(-i\delta H_j\big)$ (3.30), where in the case of instantaneously applied gates δH_j was given by $\delta H_j = H_0\Delta t$. Now each gate is realized as Euler-DCG. If the underlying Eulerian cycle is based on a decoupling scheme $\mathcal{G}$ of length $n_c = |\mathcal{G}|$

together with a set of generators $\Gamma \subset \mathcal{G}$, the corresponding gate time is now given by $t_g = n_c|\Gamma|\tau_p + n_c\tau_g$. Hence the gate error is mainly due to the first-order correction following the zeroth-order term (3.63) describing an Euler-DCG, i. e. we have $\delta H_j = \overline{H}^{(1)} = \mathcal{O}\big((H_0)^2 \cdot t_g\big)$. If the quantum algorithm U_{QA} describes a quantum map, according to equation (3.35) we expect the resulting fidelity to behave as

$$F_{e\ \mathrm{app}}^{\mathsf{Euler\text{-}DCG}}(n_a) = \exp\Big(-\frac{n_a}{t_a} - \frac{2}{d\sigma}\frac{n_a^2}{t_a}\Big), \tag{3.64}$$

where t_a is now of the order $1/t_a = \mathcal{O}\big(n_g^2 \cdot (H_0)^4 \cdot t_g^2\big)$.

3.4.2 Combining the PAREC Method with Euler-DCGs

The PAREC method can be understood as translating a n_g gate decomposition $U_{QA} = U_{n_g} \dots U_1 \cdot U_1$ of a quantum algorithm U_{QA} into a new gate decomposition containing twice as much gates. If we consider n_a iterations of the quantum algorithm, we obtain (3.14):

$$U_{QA}^{n_a} = g_{[n_a,n_g]}^{\dagger}\big(V_{2n_g}^{(n_a)} \dots V_2^{(n_a)} V_1^{(n_a)}\big) \dots \big(V_{2n_g}^{(2)} \dots V_2^{(2)} V_1^{(2)}\big)\big(V_{2n_g}^{(1)} \dots V_2^{(1)} V_1^{(1)}\big), \tag{3.65}$$

with $V_{2k}^{(\tau)} = g_{[\tau,k]} \cdot U_k \cdot g_{[\tau,k]}^{\dagger}$ and $V_{2k-1}^{(\tau)} = g_{[\tau,k]} g_{[\tau,k-1]}^{\dagger}$ for $k = 1, 2, ..., n_g$ and $\tau = 1, 2, \dots, n_a$. In subsection 3.2.4 we showed that a formula for the fidelity decay of a PAREC computation is given by (3.47),

$$F_{e\ \mathrm{app}}^{\mathsf{PAREC}}(n_a) = \exp\Big(-n_a n_g \frac{1}{d}\operatorname{tr}\big(H_0^2\big)\Delta t^2\Big), \tag{3.66}$$

where we considered the simplified scenario in which each pulse $V_{2k-1}^{(\tau)}$ and each gate $V_{2k}^{(\tau)}$ is generated instantaneously, and where $V_{2k}^{(\tau)}$ is separated from $V_{2k-1}^{(\tau)}$ by the time interval $\Delta t/2$.

In order to combine the PAREC method with the use of Euler-DCGs we simply propose to implement each of the $2n_g + 1$ gates in equation (3.65) as an Euler-DCG. As a consequence, each pulse and each gate now takes up the time $t_g = n_c|\Gamma|\tau_p + n_c\tau_g$ instead of $\Delta t/2$. In addition, the error of a pulse and/or gate is now characterized by $\overline{H}^{(1)}$ instead of H_0, where $\overline{H}^{(1)} = \mathcal{O}\big((H_0)^2 \cdot t_g\big)$ denotes the first-order correction following the zeroth-order term (3.63) in the Magnus expansion of the average Hamiltonian of the Euler-DCG. Hence we expect the fidelity of the combined stabilization method to be given by

$$F_{e\ \mathrm{app}}^{\mathsf{PAREC+Euler\text{-}DCG}}(n_a) = \exp\Big(-n_a n_g \frac{1}{d}\operatorname{tr}\big(\big(\overline{H}^{(1)}\big)^2\big) 4t_g^2\Big). \tag{3.67}$$

As it was the case for the embedded dynamical decoupling strategy (EMD) which was obtained by embedding periodic dynamical decoupling (PDD) into naive random decoupling (NRD), the combined stabilization method for computations allows us to benefit from the advantages of both underlying methods: The strong suppression $\mathcal{O}\big((H_0)^4\big)$ of the Euler-DCGs and the linear decay $\mathcal{O}\big(n_a\big)$ of the PAREC method.

4 Selective Recoupling and Randomized Decoupling

In chapter 2 we considered decoupling strategies which, with the help of instantaneously applied pulses (bang-bang pulses), suppressed the action of a system Hamiltonian describing static imperfections of a quantum memory, for instance. The performance of the fundamental decoupling strategy — called periodic dynamical decoupling (PDD) — was significantly improved by embedding it into a naive random decoupling strategy (NRD). As a result we obtained the so-called embedded decoupling strategy (EMD), which combines the advantages of both underlying strategies (strong suppression and linear fidelity decay). In an analogous fashion, by embedding the symmetrized decoupling strategy (SDD), we obtained embedded symmetric decoupling (ESDD). We are now going to show how to embed a symmetric recoupling scheme. In contrast to a decoupling scheme, a recoupling scheme leads to a non-vanishing zeroth-order average Hamiltonian describing the desired recoupling. Hence, we have to be careful not to affect this zeroth-order term when trying to eliminate residual higher order terms.

As a specific example, let us consider the recently proposed recoupling scheme for dipole-coupled nuclear spins in a crystalline solid [YLM$^+$04]. While in all previously proposed similar schemes [JK99; LCYY00; SM01; Leu02] the evolution-time overhead grows linearly with the number of spins, this particular scheme leads to an evolution-time overhead which is independent of the number of spins involved. Thus, it appears to be well suited for the stabilization of quantum information processors against unwanted inter-qubit interactions. This recoupling scheme uses particular combinations of fast broadband and slower selective radio-frequency fields to turn off all couplings except those between two particularly selected ensembles of spins. Thereby, spins within each ensemble representing a particular logical qubit are decoupled [LGYY02]. Furthermore, cross-couplings between selected ensembles are avoided by requiring that qubit couplings have to be much stronger than any other couplings within each ensemble. Unwanted couplings are suppressed up to second-order average Hamiltonian theory with the help of time-symmetric pulse sequences. Despite many advantages in this recoupling scheme the residual higher-order interactions accumulate coherently thus leading to a quadratic-in-time decay of the fidelity of any quantum state (compare with subsection 3.2.3). This restricts the achievable time scales of reliable quantum computation significantly.

In this chapter it is demonstrated that the performance of this recoupling scheme can be improved significantly by embedding it into a stochastic decoupling scheme (NRD). In contrast to a deterministic scheme which repetitively applies a certain sequence of pulses (compare with subsection 2.3.1), the corresponding stochastic scheme selects its pulses randomly (compare with subsection 2.3.2). Stochastic schemes are advantageous whenever the set these pulses are chosen from is large. In the case of an annihilator like the set of Pauli operators, for example, this set grows exponentially with the number of qubits. By a suitable embedding of the recoupling scheme into a NRD scheme based on Pauli operators, the coherent accumulation of higher-order residual interactions can be destroyed to a large extent so that the fidelity decay of any quantum state is slowed down significantly to an almost linear-in-time one. As a result, reliable quantum computation can be performed on significantly longer time scales. The results presented in this chapter have been published in [KA06].

This chapter is organized as follows: The basic ideas underlying the recently proposed deterministic recoupling scheme of reference [YLM$^+$04] are summarized briefly in section 4.1 for the sake of completeness. In section 4.2 a simple restricted embedded decoupling scheme is introduced. Though it already leads to first improvements in comparison with the deterministic selective recoupling scheme of reference [YLM$^+$04], its error suppressing properties can still be improved significantly by an additional simple symmetrization procedure. We analyze the stabilization properties of this symmetrized embed-

ded recoupling scheme for a unitary two-qubit swap gate. In section 4.3 its stabilizing properties are investigated by applying it to the iterated quantum algorithm of the quantum sawtooth map [BCMS01].

4.1 Deterministic Selective Recoupling of Qubits

In this section the basic ideas underlying the recently proposed recoupling scheme of reference [YLM+04] are summarized. In particular, the form and magnitude of the residual higher-order interaction is discussed which cannot be suppressed by the suggested pulse sequences.

Let us consider n nuclear spin-1/2 systems in a crystalline solid which are interacting with an external static magnetic field in z-direction. In the rotating wave approximation their Hamiltonian is given by [Abr61, chapter IV section II A]

$$H_0 = \underbrace{-\sum_{k=0}^{n-1} \frac{\hbar\omega_k}{2} Z_k}_{H_Z} + \underbrace{\sum_{k=0}^{n-2} \sum_{l=k+1}^{n-1} \frac{J_{kl}}{4} \big(2Z_kZ_l - X_kX_l - Y_kY_l\big)}_{H_D} \tag{4.1}$$

with the Pauli spin operators X, Y, and Z. Thereby, the Larmor frequencies ω_k of the first term characterize the interaction strengths of these spins with the external magnetic field. Using a magnetic field gradient the ω_k are adjusted in such a way that the spins can be addressed individually. The second term of the Hamiltonian (4.1) describes the dipole-dipole interaction of the nuclear spins with the coupling strength J_{kl} between spins k and l being inversely proportional to the cubic power of their distance. To keep the notation as simple as possible, we set $\hbar = 1$ for the remaining chapter.

4.1.1 Decoupling

If these nuclear spins are used as qubits of a quantum memory, for example, one has to protect them against the perturbing influence of the interaction Hamiltonian (4.1). In the framework of a deterministic decoupling scheme (chapter 2) this may be achieved by an appropriate sequence of fast electromagnetic pulses. For $\alpha \in \{X, Y, Z\}$, let us define a global $\pi/2$-pulse as

$$P_\alpha = \bigotimes_{k=0}^{n-1} \exp\big(-i\alpha_k\pi/4\big) = P_\alpha^\dagger. \tag{4.2}$$

Analogously, a global π-pulse is defined as $\alpha^{\otimes n}$. A decoupling scheme for the Zeeman term H_Z is given by the set $\{\mathcal{I}, X^{\otimes n}\}$, for instance. Hence, in order to suppress H_Z, a series of fast global $X^{\otimes n}$-pulses is applied, leaving the dipole-dipole coupling term H_D invariant. This latter term can be suppressed by the well known WHH scheme $\{\mathcal{I}, P_x, P_{\bar{y}}P_x\}$ ([WHH68], subsection 2.2.2). Using the symmetric dynamical decoupling (SDD) strategy, the WHH pulse sequence consists of four fast $\pi/2$-pulses applied at times Δt, $2\Delta t$, $4\Delta t$ and $5\Delta t$. Thus, the resulting unitary time evolution after this pulse sequence, i. e. at time $t_c = 6\Delta t$, is given by

$$\begin{aligned} U(t_c) &= \exp(-iH_D\Delta t)P_{\bar{x}} \exp(-iH_D\Delta t)P_y \exp(-iH_D 2\Delta t)P_{\bar{y}} \exp(-iH_D\Delta t)P_x \exp(-iH_D\Delta t) \\ &\equiv \exp(-i\tilde{H}_6\Delta t)\ldots\exp(-i\tilde{H}_2\Delta t)\exp(-i\tilde{H}_1\Delta t), \end{aligned} \tag{4.3}$$

with the interaction-picture (toggled) Hamiltonians $\tilde{H}_1 = \tilde{H}_6 = H_D$, $\tilde{H}_2 = \tilde{H}_5 = P_{\bar{x}}\hat{H}_D\hat{P}_x$ and $\tilde{H}_3 = \tilde{H}_4 = P_{\bar{x}}P_yH_DP_{\bar{y}}P_x$. As a consequence, in zeroth-order average Hamiltonian theory (AHT) the time-averaged Hamiltonian vanishes, i. e.

$$\overline{H}_D^{(0)} = \frac{1}{6}\sum_{j=1}^{6} \tilde{H}_j = 0. \tag{4.4}$$

Due to the time reversal symmetry of the WHH pulse sequence, i. e. $\tilde{H}(t) = \tilde{H}(t_c - t)$, in AHT all odd higher-order Hamiltonians vanish (theorem 2.1.1): $\overline{H}_D^{(2i+1)} = 0$ for $i \in \mathbb{N}_0$.

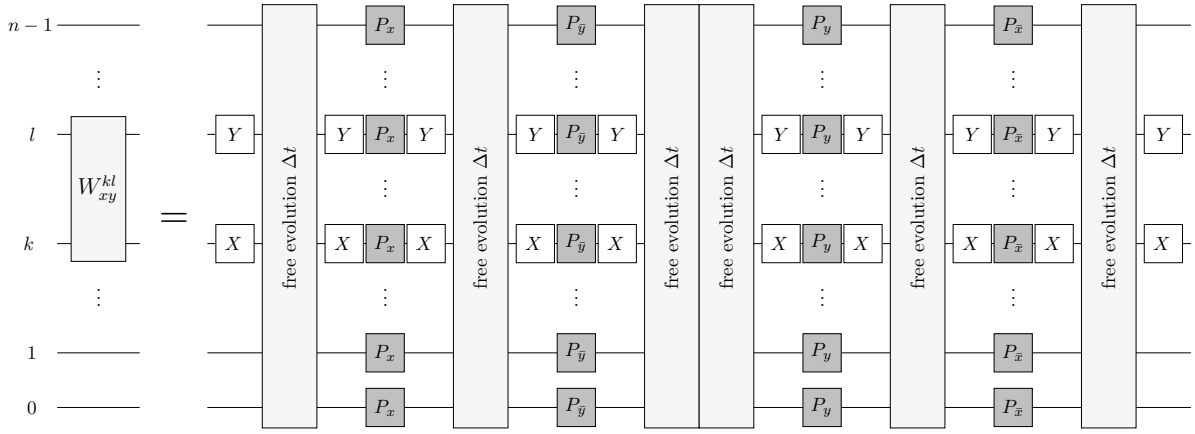


Figure 4.1: Schematic representation of the unitary W_{xy}^{kl} quantum gate acting on qubits k and l: *Free evolution* indicates time evolution according to the Hamiltonian H_D over a time interval of duration Δt.

4.1.2 Selective Recoupling

If these nuclear spins are used as qubits of a quantum information processor one also has to implement universal sets of unitary quantum gates. In particular, one needs to be able to implement two-qubit entanglement gates, such as controlled-phase gates. This can be accomplished by recoupling qubits selectively with the help of a Super-WHH pulse sequence as proposed in reference [YLM+04]. Such a Super-WHH sequence recoupling qubits k and l consists of three WHH sequences applied to the toggled Hamiltonians

$$\tilde{H}_{zz}^{kl} = Z_k Z_l H_D Z_k Z_l, \tag{4.5a}$$

$$\tilde{H}_{xy}^{kl} = X_k Y_l H_D X_k Y_l, \tag{4.5b}$$

$$\text{and } \tilde{H}_{yx}^{kl} = Y_k X_l H_D Y_k X_l, \tag{4.5c}$$

respectively. Correspondingly, there are 18 time periods of duration Δt during which the time evolution is described by the double-toggled Hamiltonians $\tilde{\tilde{H}}_1 = \tilde{H}_{zz}^{kl}$, $\tilde{\tilde{H}}_2 = \hat{P}_{\bar{x}} \tilde{H}_{zz}^{kl} \hat{P}_x$, et cetera. The appropriate WHH pulse sequence of the $\tilde{H}_{xy}^{kl}$ Hamiltonian, for example, is illustrated in figure 4.1, where *free evolution* denotes the time evolution according to the Hamiltonian H_D over a time interval of duration Δt. The quantum gates resulting from these WHH sequences are denoted by W_{xy}^{kl}, W_{zz}^{kl}, and W_{yx}^{kl}, respectively. The Super-WHH sequence is finally obtained from a combination of these latter quantum gates preceeded by the corresponding time reversed sequence (compare with the inner part of figure 4.2). As a consequence [YLM+04], this Super-WHH sequence yields the average Hamiltonian $\overline{H}_D = \overline{H}_D^{(0)} + \overline{H}_D^{(1)} + \overline{H}_D^{(2)} + \ldots$, with

$$\overline{H}_D^{(0)} = \frac{1}{t_c} \sum_{j=1}^{n_c} \tilde{\tilde{H}}_j \Delta t = J_{kl}^{(0)} \big(X_k X_l + Y_k Y_l + Z_k Z_l \big), \tag{4.6}$$

$n_c = 36$, $t_c = n_c \Delta t$, and with the renormalized zeroth-order recoupling strength $J_{kl}^{(0)} = (J_{kl}/4) \times (8/9)$. Due to the time reversal symmetry of the Super-WHH sequence, in AHT all odd-valued higher order Hamiltonians vanish, i. e. $\overline{H}_D^{(1)} = \overline{H}_D^{(3)} = 0$, et cetera. Note that in contrast to the selective decoupling schemes of subsection 2.2.3, the selective recoupling scheme presented above changes the form of the selected coupling (from $2Z_kZ_l - X_kX_l - Y_kY_l$ in (4.1) to $X_kX_l + Y_kY_l + Z_kZ_l$ in (4.6)).

With the help of the zeroth-order recoupled Hamiltonian $\overline{H}_D^{(0)}$ of equation (4.6) one can approximate

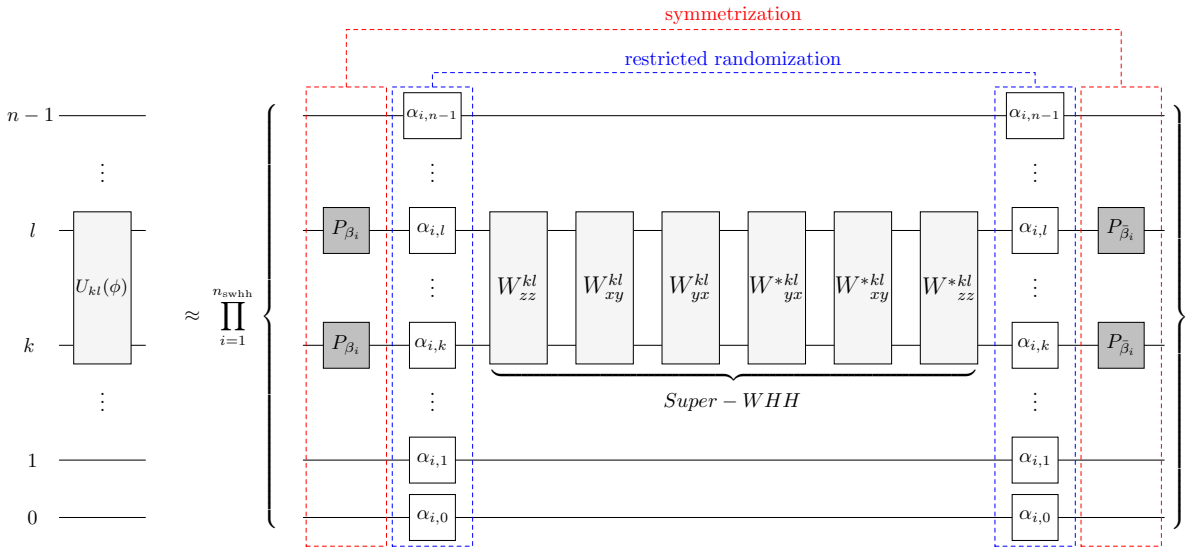


Figure 4.2: Schematic representation of the two-qubit gate $U_{kl}(\phi)$ obtained by recoupling qubits k and l according to equation (4.7): The W^* gates are obtained by reversing the order of the broadband pulses suppressing the Zeeman term. Residual second-order terms of AHT can be eliminated by the restricted randomization step accomplished by random selective π-pulses $\alpha_{i,j} \in \{\mathcal{I}, X, Y, Z\}$. Thereby $\alpha_{i,k}$ has to be equal to $\alpha_{i,l}$ to ensure that the wanted gate action is not disturbed. Still remaining terms are symmetrized by random $\pi/2$-pulses $P_{\beta_i} = \exp\bigl(-i\beta_i\pi/4\bigr)$, $\beta_i \in \{X, Y, Z\}$. Either condition (4.8) or condition (4.19) has to be fulfilled depending on whether the original Super-WHH or the symmetrized Super-WHH sequence is used.

unitary two-qubit quantum gates of the form

$$U_{kl}(\phi) = \exp\bigl(-i(X_kX_l + Y_kY_l + Z_kZ_l)\phi\bigr) \equiv \exp\bigl(-iH_g^{kl}\phi\bigr). \tag{4.7}$$

Thereby, for a particular value of the phase ϕ one has to adjust the time Δt between two successive pulses of a WHH sequence and the number of times n_{swhh} a Super-WHH sequence has to be applied according to the relation

$$J_{kl}^{(0)} \cdot n_{\mathrm{swhh}} n_c \Delta t = \phi \tag{4.8}$$

(compare with figure 4.2). However, because of the residual higher-order interactions which have not been canceled by the Super-WHH pulse sequence, this implementation of a two-qubit quantum gate is only approximate. The error resulting from these residual higher-order interactions is dominated by the second-order term of AHT which is given by (2.15c),

$$\overline{H}_D^{(2)} = -\frac{1}{6t_c} \sum_{i\geq j\geq k=1}^{n_c} \Bigl([\tilde{\tilde{H}}_i, [\tilde{\tilde{H}}_j, \tilde{\tilde{H}}_k]] + [[\tilde{\tilde{H}}_i, \tilde{\tilde{H}}_j], \tilde{\tilde{H}}_k]\Bigr)\Delta t^3 \times \begin{cases} 1/2 & \text{if } i = j \text{ or } j = k \\ 1 & \text{else} \end{cases}. \tag{4.9}$$

Therefore, the lowest-order correction to the recoupled Hamiltonian of equation (4.6) is given by

$$\begin{aligned} \overline{H}_D^{(2)} = \sum_a^{\neq k,l} \Bigl[& X_kX_l\Bigl(-322J_{al}^2J_{ak} + 446J_{ak}^2J_{al} + 3628J_{al}J_{ak}J_{kl} - 2906J_{ak}^2J_{kl} - 1370J_{al}^2J_{kl}\Bigr) + \\ & Y_kY_l\Bigl(+308J_{al}^2J_{ak} + 308J_{ak}^2J_{al} + 3208J_{al}J_{ak}J_{kl} - 2588J_{ak}^2J_{kl} - 2588J_{al}^2J_{kl}\Bigr) + \\ & Z_kZ_l\Bigl(+446J_{al}^2J_{ak} - 322J_{ak}^2J_{al} + 3580J_{al}J_{ak}J_{kl} - 1922J_{ak}^2J_{kl} - 3458J_{al}^2J_{kl}\Bigr) \\ & \Bigr]\Delta t^2/1728 + \dots . \end{aligned} \tag{4.10}$$

Thereby, only terms of the form $\alpha_k\beta_l = \alpha_k \otimes \beta_l \otimes \mathcal{I}_{\{0,1,\ldots,n-1\}\setminus\{k,l\}}$ with $\alpha, \beta \in \{X, Y, Z\}$ are indicated as all other terms are irrelevant for our subsequent discussion. As a consequence, the gate Hamiltonian resulting from recoupling qubits k and l by a Super-WHH sequence is of the form

$$\begin{aligned} H_g'^{kl} = \overline{H}_D &= \overline{H}_D^{(0)} + \overline{H}_D^{(2)} + \overline{H}_D^{(4)} + \ldots \\ &= J_{kl}^{(0)}\big(X_kX_l + Y_kY_l + Z_kZ_l\big) + \mathcal{O}\big(J(J\Delta t)^2\big). \end{aligned} \tag{4.11}$$

To estimate the resulting error affecting the unitary gate $U'_{kl}(\phi)$ generated by $H_g'^{kl}$ we study the entanglement fidelity given by (2.25),

$$F_e = \Big|\frac{1}{d}\operatorname{tr}\Big(U_{kl}^\dagger(\phi)\cdot U'_{kl}(\phi)\Big)\Big|^2, \tag{4.12}$$

comparing the action of $U'_{kl}(\phi)$ with the action of the ideal gate $U_{kl}(\phi)$ generated by H_g^{kl}. A short time expansion of F_e can be derived by using the following lemma.

Lemma 4.1.1. *Let x and y denote Hermitian operators, and let the unitaries U and U' be defined as $U = \exp(-ixt)$ and $U' = \exp\big(-i(x+y)t\big)$, respectively. Then a series expansion of $U^\dagger \cdot U'$ is given by*

$$U^\dagger \cdot U' = \mathcal{I} - iyt + \frac{1}{2}[x,y]t^2 - \frac{1}{2}y^2 + \frac{i}{6}[x,[x,y]]t^3 - \frac{i}{6}y[x,y]t^3 - \frac{i}{3}[x,y]yt^3 + \mathcal{O}(t^4). \tag{4.13}$$

By setting $x = H_g^{kl}$, $y = \overline{H}_D^{(2)} + \overline{H}_D^{(4)} + \ldots$, and the gate-time $t \equiv n_{\text{swhh}}n_c\Delta t = \phi/J_{kl}^{(0)}$ according to condition (4.8), we obtain the expression

$$F_e = \Big|\frac{1}{d}\operatorname{tr}\Big(U_{kl}^\dagger(\phi)\cdot U'_{kl}(\phi)\Big)\Big|^2 = 1 - \frac{1}{d}\operatorname{tr}\Big(\big(\overline{H}_D^{(2)}\big)^2\Big)t^2 + \mathcal{O}\big(t^4\big) \tag{4.14}$$

$$= 1 - \mathcal{O}\big((J^3\Delta t^2\cdot n_{\text{swhh}}n_c\Delta t)^2\big) = 1 - \mathcal{O}\big(\phi^6/(n_{\text{swhh}}n_c)^4\big). \tag{4.15}$$

For a fixed phase ϕ, the strength of the fidelity decay of $U'_{kl}(\phi)$ is inversely proportional to the fourth power of the number of Super-WHH iterations.

4.2 Embedded Selective Recoupling

The selective recoupling scheme of the preceding section applies the symmetric dynamical decoupling (SDD, see subsection 2.3.1) strategy in order to get a vanishing first-order term in the Magnus expansion of the average Hamiltonian describing the time evolution of a single recoupling cycle. However, in contrast to SDD the zeroth-order AHT term does not vanish and describes the desired recoupling. In this section we are going to show how the recoupling scheme can be embedded into a naive random decoupling (NRD, see subsection 2.3.2) scheme. By embedding SDD into NRD, we devised the embedded symmetric decoupling (ESDD, subsection 2.3.2) strategy combining the advantages of both underlying strategies. Now, however, we have to prevent the NRD pulses from averaging out the desired recoupling action, i. e. they should merely suppress the remaining second (and higher) order AHT term(s) and leave the zeroth-order term unaffected. As a consequence, we are not able to suppress the remaining terms entirely. Fortunately, the non-suppressible part can be cast into the form of the desired recoupling, thereby simply renormalizing the effective recoupling strength.

4.2.1 Embedding the Selective Recoupling Scheme

The residual interaction described by the Hamiltonian (4.10) can be suppressed significantly by embedding the recoupling scheme of section 4.1 into a naive random decoupling (NRD) scheme based on an annihilator as the set of Pauli operators $\mathcal{P}_2^n$. For this purpose we choose at random an n-fold tensor

product of Pauli-matrices $\alpha_{i,0} \otimes \alpha_{i,1} \otimes \cdots \otimes \alpha_{i,n-1}$, with $\alpha_{i,j} \in \mathcal{P}_2 = \{\mathcal{I}, X, Y, Z\}$ for $j = 0, 1, \ldots, n-1$, and apply it before and after the i-th Super-WHH sequence. This way each deterministic Super-WHH sequence is embedded within two statistically independent random Pauli operations. In contrast to a usual dynamical decoupling scenario ([KA05], chapter 2) in our case we have to choose the Pauli-matrices in such a way that they leave the ideally recoupled gate Hamiltonian H_g^{kl} of equation (4.7) invariant. This can be achieved by imposing the restriction that the randomly chosen statistically independent Pauli spin operators have to be identical for qubits k and l for each Super-WHH sequence, i. e. $\alpha_{i,k} = \alpha_{i,l}$ for all $i \in \{1, \ldots, n_{\mathrm{swhh}}\}$. This restriction assures that terms of the form $\alpha_k \alpha_l$ in H_g^{kl} remain invariant (compare with figure 4.2). Since $\overline{H}_D^{(2)}$ contains no terms of the form $\alpha_k \beta_l$ with $\alpha \neq \beta$ (compare with equation (4.10)) the Pauli-matrices for qubits k and l can always be omitted, i. e. chosen to be the identity, $\alpha_{i,k} = \alpha_{i,l} = \mathcal{I}$.

The only terms of the Hamiltonian $\overline{H}_D^{(2)}$ which cannot be eliminated by this constrained randomization method are the ones containing terms of the form $\alpha_k \alpha_l$ ($\alpha \in \{X, Y, Z\}$) which are shown in equation (4.10). However, by an additional symmetrization these terms can be made rotationally invariant so that they can be cast into the form of equation (4.6). Thus, for a given value of ϕ these terms lead to a renormalization of the values of the required gate parameters Δt and n_{swhh}. This rotational symmetrization can be achieved by selective $\pi/2$-pulses as defined in equation (4.2). For this purpose one chooses one of the three unitary transformations $\{P_{\beta_i,k} P_{\beta_i,l}\}_{\beta_i \in \{X,Y,Z\}}$ acting on qubits k and l at random and applies it before and the corresponding inverse transformation after the i-th Super-WHH sequence (compare with Fig. 4.2). This way the coefficients of the $\alpha_k \alpha_l$-terms are permuted in the relevant toggled Hamiltonians. As a consequence one obtains the statistically and rotationally averaged second-order contribution

$$\mathbb{E}\overline{H}_D^{(2)} = \big(X_k X_l + Y_k Y_l + Z_k Z_l\big) J_{kl}^{(2)} \Delta t^2 \tag{4.16}$$

with

$$J_{kl}^{(2)} = \sum_a^{\neq k,l} \Big(\frac{1}{12} (J_{al}^2 J_{ak} + J_{ak}^2 J_{al}) + \frac{217}{108} J_{al} J_{ak} J_{kl} - \frac{103}{72} (J_{ak}^2 J_{kl} + J_{al}^2 J_{kl}) \Big). \tag{4.17}$$

Here, $\mathbb{E}$ denotes the average taken over the $\alpha_{i,j} \in \mathcal{P}_2 = \{\mathcal{I}, X, Y, Z\}$ and the $\beta_i \in \{X, Y, Z\}$. By this combined randomization and symmetrization method the improved recoupled Hamiltonian

$$\begin{aligned} H_g''^{kl} &= \mathbb{E}\overline{H}_D^{(0)} + \mathbb{E}\overline{H}_D^{(2)} + \mathbb{E}\overline{H}_D^{(4)} + \ldots \\ &= \Big(J_{kl}^{(0)} + J_{kl}^{(2)} \Delta t^2 + \mathcal{O}\big(J(J\Delta t)^4\big) \Big) \times \big(X_k X_l + Y_k Y_l + Z_k Z_l \big) \end{aligned} \tag{4.18}$$

is obtained. In contrast to $H_g'^{kl}$ given by equation (4.11), now the effective recoupling strength is renormalized and the residual error is suppressed up to fourth order in the small coupling parameter $J\Delta t \ll 1$. Thus, in order to implement a $U_{kl}(\phi)$-gate, for example, we now have to choose the renormalized characteristic parameter $\Delta t'$ in such a way that the condition

$$\big(J_{kl}^{(0)} + J_{kl}^{(2)} \Delta t'^2 \big) n_{\mathrm{swhh}} n_c \Delta t' = \phi \tag{4.19}$$

is fulfilled. As a result, in general the required time of free evolution $\Delta t'$ depends on the chosen qubit pair (k, l).

4.2.2 Performance of a Recoupled Quantum Gate

In this section the stabilizing properties of selective recoupling by the embedded symmetric dynamical decoupling (ESDD) method of the preceding section is investigated for a unitary phase gate as described by equation (4.7). As shown in equation (4.15) the fidelity of a unitary phase gate $U_{kl}(\phi)$ which is realized by recoupling qubits k and l with the help of the average Hamiltonian of equation (4.11) (i. e.

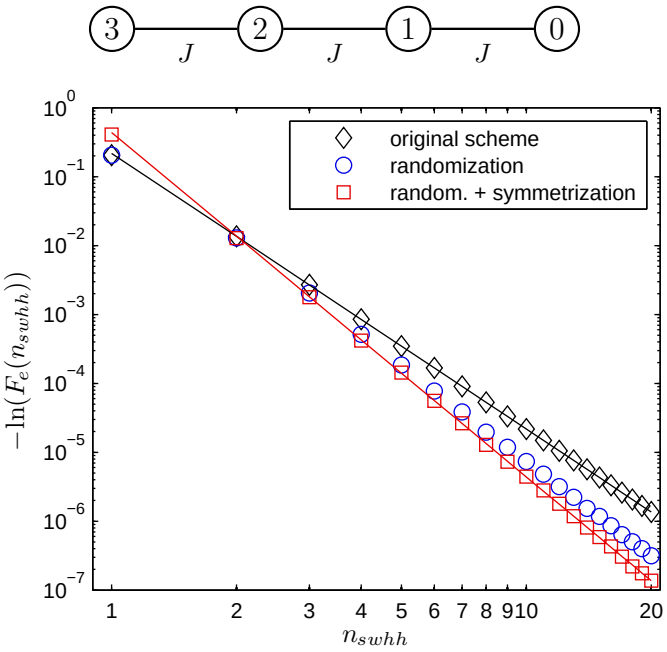


Figure 4.3: The entanglement fidelity (*bottom*) of the $U_{12}(\pi/4)$-gate on a linear four-qubit chain (*top*) as a function of the number of repetitions n_{swhh} of the Super-WHH sequence: the original Super-WHH sequence (diamonds), the unsymmetrized embedded scheme (circles), and the complete embedded scheme with adapted pulse interval $\Delta t'$ according to (4.19) (squares). The solid lines represent the fitting functions $\exp(-c/n_{\text{swhh}}^4)$ and $\exp(-c_{\text{ESDD}}/n_{\text{swhh}}^5)$ with $c = 0.22$ and $c_{\text{ESDD}} = 0.44$.

by applying the SDD strategy) deviates from unity by terms of the order of $\mathcal{O}\left(\phi^6/(n_{\text{swhh}}n_c)^4\right)$. Here, n_{swhh} denotes the number of required iterations of the Super-WHH sequence which is related to the time Δt of the intermediate free evolution and the phase ϕ as determined by relation (4.8).

In order to estimate the improvement achievable with the help of the embedded recoupling scheme, let us recall our result for the non-embedded original scheme (4.15):

$$F_e = \left|\frac{1}{d}\,\mathrm{tr}\Big(U_{kl}^{\dagger}(\phi)\cdot U'_{kl}(\phi)\Big)\right|^2 = 1 - \frac{1}{d}\,\mathrm{tr}\Big(\big(\overline{H}_D^{(2)}\big)^2\Big)t^2 + \cdots = 1 - \mathcal{O}\big(\phi^6/(n_{\text{swhh}}n_c)^4\big). \tag{4.20}$$

This expression is of the same form as the short time expansion of the fidelity of a quantum memory protected using the SDD strategy (2.84). As we found out in subsection 2.3.2, the corresponding ESDD fidelity (2.110) is obtained by replacing one power of the total time t by the time of a basic cycle. Applying these results to the recoupling case, this means that we have to replace one power of the total time $t = n_{\text{swhh}}\cdot n_c\Delta t$ in the preceding equation by the time $n_c\Delta t$ taken by a single Super-WHH cycle. As a result we obtain the estimation

$$\begin{aligned} F_e = \mathbb{E}\left|\frac{1}{d}\,\mathrm{tr}\Big(U_{kl}^{\dagger}(\phi)\cdot U'_{kl}(\phi)\Big)\right|^2 &= 1 - \frac{1}{d}\,\mathrm{tr}\Big(\big(\overline{H}_D^{(2)}\big)^2\Big)n_{\text{swhh}}n_c\Delta t\cdot n_c\Delta t + \ldots \\ &= 1 - \mathcal{O}\big(\phi^6/(n_{\text{swhh}}^5 n_c^4)\big), \end{aligned} \tag{4.21}$$

where we used condition (4.8) to approximate the relevant condition (4.19).

In figure 4.3 (*bottom*) the entanglement fidelity F_e of a unitary $U_{12}(\pi/4)$-gate and its dependence on the number of performed Super-WHH sequences n_{swhh} is depicted. In these numerical simulations this unitary quantum gate is realized by recoupling of the two central qubits 1 and 2 of a linear four-qubit chain (containing the qubits $0, 1, 2$, and 3). The coupling strength is assumed to be constant for

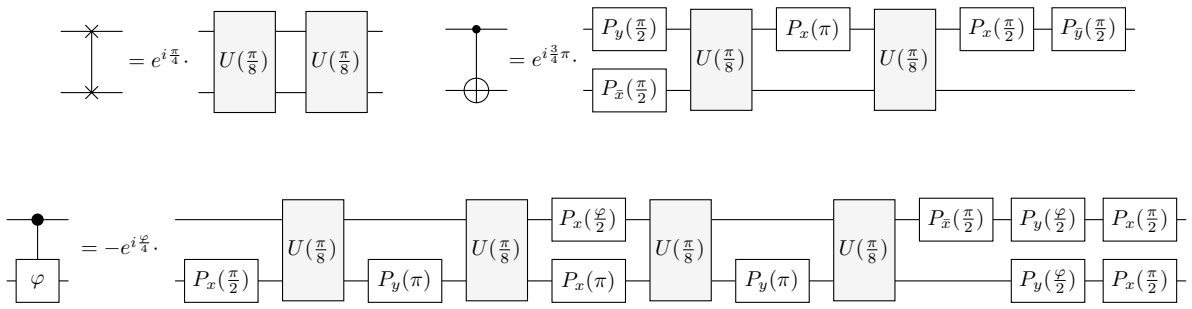


Figure 4.4: Quantum circuits implementing the SWAP, the CNOT, and the controlled-phase gate CP(φ) by using a $U(\pi/8)$ gate generated by Super-WHH recoupling. The single qubit gate $P_\alpha(\varphi)$ is defined as $P_\alpha(\varphi) = \exp\bigl(-i\alpha\varphi/2\bigr) = P_{\bar{\alpha}}^\dagger(\varphi)$ for $\alpha \in \{X, Y, Z\}$.

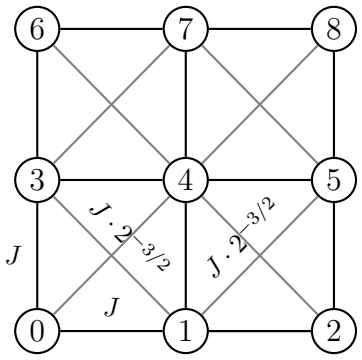
adjacent qubits and to be vanishing between all other qubits (compare with Fig. 4.3 (*top*)). Apart from an irrelevant global phase this unitary $U(\pi/4)$-gate is nothing but a SWAP-gate (compare with figure 4.4). The statistical averaging was performed over 100 runs with statistically independent realizations of the random pulses involved. Figure 4.3 (*bottom*) demonstrates that the fidelity (diamonds) resulting from non-embedded original Super-WHH pulse sequences can be fitted well by a function of the form $\exp(-c/n_{\text{swhh}}^4)$ with $c \approx 0.22$. This is consistent with the simple estimate (4.15). Using a recoupling scheme based on the embedded procedure discussed in section 4.2 while choosing Δt according to condition (4.19), we notice that the resulting fidelity (squares) is fitted well by a function of the form $\exp(-c_{\text{ESDD}}/n_{\text{swhh}}^5)$ with $c_{\text{ESDD}} \approx 0.44$, which confirms our estimate (4.21). If symmetrization is omitted an intermediate behavior is obtained (circles).

4.3 Numerical Simulation of a Quantum Algorithm

In this section the question is explored how much can be gained by stabilizing an iterative quantum algorithm by the embedded recoupling scheme of section 4.2. Using the embedded recoupling scheme to implement a quantum algorithm is reminiscent of the PAREC-method of section 3.2 in the sense that each period of imperfect evolution is suppressed using naive random decoupling (NRD). Hence, in addition to the improvement which is achieved for a single recoupled quantum gate, we expect the fidelity decay of a quantum algorithm using the embedded recoupling scheme to be linear in time instead of quadratic in time.

4.3.1 Quantum Computation with a Recoupled Quantum Gate

For purposes of quantum computation one needs to know how to perform two-qubit entanglement gates, such as the controlled-not gate (CNOT-gate) or the controlled-phase gate (CP(φ)-gate), on the basis of the recoupled Hamiltonian H_g^{kl} (4.7). Definitely, such quantum gates can be performed only between qubits k and l which are coupled, i. e. for which $J_{kl} \neq 0$. Therfore, in order to be able to entangle any two qubits of a quantum computer it is necessary to swap qubit pairs with vanishing coupling constants to neighboring positions. Fortunately, such a unitary swapping gate can be realized easily by the unitary phase gate of equation (4.7) because $\text{SWAP}_{kl} = U_{kl}(\pi/4)$. Throughout the rest of this section we will use the quantum phase gate $U_{kl}(\pi/8)$ as a basic building block for all two-qubit quantum gates. Thus, the quantum SWAP_{kl}-gate consists of the repeated application of two such gates. For the realization of other two-qubit quantum gates repeated applications of this $U_{kl}(\pi/8)$-gate in combination with single-qubit gates are required. In figure 4.4 basic gate decompositions are depicted for the CNOT-gate, the CP(φ), and for the SWAP-gate. A description of these gates can be found in appendix B.1.



(k,l)	$J_{kl}^{(2)}/J^3$
$\{(0,1),(1,2),$ $(6,7),(7,8),$ $(0,3),(3,6),$ $(2,5),(5,8)\}$	$-\frac{923}{192}+\frac{113}{54\sqrt{2}}$
$\{(1,4),(4,7),$ $(3,4),(4,5)\}$	$-\frac{2357}{288}+\frac{113}{27\sqrt{2}}$

Figure 4.5: *Left*: The qubits of the nine-qubit quantum information processor are arranged on a lattice. The lines connecting qubits i and j indicate the values of the coupling constants J_{ij}; *right*: The table shows the two different values of the coupling constants $J_{kl}^{(2)}$ for each qubit pair (k,l).

These decompositions will be used in the next section for the simulation of a quantum algorithm. The $U_{kl}(\pi/8)$-gate itself can be generated approximately either by repeated application of the original or of the embedded Super-WHH recoupling sequence using either condition (4.8) or relation (4.19) for the determination of the free evolution time Δt between successive fast pulses.

4.3.2 Lattice Model of a Quantum Computer

For the subsequent numerical simulations of a quantum algorithm we consider a quantum information processor consisting of $n = 9$ qubits which are arranged on a lattice as indicated in figure 4.5. The coupling constants of vertical or horizontal qubit pairs are assumed to be equal while the coupling constants of diagonal neighbors are smaller by a factor of $2^{-3/2}$ due to the larger distance between them. Non-neighboring qubits are assumed to be uncoupled. According to relation (4.19) this implies that in the embedded recoupling scheme two different time intervals Δt are required for the free evolutions. The values of the coupling strengths $J_{kl}^{(2)}$ (4.17) for the 9-qubit lattice used in our subsequent simulation are apparent from the table of figure 4.5.

In the following it is assumed that a quantum algorithm is performed on this quantum information processor according to the following rules:

(i) Single-qubit gates are performed instantaneously and perfectly. (Even though in the setting of [YLM$^+$04] selective gates are generated slowly using weak pulses, reference [YLM$^+$04] describes a way of implementing them in such a way that the inter-qubit couplings are decoupled during the application time. Hence, in good approximation, they might be viewed as being applied instantaneously.)

(ii) Two-qubit gates between vertical or horizontal neighboring qubits are performed by repeated applications of the unitary $U_{kl}(\pi/8)$-gate in combination with single-qubit gates as illustrated in figure 4.4. The $U_{kl}(\pi/8)$-gate itself is generated by applying Super-WHH sequences n_{swhh} times as indicated in figure 4.2.

(iii) If the target qubits of a two-qubit gate are not vertical or horizontal neighbors they are moved into such positions by applying a sequence of SWAP-gates according to the following simple strategy *: If the vertical position of the qubits is the same, move the lower qubit to the upper one. Otherwise,

*A better but more complicated strategy would be to minimize the number of SWAP-gates. Note that due to the simple strategy used in this paper the first few iterations of a quantum algorithm take different amounts of computation time because the initial positions of the logical qubits are varying and so does the number of SWAP-gates.

move the lower one to the same horizontal position and afterwards move the left one as far as necessary to the right.

(iv) A Super-WHH sequence is always applied in such a way that the qubit whose physical position has the smaller label (compare with figure 4.5) is qubit k in W_{xy}^{kl}, i. e. it is transformed by the X transformations. The gate sequence of the $\mathsf{CP}(\varphi)$-gate (compare with figure 4.4) is applied in such a way that the first single-qubit gate is applied always to the qubit with the smaller label.

4.3.3 The Quantum Algorithm

In order to investigate the stabilizing properties of the embedded recoupling scheme the quantum algorithm of the quantum sawtooth map [BCMS01] is simulated according to the rules of the preceding subsection. One iteration of the quantum sawtooth map transforms an initial n-qubit quantum state $|\Psi(0)\rangle$ to the quantum state

$$|\Psi(1)\rangle = \exp\Big(-\frac{i}{2}m^2T\Big)\exp\Big(-ikV(q)\Big)|\Psi(0)\rangle \tag{4.22}$$

with the sawtooth potential $V(q) = -\frac{1}{2}(q-\pi)^2$ $(0 \leq q < 2\pi)$ and the (dimensionless) momentum operator m whose eigenstates form the computational basis, $m|i\rangle = i|i\rangle$ for $i = 0, 1 \ldots, 2^n - 1$. The position operator q is related to the momentum operator via the quantum Fourier transform (QFT):

$$q = U_{\mathrm{QFT}}^{-1} \cdot \frac{2\pi}{d} m \cdot U_{\mathrm{QFT}}. \tag{4.23}$$

Initially the nine-qubit quantum information processor is prepared in the momentum eigenstate $|\Psi(0)\rangle = |100110011\rangle$. The (dimensionless) parameters of the sawtooth map are assumed to have the same values as in the previous simulations of reference [LS05], i. e. $T = 2\pi/2^n$ and $kT = -0.5$. Therefore, in Husimi functions[†], such as the ones presented in figure 4.8, the dynamics of the sawtooth map are restricted to a phase-space cell of size $2\pi \times 2\pi$ and its corresponding classical dynamics are integrable. In these Husimi functions the initial state corresponds to a horizontal line slightly above the middle.

Our gate decomposition of the quantum algorithm of this sawtooth map consists of $n_g = 2n^2 + 2n$ quantum gates. A detailed description can be found in appendix B.2. In particular, $2 \times n(n+1)/2$ quantum gates originate from the two quantum Fourier transforms after which the inversion of the qubit positions is taken care of by relabeling instead of swapping.

4.3.4 Numerical Results

In figures 4.6, 4.7, and 4.8 results of our numerical simulations of the pure state fidelity

$$f(t) = \big|\langle\Psi(t)|\Psi_{\mathrm{ideal}}(t)\rangle\big|^2 \tag{4.24}$$

are presented for different numbers of repetitions $n_{\mathrm{swhh}} \in \{5, 6, \ldots, 10, 16\}$ of the Super-WHH sequences. For each value of n_{swhh} we calculated the fidelity of the quantum state $|\Psi(t)\rangle$ of the quantum sawtooth map for up to $t = 300$ iterations as well as the corresponding Husimi functions.

The quadratic-in-time fidelity decay of the original recoupling scheme is clearly apparent from figures 4.6 and 4.7. (The corresponding fidelities are plotted in *black*). This decay is caused by the coherent accumulation of errors due to the second-order AHT-term of the Super-WHH sequences involved in the realizations of the unitary $U_{kl}(\pi/8)$-gates. The situation is somewhat reminiscent of the situation analyzed in subsection 3.2.3, where it was assumed that each gate is preceded by a static imperfection.

[†]The definition of a Husimi function is given in appendix B.3

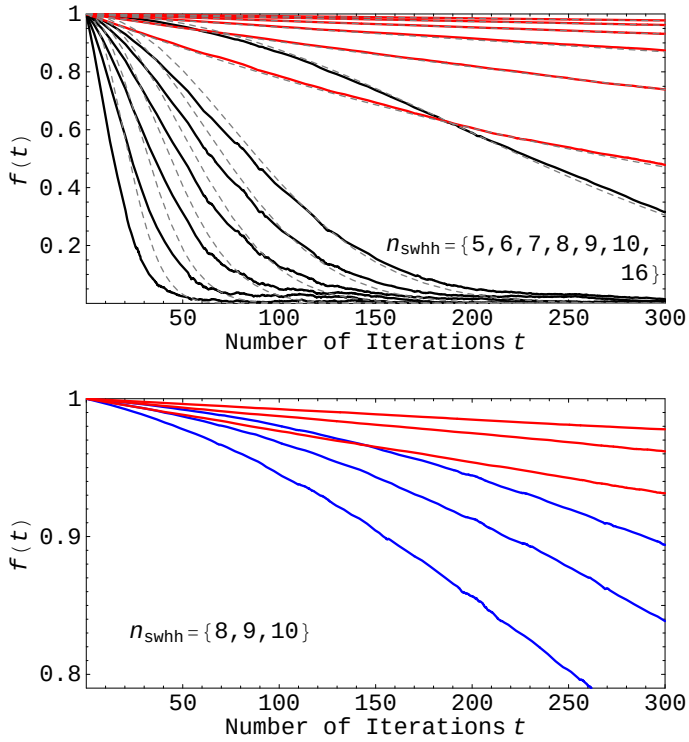


Figure 4.6: *Upper plot:* Fidelity plots of the quantum sawtooth map implemented with the original recoupling scheme of [YLM$^+$04] (lower plots, black) and the corresponding plots of the embedded recoupling scheme (upper plots, red): Dashed curves show the fidelity estimations according to equations (4.25) and (4.26). *Lower plot:* Fidelity plots of the embedded recoupling scheme (upper plots, red) and the embedded but unsymmetrized scheme (lower plots, blue).

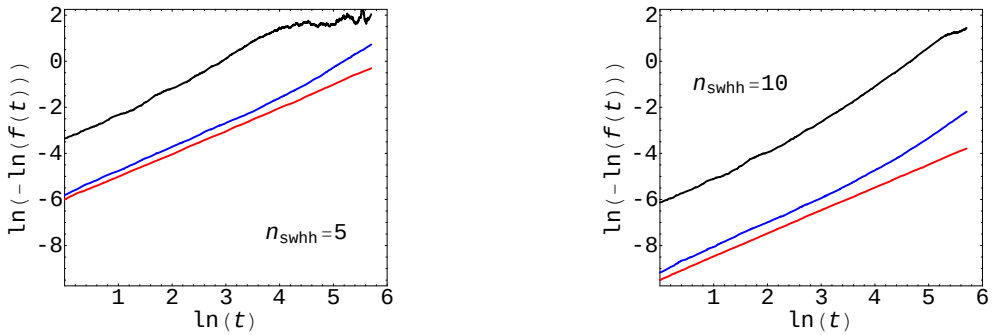


Figure 4.7: Logarithmic fidelity plots of the quantum sawtooth map with $n_{\mathrm{swhh}} = 5$ *(left)* and $n_{\mathrm{swhh}} = 10$ *(right)*: the original scheme of [YLM$^+$04] (upper curve, black), the embedded scheme (lowest curve, red), and the embedded scheme without symmetrization (middle curve, blue).

(Here the imperfection depends on the index pair (k, l).) The t-dependence of the fidelity can be fitted by the function

$$f(t) = \exp\bigl(-c \cdot t^2/n_{\mathrm{swhh}}^4\bigr) \tag{4.25}$$

with $c \approx 0.87$ (compare with the seven lowest dashed lines of the upper picture of figure 4.6). According to equation (3.35) describing the behavior of the entanglement fidelity in the presence of static imperfections (subsection 3.2.3), there should also be a linear contribution in the exponent of (4.25) which dominates the fidelity decay for small numbers of iterations. Neglecting this linear contribution is the reason for the slightly imperfect overlap of our fitted fidelities with the corresponding numerical results.

Using the embedded Super-WHH sequence together with the appropriately chosen free evolution times given by equation (4.19), it is possible to get an almost linear-in-time fidelity decay at least on time scales where errors of the order of $\mathcal{O}\bigl(J(J\Delta t)^4\bigr)$ are negligible (compare with Figs. 4.6 and 4.7 (*red* plots)). In these cases the fidelity decay can be fitted by the function

$$f(t) = \exp\bigl(-c_{\mathsf{ESDD}} \cdot t/n_{\mathrm{swhh}}^5\bigr) \tag{4.26}$$

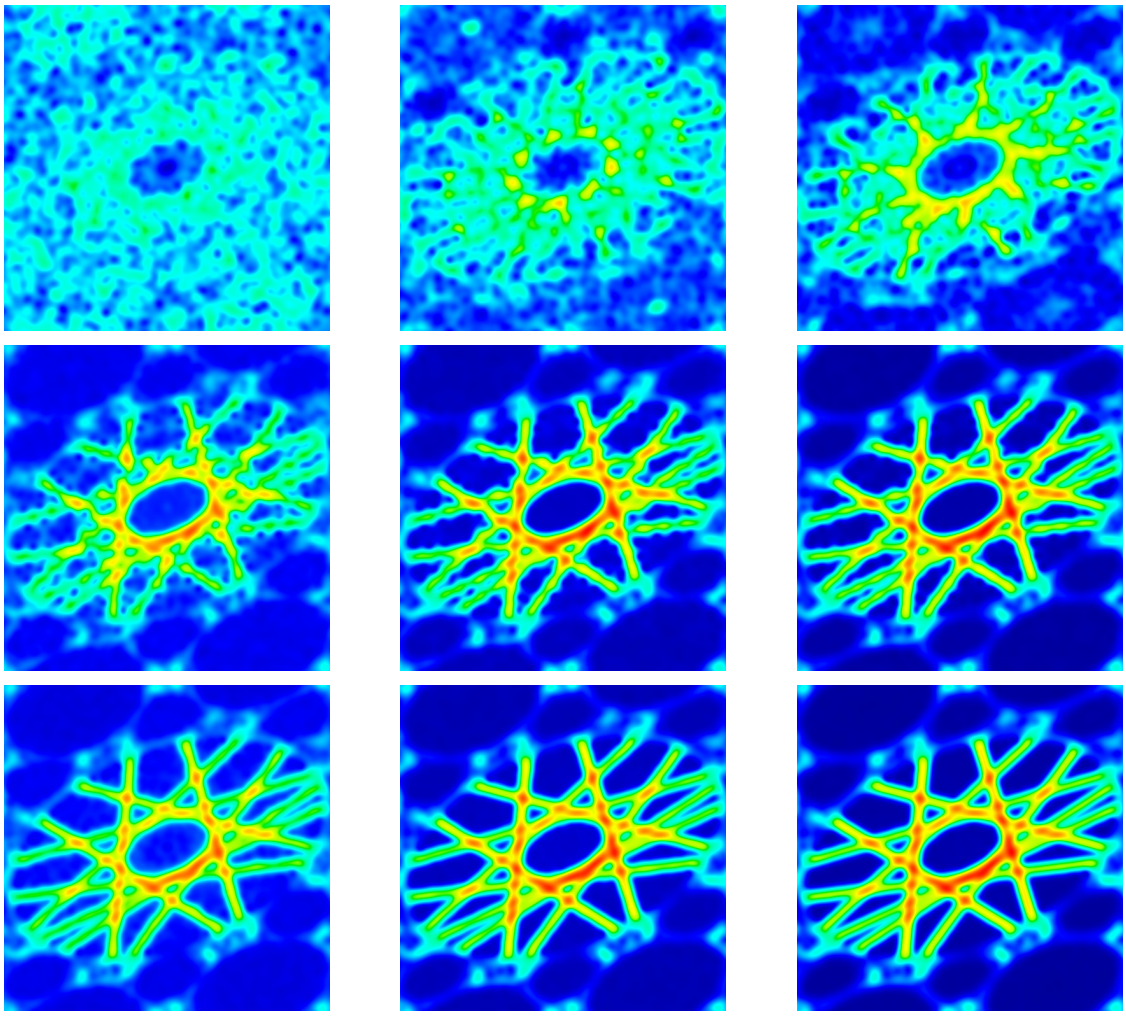
with $c_{\mathsf{ESDD}} \approx 7.85$ (compare with the six upper dashed lines of the upper picture of figure 4.6 which are almost indistinguishable from the corresponding full curves). The use of the embedded Super-WHH sequence does not only improve the action of a single $U_{kl}(\pi/8)$-gate but also prevents the residual imperfections to accumulate during the subsequent application of multiple gates. It can therefore be seen as a variant of the PAREC-method of section 3.2.

Simulations based on the embedded recoupling scheme without the symmetrization step are shown in figure 4.6 (lower part, *blue* plots) and figure 4.7 (*blue*). The fidelity decay is suppressed significantly but on the time scale of these plots it is still quadratic in time. This originates from the fact that terms of the Hamiltonian of equation (4.10) of the form $\alpha_k\alpha_l$, $\alpha \in \{X, Y, Z\}$, are not eliminated by the restricted randomization.

4.4 Conclusions

We showed how a selective recoupling scheme can be embedded into a stochastic decoupling scheme in such a way that the desired coupling remains conserved and that, in addition, the coherent accumulation of higher-order errors is suppressed significantly. While we focused on a specific example, the same general idea applies to other recoupling schemes as well. Even if computation times of a quantum information processor are so long that the residual higher-order interaction term of equation (4.18) of the order of $\mathcal{O}\bigl(J(J\Delta t)^4\bigr)$ is no longer negligible, it is possible to suppress also these errors significantly by a suitable adjustment of the free evolution time Δt involved in the realization of the relevant two-qubit gates ($U(\phi)$-gates). In generalization of the procedure discussed in section 4.2 (compare with condition (4.19)) this can be achieved either by explicitly calculating the fourth-order contribution of AHT and by solving the corresponding implicit equation of fifth order for Δt involving renormalized coupling strengths or, alternatively, adjusting the value of Δt so that the resulting fidelity decay is as small as possible.

Basic properties of our embedded scheme were analyzed for a single two-qubit gate. In particular, it was demonstrated that our proposed embedded symmetrized recoupling scheme results in an improvement of the scaling of the error of a swapping gate with n_{swhh}^{-5} instead of n_{swhh}^{-4}. Here, n_{swhh} denotes the number of repetitions of an embedded Super-WHH sequence which are required for the realization of the phase gate. Therefore, in our embedded recoupling scheme fewer numbers of repetitions of Super-WHH sequences are necessary for achieving a particular degree of error suppression. Typically, this also implies fewer pulses which are required for performing a quantum computation with a particular error tolerance. This aspect is apparent from the upper plot of figure 4.6 where at $t \approx 70$ iterations



$n_{\mathrm{swhh}} = 6$ $n_{\mathrm{swhh}} = 8$ $n_{\mathrm{swhh}} = 10$

Figure 4.8: Husimi functions of the quantum states resulting from the quantum sawtooth map: *(Upper row)* The original scheme of [YLM$^+$04], *(Middle row)* the embedded but unsymmetrized scheme, *(Lower row)* the embedded scheme. The $U(\pi/8)$-gates used in the computations consist of $n_{\mathrm{swhh}} = \{6, 8, 10\}$ Super-WHH sequences (from left to right). These functions are averaged over $290 \leq t \leq 299$ numbers t of iterations of the sawtooth map.

the fidelity of the original recoupling scheme with $n_{\mathrm{swhh}} = 16$ is the same as the one of the embedded symmetrized recoupling scheme with $n_{\mathrm{swhh}} = 6$.

While the original Super-WHH sequence makes use of selective pulses on two of the qubits at the same time, our embedded scheme also makes use of simultaneous selective pulses on all qubits. Since a selective pulse addressing a qubit with Larmor frequency ω_k induces erroneous rotations of qubits with nearby Larmor frequencies it may become important to use correction techniques as described in reference [SVC00].

Part II

Codes and Cryptography

5 Classical Error Correction

One of the fundamental quests of classical information theory is to transmit information reliably over a noisy channel. Addressed by Shannon in 1948 [Sha48], his famous noisy coding theorem associates to each channel a non-negative number C, the so-called capacity of the channel, and assures that for any rate below C, reliable information transmission over the channel is possible with the help of error-correcting codes. This chapter serves to provide the background on classical error correction which is necessary for the understanding of the forthcoming chapters on quantum error correction. For a more complete introduction to coding theory we refer to the books of MacKay [Mac03] and Welsh [Wel88].

After defining the capacity of discrete memoryless channels in section 5.1, we take a closer look at error-correcting codes in section 5.2. Linear codes form an important subclass of codes and are treated separately in section 5.3. Eventually, we show in section 5.4 that picking a linear code at random allows us to transmit information over the binary symmetric channel at a rate arbitrary close to the capacity, i. e. we prove a special case of the noisy coding theorem.

5.1 Capacity of Discrete Memoryless Channels

In classical information theory, a discrete memoryless channel is a simple model of a noisy channel used for information tranmission. It takes as input a symbol a_i from a certain input alphabet $\Sigma_1 = \{a_1, \ldots, a_s\}$ and outputs a symbol b_j from a certain output alphabet $\Sigma_2 = \{b_1, \ldots, b_r\}$ according to a fixed conditional probability distribution $p_{ji} = \Pr(b_j|a_i)$. The $r \times s$ dimensional matrix p_{ji} is called channel matrix. Most of the time we will consider channels where the input alphabet as well as the output alphabet is the set $\mathbb{F}_q$ containing the numbers from 0 to $q-1$.

The binary symmetric channel (BSC) is the most simple discrete memoryless channel. It is defined on the binary alphabet $\mathbb{F}_2$ and its channel matrix is given by $\Pr(a|a) = 1-p$ and $\Pr(a \oplus 1|a) = p$ with $a \in \mathbb{F}_2$. It is therefore completely specified by a single parameter $p \in [0,1]$.

An n-fold extension of a discrete memoryless channel corresponds to n uses of the channel. Such an extended channel takes as input a string $\vec{a} = (a_{i_1}, \ldots, a_{i_n}) \in \Sigma_1^n$ and outputs a string $\vec{b} = (b_{j_1}, \ldots, b_{j_n}) \in \Sigma_2^n$ according to the conditional probability distribution $\Pr(\vec{b}|\vec{a}) = \Pr(b_{j_1}|a_{i_1}) \cdot \Pr(b_{j_2}|a_{i_2}) \ldots \Pr(b_{j_n}|a_{i_n})$.

Definition 5.1.1 (Capacity of discrete memoryless channels)**.** Consider a discrete memoryless channel χ with input alphabet $\Sigma_1 = \{a_1, \ldots, a_s\}$, output alphabet $\Sigma_2 = \{b_1, \ldots, b_r\}$ and channel matrix $p_{ji} = \Pr(b_j|a_i)$. Let $P = \{p_1, \ldots, p_s\}$ be the probability distribution of a source S outputting symbol $a_i \in \Sigma_1$, i. e. $\Pr(a_i) = p_i$. Then the joint probability $\Pr(b_j, a_i)$ of the channel outputting symbol $b_j \in \Sigma_2$ and getting the input a_i is given by $\Pr(b_j, a_i) = \Pr(b_j|a_i)\Pr(a_i) = p_{ji}p_i$. The total probability of receiving output b_j is given by $q_j = \Pr(b_j) = \sum_{i=1}^{s} \Pr(b_j, a_i)$. The *capacity* $C(\chi)$ of the channel χ is defined as the mutual information between the source S and the receiver R, maximized over all input probability distributions P:

$$C(\chi) = \max_P I(S:R). \tag{5.1}$$

(The mutual information $I(S:R)$ was defined in equation (1.5) as $H(S) + H(R) - H(S,R)$, where $H(S) = -\sum_i p_i \log_2 p_i$ denotes the Shannon entropy of the source, $H(R) = -\sum_j q_j \log_2 q_j$ the Shannon entropy of the receiver, and $H(S,R) = -\sum_{ij} \Pr(b_j, a_i) \log_2 \Pr(b_j, a_i)$ the joint entropy of source and receiver.)

Remark. It is straightforward to calculate the capacity of the binary symmetric channel. The mutual information is maximal for a uniform input distribution and we get

$$C(\mathrm{BSC}_p) = 1 - H_2(p). \tag{5.2}$$

Naturally, the capacity of the n-fold extension of the BSC is n times its single capacity since the mutual information is additive.

5.2 Error Correction

Let us assume now that the input alphabet and the output alphabet of the noisy channel under consideration are both given by the set $\mathbb{F}_q$. When a string $\vec{x} \in \mathbb{F}_q^n$ is sent over the channel, the output will be a string $\vec{y} \in \mathbb{F}_q^n$ which is altered by the noise in the channel. What we would like to do is to deduce the original input $\vec{x}$ from the received string $\vec{y}$. Such a task becomes feasible only if we restrict the set of possible input strings.

Definition 5.2.1. A q-ary error-correcting code $\mathcal{C}$ of length n is a subset $\mathcal{C} \subset \mathbb{F}_q^n$ of all possible q-ary strings of length n. The members $\vec{x} \in \mathcal{C}$ of a code are called codewords.

The next step is to choose a decoding rule $\mathcal{D}$, which tells us which output strings have to be mapped to which codewords. The optimal decoding rule $\mathcal{D}_{\mathrm{opt}}$ decodes an output $\vec{y}$ as the codeword $\vec{x} \in \mathcal{C}$, which has the highest probability $\Pr(\vec{x}|\vec{y})$ of being sent through the channel conditioned on the event that $\vec{y}$ was received,

$$\mathcal{D}_{\mathrm{opt}}(\vec{y}) = \vec{x} \in \mathcal{C} \text{ s. t. } \Pr(\vec{x}|\vec{y}) \text{ is maximal.} \tag{5.3}$$

The probabilities $\Pr(\vec{x})$ of having codeword $\vec{x}$ as input must be known to implement such a decoder, since

$$\Pr(\vec{x}|\vec{y}) = \frac{\Pr(\vec{y}|\vec{x}) \Pr(\vec{x})}{\sum_{\vec{x}'} \Pr(\vec{y}|\vec{x}') \Pr(\vec{x}')}, \tag{5.4}$$

where $\Pr(\vec{y}|\vec{x}) = \prod_i \Pr(y_i|x_i)$ is specified by the channel matrix $\Pr(y_i|x_i)$. Hence, usually a so called maximum likelihood decoder is used, which decodes $\vec{y}$ to the codeword that maximizes $\Pr(\vec{y}|\vec{x})$,

$$\mathcal{D}_{\mathrm{mlk}}(\vec{y}) = \vec{x} \in \mathcal{C} \text{ s. t. } \Pr(\vec{y}|\vec{x}) \text{ is maximal.} \tag{5.5}$$

For the binary symmetric channel with $p < 1/2$, the maximum likelihood decoder is equivalent to a minimum distance decoder $\mathcal{D}_{\mathrm{min}}$ which decodes $\vec{y}$ as the codeword $\vec{x}$ that has minimum Hamming distance to $\vec{y}$,

$$\mathcal{D}_{\mathrm{min}}(\vec{y}) = \vec{x} \in \mathcal{C} \text{ s. t. } \mathrm{dist}(\vec{x}, \vec{y}) \text{ is minimal.} \tag{5.6}$$

For a given noisy channel, code $\mathcal{C}$ and decoding rule $\mathcal{D}$, the average error probability is given by

$$p_{\mathrm{error}} = \sum_{\vec{x}_{\mathrm{in}} \in \mathcal{C}} \Pr(\vec{x}_{\mathrm{in}}) \Pr(\vec{x}_{\mathrm{out}} \neq \vec{x}_{\mathrm{in}}|\vec{x}_{\mathrm{in}}) = \sum_{\vec{x}_{\mathrm{in}} \in \mathcal{C}} \Pr(\vec{x}_{\mathrm{in}}) \sum_{\vec{y}:\mathcal{D}(\vec{y}) \neq \vec{x}_{\mathrm{in}}} \Pr(\vec{y}|\vec{x}_{\mathrm{in}}), \tag{5.7}$$

where $\Pr(\vec{x}_{\mathrm{in}})$ denotes the probability of having $\vec{x}_{\mathrm{in}}$ as input string. In order to communicate reliably over the channel, we have to find a code $\mathcal{C}$ and decoder $\mathcal{D}$ such that this error probability, or even better the maximum error probability

$$p^*_{\mathrm{error}} = \max_{\vec{x}_{\mathrm{in}} \in \mathcal{C}} \Pr(\vec{x}_{\mathrm{out}} \neq \vec{x}_{\mathrm{in}}|\vec{x}_{\mathrm{in}}), \tag{5.8}$$

becomes very small.

The next subsection examines the conditions under which perfect error correction ($p^*_{\mathrm{error}} = 0$) becomes possible. Afterwards, the succeeding subsection deals with Shannon's noisy coding theorem, which tells us under which conditions error correction is possible if we allow some small probability of error ($p^*_{\mathrm{error}} < \varepsilon$).

5.2.1 Perfect Error Correction

If a code $\mathcal{C}$ has the property that its codewords are very distinct, it may become possible to reconstruct the originally sent codeword $\vec{x} \in \mathcal{C}$ from the received $\vec{y}$ in a perfect manner (at least as long as not to many errors occur). To formulate this idea precisely, we need the following definition.

Definition 5.2.2. The minimum distance d of an error-correcting code $\mathcal{C}$ is defined as the minimum Hamming distance between different codewords $\vec{x}, \vec{y} \in \mathcal{C}$:

$$d(\mathcal{C}) = \min_{\vec{x},\vec{y}\in\mathcal{C} \text{ s. t. } \vec{x}\neq\vec{y}} \operatorname{dist}(\vec{x}, \vec{y}). \tag{5.9}$$

Lemma 5.2.1. *Given a q-ary error-correcting code $\mathcal{C}$ of length n with minimum distance $d \geq 2e+1$, information can be sent reliably over a noisy channel as long as the channel does not introduce more than e errors. The transmission rate is given by $\log_q(|\mathcal{C}|)/n$.*

Proof. To deduce the originally sent codeword $\vec{x}$, we use minimum distance decoding. Since the e-spheres $S_e(\vec{x}) = \{\vec{s} \in \mathbb{F}_q^n | \operatorname{dist}(\vec{x}, \vec{s}) \leq e\}$ around distinct codewords of a code with distance greater than $2e$ do not overlap, the original codeword can be recovered from the received $\vec{y}$ as long as no more than e errors are made by the channel. □

At which rate can we encode information if we want to protect it perfectly against e errors, i. e if we demand a distance $d = 2e + 1$? A lower bound on this rate is given by the Gilbert Varshamov bound.

Theorem 5.2.2 (see e. g. chapter 4.2 in [Wel88])**.** *Gilbert Varshamov lower bound for q-ary codes. A lower bound on the maximum number of codewords $A_q(n, d)$ of a q-ary code of length n with minimum distance d is given by*

$$A_q(n, d) \geq q^n / \left(\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i \right). \tag{5.10}$$

Proof. Suppose $\mathcal{C}$ is a code of length n with minimum distance d and maximum number of codewords. There can be no vector in $\mathbb{F}_q^n \setminus \mathcal{C}$ that has distance greater than d from all the codewords of $\mathcal{C}$. All q^n vectors have to be included in the $d-1$ spheres around the A_q codewords. An upper bound on the number of vectors contained in these spheres is given by

$$A_q(n, d) \cdot \sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i. \qquad \square$$

Corollary 5.2.3. *For large n the Gilbert Varshamov lower bound becomes*

$$\frac{\log_q A_q(d, n)}{n} \geq 1 - H_{q[\log_q]}\Big(1 - \frac{d}{n}, \frac{d/n}{q-1}, \ldots, \frac{d/n}{q-1}\Big). \tag{5.11}$$

Proof. Setting $\lambda = (d-1)/n$ and writing the sum over i as

$$\sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i = \sum_{i=0}^{\lambda n} \binom{n}{i} \Big(\frac{q-1}{q}\Big)^i \Big(\frac{1}{q}\Big)^{n-i} \cdot q^n, \tag{5.12}$$

the Chernoff bound 1.2.1 can be applied to obtain the upper bound

$$\exp_q\Big(n H_{q[\log_q]}\big(1-\lambda, \lambda/(q-1), \ldots, \lambda/(q-1)\big)\Big) \tag{5.13}$$

if $\lambda = (d-1)/n < (q-1)/q$. The proof is completed noting that $H_{q[\log_q]}$ is monotonically increasing in λ. □

5.2.2 Shannon's Noisy Coding Theorem

Here we state Shannon's noisy coding theorem for discrete memoryless channels (see e. g. [Mac03, chapter 10] or [Wel88, section 3.5]).

Theorem 5.2.4 (Shannon's noisy coding theorem)**.** *For any $\varepsilon > 0$ and R smaller than the channel capacity C, there exists (for large enough n) a code $\mathcal{C}$ of length n and rate not smaller than R, together with a decoding rule $\mathcal{D}$, such that the maximum probability p^*_{error} of getting a decoding error is smaller than ε.*

We will give a proof for the special case of the binary symmetric channel in section 5.4. It can also be shown that transmission at rates above the capacity becomes an impossible task if we continue to demand an arbitrary low error rate, see e. g. [Wel88, section 3.6].

5.3 Linear Codes

Definition 5.3.1. A linear q-ary error-correcting code $\mathcal{C}$ of length n is a subspace of $\mathbb{F}_q^n$. If $\mathcal{C}$ is a k-dimensional subspace, we say $\mathcal{C}$ is an $[n,k]_q$ code or denote it as $\mathcal{C}_{[n,k]_q}$. If its minimum distance d is known, we say it is an $[n,k,d]_q$ code.

Remark. The minimum distance d of an $[n,k]_q$ code $\mathcal{C}$ is the minimum weight of its nonzero codewords since $d(\mathcal{C}) = \min_{\vec{x}\neq\vec{y}\in\mathcal{C}} \operatorname{dist}(\vec{x},\vec{y}) = \min_{\vec{x}\neq\vec{0}\in\mathcal{C}} \operatorname{wt}(\vec{x})$.

If $\dim(\mathcal{C}) = k$, $\mathcal{C}$ consists of q^k codewords which are linear combinations of k linearly independent generating elements $\vec{g}_i \in \mathbb{F}_q^n$ $(i = 1,\ldots,k)$. The $k \times n$ matrix G whose rows are the $\vec{g}_i$ is called generator matrix. The k row vectors of the generator matrix G can be extended to form a basis of $\mathbb{F}_q^n$ by adding $n-k$ additional linearly independent vectors $\vec{g}_j$ $(j = k+1,\ldots,n)$. Each element $\vec{x}$ in $\mathbb{F}_q^n$ can then be expressed as a linear combination of the $\vec{g}_i$: $\vec{x} = \sum_{i=1}^n u_i\vec{g}_i$, $u_i \in \mathbb{F}_q$. The string $(u_{k+1},\ldots,u_n) \in \mathbb{F}_q^{n-k}$ is called the syndrome. For a given string $\vec{x}$, the syndrome can easily be calculated by matrix multiplication with an $(n-k)\times n$ dimensional parity check matrix H whose rows $\vec{h}_i$ $(i = 1,\ldots,n-k)$ satisfy $\vec{h}_i\cdot\vec{g}_j = 0$ for $j = 1,\ldots,k$ and $\vec{h}_i\cdot\vec{g}_j = \delta_{i,j-k}$ for $j = k+1,\ldots,n$. There is a one-to-one correspondence between the cosets of $\mathcal{C}$ in $\mathbb{F}_q^n$ and the syndromes.

When a string $\vec{y}$ is received over a noisy channel, the set of possible errors is given by $\{\vec{e} = \vec{y}-\vec{x} | \vec{x}\in\mathcal{C}\}$. If $\mathcal{C}$ is a linear code, $\vec{x}\in\mathcal{C}$ implies that $-\vec{x}$ is also a member of $\mathcal{C}$. This means that the set of possible errors is given by the coset of $\mathcal{C}$ in $\mathbb{F}_q^n$ which contains $\vec{y}$ and which can be identified unambiguously by the syndrome $\vec{s} = H\vec{y}^T$ of the received string $\vec{y}$. Certain decoders are able to make use of this fact to speed up the decoding process to some extent. The minimum distance decoder $\mathcal{D}_{\min}$ for example has to find the element $\vec{e}_0$ of minimum weight in the coset of $\mathcal{C}$ which contains $\vec{y}$. For all coset members $\vec{y}' \in \vec{y}+\mathcal{C}$ the result of the decoder is given by $\mathcal{D}_{\min}(\vec{y}') = \vec{y}' - \vec{e}_0$. Therefore, knowledge of the syndrome $\vec{s} = H\vec{y}^T$ of the received $\vec{y}$ allows the use of a look-up table $\vec{e}_0(\vec{s})$ (which has to be calculated only once in the beginning) to find the required minimum-weight-element $\vec{e}_0$.

In the last section we gave a lower bound (Gilbert Varshamov bound) on the rate of codes with minimum distance d. For linear codes we can find a better lower bound (which is sometimes called Varshamov bound) by taking into account the structure of such codes. The following lemma establishes a relation between the parity check matrix H and the minimum distance d. It is then used to prove the lower bound given in the following theorem which is a generalization of [MS77, theorem 12 from chapter 1, §10] or [Wel88, problem 21 chapter 4] to q-ary codes.

Lemma 5.3.1 (Theorem 10 from chapter 1, §10 in [MS77])**.** *If H is the $(n-k)\times n$-dimensional parity check matrix of an $[n,k]_q$ code, then the code has minimum distance d iff every $d-1$ columns of H are linearly independent and some d columns are linear dependent.*

Proof. Some d columns of H are linear dependent $\Leftrightarrow H\vec{x}^T = 0$ for some $\vec{x}$ with weight $d \Leftrightarrow$ There is a codeword $\vec{x}$ of weight d. The same chain applies to the $d-1$ linearly independent columns of H with the result that there are no codewords of weight less than d. □

Theorem 5.3.2 (Varshamov lower bound for linear q-ary codes)**.** *An $[n,k,d]_q$ code exists provided that*

$$\sum_{j=0}^{d-2} \binom{n-1}{j} (q-1)^j < q^{n-k}. \tag{5.14}$$

Proof. We construct an $(n-k)\times n$ dimensional parity check matrix H such that all $d-1$ columns are linearly independent and use lemma 5.3.1. The first column can by any nonzero $n-k$ column vector. Suppose we have chosen i columns such that all $d-1$ columns are linearly independent. We can add another column and keep this property if the number of distinct linear combinations of $d-2$ or fewer of these columns is less than q^{n-k}. This number is

$$\sum_{j=0}^{d-2} \binom{i}{j} (q-1)^j. \qquad \square$$

Remark. As an example we calculate the above bound for a binary code of length $n = 11$ and distance $d = 3$ and get $A_2(11,3) \geq 128$. The Gilbert Varshamov bound for general codes given in theorem 5.2.2 assures us only that $A_2(11,3) \geq 31$.

Remark. The asymptotic version of the above bound coincides with the asymptotic version of the bound for general codes given in corollary 5.2.3 if we replace $\log_q A_q(n,d)$ by k.

We close this section by giving the definition of the dual code $\mathcal{C}^\perp$ of a code $\mathcal{C}$. Dual codes are helpful in connection with quantum CSS codes as will become clear in section 6.3.

Definition 5.3.2. The dual code $\mathcal{C}^\perp$ of a q-ary code $\mathcal{C}$ of length n is defined using the ordinary inner product of vectors modulo q,

$$\mathcal{C}^\perp = \{\vec{x} \in \mathbb{F}_q^n \mid \forall \vec{c} \in \mathcal{C},\ \vec{x}\cdot\vec{c} = 0 \pmod q\}. \tag{5.15}$$

Remark. If $\mathcal{C}$ is an $[n,k]_q$ code, its dual code $\mathcal{C}^\perp$ is an $[n,n-k]_q$ code.

5.4 Random Linear Codes and the Binary Symmetric Channel

In this section it is shown that a random linear code can — at least in principle — be used to communicate reliably over a binary symmetric channel at a rate arbitrary close to its capacity. To achieve this goal, we do not demand perfect error correction as it was done in lemma 5.2.1, but we demand only a small maximum probability p^*_{error} of getting a decoding error. Since we are going to use a typical set decoder $\mathcal{D}_{\text{typ}}$, we first need to define typical sets and discuss their relevant asymptotic properties in subsections 5.4.1 and 5.4.2. Then, in subsection 5.4.3, it is shown that taking the average over all linear $[n,k]_q$ codes leads to an arbitrary small error probability p^*_{error} (for large enough n) which proves a special case of Shannon's noisy coding theorem.

5.4.1 Typical Sets

This subsection deals with typical sequences [HK02, section 2.6]. The asymptotic properties of a set of typical sequences allows such a set to be used to construct so-called typical-set decoders.

A discrete random variable X is characterized by a set of possible outcomes $A = (a_1, \ldots, a_s)$, $s = |A|$, together with an associated probability distribution $P = (p_1, \ldots, p_s)$ such that X takes on the values

$a_i \in A$ with probability $\Pr(X = a_i) = P(a_i) = p_i$. The outcome of an ensemble of n independent and identically distributed (iid) random variables $X^n = (X_1, \ldots, X_n)$ is a sequence $\vec{x} = (x_1, \ldots, x_n) \in A^n$ where the probability of getting outcome $\vec{x}$ is given by $P^n(\vec{x}) = P(x_1)P(x_2)\ldots P(x_n)$. If the ensemble is large, the output sequence will contain about $p_1 \cdot n$ times the symbol $a_1 \in A$, about $p_2 \cdot n$ times the symbol $a_2 \in A$, etc., which motivates the definition of a subset of typical sequences:

Definition 5.4.1. The set $T^n_\delta(X)$ of strongly δ-typical sequences is defined as the collection of strings in A^n whose relative frequency distribution of the symbols A is close to the probability distribution P:

$$T^n_\delta(X) = \Big\{\vec{x} \in A^n \text{ s.t. for every } x \in A, \big|N(x|\vec{x}) - nP(x)\big| < \frac{\delta n P(x)}{\log_q |A|}\Big\}, \tag{5.16}$$

where $N(x|\vec{x})$ denotes the number of times the letter $x \in A$ occurs in $\vec{x}$ (i. e. $N(x|\vec{x}) = |\{i \mid x_i = x\}|$) and the logarithm is taken with respect to the base q.

Theorem 5.4.1 (Asymptotic equipartition property of $T^n_\delta(X)$)**.**
(a) For any length n and any $\vec{x} \in T^n_\delta(X)$,

$$\Big|\frac{1}{n}\log_q P^n(\vec{x}) + H_{[\log_q]}(X)\Big| \leq \delta, \tag{5.17}$$

or in other words, for all $\vec{x} \in T^n_\delta(X)$, $\exp_q\big(-n(H_{[\log_q]}(X) - \delta)\big) \geq P^n(\vec{x}) \geq \exp_q\big(-n(H_{[\log_q]}(X) + \delta)\big)$.
(b) For any $\Delta > 0$ and for n sufficiently large,

$$\Pr(\vec{x} \in T^n_\delta(X)) = \sum_{\vec{x} \in T^n_\delta(X)} P^n(\vec{x}) \geq 1 - \Delta. \tag{5.18}$$

(c) For any $\Delta > 0$ and for n sufficiently large, the cardinality of $T^n_\delta(X)$ is bounded by

$$(1-\Delta)\exp_q\big(n(H_{[\log_q]}(X) - \delta)\big) \leq |T^n_\delta(X)| \leq \exp_q\big(n(H_{[\log_q]}(X) + \delta)\big). \tag{5.19}$$

Proof of (a).

$$\begin{aligned}
\Big|\frac{1}{n}\log_q P^n(\vec{x}) + H_{[\log_q]}(X)\Big| &= \Big|\sum_{x\in A}\frac{N(x|\vec{x})}{n}\log_q P(x) - \sum_{x\in A} P(x)\log_q P(x)\Big| \\
&\leq \sum_{x\in A}\frac{1}{n}\big|N(x|\vec{x}) - nP(x)\big|(-\log_q P(x)) \\
&\leq \sum_{x\in A}\frac{\delta P(x)}{\log_q |A|}(-\log_q P(x)) \qquad \text{by def.} \\
&= \delta \cdot H_{[\log_q]}(X)/\log_q |A| \leq \delta \qquad \square
\end{aligned}$$

Proof of (b). For each $x \in A$, let F_x be the event that $X^n = (X_1, \ldots, X_n)$ takes on a value $\vec{x} \in A^n$ that does not satisfy

$$\big|N(x|\vec{x}) - nP(x)\big| < \frac{\delta n P(x)}{\log_q |A|}.$$

Chebyshev's inequality tells us that

$$\Pr(F_x) = \Pr\Big(\Big|\frac{N(x|\vec{x})}{n} - P(x)\Big| \geq \frac{\delta P(x)}{\log_q |A|}\Big) \leq \frac{P(x)(1-P(x))}{n}\cdot\Big(\frac{\log_q |A|}{\delta P(x)}\Big)^2.$$

If a sequence $\vec{x}$ is not in $T^n_\delta(X)$, it follows that at least one of the events $\{F_x\}_{x\in A}$ occurs and by the union bound we have

$$\Pr(\vec{x} \notin T^n_\delta(X)) \leq \sum_{x\in A}\Pr(F_x) \leq |A|\frac{(\log_q |A|)^2}{n\delta^2}\max_{x\in A}\frac{1-P(x)}{P(x)},$$

which is smaller than any $\Delta > 0$ for sufficiently large n. $\square$

Proof of (c). We prove the upper bound using (a),

$$|T_\delta^n(X)| \cdot \exp_q\big(-n(H_{[\log_q]}(X)+\delta)\big) \leq \sum_{\vec{x}\in T_\delta^n(X)} P^n(\vec{x}) \leq 1.$$

The lower bound follows from (b) and (a),

$$1-\Delta \leq \sum_{\vec{x}\in T_\delta^n(X)} P^n(\vec{x}) \leq |T_\delta^n(X)| \cdot \exp_q\big(-n(H_{[\log_q]}(X)-\delta)\big). \qquad \square$$

Remark. The set $\tilde{T}_\delta^n(X)$ of weakly δ-typical sequences is defined as the collection of strings in A^n satisfying property (a) of theorem 5.4.1,

$$\tilde{T}_\delta^n(X) = \Big\{\vec{x}\in A^n \text{ s. t. } \Big|\frac{1}{n}\log_q P^n(\vec{x}) + H_{[\log_q]}(X)\Big| < \delta\Big\}. \tag{5.20}$$

It is possible to show that the set of weakly typical sequences also satisfies the remaining asymptotic equipartition properties (b) and (c). Therefore it would be sufficient to use weakly typical sets for the purpose of typical set decoding. But since we will need the strongly typical set later on in this thesis to construct conditional typical sets for the purpose of decoding certain random quantum codes, we decided to work with strongly typical sets right from the start. In the following, when we speak of typical sets or sequences we always mean strongly typical.

5.4.2 Joint Typical Sets

In the context of random quantum codes, occasionally we'll have to work with the conditional typical sets [HK02, section 2.6] corresponding to a certain joint typical set. We present the necessary material here, since it fits in this section dealing with typical sets in general.

Definition 5.4.2. Let the joint probability distribution of two random variables X and Y taking on values in the finite alphabets A and B be given by $\{P(x,y) \mid x\in A \text{ and } y\in B\}$. The set of jointly strongly δ-typical sequences $(\vec{x},\vec{y}) = ((x_1,y_1),\ldots,(x_n,y_n))$ ($\vec{x}\in A^n$ and $\vec{y}\in B^n$) of length n is defined by

$$T_\delta^n(XY) = \Big\{(\vec{x},\vec{y}) \text{ s. t. for all } x\in A \text{ and } y\in B, \big|N(xy|\vec{x}\vec{y}) - nP(x,y)\big| \leq \frac{\delta n P(x,y)}{\log_q|A\times B|}\Big\}, \tag{5.21}$$

where $N(xy|\vec{x}\vec{y}) = |\{i \mid (x_i,y_i)=(x,y)\}|$. For a given joint typical set $T_\delta^n(XY)$, we define the set of typical X-sequences as

$$T_\delta'^n(X) = \{\vec{x}\in A^n \mid (\vec{x},\vec{y})\in T_\delta^n(XY) \text{ for some } \vec{y}\in B^n\}, \tag{5.22}$$

and we define the conditional typical set for a given $\vec{x}\in A^n$ as

$$T_\delta^n(Y|\vec{x}) = \{\vec{y}\in B^n \mid (\vec{x},\vec{y})\in T_\delta^n(XY)\}. \tag{5.23}$$

Remark. Any $\vec{x}\in T_\delta'^n(X)$ also belongs to $T_\delta^n(X)$. *Proof.* For all $x\in A$ we have

$$\big|N(x|\vec{x}) - nP(x)\big| = \Big|\sum_{y\in B}\big(N(xy|\vec{x}\vec{y}) - nP(x,y)\big)\Big| \leq \sum_{y\in B}\big|N(xy|\vec{x}\vec{y}) - nP(x,y)\big|,$$

which holds for any $\vec{y}\in B^n$. By the definition of $T_\delta'^n(X)$,

$$\sum_{y\in B}\big|N(xy|\vec{x}\vec{y}) - nP(x,y)\big| \leq \sum_{y\in B}\frac{\delta n P(x,y)}{\log_q|A\times B|} = \frac{\delta n P(x)}{\log_q|A\times B|} \leq \frac{\delta n P(x)}{\log_q|A|}. \qquad \square$$

Theorem 5.4.2 (Asymptotic equipartition property of $T_\delta^n(XY)$)**.** *(a) For any $(\vec{x},\vec{y}) \in T_\delta^n(XY)$,*

$$\left|\frac{1}{n}\log_q P^n(\vec{x},\vec{y}) + H_{[\log_q]}(XY)\right| \leq \delta, \tag{5.24a}$$

$$\left|\frac{1}{n}\log_q P^n(\vec{x}) + H_{[\log_q]}(X)\right| \leq \delta, \tag{5.24b}$$

$$\left|\frac{1}{n}\log_q P^n(\vec{y}|\vec{x}) + H_{[\log_q]}(Y|X)\right| \leq 2\delta. \tag{5.24c}$$

(b) For any $\Delta > 0$, and n sufficiently large,

$$\Pr\big((\vec{x},\vec{y}) \in T_\delta^n(XY)\big) \geq 1-\Delta, \tag{5.25a}$$

$$\Pr\big(\vec{x} \in T_\delta'^n(X)\big) \geq 1-\Delta. \tag{5.25b}$$

(c) For any $\Delta > 0$, $\vec{x} \in T_\delta'^n(X)$, and n sufficiently large,

$$(1-\Delta)\exp_q\big(n(H_{[\log_q]}(XY)-\delta)\big) \leq |T_\delta^n(XY)| \leq \exp_q\big(n(H_{[\log_q]}(XY)+\delta)\big), \tag{5.26a}$$

$$(1-\Delta)\exp_q\big(n(H_{[\log_q]}(X)-\delta)\big) \leq |T_\delta'^n(X)| \leq \exp_q\big(n(H_{[\log_q]}(X)+\delta)\big), \tag{5.26b}$$

$$|T_\delta^n(Y|\vec{x})| \leq \exp_q\big(n(H_{[\log_q]}(Y|X)+2\delta)\big). \tag{5.26c}$$

Proof of (a). The proof of the first inequality is nearly identical to the proof of part (a) of theorem 5.4.1. To prove the second inequality, note that it was shown in the above remark that $(\vec{x},\vec{y}) \in T_\delta^n(XY)$ implies $\vec{x} \in T_\delta^n(X)$. The last inequality is proven by applying the first two inequalities to the expression $P^n(\vec{y}|\vec{x}) = P^n(\vec{x},\vec{y})/P^n(\vec{x})$. □

Proof of (b). The proof of the first part is nearly identical to the proof of part (b) of theorem 5.4.1. For the proof of the second part, we note that

$$\begin{aligned}\Pr\big((\vec{x},\vec{y}) \in T_\delta^n(XY)\big) &= \sum_{(\vec{x},\vec{y})\in T_\delta^n(XY)} P^n(\vec{x},\vec{y}) = \sum_{\vec{x}} \sum_{\text{s. t. } (\vec{x},\vec{y})\in T_\delta^n(XY)} \sum_{\vec{y}} P^n(\vec{x},\vec{y}) \\ &\leq \sum_{\vec{x}\in T_\delta'^n(X)} \sum_{\vec{y}} P^n(\vec{x},\vec{y}) = \sum_{\vec{x}\in T_\delta'^n(X)} P^n(\vec{x}) = \Pr\big(\vec{x} \in T_\delta'^n(X)\big).\end{aligned}$$ □

Proof of (c). The proof goes as the proof of part (c) of theorem 5.4.1, using the results of part (a) and (b). For $|T_\delta^n(Y|\vec{x})|$ only an upper bound can be proved, since the corresponding statement in (b) which is needed to prove the lower bound does not hold. □

5.4.3 Random Coding

We are now going to prove a special case of Shannon's noisy coding theorem (theorem 5.2.4). We consider the binary symmetric channel with error probability p, the capacity of which was shown to be $1 - H_2(p)$ in equation (5.2).

Theorem 5.4.3. *Let BSC_p be the binary symmetric channel with error probability p and let $\varepsilon > 0$. Then, as long as*

$$\frac{k}{n} < C(BSC_p) = 1 - H_2(p), \tag{5.27}$$

and for large enough n, there exists an $[n,k]_q$ code $\mathcal{C}$, together with a decoder $\mathcal{D}$, such that the maximum probability p_{error}^* of getting a decoding error is smaller than ε.

Proof. A binary linear $[n,k]_2$ code $\mathcal{C}$ is a k-dimensional subspace of $\mathbb{F}_2^n$. Hence it is completely specified by an $(n-k)\times n$ dimensional parity check matrix H such that $H\cdot\vec{x}^T=0$ for all $\vec{x}\in\mathcal{C}$. If we want to use such a code to send information over a n-fold extension of the binary symmetric channel with bit flip probability p (BSC_p^n), we need to specify the decoding algorithm. Let X be a random variable representing the error of BSC_p, i. e. X takes on the values $A=\{0,1\}$ with probability $P=\{1-p,p\}$. We are going to use a typical set decoder $\mathcal{D}_{\mathrm{typ}}$ which calculates the syndrome $H\cdot\vec{y}^T$ of the received vector $\vec{y}\in\mathbb{F}_2^n$, and checks whether there is exactly one error vector $\vec{e}$ within the typical set $T_\delta^n(X)$ such that $H\cdot\vec{e}^T=H\cdot\vec{y}^T$. If this is the case, the decoder outputs $\vec{x}_{\mathrm{out}}=\vec{y}-\vec{e}$, otherwise it produces a decoding error.

We are now going to determine an upper bound on the maximum decoding error probability p^*_{error}. Since the error produced by the BSC_p^n does not depend on its input $\vec{x}_{\mathrm{in}}$, the probability $\Pr(\vec{x}_{\mathrm{out}}\neq\vec{x}_{\mathrm{in}}|\vec{x}_{\mathrm{in}})$ of getting a decoding error does not depend on the input $\vec{x}_{\mathrm{in}}$, either. Hence,

$$p^*_{\mathrm{error}}=\max_{\vec{x}_{\mathrm{in}}\in\mathcal{C}}\Pr(\vec{x}_{\mathrm{out}}\neq\vec{x}_{\mathrm{in}}|\vec{x}_{\mathrm{in}})=\Pr(\vec{x}_{\mathrm{out}}\neq\vec{x}_{\mathrm{in}}). \tag{5.28}$$

To estimate the decoding error probability $\Pr(\vec{x}_{\mathrm{out}}\neq\vec{x}_{\mathrm{in}})$, we have to sum over all possible errors produced by the BSC_p^n:

$$\begin{aligned}\Pr(\vec{x}_{\mathrm{out}}\neq\vec{x}_{\mathrm{in}})&=\sum_{\vec{e}\in\mathbb{F}_2^n}p^e(1-p)^{n-e}\cdot\begin{cases}1 & \text{if typical set decoding for } \vec{e} \text{ fails}\\ 0 & \text{else}\end{cases}\\ &\equiv\sum_{\vec{e}\in\mathbb{F}_2^n}p^e(1-p)^{n-e}\cdot\mathbb{1}[\text{typical set decoding for } \vec{e} \text{ fails}]\,.\end{aligned} \tag{5.29}$$

Here we denoted by $e=\mathrm{wt}(\vec{e})$ the number of 1s in $\vec{e}$, a notation we shall use throughout. The function $\mathbb{1}[x]$ returns 1 if the boolean expression x is true and 0 if it is false. We split up this sum into a sum over typical errors and a sum over the remaining ones. The later can be upper bounded by theorem 5.4.1b leading to

$$\Pr(\vec{x}_{\mathrm{out}}\neq\vec{x}_{\mathrm{in}})\leq\Delta+\sum_{\vec{e}\in T_\delta^n(X)}p^e(1-p)^{n-e}\cdot\mathbb{1}[\text{typical set decoding for } \vec{e} \text{ fails}]\,. \tag{5.30}$$

The sum over the typical errors can be upper bounded by

$$\sum_{\vec{e}\in T_\delta^n(X)}p^e(1-p)^{n-e}\cdot\sum_{\vec{e}'\in T_\delta^n(X)}^{\text{s. t. }\vec{e}'\neq\vec{e}}\mathbb{1}\left[H\cdot(\vec{e}-\vec{e}')^T=\vec{0}^T\right]. \tag{5.31}$$

Now we take the average of p^*_{error} over all linear codes. Let

$$A_{n,k,q}=\{\mathcal{C}\subseteq\mathbb{F}_q^n\mid\mathcal{C}\text{ is an }[n,k]_q\text{-code}\} \tag{5.32}$$

denote the set containing all $[n,k]_q$ codes and let

$$A_{n,k,q}(\vec{x})=\{\mathcal{C}\in A_{n,k,q}\mid\vec{x}\in\mathcal{C}\} \tag{5.33}$$

be the subset of codes which contain a certain nonzero codeword $\vec{x}\in\mathbb{F}_q^n$. In the following we need an upper bound for the quantity $|A_{n,k,2}(\vec{x})|/|A_{n,k,2}|$. It is proved in corollary C.1.2 in appendix C.1 that such a bound is given by

$$\frac{|A_{n,k,q}(\vec{x})|}{|A_{n,k,q}|}=\frac{q^k-1}{q^n-1}\leq\frac{1}{q^{n-k}}. \tag{5.34}$$

With the help of the above bound we obtain

$$\begin{aligned}
\left\langle p^*_{\text{error}} \right\rangle_{\mathcal{C}\in A_{n,k,2}} &\leq \Delta + \sum_{\vec{e}\in T^n_\delta(X)} p^e(1-p)^{n-e} \cdot \sum_{\vec{e}'\in T^n_\delta(X)}^{\text{s.t. } \vec{e}'\neq\vec{e}} (1/2)^{n-k} && \text{by (5.34)} \\
&\leq \Delta + (|T^n_\delta(X)| - 1)2^{k-n} \\
&\leq \Delta + 2^{n(H_2(p)+\delta)-n+k} && \text{by theorem 5.4.1c.} \qquad (5.35)
\end{aligned}$$

This quantity becomes arbitrarily small for large enough n as long as

$$\frac{k}{n} < 1 - H_2(p) - \delta. \qquad (5.36)$$

Since the above statement holds for any δ, we are free to choose δ as small as we like. Hence, for any $\varepsilon > 0$ and any rate R below the channel capacity $C(\text{BSC}_p) = 1 - H_2(p)$, there exists (for large enough n) a linear code $\mathcal{C}$ of length n and rate not smaller than R, such that the maximum probability p^*_{error} of getting a decoding error is smaller than ε. □

Remark. The achievable rate for reliable transmission over the BSC_p as proven above is given by $1 - H_{2[\log_2]}(p)$. Demanding perfect error correction of up to np errors, the Gilbert-Varshamov bound in corollary 5.2.3 assures the existence of codes with a rate of at least $1 - H_{2[\log_2]}(2p)$. Let the maximum value of tolerable noise $p_{\max}$ of the BSC_p be defined as the value of p for which the transmission rate becomes zero. By comparing the two rates we find that permitting a small decoding error probability results in a value of $p_{\max}$ twice as high as in the case of perfect error correction.

6 Quantum Error-Correcting Codes

To be of any practical use, a quantum memory has to be accessible from the outside to allow for measurements and the manipulation of the stored data. Therefore, it can never be isolated perfectly from the environment and has to be treated as an open quantum system, i. e. as part of a larger quantum system. In such a system, the most general state evolution is not unitary anymore, but is given by a trace preserving completely positive map (tpcp-map) $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ between density operators on a Hilbert space $\mathcal{H}$ describing the system. Whereas unitary evolution is — at least in theory — always reversible, an error described by a tpcp-map can in general not be reversed, i. e. there exists no tpcp-map $\mathcal{R}$ such that $\mathcal{R}(\mathcal{A}(\rho)) = \rho$ for any $\rho \in \mathcal{S}(\mathcal{H})$. To be able to perform quantum error correction, we therefore have to demand less. The trick is to restrict our attention to a subspace $\mathcal{C}$, called quantum code, of the Hilbert space of the quantum memory which has to be protected. The question is whether it is possible to undo an error described by a tpcp-map $\mathcal{A}$ at least on such a subspace $\mathcal{C}$.

In section 6.1 we present the necessary and sufficient conditions a quantum code has to fulfill in order to be able to recover from a given set of errors. An important family of quantum codes is given by the so-called stabilizer codes, which are discussed in section 6.2. CSS codes form a subclass of stabilizer codes and are treated separately in section 6.3. By encoding a quantum register which is already encoded by some 'outer' stabilizer code a second time, this time using some other 'inner' stabilizer code, one obtains a so-called concatenated code as we will discuss in section 6.4.

6.1 Reversibility of Quantum Operations

Definition 6.1.1. A quantum error-correcting code $\mathcal{C}$ is a subspace of the Hilbert space of a quantum memory which we would like to preserve. For instance, a code which protects k qubits might encode them into a 2^k dimensional subspace $\mathcal{C}$ of the Hilbert space $\mathcal{H} = \mathcal{H}_2^{\otimes n}$ of n physical qubits.

Is it possible to undo a quantum error described by a tpcp-map $\mathcal{A}$ on such a subspace $\mathcal{C}$, i. e. does there exist a recovery operation described by a tpcp-map $\mathcal{R}$ such that

$$\mathcal{R}\big(\mathcal{A}(\rho)\big) = \rho \tag{6.1}$$

for all $\rho \in \mathcal{S}(\mathcal{C})$? The necessary and sufficient condition a quantum code has to fulfill to allow for the recovery from a tpcp-map $\mathcal{A}$ was found by Knill and Laflamme [KL97]:

Theorem 6.1.1 ([KL97; NCSB98])**.** *Let $\{A_\mu\}$ be the operators in an operator sum representation of a tpcp-map $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$,*

$$\mathcal{A} : \rho \mapsto \mathcal{A}(\rho) = \sum_\mu A_\mu \rho A_\mu^\dagger. \tag{6.2}$$

Then a necessary and sufficient condition for reversibility of $\mathcal{A}$ on a quantum code $\mathcal{C}$ is given by

$$\Pi_\mathcal{C} A_\mu^\dagger A_\nu \Pi_\mathcal{C} = \Pi_\mathcal{C} C_{\mu\nu}, \tag{6.3}$$

where $\Pi_\mathcal{C}$ denotes the projection on the code space and $C_{\mu\nu}$ is a Hermitian matrix.

Remark. The operator sum representation is not unique. But since different representations $\{A_\mu\}$,$\{B_\nu\}$ of a certain tpcp-map are related as $A_\mu = \sum_\nu u_{\mu\nu} B_\nu$ with unitary $u_{\mu\nu}$, the criterion given above does not depend on the representation.

Let us introduce the set containing all n-fold tensor products of Pauli operators,

$$\mathcal{P}_q^n = \{XZ(\vec{a}) \mid \vec{a} \in \mathbb{F}_q^{2n}\}, \tag{6.4}$$

as defined in section 1.2, as a basis for quantum errors acting on a quantum memory consisting of n qudits of dimension q.

Lemma 6.1.2. *If we consider a subset $\mathcal{E} \subseteq \mathcal{P}_q^n$ of such error operators and Knill and Laflamme's condition is satisfied for all errors $E_a \in \mathcal{E}$, i. e.*

$$\Pi_{\mathcal{C}} E_a^\dagger E_b \Pi_{\mathcal{C}} = \Pi_{\mathcal{C}} C_{ab} \quad \text{for all } E_{a,b} \in \mathcal{E}, \tag{6.5}$$

then the quantum code $\mathcal{C}$ allows for the correction of all tpcp-maps whose operator sum representation contains only elements which can be written as linear combinations of the $E_a \in \mathcal{E}$.

Proof. If the elements of a operator sum representation $\{A_\mu\}$ of $\mathcal{A}$ can be written as $A_\mu = \sum_i a_{\mu i} E_i$ with $E_i \in \mathcal{E}$ and (6.5) is satisfied, then equation (6.3) is satisfied, too:

$$\Pi_{\mathcal{C}} A_\mu^\dagger A_\nu \Pi_{\mathcal{C}} = \sum_{ij} a_{\mu i}^* a_{\nu j} \Pi_{\mathcal{C}} E_i^\dagger E_j \Pi_{\mathcal{C}} = \Pi_{\mathcal{C}} \sum_{ij} a_{\mu i}^* a_{\nu j} C_{ij} = \Pi_{\mathcal{C}} C'_{\mu\nu}. \tag{6.6}$$

□

Definition 6.1.2. For a given set $\mathcal{E} \subseteq \mathcal{P}_q^n$, a code is said to be degenerate if the matrix C_{ab} in (6.5) is singular.

Definition 6.1.3. A code is said to correct t errors if (6.5) is satisfied for the set $\mathcal{E}$ containing all Pauli operators which are composed of at least $n - t$ $\mathcal{I}$'s. If we define the weight of a Pauli operator as the number of qudits on which it acts non-trivially, the statement can be reformulated as follows: A code is said to correct t errors if (6.5) is satisfied for the set $\mathcal{E} = \{E_i \in \mathcal{P}_q^n | \operatorname{wt}(E_i) \leq t\}$.

Definition 6.1.4. A quantum code is said to have minimum distance d if it detects all errors in $\mathcal{E} = \{E_i \in \mathcal{P}_q^n | \operatorname{wt}(E_i) \leq d-1\}$, i. e. if $\Pi_{\mathcal{C}} E_i \Pi_{\mathcal{C}} = \alpha_i \Pi_{\mathcal{C}}$ for all $E_i \in \mathcal{E}$ with $\alpha_i \in \mathbb{C}$.

Remark. A quantum code with distance $d \geq 2t+1$ corrects t errors because (6.5) will be satisfied for the set $\mathcal{E} = \{E_i \in \mathcal{P}_q^n | \operatorname{wt}(E_i) \leq t\}$.

6.2 Stabilizer Codes

The stabilizer code formalism has been developed mainly by Gottesman in [Got96; Got97]. It has been generalized to handle quantum systems of dimension higher than two in [Got99; Rai99]. This section deals with quantum systems of dimension q (prime), but in principle q could also be a power of a prime.

The stabilizer formalism proposes the common eigenspaces of an abelian subgroup of the Pauli group $\mathfrak{P}_q^n$ as codespaces. Since $\mathfrak{P}_q^n$ and the space $\mathbb{F}_q^{2n}$, which forms a group under addition modulo q, are related by the ray representation (1.26), stabilizer codes can be described in two equivalent ways. We will focus mainly on the description in the $\mathbb{F}_q^{2n}$ picture.

6.2.1 Stabilizers and Codespaces

Definition 6.2.1. A stabilizer is a self-orthogonal subspace $L \subset \mathbb{F}_q^{2n}$ with respect to the symplectic inner product, i. e. $L \subseteq L^\perp$ where $L^\perp = \{\vec{x} \in \mathbb{F}_q^{2n} \mid \forall \vec{l} \in L, (\vec{x}, \vec{l})_{sp} = 0\}$. Equivalently, using the $XZ(\cdot)$ representation, a stabilizer $S = \{\omega^k XZ(\vec{l}) \mid \vec{l} \in L, k \in \mathbb{F}_q\}^*$ is an abelian subgroup of the Pauli group $\mathfrak{P}_q^n$.

*If $q = 2$, ω^k should be replaced by $\mu \in \{\pm 1, \pm i\}$

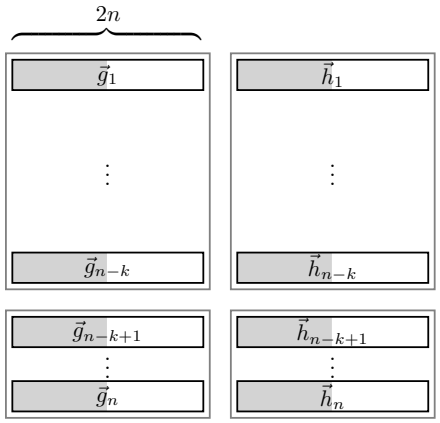


Figure 6.1: A stabilizer code is specified by the generating elements $\vec{g}_i = (\vec{g}_i^x, \vec{g}_i^z) \in \mathbb{F}_q^{2n}$ of a self-orthogonal subspace $L = \text{span}\{\vec{g}_1, \ldots, \vec{g}_{n-k}\} \subseteq L^\perp$. Any extension of these vectors to a hyperbolic basis $\mathbb{F}_q^{2n} = \text{span}\{\vec{g}_1, \ldots, \vec{g}_n, \vec{h}_1, \ldots, \vec{h}_n\}$ with $(\vec{g}_i, \vec{h}_j)_{sp} = \delta_{ij}$, $(\vec{g}_i, \vec{g}_j)_{sp} = 0$ and $(\vec{h}_i, \vec{h}_j)_{sp} = 0$, specifies a specific encoding.

Remark. An $(n-k)$-dimensional self-orthogonal subspace $L \subset \mathbb{F}_q^{2n}$ can always be specified by $n-k$ linearly independent generating elements, e. g. $L = \text{span}\{\vec{g}_1, \ldots, \vec{g}_{n-k}\}$ with $\vec{g}_i = (\vec{g}_i^x, \vec{g}_i^x) \in \mathbb{F}_q^{2n}$ for $i = 1, \ldots, n-k$.

Lemma 6.2.1 (see e. g. [Got97] or [NC00])**.** *A commutative subgroup $S \subset \mathfrak{P}_q^n$ corresponding to an $(n-k)$-dimensional self-orthogonal subspace $L \subset \mathbb{F}_q^{2n}$ divides the Hilbert space $\mathcal{H}_q^{\otimes n}$ into q^{n-k} common eigenspaces of dimension q^k.*

Proof. The construction of a basis of such a q^k-dimensional eigenspace in the next subsection implies the proof. □

Definition 6.2.2. The q^k-dimensional eigenspaces corresponding to an $(n-k)$-dimensional stabilizer $L \subseteq L^\perp \subseteq \mathbb{F}_q^{2n}$ can be labeled by a vector $\vec{s} \in \mathbb{F}_q^{n-k}$. They are defined to be the corresponding stabilizer codes $\mathcal{C}(L, \vec{s})$. We will use the notation $[[n,k]]_q$ code to denote an $(n-k)$-dimensional stabilizer code L, or strictly speaking, to denote the collection of all code spaces $\mathcal{C}(L, \vec{s})$ corresponding to a specific stabilizer L of dimension $n-k$. If the distance d of an $[[n,k]]_q$ code is known, we say the code is an $[[n,k,d]]_q$ code.

Remark. We will see below that all these code spaces are equivalent in the sense that they have identical error correcting properties.

6.2.2 Encoding Operations

Lemma 6.2.2 (see e. g. [Ham05; Ham03; WMU06])**.** *For a given set $\{\vec{g}_1, \ldots, \vec{g}_{n-k}\}$ of generating elements of some self-orthogonal $(n-k)$-dimensional subspace $L \subseteq L^\perp \subseteq \mathbb{F}_q^{2n}$, it is always possible to find vectors $\{\vec{g}_{n-k+1}, \ldots, \vec{g}_n\}$ and $\{\vec{h}_1, \ldots, \vec{h}_n\}$ such that*

$$(\vec{g}_i, \vec{h}_j)_{sp} = \delta_{ij}, \qquad (\vec{g}_i, \vec{g}_j)_{sp} = 0, \qquad (\vec{h}_i, \vec{h}_j)_{sp} = 0. \tag{6.7}$$

Vectors $\{\vec{g}_1, \ldots, \vec{g}_n, \vec{h}_1, \ldots, \vec{h}_n\}$ satisfying the above conditions are said to form a hyperbolic basis of $\mathbb{F}_q^{2n}$.

Remark. Note that $L^\perp = \text{span}\{\vec{g}_1, \ldots, \vec{g}_n, \vec{h}_{n-k+1}, \ldots, \vec{h}_n\}$. $L^\perp$ is called the normalizer.

We are now going to show that such an extension of the generating set of a stabilizer to a hyperbolic basis together with a set of phase factors to be defined below, completely specifies a unitary encoding operation. Let us define the operators

$$\overline{Z}_i = \theta_z(i) XZ(\vec{g}_i) \qquad \overline{X}_i = \theta_x(i) XZ(\vec{h}_i) \tag{6.8}$$

using some fixed set $\{\theta_\alpha(i) \in \{\omega^r | r \in \mathbb{F}_q\}\}_{\alpha\in\{x,z\},i\in\{1\ldots n\}}$ of phase factors, and let us define the abbreviations

$$\overline{X}^{\vec{u}} = \prod_{i=1}^{n} \overline{X}_i^{u_i}, \qquad X^{\vec{u}} = X^{u_1} \otimes \cdots \otimes X^{u_n} = XZ(\vec{u}, \vec{0}), \tag{6.9a}$$

$$\overline{Z}^{\vec{v}} = \prod_{i=1}^{n} \overline{Z}_i^{v_i}, \qquad Z^{\vec{v}} = Z^{v_1} \otimes \cdots \otimes Z^{v_n} = XZ(\vec{0}, \vec{v}). \tag{6.9b}$$

Since the $\{\overline{Z}_i\}_{i\in\{1,\ldots,n\}}$ commute with each other, there has to be a non-empty common eigenspace with eigenvalue list $(\lambda_1, \ldots, \lambda_n)$. Let us define $\overline{|0,\ldots,0\rangle}$ as a normalized vector in this eigenspace. By applying the operator $\overline{X}^{\vec{u}}$ to both sides of the eigenequation

$$\overline{Z}_i \overline{|0,\ldots,0\rangle} = \lambda_i \overline{|0,\ldots,0\rangle}, \tag{6.10}$$

and by making use of the fact that $\overline{Z}_i\overline{X}_i = \omega\overline{X}_i\overline{Z}_i$, we find that the state $\overline{X}^{\vec{u}}\overline{|0,\ldots,0\rangle}$ is an eigenstate of the $\{\overline{Z}_i\}_{i\in\{1,\ldots,n\}}$ with eigenvalue list $(\lambda_1\omega^{u_1}, \ldots, \lambda_n\omega^{u_n})$. Hence there have to exist at least q^n different eigenspaces, each of which must be of dimension one. In the following we will always chose $\overline{|0,\ldots,0\rangle}$ as the common eigenvector with eigenvalue list $(\lambda_1, \ldots, \lambda_n) = (1, \ldots, 1)$. The encoding operator U_{enc} is defined as the unitary which maps the states $|\vec{u}\rangle$ of the computational basis onto the states $\overline{|\vec{u}\rangle} = \overline{X}^{\vec{u}}\overline{|0,\ldots,0\rangle}$,

$$U_{\text{enc}} : X^{\vec{u}}|0,\ldots,0\rangle = |u_1,\ldots,u_n\rangle \mapsto \overline{X}^{\vec{u}}\overline{|0,\ldots,0\rangle} = \overline{|u_1,\ldots,u_n\rangle}. \tag{6.11}$$

It is straightforward to show that i)

$$\overline{Z}^{\vec{v}}\overline{|l_1,\ldots,l_n\rangle} = \omega^{\vec{v}\cdot\vec{l}}\overline{|l_1,\ldots,l_n\rangle} \tag{6.12}$$

and that ii)

$$U_{\text{enc}} X^{\vec{u}} U_{\text{enc}}^\dagger = \overline{X}^{\vec{u}} \qquad U_{\text{enc}} Z^{\vec{v}} U_{\text{enc}}^\dagger = \overline{Z}^{\vec{v}}. \tag{6.13}$$

Because of equation (6.13), we will call the operators $\{\overline{X}_i, \overline{Z}_i\}_{i\in\{1,\ldots,n\}}$ defined in (6.8) encoded X- and Z-operators.

Remark. The Clifford group consists of all operators which map Pauli operators to Pauli operators. It follows from equation (6.13) that U_{enc} is an element of the Clifford group.

The codespace with label (sometimes called syndrome) $(s_1, \ldots, s_{n-k})$ is the common eigenspace of (the generators of) the stabilizer $\{\overline{Z}_i\}_{i\in\{1,\ldots,n-k\}}$, with eigenvalue list $(\omega^{s_1}, \ldots, \omega^{s_{n-k}})$ and can be written as

$$\mathcal{C}(L, \vec{s}) = \text{span}\{\overline{|s_1,\ldots,s_{n-k},c_1,\ldots,c_k\rangle} \mid (c_1,\ldots,c_k) \in \mathbb{F}_q^k\}. \tag{6.14}$$

An encoded quantum state is given by

$$\mathcal{C}(L,\vec{s}) \ni |\psi\rangle_{\vec{s}} = \sum_{c_1,\ldots,c_k} \alpha_{c_1,\ldots,c_k} \overline{|s_1,\ldots,s_{n-k},c_1,\ldots,c_k\rangle}, \text{ with } \alpha_{c_1,\ldots,c_k} \in \mathbb{C}. \tag{6.15}$$

Operators $\{\overline{Z}_i, \overline{X}_i\}_{i\in\{n-k+1,\ldots,n\}}$ manipulate the encoded state, i. e. they perform logical Z_{i-n+k} and X_{i-n+k} operations on the $(i-n+k)$-th encoded qudit.

For a given hyperbolic basis $\{\vec{g}_1, \ldots, \vec{g}_n, \vec{h}_1, \ldots, \vec{h}_n\}$ of $\mathbb{F}_q^{2n}$, any vector $\vec{a} \in \mathbb{F}_q^{2n}$ can be expressed as linear combination of the basis elements,

$$\begin{aligned} \vec{a} &= (a_1^x, \ldots, a_n^x, a_1^z, \ldots, a_n^z) \\ &= \sum_{i=1}^{n-k} \big(s_i\vec{h}_i + n_i\vec{g}_i\big) + \sum_{i=n-k+1}^{n} \big(l_{i-(n-k)}^x \vec{h}_i + l_{i-(n-k)}^z \vec{g}_i\big), \end{aligned} \tag{6.16}$$

where $s_i = (\vec{g}_i, \vec{a})_{sp}$, et cetera. Together with equation (1.25) we obtain the following lemma.

Lemma 6.2.3. *Any Pauli operator $XZ(\vec{a} \in \mathbb{F}_q^{2n}) \in \mathcal{P}_q^n$ can be expressed (up to a phase) as product of some powers of the operators $XZ(\vec{g}_i), XZ(\vec{h}_i)$,*

$$XZ(\vec{a}) \sim \prod_{i=1}^{n-k} \big(XZ(\vec{h}_i)^{s_i} XZ(\vec{g}_i)^{n_i}\big) \prod_{i=1}^{k} \big(XZ(\vec{h}_{i+n-k})^{l_i^x} XZ(\vec{g}_{i+n-k})^{l_i^z}\big), \tag{6.17}$$

or by using the operators $\overline{Z}_i, \overline{X}_i$ defined in (6.8)*,*

$$\sim \prod_{i=1}^{n-k} \big(\overline{X}_i^{s_i} \overline{Z}_i^{n_i}\big) \prod_{i=1}^{k} \big(\overline{X}_{i+n-k}^{l_i^x} \overline{Z}_{i+n-k}^{l_i^z}\big) = \overline{X}^{(\vec{s},\vec{l}^x)} \overline{Z}^{(\vec{n},\vec{l}^z)}, \tag{6.18}$$

where the strings $\vec{s}, \vec{n} \in \mathbb{F}_q^{n-k}$ and $\vec{l}^x, \vec{l}^z \in \mathbb{F}_q^k$ are defined in (6.16)*.*

6.2.3 Correctable Errors

For which sets of errors $\mathcal{E} \subseteq \mathcal{P}_q^n$ is Knill and Laflamme's condition for reversibility satisfied on the codespaces $\mathcal{C}(L, \vec{s})$ of a stabilizer code, or in other words, what are the errors that can be corrected ? As we will see, neither does the answer depend on the label $\vec{s}$ of the codespace we have chosen to encode some information, nor does it depend on the encoding operation U_{enc}.

Lemma 6.2.4 (see e. g. [Got97])**.** *Let $\Pi_{\mathcal{C}(L,\vec{s})}$ be the projector on the codespace $\mathcal{C}(L, \vec{s})$. Then equation* (6.5) *with the substitution $\Pi_{\mathcal{C}} \mapsto \Pi_{\mathcal{C}(L,\vec{s})}$ will be satisfied for $\mathcal{E} \subseteq \mathcal{P}_q^n$ iff for each $E_a, E_b \in \mathcal{E}$ one of the following holds:*

- *$E_a^\dagger E_b$ is an element of the stabilizer S.*
- *There exists an element in S that does not commute with $E_a^\dagger E_b$.*

Proof. We are going to show that if one of the above conditions is satisfied for each $E_a, E_b \in \mathcal{E}$, equation (6.5) will be satisfied, too. If not, i. e. if there exists $E_a^\dagger E_b \notin S$ and there doesn't exist any element in S that does not commute with $E_a^\dagger E_b$, then equation (6.5) cannot be satisfied. The first point is equivalent to $\Pi_{\mathcal{C}(L,\vec{s})} E_a^\dagger E_b \Pi_{\mathcal{C}(L,\vec{s})} = \Pi_{\mathcal{C}(L,\vec{s})} C_{ab}$, with $C_{ab} = \omega^k$, $k \in \mathbb{F}_q$. Regarding the second point, let M be the non-commuting element in S and let its eigenvalue of the eigenspace $\mathcal{C}(L, \vec{s})$ be m, $M\Pi_{\mathcal{C}(L,\vec{s})} = m\Pi_{\mathcal{C}(L,\vec{s})}$. Then,

$$m\Pi_{\mathcal{C}(L,\vec{s})} E_a^\dagger E_b \Pi_{\mathcal{C}(L,\vec{s})} = \Pi_{\mathcal{C}(L,\vec{s})} E_a^\dagger E_b M \Pi_{\mathcal{C}(L,\vec{s})} = \omega^x \Pi_{\mathcal{C}(L,\vec{s})} M E_a^\dagger E_b \Pi_{\mathcal{C}(L,\vec{s})} = \omega^x m \Pi_{\mathcal{C}(L,\vec{s})} E_a^\dagger E_b \Pi_{\mathcal{C}(L,\vec{s})},$$

for some $x \neq 0 \in \mathbb{F}_q$ and it follows that (6.5) is fulfilled with $C_{ab} = 0$. The remaining possibility is that $E_a^\dagger E_b \notin S$, and there doesn't exists any element in S that does not commute with $E_a^\dagger E_b$. It follows that $E_a^\dagger E_b$ commutes with the stabilizer, but is not in the stabilizer itself. Hence, it performs a logical operation on the encoded data and equation (6.5) cannot be satisfied. □

To visualize the structure of the correctable error sets $\mathcal{E} \subseteq \mathcal{P}_q^n$, let us first define three quotient groups together with their corresponding transversals (generating sets for the coset decompositions):

- The cosets of $L^\perp$ in $\mathbb{F}_q^{2n}$ ($\mathbb{F}_q^{2n}/L^\perp$). Let a transversal of this decomposition be given by $G = \{\vec{f}_\alpha\}$, i. e. $\vec{f}_\alpha L^\perp \cap \vec{f}_\beta L^\perp = \emptyset$ if $\alpha \neq \beta$ and $\cup \vec{f}_\alpha L^\perp = \mathbb{F}_q^{2n}$. (Note that in the case under consideration $\vec{f}_\alpha L^\perp = \{\vec{f}_\alpha \times \vec{l} \mid \vec{l} \in L^\perp\}$, where the group multiplication rule $\times$ is addition modulo q.) There are $|G| = q^{2n}/q^{n+k} = q^{n-k}$ such cosets.
  [For a specific encoding specified by a hyperbolic basis $\{\vec{g}_1, \ldots, \vec{g}_n, \vec{h}_1 \ldots, \vec{h}_n\}$ we could choose $G = \text{span}\{\vec{h}_1, \ldots, \vec{h}_{n-k}\}$, for instance. Each of these cosets might be labeled unambiguously by a syndrome vector $\vec{s} \in \mathbb{F}_q^{n-k}$ such that $s_i = (\vec{g}_i, \vec{v})_{sp}$, where $\vec{v}$ is an arbitrary vector in the corresponding coset.]

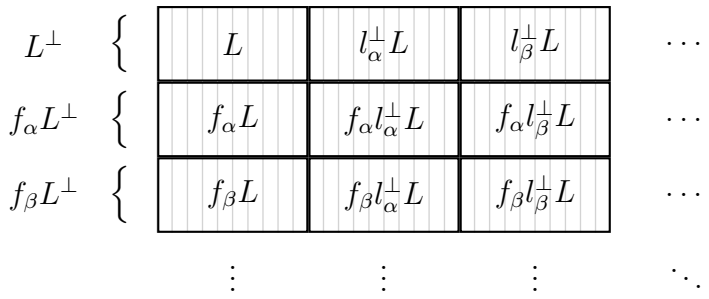


Figure 6.2: All elements of $\mathbb{F}_q^{2n}$ (gray boxes) are arranged in cosets $\mathbb{F}_q^{2n}/L$ (black boxes) generated by some stabilizer L. A corresponding stabilizer code corrects a subset $\mathcal{E} \subseteq \mathbb{F}_q^{2n}$ iff in each row of the diagram no more than one black box is populated by members of $\mathcal{E}$.

- The cosets of L in $L^\perp$ ($L^\perp/L$). Let a corresponding transversal be given by $G^\perp = \{l_\alpha^\perp\}$. There are $|G^\perp| = q^{n+k}/q^{n-k} = q^{2k}$ such cosets.
  [For a specific encoding we could choose $G^\perp = \mathrm{span}\{\vec{g}_{n-k+1}, \ldots, \vec{g}_n, \vec{h}_{n-k+1}, \ldots, \vec{h}_n\}$, for instance. Each of these cosets might be labeled by a logical error vector $\vec{l} = (\vec{l^x}, \vec{l^z}) \in \mathbb{F}_q^{2k}$ such that $l_i^x = (\vec{g}_{i+n-k}, \vec{v})_{sp}$ and $l_i^z = (\vec{v}, \vec{h}_{i+n-k})_{sp}$, where $\vec{v}$ is an arbitrary vector in the corresponding coset.]
- The cosets of L in $\mathbb{F}_q^{2n}$ ($\mathbb{F}_q^{2n}/L$). A corresponding transversal might be obtained by taking the direct product $G \otimes G^\perp$. There are $|G \otimes G^\perp| = q^{n-k}q^{2k} = q^{n+k}$ such cosets.
  [For a specific encoding we could choose $G \otimes G^\perp = \mathrm{span}\{\vec{g}_{n-k+1}, \ldots, \vec{g}_n, \vec{h}_1, \ldots, \vec{h}_n\}$, for instance. Each of these cosets can be labeled by a vector $(\vec{s}, \vec{l}) \in \mathbb{F}_q^{n+k}$.]

Obviously an error set $\mathcal{E} \subseteq \mathcal{P}_q^n$ can equivalently be expressed as a subset $\mathcal{E}_\mathbb{F} \subseteq \mathbb{F}_q^{2n}$ s. t. $\mathcal{E} = \{XZ(\vec{e}) \mid \vec{e} \in \mathcal{E}_\mathbb{F}\}$. In the following we will use the same notation $\mathcal{E}$ for both sets.

Corollary 6.2.5. *Using the coset language, a subset $\mathcal{E} \subseteq \mathbb{F}_q^{2n}$ can be corrected by a stabilizer L, iff in each of the cosets of $L^\perp$ in $\mathbb{F}_q^{2n}$, no more than one of the $\mathbb{F}_q^{2n}/L$-cosets includes elements of $\mathcal{E}$ (compare with figure 6.2).*

Proof. If a $\mathbb{F}_q^{2n}/L$-coset includes some elements $\vec{a}, \vec{b} \in \mathcal{E}$ it follows that $-\vec{a} + \vec{b} \in L$ and the first of the two conditions in lemma 6.2.4 is satisfied. If there is no more than one $\mathbb{F}_q^{2n}/L$-coset populated within a $\mathbb{F}_q^{2n}/L^\perp$-coset, it follows that for all $\vec{a}, \vec{b} \in \mathcal{E}$ s. t. $-\vec{a} + \vec{b} \notin L$, $-\vec{a} + \vec{b} \notin L^\perp$ which is equivalent to $-\vec{a} + \vec{b} \in \mathbb{F}_q^{2n} \setminus L^\perp$, and there exists an element $\vec{g} \in L$ s. t. $(\vec{g}, -\vec{a} + \vec{b})_{sp} \neq 0$ and the second condition in lemma 6.2.4 is satisfied. □

Remark. Since errors in different $\mathbb{F}_q^{2n}/L^\perp$-cosets lead to different syndromes when the stabilizer is measured, and errors in different $\mathbb{F}_q^{2n}/L$-cosets generate different encoded operations, the corollary makes the following intuitive statement: All errors having the same syndrome must act in the same way on the encoded information. Otherwise, knowing the syndrome wouldn't be enough.

Lemma 6.2.6. *A stabilizer code is degenerate if and only if more than one element in the error set $\mathcal{E} \subseteq \mathbb{F}_q^{2n}$ belongs to the same coset of L in $\mathbb{F}_q^{2n}$.*

Proof. If the code corrects $\mathcal{E}$, the condition

$$\Pi_{\mathcal{C}(L,\vec{s})} XZ(\vec{a})^\dagger XZ(\vec{b}) \Pi_{\mathcal{C}(L,\vec{s})} = \Pi_{\mathcal{C}(L,\vec{s})} C_{\vec{a},\vec{b}} \tag{6.19}$$

is satisfied for all $\vec{a}, \vec{b} \in \mathcal{E}$ and for all code spaces $\mathcal{C}(L, \vec{s})$. According to definition 6.1.2, in order to determine whether or not the code is degenerate, we have to determine whether or not $C_{\vec{a},\vec{b}}$ is singular.

We do this by examining the eigenvalues of $C_{\vec{a},\vec{b}}$. Note that each element $\vec{a}$ in $\mathcal{E}$ can be decomposed as in (6.16),

$$\vec{a} = \sum_{i=1}^{n-k} \big(s_i \vec{h}_i + n_i \vec{g}_i\big) + \sum_{i=n-k+1}^{n} \big(l^x_{i-(n-k)} \vec{h}_i + l^z_{i-(n-k)} \vec{g}_i\big), \tag{6.20}$$

and the corresponding Pauli operator $XZ(\vec{a})$ can be written as $\overline{X}^{(\vec{s},\vec{l}^x)}\overline{Z}^{(\vec{n},\vec{l}^z)}$ (see lemma 6.2.3). Let us sort the elements of $\mathcal{E}$ according to their syndrome $\vec{s} = (s_1, \ldots, s_{n-k})$. Then it is clear that $(C_{\vec{a},\vec{b}})$ becomes block-diagonal since $\Pi_{\mathcal{C}(L,\vec{s})} XZ(\vec{a})^\dagger XZ(\vec{b}) \Pi_{\mathcal{C}(L,\vec{s})} = 0$ for $\vec{s}(\vec{a}) \neq \vec{s}(\vec{b})$. We restrict our attention to one of these blocks, i. e. we consider only elements of $\mathcal{E}$ with the same syndrome $\vec{s}$. Corollary 6.2.5 tells us that there is only one coset of L in the coset of $L^\perp$ in $\mathbb{F}_q^{2n}$ characterized by $\vec{s}$ which is populated with members of $\mathcal{E}$. Let us assume first that $\mathcal{E}$ contains all q^{n-k} members of this particular coset of L. The Pauli operators of these members are given by the set $\{\overline{X}^{(\vec{s},\vec{l}^x)}\overline{Z}^{(\vec{n},\vec{l}^z)}\}_{\vec{n}\in\mathbb{F}_q^{n-k}}$ and the matrix elements $c_{\vec{n},\vec{m}}$ of the block are given by

$$\begin{aligned} c_{\vec{n},\vec{m}} \Pi_{\mathcal{C}(L,\vec{s})} &= \Pi_{\mathcal{C}(L,\vec{s})} (\overline{X}^{(\vec{s},\vec{l}^x)}\overline{Z}^{(\vec{n},\vec{l}^z)})^\dagger (\overline{X}^{(\vec{s},\vec{l}^x)}\overline{Z}^{(\vec{m},\vec{l}^z)}) \Pi_{\mathcal{C}(L,\vec{s})} = \\ &= \Pi_{\mathcal{C}(L,\vec{s})} \overline{Z}^{-\vec{n}} \overline{Z}^{\vec{m}} \Pi_{\mathcal{C}(L,\vec{s})} = \omega^{(\vec{m}-\vec{n})\cdot\vec{\mathfrak{s}}} \Pi_{\mathcal{C}(L,\vec{s})}. \end{aligned} \tag{6.21}$$

Using the fact that $\sum_{\vec{m}\in\mathbb{F}_q^{n-k}} \omega^{\vec{m}\cdot\vec{v}}/q^{n-k} = \delta_{\vec{v},\vec{0}}$ we find that the unitary

$$u = \sum_{\vec{i},\vec{n}\in\mathbb{F}_q^{n-k}} u_{\vec{i},\vec{n}} |\vec{i}\rangle\langle\vec{n}| = \sum_{\vec{i},\vec{n}\in\mathbb{F}_q^{n-k}} \omega^{\vec{i}\cdot\vec{n}}/\sqrt{q^{n-k}} |\vec{i}\rangle\langle\vec{n}| \tag{6.22}$$

diagonalizes $c = \sum_{\vec{n},\vec{m}\in\mathbb{F}_q^{n-k}} c_{\vec{n},\vec{m}} |\vec{n}\rangle\langle\vec{m}|$,

$$ucu^\dagger = \sum_{\vec{i},\vec{j}\in\mathbb{F}_q^{n-k}} |\vec{i}\rangle\langle\vec{j}| \sum_{\vec{n},\vec{m}\in\mathbb{F}_q^{n-k}} u_{\vec{i},\vec{n}} c_{\vec{n},\vec{m}} u^\dagger_{\vec{m},\vec{j}} = q^{n-k} |\vec{\mathfrak{s}}\rangle\langle\vec{\mathfrak{s}}|. \tag{6.23}$$

Hence the eigenvalues of the block c are $(q^{n-k}, 0, \ldots, 0)$ which makes the block singular. Inverting the diagonalization leads to $c = q^{n-k}|\psi\rangle\langle\psi|$ with $|\psi\rangle = u^\dagger|\vec{\mathfrak{s}}\rangle = \sum_{\vec{i}\in\mathbb{F}_q^{n-k}} \omega^{-\vec{\mathfrak{s}}\vec{i}}/\sqrt{q^{n-k}}|\vec{i}\rangle$. If the set $\mathcal{E}$ contains only a subset S of the q^{n-k} members of the coset of L, we have to consider the operator $c|_S = \sum_{\vec{n},\vec{m}\in S} c_{\vec{n},\vec{m}} |\vec{n}\rangle\langle\vec{m}|$. Since $c|_S$ can be written as $c|_S = q^{n-k}|\psi_S\rangle\langle\psi_S|$ with $|\psi_S\rangle = \sum_{\vec{i}\in S} \omega^{-\vec{\mathfrak{s}}\vec{i}}/\sqrt{q^{n-k}}|\vec{i}\rangle$, normalization of $|\psi_S\rangle$ leads to $|\tilde{\psi}_S\rangle = \sqrt{q^{n-k}}/\sqrt{|S|} \cdot |\psi_S\rangle$ and we obtain $c|_S = |S||\tilde{\psi}_S\rangle\langle\tilde{\psi}_S|$. Hence the eigenvalues of the block $c|_S$ are $(|S|, 0, \ldots, 0)$ and again the block is singular. The only possibility to obtain a non-singular block is that at most one member of the coset of L in $\mathbb{F}_q^{2n}$ is in $\mathcal{E}$, i. e. $|S| = 1$. Hence, if the code is non-degenerated, $C_{\vec{a},\vec{b}}$ is the identity matrix. □

If we want to correct the set $\mathcal{E} = \{E_i \in \mathcal{P}_q^n | \operatorname{wt}(E_i) \leq t\}$ containing error operators of weight $\leq t$, the stabilizer code has to be at least of distance $d \geq 2t + 1$. Let us first give a simple rule to calculate the distance d of a given stabilizer code.

Corollary 6.2.7. *The distance d of a stabilizer code is the minimum weight[†] of the elements in $L^\perp \setminus L$.*

Proof. It follows from lemma 6.2.4 and definition 6.1.4 that for a stabilizer code of distance d, each error operator $E \in \mathcal{P}_q^n$ of weight less than d is either in S or does not commute with some $M \in S$. This statement is equivalent to each of the following statements and to the corollary itself: Each $\vec{e} \in \mathbb{F}_q^{2n}$ of weight less than d is in $L \cup (\mathbb{F}_q^{2n} \setminus L^\perp)$; In $L^\perp \setminus L$ is no element of weight less than d; □

Now we state a quantum Gilbert Varshamov lower bound on the rate of q-ary stabilizer codes of distance d.

[†]Here the weight of an element $\vec{e} \in \mathbb{F}_q^{2n}$ is defined as the weight of $XZ(\vec{e})$.

Theorem 6.2.8 (Gilbert Varshamov bound for stabilizer codes [FM04]). *Suppose $n > k \geq 2$, $d \geq 2$ and $n = k \pmod 2$. Then there exists a stabilizer code of distance d encoding k qudits into n, provided that*

$$\frac{q^{n-k+2}-1}{q^2-1} > \sum_{i=1}^{d-1}(q^2-1)^{i-1}\binom{n}{i}. \tag{6.24}$$

Since the proof is more sophisticated, we refer to the original work [FM04]. A weaker bound is given in [KKKS06] (and [MU02] for the binary case). Note that the bound found for the binary case in [EM96] and [Got97, chapter 7.1] has been criticized (see e. g. [HNO03]). An asymptotic version of the above bound was known previously [AK01].

Corollary 6.2.9 (Asymptotic GV for stabilizer codes [AK01]). *For large n, there exist stabilizer codes of distance d encoding k qudits into n, such that*

$$\frac{k}{n} \geq 1 - 2H_{q^2[\log_{q^2}]}\Big(1-\frac{d}{n}, \frac{d/n}{q^2-1}, \ldots, \frac{d/n}{q^2-1}\Big). \tag{6.25}$$

Proof. Using the Chernoff bound 1.2.1 (as it was done in proving the asymptotic limit of theorem 5.2.2), this corollary follows from theorem 6.2.8. □

Remark. For qubits ($q = 2$) the asymptotic bound becomes [CRSS97] [Pre98, chapter 7.14]

$$\frac{k}{n} \geq 1 - H_2\Big(\frac{d}{n}\Big) - \frac{d}{n}\log_2 3. \tag{6.26}$$

6.2.4 Recovery Operation

For stabilizer codes Knill and Laflamme's criterion for reversibility of a quantum operation $\mathcal{A}$ on a codespace $\mathcal{C}$ leads to lemma 6.2.4 and corollary 6.2.5, telling us what kind of error subsets $\mathcal{E} \subseteq \mathcal{P}_q^n$ might be corrected by a certain stabilizer code. We are now going to write down the recovery operation which achieves the desired correction of such an error subset.

As discussed in the last subsection, the cosets of $L^\perp$ in $\mathbb{F}_q^{2n}$ can be labeled by a syndrome vector $\vec{s} \in \mathbb{F}_q^{n-k}$ such that $s_i = (\vec{g}_i, \vec{v})_{sp}$, where $\vec{v}$ is an arbitrary member of the corresponding coset. Let us construct a set of coset representatives (a transversal) J_0 by choosing a vector $\vec{J}_0(\vec{s})$ from each coset $\vec{s}$ of $L^\perp$ in $\mathbb{F}_q^{2n}$, $J_0 = \{\vec{J}_0(\vec{s}) | \vec{s} \in \mathbb{F}_q^{n-k}\}$. The error subset $J = J_0 + L \subseteq \mathbb{F}_q^{2n}$[‡] can obviously be corrected: Using figure 6.2, J_0 by construction has the property that each of the rows in the figure contains exactly one of its elements. Now we can easily write down the recovery operation which reverses all quantum operations $\mathcal{A}$ with support on J on the codespace $\mathcal{C}(L, \vec{s}_0)$:

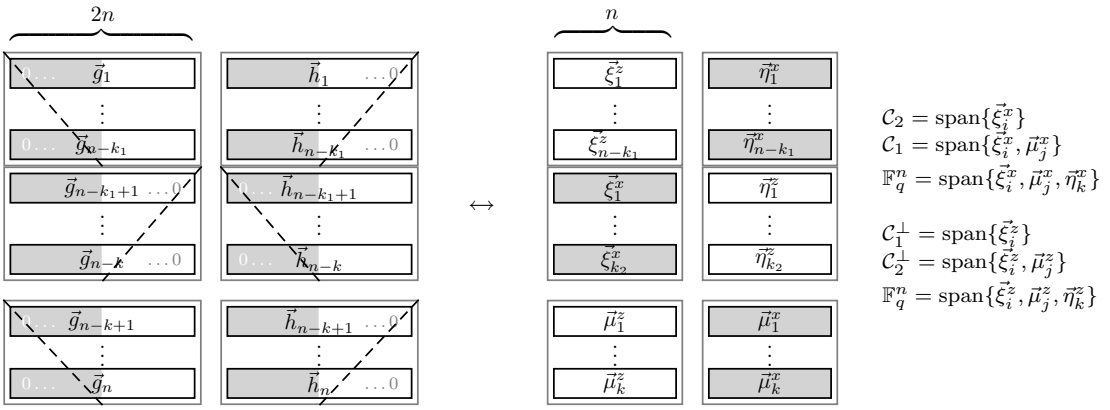
$$\mathcal{R}_{\vec{s}_0}^{(J)}\big(\mathcal{A}(\rho)\big) = \sum_{\vec{t}\in\mathbb{F}_q^{n-k}} XZ^\dagger(\vec{J}_0(\vec{t}))\,\Pi_{\mathcal{C}(L,\vec{s}_0+\vec{t})}\;\mathcal{A}(\rho)\;\Pi_{\mathcal{C}(L,\vec{s}_0+\vec{t})}\,XZ(\vec{J}_0(\vec{t})) = \rho \in \mathcal{S}(\mathcal{C}(L,\vec{s}_0)). \tag{6.27}$$

To generate this tpcp-map, we could first measure the syndrome $\vec{s}_0 + \vec{t}$, thereby projecting onto the codespace $\mathcal{C}(L, \vec{s}_0 + \vec{t})$. Afterwards we apply the Pauli-operator $XZ^\dagger(\vec{J}_0(\vec{t}))$ to go back into the original codespace $\mathcal{C}(L, \vec{s}_0)$ and to undo the remaining logical error. Note that a stabilizer code correcting J_0 is non-degenerate, but becomes degenerate when correcting J.

6.3 CSS Codes

CSS codes are constructed from two classical linear codes $\mathcal{C}_1$ and $\mathcal{C}_2$ such that $\mathcal{C}_2 \subseteq \mathcal{C}_1$. They have been developed independently by Calderbank, Shor and Steane [CS96; Ste96] in 1996. Since CSS codes also form a subclass of stabilizer codes, we will start the description of these codes from this point of view, and establish the connection with the classical codes later on in this section.

[‡] $A + B = \{a + b | a \in A, b \in B\}$



(a) CSS code in stabilizer notation

(b) CSS code in 'CSS-notation'

Figure 6.3: (a) CSS codes form a subclass of stabilizer codes: For the first $n - k_1$ generating elements $\vec{g} = (g_1^x, \ldots, g_n^x, g_1^z, \ldots, g_n^z)$, the x-part of the vector is 0, while for the next k_2 generating elements, the z-part is 0 ($k = k_1 - k_2$). An extension to a hyperbolic basis can be chosen which shows an analogous structure. (b) Since each of the vectors in $\mathbb{F}_q^{2n}$ becomes effectively a vector in $\mathbb{F}_q^n$, we refer to these n-dit vectors as indicated in the figure. Using the definition of a CSS code by the means of two classical codes $\mathcal{C}_2 \subseteq \mathcal{C}_1$, the relations between these codes and the n-dit vectors is shown on the right.

Definition 6.3.1. CSS codes form a subclass of stabilizer codes in which the generating elements of the stabilizer $L = \mathrm{span}\{\vec{g}_1, \ldots, \vec{g}_{n-k}\}$, $\vec{g}_i \in \mathbb{F}_q^{2n}$, have either a vanishing x-part ($\vec{g} = (0, \ldots, 0, g_1^z, \ldots, g_n^z)$, z-type $\vec{g}$) or a vanishing z-part ($\vec{g} = (g_1^x, \ldots, g_n^x, 0, \ldots, 0)$, x-type $\vec{g}$). Setting $k = k_1 - k_2$, we will use the convention that the first $n - k_1$ generating elements are z-type vectors, while the next k_2 generating elements are x-type vectors.

6.3.1 Encoding Operations

As discussed in the last section, an encoding for a stabilizer code L is specified by an extension of the generating elements $\{\vec{g}_1, \ldots, \vec{g}_{n-k}\}$ of L to a hyperbolic basis $\{\vec{g}_1, \ldots, \vec{g}_n, \vec{h}_1, \ldots, \vec{h}_n\}$ spanning $\mathbb{F}_q^{2n}$. The elements of a hyperbolic basis obey relations (6.7), i. e. vanishing symplectic inner products between any two $\vec{g}$'s and any two $\vec{h}$'s, and non-vanishing inter inner product: $(\vec{g}_i, \vec{h}_j)_{sp} = \delta_{ij}$ (compare with figure 6.1). According to their definition, the generating elements of CSS codes are of x-type and z-type only. Considering possible extensions to hyperbolic bases for such codes, it turns out that it is always possible to find extensions which have the same x-type/z-type structure. For example the first $n - k_1$ vectors $\{\vec{h}_1, \ldots, \vec{h}_{n-k_1}\}$ have to be x-type vectors in order to fulfill $(\vec{g}_i, \vec{h}_j)_{sp} = \delta_{ij}$, since the first $n - k_1$ generating elements are z-type vectors. The detailed form of such extensions is shown in figure 6.3a. This means that a CSS code plus an encoding is effectively specified by $2n$ vectors in $\mathbb{F}_q^n$. Each of these n-dit vectors is given a unique notation as indicated in figure 6.3b, e. g. the first $n - k_1$ generating elements $\vec{g}_i \in \mathbb{F}_q^{2n}$ (which are z-type vectors) are denoted as $\vec{\xi}_i^z \in \mathbb{F}_q^n$ now ($\vec{g}_i = (\vec{0}, \vec{\xi}_i^z)$). The basis $\{\vec{g}_1, \ldots, \vec{g}_n;\ \vec{h}_1, \ldots, \vec{h}_n\}$ becomes

$$\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{n-k_1}^z, \vec{\xi}_1^x, \ldots, \vec{\xi}_{k_2}^x, \vec{\mu}_1^z, \ldots, \vec{\mu}_k^z;\ \vec{\eta}_1^x, \ldots, \vec{\eta}_{n-k_1}^x, \vec{\eta}_1^z, \ldots, \vec{\eta}_{k_2}^z, \vec{\mu}_1^x, \ldots, \vec{\mu}_k^x\} \tag{6.28}$$

in the new notation. Since both notations are equivalent, occasionally we will use them simultaneously. The three relations (6.7) a hyperbolic basis has to fulfill, translate into nine relations the n-dit vectors (6.28) have to fulfill. Regarding the n-dit vectors as row-vectors, we can put these nine relations into

one single equation:

$$\begin{pmatrix} \vec{\xi}_1^z \\ \vdots \\ \vec{\eta}_1^z \\ \vdots \\ \vec{\mu}_1^z \\ \vdots \end{pmatrix} \cdot \left((\vec{\eta}_1^x)^T \cdots (\vec{\xi}_1^x)^T \cdots (\vec{\mu}_1^x)^T \right) = \begin{pmatrix} 1\ 0 \cdots & 0\ 0 \cdots & 0\ 0 \cdots \\ 0\ 1 & 0\ 0 & 0\ 0 \\ \vdots\ \ddots & \vdots\ \ddots & \vdots\ \ddots \\ 0\ 0 \cdots & 1\ 0 \cdots & 0\ 0 \cdots \\ 0\ 0 & 0\ 1 & 0\ 0 \\ \vdots\ \ddots & \vdots\ \ddots & \vdots\ \ddots \\ 0\ 0 \cdots & 0\ 0 \cdots & 1\ 0 \cdots \\ 0\ 0 & 0\ 0 & 0\ 1 \\ \vdots\ \ddots & \vdots\ \ddots & \vdots\ \ddots \end{pmatrix}. \tag{6.29}$$

It follows that the two matrices which are multiplied above, cannot be singular. This fact is equivalent to

$$\mathbb{F}_q^n = \mathrm{span}\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{n-k_1}^z, \vec{\eta}_1^z, \ldots, \vec{\eta}_{k_2}^z, \vec{\mu}_1^z, \ldots, \vec{\mu}_k^z\} \tag{6.30a}$$

$$\text{and } \mathbb{F}_q^n = \mathrm{span}\{\vec{\eta}_1^x, \ldots, \vec{\eta}_{n-k_1}^x, \vec{\xi}_1^x, \ldots, \vec{\xi}_{k_2}^x, \vec{\mu}_1^x, \ldots, \vec{\mu}_k^x\}, \tag{6.30b}$$

respectively.

As it is mentioned in the beginning of this section, the original construction of CSS codes makes use of two classical codes $\mathcal{C}_2 \subseteq \mathcal{C}_1$. Let $\mathcal{C}_1$ be an $[n, k_1]_q$ code code encoding k_1 dits into n, and $\mathcal{C}_2$ be an $[n, k_2]_q$ code with $k_2 \leq k_1$. Then the CSS code which is constructed using these classical codes, plus an encoding, is specified by the two lists of vectors,

$$\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{n-k_1}^z\ ,\ \vec{\eta}_1^z, \ldots, \vec{\eta}_{k_2}^z\ ,\ \vec{\mu}_1^z, \ldots, \vec{\mu}_k^z\} \text{ and}$$
$$\{\vec{\eta}_1^x, \ldots, \vec{\eta}_{n-k_1}^x\ ,\ \vec{\xi}_1^x, \ldots, \vec{\xi}_{k_2}^x\ ,\ \vec{\mu}_1^x, \ldots, \vec{\mu}_k^x\},$$

both spanning $\mathbb{F}_q^n$ and satisfying (6.29), where $\mathcal{C}_1^\perp = \mathrm{span}\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{n-k_1}^z\}$ and $\mathcal{C}_2 = \mathrm{span}\{\vec{\xi}_1^x, \ldots, \vec{\xi}_{k_2}^x\}$. It follows that $\mathcal{C}_2^\perp$ has to be spanned by $\{\vec{\xi}_i^z, \vec{\mu}_j^z\}_{i\in\{1\ldots n-k_1\}, j\in\{1\ldots k\}}$, while $\mathcal{C}_1$ has to be spanned by $\{\vec{\xi}_i^x, \vec{\mu}_j^x\}_{i\in\{1\ldots k_2\}, j\in\{1\ldots k\}}$ in order to satisfy (6.29).

Keeping in mind that a CSS code together with a corresponding encoding operation is fully specified by the two sets of n-dit vectors in equation (6.30) and by a set of phases $\{\theta_\alpha(i) \in \{\omega^r | r \in \mathbb{F}_q\}\}_{\alpha\in\{x,z\}, i\in\{1\ldots n\}}$, we are now going to explicitly construct the q^k encoded basis states for all q^{n-k} codespaces using the definition of the encoding operator U_{enc} given in (6.11). First, we have to find the common eigenvector of the set $\{\overline{Z}_i\}_{i\in\{1,\ldots,n\}}$ of encoded Z-operators with eigenvalue list $(\omega^0, \ldots, \omega^0)$. Let us set the phase factors $\theta_z(\cdot)$ and $\theta_x(\cdot)$ equal to one, i. e. we use the encoded operators $\overline{Z}_i = XZ(\vec{g}_i)$ and $\overline{X}_i = XZ(\vec{h}_i)$ for $i \in \{1, \ldots, n\}$. The only X-operators in the set of encoded Z-operators are those constructed from elements spanning $\mathcal{C}_2$. Hence, the state

$$\overline{|0\ldots0\rangle} = \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{\vec{\mathfrak{v}}\in\mathcal{C}_2} |\ \vec{\mathfrak{v}}\ \rangle \tag{6.31}$$

is certainly a common eigenstate of these operators with eigenvalue $+1$. It is also a common $+1$ eigenstate of the Z-operators in $\{\overline{Z}_i\}$, since all these operators are generated by elements of $\mathcal{C}_2^\perp$. Applying the $\{\overline{X}_j = XZ(\vec{h}_j)\}_{j\in\{1,\ldots,n\}}$-operators onto the state $\overline{|0\ldots0\rangle}$ constructs all encoded states:

$$\begin{aligned} \overline{|\vec{x}, \vec{z}, \vec{c}\rangle} &= \overline{X}^{(\vec{x},\vec{z},\vec{c})} \overline{|0\ldots0\rangle} \\ &= \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{\vec{\mathfrak{v}}\in\mathcal{C}_2} \omega^{\vec{\mathfrak{z}}\cdot\vec{\mathfrak{v}}} |\ \vec{\mathfrak{v}} + \vec{\mathfrak{c}} + \vec{\mathfrak{x}}\ \rangle, \end{aligned} \tag{6.32}$$

where the vectors $\vec{\mathfrak{x}}, \vec{\mathfrak{z}}$ and $\vec{\mathfrak{c}}$ are given by

$$\vec{\mathfrak{x}} = \sum_{i=1}^{n-k_1} x_i \vec{\eta}_i^x, \qquad \vec{\mathfrak{z}} = \sum_{i=1}^{k_2} z_i \vec{\eta}_i^z, \qquad \text{and } \vec{\mathfrak{c}} = \sum_{i=1}^{k} c_i \vec{\mu}_i^x. \tag{6.33}$$

The basis of the q^k-dimensional code space $\mathcal{C}(L,\vec{s})$ with syndrome $\vec{s}=(\vec{x},\vec{z})$ is given by the orthonormal set of states $\{\overline{|\vec{x},\vec{z},\vec{c}\rangle}\}_{\vec{c}\in\mathbb{F}_q^k}$.

As it was mentioned in section 6.2.2, any vector $\vec{a}\in\mathbb{F}_q^{2n}$ can be expressed as linear combination of the basis elements of a given hyperbolic basis $\{\vec{g}_1,\dots,\vec{g}_n,\vec{h}_1,\dots,\vec{h}_n\}$ of $\mathbb{F}_q^{2n}$,

$$\begin{aligned}\vec{a}&=(a_1^x,\dots,a_n^x,a_1^z,\dots,a_n^z)=(\vec{a}^x,\vec{a}^z)\\&=\sum_{i=1}^{n-k}\big(s_i\vec{h}_i+n_i\vec{g}_i\big)+\sum_{i=1}^{k}\big(l_i^x\vec{h}_{i+n-k}+l_i^z\vec{g}_{i+n-k}\big),\end{aligned}\tag{6.34}$$

where $s_i=(\vec{a},\vec{g}_i)_{sp}$ etc. Taking into account the special structure of such a basis in the CSS case (i. e. the fact that $\vec{g}_1=(\vec{0},\vec{\xi}_1^z)$ etc.), the x- and z-part of $\vec{a}$ can be decomposed separately,

$$\vec{a}^x=\sum_{i=1}^{n-k_1}s_i^x\vec{\eta}_i^x+\sum_{i=1}^{k_2}n_i^z\vec{\xi}_i^x+\sum_{i=1}^{k}l_i^x\vec{\mu}_i^x\tag{6.35a}$$

$$\vec{a}^z=\sum_{i=1}^{n-k_1}n_i^x\vec{\xi}_i^z+\sum_{i=1}^{k_2}s_i^z\vec{\eta}_i^z+\sum_{i=1}^{k}l_i^z\vec{\mu}_i^z,\tag{6.35b}$$

where $\vec{s}=(\vec{s}^x,\vec{s}^z)$, $\vec{n}=(\vec{n}^x,\vec{n}^z)$ and e. g. $s_1=(\vec{g}_1,\vec{a})_{sp}=\vec{\xi}_1^z\cdot\vec{a}^x=s_1^x$, et cetera. Analogous to lemma 6.2.3, expression (1.25) gives the next lemma.

Lemma 6.3.1. *Any Pauli operator $XZ(\vec{a}\in\mathbb{F}_q^{2n})\in\mathcal{P}_q^n$ can be expressed (up to a phase) as product of some powers of the operators $XZ(\vec{\eta}_i^x,\vec{0}),XZ(\vec{\xi}_j^x,\vec{0}),XZ(\vec{\mu}_l^x,\vec{0})$ and $XZ(\vec{0},\vec{\xi}_i^z),XZ(\vec{0},\vec{\eta}_j^z),XZ(\vec{0},\vec{\mu}_l^z)$,*

$$\begin{aligned}XZ(\vec{a})\sim\prod_{i=1}^{n-k_1}\big(XZ(\vec{\eta}_i^x,\vec{0})^{s_i^x}XZ(\vec{0},\vec{\xi}_i^z)^{n_i^x}\big)\prod_{i=1}^{k_2}\big(XZ(\vec{0},\vec{\eta}_i^z)^{s_i^z}XZ(\vec{\xi}_i^x,\vec{0})^{n_i^z}\big)\\\prod_{i=1}^{k}\big(XZ(\vec{\mu}_i^x,\vec{0})^{l_i^x}XZ(\vec{0},\vec{\mu}_i^z)^{l_i^z}\big),\end{aligned}\tag{6.36}$$

or by using the operators $\overline{Z}_i,\overline{X}_i$ as defined in (6.8),

$$\sim\prod_{i=1}^{n-k_1}(\overline{X}_i^{s_i^x}\overline{Z}_i^{n_i^x})\prod_{i=1}^{k_2}(\overline{X}_{i+n-k_1}^{s_i^z}\overline{Z}_{i+n-k_1}^{n_i^z})\prod_{i=1}^{k}(\overline{X}_{i+n-k}^{l_i^x}\overline{Z}_{i+n-k}^{l_i^z})\quad=\quad\overline{X}^{(\vec{s}^x,\vec{s}^z,\vec{l}^x)}\overline{Z}^{(\vec{n}^x,\vec{n}^z,\vec{l}^z)},\tag{6.37}$$

where the strings $\vec{s}^x,\vec{s}^z,\vec{l}^x$ and $\vec{n}^x,\vec{n}^z,\vec{l}^z$ are defined by (6.35).

6.3.2 Correctable Errors

Corollary 6.3.2. *The distance d of a CSS quantum code constructed from classical codes $\mathcal{C}_2\subseteq\mathcal{C}_1$ is given by*

$$d=\min\{\mathrm{wt}(\vec{c})\mid\vec{c}\in(\mathcal{C}_1\setminus\mathcal{C}_2)\cup(\mathcal{C}_2^\perp\setminus\mathcal{C}_1^\perp)\}.\tag{6.38}$$

Proof. According to corollary 6.2.7, the distance of a stabilizer code is the weight of the lightest element in $L^\perp\setminus L$. As can be seen in figure 6.3b, the weight of the lightest non-zero element in $L^\perp$ is the minimum distance of $\mathcal{C}_1$ and $\mathcal{C}_2^\perp$ since $\mathcal{C}_1=\mathrm{span}\{\vec{\xi}_i^x,\vec{\mu}_j^x\}$, $\mathcal{C}_2^\perp=\mathrm{span}\{\vec{\xi}_i^z,\vec{\mu}_j^z\}$ and $L^\perp=\mathrm{span}\{(\vec{a},\vec{0}),(\vec{0},\vec{b})\mid\vec{a}\in\mathcal{C}_1,\vec{b}\in\mathcal{C}_2^\perp\}$. It remains to subtract $L=\mathrm{span}\{(\vec{a},\vec{0}),(\vec{0},\vec{b})\mid\vec{a}\in\mathcal{C}_2,\vec{b}\in\mathcal{C}_1^\perp\}$. □

Theorem 6.3.3. *There exist CSS codes of distance d encoding k qudits into n such that (for large enough n)*

$$\frac{k}{n}\geq 1-2H_{q[\log_q]}\Big(1-\frac{d}{n},\frac{d/n}{q-1},\dots,\frac{d/n}{q-1}\Big).\tag{6.39}$$

Proof. In chapter C.2, a Gilbert-Varshamov lower bound for self-orthogonal codes is established. It guarantees the existence of $[n, n-\mathsf{k}, d]_q$ codes $\mathcal{C}^\perp$ of rate

$$\frac{n-\mathsf{k}}{n} \geq 1 - H_{q[\log_q]}\Big(1-\frac{d}{n}, \frac{d/n}{q-1}, \ldots, \frac{d/n}{q-1}\Big) \tag{6.40}$$

such that $\mathcal{C} \subseteq \mathcal{C}^\perp$. A CSS-code constructed from such a code encodes $k = k_1 - k_2 = (n-\mathsf{k}) - \mathsf{k}$ qudits into n. Hence its rate is given by (6.39). □

For CSS codes a transversal J_0 for the cosets of $L^\perp$ in $\mathbb{F}_q^{2n}$ can be specified by fixing a transversal Γ_1 of $\mathbb{F}_q^n/\mathcal{C}_1$ and a transversal Γ_2 of $\mathbb{F}_q^n/\mathcal{C}_2^\perp$. Then,

$$J_0 = \{XZ(\vec{a}^x, \vec{a}^z) \mid \vec{a}^x \in \Gamma_1,\ \vec{a}^z \in \Gamma_2\}, \tag{6.41}$$

and the correctable error set $J = J_0 + L$ is given by

$$J = \{XZ(\vec{a}^x, \vec{a}^z) \mid \vec{a}^x \in \Gamma_1 + \mathcal{C}_2,\ \vec{a}^z \in \Gamma_2 + \mathcal{C}_1^\perp\}. \tag{6.42}$$

6.4 Concatenated Codes

If a quantum register corresponding to a certain set of qudits is encoded using a stabilizer code, the resulting qudits may be encoded once more using some other stabilizer code. Equivalently, such a twofold encoding process may be considered as a single one, encoding the initial register only once using a so-called concatenated stabilizer code. We will call the code which is used first the outer code and the code used for the second encoding the inner code[§]. This section examines how such a concatenated code is obtained from its two subcodes.

6.4.1 The Outer Code

The stabilizer code used in a twofold encoding process to encode the qudits before the second encoding is applied is called the outer code. Let the outer code encode K qudits into N and let its stabilizer L^{out} be spanned by $\{\vec{G}_1, \ldots, \vec{G}_{N-K}\}$. As discussed in section 6.2.2, any extension

$$\{\vec{G}_{N-K+1}, \ldots, \vec{G}_N, \vec{H}_1, \ldots, \vec{H}_N\}$$

of the generating elements of L^{out} to a hyperbolic basis of $\mathbb{F}_q^{2N}$ together with a set of phases $\{\Theta_\alpha(i) \in \{\omega^r | r \in \mathbb{F}_q\}\}_{\alpha\in\{x,z\}, i\in\{1\ldots N\}}$ defines a unitary encoding operation U^{out} as follows:

$$U^{\text{out}}|\beta_1 \ldots \beta_N\rangle = \overline{|\beta_1 \ldots \beta_N\rangle}_{\text{out}} = \overline{X}_{\text{out}}^{\vec{\beta}} \overline{|0 \ldots 0\rangle}_{\text{out}}, \tag{6.43}$$

where $\overline{|0\ldots0\rangle}_{\text{out}}$ is defined as the common eigenvector of the operators $\{\overline{Z}_{\text{out},i}\}_{i\in 1\ldots N}$ with all the eigenvalues equal to ω^0, and the $\overline{X}_{\text{out},i}$ and $\overline{Z}_{\text{out},i}$ are defined as

$$\overline{X}_{\text{out}}^{\vec{\beta}} = \prod_i \overline{X}_{\text{out},i}^{\beta_i} \qquad \overline{X}_{\text{out},i} = \Theta_x(i) XZ(\vec{H}_i) \tag{6.44a}$$

$$\overline{Z}_{\text{out}}^{\vec{\beta}} = \prod_i \overline{Z}_{\text{out},i}^{\beta_i} \qquad \overline{Z}_{\text{out},i} = \Theta_z(i) XZ(\vec{G}_i). \tag{6.44b}$$

The encoding operator U^{out} as defined above has the property of mapping the Pauli operators $X^{\vec{u}}$ and $Z^{\vec{v}}$ onto their encoded versions $\overline{X}_{\text{out}}^{\vec{u}}$ and $\overline{Z}_{\text{out}}^{\vec{v}}$ (see (6.13)):

$$U^{\text{out}} X^{\vec{u}} U^{\text{out}\dagger} = \overline{X}_{\text{out}}^{\vec{u}} \qquad U^{\text{out}} Z^{\vec{v}} U^{\text{out}\dagger} = \overline{Z}_{\text{out}}^{\vec{v}}. \tag{6.45}$$

[§]Some authors label the codes the other way round making the first code the inner code and the second code the outer code.

6.4.2 The Inner Code

Imagine we would like to encode the N qudits resulting from the application of U^{out} once more, this time using a stabilizer code encoding k qudits into n. We will call the stabilizer code used for such a second level encoding the inner code. Then the N qudits have to be partitioned into groups of size k (we assume that N is divisible by k), and the encoding operation U^{in} of the inner code has to be applied to all of these groups. Let the stabilizer of the inner code be $L^{\text{in}} = \text{span}\{\vec{g}_1, \ldots, \vec{g}_{n-k}\}$. As it is the case for the outer code, any extension of these vectors to a hyperbolic basis $\{\vec{g}_1, \ldots, \vec{g}_n, \vec{h}_1, \ldots, \vec{h}_n\}$ of $\mathbb{F}_q^{2n}$ together with a set of phases $\{\theta_\alpha(i) \in \{\omega^r | r \in \mathbb{F}_q\}\}_{\alpha\in\{x,z\}, i\in\{1\ldots n\}}$ specifies an encoding operator U^{in}. The set of expressions (6.43), (6.44) and (6.45) applies if the token 'out' is replaced by 'in'.

6.4.3 The Concatenated Code

As a result of such a two step encoding procedure, K qudits have been encoded into $\mathfrak{n} = N/k \times n$. We are interested in the unitary encoder U^{con} of the concatenated code. For given encoding operations U^{out} and U^{in} derived from corresponding hyperbolic bases as described above, can we construct a corresponding hyperbolic basis, let's say $\{\vec{\mathfrak{g}}_1, \ldots, \vec{\mathfrak{g}}_\mathfrak{n}, \vec{\mathfrak{h}}_1, \ldots, \vec{\mathfrak{h}}_\mathfrak{n}\}$, of $\mathbb{F}_q^{2\mathfrak{n}}$ that specifies U^{con} ? Let us denote the initial state of the N qudits which are going to be encoded first by U^{out} by $|\beta_1^1 \ldots \beta_k^1, \ldots, \beta_1^{N/k} \ldots \beta_k^{N/k}\rangle$ and let us label the first group of k qudits by B_1, the second group by B_2, et cetera. Before the inner encoding is applied, additional $n-k$ qudits have to be added to each of the groups B_i. Let us label the $n-k$ qudits added to B_i by A_i and let them be in the state $|\alpha_1^i \ldots \alpha_{n-k}^i\rangle$. Then, the inner encoding operator U^{in} is applied to each of the sets $A_i \cup B_i$ and the total encoding procedure can be viewed as applying the single operator

$$U_{AB}^{\text{con}} = [U_{A_1B_1}^{\text{in}} \otimes \cdots \otimes U_{A_{N/k}B_{N/k}}^{\text{in}}] \cdot U_B^{\text{out}}, \qquad (A = \cup_i A_i,\ B = \cup_j B_j), \tag{6.46}$$

describing the encoding of the concatenated code, to the state

$$|\underbrace{\alpha_1^1 \ldots \alpha_{n-k}^1}_{A_1}, \underbrace{\beta_1^1 \ldots \beta_k^1}_{B_1}; \underbrace{\alpha_1^2 \ldots \alpha_{n-k}^2}_{A_2}, \underbrace{\beta_1^2 \ldots \beta_k^2}_{B_2}; \ldots; \underbrace{\alpha_1^{N/k} \ldots \alpha_{n-k}^{N/k}}_{A_{N/k}}, \underbrace{\beta_1^{N/k} \ldots \beta_k^{N/k}}_{B_{N/k}}\rangle. \tag{6.47}$$

A quantum circuit depicting the situation (for $|\beta_1^1, \ldots, \beta_k^{N/k}\rangle = |0 \ldots 0, \Psi_1, \ldots, \Psi_K\rangle$ and $|\alpha_j^i\rangle = |0\rangle$) is presented in figure 6.4.

We are now going to determine the elements of the hyperbolic basis $\{\vec{\mathfrak{g}}_1, \ldots, \vec{\mathfrak{g}}_\mathfrak{n}, \vec{\mathfrak{h}}_1, \ldots, \vec{\mathfrak{h}}_\mathfrak{n}\}$ of $\mathbb{F}_q^{2\mathfrak{n}}$ that specifies U^{con} by calculating the operators $\{\overline{X}_{\text{con},i} \sim XZ(\vec{\mathfrak{h}}_i), \overline{Z}_{\text{con},i} \sim XZ(\vec{\mathfrak{g}}_i)\}_{i\in 1\ldots\mathfrak{n}}$ using (6.45) and the corresponding expressions for the inner and the concatenated code. Before we proceed, let us define a map with parameter $j \in \{1, \ldots, N/k\}$ mapping a string $\vec{a} = (\vec{a}^x, \vec{a}^z) \in \mathbb{F}_q^{2n}$ to a string $\vec{\mathfrak{a}} = (\vec{\mathfrak{a}}^x, \vec{\mathfrak{a}}^z) \in \mathbb{F}_q^{2\mathfrak{n}}$ by

$$\mathbb{F}_q^{2n} \ni\ \vec{a} \mapsto \vec{a}^{(j)} = \vec{\mathfrak{a}}\ \in \mathbb{F}_q^{2\mathfrak{n}}, \tag{6.48}$$

where $\vec{\mathfrak{a}}^x = (\vec{0}, \ldots, \vec{0}, \vec{a}^x, \vec{0}, \ldots, \vec{0})$ contains $\vec{a}^x$ in position j and $\vec{\mathfrak{a}}^z$ is defined analogously. Let $i \in \{1, \ldots, n-k\}$, $j \in \{1, \ldots, N/k\}$ and let the entries of $\vec{u} \in \mathbb{F}_q^{n-k}$ be given by $u_s = \delta_{s,i}$. Then,

$$\begin{aligned}
\overline{X}_{\text{con},(j-1)(n-k)+i} &= U^{\text{con}}\, X_{A_j}^{\vec{u}}\, U^{\text{con}\dagger} \\
&= [U_{A_1B_1}^{\text{in}} \otimes \ldots] \cdot U_B^{\text{out}}\, X_{A_j}^{\vec{u}}\, U_B^{\text{out}\dagger} \cdot [U_{A_1B_1}^{\text{in}} \otimes \ldots]^\dagger \\
&= [U_{A_1B_1}^{\text{in}} \otimes \ldots]\, X_{A_j}^{\vec{u}}\, [U_{A_1B_1}^{\text{in}} \otimes \ldots]^\dagger \\
&= U_{A_jB_j}^{\text{in}}\, X_{A_j}^{\vec{u}}\, U_{A_jB_j}^{\text{in}\dagger} \otimes \mathcal{I}_{AB\setminus\{A_jB_j\}} \\
&= \theta_x(i) XZ(\vec{h}_i)_{A_jB_j} \otimes \mathcal{I}_{AB\setminus\{A_jB_j\}},
\end{aligned}$$

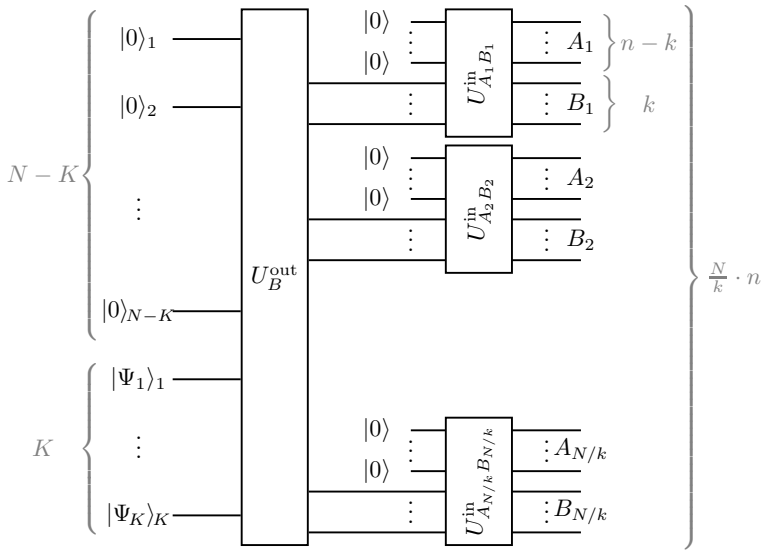


Figure 6.4: Quantum circuit of the encoder of a concatenated quantum code. First the outer code is applied which encodes K data qudits into N qudits after adding the state $|0\rangle^{\otimes N-K}$. Then the inner code encodes k qudits into n after adding $N/k \times (n-k)$ additional qudits prepared as $|0\rangle$ (we assume that N/k is an integer). Altogether the concatenated code encodes K logical qudits into $\mathfrak{n} = N/k \cdot n$ physical qudits.

and essentially the same calculation for $\overline{Z}_{\text{con},(j-1)(n-k)+i}$ leads to the conclusion that

$$\vec{\mathfrak{h}}_{(j-1)(n-k)+i} = \vec{h}_i^{(j)}, \tag{6.49a}$$

$$\vec{\mathfrak{g}}_{(j-1)(n-k)+i} = \vec{g}_i^{(j)}. \tag{6.49b}$$

To determine the remaining $2N$ elements of the hyperbolic basis let $i \in \{1, \ldots, N\}$ and let the entries of $\vec{u} \in \mathbb{F}_q^N$ be given by $u_s = \delta_{s,i}$. Denoting the entries of $\vec{H}_i = (\vec{H}_i^x, \vec{H}_i^z)$ as

$$\begin{aligned}\vec{H}_i^x &= \big((\vec{H}_i^x)_1^1 \ldots (\vec{H}_i^x)_k^1, \ldots, (\vec{H}_i^x)_1^{N/k} \ldots (\vec{H}_i^x)_k^{N/k}\big)\\ \vec{H}_i^z &= \big((\vec{H}_i^z)_1^1 \ldots (\vec{H}_i^z)_k^1, \ldots, (\vec{H}_i^z)_1^{N/k} \ldots (\vec{H}_i^z)_k^{N/k}\big),\end{aligned}$$

we obtain

$$\begin{aligned}\overline{X}_{\text{con},\mathfrak{n}-N+i} &= U^{\text{con}}\, X_B^{\vec{u}}\, U^{\text{con}\dagger}\\ &= [U^{\text{in}}_{A_1B_1} \otimes \ldots] \cdot U_B^{\text{out}}\, X_B^{\vec{u}}\, U_B^{\text{out}\dagger} \cdot [U^{\text{in}}_{A_1B_1} \otimes \ldots]^\dagger\\ &= [U^{\text{in}}_{A_1B_1} \otimes \ldots]\, \Theta_x(i) XZ(\vec{H}_i)_B\, [U^{\text{in}}_{A_1B_1} \otimes \ldots]^\dagger\\ &= \Theta_x(i) \bigotimes_{j=1}^{N/k} U^{\text{in}}_{A_jB_j}\, XZ\big((\vec{H}_i^x)_1^j \ldots (\vec{H}_i^x)_k^j, (\vec{H}_i^z)_1^j \ldots (\vec{H}_i^z)_k^j\big)_{B_j}\, U^{\text{in}\dagger}_{A_jB_j}\\ &= \Theta_x(i) \bigotimes_{j=1}^{N/k} \Big(\prod_{s=1}^{k} \big[\theta_x(n-k+s) XZ(\vec{h}_{n-k+s})\big]^{(\vec{H}_i^x)_s^j} \cdot \big[\theta_z(n-k+s) XZ(\vec{g}_{n-k+s})\big]^{(\vec{H}_i^z)_s^j} \Big)_{A_jB_j}\\ &\sim \bigotimes_{j=1}^{N/k} XZ\Big(\sum_{s=1}^{k} \big((\vec{H}_i^x)_s^j \cdot \vec{h}_{n-k+s} + (\vec{H}_i^z)_s^j \cdot \vec{g}_{n-k+s}\big) \Big)_{A_jB_j},\end{aligned}$$

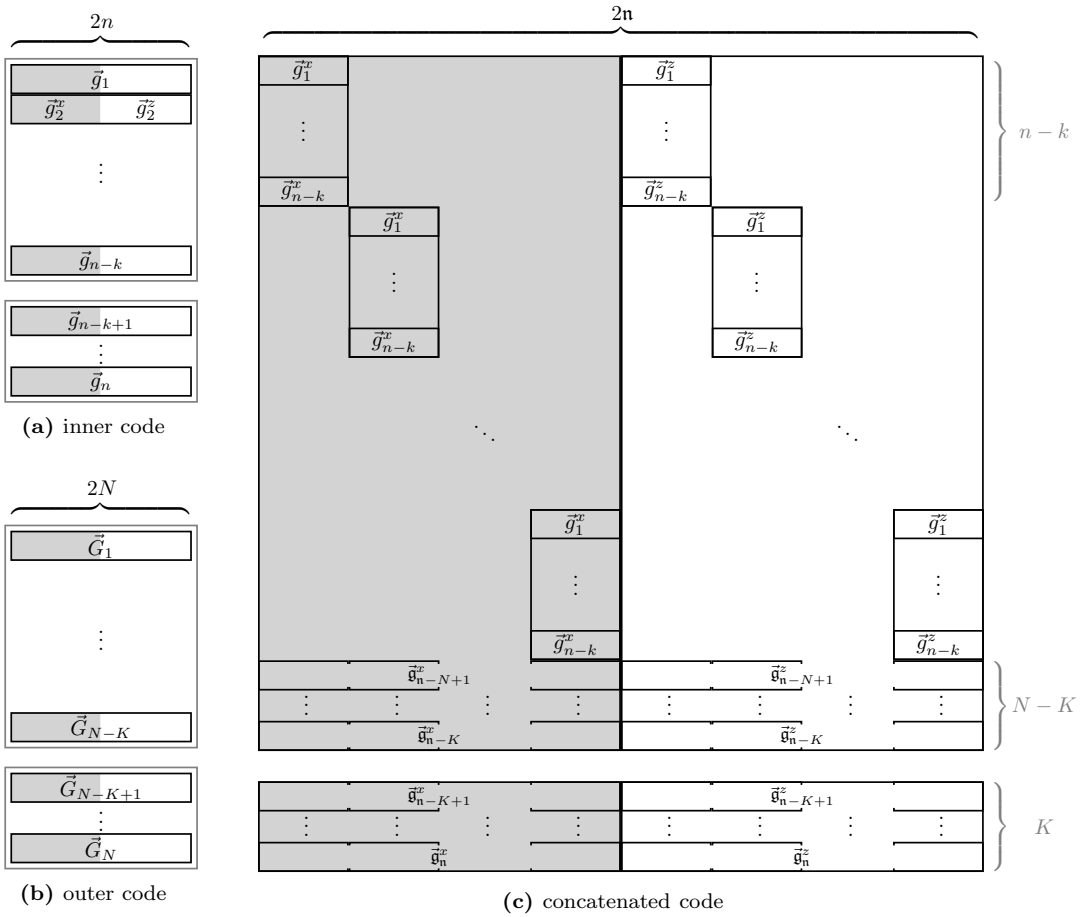


Figure 6.5: An inner $[[n,k]]_q$ code is concatenated with an outer $[[N,K]]_q$ code resulting in an $[[\mathfrak{n},K]]_q$ code with $\mathfrak{n} = N/k \times n$ (we assume N is divisible by k). If the inner code plus an encoding U^{in} is specified by the hyperbolic basis $\{\vec{g}_1,\dots,\vec{g}_n,\vec{h}_1,\dots,\vec{h}_n\}$ *(upper left part)* and the outer code plus an encoding U^{out} is specified by $\{\vec{G}_1,\dots,\vec{G}_N,\vec{H}_1,\dots,\vec{h}_N\}$ *(lower left part)*, the resulting concatenated code plus an encoding U^{con} is specified by $\{\vec{\mathfrak{g}}_1,\dots,\vec{\mathfrak{g}}_{\mathfrak{n}},\vec{\mathfrak{h}}_1,\dots,\vec{\mathfrak{h}}_{\mathfrak{n}}\}$ *(right part)* which is related to the bases of the inner and outer code via equations (6.49) (for $(\vec{\mathfrak{g}}_i,\vec{\mathfrak{h}}_i)$, $i\in\{1,\dots,\mathfrak{n}-N\}$) and (6.50) (for $(\vec{\mathfrak{g}}_i,\vec{\mathfrak{h}}_i)$, $i\in\{\mathfrak{n}-N+1,\dots,\mathfrak{n}\}$). In the figure, only the first half of the bases (i. e. $\{\vec{g}_1,\dots,\vec{g}_n\}$, etc.) is shown.

and again essentially the same calculation for $\overline{Z}_{\text{con},\mathfrak{n}-N+i}$ leads to the conclusion that

$$\vec{\mathfrak{h}}_{\mathfrak{n}-N+i} = \sum_{j=1}^{N/k}\sum_{s=1}^{k}\Big((\vec{H}_i^x)_s^j\cdot\vec{h}_{n-k+s}^{(j)} + (\vec{H}_i^z)_s^j\cdot\vec{g}_{n-k+s}^{(j)}\Big), \tag{6.50a}$$

$$\vec{\mathfrak{g}}_{\mathfrak{n}-N+i} = \sum_{j=1}^{N/k}\sum_{s=1}^{k}\Big((\vec{G}_i^x)_s^j\cdot\vec{h}_{n-k+s}^{(j)} + (\vec{G}_i^z)_s^j\cdot\vec{g}_{n-k+s}^{(j)}\Big). \tag{6.50b}$$

7 Quantum Channel Capacity

Shannon's noisy coding theorem is one of the fundamental theorems of classical information theory. As discussed in chapter 5, it assigns to each channel a non-negative number C, called channel capacity, such that for any rate below the capacity, there exists an error correcting scheme achieving reliable transmission over the channel. The channel capacity is given by the maximum mutual information between source and receiver and the theorem is proven by showing that typical set decoding using random linear codes leads to an arbitrary small decoding error probability. This chapter deals with the quantum analog of Shannon's noisy coding theorem.

It was not until Shor presented a nine qubit quantum error-correcting code in his seminal paper [Sho95], that it was known whether there exist error correction methods for quantum information at all. In the same paper, Shor stated that the ultimate goal would be to find a quantum analog of Shannon's noisy coding theorem, i. e. to define a quantum analog of the Shannon capacity for a quantum channel, and to find encoding schemes which approach this capacity. About a year later the demanded quantum noisy coding theorem was proposed by Lloyd [Llo97]. As it was conjectured by Schumacher and Nielsen [SN96], the role analogous to that played by the mutual information in the classical theory is taken by the regularized coherent information, which corresponds to the limit of the coherent information as the number of channel uses goes to infinity. A rigorous proof that the quantum capacity is upper bounded by the regularized coherent information was given by Barnum, Nielsen and coworkers in [BNS98; BKN00], while the converse part (the capacity is lower bounded by the regularized coherent information) was shown by Shor himself [Sho02] (unpublished) and Devetak [Dev05].

While the Shannon capacity of a classical channel is given by a formula involving a single use of the channel, the quantum capacity involves the limit as the number of channel uses goes to infinity and cannot be expressed by a single letter formula. Therefore, the computation of the quantum capacity for a given quantum channel remains to be a hard problem and is not feasible in general. To obtain at least a lower bound on the quantum capacity, one may calculate the achievable rate of the so-called one-way hashing entanglement distillation protocol by Bennett et al. [BDSW96], which corresponds to a quantum error correcting scheme making use of random stabilizer codes (see [Got97, section 7.6] and [Pre98, section 7.16.2] for the binary case, and [Ham02a] for the general one). The fact that the 'hashing'-rate is indeed only a lower bound on the quantum capacity was shown by Shor and Smolin in [SS96] (and later together with DiVincenzo in [DSS98]). By concatenating an outer random stabilizer code with a deterministic inner one, they found that rates above the hashing rate could be achieved for very noisy depolarizing qubit channels. This result came somewhat as a surprise since it stands in contrast to the classical case where random codes do achieve the capacity of a channel.

In section 7.1 we define the quantum capacity of a noisy quantum channel and present the quantum noisy coding theorem, i. e. the representation of the capacity in terms of the regularized coherent information. For the remaining part of the chapter, we restrict our attention to a certain subclass of quantum channels, so-called memoryless Pauli channels. As it is discussed in section 7.2, this kind of channels are especially easy to analyze and allow us to obtain lower bounds on the capacity of general channels. We present the quantum coding scheme based on random stabilizer codes and corresponding to the one-way hashing protocol in section 7.3. In addition we give a rigorous proof that the hashing-rate can be obtained by using only CSS codes, a result which has been used by Lo in [Lo01] to prove the security of the 6-state quantum key distribution protocol, but for which no elaborated proof can be found in the literature. Concatenation of random codes with deterministic ones [SS96; DSS98] allows for rates surpassing the hashing-rate under certain circumstances. We determine the achievable rate

of such concatenated coding schemes in section 7.4. Eventually we apply the results of the preceeding sections to calculate new lower bounds on the capacity of the qubit depolarizing channel in section 7.5. After giving a detailed description of the deterministic inner code used by [DSS98; SS07], we evaluate the achievable rate for this code for larger code sizes than it was done before in [SS07].

7.1 Quantum Noisy Coding Theorem

A quantum channel is a trace preserving complete positive map (tpcp-map) $\mathcal{M} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ map between density matrices on a Hilbert space $\mathcal{H}$. In this thesis we are primary concerned with discrete Hilbert spaces of dimension q. To send quantum information reliably over a noisy quantum channel, one might protect it by encoding it into a quantum error-correcting code $\mathcal{C} \subset \mathcal{H}^{\otimes n}$, which encodes say k qudits into n. The rate at which we send quantum information in this case would be given by the ratio k/n. To quantify how good the protection works, we may use the minimum pure-state fidelity which is defined for a quantum channel $\mathcal{M}$ and a quantum code $\mathcal{C}$ with corresponding recovery operation $\mathcal{R}$ as

$$F_p(\mathcal{C}, \mathcal{R}\mathcal{M}^{\otimes n}) = \min_{|\psi\rangle \in \mathcal{C}} \langle\psi|\mathcal{R}\big(\mathcal{M}^{\otimes n}(|\psi\rangle\langle\psi|)\big)|\psi\rangle. \tag{7.1}$$

The capacity of a quantum channel for transmitting quantum information was defined by Bennett et al. [BDSW96; BDS97; DSS98] with the help of the minimum pure-state fidelity as follows:

Definition 7.1.1. The quantum capacity $Q(\mathcal{M})$ of a quantum channel $\mathcal{M} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is defined as the maximum number Q, such that for any rate $R < Q$ and any $\varepsilon > 0$, there exists a quantum code $\mathcal{C}$ with rate $k/n \geq R$, together with a recovery operation $\mathcal{R}$, such that

$$F_p(\mathcal{C}, \mathcal{R}\mathcal{M}^{\otimes n}) > 1 - \varepsilon. \tag{7.2}$$

Remark. There exist quite a lot of different definitions for the quantum capacity. For example, the minimum pure-state fidelity might be replaced by the entanglement fidelity [BKN00]. As it turns out, all these definitions are equivalent. For an overview see [KW04]: 'Tema con variazioni: quantum channel capacity'.

The question raised by Shor in his seminal paper on quantum error correction [Sho95] was whether there exits a quantum analog of Shannon's noisy coding theorem relating the quantum capacity of a quantum channel to a quantity corresponding to the mutual information in the classical theory. Such a quantum noisy coding theorem was proposed by Lloyd [Llo97]. The quantity taking the role the mutual information played in the classical case is taken by the coherent information, which is defined for a quantum channel $\mathcal{M} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ and a quantum state $\rho \in \mathcal{S}(\mathcal{H})$ as

$$I_c(\rho, \mathcal{M}) = S(\mathcal{M}(\rho)) - S(\mathcal{M} \otimes \mathcal{I}(|\psi\rangle\langle\psi|)), \tag{7.3}$$

where $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}$ is a purification of ρ.

Theorem 7.1.1 (Quantum noisy coding theorem)**.** *The quantum capacity $Q(\mathcal{M})$ of a quantum channel $\mathcal{M} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is given by the regularized coherent information,*

$$Q(\mathcal{M}) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho} I_c(\rho, \mathcal{M}^{\otimes n}), \tag{7.4}$$

which is obtained by taking the limit as n goes to infinity of $I_c(\rho, \mathcal{M}^{\otimes n})/n$ maximized over all density operators on $\mathcal{H}^{\otimes n}$.

It was proved rigorously by Barnum, Nielsen and coworkers in [BNS98; BKN00], that the regularized coherent information is an upper bound on the capacity $Q(\mathcal{M})$, while the other direction of the theorem ($Q(\mathcal{M})$ is lower bounded by the regularized coherent information) was shown by Shor himself [Sho02, (unpublished)] and Devetak [Dev05].

7.2 Pauli Channels

In this section we consider a special class of tpcp-maps called Pauli channels. Pauli channels have the nice property of being easy to analyze. In addition, any more general channel may be converted into a Pauli channel by a process called discrete twirling. This allows lower bounds on the capacity of Pauli channels to be applicable to more general channels as well.

In the first subsection we give the definition of a Pauli channel. The subsequent subsection explains how a general channel may be twirled to become a Pauli channel.

7.2.1 Definitions

Definition 7.2.1. A Pauli channel $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is a tpcp-map between density operators on a q-dimensional Hilbert space $\mathcal{H}$ given by

$$\mathcal{A} : \rho \mapsto \mathcal{A}(\rho) = \sum_{\vec{e} \in \mathbb{F}_q^2} P_{\mathcal{A}}(\vec{e})\, XZ(\vec{e}) \rho XZ(\vec{e})^\dagger \tag{7.5}$$

for some probability distribution $P_{\mathcal{A}}$ on $\mathbb{F}_q^2$.

If we speak of a memoryless quantum channel, we mean a channel acting identically and independently on multiple qudits. For example, a memoryless Pauli channel $\mathcal{A}^{\otimes n} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$ between density operators on $\mathcal{H}^{\otimes n}$ is given by

$$\mathcal{A}^{\otimes n} : \rho \mapsto \mathcal{A}^{\otimes n}(\rho) = \sum_{\vec{e} \in \mathbb{F}_q^{2n}} P_{\mathcal{A}}^n(\vec{e})\, XZ(\vec{e}) \rho XZ(\vec{e})^\dagger, \tag{7.6}$$

where $P_{\mathcal{A}}^n(\vec{e} = (e_1^x, \ldots, e_n^x, e_1^z, \ldots, e_n^z)) = \prod_{i=1}^n P_{\mathcal{A}}(e_i^x, e_i^z)$. In contrast to (7.6), a general Pauli channel $\mathcal{G} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$ is defined by a probability distribution $P_{\mathcal{G}}$ on $\mathbb{F}_q^{2n}$ which is not necessarily a product distribution.

7.2.2 Discrete Twirling

We follow [Ham03, section 2.3–2.5]. First we note that there is a one-to-one map between a complete positive map $\mathcal{M} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$ and a non-negative operator $\rho_{\mathcal{M}}$ in $\mathcal{S}(\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{\otimes n})$ defined by

$$\rho_{\mathcal{M}} = [\mathcal{I} \otimes \mathcal{M}](|\Phi_{\vec{0}}\rangle\langle\Phi_{\vec{0}}|), \tag{7.7}$$

where $|\Phi_{\vec{0}}\rangle$ denotes a Bell state (compare with definition 1.2.8). Let an operator sum representation of $\mathcal{M}$ be given by $\mathcal{M} : \rho \mapsto \sum_\mu M_\mu \rho M_\mu^\dagger$. Then, by comparing the expressions

$$\begin{aligned} \rho_{\mathcal{M}} &= \sum_{\vec{y},\vec{z} \in \mathbb{F}_q^{2n}} |\Phi_{\vec{y}}\rangle \underbrace{\langle \Phi_{\vec{y}} | \rho_{\mathcal{M}} | \Phi_{\vec{z}} \rangle}_{m_{\vec{y},\vec{z}}} \langle \Phi_{\vec{z}}| \\ &= \frac{1}{q^n} \sum_{\vec{i},\vec{j} \in \mathbb{F}_q^n} |\vec{i}\rangle_A\langle\vec{j}| \otimes \sum_{\vec{y},\vec{z} \in \mathbb{F}_q^{2n}} m_{\vec{y},\vec{z}} XZ(\vec{y})_B |\vec{i}\rangle_B\langle\vec{j}| XZ(\vec{z})_B^\dagger \end{aligned} \tag{7.8}$$

and

$$\rho_{\mathcal{M}} = \frac{1}{q^n} \sum_{\vec{i},\vec{j} \in \mathbb{F}_q^n} |\vec{i}\rangle_A\langle\vec{j}| \otimes \sum_\mu M_\mu |\vec{i}\rangle_B\langle\vec{j}| M_\mu^\dagger, \tag{7.9}$$

it follows that any $\mathcal{M} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$ may be expressed as

$$\mathcal{M} : \rho \mapsto \sum_{\vec{y},\vec{z} \in \mathbb{F}_q^{2n}} m_{\vec{y},\vec{z}} XZ(\vec{y}) \rho XZ(\vec{z})^\dagger, \quad \text{with } m_{\vec{y},\vec{z}} = \langle \Phi_{\vec{y}} | [\mathcal{I} \otimes \mathcal{M}](|\Phi_{\vec{0}}\rangle\langle\Phi_{\vec{0}}|) | \Phi_{\vec{z}} \rangle. \tag{7.10}$$

Discrete twirling ([Ham03], [BBP$^+$97; BDSW96] for the binary case) converts the state

$$\mathcal{S}(\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{\otimes n}) \ni \rho_{\mathcal{M}} = \sum_{\vec{y},\vec{z}\in\mathbb{F}_q^{2n}} m_{\vec{y},\vec{z}}\, |\Phi_{\vec{y}}\rangle\langle\Phi_{\vec{z}}| \tag{7.11}$$

into a Bell diagonal one by applying one of the bilateral rotations $\{XZ(\vec{x})^* \otimes XZ(\vec{x}) | \vec{x} \in \mathbb{F}_q^{2n}\}$ (* denoting complex conjugation) at random,

$$\begin{aligned}
\tilde{\rho}_{\mathcal{M}} &= \frac{1}{q^{2n}} \sum_{\vec{x}\in\mathbb{F}_q^{2n}} \big(XZ(\vec{x})^* \otimes XZ(\vec{x})\big)\rho_{\mathcal{M}}\big(XZ(\vec{x})^* \otimes XZ(\vec{x})\big)^\dagger \\
&= \frac{1}{q^{2n}} \sum_{\vec{x},\vec{y},\vec{z}\in\mathbb{F}_q^{2n}} m_{\vec{y},\vec{z}}\big(\mathcal{I} \otimes XZ(\vec{x})XZ(\vec{y})XZ(\vec{x})^\dagger\big)|\Phi_{\vec{0}}\rangle\langle\Phi_{\vec{0}}|\big(\mathcal{I} \otimes XZ(\vec{x})XZ(\vec{z})XZ(\vec{x})^\dagger\big)^\dagger \qquad (7.12) \\
&= \frac{1}{q^{2n}} \sum_{\vec{y},\vec{z}\in\mathbb{F}_q^{2n}} m_{\vec{y},\vec{z}} \sum_{\vec{x}\in\mathbb{F}_q^{2n}} \omega^{(\vec{x},\vec{y})_{sp}-(\vec{x},\vec{z})_{sp}}\big(\mathcal{I} \otimes XZ(\vec{y})\big)|\Phi_{\vec{0}}\rangle\langle\Phi_{\vec{0}}|\big(\mathcal{I} \otimes XZ(\vec{z})\big)^\dagger \\
&= \sum_{\vec{y}\in\mathbb{F}_q^{2n}} m_{\vec{y},\vec{y}}\, |\Phi_{\vec{y}}\rangle\langle\Phi_{\vec{y}}|. \qquad (7.13)
\end{aligned}$$

To arrive at (7.12) we made use of lemma C.3.1. We obtain from (7.12) that

$$\tilde{\rho}_{\mathcal{M}} = \big[\mathcal{I} \otimes \frac{1}{q^{2n}} \sum_{\vec{x}\in\mathbb{F}_q^{2n}} \mathcal{N}_{\vec{x}}\mathcal{M}\mathcal{N}_{\vec{x}}^\dagger\big](|\Phi_{\vec{0}}\rangle\langle\Phi_{\vec{0}}|), \tag{7.14}$$

with $\mathcal{N}_{\vec{x}} : \rho \mapsto XZ(\vec{x})\rho XZ(\vec{x})^\dagger$, which leads to the central theorem of this subsection.

Theorem 7.2.1. *Any completely positive map $\mathcal{M} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$ can be converted into a general Pauli channel $\tilde{\mathcal{M}} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$ such that*

$$\tilde{\mathcal{M}} : \rho \mapsto \frac{1}{q^{2n}} \sum_{\vec{x}\in\mathbb{F}_q^{2n}} \mathcal{N}_{\vec{x}}\mathcal{M}\mathcal{N}_{\vec{x}}^\dagger(\rho) = \sum_{\vec{e}\in\mathbb{F}_q^{2n}} P_{\mathcal{M}}(\vec{e})\, XZ(\vec{e})\rho XZ(\vec{e})^\dagger, \tag{7.15}$$

with $P_{\mathcal{M}}(\vec{e}) = m_{\vec{e},\vec{e}} = \langle\Phi_{\vec{e}}|[\mathcal{I} \otimes \mathcal{M}](|\Phi_{\vec{0}}\rangle\langle\Phi_{\vec{0}}|)|\Phi_{\vec{e}}\rangle$.

To obtain a lower bound on the quantum capacity of a general memoryless channel $\mathcal{M}^{\otimes n} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$, we apply twirling to convert the channel into the memoryless Pauli channel $\tilde{\mathcal{M}}^{\otimes n} : \mathcal{S}(\mathcal{H}^{\otimes n}) \to \mathcal{S}(\mathcal{H}^{\otimes n})$. Hence, any lower bound for $\tilde{\mathcal{M}}^{\otimes n}$ is automatically a lower bound for $\mathcal{M}^{\otimes n}$.

Remark. Let an operator sum representation of $\mathcal{M} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ be given by $\mathcal{M} : \rho \mapsto \sum_\mu M_\mu \rho M_\mu^\dagger$ with $M_\mu = \sum_{w^x,w^z} a_{\mu,w^x,w^z} XZ(w^x, w^z)$. Then,

$$\begin{aligned}
P_{\mathcal{M}}(\vec{e}) &= \langle\Phi_{\vec{e}}|[\mathcal{I} \otimes \mathcal{M}](|\Phi_{\vec{0}}\rangle\langle\Phi_{\vec{0}}|)|\Phi_{\vec{e}}\rangle \\
&= \sum_\mu |a_{\mu,e^x,e^z}|^2, \qquad (7.16)
\end{aligned}$$

which coincides with the definition of a probability distribution $P_{\mathcal{M}}$ on $\mathbb{F}_q^2$ of a general memoryless channel in [Ham02b, section II].

7.3 Lower Bounds on the Capacity of Memoryless Pauli Channels

A lower bound on the quantum capacity of a binary memoryless Pauli channel was found by Bennett et al. [BBP$^+$96] by constructing the breeding entanglement distillation protocol. Imagine two distant parties, say Alice and Bob, who would like to share a set of maximally entangled states, are connected

only via a noisy quantum channel. If Alice prepares a set of maximally entangled bipartite states and sends Bobs half through the channel, they end up sharing a set of imperfect maximally entangled states. The task of an entanglement distillation protocol is now to distill a smaller set of (nearly) maximally entangled states by means of classical communication and local operations only. Since the breeding protocol has the need for some pre-distilled maximally entangled states, a revised version of this protocol, the so-called one-way hashing protocol, was proposed in [BDSW96]. Both protocols make use of one-way classical communication only and are therefore equivalent [BDSW96] to a scheme where Alice uses a quantum error correcting code to protect Bobs half of the smaller set of perfect states during transmission over the noisy quantum channel.

In the first subsection, the quantum error correcting scheme (generalized to qudits) corresponding to the one-way hashing entanglement distillation protocol is presented. It corresponds to the use of a random stabilizer code (see [Got97, section 7.6] and [Pre98, section 7.16.2] for the binary case, [Ham02a] for the general one). The achievable rate of this scheme is a lower bound on the quantum capacity of the memoryless Pauli channel (the quantum capacity is by definition the highest achievable rate). In the second subsection it is shown that the same result can be achieved using random CSS codes, which is of interest for quantum key distribution since entanglement distillation protocols based on CSS codes are reducible to prepare and measure QKD schemes ([SP00; Ham06], subsection 8.1.2). In fact this result was used by Lo in [Lo01] to prove the security of the 6-state protocol.

7.3.1 Random Stabilizer Codes

In this section we prove the following theorem due to [Got97, section 7.6] and [Pre98, section 7.16.2] (binary case) and [Ham02a] (general case).

Theorem 7.3.1. *Let $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ be a Pauli channel with probability distribution $P_{\mathcal{A}}$ on $\mathbb{F}_q^2$ and let $\varepsilon > 0$. Then, as long as*

$$\frac{k}{n} < 1 - H_{q^2[\log_q]}(P_{\mathcal{A}}), \tag{7.17}$$

and for large enough n, there exists a stabilizer L of dimension $n-k$ such that for any corresponding stabilizer code $\mathcal{C}_{(L,\vec{s})}$, there exists a recovery operation $\mathcal{R}_{(\vec{s})}$ with minimum fidelity

$$F_p\big(\mathcal{C}_{(L,\vec{s})}, \mathcal{R}_{(\vec{s})}\mathcal{A}^{\otimes n}\big) = \min_{|\psi\rangle \in \mathcal{C}_{(L,\vec{s})}} \langle\psi|\mathcal{R}_{(\vec{s})}(\mathcal{A}^{\otimes n}(|\psi\rangle\langle\psi|))|\psi\rangle > 1-\varepsilon. \tag{7.18}$$

Remark. [Ham02a] shows the following stronger result: Let integers n, k and $R \in \mathbb{R}$ satisfy $0 \le k \le Rn$ and $0 \le R < 1$. Then, the minimum fidelity of (7.18) is at least

$$1 - (n+1)^{2(q^2-1)} q^{-nE(R,P_{\mathcal{A}})}, \tag{7.19}$$

where the random coding exponent $E(R, P_{\mathcal{A}})$ stays positive as long as $R < 1 - H_{q^2[\log_q]}(P_{\mathcal{A}})$. The proof of the stronger statement is more elaborate than the simple proof of theorem 7.3.1, which uses typical set decoding as in section 5.4.

Before we start with the proof of theorem 7.3.1, we need the following lemma.

Lemma 7.3.2 (Lemma 6 of [Ham02b])**.** *Let the set of all stabilizers of dimension $n-k$ be given by*

$$\mathsf{A}_{n,k} = \{L \subset F_q^{2n} \mid L \text{ is linear }, L \subseteq L^{\perp}, \dim L = n-k\} \tag{7.20}$$

and let

$$\mathsf{A}_{n,k}(\vec{x}) = \{L \in \mathsf{A}_{n,k} \mid \vec{x} \in L^{\perp} \setminus \{\vec{0}\}\ \}. \tag{7.21}$$

Then, $|\mathsf{A}_{n,k}(\vec{0})| = 0$ and

$$\frac{|\mathsf{A}_{n,k}(\vec{x})|}{|\mathsf{A}_{n,k}|} = \frac{q^{n+k}-1}{q^{2n}-1} \le \frac{1}{q^{n-k}} \tag{7.22}$$

for any nonzero $\vec{x} \in \mathbb{F}_q^{2n}$.

Proof of theorem 7.3.1. For fixed n and k, we pick a stabilizer $L \in \mathsf{A}_{n,k}$ and encode k qudits into one of the codespaces $\mathcal{C}(L, \vec{s})$ labeled by $\vec{s} \in \mathbb{F}_q^{n-k}$. We apply the definition 5.4.1 of a typical set to the random variable X taking on values $(e^x, e^z) \in \mathbb{F}_q^2$ according to the probability distribution $P_{\mathcal{A}}$ on $\mathbb{F}_q^2$:

$$T_\delta^n = \Big\{ \vec{e} = (\vec{e^x}, \vec{e^z}) \in \mathbb{F}_q^{2n} \text{ s. t. for every } (e^x, e^z) \in \mathbb{F}_q^2, \\ \big| N((e^x, e^z)|\vec{e}) - nP_{\mathcal{A}}(e^x, e^z) \big| < \frac{\delta n P_{\mathcal{A}}(e^x, e^z)}{\log_q |\mathbb{F}_q^2|} \Big\}, \quad (7.23)$$

For a given stabilizer L, we construct a transversal $J(L)$ for the cosets of $L^\perp$ in $\mathbb{F}_q^{2n}$ according to the following rule: If a coset contains exactly one typical vector $\vec{e} \in T_\delta^n$, then add this vector to $J(L)$, else pick the corresponding representative at random. A recovery operation $\mathcal{R}_{(J(L),\vec{s})}$ which corrects an error set like $J(L)$ was defined in subsection 6.2.4. As a consequence of $J(L)$ being a transversal, our code will be non-degenerate. The minimum fidelity of our coding scheme,

$$F_p\big(\mathcal{C}_{(L,\vec{s})}, \mathcal{R}_{(J(L),\vec{s})} \mathcal{A}^{\otimes n}\big) = \min_{|\psi\rangle \in \mathcal{C}_{(L,\vec{s})}} \langle \psi | \mathcal{R}_{(J(L),\vec{s})}(\mathcal{A}^{\otimes n}(|\psi\rangle\langle\psi|)) |\psi\rangle, \quad (7.24)$$

will certainly be not less than $\sum_{\vec{e} \in J(L)} P_{\mathcal{A}}^n(\vec{e})$, since the fidelity will be one if $\vec{e} \in J(L)$. In other words,

$$1 - F_p\big(\mathcal{C}_{(L,\vec{s})}, \mathcal{R}_{(J(L),\vec{s})} \mathcal{A}^{\otimes n}\big) \leq \sum_{\vec{e} \notin J(L)} P_{\mathcal{A}}^n(\vec{e}) \\ \leq \sum_{\vec{e} \notin T_\delta^n} P_{\mathcal{A}}^n(\vec{e}) + \sum_{\vec{e} \in T_\delta^n} P_{\mathcal{A}}^n(\vec{e}) \cdot \mathbb{1}\begin{bmatrix} \exists\, \vec{e}' \in T_\delta^n \text{ with } \vec{e}' \neq \vec{e} \text{ s. t.} \\ (\vec{g}_i, \vec{e} - \vec{e}')_{sp} = 0 \text{ for } 1 \leq i \leq n-k \end{bmatrix}. \quad (7.25)$$

The first sum is upper bounded by $\Delta \sim 1/(\delta^2 n)$ (part b of theorem 5.4.1) and the latter by

$$\sum_{\vec{e} \in T_\delta^n} P_{\mathcal{A}}^n(\vec{e}) \sum_{\vec{e}' \in T_\delta^n}^{\text{s. t. } \vec{e}' \neq \vec{e}} \mathbb{1}[(\vec{g}_i, \vec{e} - \vec{e}')_{sp} = 0 \text{ for } 1 \leq i \leq n-k]\,. \quad (7.26)$$

Therefore, averaging over all stabilizers $L \in \mathsf{A}_{n,k}$ leads to

$$\begin{aligned} 1 - \overline{F}_p &\equiv \big\langle 1 - F_p\big(\mathcal{C}_{(L,\vec{s})}, \mathcal{R}_{(J(L),\vec{s})} \mathcal{A}^{\otimes n}\big) \big\rangle_{L \in \mathsf{A}_{n,k}} \\ &\leq \Delta + \sum_{\vec{e} \in T_\delta^n} P_{\mathcal{A}}^n(\vec{e}) \sum_{\vec{e}' \in T_\delta^n}^{\text{s. t. } \vec{e}' \neq \vec{e}} \frac{|\mathsf{A}_{n,k}(\vec{e} - \vec{e}')|}{|\mathsf{A}_{n,k}|} && \text{by theorem 5.4.1b and (7.26)} \\ &\leq \Delta + (|T_\delta^n| - 1) q^{k-n} && \text{by lemma 7.3.2} \\ &\leq \Delta + \exp_q\big(n(H_{q^2[\log_q]}(P_{\mathcal{A}}) + \delta) + k - n\big) && \text{by theorem 5.4.1c.} \end{aligned}$$

This quantity becomes arbitrary small for large enough n as long as

$$\frac{k}{n} < 1 - H_{q^2[\log_q]}(P_{\mathcal{A}}) - \delta. \quad (7.27)$$

Since the above statement holds for any δ, we are free to choose δ as small as we like. So far we have shown that the fidelity $\overline{F}_p$ averaged over all stabilizers is larger than $1 - \varepsilon$. It follows that there exists at least one stabilizer $L \in \mathsf{A}_{n,k}$ such that $F_p\big(\mathcal{C}_{(L,\vec{s})}, \mathcal{R}_{(J(L),\vec{s})} \mathcal{A}^{\otimes n}\big) > 1 - \varepsilon$. □

7.3.2 Random CSS Codes

In this subsection we show that using CSS codes instead of general stabilizer codes is sufficient for theorem 7.3.1 to hold, i. e. we prove the following theorem proposed by Lo in [Lo01] to prove the security of the 6-state quantum key distribution protocol.

Theorem 7.3.3. *Let $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ be a Pauli channel with probability distribution $P_{\mathcal{A}}$ on $\mathbb{F}_q^2$ and let $\varepsilon > 0$. Then, as long as*

$$\frac{k}{n} < 1 - H_{q^2[\log_q]}(P_{\mathcal{A}}), \tag{7.28}$$

and for large enough n, there exists a pair of codes $\mathcal{C}_2 \subset \mathcal{C}_1$ such that for any codespace $\mathcal{C}_{(L(\mathcal{C}_1,\mathcal{C}_2),\vec{s})}$ of the corresponding CSS code with stabilizer $L(\mathcal{C}_1, \mathcal{C}_2)$, there exists a two step recovery operation $\mathcal{R}_{(\vec{s})}$, first correcting the bit errors and then, by using the bit error syndrome to reduce the uncertainty on the phase errors, correcting the phase errors, with minimum fidelity

$$F_p\big(\mathcal{C}_{(L(\mathcal{C}_1,\mathcal{C}_2),\vec{s})}, \mathcal{R}_{(\vec{s})}\mathcal{A}^{\otimes n}\big) = \min_{|\psi\rangle \in \mathcal{C}_{(L(\mathcal{C}_1,\mathcal{C}_2),\vec{s})}} \langle\psi|\mathcal{R}_{(\vec{s})}(\mathcal{A}^{\otimes n}(|\psi\rangle\langle\psi|))|\psi\rangle > 1 - \varepsilon. \tag{7.29}$$

For the proof of theorem 7.3.3 we need the joint- and conditional typical sets from subsection 5.4.2, and corollaries C.1.2 and C.1.5 from appendix C.1.

Proof of theorem 7.3.3. We apply definition 5.4.2 of a set of jointly strongly δ-typical sequences of length n to the two random variables X and Z with joint probability distribution $P_{\mathcal{A}} = \{P_{\mathcal{A}}(x,z) = P_{\mathcal{A}}(z|x) \cdot P_{\mathcal{A}}(x)\}_{x,z\in\mathbb{F}_q}$ and obtain (i) the joint typical set:

$$T_\delta^n(XZ) = \Big\{(\vec{x},\vec{z}) \text{ s. t. for all } x,z \in \mathbb{F}_q, \big|N(xz|\vec{x}\vec{z}) - nP_{\mathcal{A}}(x,z)\big| \leq \frac{\delta n P_{\mathcal{A}}(x,z)}{\log_q |\mathbb{F}_q^4|}\Big\}, \tag{7.30}$$

(ii) the set of typical X-sequences:

$$T_\delta'^n(X) = \{\vec{x} \in \mathbb{F}_q^n \mid (\vec{x},\vec{z}) \in T_\delta^n(XZ) \text{ for some } \vec{z} \in \mathbb{F}_q^n\}, \tag{7.31}$$

and (iii) the conditional typical set of Z-sequences for a given $\vec{x} \in \mathbb{F}_q^n$:

$$T_\delta^n(Z|\vec{x}) = \{\vec{z} \in \mathbb{F}_q^n \mid (\vec{x},\vec{z}) \in T_\delta^n(XZ)\}. \tag{7.32}$$

A CSS code $\mathcal{C}$ encoding $k = k_1 - k_2$ qudits into n is a stabilizer code whose $n-k$ dimensional stabilizer L is constructed from two linear codes $\mathcal{C}_2 \subseteq \mathcal{C}_1$, where $\mathcal{C}_1$ is an $[n,k_1]_q$ code correcting bit errors and $\mathcal{C}_2^\perp$ is an $[n,n-k_2]_q$ code correcting phase errors. Let us fix n, k and an $n-k$ dimensional stabilizer $L \equiv L(\mathcal{C}_1,\mathcal{C}_2)$ and encode k qudits into one of the codespaces $\mathcal{C}(L,\vec{s})$ labeled by $\vec{s} \in \mathbb{F}_q^{n-k}$. A non-degenerate correctable error set $J(L)$ for the CSS-type stabilizer L can be specified by fixing a transversal Γ_1 of $\mathbb{F}_q^n/\mathcal{C}_1$ and a transversal Γ_2 of $\mathbb{F}_q^n/\mathcal{C}_2^\perp$,

$$J(L) = \{XZ(\vec{a}^x,\vec{a}^z) \mid \vec{a}^x \in \Gamma_1,\ \vec{a}^z \in \Gamma_2\}. \tag{7.33}$$

Let us assume now that the actual error of the Pauli channel is in the set $T_\delta^n(XZ)$. We split up the recovery operation for the correctable error set $J(L)$ into two parts. In the first step, we try identify the bit error $\vec{x} \in T_\delta'^n(X)$ by using a typical set decoder: Γ_1 is chosen in such a way that each of its coset representatives is either the only coset member which is in $T_\delta'^n(X)$, or, if there are none or multiple coset members which are in $T_\delta'^n(X)$, it is chosen at random. By measuring the bit error syndrome $\vec{s}^x$ (i. e. by measuring the eigenvalue list of the Pauli operators corresponding to the first $n-k_1$ generating elements of L), we identify a coset of $\mathcal{C}_1$ in $\mathbb{F}_q^n$ and conclude that the actual bit error $\vec{x}$ is the corresponding coset representative in Γ_1. In the next step, we use the information about the bit error $\vec{x}$ to reduce the uncertainty on the remaining phase error $\vec{z}$: Since we know that $\vec{z}$ has to be in $T_\delta^n(Z|\vec{x})$, we apply typical set decoding for the set $T_\delta^n(Z|\vec{x})$ by setting Γ_2 accordingly. The measurement of the phase error syndrome $\vec{s}^z$ (corresponding to the eigenvalue list of the Pauli operators corresponding to the last k_2 generating elements of L), identifies a coset of $\mathcal{C}_2^\perp$ in $\mathbb{F}_q^n$ and we conclude that the actual phase error $\vec{z}$ is the corresponding coset representative in Γ_2. To find a lower bound on the minimum fidelity,

$$F_p\big(\mathcal{C}_{(L,\vec{s})}, \mathcal{R}_{(J(L),\vec{s})}\mathcal{A}^{\otimes n}\big) = \min_{|\psi\rangle \in \mathcal{C}_{(L,\vec{s})}} \langle\psi|\mathcal{R}_{(J(L),\vec{s})}(\mathcal{A}^{\otimes n}(|\psi\rangle\langle\psi|))|\psi\rangle, \tag{7.34}$$

of our coding scheme, we note that the fidelity will certainly be greater or equal than the probability of success of the coding scheme. In other words, one minus the fidelity will be upper bounded by the probability of failure. We proceed by finding an upper bound on the probability of failure. Our scheme fails if (i) the actual error is not within the joint typical set $T^n_\delta(XZ)$, (ii) it is in $T^n_\delta(XZ)$, but bit error correction fails because the measured coset of $\mathcal{C}_1$ in $\mathbb{F}_q^n$ contains multiple coset members which are in $T'^n_\delta(X)$, or (iii) the actual error is in $T^n_\delta(XZ)$, bit error corrections works, but phase error correction fails because the measured coset of $\mathcal{C}_2^\perp$ in $\mathbb{F}_q^n$ contains multiple coset members which are in $T^n_\delta(Z|\vec{x})$. Conditioned on the assumption that the actual error is $(\vec{x},\vec{z}) \in T^n_\delta(XZ)$, bit error correction fails if the following boolean expression is true,

$$F_{\text{bit}} = (\exists \vec{x}' \in T'^n_\delta(X) \text{ with } \vec{x}' \neq \vec{x} \text{ s. t.} H_1(\vec{x}-\vec{x}')^T = \vec{0}^T), \quad (H_1 \text{ parity check matrix of } \mathcal{C}_1), \tag{7.35}$$

and phase error correction fails (assuming that bit error correction succeeded) if

$$F_{\text{phase}} = (\exists \vec{z}' \in T^n_\delta(Z|\vec{x}) \text{ with } \vec{z}' \neq \vec{z} \text{ s. t. } H_2(\vec{z}-\vec{z}')^T = \vec{0}^T), \quad (H_2 \text{ parity check matrix of } \mathcal{C}_2^\perp), \tag{7.36}$$

is true. Using these boolean expressions, we obtain

$$\begin{aligned}
1-F_p &\equiv 1 - F_p\big(\mathcal{C}_{(L,\vec{s})}, \mathcal{R}_{(J(L),\vec{s})}\mathcal{A}^{\otimes n}\big) \\
&\leq \sum_{\vec{e}\notin T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e}) + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e}) \cdot \mathbb{1}[F_{\text{bit}} \vee (\neg F_{\text{bit}} \wedge F_{\text{phase}})] \\
&\leq \Delta + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e}) \cdot \mathbb{1}[F_{\text{bit}} \vee F_{\text{phase}}] \qquad \text{by thm 5.4.2b} \\
&\leq \Delta + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e}) \cdot \big(\mathbb{1}[F_{\text{bit}}] + \mathbb{1}[F_{\text{phase}}]\big) \\
&\leq \Delta + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e}) \cdot \Big(\sum_{\vec{x}'\in T'^n_\delta(X)}^{\vec{x}'\neq\vec{x}} \mathbb{1}\big[H_1(\vec{x}-\vec{x}')^T = \vec{0}^T\big] + \sum_{\vec{z}'\in T^n_\delta(Z|\vec{x})}^{\vec{z}'\neq\vec{z}} \mathbb{1}\big[H_2(\vec{z}-\vec{z}')^T = \vec{0}^T\big]\Big).
\end{aligned}$$

Now we are going to take the average of $1-F_p$ over all code pairs $(\mathcal{C}_1,\mathcal{C}_2)$ which satisfy $\mathcal{C}_2 \subseteq \mathcal{C}_1$. Let us denote by

$$A_{n,k,q} = \{\mathcal{C} \subseteq \mathbb{F}_q^n \mid \mathcal{C} \text{ is a } [n,k]_q\text{-code}\} \tag{7.37}$$

the set of all $[n,k]_q$ codes. Let $\mathcal{K}$ be an $[n,\kappa]_q$ code and let $\vec{c}$ be some nonzero codeword in $\mathbb{F}_q^n$, then we denote by $A_{n,k,q}(\vec{c})$ the set of all $[n,k]_q$ codes which contain $\vec{c}$, and, in an analogous fashion, we denote by $A_{n,k,q}(\mathcal{K})$ and $A_{n,k,q}(\mathcal{K},\vec{c})$ the set of codes which contain $\mathcal{K}$ and $\mathcal{K}\cup\vec{c}$, respectively (see section C.1). We denote the average over all codes by $\big\langle\big\langle\cdot\big\rangle_{\mathcal{C}_2^\perp}\big\rangle_{\mathcal{C}_1}$ using the shorthand notation $\big\langle\cdot\big\rangle_{\mathcal{C}_2^\perp} \equiv \big\langle\cdot\big\rangle_{\mathcal{C}_2^\perp \in A_{n,n-k_2,q}(\mathcal{C}_1^\perp)}$ since we are allowed to average only over those codes $\mathcal{C}_2^\perp$ which include $\mathcal{C}_1^\perp$. With the help of corollaries C.1.2 and C.1.5 we obtain

$$\begin{aligned}
\big\langle\big\langle 1-F_p\big\rangle_{\mathcal{C}_2^\perp}\big\rangle_{\mathcal{C}_1} &\leq \Delta + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e}) \sum_{\vec{x}'\in T'^n_\delta(X)}^{\vec{x}'\neq\vec{x}} \frac{|A_{n,k_1,q}(\vec{x}-\vec{x}')|}{|A_{n,k_1,q}|} + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e}) \sum_{\vec{z}'\in T^n_\delta(Z|\vec{x})}^{\vec{z}'\neq\vec{z}} \Big\langle \frac{|A_{n,n-k_2,q}(\mathcal{C}_1^\perp, \vec{z}-\vec{z}')|}{|A_{n,n-k_2,q}(\mathcal{C}_1^\perp)|}\Big\rangle_{\mathcal{C}_1} \\
&\leq \Delta + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e})\big(|T'^n_\delta(X)|-1\big)q^{-n+k_1} + \sum_{\vec{e}\in T^n_\delta(XZ)} P^n_{\mathcal{A}}(\vec{e})\big(|T^n_\delta(Z|\vec{x})|-1\big)q^{-k_2},
\end{aligned}$$

and by part c of theorem 5.4.2,

$$\leq \Delta + \exp_q\big(n(H_{[\log_q]}(X)+\delta) - n + k_1\big) + \exp_q\big(n(H_{[\log_q]}(Z|X)+2\delta) - k_2\big).$$

This quantity becomes arbitrary small for sufficiently large n, as long as $nH_{[\log_q]}(X) - n + k_1 < 0$ and $nH_{[\log_q]}(Z|X) - k_2 < 0$, which can always be satisfied as long as

$$\frac{k_1-k_2}{n} < 1 - nH_{[\log_q]}(Z|X) - nH_{[\log_q]}(X) = 1 - H_{[\log_q]}(XZ).$$

So far we have shown that the fidelity averaged over all code pairs $(\mathcal{C}_1, \mathcal{C}_2)$ such that $\mathcal{C}_2 \subset \mathcal{C}_1$ is larger than $1-\varepsilon$,

$$\left\langle \left\langle F_p(\mathcal{C}_{(L(\mathcal{C}_1,\mathcal{C}_2),\vec{s})}, \mathcal{R}_{(J(L),\vec{s})}\mathcal{A}^{\otimes n}) \right\rangle_{\mathcal{C}_2^\perp} \right\rangle_{\mathcal{C}_1} > 1-\varepsilon.$$

It follows that there exists at least one pair of codes $(\mathcal{C}_1, \mathcal{C}_2)$ such that $F_p\big(\mathcal{C}_{(L(\mathcal{C}_1,\mathcal{C}_2),\vec{s})}, \mathcal{R}_{(J(L),\vec{s})}\mathcal{A}^{\otimes n}\big)$ is larger than $1-\varepsilon$. □

7.4 Concatenating Random and Deterministic Codes

It was shown by Shor and Smolin in [SS96] (and later together with DiVincenzo in [DSS98]) that the achievable rate for reliable quantum communication over a memoryless Pauli channel $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ using random stabilizer codes (section 7.3) is indeed only a lower bound on the quantum capacity of the channel: By concatenating a certain deterministic inner code with a random outer code, they found that reliable transmission over the depolarizing channel, a special type of Pauli channel characterized by a single noise parameter p, becomes feasible for higher values of noise than allowed by random codes alone. This result is somewhat surprising since in the classical case, random codes do achieve the capacity of discrete memoryless channels.

For a given inner code, concatenated as described above, we determine the achievable rate for reliable quantum communication over a memoryless Pauli channel in subsection 7.4.1 [DSS98; Ham05]. In the subsequent subsection 7.4.2 we show that this rate can be expressed as coherent information of a maximally mixed state in the codespace of the inner code [DSS98; Ham05]. We apply these results to the depolarizing channel using a so-called cat code as inner code in the following section.

7.4.1 Achievable Rate

We are going to determine the achievable rate for reliable quantum communication over a memoryless Pauli channel, when using a concatenated code whose outer code is chosen at random. Let the deterministic inner code be an $[[n,k]]_q$ code with stabilizer $L^{\text{in}} = \{\vec{g}_1, \ldots, \vec{g}_{n-k}\}$, and let an extension to a hyperbolic basis of $\mathbb{F}_q^{2n}$ be given by $\{\vec{g}_{n-k+1}, \ldots, \vec{g}_n, \vec{h}_1, \ldots, \vec{h}_n\}$. By writing a vector $\vec{a} \in \mathbb{F}_q^{2n}$ as linear combination of the basis elements of such a basis,

$$\begin{aligned}\vec{a} &= (a_1^x, \ldots, a_n^x, a_1^z, \ldots, a_n^z) \\ &= \sum_{i=1}^{n-k} \big(s_i \vec{h}_i + n_i \vec{g}_i\big) + \sum_{i=n-k+1}^{n} \big(l^x_{i-(n-k)} \vec{h}_i + l^z_{i-(n-k)} \vec{g}_i\big),\end{aligned}$$

with $s_i = (\vec{g}_i, \vec{a})_{sp}$, $n_i = (\vec{a}, \vec{h}_i)_{sp}$ for $i \in \{1, \ldots, n-k\}$ and $l^x_{i-(n-k)} = (\vec{g}_i, \vec{a})_{sp}$, $l^z_{i-(n-k)} = (\vec{a}, \vec{h}_i)_{sp}$ for $i \in \{n-k+1, \ldots, n\}$, we derived lemma 6.2.3, relating the corresponding Pauli operators:

$$XZ(\vec{a}) \sim \overline{X}^{(\vec{s},\vec{l}^x)} \overline{Z}^{(\vec{n},\vec{l}^z)}. \tag{7.38}$$

This relation allows us to rewrite the action of a memoryless Pauli channel $\mathcal{A}^{\otimes n}$ defined by the probability distribution $P^n_{\mathcal{A}}(\vec{a} = (a_1^x, \ldots, a_n^x, a_1^z, \ldots, a_n^z)) = \prod_{i=1}^n P_{\mathcal{A}}(a_i^x, a_i^z)$, as follows:

$$\begin{aligned}\mathcal{A}^{\otimes n} : \rho \mapsto \mathcal{A}^{\otimes n}(\rho) &= \sum_{\vec{a} \in \mathbb{F}_q^{2n}} P^n_{\mathcal{A}}(\vec{a})\, XZ(\vec{a}) \rho XZ(\vec{a})^\dagger, \\ &= \sum_{\vec{s},\vec{n} \in \mathbb{F}_q^{n-k}} \sum_{\vec{l}^x, \vec{l}^z \in \mathbb{F}_q^k} P_{\mathcal{A}}(\vec{s}, \vec{n}, \vec{l}^x, \vec{l}^z) \left(\overline{X}^{(\vec{s},\vec{l}^x)} \overline{Z}^{(\vec{n},\vec{l}^z)}\right) \rho \left(\overline{X}^{(\vec{s},\vec{l}^x)} \overline{Z}^{(\vec{n},\vec{l}^z)}\right)^\dagger.\end{aligned} \tag{7.39}$$

In addition to $P_{\mathcal{A}}(\vec{s},\vec{n},\vec{l}^x,\vec{l}^z) = P_{\mathcal{A}}^n(\vec{a})$ we define

$$P_{\mathcal{A}}(\vec{s},\vec{l}^x,\vec{l}^z) = \sum_{\vec{n}\in\mathbb{F}_q^{n-k}} P_{\mathcal{A}}(\vec{s},\vec{n},\vec{l}^x,\vec{l}^z), \qquad P_{\mathcal{A}}(\vec{s}) = \sum_{\vec{l}^x,\vec{l}^z\in\mathbb{F}_q^k} P_{\mathcal{A}}(\vec{s},\vec{l}^x,\vec{l}^z), \tag{7.40}$$

and the conditional probability $P_{\mathcal{A}}(\vec{l}^x,\vec{l}^z|\vec{s}) = P_{\mathcal{A}}(\vec{s},\vec{l}^x,\vec{l}^z)/P_{\mathcal{A}}(\vec{s})$. Note that $P_{\mathcal{A}}(\vec{s},\vec{l}^x,\vec{l}^z)$ denotes the probability of having an error in a certain coset of L^{in} in $\mathbb{F}_q^{2n}$, and $P_{\mathcal{A}}(\vec{s})$ denotes the probability of having an error in a certain coset of $L^{\text{in}\perp}$ in $\mathbb{F}_q^{2n}$. Hence, the probability distributions given by $\{P_{\mathcal{A}}(\vec{s},\vec{l}^x,\vec{l}^z)\}$ and $\{P_{\mathcal{A}}(\vec{s})\}$ do not depend on the detailed form of the hyperbolic basis (and therefore on the encoding), but depend only on the stabilizer L^{in} itself.

Let the random outer code be an $[[N,K]]_q$ code as in section 6.4 (with N divisible by k). We encode some K-qudit quantum state within one of the codespaces of the concatenated code and send the resulting $\mathfrak{n} = N/k \times n$ qudits through the Pauli channel $\mathcal{A}^{\otimes N/k\times n}$. The result of a measurement of the first $N/k\times(n-k)$ operators $\overline{Z}_{\text{con},i}$, $i\in\{1,\ldots,N/k\times(n-k)\}$ (which corresponds to a measurement of the N/k syndromes of the inner codes) can be expressed as $\vec{S} = (\vec{s}_1,\ldots,\vec{s}_{N/k})$, $\vec{s}_j \in \mathbb{F}_q^{n-k}$ for $j=1,\ldots,N/k$. Applying definition 5.4.2 of a set of jointly strongly δ-typical sequences of length N/k to the two random variables E and S taking on values $\vec{l} = (\vec{l}^x,\vec{l}^z)\in\mathbb{F}_q^{2k}$ and $\vec{s}\in\mathbb{F}_q^{n-k}$ according to the joint probability distribution $\{P_{\mathcal{A}}(\vec{l},\vec{s})\}$ given by (7.40), we obtain (i) the joint typical set:

$$T_\delta^{N/k}(ES) = \Big\{(\vec{L},\vec{S}) \text{ s.t. for all } \vec{l}\in\mathbb{F}_q^{2k},\vec{s}\in\mathbb{F}_q^{n-k},\ \Big|N(\vec{l}\vec{s}|\vec{L}\vec{S}) - \frac{N}{k}P_{\mathcal{A}}(\vec{l},\vec{s})\Big| \le \frac{\delta N P_{\mathcal{A}}(\vec{l},\vec{s})}{k\log_q|\mathbb{F}_q^{n+k}|}\Big\}, \tag{7.41}$$

with $\vec{L} = (\vec{l}_1,\ldots,\vec{l}_{N/k}) = (\vec{l}_1^x,\vec{l}_1^z,\ldots,\vec{l}_{N/k}^x,\vec{l}_{N/k}^z)\in\mathbb{F}_q^{N/k\times 2k}$, (ii) the set of typical S-sequences:

$$T_\delta'^{N/k}(S) = \{\vec{S}\in\mathbb{F}_q^{N/k\times(n-k)} \mid (\vec{L},\vec{S})\in T_\delta^{N/k}(ES) \text{ for some } \vec{L}\in\mathbb{F}_q^{N/k\times 2k}\}, \tag{7.42}$$

and (iii) the conditional typical set of E-sequences for a given $\vec{S}\in\mathbb{F}_q^{N/k\times(n-k)}$:

$$T_\delta^{N/k}(E|\vec{S}) = \{(\vec{L},\vec{S}) \mid (\vec{L},\vec{S})\in T_\delta^{N/k}(ES)\}. \tag{7.43}$$

Let us assume now that the actual error of $\mathcal{A}^{\otimes N/k\times n}$ is in $T_\delta^{N/k}(ES)$. This assumption is satisfied, since for N/k sufficiently large, the probability of the error being in $T_\delta^{N/k}(ES)$ is larger than $1-\Delta$ for any $\Delta>0$ (part b of theorem 5.4.2). Conditioned on the result $\vec{S}$ of the measurement described above, the situation is equivalent to a scenario where only an $[[N,K]]_q$ code is used to protect against the Pauli channel

$$\mathcal{A}_{\text{eff}} = \bigotimes_{j=1}^{N/k}\mathcal{G}_j, \qquad \text{with } \mathcal{G}_j : \mathcal{S}(\mathcal{H}^{\otimes k}) \to \mathcal{S}(\mathcal{H}^{\otimes k}), \tag{7.44}$$

where $\mathcal{G}_j$ is a general Pauli channel whose probability distribution $P_{\mathcal{G}_j} = \{P_{\mathcal{A}}(\vec{l}|\vec{s}_j)\}$ depends on the value of $\vec{s}_j$ in $\vec{S} = (\vec{s}_1,\ldots,\vec{s}_{N/k})$. Since the errors of $\mathcal{A}_{\text{eff}}$ are known to be in $T_\delta^{N/k}(E|\vec{S})$, we are going to use corresponding typical set decoding. It is known from the proof of theorem 7.3.1 that taking the average over all $[[N,K]]_q$ codes results in an average minimum fidelity which is greater than $1-\varepsilon$ for any $\varepsilon>0$, as long as the exponent of

$$(T_\delta^n(E|\vec{S})-1)\cdot q^{K-N} \le \exp_q\Big(\frac{N}{k}(H_{[\log_q]}(E|S)+2\delta) - (N-K)\Big) \tag{7.45}$$

is negative and N/k is sufficiently large. Since

$$H_{[\log_q]}(E|S) = \sum_{\vec{s}\in\mathbb{F}_q^{n-k}} P_{\mathcal{A}}(\vec{s}) H_{q^{2k}[\log_q]}\big(\{P_{\mathcal{A}}(\vec{l}^x,\vec{l}^z|\vec{s})\}\big), \tag{7.46}$$

we have proven the following theorem due to [DSS98, for $k=1$ and $q=2$] and [Ham05].

Theorem 7.4.1. *Let $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ be a Pauli channel with probability distribution $P_{\mathcal{A}}$ on $\mathbb{F}_q^2$, let some inner $[[n,k]]_q$ code be fixed, and let $\varepsilon > 0$. Then there exists an outer $[[N,K]]_q$ code such that for any codespace of the corresponding concatenated $[[\mathfrak{n},K]]$ code (with $\mathfrak{n} = N/k \cdot n$), there exists a recovery operation with minimum fidelilty larger than $1-\varepsilon$, as long as the total rate $K/\mathfrak{n}$ satisfies*

$$\frac{K}{\mathfrak{n}} < \frac{1}{n}\Big(k - \sum_{\vec{s}\in\mathbb{F}_q^{n-k}} P_{\mathcal{A}}(\vec{s}) H_{q^{2k}[\log_q]}\big(\{P_{\mathcal{A}}(\vec{l}^x,\vec{l}^z|\vec{s})\}\big)\Big), \tag{7.47}$$

where $\{P_{\mathcal{A}}(\vec{s})\}$ and $\{P_{\mathcal{A}}(\vec{l}^x,\vec{l}^z|\vec{s})\}$ are defined by (7.40).

Remark (i). Hamada shows the stronger result that one minus the minimum fidelity is upper bounded by epsilon, where epsilon drops exponentially in N/k as long as condition (7.47) is satisfied [Ham05].

Remark (ii). If the deterministic inner code is a CSS code, we might concatenate it with random outer CSS codes as in theorem 7.3.3. Since the resulting code will also be a CSS code, this means we could achieve the rate in equation (7.47) by using only CSS codes. This result has been used by Lo [Lo01] to improve the security proof of the 6-state protocol: While the standard security proof obtains the maximum tolerable bit error rate from the hashing rate of theorem 7.3.3, Lo used the CSS analog of theorem 7.4.1 to obtain a maximum tolerable bit error rate given by equation (7.47). By using an inner CSS code whose stabilizer consists entirely of Z-type operators, the protocol remains to be reducible to a prepare and measure scheme. (The inner code used by Lo is the so-called cat code which is treated in detail in section 7.5.)

7.4.2 Achievable Rate and Coherent Information

In the preceding subsection we showed that by concatenating a deterministic inner code with a random outer one, we can achieve reliable quantum communication over a memoryless Pauli channel up to a rate given by theorem 7.4.1. We are now going to express this rate in terms of the coherent information of a maximally mixed state defined on one of the codespaces of the inner code. This way a relationship with the quantum capacity $Q(\mathcal{A})$ of a memoryless Pauli channel $\mathcal{A} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is established, which can be expressed as regularized coherent information (7.4),

$$Q(\mathcal{A}) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho} I_c(\rho, \mathcal{A}^{\otimes n}).$$

We prove the following theorem due to [DSS98, for $k=1$ and $q=2$] and [Ham05]:

Theorem 7.4.2. *Concatenation of a random outer $[[N,K]]_q$ code with an inner $[[n,k]]_q$ code with stabilizer L allows for reliable quantum communication over a Pauli channel defined by a probability distribution $P_{\mathcal{A}}$ on $\mathbb{F}_q^2$ as long as the total rate $K/\mathfrak{n}$ satisfies (theorem 7.4.1)*

$$\frac{K}{\mathfrak{n}} < \frac{1}{n}\Big(k - \sum_{\vec{s}\in\mathbb{F}_q^{n-k}} P_{\mathcal{A}}(\vec{s}) H_{q^{2k}[\log_q]}\big(\{P_{\mathcal{A}}(\vec{l}^x,\vec{l}^z|\vec{s})\}\big)\Big). \tag{7.48}$$

This rate can be expressed as the coherent information of a maximally mixed state defined on one of the codespaces $\mathcal{C}(L,\vec{s})$ of the inner code,

$$= \frac{1}{n} I_c\Big(\frac{1}{q^k}\Pi_{\mathcal{C}(L,\vec{s})}, \mathcal{A}^{\otimes n}\Big). \tag{7.49}$$

Proof. Let ρ be a maximally mixed state defined on one of the q^{n-k} codespaces $\mathcal{C}(L,\vec{s})$ of the inner code,

$$\rho = \frac{1}{q^k}\Pi_{\mathcal{C}(L,\vec{s}_0)} = \frac{1}{q^k}\sum_{\vec{c}\in\mathbb{F}_q^k} \overline{|\vec{s}_0,\vec{c}\rangle\langle\vec{s}_0,\vec{c}|},$$

where $\Pi_{\mathcal{C}(L,\vec{s}_0)}$ denotes the projector on codespace $\mathcal{C}(L,\vec{s}_0)$, and let

$$|\psi_{\vec{s}_0}\rangle = \frac{1}{\sqrt{q^k}} \sum_{\vec{c}\in\mathbb{F}_q^k} \overline{|\vec{s}_0,\vec{c}\rangle} \otimes |\vec{c}\rangle$$

be a corresponding purification of ρ. Since the coherent information $I_c(\rho,\mathcal{A}^{\otimes n})$ is defined as the difference between $S(\mathcal{A}^{\otimes n}(\rho))$ and $S\big(\mathcal{A}^{\otimes n}\otimes\mathcal{I}(|\psi_{\vec{s}_0}\rangle\langle\psi_{\vec{s}_0}|)\big)$ in equation (7.3), we proceed by calculating these quantities. We start with the von Neumann entropy of $\mathcal{A}^{\otimes n}(\rho)$: By making use of the channel representation in equation (7.39), we obtain

$$\rho = \frac{1}{q^k}\sum_{\vec{c}\in\mathbb{F}_q^k} \overline{|\vec{s}_0,\vec{c}\rangle\langle\vec{s}_0,\vec{c}|} \mapsto \mathcal{A}^{\otimes n}(\rho) = \frac{1}{q^k}\sum_{\vec{s}\in\mathbb{F}_q^{n-k}} P_{\mathcal{A}}(\vec{s})\Pi_{\mathcal{C}(L,\vec{s})},$$

and eventually

$$S(\mathcal{A}^{\otimes n}(\rho)) = k + H_{q^{n-k}[\log_q]}\big(\{P_{\mathcal{A}}(\vec{s})\}\big). \tag{7.50}$$

To determine the von Neumann entropy resulting from a channel application to a purification of ρ, we use again the channel representation in (7.39) and obtain

$$\mathcal{A}^{\otimes n}\otimes\mathcal{I}(|\psi_{\vec{s}_0}\rangle\langle\psi_{\vec{s}_0}|) = \sum_{\vec{s}\in\mathbb{F}_q^{n-k}}\sum_{\vec{l}=(\vec{l}^x,\vec{l}^z)\in\mathbb{F}_q^{2k}} P_{\mathcal{A}}(\vec{s},\vec{l})\big|\psi^{\vec{l}^z}_{\vec{s}_0+\vec{s},\vec{l}^x}\big\rangle\big\langle\psi^{\vec{l}^z}_{\vec{s}_0+\vec{s},\vec{l}^x}\big|$$

with

$$\big|\psi^{\vec{l}^z}_{\vec{s}_0+\vec{s},\vec{l}^x}\big\rangle = \frac{1}{\sqrt{q^k}}\sum_{\vec{c}\in\mathbb{F}_q^k}\omega^{\vec{c}\cdot\vec{l}^z}\overline{|\vec{s}_0+\vec{s},\vec{c}+\vec{l}^x\rangle}\otimes|\vec{c}\rangle.$$

One can easily check that the set of kets

$$\big\{\big|\psi^{\vec{l}^z}_{\vec{s}_0+\vec{s},\vec{l}^x}\big\rangle \mid \vec{s}\in\mathbb{F}_q^{n-k},\vec{l}=(\vec{l}^x,\vec{l}^z)\in\mathbb{F}_q^{2k}\big\}$$

forms an orthonormal basis of $\mathcal{H}^{\otimes(n+k)}$. Therefore, the von Neumann entropy of $\mathcal{A}^{\otimes n}\otimes\mathcal{I}(|\psi_{\vec{s}_0}\rangle\langle\psi_{\vec{s}_0}|)$ is given by the corresponding Shannon entropy,

$$S\big(\mathcal{A}^{\otimes n}\otimes\mathcal{I}(|\psi_{\vec{s}_0}\rangle\langle\psi_{\vec{s}_0}|)\big) = H_{q^{n+k}[\log_q]}\big(\{P_{\mathcal{A}}(\vec{s},\vec{l})\}\big). \tag{7.51}$$

The proof is finished by subtracting (7.51) from (7.50). □

7.5 Concatenated Codes and the Depolarizing Channel

The depolarizing channel is a special type of Pauli channel which is characterized by a single noise parameter $p\in[0,1]$. It can be interpreted as a quantum channel which transmits a qudit of dimension q unperturbed with probability $1-\tilde{p}=1-pq^2/(q^2-1)$, while exchanging it with the completely mixed state $\mathcal{I}/q$ with probability $\tilde{p}$. By concatenating certain inner codes with random outer codes, it was shown by Shor and Smolin in [SS96] (and later together with DiVincenzo in [DSS98]) that for very noisy depolarizing channels, the hashing rate given by theorem 7.3.1 (representing the achievable rate for reliable quantum communication using random codes alone) can be exceeded by the rate in theorem 7.4.1 (representing the achievable rate using concatenated codes). In this section we present these inner codes and determine the resulting rates. For the depolarizing channel, the maximum amount of noise $p_{\max}$ is defined as the level of noise for which the quantum capacity becomes zero. The best lower bound on $p_{\max}$ of the qubit depolarizing channel known so far was found in [SS07] by using an inner $[[5\times 16,1]]_2$ code. We improve this bound by presenting the results of numerical calculations up to an inner $[[5\times 22,1]]_2$ code.

First, we define the depolarizing channel in subsection 7.5.1. Then, in subsection 7.5.2, we briefly review how the action of a Pauli channel is rewritten for a fixed (inner) code as it was done in subsection 7.4.1. Subsection 7.5.3 presents the so-called cat code, the inner code used in [SS96] and [DSS98]. The succeeding subsection 7.5.4 deals with the concatenated cat code of [DSS98] and [SS07]. This code results from concatenating an outer 'flipped'-type cat code with an inner 'standard' cat code and leads to the best known lower bound on the maximum tolerable noise $p_{\max}$ of the qubit depolarizing channel.

7.5.1 Depolarizing Channel

Definition 7.5.1. The depolarizing channel $\mathcal{D}_p : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is a Pauli channel between density operators on a q-dimensional Hilbert space $\mathcal{H}$, whose probability distribution on $\mathbb{F}_q^2$ is characterized by a single parameter $p \in [0,1]$:

$$\mathcal{D}_p : \rho \mapsto \mathcal{D}_p(\rho) = (1-p)\rho + \sum_{\vec{e}\in\mathbb{F}_q^2}^{\vec{e}\neq(0,0)} \frac{p}{q^2-1}\, XZ(\vec{e})\rho XZ(\vec{e})^\dagger. \tag{7.52}$$

Remark. The depolarizing channel $\mathcal{D}_p$ can be written as

$$\mathcal{D}_{\tilde{p}} : \rho \mapsto \mathcal{D}_{\tilde{p}}(\rho) = (1-\tilde{p})\cdot\rho + \tilde{p}\cdot\frac{1}{q}\mathcal{I}, \tag{7.53}$$

with $\tilde{p} = p\cdot q^2/(q^2-1)$ by using the fact that

$$\frac{1}{q^2}\sum_{\vec{e}\in\mathbb{F}_q^2} XZ(\vec{e})\rho XZ(\vec{e})^\dagger = \frac{1}{q}\mathcal{I}, \tag{7.54}$$

for any normalized $\rho \in \mathcal{S}(\mathcal{H})$.

The highest value of p up to which the quantum capacity of the depolarizing channel $\mathcal{D}_p$ remains non-zero is defined as the channels maximal tolerable level of noise $p_{\max}$,

$$Q(\mathcal{D}_{p_{\max}}) = 0. \tag{7.55}$$

For the qubit depolarizing channel ($q = 2$) we get a lower bound on $p_{\max}$ by calculating the value of p for which the hashing rate of theorems 7.3.1 and 7.3.3 becomes zero. We obtain $p_{\max} > p_{\max}^{\text{hash}} = 18.9290\%$.

While taking the limit as the number of channel uses goes to infinity prevents us from calculating the quantum capacity of the depolarizing channel using the regularized coherent information in equation (7.4),

$$Q(\mathcal{D}_p) = \lim_{n\to\infty}\frac{1}{n}\max_{\rho} I_c(\rho, \mathcal{D}_p^{\otimes n}), \tag{7.56}$$

we are going to calculate the one-shot capacity of the qubit depolarizing channel,

$$Q^{(1)}(\mathcal{D}_p) = \max_{\rho} I_c(\rho, \mathcal{D}_p). \tag{7.57}$$

Lemma 7.5.1. *The one-shot capacity of the qubit depolarizing channel is given by*

$$Q^{(1)}(\mathcal{D}_p) = 1 - H_{4[\log_2]}\big(\{1-p, p/3, p/3, p/3\}\big), \tag{7.58}$$

which equals the hashing rate of theorem 7.3.1.

Proof. The representation of the depolarizing channel in (7.53) shows us that the depolarizing channel does not depend on the basis in which the Pauli operators $XZ(\cdot)$ are defined. Therefore, we can assume without restriction of any kind that the state ρ which maximizes $Q^{(1)}(\mathcal{D}_p)$ is given by $\rho = c|0\rangle\langle 0| + (1-$

$c)|1\rangle\langle 1|$, where $\{|i\rangle\}_{i=0,1}$ is the basis of the qubit Hilbert space which defines the Pauli operators (i. e. $Z|1\rangle = -|1\rangle$ for example). A purification of ρ is given by $|\psi\rangle = \sqrt{c}|0\rangle \otimes |0\rangle + \sqrt{1-c}|1\rangle \otimes |1\rangle$. Now we follow the proof given in [AC97, section V] which shows by a straightforward calculation of

$$f(c,p) = I_c(\rho, \mathcal{D}_p) = S(\mathcal{D}_p(\rho)) - S([\mathcal{D}_p \otimes \mathcal{I}](|\psi\rangle\langle\psi|)) \tag{7.59}$$

that for all values of p the maximum of $f(c,p)$ is obtained for $c = 1/2$. □

7.5.2 Pauli Channel Representation for a CSS Code

In subsection 7.4.1 we determined the achievable rate for reliable quantum communication over a memoryless Pauli channel for a concatenated code whose outer code is chosen at random. We repeat briefly how we rewrote the action of a Pauli channel $\mathcal{A}$ with probability distribution $\{P_{\mathcal{A}}(a^x, a^z)\}$, $(a^x, a^z) \in \mathbb{F}_q^2$, for some given inner $[[n,k]]_q$ code to arrive at a channel with probability distribution $\{P_{\mathcal{A}}(\vec{l}, \vec{s})\}$. Since all inner codes considered in this section are CSS codes, this time we specialize in an inner CSS code.

As discussed in section 6.3.1, an $[[n,k]]_q$ CSS code together with an encoding may be specified by two bases of $\mathbb{F}_q^n$,

$$\mathbb{F}_q^n = \operatorname{span}\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{n-k_1}^z, \vec{\eta}_1^z, \ldots, \vec{\eta}_{k_2}^z, \vec{\mu}_1^z, \ldots, \vec{\mu}_k^z\} \tag{7.60a}$$

$$\text{and } \mathbb{F}_q^n = \operatorname{span}\{\vec{\eta}_1^x, \ldots, \vec{\eta}_{n-k_1}^x, \vec{\xi}_1^x, \ldots, \vec{\xi}_{k_2}^x, \vec{\mu}_1^x, \ldots, \vec{\mu}_k^x\}, \tag{7.60b}$$

with $k = k_1 - k_2$, fulfilling conditions (6.29). By writing the x-component [z-component] of a vector $\vec{a} = (a_1^x, \ldots, a_n^x, a_1^z, \ldots, a_n^z) \in \mathbb{F}_q^{2n}$ as linear combination of the basis elements of the $\{\vec{\xi}_1^z \ldots, \vec{\eta}_1^z \ldots, \vec{\mu}_1^z \ldots\}$ [$\{\vec{\eta}_1^x \ldots, \vec{\xi}_1^x \ldots, \vec{\mu}_1^x \ldots\}$] basis, we obtained (6.35),

$$\vec{a}^x = \sum_{i=1}^{n-k_1} s_i^x \vec{\eta}_i^x + \sum_{j=1}^{k_2} n_j^z \vec{\xi}_j^x + \sum_{r=1}^{k} l_r^x \vec{\mu}_r^x \tag{7.61}$$

$$\vec{a}^z = \sum_{i=1}^{n-k_1} n_i^x \vec{\xi}_i^z + \sum_{j=1}^{k_2} s_j^z \vec{\eta}_j^z + \sum_{r=1}^{k} l_r^z \vec{\mu}_r^z, \tag{7.62}$$

with $s_i^x = \vec{\xi}_i^z \cdot \vec{a}^x$, $n_j^z = \vec{\eta}_j^z \cdot \vec{a}^x$, $l_r^x = \vec{\mu}_r^z \cdot \vec{a}^x$ and $n_i^x = \vec{\eta}_i^x \cdot \vec{a}^z$, $s_j^z = \vec{\xi}_j^x \cdot \vec{a}^z$, $l_r^z = \vec{\mu}_r^x \cdot \vec{a}^z$, which led to lemma 6.3.1, relating the corresponding Pauli operators:

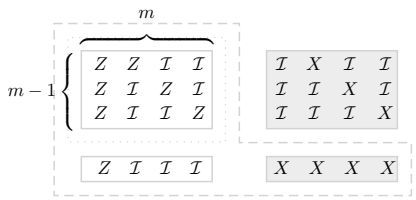
$$XZ(\vec{a}) \sim \overline{X}^{(\vec{s}^x, \vec{s}^z, \vec{l}^x)} \overline{Z}^{(\vec{n}^x, \vec{n}^z, \vec{l}^z)}. \tag{7.63}$$

This relation allows us to rewrite the action of a memoryless Pauli channel $\mathcal{A}^{\otimes n}$ with probability distribution $P_{\mathcal{A}}^n(\vec{a} = (a_1^x, \ldots, a_n^x, a_1^z, \ldots, a_n^z)) = \prod_{i=1}^n P_{\mathcal{A}}(a_i^x, a_i^z)$ as

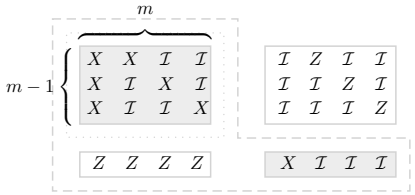
$$\begin{aligned} \mathcal{A}^{\otimes n}(\rho) &= \sum_{\vec{a} \in \mathbb{F}_q^{2n}} P_{\mathcal{A}}^n(\vec{a})\, XZ(\vec{a}) \rho XZ(\vec{a})^\dagger \\ &= \sum_{\vec{s}, \vec{n} \in \mathbb{F}_q^{n-k}} \sum_{\vec{l} \in \mathbb{F}_q^{2k}} P_{\mathcal{A}}(\vec{s}, \vec{n}, \vec{l}) \left(\overline{X}^{(\vec{s}^x, \vec{s}^z, \vec{l}^x)} \overline{Z}^{(\vec{n}^x, \vec{n}^z, \vec{l}^z)}\right) \rho \left(\overline{X}^{(\vec{s}^x, \vec{s}^z, \vec{l}^x)} \overline{Z}^{(\vec{n}^x, \vec{n}^z, \vec{l}^z)}\right)^\dagger, \end{aligned} \tag{7.64}$$

with $\vec{s} = (\vec{s}^x \in \mathbb{F}_q^{n-k_1}, \vec{s}^z \in \mathbb{F}_q^{k_2}) \in \mathbb{F}_q^{n-k}$, $\vec{n} = (\vec{n}^x \in \mathbb{F}_q^{n-k_1}, \vec{n}^z \in \mathbb{F}_q^{k_2}) \in \mathbb{F}_q^{n-k}$ and $\vec{l} = (\vec{l}^x \in \mathbb{F}_q^k, \vec{l}^z \in \mathbb{F}_q^k) \in \mathbb{F}_q^{2k}$. In addition to $P_{\mathcal{A}}(\vec{s}, \vec{n}, \vec{l}) = P_{\mathcal{A}}^n(\vec{a})$ we define the probabilities

$$P_{\mathcal{A}}(\vec{s}, \vec{l}) = \sum_{\vec{n} \in \mathbb{F}_q^{n-k}} P_{\mathcal{A}}(\vec{s}, \vec{n}, \vec{l}), \qquad P_{\mathcal{A}}(\vec{s}) = \sum_{\vec{l} \in \mathbb{F}_q^{2k}} P_{\mathcal{A}}(\vec{s}, \vec{l}), \qquad P_{\mathcal{A}}(\vec{l}|\vec{s}) = P_{\mathcal{A}}(\vec{s}, \vec{l}) / P_{\mathcal{A}}(\vec{s}), \tag{7.65}$$



(a) 'Standard' cat code. The operators on the left hand side are $\{XZ(\vec{0}, \vec{\xi}_i^z)\}$ $(i = 1 \dots m-1$ from top to bottom) and $XZ(\vec{0}, \vec{\mu}^z)$, those on the right hand side are $\{XZ(\vec{\eta}_i^x, \vec{0})\}$ and $XZ(\vec{\mu}^x, \vec{0})$.



(b) 'Flipped' cat code. The operators on the left hand side are $\{XZ(\vec{\xi}_i^x, \vec{0})\}$ $(i = 1 \dots m-1$ from top to bottom) and $XZ(\vec{0}, \vec{\mu}^z)$, those on the right hand side are $\{XZ(\vec{0}, \vec{\eta}_i^z)\}$ and $XZ(\vec{\mu}^x, \vec{0})$.

Figure 7.1: The encoded Pauli operators corresponding to a certain encoding of (a) the 'standard' cat code and (b) the 'flipped' cat code, both encoding one qubit into $m = 4$. Operators corresponding to (generators of) the stabilizer are within the dotted line, operators corresponding to (generators of) the normalizer within the dashed one.

as in equation (7.40). The achievable rate R for reliable quantum communication over a memoryless Pauli channel characterized by $\{P_{\mathcal{A}}(a^x, a^z)\}$, $(a^x, a^z) \in \mathbb{F}_q^2$, is given by theorem 7.4.1:

$$R = \frac{1}{n}\Big(k - \sum_{\vec{s} \in \mathbb{F}_q^{n-k}} P_{\mathcal{A}}(\vec{s}) H_{q^{2k}[\log_q]}\big(\{P_{\mathcal{A}}(\vec{l}|\vec{s})\}\big)\Big). \tag{7.66}$$

To determine R for the inner $[[n, k,]]_q$ code specified by (7.60), we obviously have to know the corresponding probability distributions $\{P_{\mathcal{A}}(\vec{s})\}$ and $\{P_{\mathcal{A}}(\vec{l}|\vec{s})\}$. In the following subsections we determine these distributions for various inner codes.

7.5.3 The Cat Code

The cat code used in [SS96; DSS98] is an $[[m, k = 1]]_2$ CSS code with $k = k_1 = 1$ and $k_2 = 0$. It is specified by a classical code $\mathcal{C}_1^{\perp} = \mathrm{span}\{\vec{\xi}_1^z, \dots, \vec{\xi}_{m-k_1}^z\}$, where the entries of the vector $\vec{\xi}_i^z$ are given by $(\vec{\xi}_i^z)_j = \delta_{j,1} + \delta_{j,i+1}$ for $j \in \{1, \dots, m\}$. The corresponding stabilizer is $L = \mathrm{span}\{(\vec{0}, \vec{\xi}_1^z), \dots, (\vec{0}, \vec{\xi}_{m-1}^z)\}$. We consider an extension to the bases

$$\begin{aligned} \mathbb{F}_2^n &= \{\vec{\xi}_1^z, \dots, \vec{\xi}_{m-1}^z\,,\, \vec{\mu}^z\} \\ \text{and } \mathbb{F}_2^n &= \{\vec{\eta}_1^x, \dots, \vec{\eta}_{m-1}^x\,,\, \vec{\mu}^x\} \end{aligned} \tag{7.67}$$

as shown in figure 7.1(a). To construct the encoding U_{enc} associated with these extensions, we set the phase factors $\theta_z(\cdot)$ and $\theta_x(\cdot)$ equal to one as it was done in subsection 6.3.1. Then, the corresponding encoded states of equation (6.32) become

$$\overline{|\vec{s}^x, l^x\rangle} = \overline{X}^{(\vec{s}^x, l^x)} \overline{|0 \dots 0\rangle} = \Big|\; l^x \cdot \vec{\mu}^x + \sum_{i=1}^{m-1} s_i^x \cdot \vec{\eta}_i^x \;\Big\rangle. \tag{7.68}$$

The code is called cat code because a pure one qubit state $\alpha|0\rangle + \beta|1\rangle$ encoded in the codespace $\mathcal{C}(L, \vec{0})$ becomes a cat state,

$$U_{\mathrm{enc}}|\vec{0}\rangle \otimes \big(\alpha|0\rangle + \beta|1\rangle\big) = \alpha|0, \dots, 0, 0\rangle + \beta|1, \dots, 1, 1\rangle. \tag{7.69}$$

We do not calculate the probability distributions $\{P_{\mathcal{A}}(\vec{s})\}$ and $\{P_{\mathcal{A}}(\vec{l}|\vec{s})\}$ for the cat code, since they emerge as a special case of the corresponding distributions of the concatenated cat code in subsection 7.5.4 (see equation (7.86)).

The cat code improves the hashing rate lower bound $p_{\max}^{\mathrm{hash}} = 18.9290\%$ on the maximum tolerable level of noise $p_{\max}$ of the qubit depolarizing channel. By setting the rate $R_m(p)$ of (7.66) for a cat

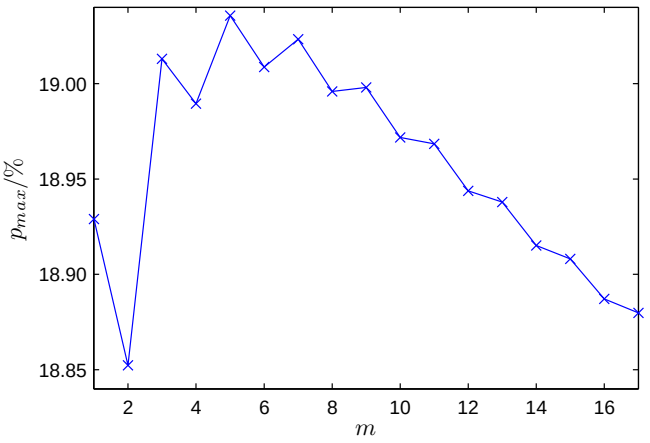


Figure 7.2: The maximum tolerable value of noise p for the qubit depolarizing channel $\mathcal{D}_p$ as a function of the size of the inner $[[m,1]]_2$ cat code. The highest value $p^{\text{cat}}_{\max}(m=5) = 19.0356\%$ is obtained for $m=5$.

code of size m equal to zero, we obtain the values $p^{\text{cat}}_{\max}(m)$ shown in figure 7.2. The highest value (and therefore the best lower bound on $p_{\max}$) is obtained for $m=5$, $p^{\text{cat}}_{\max}(m=5) = 19.0356\%$. The rates $R_m(p)$ for $m=3$ and $m=5$ are shown in figure 7.5 *(blue)*.

Remark. As discussed in the remark following theorem 7.4.1, the value $p^{\text{cat}}_{\max}(m=5)$ was used by Lo in [Lo01] to improve the security of 6-state quantum key distribution protocol. Since the 6-state protocol corresponds to a qubit depolarizing channel $\mathcal{D}_{\frac{3}{2}p}$, he improved the maximum tolerable bit error rate of the 6-state protocol from $\frac{2}{3}\cdot p^{\text{hash}}_{\max} = 12.6193\%$ to $\frac{2}{3}\cdot p^{\text{cat}}_{\max}(m=5) = 12.6904\%$.

The concatenated cat code presented in subsection 7.5.4 is obtained by concatenating an inner cat code with an outer 'flipped' version of the cat code. We proceed by presenting this 'flipped' cat code, whose stabilizer is obtained from the stabilizer of the 'standard' cat code described above by exchanging the Z operators with X operators.

'Flipped' Cat Code

The 'flipped' cat code is an $[[m, k=1]]_2$ CSS code with $k_1 = m$, $k_2 = m-1$ and $k = k_1 - k_2 = 1$. It is specified by a classical code $\mathcal{C}_2 = \text{span}\{\vec{\xi}^x_1, \ldots, \vec{\xi}^x_{k_2}\}$, where the entries of the vector $\vec{\xi}^x_i$ are given by $(\vec{\xi}^x_i)_j = \delta_{j,1} + \delta_{j,i+1}$ for $j \in \{1,\ldots,m\}$. The corresponding stabilizer is $L = \text{span}\{(\vec{\xi}^x_1, \vec{0}), \ldots, (\vec{\xi}^x_{m-1}, \vec{0})\}$. We consider an extension to the bases

$$\begin{aligned} \mathbb{F}_2^n &= \{\vec{\eta}^z_1, \ldots, \vec{\eta}^z_{m-1}\,,\, \vec{\mu}^z\} \\ \text{and } \mathbb{F}_2^n &= \{\vec{\xi}^x_1, \ldots, \vec{\xi}^x_{m-1}\,,\, \vec{\mu}^x\} \end{aligned} \tag{7.70}$$

as shown in figure 7.1(b). To construct the encoding U_{enc} associated with these extensions, we set the phase factors $\theta_z(\cdot)$ and $\theta_x(\cdot)$ equal to one as it was done in subsection 6.3.1. Then, the corresponding encoded states of equation (6.32) become

$$\overline{|\vec{s}^z, \vec{l}^x\rangle} = \overline{X}^{(\vec{s}^z, \vec{l}^x)}\overline{|0\ldots0\rangle} = \frac{1}{\sqrt{|\mathcal{C}_2|}} \sum_{\vec{\mathfrak{v}}\in\mathcal{C}_2} \omega^{\vec{\mathfrak{z}}\cdot\vec{\mathfrak{v}}} |\, \vec{\mathfrak{v}} + l^x \cdot \vec{\mu}^x \,\rangle, \text{ with } \vec{\mathfrak{z}} = \sum_{i=1}^{m-1} s^z_i \vec{\eta}^z_i. \tag{7.71}$$

7.5.4 The Concatenated Cat Code

By concatenating an outer $[[m_2,1]]_q$ 'flipped' cat code with an inner $[[m_1,1]]_2$ 'standard' cat code, we obtain the $[[m_1 \times m_2, 1]]_2$ code used in [DSS98; SS07] and shown in figure 7.3. The $[[m_1 \times m_2, 1]]_2$

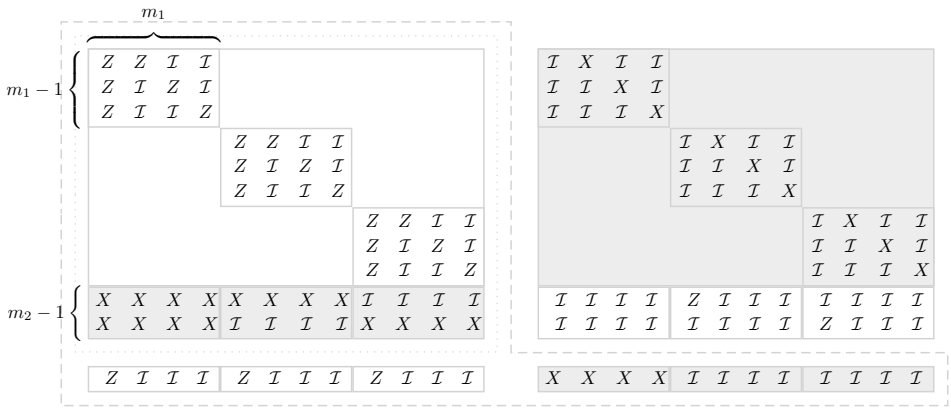


Figure 7.3: The encoded Pauli operators corresponding to a hyperbolic basis of the concatenated cat code. Here $m_1 = 4$ and $m_2 = 3$, so that one qubit is encoded into $n = m_1 \times m_2$. The first $m_2 \times (m_1 - 1)$ operators on the left hand side are the $\{XZ(\vec{0}, \vec{\xi}_i^z)\}$, the next $m_2 - 1$ the $\{XZ(\vec{\xi}_i^x, \vec{0})\}$ and the last one is $XZ(\vec{0}, \vec{\mu}^z)$. Those on the right hand side are $\{XZ(\vec{\eta}_i^x, \vec{0})\}$, $\{XZ(\vec{0}, \vec{\eta}_i^z)\}$ and $XZ(\vec{\mu}^x, \vec{0})$ accordingly.

code is a CSS code with parameters $n = m_1 m_2$, $k_1 = m_2$, $k_2 = m_2 - 1$ and $k = k_1 - k_2 = 1$. We proceed by calculating the corresponding probability distributions $\{P_{\mathcal{A}}(\vec{s})\}$ and $\{P_{\mathcal{A}}(\vec{l}|\vec{s})\}$ which allow us to evaluate the achievable transmission rate given in equation (7.66).

Joint Probabilities of Logical Errors and Syndrome

We are going to calculate the joint probabilities $\{P_{\mathcal{A}}(\vec{l}, \vec{s})\}$ with $\vec{l} \in \mathbb{F}_2^2$ and $\vec{s} \in \mathbb{F}_2^{m_1 m_2 - 1}$ defined in equation (7.65) for the $[[m_1 \times m_2, 1]]_2$ code described above. From $\{P_{\mathcal{A}}(\vec{l}, \vec{s})\}$ we will obtain $\{P_{\mathcal{A}}(\vec{s})\}$ by summation over $\vec{l}$ and $\{P_{\mathcal{A}}(\vec{l}|\vec{s})\}$ by $P_{\mathcal{A}}(\vec{l}|\vec{s}) = P_{\mathcal{A}}(\vec{l}, \vec{s})/P_{\mathcal{A}}(\vec{s})$. Neither the details of these calculations nor formulas expressing the resulting probabilities have been presented in the literature [DSS98; SS07].

We denote the elements of the probability distribution $\{P_{\mathcal{A}}(a^x, a^z)\}$, $(a^x, a^z) \in \mathbb{F}_2^2$, of the qubit Pauli channel $\mathcal{A}$ as $\{p_e, p_x, p_y, p_z\}$, i. e.

$$\mathcal{A}(\rho) = p_e \rho + p_x X \rho X^\dagger + p_y (XZ)\rho(XZ)^\dagger + p_z Z \rho Z^\dagger. \tag{7.72}$$

Then, by definition, $P_{\mathcal{A}}(\vec{s}, \vec{n}, \vec{l}) = P_{\mathcal{A}}^n(\vec{a})$, where $\vec{a} \in \mathbb{F}_2^{2n}$ depends on $\vec{s}, \vec{n}, \vec{l}$ via the bases decomposition given in equation (7.61). $P_{\mathcal{A}}(\vec{s}, \vec{l})$ was defined in equation (7.65) as

$$P_{\mathcal{A}}(l^x, l^z, \vec{s}^x, \vec{s}^z) = \sum_{\vec{n}^x \in \mathbb{F}_q^{n-k_1}} \sum_{\vec{n}^z \in \mathbb{F}_q^{k_2}} P_{\mathcal{A}}(\vec{s}, \vec{n}, \vec{l}). \tag{7.73}$$

Do we have to calculate this sum for all $2^{m_1 m_2 - 1}$ distinct syndromes $\vec{s} = (\vec{s}^x \in \mathbb{F}_2^{m_2(m_1-1)}, \vec{s}^z \in \mathbb{F}_2^{m_2 - 1})$? A moment's thought shows that $P_{\mathcal{A}}(\vec{l}, \vec{s})$ actually depends only on the value of

$$((0, \beta_1), (\alpha_2, \beta_2), \ldots, (\alpha_{m_2}, \beta_{m_2})), \tag{7.74}$$

where $\alpha_i = s_{i-1}^z$ and β_j is total number of ones in the j-th $(m_1 - 1)$-bit block in $\vec{s}^x$ (compare with figure 7.4). In addition, only the frequency distribution of the (α_i, β_i) matters.

For some $\vec{s}$ which has the properties expressed by (7.74), we have

$$\begin{aligned} P_{\mathcal{A}}(l^x, l^z, \vec{s}^x, \vec{s}^z) &\equiv P_{\mathcal{A}}\big(l^x, l^z, (0, \beta_1), (\alpha_2, \beta_2), \ldots, (\alpha_{m_2}, \beta_{m_2})\big) \\ &= \sum_{\vec{n}^x \in \mathbb{F}_q^{m_2(m_1-1)}} \sum_{\vec{n}^z \in \mathbb{F}_q^{m_2 - 1}} P_{\mathcal{A}}^n\big(\vec{a}^x(\vec{s}^x, \vec{n}^z, l^x), \vec{a}^z(\vec{s}^z, \vec{n}^x, l^z)\big). \end{aligned} \tag{7.75}$$

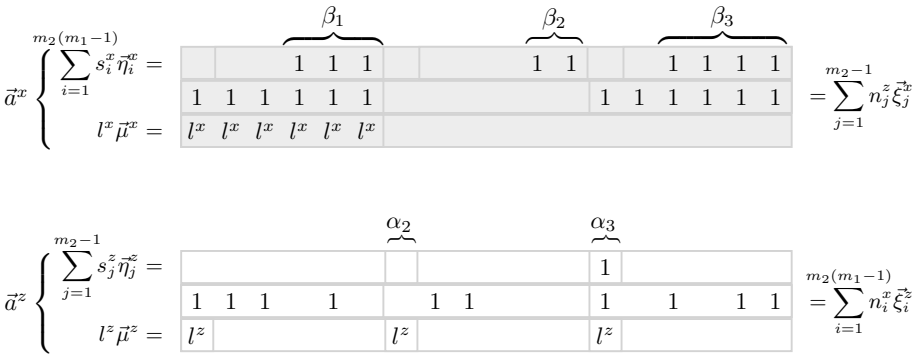


Figure 7.4: Graphical representation of the strings $\vec{a}^x = \sum_i s_i^x \vec{\eta}_i^x + \sum_j n_i^z \vec{\xi}_j^x + l^x \vec{\mu}^x$ and $\vec{a}^z = \sum_i n_i^x \vec{\xi}_i^z + \sum_j s_i^z \vec{\eta}_j^z + l^z \vec{\mu}^z$ for the concatenated cat code with $m_1 = 6$ and $m_2 = 3$. The structure of the $\vec{\xi}_j^x$ leads to an even number of completely filled blocks of size m_1 in the middle part of $\vec{a}^x$. Similarly, the structure of the $\vec{\xi}_i^z$ leads to an even number of ones in each of the m_2 blocks of size m_1 in the middle part of $\vec{a}^z$.

The sum over $\vec{n}^z$ can be written as

$$\sum_{b_1=0}^{1} \cdots \sum_{b_{m_2}=0}^{1} \frac{1+(-1)^{\sum_i b_i + l^x}}{2}, \tag{7.76}$$

which assures that the total number of completely filled blocks of size m_1 in $\vec{a}^x$ (compare with figure 7.4) is even for $l^x = 0$ and odd for $l^x = 1$. The sum over $\vec{n}^x$ is decomposed into m_2 sums each of which is written using the shorthand notation

$$\sum_{l_i,t_i}^{(\alpha_i,\beta_i)} \equiv \sum_{l_i=0}^{\beta_i} \sum_{t_i=0}^{m_1-\beta_i} \frac{1+(-1)^{l_i+t_i+l^z+\alpha_i}}{2} \binom{\beta_i}{l_i} \binom{m_1-\beta_i}{t_i}. \tag{7.77}$$

Here, l_i denotes the number of ones which are placed in a region of $\vec{a}^z$ where $\vec{a}^x$ contains ones counted by β_i, and t_i denotes the number of ones which are placed in the remaining regions of $\vec{a}^z$. Therefore, there are l_i Y-errors, $\beta_i - l_i$ X-errors and t_i Z-errors if $b_i = 0$, while there are l_i Z-errors, t_i Y-errors and $m_1 - \beta_1 - t_i$ X-errors if $b_i = 1$. Altogether we obtain

$$P_{\mathcal{A}}(l^x, l^z, \vec{s}) = \sum_{b_1=0}^{1} \cdots \sum_{b_{m_2}=0}^{1} \frac{1+(-1)^{\sum_i b_i + l^x}}{2} \sum_{l_1,t_1}^{(0,\beta_1)} \sum_{l_2,t_2}^{(\alpha_2,\beta_2)} \cdots \sum_{l_{m_2},t_{m_2}}^{(\alpha_{m_2},\beta_{m_2})} \prod_{i=1}^{m_2} \Big(p_y^{l_i} p_z^{t_i} p_x^{\beta_i - l_i} p_e^{m_1-\beta_i-t_i} \Big)^{1-b_i} \Big(p_z^{l_i} p_y^{t_i} p_e^{\beta_i - l_i} p_x^{m_1-\beta_i-t_i} \Big)^{b_i}, \tag{7.78}$$

which can be simplified by applying the following binomial series identity,

$$\sum_{k=0}^{n} \binom{n}{k} \frac{1+(-1)^{k+l}}{2} x^k y^{n-k} = \frac{1}{2}\big((x+y)^n + (-1)^l (y-x)^n\big), \tag{7.79}$$

first to each sum over l_i and then to each sum over t_i, leading to

$$P_{\mathcal{A}}(l^x, l^z, \vec{s}) = \sum_{b_1=0}^{1} \cdots \sum_{b_{m_2}=0}^{1} \frac{1+(-1)^{\sum_i b_i + l^x}}{2} F_{b_1}(l^z, 0, \beta_1) F_{b_2}(l^z, \alpha_2, \beta_2) \ldots F_{b_{m_2}}(l^z, \alpha_{m_2}, \beta_{m_2}), \tag{7.80}$$

with

$$F_0(l^z,\alpha,\beta)=\frac{1}{2}\big[(p_x+p_y)^\beta(1-p_x-p_y)^{m_1-\beta}+(-1)^{l_z+\alpha}(p_x-p_y)^\beta(1-p_x-p_y-2p_z)^{m_1-\beta}\big] \quad (7.81a)$$

$$F_1(l^z,\alpha,\beta)=\frac{1}{2}\big[(1-p_x-p_y)^\beta(p_x+p_y)^{m_1-\beta}+(-1)^{l_z+\alpha}(1-p_x-p_y-2p_z)^\beta(p_x-p_y)^{m_1-\beta}\big]. \quad (7.81b)$$

In the above expressions we replaced p_e by $1-p_x-p_y-p_z$. By adding up the last remaining sums over the b_i, eventually we arrive at the final result,

$$P_{\mathcal{A}}(l^x,l^z,\vec{s}) = \frac{1}{2}\left[\prod_{i=1}^{m_2}\big(F_0(l^z,\alpha_i,\beta_i)+F_1(l^z,\alpha_i,\beta_i)\big)+(-1)^{l_x}\prod_{i=1}^{m_2}\big(F_0(l^z,\alpha_i,\beta_i)-F_1(l^z,\alpha_i,\beta_i)\big)\right], \quad (7.82)$$

where α_1 is always assumed to be zero.

The observation that only the frequency distribution of the (α_i,β_i) matters allows us to speed up the summation over all possible syndromes drastically. To calculate the total probability of getting a certain logical error, we have to evaluate the sum over all $2^{m_1m_2-1}$ syndromes $\vec{s}$,

$$P_{\mathcal{A}}(l^x,l^z)=\sum_{\vec{s}\in\mathbb{F}_2^{m_2m_1-1}}P_{\mathcal{A}}(l^x,l^z,\vec{s}) \quad (7.83)$$

$$=\sum_{\beta_1=0}^{m_1-1}\binom{m_1-1}{\beta_1}\sum_{(\alpha_2,\beta_2),\ldots,(\alpha_{m_2},\beta_{m_2})}\binom{m_1-1}{\beta_2}\cdots\binom{m_1-1}{\beta_{m_2}}\times P_{\mathcal{A}}\big(l^x,l^z,((0,\beta_1),(\alpha_2,\beta_2),\ldots,(\alpha_{m_2},\beta_{m_2}))\big). \quad (7.84)$$

Since (α_i,β_i) takes on $2m_1$ different values, this expression simplifies to

$$=\sum_{\beta_1=0}^{m_1-1}\binom{m_1-1}{\beta_1}\sum_{\substack{a_1,a_2,\ldots,a_{2m_1}=0\\ \text{s.t. }\sum_i a_i=m_2-1}}^{m_2-1}\frac{(m_2-1)!}{a_1!a_2!\ldots a_{2m_1}!}\prod_{i=1}^{2m_1}\binom{m_1-1}{\beta(i)}^{a_i}\times P_{\mathcal{A}}\big(l^x,l^z,((0,\beta_1),\{(\alpha(j),\beta(j))^{a_j}\}_{j=1\ldots2m_1})\big). \quad (7.85)$$

Instead of adding up $2^{m_1m_2-1}$ terms as in (7.83), we only have to consider $m_1\cdot\binom{m_2-1+2m_1-1}{m_2-1}$ terms.

Joint Probabilities for the Cat Code

By setting $m_2=1$ and $m_1=m$ in equation (7.82), we get the joint probabilities for the $[[m,1]]_2$ cat code of subsection 7.5.3,

$$P_{\mathcal{A}}(l^x,l^z,\vec{s}^{\,x})=\frac{1}{2}\Big[(p_x+p_y)^{l^x(m-2\beta)+\beta}(1-p_x-p_y)^{(1-l^x)(m-2\beta)+\beta}+(-1)^{l^z}(p_x-p_y)^{l^x(m-2\beta)+\beta}(1-p_x-p_y-2p_z)^{(1-l^x)(m-2\beta)+\beta}\Big]. \quad (7.86)$$

Here, β denotes the number of ones in $\vec{s}=\vec{s}^{\,x}$.

Remark. If we calculate expressions like (7.81) or (7.86) for the depolarizing channel $\mathcal{D}_p$, we have $p_x=p_y=p_z=p/3$ and therefore some of the products in these expressions become zero. If such a product is exponentiated, as it is the case for the term $(p_x-p_z)^\beta$ for instance, one has to take special care of the case $\beta=0$ in which the term is equal to one.

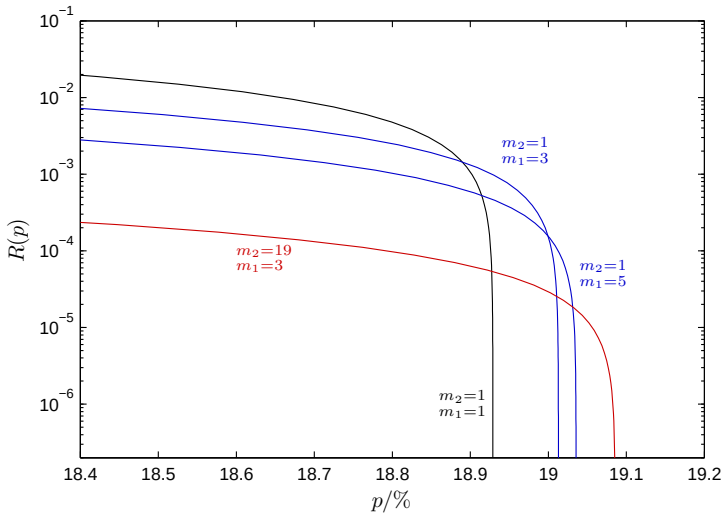


Figure 7.5: Achievable transmission rates for various $[[m_1 \times m_2, 1]]_2$ codes over the qubit depolarizing channel $\mathcal{D}_p$ plotted as function of the noise p: The hashing rate (corresponding to $m_1 = m_2 = 1$) *(black)*, the cat code ($m_2 = 1$) with $m_1 = 3$ and $m_1 = 5$ *(blue)*, the concatenated cat code with parameters $m_1 = 3, m_2 = 19$ *(red)* and $m_1 = 5, m_2 = 15$ *(orange)*.

Results for the Depolarizing Channel

We use (7.82) and (7.85) to evaluate the achievable transmission rate of equation (7.66) for various inner $[[m_1 \times m_2, 1]]_2$ concatenated cat codes concatenated with random outer codes over the qubit ($q = 2$) depolarizing channel $\mathcal{D}_p$. The hashing rate (corresponding to $m_1 = m_2 = 1$) is compared with the rates of various concatenated cat codes in figure 7.5. It can be seen that the hashing rate, which equals the one-shot capacity as shown in lemma 7.5.1,

$$Q^{(1)}(\mathcal{D}_p) = \max_{\rho} I_c(\rho, \mathcal{D}_p), \tag{7.87}$$

is surpassed e. g. by the rate of the cat code ($m_2 = 1$) of size $m = m_1 = 5$ for high values of noise ($p \approx 0.19$). Since this rate may be expressed as

$$\frac{1}{5} I_c\Big(\frac{1}{2}\Pi_{\mathcal{C}}, \mathcal{D}_p^{\otimes 5}\Big), \tag{7.88}$$

where $\Pi_{\mathcal{C}}$ denotes the projector on one of the codespaces of the $[[5, 1]]_2$ cat code (see subsection 7.4.2), it is clear that the limit as n goes to infinity in the regularized coherent information expressing the quantum capacity of a quantum channel $\mathcal{A}$,

$$Q(\mathcal{A}) = \lim_{n\to\infty} \frac{1}{n} \max_{\rho} I_c(\rho, \mathcal{A}^{\otimes n}), \tag{7.89}$$

is crucial since this example shows that in general $Q^{(n)}(\mathcal{A}) = \max_\rho I_c(\rho, \mathcal{A}^{\otimes n})/n$ might be larger than $Q^{(1)}(\mathcal{A})$.

So far lower bounds on the maximum tolerable noise $p_{\max}$ of the qubit depolarizing channel have been determined by (i) setting the hashing rate of theorem 7.3.1 and 7.3.3 equal to zero $\Rightarrow p_{\max}^{\text{hash}} = 18.9290\%$ and (ii) by setting the rate of (7.66) for a cat code of size m equal to zero $\Rightarrow p_{\max}^{\text{cat}}(m = 5) = 19.0356\%$. Now we set the rate of (7.66) for various $[[m_1 \times m_2, 1]]_2$ concatenated cat codes equal to zero. The corresponding tolerable values of noise are plotted in figure 7.6 as a function of m_2 for various values of m_1. It can be seen that for $m_1 = 3$ the best lower bound is obtained for $m_2 = 19$, $p_{\max}^{\text{conc-cat}}(m_1 = 3, m_2 =$

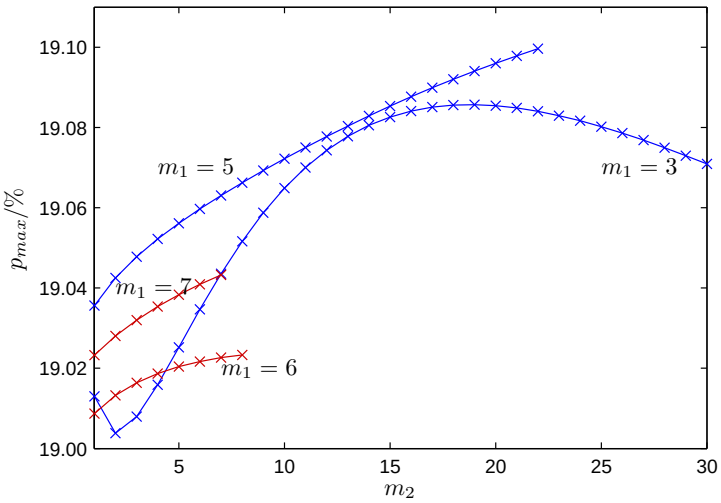


Figure 7.6: The maximum tolerable value of noise p for the qubit depolarizing channel $\mathcal{D}_p$ as a function of the size m_2 of the inner $[[m_1 \times m_2, 1]]_2$ concatenated cat code for various values of m_1. For $m_1 = 3$ the highest value is obtained for $m_2 = 19$, $p_{\max}^{\text{conc-cat}}(m_1 = 3, m_2 = 19) = 19.0857\%$. For $m_1 = 5$ the highest value shown is $p_{\max}^{\text{conc-cat}}(m_1 = 5, m_2 = 22) = 19.0996\%$.

19) = 19.0857%. Due to computational limitations (calculation of the $m_1 = 5, m_2 = 22$ point took roughly a week on a Intel core 2 duo E8500 CPU), the $m_1 = 5$ curve was calculated only up to $m_2 = 22$ leading to the best lower bound known to date of $p_{\max}^{\text{conc-cat}}(m_1 = 5, m_2 = 22) = 19.0996\%$. This beats the highest previously known lower bound of [SS07] which was $p_{\max}^{\text{conc-cat}}(m_1 = 5, m_2 = 16) = 19.0877\%$. While the optimal value of m_2 for $m_1 = 5$ was conjectured in [SS07] to be $m_2 \approx 25$, according to our new data we expect it to lie slightly higher (maybe $m_2 \approx 30$).

8 Quantum Cryptography

Quantum key distribution (QKD) protocols try to establish a secure and random key between two distant parties usually called Alice and Bob. While the security of corresponding classical protocols relies on the assumption that an eavesdropper has limited computational power, the security of a QKD protocol is guaranteed by the validity of quantum mechanics. Quantum cryptography was initiated by Bennett and Brassard in 1984 who developed the first QKD protocol, which is now called BB84 protocol [BB84]. A natural extension of BB84 which makes use of four different quantum states is the 6-state protocol [Bru98] which makes use of two additional quantum states. To prove the security of a QKD protocol, one makes the worst case assumption that the quantum channel connecting the two parties is under complete control of an eavesdropper, usually named Eve. Since non-orthogonal quantum states cannot be cloned perfectly [Die82; WZ82], the two users Alice and Bob are able to detect the presence of an eavesdropper by comparing some of Bob's measurement results with Alice's preparations in a step called parameter estimation. Depending on the result, they might either abort the protocol, or, if the action of the eavesdropper seems harmless enough, proceed with an error correction and privacy amplification step to obtain a random and private key.

Using a quantum channel to create a secret key between two parties is closely related to using the channel to send quantum information, with many results found in one area applicable in the other. For instance, by treating the steps in a quantum key distribution (QKD) protocol coherently and viewing the entire process as an entanglement distillation scheme, one can use properties of random quantum error-correcting codes to prove the security of the BB84 and 6-state protocols up to bit error rates of $p_{\max}^{\text{BB84}} = 11.0028\%$ [SP00] and $p_{\max}^{\text{6-st.}} = 12.6193\%$ [Lo01], respectively. Conversely, the formula for the quantum channel capacity can be obtained by importing the key rate resulting from a general approach to secret key generation over a known channel [DW04; Dev05; DW05].

One of the surprising results related to quantum capacity is the non-optimality of random codes, in contrast to the classical case. As it was shown in chapter 5, the classical capacity of a channel can be achieved by using randomly-constructed block codes, and the independence of one input to the channel from the next results in a so-called single-letter formula for the capacity. While random coding can be used to create quantum error-correcting codes as well (compare with section 7.3), these do not always achieve the capacity. Better performance can be achieved by structured codes which exploit the ability of quantum error-correcting codes to correct errors without precisely identifying them, a property called degeneracy (compare with section 7.4).

By appealing to the coherent formulation of the protocol, degenerate codes should also be useful in QKD. This was shown to be the case in the original security proof of the 6-state protocol [Lo01], as the results of [DSS98] were used to improve the error rate threshold to $p_{\max}^{\text{6-st.}} = 12.6904\%$. More striking threshold improvements are possible, if counterintuitive, by simply adding noise to the raw key bits before they are processed into the final key, a procedure known as local randomization [KGR05; RGK05]. This improves the error rate thresholds for the two protocols to $p_{\max}^{\text{BB84}} = 12.4120\%$ and $p_{\max}^{\text{6-st.}} = 14.1119\%$, respectively. At first glance, these results make no sense in the coherent picture of QKD, since adding more noise to already noisy entangled pairs only decreases the amount of pure entanglement which can be extracted. The entanglement/secret-key analogy does not hold perfectly, however; entangled states are sufficient, but not necessary, for creation of secret keys. A broader class of states, called private states, leads to secret keys when measured [HHHO05], and these should properly be the target output of the coherent version of the QKD protocol. Indeed, the exact error thresholds are recovered in the coherent picture when the QKD protocols with local randomization are analyzed in these terms [RS07].

With a systematic understanding of how degenerate codes and local randomization boost the key rate, it becomes sensible to combine the two methods to look for even higher thresholds. Recently it was shown in [SRS08] that doing so improves the error threshold of the BB84 protocol up to at least $p_{\max}^{\mathrm{BB84}} \approx 12.92\%$ by using the same type of structured code studied in [SS96; DSS98; SS07] and subsection 7.5.3. These specific codes consist of the concatenation of two codes, the first a simple repetition code and the second a random code. The repetition code, sometimes called a cat code in the context of quantum information theory since the codewords are $|0\rangle^{\otimes m}$ and $|1\rangle^{\otimes m}$, induces degeneracy in the overall code since a phase flip on any of the physical qubits leads to the same logical error, and is corrected in the same way. In particular, blocklength $m = 400$ corresponds to the threshold stated above. Since the random code portion of the protocol corresponds to information reconciliation and privacy amplification in the classical view, the local randomization and the repetition code together become a type of *preprocessing* performed before these "usual" steps.

In this chapter we show that the same preprocessing protocol as used in [SRS08] can also be used to improve the maximum tolerable bit error rate for the 6-state protocol, up to at least $p_{\max}^{\text{6-st.}} = 14.5930\%$ for a blocksize of $m = 300$. This is already quite close to the upper bound of 14.6447% [FGG$^+$97; KGR05; MCL06] on the tolerable error rate for the BB84 protocol, and since the error threshold grows with blocklength, the bound is presumably exceeded at larger blocklengths, indicating the higher robustness of the 6-state protocol. We also improve the lower bounds for the BB84 protocol presented in [SRS08]. In addition we investigate iterating the preprocessing scheme in the BB84 protocol, and show an improvement both in rate and error threshold over single-round preprocessing for even modest blocklengths. The results presented in this chapter have been obtained in collaboration with J. Renes and have been published in [KR08].

To begin, section 8.1 explains the BB84 and 6-state QKD protocols and summarizes Shor and Preskill's security proof [SP00] which uses the structure of CSS codes to show the equivalence between these protocols and corresponding entanglement distillation protocols. Section 8.2 describes the preprocessing scheme in more depth and then derives secret key rate expressions for the BB84 and the 6-state protocols. Numerical calculations for blocklengths into the hundreds are then presented for the two protocols. We explain how representation theory is helpful for the numerical evaluation of such key rates in both cases. Section 8.3 examines the advantages of iterating the preprocessing protocol to achieve higher rates and thresholds for the same amount of effort in noise addition and block coding.

8.1 BB84 and 6-State Protocols

The BB84 [BB84] and the 6-state [Bru98] protocol are QKD protocols of the prepare and measure type. Their goal is to establish a random and secret key between two parties — usually called Alice and Bob — which are connected via a quantum channel and a classical channel. The quantum channel is fully accessible to an eavesdropper — traditionally called Eve — while the classical channel is assumed to be authenticated, i. e. Eve can only listen to the messages, but cannot interfere. (To authenticate the classical channel, Alice and Bob need to share a small secret key in advance. Hence, strictly speaking, QKD protocols are secret key growing protocols.)

Remark. While Alice and Bob have to use two-way classical communication for the parameter estimation step of the protocol, this chapter deals only with protocols using one-way communication during the error correction and privacy amplification steps. The use of two-way communication during these steps allows them to obtain a secure key for even higher levels of noise [GL03], which we assume is caused by Eve.

8.1.1 Description of the Protocols

Let $s = 2$ for the BB84 protocol and $s = 3$ for the 6-state protocol. If we denote the eigenstates corresponding to eigenvalues $+1$ and -1 of the Pauli Z matrix by $|0\rangle_z$ and $|1\rangle_z$, the corresponding eigenstates of the Pauli X and Y matrices are given by

$$|0\rangle_x = (|0\rangle_z + |1\rangle_z)/\sqrt{2} \qquad |1\rangle_x = (|0\rangle_z - |1\rangle_z)/\sqrt{2} \tag{8.1}$$

$$|0\rangle_y = (|0\rangle_z + i|1\rangle_z)/\sqrt{2} \qquad |1\rangle_y = (|0\rangle_z - i|1\rangle_z)/\sqrt{2}. \tag{8.2}$$

In addition, let $B(0) = z$, $B(1) = x$ and $B(2) = y$.

Alice chooses a random sequence of zeros and ones $\vec{x} = (x_1, x_2, \ldots, x_N) \in \mathbb{F}_2^N$ of length $N \gtrsim 2 \cdot s \cdot n$ and a random sequence $\vec{b} = (b_1, b_2, \ldots, b_N) \in \mathbb{F}_s^N$. Then she prepares the sequence of quantum states $\bigotimes_{i=1}^N |x_i\rangle_{B(b_i)}$ and sends them to Bob. Bob chooses a random sequence $\vec{b'} = (b'_1, b'_2, \ldots, b'_N) \in \mathbb{F}_s^N$ and measures the i-th qubit in the basis $B(b'_i)$ denoting the result as $y_i \in \mathbb{F}_2$. After Bob finished his measurements, he announces this fact and both parties compare their strings $\vec{b}$ and $\vec{b'}$. If $b_i \neq b'_i$ they remove the i-th entry from their strings $\vec{x}$ and $\vec{y}$. The resulting strings $\vec{x}_{\text{sifted}}$ and $\vec{y}_{\text{sifted}}$ form the sifted key and are of length $2 \cdot n$ approximately. If the quantum states had been transmitted unperturbed, the sifted keys of Alice and Bob coincide, $\vec{x}_{\text{sifted}} = \vec{y}_{\text{sifted}}$. To check whether this is the case, Alice selects half of the bits to serve as check bits, submits her choice to Bob, and both parties compare this part of their sifted key. The resulting error rate is called the bit-error rate p. If the bit-error rate p is zero, they can be confident that no eavesdropper was present and may use the remaining n bits $\vec{x}'_{\text{sifted}}$ and $\vec{y}'_{\text{sifted}}$ as a secure and random key.

In practice there will always be a bit-error rate $p > 0$ due to imperfections of the quantum channel or the presence of an eavesdropper. Hence the task is to proof the security of the protocols up to a certain bit-error rate $p_{\max}$. As long as $p < p_{\max}$, Alice and Bob should be able to perform error correction and privacy amplification to obtain a secure key $\vec{k}$ of length $k < n$ from $\vec{x}'_{\text{sifted}} \in \mathbb{F}_2^n$ and from $\vec{y}'_{\text{sifted}} \in \mathbb{F}_2^n$. The first simple proof of security was given by Shor and Preskill [SP00]: By treating the steps in a QKD protocol coherently and viewing the entire process as an entanglement distillation scheme, one can use properties of random quantum error-correcting codes to prove the security of the BB84 and 6-state protocols up to bit-error rates of $p_{\max}^{\text{BB84}} = 11.0028\%$ [SP00] and $p_{\max}^{\text{6-st.}} = 12.6193\%$ [Lo01], respectively.

8.1.2 Shor and Preskill's Security Proof

The security proof of Shor and Preskill is based on the observation of Deutsch et al. [DEJ+96] and Lo and Chau [LC99] that entanglement distillation protocols provide a way to establish a secret key between the two parties Alice and Bob. If, as a result of an entanglement distillation protocol, Alice and Bob share (near) perfect states $|\Phi^+\rangle_{AB} = (|00\rangle_{AB} + |11\rangle_{AB})/\sqrt{2}$, a bipartite measurement of $|\Phi^+\rangle_{AB}$ in the z-basis results in a shared secret bit*. Shor and Preskill [SP00] (see also [GP01] for a more elaborate version of the proof) realized that an entanglement distillation protocol making use of CSS codes is equivalent to the BB84 protocol. Their proof was adapted to the 6-state protocol by Lo [Lo01]. In the following we describe the corresponding entanglement distillation protocol and its reduction to a prepare and measure scheme. For BB84, let $T = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ denote the Hadamard matrix mapping the z-basis onto the x-basis and vice versa. For the 6-state protocol, let

$$T = \exp\Big(-\frac{i}{2}(X + Y + Z)/\sqrt{3} \cdot \frac{2\pi}{3}\Big) \cdot e^{i\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} \tag{8.3}$$

denote the rotation of angle $2\pi/3$ around the axis $(1, 1, 1)/\sqrt{3}$, mapping the z-axis to the x-axis, the x-axis to the y-axis, and the y-axis to the z-axis.

*A maximally entangled state like $|\Phi^+\rangle_{AB}$ is not necessary to provide a secret bit; so-called private states are necessary and sufficient [HHHO05].

Entanglement Distillation Protocol

Alice prepares $N = 2n$ maximally entangled pairs $|\Phi^+\rangle_{AB}$ and chooses a random string $\vec{b} = (b_1, \ldots, b_N) \in \mathbb{F}_s^N$. After applying the operation $T_B^{b_i}$ onto Bob's part of the i-th pair, she sends him his half of the states. Bob acknowledges the reception of his qubits. Alice picks out n pairs which have to serve as check pairs and tells Bob the string $\vec{b}$ together with her choice of the check pairs. Bob applies the operation $T_B^{-b_i}$ onto his i-th qubit. Both parties measure the check pairs in the z-basis, share their results and obtain the bit-error rate p. Since there is no way for Eve to know the check pairs in advance, the bit-error rate of the check bits should be a pretty good estimate for the bit error rate of the remaining n pairs.

Let us assume now that Eve's attack can be described by a memoryless Pauli channel $\mathcal{E}^{\otimes N}$ where $\mathcal{E}$ is characterized by the probability distribution $\{q_I, q_x, q_y, q_z\}$. Of course Eve might apply any completely positive map, but, as it was pointed out in [LC99], the entanglement distillation protocol which will be used to generate $k < n$ (near) perfect pairs from the remaining n, commutes with a measurement of each pair in the Bell basis. Hence the most general attack of Eve can be described by a general Pauli channel which corresponds to the twirled version of Eve's attack (compare with theorem 7.2.1). Furthermore, it can be shown that if the entanglement distillation protocol is capable of correcting an uncorrelated Pauli attack, it is also capable of correcting a correlated one (if Alice and Bob apply a random permutation to their qubits; see e. g. [GL03]). As a result of the application of the $T_B^{b_i}$ with $b_i \in \mathbb{F}_s$, parameter estimation assures us that the effective Pauli channel

$$\mathcal{E}_{\text{eff}}(\rho) = \frac{1}{s} \sum_{j=0}^{s-1} T^{-j} \mathcal{E}(T^j \rho T^{-j}) T^j \tag{8.4}$$

is characterized by the probability distribution $\{p_{uv}\} \equiv \{p_{00}, p_{10}, p_{11}, p_{01}\}$ s. t.

$$\{p_{uv}\} = \begin{cases} \{1-2p+t, p-t, t, p-t\}, t \in [0,p], & \text{in case of the BB84 protocol.} \\ \{1-\frac{3}{2}p, \frac{p}{2}, \frac{p}{2}, \frac{p}{2}\}, & \text{in case of the 6-state protocol.} \end{cases} \tag{8.5}$$

We are now going to describe the entanglement distillation protocol which is capable of distilling $k < n$ (near) perfect $|\Phi^+\rangle_{AB}$ pairs from the remaining state $\mathcal{I}_A \otimes \mathcal{E}_{\text{eff},B}^{\otimes n}\big((|\Phi^+\rangle\langle\Phi^+|)^{\otimes n}\big)$ as long as the bit error rate p is not too high. Let us fix a CSS code encoding $k = k_1 - k_2$ qubits into n. As explained in section 6.3, together with an encoding U_{enc} such a code is specified by the two lists of vectors

$$\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{n-k_1}^z, \vec{\eta}_1^z, \ldots, \vec{\eta}_{k_2}^z, \vec{\mu}_1^z, \ldots, \vec{\mu}_k^z\} \text{ and}$$
$$\{\vec{\eta}_1^x, \ldots, \vec{\eta}_{n-k_1}^x, \vec{\xi}_1^x, \ldots, \vec{\xi}_{k_2}^x, \vec{\mu}_1^x, \ldots, \vec{\mu}_k^x\},$$

both spanning $\mathbb{F}_2^n$ and satisfying (6.29), where $\mathcal{C}_1^\perp = \text{span}\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{n-k_1}^z\}$ and $\mathcal{C}_2 = \text{span}\{\vec{\xi}_1^x, \ldots, \vec{\xi}_{k_2}^x\}$ are classical linear codes satisfying $\mathcal{C}_2 \subseteq \mathcal{C}_1$. Note that because of lemma C.3.2,

$$|\Phi^+\rangle_{AB}^{\otimes n} = U_{\text{enc},A}^* \otimes U_{\text{enc},B} |\Phi^+\rangle_{AB}^{\otimes n} = \frac{1}{\sqrt{2^{n-k_1}}} \sum_{\vec{x} \in \mathbb{F}_2^{n-k_1}} \frac{1}{\sqrt{2^{k_2}}} \sum_{\vec{z} \in \mathbb{F}_2^{k_2}} \frac{1}{\sqrt{2^k}} \sum_{\vec{c} \in \mathbb{F}_2^k} \overline{|\vec{x},\vec{z},\vec{c}\rangle}_A^* |\overline{\vec{x},\vec{z},\vec{c}}\rangle_B. \tag{8.6}$$

Alice measures her stabilizers $\{\overline{Z}_i^*\}_{i=1,\ldots,n-k}$, sends her resulting syndrome $\vec{s}_A = (\vec{x}, \vec{z})$ to Bob, who, by measuring his stabilizers $\{\overline{Z}_i\}_{i=1,\ldots,n-k}$, obtains the syndrome $\vec{s}_B = \vec{s}_A + \vec{s}$ and calculates the relative syndrome $\vec{s}$. Depending on $\vec{s}$, Bob performs error correction. Eventually, Alice and Bob both measure $\{\overline{Z}_i^*\}_{i=n-k+1,\ldots,n}$ and $\{\overline{Z}_i\}_{i=n-k+1,\ldots,n}$, respectively, to obtain the k bit key. (Alternatively they might also decode, obtain $|\Phi^+\rangle_{AB}^{\otimes k}$, and measure in the z-basis to obtain the key.)

Protocol based on Quantum Error Correction

Since Alice might perform her measurements immediately after the preparation of $|\Phi^+\rangle_{AB}^{\otimes N}$, the following procedure is equivalent: She chooses the syndrome $(\vec{x},\vec{z})$, the key $\vec{c}$, and the values of the check bits at random. Then she prepares the n-qubit state $\overline{|\vec{x},\vec{z},\vec{c}\rangle}$ and inserts the n check states prepared as $|0\rangle_z$ or $|1\rangle_z$ in random positions. After choosing a random string $\vec{b}=(b_1,\ldots,b_N)\in\mathbb{F}_s^N$, she applies the operation T^{b_i} onto the i-th qubit, and sends her $N=2n$ qubits to Bob. Bob acknowledges the reception of the qubits. Alice tells Bob the string $\vec{b}$ together with the positions of the check qubits. Bob applies the operation T^{-b_i} onto his i-th qubit. He measures the check qubits in the z-basis, they share their check bit data, and, as a result, obtain the bit-error rate p. At this point Bob is left with the state $\mathcal{E}_{\text{eff}}^{\otimes n}\big(\overline{|\vec{x},\vec{z},\vec{c}\rangle\langle\vec{x},\vec{z},\vec{c}|}\big)$. Alice tells him the syndrome $\vec{s}_A=(\vec{x},\vec{z})$, and Bob knows that the key is encoded in the codespace $\mathcal{C}\big(L(\mathcal{C}_1,\mathcal{C}_2),\vec{s}_A\big)$ of the CSS code. He applies the appropriate recovery operation $\mathcal{R}_{\vec{s}_A}$ by measuring the stabilizers $\{\overline{Z}_i\}_{i=1,\ldots,n-k}$ followed by error correction. Eventually, Bob measures $\{\overline{Z}_i\}_{i=n-k+1,\ldots,n}$ to obtain the k bit key.

The rate k/n of the key they can generate this way depends only on the form of the memoryless Pauli channel $\mathcal{E}_{\text{eff}}$ which in turn depends only on the bit error rate p. Hence, lower bounds on the rates are given by theorem 7.3.3 which states that, as long as

$$\frac{k}{n}<1-H_{4[\log_2]}(\{p_{uv}\}), \tag{8.7}$$

and for large enough n, there exists a pair of codes $\mathcal{C}_2\subset\mathcal{C}_1$ such that for any codespace $\mathcal{C}_{(L(\mathcal{C}_1,\mathcal{C}_2),\vec{s})}$ of the corresponding CSS code with stabilizer $L(\mathcal{C}_1,\mathcal{C}_2)$, there exists a recovery operation with minimum fidelity larger than $1-\varepsilon$ for any $\varepsilon>0$. To obtain higher rates, they might also use concatenated CSS codes as it was done by Lo [Lo01] (see the second remark following theorem 7.4.1).

Remark. In the case of the BB84 protocol the set $\{p_{uv}\}$ is not completely known and we have to assume the worst case, i. e. we have to minimize the key rates over the unknown parameter $t\in[0,p]$.

BB84 and 6-state Protocol

Finally we are going to show that the protocol based on quantum error correction is equivalent to the BB84 and the 6-state protocol, respectively. The crucial observation is that the recovery operation for CSS codes decomposes into bit and phase error correction. Since Bob obtains the key by measuring the operators $\{\overline{Z}_i\}_{i=n-k+1,\ldots,n}$, where $\overline{Z}_{n-k+j}=XZ(\vec{0},\vec{\mu}_j^z)$ for $j=1\ldots k$, he does not need to perform phase error correction. Hence, he only needs to know the absolute bit syndrome $\vec{x}$ and the relative bit syndrome obtained by measuring the $\overline{Z}_j=XZ(\vec{0},\vec{\xi}_j^z)$, $j=1\ldots n-k_1$. To obtain his measurement results, he might simply measure all qubits in the z-Basis, obtain a string $\vec{y}\in\mathbb{F}_2^n$ and reconstruct them via $\vec{\mu}_j^z\cdot\vec{y}$, $j=1\ldots k$, and $\vec{\xi}_j^z\cdot\vec{y}$, $j=1\ldots n-k_1$, respectively. Alice, who in turn does not need to send the phase error syndrome $\vec{z}$, prepares on average the state

$$\begin{aligned}\frac{1}{2^{k_2}}\sum_{\vec{z}\in\mathbb{F}_2^{k_2}}\overline{|\vec{x},\vec{z},\vec{c}\rangle\langle\vec{x},\vec{z},\vec{c}|}&=\frac{1}{|\mathcal{C}_2|}\sum_{\vec{\mathfrak{v}}_1,\vec{\mathfrak{v}}_2\in\mathcal{C}_2}\frac{1}{2^{k_2}}\sum_{\vec{z}\in\mathbb{F}_2^{k_2}}(-1)^{\vec{\mathfrak{z}}\cdot(\vec{\mathfrak{v}}_1-\vec{\mathfrak{v}}_2)}|\vec{\mathfrak{v}}_1+\vec{\mathfrak{c}}+\vec{\mathfrak{x}}\rangle\langle\vec{\mathfrak{v}}_1+\vec{\mathfrak{c}}+\vec{\mathfrak{x}}|\\&=\frac{1}{|\mathcal{C}_2|}\sum_{\mathfrak{v}\in\mathcal{C}_2}|\vec{\mathfrak{v}}+\vec{\mathfrak{c}}+\vec{\mathfrak{x}}\rangle\langle\vec{\mathfrak{v}}+\vec{\mathfrak{c}}+\vec{\mathfrak{x}}|,\end{aligned} \tag{8.8}$$

where $\vec{\mathfrak{x}}$, $\vec{\mathfrak{z}}$ and $\vec{\mathfrak{c}}$ had been defined in (6.33) as

$$\vec{\mathfrak{x}}=\sum_{i=1}^{n-k_1}x_i\vec{\eta}_i^x,\qquad\qquad\vec{\mathfrak{z}}=\sum_{i=1}^{k_2}z_i\vec{\eta}_i^z,\qquad\qquad\text{and }\vec{\mathfrak{c}}=\sum_{i=1}^{k}c_i\vec{\mu}_i^x. \tag{8.9}$$

Note that $\vec{\mathfrak{v}}+\vec{\mathfrak{c}}\in\mathcal{C}_1$ and $\vec{\mathfrak{v}}+\vec{\mathfrak{c}}+\vec{\mathfrak{x}}\in\mathbb{F}_2^n$ so that Alice just prepares a sequence of n random states taken from the set $\{|0\rangle_z,|1\rangle_z\}$.

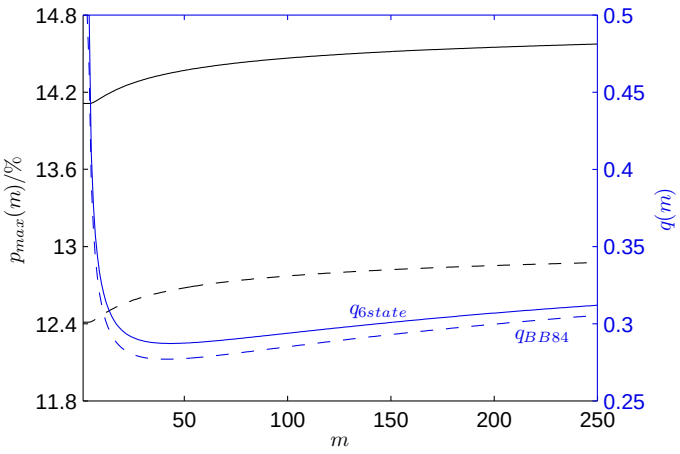


Figure 8.1: Maximum tolerable bit error rate $p_{\max}$ *(left y-axis, black)* and the corresponding rate q of the added noise for which it is achieved *(right y-axis, blue)* versus block length m. Dashed lines correspond to the BB84 protocol, solid lines to the 6-state protocol.

In summary, we have the following secure protocol: Alice and Bob implement the corresponding QKD protocol as described in subsection 8.1.1. As a result they end up with Alice having the n bits $\vec{x}'_{\text{sifted}}$, Bob having the n bits $\vec{y}'_{\text{sifted}}$, and both knowing the bit error rate p. They decide on a CSS code encoding k qubits into n which is able to correct the memoryless Pauli channel $\mathcal{E}_{\text{eff}}^{\otimes n}$ characterized by the probability distribution of equation (8.5). Alice interprets $\vec{x}'_{\text{sifted}}$ as $(\vec{\mathfrak{v}}+\vec{\mathfrak{c}})+\vec{\mathfrak{x}}$ with random $(\vec{\mathfrak{v}}+\vec{\mathfrak{c}}) \in \mathcal{C}_1$ and random syndrome $\vec{\mathfrak{x}}$, and tells Bob the syndrome. Bob's data $\vec{y}'_{\text{sifted}}$ can be written as the sum of Alice's string plus an error, $\vec{y}'_{\text{sifted}} = \vec{x}'_{\text{sifted}} + \vec{e}$. Bob subtracts the syndrome, obtains $(\vec{\mathfrak{v}}+\vec{\mathfrak{c}})+\vec{e}$, and performs bit error correction with the classical code $\mathcal{C}_1$ to obtain $(\vec{\mathfrak{v}}+\vec{\mathfrak{c}})$. To obtain the key, he extracts the coset of $\mathcal{C}_2$ in $\mathcal{C}_1$, $\vec{\mu}_i^z \cdot (\vec{\mathfrak{v}}+\vec{\mathfrak{c}}) = \vec{\mu}_i^z \cdot \vec{\mathfrak{c}} = c_i$. The last step can be viewed as privacy amplification: The correct k_1 bits included in $(\vec{\mathfrak{v}}+\vec{\mathfrak{c}})$ are shrunk into $k = k_1 - k_2$ private bits.

8.2 Combined Preprocessing

The preprocessing protocol proposed in [SRS08] combines local randomization with the use of a degenerated quantum code. It begins after Bob has received the quantum signals from Alice and they have sifted their raw keys to throw out mismatches between the preparation and measurement basis. Alice then flips each of her sifted key bits $(x_1, \ldots, x_n)$ with probability q, resulting in new bits $(\tilde{x}_1, \ldots, \tilde{x}_n)$. These are partitioned into blocks of size m, and for each block she computes the syndrome $(\tilde{x}_1 \oplus \tilde{x}_2, \tilde{x}_1 \oplus \tilde{x}_3, \ldots, \tilde{x}_1 \oplus \tilde{x}_m)$ and sends this information to Bob. He computes the *relative* syndrome of their blocks by adding his corresponding syndrome to Alice's, modulo two. Alice's message is public knowledge, but the first bit of each block is still secret, so it is kept as a potential key bit. The protocol then proceeds with the usual error correction and privacy amplification steps to transform these kept bits into a secret key, now aided by the relative syndrome of each block and knowledge of the probability q of local randomization. Without local randomization, it turns out that $m = 5$ is the optimal blocklength for improving the error threshold in the 6-state protocol — longer blocklengths have worse thresholds (compare with figure 7.2 of section 7.5). However, the results in [SRS08] indicate that with the addition of noise, the highest tolerable bit error rate of BB84 grows with the blocksize m, and we find a similar result in the 6-state case (see figure 8.1).

8.2.1 Security Proof

We determine the secure key rates of the BB84 and 6-state one-way key distillation protocols involving the preprocessing protocol described above using the security proof of Renner [Ren05, corollary 6.5.2]. This proof states that the secure key rate of such a protocol is given by

$$r = \frac{1}{m} \min_{\sigma_{AB} \in \Gamma} \big(S(X|E) - S(X|Y) \big) \tag{8.10}$$

where the minimum ranges over the set of states Γ of all density operators on the 2×2 dimensional Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ such that the measurement performed during the parameter estimation phase of the protocol leads to a certain bit error rate p. The conditional von Neumann entropies in (8.10) are calculated for the states

$$\sigma_{XY\overline{E}} = \mathcal{E}_{XY\overline{E} \leftarrow A^m B^m E^m} (\sigma_{ABE}^{\otimes m}) \tag{8.11}$$

which describe the processing of each block, including local randomization and syndrome calculation, and eventual measurement of the output qubits of the repetition code. That is, the preprocessing is treated quantum-mechanically or coherently, but the usual processing classically. Here X denotes Alice's key outcome when measuring the output bits and Y Bob's key and syndrome outcomes.

For the BB84 protocol the set Γ contains the states

$$\sigma_{AB} = \sum_{u,v} p_{uv} XZ_B(u,v) |\Phi^+\rangle\langle\Phi^+| XZ_B^\dagger(u,v), \tag{8.12}$$

where $|\Phi^+\rangle_{AB} = \frac{1}{\sqrt{2}} \sum_k |kk\rangle_{AB}$ and $\{p_{uv}\} \equiv \{p_{00}, p_{10}, p_{11}, p_{01}\} = \{1-2p+t, p-t, t, p-t\}, t \in [0,p]$. In the 6-state protocol, meanwhile, parameter estimation assures us that Γ contains only the single state σ_{AB} with $\{p_{uv}\} = \{1 - \frac{3}{2}p, \frac{p}{2}, \frac{p}{2}, \frac{p}{2}\}$.

Using Renner's proof allows us to include the preprocessing but still only minimize over the quantum states σ corresponding to individual signals. The crucial simplification is that the quantum state of the block can be taken to be the product $\sigma^{\otimes m}$ without loss of generality. Other proof techniques would require minimization over all possible (potentially-entangled) block states, or an additional step in the parameter estimation procedure to ensure that the state does have this power form.

8.2.2 Computation of the Secure Key Rate

To compute the secure key rates we make use of the fact that the difference of entropies in (8.10) can also be written as difference of corresponding quantum mutual informations, i. e. $S(X|E) - S(X|Y) = I(X:Y) - I(X:E)$. In order to calculate these quantities, we need to determine the states $\sigma_{XY\overline{E}}$ defined in (8.11) for both protocols, i. e. for σ_{AB} being a member of the two different sets Γ defined in the paragraph including equation (8.12). We are going to perform the rate calculation for a general σ_{AB} and specialize in the two different cases in the succeeding subsections.

An m-fold tensor product of a purification of a general Bell diagonal σ_{AB} is given by

$$|\sigma\rangle_{ABE} \equiv |\sigma\rangle_{ABE_1E_2}^{\otimes m} = \sum_{\vec{u},\vec{v}} \sqrt{p_{\vec{u},\vec{v}}}\, XZ_B(\vec{u},\vec{v}) |\Phi^+\rangle_{AB}^{\otimes m} |\vec{u}\rangle_{E_1} |\vec{v}\rangle_{E_2}, \tag{8.13}$$

where $p_{\vec{u},\vec{v}} = \prod_{i=1}^m p_{u_i,v_i}$. We now need to calculate the state resulting from noisy preprocessing followed by a blockwise stabilizer code measurement in which the stabilizers contain Pauli $\mathcal{I}$ and Z operators only.

Local Randomization

The first step, local randomization, can be described in a coherent way by adding a classical register $\mathbf{A'}$ (such systems will be denoted with boldface type) in the state $\big((1-q)|0\rangle\langle 0| + q|1\rangle\langle 1|\big)^{\otimes m}$ and then

applying controlled not gates from the individual register states to the bits A. This leads to

$$|\sigma'\rangle_{ABE} = \sum_{\vec{u},\vec{v},\vec{f}} \sqrt{p_{\vec{u},\vec{v}} q_{\vec{f}}}\, XZ_B(\vec{u}+\vec{f},\vec{v}) |\Phi^+\rangle_{AB}^{\otimes m} |\vec{f}\rangle_{\mathbf{A}'} |\vec{u}\rangle_{E_1} Z_{E_2}^{\vec{f}} |\vec{v}\rangle_{E_2}, \tag{8.14}$$

where $\vec{f} \in \mathbb{F}_2^m$ and $q_{\vec{f}} = q^f (1-q)^{m-f}$ for $f = |\vec{f}|$, the number of 1s in $\vec{f}$, a notation we shall use throughout. Here we have used the fact that $X_A|\Phi^+\rangle_{AB} = X_B|\Phi^+\rangle_{AB}$ (compare with lemma C.3.1) to simplify the expression; this move is responsible for the $Z^{\vec{f}}$ operation applied to E_2.

Syndrome Measurement

In the second step, Alice and Bob both measure the $m-1$ (generators of the) stabilizers of a $\mathcal{I}/Z$-only stabilizer code which encodes one logical qubit into m physical qubits. Using a public (authenticated) channel, Alice sends her syndrome to Bob who calculates the relative syndrome $\vec{s}$ by adding Alice's string to his measurement outcome modulo two. Afterwards both decode their encoded state. Such a stabilizer code is a CSS code constructed from classical linear codes $\mathcal{C}_2 \subset \mathcal{C}_1$, where $\mathcal{C}_2 = \{\vec{0}\}$ contains only the zero codeword and $\mathcal{C}_1 = \{\vec{0}, \vec{\mu}_1^x\}$ is spanned by a single codeword $\vec{\mu}_1^x$ (compare with section 6.3). Together with an encoding $U_{\text{enc}}|\vec{e},c\rangle = \overline{|\vec{e},c\rangle}$, where

$$\overline{|\vec{e},c\rangle} = \big|c\cdot\vec{\mu}_1^x + \sum_{j=1}^{m-1} e_j \cdot \vec{\eta}_j^x\big\rangle, \tag{8.15}$$

our CSS code is completely specified by defining two bases $\{\vec{\xi}_1^z,\ldots,\vec{\xi}_{m-1}^z,\vec{\mu}_1^z\}$ and $\{\vec{\eta}_1^x,\ldots,\vec{\eta}_{m-1}^x,\vec{\mu}_1^x\}$ both spanning $\mathbb{F}_2^m$ and satisfying condition (6.29) (see section 6.3.1). In this case the stabilizers are given by $\overline{Z}_i = XZ(\vec{0},\vec{\xi}_i^z)$, $i = 1\ldots m-1$, and a measurement of these stabilizers on the encoded state (8.15) will give the syndrome $\vec{e}$. Measurement of the logical Z operator $\overline{Z}_m = XZ(\vec{0},\vec{\mu}_1^z)$ gives the value of the encoded bit c. Applying one of the $\overline{X}_i = XZ(\vec{\eta}_i^x,\vec{0})$, $i = 1\ldots m-1$, operators on a encoded state results in a flip of the i-th bit of the syndrome, while applying the logical X operator $\overline{X}_m = XZ(\vec{\mu}_1^x,\vec{0})$ flips the encoded bit, $c \mapsto c\oplus 1$. Both the set of all $\overline{Z}_i$ and the set of all $\overline{X}_j$ are complete sets of commuting observables. Note that because of lemma C.3.2,

$$|\Phi^+\rangle_{AB}^{\otimes m} = U_{\text{enc},A}^* \otimes U_{\text{enc},B} |\Phi^+\rangle_{AB}^{\otimes m} = \frac{1}{\sqrt{2^{m-1}}} \sum_{\vec{e}\in\mathbb{F}_2^{m-1}} \frac{1}{\sqrt{2}} \sum_{c\in\mathbb{F}_2} \overline{|\vec{e},c\rangle}_A^* \overline{|\vec{e},c\rangle}_B. \tag{8.16}$$

In other words, the maximally-entangled state of m physical qubits is the equal superposition of a logical maximally-entangled state in all the possible encodings. Also note that lemma 6.3.1 tells us that any m fold Pauli operator can be decomposed as

$$XZ(\vec{u}',\vec{v}) = \overline{X}_m^{l^x} \overline{Z}_m^{l^z} \prod_{i=1}^{m-1} \overline{X}_i^{s_i^x} \overline{Z}_i^{n_i^x}, \tag{8.17}$$

where $s_i^x = \vec{\xi}_i^z \cdot \vec{u}'$, $n_i^x = \vec{\eta}_i^x \cdot \vec{v}$, and $l^x = \vec{\mu}_1^z \cdot \vec{u}'$ and $l^z = \vec{\mu}_1^x \cdot \vec{v}$ are the logical bit and phase flip errors resulting when this Pauli operator is applied to an encoded state like (8.15). Using these two facts we find that, after Bob's calculation of the relative syndrome $\vec{s}$, the tripartite state can be expressed as (up to a local unitary acting only on Eve's systems)

$$|\sigma''\rangle_{ABE} = \sum_{\vec{u},\vec{v},\vec{f}} \sqrt{p_{\vec{u},\vec{v}} q_{\vec{f}}}\, XZ_B\big(\vec{\mu}_1^z\cdot(\vec{u}+\vec{f}),\vec{\mu}_1^x\cdot\vec{v}\big) |\Phi^+\rangle_{AB} |\vec{f}\rangle_{\mathbf{A}'} |\vec{u}\rangle_{E_1} Z_{E_2}^{\vec{f}} |\vec{v}\rangle_{E_2} |\vec{s}\rangle_{\mathbf{B}'}, \tag{8.18}$$

where $\vec{s} = (\vec{\xi}_1^z\cdot(\vec{u}+\vec{f}),\ldots,\vec{\xi}_{m-1}^z\cdot(\vec{u}+\vec{f}))$. While the registers A and B in equation (8.14) have been m-qubit registers, here they contain only a single qubit each. Alice missing $(m-1)$-qubits have been traced out since they contained only classical information about her absolute syndrome (accessible to all parties). The rest of Bob's m-qubit register now contains classical information about the relative syndrome $\vec{s}$ and is labeled $\mathbf{B}'$.

Key Bit Measurement

Finally, Alice and Bob both measure their key bit. Alice forgets about which bits she flipped by tracing out the $\mathbf{A}'$ register. The correlations between Alice, Bob, and Eve are described by the following semiclassical state:

$$\sigma_{XY\overline{E}} = \frac{1}{2}\sum_x [x]_A \otimes \sum_{\vec{u},\vec{f}} \sum_{\vec{v}_1,\vec{v}_2} \sqrt{p_{\vec{u},\vec{v}_1} p_{\vec{u},\vec{v}_2}}\, q_{\vec{f}}\, [x + \vec{\mu}_1^z \cdot (\vec{u}+\vec{f})]_B \\ \otimes [\vec{s}]_{B'} \otimes [\vec{u}]_{E_1} \otimes (Z^{\vec{\mu}_1^x})^x_{E_2} Z^{\vec{f}}_{E_2} |\vec{v}_1\rangle\langle\vec{v}_2| Z^{\vec{f}}_{E_2} (Z^{\vec{\mu}_1^x})^x_{E_2}, \quad (8.19)$$

where $[x]_B = |x\rangle\langle x|_B$, etc. Note that the state is diagonal in E_1 since the quantities $\vec{\xi}_i^z \cdot (\vec{u}+\vec{f})$, $i = 1, \ldots, m-1$, and $\vec{\mu}_1^z \cdot (\vec{u}+\vec{f})$ are all classical: The former are already classical in (8.18), the latter became classical after the key bit measurements by Alice and Bob. The $\{\vec{\xi}_1^z, \ldots, \vec{\xi}_{m-1}^z, \vec{\mu}_1^z\}$ span $\mathbb{F}_2^m$ thereby completely fixing the string $\vec{u}+\vec{f}$.

The Mutual Information between Alice and Bob and Alice and Eve

To calculate the quantum mutual information between Alice and Bob we trace out Eve and obtain

$$\begin{aligned}
\sigma_{XY} &= \frac{1}{2}\sum_x [x]_A \otimes \sum_{\vec{u},\vec{f}} p_{\vec{u}} q_{\vec{f}}\, [x + \vec{\mu}_1^z \cdot (\vec{u}+\vec{f})]_B \otimes \left[\left(\vec{\xi}_1^z \cdot (\vec{u}+\vec{f}), \ldots\right)\right]_{B'} \\
&= \frac{1}{2}\sum_x [x]_A \otimes \sum_{\vec{u}} \tilde{p}_{\vec{u}}\, [x + \vec{\mu}_1^z \cdot \vec{u}]_B \otimes [(\vec{\xi}_1^z \cdot \vec{u}, \vec{\xi}_2^z \cdot \vec{u}, \ldots)]_{B'} \\
&= \frac{1}{2}\sum_x [x]_A \otimes \sum_{l^x,\vec{s}} \tilde{P}(l^x, \vec{s})[x + l^x]_B \otimes [\vec{s}]_{B'}, \qquad (8.20)
\end{aligned}$$

where $\tilde{p}_{\vec{u}}$ is defined as $\tilde{p}_{\vec{u}} = \tilde{p}^u (1-\tilde{p})^{m-u}$ with $\tilde{p} = p(1-q) + (1-p)q$. In the last step we used $\vec{u} = l^x \vec{\mu}_1^x + \sum_{i=1}^{m-1} s_i \vec{\eta}_i^x$ to write the sum over $\vec{u}$ as a sum over l^x and $\vec{s}$, where l^x is the logical X error, i.e. X error on the first qubit in the block; i. e. we have $\tilde{P}(l^x, \vec{s}) = \tilde{p}_{\vec{u}(l^x,\vec{s})}$. This immediately yields

$$I(X:Y) = 1 - \sum_{\vec{s}\in\mathbb{F}_2^{m-1}} \tilde{P}(\vec{s}) H_2(\tilde{P}(l^x|\vec{s})), \qquad (8.21)$$

using the binary entropy $H_2(x) = -x\log_2 x - (1-x)\log_2(1-x)$. Note that $I(X:Y)$ does not depend on the particular values $\{p_{uv}\}$ in σ_{AB} (see (8.12)), but only depends on the bit error rate $p = p_{10} + p_{11}$. The form of the mutual information indicates the advantage provided by the syndrome. If Alice did not send any information, Bob's state would be averaged over the possible syndromes, and the mutual information would involve the entropy of the average of the $\tilde{P}(l^x|\vec{s})$ rather than the average of the entropies. By concavity of entropy, the latter rate is larger.

To calculate the quantum mutual information between Alice and Eve, we trace out Bob's systems and obtain

$$\sigma_{X\overline{E}} = \frac{1}{2}\sum_x [x]_A \otimes \rho^{(x)}_{E_1E_2}, \qquad (8.22)$$

$$\rho^{(x)}_{E_1E_2} = \sum_{\vec{u}} p_{\vec{u}}\, [\vec{u}]_{E_1} \otimes \rho^{(x),\vec{u}}_{E_2}, \qquad \text{and} \qquad (8.23)$$

$$\rho^{(x),\vec{u}}_{E_2} = (Z^{\vec{\mu}_1^x})^x \sum_{\vec{f}} q_{\vec{f}} Z^{\vec{f}} |\Psi_{|\vec{u}}\rangle\langle\Psi_{|\vec{u}}| Z^{\vec{f}} (Z^{\vec{\mu}_1^x})^x, \qquad (8.24)$$

with

$$|\Psi_{|\vec{u}}\rangle = \sum_{\vec{v}} \sqrt{p_{\vec{v}|\vec{u}}}\, |\vec{v}\rangle. \qquad (8.25)$$

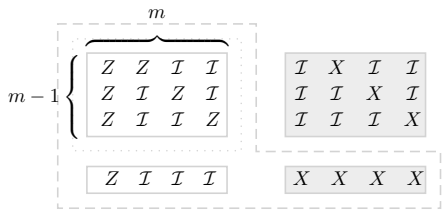


Figure 8.2: Cat code encoding one qubit into $m = 4$. The operators on the left hand side are the $\{XZ(\vec{0}, \vec{\xi}_i^z)\}$ ($i = 1 \dots m-1$ from *top* to *bottom*) and $XZ(\vec{0}, \vec{\mu}_1^z)$, those on the right hand side are $\{XZ(\vec{\eta}_i^x, \vec{0})\}$ and $XZ(\vec{\mu}_1^x, \vec{0})$. The (generators of the) stabilizers are within the dotted line, the (generators of the) normalizers within the dashed one.

It follows that the quantum mutual information between Alice and Eve is given by

$$I(X:E) = \sum_{\vec{u} \in \mathbb{F}_2^m} p_{\vec{u}} \Big[S\Big(\frac{1}{2}\rho_{E_2}^{(0),\vec{u}} + \frac{1}{2}\rho_{E_2}^{(1),\vec{u}}\Big) - S\Big(\rho_{E_2}^{(0),\vec{u}}\Big)\Big]. \tag{8.26}$$

We now restrict ourselves to the cat code presented in subsection 7.5.3, which is given by $(\vec{\xi}_i^z)_j = \delta_{1j} + \delta_{i+1,j}$ for $i = 1 \dots m-1$, $(\vec{\mu}_1^z)_j = \delta_{1j}$ and $(\vec{\eta}_i^x)_j = \delta_{i+1,j}$ for $i = 1 \dots m-1$, $(\vec{\mu}_1^x)_j = 1$ (see figure 8.2 which is the same as figure 7.1a). This code leads to the correct coherent description of the syndrome calculation of the combined preprocessing scheme. The name comes from the fact that $\alpha|\overline{\vec{0}, 0}\rangle + \beta|\overline{\vec{0}, 1}\rangle = \alpha|00\dots0\rangle + \beta|11\dots1\rangle$, a Schrödinger cat state when $\alpha = \beta = \frac{1}{\sqrt{2}}$. For the cat code we obtain the probability distribution $\tilde{P}(l^x, \vec{s})$ in the mutual information between Alice and Bob by summing equation (7.86) over $l^z \in \{0, 1\}$,

$$\begin{aligned} P(l^x, \vec{s}) &= (p_{10} + p_{11})^{l^x(m-2s)+s}(1 - p_{10} - p_{11})^{(1-l^x)(m-2s)+s} \\ &= (p^s(1-p)^{m-s})^{1-l^x}(p^{m-s}(1-p)^s)^{l^x}, \end{aligned} \tag{8.27}$$

and by replacing p with $\tilde{p}$.

We proceed with the computation of the mutual information between Alice and Eve for the BB84 and the 6-state protocol separately in the following two subsections. Before we step into these calculations, let us examine the special case $q = 0$ which can be treated without specifying the protocols: In expression (8.21) for the mutual information between Alice and Bob we simply have to replace $\tilde{p}$ with p. To calculate the mutual information between Alice and Eve given by (8.26), we note that $\rho_{E_2}^{(x),\vec{u}}$ is now a pure state. Using the fact that $\langle \Psi_{|\vec{u}} | Z^{\vec{\mu}_1^x} | \Psi_{|\vec{u}} \rangle = 1 - 2P(l^z = 0|\vec{u})$, we find that

$$I_{q=0}(X:E) = \sum_{\vec{s} \in \mathbb{F}_2^{m-1}, l^x \in \mathbb{F}_2} P(\vec{s}, l^x) H_2\big(\{P(l^z|l^x, \vec{s})\}\big). \tag{8.28}$$

Hence, we have the following theorem which already emerged as a result of Shor and Preskill's security proof in subsection 8.1.2.

Theorem 8.2.1. *The secure key rate of the BB84 protocol [6-state protocol] involving only the syndrome calculation part of the combined preprocessing scheme is given by*

$$r_{q=0}(m,p) = \frac{1}{m} \min_{\sigma_{AB} \in \Gamma} \Big(1 - \sum_{\vec{s} \in \mathbb{F}_2^{m-1}} P(\vec{s})\, H_{4[\log_2]}\big(\{P(l^x, l^z|\vec{s})\}\big)\Big), \tag{8.29}$$

where $P(l^x, l^z|\vec{s}) = P(l^x, l^z, \vec{s})/P(\vec{s})$ is the conditional error probability for the cat code, the joint probability $P(l^x, l^z, \vec{s})$ of which is given by equation (7.86), and the set Γ contains the Bell diagonal states characterized by the probability distribution $\{p_{uv}\} = \{1-2p+t, p-t, t, p-t\}_{t \in [0,p]}$ [$\{p_{uv}\} = \{1-\frac{3}{2}p, \frac{p}{2}, \frac{p}{2}, \frac{p}{2}\}$].

Remark (i). Note that (apart from the minimization) the secure rate of the above theorem is exactly the rate at which we can send quantum information reliably over a Pauli channel characterized by the probability distribution $\{p_{uv}\}$ when using a concatenation of a random outer CSS code with an inner cat code (see theorem 7.4.1 and the following remarks). Therefore, as we already mentioned in subsection 7.5.3, results on the maximum tolerable noise of the qubit depolarizing channel characterized by $\{1-p, \frac{p}{3}, \frac{p}{3}, \frac{p}{3}\}$ can be applied to the 6-state protocol if the factor 2/3 is taken into account [Lo01]. In particular it was shown in subsection 7.5.3 that the highest robustness is obtained for $m = 5$ leading to maximal tolerable bit error rate of $p_{\max}^{\text{6-st.}}(m = 5, q = 0) = 12.6904\%$. As it will be shown later, the minimum for the BB84 protocol is achieved for independent errors, $\{(1-p)^2, p(1-p), p^2, p(1-p)\}$, and it turns out that the optimal block length is $m = 7$ leading to $p_{\max}^{\text{BB84}}(m = 7, q = 0) = 11.2107\%$.

Remark (ii). If we use no preprocessing at all, we obtain the secure key rates from (8.29) by setting $m = 1$,

$$r_{q=0}(m = 1, p) = \min_{\sigma_{AB} \in \Gamma} \Big(1 - H_{4[\log_2]}\big(\{p_{uv}\}\big) \Big). \tag{8.30}$$

If we leave aside the minimization, this is exactly the rate at which we can send quantum information reliably over a Pauli channel characterized by the probability distribution $\{p_{uv}\}$ when using a random CSS code (see theorem 7.3.3 and (8.7)). For the BB84 protocol, the minimum is achieved for independent errors and we obtain the rate [SP00]

$$r_{\text{SP}}(p) = 1 - 2H_2(p). \tag{8.31}$$

Secure key generation becomes impossible for bit error rates higher than $p_{\max}^{\text{BB84}}(m = 1, q = 0) = 11.0028\%$. For the 6-state protocol, the minimization is obsolete. We obtain the rate [Lo01]

$$r_{\text{Lo}}(p) = 1 - H_2(3p/2) + \frac{3p}{2} \log_2 3, \tag{8.32}$$

and secure key generation becomes impossible for bit error rates higher than $p_{\max}^{\text{6-st.}}(m = 1, q = 0) = 12.6193\%$.

BB84

To calculate the secure key rate of the combined preprocessing scheme for the BB84 protocol, we must find the minimum over all σ_{AB} of the difference between the quantum mutual information between Alice and Bob and Alice and Eve. Since $I(X : Y)$ does not depend on the particular structure of $\{p_{uv}\} = \{1 - 2p + t, p - t, t, p - t\}$, $t \in [0, p]$, in σ_{AB}, but only depends on the bit error rate $p = p_{10} + p_{11}$, this corresponds to finding the maximum of $I(X : E)$. Let us assume for a moment that this maximum is achieved for independent bit and phase errors, i. e. we consider the state σ_{AB} with $\{p_{uv}\} = \{1 - 2p + t, p - t, t, p - t\}$ and $t = p^2$. In this case $|\Psi_{|\vec{u}}\rangle$ does not depend on $\vec{u}$, and we get

$$\rho_{E_2}^{(x),\vec{u}} = (Z^{\vec{\mu}_1^x})^x \rho_{pq}^{\otimes m} (Z^{\vec{\mu}_1^x})^x \tag{8.33}$$

with $\rho_{pq} = (1-q)|\varphi_+\rangle\langle\varphi_+| + q|\varphi_-\rangle\langle\varphi_-|$ and $|\varphi_\pm\rangle = \sqrt{1-p}|0\rangle \pm \sqrt{p}|1\rangle$. Part E_1 and E_2 of the state $\rho_{E_1E_2}^{(x)}$ in (8.23) are now completely decoupled. As it was shown in [SRS08], the fact that E_1 is classical allows the corresponding state describing dependent errors to be reconstructed from this state: After tracing out the E_1 part, we add an ancilla $[0]_{E_3}$, apply the isometry $\sum_{\vec{u},\vec{v}} \sqrt{p_{\vec{u}|\vec{v}}} |\vec{u}\rangle_{E_3} \langle 0| \otimes [\vec{v}]_{E_2}$ and eventually dephase the ancilla. Since quantum mutual information never increases under local operations, the maximum of $I(X : E)$ is indeed achieved for independent errors and (8.26) becomes

$$I(X : E) = S\Big(\frac{1}{2}\rho_{pq}^{\otimes m} + \frac{1}{2}(Z\rho_{pq}Z)^{\otimes m}\Big) - mS\big(\rho_{pq}\big). \tag{8.34}$$

Subtraction of (8.34) from (8.21) gives the secure key rate:

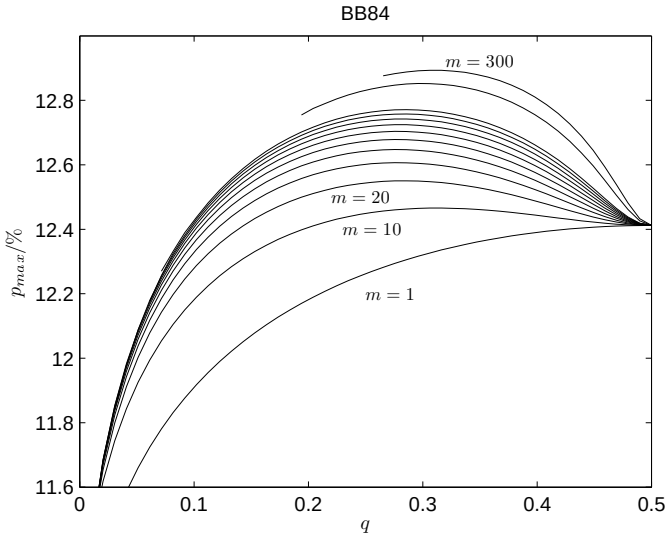


Figure 8.3: Highest tolerable bit error rate $p_{\max}^{\mathrm{BB84}}$ of the BB84 protocol as a function of the added noise q for different block lengths $m \in \{1, 10, 20, \ldots, 90, 100, 200, 300\}$.

Theorem 8.2.2. *The secure key rate of the BB84 protocol involving the combined preprocessing scheme is given by*

$$r(m,p) = \max_q \frac{1}{m}\Big[1 - \sum_{s=0}^{m-1} \binom{m-1}{s} \tilde{P}(s) H_2\Big(\tilde{P}(l^x|s)\Big) - S\Big(\frac{1}{2}\rho_{pq}^{\otimes m} + \frac{1}{2}(Z\rho_{pq}Z)^{\otimes m}\Big) + mH_2\Big(\tfrac{1}{2}(1+\sqrt{1-16p(1-p)q(1-q)})\Big)\Big]. \quad (8.35)$$

Remark. Without the use of the cat code (i. e. if we take $m = 1$) the rate reduces to [KGR05; RGK05]

$$r(p) = \max_q \Big[1 - H_2(\tilde{p}) - H_2(p) + H_2\Big(\frac{1}{2}(1+\sqrt{1-16p(1-p)q(1-q)})\Big)\Big]. \quad (8.36)$$

Omitting the maximization over q, the above formula (8.35) gives the key rate $r_{m,q}(p)$ for some fixed values of m and q as a function of the bit error rate p. By setting $r_{m,q}(p)$ equal to zero, we find $p_{\max}^{\mathrm{BB84}}(m,q)$, the maximum tolerable bit error rate for given m and q. For very high levels of added noise, i. e. for $q = \frac{1}{2} - \epsilon$, we find that for all values of m, the key rate becomes zero at the bit error rate $p_{\max}^{\mathrm{BB84}}(m, q = \frac{1}{2} - \epsilon) = 12.4120\%$, but by adding less noise at higher values of m, secret keys can be generated for even larger bit error rates (compare with figure 8.3). Figure 8.4 shows plots of the key rates given by (8.31) and (8.36) (black) and the maximum over the key rates given by (8.35) (red) for values of m up to 250. The increase of the maximal tolerable bit error rate with the block length m is illustrated in figure 8.1. The highest value of m for which we maximized the tolerable bit error rate as function of the added noise q was $m = 500$ leading to $p_{\max}^{\mathrm{BB84}}(m = 500, q = 0.32656) = 12.9379\%$.

By far the most difficult part in the numerical evaluation of (8.35) is computing the von Neumann entropy, as it contains a sum of two m-fold tensor products of different one qubit density operators. Such an expression can be more efficiently calculated by taking into account its block diagonal structure which follows from permutation invariance, as detailed in the next subsection.

6-State

Since the set Γ only contains the single state $\{p_{uv}\} = \{1 - \frac{3}{2}p, \frac{p}{2}, \frac{p}{2}, \frac{p}{2}\}$, minimization over σ_{AB} is unnecessary and the secure key rate is directly given by the difference of the quantum mutual informations

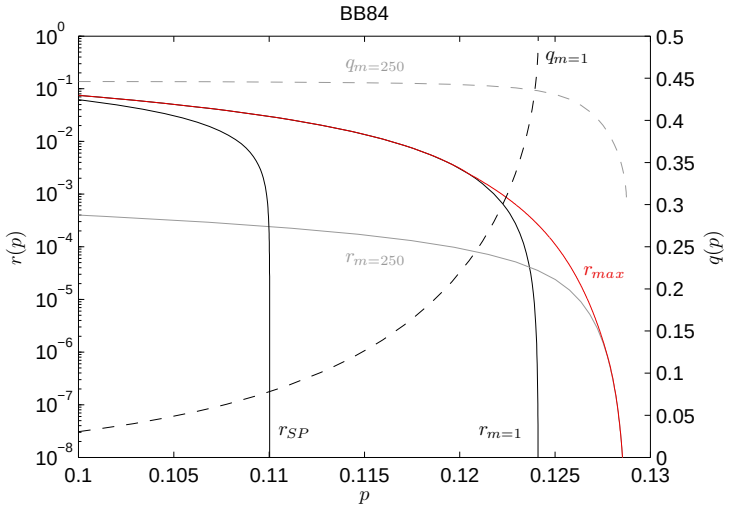


Figure 8.4: Secure key rate r of BB84 for various types of preprocessing versus bit error rate p. No preprocessing corresponds to r_{SP}, noisy preprocessing to $r_{m=1}$, and the maximum over all block lengths $m \leq 250$ to $r_{\max}$, shown in red. For the rates achieved by the blocklengths $m = 1$ and $m = 250$, the corresponding rate of the added noise is shown on the right y axis.

between Alice and Bob (8.21) and Alice and Eve. Despite the simplicity of Γ, this calculation is more difficult than BB84 due to the correlation between bit and phase errors. The corresponding conditional probabilities are given by $p_{v=1|u=0} = \frac{p}{2(1-p)} = p'$, $p_{v=0|u=0} = 1-p'$ and $p_{v|u=1} = \frac{1}{2}$. Therefore, denoting the number of ones in $\vec{u}$ as u and by reordering the qubits in such a way that the first u qubits are the ones with $u_i = 1$, we get

$$|\Psi_{|\vec{u}}\rangle = \sum_{\vec{v}} \sqrt{p_{\vec{v}|\vec{u}}}|\vec{v}\rangle = |+\rangle^{\otimes u} \otimes |\varphi'_+\rangle^{\otimes m-u} = |\Psi_{|u}\rangle \tag{8.37}$$

with $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and $|\varphi'_\pm\rangle = \sqrt{p'}|0\rangle \pm \sqrt{1-p'}|1\rangle$, leading to

$$\begin{aligned}\rho_{E_2}^{(x),u} &= (Z^{\vec{\mu}_1^x})^x \sum_{\vec{f}} q_{\vec{f}} Z^{\vec{f}} [+]^{\otimes u} \otimes [\varphi'_+]^{\otimes m-u} Z^{\vec{f}} (Z^{\vec{\mu}_1^x})^x \\ &= (Z^{\vec{\mu}_1^x})^x \sigma^{\otimes u} \otimes \gamma^{\otimes m-u} (Z^{\vec{\mu}_1^x})^x \end{aligned} \tag{8.38}$$

with $\sigma = (1-q)[+] + q[-]$ and $\gamma = (1-q)[\varphi'_+] + q[\varphi'_-]$. Reordering the state in this manner does not change the entropy, and so will not alter the rate. Using these results the quantum mutual information between Alice and Eve (8.26) can be expressed as

$$\begin{aligned} I(X:E) = \sum_{u=0}^{m} \binom{m}{u} p^u (1-p)^{m-u} \Big[S\Big(\frac{1}{2}\sigma^{\otimes u} \otimes \gamma^{\otimes m-u} + \frac{1}{2}(Z\sigma Z)^{\otimes u} \otimes (Z\gamma Z)^{\otimes m-u}\Big) \\ - uH_2(q) - (m-u)H_2\Big(\tfrac{1}{2}(1+\sqrt{1-16p'(1-p')q(1-q)})\Big)\Big]. \end{aligned} \tag{8.39}$$

Since σ and $Z\sigma Z$ are diagonal in the same basis we are able to write the von Neumann entropy as

$$\sum_{k=0}^{u} \binom{u}{k} S\Big(\frac{q^k(1-q)^{u-k}}{2} \gamma^{\otimes m-u} + \frac{(1-q)^k q^{u-k}}{2} (Z\gamma Z)^{\otimes m-u}\Big) \tag{8.40}$$

which is of the same form as the von Neumann entropy in (8.34). Therefore the same methods for evaluation can be applied; see the next subsection.

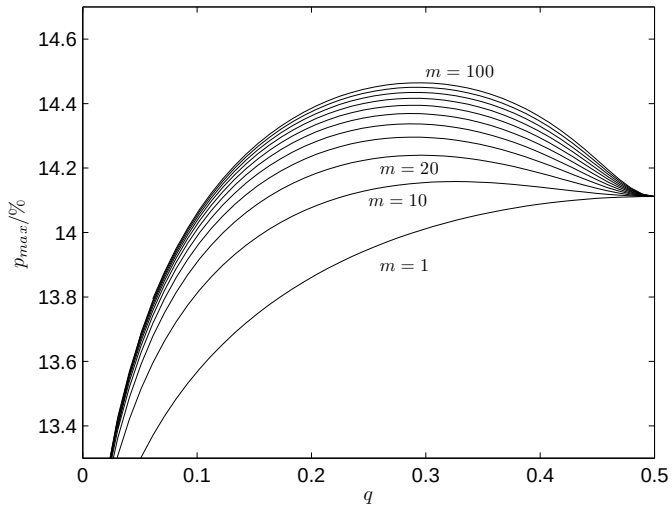


Figure 8.5: Highest tolerable bit error rate $p_{\max}^{\text{6-st.}}$ of the 6-state protocol as a function of the added noise q for different block lengths $m \in \{1, 10, 20, \ldots, 90, 100\}$.

Theorem 8.2.3. *The secure key rate of the 6-state protocol involving the combined preprocessing scheme is given by subtracting* (8.39) *from* (8.21)*,*

$$r(m,p) = \max_q \frac{1}{m}\Big[1 - \sum_{s=0}^{m-1} \binom{m-1}{s} \tilde{P}(s) H_2\big(\tilde{P}(l^x|s)\big) - I(X:E)\Big]. \tag{8.41}$$

Remark. For $m = 1$ the rate (8.41) reduces to [KGR05; RGK05],

$$r(p) = \max_q \Big[1 - H_2(\tilde{p}) - \sum_u p_u \left(H_2\big(p_{v|u}\big) - H_2\big(\tfrac{1}{2}(1 + \sqrt{1 - 16 p_{1|u}(1-p_{1|u}) q(1-q)})\big)\right)\Big], \tag{8.42}$$

As it is the case for the BB84 protocol, the key rate becomes zero for all values of m for $q \to \frac{1}{2}$ (this time at bit error rate $p_{\max}^{\text{6-st.}}(m, q = \frac{1}{2} - \epsilon) = 14.1119\%$), but again adding less noise at higher values of m gives rise to secret keys for even higher bit error rates (compare with figure 8.5). In figure 8.6 we show the key rates in these special cases as well as the general case for optimal noise and blocklengths up to $m = 125$. Included are $q = 0, m = 1$ (black), $q = 0, m = 5$ (dotted), and $m = 1$ for the optimal q (black). The maximum over the key rates given by (8.41) for values of m up to 125 is shown in red, along with the specific case of $m = 125$. The increase of the maximal tolerable bit error rate with the block length m is illustrated in figure 8.1. The highest value of m for which we maximized the tolerable bit error rate as function of the added noise q was $m = 250$ leading to $p_{\max}^{\text{6-st.}}(m = 250, q = 0.31210) = 14.5741\%$. Since the computation for larger blocksizes becomes rather slow, we extrapolated the value for the optimum noise leading to $q \approx 0.31650$ for $m = 300$. By calculating the highest tolerable bit error for this value of noise we get the best lower bound $p_{\max}^{\text{6-st.}}(m = 300, q = 0.31650) = 14.5930\%$. It seems likely that for large blocklength ($m \approx 500$) the threshold of the 6-state protocol exceeds the lowest known upper bound on the threshold for the BB84 protocol (14.6447%).

8.2.3 Evaluation of the Key Rates

To evaluate the secure key rates of the BB84 and 6-state protocols given in theorems 8.2.2 and 8.2.3 for a certain set of parameters m, p and q, in both cases a von Neumann entropy of the form

$$S\big(\alpha \cdot \rho^{\otimes n} + \beta \cdot (Z\rho Z)^{\otimes n}\big) \tag{8.43}$$

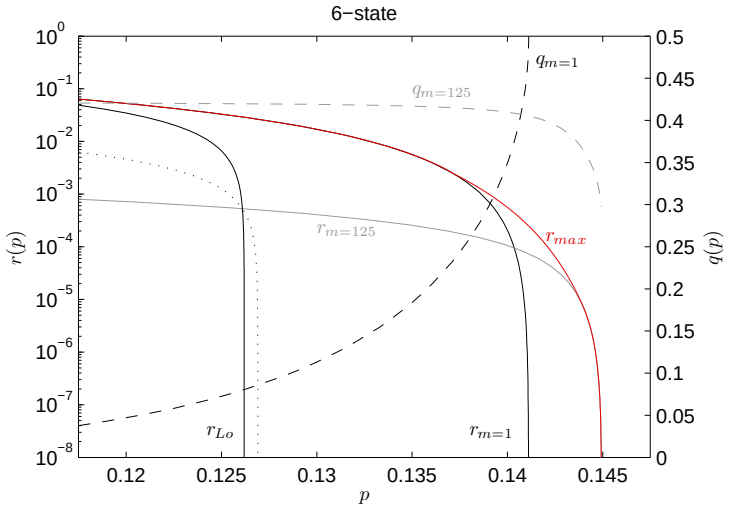


Figure 8.6: Secure key rate r of the 6-state protocol for various types of preprocessing versus bit error rate p. No preprocessing corresponds to r_{Lo}, noisy preprocessing to $r_{m=1}$, and the maximum achievable rate over all blocklengths $m \leq 125$, to $r_{\max}$, shown in red. For the rates achieved by the blocklengths $m = 1$ and $m = 125$, the corresponding rate of the added noise is shown on the right y-axis. The dotted rate with $p_{\max}^{\text{6-st.}} = 12.6904\%$ is due to Lo, corresponding to use of a repetition code of blocklength $m = 5$ and no noisy preprocessing.

with $\alpha, \beta \in \mathbb{R}$, $n \in \{1, 2, \ldots, m\}$, and

$$\rho = \begin{cases} \rho_{pq} = (1-q)[\varphi_+] + q[\varphi_-] & \text{for the BB84 protocol} \\ \gamma = (1-q)[\varphi'_+] + q[\varphi'_-] & \text{for the 6-state protocol} \end{cases} \tag{8.44}$$

with $|\varphi_\pm\rangle = \sqrt{p}|0\rangle \pm \sqrt{1-p}|1\rangle$, $|\varphi'_\pm\rangle = \sqrt{p'}|0\rangle \pm \sqrt{1-p'}|1\rangle$ and $p' = p/(2(1-p))$, has to be computed. In the following we restrict ourselves to the BB84 protocol, the corresponding results for the 6-state protocol are obtained simply by replacing p with p'. In the Bloch sphere representation, the density matrices ρ and $\sigma \equiv Z\rho Z$ are represented by non-normalized vectors

$$\vec{r}_\pm = \big(\pm 2\sqrt{p(1-p)}(1-2q), 0, 1-2p\big), \tag{8.45}$$

with $r = |\vec{r}_\pm| = \sqrt{1 - 16p(1-p)q(1-q)}$, such that

$$\rho = \frac{1}{2}\big(\mathcal{I} + \vec{r}_+ \cdot \vec{s}\big), \tag{8.46}$$

$$\sigma = \frac{1}{2}\big(\mathcal{I} + \vec{r}_- \cdot \vec{s}\big), \tag{8.47}$$

where $\vec{s} = (X, Y, Z)$ denotes a vector containing the Pauli spin matrices. The vector $\vec{r}_+$ is obtained from $\vec{r}_-$ by rotating $\vec{r}_-$ around the y-axis by the angle θ,

$$\frac{\vec{r}_+}{r} \cdot \frac{\vec{r}_-}{r} = \frac{1 - 8p(1-p)(1-2q(1-q))}{r^2} = \cos\theta. \tag{8.48}$$

If we diagonalize ρ and σ, we obtain

$$\rho = U_\rho \varrho U_\rho^\dagger, \qquad \varrho = \mathsf{diag}\{\rho_1, \rho_2\}, \tag{8.49}$$

$$\sigma = U_\sigma \varsigma U_\sigma^\dagger, \qquad \varsigma = \mathsf{diag}\{\sigma_1, \sigma_2\}, \tag{8.50}$$

and the eigenvalues $\{\rho_1, \rho_2\}$ and $\{\sigma_1, \sigma_2\}$ of ρ and σ are both given by $\{(1+r)/2, (1-r)/2\}$.

To speed up the computation of von Neumann entropies of expressions like $\alpha \cdot \rho^{\otimes n} + \beta \cdot (Z\rho Z)^{\otimes n}$, we make use of their permutation invariance. As it is discussed in subsection D.2.1 of appendix D, operators like $\rho^{\otimes n}$ become block-diagonal when expressed in the Schur basis. In other words, the reducible representation $D(\rho) = \rho^{\otimes n}$ decomposes into a direct sum of inequivalent irreducible representations $D^{(\nu)}(\rho)$ labeled by a Young diagram ν, where the irrep $D^{(\nu)}(\rho)$ occurs $h_\nu(\mathsf{S}_n)$ times and is of dimension $h_\nu(\mathsf{GL}_2)$:

$$D(\rho) \equiv \rho^{\otimes n} = \bigoplus_\nu D^{(\nu)}(\rho) \otimes \mathcal{I}_{h_\nu(\mathsf{S}_n)}. \tag{8.51}$$

In the qubit case, the summation over the Young diagrams ν becomes a summation over the index j which ranges from $0 \ldots \frac{n}{2}$ for even n and $\frac{1}{2} \ldots \frac{n}{2}$ for odd n. The dimension of the irreps $D^{(j)}(\rho)$ is given by $h_j(\mathsf{GL}_2) = 2j + 1$ and they are spanned by basis states labeled by a 'Weyl tableau' $k = -j, \ldots, +j$. Their degeneracy is given by

$$h_j(\mathsf{S}_n) = \binom{n}{n/2 - j} \frac{2j+1}{n/2 + j + 1}. \tag{8.52}$$

Diagonal density operators like ϱ and ς can easily be expressed in the j-th representation, since they are diagonal in all these representations, too. The action of $\varrho^{\otimes n}$ on basis states of the Schur basis becomes simply a multiplication by powers of the two eigenvalues because of the symmetry properties of these basis states: Each basis state of the Schur basis labeled by a certain Young diagram j and Weyl tableaux k consists of a superposition of computational basis states which are permutations of $|01\rangle^{\otimes(m/2-j)}|0\rangle^{\otimes(j-k)}|1\rangle^{\otimes(j+k)}$ independently of the Young tableaux (specifying degeneracy). Hence we obtain

$$D^{(j)}(\varrho) = \mathsf{diag}\{\rho_1^{j-k}\rho_2^{j+k}(\rho_1\rho_2)^{m/2-j}\}_{k=-j\ldots j}, \tag{8.53}$$

and an analogous expression for $D^{(j)}(\varsigma)$. To obtain the desired non-diagonal block matrices $D^{(j)}(\rho)$ $[D^{(j)}(\sigma)]$, we have to apply the unitary $U_\rho \in \mathsf{SU}_2$ $[U_\sigma \in \mathsf{SU}_2]$ in the irrep j onto $D^{(j)}(\varrho)$ $[D^{(j)}(\varsigma)]$,

$$D^{(j)}(\rho) = D^{(j)}(U_\rho) \cdot D^{(j)}(\varrho) \cdot D^{\dagger(j)}(U_\rho). \tag{8.54}$$

Since the $\mathsf{SU}_2 \subset \mathsf{GL}_2$ is locally equivalent to SO_3 (see e. g. [Tun85]), the matrices $D^{(j)}(U_\rho)$ and $D^{(j)}(U_\sigma)$ are Wigner rotation matrices. In our case these Wigner matrices describe a rotation of $\pm\theta/2$ around the y-axis (where θ is defined by eq. (8.48)) and are given simply by matrix exponentiation,

$$D^{(j)}(U_\rho) = \exp\bigl(-iJ_y \cdot \theta/2\bigr) \qquad\qquad D^{(j)}(U_\sigma) = \exp\bigl(+iJ_y \cdot \theta/2\bigr), \tag{8.55}$$

where $J_y = (J_+ - J_-)/(2i)$ and $J_\pm$ denotes the usual angular momentum ladder operators,

$$J_\pm|j,k\rangle = \sqrt{j(j+1) - k(k\pm1)}|j,k\pm1\rangle. \tag{8.56}$$

This way,

$$S\bigl(\alpha \cdot \rho^{\otimes n} + \beta \cdot \sigma^{\otimes n}\bigr) = \sum_{j=0,1/2}^{n/2} h_j(\mathsf{S}_n) \cdot S\bigl(\alpha \cdot D^{(j)}(\rho) + \beta \cdot D^{(j)}(\sigma)\bigr), \tag{8.57}$$

and it becomes feasible to calculate such expressions for values of n up to several hundreds. (Since we are only interested in the eigenvalues of $\alpha D^{(j)}(\rho) + \beta D^{(j)}(\sigma)$, in practice we might apply only a unitary which rotates by $2 \times \theta/2$ to $D^{(j)}(\varrho)$ and leave $D^{(j)}(\varsigma)$ in the diagonal form.)

8.3 Iterated Preprocessing

By combining local randomization with the cat code of size m, Alice and Bob gain an advantage over Eve and intuitively it seems this advantage might be even bigger by performing the procedure twice. In this section we discuss such a twofold iterated protocol where Alice adds noise at a rate q to m_2

blocks of size m_1 each, and then after measuring the syndromes of these blocks, adds further noise at another rate Q to the m_2 'key' bits of these blocks. Then the syndrome of these m_2 bits is measured and the remainder of the protocol proceeds as usual. We restrict ourselves to the BB84 protocol for simplicity. Using essentially the same argument as in section 8.2.2, we find that we only need to consider independent bit and phase errors described by the state σ_{AB} with $\{p_{uv}\} = \{1-2p+t, p-t, t, p-t\}$ and $t = p^2$: (i) $I(X:Y)$ depends only on the bit error rate $p = p_{10} + p_{11}$, (ii) therefore we have to find the maximum of $I(X:E)$, (iii) which is achieved for independent errors. The proof of (iii) works as in section 8.2.2, since, as we will see, E_1 of $\sigma_{X\overline{E}}$ is again classical.

8.3.1 Rate Calculation

We start with an $m_2 \times m_1$-fold tensor product of a purification of σ_{AB}, $|\sigma\rangle_{ABE}^{\otimes m_2}$, where $|\sigma\rangle_{ABE}$ is the m_1-fold tensor product which was defined in equation (8.13), (we now denote m as m_1).

First Iteration

The first step of the iterated preprocessing protocol is to apply the combined preprocessing protocol of the preceeding section to each of the m_2 blocks of size m_1. As explained in subsection 8.2.2, after this step, the i-th block of size m_1 is given by (8.18),

$$|\sigma''\rangle_{ABE}^{(i)} = \sum_{\vec{u}_i,\vec{v}_i,\vec{f}_i} \sqrt{p_{\vec{u}_i,\vec{v}_i} q_{\vec{f}_i}}\, XZ_B\big(\vec{\mu}_1^z \cdot (\vec{u}_i + \vec{f}_i), \vec{\mu}_1^x \cdot \vec{v}_i\big) |\Phi^+\rangle_{AB} |\vec{f}_i\rangle_{\mathbf{A}'} |\vec{u}_i\rangle_{E_1} Z_{E_2}^{\vec{f}_i} |\vec{v}_i\rangle_{E_2} |\vec{s}_i\rangle_{\mathbf{B}'}, \quad (8.58)$$

where $\vec{s}_i = (\vec{\xi}_1^z \cdot (\vec{u}_i + \vec{f}_i), \ldots, \vec{\xi}_{m_1-1}^z \cdot (\vec{u}_i + \vec{f}_i))$ and we added the index $i \in \{1, \ldots, m_2\}$.

Second Iteration

After adding additional noise at rate Q to the key bit of each of the m_2 blocks the state is described as

$$|\sigma'''\rangle_{ABE}^{(i)} = \sum_{\vec{f}_i,\vec{u}_i,\vec{v}_i} \sum_{F_i} \sqrt{p_{\vec{u}_i,\vec{v}_i} q_{\vec{f}_i}\, Q_{F_i}}\, XZ_B\big(\vec{\mu}_1^z \cdot (\vec{u}_i + \vec{f}_i) + F_i, \vec{\mu}_1^x \cdot \vec{v}_i\big) |\Phi^+\rangle_{AB}$$
$$\otimes |\vec{u}_i\rangle_{E_1} (Z^{\vec{\mu}_1^x})^{F_i} Z^{\vec{f}_i} |\vec{v}_i\rangle_{E_2} |\vec{s}_i\rangle_{\mathbf{B}'} |\vec{f}_i\rangle_{\mathbf{A}'} |F_i\rangle_{\mathbf{A}''} \quad (8.59)$$

with classical registers $\mathbf{B}'$, $\mathbf{A}'$ and $\mathbf{A}''$. Now we define the abbreviations $\vec{U} = \big(\vec{\mu}_1^z \cdot (\vec{u}_1 + \vec{f}_1), \ldots, \vec{\mu}_1^z \cdot (\vec{u}_{m_2} + \vec{f}_{m_2})\big)$ and $\vec{V} = \big(\vec{\mu}_1^x \cdot \vec{v}_1, \ldots, \vec{\mu}_1^x \cdot \vec{v}_{m_2}\big)$. Again Alice and Bob both measure their stabilizers (this time the cat code is of length m_2), and Alice sends her result to Bob, who calculates the relative syndrome $\vec{S} = (\vec{\xi}_1 \cdot (\vec{U} + \vec{F}), \ldots, \vec{\xi}_{m_2-1} \cdot (\vec{U} + \vec{F}))$. Both then measure their key bit. The tripartite semiclassical state describing the correlations is now given by

$$\sigma_{XY\overline{E}} = \frac{1}{2} \sum_{\vec{F}} Q_{\vec{F}} \sum_{\vec{f}_1,\ldots,\vec{f}_{m_2}} q_{\vec{f}_1} \ldots q_{\vec{f}_{m_2}} \sum_{\vec{u}_1,\ldots,\vec{u}_{m_2}} p_{\vec{u}_1} \ldots p_{\vec{u}_{m_2}}$$
$$\times \sum_x [x]_A \otimes [x + L^x]_B \otimes [\vec{s}_1, \ldots, \vec{s}_{m_2}, \vec{S}]_{B'} \otimes [\vec{u}_1, \ldots, \vec{u}_{m_2}]_{E_1}$$
$$\otimes\ (Z^{\otimes m_1 m_2})^x \bigotimes_{i=1}^{m_2} \big((Z^{\otimes m_1})^{F_i} Z^{\vec{f}_i} |\Psi\rangle\langle\Psi| Z^{\vec{f}_i} (Z^{\otimes m_1})^{F_i}\big) (Z^{\otimes m_1 m_2})^x, \quad (8.60)$$

where $|\Psi\rangle = \sum_{\vec{v}} \sqrt{p_{\vec{v}}} |\vec{v}\rangle$ and $\vec{s}_i = (\vec{\xi}_1 \cdot (\vec{u}_i + \vec{f}_i), \ldots)$, $\vec{S} = (\vec{\xi}_1 \cdot (\vec{U} + \vec{F}), \ldots)$, and $L^x = \vec{\mu}_1^z \cdot (\vec{U} + \vec{F})$. Note that, as it was the case for (8.19), the state is classical in E_1 since the quantities $\{\vec{s}_1, \ldots, \vec{s}_{m_2}, \vec{S}, L^x\}$ are all classical; $\{\vec{S}, L^x\}$ fixes $\vec{U} + \vec{F}$, and, since $(\vec{s}_i)_j = \vec{\xi}_j \cdot (\vec{u}_i + \vec{f}_i) = \vec{\xi}_j \cdot (\vec{u}_i + \vec{f}_i + F_i \cdot \vec{1})$, together with $\{\vec{s}_1, \ldots, \vec{s}_{m_2}\}$ the string $(\vec{u}_1 + \vec{f}_1 + F_1 \cdot \vec{1}, \ldots, \vec{u}_{m_2} + \vec{f}_{m_2} + F_{m_2} \cdot \vec{1})$ is fixed (compare with figure 8.7 which shows the stabilizers of the corresponding concatenated cat code).

The Quantum Mutual Informations

To calculate the quantum mutual information between Alice and Bob we trace out Eve's systems and obtain

$$\sigma_{XY} = \frac{1}{2}\sum_{\vec{F}} Q_{\vec{F}} \sum_{\vec{u}_1 \dots \vec{u}_{m_2}} \tilde{p}_{\vec{u}_1} \dots \tilde{p}_{\vec{u}_{m_2}} \sum_x [x]_A \otimes [x + L^x]_B \otimes [\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}]_{B'}, \tag{8.61}$$

using $\tilde{p} = p(1-q) + (1-p)q$. Since Alice's additional noise $\vec{f}$ is now combined with Eve's noise $\vec{u}$, $\vec{f}$ no longer appears in the the syndromes $\vec{s}_i$ and $\vec{S}$: $\vec{s}_i = (\vec{\xi}_1 \cdot \vec{u}_i, \dots) = (\vec{\xi}_1 \cdot (\vec{u}_i + F_i \cdot \vec{1}), \dots)$, $\vec{S} = (\vec{\xi}_1 \cdot (\vec{U}' + \vec{F}), \dots)$. Additionally, L^x is now $L^x = \vec{\mu}_1^z \cdot (\vec{U}' + \vec{F})$, with $\vec{U}' = (\vec{\mu}_1^z \cdot \vec{u}_1, \dots, \vec{\mu}_1^z \cdot \vec{u}_{m_2})$. Hence,

$$\sigma_{XY} = \frac{1}{2}\sum_x [x]_A \otimes \sum_{\vec{u}_1 \dots \vec{u}_{m_2}} \tilde{P}'(\vec{u}_1, \dots, \vec{u}_{m_2})[x + L^x]_B \otimes [\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}]_B, \tag{8.62}$$

with

$$\tilde{P}'(\vec{u}_1, \dots, \vec{u}_{m_2}) = \prod_{i=1}^{m_2} \left[(1-Q)\tilde{p}_{\vec{u}_i} + Q\tilde{p}_{\vec{u}_i + \vec{1}}\right] \tag{8.63}$$

and $\vec{s}_i = (\vec{\xi}_1 \cdot \vec{u}_i, \dots)$, $\vec{S} = (\vec{\xi}_1 \cdot \vec{U}', \dots)$ and $L^x = \vec{\mu}_1^z \cdot \vec{U}'$, or,

$$\sigma_{XY} = \frac{1}{2}\sum_x [x]_A \otimes \sum_{\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}, L^x} \tilde{P}'(\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}, L^x)[x + L^x]_B \otimes [\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}]_B, \tag{8.64}$$

where the probability distribution $\tilde{P}'(\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}, L^x)$ only depends on the number of ones in each of the syndromes $\vec{s}_i$ and $\vec{S}$ (we assume that the zeros and ones in $\vec{S}$ are ordered such that the syndromes $\vec{s}_i$, $i \in \{1, \dots, m_2 - S\}$, correspond to $S_i = 0$):

$$\begin{aligned}\tilde{P}'(L^x = 0, s_1 \dots s_{m_2}, S) = \prod_{i=1}^{m_2-S} &[(1-\tilde{p})^{m_1 - s_i}\tilde{p}^{s_i}(1-Q) + (1-\tilde{p})^{s_i}\tilde{p}^{m_1-s_i}Q]\times \\ &\prod_{i=m_2-S+1}^{m_2} [(1-\tilde{p})^{s_i}\tilde{p}^{m_1-s_i}(1-Q) + (1-\tilde{p})^{m_1-s_i}\tilde{p}^{s_i}Q].\end{aligned} \tag{8.65}$$

The mutual information can therefore be written as

$$I(X:Y) = 1 - \sum_{\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}} \tilde{P}'(\vec{s}_1 \dots \vec{s}_{m_2}, \vec{S}) H_2\big(\tilde{P}'(L^x | \vec{s}_1 \dots \vec{s}_{m_2}, \vec{S})\big). \tag{8.66}$$

In addition we see by the means of (8.65) that for a given value of S only the frequency distribution of the s_i, $i \in \{1, \dots, m_2 - S\}$, and the s_j, $j \in \{m_2 - S + 1, \dots, m_2\}$, matters. This fact can be used to speed up the calculation of the sum over the syndromes in (8.66),

$$\begin{aligned}I(X:Y) = 1 - \sum_{S=0}^{m_2-1} \sum_{\substack{c_0,\dots,c_{m_1-1}=0 \\ \text{s. t. } \sum_i c_i = m_2 - S}}^{S} \prod_{j=0}^{m_2-1} \binom{m_1 - 1}{j}^{c_j} \sum_{\substack{a_0,\dots,a_{m_1-1}=0 \\ \text{s. t. } \sum_i a_i = S}}^{S} \prod_{j=0}^{m_2-1} \binom{m_1-1}{j}^{a_j} \\ \tilde{P}'(s_1 \dots s_{m_2}, S) H_2\big(\tilde{P}'(L^x | s_1 \dots s_{m_2}, S)\big),\end{aligned} \tag{8.67}$$

where $(s_1, \dots, s_{m_2-S})$ contains $c_0 \times 0, \dots, c_{m_1-1} \times m_1 - 1$, and $(s_{m_2-S+1}, \dots, s_{m_2})$ contains $a_0 \times 0, \dots, a_{m_1-1} \times m_1 - 1$.

Tracing out Bob's systems from (8.60), and writing the resulting state as in (8.22)-(8.24), we obtain

$$\rho_{E_2}^{(x), \vec{u}_1 \dots \vec{u}_{m_2}} = (Z^{\otimes m_1 m_2})^x \left[(1-Q)\rho_{pq}^{\otimes m_1} + Q(Z\rho_{pq}Z)^{\otimes m_1}\right]^{\otimes m_2} (Z^{\otimes m_1 m_2})^x, \tag{8.68}$$

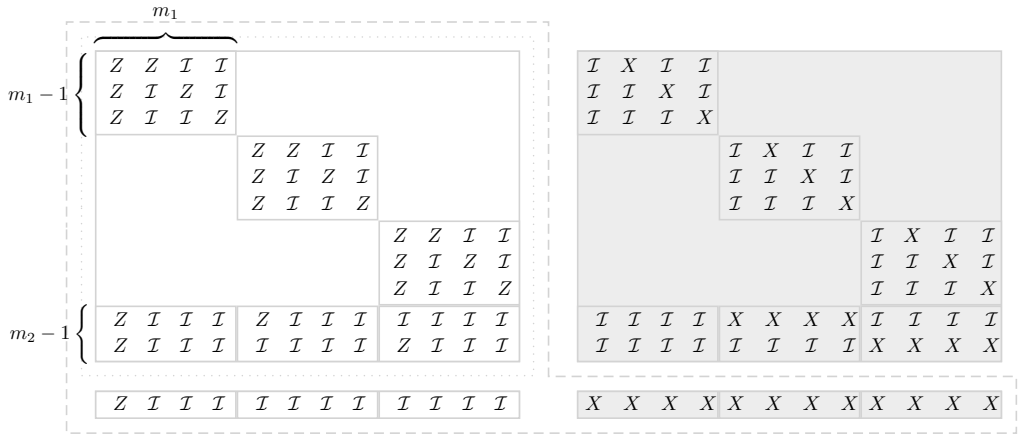


Figure 8.7: The concatenated code for the iterated preprocessing of size $m_1 = 4$ and $m_2 = 3$ encoding one qubit into $n = m_1 \times m_2$. The operators on the left hand side are the $\{\overline{Z}_i\}_{i=1\dots n}$ with $\overline{Z}_i = XZ(\vec{0}, \vec{\xi}_i^z)$ for $1 \leq i \leq n-1$ and $\overline{Z}_n = XZ(\vec{0}, \vec{\mu}^z)$, those on the right hand side the $\{\overline{X}_i\}_{i=1\dots n}$ with $\overline{X}_i = XZ(\vec{\eta}_i^x, \vec{0})$ for $1 \leq i \leq n-1$ and $\overline{X}_n = XZ(\vec{\mu}^x, \vec{0})$. The (generators of the) stabilizers are within the dotted line, the (generators of the) normalizers within the dashed one.

which does not depend on the strings $\vec{u}_1 \dots \vec{u}_{m_2}$ and the mutual information between Alice and Eve (8.26) can be seen to be

$$I(X:E) = S\Big(\frac{1}{2}\big[(1-Q)\rho_{pq}^{\otimes m_1} + Q(Z\rho_{pq}Z)^{\otimes m_1}\big]^{\otimes m_2} + \frac{1}{2}\big[Q\rho_{pq}^{\otimes m_1} + (1-Q)(Z\rho_{pq}Z)^{\otimes m_1}\big]^{\otimes m_2}\Big) - m_2 S\big((1-Q)\rho_{pq}^{\otimes m_1} + Q(Z\rho_{pq}Z)^{\otimes m_1}\big) \quad (8.69)$$

Once more the secure key rate is given by the difference of these mutual informations.

Theorem 8.3.1. *The secure key rate of the BB84 protocol involving the iterated preprocessing protocol of size $m_2 \times m_1$ is given by*

$$r(m_1, m_2, p) = \max_{q,Q} \frac{1}{m_1 m_2}\big(I(X:Y) - I(X:E)\big), \quad (8.70)$$

where the mutual informations are defined in (8.66) *and* (8.69)*.*

Again the hardest part in the numerical evaluation of (8.70) comes from the von Neumann entropies. One contains a sum of two m_2-fold tensor products of different density operators, but this time these density operators are m_1-qubit density operators. For more details on the evaluation of (8.70) see the next subsection.

We compare the resulting key rate of the $m_1 \times m_2 = 3 \times 3$ iterated code with the key rates of the non-iterated codes of blocksizes $m \in \{9, 10, 11\}$ in figure 8.8. The entire rate curve of the 3×3 code shifts to higher values than the single round $m = 9$ code, while the total amount of noise $q_{\text{tot}} = q(1-Q)+(1-q)Q$ added to the sifted key bits is essentially the same as in the case of one round, showing that the improvement comes from making better use of the same amount of noise.

8.3.2 Rate Evaluation

As it was the case for the non-iterated preprocessing protocol, to evaluate the mutual information between Alice and Eve (given by (8.70)) von Neumann entropies of the form

$$S\big(\alpha \cdot \rho^{\otimes n} + \beta \cdot (Z\rho Z)^{\otimes n}\big) \quad (8.71)$$

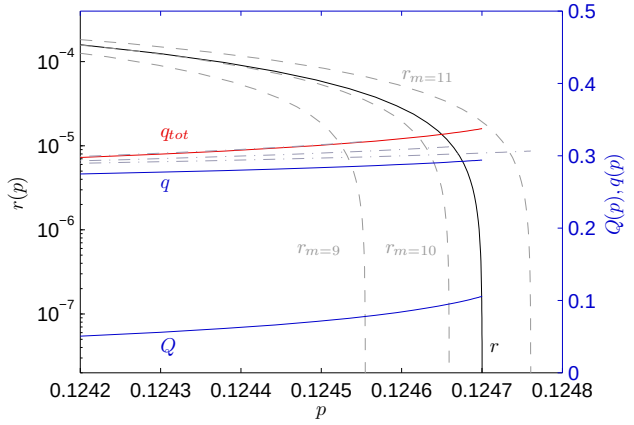


Figure 8.8: Secure key rate r of BB84 with iterated preprocessing of size $m_1 \times m_2 = 3 \times 3$ versus bit error rate p. The *right* y-axis shows the corresponding values of added noise in the first (q) and second iteration (Q) as well as values of the total amount of added noise ($q_{tot} = q(1-Q) + (1-q)Q$, red). For comparison, the rates of the non-iterated protocol are shown for blocksizes $m \in \{9, 10, 11\}$ (dashed lines). The corresponding values of added noise for these cases are also shown (dash-dot lines).

have to be evaluated. This time, in addition to the case where ρ is a qubit density operator, there is also the case where ρ is a qudit density operator of dimension 2^{m_1}. Such an expression can also be calculated more efficiently by taking into account its permutation invariance (see subsection 8.2.3),

$$S\big(\alpha \cdot \rho^{\otimes n} + \beta \cdot \sigma^{\otimes n}\big) = \sum_{\nu} h_\nu(\mathsf{S}_n) \cdot S\big(\alpha \cdot D^{(\nu)}(\rho) + \beta \cdot D^{(\nu)}(\sigma)\big), \tag{8.72}$$

(with $\sigma \equiv Z\rho Z$) but now we cannot determine the irreducible representations $D^{(\nu)}(\rho)$ by matrix multiplication from their diagonal counterparts $D^{(\nu)}(\varrho)$, because there is no simple way to determine the representation matrices of $\mathsf{SU}_{2^{m_1}}$ for $m_1 > 1$. Therefore, we explicitly calculate the Schur basis $\{|W_k^{(\nu)} Y_m^{(\nu)}\rangle\}$ (see section D.2) of n qudits of dimension 2^{m_1} with the help of the eigenfunction method [CPW02], and obtain

$$|W_k^{(\nu)} Y_m^{(\nu)}\rangle = \sum_{i_1 \ldots i_n} [U_{\mathrm{Sch}}]_{i_1 \ldots i_n}^{W_k^{(\nu)} Y_m^{(\nu)}} |i_1, \ldots, i_n\rangle. \tag{8.73}$$

Then we determine the matrix elements of both the $D^{(\nu)}(\rho)$ and the $D^{(\nu)}(\sigma)$ blocks by using the Schur basis states (8.73),

$$D_{kk'}^{(\nu)}(\rho) = \langle W_k^{(\nu)} Y_m^{(\nu)} | \rho^{\otimes n} | W_{k'}^{(\nu)} Y_m^{(\nu)}\rangle = \sum_{j_1 \ldots j_n} \sum_{i_1 \ldots i_n} [U_{\mathrm{Sch}}^*]_{j_1 \ldots j_n}^{W_k^{(\nu)} Y_m^{(\nu)}} [U_{\mathrm{Sch}}]_{i_1 \ldots i_n}^{W_{k'}^{(\nu)} Y_m^{(\nu)}} \langle j_1, \ldots, j_n | \rho^{\otimes n} | i_1, \ldots, i_n\rangle, \tag{8.74}$$

for some arbitrary Young tableau $Y_m^{(\nu)}$ which specifies the degeneracy of the irreps ν of $\mathsf{GL}_{2^{m_1}}$.

For example, to calculate the key rate of the $m_1 \times m_2 = 3 \times 3$ case presented in the last subsection, we calculated the Schur basis of

$$\begin{aligned} \mathcal{H}_8^{\otimes 3} = \mathrm{span}\Big\{|W_{k_j}^{([3])}\rangle\Big\}_{j=1\ldots120} \otimes |Y_{m_1}^{([3])}\rangle \bigoplus \\ \mathrm{span}\Big\{|W_{k_j}^{([2,1])}\rangle\Big\}_{j=1\ldots168} \otimes \mathrm{span}\Big\{|Y_{m_i}^{([2,1])}\rangle\Big\}_{i=1\ldots2} \bigoplus \\ \mathrm{span}\Big\{|W_{k_j}^{([1,1,1])}\rangle\Big\}_{j=1\ldots56} \otimes |Y_{m_1}^{([1,1,1])}\rangle, \end{aligned} \tag{8.75}$$

and the calculation of the eigenvalues of a 512×512 dimensional matrix in (8.71) reduces to a calculation of the eigenvalues of three matrices of dimension 120×120, 168×168 and 56×56 in (8.72).

It may be possible to further streamline the calculation by taking into account the fact that the qudit inputs to the second round are block-diagonal themselves. Hence more sophisticated representation-theoretic methods, in particular a Clebsch-Gordon decomposition of the states input to the second preprocessing round, should make the analysis of more rounds and larger blocksizes tractable.

Appendix

A Tables of Difference Schemes and Orthogonal Arrays

In this chapter of the appendix we list difference schemes based on $\mathbb{F}_2^2$ and orthogonal arrays with four levels. Orthogonal arrays $OA(n_c, n, 2, 4)$ with four levels and strength two can be used to build decoupling schemes of length n_c for any Hamiltonian H_0 describing a network of up to n qubits with arbitrary qubit-qubit couplings. If the couplings involve only terms of the form $J_x^{ij} X_i \otimes X_j + J_y^{ij} Y_i \otimes Y_j + J_z^{ij} Z_i \otimes Z_j$, decoupling schemes of smaller length can be obtained from difference schemes $D(n_c, n, 4)$ based on $\mathbb{F}_2^2$. The decoupling schemes $\{g_j\}_{j=0}^{n_c-1}$ are obtained by setting $g_j = u_{m_{1,j+1}} \otimes u_{m_{2,j+1}} \otimes \cdots \otimes u_{m_{n,j+1}}$, where m_{ij} denotes the matrix elements of the corresponding orthogonal array or difference scheme, and the set $\{u_i\}_{i=0}^3$ denotes the set of Pauli operators $\mathcal{P}_2 = \{\mathcal{I}, X, Y, Z\}$. (Alternatively, in the case of an orthogonal array, any nice error basis may be chosen to form the set $\{u_i\}_{i=0}^3$).

A.1 Difference Schemes

For an overview over construction methods and lower bounds on the maximal number $c \in \{2, 3, \ldots, 4\lambda\}$ for which a difference scheme $D(4\lambda, c, 4)$, $\lambda \in \mathbb{N}$, exists, we refer to [HSS99, chapter 6]. We list difference schemes $D(4\lambda, 4\lambda, 4)$ for $\lambda = \{1, 2, 3, 4\}$ in tables A.1 – A.4. The entries $\{0, 1, 2, 3\}$ are to be understood as elements in $\mathbb{F}_2^2$, $0 = (0,0), 1 = (1,0), 2 = (1,1), 3 = (0,1)$. Note that all schemes are symmetric with respect to their matrix indices, i. e. $m_{ij} = m_{ji}$. The difference schemes in tables A.1 and A.2 are the same as those presented in [SM01], the schemes in tables A.3 and A.4 have been obtained by the author via a computer search. It was conjectured in [SM01] that schemes $D(4\lambda, 4\lambda, 4)$ may exist for all $\lambda \in \mathbb{N}$. For $\lambda = 5$ at the current time only a lower bound of $c \geq 10$ is known.

A.2 Orthogonal Arrays

We list orthogonal arrays $OA(16, 5, 2, 4)$, $OA(32, 9, 2, 4)$, and $OA(48, 13, 2, 4)$ in tables A.5 – A.7. The arrays are constructed using the difference schemes listed in tables A.1 – A.3 in connection with the construction method described in [HSS99, corollary 6.20], which, for a given difference scheme $D(n_c, n, 4)$, leads to an $OA(4n_c, n+1, 2, 4)$. As a result, the upper left $(n-1) \times (n-1)$ corner of any of the listed orthogonal arrays of the form $OA(n_c, n, 2, 4)$ is identical to the corresponding difference scheme.

0	0	0	0
0	1	2	3
0	2	3	1
0	3	1	2

Table A.1: $D(4, 4, 4)$

0	0	0	0	0	0	0	0
0	0	1	1	2	2	3	3
0	1	2	3	0	1	2	3
0	1	3	2	2	3	1	0
0	2	0	2	3	1	3	1
0	2	1	3	1	3	0	2
0	3	2	1	3	0	1	2
0	3	3	0	1	2	2	1

Table A.2: $D(8, 8, 4)$

0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	1	1	2	2	2	3	3	3
0	0	0	2	2	2	3	3	3	1	1	1
0	1	2	1	2	3	0	1	3	0	2	3
0	1	2	2	3	1	1	3	0	3	0	2
0	1	2	3	1	2	3	0	1	2	3	0
0	2	3	0	1	3	2	3	1	0	1	2
0	2	3	1	3	0	3	1	2	2	0	1
0	2	3	3	0	1	1	2	3	1	2	0
0	3	1	0	3	2	0	2	1	3	2	1
0	3	1	2	0	3	1	0	2	2	1	3
0	3	1	3	2	0	2	1	0	1	3	2

Table A.3: $D(12, 12, 4)$

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3
0	0	0	0	2	2	2	2	3	3	3	3	1	1	1	1
0	0	0	0	3	3	3	3	1	1	1	1	2	2	2	2
0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
0	1	2	3	1	0	3	2	2	3	0	1	3	2	1	0
0	1	2	3	2	3	0	1	3	2	1	0	1	0	3	2
0	1	2	3	3	2	1	0	1	0	3	2	2	3	0	1
0	2	3	1	0	2	3	1	0	2	3	1	0	2	3	1
0	2	3	1	1	3	2	0	2	0	1	3	3	1	0	2
0	2	3	1	2	0	1	3	3	1	0	2	1	3	2	0
0	2	3	1	3	1	0	2	1	3	2	0	2	0	1	3
0	3	1	2	0	3	1	2	0	3	1	2	0	3	1	2
0	3	1	2	1	2	0	3	2	1	3	0	3	0	2	1
0	3	1	2	2	1	3	0	3	0	2	1	1	2	0	3
0	3	1	2	3	0	2	1	1	2	0	3	2	1	3	0

Table A.4: $D(16, 16, 4)$

0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3
0	1	2	3	1	0	3	2	2	3	0	1	3	2	1	0
0	2	3	1	1	3	2	0	2	0	1	3	3	1	0	2
0	3	1	2	1	2	0	3	2	1	3	0	3	0	2	1
0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3

Table A.5: $OA(16, 5, 2, 4)$

0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	0	1	1	2	2	3	3	1	1	0	0	3	3	2	2
0	1	2	3	0	1	2	3	1	0	3	2	1	0	3	2
0	1	3	2	2	3	1	0	1	0	2	3	3	2	0	1
0	2	0	2	3	1	3	1	1	3	1	3	2	0	2	0
0	2	1	3	1	3	0	2	1	3	0	2	0	2	1	3
0	3	2	1	3	0	1	2	1	2	3	0	2	1	0	3
0	3	3	0	1	2	2	1	1	2	2	1	0	3	3	0
0	0	1	1	2	2	3	3	0	0	1	1	2	2	3	3

2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3
2	2	3	3	0	0	1	1	3	3	2	2	1	1	0	0
2	3	0	1	2	3	0	1	3	2	1	0	3	2	1	0
2	3	1	0	0	1	3	2	3	2	0	1	1	0	2	3
2	0	2	0	1	3	1	3	3	1	3	1	0	2	0	2
2	0	3	1	3	1	2	0	3	1	2	0	2	0	3	1
2	1	0	3	1	2	3	0	3	0	1	2	0	3	2	1
2	1	1	2	3	0	0	3	3	0	0	3	2	1	1	2
0	0	1	1	2	2	3	3	0	0	1	1	2	2	3	3

Table A.6: $OA(32, 9, 2, 4)$

0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	1	1	1	2	2	2	3	3	3	1	1	1	0	0	0	3	3	3	2	2	2
0	0	0	2	2	2	3	3	3	1	1	1	1	1	1	3	3	3	2	2	2	0	0	0
0	1	2	1	2	3	0	1	3	0	2	3	1	0	3	0	3	2	1	0	2	1	3	2
0	1	2	2	3	1	1	3	0	3	0	2	1	0	3	3	2	0	0	2	1	2	1	3
0	1	2	3	1	2	3	0	1	2	3	0	1	0	3	2	0	3	2	1	0	3	2	1
0	2	3	0	1	3	2	3	1	0	1	2	1	3	2	1	0	2	3	2	0	1	0	3
0	2	3	1	3	0	3	1	2	2	0	1	1	3	2	0	2	1	2	0	3	3	1	0
0	2	3	3	0	1	1	2	3	1	2	0	1	3	2	2	1	0	0	3	2	0	3	1
0	3	1	0	3	2	0	2	1	3	2	1	1	2	0	1	2	3	1	3	0	2	3	0
0	3	1	2	0	3	1	0	2	2	1	3	1	2	0	3	1	2	0	1	3	3	0	2
0	3	1	3	2	0	2	1	0	1	3	2	1	2	0	2	3	1	3	0	1	0	2	3
0	0	0	1	1	1	2	2	2	3	3	3	0	0	0	1	1	1	2	2	2	3	3	3

2	2	2	2	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3
2	2	2	3	3	3	0	0	0	1	1	1	3	3	3	2	2	2	1	1	1	0	0	0
2	2	2	0	0	0	1	1	1	3	3	3	3	3	3	1	1	1	0	0	0	2	2	2
2	3	0	3	0	1	2	3	1	2	0	1	3	2	1	2	1	0	3	2	0	3	1	0
2	3	0	0	1	3	3	1	2	1	2	0	3	2	1	1	0	2	2	0	3	0	3	1
2	3	0	1	3	0	1	2	3	0	1	2	3	2	1	0	2	1	0	3	2	1	0	3
2	0	1	2	3	1	0	1	3	2	3	0	3	1	0	3	2	0	1	0	2	3	2	1
2	0	1	3	1	2	1	3	0	0	2	3	3	1	0	2	0	3	0	2	1	1	3	2
2	0	1	1	2	3	3	0	1	3	0	2	3	1	0	0	3	2	2	1	0	2	1	3
2	1	3	2	1	0	2	0	3	1	0	3	3	0	2	3	0	1	3	1	2	0	1	2
2	1	3	0	2	1	3	2	0	0	3	1	3	0	2	1	3	0	2	3	1	1	2	0
2	1	3	1	0	2	0	3	2	3	1	0	3	0	2	0	1	3	1	2	3	2	0	1
0	0	0	1	1	1	2	2	2	3	3	3	0	0	0	1	1	1	2	2	2	3	3	3

Table A.7: $OA(48, 13, 2, 4)$

B Quantum Algorithms for Quantum Maps

This chapter presents quantum algorithms implementing quantum maps like the quantum sawtooth map [BCMS01] and the quantum tent map [FFS04]. These algorithms have been used in this thesis to study the error suppressing properties of the PAREC method in section 3.2 and the embedded recoupling scheme in chapter 4 by means of numerical simulations. A more elaborated discussion of such algorithms can be found in the author's diploma thesis [Ker04, chapter 2]. Furthermore, we define a discrete Husimi function which can be understood as the coherent state representation of a quantum state, and which can be used to illustrate quantum states.

B.1 Quantum Gates

Before we are going to derive a decomposition of a quantum map into a sequence of elementary one- and two-qubit gates, we have to define these gates. Each of the one- and two-qubit gates will be represented in the standard computational basis $\{|0\rangle, |1\rangle\}$ and $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$, respectively. Let us start with the one-qubit gates.

B.1.1 One-Qubit Gates

Phase Gate

The phase gate $\mathsf{P}_t(\varphi)$ applies a phase φ if the t-th qubit is in the state $|1\rangle$.

$$-\boxed{\varphi}- \Leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & e^{i\varphi} \end{pmatrix}$$

Hadamard Gate

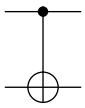
The Hadamard gate H_t generates a superposition of $|0\rangle$ and $|1\rangle$.

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

B.1.2 Two-Qubit Gates

The Controlled-Not Gate

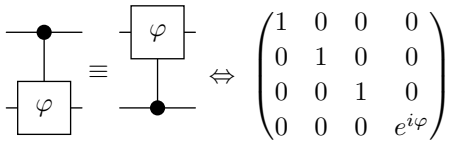
The controlled-not gate CNOT_{ct} flips the state of the target qubit t if the control qubit c is in the state $|1\rangle$.



$$\Leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The Controlled-Phase Gate

The controlled-phase gate $\mathsf{CP}_{c_1c_2}$ applies a phase φ if the control qubits c_1 and c_2 are both in the state $|1\rangle$.

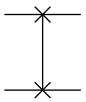


$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & e^{i\varphi} \end{pmatrix}$$

A three-qubit controlled phase gate $\mathsf{CCP}_{c_1c_2c_3}$ might be defined in a similar fashion.

The Swap Gate

The swap gate $\mathsf{SWAP}_{t_1t_2}$ exchanges the state of the target qubits t_1 and t_2.



$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

B.2 Gate Decompositions for Quantum Maps

Let us consider a quantum computer consisting of n qubits. The Hilbert space $\mathcal{H} = \mathcal{H}_2^{\otimes n}$ spanned by the computational basis $\{|i_0, i_1, \ldots, i_{n-1}\rangle\}$, with $i_j \in \{0, 1\}$ for $j = 0, 1, \ldots, n-1$, is of dimension $d = 2^n$. A short hand notation of the basis states is given by $|i\rangle = |i_0, i_1, \ldots, i_{n-1}\rangle$ with $i = \sum_{j=0}^{n-1} i_j \cdot 2^j$. We are going to construct a decomposition of a quantum map

$$U = \exp\Big(-\frac{i}{2}m^2 T\Big)\exp\Big(-ikV(q)\Big), \tag{B.1}$$

characterized by the parameters $T = 2\pi/d$ and $k \in \mathbb{R}$, into a sequence of elementary one- and two-qubit gates defined in the preceding section. Here, m denotes the momentum operator whose eigenstates form the computational basis, $m|i\rangle = i|i\rangle$, and q denotes the position operator which is related to the momentum operator via the quantum Fourier transform (QFT):

$$q = U_{\mathrm{QFT}}^{-1} \cdot \frac{2\pi}{d} m \cdot U_{\mathrm{QFT}}. \tag{B.2}$$

As a consequence, the quantum map can be written as the product of four unitaries

$$U = \exp\Big(-\frac{i}{2}m^2 T\Big) \cdot U_{\mathrm{QFT}}^{-1} \cdot \exp\Big(-ikV\big(\frac{2\pi}{d}m\big)\Big) \cdot U_{\mathrm{QFT}}. \tag{B.3}$$

Each of these unitaries, the QFT U_{QFT}, the kick operator $\exp\big(-ikV(2\pi m/d)\big)$, the inverse QFT and the free evolution operator $\exp\big(-\frac{i}{2}m^2T\big)$, can be decomposed into a sequence of elementary one- and two-qubit gates. We present gate decompositions for the kick operator employing the sawtooth-potential

$$V_{\mathrm{saw}}(q) = -\frac{1}{2}(q-\pi)^2 \tag{B.4}$$

and the tent-potential

$$V_{\mathrm{tent}}(q) = \begin{cases} -\frac{1}{2}q(q-\pi) & ,0 \le q < \pi \\ \frac{1}{2}(q-\pi)(q-2\pi) & ,\pi \le q < 2\pi \end{cases}. \tag{B.5}$$

The classical map corresponding to the quantum map (B.1) is given by

$$\begin{aligned} p' &= p - KV'(q) \pmod{2\pi} \\ q' &= q + p' \pmod{2\pi} \end{aligned} \tag{B.6}$$

and depends only on the single parameter $K = kT$.

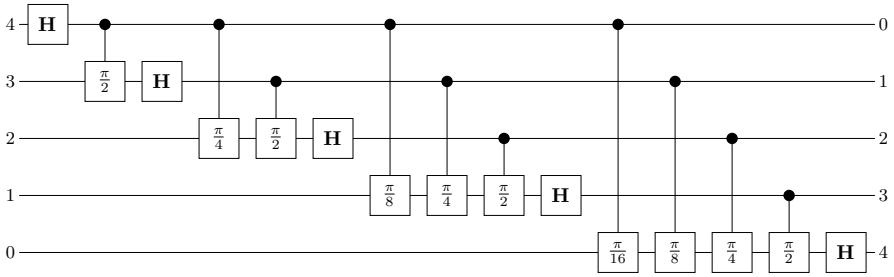


Figure B.1: Quantum circuit of the quantum Fourier transform for $n = 5$ qubits.

B.2.1 The Quantum Fourier Transform

If we let the quantum Fourier transform (QFT) reverse the order of the qubits, i. e. if

$$U_{\mathrm{QFT}}|m_0\rangle \otimes |m_1\rangle \otimes \cdots \otimes |m_{n-1}\rangle = \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} \exp\Big(i\frac{2\pi}{d}mx\Big)|x_{n-1}\rangle \otimes \cdots \otimes |x_1\rangle \otimes |x_0\rangle, \tag{B.7}$$

a decomposition of U_{QFT} into $n(n+1)/2$ quantum gates (Hadamard gates and controlled-phase gates) is given by [EJ96]

$$U_{\mathrm{QFT}} = \prod_{j=n-1}^{0} \left(\Big(\prod_{i=n-1}^{i>j} \mathsf{CP}_{ji}\Big(\frac{\pi}{2^{i-j}}\Big)\Big)\mathsf{H}_j \right), \tag{B.8}$$

where the product over j is non-commutative and has to be applied starting with $j = n-1$. A corresponding quantum circuit for $n = 5$ qubits is depicted in figure B.1. The inverse operation U_{QFT}^{-1} is obtained from (B.8) by multiplying each phase by the factor minus one.

B.2.2 The Free Evolution Operator

The free evolution operator $\exp\big(-\frac{i}{2}m^2T\big)$ is implemented by a series of controlled- and uncontrolled-phase gates. Using the binary representation $m = \sum_{j=0}^{n-1} m_j \cdot 2^j$, we obtain

$$\begin{aligned} \exp\Big(-\frac{i}{2}m^2T\Big) &= \exp\Big(-iT \sum_{v,w=0}^{n-1} m_v m_w 2^{v+w-1}\Big) \\ &= \prod_{v=0}^{n_q-1} \exp\left(-iTm_v 2^{2v-1}\right) \prod_{v<w}^{n_q-1} \exp\left(-iTm_v m_w 2^{v+w}\right), \end{aligned} \tag{B.9}$$

which translates to a series of $n(n+1)/2$ phase gates as follows:

$$\prod_{v=0}^{n-1} \mathsf{P}_v\big(-T2^{2v-1}\big) \prod_{v<w}^{n-1} \mathsf{CP}_{vw}\big(-T2^{v+w}\big). \tag{B.10}$$

B.2.3 The Kick Operator

The gate decomposition of the kick operator $\exp\big(-ikV(2\pi m/d)\big)$ depends on the detailed form of the potential V. We start with the sawtooth potential given by (B.4) and proceed with the tent potential given by (B.5).

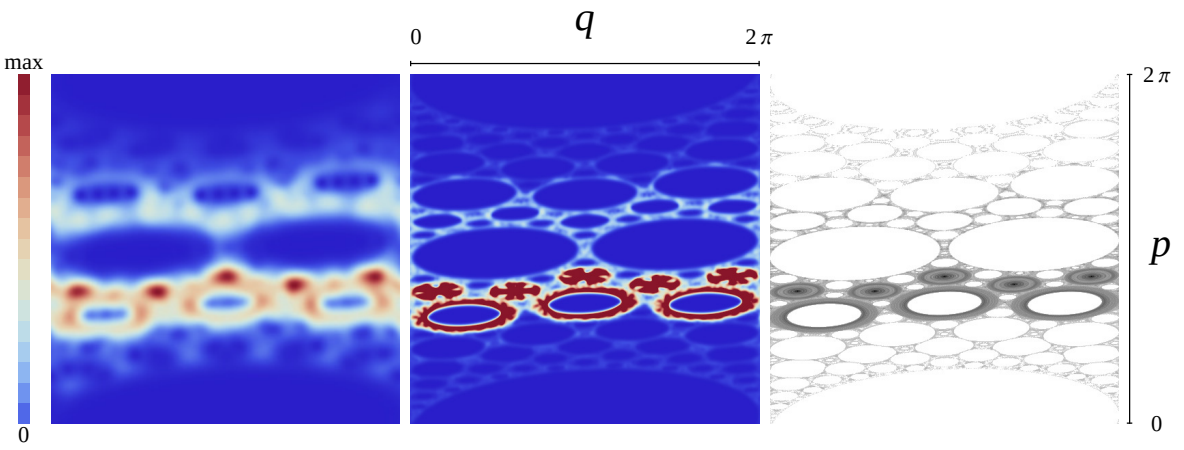


Figure B.2: Husimi function of the quantum sawtooth map with parameters $K = kT = -0.1$ and $T = 2\pi/2^n$: for $n = 8$ qubits (*left*) and $n = 12$ qubits (*middle*). Classical trajectories (*right*).

Sawtooth Map

The kick operator of the sawtooth map is given by

$$\exp\Big(i\frac{k}{2}\big(\frac{2\pi}{d}m-\pi\big)^2\Big) = \exp\Big(i\frac{2k\pi^2}{d^2}m^2\Big)\exp\Big(-i\frac{2k\pi^2}{d}m\Big)\exp\Big(i\frac{k\pi^2}{2}\Big). \tag{B.11}$$

Omitting the global phase, this translates into the sequence

$$\prod_{v=0}^{n-1}\mathsf{P}_v\Big(-\frac{2k\pi^2}{d}2^v+\frac{2k\pi^2}{d^2}2^{2v}\Big)\prod_{v<w}^{n-1}\mathsf{CP}_{vw}\Big(\frac{4k\pi^2}{d^2}2^{v+w}\Big) \tag{B.12}$$

consisting of $n(n+1)/2$ phase gates.

Hence, in total, the quantum algorithm implementing the quantum sawtooth map consists of $n_g = 4\times n(n+1)/2 = 2n(n+1)$ elementary quantum gates. The algorithm presented in this section is an improved version of the algorithm proposed by Benenti et. al. in [BCMS01] and [BCMS03] which makes use of a four-phase two-qubit gate which applies an individual phase to each of the states $\{|00\rangle,|01\rangle,|10\rangle,|11\rangle\}$ and consists of the larger number of $3n^2+n$ quantum gates in total. As an example, let us apply the sawtooth map U with parameters $K = kT = -0.1$ and $T = 2\pi/2^n$ on the initial state $|\Psi\rangle = |\,[0.38\cdot 2^n]\,\rangle$. Figure B.2 shows the average of the Husimi function of the state $U^t|\Psi\rangle$ taken over $950 \le t \le 1000$. The calculation was performed for $n = 8$ (*left part*) and $n = 12$ qubits (*middle part*). The color gradient encodes the function values ranging from 0 (*blue*) up to the maximal value (*red*). For comparison, there are also 1000 classical trajectories depicted (*right part*) starting in the range $(0 \le q < 2\pi, p \approx 0.38\cdot 2\pi)$ and resulting from 2000 iterations of the classical map (B.6).

Tent Map

Setting $\bar{q}(q) = q$ if $0 \le q < \pi$ and $\bar{q}(q) = q-\pi$ if $\pi \le q < 2\pi$, the tent-potential becomes

$$V_{\text{tent}}\big(\bar{q}(q)\big) = \begin{cases} -\frac{1}{2}\bar{q}(\bar{q}-\pi) & ,0\le q<\pi \\ +\frac{1}{2}\bar{q}(\bar{q}-\pi) & ,\pi\le q<2\pi \end{cases}. \tag{B.13}$$

In order to implement the kick operator $\exp\big(-ikV_{\text{tent}}(2\pi m/d)\big)$, we start by applying the operator $\exp\big(ik\frac{1}{2}\bar{q}(\bar{q}-\pi)\big) = \exp\big(i\frac{2k\pi^2}{d^2}\overline{m}^2\big)\exp\big(-i\frac{k\pi^2}{d}\overline{m}\big)$, where $\overline{m} = \sum_{j=0}^{n-2} m_j\cdot 2^j$ does not depend on the most

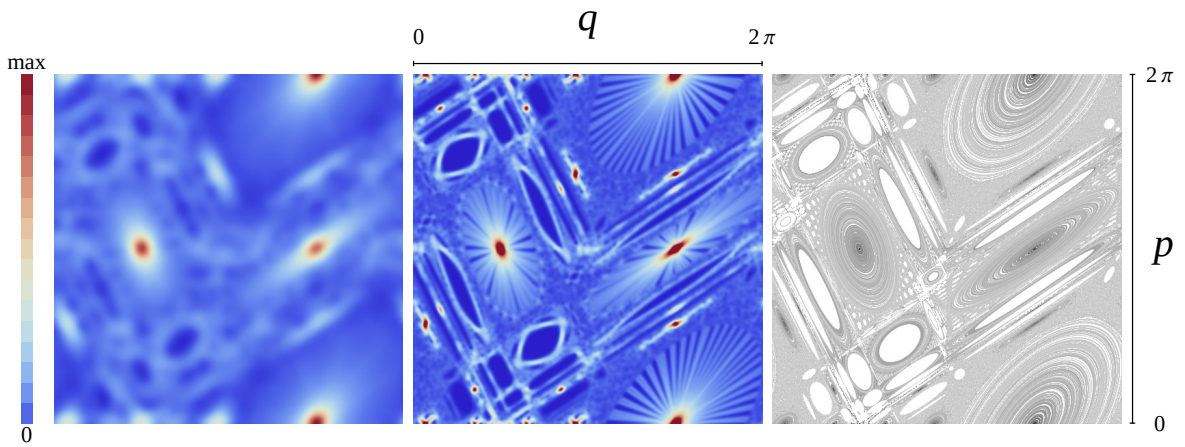


Figure B.3: Husimi function of the quantum tent map with parameters $K = kT = 4/3$ and $T = 2\pi/2^n$: for $n = 8$ qubits (*left*) and $n = 12$ qubits (*middle*). Classical trajectories (*right*).

significant qubit in position $n-1$. This operator translates into the following sequence of phase gates:

$$\prod_{v=0}^{n-2} \mathsf{P}_v\Big(-\frac{k\pi^2}{d}2^v + \frac{2k\pi^2}{d^2}2^{2v}\Big) \prod_{v<w}^{n-2} \mathsf{CP}_{vw}\Big(\frac{4k\pi^2}{d^2}2^{v+w}\Big) \tag{B.14}$$

Since states with $q \geq \pi$ should have been multiplied with $\exp\big(-ik\frac{1}{2}\bar{q}(\bar{q}-\pi)\big)$ instead, the next step is to apply the operator $\exp\big(-ik\bar{q}(\bar{q}-\pi)\big)$ onto all such states. This can be done by using the same gate sequence as in (B.14), if each phase is multiplied by the factor -2, and each gate is additionally controlled by the most significant qubit in position $n-1$:

$$\prod_{v=0}^{n-2} \mathsf{CP}_{n-1,v}\Big(\frac{2k\pi^2}{d}2^v - \frac{4k\pi^2}{d^2}2^{2v}\Big) \prod_{v<w}^{n-2} \mathsf{CCP}_{n-1,v,w}\Big(-\frac{8k\pi^2}{d^2}2^{v+w}\Big). \tag{B.15}$$

The three-qubit gate $\mathsf{CCP}_{c_1c_2c_3}$ can be implemented by the following five qubit sequence:

$$\mathsf{CCP}_{c_1c_2c_3} = \mathsf{CP}_{c_2c_1}\Big(\frac{\varphi}{2}\Big)\mathsf{CP}_{c_2c_3}\Big(\frac{\varphi}{2}\Big)\mathsf{CNOT}_{c_1c_3}\mathsf{CP}_{c_2c_3}\Big(-\frac{\varphi}{2}\Big)\mathsf{CNOT}_{c_1c_3}. \tag{B.16}$$

As a consequence, the kick operator of the tent map is decomposed into $3n^2 - 7n + 4$ elementary one- and two-qubit quantum gates.

In total, the quantum algorithm implementing the quantum tent map consists of $n_g = 3 \times n(n+1)/2 + 3n^2 - 7n + 4 = \frac{9}{2}n^2 - \frac{11}{2}n + 4$ quantum gates. It was originally proposed by Frahm et. al. in [FFS04]. As an example, let us apply the tent map U with parameters $K = kT = 4/3$ and $T = 2\pi/2^n$ on the initial state $|\Psi\rangle = (|\ 0\ \rangle + |\ 2^{n-1}\ \rangle)/\sqrt{2}$. Figure B.3 shows the average of the Husimi function of the state $U^t|\Psi\rangle$ taken over $950 \leq t \leq 1000$. The calculation was performed for $n = 8$ (*left part*) and $n = 12$ qubits (*middle part*). For comparison, there are 1000 classical trajectories shown (*right part*) starting in the range $(0 \leq q < 2\pi, p \in \{0, \pi\})$ and resulting from 2000 iterations of the classical map (B.6).

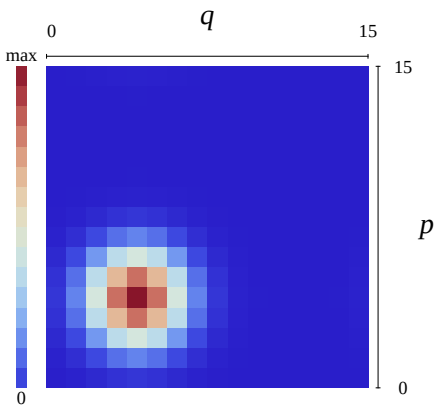


Figure B.4: Husimi function $H(q,p)$ of the coherent $n=4$ qubit state $|\Phi(4,4)\rangle$.

B.3 Coherent States and the Husimi Function

Let us consider a quantum register consisting of n qubits described by a Hilbert space $\mathcal{H} = \mathcal{H}_2^{\otimes n}$ of dimension $d = 2^n$. A coherent state in position $(0 \le q < d, 0 \le p < d)$ is defined as

$$|\Phi(q,p)\rangle = \Big(\frac{2}{d}\Big)^{\frac{1}{4}} \sum_{j=0}^{d-1} \exp\Big(-i\frac{2\pi}{d}jq - \frac{\pi}{d}D^2(j,p)\Big)|\ j\ \rangle. \tag{B.17}$$

Here, $D(j,p)$ denotes the difference $j-p$ mapped to the range $-d/2 \le D < d/2$:

$$D(j,p) = \Big(j - p + \frac{d}{2} \pmod d\Big) - \frac{d}{2}. \tag{B.18}$$

The state $|\Phi(q,p)\rangle$ is normalized in the limit of large d. A quantum algorithm which prepares a coherent state in good approximation can be found in [PRS04].

The Husimi function $H(q,p)$ of a quantum state $|\Psi\rangle = \sum_{j=0}^{d-1} \Psi_j |j\rangle$ is defined as the absolute square of the inner product between $|\Psi\rangle$ and a coherent state $|\Phi(q,p)\rangle$:

$$\begin{aligned} H(q,p) &= \frac{1}{d}\big|\langle\Phi(q,p)|\Psi\rangle\big|^2 \\ &= \Big(\frac{2}{d^3}\Big)^{\frac{1}{2}} \Big|\sum_{j=0}^{d-1} \exp\Big(i\frac{2\pi}{d}jq - \frac{\pi}{d}D^2(j,p)\Big)\Psi_j\Big|^2 \end{aligned} \tag{B.19}$$

According to the above formula, a calculation of all d^2 values of $H(q,p)$ takes $\mathcal{O}(d^3)$ steps. As it was recognized in [FFS04], this calculation can be accelerated substantially by noting that a Fourier transformation is involved in expression (B.19): Let us rewrite the equation as

$$\begin{aligned} H(q,p) &= \Big|\sum_{q'=0}^{d-1} \langle q| \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} \exp\Big(\frac{2\pi i}{d}jq'\Big)\Big[\Psi_j\Big(\frac{2}{d}\Big)^{\frac{1}{4}} \exp\Big(-\frac{\pi}{d}D^2(j,p)\Big)\Big]|q'\rangle\Big|^2 \\ &= \Big|\langle q| \sum_{q'=0}^{d-1} \tilde{\Psi}_{q'}|q'\rangle\Big|^2, \end{aligned} \tag{B.20}$$

where the vector with entries $\tilde{\Psi}_{q'}$ denotes the Fourier transform of

$$|\Psi'\rangle = \sum_{j=0}^{d-1}\Big[\Psi_j\Big(\frac{2}{d}\Big)^{\frac{1}{4}} \exp\Big(-\frac{\pi}{d}D^2(j,p)\Big)\Big]|j\rangle. \tag{B.21}$$

By calculating the fast Fourier transformation for the d vectors $|\Psi'\rangle$ associated with $p \in \{0, 1, \ldots, d-1\}$, all values of $H(q, p)$ can be obtained in only $\mathcal{O}(d^2 \log_2 d)$ steps. In the limit of large d, the function values of the Husimi function add up to one:

$$\sum_{p=0}^{d-1} \sum_{q=0}^{d-1} H(p, q) = 1. \tag{B.22}$$

As an example, the Husimi function of the coherent $n = 4$ qubit state $|\Phi(d/4, d/4)\rangle$ is depicted in figure B.4.

C Technical Results

This chapter of the appendix contains various technical results which are referred to in part II of this thesis. The first section proves some counting lemmas for linear codes, the second section proves the existence of good self-orthogonal codes, and the third section proves some lemmas concerning a Bell state.

C.1 Linear Codes

This section provides two corollaries which are needed for the proof the random coding arguments in subsections 5.4.3 and 7.3.2.

Let us denote the set containing all $[n,k]_q$ codes by

$$A_{n,k,q} = \{\mathcal{C} \subseteq \mathbb{F}_q^n \,|\, \mathcal{C} \text{ is an } [n,k]_q\text{-code}\}, \tag{C.1}$$

and let us denote the subset of codes in $A_{n,k,q}$ which contain a certain nonzero codeword $\vec{x} \in \mathbb{F}_q^n$ by

$$A_{n,k,q}(\vec{x}) = \{\mathcal{C} \in A_{n,k,q} \,|\, \vec{x} \in \mathcal{C}\}. \tag{C.2}$$

Lemma C.1.1. *The total number of $[n,k]_q$ codes is given by*

$$|A_{n,k,q}| = \frac{\prod_{i=0}^{k-1}(q^n - q^i)}{\prod_{i=0}^{k-1}(q^k - q^i)} \qquad (1 \leq k \leq n) \tag{C.3}$$

and $|A_{n,0,q}| = 1$. The number of $[n,k]_q$ codes which contain a certain nonzero vector $\vec{x}$ is given by

$$|A_{n,k,q}(\vec{x})| = \frac{\prod_{i=1}^{k-1}(q^n - q^i)}{\prod_{i=1}^{k-1}(q^k - q^i)} \qquad (1 < k \leq n), \tag{C.4}$$

and $|A_{n,1,q}(\vec{x})| = 1$, $|A_{n,0,q}(\vec{x})| = 0$ independently of $\vec{x} \neq \vec{0}$.

Proof. To determine the total number of $[n,k]_q$ codes, we have to count all possibilities to choose k linearly independent vectors from $\mathbb{F}_q^n$. There are $q^n - 1$ candidates for the first vector, there remain $q^n - q$ for the second, $q^n - q^2$ for the third, and so on. We therefore get in total $N = (q^n - q^0)(q^n - q^1)\ldots(q^n - q^{k-1})$ possibilities. Since many of these selections of k vectors span the same codespace, we have to divide N by the number of ways a set of k generating vectors can be found for a k-dimensional subspace. This number is $(q^k - q^0)(q^k - q^1)\ldots(q^k - q^{k-1})$. The number of linear codes containing a particular nonzero $\vec{x}$ can be found in a similar fashion, but now as first independent vector we choose $\vec{x}$ itself. □

Corollary C.1.2. *One obtains from the above lemma that for any nonzero $\vec{x} \in \mathbb{F}_q^n$*

$$\frac{|A_{n,k,q}(\vec{x})|}{|A_{n,k,q}|} = \frac{q^k - 1}{q^n - 1} \leq \frac{1}{q^{n-k}}. \tag{C.5}$$

The following lemmas are slight generalizations of lemma C.1.1.

Lemma C.1.3. *Let $\mathcal{K}$ be an $[n,\kappa]_q$ code and let*

$$A_{n,k,q}(\mathcal{K}) = \{\mathcal{C} \in A_{n,k,q} \mid \mathcal{K} \subseteq \mathcal{C}\} \tag{C.6}$$

be the set of all $[n,k]_q$ codes which contain $\mathcal{K}$. Then,

$$|A_{n,k,q}(\mathcal{K})| = \frac{\prod_{i=\kappa}^{k-1}(q^n - q^i)}{\prod_{i=\kappa}^{k-1}(q^k - q^i)} \qquad (\kappa < k \leq n), \tag{C.7}$$

and $|A_{n,\kappa,q}(\mathcal{K})| = 1$, $|A_{n,k,q}(\mathcal{K})| = 0$ for $(k < \kappa)$.

Lemma C.1.4. *Let $\mathcal{K}$ be an $[n,\kappa]_q$ code and let*

$$A_{n,k,q}(\mathcal{K},\vec{x}) = \{\mathcal{C} \in A_{n,k,q} \mid \mathcal{K} \subseteq \mathcal{C} \text{ and } \vec{x} \in \mathcal{C}\} \tag{C.8}$$

be the set of all $[n,k]_q$ codes which contain $\mathcal{K}$ and a certain nonzero vector $\vec{x} \in \mathbb{F}_q^n$. Then,

$$|A_{n,k,q}(\mathcal{K},\vec{x})| = \begin{cases} |A_{n,k,q}(\mathcal{K})| & \text{if } \vec{x} \in \mathcal{K} \\ \frac{\prod_{i=\kappa+1}^{k-1}(q^n - q^i)}{\prod_{i=\kappa+1}^{k-1}(q^k - q^i)} & \text{if } \vec{x} \notin \mathcal{K} \text{ and } \kappa+1 < k \leq n \\ 1 & \text{if } \vec{x} \notin \mathcal{K} \text{ and } \kappa+1 = k \\ 0 & \text{else} \end{cases}. \tag{C.9}$$

Corollary C.1.5. *Let $\mathcal{K}$ be an $[n, n-k_1]_q$ code and let $\langle \cdot \rangle_{\mathcal{K} \in A_{n,n-k_1,q}}$ denote the average over all such codes. Then,*

$$\Big\langle \frac{|A_{n,n-k_2,q}(\mathcal{K},\vec{x})|}{|A_{n,n-k_2,q}(\mathcal{K})|} \Big\rangle_{\mathcal{K} \in A_{n,n-k_1,q}} = \frac{q^{n-k_2}-1}{q^n - q} \leq \frac{1}{q^{k_2}}. \tag{C.10}$$

Proof.

$$\Big\langle \frac{|A_{n,n-k_2,q}(\mathcal{K},\vec{x})|}{|A_{n,n-k_2,q}(\mathcal{K})|} \Big\rangle_{\mathcal{K} \in A_{n,n-k_1,q}} = \frac{1}{|A_{n,n-k_1,q}|} \sum_{\mathcal{K} \in A_{n,n-k_1,q}} \frac{|A_{n,n-k_2,q}(\mathcal{K},\vec{x})|}{|A_{n,n-k_2,q}(\mathcal{K})|}$$

We use lemma C.1.3 and C.1.4 and obtain

$$\begin{aligned} &= \frac{1}{|A_{n,n-k_1,q}|} \sum_{\mathcal{K} \in A_{n,n-k_1,q}} \begin{cases} 1 & \text{, if } \vec{x} \in \mathcal{K} \\ \frac{q^{n-k_2}-q^{n-k_1}}{q^n - q^{n-k_1}} & \text{, else} \end{cases} \\ &= 1 \cdot \frac{|A_{n,n-k_1,q}(\vec{x})|}{|A_{n,n-k_1,q}|} + \frac{q^{n-k_2}-q^{n-k_1}}{q^n - q^{n-k_1}} \cdot \Big(1 - \frac{|A_{n,n-k_1,q}(\vec{x})|}{|A_{n,n-k_1,q}|}\Big). \end{aligned}$$

Corollary C.1.2 tells us that $|A_{n,n-k_1,q}(\vec{x})|/|A_{n,n-k_1,q}| = (q^{n-k_1}-1)/(q^n-1)$ which leads to the desired result. □

C.2 Self-Orthogonal Codes

In this section it is shown that good self-orthogonal codes do exist. A self-orthogonal q-ary linear $[n,\mathsf{k}]_q$ code $\mathcal{C}$ over the field $\mathbb{F}_q^n$ is a code which is contained in its dual $[n, n-\mathsf{k}]_q$ code $\mathcal{C}^\perp$. A code $\mathcal{C}$ is called self-dual provided that $\mathcal{C} = \mathcal{C}^\perp$ (in this case n has to be even and $\mathsf{k} = n/2$). Self-orthogonal $[n,\mathsf{k}]_q$ codes can be used to construct quantum CSS-codes encoding $k = n - 2\mathsf{k}$ qudits into n. If $\mathcal{C}^\perp$ has minimum distance d, $\mathcal{C}$ has to be at least of the same minimum distance. Hence the quantum CSS-code will be of distance d.

In the following subsections, a Gilbert-Varshamov lower bound is established, which guarantees the existence of self-orthogonal $[n, \mathsf{k}, d]_q$ codes such that the dual $[n, n-\mathsf{k}, d]_q$ code has minimum distance d and rate

$$\frac{n-\mathsf{k}}{n} \geq 1 - H_{q[\log_q]}\Big(1 - \frac{d}{n}, \frac{d/n}{q-1}, \ldots, \frac{d/n}{q-1}\Big). \tag{C.11}$$

For the binary case ($q = 2$) this result was found by Calderbank and Shor [CS96]. The corresponding proof is given in the first section. The nonbinary case ($q \geq 3$) has to be treated separately. It is proven in the second section using results presented in [Ham04].

C.2.1 The Binary Case

Lemma C.2.1 ([CS96])**.** *For even $n \geq 2$ and $0 < \mathsf{k} \leq n/2$, let*

$$A(n, \mathsf{k}) = \big\{\mathcal{C} \subseteq F_2^n \mid \mathcal{C} \text{ is a } [n, \mathsf{k}]_2\text{-code}, \{\vec{0}, \vec{1}\} \subseteq \mathcal{C} \subseteq \mathcal{C}^\perp\big\} \tag{C.12}$$

be the set of all self-orthogonal $[n, \mathsf{k}]_2$ codes which include the $[n, 1]_2$ subcode $\{\vec{0}, \vec{1}\}$, and let

$$A_{\vec{x}} = \big\{\mathcal{C} \in A(n, \mathsf{k}) \mid \vec{x} \in \mathcal{C}^\perp\big\} \tag{C.13}$$

be the subset of $A(n, \mathsf{k})$ including only those codes whose dual codes include $\vec{x} \in \mathbb{F}_2^n$. Then, there exists a constant T_0 satisfying $|A_{\vec{x}}| = T_0$ for any $\vec{x} \in \mathbb{F}_2^n$ with $\vec{x} \neq \vec{0}$, $\vec{x} \neq \vec{1}$ and $\vec{x} \cdot \vec{x} = 0 \pmod 2$.

Remark. For a proof we refer to [CS96]. Note that for all $\vec{x} \in \mathcal{C}^\perp$, $\mathrm{wt}(\vec{x}) = 0 \pmod 2$ which follows from $\vec{1} \cdot \vec{x} = 0 \pmod 2$. (By $\vec{1}$ we denote the vector $(1, 1, \ldots, 1) \in \mathbb{F}_2^n$ and analogously $\vec{0} = (0, 0, \ldots, 0) \in \mathbb{F}_2^n$.)

Theorem C.2.2. *Consider the set of codes $\Phi = \{\mathcal{C}^\perp \mid \mathcal{C} \in A(n, \mathsf{k})\}$. Then, as long as*

$$\sum_{s=1}^{2s \leq d-1} \binom{n}{2s} < \frac{2^{n-1} - 2}{2^{n-\mathsf{k}} - 2}, \tag{C.14}$$

there exist codes of minimum distance d in Φ.

Proof. Counting all vectors $\vec{x}$ (except $\vec{x} = \vec{0}$ and $\vec{x} = \vec{1}$) in Φ in two different ways, we get (by noting that $|\Phi| = |A(n, \mathsf{k})|$)

$$|A(n, \mathsf{k})| \cdot (2^{n-\mathsf{k}} - 2) = (2^{n-1} - 2) \cdot T_0. \tag{C.15}$$

There are $\sum_{s=1}^{2s \leq d-1} \binom{n}{2s}$ nonzero vectors of even weight less than d. These vectors are distributed over $\sum_{s=1}^{2s \leq d-1} \binom{n}{2s} \cdot T_0$ codes at most. As long as this number of codes is smaller than $|A(n, \mathsf{k})|$ (the total number of codes in Φ), there have to be codes in Φ which are at least of minimum distance d. □

Corollary C.2.3. *Consider the set codes $\Phi = \{\mathcal{C}^\perp \mid \mathcal{C} \in A(n, \mathsf{k})\}$. Then, as long as*

$$\frac{n-\mathsf{k}}{n} < 1 - H_2(d/n), \tag{C.16}$$

there exist codes of minimum distance d in Φ.

Proof. The tail inequality gives an upper bound for the left hand side of (C.14):

$$\sum_{s=1}^{2s \leq d-1} \binom{n}{2s} < \sum_{j=0}^{d-1} \binom{n}{j} \leq 2^{nH_2((d-1)/n)} < 2^{nH_2(d/n)}.$$

A lower bound for the right hand side of (C.14) is given by $2^{n-1}/2^{n-\mathsf{k}}$. Hence, as long as $nH_2(d/n) + n - \mathsf{k} < n - 1$ condition (C.14) will be satisfied, too. For large n this leads to condition (C.16). □

C.2.2 The Higher Dimensional Case

Lemma C.2.4 ([Ham04]). *For $q \geq 3$ let*

$$A(n,k) = \left\{\mathcal{C} \subseteq F_q^n \mid \mathcal{C} \text{ is a } [n,k]_q\text{-code}, \mathcal{C} \subseteq \mathcal{C}^\perp\right\} \tag{C.17}$$

be the set of all self-orthogonal $[n,k]_q$ codes, and let

$$A_{\vec{x}} = \left\{\mathcal{C} \in A(n,k) \mid \vec{x} \in \mathcal{C}^\perp\right\} \tag{C.18}$$

be the subset of $A(n,k)$ including only those codes whose dual code includes $\vec{x} \in \mathbb{F}_q^n$. Then, for any $u \in \mathbb{F}_q$, there exists a constant T_u satisfying $|A_{\vec{x}}| = T_u$ for any nonzero $\vec{x} \in \mathbb{F}_q^n$ with $\vec{x} \cdot \vec{x} = u \pmod q$.

Remark. For a proof of the above lemma we refer to [Ham04, Lemma 1]. The following theorem is proven using results from [Ham04, Corollary 1].

Theorem C.2.5. *Consider the set of codes $\Phi = \{\mathcal{C}^\perp \mid \mathcal{C} \in A(n,k)\}$. Then, as long as*

$$\sum_{j=1}^{d-1} \binom{n}{j} (q-1)^j < \frac{q^{n-q+1}-1}{q^{n-k}-1}, \tag{C.19}$$

there exist codes of minimum distance d in Φ.

Proof. Let $S_u = \{\vec{x} \in \mathbb{F}_q^n | \vec{x} \cdot \vec{x} = u \pmod q, \vec{x} \neq \vec{0}\}$ for $u \in \mathbb{F}_q$. It follows that $|S_u| \geq q^{n-q+1} - 1$ since the first $n-q+1$ digits of any $\vec{x} \in S_u$ can be set in arbitrary manner (except to $(0,\ldots,0)$). Counting pairs $(\vec{x}, \mathcal{C})$ such that $\vec{x} \cdot \vec{x} = u \pmod q$, $\vec{x} \neq \vec{0}$, and $\vec{x} \in \mathcal{C}^\perp \in \Phi$, we find that (noting that $|\Phi| = |A(n,k)|$)

$$|S_u| \cdot T_u \leq |A(n,k)| \cdot (q^{n-k} - 1) \tag{C.20}$$

and we get (using the upper bound on $|S_u|$)

$$\frac{q^{n-q+1}-1}{q^{n-k}-1} \leq \frac{|A(n,k)|}{T_u}. \tag{C.21}$$

There are $\sum_{j=1}^{d-1} \binom{n}{j}(q-1)^j$ nonzero vectors of weight less than d. These vectors are distributed over $\sum_{j=1}^{d-1} \binom{n}{j}(q-1)^j \cdot \max_u\{T_u\}$ codes at most. As long as this number is smaller than $|A(n,k)|$, there have to be codes in Φ which are at least of minimum distance d. Because of (C.21), equation (C.19) is a sufficient condition. □

Corollary C.2.6. *Consider the set of codes $\Phi = \{\mathcal{C}^\perp \mid \mathcal{C} \in A(n,k)\}$. Then, for large enough n, as long as*

$$\frac{n-k}{n} < 1 - H_{q[\log_q]}\Big(1 - \frac{d}{n}, \frac{d/n}{q-1}, \ldots, \frac{d/n}{q-1}\Big) \tag{C.22}$$

there exist codes of minimum distance d in Φ.

Proof. By using the Chernoff bound 1.2.1 it was shown in the proof of corollary 5.2.3 that an upper bound for the left hand side of (C.19) is given by

$$\sum_{j=0}^{d-1} \binom{n}{j} (q-1)^j < \exp_q\left(nH_{q[\log_q]}\Big(1 - \frac{d}{n}, \frac{d}{n(q-1)}, \ldots, \frac{d}{n(q-1)}\Big)\right). \tag{C.23}$$

A lower bound for the right hand side of (C.19) is given by

$$\frac{q^{n-q+1}}{q^{n-k}} < \frac{q^{n-q+1}-1}{q^{n-k}-1}. \tag{C.24}$$

Therefore, as long as

$$\frac{n-k}{n} < 1 - H_{q[\log_q]}\Big(1 - \frac{d}{n}, \frac{d/n}{q-1}, \ldots, \frac{d/n}{q-1}\Big) - \frac{q-1}{n}, \tag{C.25}$$

condition (C.19) will be satisfied, too. For large n we can neglect the $q-1$ term. □

C.3 Bell State Lemmas

We are going to prove two simple lemmas concerning the Bell state $|\Phi\rangle_{AB} = q^{-\frac{1}{2}} \sum_{j=0}^{q-1} |j\tilde{j}\rangle_{AB}$ that are relevant in chapter 8. Here, $|i\tilde{j}\rangle_{AB} = |i\rangle_A \otimes |\tilde{j}\rangle_B$, where $\{|i\rangle_A\}_{i=0,\ldots,q-1}$ and $\{|\tilde{j}\rangle_B\}_{j=0,\ldots,q-1}$ denote orthonormal bases of the q-dimensional Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively.

Lemma C.3.1. *Let* $|\Phi\rangle_{AB} = \frac{1}{\sqrt{q}} \sum_{j=0}^{q-1} |j\tilde{j}\rangle_{AB}$. *Then,*

$$O_A^T \otimes \mathcal{I}_B |\Phi\rangle_{AB} = \mathcal{I}_A \otimes O_B |\Phi\rangle_{AB}, \tag{C.26}$$

if the transposition is with respect to the $\{|j\rangle_A\}$ *basis and* O_B *has the same matrix elements with respect to the* $\{|\tilde{j}\rangle_B\}$ *basis as* O_A *with respect to the* $\{|j\rangle_A\}$ *basis, i. e.* $O_A = \sum_{ij} O_{ij} |i\rangle_{AA}\langle j|$ *and* $O_B = \sum_{ij} O_{ij} |\tilde{i}\rangle_{BB}\langle\tilde{j}|$.

Proof. We obtain

$$\begin{aligned} O_A^T \otimes \mathcal{I}_B |\Phi\rangle_{AB} &= \frac{1}{\sqrt{d}} \sum_{ij} O_{ij} |j\rangle_{AA}\langle i| \sum_k |k\rangle_A |\tilde{k}\rangle_B \\ &= \frac{1}{\sqrt{d}} \sum_{ij} O_{ij} |j\rangle_A |\tilde{i}\rangle_B \\ &= \frac{1}{\sqrt{d}} \sum_{ij} O_{ij} |\tilde{i}\rangle_{BB}\langle\tilde{j}| \sum_k |k\rangle_A |\tilde{k}\rangle_B \\ &= \mathcal{I}_A \otimes O_B |\Phi\rangle_{AB}. \end{aligned}$$

□

Lemma C.3.2. *Let* $|\Phi\rangle_{AB} = \frac{1}{\sqrt{q}} \sum_{j=0}^{q-1} |j\tilde{j}\rangle_{AB}$. *Then, for any unitary* U,

$$U_A^* \otimes U_B |\Phi\rangle_{AB} = |\Phi\rangle_{AB} \tag{C.27}$$

if the conjugation is with respect to the $\{|j\rangle_A\}$ *basis and* U_B *has the same matrix elements with respect to the* $\{|\tilde{j}\rangle_B\}$ *basis as* U_A *with respect to the* $\{|j\rangle_A\}$ *basis, i. e.* $U_A = \sum_{ij} U_{ij} |i\rangle_{AA}\langle j|$ *and* $U_B = \sum_{ij} U_{ij} |\tilde{i}\rangle_{BB}\langle\tilde{j}|$.

Proof. We obtain

$$\begin{aligned} U_A^* \otimes U_B |\Phi\rangle_{AB} &= \frac{1}{\sqrt{d}} \sum_{ijmnk} U_{ij}^* |i\rangle_{AA}\langle j|U_{mn}|\tilde{m}\rangle_{BB}\langle\tilde{n}||k\rangle_A |\tilde{k}\rangle_B \\ &= \frac{1}{\sqrt{d}} \sum_{ijmn} U_{ij}^* |i\rangle_A U_{mn} |\tilde{m}\rangle_{BB}\langle\tilde{n}|\tilde{j}\rangle_B \\ &= \frac{1}{\sqrt{d}} \sum_{ijm} U_{ij}^* |i\rangle_A U_{mj} |\tilde{m}\rangle_B \\ &= \frac{1}{\sqrt{d}} \sum_{im} \delta_{im} |i\rangle_A |\tilde{m}\rangle_B = |\Phi\rangle_{AB}. \end{aligned}$$

□

D Schur Transform and Eigenfunction Method

The Schur transform is a unitary transformation relating the standard computational basis of n qudits of dimension q to a basis associated with the representation theory of the symmetric and general linear groups. This chapter explains how the eigenfunction method [CPW02] can be used to obtain a computer program which calculates the Schur transform for given values of n and q. As explained in section D.1, the eigenfunction method decomposes a given group representation into its irreducible parts. It is shown in section D.2 how the Schur transform can be obtained with the help of the eigenfunction method applied to the natural representation of the symmetric group S_n. In addition we present some examples and discuss how the Schur transform allows for efficient communication in the absence of a shared reference frame.

D.1 The Eigenfunction Method

This section summarizes the eigenfunction method (EFM) of Chen, Ping and Wang [CPW02]. Let $R(G)$ be a d-dimensional representation of a finite group G on an inner product space $\mathcal{V}$ over the field $\mathbb{C}$. The EFM can be used to decompose $\mathcal{V}$ into a direct sum of irreducible subspaces and to construct a basis for each of these subspaces which corresponds to a given canonical subgroup chain. To achieve this decomposition of $\mathcal{V}$, a complete set of commuting observables (CSCO) $\mathfrak{C}$ is constructed, whose eigenvectors (eigenfunctions) are the desired basis vectors. They can be identified by their eigenvalue list. All the results presented in this section are taken from [CPW02]. While we tried to supply the proofs for the fundamental results, we sometimes give the remark 'it can be shown'. These missing proofs can be found in [CPW02].

We start with a description of the EFM for general finite groups in subsection D.1.1 and specialize in the symmetric group in subsection D.1.2.

D.1.1 General Finite Groups

We begin with the construction of the CSCO $\mathfrak{C}$ decomposing the representation space of the regular representation of a finite group G. Let $\mathcal{V}$ be an inner product space of dimension $d = n_G$ over the field $\mathbb{C}$, where $n_G = |G|$ denotes the order of the finite group G, and fix an orthonormal basis $\{|i\rangle\}$ $(i = 0, \ldots, d-1)$. The elements of the regular representation $R(G)$ of G have the property that $|i\rangle = R_i|0\rangle$ for all $R_i \equiv R(i)$ with $i \in G$, with R_0 denoting the identity. In other words,

$$\langle i|R_k|j\rangle \equiv D_{ij}(k) = \begin{cases} 1 & \text{if } R_k R_j = R_i \\ 0 & \text{else} \end{cases} . \tag{D.1}$$

The state $|0\rangle$ is said to possess no symmetry with respect to G. A state $|0\rangle'$ which remains invariant under G is called totally symmetric with respect to G (it would generate a one-dimensional representation). States showing an intermediate behavior are said to possess partial symmetry.

Subsequently, we show how the construction of the CSCO $\mathfrak{C}$ has to be adjusted when dealing with non-regular representations $R(G)$. In this case the state $|0\rangle$ is invariant under a set of elements G_{in} forming a non-trivial subgroup of G, i. e. $R_a|0\rangle = |0\rangle$ for all $a \in G_{\text{in}}$, and is said to possess at least partial symmetry with respect to G. Naturally, the dimension d of a non-regular rep space spanned by the linearly independent $|i\rangle = R_i|0\rangle$, $R_i \in R(G)$, is smaller than n_G.

Reduction of the Regular Representation

Let us define a class operator C_i for each of the n_ζ conjugacy classes of G as the sum over all operators in the corresponding class,

$$C_i = \sum_{j=1}^{n_i} R(a_j^{(i)}), \quad i = 1 \dots n_\zeta, \tag{D.2}$$

where $a_j^{(i)}$ denotes the j-th element of the i-th class and n_i denotes the total number of elements in the i-th class. The class operators commute with all elements in $R(G)$, $[C_i, R_a] = 0$ for all $a \in G$, and therefore with one another, $[C_i, C_j] = 0$ for $i, j = 1 \dots n_\zeta$. We assume that the $\{C_i\}_{i=1}^{n_\zeta}$ are self-adjoint (they are if the classes are ambivalent), otherwise an equivalent set of n_ζ self-adjoint operators $\{C_i'\}_{i=1}^{n_\zeta}$ can be obtained by taking suitable linear combinations of the non-ambivalent C_i. The class space is defined as the n_ζ-dimensional subspace of the regular rep space $\mathcal{V}$ spanned by the orthogonal set of states

$$\left\{ |C_i\rangle = \sum_{j=1}^{n_i} R(a_j^{(i)})|0\rangle \right\}_{i=1}^{n_\zeta} \tag{D.3}$$

with $\langle C_j | C_i \rangle = n_i \delta_{ij}$. It can be shown that the class space forms a so-called natural representation space of the class operators, and that the set of n_ζ class operators $(C_1, \dots, C_{n_\zeta})$ is a CSCO of the natural rep, reducing the natural rep to a sum of n_ζ one-dimensional irreps via the eigenvector equation

$$(C_1, \dots, C_{n_\zeta})|Q^{(\nu)}\rangle = (\lambda_1^{(\nu)}, \dots, \lambda_{n_\zeta}^{(\nu)})|Q^{(\nu)}\rangle \equiv \lambda^{(\nu)}|Q^{(\nu)}\rangle, \tag{D.4}$$

with $|Q^{(\nu)}\rangle = \sum_{j=1}^{n_\zeta} q_j^{(\nu)} |C_j\rangle$ and $q_j^{(\nu)} \in \mathbb{C}$. In general $(C_1, \dots, C_{n_\zeta})$ is over-complete. If a subset $C = (C_{i_1}, \dots, C_{i_l})$ of the class operators $(C_1, \dots, C_{n_\zeta})$ is a CSCO of the class space, then C is called CSCO of the first kind (CSCO-I) of G (different CSCO's are equivalent in the sense that they lead to the same eigenvectors $|Q^{(\nu)}\rangle$). It can be shown that in any representation space $\mathcal{V}$ the eigenvalues $\lambda^{(\nu)}$ of C do not go beyond the n_ζ values determined in class space, and that in the regular representation space there are n_ζ and only n_ζ distinct eigenvalues $\lambda^{(\nu)}$. By theorem 1.2.3, the eigenspaces of C in a rep space $\mathcal{V}$ are representation spaces and the regular rep space $\mathcal{V}$ is reduced to a direct sum of n_ζ mutually orthogonal subspaces,

$$\mathcal{V} = \bigoplus_{\nu=1}^{n_\zeta} \mathcal{V}_\nu, \tag{D.5}$$

where $C\mathcal{V}_\nu = \lambda^{(\nu)}\mathcal{V}_\nu$ (in a non-regular representation space $\mathcal{V}$ one or more of the $\mathcal{V}_\nu$ might be trivial subspaces containing only the zero vector). Using the fact that the representative of C on $\mathcal{V}_\nu$ must be equal to the identity times the eigenvalue $\lambda^{(\nu)}$, it can be seen that representation spaces belonging to different eigenvalues are inequivalent. A representation space $\mathcal{V}_\nu$ may still be reducible,

$$\mathcal{V}_\nu = \mathcal{V}_{\nu,1} \oplus \dots \oplus \mathcal{V}_{\nu,\tau_\nu}, \tag{D.6}$$

where irreps $\mathcal{V}_{\nu,k}$ with the same label ν are equivalent. The results presented so far lead to the well known result that a finite group with n_ζ classes has n_ζ and only n_ζ inequivalent irreps. The irreps can be labeled uniquely by the eigenvalue list $\lambda^{(\nu)}$ of a CSCO-I C (we use the symbol ν as label). If a vector $|\psi^{(\nu)}\rangle$ belongs to the eigenspace $\mathcal{V}_\nu$ of a CSCO-I C of G, the vector is said to belong to the irrep ν of G.

Theorem D.1.1. *A necessary and sufficient condition for a vector $|\psi^{(\nu)}\rangle$ to belong to the irrep ν of G is that*

$$C|\psi^{(\nu)}\rangle = \lambda^{(\nu)}|\psi^{(\nu)}\rangle. \tag{D.7}$$

Proof. The sufficiency is trivial. We prove that the condition is a necessary one. Suppose $|\psi^{(\nu)}\rangle$ is a vector in an irreducible subspace $\mathcal{V}_\nu$ of G. It follows that $\mathcal{V}_\nu$ is an invariant subspace of C and by Schur's lemma we obtain that $\mathcal{V}_\nu$ is necessarily an eigenspace of C. □

This theorem is the corner stone of the EFM. It allows the problem of finding the irreps of G to be converted into the problem of finding the eigenspaces of a CSCO-I C of G (i. e. we have to diagonalize the operator C^* in the reducible basis $|0\rangle, \dots, |d-1\rangle$ spanning $\mathcal{V}$).

Let us now consider a canonical subgroup chain $G \supset G(s_1) \supset G(s_2) \dots$, or by using the the abbreviation $G(s) = G(s_1) \supset G(s_2) \dots$, $G \supset G(s)$. Analogous to theorem D.1.1 we obtain the following theorem.

Theorem D.1.2. *A necessary and sufficient condition for a vector $|\psi^{(\nu)}_{\lambda(s_1),\lambda(s_2),\dots}\rangle$ in a rep space $\mathcal{V}$ to belong to the irreps $\nu, \lambda(s_1), \lambda(s_2), \dots$ of a subgroup chain $G \supset G(s)$ is that the vector satisfies the following eigenequations,*

$$\begin{pmatrix} C \\ C(s_1) \\ C(s_2) \\ \vdots \end{pmatrix} |\psi^{(\nu)}_{\lambda(s_1),\lambda(s_2),\dots}\rangle = \begin{pmatrix} \nu \\ \lambda(s_1) \\ \lambda(s_1) \\ \vdots \end{pmatrix} |\psi^{(\nu)}_{\lambda(s_1),\lambda(s_2),\dots}\rangle, \tag{D.8}$$

where C is a CSCO-I of G and $C(s_i)$ is a CSCO-I of $G(s_i)$.

Remark. Using the abbreviations $C(s) = (C(s_1), C(s_2), \dots)$ and $m = (\lambda(s_1), \lambda(s_2), \dots)$, the eigenequation of the above theorem becomes

$$\begin{pmatrix} C \\ C(s) \end{pmatrix} |\psi^{(\nu)}_m\rangle = \begin{pmatrix} \nu \\ m \end{pmatrix} |\psi^{(\nu)}_m\rangle. \tag{D.9}$$

If the subgroup chain is canonical, the set $(C, C(s))$ is called CSCO-II of G.

Suppose the eigenspace $\mathcal{V}_\nu$ is an irreducible rep space of G. Than the degeneracy of the eigenvalue $\lambda^{(\nu)}$ in (D.7) is equal to the dimension h_ν of the irrep and is totally lifted by the eigenequations of the $C(s)$ (i. e. the degeneracy of the eigenvalues $\{(\nu, m_i)\}_{i=1}^{h_\nu}$ is one). If $\mathcal{V}_\nu$ is a reducible rep space of G, the degeneracy of $\lambda^{(\nu)}$ is given by $\tau_\nu \times h_\nu$ and for each value (ν, m_i) there are τ_ν linearly independent eigenvectors $|\psi^{(\nu)\tau}_m\rangle$, $\tau = 1 \dots \tau_\nu$, $\tau_\nu \in \{2, 3, 4, \dots\}$.

We now introduce the intrinsic group $\bar{G}$ of G which is used to complete the set CSCO-II to a complete set of commuting observables (CSCO-III) $\mathfrak{C}$ on the representation space $\mathcal{V}$.

Definition D.1.1. For each operator g in G, we define a super-operator $\bar{g}$ acting on the elements of the group algebra $\mathcal{A} = \mathbb{C}G$ (any element a in $\mathcal{A}$ can be written as $a = \sum_{g \in G} a_g g$ with $a_g \in \mathbb{C}$) by

$$\bar{g}a = ag \quad \text{for all } a \in \mathcal{A}. \tag{D.10}$$

The group formed by all $\bar{g}$ is called the intrinsic group $\bar{G}$ of G.

We proceed by proving two important lemmas concerning the intrinsic group.

Lemma D.1.3. *The operators in $\bar{G}$ commute with those in G.*

Proof. We have $s\bar{r}t = str = \bar{r}st$ for all $t \in \mathcal{A}$ and therefore $[\bar{r}, s] = 0$ for all $s \in G$ and $\bar{r} \in \bar{G}$. □

Lemma D.1.4. *The group $\bar{G}$ is anti-isomorphic to G.*

Proof. Suppose the multiplication relation in G is $rs = u$ for $r, s, u \in G$. Then $\bar{s}\bar{r}t = \bar{s}tr = trs = tu = \bar{u}t$ for all $t \in \mathcal{A}$ and we have $\bar{s}\bar{r} = \bar{u}$. □

*C is a set of commuting operators, but by taking a suitable linear combination of these operators, it suffices to diagonalize only one single operator.

If we consider the action of the elements of the intrinsic group on the representation space $\mathcal{V}$ of a representation $R(G)$ with basis $\{|i\rangle = R_i|0\rangle\}$, we have to define a state, say $|0\rangle$, as the intrinsic state, i. e. the elements of $\bar{G}$ act on the basis states as

$$\bar{R}_b|a\rangle = \bar{R}_b R_a|0\rangle = R_a R_b|0\rangle. \tag{D.11}$$

Note that if the intrinsic state is invariant under a symmetry group $G_{\text{in}} \subset G$, we have $\bar{R}_b|0\rangle = \bar{R}_b T|0\rangle = TR_b|0\rangle = T|b\rangle$ for all $T \in G_{\text{in}}$ and on the other hand $\bar{R}_b|0\rangle = R_b|0\rangle = |b\rangle$ which is a contradiction. Therefore, the following only holds for the regular representation $R(G)$ for which G_{in} contains only the identity. The anti-isomorphism between G and $\bar{G}$ assures that the conclusions about G apply to $\bar{G}$ as well:

(i) If $C = (C_{i_1}, \dots, C_{i_l})$ is a CSCO-I of G, then $\bar{C} = (\bar{C}_{i_1}, \dots, \bar{C}_{i_l})$ is a CSCO-I of $\bar{G}$ with

$$\bar{C}_i = \sum_{j=1}^{n_i} \bar{R}(a_j^{(i)}). \tag{D.12}$$

Note that the CSCO-I of G and $\bar{G}$ are equal, since

$$\bar{C}_i R_k = \big(\sum_{j=1}^{n_i} \bar{R}(a_j^{(i)})\big) R_k = R_k \big(\sum_{j=1}^{n_i} R(a_j^{(i)})\big) = R_k C_i = C_i R_k, \tag{D.13}$$

where the last identity holds because $[C_i, R] = 0$ for all $R_k \in R(G)$.

(ii) If G has a canonical subgroup chain $G \supset G(s)$, $G(s) = G(s_1) \supset G(s_2) \supset \dots$, with CSCO-II $\big(C, C(s) = (C(s_1), C(s_2), \dots)\big)$, $\bar{G}$ has a canonical subgroup chain $\bar{G} \supset \bar{G}(s)$, $\bar{G}(s) = \bar{G}(s_1) \supset \bar{G}(s_2) \supset \dots$, with CSCO-II $\big(\bar{C}, \bar{C}(s) = (\bar{C}(s_1), \bar{C}(s_2), \dots)\big)$.

Because of lemma D.1.3 $[C(s), \bar{C}(s)] = 0$, which allows the $\bar{C}(s)$ to be added to a CSCO-II of G and the following theorem to be proved.

Theorem D.1.5. *The set $\mathfrak{C} = (C, C(s), \bar{C}(s))$ defined on the regular rep space $\mathcal{V}$ of a group G with canonical subgroup chain $G(s) = G(s_1) \supset G(s_2) \supset \dots$ is a CSCO on $\mathcal{V}$ (called CSCO-III). The corresponding eigenequation is given by*

$$\begin{pmatrix} C \\ C(s) \\ \bar{C}(s) \end{pmatrix} |\psi_m^{(\nu)k}\rangle = \begin{pmatrix} \nu \\ m \\ k \end{pmatrix} |\psi_m^{(\nu)k}\rangle, \tag{D.14}$$

with $k = (\bar{\lambda}(s_1), \bar{\lambda}(s_2), \dots)$.

Because $(C, \bar{C}(s))$ commutes with all the elements in $R(G)$, the eigenspaces $\mathcal{V}_{\nu,k} = \text{span}\big\{|\psi_{m_i}^{(\nu)k}\rangle\big\}$, $i = 1 \dots h_\nu$ of $(C, \bar{C}(s))$ are necessarily representation spaces of $R(G)$ and the degeneracy of m_i is necessarily independent of i. Since in addition $\mathfrak{C} = (C, C(s), \bar{C}(s))$ is a CSCO of $\mathcal{V}$, $(C(s), \bar{C}(s))$ is necessarily a CSCO in each eigenspace $\mathcal{V}_\nu$, $\nu = 1 \dots n_\zeta$, and the degeneracy of m_i, $i = 1 \dots h_\nu$, in $\mathcal{V}_\nu$ is completely lifted by the eigenvalue k of $\bar{C}(s)$. It can be shown that the representatives of the operators $C(s_i)$ and $\bar{C}(s_i)$ in $\mathcal{V}_\nu$ are similar matrices. Therefore, the characteristic equations of $C(s)$ and $\bar{C}(s)$ in $\mathcal{V}_\nu$ are identical and it follows that the eigenvalue k takes on the values $k_i = m_i$ for $i = 1 \dots h_\nu$. Equation (D.6) in the regular rep case becomes

$$\mathcal{V}_\nu = \bigoplus_{i=1}^{\tau_\nu = h_\nu} \mathcal{V}_{\nu,k_i}, \tag{D.15}$$

and we have $d = n_G = \sum_{\nu=1}^{n_\zeta} h_\nu^2$.

Since the normalized vectors $|\psi_m^{(\nu)k}\rangle$ are obtained by solving an eigenequation, they are determined only up to a phase factor. Let the eigenvectors be expressed as

$$|\psi_m^{(\nu)k}\rangle = \sum_{i=0}^{n_G-1} u_{\nu mk,i}|i\rangle, \quad u_{\nu mk,i} \in \mathbb{C}, \tag{D.16}$$

or, in the basis of the $|i\rangle$, as column vector $\vec{u}_{\nu mk}$. The standard phase choice is the convention to choose the $u_{\nu mm,0}$ to be real and positive for all m (in fact it can be shown that in this case $u_{\nu mm,0} = \sqrt{h_\nu/n_G}$). Starting with the first eigenvalue of k (denoted now simply as $k = 1$), the phases of the vectors $\vec{u}_{\nu m1}$ for $m = 2 \dots h_\nu$ can be chosen arbitrarily. This choice fixes the representation matrices (we demand them to be identical in equivalent representations), which are now given by

$$D_{ij}^{(\nu)}(a) = \langle \psi_i^{(\nu)1}|R_a|\psi_j^{(\nu)1}\rangle, \tag{D.17}$$

for all $a \in G$. When using the standard phase choice, the representation matrices can be shown to be directly related with the vectors $|\psi_i^{(\nu)j}\rangle$ via

$$D_{ij}^{(\nu)}(a) = \sqrt{\frac{n_G}{h_\nu}} u_{\nu ij,a}^\star. \tag{D.18}$$

This expression allows the phases of the remaining vectors $\vec{u}_{\nu mk}$ with $k > 1$ to be fixed by demanding that $u_{\nu mk,a}$ is equal to $\sqrt{h_\nu/n_G} \cdot D_{mk}^{(\nu)}(a)^\star$ for all m.

Reduction of Non-Regular Reps

The construction of the CSCO-II of G is the same for regular and non-regular representations. For non-regular reps, only the completion of the CSCO-II to the CSCO-III $\mathfrak{C}$ has to be adjusted. As it was shown in the paragraph following equation (D.11), an intrinsic state which is invariant under a non-trivial symmetry group $G_{\text{in}} \subset G$, leads to a contradiction which makes the definition of intrinsic group elements meaningless. The observation which saves the day is the following.

Lemma D.1.6. *If a class operator $C_i(s_j)$ of a subgroup $G(s_j)$ of G commutes with the symmetry group $G_{in} \subset G$, then the class operator $\bar{C}_i(s_j)$ of the corresponding intrinsic group $\bar{G}(s_j)$ has a well defined meaning.*

Proof. We repeat the calculation which led to the contradiction. On the one hand we have $\bar{C}_i(s_j)|a\rangle = \bar{C}_i(s_j)R_a|0\rangle = R_aC_i(s_j)|0\rangle$, on the other hand we have $\bar{C}_i(s_j)|a\rangle = \bar{C}_i(s_j)R_aT|0\rangle = R_aTC_i(s_j)|0\rangle$. But if the class operator and the symmetry group commute, we continue the calculation and obtain $\dots = R_aC_i(s_j)T|0\rangle = R_aC_i(s_j)|0\rangle$ for all $T \in G_{\text{in}}$. The contradiction vanishes. □

If we remove all subgroups from the canonical subgroup chain $G(s) = G(s_1) \supset G(s_2) \supset \dots$ whose class operators do not commute with G_{in}, we obtain the (non-canonical) subgroup chain $G(s') = G(s_{i_1}) \supset G(s_{i_2}) \supset \dots$ (with $i_1 < i_2$) and the lemma tells us that the CSCO-I's $\bar{C}(s') = (\bar{C}(s_{i_1}), \bar{C}(s_{i_2}), \dots)$ of $\bar{G}(s')$ still have a definite meaning. Theorem D.1.5 is replaced by:

Theorem D.1.7. *Let $\mathcal{V} = \text{span}\{R_a|0\rangle \mid a \in G\}$ be the rep space of a rep $R(G)$ of a group G with canonical subgroup chain $G(s)$ and symmetry group G_{in}. Then the set $\mathfrak{C} = (C, C(s), \bar{C}(s'))$ (with $\bar{C}(s')$ as defined above) is a CSCO on $\mathcal{V}$ (called CSCO-III). The corresponding eigenequation is given by*

$$\begin{pmatrix} C \\ C(s) \\ \bar{C}(s') \end{pmatrix} |\psi_m^{(\nu)\kappa}\rangle = \begin{pmatrix} \nu \\ m \\ \kappa \end{pmatrix} |\psi_m^{(\nu)\kappa}\rangle, \tag{D.19}$$

with $\kappa = (\bar{\lambda}(s_{i_1}), \bar{\lambda}(s_{i_2}), \dots)$.

To set the phases of the vectors $|\psi_m^{(\nu)\kappa}\rangle$ in such a way that the representation matrices

$$D_{ij}^{(\nu)\kappa}(a) = \langle\psi_i^{(\nu)\kappa}|R_a|\psi_j^{(\nu)\kappa}\rangle \tag{D.20}$$

do not depend on κ and agree with those of the regular rep, we choose the phases of $|\psi_1^{(\nu)\kappa}\rangle$ arbitrarily for all κ and use the known matrix elements of the regular rep matrices to determine the phases of the $|\psi_m^{(\nu)\kappa}\rangle$ with $m > 1$.

D.1.2 Symmetric Groups

The results of the preceding subsection are now specialized for the case that the group G under consideration is the symmetric group S_n. Using standard results of the representation theory of S_n, the construction of the CSCO $\mathfrak{C}$ can be simplified.

Representation Spaces

The elements of S_n are permutations which are denoted as

$$p = \begin{pmatrix} 1 & 2 & \dots & n \\ p(1) & p(2) & \dots & p(n) \end{pmatrix}. \tag{D.21}$$

The inverse of $p \in \mathsf{S}_n$ is given by

$$p^{-1} = \begin{pmatrix} p(1) & p(2) & \dots & p(n) \\ 1 & 2 & \dots & n \end{pmatrix} = \begin{pmatrix} 1 & 2 & \dots & n \\ p^{-1}(1) & p^{-1}(2) & \dots & p^{-1}(n) \end{pmatrix}, \tag{D.22}$$

where the right-hand side is obtained by permuting the columns of the matrix of the left-hand side. Let $\mathcal{H}_q$ be the Hilbert space of a qudit of dimension q and let an orthonormal basis $\{|0\rangle, |1\rangle, \dots, |q-1\rangle\}$ be fixed. In this section we will usually label these basis states using the Greek alphabet, i. e. $|0\rangle \equiv |\alpha\rangle$, $|1\rangle \equiv |\beta\rangle$, $|2\rangle \equiv |\gamma\rangle$, and so on. A representation $D(\mathsf{S}_n)$ of S_n on the q^n-dimensional linear vector space $\mathcal{H}_q^{\otimes n}$ is given by defining the action of a permutation $p \in \mathsf{S}_n$ on a n-fold tensor-product of one qudit basis states by

$$D(p)|i_1\,,\, i_2\,,\, \dots\,,\, i_n\rangle = |i_{p^{-1}(1)}\,,\, i_{p^{-1}(2)}\,,\, \dots\,,\, i_{p^{-1}(n)}\rangle. \tag{D.23}$$

We define the configuration of a standard basis vector in $\mathcal{H}_q^{\otimes n}$ as a string of integers of length q counting the number of times a certain one-qudit basis state appears in the vector, e. g. for $n = 5$ and $q = 4$ we have

$$\mathsf{config}(|\alpha\,\alpha\,\delta\,\alpha\,\beta\rangle) = (3, 1, 0, 1). \tag{D.24}$$

Since the configuration of basis vectors in $\mathcal{H}_q^{\otimes n}$ remains invariant under S_n, the representation space $\mathcal{H}_q^{\otimes n}$ of S_n decomposes into a direct sum of representation spaces each of which is characterized by a certain configuration string. The dimension of a rep space $\mathcal{V}$ with configuration $\mathsf{config} = (n_0, n_1, \dots, n_{q-1})$, $\sum_i n_i = n$, is given by the multinomial coefficient $n!/(\prod_i n_i!)$. Altogether there are $\binom{n+q-1}{q-1}$ different configurations.

The regular representation space $\mathcal{V}$ occurs only if $q = n$ and coincides with the rep space with configuration $(1, 1, \dots, 1)$. Its basis vectors $\{|p\rangle\}$ are obtained by applying the $g = n!$ elements of S_n to the generating state $|0\rangle := |0, 1, \dots, q-1\rangle$, i. e. $|p\rangle = D(p)|0\rangle$ with $p \in \mathsf{S}_n$.

For a non-regular representation space $\mathcal{V}$ with configuration $(n_0, n_1, \dots, n_{q-1}) \neq (1, 1, \dots, 1)$, we define the generating state $|0\rangle$ by

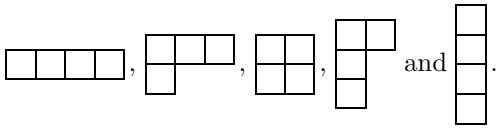
$$|0\rangle := \big|\underbrace{0, \dots, 0}_{n_0}\,,\, \underbrace{1, \dots, 1}_{n_1}\,,\, \dots\,,\, \underbrace{q-1, \dots, q-1}_{n_{q-1}}\big\rangle. \tag{D.25}$$

The generating state $|0\rangle$ is obviously invariant under a non-trivial symmetry group $G_{\text{in}} \subset \mathsf{S}_n$ and the dimension $d = n!/(\prod_i n_i!)$ of the non-regular rep space $\mathcal{V} = \text{span}\{D(p)|0\rangle \mid p \in \mathsf{S}_n\}$ is smaller than $n_G = n!$.

In the remaining part of this subsection we explain how the EFM described in the last subsection is applied to a rep space $\mathcal{V} \subset \mathcal{H}_q^{\otimes n}$ of S_n characterized by a certain configuration string.

Young Diagrams & CSCO-I

Each permutation can be decomposed into a product of disjoint cycles, for example $p = \left(\begin{smallmatrix}1&2&3&4&5&6\\3&5&1&2&4&6\end{smallmatrix}\right)$ can be written as the product $p = \left(\begin{smallmatrix}1&3\\3&1\end{smallmatrix}\right) \times \left(\begin{smallmatrix}2&4&5\\5&2&4\end{smallmatrix}\right) \times \left(\begin{smallmatrix}6\\6\end{smallmatrix}\right) \equiv (13)(254)(6)$. The conjugacy classes of the symmetric group S_n are characterized by a certain cycle structure: Each class contains only elements of one particular cycle structure. A cycle structure corresponds to a partition of n. A partition ν of n is given by a set of positive integers $\nu = [\nu_1, \nu_2, \ldots, \nu_v]$ such that $\sum_i \nu_i = n$ and $\nu_i \geq \nu_{i+1}$. It can be depicted as a Young diagram in which the i-th row contains ν_i boxes. For instance, for $n = 4$ there are the partitions $[4]$, $[3,1]$, $[2,2]$, $[2,1,1]$ and $[1,1,1,1]$ which correspond to the Young diagrams



Since the number of inequivalent representations of a group G is equal to the number of its conjugacy classes, the inequivalent reps of S_n may be labeled by Young diagrams corresponding to the partitions of n. This means that there is a one-to-one correspondence between the eigenvalues of the CSCO-I C of S_n and the Young diagrams corresponding to the partitions of n.

If a state $|\psi^{(\nu)}\rangle$ is in a rep space $\mathcal{V}_\nu \subset \mathcal{V}$ labeled by the Young diagram $\nu = [\nu_1, \nu_2, \ldots, \nu_v]$, it is necessarily an eigenstate of the class operators of S_n. It can be shown that the eigenvalues λ_2^n and λ_3^n of the 2- and 3-cycle class operators C_2^n and C_3^n can be expressed as functions of the Young diagram ν as follows,

$$\lambda_2^n = \frac{n}{2} + \frac{1}{2}\sum_{i=1}^{v} \nu_i(\nu_i - 2i) \tag{D.26a}$$

$$\lambda_3^n = \frac{2}{3}n - \frac{1}{2}n^2 + \frac{1}{3}\sum_{i=1}^{v} \nu_i\big[\nu_i^2 - (3i - 3/2)\nu_i + 3i(i-1)\big]. \tag{D.26b}$$

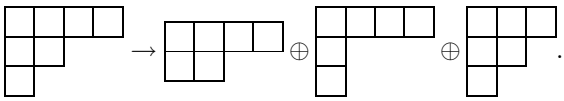
For $n < 6$ the 2-cycle class operator alone forms a CSCO-I, but for $n = 6$ degeneracy occurs which has to be lifted by adding for example the 3-cycle class operator. For $n < 15$ a CSCO-I C of S_n is given by the 2- and 3-cycle class operators, $C = (C_2^n, C_3^n)$. The eigenvalues λ_2^n and λ_3^n of C_2^n and C_3^n are listed in the form $\begin{smallmatrix}\lambda_2^n\\ \lambda_3^n\end{smallmatrix}$ for $n = 1 \ldots 7$ in figure D.1.

Young Tableaux & CSCO-II

A canonical subgroup chain $\mathsf{S}_n \supset G(s)$ of S_n is given by $G(s) = \mathsf{S}_{n-1} \supset \cdots \supset \mathsf{S}_3 \supset \mathsf{S}_2$. According to theorem D.1.2, the CSCO-II of S_n is given by $\big(C(\mathsf{S}_n), C(s) = (C(\mathsf{S}_{n-1}), \ldots, C(\mathsf{S}_2))\big)$, where the operator $C(\mathsf{S}_i)$ denotes the CSCO-I of S_i. It can be shown that this set of commuting observables is over-complete and that a simpler set is given by $\big(C_2^n, C_2^{n-1}, \ldots, C_2^2\big)$ which contains only 2-cycle class operators. While the 2-cycle eigenvalue of, say, C_2^i alone is not necessarily enough to deduce the eigenvalue $\lambda(\mathsf{S}_i)$ of $C(\mathsf{S}_i)$, the whole set of 2-cycle eigenvalues allows us to deduce the eigenvalues ν and $m = (\lambda(\mathsf{S}_{n-1}), \ldots, \lambda(\mathsf{S}_2))$ of $C(\mathsf{S}_n)$ and $C(s)$. The reason behind this fact is the branching law, which states that a subduced rep $D^{(\nu)}(\mathsf{S}_j) \downarrow \mathsf{S}_{j-1}$ of an irrep ν of S_j decomposes into

$$D^{(\nu)}(\mathsf{S}_j) \downarrow \mathsf{S}_{j-1} = \bigoplus_{\nu'} D^{(\nu')}(\mathsf{S}_{j-1}), \tag{D.27}$$

where the ν' are obtained from the Young diagram ν by removing a single box in all possible ways, e. g.



(D.28)

In figure D.1 the Young diagrams of S_n are shown in rows from $n = 1$ (*top row*) to $n = 7$ (*bottom row*). The branching law is indicated in the figure by gray arrows. Under each Young diagram the eigenvalues of λ_2^n and λ_3^n of C_2^n and C_3^n are shown in the form $\begin{smallmatrix}\lambda_2^n\\ \lambda_3^n\end{smallmatrix}$. As an example of how the eigenvalue list of $\big(C_2^n, C_2^{n-1}, \ldots, C_2^2\big)$ determines all Young diagrams (i. e. all eigenvalues ν and m of C and $C(s)$), let us consider the case $n = 7$ with

$$\big(C_2^7, C_2^6, C_2^5, C_2^4, C_2^3, C_2^2, C_2^1\big)|\psi_m^{(\nu)}\rangle = \big(3, 3, 0, 2, 0, 1, 0\big)|\psi_m^{(\nu)}\rangle \tag{D.29}$$

(we added C_2^1 whose only eigenvalue is zero) which is shown in figure D.1 in red. Starting at the top of the tree diagram with the eigenvalue 0 of C_2^1, we follow the gray arrow (in the opposite direction from top to bottom) which leads to the next eigenvalue 1 of C_2^2 (the resulting path is shown in red), and so on. By following a path provided by the branching law, any possible degeneracy of the 2-cycle eigenvalues (in our case the degeneracy of the eigenvalue 3 of C_2^6) is artificially lifted since it can be shown that only one of them will be accessible by the preceding path. Therefore, the eigenvalue list of the CSCO-II $\big(C_2^n, C_2^{n-1}, \ldots, C_2^2\big)$ describes a unique path connecting Young diagrams of $\mathsf{S}_n, \mathsf{S}_{n-1}, \ldots \mathsf{S}_2, \mathsf{S}_1$ which correspond to the eigenvalues (ν, m) of the original CSCO-II $\big(C(\mathsf{S}_n), C(s) = (C(\mathsf{S}_{n-1}), \ldots, C(\mathsf{S}_2))\big)$. Instead of writing down all n Young diagrams, they are summarized in a Young tableau $Y_m^{(\nu)}$ which is obtained from the Young diagram ν corresponding to the eigenvalue of $C(\mathsf{S}_n)$ by filling its boxes with the numbers $1, 2, \ldots n$ in such a way that if the box with number n is removed, we obtain a Young tableau which is shaped like the Young diagram corresponding to the eigenvalue $\lambda(\mathsf{S}_{n-1})$ of $C(\mathsf{S}_{n-1})$, and so on. For our example we obtain

$$(3,3,0,2,0,1,0) \to Y_m^{(\nu)} = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 4 & 6 \\ \hline 3 & 7 \\ \cline{1-2} 5 \\ \cline{1-1}\end{array} =$$

$$\left(\nu = \begin{array}{|c|c|c|c|}\hline & & & \\ \hline & \\ \cline{1-2} \\ \cline{1-1}\end{array}, m = \begin{array}{|c|c|c|c|}\hline & & & \\ \hline \\ \cline{1-1}\end{array}, \begin{array}{|c|c|c|}\hline & & \\ \hline \\ \cline{1-1}\end{array}, \begin{array}{|c|c|c|}\hline & & \\ \hline & \\ \cline{1-2}\end{array}, \begin{array}{|c|c|}\hline & \\ \hline \\ \cline{1-1}\end{array}, \begin{array}{|c|c|}\hline & \\ \hline & \\ \cline{1-2}\end{array}, \begin{array}{|c|}\hline \\ \hline \\ \cline{1-1}\end{array}, \begin{array}{|c|}\hline \\ \hline \\ \cline{1-1}\end{array}\right) . \tag{D.30}$$

As a consequence, a Young tableaux is always filled in such a way that the successive removal of boxes corresponding to the numbers n, $n-1$, etc., results in valid Young diagrams: In a Young tableau, the numbers always increase to the right and downwards.

Weyl Tableaux & CSCO-III

Let us consider the representation space $\mathcal{V} \subset \mathcal{H}_q^{\otimes n}$ with configuration $\mathsf{config} = (n_0, n_1, \ldots, n_{q-1})$, $\sum_i n_i = n$. The generating state $|0\rangle$ defined in equation (D.25),

$$|0\rangle := \big|\underbrace{0, \ldots, 0}_{n_0}\ ,\ \underbrace{1, \ldots, 1}_{n_1}\ ,\ \ldots\ ,\ \underbrace{q-1 \ldots q-1}_{n_{q-1}}\big\rangle, \tag{D.31}$$

is invariant under the action of the subgroup $G_{\text{in}} \subset \mathsf{S}_n$ containing $|G_{\text{in}}| = \prod_i n_i!$ elements. The group G_{in} decomposes S_n into a disjoint set of left cosets, $\mathsf{S}_n = G_{\text{in}} \cup a' G_{\text{in}} \cup b' G_{\text{in}} \cup \ldots$, where each coset contains $|G_{\text{in}}|$ elements and $\{a', b', \ldots\}$ denotes a set of coset representatives. An orthonormal basis $\{|i\rangle\}_{i=0\ldots d-1}$ of $\mathcal{V}$ is obtained by applying the $d = |\mathsf{S}_n|/|G_{\text{in}}| = n!/\prod_i n_i!$ coset representatives to the generating state $|0\rangle$, i. e. $|i\rangle = D(p_i)|0\rangle$, with $p_i \in \{a', b', \ldots\}$. The basis $\{|i\rangle\}_{i=0\ldots d-1}$ forms a subset

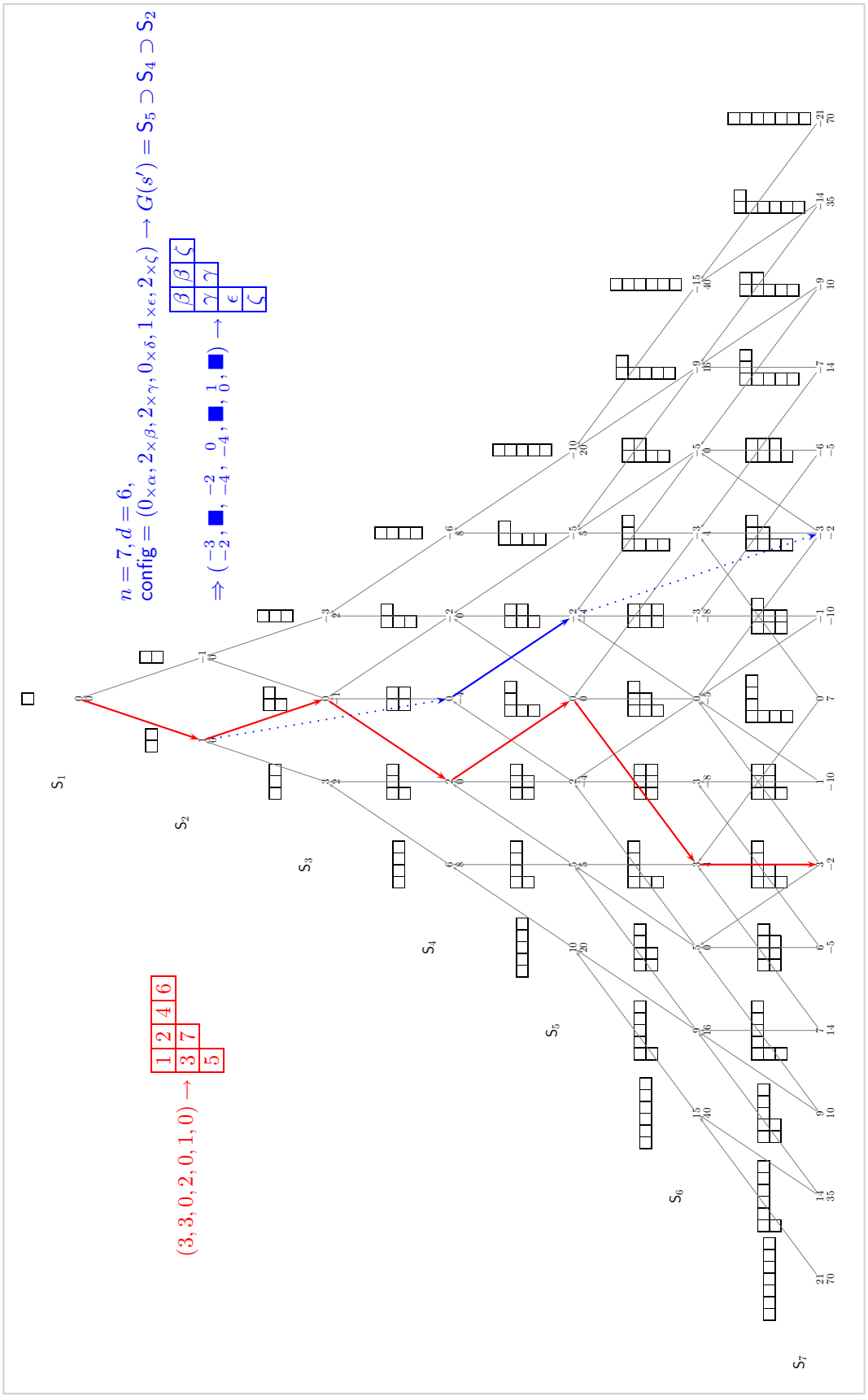


Figure D.1: Tree diagram showing the Young-diagrams corresponding to the partitions of $n = 1$ (*top row*) up to $n = 7$ (*bottom row*). The gray arrows indicate the branching law. Under each Young-diagram the corresponding eigenvalues λ_2^n of the 2-cycle class operator C_2^n and λ_3^n of the 3-cycle class operator C_3^n are shown in the form $\frac{\lambda_2^n}{\lambda_3^n}$.

of the computational basis of $\mathcal{H}_q^{\otimes n}$. We define the string of integers $\mathsf{config}' = (n_{i_1}, n_{i_2}, \ldots, n_{i_l})$ by removing all zeros from config. Then the structure of the symmetry group $G_{\text{in}} \subset \mathsf{S}_n$ is given by

$$G_{\text{in}} = \mathsf{S}_{n_{i_1}} \otimes \mathsf{S}_{n_{i_2}} \otimes \cdots \otimes \mathsf{S}_{n_{i_l}}. \tag{D.32}$$

It is easy to see that the class operators $C_i^{n(j)}$ of the subgroups $\{\mathsf{S}_{n(j)}\}_{j=1\ldots l}$ with $n(j) = \sum_{c=1}^{j} n_{i_c}$ commute with all the elements in G_{in}. Therefore, according to lemma D.1.6, the corresponding class operators $\bar{C}_i^{n(j)}$ of the intrinsic group are well defined and according to theorem D.1.7, the CSCO-II of $\mathcal{V}$ can be extended to a CSCO $\mathfrak{C}$ on $\mathcal{V}$ (called CSCO-III) by adding the set $\bar{C}(s') = (\bar{C}(\mathsf{S}_{n(l-1)}), \ldots, \bar{C}(\mathsf{S}_{n(2)}), \bar{C}(\mathsf{S}_{n(1)}))$ of CSCO-I's of $\bar{G}(s') = \bar{\mathsf{S}}_{n(l-1)} \supset \cdots \supset \bar{\mathsf{S}}_{n(2)} \supset \bar{\mathsf{S}}_{n(1)}$. Since $\bar{G}(s')$ is not a canonical subgroup chain of S_n anymore, the set $(\bar{C}(\mathsf{S}_n), \bar{C}(s'))$ cannot be replaced by a set of 2-cycle operators as it was done for the set $(C(\mathsf{S}_n), C(s))$.

To give an example, let $n = 7$, $q = 6$, and let us decompose the rep space $\mathcal{V} \subset \mathcal{H}_q^{\otimes n}$ with $\mathsf{config} = (0_{\times\alpha}, 2_{\times\beta}, 2_{\times\gamma}, 0_{\times\delta}, 1_{\times\epsilon}, 2_{\times\zeta})$. For this configuration we have $G(s') = \mathsf{S}_5 \supset \mathsf{S}_4 \supset \mathsf{S}_2$ and we consider an eigenvector $|\psi^{(\nu)\kappa}\rangle$ with eigenequation

$$(\bar{C}(\mathsf{S}_n), \bar{C}(s'))|\psi^{(\nu)\kappa}\rangle = (\tbinom{-3}{-2}, \blacksquare, \tbinom{-2}{-4}, \tbinom{0}{-4}, \blacksquare, \tbinom{1}{0}, \blacksquare)|\psi^{(\nu)\kappa}\rangle \equiv (\nu, \kappa)|\psi^{(\nu)\kappa}\rangle. \tag{D.33}$$

(We expanded the eigenvalue list to the length n by inserting black squares in the places where subgroups have been removed from $G(s)$ to obtain $G(s')$.) Since $n < 15$, $\bar{C}(\mathsf{S}_{n(j)}) = (\bar{C}_2^{n(j)}, \bar{C}_3^{n(j)})$ and we wrote the corresponding eigenvalues on top of each other. Our eigenvalue list (ν, κ) corresponds to a set of Young diagrams (indicated in blue in figure D.1),

$$(\tbinom{-3}{-2}, \blacksquare, \tbinom{-2}{-4}, \tbinom{0}{-4}, \blacksquare, \tbinom{1}{0}, \blacksquare) \to W_\kappa^{(\nu)} = \begin{array}{|c|c|c|}\hline \beta & \beta & \zeta \\ \hline \gamma & \gamma \\ \cline{1-2} \epsilon \\ \cline{1-1} \zeta \\ \cline{1-1} \end{array} = \left(\nu = \begin{array}{|c|c|c|}\hline & & \\ \hline & \\ \cline{1-2} \\ \cline{1-1} \\ \cline{1-1} \end{array}, \kappa = \begin{array}{|c|c|}\hline & \\ \hline & \\ \hline \\ \cline{1-1} \end{array}, \begin{array}{|c|c|}\hline & \\ \hline & \\ \hline \end{array}, \begin{array}{|c|c|}\hline & \\ \hline \end{array} \right), \tag{D.34}$$

which can be summarized in a so-called Weyl tableau $W_\kappa^{(\nu)}$ as follows: The Weyl tableau $W_\kappa^{(\nu)}$ is the Young diagram of $\mathsf{S}_n = \mathsf{S}_{n(l)}$ corresponding to the eigenvalue ν of $\bar{C}(\mathsf{S}_n) = C(\mathsf{S}_n)$ (compare with equation (D.13)) in which n_i boxes are filled with the i-th letter (basis state) of the Greek alphabet, and where the filling is done in such a way, that removing the n_{i_l} boxes filled with the i_l-th letter results in a Weyl tableau which is shaped like the Young diagram corresponding to the eigenvalue of $\bar{C}(\mathsf{S}_{n(l-1)})$, and so on. It follows that a Weyl tableaux is always filled in such a way that the successive removal of boxes corresponding to the i_l-th, i_{l-1}-th, etc., letter results in valid Young diagrams: In a Weyl tableau, letters have to increase downwards and never decrease to the right. Because of the former restriction, the maximum number of rows of a Young diagram is given by the number of letters (basis states) q.

Final Remarks

As a summary, the complete CSCO $\mathfrak{C}$ of S_n on a rep space $\mathcal{V} \subset \mathcal{H}_q^{\otimes n}$ characterized by a configuration $\mathsf{config} = (n_0, n_1, \ldots, n_{q-1})$ is given by the set of operators

$$\mathfrak{C} = \big(C_2^n, \quad C_2^{n-1}, \ldots, C_2^2, \quad \bar{C}(\mathsf{S}_{n(l-1)}), \ldots, \bar{C}(\mathsf{S}_{n(3)}), \bar{C}(\mathsf{S}_{n(2)})\big)^\dagger. \tag{D.35}$$

(Note that $\bar{C}(\mathsf{S}_{n(l)})$ and $\bar{C}(\mathsf{S}_{n(1)})$ are obsolete since $\bar{C}(\mathsf{S}_{n(l)}) = \bar{C}(\mathsf{S}_n) = C(\mathsf{S}_n)$ and $\bar{C}(\mathsf{S}_{n(1)})$ always corresponds to the Young diagram of the form $\nu = [n(1)]$). The eigenvectors $|\psi_m^{(\nu)\kappa}\rangle$ of the corresponding eigenequation

$$\begin{aligned} \mathfrak{C}|\psi_m^{(\nu)\kappa}\rangle &= (\nu, \kappa, m)|\psi_m^{(\nu)\kappa}\rangle \\ &\equiv (W_\kappa^{(\nu)}, Y_m^{(\nu)})|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle \end{aligned} \tag{D.36}$$

[†]Instead of constructing the operators $\bar{C}(\mathsf{S}_{n(j)})$, in practice it is of advantage to construct the operators $C(\mathcal{S}_{n(j)})$ which correspond to the so-called state permutation group $\mathcal{S}_{n(j)}$ and which are identical to the $\bar{C}(\mathsf{S}_{n(j)})$.

are labeled by (i) a Young diagram ν labeling inequivalent rep spaces $\mathcal{V}_\nu$, (ii) a Weyl tableau $W_\kappa^{(\nu)}$ labeling equivalent irreducible rep spaces $\mathcal{V}_{\nu,\kappa} \subset \mathcal{V}_\nu$, (iii) a Young tableau $Y_m^{(\nu)}$ labeling the basis states of an irreducible rep space.

The orthonormal $|\psi_m^{(\nu)\kappa}\rangle$ obtained from equation (D.36) are determined only up to a phase. The Yamanouchi phase convention demands off-diagonal matrix elements of adjacent transpositions to be positive. It can be shown that as a result of this convention, the following rule determines the phase of a vector

$$|\psi_m^{(\nu)\kappa}\rangle = \sum_{p\in\{a',b',\dots\}} u_{\nu m\kappa,p}|p\rangle. \tag{D.37}$$

Lemma D.1.8. *To satisfy the Yamanouchi phase convention, the phase of a vector $|\psi_m^{(\nu)\kappa}\rangle$ has to be chosen in such a way that $u_{\nu m\kappa,\mathfrak{p}} > 0$, where $\mathfrak{p}$ is called the principal term. The corresponding principal state $|\mathfrak{p}\rangle = |\mathfrak{p}_1, \mathfrak{p}_2, \dots, \mathfrak{p}_n\rangle$ is constructed by setting $\mathfrak{p}_i$ equal to the Greek letter in the box of the Weyl tableau $W_\kappa^{(\nu)}$, which is in the same position as the box in the Young tableau $Y_m^{(\nu)}$ containing the number i. For example,*

$$|\psi_m^{(\nu)\kappa}\rangle \equiv |W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle = \left|\begin{array}{|c|c|c|c|}\hline \alpha&\alpha&\alpha&\beta\\\hline \beta&\delta\\\cline{1-2}\gamma\\\cline{1-1}\end{array}\;\begin{array}{|c|c|c|c|}\hline 1&2&4&5\\\hline 3&7\\\cline{1-2}6\\\cline{1-1}\end{array}\right\rangle \rightarrow |\mathfrak{p}\rangle = |\alpha\alpha\beta\alpha\beta\gamma\delta\rangle. \tag{D.38}$$

The basis $\{|\psi_m^{(\nu)\kappa}\rangle\}$ of $\mathcal{V}$ obeying the Yamanouchi phase convention is called quasi-standard basis (it is called standard basis or Young-Yamanouchi basis for the special case of $\mathcal{V}$ being the regular rep space).

We close this subsection by defining an order of the Young tableaux $Y_m^{(\nu)}$ and Weyl tableaux $W_\kappa^{(\nu)}$ corresponding to a certain Young diagram $\nu = [\nu_1, \dots, \nu_v]$, $\sum_i \nu_i = n$. Before we start, note that the Young diagrams $\{\nu\}$ corresponding to partitions of n are ordered by sorting the q-digit strings given by a partition and supplemented with zeros if $v < q$ (note that $v \leq q$),

$$(\nu_1, \nu_2, \dots, \nu_q) = \big([\nu_1, \dots, \nu_v], \underbrace{0, \dots 0,}_{q-v}\big), \tag{D.39}$$

in descending order. The Young tableaux $\{Y_{m_i}^{(\nu)}\}_{i=1\dots h_\nu}$ are enumerated by their eigenvalue list of $(C_2^n, \dots C_2^2, C_2^1)$ and are sorted in descending order. The total number of Young tableaux (for a given Young diagram ν) is equal to the dimension h_ν of the irrep labeled by ν and is given by the hook length formula [CPW02, page 120]. To define an order of the Weyl tableaux $\{W_{\kappa i}^{(\nu)}\}_{i=1\dots\tau_\nu}$[‡], we enumerate them by their corresponding Gel'fand symbols which are then sorted in descending order. The Gel'fand symbol corresponding to a Weyl tableau $W_\kappa^{(\nu)}$ is defined as the list of non-negative integers

$$\big[\nu_1^{(1)}, \nu_2^{(1)}, \dots, \nu_q^{(1)}\,;\, \nu_1^{(2)}, \nu_2^{(2)}, \dots, \nu_{q-1}^{(2)}\,;\, \dots\,;\, \nu_1^{(q-1)}, \nu_2^{(q-1)}\,;\, \nu_1^{(q)}\big], \tag{D.40}$$

where $\nu^{(1)}$ denotes the Young diagram ν (supplemented with zeros as in equation (D.39)) and $\nu^{(j)} = [\nu_1^{(j)}, \dots, \nu_{q-j+1}^{(j)}]$ for $j = 2 \dots q$ denotes the Young diagram which results after removing the boxes with letters $q-1, \dots, q-j+1$ from $W_\kappa^{(\nu)}$. For instance,

$$\begin{array}{|c|c|c|c|}\hline \beta&\beta&\gamma&\gamma\\\hline \gamma&\delta\\\cline{1-2}\delta\\\cline{1-1}\end{array} \leftrightarrow \begin{pmatrix} 4\;2\;1\;0\\ 4\;1\;0\\ 2\;0\\ 0\end{pmatrix}. \tag{D.41}$$

Example

To give an example, we apply the EFM to the 12-dimensional rep space $\mathcal{V} \subset \mathcal{H}_4^{\otimes 4}$ with configuration $\mathsf{config} = (1,0,1,2)$. The computational basis of $\mathcal{V}$ is given by $\{|i\rangle\}_{i=0\dots11} = \{|\alpha\gamma\delta\delta\rangle, \dots, |\delta\delta\alpha\gamma\rangle, |\delta\delta\gamma\alpha\rangle\}$.

[‡]Note that τ_ν depends on the configuration of $\mathcal{V}$.

ν	$W_\kappa^{(\nu)}$	$Y_m^{(\nu)}$	$\lvert\alpha\gamma\delta\delta\rangle$	$\lvert\alpha\delta\gamma\delta\rangle$	$\lvert\alpha\delta\delta\gamma\rangle$	$\lvert\gamma\alpha\delta\delta\rangle$	$\lvert\gamma\delta\alpha\delta\rangle$	$\lvert\gamma\delta\delta\alpha\rangle$	$\lvert\delta\alpha\gamma\delta\rangle$	$\lvert\delta\alpha\delta\gamma\rangle$	$\lvert\delta\gamma\alpha\delta\rangle$	$\lvert\delta\gamma\delta\alpha\rangle$	$\lvert\delta\delta\alpha\gamma\rangle$	$\lvert\delta\delta\gamma\alpha\rangle$
			$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$
			$1/6\sqrt{3}$	$1/6\sqrt{3}$	$-1/6\sqrt{3}$	$1/6\sqrt{3}$	$1/6\sqrt{3}$	$-1/6\sqrt{3}$	$1/6\sqrt{3}$	$-1/6\sqrt{3}$	$1/6\sqrt{3}$	$-1/6\sqrt{3}$	$-1/6\sqrt{3}$	$-1/6\sqrt{3}$
			$1/6\sqrt{6}$	$-1/12\sqrt{6}$	$1/12\sqrt{6}$	$1/6\sqrt{6}$	$-1/12\sqrt{6}$	$1/12\sqrt{6}$	$-1/12\sqrt{6}$	$1/12\sqrt{6}$	$-1/12\sqrt{6}$	$1/12\sqrt{6}$	$-1/6\sqrt{6}$	$-1/6\sqrt{6}$
			0	$1/4\sqrt{2}$	$1/4\sqrt{2}$	0	$1/4\sqrt{2}$	$1/4\sqrt{2}$	$-1/4\sqrt{2}$	$-1/4\sqrt{2}$	$-1/4\sqrt{2}$	$-1/4\sqrt{2}$	0	0
			0	0	$1/6\sqrt{6}$	0	0	$-1/6\sqrt{6}$	0	$1/6\sqrt{6}$	0	$-1/6\sqrt{6}$	$1/6\sqrt{6}$	$-1/6\sqrt{6}$
			0	$1/4\sqrt{3}$	$1/12\sqrt{3}$	0	$-1/4\sqrt{3}$	$-1/12\sqrt{3}$	$1/4\sqrt{3}$	$1/12\sqrt{3}$	$-1/4\sqrt{3}$	$-1/12\sqrt{3}$	$-1/6\sqrt{3}$	$1/6\sqrt{3}$
			$1/2$	$1/4$	$1/4$	$-1/2$	$-1/4$	$-1/4$	$-1/4$	$-1/4$	$1/4$	$1/4$	0	0
			$1/6\sqrt{6}$	$-1/12\sqrt{6}$	$-1/12\sqrt{6}$	$1/6\sqrt{6}$	$-1/12\sqrt{6}$	$-1/12\sqrt{6}$	$-1/12\sqrt{6}$	$-1/12\sqrt{6}$	$-1/12\sqrt{6}$	$-1/12\sqrt{6}$	$1/6\sqrt{6}$	$1/6\sqrt{6}$
			0	$1/4\sqrt{2}$	$-1/4\sqrt{2}$	0	$1/4\sqrt{2}$	$-1/4\sqrt{2}$	$-1/4\sqrt{2}$	$1/4\sqrt{2}$	$-1/4\sqrt{2}$	$1/4\sqrt{2}$	0	0
			0	$1/4$	$-1/4$	0	$-1/4$	$1/4$	$1/4$	$-1/4$	$-1/4$	$1/4$	$1/2$	$-1/2$
			$1/6\sqrt{3}$	$1/12\sqrt{3}$	$-1/4\sqrt{3}$	$-1/6\sqrt{3}$	$-1/12\sqrt{3}$	$1/4\sqrt{3}$	$-1/12\sqrt{3}$	$1/4\sqrt{3}$	$1/12\sqrt{3}$	$-1/4\sqrt{3}$	0	0
			$1/6\sqrt{6}$	$-1/6\sqrt{6}$	0	$-1/6\sqrt{6}$	$1/6\sqrt{6}$	0	$1/6\sqrt{6}$	0	$-1/6\sqrt{6}$	0	0	0

Table D.1: The quasi-standard basis $\{|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle\}$ of $\mathcal{V} \subset \mathcal{H}_4^{\otimes 4}$ with $\mathsf{config} = (1,0,1,2)$.

$$D^{([3,1])}(p_{(12)}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad D^{([3,1])}(p_{(23)}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \qquad D^{([3,1])}(p_{(34)}) = \begin{pmatrix} -\frac{1}{3} & \frac{\sqrt{8}}{3} & 0 \\ \frac{\sqrt{8}}{3} & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$D^{([2,1,1])}(p_{(12)}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad D^{([2,1,1])}(p_{(23)}) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad D^{([2,1,1])}(p_{(34)}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -\frac{1}{3} & \frac{\sqrt{8}}{3} \\ 0 & \frac{\sqrt{8}}{3} & \frac{1}{3} \end{pmatrix}$$

The sorted basis vectors $|\psi_m^{(\nu)\kappa}\rangle \equiv |W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle$ of the quasi-standard basis of $\mathcal{V}$ obtained via the EFM are shown in table D.1. As it can be seen from the table, $\mathcal{V}$ decomposes into 4 irreducible subspaces,

$$\mathcal{V} = \mathcal{V}_{[4]} \oplus \mathcal{V}_{[3,1],1} \oplus \mathcal{V}_{[3,1],2} \oplus \mathcal{V}_{[2,2]} \oplus \mathcal{V}_{[2,1,1]}, \tag{D.42}$$

where the irrep $[3,1]$ is two-fold degenerated. The dimensions of the irreps are given by $h_{[4]} = 1$, $h_{[3,1]} = 3$, $h_{[2,2]} = 2$ and $h_{[2,1,1]} = 3$. Since $\tau_{[3,1]} = 2$ and $\tau_\nu = 1$ for the remaining ν, we can easily check that $\sum_\nu \tau_\nu \times h_\nu = \dim(\mathcal{V})$. Below the table, the representation matrices of adjacent transpositions are shown for the two 3-dimensional irreps $[3,1]$ and $[2,1,1]$.

D.2 Schur Transform

The Schur transform is a unitary transformation relating the standard computational basis of n qudits of dimension q to the Schur basis, a basis associated with the representation theory of the symmetric and general linear groups. In the preceding subsection it was shown that the vector space of n qudits decomposes into a direct sum of representation spaces $\mathcal{V}$ of the symmetric group, each of which is characterized by the frequency distribution of the one-qudit basis states. In this section we show that the Schur basis is given by the collection of the quasi-standard bases of the symmetric group of all the rep spaces $\mathcal{V}$.

D.2.1 The Schur Basis

Let $\mathcal{H}_q$ denote the Hilbert space of a qudit of dimension q and let an orthonormal basis $\{|0\rangle, |1\rangle, \ldots, |q-1\rangle\}$ be fixed. Occasionally we label these q basis states using letters from the Greek alphabet, i. e. $|0\rangle \equiv |\alpha\rangle$, $|1\rangle \equiv |\beta\rangle$, $|2\rangle \equiv |\gamma\rangle$, and so on. The set of invertible linear transformations on $\mathcal{H}_q$ is called the general linear group $\mathsf{GL}(q,\mathbb{C}) = \mathsf{GL}_q$. An element $\rho \in \mathsf{GL}_q$ is defined by $q \times q$ complex numbers ρ_{ij} and transforms the basis states according to

$$\rho|j\rangle = \sum_{i=0}^{q-1} \rho_{ij} |i\rangle. \tag{D.43}$$

The computational basis of the Hilbert space $\mathcal{H}_q^{\otimes n}$ of n qudits of dimension q is given by the set of n-fold product states of the one-qudit basis states,

$$\mathcal{H}_q^{\otimes n} = \mathrm{span}\big\{|i_1, i_2, \ldots, i_n\rangle\big\}, \tag{D.44}$$

with $0 \leq i_j < q$ for $j \in \{1, 2, \ldots, n\}$. A representation $D(\mathsf{GL}_q)$ of GL_q on $\mathcal{H}_q^{\otimes n}$ is defined by

$$D(\rho)|i_1, i_2, \ldots, i_n\rangle = \rho \otimes \rho \cdots \otimes \rho |i_1, i_2, \ldots, i_n\rangle \tag{D.45}$$

for any $\rho \in \mathsf{GL}_q$. For the symmetric group S_n a representation $D(\mathsf{S}_n)$ on $\mathcal{H}_q^{\otimes n}$ was defined by equation (D.23), where the action of $D(p)$ on a computational basis state was defined as

$$D(p)|i_1, i_2, \ldots, i_n\rangle = |i_{p^{-1}(1)}, i_{p^{-1}(2)}, \ldots, i_{p^{-1}(n)}\rangle \tag{D.46}$$

for any $p \in \mathsf{S}_n$. Hence, the q^n-dimensional vector space $\mathcal{H}_q^{\otimes n}$ forms a representation space for both the symmetric group S_n and the general linear group GL_q. An important observation is the following lemma.

Lemma D.2.1. *Elements of $D(\mathsf{S}_n)$ and $D(\mathsf{GL}_q)$ commute, i. e.*

$$[D(p), D(\rho)] = 0, \tag{D.47}$$

for all $p \in \mathsf{S}_n$ and all $\rho \in \mathsf{GL}_q$.

As it was discussed in subsection D.1.2, $\mathcal{H}_q^{\otimes n}$ is a direct sum of rep spaces $\mathcal{V}$ of S_n, each of which is spanned by a subset of the computational basis which is characterized by a configuration string $\mathsf{config} = (n_0, n_1, \ldots, n_{q-1})$ of length q (with $\sum_i n_i = n$) specifying the number of one-qudit basis states (see eq. (D.24)). Let us now calculate the quasi-standard basis of S_n for each of the $\binom{n+q-1}{q-1}$ different representation spaces $\mathcal{V} \subset \mathcal{H}_q^{\otimes n}$ by solving the eigenvalue equation (D.36) and applying the Yamanouchi phase convention[§]. The collection of all basis states obtained this way,

$$\Big\{|W_{\kappa_j}^{(\nu)} Y_{m_i}^{(\nu)}\rangle\Big\}, \text{ with } \nu = \{[n], [n-1,1], \ldots\},\ j = \{1, \ldots, h_\nu(\mathsf{GL}_q)\},\ i = \{1, \ldots, h_\nu(\mathsf{S}_n)\}, \tag{D.48}$$

forms the Schur basis which has the following properties:

Lemma D.2.2 (Properties of the Schur basis)**.**

(i) The subspaces $\mathcal{V}_{\nu,\kappa}$ which are spanned by the $\big\{|W_\kappa^{(\nu)} Y_{m_i}^{(\nu)}\rangle\big\}_{i=1,\ldots,h_\nu(\mathsf{S}_n)}$ are irreducible rep spaces of S_n. For $p \in \mathsf{S}_n$ we have

$$D(p)|W_\kappa^{(\nu)} Y_{m_i}^{(\nu)}\rangle = \sum_{j=1}^{h_\nu(\mathsf{S}_n)} D_{ji}^{(\nu)}(p)|W_\kappa^{(\nu)} Y_{m_j}^{(\nu)}\rangle. \tag{D.49}$$

The dimension $h_\nu(\mathsf{S}_n)$ of these irreps is given by the hook length formula (see e. g. [CPW02, page 120]) and depends only on ν.

(ii) The subspaces $\mathcal{V}_\nu^m$ which are spanned by the $\big\{|W_{\kappa_j}^{(\nu)} Y_m^{(\nu)}\rangle\big\}_{j=1,\ldots,h_\nu(\mathsf{GL}_q)}$ are irreducible rep spaces of GL_q. For $\rho \in \mathsf{GL}_q$ we have

$$D(\rho)|W_{\kappa_i}^{(\nu)} Y_m^{(\nu)}\rangle = \sum_{j=1}^{h_\nu(\mathsf{GL}_q)} D_{ji}^{(\nu)}(\rho)|W_{\kappa_j}^{(\nu)} Y_m^{(\nu)}\rangle. \tag{D.50}$$

The dimension $h_\nu(\mathsf{GL}_q)$ of these irreps is given by the Robinson formula (see e. g. [CPW02, page 319]) and depends on ν and q.

Proof. Part (i) was shown in detail in subsection D.1.2. To give a (partial) prove of part (ii), we recall that an over-complete CSCO-III of S_n on a subspace $\mathcal{V} \subset \mathcal{H}_q^{\otimes n}$ characterized by a certain configuration config is given by the union of $(C(\mathsf{S}_n), C(s))$ and $(C(\mathsf{S}_n), \bar{C}(s'))$ (see subsection D.1.2), and that the $|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle$ are the eigenstates of the CSCO-III with eigenvalues $Y_m^{(\nu)}$ and $W_\kappa^{(\nu)}$,

$$\begin{aligned}\big(C(\mathsf{S}_n), C(s)\big)|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle &= Y_m^{(\nu)}|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle \\ \big(C(\mathsf{S}_n), \bar{C}(s')\big)|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle &= W_\kappa^{(\nu)}|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle.\end{aligned}$$

Let us now assume that the operators $(C(\mathsf{S}_n), C(s))$ are extended to the corresponding operators on $\mathcal{H}_q^{\otimes n}$. By lemma D.2.1, any $D(\rho)$ with $\rho \in \mathsf{GL}_q$ commutes with $(C(\mathsf{S}_n), C(s))$,

$$\big(C(\mathsf{S}_n), C(s)\big) D(\rho)|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle = Y_m^{(\nu)} D(\rho)|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle,$$

and we conclude that

$$D(\rho)|W_{\kappa_i}^{(\nu)} Y_m^{(\nu)}\rangle = \sum_{j=1}^{h_\nu(\mathsf{GL}_q)} D_{ji}^{(\nu)m}(\rho)|W_{\kappa_j}^{(\nu)} Y_m^{(\nu)}\rangle.$$

Eventually, it can be shown that the representations $D^{(\nu)m}(\rho)$ do not depend on m and that they are irreducible. □

[§]Actually we do not have to perform this calculation for all the spaces $\mathcal{V} \subset \mathcal{H}_q^{\otimes n}$. If the non-zero elements config' and $\widetilde{\mathsf{config}'}$ of the configurations config and $\widetilde{\mathsf{config}}$ of rep spaces $\mathcal{V}$ and $\tilde{\mathcal{V}}$ are the same, the CSCO-III of $\mathcal{V}$ and $\tilde{\mathcal{V}}$ is identical and we can adopt solutions already known by relabeling the basis states and Weyl tableaux.

It follows from lemma D.2.2 that the common representation space $\mathcal{H}_q^{\otimes n}$ decomposes into a direct sum of tensor spaces,

$$\mathcal{H}_q^{\otimes n} = \bigoplus_{\nu} \operatorname{span}\Big\{|W_{\kappa_j}^{(\nu)}\rangle\Big\}_{j=1,\ldots,h_\nu(\mathsf{GL}_q)} \bigotimes \operatorname{span}\Big\{|Y_{m_i}^{(\nu)}\rangle\Big\}_{i=1,\ldots,h_\nu(\mathsf{S}_n)}, \tag{D.51}$$

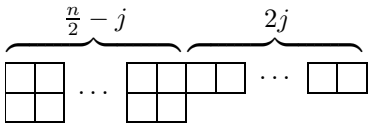
where the sum over the Young diagrams ν runs over all diagrams with at most q rows. Any product of operators $D(p)$ and $D(\rho)$ becomes block-diagonal,

$$D(\rho)D(p) = \bigoplus_{\nu} D^{(\nu)}(\rho) \otimes D^{(\nu)}(p), \tag{D.52}$$

for any $p \in \mathsf{S}_n$ and any $\rho \in \mathsf{GL}_q$.

The Special Case of Qubits

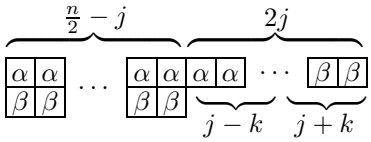
For $q = 2$ things are simpler, as Young diagrams $\{\nu = [\nu_1, \nu_2]\}$, with $\nu_1 \geq \nu_2 \geq 0$ and $\nu_1 + \nu_2 = n$, consist of at most two rows and can be labeled by an index j, where $2j$ is the number of columns consisting of one row only ($j = 0 \ldots \frac{n}{2}$ if n is even, $j = \frac{1}{2} \ldots \frac{n}{2}$ if n is odd).

$j \leftrightarrow \nu = [\frac{n}{2} + j, \frac{n}{2} - j] =$ 

The dimension of the irrep j of S_n is given by the hook length formula, which in this case yields

$$h_j(\mathsf{S}_n) = \binom{n}{n/2 - j} \frac{2j+1}{n/2 + j + 1}. \tag{D.53a}$$

The Weyl tableaux $W_k^{(\nu=j)}$ are now labeled by $k = -j, \ldots, j$, where $j + k$ denotes the number of β's (ones) in the first row of the Weyl tableaux:



The total number of Weyl tableaux for a given j is

$$h_j(\mathsf{GL}_2) = 2j + 1. \tag{D.53b}$$

Equation (D.51) becomes

$$\mathcal{H}_2^{\otimes n} = \bigoplus_{j=0,1/2}^{n/2} \operatorname{span}\Big\{|W_k^{(j)}\rangle\Big\}_{k=-j\cdots+j} \otimes \operatorname{span}\Big\{|Y_{m_i}^{(j)}\rangle\Big\}_{i=1\ldots h_j(\mathsf{S}_n)}. \tag{D.54}$$

D.2.2 Examples

Within the framework of this theses, a matlab program has been developed which obtains the Schur basis for given values of n and q by implementing the ideas presented in subsections D.1.2 and D.2.1. To give some examples, we present some of the Schur bases obtained by the program.

The first example is the Schur basis for $q = 3$ and $n = 3$. The Hilbert space of the three qudits decomposes as

$$\mathcal{H}_3^{\otimes 3} = \operatorname{span}\Big\{|W_{\kappa_j}^{([3])}\rangle\Big\}_{j=1\dots10} \otimes |Y_{m_1}^{([3])}\rangle \bigoplus \operatorname{span}\Big\{|W_{\kappa_j}^{([2,1])}\rangle\Big\}_{j=1\dots8} \otimes \operatorname{span}\Big\{|Y_{m_i}^{([2,1])}\rangle\Big\}_{i=1\dots2} \bigoplus |W_{\kappa_1}^{([1,1,1])}\rangle \otimes |Y_{m_1}^{([1,1,1])}\rangle, \quad \text{(D.55)}$$

and the Schur-basis-vectors are listed in table D.2.

As a second example, we consider $q = 2$ and $n = 1, 2, 3, 4, 5$. The resulting Schur-basis vectors are listed in table D.3 for $n = 2, 3, 4$ and in table D.4 for $n = 5$. The Schur basis of $\mathcal{H}_2^{\otimes 1}$ coincides with the computational basis, i. e. we have $\left|\boxed{\alpha}\,\boxed{1}\right\rangle = |\alpha\rangle$ and $\left|\boxed{\beta}\,\boxed{1}\right\rangle = |\beta\rangle$. The Schur basis of $\mathcal{H}_2^{\otimes 2}$ is given by

$$\begin{aligned}
\left|\boxed{\alpha}\boxed{\alpha}\,\boxed{1}\boxed{2}\right\rangle &= |\alpha\alpha\rangle, \\
\left|\boxed{\alpha}\boxed{\beta}\,\boxed{1}\boxed{2}\right\rangle &= (|\alpha\beta\rangle + |\beta\alpha\rangle)/\sqrt{2}, \\
\left|\boxed{\beta}\boxed{\beta}\,\boxed{1}\boxed{2}\right\rangle &= |\beta\beta\rangle, \text{ and} \\
\left|\begin{array}{|c|}\hline\alpha\\\hline\beta\\\hline\end{array}\;\begin{array}{|c|}\hline1\\\hline2\\\hline\end{array}\right\rangle &= (|\alpha\beta\rangle - |\beta\alpha\rangle)/\sqrt{2}.
\end{aligned} \quad \text{(D.56)}$$

The Hilbert space of the $n = 5$ qubits decomposes as in equation (D.54) with $h_{1/2}(\mathsf{S}_5) = 5$, $h_{3/2}(\mathsf{S}_5) = 4$ and $h_{5/2}(\mathsf{S}_5) = 1$.

D.2.3 Application: Communication without a Shared Reference Frame

An important application for the Schur transform in the context of quantum information theory is classical and quantum communication without a shared reference frame [BRS03; BRS07]. Let us restrict ourselves to the case where two parties, say Alice and Bob, are connected via an ideal quantum channel transmitting qubits. If they don't share a common reference frame, the action of the quantum channel is to apply a random change of the computational basis spanning the Hilbert space $\mathcal{H}_2$ of the qubits. When Alice sends n qubits in the state $\rho \in \mathcal{S}(\mathcal{H}_2^{\otimes n})$, Bob receives the state

$$\mathcal{M}_n(\rho) = \int U^{\otimes n}\, \rho\, U^{\dagger\otimes n}\, dU. \quad \text{(D.57)}$$

Using the Schur basis, the Hilbert space of n qubits decomposes as in equation (D.54) and Schur's lemma assures that the action of $\mathcal{M}_n$ can be written as

$$\mathcal{M}_n = \sum_{j=0,1/2}^{n/2} \left(\mathcal{D}_{2j+1}^{(j)} \otimes \mathcal{I}_{h_j(\mathsf{S}_n)}^{(j)}\right) \cdot \Pi_j, \quad \text{(D.58)}$$

where Π_j denotes the projection on the Young diagram j and $\mathcal{D}_{2j+1}^{(j)}$ denotes the complete depolarizing channel on the $2j+1$ dimensional tensor space of the irreps of SU_2.

To transmit classical information to Bob, Alice chooses a normalized state

$$|\Gamma_{m_i}^{(j)}\rangle = \sum_{k=-j}^{+j} \alpha_k |W_k^{(j)}\rangle |Y_{m_i}^{(j)}\rangle, \qquad \alpha_k \in \mathbb{C}, \quad \text{(D.59)}$$

for each Young diagram $j = 0, 1/2 \dots n$ and Young tableau $Y_{m_i}^{(j)}$, $i = 1 \dots h_j(\mathsf{S}_n)$ (for example $\alpha_k = \delta_{k,+j}$). Altogether there are

$$c_n = \sum_{j=0,1/2}^{n/2} h_j(\mathsf{S}_n) = \begin{cases} \binom{n}{n/2} & \text{if } n \text{ is even} \\ \binom{n}{n/2+1/2} & \text{if } n \text{ is odd} \end{cases} \quad \text{(D.60)}$$

ν	$W_\kappa^{(\nu)}$	$Y_m^{(\nu)}$						
			$\lvert\alpha\alpha\alpha\rangle$ 1					
			$\lvert\alpha\alpha\beta\rangle$ $1/3\sqrt{3}$	$\lvert\alpha\beta\alpha\rangle$ $1/3\sqrt{3}$	$\lvert\beta\alpha\alpha\rangle$ $1/3\sqrt{3}$			
			$\lvert\alpha\beta\beta\rangle$ $1/3\sqrt{3}$	$\lvert\beta\alpha\beta\rangle$ $1/3\sqrt{3}$	$\lvert\beta\beta\alpha\rangle$ $1/3\sqrt{3}$			
			$\lvert\beta\beta\beta\rangle$ 1					
			$\lvert\alpha\alpha\gamma\rangle$ $1/3\sqrt{3}$	$\lvert\alpha\gamma\alpha\rangle$ $1/3\sqrt{3}$	$\lvert\gamma\alpha\alpha\rangle$ $1/3\sqrt{3}$			
			$\lvert\alpha\beta\gamma\rangle$ $1/6\sqrt{6}$	$\lvert\alpha\gamma\beta\rangle$ $1/6\sqrt{6}$	$\lvert\beta\alpha\gamma\rangle$ $1/6\sqrt{6}$	$\lvert\beta\gamma\alpha\rangle$ $1/6\sqrt{6}$	$\lvert\gamma\alpha\beta\rangle$ $1/6\sqrt{6}$	$\lvert\gamma\beta\alpha\rangle$ $1/6\sqrt{6}$
			$\lvert\beta\beta\gamma\rangle$ $1/3\sqrt{3}$	$\lvert\beta\gamma\beta\rangle$ $1/3\sqrt{3}$	$\lvert\gamma\beta\beta\rangle$ $1/3\sqrt{3}$			
			$\lvert\alpha\gamma\gamma\rangle$ $1/3\sqrt{3}$	$\lvert\gamma\alpha\gamma\rangle$ $1/3\sqrt{3}$	$\lvert\gamma\gamma\alpha\rangle$ $1/3\sqrt{3}$			
			$\lvert\beta\gamma\gamma\rangle$ $1/3\sqrt{3}$	$\lvert\gamma\beta\gamma\rangle$ $1/3\sqrt{3}$	$\lvert\gamma\gamma\beta\rangle$ $1/3\sqrt{3}$			
			$\lvert\gamma\gamma\gamma\rangle$ 1					
			$\lvert\alpha\alpha\beta\rangle$ $1/3\sqrt{6}$ 0	$\lvert\alpha\beta\alpha\rangle$ $-1/6\sqrt{6}$ $1/2\sqrt{2}$	$\lvert\beta\alpha\alpha\rangle$ $-1/6\sqrt{6}$ $-1/2\sqrt{2}$			
			$\lvert\alpha\beta\beta\rangle$ $1/6\sqrt{6}$ $1/2\sqrt{2}$	$\lvert\beta\alpha\beta\rangle$ $1/6\sqrt{6}$ $-1/2\sqrt{2}$	$\lvert\beta\beta\alpha\rangle$ $-1/3\sqrt{6}$ 0			
			$\lvert\alpha\alpha\gamma\rangle$ $1/3\sqrt{6}$ 0	$\lvert\alpha\gamma\alpha\rangle$ $-1/6\sqrt{6}$ $1/2\sqrt{2}$	$\lvert\gamma\alpha\alpha\rangle$ $-1/6\sqrt{6}$ $-1/2\sqrt{2}$			
			$\lvert\alpha\beta\gamma\rangle$ $1/3\sqrt{3}$ 0	$\lvert\alpha\gamma\beta\rangle$ $-1/6\sqrt{3}$ $1/2$	$\lvert\beta\alpha\gamma\rangle$ $1/3\sqrt{3}$ 0	$\lvert\beta\gamma\alpha\rangle$ $-1/6\sqrt{3}$ $1/2$	$\lvert\gamma\alpha\beta\rangle$ $-1/6\sqrt{3}$ $-1/2$	$\lvert\gamma\beta\alpha\rangle$ $-1/6\sqrt{3}$ $-1/2$
			$\lvert\beta\beta\gamma\rangle$ $1/3\sqrt{6}$ 0	$\lvert\beta\gamma\beta\rangle$ $-1/6\sqrt{6}$ $1/2\sqrt{2}$	$\lvert\gamma\beta\beta\rangle$ $-1/6\sqrt{6}$ $-1/2\sqrt{2}$			
			$\lvert\alpha\beta\gamma\rangle$ 0 $1/3\sqrt{3}$	$\lvert\alpha\gamma\beta\rangle$ $1/2$ $1/6\sqrt{3}$	$\lvert\beta\alpha\gamma\rangle$ 0 $-1/3\sqrt{3}$	$\lvert\beta\gamma\alpha\rangle$ $-1/2$ $-1/6\sqrt{3}$	$\lvert\gamma\alpha\beta\rangle$ $1/2$ $-1/6\sqrt{3}$	$\lvert\gamma\beta\alpha\rangle$ $-1/2$ $1/6\sqrt{3}$
			$\lvert\alpha\gamma\gamma\rangle$ $1/6\sqrt{6}$ $1/2\sqrt{2}$	$\lvert\gamma\alpha\gamma\rangle$ $1/6\sqrt{6}$ $-1/2\sqrt{2}$	$\lvert\gamma\gamma\alpha\rangle$ $-1/3\sqrt{6}$ 0			
			$\lvert\beta\gamma\gamma\rangle$ $1/6\sqrt{6}$ $1/2\sqrt{2}$	$\lvert\gamma\beta\gamma\rangle$ $1/6\sqrt{6}$ $-1/2\sqrt{2}$	$\lvert\gamma\gamma\beta\rangle$ $-1/3\sqrt{6}$ 0			
			$\lvert\alpha\beta\gamma\rangle$ $1/6\sqrt{6}$	$\lvert\alpha\gamma\beta\rangle$ $-1/6\sqrt{6}$	$\lvert\beta\alpha\gamma\rangle$ $-1/6\sqrt{6}$	$\lvert\beta\gamma\alpha\rangle$ $1/6\sqrt{6}$	$\lvert\gamma\alpha\beta\rangle$ $1/6\sqrt{6}$	$\lvert\gamma\beta\alpha\rangle$ $-1/6\sqrt{6}$

Table D.2: Schur basis $\{|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle\}$ of $\mathcal{H}_3^{\otimes 3}$.

j/ν	$k/W_\kappa^{(\nu)}$	$Y_m^{(\nu)}$	
1	1		$\lvert\alpha\alpha\rangle$
	$\alpha\alpha$	$\begin{smallmatrix}1&2\end{smallmatrix}$	1
	0		$\lvert\alpha\beta\rangle$ $\lvert\beta\alpha\rangle$
	$\begin{smallmatrix}\alpha&\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2\end{smallmatrix}$	$1/2\sqrt{2}$ $1/2\sqrt{2}$
	-1		$\lvert\beta\beta\rangle$
	$\begin{smallmatrix}\beta&\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2\end{smallmatrix}$	1
0	0		$\lvert\alpha\beta\rangle$ $\lvert\beta\alpha\rangle$
	$\begin{smallmatrix}\alpha\\\beta\end{smallmatrix}$	$\begin{smallmatrix}1\\2\end{smallmatrix}$	$1/2\sqrt{2}$ $-1/2\sqrt{2}$

j/ν	$k/W_\kappa^{(\nu)}$	$Y_m^{(\nu)}$	
3/2	3/2		$\lvert\alpha\alpha\alpha\rangle$
	$\alpha\alpha\alpha$	$\begin{smallmatrix}1&2&3\end{smallmatrix}$	1
	1/2		$\lvert\alpha\alpha\beta\rangle$ $\lvert\alpha\beta\alpha\rangle$ $\lvert\beta\alpha\alpha\rangle$
	$\alpha\alpha\beta$	$\begin{smallmatrix}1&2&3\end{smallmatrix}$	$1/3\sqrt{3}$ $1/3\sqrt{3}$ $1/3\sqrt{3}$
	$-1/2$		$\lvert\alpha\beta\beta\rangle$ $\lvert\beta\alpha\beta\rangle$ $\lvert\beta\beta\alpha\rangle$
	$\alpha\beta\beta$	$\begin{smallmatrix}1&2&3\end{smallmatrix}$	$1/3\sqrt{3}$ $1/3\sqrt{3}$ $1/3\sqrt{3}$
	$-3/2$		$\lvert\beta\beta\beta\rangle$
	$\beta\beta\beta$	$\begin{smallmatrix}1&2&3\end{smallmatrix}$	1
1/2	1/2		$\lvert\alpha\alpha\beta\rangle$ $\lvert\alpha\beta\alpha\rangle$ $\lvert\beta\alpha\alpha\rangle$
	$\begin{smallmatrix}\alpha&\alpha\\\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2\\3\end{smallmatrix}$	$1/3\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$
		$\begin{smallmatrix}1&3\\2\end{smallmatrix}$	0 $1/2\sqrt{2}$ $-1/2\sqrt{2}$
	$-1/2$		$\lvert\alpha\beta\beta\rangle$ $\lvert\beta\alpha\beta\rangle$ $\lvert\beta\beta\alpha\rangle$
	$\begin{smallmatrix}\alpha&\beta\\\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2\\3\end{smallmatrix}$	$1/6\sqrt{6}$ $1/6\sqrt{6}$ $-1/3\sqrt{6}$
		$\begin{smallmatrix}1&3\\2\end{smallmatrix}$	$1/2\sqrt{2}$ $-1/2\sqrt{2}$ 0

j/ν	$k/W_\kappa^{(\nu)}$	$Y_m^{(\nu)}$	
2	2		$\lvert\alpha\alpha\alpha\alpha\rangle$
	$\alpha\alpha\alpha\alpha$	$\begin{smallmatrix}1&2&3&4\end{smallmatrix}$	1
	1		$\lvert\alpha\alpha\alpha\beta\rangle$ $\lvert\alpha\alpha\beta\alpha\rangle$ $\lvert\alpha\beta\alpha\alpha\rangle$ $\lvert\beta\alpha\alpha\alpha\rangle$
	$\alpha\alpha\alpha\beta$	$\begin{smallmatrix}1&2&3&4\end{smallmatrix}$	1/2 1/2 1/2 1/2
	0		$\lvert\alpha\alpha\beta\beta\rangle$ $\lvert\alpha\beta\alpha\beta\rangle$ $\lvert\alpha\beta\beta\alpha\rangle$ $\lvert\beta\alpha\alpha\beta\rangle$ $\lvert\beta\alpha\beta\alpha\rangle$ $\lvert\beta\beta\alpha\alpha\rangle$
	$\alpha\alpha\beta\beta$	$\begin{smallmatrix}1&2&3&4\end{smallmatrix}$	$1/6\sqrt{6}$ $1/6\sqrt{6}$ $1/6\sqrt{6}$ $1/6\sqrt{6}$ $1/6\sqrt{6}$ $1/6\sqrt{6}$
	-1		$\lvert\alpha\beta\beta\beta\rangle$ $\lvert\beta\alpha\beta\beta\rangle$ $\lvert\beta\beta\alpha\beta\rangle$ $\lvert\beta\beta\beta\alpha\rangle$
	$\alpha\beta\beta\beta$	$\begin{smallmatrix}1&2&3&4\end{smallmatrix}$	1/2 1/2 1/2 1/2
	-2		$\lvert\beta\beta\beta\beta\rangle$
	$\beta\beta\beta\beta$	$\begin{smallmatrix}1&2&3&4\end{smallmatrix}$	1
1	1		$\lvert\alpha\alpha\alpha\beta\rangle$ $\lvert\alpha\alpha\beta\alpha\rangle$ $\lvert\alpha\beta\alpha\alpha\rangle$ $\lvert\beta\alpha\alpha\alpha\rangle$
	$\begin{smallmatrix}\alpha&\alpha&\alpha\\\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2&3\\4\end{smallmatrix}$	$1/2\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$
		$\begin{smallmatrix}1&2&4\\3\end{smallmatrix}$	0 $1/3\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$
		$\begin{smallmatrix}1&3&4\\2\end{smallmatrix}$	0 0 $1/2\sqrt{2}$ $-1/2\sqrt{2}$
	0		$\lvert\alpha\alpha\beta\beta\rangle$ $\lvert\alpha\beta\alpha\beta\rangle$ $\lvert\alpha\beta\beta\alpha\rangle$ $\lvert\beta\alpha\alpha\beta\rangle$ $\lvert\beta\alpha\beta\alpha\rangle$ $\lvert\beta\beta\alpha\alpha\rangle$
	$\begin{smallmatrix}\alpha&\alpha&\beta\\\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2&3\\4\end{smallmatrix}$	$1/6\sqrt{6}$ $1/6\sqrt{6}$ $-1/6\sqrt{6}$ $1/6\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$
		$\begin{smallmatrix}1&2&4\\3\end{smallmatrix}$	$1/3\sqrt{3}$ $-1/6\sqrt{3}$ $1/6\sqrt{3}$ $-1/6\sqrt{3}$ $1/6\sqrt{3}$ $-1/3\sqrt{3}$
		$\begin{smallmatrix}1&3&4\\2\end{smallmatrix}$	0 1/2 1/2 −1/2 −1/2 0
	−1		$\lvert\alpha\beta\beta\beta\rangle$ $\lvert\beta\alpha\beta\beta\rangle$ $\lvert\beta\beta\alpha\beta\rangle$ $\lvert\beta\beta\beta\alpha\rangle$
	$\begin{smallmatrix}\alpha&\beta&\beta\\\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2&3\\4\end{smallmatrix}$	$1/6\sqrt{3}$ $1/6\sqrt{3}$ $1/6\sqrt{3}$ $-1/2\sqrt{3}$
		$\begin{smallmatrix}1&2&4\\3\end{smallmatrix}$	$1/6\sqrt{6}$ $1/6\sqrt{6}$ $-1/3\sqrt{6}$ 0
		$\begin{smallmatrix}1&3&4\\2\end{smallmatrix}$	$1/2\sqrt{2}$ $-1/2\sqrt{2}$ 0 0
0	0		$\lvert\alpha\alpha\beta\beta\rangle$ $\lvert\alpha\beta\alpha\beta\rangle$ $\lvert\alpha\beta\beta\alpha\rangle$ $\lvert\beta\alpha\alpha\beta\rangle$ $\lvert\beta\alpha\beta\alpha\rangle$ $\lvert\beta\beta\alpha\alpha\rangle$
	$\begin{smallmatrix}\alpha&\alpha\\\beta&\beta\end{smallmatrix}$	$\begin{smallmatrix}1&2\\3&4\end{smallmatrix}$	$1/3\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $1/3\sqrt{3}$
		$\begin{smallmatrix}1&3\\2&4\end{smallmatrix}$	0 1/2 −1/2 −1/2 1/2 0

Table D.3: Schur bases $\{|W_\kappa^{(\nu)}Y_m^{(\nu)}\rangle\}$ of $\mathcal{H}_2^{\otimes 2}$ (*upper left corner*), $\mathcal{H}_2^{\otimes 3}$ (*upper right corner*), and $\mathcal{H}_2^{\otimes 4}$ (*bottom*).

j/ν	$k/W_\kappa^{(\nu)}$	$Y_m^{(\nu)}$	
5/2	$-5/2$		$\lvert\alpha\alpha\alpha\alpha\alpha\rangle$
			1
	$-3/2$		$\lvert\alpha\alpha\alpha\alpha\beta\rangle \lvert\alpha\alpha\alpha\beta\alpha\rangle \lvert\alpha\alpha\beta\alpha\alpha\rangle \lvert\alpha\beta\alpha\alpha\alpha\rangle \lvert\beta\alpha\alpha\alpha\alpha\rangle$
			$1/5\sqrt{5}$ $1/5\sqrt{5}$ $1/5\sqrt{5}$ $1/5\sqrt{5}$ $1/5\sqrt{5}$
	$-1/2$		$\lvert\alpha\alpha\alpha\beta\beta\rangle \lvert\alpha\alpha\beta\alpha\beta\rangle \lvert\alpha\alpha\beta\beta\alpha\rangle \lvert\alpha\beta\alpha\alpha\beta\rangle \lvert\alpha\beta\alpha\beta\alpha\rangle \lvert\alpha\beta\beta\alpha\alpha\rangle \lvert\beta\alpha\alpha\alpha\beta\rangle \lvert\beta\alpha\alpha\beta\alpha\rangle \lvert\beta\alpha\beta\alpha\alpha\rangle \lvert\beta\beta\alpha\alpha\alpha\rangle$
			$1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$
	$1/2$		$\lvert\alpha\alpha\beta\beta\beta\rangle \lvert\alpha\beta\alpha\beta\beta\rangle \lvert\alpha\beta\beta\alpha\beta\rangle \lvert\alpha\beta\beta\beta\alpha\rangle \lvert\beta\alpha\alpha\beta\beta\rangle \lvert\beta\alpha\beta\alpha\beta\rangle \lvert\beta\alpha\beta\beta\alpha\rangle \lvert\beta\beta\alpha\alpha\beta\rangle \lvert\beta\beta\alpha\beta\alpha\rangle \lvert\beta\beta\beta\alpha\alpha\rangle$
			$1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$ $1/10\sqrt{10}$
	$3/2$		$\lvert\alpha\beta\beta\beta\beta\rangle \lvert\beta\alpha\beta\beta\beta\rangle \lvert\beta\beta\alpha\beta\beta\rangle \lvert\beta\beta\beta\alpha\beta\rangle \lvert\beta\beta\beta\beta\alpha\rangle$
			$1/5\sqrt{5}$ $1/5\sqrt{5}$ $1/5\sqrt{5}$ $1/5\sqrt{5}$ $1/5\sqrt{5}$
	$5/2$		$\lvert\beta\beta\beta\beta\beta\rangle$
			1
3/2	$-3/2$		$\lvert\alpha\alpha\alpha\alpha\beta\rangle \lvert\alpha\alpha\alpha\beta\alpha\rangle \lvert\alpha\alpha\beta\alpha\alpha\rangle \lvert\alpha\beta\alpha\alpha\alpha\rangle \lvert\beta\alpha\alpha\alpha\alpha\rangle$
			$2/5\sqrt{5}$ $-1/10\sqrt{5}$ $-1/10\sqrt{5}$ $-1/10\sqrt{5}$ $-1/10\sqrt{5}$
			0 $1/2\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$
			0 0 $1/3\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$
			0 0 0 $1/2\sqrt{2}$ $-1/2\sqrt{2}$
	$-1/2$		$\lvert\alpha\alpha\alpha\beta\beta\rangle \lvert\alpha\alpha\beta\alpha\beta\rangle \lvert\alpha\alpha\beta\beta\alpha\rangle \lvert\alpha\beta\alpha\alpha\beta\rangle \lvert\alpha\beta\alpha\beta\alpha\rangle \lvert\alpha\beta\beta\alpha\alpha\rangle \lvert\beta\alpha\alpha\alpha\beta\rangle \lvert\beta\alpha\alpha\beta\alpha\rangle \lvert\beta\alpha\beta\alpha\alpha\rangle \lvert\beta\beta\alpha\alpha\alpha\rangle$
			$1/10\sqrt{15}$ $1/10\sqrt{15}$ $-1/15\sqrt{15}$ $1/10\sqrt{15}$ $-1/15\sqrt{15}$ $-1/15\sqrt{15}$ $1/10\sqrt{15}$ $-1/15\sqrt{15}$ $-1/15\sqrt{15}$ $-1/15\sqrt{15}$
			$1/2$ $-1/6$ $1/3$ $-1/6$ $1/3$ $-1/3$ $-1/6$ $1/3$ $-1/3$ $-1/3$
			0 $1/3\sqrt{2}$ $1/3\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $1/6\sqrt{2}$ $-1/3\sqrt{2}$
			0 0 0 $1/6\sqrt{6}$ $1/6\sqrt{6}$ $1/6\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$ 0
	$1/2$		$\lvert\alpha\alpha\beta\beta\beta\rangle \lvert\alpha\beta\alpha\beta\beta\rangle \lvert\alpha\beta\beta\alpha\beta\rangle \lvert\alpha\beta\beta\beta\alpha\rangle \lvert\beta\alpha\alpha\beta\beta\rangle \lvert\beta\alpha\beta\alpha\beta\rangle \lvert\beta\alpha\beta\beta\alpha\rangle \lvert\beta\beta\alpha\alpha\beta\rangle \lvert\beta\beta\alpha\beta\alpha\rangle \lvert\beta\beta\beta\alpha\alpha\rangle$
			$1/15\sqrt{15}$ $1/15\sqrt{15}$ $1/15\sqrt{15}$ $-1/10\sqrt{15}$ $1/15\sqrt{15}$ $1/15\sqrt{15}$ $-1/10\sqrt{15}$ $1/15\sqrt{15}$ $-1/10\sqrt{15}$ $-1/10\sqrt{15}$
			$1/3$ $1/3$ $-1/3$ $1/6$ $1/3$ $-1/3$ $1/6$ $-1/3$ $1/6$ $-1/2$
			$1/3\sqrt{2}$ $-1/6\sqrt{2}$ $1/6\sqrt{2}$ $1/6\sqrt{2}$ $-1/6\sqrt{2}$ $1/6\sqrt{2}$ $1/6\sqrt{2}$ $-1/3\sqrt{2}$ $-1/3\sqrt{2}$ 0
			0 $1/6\sqrt{6}$ $1/6\sqrt{6}$ $1/6\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$ $-1/6\sqrt{6}$ 0 0 0
	$3/2$		$\lvert\alpha\beta\beta\beta\beta\rangle \lvert\beta\alpha\beta\beta\beta\rangle \lvert\beta\beta\alpha\beta\beta\rangle \lvert\beta\beta\beta\alpha\beta\rangle \lvert\beta\beta\beta\beta\alpha\rangle$
			$1/10\sqrt{5}$ $1/10\sqrt{5}$ $1/10\sqrt{5}$ $1/10\sqrt{5}$ $-2/5\sqrt{5}$
			$1/6\sqrt{3}$ $1/6\sqrt{3}$ $1/6\sqrt{3}$ $-1/2\sqrt{3}$ 0
			$1/6\sqrt{6}$ $1/6\sqrt{6}$ $-1/3\sqrt{6}$ 0 0
			$1/2\sqrt{2}$ $-1/2\sqrt{2}$ 0 0 0
1/2	$-1/2$		$\lvert\alpha\alpha\alpha\beta\beta\rangle \lvert\alpha\alpha\beta\alpha\beta\rangle \lvert\alpha\alpha\beta\beta\alpha\rangle \lvert\alpha\beta\alpha\alpha\beta\rangle \lvert\alpha\beta\alpha\beta\alpha\rangle \lvert\alpha\beta\beta\alpha\alpha\rangle \lvert\beta\alpha\alpha\alpha\beta\rangle \lvert\beta\alpha\alpha\beta\alpha\rangle \lvert\beta\alpha\beta\alpha\alpha\rangle \lvert\beta\beta\alpha\alpha\alpha\rangle$
			$1/2\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $1/6\sqrt{2}$ $1/6\sqrt{2}$
			0 $2/3$ $-1/3$ $-1/3$ $1/6$ $-1/6$ $-1/3$ $1/6$ $-1/6$ $1/3$
			0 0 0 $1/3\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $-1/3\sqrt{3}$ $1/6\sqrt{3}$ $1/6\sqrt{3}$ 0
			0 0 $1/3\sqrt{3}$ 0 $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ 0 $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $1/3\sqrt{3}$
			0 0 0 0 $1/2$ $-1/2$ 0 $-1/2$ $1/2$ 0
	$1/2$		$\lvert\alpha\alpha\beta\beta\beta\rangle \lvert\alpha\beta\alpha\beta\beta\rangle \lvert\alpha\beta\beta\alpha\beta\rangle \lvert\alpha\beta\beta\beta\alpha\rangle \lvert\beta\alpha\alpha\beta\beta\rangle \lvert\beta\alpha\beta\alpha\beta\rangle \lvert\beta\alpha\beta\beta\alpha\rangle \lvert\beta\beta\alpha\alpha\beta\rangle \lvert\beta\beta\alpha\beta\alpha\rangle \lvert\beta\beta\beta\alpha\alpha\rangle$
			$1/6\sqrt{2}$ $1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $-1/6\sqrt{2}$ $1/2\sqrt{2}$
			$1/3$ $-1/6$ $1/6$ $-1/3$ $-1/6$ $1/6$ $-1/3$ $-1/3$ $2/3$ 0
			0 $1/6\sqrt{3}$ $1/6\sqrt{3}$ $-1/3\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ $1/3\sqrt{3}$ 0 0 0
			$1/3\sqrt{3}$ $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ 0 $-1/6\sqrt{3}$ $-1/6\sqrt{3}$ 0 $1/3\sqrt{3}$ 0 0
			0 $1/2$ $-1/2$ 0 $-1/2$ $1/2$ 0 0 0 0

Table D.4: Schur basis $\{|W_\kappa^{(\nu)} Y_m^{(\nu)}\rangle\}$ of $\mathcal{H}_2^{\otimes 5}$.

such states. Bob can identify these states by a measuring the Young diagram and Young tableau since

$$\mathcal{M}_n\big(|\Gamma^{(j)}_{m_i}\rangle\langle\Gamma^{(j)}_{m_i}|\big) = \frac{1}{2j+1}\mathcal{I}\otimes|Y^{(j)}_{m_i}\rangle\langle Y^{(j)}_{m_i}|. \tag{D.61}$$

Asymptotically, the rate at which Alice is able to send classical information to Bob tends to one,

$$\lim_{n\to\infty}\frac{\log_2(c_n)}{n} \approx 1 - \frac{1}{2n}\log_2(n). \tag{D.62}$$

As an example, consider the Schur basis of five qubits in table D.4. Here, $c_5 = 1+4+5 = 10$ and Alice is able to send classical information at a rate ≈ 0.66.

To transmit quantum information to Bob, Alice encodes the information into the subsystem spanned by the $\{|Y^{(j)}_{m_i}\rangle\}_{i=1\ldots h_j(\mathsf{S}_n)}$ with the largest dimension $h_j(\mathsf{S}_n)$, i. e. she prepares a state

$$\sigma\otimes\rho = \sum_{k,k'=-j}^{+j}\sigma_{kk'}|W^{(j)}_k\rangle\langle W^{(j)}_{k'}|\otimes\sum_{i,i'=1}^{h_j(\mathsf{S}_n)}\rho_{ii'}|Y^{(j)}_{m_i}\rangle\langle Y^{(j)}_{m_{i'}}| \tag{D.63}$$

with arbitrary $\sigma_{kk'}$. Bob receives the state

$$\mathcal{M}_n(\sigma\otimes\rho) = \frac{1}{2j+1}\sum_{k=-j}^{+j}|W^{(j)}_k\rangle\langle W^{(j)}_k|\otimes\sum_{i,i'=1}^{h_j(\mathsf{S}_n)}\rho_{ii'}|Y^{(j)}_{m_i}\rangle\langle Y^{(j)}_{m_{i'}}| = \frac{1}{2j+1}\mathcal{I}\otimes\rho. \tag{D.64}$$

For large n, $h_j(\mathsf{S}_n)$ becomes maximal for $j_{\max} = \sqrt{n}/2$. Again the rate at which Alice is able to send quantum information to Bob asymptotically tends to one,

$$\lim_{n\to\infty}\frac{\log_2\big(h_{j_{\max}}(\mathsf{S}_n)\big)}{n} \approx 1 - \frac{1}{2n}\log_2(n). \tag{D.65}$$

For our example of five qubits, the largest dimension is $h_{1/2}(\mathsf{S}_5) = 5$ and Alice is able to send qubits at a rate ≈ 0.46.

Bibliography

[Abr61] A. Abragam. *The Principles of Nuclear Magnetism*, volume 32 of *International Series of Monographs on Physics*. Oxford University Press (1961).

[AC97] C. Adami and N. J. Cerf. von Neumann capacity of noisy quantum channels. *Phys. Rev. A*, **56**(5), 3470–3483 (1997). arXiv:quant-ph/9609024v3.

[AK01] A. Ashikhmin and E. Knill. Nonbinary quantum stabilizer codes. *IEEE Trans. Inf. Theory*, **47**(7), 3065–3072 (2001). arXiv:quant-ph/0005008v1.

[BB84] Charles H. Bennett and Gilles Brassard. Quantum Cryptography: Public Key Distribution and Coin Tossing. In *Proceedings of IEEE International Conference on Computers, Systems & Signal Processing, Bangalore, India*, pages 175–179 (1984).

[BBC+93] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, **70**(13), 1895–1899 (1993).

[BBP+96] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels. *Phys. Rev. Lett.*, **76**(5), 722–725 (1996). arXiv:quant-ph/9511027v2.

[BBP+97] —. Erratum: Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels. *Phys. Rev. Lett.*, **78**(10), 2031 (1997).

[BCMS01] Giuliano Benenti, Giulio Casati, Simone Montangero, and Dima L. Shepelyansky. Efficient Quantum Computing of Complex Dynamics. *Phys. Rev. Lett.*, **87**, 227901 (2001). arXiv:quant-ph/0107036v1.

[BCMS03] —. Statistical Properties of Eigenvalues for an Operating Quantum Computer with Static Imperfections. *Eur. Phys. J. D*, **22**, 285–293 (2003). arXiv:quant-ph/0206130v1.

[BDNB04] Michael J. Bremner, Jennifer L. Dodd, Michael A. Nielsen, and Dave Bacon. Fungible dynamics: There are only two types of entangling multiple-qubit interactions. *Phys. Rev. A*, **69**(1), 012313 (2004). arXiv:quant-ph/0307148v1.

[BDS97] Charles H. Bennett, David P. DiVincenzo, and John A. Smolin. Capacities of Quantum Erasure Channels. *Phys. Rev. Lett.*, **78**(17), 3217–3220 (1997). arXiv:quant-ph/9701015v2.

[BDSW96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, **54**(5), 3824–3851 (1996). arXiv:quant-ph/9604024v2.

[BKN00] H. Barnum, E. Knill, and M.A. Nielsen. On quantum fidelities and channel capacities. *IEEE Trans. Inf. Theory*, **46**(4), 1317–1329 (2000). arXiv:quant-ph/9809010v1.

[BNS98] Howard Barnum, M. A. Nielsen, and Benjamin Schumacher. Information transmission through a noisy quantum channel. *Phys. Rev. A*, **57**(6), 4153–4175 (1998). arXiv:quant-ph/9702049v1.

[BRS03] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Classical and Quantum Communication without a Shared Reference Frame. *Phys. Rev. Lett.*, **91**(2), 027901 (2003). arXiv:quant-ph/0302111v3.

[BRS07] —. Reference frames, superselection rules, and quantum information. *Reviews of Modern Physics*, **79**(2), 555 (2007). arXiv:quant-ph/0610030v3.

[Bru98] Dagmar Bruß. Optimal Eavesdropping in Quantum Cryptography with Six States. *Phys. Rev. Lett.*, **81**(14), 3018–2021 (1998). arXiv:quant-ph/9805019v2.

[Bur81] D. P. Burum. Magnus expansion generator. *Phys. Rev. B*, **24**(7), 3684–3692 (1981).

[Che06] Pochung Chen. Geometric continuous dynamical decoupling with bounded controls. *Phys. Rev. A*, **73**(2), 022343 (2006). arXiv:quant-ph/0507265v1.

[CPW02] Jin-Quan Chen, Jialun Ping, and Fan Wang. *Group Representation Theory for Physicists*. World Scientific, Singapore, 2nd edition (2002).

[CRSS97] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum Error Correction and Orthogonal Geometry. *Phys. Rev. Lett.*, **78**(3), 405–408 (1997). arXiv:quant-ph/9605005v3.

[CS96] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Phys. Rev. A*, **54**(2), 1098–1105 (1996). arXiv:quant-ph/9512032v2.

[CTDL77] Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloe. *Quantum Mechanics (Vol. 1 and 2)*. Wiley (1977).

[DEJ+96] David Deutsch, Artur Ekert, Richard Jozsa, Chiara Macchiavello, Sandu Popescu, and Anna Sanpera. Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels. *Phys. Rev. Lett.*, **77**(13), 2818–2821 (1996). arXiv:quant-ph/9604039.

[Dev05] I. Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Trans. Inf. Theory*, **51**(1), 44–55 (2005). arXiv:quant-ph/0304127v6.

[DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Trans. Inf. Theory*, **22**(6), 644–654 (1976).

[Die82] D. Dieks. Communication by EPR devices. *Phys. Lett. A*, **92**(6), 271–272 (1982).

[DJ92] David Deutsch and Richard Jozsa. Rapid Solution of Problems by Quantum Computation Rapid Solution of Problems by Quantum Computation. *Proc. R. Soc. A*, **439**(1907), 553–558 (1992).

[DSS98] David P. DiVincenzo, Peter W. Shor, and John A. Smolin. Quantum-channel capacity of very noisy channels. *Phys. Rev. A*, **57**(2), 830–839 (1998). arXiv:quant-ph/9706061v3.

[DW04] I. Devetak and A. Winter. Relating Quantum Privacy and Quantum Coherence: An Operational Approach. *Phys. Rev. Lett.*, **93**(8), 080501 (2004). arXiv:quant-ph/0307053v1.

[DW05] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proc. R. Soc. A*, **461**(2053), 207–235 (2005). arXiv:quant-ph/0306078v1.

[EBW87] Richard R. Ernst, Geoffrey Bodenhausen, and Alexander Wokaun. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, volume 14 of *International Series of Monographs on Chemistry*. Oxford Science Publications (1987).

[EJ96] Artur Ekert and Richard Jozsa. Quantum computation and Shor's factoring algorithm. *Reviews of Modern Physics*, **68**(3), 733–753 (1996).

[Elg85] Taher Elgamal. A public key cryptosystem and a signature scheme based on discrete logarithms. *IEEE Trans. Inf. Theory*, **31**(4), 469–472 (1985).

[EM96] Artur Ekert and Chiara Macchiavello. Quantum Error Correction for Communication. *Phys. Rev. Lett.*, **77**(12), 2585–2588 (1996). arXiv:quant-ph/9602022v1.

[FFS04] Klaus M. Frahm, Robert Fleckinger, and Dima L. Shepelyansky. Quantum chaos and random matrix theory for fidelity decay in quantum computations with static imperfections. *Eur. Phys. J. D*, **29**, 139–155 (2004). arXiv:quant-ph/0312120v2.

[FGG+97] Christopher A. Fuchs, Nicolas Gisin, Robert B. Griffiths, Chi-Sheng Niu3, and Asher Peres. Optimal eavesdropping in quantum cryptography. I. Information bound and optimal strategy. *Phys. Rev. A*, **56**(2), 1163–1172 (1997). arXiv:quant-ph/9701039v1.

[FM04] Keqin Feng and Zhi Ma. A finite Gilbert-Varshamov bound for pure stabilizer quantum codes. *IEEE Trans. Inf. Theory*, **50**(12), 3323–3325 (2004).

[GKAJ08] D. Geberth, O. Kern, G. Alber, and I. Jex. Stabilization of quantum information by combined dynamical decoupling and detected-jump error correction. *Eur. Phys. J. D*, **46**(2), 381–394 (2008). arXiv:0712.1480v1.

[GL03] D. Gottesman and Hoi-Kwong Lo. Proof of security of quantum key distribution with two-way classical communications. *IEEE Trans. Inf. Theory*, **49**(2), 457–475 (2003). arXiv:quant-ph/0105121v2.

[Got96] Daniel Gottesman. Class of quantum error-correcting codes saturating the quantum Hamming bound. *Phys. Rev. A*, **54**(3), 1862–1868 (1996). arXiv:quant-ph/9604038v2.

[Got97] —. *Stabilizer Codes and Quantum Error Correction*. Ph.D. thesis, California Institute of Technology, Pasadena, California (1997). arXiv:quant-ph/9705052v1.

[Got99] —. Fault-Tolerant Quantum Computation with Higher-Dimensional Systems. In *Proceedings of the 1st NASA International Conference on Quantum Computing and Quantum Communications, QCQC'98 Palm Springs, California*, pages 302–313. Springer (1999). arXiv:quant-ph/9802007v1.

[GP01] Daniel Gottesman and John Preskill. Secure quantum key distribution using squeezed states. *Phys. Rev. A*, **63**(2), 022309 (2001). arXiv:quant-ph/00080462v2.

[Gro97] Lov K. Grover. Quantum Mechanics Helps in Searching for a Needle in a Haystack. *Phys. Rev. Lett.*, **79**, 325–328 (1997). arXiv:quant-ph/9706033v2.

[Ham02a] Mitsuru Hamada. Exponential lower bound on the highest fidelity achievable by quantum error-correcting codes. *Phys. Rev. A*, **65**(5), 052305 (2002). arXiv:quant-ph/0109114v5.

[Ham02b] —. Lower bounds on the quantum capacity and highest error exponent of general memoryless channels. *IEEE Trans. Inf. Theory*, **48**(9), 2547–2557 (2002). arXiv:quant-ph/0112103v3.

[Ham03] —. Notes On The Fidelity Of Symplectic Quantum Error-Correcting Codes. *IJQI*, **1**(4), 443–463 (2003). arXiv:quant-ph/0311003v2.

[Ham04] —. Reliability of Calderbank-Shor-Steane codes and security of quantum key distribution. *J. Phys. A: Math. Gen.*, **37**(34), 8303–8328 (2004). arXiv:quant-ph/0308029v6.

[Ham05] —. Information rates achievable with algebraic codes on quantum discrete memoryless channels. *IEEE Trans. Inf. Theory*, **51**(12), 4263–4277 (2005). arXiv:quant-ph/0207113v3.

[Ham06] —. Conjugate Codes and Applications to Cryptography. *Tamagawa University Research Review*, **12**, 19–25 (2006). arXiv:quant-ph/0610193v1.

[HHH99] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. *Phys. Rev. A*, **60**(3), 1888–1898 (1999). arXiv:quant-ph/9807091v2.

[HHHO05] Karol Horodecki, Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Secure Key from Bound Entanglement. *Phys. Rev. Lett.*, **94**(16), 160502 (2005). arXiv:quant-ph/0309110v2.

[HK02] Te Sun Han and Kingo Kobayashi. *Mathematics of Information and Coding*, volume 203 of *Translations of Mathematical Monographs*. American Mathematical Society (2002).

[HNO03] Henry L. Haselgrove, Michael A. Nielsen, and Tobias J. Osborne. Quantum States far from the Energy Eigenstates of Any Local Hamiltonian. *Phys. Rev. Lett.*, **91**(21), 210401 (2003). arXiv:quant-ph/0303022v1.

[HSS99] A. Hedayat, N. J. A. Sloane, and John Stufken. *Orthogonal Arrays: Theory and Applications*. Springer Series in Statistics. Springer (1999).

[JK99] J. A. Jones and E. Knill. Efficient Refocusing of One-Spin and Two-Spin Interactions for NMR Quantum Computation. *J. Magn. Reson.*, **141**(2), 322–325 (1999). arXiv:quant-ph/9905008v1.

[KA05] O. Kern and G. Alber. Controlling Quantum Systems by Embedded Dynamical Decoupling Schemes. *Phys. Rev. Lett.*, **95**(25), 250501 (2005). arXiv:quant-ph/0506038v1.

[KA06] —. Stabilizing selective recoupling schemes by randomization. *Phys. Rev. A*, **73**(6), 062302 (2006). arXiv:quant-ph/0602167v1.

[KAS05] O. Kern, G. Alber, and D. L. Shepelyansky. Quantum error correction of coherent errors by randomization. *Eur. Phys. J. D*, **32**(1), 153–156 (2005). arXiv:quant-ph/0407262v1.

[Ker04] Oliver Kern. *Quantenalgorithmen und Quantenabbildungen — Implementation und Fehlerkorrektur*. Diploma thesis, TU-Darmstadt (2004).

[KGR05] B. Kraus, N. Gisin, and R. Renner. Lower and Upper Bounds on the Secret-Key Rate for Quantum Key Distribution Protocols Using One-Way Classical Communication. *Phys. Rev. Lett.*, **95**(8), 080501 (2005). arXiv:quant-ph/0410215v2.

[KKKS06] A. Ketkar, A. Klappenecker, S. Kumar, and P.K. Sarvepalli. Nonbinary Stabilizer Codes Over Finite Fields. *IEEE Trans. Inf. Theory*, **52**(11), 4892–4914 (2006). arXiv:quant-ph/0508070v2.

[KL97] Emanuel Knill and Raymond Laflamme. A Theory of Quantum Error-Correcting Codes. *Phys. Rev. A*, **55**, 900–911 (1997). arXiv:quant-ph/9604034v1.

[KL05] K. Khodjasteh and D. A. Lidar. Fault-Tolerant Quantum Dynamical Decoupling. *Phys. Rev. Lett.*, **95**(18), 180501 (2005). arXiv:quant-ph/0408128v3.

[KL07] Kaveh Khodjasteh and Daniel A. Lidar. Performance of deterministic dynamical decoupling schemes: Concatenated and periodic pulse sequences. *Phys. Rev. A*, **75**(6), 062310 (2007). arXiv:quant-ph/0607086v2.

[KL08] K. Khodjasteh and D.A. Lidar. Rigorous Bounds on the Performance of a Hybrid Dynamical Decoupling-Quantum Computing Scheme. *Phys. Rev. A*, **78**(1), 012355 (2008). arXiv:0803.4320v1.

[Kni96] E. Knill. Non-binary Unitary Error Bases and Quantum Codes. *Technical Report LAUR-96-2717, Los Alamos National Laboratory* (1996). arXiv:quant-ph/9608048v2.

[KR08] O. Kern and J. M. Renes. Improved one-way rates for BB84 and 6-state protocols. *Quant. Inf. & Comp.*, **8**(8/9), 0756–0772 (2008). arXiv:0712.1494v2.

[KV09] Kaveh Khodjasteh and Lorenza Viola. Dynamically Error-Corrected Gates for Universal Quantum Computation. *Phys. Rev. Lett.*, **102**(8), 080501 (2009). arXiv:0810.0698v2.

[KW04] Dennis Kretschmann and Reinhard F Werner. Tema con variazioni: quantum channel capacity. *New J. Phys.*, **6**, 26 (2004). arXiv:quant-ph/0311037v1.

[LC99] Hoi-Kwong Lo and H. F. Chau. Unconditional Security Of Quantum Key Distribution Over Arbitrarily Long Distances. *Science*, **283**, 2050–2056 (1999). arXiv:quant-ph/9803006v5.

[LCW98] D. A. Lidar, I. L. Chuang, and K. B. Whaley. Decoherence-Free Subspaces for Quantum Computation. *Phys. Rev. Lett.*, **81**(12), 2594–2597 (1998). arXiv:quant-ph/9807004v2.

[LCYY00] Debbie W. Leung, Isaac L. Chuang, Fumiko Yamaguchi, and Yoshihisa Yamamoto. Efficient implementation of coupled logic gates for quantum computation. *Phys. Rev. A*, **61**(4), 042310 (2000). arXiv:quant-ph/9904100v1.

[Leu02] D. Leung. Simulation and reversal of n-qubit Hamiltonians using Hadamard matrices. *J. Mod. Opt.*, **49**(8), 1199–1217 (2002). arXiv:quant-ph/0107041v2.

[LGYY02] T. D. Ladd, J. R. Goldman, F. Yamaguchi, and Y. Yamamoto. All-Silicon Quantum Computer. *Phys. Rev. Lett.*, **89**(1), 017901 (2002). arXiv:quant-ph/0109039v1.

[Llo97] Seth Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, **55**(3), 1613–1622 (1997). arXiv:quant-ph/9604015v2.

[Lo01] Hoi-Kwong Lo. Proof of unconditional security of six-state quantum key distribution scheme. *Quant. Inf. & Comp.*, **1**(2), 81–94 (2001). arXiv:quant-ph/0102138v5.

[LS05] Jae Weon Lee and Dima L. Shepelyansky. Quantum chaos algorithms and dissipative decoherence with quantum trajectories. **71**(5), 056202 (2005). arXiv:quant-ph/0501120v1.

[Mac03] David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, England (2003). Available from `http://www.inference.phy.cam.ac.uk/mackay/itila/`.

[MCL06] Tobias Moroder, Marcos Curty, and Norbert Lütkenhaus. One-way quantum key distribution: Simple upper bound on the secret key rate. *Phys. Rev. A*, **74**(5), 052301 (2006). arXiv:quant-ph/0603270v1.

[MS77] F. J. MacWilliams and N. J. A. Sloane. *The Theory of Error-Correcting Codes*, volume 16 of *North-Holland Mathematical Library*. North-Holland (1977).

[MU02] Ryutaroh Matsumoto and Tomohiko Uyematsu. Lower bound for the quantum capacity of a discrete memoryless quantum channel. *J. Math. Phys.*, **43**(9), 4391 (2002). arXiv:quant-ph/0105151v4.

[NC00] Michael A. Nielsen and Issac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, England (2000).

[NCSB98] M. A. Nielsen, Carlton M. Caves, Benjamin Schumacher, and Howard Barnum. Information-Theoretic Approach to Quantum Error Correction and Reversible Measurement. *Proceedings: Mathematical, Physical and Engineering Sciences*, **454**(1969), 277–304 (1998). arXiv:quant-ph/9706064v1.

[Nie02] Michael A. Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Phys. Lett. A*, **303**(4), 249–252 (2002). arXiv:quant-ph/0205035v2.

[Pre98] John Preskill. *Physics 229: Advanced Mathematical Methods of Physics – Quantum Computation and Information*. California Institute of Technology, Pasadena, CA (1998). Available from `http://www.theory.caltech.edu/people/preskill/ph229/`.

[Pro02] Tomaž Prosen. General relation between quantum ergodicity and fidelity of quantum dynamics. *Phys. Rev. E*, **65**, 036208 (2002). arXiv:quant-ph/0106149v2.

[PRS04] Juan Pablo Paz, Augusto José Roncaglia, and Marcos Saraceno. Quantum algorithms for phase-space tomography. *Phys. Rev. A*, **69**, 032312 (2004). arXiv:quant-ph/0310126v1.

[PŽ01] Tomaž Prosen and Marko Žnidarič. Can quantum chaos enhance the stability of quantum computation ? *J. Phys. A: Math. Gen.*, **34**, L681–L687 (2001). arXiv:quant-ph/0106150v1.

[PZ03] J. Proos and Ch. Zalka. Shor's discrete logarithm quantum algorithm for elliptic curves. *Quant. Inf. & Comp.*, **3**(4), 317–344 (2003). arXiv:quant-ph/0301141v2.

[QKD] `http://www.heise.de/newsticker/meldung/25639`: Quanten-Kryptographie aus der Schweiz, (2002). `http://www.heise.de/newsticker/meldung/41778`: Quanten-Kryptografie made in USA, (2003). `http://www.heise.de/newsticker/meldung/112909`: Quantenkryptografie-Chip von Siemens, (2008).
`http://www.magiqtech.com`, `http://www.idquantique.com/`.

[Rai99] E. M. Rains. Nonbinary quantum codes. *IEEE Trans. Inf. Theory*, **45**(6), 1827–1832 (1999). arXiv:quant-ph/9703048v1.

[Ren05] Renato Renner. *Security of Quantum Key Distribution*. Ph.D. thesis, Swiss Federal Institute of Technology Zurich (2005). arXiv:quant-ph/0512258v2.

[RGK05] Renato Renner, Nicolas Gisin, and Barbara Kraus. Information-theoretic security proof for quantum-key-distribution protocols. *Phys. Rev. A*, **72**(1), 012332 (2005). arXiv:quant-ph/0502064v1.

[Rom92] Steven Roman. *Coding and Information Theory*, volume 134 of *Graduate Texts in Mathematics*. Springer (1992).

[RS07] Joseph M. Renes and Graeme Smith. Noisy Processing and Distillation of Private Quantum States. *Phys. Rev. Lett.*, **98**(2), 020502 (2007). arXiv:quant-ph/0603262v2.

[RSA78] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, **21**(2), 120–126 (1978).

[RW06] M. Rötteler and P. Wocjan. Equivalence of Decoupling Schemes and Orthogonal Arrays. *IEEE Trans. Inf. Theory*, **52**(9), 4171–4181 (2006). arXiv:quant-ph/0409135v1.

[Sch35] Erwin Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. *Die Naturwissenschaften*, **23**(48), 807–812 (1935).

[Sha48] C. E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, **27**, 379–423 & 623–656 (1948).

[Sho94] Peter W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *Proceedings of the 35th Annual Symposium on Foundations of Computer Science, Santa Fe, NM* (1994). arXiv:quant-ph/9508027v2.

[Sho95] —. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, **52**, R2493–R2496 (1995).

[Sho02] P. W. Shor. The quantum channel capacity and coherent information (Lecture Notes, MSRI Workshop on Quantum Computation). eprint (2002). `http://www.msri.org/publications/ln/msri/2002/quantumcrypto/shor/1/`.

[SM01] Marcus Stollsteimer and Günter Mahler. Suppression of arbitrary internal coupling in a quantum register. *Phys. Rev. A*, **64**(5), 052301 (2001). arXiv:quant-ph/0107059v1.

[SN96] Benjamin Schumacher and M. A. Nielsen. Quantum data processing and error correction. *Phys. Rev. A*, **54**(4), 2629–2635 (1996). arXiv:quant-ph/9604022v1.

[SP00] Peter W. Shor and John Preskill. Simple Proof of Security of the BB84 Quantum Key Distribution Protocol. *Phys. Rev. Lett.*, **85**(2), 441–444 (2000).

[SRS08] Graeme Smith, Joseph M. Renes, and John A. Smolin. Structured Codes Improve the Bennett-Brassard-84 Quantum Key Rate. *Phys. Rev. Lett.*, **100**(17), 170502 (2008). arXiv:quant-ph/0607018v2.

[SS96] Peter W. Shor and John A. Smolin. Quantum Error-Correcting Codes Need Not Completely Reveal the Error Syndrome. eprint (1996). arXiv:quant-ph/9604006v2.

[SS07] Graeme Smith and John A. Smolin. Degenerate Quantum Codes for Pauli Channels. *Phys. Rev. Lett.*, **98**(3), 030501 (2007). arXiv:quant-ph/0604107v2.

[Ste96] Andrew Steane. Multiple-Particle Interference and Quantum Error Correction. *Proc. R. Soc. A*, **452**(1954), 2551–2577 (1996). arXiv:quant-ph/9601029v2.

[Suz91] Masuo Suzuki. General theory of fractal path integrals with applications to many-body theories and statistical physics. *J. Math. Phys.*, **32**(2), 400 (1991).

[SV05] Lea F. Santos and Lorenza Viola. Dynamical control of qubit coherence: Random versus deterministic schemes. *Phys. Rev. A*, **72**, 062303 (2005). arXiv:quant-ph/0511121v1.

[SV06] —. Enhanced Convergence and Robust Performance of Randomized Dynamical Decoupling. *Phys. Rev. Lett.*, **97**(15), 150501 (2006). arXiv:quant-ph/0602168v3.

[SV08] —. Advantages of Randomization in Coherent Quantum Dynamical Control. *New J. Phys.*, **1**(1), 1–1 (2008). arXiv:0804.0890v1.

[SVC00] Matthias Steffen, Lieven M. K. Vandersypen, and Isaac L. Chuang. Simultaneous Soft Pulses Applied at Nearby Frequencies. *J. Magn. Reson.*, **146**(2), 369–374 (2000).

[Tun85] Wu-Ki Tung. *Group Theory in Physics*. World Scientific Publishing, Singapore (1985).

[Vio05] L. Viola. Randomized control of open quantum systems. In *CDC-ECC '05. 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference.*, pages 1794–1799 (2005). arXiv:quant-ph/0601106v1.

[VK03] Lorenza Viola and Emanuel Knill. Robust Dynamical Decoupling of Quantum Systems with Bounded Controls. *Phys. Rev. Lett.*, **90**(3), 037901 (2003). arXiv:quant-ph/0208056v1.

[VK05] —. Random Decoupling Schemes for Quantum Dynamical Control and Error Suppression. *Phys. Rev. Lett.*, **94**(6), 060502 (2005). arXiv:quant-ph/0511120v1.

[VKL99] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical Decoupling of Open Quantum Systems. *Phys. Rev. Lett.*, **82**(12), 2417–2421 (1999).

[VKL00] —. Dynamical Generation of Noiseless Quantum Subsystems. *Phys. Rev. Lett.*, **85**(16), 3520–3523 (2000). arXiv:quant-ph/0002072v1.

[VL98] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. *Phys. Rev. A*, **58**(4), 2733–2744 (1998). arXiv:quant-ph/9803057v1.

[VLK99] Lorenza Viola, Seth Lloyd, and Emanuel Knill. Universal Control of Decoupled Quantum Systems. *Phys. Rev. Lett.*, **83**(23), 4888–4891 (1999).

[VS06] L. Viola and L. F. Santos. Randomized dynamical decoupling techniques for coherent quantum control. *J. Mod. Opt.*, **53**(16&17), 2559–2568 (2006). arXiv:quant-ph/0602175v2.

[VSB+01] Lieven M. K. Vandersypen, Matthias Steffen, Gregory Breyta, Costantino S. Yannoni, Mark H. Sherwood, and Isaac L. Chuang. Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. *Nature*, **414**(6866), 883–887 (2001). arXiv:quant-ph/0112176v1.

[Wel88] Dominic Welsh. *Codes and Cryptography*. Oxford University Press (1988).

[WHH68] J. S. Waugh, L. M. Huber, and U. Haeberlen. Approach to High-Resolution nmr in Solids. *Phys. Rev. Lett.*, **20**(5), 180–182 (1968).

[Wie83] Stephen Wiesner. Conjugate coding. *ACM SIGACT News*, **15**(1), 78–88 (1983).

[WMU06] Shun Watanabe, Ryutaroh Matsumoto, and Tomohiko Uyematsu. Improvement of stabilizer-based entanglement distillation protocols by encoding operators. *J. Phys. A: Math. Gen.*, **39**(16), 4273–4290 (2006). arXiv:quant-ph/0506054v2.

[Woc06] Pawel Wocjan. Efficient decoupling schemes with bounded controls based on Eulerian orthogonal arrays. *Phys. Rev. A*, **73**(6), 062317 (2006). arXiv:quant-ph/0410107v1.

[WRJB02a] P. Wocjan, M. Rötteler, D. Janzing, and T. Beth. Universal simulation of Hamiltonians using a finite set of control operations. *Quant. Inf. & Comp.*, **2**(2), 133–150 (2002). arXiv:quant-ph/0109063v1.

[WRJB02b] Pawel Wocjan, Martin Rötteler, Dominik Janzing, and Thomas Beth. Simulating Hamiltonians in quantum networks: Efficient schemes and complexity bounds. *Phys. Rev. A*, **65**(4), 042309 (2002). arXiv:quant-ph/0109088v1.

[WZ82] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, **299**(5886), 802–803 (1982).

[YLM+04] Fumiko Yamaguchi, Thaddeus D. Ladd, Cyrus P. Master, Yoshihisa Yamamoto, and Navin Khaneja. Efficient decoupling and recoupling in solid state NMR for quantum computation. eprint (2004). arXiv:quant-ph/0411099v1.

[Zan99] Paolo Zanardi. Symmetrizing evolutions. *Phys. Lett. A*, **258**(2–3), 77–82 (1999). arXiv:quant-ph/9809064v2.

[Zan00] —. Stabilizing quantum information. *Phys. Rev. A*, **63**(1), 012301 (2000). arXiv:quant-ph/9910016v2.

[ZR97a] P. Zanardi and M. Rasetti. Noiseless Quantum Codes. *Phys. Rev. Lett.*, **79**(17), 3306–3309 (1997). arXiv:quant-ph/9705044v2.

[ZR97b] Paolo Zanardi and Mario Rasetti. Error avoiding quantum codes. *Mod. Phys. Lett. B*, **11**, 1085–1093 (1997). arXiv:quant-ph/9710041v1.

Danksagung

Die vorliegende Arbeit wurde in der Arbeitsgruppe von Herrn Prof. Gernot Alber angefertigt, dem ich an dieser Stelle dafür danken möchte, mir die Gelegenheit gegeben zu haben, in seiner Arbeitsgruppe mitzuarbeiten.

Des Weiteren gilt mein Dank Herrn Prof. Dima L. Shepelyansky für die produktive Zusammenarbeit im Rahmen des EU Projekts EDIQIP, und für die Gelegenheit neben seiner Arbeitsgruppe „Quantware“ in Toulouse auch die folgenden Veranstaltungen zu besuchen: Im Rahmen der International School of Physics „Enrico Fermi“ in Varenna das Programm „Quantum Computers, Algorithms and Chaos“ vom 5. bis 15. Juli 2005, und das Trimester „Quantum information, computation, and complexity“ am Institut Henri Poincaré in Paris vom 4. Januar bis 7. April 2006.

Bedanken möchte ich mich auch bei Herrn Prof. Igor Jex, dessen Arbeitsgruppe in Prag ich mehrfach besuchen konnte, und insbesondere bei seinen Studenten Stanislav Vymětal und Pavel Bažant für interessante Diskussionen.

Herrn Prof. Thomas H. Seligman gilt mein Dank für die Einladung zur Konferenz „Decoherence: Measures, models and semi-classics“ in Cuernavaca, Mexiko, vom 9. bis 22. September 2007.

Mein besonderer Dank gilt natürlich allen Mitgliedern meiner Arbeitsgruppe für die nette Zusammenarbeit und die zahlreichen Diskussionen. Für das Korrekturlesen samt hilfreichen Kommentaren seien (in alphabetischer Reihenfolge) Kedar Ranade, Joseph Renes und Ulrich Seyfarth nochmal gesondert erwähnt. Ebenfalls besonderer Dank gilt Herrn Prof. Jürgen Berges für die Übernahme des Korreferats.

Curriculum Vitae

Personal Data

Oliver Kern

Email: oliver.kern@physik.tu-darmstadt.de

Born: March 6th, 1978 in Mainz (Germany)
German citizen

Education

06/1984–07/1988 Friedrich Fröbel Schule, Primary School, (Grundschule des Kreises Offenbach)

08/1988–07/1994 Hermann Hesse Schule, Secondary School, (Gesamtschule des Kreises Offenbach)

08/1994–06/1997 Claus von Stauffenberg Schule, Secondary School, (Gymnasiale Oberstufenschule des Kreises Offenbach),
Higher Education Entrance Qualification
(Main Subjects: Mathematics and Physics)

Civilian Service

09/1997–09/1998 German Red Cross Blood Donation Service

Higher Education

10/1998–09/2000 Pre-Diploma in Physics, Technical University Darmstadt

10/2000–09/2004 Diploma in Physics, Technical University Darmstadt, Institute of Applied Physics
(Main focus: Quantum Information Theory)

Ph. D. Studies

09/2004 Beginning of Ph. D. Studies at the Technical University Darmstadt under supervision of Prof. Dr. G. Alber.

01/2006–04/2006 Marie Curie fellowship within the program 'quantum information, computation, and complexity' which took place at the Institut Henri Poincaré in Paris.

List of Publications

[1] Oliver Kern. *Quantenalgorithmen und Quantenabbildungen — Implementation und Fehlerkorrektur.* Diploma thesis, TU Darmstadt (2004).

[2] O. Kern, G. Alber, and D. L. Shepelyansky. Quantum error correction of coherent errors by randomization. *Eur. Phys. J. D*, **32**(1), 153–156 (2005). arXiv:quant-ph/0407262v1.

[3] O. Kern, and G. Alber. Suppressing decoherence of quantum algorithms by jump codes. *Eur. Phys. J. D*, **36**(2), 241–248 (2005). arXiv:quant-ph/0506037v1.

[4] O. Kern and G. Alber. Controlling Quantum Systems by Embedded Dynamical Decoupling Schemes. *Phys. Rev. Lett.*, **95**(25), 250501 (2005). arXiv:quant-ph/0506038v1.

[5] O. Kern and G. Alber. Stabilizing selective recoupling schemes by randomization. *Phys. Rev. A*, **73**(6), 062302 (2006). arXiv:quant-ph/0602167v1.

[6] D. Geberth, O. Kern, G. Alber, and I. Jex. Stabilization of quantum information by combined dynamical decoupling and detected-jump error correction. *Eur. Phys. J. D*, **46**(2), 381–394 (2008). arXiv:0712.1480v1.

[7] O. Kern and J. M. Renes. Improved one-way rates for BB84 and 6-state protocols. *Quant. Inf. & Comp.*, **8**(8/9), 0756–0772 (2008). arXiv:0712.1494v2.

Erklärung

Hiermit erkläre ich an Eides Statt, daß ich die vorliegende Dissertation selbständig, nur unter Verwendung der angegebenen Quellen und Hilfsmittel verfaßt habe. Ich habe bisher keinen Versuch unternommen, an einer anderen Hochschule das Promotionsverfahren einzuleiten.

Darmstadt, den 29. Januar 2009 Oliver Kern