Disturbed Witnesses in Quantum Complexity Theory

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des akademischen Grades
Doktorin der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation von

M.Sc. Friederike Anna Dziemba

2019

Referent:	Prof. Dr. Tobias J. Osborne	Leibniz Universität Hannover
Korreferenten:	Prof. Dr. Heribert Vollmer	Leibniz Universität Hannover
	Ass.-Prof. Dr. Wolfgang Lechner	Universität Innsbruck
Tag der Promotion:	15. Februar 2019	

Abstract

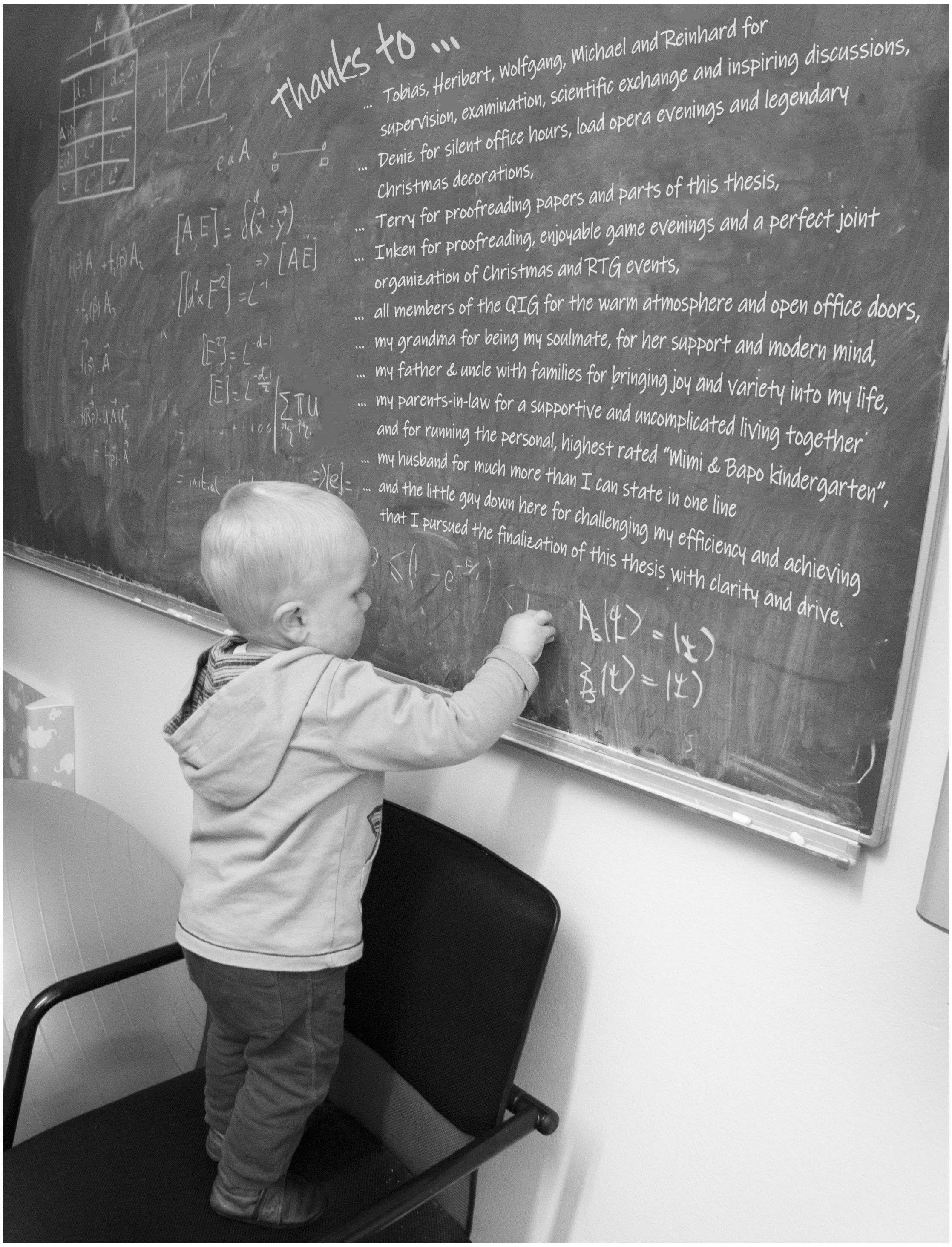
QMA is the complexity class of computational problems that are efficiently verifiable by a quantum algorithm with the help of a witness in contrast to the smaller class BQP of problems efficiently solvable by a quantum algorithm without a witness. Like their classical counterparts NP and P, the class QMA is believed to be strictly larger than BQP, but the definitive answer remains one of the most fundamental open problems of complexity theory. An equality of QMA and BQP would imply that quantum computers can solve many physically and logically relevant problems efficiently, including the Local Hamiltonian problem and the Satisfiability problem for Boolean formulas. New approaches to gain more insight into the structure of BQP and QMA as well as which witness forms are sufficient for QMA are hence worth pursuing.

This thesis comprises three research focuses: Firstly, we extend the uniform diagonalization theorem to complexity classes of promise problems in order to construct strictly intermediate problems between QMA and BQP under the assumption that these classes are unequal. The existence of strictly intermediate problems motivates our definition of noisy QMA classes, which form hierarchies of intermediate classes between QMA and BQP by restricting the witnesses to outputs of certain quantum channels.

In our second research focus we apply the tool of concatenated coding to prove a bound on the witness noise up to which QMA stays robust. Besides a bound for general i.i.d. channels, we can prove that QMA stays robust if each witness qubit is disturbed by 18% depolarizing or 27% dephasing noise, while for complete depolarization or dephasing the noisy class obviously collapses to BQP and QCMA, respectively.

In the third research focus we interpret the famous QPCP conjecture as robustness of the class QMA against high witness disturbance. Moreover, we consider a multiprover protocol by Fitzsimons and Vidick that constitutes a first step towards an important alternative formulation of the QPCP conjecture and achieve a reduction of the number of provers and an improvement of the acceptance probability for this protocol.

Keywords: quantum information, quantum complexity theory, QMA, BQP, verifying protocols, quantum Arthur-Merlin games, uniform diagonalization theorem, concatenated coding, quantum PCP conjecture, multiprover games


$A_s|\psi\rangle = |\psi\rangle$
$B_p|\psi\rangle = |\psi\rangle$

Thanks to ...

... Tobias, Heribert, Wolfgang, Michael and Reinhard for supervision, examination, scientific exchange and inspiring discussions,

... Deniz for silent office hours, load opera evenings and legendary Christmas decorations,

... Terry for proofreading papers and parts of this thesis,

... Inken for proofreading, enjoyable game evenings and a perfect joint organization of Christmas and RTG events,

... all members of the QIG for the warm atmosphere and open office doors,

... my grandma for being my soulmate, for her support and modern mind,

... my father & uncle with families for bringing joy and variety into my life,

... my parents-in-law for a supportive and uncomplicated living together and for running the personal, highest rated "Mimi & Bapo kindergarten",

... my husband for much more than I can state in one line

... and the little guy down here for challenging my efficiency and achieving that I pursued the finalization of this thesis with clarity and drive.

Table of contents

1 Introduction **1**

2 Introduction to quantum computing **5**

2.1 Quantum circuits 5

2.2 Quantum measurements 8

2.3 Pauli group 11

2.4 Quantum channels 13

2.5 Norms on operators and superoperators 15

3 Introduction to classical and quantum complexity theory **19**

3.1 The classical computing model 19

3.2 Randomized and quantum computing models 24

3.3 Standard complexity classes 26

3.4 Amplification 30

3.5 Reductions and complete problems 33

3.6 The Satisfiability and the Local Hamiltonian problem 37

4 Introduction to quantum coding **43**

4.1 Quantum coding 43

4.2 Quantum error correction 45

4.3 Quantum error detection 49

4.4 Error sets, code distance and performance bounds 52

4.5 Stabilizer codes 55

4.5.1 Definition and stabilizer groups 55

4.5.2 Check matrix representation and syndrome spaces 56

4.5.3 Normalizer 59

4.5.4 Error correction in the stabilizer formalism 61
4.5.5 Examples of common stabilizer codes 63
4.5.6 Encoding and decoding of stabilizer codes 64

5 Existence of intermediate problems 67
5.1 Introduction 67
5.2 A framework for promise problems 70
5.2.1 Extremal problems and closure properties of complexity classes 70
5.2.2 Total decidability 72
5.2.3 Recursive (re)presentation 77
5.3 An extended uniform diagonalization theorem 80
5.4 Implications 85

6 QMA with noisy witness 89
6.1 Definition 89
6.2 Applicability of the UDT and intermediateness of noisy QMA classes . 93
6.3 Interpolation of standard complexity classes by physical witness noise . 95

7 Robustness of QMA against constant witness noise 101
7.1 Introduction 101
7.2 Concatenation of general codes 104
7.3 Concatenation of stabilizer codes 107
7.4 Channel properties in Stokes representation 110
7.5 Convergence of diagonal noise 113
7.6 Convergence of general noise under stabilizer coding 115
7.7 Robustness results 119

8 QPCP and a 3-prover protocol for the Local Hamiltonian problem 127
8.1 Introduction 127
8.2 The classical PCP theorem 129
8.3 The quantum PCP conjecture 133
8.4 Noisy QMA version of the quantum PCP conjecture 136
8.5 A multiprover version of the quantum PCP conjecture 140
8.6 Improving the multiprover protocol 145
8.7 Adaption of the soundness proof 150

8.7.1 Sketch of the extraction protocol for a low energy state σ 150
8.7.2 Closeness of the extraction operators $D^t_{i,S}$ and $D^t_{i,\{i\}}$ 153
8.7.3 Energy of the extraction state σ 158

9 Conclusion 163

Notation

$[n]$	set of all natural numbers from 1 to n
Σ^*	set of all strings over $\Sigma = \{0, 1\}$
$\mathcal{L}(\mathbb{C}^d)$	set of all linear operators on the Hilbert space $\mathbb{C}^d$
$\mathcal{D}(\mathbb{C}^d)$	set of all density operators on the Hilbert space $\mathbb{C}^d$
$\mathcal{U}(\mathbb{C}^d)$	set of all unitary operators on the Hilbert space $\mathbb{C}^d$
Id	identity channel of arbitrary dimension
$\\|A\\|_\diamond$	diamond norm of superoperator A
$\\|A\\|_1$	trace norm of operator or superoperator A
$\mathrm{tr}_R(\rho)$	partial trace of ρ discarding register R
$(A)_R$	operator or superoperator A acting on register R
Π_{acc}	projection of the first qubit of a quantum circuit onto $\|1\rangle \langle 1\|$
$\mathbb{I}$, X, Y, Z	single qubit Pauli operators
$\mathcal{P}_N$	Pauli group on N qubits
$\mathcal{P}^{\otimes N}$	Pauli operator on N qubits (without prefactor in contrast to $\mathcal{P}_N$)
$(N, K, \delta)_d$	code of distance δ encoding K orthonormal states into N qudits
$[N, k, \delta]$	stabilizer code of distance δ encoding k qubits into N qubits
P_C	projection onto the code space of the code C
V_C	code space of the code C
V_S	code space of the stabilizer code with stabilizer group S
$N(S)$	normalizer of the stabilizer group S
$A \sim_{S^\pm} B$	$A = \sigma B$ for a $\sigma \in \{ps \mid p \in \{\pm 1, \pm i\}, s \in S\}$ and a stabilizer group S
$\|\bar{0}\rangle$, $\bar{\sigma}$	logical $\|0\rangle$ state and logical Pauli σ operator
$\mathbb{P}[A]$	probability of event A
$\leqslant_m^P$	Karp reduction
$\leqslant_T^P$	Cook reduction
C-c_m	Karp-complete problems of the complexity class C
C-c_T	Cook-complete problems of the complexity class C
$A \oplus B$	marked union of the problems A and B
$\mathcal{O}(f(n))$	set of all real functions g with $\lim_{n\to\infty} \frac{g(n)}{f(n)} \leqslant C$ for a constant C

List of Tables

Table **Page**

3.1 Karp-complete problems for standard complexity classes. 36

4.1 Common stabilizer codes. 63

8.1 Multiprover protocol for the k-Local Hamiltonian problem based on the protocol by Fitzsimons and Vidick. 143

8.2 Conditions on the $(N, 2)$ code required by the multiprover protocol. . . 146

List of Figures

Figure **Page**

2.1 A quantum circuit. 7

3.1 Action of a Turing machine with transition rule $\delta(s_0, 1) = (s_1, 0, R)$. . . . 21
3.2 Hierarchy of complexity classes (classes contain those connected below). 28

4.1 Concept of coding according to Knill and Laflamme. 44
4.2 Vector representation of a Pauli group element p. 56

5.1 Subset relations between the difference notions for promise problems. . 72
5.2 Mixture of the problems A and A' covering all necessary elements to place the resulting problem B outside C and C'. 83

6.1 Attacking complexity class separations by disturbed witnesses. 96

7.1 3-fold concatenation of a $(2, 2)$ code against the i.i.d. noise channel T. . 103
7.2 Bloch sphere representation of 18% depolarizing noise (left) and 27% dephasing noise (right), which does not diminish the power of QMA when applied to each witness qubit. 124

8.1 Questions posed to the (randomly labelled) provers by protocol 8.1. . . 145
8.2 Registers of the extraction algorithm for each prover $t \in [r]$. 151
8.3 The operation $D^t_{i,S}$. 153

Chapter 1

Introduction

Complexity theory is a well established field of classical computer science. It categorizes computational problems into so-called complexity classes depending on their "difficulty", which is reflected by the ressources that a classical computer needs to find a solution. Quantum complexity theory is about a similar categorization according to the quantum computing model. In both classical and quantum complexity theory, the two most important complexity classes are the class of efficiently solvable problems (called P in the classical and BQP in the quantum case) and the broader class of efficiently verifiable problems (called NP in the classical and QMA in the quantum case).

For an efficiently solvable problem there exists an algorithm whose runtime is upper bounded by a polynomial in the input length. A polynomial scaling has the advantage that the runtime only increases modestly with the input length, which is a relevant criteria for feasibility independent of the processor speed. In contrary, a non-polynomial runtime quickly reaches practically infeasible absolute values for only slightly longer inputs. Luckily, many standard problems of linear algebra, list manipulation or data processing can be solved in polynomial time on a classical computer.

Still, there exist important problems that are only known to be efficiently verifiable, i.e. efficiently solvable given extra information, the so-called witness. The witness has a specific property that can be verified efficiently if and only if the answer to the problem instance is positive. An exemplary problem in the class NP which is not known to lie in P is the question of whether a graph possesses a path covering each vertex exactly once. In this case the witness is supposed to describe such a path if existent, since the property of covering each vertex exactly once can obviously be checked efficiently. Famous problems in QMA that are not known to lie in BQP are the estimation of the ground state energy of local Hamiltonians, non-equality checks of unitary circuits and checking if local density matrices are consistent with a global quantum state.

Despite the relevance of efficiently solvable and efficiently verifiable problems, it surprisingly remains unproven if NP and QMA really consist of strictly more problems than P and BQP, respectively. The question if P = NP is of such fundamental interest that in the year 2000 the Clay Mathematics Institute announced it as one of the seven millenium problems with a prize of 1 million dollars for its solution.

While clearly P $\subseteq$ BQP and NP $\subseteq$ QMA, the relationship between NP and BQP remains unclear as well. Yet, it is a big motivation for the development of quantum computers that at least some nontrivial NP problems have been shown to lie in BQP. The most famous example is the problem of integer factorization for which no efficient classical algorithm is known but which can be solved efficiently on a quantum computer by Shor's algorithm [1].

Studying the complexity classes BQP and QMA is hence of large interest from different perspectives. The prevailing belief among complexity theorists is that these classes like their classical counterparts are unequal (see the survey in [2]). Assuming this, one might wonder if the problems in QMA can only have one of two extreme complexities (either in BQP or QMA-complete, i.e. belonging to the most difficult problems in QMA) or if QMA also contains problems of intermediate complexity. The latter possibility is suggested by several QMA problems of high physical relevance that are not known to be QMA-complete. Some problems essentially define their own complexity classes like QCMA, StoqMA or TIM, but their strict intermediateness remains unproven.

After providing introductionary material in chapters 2 – 4, in chapter 5 we prove the existence of strictly QMA-intermediate problems by extending the uniform diagonalization theorem, which originally shows intermediate problems for NP and other classical complexity classes. Since the constructed intermediate problems are rather abstract and lack physical relevance, in chapter 6 we pursue a more physical approach and define hierarchies of intermediate classes by restricting the form of the witness for a verifying protocol. This is an obvious idea since within this framework the standard classes QMA, QCMA and BQP can be expressed as special cases with a witness of trivial, classical and maximal restriction.

A reasonable choice for expressing restriction is the disturbance by a physical quantum channel. This mathematical concept is easy to work with and gives the "noisy QMA" classes a physical meaning. By considering parameter-dependent channels such as the partly depolarizing or partly dephasing channel, which interpolate between QMA – BQP and QMA – QCMA, respectively, we hope to gain some new insight into these standard complexity classes.

In capter 7 we will study under which small disturbances QMA is invariant. For this we exploit the tool of concatenated coding from the theory of fault tolerance to make the information carried by the witness robust against the channel noise. We will derive a bound for general physical channels as well as for the specific cases of the partly depolarizing and dephasing channel.

Instead of formulating collapse criteria for the high error side, in chapter 8 we will study the interpretation of the famous QPCP conjecture within our noisy QMA framework. This widely believed but open conjecture states that QMA stays invariant even if the verifying protocol only has access to constantly many witness qubits, which can be interpreted as a high disturbance. After the quantum analogue of the $\mathrm{P} \overset{?}{=} \mathrm{NP}$ question, the QPCP conjecture is the second most important open problem of quantum complexity theory and would prove that QMA tolerates a highly disturbed witness.

Progress on the QPCP conjecture involves finding alternative formulations similar to those existing for the proven classical PCP theorem. One step towards a relevant equivalent formulation is accomplished by a multiprover protocol of Fitzsimons and Vidick [3] deciding the Local Hamiltonian problem. The work horse for this protocol are quantum codes. After a discussion of the right formulation of the QPCP conjecture, we will use coding tools to reduce the number of provers and improve the acceptance probability gap of this protocol.

Chapter 2

Introduction to quantum computing

2.1 Quantum circuits

The concept of quantum information outlined in this chapter arises from the axioms of quantum mechanics and can be looked up in any standard textbook, e.g. [4].

Quantum information is stored in quantum states, which are normalized elements of a Hilbert space. A Hilbert space is a real or complex complete vector space whose norm is induced by a scalar product. Quantum computation is usually realized in finite-dimensional complex Hilbert spaces, which are isomorphic to $\mathbb{C}^d$ with the Euclidean scalar product. Hence, this thesis will restrict to the Hilbert spaces $\mathbb{C}^d$. We denote an orthonormal basis of this space by $|0\rangle$, $|1\rangle$, $\ldots|d-1\rangle$ and associate

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots \quad |d-1\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

$|0\rangle, |1\rangle, \dots, |d-1\rangle$ is also called the *computational basis* and represents classical data, since only orthonormal states can be distinguished perfectly by a quantum measurement. Since it can store d distinct classical states, we call $\mathbb{C}^d$ a *qudit* system. Like in classical computation it is common to restrict to binary systems and hence consider

$$\mathbb{C}^{2^n} = \underbrace{\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \dots \mathbb{C}^2}_{n\text{-times}}$$

as an n-qubit system and denote the computational basis by $\{|x\rangle\}_{x\in\{0,1\}^n}$.

A composite system is mathematically realized via the tensor product. Regarding a certain partition into *registers*, a state can either be a product state or *entangled*, i.e. a *superposition* of different product states.

We will use the terminology "state" also for *density matrices* $\rho \in \mathcal{D}(\mathbb{C}^d) \subset \mathcal{L}(\mathbb{C}^d)$ to express the probabilistic mixture $\rho = \sum_{i\in[t]} p_i |\phi_i\rangle \langle\phi_i|$ of the *pure* states $\{|\phi_i\rangle\}_{i\in[t]}$. A linear operator $\rho \in \mathcal{L}(\mathbb{C}^d)$ is a valid density matrix iff it is positive semi-definite and of trace 1. It is pure iff $\mathrm{tr}(\rho^2) = 1$.

The most general transformation of quantum states is described by quantum channels, which will be discussed in section 2.4. Since the action of a quantum channel is equivalent to a unitary transformation on an extended Hilbert space, a standard quantum computing model is provided by unitary quantum circuits in analogue to classical Boolean circuits.

Definition 2.1. *A* quantum circuit *on n qubits is a sequence of unitary operators U_1, U_2, $\dots U_L \in \mathcal{G}$ from a fixed* gate set $\mathcal{G} \subseteq \mathcal{U}(\mathbb{C}^{2^n})$ *with each gate applied to a subset of the n qubits. L is called the* length *of the circuit.*

Like in figure 2.1 a circuit is represented graphically by drawing a horizontal wire for each qubit and indicating the gates by boxes on the affected qubits.

The gates of the set $\mathcal{G}$ should be simple enough to allow an easy physical implementation and hence deserve the interpretation of elementary operations, while on the other hand they should be able to approximate any unitary to allow universal quantum computation. The accuracy ϵ^{-1} of an approximation $\tilde{U}$ for an operator U is defined by $\|U - \tilde{U}\| = \epsilon$ with

$$\|A\| := \sup_{\substack{|\phi\rangle \in \mathbb{C}^d, \\ \| |\phi\rangle \| = 1}} \|A\, |\phi\rangle\|$$

denoting the operator norm of $A \in \mathcal{L}(\mathbb{C}^d)$.

A gate set which fulfills these requirements and which we hence assume as canonical gate set for quantum circuits in this thesis is $\mathcal{G} = \{T, H, \mathrm{CNOT}\}$ with

$$\begin{aligned}
T &:= \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} = |0\rangle\langle 0| + e^{i\frac{\pi}{4}} |1\rangle\langle 1|, \\
H &:= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \langle 0| + \frac{|0\rangle - |1\rangle}{\sqrt{2}} \langle 1|, \\
\mathrm{CNOT} &:= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = |0\rangle\langle 0| \otimes \mathbb{I} + |1\rangle\langle 1| \otimes X, \\
X &:= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = |1\rangle\langle 0| + |0\rangle\langle 1|.
\end{aligned}$$

H is called the Hadamard gate; CNOT the controlled-NOT gate. This name is self-explanatory, since the first qubit obviously serves as *control qubit* while the second serves as *target* qubit, to which the NOT operation X is applied iff the first qubit is in

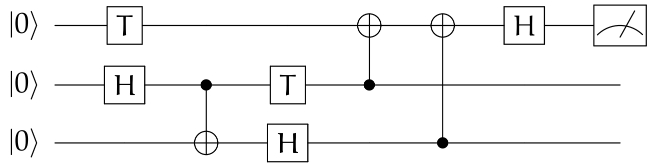


Figure 2.1: A quantum circuit.

the state $|1\rangle$. When drawing a quantum circuit the control qubit of a CNOT operation is usually indicated by a solid dot and the target qubit by a circle like in figure 2.1.

Theorem 2.2. *Any unitary on n qubits can be decomposed into* CNOT *operations and single qubit unitaries. Single qubit unitaries can be approximated to arbitrary finite accuracy by sequences of* T- *and* H-*gates.*

Proof. [4, §4.5]. □

The famous Solovay-Kitaev theorem [5] states that the length of the gate sequence necessary to approximate a unitary scales polylogarithmically with the accuracy for any universal gate set containing inverse gates (note that CNOT and H are their own inverses and T^7 is the inverse of T). Thus, only polynomially many gates are necessary to approximate a unitary on a constant number of qubits to exponential accuracy. Since a polynomial runtime overhead and an inverse exponentially small error – even occuring polynomiallly many times successively – is usually acceptable in quantum computation (e.g. it does not change the complexity class BQP of problems efficiently solvable by a quantum algorithm), constantly sized unitaries are usually not decomposed into elementary gates when describing a quantum algorithm.

Yet, from a complexity theoretic viewpoint it is also important to note that for any finite gate set a simple covering argument of the unitary group reveals unitaries on n qubits that can only be approximated by gate sequences of length exponentially in n [4, §4.5.4]. Consequently, not all quantum transformations on n qubits can be realized by a quantum circuit of polynomial length.

2.2 Quantum measurements

A quantum computation usually ends with a measurement to obtain a classical, probabilistic output. In figure 2.1 a final measurement of the first qubit is represented by the meter symbol at the end of the illustrated circuit. A quantum measurement is described by measurement operators $M_i \in \mathcal{L}(\mathbb{C}^d)$, $i \in [t]$, fulfilling the completeness relation $\sum_{i\in[t]} M_i^\dagger M_i = \mathbb{I}$ with $\mathbb{I}$ denoting the identity operator. For a state $\rho \in \mathcal{D}(\mathbb{C}^d)$ the outcome $i \in [t]$ is measured with probability

$$\mathrm{tr}(M_i \rho M_i^\dagger)$$

and the state afterwards equals

$$\frac{M_i \rho M_i^\dagger}{\mathrm{tr}(M_i \rho M_i^\dagger)}.$$

For pure states $|\psi\rangle$ the probability simplifies to

$$\mathrm{tr}(M_i |\phi\rangle \langle\phi| M_i^\dagger) = \langle\phi| M_i^\dagger M_i |\phi\rangle = \|M_i |\phi\rangle\|^2$$

and the state afterwards equals

$$\frac{M_i |\phi\rangle}{\|M_i |\phi\rangle\|}.$$

Since the trace is multiplicative with respect to the tensor product, it is clear why discarding a system register A is realized by the *partial trace* $\mathrm{tr}_A(\rho)$, which is defined via $\mathrm{tr}_A(\rho_A \otimes \rho_B) = \mathrm{tr}(\rho_A)\rho_B$ and linearity. This operation does not influence the following quantum computation on the remaining registers with regard to a final measurement.

A measurement is called *projective*, iff the measurement operators are projection, i.e. $M_i^2 = M_i$. Often we will measure an n-qubit system *in the computational basis*. In this case the measurement operators equal the projections $\{|x\rangle \langle x|\}_{x\in\{0,1\}^n}$ and the respective post-measurement states are the computational basis states $\{|x\rangle\}_{x\in\{0,1\}^n}$.

Physical quantities, such as energy, spin, occupation number etc., are expressed via an *observable* A, which is a hermitian operator on the Hilbert space $\mathbb{C}^d$. Let P_i, $i \in [l]$, be the projections onto the different eigenspaces of A to the l distinct eigenvalues λ_i. *Measuring with regard to the observable* A refers to the projective measurement $\{P_i\}_{i\in[l]}$ with respective measurement outcomes λ_i. Given a system in the quantum state ρ, the expectation value of the physical quantity described by the observable A is hence given by

$$\operatorname{tr}(A\rho).$$

It is convenient to allow quantum measurements and the classical evaluation of their outcomes at any time of a quantum algorithm and not just at its end. The next lemma adapted from [4][§2.2.8] shows how such an algorithm can be simulated by a unitary quantum circuit with one final measurement. From a complexity theoretic piont of view it is important to note that this substitution requires not more than polynomially many gates for simple measurements such as measurements in the computational basis.

Lemma 2.3. *Consider a measurement with measurement operators* $\{M_i\}_{i=0}^{t-1}$ *followed by a unitary* U_m *depending on the outcome* m. *The same probability distribution of output states is realized by adding an ancilla register of dimension* t *initialized in the state* $|0\rangle$ *to the system, applying a unitary* U_M *with*

$$U_M\left(|\psi\rangle \otimes |0\rangle\right) = \sum_{i=0}^{t-1} M_i\,|\psi\rangle \otimes |i\rangle\,,$$

followed by the controlled unitary

$$U = \sum_{i=0}^{t-1} U_i \otimes |i\rangle\,\langle i|$$

and discarding the ancilla after having it measured in the computational basis with outcome m.

Proof. Note that a unitary U_M fulfilling the above equation exists, since the completeness relation $\sum_{i=0}^{t-1} M_i^\dagger M_i = \mathbb{I}$ ensures that the above equation is consistent with U_M preserving the scalar product:

$$\big(\langle\phi| \otimes \langle 0|\big)U_M^\dagger U_M\big(|\psi\rangle \otimes |0\rangle\big) = \langle\phi|\psi\rangle\,.$$

In the original setting we measure the output m given a state $|\psi\rangle$ with probability

$$p_m \coloneqq \langle\psi|\,M_m^\dagger M_m\,|\psi\rangle$$

and obtain after the application of U_m the state

$$|\psi_m\rangle := \frac{U_m M_m |\psi\rangle}{\sqrt{p_m}}.$$

The state in our alternative setting before the ancilla measurement equals

$$|\psi^{\text{pre}}\rangle := U U_M |\psi\rangle \otimes |0\rangle = \sum_{i=0}^{t-1} U_i M_i |\psi\rangle \otimes |i\rangle .$$

It is easy to see that a measurement of the ancilla in the computational basis gives output m with probability p_m and that the post-measurement state reduced by the ancilla system equals $|\psi_m\rangle$. □

2.3 Pauli group

Definition 2.4. *A* Pauli operator *is any tensor product of the single qubit Pauli operators*

$$\{\mathbb{I}, X, Y, Z\} =: \mathcal{P}$$

with the matrix representation

$$X := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It is easy to check that all Pauli operators are unitary and hermitian. Pauli operators are frequently used in quantum computation due to the following important fact:

Lemma 2.5. *The Pauli operators $\{\mathbb{I}, X, Y, Z\}$ form a basis for the complex 2x2 matrices, i.e. every operator $\rho \in \mathcal{L}(\mathbb{C}^2)$ can be written in the form*

$$\rho = \frac{1}{2}(r_0 \, \mathbb{I} + \vec{r} \cdot \vec{\sigma})$$

with $\vec{\sigma} = (X, Y, Z)^T$ and $r_0 \in \mathbb{C}$, $\vec{r} \in \mathbb{C}^3$.

The following equivalences hold:

1. ρ *is hermitian* $\Longleftrightarrow (r_0, \vec{r})$ *is real.*
2. ρ *is positive semi-definite* $\Longleftrightarrow |\vec{r}| \leqslant r_0$.
3. $\mathrm{tr}(\rho) = 1 \Longleftrightarrow r_0 = 1$.
4. ρ *is a pure state of the Hilbert space* $\mathbb{C}^2 \Longleftrightarrow |\vec{r}| = 1$.

Proof. [6, §2.1]. □

The above lemma allows to represent a quantum state via the vector $\vec{r}$ in the so-called *Bloch sphere*, a unit sphere in $\mathbb{R}^3$ with the axes X, Y and Z (see e.g. figure 7.2). Since $|\vec{r}| \leqslant 1$ and equality holds iff ρ is pure, all states lie within the sphere with the pure states on the surface.

Like the single qubit Pauli operators form a basis for $\mathcal{L}(\mathbb{C}^2)$, the N-fold tensor products of single qubit Pauli operators obviously form a basis for $\mathcal{L}(\mathbb{C}^{2^N})$. Since

$$XY = iZ, \qquad YZ = iX, \qquad ZX = iY,$$

the N-qubit Pauli operators with a prefactor of $\{\pm 1, \pm i\}$ form a group.

Definition 2.6. *The* Pauli group on N qubits *is defined as*

$$\mathcal{P}_N := \left\{ c \cdot \sigma \mid c \in \{\pm 1, \pm i\},\ \sigma \in \mathcal{P}^{\otimes N} \right\}.$$

Definition 2.7. *The number of the Pauli operator* $\sigma \in \mathcal{P}$ *appearing in the tensor product of a Pauli group element* $\mu \in \mathcal{P}_N$ *is called the the* σ-weight *of* μ *and denoted by* $w_\sigma(\mu)$. *The total number of non-identity operators in the tensor product is called the* weight *of* μ *and denoted by* $w(\mu)$.

Since two elements of $\mathcal{P}_N$ either commute or anticommute, we introduce the following η-function:

Definition 2.8. *For any two elements* $\mu, \nu \in \mathcal{P}_N$ *we write*

$$\eta(\mu,\nu) = \begin{cases} +1 & \textit{if } \mu \textit{ and } \nu \textit{ commute} \\ -1 & \textit{if } \mu \textit{ and } \nu \textit{ anticommute.} \end{cases}$$

2.4 Quantum channels

Definition 2.9. *A* superoperator *is a linear map from* $\mathcal{L}(\mathbb{C}^d)$ *to* $\mathcal{L}(\mathbb{C}^{d'})$.

A superoperator $\mathcal{N}$ *is* completely positive *iff* $\mathcal{N} \otimes \mathrm{Id}$ *preserves positive semi-definiteness with* Id *denoting the identity channel for any dimension.*

A quantum channel *is a completely positive, trace-preserving* (cpt) *superoperator.*

Quantum channels offer the most general transformation of quantum states. The properties of complete positivity and trace-preservation are necessary and sufficient requirements to ensure that a superoperator $\mathcal{N}$ and its trivial extensions $\mathcal{N} \otimes \mathrm{Id}$ map density matrices to density matrices. Since these properties are not always easy to check for a given superoperator, the following equivalent conditions are useful:

Theorem 2.10. *The following statement are equivalent for* $\mathcal{N} : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$*:*

1. $\mathcal{N}$ *is a quantum channel.*
2. *There exist so-called* Kraus operators *or* operation elements $\{N_i\}_{i\in[t]} \subset \mathcal{L}(\mathbb{C}^d, \mathbb{C}^{d'})$ *with* $\sum_{i\in[t]} N_i^\dagger N_i = \mathbb{I}$ *and*

$$\mathcal{N}(\rho) = \sum_{i\in[t]} N_i \rho N_i^\dagger.$$

3. (Stinespring dilation) *There exists a unitary* $U \in \mathcal{U}(\mathbb{C}^d \otimes \mathbb{C}^{(d')^2})$ *such that*

$$\mathcal{N}(\rho) = \mathrm{tr}_{\mathbb{C}^{dd'}}\left(U\left(\rho \otimes |0\rangle\langle 0|^{\otimes (d')^2}\right) U^\dagger\right).$$

Proof. [7, §2.2]. □

Obviously, Kraus operators of quantum channels fulfill the same requirement as measurement operators. Hence, a quantum channel with Kraus operators $\{N_i\}_{i\in[t]}$ outputs the probabilistic mixture of post-measurement states with regard to the measurement operators $\{N_i\}_{i\in[t]}$.

Note that the set of Kraus operators for a quantum channel is not unique:

Lemma 2.11. *Given a set $\{N_i\}_{i\in[t]}$ of Kraus operators for a quantum channel $\mathcal{N}$ all other valid sets of Kraus operators are exactly those of the form $\{M_i\}_{i\in[t]}$ with*

$$M_i = \sum_{k\in[t]} U_{ki} N_k$$

for a unitary $U \in \mathcal{U}(\mathbb{C}^t)$. Note that the sets $\{N_i\}_{i\in[t]}$ and $\{M_i\}_{i\in[t]}$ can be assumed to be of same size by padding the eventually smaller with zero operators.

Proof. [4, §8.2.4]. □

We end this section by introducing some common single qubit quantum channels. Their Kraus operators can be read off easily from the following definition for arbitrary linear operators. Yet, the physical meaning of the partly depolarizing and erasing channel is better reflected by the special form they take for quantum states presented in the subsequent lemma.

Definition 2.12. *The* partly dephasing channel $T : \mathcal{L}(\mathbb{C}^2) \to \mathcal{L}(\mathbb{C}^2)$ *is defined as*

$$T_\epsilon^{\text{deph}}(\rho) := (1-\epsilon)\rho + \epsilon\frac{\rho + Z\rho Z}{2}.$$

The partly depolarizing channel $T : \mathcal{L}(\mathbb{C}^2) \to \mathcal{L}(\mathbb{C}^2)$ *is defined as*

$$T_\epsilon^{\text{depol}}(\rho) = (1-\epsilon)\rho + \epsilon\frac{\mathbb{I}\,\rho\,\mathbb{I} + X\rho X + Y\rho Y + Z\rho Z}{4}.$$

The partly erasing channel $T : \mathcal{L}(\mathbb{C}^2) \to \mathcal{L}(\mathbb{C}^{2^2})$ *is defined as*

$$T_\epsilon^{\text{eras}}(\rho) = (1-\epsilon)\rho \otimes |0\rangle\langle 0| + \epsilon\frac{\mathbb{I}\,\rho\,\mathbb{I} + X\rho X + Y\rho Y + Z\rho Z}{4} \otimes |1\rangle\langle 1|.$$

Lemma 2.13. *For density matrices* $\rho \in \mathcal{D}(\mathbb{C}^2)$ *the output of the partly depolarizing channel equals*

$$T_\epsilon^{\text{depol}}(\rho) = (1-\epsilon)\rho + \epsilon\frac{\mathbb{I}}{2}$$

and the output of the partly erasing channel

$$T_\epsilon^{\text{eras}}(\rho) = (1-\epsilon)\rho \otimes |0\rangle\langle 0| + \epsilon\frac{\mathbb{I}}{2} \otimes |1\rangle\langle 1|.$$

Proof. By using the qubit representation $\rho = \frac{\mathbb{I}}{2} + \vec{r}\cdot\vec{\sigma}$ of lemma 2.5 and the relations $XY = iZ$, $YZ = iX$ and $ZX = iY$ one can easily show that

$$\frac{1}{4}(\mathbb{I}\,\rho\,\mathbb{I} + X\rho X + Y\rho Y + Z\rho Z) = \frac{\mathbb{I}}{2}. \qquad \square$$

In all three channels above ϵ can be interpreted as error parameter. With probabiltiy $1-\epsilon$ the input state remains invariant, while with probabiltiy ϵ the state is disturbed. In case of dephasing disturbance means that a superposition $\rho = (\alpha|0\rangle + \beta|1\rangle)(\alpha^*\langle 0| + \beta^*\langle 1|)$ is replaced by the mixture $|\alpha|^2|0\rangle\langle 0| + |\beta|^2|1\rangle\langle 1|$. Hence, the completely dephasing channel only outputs mixtures of the classical states $|0\rangle$ and $|1\rangle$. In case of depolarizing and erasing disturbance means that the input state is replaced by the *completely mixed state* $\frac{\mathbb{I}}{2}$. This state does not carry any information about the input and is the equal mixture of all basis states for any orthonormal basis. The partly erasing channel indicates the disturbance moreover by an additional *flag qubit*.

2.5 Norms on operators and superoperators

The operator norm $\|A\|$ is not the only useful norm for operators $A \in \mathcal{L}(\mathbb{C}^d)$. This section introduces some additional frequently used norms on operators and superoperators.

Definition 2.14. *The* trace norm *of an operator* $A \in \mathcal{L}(\mathbb{C}^d)$ *is defined as*

$$\|A\|_1 := \operatorname{tr}\sqrt{A^\dagger A}.$$

$\|A - B\|_1$ *is called the* trace distance *between the operators* A *and* $B \in \mathcal{L}(\mathbb{C}^d)$.

Lemma 2.15. *The trace norm is a valid norm and fulfills the following properties for all operators* $A, B \in \mathcal{L}(\mathbb{C}^d)$*:*

1. $\|A\|_1 = \sup_{\substack{B \in \mathcal{L}(\mathbb{C}^d) \\ B \neq 0}} \frac{|\operatorname{tr}(AB)|}{\|B\|}$.
2. $\|AB\|_1, \|BA\|_1 \leqslant \|A\| \|B\|_1$.
3. $\|T(A)\|_1 \leqslant \|A\|_1$ *for all quantum channels* $T : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$.

Proof. Property 3 is shown in [8, proposition 4.1]; the other properties in [9, §5.2]. □

Lemma 2.16. *For any hermitian operator* $\omega \in \mathcal{L}(\mathbb{C}^d)$ *and states* $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$ *it holds that*

$$\begin{aligned} \|\omega\|_1 &= \max_{-\mathbb{I} \leqslant \Lambda \leqslant \mathbb{I}} \operatorname{tr}[\Lambda\omega], \\ \|\rho - \sigma\|_1 &= 2 \max_{0 \leqslant \Lambda \leqslant \mathbb{I}} \operatorname{tr}[\Lambda(\rho - \sigma)] \end{aligned}$$

with the maximization over hermitian operators $\Lambda \in \mathcal{L}(\mathbb{C}^d)$ *with all eigenvalues in the interval* $[-1, 1]$ *and* $[0, 1]$*, respectively.*

Proof. [10, exercise 9.1.4] and [10, lemma 9.1.1]. □

For a norm on superoperators the first choice is to consider the norm induced by the trace norm on operators:

Definition 2.17. *The* trace norm *of a superoperator* $T : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$ *is defined as*

$$\|T\|_1 \coloneqq \sup_{\substack{A \in \mathcal{L}(\mathbb{C}^d) \\ \|A\|_1 = 1}} \|T(A)\|_1.$$

Unfortunately, the trace norm on superoperator does not always stay invariant if the superoperator is tensored with the identity [9, §5.3]. This disadvantage is overcome by the stabilized version of the trace norm, the so-called *diamond norm*. Note that we state here the definition of [11, §18.2.2], which is equivalent to the one in [9, §5.3]:

Definition 2.18. *The* diamond norm *of a superoperator* $\mathsf{T} : \mathcal{L}(\mathbb{C}^{\mathrm{d}}) \to \mathcal{L}(\mathbb{C}^{\mathrm{d}'})$ *is defined as*

$$\|\mathsf{T}\|_{\Diamond} \coloneqq \|\mathsf{T} \otimes \mathrm{Id}_{\mathbb{C}^{\mathrm{d}}}\|_1.$$

Lemma 2.19. *The following holds for any superoperators* T, T_1, T_2, T_1', T_2' *and any linear operator* A *of suitable dimensions:*

1. $\|\mathsf{T}\|_{\Diamond} \geqslant \|\mathsf{T}\|_1$.
2. $\|\mathsf{T}(A)\|_{\Diamond} \leqslant \|\mathsf{T}\|_{\Diamond}\|A\|_1$.
3. $\|\mathsf{T}_1 \circ \mathsf{T}_2\|_{\Diamond} \leqslant \|\mathsf{T}_1\|_{\Diamond}\|\mathsf{T}_2\|_{\Diamond}$.
4. $\|\mathsf{T}_1 \otimes \mathsf{T}_2\|_{\Diamond} = \|\mathsf{T}_2\|_{\Diamond}\|\mathsf{T}_2\|_{\Diamond}$.
5. T *cpt* $\implies$ $\|\mathsf{T}\|_{\Diamond} = 1$.
6. *If* T_1, T_2, T_1' *and* T_2' *are of norm at most* 1 *and* $\|\mathsf{T}_1' - \mathsf{T}_1\| \leqslant \epsilon_1$ *and* $\|\mathsf{T}_2' - \mathsf{T}_2\| \leqslant \epsilon_2$, *then* $\|\mathsf{T}_2' \circ \mathsf{T}_1' - \mathsf{T}_2 \circ \mathsf{T}_1\| \leqslant \epsilon_1 + \epsilon_2$.

Proof. Properties 1 – 5 correspond to [9, lemma 12]; property 6 to [9, lemma 13]. □

Chapter 3

Introduction to classical and quantum complexity theory

3.1 The classical computing model

In this chapter we present foundations of complexity theory that are covered by standard textbooks such as [12], [13] or [14]. As a reference for quantum complexity theory consider the book by [15] and the review by [16].

Definition 3.1. *A* promise problem*, or briefly* problem*, is a tuple* $A = (A_{\text{yes}}, A_{\text{no}})$ *with*

$$\begin{aligned} A_{\text{yes}} \cap A_{\text{no}} &= \varnothing, \\ A_{\text{yes}} \cup A_{\text{no}} &\subseteq \Sigma^* \end{aligned}$$

with Σ^* *denoting the set of all strings over the binary alphabet* $\Sigma := \{0, 1\}$*.*

$A_{\text{yes}} \cup A_{\text{no}}$ *is called the* promise*. Problems with* $A_{\text{yes}} \cup A_{\text{no}} = \Sigma^*$ *are called* decision probems*.*

Without stating it each time explicitely, we will denote the problem instance inputted to a protocol always by the variable $x \in \Sigma^n$ and its length by the variable n. When using functional terminology in the description of protocols, such as "polynomially

/ exponentially many bits / steps / gates" etc., the input length n is considered as function variable.

Informally, problems are often defined as a partition of other objects than binary strings such as Boolean formulas, graphs, local Hamiltonians or density matrices. We then assume that these objects are represented via a natural binary encoding. Binary strings that do not correspond to a valid encoding according to such a scheme can be considered as no-instances by default.

Theoretical computer science mainly deals with two kinds of computational tasks:

1. Computing a function $\Sigma^* \to \Sigma^*$.
2. Deciding a bipartite question in form of a problem $A = (A_{\text{yes}}, A_{\text{no}})$.

The fundamental computational model for both is the deterministic Turing machine.

Definition 3.2. *A* (deterministic) Turing machine (DTM) M *is defined by a finite set of states* $S = \{s_i\}_{i=0}^{l} \cup \{s_f\}$ *and a transition function*

$$\delta : S \times \{0, 1, \square\} \to S \times \{0, 1, \square\} \times \{L, R, N\}.$$

A Turing machine carries out its computation on an endless tape initialized with the input $x \in \Sigma^*$ *padded by infinitely many blank symbols* $\square$ *on both sides. The Turing machine starts with the initial state* s_0 *and a reading / writing head located on the first symbol of the input. A computational step of the machine is described by the transition function. If* $\delta(s_i, a) = (s_j, b, X)$ *and the machine is in the state* s_i *and reads the symbol* a *at the head position, it transfers into the state* s_j*, overwrites the symbol by* b *and moves the head left, right or neutral depending whether* $X = L$*,* $X = R$ *or* $X = N$*, respectively.*

If the final state s_f *is reached, the machine* halts *and outputs the string that is written between the head position and the next blank symbol.*

For an input x *for that the machine halts we consider the total number of executed steps as* runtime *of the machine and denote the output by* $M(x)$*. For a multipartite input* $x_1, x_2, \ldots \in \Sigma^*$*, which is written onto the tape successivly separated by blank symbols, we denote the output accordingly by* $M(x_1, x_2, \ldots)$*.*

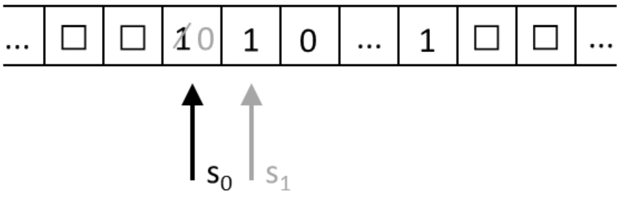


Figure 3.1: Action of a Turing machine with transition rule $\delta(s_0, 1) = (s_1, 0, R)$.

Figure 3.1 illustrates a Turing machine that carries out its first head movement on the input $x = 110 \ldots 1$ according to a transition function δ with $\delta(s_0, 1) = (s_1, 0, R)$.

A Turing machine can be represented by its *Gödel number*, a binary encoding of its transition function. Note that a finite encoding is possible since the transition function is determined by finitely many transition rules due to the state set being finite. If a binary number does not have the form of a valid Gödel number it is interpreted as the encoding of a trivial machine that always outputs 0. If we state in this paper that "a Turing machine is given" we mean that the Gödel number of the machine is supplied. Analogously, a computable function is given via the Gödel number of the machine that computes it.

Definition 3.3. *A function* $f : \Sigma^* \to \Sigma^*$ *is called* computable *iff there exists a DTM that for every input* $x \in \Sigma^*$ *outputs* $f(x)$.

A function $f : \mathbb{N}_0 \to \Sigma^*$ *is called* computable *iff there exists a DTM that for every input of length* n *outputs* $f(n)$.

A function of one of the above kinds is called polynomial-time computable *iff the runtime of the DTM is bounded by a polynomial in the input length.*

Like for problems, we informally allow other sets as function ranges than Σ^*, while strictly speaking they are considered via a binary encoding. When we later speak of computable functions that obviously map to the complex or real plane, such as computable completeness and soundness functions of complexity theoretic protocols, we assume that their range is restricted to algebraic numbers represented by a binary encoding of their minimal rational polynomial and an isolating rational interval. The work of [17] shows how algebraic standard operations such as addition, substraction multiplication, division and powering can be realized in polynomial time in this representation. Moreover, there exists a DTM that for input 1^n and an algebraic number b can compute the n-th digit of the real and imaginary part of b in time polynomial in n and the representation length of b.

The largest subset of complex numbers for that a reasonable definition of computability is possible consists of all numbers for that there exists a DTM M outputting the n-th digit of the real and imaginary part for an input of length n. Such a number could be represented simply by the Gödel number of M. Polynomial-time computability of a function mapping to this number would require then that the runtime of M is also upper bounded by a polynomial.

Still, we assume in this thesis that computable functions are limited to algebraic numbers. These are sufficient for all purposes and moreover, their representation by a minimal rational polynomial and an isolating rational interval has the advantage that equality of two algebraic numbers is decidable in polynomial time, while this is undecidable for two complex numbers represented by general Turing machines computing their digits.

While the function range can be handled very informally, one needs to strictly distinguish the two different options for the domain in the above definition. For the property of polynomial-time computability is makes a huge difference if a function is considered to be defined on $\mathbb{N}_0$ or on Σ^*. The former is computed within a runtime upper bounded by a polynomial in n, the latter – when the input is interpreted as a binary encoding of a natural number n – within a runtime upper bounded by a polynomial in $\log(n)$.

Definition 3.4. *A function* $f : \mathbb{N}_0 \to \mathbb{N}_0$ *is called* time-constructible, *iff there exists a DTM that for each input of length* n *halts after exactly* $f(n)$ *steps.*

Lemma 3.5. *The following is true for any function* $f : \mathbb{N}_0 \to \mathbb{N}_0$*:*

1. f *is time-constructible* $\Rightarrow$ f *is computable.*

2. f *is computable* $\Rightarrow \exists$ *time-constructible* $f' \geqslant f$.

Proof.

1. Let M_f^{time} be a DTM that time-constructs f. Then M_f^{time} with the following adaption computes f: Interrupt after each original computational step by changing into a new "interruption state", let the head run to the end of the written tape and increment a counter there starting at 0. Note that it is easily feasible to shift the counter whenever a new blank cell at the end of the written tape is supposed to be overwritten.

   If the originally following state is the final state, let the head remain on the beginning of the counter and change into the final state, i.e. output the counter. Otherwise, let the head run back to its original position and change into the following state.

2. Let M_f be a DTM that computes f. If the output of M_f is changed into the unary representation $1^{f(n)}$ before changing into the final state, the Turing machine needs at least time $f(n)$ and hence time-constructs a function f' with $f' \geqslant f$. □

With the definition of computability for functions we specified the first computational task stated at the beginning of the section. It remains to specify the sencond task, the decidability of problems:

Definition 3.6. *A problem* $A = (A_{\text{yes}}, A_{\text{no}})$ *is* decidable *iff there exists a DTM* M *such that*

$$\begin{aligned} \forall x \in A_{\text{yes}}\ M(x) = 1 \text{ (“M accepts input } x\text{”)}, \\ \forall x \in A_{\text{no}}\ M(x) = 0 \text{ (“M rejects input } x\text{”)}. \end{aligned}$$

It is possible to define problems that are undecidable. The most famous one is the Halting problem, whose undecidability proof works via a simple contradiction argument.

Definition 3.7. *The Halting problem* $(A_{\text{yes}}, A_{\text{no}})$ *is defined via*

$$\begin{aligned} A_{\text{yes}} &= \{x \mid \text{ the DTM with Gödel number } x \text{ halts for input } x\}, \\ A_{\text{no}} &= \{x \mid \text{ the DTM with Gödel number } x \text{ does not halt for input } x\}. \end{aligned}$$

Lemma 3.8. *The Halting problem is undecidable.*

Proof. Assume the Halting problem $(A_{\mathrm{yes}}, A_{\mathrm{no}})$ is decidable via the DTM M. Let x be the Gödel number of the DTM M' that simulates M and halts iff M outputs 0. Then

$$\begin{aligned} x \in A_{\mathrm{yes}} &\Leftrightarrow M \text{ outputs 1 for input } x \\ &\Leftrightarrow M' \text{ does not halt for input } x \\ &\Leftrightarrow x \in A_{\mathrm{no}}. \end{aligned}$$

This is a contradiction. Hence, our initial assumption was wrong and the Halting problem is undecidable. □

3.2 Randomized and quantum computing models

Besides the deterministic Turing machine theoretical computer science developed alternative computational models, but so far all models, including quantum computing, proved to be equivalent to the concept of deterministic Turing machines. The *Church-Turing thesis* asserts that this is the only possibility. The *extended Church-Turing thesis* moreover asserts that any model can simulate another one only with a polynomial runtime overhead, which is relevant since any polynomially bounded runtime corresponds to a practically feasible runtime.

Until today no definitive contradiction of the extended Church-Turing thesis has been found, although it is widely believed that quantum computing violates it. An indication is Shor's polynomial-time quantum algorithm [1] for factorization, since no classical algorithm of polynomial runtime for this problem has been found yet.

Before defining the quantum computing model, we introduce the intermediate concept of classical, randomized computing via the probabilistic Turing machine.

Definition 3.9. *A* probabilistic Turing machine *extends the model of the deterministic Turing machine by allowing a binary probabilistic branching at every computational step, which is mathematically reflected by a transition function of the form*

$$\delta : S \times \{0, 1, \square\} \to \big(S \times \{0, 1, \square\} \times \{L, R, N\}\big)^{\times 2}.$$

The two transitions possible in the state s_i with a read tape symbol a correspond to the two 3-tuple described by $\delta(s_i, a)$ and each occur with probability $\frac{1}{2}$.

The runtime of a PTM for an input x is the longest runtime over all execution branches.

A PTM accepts and rejects a problem input x probabilistically. The acceptance probability corresponds to the fraction of accepting branches and the rejection probability to the fraction of rejecting branches.

We will use all terminology introduced for deterministic Turing machines and the formalism of Gödel numbers accordingly for probabilistic Turing machines. It is an open question, if probabilistic Turing machines with reasonable output probabilities and a polynomial runtime already violate the extended Church-Turing thesis.

In section 2.1 we introduced the concept of quantum circuits. Since a circuit acts on a fixed number of qubits, it cannot describe an algorithm for variable input length. For this a whole family of quantum circuits $(V_x)_{x \in \Sigma^*}$ is needed. Constructing the circuit V_x for an input x needs nothing else than an algorithm. For a reasonable quantum computing model we therefore require the existence of a Turing machine that computes the quantum circuit V_x given input x.

As discussed before "computing" something like a quantum circuit strictly means computing a natural binary representation of it. Since quantum circuits can be assumed to be restricted to the finite gate set $\{T, H, CNOT\}$, it is easy to fix a simple binary description scheme for them, e.g. by concatenating respective gate numbers and affected qubit numbers in unary encoding separated by zeros. We will assume in the following that such a sequential, efficient *Gödel scheme* for quantum circuits is fixed. A binary number not corresponding to a valid circuit according to this scheme is interpreted as trivial circuit without gates.

Since quantum gates allow the realization of the classical universal operations NOT ($\neg$), AND ($\wedge$) and OR ($\vee$) as well as the equal mixture of the classical states $|0\rangle$ and $|1\rangle$, the quantum computing model does not only cover classical preprocessing but possesses the ability for classical and randomized subroutines at any time of the circuit.

Definition 3.10. *We call the Turing machine* M *computing a quantum circuit family* $V = (V_x)_{x\in\Sigma^*}$*, also the* generating *Turing machine of* V.

The quantum circuit family V *is* polynomial-time generated *iff the runtime of* M *is upper bounded by a polynomial.*

The quantum circuit family V accepts *an input* $x \in \Sigma^*$ *with probability*

$$\mathbb{P}_{\text{acc}}(x) = \text{tr}\left(\Pi_{\text{acc}} V_x |0\rangle \langle 0|^{\otimes z} V_x^\dagger\right)$$

and $\Pi_{\text{acc}} = |1\rangle\langle 1| \otimes \mathbb{I}^{\otimes z-1}$ *denoting the projection of the first of the* z *circuit qubits onto the* $|1\rangle$ *state.*

Note that later we will use the terminology of acceptance and rejection also for quantum circuit families for that complexity class specifications require a different input state than $|0\rangle^{\otimes z}$.

As an alternative to the above definition some literature describes the quantum computing model by a computable circuit family $(U_n)_{n\in\mathbb{N}_0}$ that just depends on the length n of the computational input x, but therefore obtains $|x\rangle \otimes |0\rangle^{\otimes z}$ as circuit input instead of the all zero state. For a polynomial runtime of the generating Turing machine these circuit families are usually called *uniform*. A uniform quantum circuit family $(U_n)_{n\in\mathbb{N}_0}$ can easily be transformed into a polynomial-time generated family $(V_x)_{x\in\Sigma^*}$ by prepending X gates to the circuit on the qubits that correspond to a 1 in the input string x. A family of quantum circuits $(V_x)_{x\in\Sigma^*}$ generated in polynomial-time by a Turing machine M on the other hand corresponds to the uniform family $(U_n)_{n\in\mathbb{N}_0}$ with the Turing machine that outputs the quantum circuit that first measures the input state in the computational basis, simulates M on the classical measurement outcome x and then carries out the quantum gates described by $M(x)$.

3.3 Standard complexity classes

A complexity class is a set of problems. Usually it comprises all problems decidable by a certain machine model, which attributes the problems a similar “complexity”. We will list here the definitions of the most important complexity classes, namely those of efficiently solvable and efficiently verifiable problems with regard to the classical, randomized and quantum computing model.

Definition 3.11. *The complexity class* P *("polynomial time") is the set of all decision problems that can be decided by a deterministic Turing machine of polynomial runtime.*

Definition 3.12. *The complexity class* NP *("non-deterministic polynomial time") is the set of all decision problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *for that there exists a deterministic Turing machine M of polynomial runtime and a polynomial* n_w *such that*

$$\begin{aligned} \forall x \in A_{\text{yes}}\ \exists y \in \Sigma^{n_w} &: M(x,y) = 1, \\ \forall x \in A_{\text{no}}\ \forall y \in \Sigma^{n_w} &: M(x,y) = 0. \end{aligned}$$

Definition 3.13. *The complexity class* $\text{PromiseBPP}(c,s)$ *(*$\text{BPP}(c,s)$*) ("bounded error probabilistic polynomial time") is the set of all (decision) problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *for that there exists a probabilistic Turing machine M of polynomial runtime such that*

$$\begin{aligned} \forall x \in A_{\text{yes}} &: \mathbb{P}[M(x) = 1] \geqslant c, \\ \forall x \in A_{\text{no}} &: \mathbb{P}[M(x) = 1] \leqslant s. \end{aligned}$$

Definition 3.14. *The complexity class* $\text{PromiseMA}(c,s)$ *(*$\text{MA}(c,s)$*) ("Merlin-Arthur") is the set of all (decision) problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *for that there exists a probabilistic Turing machine M of polynomial runtime and a polynomial* n_w *such that*

$$\begin{aligned} \forall x \in A_{\text{yes}}\ \exists y \in \Sigma^{n_w} &: \mathbb{P}[M(x,y) = 1] \geqslant c, \\ \forall x \in A_{\text{no}}\ \forall y \in \Sigma^{n_w} &: \mathbb{P}[M(x,y) = 1] \leqslant s. \end{aligned}$$

Definition 3.15. *The complexity class* $\text{BQP}(c,s)$ *("bounded error quantum polynomial time") is the set of all problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *for that there exists a polynomial-time generated familiy of quantum circuits* $V = (V_x)_{x \in \Sigma^*}$ *on z qubits with z polynomial such that*

$$\begin{aligned} \forall x \in A_{\text{yes}} &: \ \operatorname{tr}\left(\Pi_{\text{acc}} V_x \left|0\right\rangle \left\langle 0\right|^{\otimes z} V_x^\dagger\right) \geqslant c, \\ \forall x \in A_{\text{no}} &: \ \operatorname{tr}\left(\Pi_{\text{acc}} V_x \left|0\right\rangle \left\langle 0\right|^{\otimes z} V_x^\dagger\right) \leqslant s. \end{aligned}$$

Definition 3.16. *The complexity class* $\text{QCMA}(c,s)$ *("quantum-classical Merlin-Arthur") is* *the set of all problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *for that there exists a polynomial-time generated familiy of quantum circuits* $V = (V_x)_{x \in \Sigma^*}$ *on* $z + n_w$ *qubits with* z *and* n_w *polynomial such that*

$$\forall x \in A_{\text{yes}}\ \exists y \in \Sigma^{n_w} :\ \operatorname{tr}\left(\Pi_{\text{acc}} V_x \left(|0\rangle \langle 0|^{\otimes z} \otimes |y\rangle \langle y|\right) V_x^{\dagger}\right) \geqslant c,$$
$$\forall x \in A_{\text{no}}\ \forall y \in \Sigma^{n_w} :\ \operatorname{tr}\left(\Pi_{\text{acc}} V_x \left(|0\rangle \langle 0|^{\otimes z} \otimes |y\rangle \langle y|\right) V_x^{\dagger}\right) \leqslant s.$$

Definition 3.17. *The complexity class* QMA(c, s) *("quantum Merlin-Arthur") is the set of all problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *for that there exists a polynomial-time generated familiy of quantum circuits* $V = (V_x)_{x \in \Sigma^*}$ *on* $z + n_w$ *qubits with* z *and* n_w *polynomial such that*

$$\forall x \in A_{\text{yes}}\ \exists \rho \in \mathcal{D}(\mathbb{C}^{2^{n_w}}) :\ \operatorname{tr}\left(\Pi_{\text{acc}} V_x \left(|0\rangle \langle 0|^{\otimes z} \otimes \rho\right) V_x^{\dagger}\right) \geqslant c,$$
$$\forall x \in A_{\text{no}}\ \forall \rho \in \mathcal{D}(\mathbb{C}^{2^{n_w}}) :\ \operatorname{tr}\left(\Pi_{\text{acc}} V_x \left(|0\rangle \langle 0|^{\otimes z} \otimes \rho\right) V_x^{\dagger}\right) \leqslant s.$$

Definition 3.18. *In the definitions of* PromiseBPP, BPP, PromiseMA, MA, BQP, QCMA *and* QMA *the parameter* c *is called* completeness*; the parameter* s soundness. *When the parameters are omitted, the default values* $c = \frac{2}{3}$ *and* $s = \frac{1}{3}$ *are assumed.*

Running a protocol for a yes-instance is accordingly called the completeness case*; running it for a no-instance the* soundness case*.*

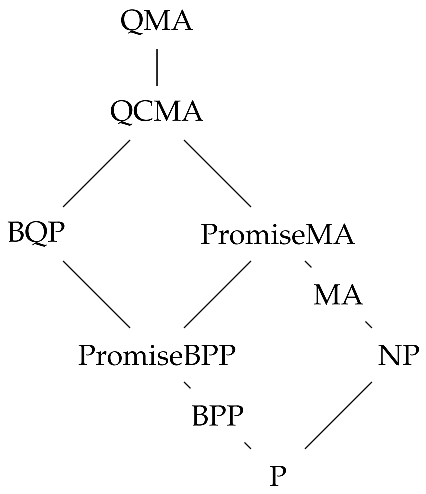


Figure 3.2: Hierarchy of complexity classes (classes contain those connected below).

The introduced complexity classes form the hierarchy depicted in figure 3.2.

We will later use the notion "C-protocol" for any Turing machine or quantum circuit family obeying the requirements for one of the above introduced complexity classes C. In extension of the notion "decidability" we say that the "C-protocol V decides the problem A" if V is the protocol for that the membership $A \in C$ is proven. Note that the runtime of a PromiseBPP-, PromiseMA-, BQP-, QCMA- and QMA-protocol also has to be polynomially bounded for non-promised inputs. Weakening this requirement would not change the complexity class, but the specification is important for the structural studies in chapter 5.

The binary string y in the definitions of NP, PromiseMA, MA and QCMA and the quantum state ρ in the definition of QMA are called the *witness* that the protocol – in this case also called the *verifier* – receives. The difference between QCMA and QMA is that a QCMA verifier receives a classical witness y and a QMA verifier a quantum witness ρ. Note, that the definition of QMA remains unchanged if the witness is restricted to pure states.

The verifier owes its name to the fact that he can decide a problem instance by simply "verifying" a condition that the witness only obeys for yes-instances. Considering for example the NP-problem Satisfiability of Boolean formulas, the verifier checks if the witness, interpreted as assigment, satisfies the Boolean formula. Such an assignment exists in case of a yes-instance, while for a no-instance no assigment has this property.

The quantifier expression with the witness allows to interpret a verifying protocol as a game in that an adversary party, historically called *Merlin* or *prover*, supplies the witness in order to convince the verifier, called *Arthur*, that the input is a yes-instance. In the completeness case we call Merlin therefore *honest*, in the soundness case *malicious*. The terminology of honest and malicious Merlins is also adapted for other suitable complexity classes like the multiprover classes in chapter 8.

It is conventional to consider the classical classes P and NP as sets of decision problems and the quantum classes BQP, QCMA and QMA as sets of promise problems. There is no need to extend P and NP to promise problems, since P- and NP-protocols split the set of all binary strings into yes- and no-instances according to their two possible outputs. On the contrary, for the respective quantum protocols there often exist instances that do neither fulfill the yes- nor the no-instance condition on the acceptance probability. In order to not rule out too many logically and physically interesting problems from these complexity classes, they are defined as sets of promise problems.

Historically, the randomized classes BPP and MA are defined restricted to decision problems like their non-randomized counterparts P and NP. But latest with the upcoming of quantum complexity theory with its strong analogies, the debate started if these classes should be broadened to promise problems. Since the difference is relevant for structural studies like in chapter 5, we call the classes restricted to decision problems BPP and MA and extended to promise problems PromiseBPP and PromiseMA.

3.4 Amplification

The mean of the binomial distribution with m tosses of success probability p equals mp. Chernoff bound is a useful tool to bound the tails of the binomial distribution:

Lemma 3.19 (Chernoff bound)**.** *Let* $p \in [0,1]$ *and* $m, k \in \mathbb{N}$ *with* $k \leqslant mp$*. Then it holds*

$$\sum_{i=0}^{k} \binom{m}{i} p^i (1-p)^{m-i} \leqslant e^{-\frac{(mp-k)^2}{2mp}} .$$

For $k \in \mathbb{N}$ *with* $mp \leqslant k \leqslant m$ *we have equivalently that*

$$\sum_{i=k}^{m} \binom{m}{i} p^i (1-p)^{m-i} \leqslant e^{-\frac{(mp-k)^2}{2mp}} .$$

Lemma 3.20 (Amplification)**.** *For all polynomial-time computable functions* c *and* s *with* $e^{-q} \leqslant s, c \leqslant 1 - e^{-q}$*, gap* $c - s \geqslant 1/q$ *and* q *polynomial it holds that*

$$\begin{aligned} \mathrm{BPP}(c,s) &= \mathrm{BPP} & \mathrm{MA}(c,s) &= \mathrm{MA} \\ \mathrm{PromiseBPP}(c,s) &= \mathrm{PromiseBPP} & \mathrm{PromiseMA}(c,s) &= \mathrm{PromiseMA} \\ \mathrm{BQP}(c,s) &= \mathrm{BQP} & \mathrm{QCMA}(c,s) &= \mathrm{QCMA}\,. \end{aligned}$$

Proof. For the proof of the statement it is sufficient to show that any completeness and soundness parameters c and s obeying the above restrictions can be amplified to $1-e^{-r}$ and e^{-r} for an arbitrary polynomial r.

Let V be a $C(c, s)$ protocol for a problem A and C any of the complexity classes BPP, MA, PromiseBPP, PromiseMA, BQP or QCMA. The following is a $C(1-e^{-r}, e^{-r})$ protocol for A:

1. Simulate $m := 8q^2r$ copies of V.
2. If more than $k := m(c+s)/2$ original protocol executions lead to acceptance then accept, otherwise reject.

Note that for the simulation in step 1 a possible witness has to be copied initially m-times. In lemma 2.3 we discussed how the final acceptance measurement of a BQP or QCMA protocol can be simulated unitarily.

Completeness: Let $c' \geqslant c$ denote the acceptance probability of the protocol V. The new protocol rejects iff at most k measurements in step 2 output acceptance, which occurs with probability

$$\mathbb{P}_{\text{rej}} = \sum_{i=0}^{\lfloor k \rfloor} \binom{m}{i} (c')^i (1-c')^{m-i}.$$

Since $k \leqslant mc - \frac{m}{2q}$ is smaller than the mean mc' of the binomial distribution, the tail becomes larger by shifting the mean to the smaller value mc. This tail can then be bounded by Chernoff bound:

$$\begin{aligned}\mathbb{P}_{\text{rej}} &\leqslant \sum_{i=0}^{\lfloor k \rfloor} \binom{m}{i} c^i (1-c)^{m-i} \\ &\leqslant \exp\left(-\frac{(mc-k)^2}{2cm}\right) \\ &\leqslant \exp\left(-\frac{1}{c}\frac{m}{2(2q)^2}\right) \\ &\leqslant e^{-r}.\end{aligned}$$

Soundness: Let $s' \leqslant s$ denote the acceptance probability of the protocol V. Since $k \geqslant ms + \frac{m}{2q}$ we can bound the acceptance probability $\mathbb{P}_{\mathrm{acc}}$ of the new protocol by the same argumentation as before:

$$\begin{aligned}\mathbb{P}_{\mathrm{acc}} &\leqslant \sum_{i=\lceil k \rceil}^{m} \binom{m}{i} (s')^i (1-s')^{m-i} \\ &\leqslant \sum_{i=\lceil k \rceil}^{m} \binom{m}{i} s^i (1-s)^{m-i} \\ &\leqslant \exp\left(-\frac{(ms-k)^2}{2sm}\right) \\ &\leqslant e^{-r}.\end{aligned}$$

□

Amplification via repetition according to the above lemma works for MA, PromiseMA and QCMA protocols, since their witness can simply be copied. For general quantum states this is however not possible. The no-cloning theorem shows that copying arbitrary quantum states contradicts unitarity (see e.g. [4, §12.1.1]). An easy workaround to make amplification also work QMA is to expect as new witness polynomially many copies of the original witness. In the proof we then only have to deal with the possibility of receiving an entangled witness:

Lemma 3.21. *[Weak amplification] For all polynomial-time computable functions* c *and* s *with* $e^{-q} \leqslant s, c \leqslant 1 - e^{-q}$, *gap* $c - s \geqslant 1/q$ *and* q *polynomial it holds that*

$$\mathrm{QMA}(c, s) = \mathrm{QMA}\,.$$

Proof. Assume that we carry out m copies of a given $\mathrm{QMA}(c, s)$ protocol as in lemma 3.20 with the difference that the length of the received witness is increased by a factor m and interpreted as m copies of original witness. Clearly, the the new protocol has completeness at least $1 - e^{-r}$, since this acceptance probability can be achieved by an m-fold tensor product of the original witness.

To prove a soundness value of e^{-r} as in lemma 3.20 we have to rule out that entanglement between the m blocks of the witness increases the acceptance probability and that the new worst case witness is indeed a product state (and hence clearly the

m-fold tensor product of the original worst case witness). To see this, note that each of the m protocol simulations can be run on distinct registers. For a witness ρ the final measurement of any simulation block i has hence the same probability distribution as if the other registers were ignored and just the partial witness $\mathrm{tr}_{[m]\setminus\{i\}}(\rho)$ supplied. Consequently, any witness ρ achieves the same acceptance probability as the product witness

$$\mathrm{tr}_{2,3,\dots,m}(\rho) \otimes \mathrm{tr}_{1,3,\dots m}(\rho) \otimes \cdots \otimes \mathrm{tr}_{1,2,\dots,m-1}(\rho). \qquad \square$$

Note that there also exists the tool of *strong amplification* [18] for QMA which avoids a lengthening of the witness compared to the above presented *weak amplification* via parallel repetition. The strong amplification method bases on the fact that the final acceptance measurement changes the output state of the circuit only slightly because it is promised to be either close to an acceptance or a rejection state. The polynomially many simulations of the original protocol necessary for weak amplification can therefore simply be replaced by polynomially loops of circuit unitary, final measurement, inverse circuit unitary and initialization measurement, which only require one witness register of original size.

3.5 Reductions and complete problems

Complexity classes give a coarse categorization of problems' complexity. Reduction notions are a useful tool for a finer and also complexity class independent comparision of complexity. A problem A reducible onto a problem B is considered simpler as B since it can be decided easily having knowledge about B.

Definition 3.22. *A problem* A *is* Karp- *or* m-reducible *to a problem* B *(notation:* $A \leqslant_m^P B$*) iff there exists a polynomial-time computable function* $f : \Sigma^* \to \Sigma^*$ *such that*

$$x \in A_{\mathrm{yes}} \Rightarrow f(x) \in B_{\mathrm{yes}},$$
$$x \in A_{\mathrm{no}} \Rightarrow f(x) \in B_{\mathrm{no}}.$$

Definition 3.23. *A promise problem* A *is* Cook- *or* T-reducible *to a promise problem* B *(notation:* $A \leqslant_T^P B$*) iff* A *can be decided by a DTM of polynomial runtime with oracle* B *(the DTM has access to an oracle state that upon entering replaces every* $x\square$*,* $x \in B_{yes}$*, at the head position instantaneously by* 1 *and every* $x\square$*,* $x \in B_{no}$*, by* 0*).*

Note that an oracle is only allowed to be queried for elements of B_{yes} and B_{no}. In the case of promise problems one has to ensure that the DTM does not query the oracle for any non-promised inputs of B.

Lemma 3.24. $A \leqslant_m^P B \Rightarrow A \leqslant_T^P B$.

Proof. Let f be the polynomial-time computable function that reduces A to B. Then A can also be solved by a polynomial-time DTM first simulating the computation of f and then querying the B-oracle on the function output. □

Both introduced reduction notions form a pre-order obeying reflexivity and transitivity on the set of problems. For being a partial order the antisymmetric property is missing: Problems that can be reduced onto each other can still be different.

A special role is assumed by those problems that are the most difficult ones in a complexity class:

Definition 3.25. *A promise problem* A *is called* m-hard *(*T-hard*) for a complexity class* C *iff all problems in* C *can be Karp-reduced (Cook-reduced) to* A*. If* A *is a problem of* C *itself,* A *is called* m-complete *(*T-complete*) for the complexity class* C*.*

We denote the set of all m- and T-complete problems for C *by* $C\text{-}c_m$ *and* $C\text{-}c_T$*, respectively.*

Lemma 3.26. *Let* C *be any of the complexity classes* PromiseBPP, PromiseMA, BQP, QCMA *or* QMA *(*P, NP, BPP *or* MA*).*

If $B \in C$ *and* $A \leqslant_m^P B$ *for a (decision) problem* A*, then* $A \in C$*.*

Proof. A C-protocol for A is easily obtained by first computing the polynomial-time reduction function from A to B and then simulating the C-protocol for B on the function output. □

The above statement also holds for Cook reducibility and the complexity class P. But for other complexity classes we do not know an analogous result. For example the validity of the implication $A \leqslant_T^P B \in \mathrm{NP} \Rightarrow A \in \mathrm{NP}$ for decision problems A and B is considered as unlikely, since it would directly imply NP = co-NP and hence a collapse of the famous polynomial hierarchy. For an introduction to the polynomial hierarchy – a multiquantifier extension of the P-NP hierarchy – see any standard textbook such as [14, §3.2].

The above lemma is the reason why Karp reducibility is considered as standard reduction notion for complexity classes above P. When refering to a problem simply as "reducible" or "complete" this is meant with regard to Karp reducibility. Even for quantum complexity classes, for which it seems natural to allow a broader reduction notion of quantum polynomial time, it is usual to work with Karp complete problems as representatives of the classes. Exploiting the capability of quantum polynomial time reductions like in the equivalence proof of QPCP formulations in proposition 8.9 is rather the exception than the rule.

Actually, to our knowledge, this thesis contains the only other application of a quantum polynomial time reduction: In section 6.2 we show that the strictly QMA-intermediate problems of chapter 5 are complete under quantum polynomial time reductions for a so-called noisy QMA class. This is the reason why we give here a formal definition of quantum polynomial time computability for reduction functions. Similar to the error bound of BQP we require that the correct function output is obtained by an efficient quantum protocol with probability at least $\frac{2}{3}$:

Definition 3.27. *A function* $f : \Sigma^* \to \Sigma^*$ *is* quantum polynomial-time computable *iff there exists a polynomial-time generated family of quantum circuits* $(V_x)_{x \in \Sigma^*}$ *on* $z \geqslant |f(x)|$ *qubits with* z *polynomial such that a final measurement of the state* $V_x |0\rangle^{\otimes z}$ *in the computational basis gives the output* $0 \dots 0 f(x)$ *with probability at least* $\frac{2}{3}$.

Definition 3.28. *A problem* A *is* quantum polynomial-time reducible *to a problem* B *(notation:* $A \leqslant_m^{QP} B$*) iff there exists a quantum polynomial-time computable function* $f : \Sigma^* \to \Sigma^*$ *such that*

$$x \in A_{\mathrm{yes}} \Rightarrow f(x) \in B_{\mathrm{yes}},$$
$$x \in A_{\mathrm{no}} \Rightarrow f(x) \in B_{\mathrm{no}}.$$

Class	Complete problem	Reference
P	all problems with at least one yes- and no-instance	
NP	k-Satisfiability with $k \geqslant 3$	[19]
BPP	?	
MA	?	
PromiseBPP	Acceptance Ratio of PTMs (canonical)	[20]
PromiseMA	Stoquastic 6-Satisfiability	[21]
BQP	Quadratically Signed Weight Enumerator	[22]
QCMA	Ground State Connectivity	[23]
QMA	k-Local Hamiltonian with $k \geqslant 2$	[24, 25]

Table 3.1: Karp-complete problems for standard complexity classes.

Lemma 3.29. *If* $B \in C$ *and* $A \leqslant_m^{QP} B$*, then* $A \in C$ *for any class* $C \in \{\mathrm{BQP}, \mathrm{QCMA}, \mathrm{QMA}\}$*.*

Proof. Let $V = (V_x)_{x \in \Sigma^*}$ be a $C\left(\frac{8}{9}, \frac{1}{9}\right)$-protocol for the problem B. The C-protocol that first simulates the quantum polynomial time algorithm for $f(x)$ including the final measurement and then simulates V_y for the measured function value y decides A with completeness $\frac{2}{3} \cdot \frac{8}{9} = \frac{16}{27}$ and soundness $\frac{2}{3} \cdot \frac{1}{9} + \frac{1}{3} = \frac{12}{27}$. These values can be amplified to the usual values of $\frac{2}{3}$ and $\frac{1}{3}$. □

The previous lemma is the analogue of lemma 3.26 and hence the reason why quantum polynomial time reductions can be considered as an equally good canonical reduction notion for quantum complexity classes above BQP as Karp reductions. However, up to this day Karp reducibility is still the prevailing reduction notion even in quantum complexity theory.

Table 3.1 lists complete problems for the standard complexity classes introduced in section 3.3. The table just contains one examplary complete problem for each complexity class, though usually several complete problems are known. For QMA these are meanwhile several dozens; for NP several thousands. Remarkably, no complete (decision) problems are known for BPP and MA. This is one of the reasons why some computer scientists believe that BPP = P and MA = NP. Moreover, this justifies why rather PromiseBPP and PromiseMA should be considered as the proper randomized analogues of P and NP.

Let us close this section by mentioning that some problems A have such high logical or physical relevance that it is worth defining all problems reducible to A as new complexity class with an own name. The complexity class TIM $\subseteq$ QMA is such an example [26], which consists of all problems reducible to a restricted Local Hamiltonian problem of transverse Ising model form. The notion of completeness hence allows an alternative, non-computing-model based approach to define complexity classes.

3.6 The Satisfiability and the Local Hamiltonian problem

In this section we briefly introduce the two most important complete problems of NP and QMA: the Satisfiability problem of Boolean formulas (SAT) and the Local Hamiltonian problem (LH). They are considered as the canonical complete problems of NP and QMA, respectively, since they were the first discovered and the completeness of all other complete problems known today is usually proven via a reduction from SAT and LH.

A *Boolean formula* ϕ is a logical expression in binary variables x_i connected via the logical operations NOT ($\neg$), AND ($\wedge$, *conjunction*) and OR ($\vee$, *disjunction*), e.g.

$$\phi = (x_1 \vee x_2) \wedge \big((x_3 \wedge \neg x_1) \vee (\neg x_2 \wedge x_3)\big).$$

A *literal* is a variable x_i or a negated variable $\neg x_i$. An assignment of binary values to the variables x_i for that the Boolean formula evaluates to 1, is called a *satisfying assignment*, e.g. $\{x_1 \to 1, x_2 \to 0, x_3 \to 1\}$ is a satisfying assignment for the above formula ϕ.

A Boolean formula is of k-*conjunctive normal form (*k*-CNF)*, iff it is a conjunction of *clauses*, each corresponding to a disjunction of at most k literals. By applying the distributive and de Morgan's laws it is possible to transform any Boolean formula into an equivalent formula in 3-CNF. For example,

$$(x_1 \vee x_2) \wedge (\neg x_1 \vee \neg x_2) \wedge x_3$$

is a formula in 3-CNF equivalent to the above defined formula ϕ. For general Boolean formuls this conversion needs exponential time. However, it is also possible to convert a general Boolean formula ϕ in polynomial-time into a formula in 3-CNF that is satisfiable iff ϕ is satisfiable.

Definition 3.30. *The* Satisfiability problem (SAT*) is the decision problem whose yes-instances are the satisfiable Boolean formulas.*

The problem k-SAT *is the decision problem whose yes-instances are the satisfiable Boolean formulas in* k*-CNF.*

Theorem 3.31 (Cook / Cook-Levin Theorem)**.** SAT *and hence* 3-SAT *are* NP*-complete.*

Proof. Original proof by S. A. Cook in 1971 [19]. Independent proof by L. Levin in 1973 with English translation in [27]. □

The NP-membership of SAT and hence 3-SAT is trivial, since a Boolean formula can be evaluated in polynomial-time given an assignment as witness. The hardness proof instead is very sophisticated, since it has to map each tuple of NP protocol and input into a Boolean formula which is satisfiable iff the Turing machine accepts the input.

The natural QMA-complete problem and analogue of k-SAT is the so-called Local Hamiltonian problem (with this shortened terminology we refer to the special k-$\mathrm{LH}_{a,b}$ problem that is QMA-complete according to the proof of the quantum Cook-Levin theorem 3.34).

A Hamiltonian H is a hermitian operator representing the energy observable of a quantum system, i.e. a quantum system in state ρ has energy

$$\mathrm{tr}(H\rho).$$

The state with the lowest energy, i.e. the eigenstate of H with the lowest eigenvalue λ_0, is called the *ground state* of H; λ_0 the *ground state energy*. Eigenstates to higher energy eigenvalues are called *excited states*.

A Hamiltonian on n qubits is k-*local* iff it equals a sum of hermitian *interaction terms* that act non-trivially only on a register S of k qubits, i.e. are of the form $(A)_S \otimes (\mathbb{I}^{\otimes n-k})_{\bar{S}}$ for a k-qubit operator A.

For quantitatively comparable energy discussions we assume in this thesis that the form of k-local Hamiltonians always fulfills the promise of the following k-LH$_{a,b}$ problem. Note that a general k-local Hamiltonian can be brought into this form by a simple offset and rescaling.

Definition 3.32. *The k-Local Hamiltonian problem with low energy value a and high energy value $b > a$ (k-LH$_{a,b}$) is a promise problem with the promise that yes- and no-instances are local Hamiltonians*

$$H = \sum_{S \in \mathcal{C}} H_S$$

on n qubits whose interaction terms H_S described by polynomially many digits are positive semi-definite, obey $\|H_S\| \leqslant 1$ and act non-trivially only on k qubits described by the set $S \subseteq [n]$.

The ground state energy equals for yes-instances at most a; for no-instances at least b.

The difference $a - b$ is called the absolute energy gap*. The absolute energy gap divided by the number of interaction terms $|\mathcal{C}|$ is called the* relative energy gap*.*

Note that in slight abuse of our previous convention, we use the variable n in the context of k-local Hamiltonians to denote the number of qubits on that the Hamiltonian is defined instead of the length of its binary encoding. But this is irrelevant for any polynomially related quantities, since the Hamiltonian description is bounded by a polynomial in n and hence all quantities bounded by a polynomial in the input length are also bounded by a polynomial in the number of Hamiltonian qubits.

The classical k-SAT problem can be considered as a special k-LH$_{0,1}$ problem whose interaction terms are projections onto computational basis states. $H_S = (|y\rangle\langle y|)_S$

represents the clause C_S on the (qu)bits of S with $y \in \Sigma^k$ as unique non-satisfying assignment:

$$H_S = (|y\rangle \langle y|)_S \quad \longleftrightarrow \quad C_S = \bigvee_{i \in S} (\neg)^{y_i} x_i.$$

Analogously to the Cook-Levin proof establishing k-SAT as natural NP-complete problem, a k-Local Hamiltonian problem with inverse polynomial energy gap can be proven to be QMA-complete. The reduction proof of Alexei Kitaev has been written down neatly by [24] and is usually called the *quantum Cook Levin proof*. We will not repeat the whole reduction proof here, but restrict ourselves to show the QMA-membership for k-Local Hamiltonian problems with inverse polynomial energy gap using the ground state as witness:

Proposition 3.33. *The* k*-Local Hamiltonian problem with low energy value* a *and high energy value* b *such that* $a - b \geqslant 1/q$ *for a polynomial* q *is in* QMA*.*

Proof. For a promised input $H = \sum_{S \in \mathcal{C}} H_S$ denote by $(|\phi_j^S\rangle)_{j \in [2^k]}$ the orthonormal eigenvectors of the 2^k-dimensional operator H_S with eigenvalues α_j^S. Let β_j^S be the coefficients such that we can express the ground state $|\eta\rangle$ as

$$|\eta\rangle = \sum_{j \in [2^k]} \beta_j^S (|\phi_j^S\rangle)_S \otimes (|\psi_j^S\rangle)_{\bar{S}}$$

for normalized states $|\psi_j^S\rangle$.

For every $S \in \mathcal{C}$ we can compute efficiently a finite-dimensional unitary T_S that maps the orthonormal states $|\phi_j^S\rangle \otimes |0\rangle$ to the orthonormal states

$$|\phi_j^S\rangle \otimes \left(\sqrt{\alpha_j^S}\,|0\rangle + \sqrt{1 - \alpha_j^S}\,|1\rangle\right).$$

Consider now the QMA protocol that first picks an interaction term H_S of the input Hamiltonian H uniformly at random, then applies T_S to the qubits of S and an ancilla qubit $|0\rangle$ and finally measures the ancilla qubit in the computational basis and accepts iff $|1\rangle$ is measured. We can show that this decides the k-$\mathrm{LH}_{a,b}$ problem with completeness

$1 - \frac{a}{m}$ and soundness $1 - \frac{b}{m}$, $m = |\mathcal{C}| \in \mathcal{O}(n^k)$, and that hence k-$\mathrm{LH}_{a,b} \in$ QMA by amplification:

Having picked a specific $S \in \mathcal{C}$, the state before the final measurement of the above described protocol equals

$$\sum_{j \in [2^k]} T_S(|\eta\rangle \otimes |0\rangle) = \sum_{j \in [2^k]} \beta_j^S (|\phi_j^S\rangle)_S \otimes (|\psi_j^S\rangle)_{\bar{S}} \left(\sqrt{\alpha_j^S}\, |0\rangle + \sqrt{1 - \alpha_j^S}\, |1\rangle \right).$$

Hence, the average acceptance probability equals

$$\begin{aligned} \mathbb{P}_{\mathrm{acc}} &= 1 - \frac{1}{m} \sum_{S \in \mathcal{C}} \sum_{j \in [2^k]} |\beta_j^S|^2 \alpha_j^S \\ &= 1 - \frac{1}{m} \sum_{S \in \mathcal{C}} \langle \eta | H_S | \eta \rangle . \end{aligned}$$ □

Theorem 3.34 (Quantum Cook-Levin theorem). *There exist a and b with $a - b \geqslant 1/q$ for a polynomial q such that the k-Local Hamiltonian problem with low energy value a and high energy value b is* QMA*-hard.*

Proof. Original proof for 5-$\mathrm{LH}_{a,b}$ in [24]. Improvement to 2-$\mathrm{LH}_{a,b}$ in [25]. □

Although we do not present the proof of the Quantum Cook-Levin theorem, we want to describe at least the form of the reduction Hamiltonian and its ground state according to the original proof [24] to give an intuition how a QMA verifier is mapped to a 5-local Hamiltonian. Let $V = (V_x)_{x \in \Sigma^*}$ be the QMA$(1 - e^{-n}, e^{-n})$ verifier for an arbitrary QMA problem and let V_x be a circuit on z ancilla and n_w witness qubits with 1- and 2-local gates $V_x^1, V_x^2, \dots V_x^T$. The quantum Cook-Levin proof reduces a problem instance x onto the Hamiltonian H acting on a first register of $z + n_w$ qubit and a second clock register with

$$\begin{aligned} H &:= H_{\mathrm{in}} + H_{\mathrm{out}} + \sum_{t=1}^{T} H_{\mathrm{prop}}(t), \\ H_{\mathrm{in}} &:= (|1\rangle \langle 1|)^{\otimes z} \otimes \mathbb{I} \otimes |0\rangle \langle 0| , \\ H_{\mathrm{out}} &:= (\mathbb{I} - \Pi_{\mathrm{acc}}) \otimes \mathbb{I} \otimes |T\rangle \langle T| , \end{aligned}$$

$$H_{\text{prop}}(t) \coloneqq \frac{1}{2}\big(\mathbb{I} \otimes |t\rangle\langle t| + \mathbb{I} \otimes |t-1\rangle\langle t-1| - V_x^t \otimes |t\rangle\langle t-1| - (V_x^t)^\dagger \otimes |t-1\rangle\langle t|\big).$$

The Hamiltonian is 5-local since the gates V_x^t are at most 2-local and a unary encoding of the clock realizes the transformations $|t-1\rangle\langle t|$, $|t\rangle\langle t|$ and $|t\rangle\langle t-1|$ as 3-local. The reduction Hamiltonian has ground state energy at most $a = \frac{1}{T^{10}}$ if x is a yes-instance and at least $b = \frac{1}{4(T+1)^3}$ if x is a no instance.

The witness in the completeness case is the so-called *history state*

$$|\eta\rangle \coloneqq \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} V_x^t V_x^{t-1} \dots V_x^1 |0\rangle^{\otimes z} \otimes |\psi\rangle \otimes |t\rangle$$

with $|\psi\rangle$ the witness for that V accepts x.

As the canonical QMA-complete problem much research has been conducted into the Local Hamiltonian problem and its variants. So far the authors of [26] conducted the probably most extensive study of variants of the Local Hamiltonian problem and their complexity class classifications.

Chapter 4

Introduction to quantum coding

4.1 Quantum coding

This chapter summarizes mainly standard textbook material such as [4, §10].

Quantum codes allow to recover quantum information from disturbances by representing it in a robust form. Although nowadays the definition of a code is sometimes used in a wider sense, this section introduces the original concept of Knill and Laflamme [28] which is illustrated in figure 4.1 and demands a lossless forth and back transformation between the original state and its robust, encoded form. This implies that the *encoding* $\mathcal{E}$ is an isometric transformation (which is equivalent to a unitary transformation on the input and an ancilla). Only then a quantum channel and coisometric transformation $\mathcal{D}$, the *decoding*, exists such that $\mathcal{D} \circ \mathcal{E} = \mathrm{Id}$. The encoding maps states of an input Hilbert space into a so-called *code space* which is a subspace within a larger Hilbert space. This ensures that certain disturbances, which kick an encoded state out of the code space, can be detected and fixed by an error correction channel.

A bit counter-intuitively a quantum code is normally not defined via an en- and decoding but only via a code space. The reason is that the specific en- and decoding is irrelevant for the question if a code space can be corrected from a given noise. But

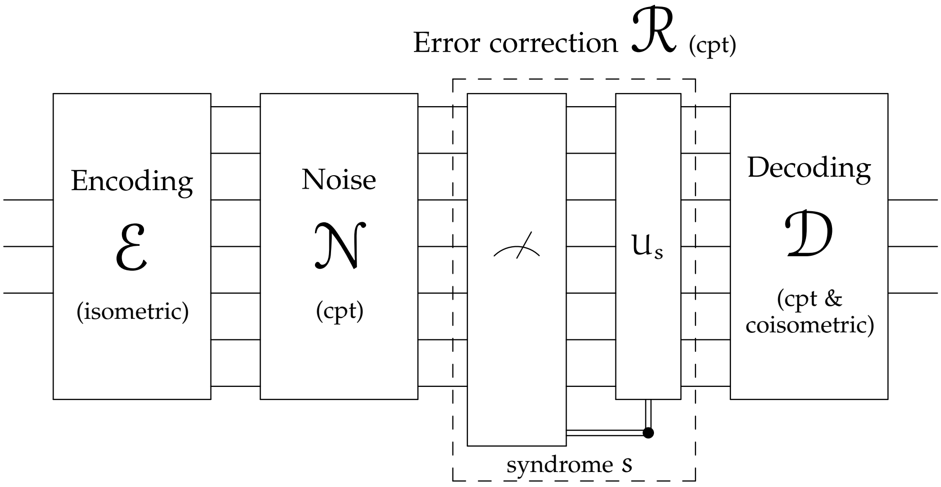


Figure 4.1: Concept of coding according to Knill and Laflamme.

note that the specific en- and decoding plays a role when we relieve the demand for perfect error correction in chapter 7 and study instead how much the concatenation of encoding, noise, error correction and decoding deviates from the identity channel.

Definition 4.1. *A* quantum code C *is defined by a* code space V_C *that is a subspace of a larger Hilbert space* $\mathbb{C}^{d'}$*. The* code space projection *is denoted by* P_C.

Definition 4.2. *Let* V_C, $\mathbb{C}^d$ *and* $\mathbb{C}^{d'}$ *be Hilbert spaces such that* $\mathbb{C}^d \simeq V_C \subseteq \mathbb{C}^{d'}$*. A map* $\mathcal{E} : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$ *with*

$$\mathcal{E}(\rho) = E\rho E^\dagger$$

for all $\rho \in \mathcal{L}(\mathbb{C}^d)$ *and an isometry* E *is called an* encoding *of the quantum code* C*. The adjoint operation* $\mathcal{D}(\rho) := E^\dagger \rho E$ *is called the corresponding* decoding.

Definition 4.3. *Given an encoding* $\mathcal{E} : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$ *we write*

$$|\bar{\phi}\rangle \langle\bar{\phi}| := \mathcal{E}(|\phi\rangle \langle\phi|)$$

for the encoding of a state $|\phi\rangle \in \mathbb{C}^{d'}$ *and call* $|\bar{\phi}\rangle$ *the* logical $|\phi\rangle$-state.

For qubit codes ($\mathbb{C}^d = \mathbb{C}^{2^k}$ *and* $\mathbb{C}^{d'} = \mathbb{C}^{2^N}$) *we denote analogously a specific operator that* *behaves on the code space like the Pauli operator* $\sigma \in \mathcal{P}^{\otimes k}$ *on the original space by* $\bar{\sigma}$ *and call it the respective* logical Pauli operator*:*

$$\bar{\sigma}\mathcal{E}(\rho)\bar{\sigma}^{\dagger} = \mathcal{E}\left(\sigma\rho\sigma^{\dagger}\right) \quad \forall \rho \in \mathcal{L}(\mathbb{C}^{d}).$$

Note that a choice of logical Pauli operators $\{\bar{\sigma} \mid \sigma \in \mathcal{P}^{\otimes k}\}$ fixes the encoding of a qubit code, because every state equals a unique linear combination of Pauli operators. If we choose instead an initial encoding, we may still have a true choice for the logical Pauli operators. In the case of stabilizer codes for example, which will be discussed in section 4.5, the degree of freedom in choosing logical Pauli operators is the multiplication by stabilizer operators.

4.2 Quantum error correction

Definition 4.4. *A noise channel* $\mathcal{N}$ *acting on a code space* V_C *is called* correctable *for the quantum code* C*, iff there exists an* error correction channel $\mathcal{R}$ *that reverses the action of* $\mathcal{N}$ *on the code space, i.e.*

$$\mathcal{R}\big(\mathcal{N}(|\psi\rangle\langle\psi|)\big) = |\psi\rangle\langle\psi| \quad \forall\, |\psi\rangle \in V_C.$$

The error correction theorem 4.8, originally phrased by Knill and Laflamme [28], will provide an easy check for the existence of an error correction channel as well a an explicit candidate in the case of existence. The explicit form shows that we can restrict our consideration to standard error correction channels which consist of a projective measurement followed by a unitary recovery operation conditioned on the measurement outcome called the *error syndrome*. The standard error correction channel is pictured in figure 4.1.

Definition 4.5. *A* standard error correction channel $\mathcal{R}$ *is defined via*

$$\mathcal{R}(\rho) = \sum_{i \in [r]} R_i P_i \rho P_i R_i^{\dagger}$$

with a projective measurement $\{P_i\}_{i\in[r]}$ *and a set of unitary* recovery operators $\{R_i\}_{i\in[r]}$.

Before proving the actual theorem, we recall the polar decomposition of linear operators and give a helpful equivalence for the later error correction condition.

Lemma 4.6 (Polar decomposition). *Every operator $A \in \mathcal{L}(\mathbb{C}^2)$ can be written in the form*

$$A = UJ = KU$$

with U unitary and J and K unique positive semi-definite operators defined by $J = \sqrt{A^\dagger A}$ and $K = \sqrt{AA^\dagger}$. If A is invertible, U is unique, too.

Proof. [4, theorem 2.3]. □

Lemma 4.7. *Let $\mathcal{N}$ be a quantum channel on a Hilbert space $\mathbb{C}^d$ and $P \in \mathcal{L}(\mathbb{C}^d)$. The following statements are equivalent:*

1. *For every set $\{N_i\}_{i\in[t]}$ of Kraus operators for $\mathcal{N}$ there is a hermitian operator $H \in \mathcal{L}(\mathbb{C}^d)$ such that $PN_i^\dagger N_j P = H_{ij}P$.*
2. *There is a set $\{N_i\}_{i\in[t]}$ of Kraus operators for $\mathcal{N}$ such that $PN_i^\dagger N_j P = H_{ij}P$ for a hermitian operator $H \in \mathcal{L}(\mathbb{C}^d)$.*
3. *There is a set $\{N_i\}_{i\in[t]}$ of Kraus operators for $\mathcal{N}$ such that $PN_i^\dagger N_j P = D_{ij}P$ for a diagonal operator $D \in \mathcal{L}(\mathbb{C}^d)$.*

Proof. The equivalences are a simple consequence from the unitary freedom auf Kraus operators stated in lemma 2.11: Given a set $\{N_i\}_{i\in[t]}$ of Kraus operators for a quantum channel $\mathcal{N}$ all other valid sets, possibly padded by zero operators, are exactly those of the form $\{M_i\}_{i\in[t]}$ with $M_i = \sum_{k\in[t]} U_{ki}N_k$ for a unitary $U \in \mathcal{U}(\mathbb{C}^d)$.

Let the operator set $\{N_i\}_{i\in[t]}$ fulfill the condition $PN_k^\dagger N_l P = H_{kl}$ for a hermitian $H \in \mathcal{L}(\mathbb{C}^d)$. Then the operators from the set $\{M_i\}_{i\in[t]}$ fulfill:

$$\begin{aligned} PM_i^\dagger M_j P &= \sum_{k,l\in[t]} U_{lj}U_{ki}^* PN_k^\dagger N_l P \\ &= \sum_{k,l\in[t]} U_{lj}(U^\dagger)_{ik} H_{kl} P \\ &= (U^\dagger H U)_{ij} P. \end{aligned}$$

The equivalences hold since $U^\dagger H U$ is always hermitian and even diagonal for U the diagonalizing unitary of H. □

Theorem 4.8 (Error correction theorem)**.** *A noise channel $\mathcal{N}$ is correctable for a quantum code C iff it has a set of Kraus operators $\{N_i\}_{i\in[t]}$ such that*

$$P_C N_i^\dagger N_j P_C = D_{ij} P_C \quad \forall i, j \in [t]$$

for a diagonal operator D.

For $i \in [t]$ let $R_i^\dagger$ be the unitary from a polar decomposition of $N_i P_C$ and $P_i := R_i^\dagger P_C R_i$. Extend the sets $\{R_i\}_{i\in[r]}$ and $\{P_i\}_{i\in[r]}$ by arbitrary unitaries and projections such that the completeness relation $\sum_{i\in[r]} P_i = \mathbb{I}$ is fulfilled. Then error correction is achieved by the standard error correction channel $\mathcal{R}$ with recovery operators $\{R_i\}_{i\in[r]}$ and projection operators $\{P_i\}_{i\in[r]}$.

$\mathcal{R}$ also corrects any channel whose Kraus operators are linear combinations of the $\{N_i\}_{i\in[t]}$.

Proof. *"$\Longrightarrow$".* To prove necessity assume that $\mathcal{R}$ is a valid quantum channel with Kraus operators $\{V_i\}_{i\in[r]}$ correcting the channel $\mathcal{N}$ with Kraus operators $\{M_j\}_{j\in[t]}$ on all states of the code space, i.e. for all quantum states ρ

$$\sum_{k\in[r]} \sum_{j\in[t]} V_k M_j P_C \rho P_C M_j^\dagger V_k^\dagger = P_C \rho P_C.$$

Hence, the quantum operation given by the Kraus operators $\{V_k M_j P_C\}_{k\in[r], j\in[t]}$ is identical to the one given by the single Kraus operator P_C. Lemma 2.11 on the unitary equivalence of Kraus operators implies therefore the existence of complex numbers c_{kj} such that $V_k M_j P_C = c_{kj} P_C$. With this

$$\begin{aligned} P_C M_i^\dagger M_j P_C &= \sum_{k\in[r]} P_C M_i^\dagger V_k^\dagger V_k M_j P_C \\ &= \sum_{k\in[r]} c_{ki}^* c_{kj} P_C \\ &= H_{ij} P_C \end{aligned}$$

with $H_{ij} := \sum_{k\in[r]} c_{ki}^* c_{kj}$ the matrix elements of a hermitian operator. According to lemma 4.7 this relation is equivalent to the desired diagonal relation.

"$\Longleftarrow$". To show sufficiency assume that $P_C N_i^\dagger N_j P_C = D_{ij} P$ is fulfilled for a diagonal operator D and a set $\{N_i\}_{i\in[t]}$ of Kraus operators for $\mathcal{N}$. With the help of the polar decomposition we can simplify $N_i P_C = R_i^\dagger \sqrt{P_C N_i^\dagger N_i P_C} = \sqrt{D_{ii}} R_i^\dagger P_C$. The consequence

$$N_i P_C N_i^\dagger = D_{ii} R_i^\dagger P_C R_i$$

shows us that the transformation of the code space projector P_C by N_i is proportional to the transformation by the unitary $R_i^\dagger$. Hence, depending on N_i the code space is transformed into subspaces characterized by the projections $P_i = R_i^\dagger P_C R_i$ fulfilling

$$P_i P_j \propto N_i \left(P_C N_i^\dagger N_j P_C\right) N_j^\dagger \propto \delta_{ij} N_i P_C N_j^\dagger.$$

Since the subspaces are orthogonal, we can distinguish them by a syndrome measurement described by projection operators $\{P_i\}_{i\in[r]}$ with P_i for $i \in [r]\backslash[t]$ choosen such that the completeness relation $\sum_{i\in[r]} P_i$ is satisfied. After the subspace check recovery is carried out by R_i, hence the desired recovery operation has the explicit form:

$$\mathcal{R}(\rho) = \sum_{i\in[r]} R_i P_i \rho P_i R_i^\dagger$$

with R_i arbitrary unitaries for $i \in [r]\backslash[t]$.

Extended correction: It remains to prove that the quantum channel $\mathcal{R}$ also corrects a quantum channel $\mathcal{M}$ with Kraus operators $\{M_i\}_{i\in[m]}$ that equal linear combinations

$$M_i = \sum_{j\in[t]} a_{ij} N_j, \quad a_{ij} \in \mathbb{C}.$$

Using previously derived equalities and the error correction condition we first simplify

$$\begin{aligned} R_k P_k M_i P_C &= \left(R_k P_k R_k^\dagger\right) R_k M_i P_C \\ &= \sum_{j\in[t]} a_{ij} (P_C R_k) N_j P_C \\ &= \sum_{j\in[t]} a_{ij} \frac{1}{\sqrt{D_{kk}}} P_C N_k^\dagger N_j P_C \\ &= a_{ik} \sqrt{D_{kk}} P_C. \end{aligned}$$

Now it is easy to check that $\mathcal{R}$ also corrects $\mathcal{M}$ acting on the code space:

$$\begin{aligned}\mathcal{R}\left(\mathcal{M}(P_C \rho P_C)\right) &= \sum_{k\in[r]} \sum_{i\in[m]} \big(R_k P_k M_i P_C\big)\rho\big(P_C M_i^\dagger P_k R_k^\dagger\big) \\ &= \left(\sum_{k\in[r]} \sum_{i\in[m]} |a_{ik}|^2 D_{kk}\right) P_C \rho P_C \\ &= P_C \rho P_C\end{aligned}$$

with the coefficient equalling 1 due to trace-preservation of $\mathcal{R}$ and $\mathcal{M}$. □

4.3 Quantum error detection

The next definition states what we operationally understand by error detection:

Definition 4.9. *A noise channel $\mathcal{N} : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^d)$ is* detectable *by a quantum code* C *with $V_C \subseteq \mathbb{C}^d$, iff there exists an* error detection channel $\mathcal{R} : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^d \otimes \mathbb{C}^2)$ *that for all $|\psi\rangle \in V_C$ fulfills*

$$\begin{aligned}\langle 0|\, \mathcal{R}(|\psi\rangle\,\langle\psi|)\, |0\rangle &= |\psi\rangle\,\langle\psi|\,, \\ \langle 0|\, \mathcal{R}\big(\mathcal{N}(|\psi\rangle\,\langle\psi|)\big)\, |0\rangle &\sim |\psi\rangle\,\langle\psi|\end{aligned}$$

with the projection $\langle 0| \ldots |0\rangle$ applying to the additional flag qubit.

An error detection channel introduces an additional flag qubit that indicates an error when found in state $|1\rangle$. In this case we impose no condition on the code register while we require it to contain the original code state if the flag qubit indicates no error by $|0\rangle$. To avoid that the trivial channel outputting the flag $|1\rangle$ with certainty is a valid error detection channel, the first equation in definition 4.9 requires additionally that the error detection channel has to output the no error flag $|0\rangle$, if the noise channel did not act on the code state.

Note that for an error correction channel we do not require any specific behaviour if the identity channel acts instead of the noise channel. That is also the reason why strictly speaking there exist noise channels which are correctable but not detectable for a given code, namely any noise that equals a single logical operation on the code space.

The next theorem shows that our definition of error detectibility is equivalent to the usual error detection condition used e.g. in [29]:

Theorem 4.10 (Error detection theorem). *A noise channel $\mathcal{N}$ with Kraus operators $\{N_i\}_{i\in[t]}$ is detectable for a quantum code C iff there exist $d_i \in \mathbb{C}$ such that*

$$P_C N_i P_C = d_i P_C$$

or, equivalently, iff for every choice of orthogonal code states $\{c_i\}_{i\in[K]}$ we have that

$$\langle c_k| N_i |c_l\rangle = d_i \delta_{kl}.$$

Error detection can be achieved by the channel $\mathcal{R}$ consisting of the projective measurement $\{P_C, \mathbb{I} - P_C\}$ and an additional flag qubit that stores the measurement outcome by $|0\rangle$ or $|1\rangle$, respectively. This error detection channel also detects any noise channel whose Kraus operators are linear combinations of the operators $\{N_i\}_{i\in[t]}$.

Proof. The equivalence of the two conditions can easily be seen by sandwiching the first equation by $\langle c_k| \ldots |c_l\rangle$.

"$\Longrightarrow$": Assume there exists an error detection channel $\mathcal{R}$ for the code C and the noise $\mathcal{N}$. Moreover, assume there exist j and $k \neq l$ sucht that $\alpha := \langle c_k| N_j |c_l\rangle \neq 0$.

Then the two properties of error detection will contradict each other: $\mathcal{R} \circ \mathcal{N}$ followed by the zero flag projection should rescale $|c_l\rangle\langle c_l|$, while on the other hand $|c_k\rangle\langle c_k|$ – into which $|c_l\rangle\langle c_l|$ is partially turned by $\mathcal{N}$ – should stay invariant under application of $\mathcal{R}$ followed by the zero flag projection. Mathematically, this contradition simply results by linearity and trace-preservation of $\mathcal{R}$:

$$\begin{aligned}\langle 0| \mathcal{R}\big(\mathcal{N}(|c_l\rangle\langle c_l|)\big) |0\rangle &= \langle 0| \mathcal{R}\left(|\alpha|^2 |c_k\rangle\langle c_k| + \text{terms of non-negative trace}\right) |0\rangle \\ &= |\alpha|^2 |c_k\rangle\langle c_k| + \text{terms of non-negative trace} \\ &\not\propto |c_l\rangle\langle c_l|.\end{aligned}$$

Hence, our assumption was wrong and $\langle c_k| N_i |c_l\rangle = d_{il}\delta_{kl}$.

It remains to prove that d_{il} is independent of l. For this assume $d_{ik} \neq d_{il}$. Replace the orthonormal basis states $|c_k\rangle$ and $|c_l\rangle$ by

$$|c'_k\rangle \coloneqq \frac{1}{\sqrt{2}}(|c_k\rangle + |c_l\rangle)$$
$$|c'_l\rangle \coloneqq \frac{1}{\sqrt{2}}(|c_k\rangle - |c_l\rangle).$$

Then

$$\langle c'_k| N_i |c'_l\rangle = \langle c'_k| \left(\frac{d_{ik} - d_{il}}{2} |c'_k\rangle + \frac{d_{ik} + d_{il}}{2} |c'_l\rangle\right) \neq 0$$

contradicts the already proven fact $\langle c'_k| N_i |c'_l\rangle = 0$. Hence, our assumption was wrong and d_{il} is in fact independent of l.

"$\Longleftarrow$": Assume $P_C N_i P_C = d_i P_C$ for all $i \in [t]$. Let $\mathcal{R}$ be the error detection channel defined in the theorem. Then clearly, $\langle 0| \mathcal{R}(|\psi\rangle\langle\psi|) |0\rangle = |\psi\rangle\langle\psi|$ for all $|\psi\rangle \in V_C$ and

$$\begin{aligned}\langle 0| \mathcal{R}\big(\mathcal{N}(|c_k\rangle\langle c_l|)\big) |0\rangle &= \sum_{i\in[t]} P_C N_i |c_k\rangle\langle c_l| N_i^\dagger P_C \\ &= \sum_{i\in[t]} |d_i|^2 |c_k\rangle\langle c_l|\end{aligned}$$

for all $k, l \in [t]$. Since the prefactor $\sum_{i\in[t]} |d_i|^2$ is independent of k and l and $\mathcal{R}$ and the flag qubit projection are linear operations, it follows that

$$\langle 0| \mathcal{R}\big(\mathcal{N}(|\psi\rangle\langle\psi|)\big) |0\rangle \sim |\psi\rangle\langle\psi| \quad \forall\, |\psi\rangle \in V_C.$$

Extended detection: $\mathcal{R}$ also detects any channel $\mathcal{M}$ with Kraus operators $\{M_i\}_{i\in[m]}$, $M_i = \sum_{j\in[t]} a_{ij} N_j$, $a_{ij} \in \mathbb{C}$, because for all $k, l \in [t]$ we have that

$$\begin{aligned}\langle 0| \mathcal{R}\big(\mathcal{M}(|c_k\rangle\langle c_l|)\big) |0\rangle &= \sum_{i\in[m]} \sum_{s,p\in[t]} \alpha_{is}\alpha_{ip}^* P_C N_s |c_k\rangle\langle c_l| N_p^\dagger P_C \\ &= \beta |c_k\rangle\langle c_l|\end{aligned}$$

with the prefactor

$$\beta \coloneqq \left(\sum_{i \in [m]} \sum_{s,p \in [t]} \alpha_{is} \alpha^*_{ip} d_s d^*_p \right)$$

being independent of k and l. □

4.4 Error sets, code distance and performance bounds

Due to the fact that the error correction channel of theorem 4.8 and the error detection channel of theorem 4.10 also correct and detect any error channel whose Kraus operators are linear combinations of the original elements, it is sufficient to consider the correctibility and detectability of so-called error sets:

Definition 4.11. *A set* $\{N_i\}_{i=[t]}$ *of linear operators* $\mathcal{L}(\mathbb{C}^d)$ *is called a* correctable (detectable) error set *for a code* C *with* $V_C \subseteq \mathbb{C}^d$ *iff the operators* $\{N_i\}_{i \in [t]}$ *fulfill the error correction condition of theorem 4.8 (error detection condition of theorem 4.10).*

Corollary 4.12. *An error set is correctable (detectable) by a quantum code* C *iff each channel whose Kraus operators equal linear combinations of the errors is correctable (detectable) by* C.

Lemma 4.13. *An error set* $\{N_i\}_{i \in [t]}$ *is correctable iff the error set* $\{N_j^\dagger N_i\}_{i,j \in [t]}$ *is detectable.*

Proof. Since $P_C N_i^\dagger N_j P_C = H_{ij} P_C$ for a complex matrix H implies directly that H is hermitian, the error correction condition of theorem 4.8 for $\{N_i\}_{i \in [t]}$ becomes equivalent to the error detection condition of theorem 4.10 for $\{N_j^\dagger N_i\}_{i,j \in [t]}$. □

Often we are seeking for qudit codes ($P_C \subseteq \mathbb{C}^{d^N}$) that are capable of correcting or detecting any error that only affects up to a certain number of qudits. This property is expressed by the notion of the code distance:

Definition 4.14. *The* distance δ *of a qudit quantum code C with* $V_C \subseteq \mathbb{C}^{d^N}$ *is the minimum number of qudits on that an operator* $N_i \in \mathcal{L}(\mathbb{C}^{d^N})$ *acts non-trivially that does not fulfill the quantum error detection condition of theorem 4.10.*

A quantum code of distance δ *encoding* K *orthogonal states into* N *qudits is called an*

$$(N, K, \delta)_d \text{ code.}$$

For qubit codes we omit the dimension index d.

Lemma 4.15. *For a quantum code of distance* δ *there exists an error correction (detection) channel that can correct* $\lfloor\frac{\delta-1}{2}\rfloor$ *(detect* $\delta - 1$*) arbitrary qudit errors but none that can correct (detect) more arbitrary qudit errors.*

Proof. The (non)-existence of error detection channels is a straightforward consequence from the definition of the distance. The statement about error correction follows by the argument of lemma 4.13. □

Due to the linearity of correctable and detectable errors, determining the code distance only requires checking the error correction and detection condition for basis operators of limited weight. Recall for example that the Pauli operators $\mathcal{P} = \{\mathbb{I}, X, Y, Z\}$ form an orthonormal basis of the single qubit operators. Consequently, a qubit code has at least distance δ iff the Pauli operators of weight less than δ fulfill the error detection condition 4.10.

After having introduced the notion of distance, it is natural to ask for the minimum number of physical qubits N for that there exists a code of distance δ capable of encoding K orthogonal states. This question cannot be answered exactly, but already a simple counting argument [30] leads to a lower, so-called *quantum Hamming bound* as well as an upper, so-called *quantum Gilbert-Varshamov bound*:

$$K \sum_{j=0}^{\lfloor\frac{\delta-1}{2}\rfloor} 3^j \binom{N}{j} \leqslant 2^N \leqslant K \sum_{j=0}^{\delta-1} 3^j \binom{N}{j}.$$

The simplicity of the derivation limits the validity of the bounds to pure codes. A code of distance δ is called pure iff distinct Pauli operators of weight at most $\lfloor\frac{\delta-1}{2}\rfloor$ map each code state to orthogonal states [31]. This property is generally stricter than non-degeneracy which only requires that such distinct Pauli operators map each code state to linearly independent states. Note, that for the later introduced stabilizer codes the notions of purity and non-degeneracy are equivalent due to lemma 4.34.

Regarding the Gilbert-Varshamov bound the restriction to pure codes can be seen as irrelevant, since it only gives us the additional information that the existing (N, K, δ) code fulfilling the bound is a pure one. It might just be the case that there exists an impure code with an even lower number of physical qubits. The quantum Hamming bound, on contrast, bears indeed the problem that it could be violated by impure codes, although such codes have not been discovered yet.

Fortunately, this shortcoming is overcome by another lower bound, the *quantum Singleton bound*, which holds for all quantum codes.

Lemma 4.16 (Quantum Singleton bound)**.** *For an* $(N, K, \delta)_d$ *quantum code it holds that*

$$N \geqslant 2(\delta - 1) + \log_d K.$$

Proof. Originally proven for qubits by [28]; adapted for qudits by [32]. □

The quantum Singleton bound proves that any qubit code encoding one logical qubit and capable of correcting any single qubit error (i.e. of distance at least 3) needs at least $2(\delta - 1) + 1 = 5$ physical qubits. A code meeting this bound is the famous 5-qubit stablizer code presented in section 4.5.5.

We quoted the higher dimensional version of the quantum Singleton bound, since for the construction in section 8.6 we require a code with a partition of the physical qubits into equally sized blocks such that each error constrained to a block has to be detectable. If we consider a block of r qubits simply as a qudit with $d = 2^r$, this translates into the search of a qudit code of distance 2 encoding at least $K = 2$ states. The quantum Singleton bound then tells us that we need at least $N = 3$ blocks.

4.5 Stabilizer codes

4.5.1 Definition and stabilizer groups

Stabilizer codes form the most famous class of quantum codes since they provide a strong algebraic structure that allows a simple and straightforward analysis. The crucial ingredient is the focus on the Pauli group as operator basis. A stabilizer code is defined via a subgroup of the Pauli group with pairwise commuting elements. Commuting hermitian operators have the advantage of having a common eigenvector basis.

Definition 4.17. *A* stabilizer group $S \subseteq \mathcal{P}_N$ *consists of commuting operators not including* $-\mathbb{I}$*. It defines a* stabilizer code $C(S)$ *via the code space* V_S *given by the common +1 eigenspace of all elements of* S*.*

Some literature allows the element $-\mathbb{I}$ in a stabilizer group, but this would lead to the trivial empty code space. Hence we restrict the definition of a stabilizer group to the one above. Moreover this saves us a lot of case differentiations and we can derive some useful properties like the following:

Lemma 4.18. *All elements* S_i *of a stabilizer group* S *are of the form* $\pm p$ *for a* $p \in \mathcal{P}^{\otimes N}$ *and hence hermitian.*

Proof. Assume S has an element S_j of the form $S_j = \pm i p$ for a $p \in \mathcal{P}^{\otimes N}$. Then $S_j^\dagger = -S_j$. But this leads to the contradiction $(S_j^\dagger)^2 = -S_j S_j^\dagger = -\mathbb{I} \in S$. □

Often it is useful to work just with a set of *generators* for the stabilizer group. Generators are elements of a group such that the group is given by all possible products between them. Remember from linear algebra that a set of m generators is called *independent* if no subset of $m-1$ elements is sufficient to generate the group.

Independent generators allow a simple characterization of the whole stabilizer group:

Proposition 4.19. *The stabilizer group* S *generated by* m *independent generators* $\{g_i\}_{i\in[m]}$ *contains* 2^m *elements, which correspond to the* 2^m *distinct, non-multiple, ordered products of generators.*

Proof. Since generators commute and square to identity, any product of them can be written ordered according to their indices and without double appearances. An ordered, non-multiple product of generators cannot be proportional to another one, because otherwise one generator can be expressed as a product of other generators and is therefore not independent. Hence, the cardinality of the stabilizer group equals the number of 2^m ordered, non-multiple products of generators. □

Corollary 4.20. *The code space projector of a stabilizer code* $C(S)$ *with independent generators* $\{g_i\}_{i\in[m]}$ *is given by*

$$P_C = \frac{1}{2^m}\sum_{S_i\in S} S_i = \frac{1}{2^m}\prod_{i\in[m]}(\mathbb{I}+g_i).$$

4.5.2 Check matrix representation and syndrome spaces

To compute the dimension of a stabilizer code space, it is convenient to represent Pauli group elements in a vector form omitting their prefactors and just indicating the positions of X-, Y- and Z-operators in their tensor product:

Definition 4.21. *The* vector representation *of a Pauli group element* $p \in \mathcal{P}_N$ *is a row vector* $\vec{p}^T$ *of length* $2N$ *over the field* $\mathbb{Z}_2$ *with a* 1 *in position j iff* p *contains a Pauli X- or Y-operator in the corresponding tensor product position and a* 1 *in position* $N + j$ *iff* p *contains a Pauli Z- or Y-operator in the* j*-th tensor product position.*

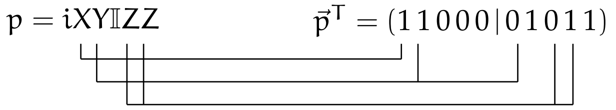


Figure 4.2: Vector representation of a Pauli group element p.

Lemma 4.22. *Two elements* $p_1, p_2 \in \mathcal{P}_N$ *commute iff* $\vec{p}_1^T\Lambda\vec{p}_2 = 0$ *with*

$$\Lambda = \begin{pmatrix} 0 & \mathbb{I} \\ \mathbb{I} & 0 \end{pmatrix}.$$

Proof. When multiplying out $\vec{p}_1^T \Lambda \vec{p}_2$, the sum contains a summand of 1 for each position $j \leqslant N$ in which the two Pauli group elements p_1 and p_2 contain different, non-identity and hence anti-commuting Pauli operators. □

Definition 4.23. *The* check matrix *of a generator set* $\{g_i\}_{i \in [m]}$ *for a stabilizer group on* N *qubits is the* $m \times 2n$ *matrix over* $\mathbb{Z}_2$ *whose rows equal the vector representations of the generators.*

Lemma 4.24. *Generators are independent iff the rows of their check matrix are linearly independent.*

Proof. This holds, since multiplying two Pauli group elements corresponds to adding their vector representations modulo 2. □

The previous lemma offers the possibility to extract a set of independent generators from any generating set of a stabilizer group by Gaussian elimination of the check matrix. By elimation steps, omitting zero rows and renumbering qubits (swapping columns) it is furthermore always possible to obtain a check matrix of independent generators in the so-called *normal form* [4, §10.5.7]:

$$\begin{array}{rccc|ccc l}
 & r & (N-k-r) & k & r & (N-k-r) & k & \text{columns} \\
r \text{ rows} & \mathbb{I} & A_1 & A_2 & B & 0 & C & \\
(N-k-r) \text{ rows} & 0 & 0 & 0 & D & \mathbb{I} & E &
\end{array} .$$

Later we will see that the independent generators $\{g_i\}_{i \in [m]}$ of a stabilizer group S play an important role in the error correction for the stabilizer code $C(S)$ since their simultaneous measurement, i.e. the projective measurement in the common eigenvector basis, will correspond to the projective measurement in their standard error correction procedure. The states that are common eigenvectors of all generators g_i to eigenvalues j_i form the *syndrome space* to the generator syndrome $j_1 \dots j_m$.

Lemma 4.25. *Consider a stabilizer code* $C(S)$ *on* N *qubits and an arbitary syndrome* $j \in \{-1, 1\}^m$ *for the independent generators* $\{g_i\}_{i \in [m]}$ *of the stabilizer group. Then there exists an operator* $E_j \in \mathcal{P}_N$ *such that* $E_j P_C E_j^\dagger$ *equals the projection onto the* j*-th syndrome space.*

Proof. Let $\vec{x} \in (\mathbb{Z}_2)^m$ be such that $(-1)^{x_i} = j_i$. Since the rows of the check matrix G are linearly independent there exists a vector $\vec{E}_j \in (\mathbb{Z}_2)^{2N}$ with $G\Lambda\vec{E}_j = \vec{x}$. According to lemma 4.22 the Pauli operator E_j corresponding to the vector representation $\vec{E}_i^T$ commutes with generator g_i iff $x_i = 0$. Consequently, the projection P_j onto the j-th syndrom space can be written as

$$P_j = \frac{1}{2^m} \prod_{i \in [m]} (\mathbb{I} + j_i g_i) = E_j P_C E_j^\dagger. \qquad \square$$

Lemma 4.26. *The projection onto the* j*-th syndrome space for a stabilizer code* $C(S)$ *with independent generators* $\{g_i\}_{i\in[m]}$ *is given by*

$$P_j = \frac{1}{2^m} \prod_{i \in [m]} (\mathbb{I} + j_i g_i) = \frac{1}{2^m} \sum_{S_i \in S} \eta(E_j, S_i) S_i$$

with $E_j \in \mathcal{P}_N$ *such that* $P_j = E_j P_C E_j^\dagger$.

Proof. According to proposition 4.19 every stabilizer corresponds to an ordered, non-multiple product of the generators $\{g_i\}_{i\in[m]}$. For $k \in \{-1,1\}^m$ let S_k denote the product of the generators g_i and J_k the product of the j_i for that $k_i = 1$. Multipliying out leads to $\prod_{i\in[m]}(\mathbb{I} + j_i g_i) = \sum_{k\in\{-1,1\}^m} J_k S_k$. Every S_k either commutes or anticommutes with E_j. Since $E_i |\psi\rangle$, $|\psi\rangle \in V_S$, is an eigenstate of P_j to the eigenvalue 1, it holdd that $J_k S_k E_j = E_j S_k$, i.e. $\eta(E_j, S_k) = J_k$. □

Proposition 4.27. *Let* S *be a stabilizer group on* N *qubits with* m *independent generators* $\{g_i\}_{i\in[m]}$. *Then* $\dim(V_S) = 2^{N-m}$.

Proof. Every generator syndrome space has the same dimension since for every syndrome j there exists a unitary $E_j \in \mathcal{P}_N$ such that $E_j P_C E_j^\dagger = P_j$ according to lemma 4.25. Different P_j are furthermore orthogonal and add up to the identity $\mathbb{I} = \sum_j P_j$. Consequently $2^N = 2^m \dim(V_S)$. □

Knowledge the above expression for the code space dimension it is easy to prove that a stabilizer group is maximal in that sense that there is no larger group stabilizing the same code space:

Lemma 4.28. *Let* $p \in \mathcal{P}_N$ *be an operator for which all elements of a stabilizer code space* V_S *are eigenstates to the eigenvalue* 1*. Then* $p \in S$.

Proof. Let $\dim(V_S) = 2^{N-m}$. Then the set $S^\circ \subseteq \mathcal{P}_N$ of all operators for which elements of V_S are eigenstates to the eigenvalue 1, has to be a group of cardinality at most 2^m. Since it has to contain S, it equals S. □

Since a stabilizer code space has a dimension of a power of 2, it can be considered as qubit space. The following notation is common in literature:

Definition 4.29. *A stabilizer code* C *of distance* δ *encoding* k *qubits into* N *qubits is called an*

$$[N, k, \delta] \text{ code.}$$

The notation is also used without stating the distance.

4.5.3 Normalizer

Definition 4.30. *The* normalizer $N(S)$ *of a subgroup* $S \subseteq \mathcal{P}_N$ *ist defined as the set of all elements* $p \in \mathcal{P}_N$ *such that* $psp^\dagger \in S$ *for all* $s \in S$.

Lemma 4.31. $S \subseteq N(S)$ *for any subgroup* $S \subseteq \mathcal{P}_N$.

Proof. Since S is a group, it contains $s_1 s_2 s_1^{-1} = s_1 s_2 s_1^\dagger$ for all $s_1, s_2 \in S$. □

Definition 4.32. *The* centralizer $Z(S)$ *of a subgroup* $S \subseteq \mathcal{P}_N$ *is defined as the set of all elements* $p \in \mathcal{P}_N$ *such that* $psp^\dagger = s$ *for all* $s \in S$.

Lemma 4.33. $N(S) = Z(S)$ *for a stabilizer group* $S \subseteq \mathcal{P}_N$.

Proof. Trivially, $Z(S) \subseteq N(S)$. An element $p \in N(S)$ either commutes or anti-commutes with any element $S_i \in S$. Hence $pS_ip^\dagger = \pm S_i \in S$. Since $-\mathbb{I} \notin S$, it only remains the option $pS_ip^\dagger = S_i$ and hence $p \in Z(S)$. □

Lemma 4.34. *Normalizer elements of a stabilizer code map code states to code states while any non-normalizer Pauli operator maps the code space to the orthogonal space.*

Proof. Since $N \in N(S)$ commutes with all stabilizers, recalling corollary 4.20 leads to

$$P_C N P_C = \left(\frac{1}{2^m} \prod_{k \in [m]} (\mathbb{I} + g_k) \right) N = N \left(\frac{1}{2^m} \prod_{k \in [m]} (\mathbb{I} + g_k) \right) = N P_C,$$

while $N \in \mathcal{P}_N \backslash N(S)$ implies that N anticommutes with at least one generator g_l of the code space resulting in

$$P_C N P_C = P_C N \left(\frac{1}{2^m} \prod_{k \in [m]} (\mathbb{I} + g_k) \right) = \underbrace{P_C (\mathbb{I} - g_l)}_{=0} N \left(\frac{1}{2^m} \prod_{\substack{k \in [m], \\ i \neq l}} (\mathbb{I} + g_k) \right) = 0. \quad \square$$

Since two errors differing by a normalizer element cannot be distinguished by an error correction procedure, the normalizer will play a central role in the stabilizer version of the error correction theorem 4.36 and will provide us with an easy expression for the code distance in corollary 4.37.

Usually normalizer elements obeying the respective commutation relations are chosen as logical Pauli operators for a stabilizer code. Clearly, stabilizer multiplication is a degree of freedom for the choice of logical Pauli operators that does not change the encoding of the code. Having a choice fixed, a product of stabilizer and logical Pauli operator is unique:

Lemma 4.35. *Let* $C(S)$ *be an* $[N, k]$ *stabilizer code with a fixed set of logical Pauli operators* $\{\bar{\sigma} \mid \sigma \in \mathcal{P}^{\otimes k}\} \subseteq N(S)$. *Then a product* $s\bar{\sigma}$ *uniquely determines* $\bar{\sigma}$ *with* $\sigma \in \mathcal{P}^{\otimes k}$ *and* $s \in S^{\pm}$ *with* $S^{\pm} := \{sS_i | s \in \{\pm 1, \pm i\}, S_i \in S\}$.

Proof. Assume $s_1\bar{\sigma}_1 = s_2\bar{\sigma}_2$ with s_1, $s_2 \in S^{\pm}$ and $\sigma \in \mathcal{P}^{\otimes k}$. Then $s_2^{\dagger} s_1 = \bar{\sigma}_2 \bar{\sigma}_1^{\dagger}$. The right hand side equals a logical operator times a prefactor. But the only logical operator that is an element of $S^{\pm}$ is the identity, hence, $\bar{\sigma}_2 = \bar{\sigma}_1$ and consequently $s_1 = s_2$. □

4.5.4 Error correction in the stabilizer formalism

The whole power of stabilizer codes shows off in the adaption of the error correction theorem 4.8. As discussed in section 4.4, the check if a code can correct arbitrary noise on t qubits can be restricted to Pauli errors on t qubits. In the case of stabilizer codes we will profit from their strength in handling Pauli operators to derive a simplified version of the error correction theorem.

Note that in literature the condition $E_i^\dagger E_j \notin N(S)\backslash S^\pm$ of the following theorem is often sloppily written as $E_i^\dagger E_j \notin N(S)\backslash S$ or justified by just providing the sufficiency statement.

Theorem 4.36 (Error correction for stabilizer codes). *An error set $\{E_i\}_{i\in[t]} \subseteq \mathcal{P}_N$ is correctable for a stabilizer code $C(S)$ iff*

$$E_i^\dagger E_j \notin N(S)\backslash S^\pm \quad \forall i, j \in [t].$$

Let $\{g_i\}_{i\in[m]}$ denote a set of independent generators of S. Error correction for any noise channel with linear combinations of the operators $\{E_i\}_{i\in[d]}$ as Kraus operators is achieved by a standard error correction channel with the syndrome measurement provided by the simultaneous measurement of the generators $\{g_i\}_{i\in[m]}$ followed by the recovery operation $E_i^\dagger$ with $E_i P_C E_i^\dagger$ equaling the projection onto the respective syndrome space. (If such an error does not exist, the respective syndrome will not occur and the recovery operation can be defined arbitrarily).

Proof. "$\Longrightarrow$": Assume the error set $\{E_i\}_{i\in[t]} \subseteq \mathcal{P}_N$ is correctable for a stabilizer code $C(S)$ and $E_i^\dagger E_j \in N(S)$ for an $i, j \in [t]$. Since $E_i^\dagger E_j$ commutes with all stabilizers, it also commutes with the code space projector P_C and according to theorem 4.8 and lemma 4.7 it holds for all $|\psi\rangle \in V_C$ that

$$E_i^\dagger E_j |\psi\rangle = P_C E_i^\dagger E_j P_C |\psi\rangle = H_{ij} |\psi\rangle$$

for a hermitian operator H. In other words, all code space states are eigenstates of $H_{ij}^\dagger E_i^\dagger E_j$ to the eigenvalue 1 with $H_{ij} \in \{\pm 1, \pm i\}$. According to lemma 4.28 this implies $E_i^\dagger E_j \in S^\pm$.

"$\Longleftarrow$": Consider a stabilizer code $C(S)$ and an error set $\{E_i\}_{i\in[t]}$ with $E_i^\dagger E_j \notin N(S)\backslash S^\pm$ for all $i, j \in [t]$. For $E_i^\dagger E_j \in S^\pm$ it holds that $P_C E_i^\dagger E_j P_C = H_{ij} P_C$ with $H_{ij} \in \{\pm 1, \pm i\}$ and $H_{ij} = H_{ji}^*$, while for $E_i^\dagger E_j \notin N(S)$ it holds that $P_C E_i^\dagger E_j P_C = 0$ according to lemma 4.34. Hence, the error correction condition of theorem 4.8 is fulfilled.

Error correction channel: If we reformulate the original error correction channel $\mathcal{R}$ of theorem 4.8 for the noise channel with Kraus operators $\{E_i/\sqrt{t}\}_{i\in[t]}$ (note that for Pauli group elements E_i these define indeed a valid channel), we obtain an error correction channel that can correct any channel whose Kraus operators are linear combinations of the errors $\{E_i\}_{i\in[t]}$.

Since $E_i P_C/\sqrt{t}$ already corresponds to a product of a unitary and a positive semidefinite operator, we can choose $R_i^\dagger = E_i$ for the unitary of the polar decomposition. After having measured that the disturbed code state lies in the support of one of the orthogonal projections $P_i = E_i P_C E_i^\dagger$ for $i \in [t]$ we thus apply R_i for recovery.

Let the string $j^i \in \{-1, 1\}^m$ indicate in position k if E_i anticommutes or commutes with the k-th generator. Then

$$P_i = E_i P_C E_i^\dagger = E_i \left(\frac{1}{2^m} \prod_{k\in[m]} (\mathbb{I} + g_k) \right) E_i^\dagger = \frac{1}{2^m} \prod_{k\in[m]} (\mathbb{I} + (j^i)_k g_k) = P_{j^i},$$

the projection onto the generator syndrome space with the syndrome j^i.

The extended set $\{P_i\}_{i\in[2^m]}$, comprising the projections onto all 2^m generator syndrome spaces, hence fulfills the specification for the projection operators of a standard error correction channel according to theorem 4.8. The remaining recovery operators R_i with $i \in [2^m]\backslash[t]$, can be chosen arbitrarily. □

According to definition 4.14 the distance of a code equals the minimum weight of an operator contradicting the error detection condition. Since we can restrict the minimum to the weight of Pauli operators we obtain:

Corollary 4.37. *The distance of a stabilizer code $C(S)$ equals the minimum weight over all elements of $N(S)\backslash S^\pm$.*

4.5.5 Examples of common stabilizer codes

Table 4.1 comprises three famous $[N, 1, 3]$ stabilizer codes, i.e. these codes encode one logical qubit into N physical qubits while being capable of correcting any single qubit error. The number of one logical qubit follows from the fact that the codes are defined by $N - 1$ independent generators (recall proposition 4.27). The distance of 3 can be computed after having determined the respective normalizers (recall corollary 4.37). The listed recovery operators are exactly those that achieve correction of any single qubit error.

By listing additionally a choice of logical $\bar{X}$ and $\bar{Z}$ operators the table also fixes the standard encoding for the codes. As mentioned before, the choice of encoding becomes relevant for the performance of the code in a noise setting beyond its perfect correctibility, e.g. in the setting of chapter 7 where concatenated coding is applied against probabilistic noise.

	5-qubit code	Steane 7-qubit code	Shor code
Independent	$XZZX\mathbb{I}$	$\mathbb{I}\mathbb{I}\mathbb{I}XXXX$	$ZZ\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}$
generators	$\mathbb{I}XZZX$	$\mathbb{I}XX\mathbb{I}\mathbb{I}XX$	$\mathbb{I}ZZ\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}$
	$X\mathbb{I}XZZ$	$X\mathbb{I}X\mathbb{I}X\mathbb{I}X$	$\mathbb{I}\mathbb{I}\mathbb{I}ZZ\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}$
	$ZX\mathbb{I}XZ$	$\mathbb{I}\mathbb{I}\mathbb{I}ZZZZ$	$\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}ZZ\mathbb{I}\mathbb{I}\mathbb{I}$
		$\mathbb{I}ZZ\mathbb{I}\mathbb{I}ZZ$	$\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}ZZ\mathbb{I}$
		$Z\mathbb{I}Z\mathbb{I}Z\mathbb{I}Z$	$\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}\mathbb{I}ZZ$
			$XXXXXX\mathbb{I}\mathbb{I}\mathbb{I}$
			$\mathbb{I}\mathbb{I}\mathbb{I}XXXXXX$
$\bar{X}$	$XXXXX$	$XXXXXXX$	$XXXXXXXXX$
$\bar{Z}$	$ZZZZZ$	$ZZZZZZZ$	$ZZZZZZZZZ$
Recovery	$\mathbb{I}, X_i, Y_i, Z_i$	$\mathbb{I}, X_i, Z_j, X_iZ_j$	$\{\mathbb{I}, X_1, X_2, X_3\} \cdot \{\mathbb{I}, X_4, X_5, X_6\}$
operators	$1 \leqslant i \leqslant 5$	$1 \leqslant i \leqslant 7$	$\cdot\{\mathbb{I}, X_7, X_8, X_9\}$
			$\cdot\{\mathbb{I}, Z_1Z_2Z_3, Z_4Z_5Z_6, Z_7Z_8Z_9\}$

Table 4.1: Common stabilizer codes.

The three listed stabilizer codes are each famous for a particular reason. The number N of physical qubits of the 5-qubit code matches the minimum number necessary for an $[N, 1, 3]$ code according to the quantum Singleton bound stated in lemma 4.16. The Steane 7-qubit code has the nice structure of a self-dual CSS code [4, §10.4.2]. A self-dual CSS code is defined by a classical linear code and inherits its error correction ability. Lastly, the Shor code can be regarded as concatenation of a bit flip and a phase flip code and hence has a very intuitive encoding and error correction algorithm [4, §10.2].

4.5.6 Encoding and decoding of stabilizer codes

Another advantage of stabilizer codes besides their structured form and simple error correction properties is their efficient en- and decoding. The standard network for encoding and decoding is described in [33, §4]. For an $[N, k]$ stabilizer code this network needs N qubits and not more than $N - k$ single qubit gates (Hadamard and Pauli operations) and $N(N - k)$ two qubit gates (controlled Pauli operations). Notice, that these parameters scale indeed efficiently in the number of logical and physical qubits.

For our purposes this is rather irrelevant, since in the applications of this thesis N qubits will always be encoded seperatedly by the same stabilizer code, e.g. the 5 qubit code. In this case the en- and decoding is efficient anyways, since we simply have to apply N times the same constant size decoding scheme. Furthermore, for the purposes of this thesis only a decoding scheme is needed. We present here the alternative decoding method by [33, §4], which only needs up to $2k(N - k + 1)$ one- and two-qubit gates and k ancilla qubits initialized in $|0\rangle^{\otimes k}$ which can store the decoded state:

Assume w.l.o.g. that the logical $\bar{Z}_i$ operations, $i \in [k]$, are tensor products of $\{\mathbb{I}, Z\}$. From the normal form one can see that this is always possible by stabilizer multiplication (note that the $\bar{Z}$ operators of the codes in table 4.1 also have this form). Clearly, the following equivalences hold:

$$\begin{aligned} & \text{the } i\text{-th logical qubit is in the state } |\bar{1}\rangle \\ \Leftrightarrow \quad & \text{the encoded state is an eigenstate of } \bar{Z}_i \text{ with eigenvalue } -1 \end{aligned}$$

$\Leftrightarrow$ the encoded state is a computational basis state with an odd number of qubits for that $\bar{Z}_i$ contains a Z operator in the state $|1\rangle$.

Therefore, a logical CNOT targeted at ancilla qubit i from the logical qubit i is implemented by the collection of CNOTs from every physical qubit j for that $\bar{Z}_i$ contains a Z operator. If we apply afterwards a controlled $\bar{X}_i$ on the encoded state controlled on the ancilla qubit, we turn the first logical qubit into the logical $|\bar{0}\rangle$ state. Applying this procedure for every logical qubit will give us the decoded state in the ancilla register disentangled from the logical $|\overline{0^k}\rangle$ state in the code register.

We now also see why encoding needs some additional work. In principle, we can apply the above described circuit backwards for encoding, but for this an initial preparation of the logical $|\overline{0^k}\rangle$ state is necessary.

Chapter 5

Existence of intermediate problems

5.1 Introduction

Despite being the most studied classes in quantum complexity theory it is still unkown if QMA is strictly larger than BQP or perhaps equal to it like it is unknown for the analogous classical pair NP and P. For an equality proof it is sufficient to focus on complete problems of the larger class while inequality implies the possibility that also so-called *intermediate problems* exist. These problems lie outside the smaller within the larger class but are not complete for it.

Studying intermediate problems may lead to a better understanding of the structure and relationship between complexity classes. Moreover, some intermediate problems have high practice relevance. Many cryptography schemes for example are based on NP-intermediate problems and not, as one may assume, on NP-complete problems, since those usually possess a too large ratio of easily solvable instances. One (probably) NP-intermediate problem used in cryptography is Factoring, which is also famous for being in BQP due to Shor's algorithm [1]. The word "probably" indicates that Factoring is only believed to be NP-intermediate due to a missing completeness proof.

The practically relevant problem Graph Isomorphism can instead be proven to be NP-intermediate under the condition that the polynomial hierarchy does not collapse to the second level. The reason is that the complement problem Graph Nonisomorphism is contained in the class AM (for a protocol see [34]), which is the highest class in the collapsing Arthur-Merlin hierarchy [35] and which can be placed within the second level Π_2 of the polynomial hierarchy (analogously to the Sipser-Gacs-Lautemann proof [36]).

Quantum complexity theory also offers some natural candidates for QMA-intermediate problems. Some are so natural that own class definitions have been introduced for them: the problem of transverse Ising model Hamiltonians defining the class TIM [26], the problem of Stoquastic Hamiltonians defining the class StoqMA [21], and of course the complete problems for QCMA such as Ground State Connectivity [23]. Note that there is no formal proof for the intermediateness of these problems, not even under a condition like in the case of Graph Isomorphism. Finding a "natural" QMA-intermediate problem under such a condition is an interesting open task. Note that the quantum Arthur Merlin hierarchy collapses as well (though to the third level QMAM [18] instead of the second level AM classically) and that recently the first approach towards a reasonable definition of a quantum polynomial hierarchy was made [37].

For both cases, quantum and classical, the question remains if the existence of intermediate probems can be proven without any condition besides QMA $\neq$ BQP and NP $\neq$ P, respectively. The positive answer for the classical case was given by *Ladner's theorem* [38] in the seventies. Actually, Ladner proves a whole hierachy of intermediate problems by constructing for every problem outside P a new problem that is strictly simpler but still outside P. Ladner formulates this as "there is no minimal problem above P".

About the same time as Ladner's result, *Schaefer's dichotomy theorem* [39] revealed the sligthly contrary fact that every naturally restricted SAT problem is either in P or NP-complete. Both statements are correct, since Ladner's intermediate problems simply do not have the form of Schaefer's "naturally" restricted SAT problems. In fact, Ladner's intermediate problems are rather unnatural and based on the specific binary encodings chosen for Boolean formulas and Turing machines whose large degrees of freedom are not reflected in any reasonable logical structure.

In quantum complexity theory the categorization of variants of the Local Hamiltonian problem by [26] can be regarded as a quantum analogue of Schaefer's dichotomy result. In contrast to this, no quantum version of Ladner's theorem has been formulated before our work [40], which we will cover in this chapter. Ladner's theorem has been generalized before, but only to complexity classes of decision problems including classes below P by an adapted reduction notion [41, 42, 43]. Moreover, the extension of the statement to two complexity classes and two problems, known as *uniform diagonalization theorem* [44, 45], implies besides Ladner's result also the undecidiability if a problem of the larger class is contained in the smaller class.

Unfortunately, also the formulation of the uniform diagonalization theorem only covers classes of decision problems, whereas quantum complexity classes such as BQP and QMA as well as their classical randomized counterparts PromiseBPP and PromiseMA are defined to contain promise problems. The main purpose of the promise is to guarantee that an algorithm of reasonable runtime can differentiate yes- and no-instances well enough, i.e. adheres the probabilistic error allowed by the definition of the complexity class. When restricting to decision problems without promises neither PromiseBPP, PromiseMA, BQP nor QMA are known to possess a complete problem. But in table 3.1 we saw that all these classes contain complete promise problems of physical or logical relevance.

In the beginning of complexity theory the concept of promise problems did not offer much advantage in studying classical, non-randomized classes such as P and NP. Because of this and due to the "untotal" property that algorithms for promise problems can have an arbitrary behaviour on some inputs, complexity theory avoided and therefore neglected promise problems for a long time. But with the introduction of randomized complexity classes and at the latest with the upcoming of quantum computing the concept of promise problems clearly deserves more respect.

The adaption of the uniform diagonalization theorem to randomized and quantum complexity classes requires two steps: In section 5.2 we first extend some necessary terminology originally defined in the context of decision problems to the context of promise problems like (total) decidability and recursive (re-)presentation and show that these properties are obeyed by standard randomized and quantum complexity classes.

Afterwards we can adapt the proof of the uniform diagonalization theorem to promise problems and their complexity classes in section 5.3.

In the final section 5.4 we present the standard implications of the uniform diagonalization theorem in terms of intermediate problems and undecidability results, which now also hold for a large variety of promise problem classes. Note that the results hold for any combinations of these classes, e.g. they also prove the existence of an infinite hierarchy of intermediate problems between BQP and PromiseBPP assuming BQP $\neq$ PromiseBPP.

5.2 A framework for promise problems

5.2.1 Extremal problems and closure properties of complexity classes

The uniform diagonalization theorem allows statements about the structure of protocol sets and not directly of problems. Since protocols of decision problem classes such as P and NP correspond to exactly one decision problem, this sublety is irrelevant, but it becomes relevant for classes of promise problems which are usually defined to consist of all *subproblems* decidable by a certain protocol type.

Definition 5.1. *A complexity class* C *that for every promise problem* $(A_{\text{yes}}, A_{\text{no}}) \in C$ *also contains every* subproblem $(A'_{\text{yes}}, A'_{\text{no}})$ *with*

$$\begin{aligned} A'_{\text{yes}} &\subseteq A_{\text{yes}} \\ A'_{\text{no}} &\subseteq A_{\text{no}} \end{aligned}$$

is called closed under promise restriction.

Clearly, the standard complexity classes of

$$\mathcal{C} := \{\text{PromiseBPP}, \text{PromiseMA}, \text{BQP}, \text{QCMA}, \text{QMA}\}$$

are closed under promise restriction.

Definition 5.2. *We call the problem with the smallest promise decided by a DTM or a* C*-protocol* M*,* $\mathsf{C} \in \mathcal{C}$*, the* extremal problem *of* M *and denote it by* $\mathsf{P}(\mathsf{M})$*.*

Accordingly, we call the decision problem decided by a P*- or* NP*-protocol* M *extremal and denote it by* $\mathsf{P}(\mathsf{M})$*.*

Definition 5.3. *We denote by* C^**,* $\mathsf{C} \in \mathcal{C}$*, the restriction of a complexity class* C *to its extremal problems.*

One might wonder why C^*, $\mathsf{C} \in \mathcal{C}$, is not used as the proper definition for the according randomized or quantum complexity class. Indeed, logically there is no reason to artificially demand a larger promise from a problem than necessary, but practically one usually starts by defining a logically or physically interesting problem and then aims at proving membership for this problem in a certain complexity class. These proofs often involve many implication arguments and approximations to finally show that an algorithm accepts with a sufficiently high or low probability. But this does usually not rule out that the algorithm also accepts some other, non-promised instances with the same high or low probability.

Even in the case of the Local Hamiltonian problem, in which the fundamental algorithm accepts with a probability that trivially relates to the promise on the energy gap (recall proposition 3.33), the final amplification in order to achieve the standard completeness and soundness parameters involves Chernoff bound. Hence, even the Local Hamiltonian problem is probably not an extremal problem according to its usual definition. But the advantage of its usual definition is that the promise on the ground state energy is simply physically describable.

We end this subsection by defining another closure property that applies to all standard complexity classes irrespective whether they consist of promise or decision problems: the closure under finite variations. The property bases on the notion of the *symmetric difference* of two problems. We see two reasonable possibilities to extend the existing definition for decision problems [45] to promise problem. We decide to define closure under finite variations via the wider notion of the two, the *total symmetric difference*, since this will become the relevant notion in the proof of the uniform diagonalization theorem.

Definition 5.4. *For promise problems* A *and* B *we define*

$$
\begin{aligned}
A\blacktriangle B &:= \{A_{\text{yes}} \cap B_{\text{no}}\} \cup \{A_{\text{no}} \cap B_{\text{yes}}\} && \textit{(symmetric difference)}\\
A\backslash B &:= \{A_{\text{yes}} \backslash B_{\text{yes}}\} \cup \{A_{\text{no}} \backslash B_{\text{no}}\} && \textit{(difference)}\\
A\triangle B &:= (A\backslash B) \cup (B\backslash A) && \textit{(total symmetric difference).}
\end{aligned}
$$

*We say "*A *equals* B almost everywhere (a.e.)*" iff* $A\triangle B$ *is finite.*

Note that the right side of the above expressions equals a set despite the fact that the problem on the left side corresponds to a pair of sets.

Obviously, for decision problems it holds that

$$A\blacktriangle B = A\backslash B = B\backslash A = A\triangle B,$$

while for general promise problems $A\backslash B \neq B\backslash A$ and the symmetric difference $A\blacktriangle B$ is only a subset of the total symmetric difference $A\triangle B$ as displayed in figure 5.1.

Definition 5.5. *A complexity class* C *(of decision problems) is* closed under finite variations (c.f.v.) *iff* $A \in C$ *implies* $B \in C$ *for every (decision) problem B that equals A almost everywhere.*

$$\begin{tikzcd}
 & A\triangle B \arrow[dl, dash] \arrow[dr, dash] & \\
A\backslash B \arrow[dr, dash] & & B\backslash A \arrow[dl, dash] \\
 & A\blacktriangle B &
\end{tikzcd}$$

Figure 5.1: Subset relations between the difference notions for promise problems.

5.2.2 Total decidability

Some standard literature such as [12] uses the notion of decidability only for decision problems. Those who use the notion in the context of promise problems (e.g. [13, 14]), agree on the one we gave in definition 3.6. The disadvantage of this definition is the arbitrary behaviour of the protocol on non-promised inputs. It is therefore reasonable to introduce a stricter version of decidability that we call *total decidability*:

Definition 5.6. *A promise problem* $A = (A_{\text{yes}}, A_{\text{no}})$ *is* totally decidable *iff there exists a DTM* M *such that*

$$\begin{aligned}
&\forall x \in A_{\text{yes}}\ M(x) = 1\\
&\forall x \in A_{\text{no}}\ M(x) = 0\\
&\forall x \in \Sigma^* \backslash (A_{\text{yes}} \cup A_{\text{no}})\ M(x) = 10.
\end{aligned}$$

Obviously total decidability implies decidability and the two notions are identical in the case of decision problems. There are reasonable examples for both promise problems that are totally decidable and promise problems that are decidable but not totally decidable:

Example 5.7. *The promise problem* $A = (A_{\text{yes}}, A_{\text{no}})$ *with*

$$\begin{aligned}
A_{\text{yes}} &:= \{M \mid \textit{the DTM } M \textit{ has even runtime for its Gödel number as input}\}\\
A_{\text{no}} &:= \{M \mid \textit{the DTM } M \textit{ has odd runtime for its Gödel number as input}\}
\end{aligned}$$

is decidable but not totally decidable.

Proof. Clearly, A is decidable by simply counting the runtime of the given machine.

The set of non-promised inputs consists of exactly those DTMs that do not halt on their own Gödel number as input. If A was totally decidable, one could hence decide the Halting problem. □

The purpose of the promise for the above problem is to avoid undecidable instances, which prevents the problem from being totally decidable. However, the original purpose of promise problems is to exclude instances for that a property check with a certain accuracy would exceed the runtime restriction of a complexity class. Take for example the Local Hamiltonian problem: In the standard protocol described in proposition 3.33 the promised gap on the ground state energy of the Hamiltonian directly relates to the acceptance probability or – if this is improved to obey the standard completeness and soundness parameter – to the runtime overhead caused by amplification.

Luckily we do not care about runtime when asking for total decidability, hence, the extremal problems of most standard complexity classes are totally decidable.

Lemma 5.8. *The extremal problem of a* PromiseBPP *or* PromiseMA *protocol is totally decidable.*

Proof. The extremal problem of a PromiseBPP-machine is totally decided by a DTM that simulates all branches of the PTM (which always halt per definition) and checks the fraction of accepting branches. In case of a PromiseMA-machine this algorithm simply has to be repeated for each of the possible 2^{n_w} witnesses. □

The above proof also holds for complexity classes based on a PTM with different requirements on the runtime r and completeness and soundness values c and s, as long as they are computable functions. Since the acceptance probability of a halting PTM is always an exactly computable rational number, the statement even holds for classes without a gap between completeness and soundness such as PP, where yes-instances are accepted with probability at least $\frac{1}{2}$ and no-instances with probability less $\frac{1}{2}$. These generalizations of runtime, completeness and soundness are also possible for the quantum analogue of the lemma, which we prove next.

Lemma 5.9. *The extremal problem of a* BQP, QCMA *or* QMA *protocol is totally decidable.*

Proof. By simulating the circuit generating DTM we obtain the Gödel number of a quantum circuit V_x on z ancilla and – in case of QCMA or QMA – on n_w witness qubits. We have to differentiate if the acceptance probability $\mathbb{P}_{\text{acc}}$ is at most s, at least c or in between with $s := \frac{1}{3}$, $c := \frac{2}{3}$ and

$$\mathbb{P}_{\text{acc}} := \langle 0^z | V_x^\dagger \Pi_{\text{acc}} V_x | 0^z \rangle$$

in the case of BQP,

$$\mathbb{P}_{\text{acc}} := \max_{y \in \Sigma^{n_w}} \left((\langle 0^z | \otimes \langle y |) V_x^\dagger \Pi_{\text{acc}} V_x (|0^z\rangle \otimes |y\rangle) \right)$$

in the case of QCMA and

$$\mathbb{P}_{\mathrm{acc}} := \text{highest eigenvalue of } Q := \langle 0^z | \, V_x^\dagger \Pi_{\mathrm{acc}} V_x \, | 0^z \rangle$$

in the case of QMA.

We recall that the generated quantum circuit just consists of H-, T- and CNOT-gates. In case of BQP the acceptance probability $\mathbb{P}_{\mathrm{acc}}$ equals hence a sum of products of elements from the field

$$\mathbb{Q}\left(\frac{1}{\sqrt{2}}, i\right)$$

(notice that the phase $e^{i\pi/4}$ of the T-gate can be written as $\frac{1}{\sqrt{2}}(1+i)$).

This finite field extension can be handled as 4-dimensional vector space W over the rational numbers with the abstract basis vectors

$$1, \; \frac{1}{\sqrt{2}}, \; i, \; \frac{i}{\sqrt{2}}.$$

A DTM can compute operations on the coefficients exactly by storing two integers and a sign for each rational number and it can carry out the finitely many different products of basis vectors abstractly. There is consequently a DTM that can compute $\mathbb{P}_{\mathrm{acc}}$ as a rational linear combination of the abstract vectors 1 and $\frac{1}{\sqrt{2}}$ (the imaginary vectors obviously vanish in the acceptance probability).

If the coefficient of the $\frac{1}{\sqrt{2}}$-vector vanishes and the coefficient of the 1-vector equals c or s, the DTM can accept or reject directly. Otherwise, the DTM can compute a monotonously decreasing upper bound and a monotonously increasing lower bound of the acceptance probability by approximating the square root with the babylonian / Heron method until two of the three cases $\mathbb{P}_{\mathrm{acc}} < s$, $\mathbb{P}_{\mathrm{acc}} > c$ and $\mathbb{P}_{\mathrm{acc}} \in]s, c[$ can be disclosed.

To totally decide the extremal problem of a QCMA circuit family the above algorithm simply has to be run for all possible witnesses $y \in \Sigma^{n_w}$.

In case of a QMA-problem the following equivalences hold:

$$\begin{aligned} s\,\mathbb{I}-Q \text{ positive semi-definite } &\Longleftrightarrow \mathbb{P}_{\mathrm{acc}} \leqslant s \\ c\,\mathbb{I}-Q \text{ not positive definite } &\Longleftrightarrow \mathbb{P}_{\mathrm{acc}} \geqslant c \end{aligned}$$

and consequently $\mathbb{P}_{\mathrm{acc}} \in \,]s, c[$ if none of the two conditions holds. Consequently, we have to argue that the positive semi-definiteness of $s\,\mathbb{I}-Q$ and the positive definiteness of $c\,\mathbb{I}-Q$ are decidable. Sylvester's criterion states the positive definiteness of a matrix is equivalent to the positivity of all its principal minors and positive semi-definiteness to the non-negativity of all its leading principal minors. It is simple to decide the positivity and non-negativity of minors (determinants of submatrices) by computing improving bounds in the vector space W as described above. □

As before the above lemma still holds when runtime, completeness and soundness of the complexity classes are changed to different computable functions. Recall from the discussion about the computability definition 3.3 that we restrict computable functions to algebraic number, since these do not only allow the computability of standard algebraic operations but also the comparision of two numbers, which is important for the above proof. The above proof also remains valid if the gate set is extended to any gates with algebraic matrix elements.

One case in which a complexity class crucially changes with the form of the gate set is a one-sided complexity class such as QMA_1 which equals QMA with $c := 1$. To achieve the perfect completeness a QMA_1 verifier is usually informally allowed to contain gates with matrix element from the field the problem is formulated in. But if one naively thought to allow any field that can be "described by words", one could construct a circuit whose extremal problem is not decidable at all. Consider e.g. a field containing a Chaitin's number whose i-th digit equals 1 if i is a yes-instance of the Halting problem and 0 otherwise. This is the reason why the authors of [46] advocate the algebraic numbers as largest reasonable field on which the matrix elements of gates in QMA_1 protocols should be defined. With this formal definition the extremal problem of a QMA_1 protocol is also totally decidable.

5.2.3 Recursive (re)presentation

The terminology “uniform diagonalization theorem” originates from the fact that the proof of the theorem constructs a problem outside two complexity classes similar to Cantor’s diagonal argument, i.e. by running over all inputs and all protocols of the two classes to construct a problem outside. The theorem will therefore only apply to complexity classes whose protocols are enumerable and the respective extremal problems totally decidable. These two properties together define a class as *recursively (re)presentable*, which we introduce below as an extension of the well-known notion for decision problems [45, 44]. In order not to overload the proofs we will show recursive (re)presentability again only for the classes P, NP, PromiseBPP, PromiseMA, BQP, QCMA and QMA, but the argumentation is adaptable to many natural complexity classes including those without completeness-soundness gap, different time- or space-restrictions or an interactive or multi-witness extension, but not necessarily to all. We will discuss below that BPP and MA are not known to be recursively presentable. The same holds for QMIP*, the class of quantum multiprover interactive proofs with unlimited entanglement, since it might even contain undecidable problems as we will discuss in section 8.5.

Definition 5.10. *A complexity class* C *is* recursively presentable, *iff it consists of all problems that are totally decidable by a DTM* M_i *from a computable series* $M_0, M_1, M_2 \ldots$ *of halting DTMs.*

We call a class C' recursively representable *iff it equals the closure of a recursively presentable class* C *under promise restriction.*

Computability of the series $M_0, M_1, M_2, \ldots$ means of course the computability of the function $i \to M_i$. Expressing it as a series just reflects better the enummerability property.

The property of recursive representability can obviously only be held by complexity classes of promise problems, since classes of decision problems are not closed under promise restriction by definition. Reversely, recursive presentability can not only apply to classes of decision problems but also to classes of promise problems, especially to those that are restricted to extremal problems.

Lemma 5.11. *The complexity classes* P, NP *and any* C^* *with* $C \in \mathcal{C}$ *are recursively presentable.*

Proof. The polynomials over $\mathbb{N}_0$ form a computable series $(p_i)_{i \in \mathbb{N}_0}$, since polynomials of a fixed degree and coefficient sum form a finite set and these sets are obviously enummerable. Hence, one can define a computable series $(M_i)_{i \in \mathbb{N}_0}$ for all P protocols with M_i – i interpreted as pair (j, k) – the P protocol simulating the DTM with Gödel number j up to runtime p_k and returning a default value if the DTM does not halt on 0 or 1 within this time.

A computable series $(M_i)_{i \in \mathbb{N}_0}$ of all DTMs that decide NP problems is realized by defining M_i, i interpreted as tuple (j, k, l), as the DTM that checks if the DTM with Gödel number j limited to time p_k accepts any witness of length p_l.

By interpreting Gödel numbers as encodings of probabilistic or quantum circuit generating Turing machines we can construct in the same manner a computable series of all C-protocols, $C \in \mathcal{C}$. Since the extremal problems of these protocols are totally decidable according to lemmata 5.8 and 5.9, we obtain a recursive presentation of the complexity class C^* by replacing the C-protocol in these series by the DTMs that totally decide the respective extremal problem. □

Corollary 5.12. *The complexity classes of* $\mathcal{C}$ *are recursively representable.*

Note that we do not know how to recursively present BPP and MA. Running over all polynomial-time PTMs is not conducive since we do not know how to decide if their extremal problems are decision problems. Another way to recursively present a class is via a complete problem as we see in the next lemma, but unfortunately also no complete problems for BPP and MA are known.

The next lemma allows to prove recursive (re)presentation for complexity classes that are defined via a complete problem like TIM, independently from a computing-model based definition. Hence, it is also irrelevant that we have no proper extremality definition for TIM problems.

Lemma 5.13. *For a decidable decision problem* A *the set*

$$A^{\geqslant^P_m} := \{\textit{decision problem } B \mid B \leqslant^P_m A\}$$

is recursively presentable and for a totally decidable promise problem A *that is not a decision problem the set*

$$A^{\geqslant^P_m} := \{\textit{promise problem } B \,|\, B \leqslant^P_m A\}$$

is recursively representable.

Proof. Similarly to the proof of lemma 5.11 we can define a computable series $(f_i)_{i \in \mathbb{N}_0}$ of all polynomial-time computable functions $\Sigma^* \to \Sigma^*$ by defining f_i as the function computed by the DTM M_i – i interpreted as pair (j, k) – simulating the DTM with Gödel number j up to runtime p_k and returning a default value if the DTM does not halt within this time.

Let M_A be the DTM that totally decides A. Then $(M_i)_{i \in \mathbb{N}_0}$ with $M_i(x) \coloneqq M_A\big(f_i(x)\big)$ is a recursive representation (presentation) of all (decision) problems B that are reducible to the (decision) problem A. □

The enumerability of reduction functions as well as of polynomial time DTMs with oracle state can also be used to prove a result the other way around: all problems of standard complexity classes that are more difficult than a problem A are also recursively presentable. The proof resembles the proof for Cook-complete decision problems by [44].

Lemma 5.14. *Let* C *be a recursively presentable complexity class c.f.v. and* $A \in C$ *be totally decidable. Then*

$$A_C^{\leqslant^P_m} := \{B \in C \,|\, A \leqslant^P_m B\},$$
$$A_C^{\leqslant^P_T} := \{B \in C \,|\, A \leqslant^P_T B\}$$

are recursively presentable.

Proof. Let $(M_i)_{i \in \mathbb{N}_0}$ be a recursive presentation of C, $(f_i)_{i \in \mathbb{N}_0}$ the computable series of all polynomial-time computable functions $\Sigma^* \to \Sigma^*$ and $(O_i)_{i \in \mathbb{N}_0}$ the computable series of all polynomial time DTMs with oracle state. Let M_A be the DTM that totally decides A.

For the recursive presentation of the set $A_C^{\leqslant_m^P}$ we define the DTM N_i, $i = (j,k)$, that checks for input x if all $y \in \Sigma^*$ with $|y| \leqslant |x|$ fulfill

$$\begin{aligned} y \in A_{\text{yes}} \quad &\Rightarrow \quad f_j(y) \in P(M_k)_{\text{yes}} \\ y \in A_{\text{no}} \quad &\Rightarrow \quad f_j(y) \in P(M_k)_{\text{no}}. \end{aligned}$$

If yes, it outputs $M_k(x)$, otherwise $M_A(x)$.

For the recursive presentation of the set $A_C^{\leqslant_T^P}$ we define the DTM N_i, $i = (j,k)$, that checks for input x if all $y \in \Sigma^*$ with $|y| \leqslant |x|$ fulfill

$$\begin{aligned} y \in A_{\text{yes}} \quad &\Rightarrow \quad \text{O_j with oracle $P(M_k)$ on input y only queries the oracle} \\ & \quad \text{for promised inputs and accepts} \\ y \in A_{\text{no}} \quad &\Rightarrow \quad \text{O_j with oracle $P(M_k)$ on input y only queries the oracle} \\ & \quad \text{for promised inputs and rejects.} \end{aligned}$$

If yes, it outputs $M_k(x)$, otherwise $M_A(x)$.

$(N_i)_{i \in \mathbb{N}_0}$ is a recursive presentation of $A_C^{\leqslant_m^P}$ ($A_C^{\leqslant_T^P}$) since $P(N_i) = P(M_k)$ if $P(M_k)$ for $i = (j,k)$ is a problem of C on that A can be m-reduced (T-reduced) while otherwise $P(N_i) = A$ almost everywhere. □

Corollary 5.15. *Let* C *be a recursively presentable complexity class c.f.v. with at least one totally decidable m-complete (T-complete) problem. Then* $C\text{-c}_m$ *(*$C\text{-c}_T$*) is recursively presentable.*

5.3 An extended uniform diagonalization theorem

Definition 5.16. *The* marked union $A \oplus A'$ *of two promise problems* A *and* A' *is defined as the promise problem* D *with*

$$\begin{aligned} D_{\text{yes}} &:= \{0x | x \in A_{\text{yes}}\} \cup \{1x | x \in A'_{\text{yes}}\} \\ D_{\text{no}} &:= \{0x | x \in A_{\text{no}}\} \cup \{1x | x \in A'_{\text{no}}\}. \end{aligned}$$

The uniform diagonalization theorem constructs for two complexity classes C and C' and two problems $A \notin C$ and $A' \notin C'$ another problem B which inherits the property not to belong to any of the two complexity classes while still being reducible to the marked union of A and A'. Note that the complexity bound on B is crucial, since just finding a problem outside two complexity classes is clearly trivial if it can be chosen arbitrarily more difficult.

For an efficient problem A such as the constant-no problem and a more diffcult problem A' such as k-LH, the marked union implies that B is reducible to the more difficult one. In the implication section 5.4 this choice together with $C := \text{QMA-c}_m$ and $C' := \text{BQP}$ will give us the extended Ladner theorem revealing B a QMA-intermediate problem.

Before we actually prove the uniform diagonalization theorem we first give the definition and an efficiency condition for the so-called *gap language* $G[r]$, which allows us later to mix the two problems stated in the uniform diagonalization theorem by restricting them to alternating intervals:

Definition 5.17. *Let* $r : \mathbb{N}_0 \to \mathbb{N}_0$ *be a computable function with* $r(n) > n$ *for all* $n \in \mathbb{N}_0$*. The* gap language generated by r *is defined as the set*

$$G[r] := \{x \in \Sigma^* \mid r^m(0) \leqslant |x| < r^{m+1}(0) \textit{ for } m \textit{ even}\}$$

with r^m *denoting the* m*-fold concatenation of* r.

Lemma 5.18. *If* $r : \mathbb{N}_0 \to \mathbb{N}_0$ *with* $r(n) > n$ *is time-constructible, then* $(G[r], \overline{G[r]}) \in P$.

Proof. Compute iteratively $r(0)$, $r^2(0)$, $r^3(0)$ ... like in the proof for time-constructible functions in lemma 3.5. Abort the iteration if the counter during the computation of $r^k(0)$ reaches $|x|$. Accept if $k - 1 = n$ is even, otherwise reject. Clearly, the computation of $r(n)$ is efficient in $r(n)$ and so is an aborted simulation in the final counter. Hence every iteration step is efficient in $|x|$. Since $r^m(0) \geqslant m$, the number of iteration steps is limited by $|x| + 1$ and the above algorithm is an efficient decision algorithm for $(G[r], \overline{G[r]})$. □

With the extended definitions and new notations introduced in the last section the following proof of the extended uniform diagonalization theorem resembles the original one limited to decision problems [45, 44].

Theorem 5.19 (Uniform diagonalization theorem)**.** *Let* C, C' *be complexity classes closed unter finite variations of which each is recursively presentable or recursively representable. Let* $A \notin C$, $A' \notin C'$ *be totally decidable promise problems. Then there exists a totally decidable promise problem* B *such that*

$$B \notin C \cup C' \textit{ and } B \leqslant_m^P A \oplus A'.$$

If A *and* A' *are decision problems or extremal for one of the complexity classes from* $\mathcal{C}$*, then so is* B*.*

Proof. Let $M_0, M_1, M_2, \ldots$ and $M_0', M_1', M_2', \ldots$ be recursive representations (presentations) for the complexity classes C and C', respectively. Due to $A \notin C$, every M_i does not (totally) decide correctly some instance of the problem A. The same holds for C' and A'. The construction idea for the new problem B is to mix A and A' such that B inherits such an instance for each M_i and M_i'.

To define a valid promise problem we mix A and A' by restricting them to alternating intervals via the previously defined gap language, i.e.

$$\begin{aligned} B_{\text{yes}} &\coloneqq (G[r] \cap A_{\text{yes}}) \cup (\overline{G[r]} \cap A'_{\text{yes}}) \\ B_{\text{no}} &\coloneqq (G[r] \cap A_{\text{no}}) \cup (\overline{G[r]} \cap A'_{\text{no}}). \end{aligned}$$

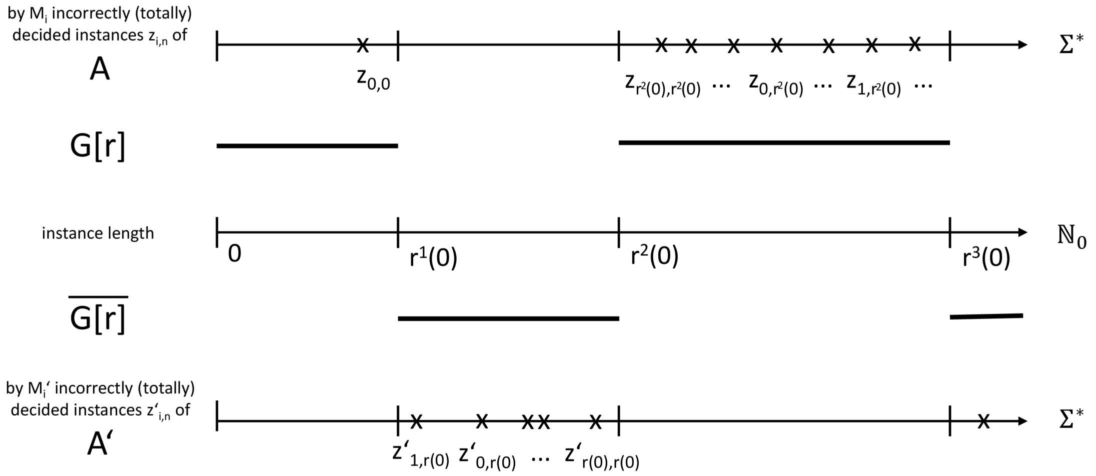
We want to define the function r such that the even intervals $G[r]$ contain for each M_i an incorrectly (totally) decided instance of A and the odd intervals $\overline{G[r]}$ for each M_i' an incorrectly (totally) decided instance of A' (see figure 5.2).

We achieve this by defining the function $q : \mathbb{N}_0 \to \mathbb{N}_0$,

$$q(n) := \max_{i \leqslant n}\{|z_{i,n}|\} + 1$$

with $z_{i,n} \in \Sigma^*$ the first string in the usual binary order such that $|z_{i,n}| > n$ and

$$z_{i,n} \in \begin{cases} A \backslash P(M_i) & \text{if } C \text{ is recursively representable} \\ A \triangle P(M_i) & \text{if } C \text{ is recursively presentable.} \end{cases}$$



by M_i incorrectly (totally) decided instances z_{i,n} of
A
Σ*
z_{0,0}
z_{r²(0),r²(0)} ··· z_{0,r²(0)} ··· z_{1,r²(0)} ···
G[r]
instance length
ℕ_0
0
r¹(0)
r²(0)
r³(0)
G[r] (overlined)
by M_i' incorrectly (totally) decided instances z'_{i,n} of
A'
Σ*
z'_{1,r(0)} z'_{0,r(0)} ··· z'_{r(0),r(0)}

Figure 5.2: Mixture of the problems A and A' covering all necessary elements to place the resulting problem B outside C and C'.

Notice that $z_{i,n}$ always exists. Otherwise,

$$\begin{aligned} A &\subseteq P(M_i) \text{ a.e.} \quad && \text{if } C \text{ is recursively representable} \\ A &= P(M_i) \text{ a.e.} \quad && \text{if } C \text{ is recursively presentable} \end{aligned}$$

and since C is closed under promise restriction in the first case and under finite variations in both cases this implies $A \in C$, which contradicts the hypothesis of the theorem.

The total decidability of A and the existence of $z_{i,n}$ imply the computability of q. Analogously, the function

$$q'(n) := \max_{i \leqslant n}\{|z'_{i,n}|\} + 1$$

is computable with $z'_{i,n} \in \Sigma^*$ defined accordingly for A' and M'_i.

We now choose our desired function r as the time-constructible function

$$r(n) \geqslant \max\{q(n), q'(n)\}$$

that exists according to lemma 3.5. The definitions of $q(n)$ and $q'(n)$ imply $r(n) > n$. Hence, the gap language $G[r]$ is well-defined. And since it is decidable, B is totally decidable.

Since $(G[r], \overline{G[r]}) \in P$ according to lemma 5.18, the function $f : \Sigma^* \to \Sigma^*$ with

$$f(x) := \begin{cases} 0x & \text{if } x \in G[r] \\ 1x & \text{if } x \in \overline{G[r]} \end{cases}$$

is a valid Karp-reduction from B to $A \oplus A'$.

$G[r] \in P$ also implies that B is extremal if A and A' are extremal for a complexity class $C \in \mathcal{C}$, since every class $C \in \mathcal{C}$ is capable of performing the polynomial-time decision algorithm for $G[r]$ as initial subroutine before simulating the algorithm for A or A'.

It only remains to show rigorously that indeed $B \notin C \cup C'$. For this assume $B \in C$. Then there exists an $i \in \mathbb{N}_0$ such that

$$\begin{aligned} B \backslash P(M_i) &= \varnothing \quad \text{if C recursively representable} \\ B \triangle P(M_i) &= \varnothing \quad \text{if C recursively presentable.} \end{aligned}$$

Let m be an even integer such that $n := r^m(0) \geqslant i$. Then there exists $z_{i,n}$ with $|z_i, n| \in [r^m(0), r^{m+1}(0)[$ with $z_{i,n} \in A \backslash P(M_i)$ ($z_{i,n} \in A \triangle P(M_i)$). Since $z_{i,n} \in G[r]$, this implies $z_{i,n} \in B \backslash P(M_i)$ ($z_{i,n} \in B \triangle P(M_i)$) which is impossible. Hence, our initial assumption is wrong and $B \notin C$. Analogously, it can be proven that $B \notin C'$. □

Notice that the most important step of the proof is indeed to choose the time-constructible function r instead of q and q' for defining the interval jumps of the gap language. If we had chosen the interval jumps at $q(n)$ and $q'(n)$, i.e. exactly at the highest necessarily contained instance, the determination of the interval that contains the input x could be far from efficient. One would have to compute all necessarily contained instances which requires a simulation of all machines M_i with $i \leqslant n$ for which we cannot fix a general polynomial runtime bound. Recalling the proof of lemma 3.5 we see that $r(n)$ is instead defined larger than the maximum of all necessarily contained instances *and* the runtime that it needs to compute them. So usually $r(n) \gg q(n)$ and the crosses indicating the necessarily contained instances in figure 5.2 should be drawn cumulated at the lower interval limits. This is also the reason why the proof technique is sometimes referred to as "delayed diagonalization" [47]. The check if the next interval limit lies above the input x can thus not only be answered positively by the output of the iterative computation of the necessarily contained instances but also when this computation exceeds a runtime of $|x|$, which is obviously efficient.

Note that it would actually be sufficient if each interval $[r^m(0), r^{m+1}(0)[$ only covered incorrectly (totally) decided instances of A (A') for the machines M_i (M'_i) with i from $r^m(0)$ down to $r^{m-2}(0)+1$, the lowest index that was not covered by the previous interval of same parity. But then the definition of the function r would get slightly more complicated, since the right interval bound $r^{m+1}(0)$ would not only depend on the left interval bound $r^m(0)$ but also on another previous interval bound.

5.4 Implications

This section briefly lists the most important implications of the uniform diagonalization theorem formulated for the classes QMA and BQP. This list is far from being comprehensive. While QMA and BQP can be substituted by many other pairs of recursively (re)presentable classes, we like to stress again that these implication are not known to hold for BPP and MA due to lacking knowledge about their recursive presentability and complete problems.

We first use the uniform diagonalization theorem to prove that the difference of standard complexity classes is not recursively presentable which then implies some undecidiability results.

Corollary 5.20. *The class* $\mathrm{QMA}^* \setminus \mathrm{BQP}^*$ *is not recursively presentable.*

Proof. If $\mathrm{BQP}^* = \mathrm{QMA}^*$, then $\mathrm{QMA}^* \setminus \mathrm{BQP}^*$ is empty and therefore by definition not recursively presentable. Let us hence consider the case $\mathrm{BQP}^* \subsetneq \mathrm{QMA}^*$. Then there exists a problem $A' \in \mathrm{QMA}^* \setminus \mathrm{BQP}^*$. Assume that $\mathrm{QMA}^* \setminus \mathrm{BQP}^*$ is recursively presentable. Clearly, $A := (\varnothing, \Sigma^*)$, A', $C := \mathrm{QMA}^* \setminus \mathrm{BQP}^*$ and $C' := \mathrm{BQP}^*$ fulfill the hypothesis of the uniform diagonalization theorem. The problem B constructed in the uniform diagonalization theorem is Karp-reducible to $A \oplus A'$ and hence in QMA^*. On the other hand, the uniform diagonalization theorem tells us that $B \notin C \cup C' = \mathrm{QMA}^*$ which is a contradiction. Hence, $\mathrm{QMA}^* \setminus \mathrm{BQP}^*$ is not recursively presentable. □

Corollary 5.21. *The classes* $\mathrm{QMA}^* \setminus \mathrm{QMA}^*\text{-}c_m$ *and* $\mathrm{QMA}^* \setminus \mathrm{QMA}^*\text{-}c_T$ *are not recursively presentable under the assumption that* QMA *does not equal* P *closed under promise restriction.*

Proof. This follows analogously to the proof of corollary 5.20 by substituting the problem $A' \coloneqq \mathrm{LH}^*$ (the extremal problem of the QMA protocol that decides the LH problem) and the respective complexity classes

$$\begin{aligned} C &\coloneqq \mathrm{QMA}^*\text{-}c_m & C' &\coloneqq \mathrm{QMA}^* \setminus \mathrm{QMA}^*\text{-}c_m, \\ C &\coloneqq \mathrm{QMA}^*\text{-}c_T & C' &\coloneqq \mathrm{QMA}^* \setminus \mathrm{QMA}^*\text{-}c_T. \end{aligned}$$

Notice that $\mathrm{QMA}^*\text{-}c_m$ and $\mathrm{QMA}^*\text{-}c_T$ are recursively presentable by corollary 5.15 and closed under finite variations since all m- and T-complete problems have infinitely many yes- and no-instances due to the assumption that QMA is strictly more powerful than the closure of P under promise restriction. □

The above results on their own might not seem very intriguing, but they reveal their whole power in the following two implications:

Corollary 5.22. *Given a* QMA *protocol it is undecidable if its extremal problem is in* BQP *assuming* $\mathrm{BQP} \subsetneq \mathrm{QMA}$.

Proof. Assume $\mathrm{BQP} \subsetneq \mathrm{QMA}$ and that it is decidable if the extremal problem of a QMA protocol is in BQP.

Consider the procedure of lemma 5.11 that constructs a recursive presentation of QMA^* by replacing each QMA protocol in the computable series of all QMA protocols by the DTM that totally decides its extremal problem. Before replacing it by a DTM decide if the extremal problem of a QMA protocol is in BQP. If this is the case, then first replace it by the QMA protocol deciding the LH^* problem. This way we obtain a recursive presentation of $\mathrm{QMA}^* \setminus \mathrm{BQP}^*$ which is a contradiction to corollary 5.20.

Hence, our assumption was wrong and it is undecidable if the extremal problem of a QMA protocol is in BQP. □

Corollary 5.23. *Given a* QMA *protocol it is undecidable if its extremal problem is m-complete (T-complete) under the assumption that* QMA *does not equal* P *closed under under promise restriction.*

Proof. Analogously. □

Originally [44] proved the above corollaries for combinations of the complexity classes NP P, PSPACE and PH as well as complement classes such as co-NP. We omit a corresponding version here, since complements are not very common to consider for classes of promise problems. One reason might be that many structural consequences that hold for complement classes of decision problems do not hold for promise problems. For example, a co-NP problem that turned out to be NP-complete under Cook reductions would imply NP = co-NP [48, 14] and hence the collapse of the polynomial hierarchy, whereas no analogous implication is known if a co-QMA problem turned out to be QMA-complete under Cook reductions.

Undecidability results form one branch of implications of the uniform diagonalization theorem. The second branch of implications – proving the existence of intermediate problems – is established by Ladner's simplification of the theorem:

Theorem 5.24 (Extended Ladner theorem)**.** *Let* A *be a problem in* $\mathrm{QMA}^* \setminus \mathrm{BQP}^*$. *Then there exists a problem* $B \in \mathrm{QMA}^* \setminus \mathrm{BQP}^*$ *with* $B \leqslant_m^P A$ *and* $A \nleqslant_T^P B$.

Proof. $C := \mathrm{BQP}^*$ is recursively presentable according to lemma 5.11 and so is $C' := \{D \in \mathrm{QMA}^* \mid A \leqslant_T^P D)$ according to lemma 5.14. These complexity classes and A and $A' := (\varnothing, \Sigma^*)$ hence fulfill the hypothesis of the uniform diagonalization theorem. Moreover, A and A' are extremal for QMA.

Consequently, there exists a problem $B \in \mathrm{QMA}^* \setminus \mathrm{BQP}^*$, $A \nleqslant_T^P B$ and $B \leqslant_m^P A \oplus A'$. The last condition simplifies to $B \leqslant_m^P A$, since the previous reduction function can be concatenated by the polynomial-time computable function that maps every string with an initial $0x$ to x and every 1 to a default no-instance of A (which has to exist due to $A \notin \mathrm{BQP}^*$). □

Corollary 5.25. *If* BQP $\subsetneq$ QMA*, then there exists an infinite hierarchy of intermediate problems between* QMA* *and* BQP* *(regarding both Karp- and Cook-reductions).*

The reader might tend to replace BQP* and QMA* in the extended Ladner theorem 5.24 and corollary 5.25 by BQP and QMA and indeed, the uniform diagonalization theorem allows this. But this statement could be meaningless, since we cannot rule out that the problem B equals the problem A up to some additional promise which makes it simpler, but such that the only QMA protocols deciding B are also those deciding A. Hence, thinking about it twice, it becomes clear that the above statements in terms of the "protocol classes" BQP* and QMA* are actually those that we pursued initially.

The extended Ladner theorem constructs the problem B as a mixture of the respective problem A and the constant-no problem A'. Thus, the hierarchy of intermediate problems between QMA and BQP constructed from the LH* problem are variants of the Local Hamiltonian problem with more and more yes-instances turned into no-instances. This is why Ladner's original proof for NP and the LH* analogue SAT is also called the method of "blowing holes into SAT" [47].

Does this descriptive property of the intermediate problems tell us something about the difficulty of specific Local Hamiltonian instances? Unfortunately, the criteria of kicking certain Local Hamiltonian instances out of the set of yes-instances is far from having any physical meaning. If a yes-instance of the Local Hamiltonian problem remains a yes-instance of the intermediate problem depends on the behaviour of specific Turing machines on specific inputs, determined by the chosen Gödel numbering for Turing machines and the chosen encoding for Local Hamiltonian instances. Due to the large degree of freedom in both encoding schemes, the holes blown into the Local Hamiltonian problem are rather artificial than physically meaningful.

We close by noting that the implications in this sections hold accordingly for other complexity classes:

Corollary 5.26. *The above corollaries 5.20, 5.22, 5.25 and the extended Ladner theorem 5.24 hold accordingly for every pair of the complexity classes* P, NP, PromiseBPP, PromiseMA, BQP, QCMA *and* QMA*.*

Corollaries 5.21 and 5.23 also hold for any of the these classes instead of QMA*.*

Chapter 6

QMA with noisy witness

6.1 Definition

In the last chapter we proved the existence of an infinite hierarchy of intermediate problems and therefore intermediate classes between QMA and BQP assuming QMA $\neq$ BQP. Unfortunately, the constructed intermediate classes are very abstract. Seeking a more physical approach, we observe that we can interpret the classes QCMA and BQP as a QMA variant with restricted access to the witness (QCMA: classical, BQP: none). It is therefore persuasive to generalize this restricted witness accesss and to define intermediate classes by introducing arbitrary noise channels on the witness.

We first define the most general case of such a *noisy QMA* class. The next section will show that this definition also covers classes for which the strictly QMA-intermediate problems of last chapter are complete under quantum polynomial time reductions. In section 6.3 and chapter 7 we will then concentrate on the more physical model of i.i.d. noise which is sufficient to express an infinite parameter interpolation between QMA – BQP and QMA – QCMA with the drawback that these noisy classes lack a proof of strict QMA-intermediateness.

Definition 6.1. *Let* $\mathcal{T} = (T_{n_w})_{n_w \in \mathbb{N}_0}$ *be a family of quantum circuits*

$$T_{n_w} : \mathcal{L}(\mathbb{C}^{2^{n_w}}) \to \mathcal{L}(\mathbb{C}^{2^{n'_w}})$$

with n'_w *polynomial in* n_w*. The complexity class* $\mathrm{QMA}_{\mathcal{T}}(c, s)$ *is the set of all problems* $A = (A_{\text{yes}}, A_{\text{no}})$ *for that there exists a polynomial-time generated familiy of quantum circuits* $V = (V_x)_{x \in \Sigma^*}$ *on* $z + n'_w$ *qubits with* z *polynomial and* $n_w = An^a$ *for an* A*,* $a \in \mathbb{N}$ *such that*

$$\forall x \in A_{\text{yes}} \; \exists \rho \in \mathcal{D}(\mathbb{C}^{2^{n_w}}) : \; \mathrm{tr}\left(\Pi_{\text{acc}} V_x \left(|0\rangle \langle 0|^{\otimes z} \otimes T_{n_w}(\rho) \right) V_x^\dagger \right) \geqslant c,$$
$$\forall x \in A_{\text{no}} \; \forall \rho \in \mathcal{D}(\mathbb{C}^{2^{n_w}}) : \; \mathrm{tr}\left(\Pi_{\text{acc}} V_x \left(|0\rangle \langle 0|^{\otimes z} \otimes T_{n_w}(\rho) \right) V_x^\dagger \right) \leqslant s.$$

We will use standard QMA terminology accordingly for $\mathrm{QMA}_{\mathcal{T}}$ protocols. The variable n_w always denotes the number of witness qubits before the channel application.

Despite the general definition above we will soon restrict to classes with more physical noise channels of *i.i.d.* form (*independently and identically distributed*), i.e. every witness qubit will be disturbed by the same single qubit channel T. Note that the next definition allows a dependence of the channel T on the number of witness qubits n_w. One can argue that only those channels T are really physical that are not influenced by the number of witness qubits. But the broader definition turns out to be useful for a first robustness and amplification result in section 6.3, for a formulation of the QPCP conjecture in chapter 8 and for the next chapter 7, in which the effective channels of concatenated coding will exhibit a dependence on n_w.

Definition 6.2. *For a single qubit channel* $T : \mathcal{L}(\mathbb{C}^2) \to \mathcal{L}(\mathbb{C}^{2^l})$ *we use the shortened notation*

$$\mathrm{QMA}_T \coloneqq \mathrm{QMA}_{(T_{n_w})_{n_w \in \mathbb{N}_0}}$$

for the the i.i.d. noise channels $T_{n_w} \coloneqq T^{\otimes n_w}$*.*

The single qubit channel T *is allowed to depend on the number* n_w *of witness qubits.*

Definition 6.3. *We define* $\mathrm{QMA}_{\mathcal{T}} \coloneqq \mathrm{QMA}_{\mathcal{T}}(2/3, 1/3)$ *and* $\mathrm{QMA}_T \coloneqq \mathrm{QMA}_T(2/3, 1/3)$*.*

In the chosen definition of noisy QMA the witness undergoes the channel in both the completeness and soundness case. We call this the *non-suspicious* definition of noisy

QMA, since the verifier can even trust in the soundness case that he receives a valid channel output. This definition allows a physical interpretation and a rephrasing of the famous QPCP conjecture in chapter 8 in terms of noisy QMA. But clearly, one could as well define noisy QMA in a *suspicious* manner in that the soundness performance has to hold for arbitrary witnesses. The following suspicious definition has been introduced in [49] and our own work [50]:

Definition 6.4. *The* suspicous *noisy QMA class* $\mathrm{QMA}_{\mathcal{T}}^{s}$ *is defined like the complexity class* $\mathrm{QMA}_{\mathcal{T}}$ *in definition 6.1 with the change that in case of a no-instance not just channel outputs but all witnesses on* n'_w *qubits fulfill the soundness condition, i.e.*

$$\forall x \in A_{no}\ \forall \rho \in \mathcal{D}\left(\mathbb{C}^{2^{n'_w}}\right): \ \mathrm{tr}\left(\Pi_{acc} V_x\big(\left|0\right\rangle\left\langle 0\right|^{\otimes z} \otimes \rho\big) V_x^{\dagger}\right) \leqslant s.$$

Note that the robustness results of QMA against decreasing noise in section 6.3 and against constant noise in chapter 7 would not be affected if we worked with the suspicious instead of the non-suspicious definition of noisy QMA.

The main advantage of the suspicious class $\mathrm{QMA}_{\mathcal{T}}^{s}$ is that it is a subset of QMA per definition, whereas this is unclear for general $\mathrm{QMA}_{\mathcal{T}}$ classes due to the weaker condition imposed on the verifier in the soundness case. But this does not prevent us to work with the non-suspicious variant as our main definition of noisy QMA, since all channel families we consider ensure the subset relation to QMA nevertheless due to their efficient simulability:

Lemma 6.5. $\mathrm{QMA}_{(T_{n_w})_{n_w \in \mathbb{N}_0}}(c, s) \subseteq \mathrm{QMA}(c, s)$ *if there exists a polynomial time Turing machine that for input* 1^{n_w} *outputs a quantum circuit* U_{n_w} *on* $z + n_w \geqslant n'_w$ *qubits, z polynomial, that simulates* T_{n_w}*, i.e. for all* $\rho \in \mathcal{D}\left(\mathbb{C}^{2^{n_w}}\right)$

$$T_{n_w}(\rho) = \mathrm{tr}_{[n_w + z - n'_w]}\left(U_{n_w}\big((\left|0\right\rangle\left\langle 0\right|)^{\otimes z} \otimes \rho\big) U_{n_w}^{\dagger}\right).$$

Proof. Let $(V_x)_{x \in \Sigma^*}$ be a $\mathrm{QMA}_{(T_{n_w})_{n_w \in \mathbb{N}_0}}(c, s)$ verifier of a problem A. Then A can be decided by the $\mathrm{QMA}(c, s)$ verifier whose generating Turing machine outputs the concatenation of the circuits U_{n_w} and V_x with distinct ancilla qubits and the last n'_w qubits of U_{n_w} serving as witness register for V_x. □

We end this section with some closing remarks on complete problems. Clearly, each complexity class $\mathrm{QMA}_{(T_{n_w})_{n_w \in \mathbb{N}_0}}(c, s)$ with $T_{n_w} : \mathcal{L}(\mathbb{C}^{2^{n_w}}) \to \mathcal{L}(\mathbb{C}^{2^{n'_w}})$ has a trivial complete problem:

> Given a polynomial-time generated family of quantum circuits $(V_x)_{x \in \Sigma^*}$ on $z + n'_w$ qubits, z polynomial, and an input $x \in \Sigma^*$, decide whether there exists a state $\rho \in \mathcal{D}(\mathbb{C}^{2^{n_w}})$ such that V_x accepts $|0\rangle\langle 0|^{\otimes z} \otimes T_{n_w}(\rho)$ with probability at least c (yes-instance) or if for all $\rho \in \mathcal{D}(\mathbb{C}^{2^{n_w}})$ V_x accepts $|0\rangle\langle 0|^{\otimes z} \otimes T_{n_w}(\rho)$ with probability at most s (no-instance).

One might wonder if besides this problem there exists a more sophisticated complete problem for noisy QMA classes. The most obvious idea is to study adaptions of QMA-complete problems, for example the variant of the Local Hamiltonian problem restricted to the question if there exists a channel output of high or low energy. This is a physically relevant problem and lies indeed in the respective noisy QMA class and unlikely in BQP for non-trivial channel families. Still, we cannot prove its $\mathrm{QMA}_{\mathcal{T}}$-hardness, since the witnesss in the reduction proof – the history state mentioned in section 3.6 – obeys a structure that is not reflected by the channel output (the channel would also disturbs the clock register).

For most QMA-complete problems a $\mathrm{QMA}_{\mathcal{T}}$ adaption fails in the hardness proof. One exception is the problem Non-Identity Check (sometimes also called misleadingly Identity Check) [51]. This QMA-complete problem is about deciding if a given unitary behaves far from identity on at least one state (yes-instance) or if it works almost as identity on all states (no-instance). A restriction of the question to computational basis states is a QCMA-complete problem [52]. It is not hard to see that the reduction proof works analogously for "Non-Identity Check on Channel Output States" as long as the channel family is able to output one pure state which is needed to represent the ancilla of a quantum circuit.

Unfortunately, for this problem we cannot show membership in $\mathrm{QMA}_{\mathcal{T}}$. For the QMA-complete problem Non-Identity Check and the QCMA-complete problem Non-Identity Check on Basis States the membership proof works by comparing the doubly supplied witness state with the respective state after the application of the unitary. Unfortunately, the SWAP test [53] needed to compare quantum states only works for pure states, but not for mixed states usually outputted by channels.

Hence, it remains an open qustion if there exists an interesting, non-trivial $\mathrm{QMA}_{\mathcal{T}}$-problem for a reasonable large range of channel families $\mathcal{T}$.

6.2 Applicability of the uniform diagonalization theorem and intermediateness of noisy QMA classes

We use the definition 5.2 of extremal problems and the notation $\mathrm{QMA}^*_{\mathcal{T}}$ of definition 5.3 accordingly for noisy QMA classes. In analogy to section 5.2 we briefly argue in this section that extremal problems of $\mathrm{QMA}_{\mathcal{T}}$ machines are totally decidable and that the class $\mathrm{QMA}^*_{\mathcal{T}}$ is recursively presentable for almost all channel families $\mathcal{T}$. As a consequence the uniform diagonalization theorem is applicable to these noisy QMA classes.

Furthermore, it can be shown that the class defined by all problems quantum polynomial-time reducible onto a $\mathrm{QMA}_{\mathcal{T}}$-intermediate problem with regard to BQP constructed by the extended Ladner theorem is again a noisy QMA class. Hence, as additional comment to the discussion at the end of last section, we stress that at least some noisy QMA classes have a complete problem besides the canonical one. Yet, these problems are rather artifical and complete only with regard to quantum polynomial time reductions. Note that up to our knowledge this is the only other occurence of quantum polynomial time reducibility in complexity theory besides the equivalence proof of standard QPCP formulations in proposition 8.9.

Lemma 6.6. *The extremal problem of a* $\mathrm{QMA}_{(T_{n_w})_{n_w \in \mathbb{N}_0}}$ *protocol is totally decidable if* T_{n_w} *can be simulated by the quantum circuit* U_{n_w} *of a computable family* $(U_{n_w})_{n_w \in \mathbb{N}_0}$.

Proof. Given a $\mathrm{QMA}_{(T_{n_w})_{n_w \in \mathbb{N}_0}}$ protocol $(V_x)_{x \in \Sigma^*}$, its extremal problem can be totally decided by a DTM similar to the one for QMA protocols defined in the proof of lemma 5.9 by simply assuming that the circuit V_x is preceded by the gates of U_{n_w} with distinct ancilla qubits and such that the channel output register of U_{n_w} is the witness register of V_x. □

Corollary 6.7. *The complexity class* $\mathrm{QMA}^*_{\mathcal{T}}$, $\mathcal{T} = (T_{n_w})_{n_w \in \mathbb{N}_0}$, *is recursively presentable and* $\mathrm{QMA}_{\mathcal{T}}$ *hence recursively representable if* T_{n_w} *can be simulated by the quantum circuit* U_{n_w} *of a computable family* $(U_{n_w})_{n_w \in \mathbb{N}_0}$.

Due to the above lemma and corollary and the existence of an extremal trivial complete problem, $\mathrm{QMA}_{\mathcal{T}}$ with $\mathcal{T}$ having the above stated simulability property fulfills the requirements of the uniform diagonalization theorem 5.19. Hence, also the implications of section 5.4 are applicable to these $\mathrm{QMA}_{\mathcal{T}}$ classes.

We end this section by proving that the intermediate problem between a class $\mathrm{QMA}_{\mathcal{T}}$ and BQP constructed by the extended Ladner theorem is quantum polynomial time complete for another noisy class $\mathrm{QMA}_{\tilde{\mathcal{T}}}$. Expressing the intermediate problem as element of $\mathrm{QMA}_{\tilde{\mathcal{T}}}$ is possible since the decision if a yes-instance of the $\mathrm{QMA}_{\mathcal{T}}$ complete problem is turned into a no-instance in Ladner's construction just depends on its length n and hence directly on the length $n_w = Cn^c$ of the witness which can be reflected in the structure of the channel family $\tilde{\mathcal{T}}$.

Lemma 6.8. *Let* $\mathrm{QMA}_{\mathcal{T}}$ *with* $\mathcal{T} = (T_{n_w})_{n_w \in \mathbb{N}_0}$ *be such that* T_{n_w} *can be simulated by the quantum circuit* U_{n_w} *of a computable family* $(U_{n_w})_{n_w \in \mathbb{N}_0}$. *Then there exists a problem* $B \in \mathrm{QMA}^*_{\mathcal{T}} \setminus \mathrm{BQP}$ *that is not* $\mathrm{QMA}_{\mathcal{T}}$*-complete but complete under quantum polynomial time reductions for the class* $\mathrm{QMA}^*_{\tilde{\mathcal{T}}}$, $\tilde{\mathcal{T}} = (\tilde{T}_{n_w})_{n_w \in \mathbb{N}_0}$ *with*

$$\tilde{T}_{n_w}(\rho) := \begin{cases} T_{n_w}(\rho) & \text{if } n \in G[r] \\ \left(\frac{\mathbb{I}}{2}\right)^{\otimes n'_w} & \text{if } n \in \overline{G[r]} \end{cases}$$

and $G[r]$ *the gap language from the proof of the uniform diagonalization theorem 5.19 with* $C := \mathrm{BQP}^*$, $C' := \mathrm{QMA}^*_{\mathcal{T}}\text{-}c_m$, A *the trivial complete problem for* $\mathrm{QMA}_{\mathcal{T}}$ *and* $A' := (\emptyset, \Sigma^*)$.

Proof. Membership: The problem B constructed in the proof of the uniform diagonalization theorem for the above choices of classes and problems lies in $\mathrm{QMA}^*_{\mathcal{T}} \setminus \mathrm{BQP}$, but is not $\mathrm{QMA}_{\mathcal{T}}$-complete. The $\mathrm{QMA}^*_{\tilde{\mathcal{T}}}$ protocol for B that first decides efficiently if the input lies in $G[r]$ or $\overline{G[r]}$ to carry out then the $\mathrm{QMA}^*_{\mathcal{T}}$ protocol for A or to reject immediately, can obviously also be regarded as valid $\mathrm{QMA}^*_{\tilde{\mathcal{T}}}$ protocol for B.

Hardness: Since $B \notin \mathrm{BQP}$, it has to contain at least one yes-instance b_{yes} and one no-instance b_{no}. For every problem $D = (D_{\mathrm{yes}}, D_{\mathrm{no}}) \in \mathrm{QMA}_{\tilde{\mathcal{T}}} \subseteq \mathrm{QMA}_{\mathcal{T}}$ there exists a polynomial-time computable function $f : \Sigma^* \to \Sigma^*$ reducing D to A. The function $f' : \Sigma^* \to \Sigma^*$ with

$$f'(x) := \begin{cases} f(x) & \text{if } |x| \in G[r] \\ b_{\mathrm{yes}} & \text{if } |x| \in \overline{G[r]} \text{ and } x \in D_{\mathrm{yes}} \\ b_{\mathrm{no}} & \text{if } |x| \in \overline{G[r]} \text{ and } x \in D_{\mathrm{no}} \end{cases}$$

then reduces D to B.

The function f' is quantum polynomial-time computable, since $(G[r], \overline{G[r]}) \in \mathrm{P}$ and the simulation of the $\mathrm{QMA}_{\tilde{\mathcal{T}}}$ protocol for D on the witness state $\left(\frac{\mathrm{Id}}{2}\right)^{\otimes n'_w}$ is a valid BQP algorithm that for $|x| \in \overline{G[r]}$ decides if $x \in D_{\mathrm{yes}}$ or $x \in D_{\mathrm{no}}$. □

6.3 Interpolation of standard complexity classes by physical witness noise

We introduced the concept of noisy QMA classes with the motivation to express intermediate classes between the standard classes QMA, QCMA and BQP. For this it is sufficient to consider noisy QMA classes with i.i.d. noise channels $T^{\otimes n_w}$ for which we introduced the shortened notation QMA_T in definition 6.2.

We can "interpolate" between the complexity classes QMA – BQP and QMA – QCMA for example by changing the error parameter ϵ of the partly depolarizing channel

$$T_\epsilon^{\mathrm{depol}}(\rho) = (1-\epsilon)\rho + \epsilon\frac{\mathbb{I}}{2}$$

and the partly dephasing channel

$$T_\epsilon^{\mathrm{deph}}(\rho) = (1-\epsilon)\rho + \epsilon\frac{\rho + Z\rho Z}{2},$$

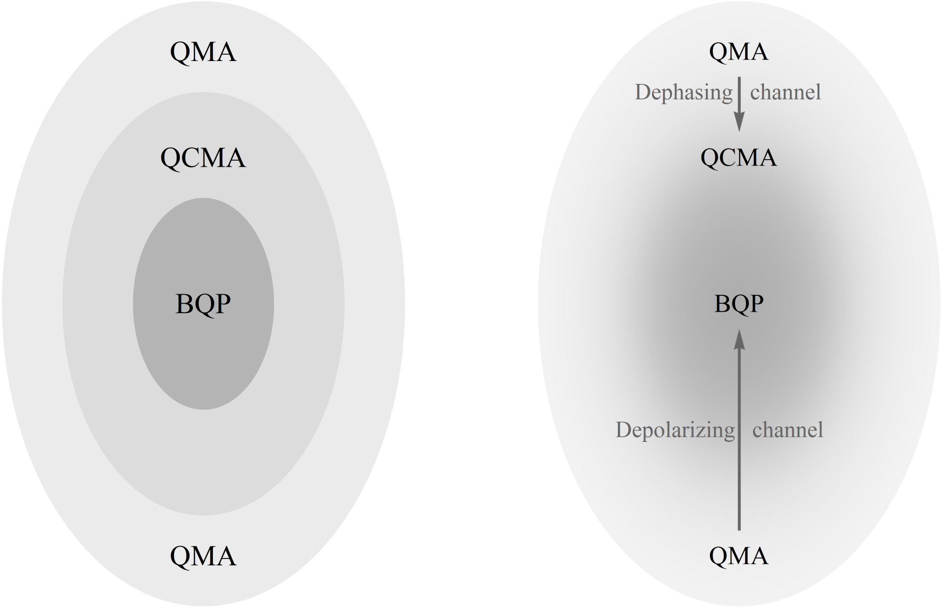


Figure 6.1: Attacking complexity class separations by disturbed witnesses.

respectively. Clearly, $\mathrm{QMA}_{T_0^{\mathrm{depol}}} = \mathrm{QMA}_{T_0^{\mathrm{deph}}} = \mathrm{QMA}$ on the one hand and $QMA_{T_1^{\mathrm{depol}}} = \mathrm{BQP}$ and $\mathrm{QMA}_{T_1^{\mathrm{deph}}} = \mathrm{QCMA}$ on the other hand.

Note that it is not clear if the error parameters of these channels offer a smooth interpolation it that sense that they allows to represent infinite many intermediate classes like Ladner's theorem does. The noisy classes might also jump at some value directly from QMA to BQP or QCMA.

In the next chapter we will see that a quite high constant value of ϵ allows that the respective noisy QMA class still equals QMA. Besides deriving concrete upper bounds on the error parameter ϵ for the partly depolarizing and the partly dephasing channel we will also derive an upper bound for the noise of a general single qubit channel. As a general noise quantity of a channel one can consider the diamond norm on its deviation from the identity channel. This is compatible with the consideration of ϵ as noise parameter for the channels $T = T_\epsilon^{\mathrm{depol}}$ or $T = T_\epsilon^{\mathrm{deph}}$, since it holds $\|T - \mathrm{Id}\,\|_\Diamond = C\epsilon$ for C in a constant range according to lemma 7.28.

The basis for deriving constant error bounds up to which QMA is robust, is that QMA can at least resist some small, i.e. decreasing noise. The next proposition shows that $\mathrm{QMA}_T = \mathrm{QMA}$ is true if the noise of T decreases with the witness length due to amplification and simple norm inequalities.

Proposition 6.9. $\mathrm{QMA}_T = \mathrm{QMA}$ *for every constant* $\delta > 0$ *and every quantum channel* T *with*

$$\|T - \mathrm{Id}\,\|_\diamond \leqslant \frac{2(1-\delta)}{n_w}.$$

Furthermore, $\mathrm{QMA}_T(c,s) = \mathrm{QMA}_T$ *for the above channels and for all polynomial time computable functions* c *and* s *with* $e^{-q} \leqslant s, c \leqslant 1 - e^{-q}$*, gap* $c - s \geqslant 1/q$ *and* q *polynomial.*

Proof. Since $\mathrm{QMA}_T \subseteq \mathrm{QMA}$ is trivial, we only have to prove the opposite subset relation. For this consider a QMA $(1 - e^{-p}, e^{-p})$ verifier $V = (V_x)_{x \in \Sigma^*}$, p polynomial, for a problem A as QMA_T verifier. The soundness value obviously remains unchanged.

To compute the completeness probability of the QMA_T protocol let $\Pi_x := V_x^\dagger \Pi_{acc} V_x$ denote the projection operator for the final output measurement preceded by the verifier's circuit. Since projective measurement operators are included in the maximizing set of lemma 2.16, the difference in the acceptance probabilities supplied an undisturbed witness ρ compared to the disturbed witness $T^{\otimes n_w}(\rho)$ can be estimated as follows:

$$\begin{aligned}
&\left|\mathrm{tr}\left[\Pi_x\big(\,|0\rangle\,\langle 0|^{\otimes z} \otimes (T^{\otimes n_w}(\rho) - \rho)\big)\right]\right| \\
&\qquad\overset{2.16}{\leqslant} \frac{1}{2}\|\,|0\rangle\,\langle 0|^{\otimes z} \otimes (T^{\otimes n_w}(\rho) - \rho)\|_1 \\
&\qquad\overset{2.17}{\leqslant} \frac{1}{2}\|\,\mathrm{Id}^{\otimes z} \otimes (T^{\otimes n_w} - \mathrm{Id}^{\otimes n_w})\|_1 \\
&\qquad\overset{2.19}{\leqslant} \frac{1}{2}\|T^{\otimes n_w} - \mathrm{Id}^{\otimes n_w}\,\|_\diamond \\
&\qquad= \frac{1}{2}\|(T \otimes \mathrm{Id} \otimes \mathrm{Id} \otimes \dots) \circ (\mathrm{Id} \otimes T \otimes \mathrm{Id} \otimes \dots) \circ \dots - \mathrm{Id}^{\otimes n_w}\,\|_\diamond \\
&\qquad\overset{2.19}{\leqslant} \frac{1}{2} n_w \|T - \mathrm{Id}\,\|_\diamond \\
&\qquad\leqslant 1 - \delta.
\end{aligned}$$

The completeness probability might hence decrease to $\delta - e^{-p}$. Assume w.l.o.g. that V outputs the correct answer deterministically for small inputs with $e^{-p} \geqslant \frac{\delta}{4}$. Then V is a $\mathrm{QMA}_T\left(c', e^{-p}\right)$ verifier for the problem A with $c' - e^{-p} \geqslant \frac{\delta}{2}$. The statement $\mathrm{QMA} \subseteq \mathrm{QMA}_T\left(\frac{2}{3}, \frac{1}{3}\right)$ is implied by amplification which we show next.

Amplification is shown if we can construct a $\mathrm{QMA}_T\left(1 - e^{-r}, e^{-r}\right)$ protocol for any problem $A \in \mathrm{QMA}_T(c, s)$ and polynomial r with c and s obeying the restrictions of the theorem. This works by parallelization as in the proof of lemma 3.21. We only have to take the detour $\mathrm{QMA}_T(c, s) \subseteq \mathrm{QMA}\left(1 - e^{-p}, e^{-p}\right)$ and make a sufficient m-fold parallelization of the existing $\mathrm{QMA}(1 - e^{-p}, e^{-p})$ verifier for A that is a $\mathrm{QMA}_T\left(c', e^{-p}\right)$ verfier at the same time. Define $\tilde{T}_{n_w} := T_{mn_w}$ with the lower indices denoting the parameter of the channels which we normally do not spell out. The sublety that each of the verifiers in parallel receives a witness disturbed by $\tilde{T}$ i.i.d. noise instead of T i.i.d. noise due to the increased witness length is now irrelevant, since the verifier still possesses a completeness probability of c' and soundness probability of e^{-p} due to

$$\|\tilde{T} - \mathrm{Id}\,\|_{\diamond} \leqslant \frac{2(1-\delta)}{mn_w} \leqslant \frac{2(1-\delta)}{n_w}$$

and the same argumentation as before. □

Since the noise T in the above amplification proof changes with the parallelization of the witness, one might favor a variant of strong amplification [18] that works for QMA without lengthening of the witness. Unfortunately, this method achieves the desired completeness probability by a specific pure witness state which is not generally known to be a valid channel output.

In the next chapter we will use the tool of concatenated coding to turn channels T of small constant noise into effective channels T' whose noise decreases with the witness length. With the argument from the previous proposition this implies that $\mathrm{QMA}_T = \mathrm{QMA}$ for small constant noise T.

This thesis does not comprise a comprehensive study of noisy QMA classes for high noise. Clearly, QMA_T should become strictly weaker than QMA if the noise of the channel T is so disturbant that it prevents any transmission of quantum information. A standard quantity for expressing the amount of transmitted quantum information is the quantum capacity of a channel (see e.g. [10]). Unfortunately, this well-studied quantity from quantum Shannon theory is not suitable for any collapse argument of noisy QMA, since its definition as limes of transmitted qubits per number of channel uses does note rule out the possibility that states of n_w qubits can be transmitted via polynomially many channels uses despite a vanishing quantum capacity.

The 50-50-erasure channel

$$T^{\text{eras}}_{\frac{1}{2}}(\rho) = \frac{1}{2}\left(\rho \otimes |0\rangle\right) + \frac{1}{2}\left(\frac{\mathbb{I}}{2} \otimes |1\rangle\right)$$

for example has zero quantum capacity [54], but can nevertheless transmit – at least with some average error – states of n_w qubits via $\Theta(n_w^2)$ channel uses [55, 56]. If additionally, the extraction of the n_w qubits from the channel output can be accomplished efficiently (this is unclear since the necessary code is not known explicitly) and the error on the actual witness state does not deviate to much from the average error to allow compensation by amplification, then it holds that $\text{QMA}_{T^{\text{eras}}_{1/2}} = \text{QMA}$. Although ultimately this knowledge is missing, the 50-50-erasure channel is a strong indication for the existence of channels without quantum capacity that do not diminish the power of QMA.

Regarding the quantum erasure channel one might very well even believe in $\text{QMA}_{T^{\text{eras}}_{p}} = \text{QMA}$ for increasing $p = 1 - \frac{k}{n_w}$, k constant, as section 8.4 will reveal this as equivalent to the famous QPCP conjecture.

CHAPTER 7

Robustness of QMA against constant witness noise

7.1 Introduction

In the last chapter we introduced QMA classes with disturbed witnesses. At the end we concentrated on physically realistic noise channels that disturb each qubit independently and equally. Channels of this form are sufficient to interpolate between standard complexity classes such as QMA, QCMA and BQP. The last chapter ended with the proof that QMA stays invariant if the noise on each single qubit decreases with the witness length. This was a simple consequence of amplification by parallel repetition.

In this chapter we will improve this result and show that QMA even stays invariant if each qubit is affected by a quantum channel T of small constant noise. We will derive a general bound on the single qubit noise in terms of the diamond norm $\|T - \mathrm{Id}\,\|_\diamond$ and an improved bound on the error parameter $\epsilon \leqslant \|T - \mathrm{Id}\,\|_\diamond$ for the partly depolarizing channel

$$T_\epsilon^{\mathrm{depol}}(\rho) = (1-\epsilon)\rho + \epsilon\frac{\mathbb{I}}{2}$$

that interpolates between QMA and BQP and the partly dephasing channel

$$T_{\epsilon}^{\text{deph}}(\rho) = (1-\epsilon)\rho + \epsilon\frac{\rho + Z\rho Z}{2}$$

that interpolates between QMA and QCMA. This work was published in [50] around the same time as [49] stated a similar existence statement without explicit proof and calculation of conrete bounds.

The tool to prove robustness of QMA against constantly disturbed witness qubits is provided by concatenated coding. Instead of an undisturbed witness the noisy QMA protocol expects to receive a disturbed, encoded witness. By first applying an error correction and decoding procedure, the protocol can remove enough noise to simulate afterwards the original QMA algorithm with a sufficiently close outcome. Note that no code will be capable of removing the noise completely, because i.i.d. noise includes the possibility that all witness qubits are non-trivially disturbed at the same time, e.g. with probability ϵ^{n_w} for the two channels stated above. In such a setting the specific encoding, error correction and decoding chosen along with the code becomes relevant for its performance. Since these operations are not fixed in the original definition 4.1 of a quantum code, we agree on the following convention for this chapter:

Definition 7.1. *In this chapter every quantum code encodes* 1 *qubit into* N *qubits via a specific encoding* $\mathcal{E}$*. Furthermore, specific logical Pauli operators* $\bar{\mathbb{I}}, \bar{X}, \bar{Y}, \bar{Z}$*, a specific decoding* $\mathcal{D}$ *and a specific standard error correction channel* $\mathcal{R}$ *with projective measurement* $\{P_j\}_{j\in[l]}$ *and recovery operators* $\{R_j\}_{j\in[l]}$ *is assumed.*

The idea of concatenated coding is not to encode each qubit only once, but to consider the combination of encoding, noise, error correction and decoding as a new *effective noise channel* for that another layer of coding is applied (see figure 4.1). After k levels of coding against the single qubit noise T the *effective channel* $T^{(k)}$ hence equals

$$\begin{aligned} T^{(0)} &= T, \\ T^{(k)} &= \mathcal{D} \circ \mathcal{R} \circ \left(T^{(k-1)}\right)^{\otimes N} \circ \mathcal{E}. \end{aligned}$$

The proof technique of concatenated coding led 1999 to the theorem of quantum fault tolerance [57]. This theorem depicts a milestone of quantum computation by promising

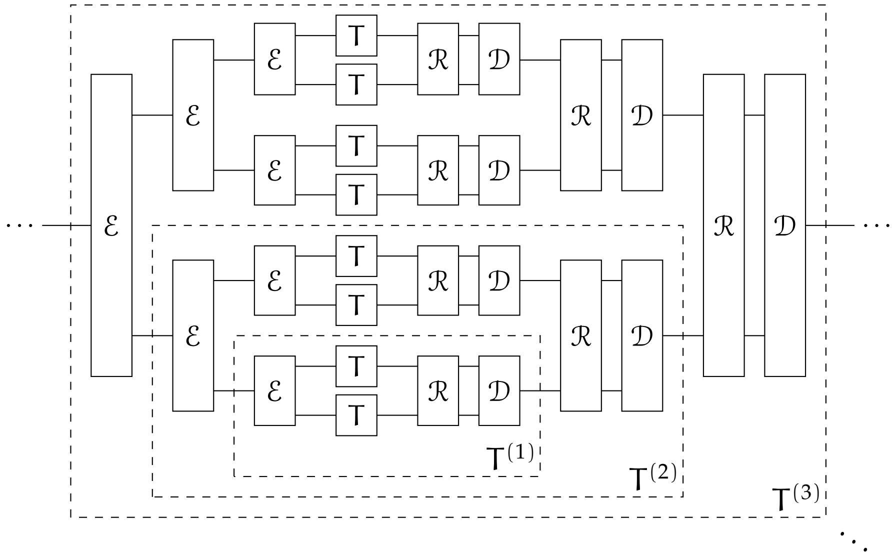


Figure 7.1: 3-fold concatenation of a $(2,2)$ code against the i.i.d. noise channel T.

its physical feasibility. The construction introduces a polynomial overhead of gates into any quantum circuit with the result that constant noise below a non-trivial threshold on each single gate only results in an exponentially small output error.

We revisit this technique now for our scenario in which each witness qubit is affected by constant noise. If we only worked with a fixed number of coding layers the effective channel $\mathsf{T}^{(k)}$ would simply equal another constant noise and the error on the witness state would increase with the number of witness qubits n_w (recall the optimal bound $\|(\mathsf{T}^{(k)})^{\otimes n_w} - \mathrm{Id}^{\otimes n_w}\|_\diamond \leqslant n_w \|\mathsf{T}^{(k)} - \mathrm{Id}\,\|_\diamond$ from lemma 2.19). Hence, the centerpiece of concatenated coding is that the number of coding levels depends on the length of the input state. The simple structure of concatenated single qubit codes against i.id. qubit noise will allow us a feasible analysis with the result that the overall error decreases super-exponentially with the number of coding levels if the constant single qubit noise lies below a non-trivial threshold. Note that the decreasing of the error has to be at least exponentially in the number of coding levels for our purpose since an efficient protocol can decode only logarithmically many levels.

The analysis breaks down into studying the fixed points and convergence behaviour of the *coding map* which describes the transformation of a noise channel T into the respective effective channel after one coding level:

Definition 7.2. *The* coding map Ω^C *for an* $(N, 2)$ *code* C *is defined via*

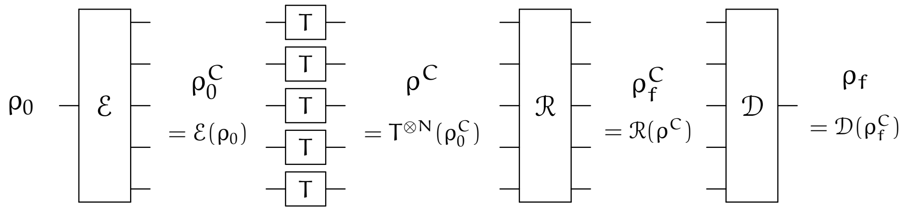
$$\begin{aligned}\Omega^C &: \mathcal{L}\big(\mathcal{L}(\mathbb{C}^2)\big) \to \mathcal{L}\big(\mathcal{L}(\mathbb{C}^2)\big)\\ \Omega^C(T) &= \mathcal{D} \circ \mathcal{R} \circ T^{\otimes N} \circ \mathcal{E}.\end{aligned}$$

Notice that we can regard an effective channel as new noise for another coding level only because our initial assumption of i.i.d. noise ensures that after k coding levels each of the N^k physical qubits for an original qubit is still disturbed by the same noise T. If this assumption was dropped, we would need knowledge about how the noise extends to larger systems and could not simply reduce our study to the fixed point analysis of the coding map Ω^C.

In the following two sections 7.2 and 7.3 we derive an expression for the coding map of general single qubit codes and stabilizer codes. These expressions were first formulated by [58]. The fixed point analysis of these coding maps in sections 7.5 and 7.6 is based on the work by [59].

7.2 Concatenation of general codes

Definition 7.3. *We use the following notation for the different stages an initial operator* $\rho_0 \in \mathcal{L}(\mathbb{C}^2)$ *runs through while passing one level of coding:*



Given the initial and final operator we denote the expectation values of the Pauli operators $\sigma \in \mathcal{P}$ *by* $\langle\sigma\rangle_0 = \mathrm{tr}(\sigma\rho_0)$ *and* $\langle\sigma\rangle_f = \mathrm{tr}(\sigma\rho_f)$*, respectively.*

Note that we use the above notation not only for valid density matrices but for any operator $\rho \in \mathcal{L}(\mathbb{C}^2)$. Since an operator in $\mathcal{L}(\mathbb{C}^2)$ can be written as a linear combination of Pauli operators, a superoperator $\mathsf{T} \in \mathcal{L}\big(\mathsf{L}(\mathbb{C}^2)\big)$ is fully described by its *Stokes parametrization*: a complex 4×4 matrix with the matrix entry $\mathsf{T}_{\sigma\sigma'}$ representing the prefactor of $\sigma \in \mathcal{P}$ in the output operator given the input operator $\sigma' \in \mathcal{P}$. Observe that the prefactor of σ given the linear combination $\rho \in \mathcal{L}(\mathbb{C}^2)$ equals exactly half of the expectation value of σ:

$$\rho = \frac{1}{2} \sum_{\sigma \in \mathcal{P}} \langle \sigma \rangle \sigma.$$

Consequently, the matrix element $\mathsf{T}_{\sigma\sigma'}$ can also be considered as the expectation value of the Pauli operator σ given the operator $\mathsf{T}\left(\frac{1}{2}\sigma'\right)$:

$$\mathsf{T}_{\sigma\sigma'} = \operatorname{tr}\left(\sigma \mathsf{T}\left(\frac{1}{2}\sigma'\right)\right).$$

We will be able to derive an expression for the coding map with the help of two simple expressions relating the initial and final expectation values $\langle \sigma \rangle_0$ and $\langle \sigma \rangle_f$ to the so-called en- and decoding operators of the code. One might find the terminology "decoding operators" a bit misleading since the operators rather describe the action of the error correction channel than the actual decoding. However, we stick to this terminology to stay consistent with the literature [58] from which we adapted the following definitions and theorem 7.8.

Definition 7.4. *The* encoding operators E_σ, $\sigma \in \mathcal{P}$, *for a quantum code* C *are defined via*

$$\mathsf{E}_\sigma := \frac{1}{2} \mathsf{P}_\mathsf{C} \bar{\sigma}.$$

Since encoding operators act like the respective logical Pauli operators times 1/2 on the codespace and vanish on the orthogonal space, an encoded state can be expressed as following:

Corollary 7.5. *A quantum code* C *encodes an initial operator* ρ_0 *into*

$$\rho_0^\mathsf{C} = \sum_{\sigma \in \mathcal{P}} \langle \sigma \rangle_0 \mathsf{E}_\sigma.$$

Definition 7.6. *The* decoding operators D_σ*,* $\sigma \in \mathcal{P}$*, of a quantum code* C *are defined via*

$$D_\sigma := 2 \sum_{j \in [l]} P_j^\dagger R_j^\dagger E_\sigma R_j P_j .$$

Corollary 7.7. *The Pauli expectation values of the output operator* ρ_f *for a quantum code* C *are given by*

$$\langle \sigma \rangle_f = \operatorname{tr}\left(D_\sigma \rho^C \right).$$

Theorem 7.8. *Consider a quantum code* C *with encoding and decoding operators*

$$\begin{aligned} E_{\sigma'} &= \sum_{\mu \in \mathcal{P}^{\otimes N}} \alpha_\mu^{\sigma'} \left(\frac{1}{2} \mu_1 \right) \otimes \cdots \otimes \left(\frac{1}{2} \mu_N \right) \\ D_\sigma &= \sum_{\nu \in \mathcal{P}^{\otimes N}} \beta_\nu^\sigma \nu_1 \otimes \cdots \otimes \nu_N . \end{aligned}$$

Then a coding level transforms a single qubit superoperator T *into to the effective superoperator* $\tilde{T} = \Omega^C(T)$ *with*

$$\tilde{T}_{\sigma\sigma'} = \sum_{\mu,\nu \in \mathcal{P}^{\otimes N}} \beta_\nu^\sigma \alpha_\mu^{\sigma'} \prod_{j \in [N]} T_{\nu_j \mu_j} .$$

Proof. The matrix element $\tilde{T}_{\sigma\sigma'}$ equals the final expectation value $\langle \sigma \rangle_f$ given the initial operator $\rho_0 = \frac{1}{2}\sigma'$. Hence we obtain:

$$\begin{aligned} \tilde{T}_{\sigma\sigma'} &= \langle \sigma \rangle_f \\ &\overset{7.7}{=} \operatorname{tr}(D_\sigma \rho^C) \\ &= \operatorname{tr}\left(D_\sigma T^{\otimes N}(\rho_0^C) \right) \\ &\overset{7.5}{=} \operatorname{tr}\left(D_\sigma T^{\otimes N}(E_{\sigma'}) \right) \\ &= \sum_{\mu,\nu \in \mathcal{P}^{\otimes N}} \beta_\nu^\sigma \alpha_\mu^{\sigma'} \prod_{j \in [N]} \operatorname{tr}\left(\nu_j T \left(\frac{1}{2} \mu_j \right) \right) \\ &= \sum_{\mu,\nu \in \mathcal{P}^{\otimes N}} \beta_\nu^\sigma \alpha_\mu^{\sigma'} \prod_{j \in [N]} T_{\nu_j \mu_j} . \end{aligned}$$ □

The previous theorem shows that the matrix elements of the efficient channel $\tilde{T}$ are polynomials of degree N in the matrix entries of T. A smaller physical qubit number N of the chosen code might hence simplify the analysis of concatenated quantum coding. However, in many cases simplification is rather achieved by codes preserving a specific structure of the effective channels. Concatenated stabilizer codes provide for example the advantage to be diagonality preserving as we will see in the next section.

7.3 Concatenation of stabilizer codes

In this section we adapt the results of the previous section for stabilizer codes. With the following notation the expressions for decoding operators and coding map get particularly simple:

Definition 7.9. *In this chapter the stabilizer group of a code* $C(S)$ *is denoted by* $S = \{S_i\}_{i=1}^{2^m}$ *and its independent generators by* $\{g_i\}_{i=1}^{m}$.

The stabilizer code is always provided with a standard error correction channel as described in theorem 4.36, i.e. the projection operators $\{P_j\}_{j\in[2^m]}$ *project onto the generator syndrome spaces and the recovery operators* $\{R_j\}_{j\in[2^m]}$ *are Pauli group elements turning the syndrome spaces into the codespace.*

We characterize the en- and decoding operators by coefficients $\alpha_\mu^{\sigma'}$ *and* β_ν^σ *as in theorem 7.8. We write* $|\mu|$ *for the Pauli operator equalling* $\mu \in \mathcal{P}_N$ *without prefactor and* $|\mu|_k$ *to refer to the single qubit Pauli operator in the* k*-th tensor product position of* μ.

We can derive a short expression for the decoding operators of a stabilizer code making use of the η-function introduced in definition 2.8 indicating the commutation relation of two Pauli group elements:

Definition 7.10. *For a stabilizer code* $C(S)$ *we define for all* $i \in [2^m]$ *and* $\sigma \in \mathcal{P}$ *the* f-coeffcient

$$f_{i\sigma} := \sum_{j\in[2^m]} \eta(R_j, S_i)\eta(R_j, \bar{\sigma}).$$

Lemma 7.11. *The decoding operators of a stabilizer code can be written as*

$$D_\sigma = \frac{1}{|S|} \sum_{i \in [2^m]} f_{i\sigma} S_i \bar{\sigma}.$$

Proof. Since the code space projection for stabilizer codes equals $P_C = \frac{1}{|S|} \sum_{i \in [2^m]} S_i$ we can insert $E_\sigma = \frac{1}{2|S|} \sum_{i \in [2^m]} S_i \bar{\sigma}$ into the expression for the decoding operators by definition 7.6:

$$\begin{aligned} D_\sigma &= \frac{1}{|S|} \sum_{i \in [2^m]} \sum_{j \in [2^m]} P_j R_j^\dagger S_i \bar{\sigma} R_j P_j \\ &= \frac{1}{|S|} \sum_{i \in [2^m]} \sum_{j \in [2^m]} P_j \eta(R_j, S_i) \eta(R_j, \bar{\sigma}) S_i \bar{\sigma} P_j. \end{aligned}$$

According to lemma 4.26 the syndrome space projections P_j equal $\frac{1}{|S|} \sum_{i \in [2^m]} \eta(R_j, S_i) S_i$. Since the P_j furthermore commute with all stabilizers and logical operators we obtain

$$\begin{aligned} D_\sigma &= \frac{1}{|S|} \sum_{i \in [2^m]} \sum_{j \in [2^m]} \eta(R_j, S_i) \eta(R_j, \bar{\sigma}) S_i \bar{\sigma} \\ &= \frac{1}{|S|} \sum_{i \in [2^m]} f_{i\sigma} S_i \bar{\sigma}. \end{aligned}$$ □

Lemma 7.12. *The coefficients of the en- and decoding operators of a stabilizer code fulfill*

$$\alpha_\mu^\sigma = \begin{cases} \pm 1 & \textit{if } \mu = |S_i \bar{\sigma}| \\ 0 & \textit{otherwise,} \end{cases}$$

$$\beta_\mu^\sigma = \begin{cases} \frac{f_{i\sigma}}{|S|} \alpha_\mu^\sigma & \textit{if } \mu = |S_i \bar{\sigma}| \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. From the expression

$$E_\sigma = \frac{1}{2|S|} \sum_{i \in [2^m]} S_i \bar{\sigma}$$

it is clear that $\alpha_\mu^\sigma = 0$ if $\mu \neq |S_i \bar{\sigma}|$ for all stabilizers $S_i \in S$.

Let's now consider the case $\mu = |S_i\bar{\sigma}|$. First it is important to observe that the stabilizer S_i and the Pauli operator σ are uniquely determined by the product $|S_i\bar{\sigma}|$ according to lemma 4.35.

Since stabilizers and logical operators are hermitian and commute, $S_i\bar{\sigma}$ is hermitian, too. As a consequence each summand in the encoding operator E_σ equals a distinct tensor product of Pauli operators with prefactor $\pm\frac{1}{2^N}$ and thus $\alpha^\sigma_{|S_i\bar{\sigma}|} = \pm 1$.

The expression for β^σ_μ follows directly from the coefficient comparision of E_σ and

$$D_\sigma = \frac{1}{|S|}\sum_{i\in[2^m]} f_{i\sigma} S_i\bar{\sigma}. \qquad \square$$

The above properties of the α and β coefficients allow an easy computation of the coding map based on the expression by theorem 7.8. The expression gets even simpler if the original superoperator T is *diagonal* in the Stokes representation. Due to trace preservation these channels are fully described by the notation

$$T = [T_{XX}, T_{YY}, T_{ZZ}].$$

The next theorem, adapted from [58], shows that concatenated stabilizer codes preserve the diagonality of superoperators. Note that the arXiv publication of [58] contains wrong exponents.

Theorem 7.13. *A stabilizer code* $C(S)$ *transforms a single qubit superoperator* $T = \mathrm{diag}(e, x, y, z)$ *into the effective superoperator* $\tilde{T} = \Omega^C(T)$ *with*

$$\tilde{T}_{\sigma\sigma'} = \delta_{\sigma\sigma'}\frac{1}{|S|}\sum_{S_i\in S} f_{i\sigma} e^{w_{\mathbb{I}}(S_i\bar{\sigma})} x^{w_X(S_i\bar{\sigma})} y^{w_Y(S_i\bar{\sigma})} z^{w_Z(S_i\bar{\sigma})}.$$

Proof. Due to the diagonality of T the statement of theorem 7.8 simplifies to

$$\begin{aligned}\tilde{T}_{\sigma\sigma'} &= \sum_{\mu\in\mathcal{P}^{\otimes n}} \beta^\sigma_\mu \alpha^{\sigma'}_\mu \prod_{j\in[N]} T_{\mu_j\mu_j} \\ &= \sum_{\mu\in\mathcal{P}^{\otimes n}} \beta^\sigma_\mu \alpha^{\sigma'}_\mu e^{w_{\mathbb{I}}(\mu)} x^{w_X(\mu)} y^{w_Y(\mu)} z^{w_Z(\mu)}.\end{aligned}$$

Lemma 7.12 allows us to replace the β coefficients and reduce the summation to operators $\mu \in \mathcal{P}^{\otimes n}$ of the form $|S_i\bar{\sigma}|$:

$$\begin{aligned}\tilde{T}_{\sigma\sigma'} &= \frac{1}{|S|}\sum_{S_i\in S} f_{i\sigma}\alpha^{\sigma}_{|S_i\bar{\sigma}|}\alpha^{\sigma'}_{|S_i\bar{\sigma}|} e^{w_{\mathbb{I}}(S_i\bar{\sigma})}x^{w_X(S_i\bar{\sigma})}y^{w_Y(S_i\bar{\sigma})}z^{w_Z(S_i\bar{\sigma})} \\ &= \delta_{\sigma\sigma'}\frac{1}{|S|}\sum_{S_i\in S} f_{i\sigma}e^{w_{\mathbb{I}}(S_i\bar{\sigma})}x^{w_X(S_i\bar{\sigma})}y^{w_Y(S_i\bar{\sigma})}z^{w_Z(S_i\bar{\sigma})}.\end{aligned}$$ □

7.4 Channel properties in Stokes representation

Before studying how the effective channel series generated by concatenated coding converges, this section provides a short overview of channel properties in Stokes representation. Note that a superoperator T is called *unital* iff $T(\mathbb{I}) = \mathbb{I}$ and that self-adjointness is meant with respect to the Hilbert-Schmidt inner product, i.e.

$$\operatorname{tr}\left(\rho^\dagger T(\rho')\right) = \operatorname{tr}\left(T(\rho)\rho'\right)$$

for all $\rho, \rho' \in \mathcal{L}(\mathbb{C}^2)$.

Lemma 7.14. *The following correspondances hold between a superoperator $T \in \mathcal{L}\big(\mathcal{L}(\mathbb{C}^2)\big)$ and its Stokes matrix representation*

$$T = \begin{pmatrix} z & s \\ t & A \end{pmatrix}$$

with A a complex 3x3 matrix and s and t row and column vector of length 3, respectively:

linear function on $\mathcal{L}(\mathbb{C}^2)$	*Stokes representation*
hermiticity preserving	*real*
trace preserving	$(z, s) = (1, 0, 0, 0)$
self-adjoint	*hermitian*
unital	$t = (0, 0, 0)^T$

Proof. [6, §2.1]. □

Lemma 7.15. *For a valid quantum channel* T *and all* $\sigma \in \{X, Y, Z\}$ *it holds that*

1. $T_{\sigma X}^2 + T_{\sigma Y}^2 + T_{\sigma Z}^2 \leqslant (1 - |T_{\sigma \mathbb{I}}|)^2$,
2. $T_{X\sigma}^2 + T_{Y\sigma}^2 + T_{Z\sigma}^2 \leqslant 1 - T_{X\mathbb{I}}^2 - T_{Y\mathbb{I}}^2 - T_{Z\mathbb{I}}^2$.

Proof. [59, lemma 5.1]. □

Corollary 7.16. $T_{\sigma\sigma'} \in [-1, 1]$ *for all matrix entries of a valid quantum channel* T.

Proof. First, we note that all matrix entries of T are real according to lemma 7.14, since complete positivity also implies that quantum channels preserve hermiticity. Moreover, the completely mixed state is only mapped to a valid quantum state if the entries of the first column of T lie within the interval $[-1, 1]$. The first property of lemma 7.15 implies then that all matrix entries are of norm at most 1. □

An even stricter range is provable for the matrix entries of diagonal channels. It allows us later to prove convergence of a diagonal channel towards identity by only showing the convergence towards 1 of two out of the three matrix entries T_{XX}, T_{YY} and T_{ZZ}:

Lemma 7.17. *A diagonal quantum channel* $\mathsf{T} = [T_{XX}, T_{YY}, T_{ZZ}]$ *fulfills*

$$|T_{XX} \pm T_{YY}| \leqslant |1 \pm T_{ZZ}|,$$

i.e. the point (T_{XX}, T_{YY}, T_{ZZ}) *lies within the tetrahedron* Δ *defined by the corners* $(1, 1, 1)$, $(1, -1, -1)$, $(-1, 1, -1)$ *and* $(-1, -1, 1)$.

Proof. [6, Appendix B]. □

Lemma 7.18. *A quantum channel* T *is a* Pauli channel, *i.e.*

$$T(\rho) = (1 - p_X - p_Y - p_Z)\rho + p_X X\rho X + p_Y Y\rho Y + p_Z Z\rho Z$$

iff T *is diagonal with*

$$\mathsf{T} = [1 - 2(p_Y + p_Z), 1 - 2(p_X + p_Z), 1 - 2(p_X + p_Y)].$$

Proof. [59, §II]. □

We end this section with another result by [59, theorem 6.2] showing that every single qubit channel corresponds to a channel with a simplified Stokes representation padded by unitary channels:

Lemma 7.19 (SVD standard form of channels)**.** *For every single qubit channel* T *there exist unitary single qubit channels* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *such that all off-diagonal, unital entries of* $\mathcal{U}_2 \circ \mathsf{T} \circ \mathcal{U}_1$ *vanish.*

Proof. Consider an arbitrary single qubit channel

$$\mathsf{T} = \begin{pmatrix} 1 & 0 \quad 0 \quad 0 \\ t & A \end{pmatrix},$$

$t \in \mathbb{R}^3$, $A \in \mathbb{R}_{3,3}$, and the singluar value decomposition $A = O_2^\dagger D O_1^\dagger$. Then $A = R_2^\dagger(\pm D)R_1^\dagger$ with R_i a rotation because an orthogonal operator O_i can always be written as $O_i = \pm R_i$.

Since a rotation in the Bloch sphere corresponds to a unitary operator on the state in its $\mathbb{C}^2$-representation

$$\mathcal{U}_i = \begin{pmatrix} 1 & 0 \\ 0 & R_i \end{pmatrix}$$

is a valid unitary channel.

We obtain the SVD standard form

$$\mathcal{U}_2 \circ \mathsf{T} \circ \mathcal{U}_1 = \begin{pmatrix} 1 & 0 \quad 0 \quad 0 \\ R_2 t & \pm D \end{pmatrix}.$$ □

7.5 Convergence of diagonal noise

The motivation for studying concatenated coding is the idea that the effective channel gets less noisy with every coding level and eventually converges to the identity. But of course this will only happen if both the code and the initial noise fulfill certain minimum requirements. Mathematically speaking, the question is if the identity is an attracting fixed point of the coding map and if the initial noise lies within its basin of attraction.

In this section we restrict our studies to the convergence of diagonal, trace-preserving superoperators T under diagonality preserving codes. For such superoperators the coding map is fully described by its *diagonal-reduced coding map* variant

$$\Omega_d^C : \mathbb{R}^3 \to \mathbb{R}^3$$
$$\Omega_d^C\big([T_{XX}, T_{YY}, T_{ZZ}]\big) := [\tilde{T}_{XX}, \tilde{T}_{YY}, \tilde{T}_{ZZ}]$$

with $\tilde{T} = \Omega^C(T)$.

With the next definition we introduce some standard analysis vocabulary for functions $\mathbb{R}^k \to \mathbb{R}^k$ with an arbitrary choice of norm on $\mathbb{R}^k$.

Definition 7.20. *Let* $f : \mathbb{R}^k \to \mathbb{R}^k$.

- *A point* $p = f(p)$ *is called a* fixed point *of* f.
- *The* orbit *of a point* $x \in \mathbb{R}^k$ *is the series* $\big(x, f(x), f^2(x), f^3(x), \dots\big)$ *with* f^n *indicating the* n*-times concatenation of* f.
- *A fixed point* p *is called* locally attracting *iff it has a neighborhood* $U \subseteq \mathbb{R}^k$ *such that the orbit of every* $x \in U$ *converges towards* p.
- *The* basin of attraction $B(p)$ *of an attracting fixed point* p *is its largest neighborhood such that the orbit of every point in* $B(p)$ *converges towards* p.

We write $\|Df(p)\|$ to denote the induced norm of the Jacobian matrix Df of a differentiable function $f : \mathbb{R}^k \to \mathbb{R}^k$ in the point p:

$$\|Df(p)\| = \sup\left\{ \|(Df(p))r\| \,\middle|\, r \in \mathbb{R}^k, \|r\| \leqslant 1 \right\}.$$

Theorem 7.8 tells us that the diagonal-reduced coding map $\Omega_d^C : \mathbb{R}^3 \to \mathbb{R}^3$ maps to a vector of polynomials and is therefore a C_1-function, i.e. its Jacobian exists and is continuous. This allows us to apply a standard result from analysis giving a sufficient condition for a fixed point p to be attracting:

Lemma 7.21. *Let $f : \mathbb{R}^k \to \mathbb{R}^k$ be a C_1-function with fixed point p. If $\lambda_0 := \|Df(p)\| < 1$ then p is locally attracting.*

Proof. Choose a λ with $\lambda_0 < \lambda < 1$. Since $Df(x)$ is continuous there exists a neighbourhood U of p such that $\|Df(x)\| \leqslant \lambda$ for all $x \in U$. By the mean value theorem the orbits of all $x \in U$ converge to $p = f^n(p)$:

$$\lim_{n\to\infty} \|f^n(x) - f^n(p)\| \leqslant \lim_{n\to\infty} \lambda^n \|x - p\| = 0. \qquad \square$$

The next proposition corresponding to [59, theorem 3.3] shows that $\|D\Omega_d^C([1,1,1])\| = 0$ for codes with a certain minimum error-correction capability. With the above lemma this indeed reveals the identity channel as an attracting fixed point of the diagonal-reduced coding map.

Proposition 7.22. *Let C be a diagonality-preserving quantum code correcting any single qubit error. Then the Jacobian of the diagonal-reduced coding map Ω_d^C vanishes at $[1,1,1]$, exposing it as an attracting fixed point.*

Proof. Consider the single qubit noise $T = [1, 1-2\epsilon, 1-2\epsilon]$ which imposes an X-error with probability ϵ and no error otherwise according to lemma 7.18. Since C can correct any single qubit error the probability that a coding level on the product noise $T^{\otimes N}$ returns the identity channel is $1 - \mathcal{O}(\epsilon^2)$. Hence

$$\Omega_d^C(T) = [1 - \mathcal{O}(\epsilon^2), 1 - \mathcal{O}(\epsilon^2), 1 - \mathcal{O}(\epsilon^2)].$$

Because the effective channel does not contain a first order term in ϵ, the derivative in the direction of $\nu_X = [0, -2, -2]$ vanishes:

$$\begin{aligned} D\Omega_d^C\big([1,1,1]\big) \cdot \nu_x &= \frac{d}{d\epsilon} \Omega_d^C\big(\underbrace{[1,1,1] + \epsilon\nu_X}_{=T}\big)_{|\epsilon=0} \\ &= [0,0,0]. \end{aligned}$$

The argument runs analogously for $\nu_Y = [-2, 0, -2]$ and $\nu_Z = [-2, -2, 0]$. As ν_X, ν_Y, ν_Z are linearly independent it follows that $D\Omega_d^C\big([1,1,1]\big) = 0$. □

Notice that the above proposition only states that small diagonal noise will converge towards the identity by suitable concatenated coding. This can be concluded from the derivative of the restricted function Ω_d^C, because the starting point is assumed to be diagonal and diagonality of the channel series is preserved. It is important to stress that so far the proposition does not allow any conclusion for non-diagonal noise. In such a case the derivative of the full coding map Ω^C has to be checked. So far it is neither clear that this full derivative vanishes for the identity channel nor that its norm is at all smaller than 1. But restricting to stabilizer codes of distance at least 3 in the next section, we can prove directly that the identity channel is indeed an attracting fixed point of their full coding map.

7.6 Convergence of general noise under stabilizer coding

In this section we extend our convergence study to non-diagonal noise but under the restriction of concatenated stabilizer codes. Since the minimum weight over all non-identity stabilizers will turn out to be one of the parameters determining the convergence behaviour, we introduce the following definition:

Definition 7.23. *We call a stabilizer code* $C(S)$ *an* $[N, k, \delta, w]$ *code if it is an* $[N, k, \delta]$ *code with*

$$w := \min_{S_i \in S \setminus \{\mathbb{I}\}} w(S_i).$$

In the following theorems often an $[N, 1, \delta, w]$ code with distance $\delta \geqslant 3$ and $w \geqslant 2$ is assumed. Notice that such codes exists, e.g. the 5-qubit code, the Steane 7-qubit code and the Shor code have these parameters.

Stabilizer codes have the advantage that noise channels with small off-diagonal terms almost behave like diagonal channels as the next theorem adapted from [59, theorem 5.5] shows:

Theorem 7.24. *Consider coding with an* $[N, 1, \delta, w]$ *stabilizer code against a noise channel* $T = D + \epsilon F$, $|F_{\sigma\sigma'}| \leqslant 1$ *for all* $\sigma, \sigma' \in \mathcal{P}$, *with* D *containing the diagonal and* ϵF *the off-diagonal elements in Stokes representation. Then*

$$\left|\left(\Omega^C(T) - \Omega^C(D)\right)_{\sigma\sigma'}\right| \leqslant \begin{cases} c_F \epsilon^\delta & \textit{if } \sigma \neq \sigma' \\ c_F \epsilon^w & \textit{if } \sigma = \sigma' \end{cases}$$

with $2^m \leqslant c_F := 2^m \max_{\sigma \in \mathcal{P}} \sum_{S_i \in S} \left|\beta^{\sigma}_{|S_i \bar{\sigma}|}\right| \leqslant 4^m$.

Proof. The off-diagonal elements can be bounded by

$$\begin{aligned} \left|\left(\Omega^C(T)\right)_{\sigma\sigma'}\right| &\overset{7.8}{=} \left| \sum_{\mu,\nu \in \mathcal{P}^{\otimes N}} \beta^{\sigma}_{\nu} \alpha^{\sigma'}_{\mu} \prod_{k \in [N]} T_{\nu_k \mu_k} \right| \\ &\overset{7.12}{\leqslant} \sum_{S_i, S_j \in S} \left|\beta^{\sigma}_{|S_i \bar{\sigma}|}\right| \left| \prod_{k \in [N]} T_{|S_i \bar{\sigma}|_k |S_j \bar{\sigma}'|_k} \right| \\ &\leqslant 2^m \sum_{S_i \in S} \left|\beta^{\sigma}_{|S_i \bar{\sigma}|}\right| \epsilon^\delta \\ &\leqslant c_F \epsilon^\delta. \end{aligned}$$

The third line follows from the second line, since $|S_i\bar{\sigma}|$ and $|S_j\bar{\sigma}'|$ are multiplicative related by an element from $N(S)\backslash S^{\pm}$ and hence differ in at least δ qubit positions. Thus the product contains at least δ off-diagonal matrix entries of T while the remaining factors are upper-bounded by 1 according to corollary 7.16.

For the diagonal elements we first derive analogously

$$\begin{aligned}\left|\left(\Omega^C(T)-\Omega^C(D)\right)_{\sigma\sigma}\right| &\overset{7.8}{=} \left|\sum_{\mu,\nu\in\mathcal{P}^{\otimes N}} \beta^{\sigma}_{\nu}\alpha^{\sigma}_{\mu}\left(\prod_{k\in[N]} T_{\nu_k\mu_k} - \prod_{k\in[N]} D_{\nu_k\mu_k}\right)\right| \\ &\overset{7.12}{\leqslant} \sum_{S_i,S_j\in S}\left|\beta^{\sigma}_{|S_i\bar{\sigma}|}\right|\left|\prod_{k\in[N]} T_{|S_i\bar{\sigma}|_k|S_j\bar{\sigma}|_k} - \prod_{k\in[N]} D_{|S_i\bar{\sigma}|_k|S_j\bar{\sigma}|_k}\right|.\end{aligned}$$

Since the two products are the same and cancel if $S_i = S_j$, we can restrict the sum to $S_i \neq S_j$. Given this condition the second product vanishes since it contains at least one off-diagonal factor. Recalling finally that $S_i\bar{\sigma}$ and $S_j\bar{\sigma}$ are multiplicative related by a non-identity stabilizer whose weight is greater or equal to w, it follows that

$$\begin{aligned}\left|\left(\Omega^C(T)-\Omega^C(D)\right)_{\sigma\sigma}\right| &\leqslant \sum_{\substack{S_i,S_j\in S\\ S_i\neq S_j}}\left|\beta^{\sigma}_{|S_i\bar{\sigma}|}\right|\left|\prod_{k\in[N]} T_{|S_i\bar{\sigma}|_k|S_j\bar{\sigma}|_k}\right| \\ &\leqslant 2^m\sum_{S_i\in S}\left|\beta^{\sigma}_{(S_i\bar{\sigma})}\right|\epsilon^{w} \\ &\leqslant c_F\epsilon^{w}.\end{aligned}$$

It remains to prove the bounds on the constant c_F. Lemma 7.12 shows $\left|\beta^{\sigma}_{|S_i\bar{\sigma}|}\right| \leqslant 1$ for all $S_i \in S$ and $\sigma \in \mathcal{P}$. Together with $\left|\beta^{\mathbb{I}}_{\mathbb{I}}\right| = 1$ it follows directly that

$$2^m \leqslant c_F = 2^m \max_{\sigma\in\mathcal{P}}\sum_{S_i\in S}\left|\beta^{\sigma}_{|S_i\bar{\sigma}|}\right| \leqslant 4^m. \qquad \square$$

The previous theorem tells us how a non-diagonal noise channel converges towards a diagonal channel under suitable concatenated stabilizer coding. But we are interested particularly in the convergence towards a specific diagonal channel, namely the identity.

For the off-diagonal elements the previous theorem already provides a bound, but for the diagonal elements we have to combine the above result via triangle inequality with our knowledge from theorem 7.22 about the convergence of diagonal superoperators towards the identity.

Theorem 7.25. *Consider coding with an* $[N, 1, \delta, w]$ *stabilizer code with distance* $\delta \geqslant 3$ *and* $w \geqslant 2$ *against a noise channel* T *obeying* $|(T - \mathrm{Id})_{\sigma\sigma'}| \leqslant \epsilon \leqslant 1$ *for all* $\sigma, \sigma' \in \mathcal{P}$*. Then there exists a constant* c *such that*

$$\left|\left(\Omega^C(T) - \mathrm{Id}\right)_{\sigma\sigma'}\right| \leqslant c\epsilon^2 \qquad \forall \sigma, \sigma' \in \mathcal{P}.$$

Proof. Due to theorem 7.24 and the triviality of the case $\sigma = \sigma' = \mathbb{I}$, it only remains to show the statement for the cases $\sigma = \sigma' \in \{X, Y, Z\}$. For this write

$$T = \mathrm{Id} + \epsilon M + \epsilon F$$

with $D := \mathrm{Id} + \epsilon M$ just containing the diagonal and ϵF the off-diagonal entries.

From theorem 7.22 we know that the Jacobian of the diagonal-reduced coding map Ω^C_d vanishes at the point $[1, 1, 1]$ and that hence the Taylor series expansion of $(\Omega^C_d)_\sigma$ around this point does not contain a linear term. Since the Taylor series is finite and $\epsilon \leqslant 1$, there exists a constant c_M such that

$$\left|\left(\Omega^C(D) - \mathrm{Id}\right)_{\sigma\sigma}\right| = \left|(\Omega^C_d)_\sigma(D) - 1\right| \leqslant c_M \epsilon^2.$$

Triangle inequality and the previous theorem 7.24 then lead to

$$\begin{aligned}\left|\left(\Omega^C(T) - \mathrm{Id}\right)_{\sigma\sigma}\right| &\leqslant \left|\left(\Omega^C(T) - \Omega^C(D)\right)_{\sigma\sigma}\right| + \left|\left(\Omega^C(D) - \mathrm{Id}\right)_{\sigma\sigma}\right| \\ &\leqslant (c_F + c_M)\epsilon^2. \qquad \square\end{aligned}$$

So far we derived an expression for how noise decreases by one level of coding. Complete induction helps us to derive a non-recursive bound on the noise that remains after k levels of concatenated coding:

Lemma 7.26. *The recursive series* $\epsilon_{k+1} = \alpha\epsilon_k^2$ *has the explicit form*

$$\epsilon_k = \frac{1}{\alpha}(\alpha\epsilon_0)^{2^k}.$$

Proof. The statement is proven by complete induction. For $k = 0$ the statement is obviously fulfilled. Now assume the statement is true for a specific k. Then

$$\begin{aligned}\epsilon_{k+1} &= \alpha\epsilon_k^2 \\ &\overset{\text{i.h.}}{=} \alpha\left(\frac{1}{\alpha}(\alpha\epsilon_0)^{2^k}\right)^2 \\ &= \frac{1}{\alpha}(\alpha\epsilon_0)^{2^{k+1}}.\end{aligned}$$

□

Corollary 7.27. *Consider the setting of theorem 7.25 and denote by* $T^{(k)}$ *the effective channel after* k *levels of concatenated coding against the noise channel* $T = T^{(0)}$. *Then*

$$\left|\left(T^{(k)} - \mathrm{Id}\right)_{\sigma\sigma'}\right| \leqslant \frac{1}{c}(c\epsilon)^{2^k} \qquad \forall\sigma, \sigma' \in \mathcal{P}.$$

7.7 Robustness results

In this section we combine the previous results to prove that $\mathrm{QMA}_T = \mathrm{QMA}$ even for channels T of constant noise. The central idea is to verify a QMA problem by a QMA_T protocol receiving a disturbed encoded version of the witness that the QMA protocol would receive. The QMA_T protocol carries out error correction and decoding to extract the original witness and then simulates the original QMA protocol. Despite the efficiency restriction to maximally polylogarithmic many levels of coding the extracted witness will be sufficiently close to the original witness to guarantee any acceptance probability in the usual range.

The main theorem is preceded by a lemma allowing us to convert between the bound on the diamond norm of $T-\mathrm{Id}$ and bounds on the matrix entries in Stokes representation.

Lemma 7.28. *For all superoperators* T *and all* $\sigma, \sigma' \in \mathcal{P}$ *it holds that*

$$\frac{1}{16c_\Diamond}\|\mathsf{T} - \mathrm{Id}\,\|_\Diamond \leqslant |(\mathsf{T} - \mathrm{Id})_{\sigma\sigma'}| \leqslant \|\mathsf{T} - \mathrm{Id}\|_\Diamond$$

with $c_\Diamond$ *the maximum diamond norm over all superoperators whose matrix entries in Stokes representation all vanish except from one entry equalling* 1.

Proof. The first inequality is derived simply by applying the triangle inequality to the 16 matrix entries of $\mathsf{T} - \mathrm{Id}$ in Stokes representation. Observe, that for trace-preserving T the number 16 can be replaced by 12 and that the constant $c_\Diamond$ can be easily computed via semi-definite programming [60]. For our purposes the actual value of the positive constant is not important. Of course, $c_\Diamond \neq 0$, since $\|\cdot\|_\Diamond$ is a valid norm.

The second inequality follows from

$$\begin{aligned}
\|\mathsf{T} - \mathrm{Id}\,\|_\Diamond &\overset{2.19}{>} \|\mathsf{T} - \mathrm{Id}\,\|_1 \\
&\overset{2.17}{\geqslant} \frac{1}{2}\|(\mathsf{T} - \mathrm{Id})(\sigma')\|_1 \\
&= \frac{1}{2}\left\| \sum_{\sigma'' \in \mathcal{P}} (\mathsf{T} - \mathrm{Id})_{\sigma''\sigma'}\sigma'' \right\|_1 \\
&\overset{2.16}{=} \frac{1}{2} \max_{-\mathbb{I} \leqslant \Lambda \leqslant \mathbb{I}} \operatorname{tr}\left[\Lambda\left(\sum_{\sigma'' \in \mathcal{P}} (\mathsf{T} - \mathrm{Id})_{\sigma''\sigma'}\sigma''\right)\right] \\
&\geqslant |(\mathsf{T} - \mathrm{Id})_{\sigma\sigma'}|
\end{aligned}$$

with the last line implied by the previous one since the maximizing set comprises the operators $\Lambda = \pm\sigma$. □

Theorem 7.29. *There is a constant* $\epsilon_\Diamond > 0$ *such that* $\mathrm{QMA}_\mathsf{T} = \mathrm{QMA}$ *for every quantum channel* T *with* $\|\mathsf{T} - \mathrm{Id}\,\|_\Diamond \leqslant \epsilon_\Diamond$.

Furthermore, $\mathrm{QMA}_\mathsf{T}(c, s) = \mathrm{QMA}_\mathsf{T}$ *for a channel* T *with the above property and for all polynomial-time computable functions* c *and* s *with* $e^{-q} \leqslant s, c \leqslant 1 - e^{-q}$, *gap* $c - s \geqslant 1/q$ *and* q *polynomial.*

Proof. To prove the equality statement and amplification at the same time, we show that $\mathrm{QMA} \subseteq \mathrm{QMA}_T(1-e^{-r}, e^{-r})$ for an arbitrary polynomial r.

Assume the setting of concatenated coding against the noise channel T as in proposition 7.25 and corollary 7.27. Since $|(T-\mathrm{Id})_{\sigma\sigma'}| \leqslant \epsilon_\Diamond$ according to lemma 7.28, choosing $\epsilon_\Diamond < \min\{1, c^{-1}\}$ ensures that the matrix entries decrease superexponentially with the number of coding levels k:

$$|(T^{(k)}-\mathrm{Id})_{\sigma\sigma'}| \leqslant \frac{1}{c}\left(c\epsilon_\Diamond\right)^{2^k} \qquad \forall \sigma, \sigma' \in \mathcal{P}.$$

Now let A be a problem in QMA and V a $\mathrm{QMA}(1-\frac{1}{2}e^{-r}, e^{-r})$ verifier for it. Let V' be the verifier V preceded by k levels of error correction and decoding. Clearly, the soundness value of V' does not increase over the soundness of V. Let $\Pi_x := (V_x)^\dagger \Pi_{\mathrm{acc}} V_x$ be the final projective measurement preceded by the circuit V_x. The completeness c' of V' can be lower bounded analogously to the derivation in proposition 6.9 with the help of the witness ρ of length n_w that achieves the highest acceptance probability for protocol V:

$$\begin{aligned}
c' &\geqslant \mathrm{tr}\left[\Pi_x\left(|0\rangle\langle 0|^{\otimes z} \otimes \left(T^{(k)}\right)^{\otimes n_w}(\rho)\right)\right] \\
&\geqslant c - \left|\mathrm{tr}\left[\Pi_x\left(|0\rangle\langle 0|^{\otimes z} \otimes \left(\left(T^{(k)}\right)^{\otimes n_w}(\rho) - \rho\right)\right)\right]\right| \\
&\overset{2.16}{\geqslant} c - \frac{1}{2}\left\||0\rangle\langle 0|^{\otimes z} \otimes \left(\left(T^{(k)}\right)^{\otimes n_w}(\rho) - \rho\right)\right\|_1 \\
&\overset{2.17}{\geqslant} c - \frac{1}{2}\left\|\mathrm{Id}^{\otimes z} \otimes \left(\left(T^{(k)}\right)^{\otimes n_w} - \mathrm{Id}^{\otimes n_w}\right)\right\|_1 \\
&\overset{2.19}{\geqslant} c - \frac{1}{2}\left\|\left(T^{(k)}\right)^{\otimes n_w} - \mathrm{Id}^{\otimes n_w}\right\|_\Diamond \\
&= c - \frac{1}{2}\left\|\left(\left(T^{(k)}\right) \otimes \mathrm{Id} \otimes \mathrm{Id} \otimes \dots\right) \circ \left(\mathrm{Id} \otimes \left(T^{(k)}\right) \otimes \mathrm{Id} \otimes \dots\right) \circ \dots - \mathrm{Id}^{\otimes n_w}\right\|_\Diamond \\
&\overset{2.19}{\geqslant} c - \frac{1}{2} n_w \left\|T^{(k)} - \mathrm{Id}\right\|_\Diamond \\
&\geqslant 1 - \frac{1}{2}e^{-r} - \frac{8c_\Diamond}{c} n_w \left(c\epsilon_\Diamond\right)^{2^k}.
\end{aligned}$$

Clearly, $k := \log n$ implies $c' \geqslant 1 - e^{-r}$ for large enough input lengths n. For smaller input lengths redefine V' to give the correct outputs deterministically. Each of the polynomially many identical error correction and decoding operations of the $[N,1]$ stabilizer code can be realized by constantly many gates. Therefore the overall error correction and decoding can be accomplished efficiently and V' is a valid $\mathrm{QMA}_T(1-e^{-r}, e^{-r})$ verifier for A. □

Note that the superexponential decreasing $\propto (c\epsilon)^{2^k}$ of the effective single qubit noise achieved by our specific concatenated stabilizer coding is even more than needed for the proof. An exponential decrease $\|T^{(k)} - \mathrm{Id}\,\|_\diamond \leqslant \alpha^k$ for a constant $\alpha \in [0,1[$ would already be sufficient to achieve a completeness of $c' = c - \frac{1}{2n^a}$ with arbitrary a for the QMA_T protocol chosing $k = \frac{\log(n_w n^a)}{\log 1/\alpha}$ coding levels. Afterwards amplification via parallel repetition can improve the completeness to a function inverse exponentially close to 1.

We mention this, since one might want to consider more general codes with a different convergence behaviour, when aiming at optimizing the noise value tolerable for $\mathrm{QMA}_T = \mathrm{QMA}$. Regarding stabilizer codes it is also possible to extend the consideration to codes with $w = 1$, if the off-diagonal elements of the noise channel are smaller than the square of the diagonal elements, since then the proofs of theorem 7.25 and corollary 7.27 still hold. Moreover, the SVD standard form from lemma 7.19 allows to restrict to noise channels with few or – in case of unital channels - even none off-diagonal entries, since a verifier can apply the respective finite dimensional unitaries between the decoding layers. All these ideas are worth considering when optimizing the bound on the noise for that $\mathrm{QMA}_T = \mathrm{QMA}$.

We will spend the rest of this section on computing (suboptimal) bounds for general channels and for the partly depolarizing and dephasing channel. Recall that these channels are of special interest since they lead to the complexity classes $\mathrm{BQP} = \mathrm{QMA}_{T_1^{\mathrm{depol}}}$ and $\mathrm{QCMA} = \mathrm{QMA}_{T_1^{\mathrm{deph}}}$ for the highest error parameter $\epsilon = 1$. Rewritten in Stokes representation the partly depolarizing and the partly dephasing channel are diagonal and equal

$$T_\epsilon^{\mathrm{depol}} = [1-\epsilon, 1-\epsilon, 1-\epsilon],$$
$$T_\epsilon^{\mathrm{deph}} = [1-\epsilon, 1-\epsilon, 1].$$

Our derived noise bounds are probably far off from optimal, since we only compare the performances of the 5-qubit, the Steane 7-qubit and the Shor code. Thanks to the structure of these codes some of the diagonal coding map entries separate. This reduces our task to computing a convergence region around the attractive fixed point 1 for continous, one-dimensional functions.

Theorem 7.30. $\mathrm{QMA}_{\mathsf{T}} = \mathrm{QMA}$ *if* T *is diagonal in Stokes representation and* $\|\mathsf{T}-\mathrm{Id}\,\|_{\diamond} \leqslant 0.18$ *or if* T *is arbitrary and* $\|\mathsf{T}-\mathrm{Id}\,\|_{\diamond} \leqslant 0.014$.

Proof. The values can be derived using the 5-qubit code as specified in table 4.1. The diagonal-reduced coding map of the 5-qubit code is

$$\begin{aligned}\Omega^{\mathsf{C}}_{\mathsf{d}}([x,y,z]) &= [f(x,y,z), f(y,z,x), f(z,x,y)]\\ f(x,y,z) &= -\frac{1}{4}x^5 + \frac{5}{4}xy^2 + \frac{5}{4}xz^2 - \frac{5}{4}xy^2z^2.\end{aligned}$$

For the partly depolarizing channel with $x = y = z$ the function f simpifies to the one-dimensional function

$$f_{1D}(x) := f(x,x,x) = \frac{5}{2}x^3 - \frac{3}{2}x^5$$

whose next fixed point below 1 turns out to equal $1 - \epsilon_1$ with $\epsilon_1 := 1 - \sqrt{\frac{2}{3}} > 0.18$. Since f is continuous and maps valid channels to valid channel matrix elements, it holds $x < f(x) \leqslant 1$ for all $x \in\,]1-\epsilon_1, 1[$. Hence, any depolarizing channel $\mathsf{T}^{\mathrm{depol}}_{\epsilon}$ with $\epsilon < \epsilon_1$ will converge towards the identity under concatenated 5-qubit codes.

Regarding the convergence speed, we know by corollary 7.27 that there exists an $\epsilon_0 > 0$ such that an effective channel series of $\mathsf{T}^{\mathrm{depol}}_{\epsilon}$ with $\epsilon \leqslant \epsilon_0$ converges superexponentially towards the identity channel, which is sufficient for theorem 7.29 to prove $\mathrm{QMA}_{\mathsf{T}^{\mathrm{depol}}_{\epsilon}} = \mathrm{QMA}$. A channel $\mathsf{T}^{\mathrm{depol}}_{\epsilon}$ with $\epsilon \in\,]\epsilon_0, \epsilon_1[$ is turned by constantly many coding layers into an effective channel $\mathsf{T}^{\mathrm{depol}}_{\epsilon'}$ with $\epsilon' \leqslant \epsilon_0$.

Regarding an arbitrary diagonal, trace preserving superoperator T with elements within the interval $[1-\epsilon, 1]$ for any $\epsilon < \epsilon_1$, it is important to observe that f guarantees

$$\left(\Omega^{\mathsf{C}}(\mathsf{T}^{\mathrm{depol}}_{\epsilon})\right)_{\sigma\sigma} \leqslant \left(\Omega^{\mathsf{C}}(\mathsf{T})\right)_{\sigma\sigma} \leqslant 1 \qquad \forall \sigma \in \mathcal{P}.$$

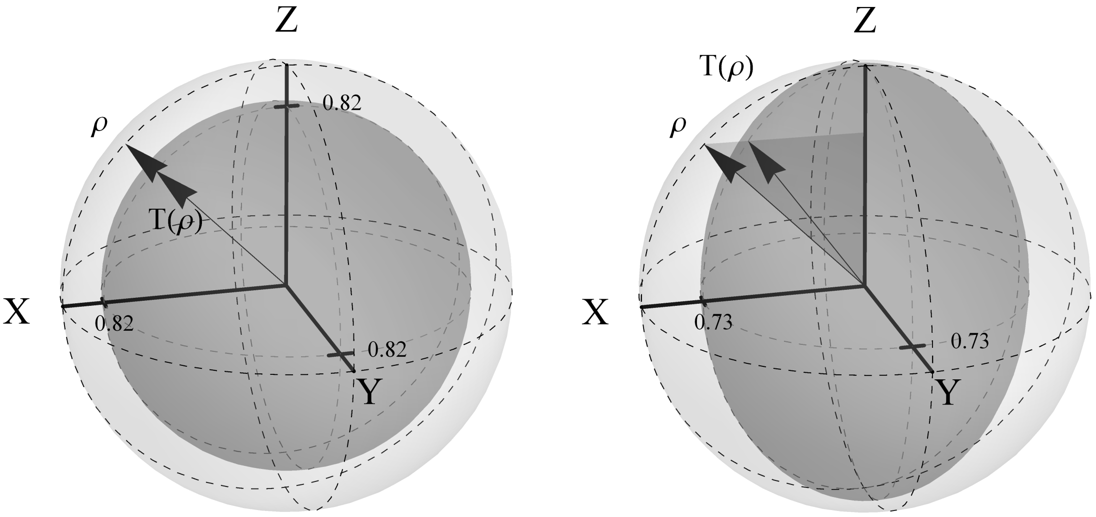


Figure 7.2: Bloch sphere representation of 18% depolarizing noise (left) and 27% dephasing noise (right), which does not diminish the power of QMA when applied to each witness qubit.

Hence, $\mathrm{QMA}_\mathsf{T} = \mathrm{QMA}$ for every diagonal channel T with $|(\mathsf{T} - \mathrm{Id})_{\sigma\sigma}| < \epsilon_1$. According to theorem 7.25 this is also true for every non-diagonal channel T with $|(\mathsf{T} - \mathrm{Id})_{\sigma\sigma}| \leqslant 0.014 < (c_\mathsf{F} + \epsilon_1^{-1})^{-1}$ with the constant $c_\mathsf{F} \leqslant 64$ of theorem 7.24.

The bounds on the diamond norm follow by $|(\mathsf{T} - \mathrm{Id})_{\sigma\sigma}| \leqslant \|\mathsf{T} - \mathrm{Id}\,\|_\diamond$. □

Theorem 7.31.

$$\mathrm{QMA}_{\mathsf{T}_\epsilon^{\mathrm{depol}}} = \begin{cases} \mathrm{QMA} & \textit{for } \epsilon \leqslant 0.18 \\ \mathrm{BQP} & \textit{for } \epsilon = 1 \end{cases}$$

$$\mathrm{QMA}_{\mathsf{T}_\epsilon^{\mathrm{deph}}} = \begin{cases} \mathrm{QMA} & \textit{for } \epsilon \leqslant 0.27 \\ \mathrm{QCMA} & \textit{for } \epsilon = 1. \end{cases}$$

Proof. The bound on ϵ guaranteeing $\mathrm{QMA}_{\mathsf{T}_\epsilon^{\mathrm{depol}}} = \mathrm{QMA}$ has already been derived via the 5-qubit code in the proof of theorem 7.30.

The higher bound on ϵ guaranteeing $\mathrm{QMA}_{\mathsf{T}_\epsilon^{\mathrm{deph}}} = \mathrm{QMA}$ can be derived by the Shor code with interchanged logical $\bar{X}$ and $\bar{Z}$ operation compared to table 4.1. Its diagonal-reduced coding map is of the form

$$\Omega_d^C([x, y, z]) = [f(x), g(x, y, z), h(z)]$$
$$f(x) = \left(\frac{3x}{2} - \frac{x^3}{2}\right)^3 .$$

The next fixed point below 1 of the function f is slightly below 0.73. Furthermore, we already start at the attracting point 1 of the function h with the dephasing channel $T_\epsilon^{\mathrm{deph}} = [1-\epsilon, 1-\epsilon, 1]$. By lemma 7.18 the convergence of two diagonal channel elements towards 1 is a sufficient criterion for all three converging towards 1. With the same argumentation as before a channel $T_\epsilon^{\mathrm{deph}}$ with $\epsilon < 0.27$ therefore converges superexponentially towards the identity under concatenated Shor codes, proving $\mathrm{QMA}_{T_\epsilon^{\mathrm{deph}}} = \mathrm{QMA}$ via theorem 7.29. □

CHAPTER 8

QPCP and a 3-prover protocol for the Local Hamiltonian problem

8.1 Introduction

The intention behind the introduction of the noisy QMA framework was to study the significance of the witness for verification protocols. The famous concept of PCP ("probabilistically checkable proofs") can also be phrased in these terms: How well can a problem be verified if only a certain number of (qu)bits of the witness are accessed probabilistically? Clearly, the problems SAT and k-LH can be verified probabilistically by simply evaluating one clause or one interaction term, respectively. The probability gap between completeness and soundness scales inverse polynomially with the input length in this case. What is extremely surprising is that SAT – and hence any NP problem – can be verified even with a constant probability gap by accessing only constantly many witness bits! This is the statement of the famous PCP theorem which was proven after a long line of work in the 90s by Arora and Safra [61] and reproven with a new ansatz in 2007 by Irit Dinur [62].

What about an analogous result for QMA, the quantum analogue of NP? In 2006 the well-known complexity theorist Scott Aaronson first raised the question about the existence of a quantum PCP theorem on his blog [63]:

> "I'm 99% sure that this theorem (alright, conjecture) or something close to it is true. I'm 95% sure that the proof will require a difficult adaptation of classical PCP machinery (whether Iritean or pre-Iritean), in much the same way that the Quantum Fault-Tolerance theorem required a difficult adaptation of classical fault-tolerance machinery. I'm 85% sure that the proof is achievable in a year or so, should enough people make it a priority. I'm 75% sure that the proof, once achieved, will open up heretofore undreamt-of vistas of understanding and insight. I'm 0.01% sure that I can prove it. And that is why I hereby bequeath the actual proving part to you, my readers."

Now, over a decade later, it is clear that the proof of a QPCP theorem did not take a year; despite much effort, it is still open until today. Many steps in the two proofs of the classical PCP theorem seem difficult or even impossible to quantize, as also Aaronson commented later on his own blog entry:

> "I'm quite certain that a Quantum PCP theorem will require significant new ideas. Recently I spent a day or two studying Irit's proof of the classical PCP theorem (which I hadn't done before), and I found about 20 violations of the No-Cloning theorem on every page."

We can fomulate the QPCP conjecture as QMA $=$ QMA$_{\mathcal{T}}$ with $\mathcal{T}$ the channel family that erases all but constantly many random witness qubits. We will soon see that one can also equivalently consider the situation in that every witness qubit is disturbed independently by identical erasure channels with error parameter $1-\frac{k}{n_w}$. In this situation only the expected number of undisturbed qubits equals the constant k. This formulation fits nicely into our line of arguments according to which QMA remains invariant

> with decreasing error parameter $\frac{k}{n_w}$ due to amplifcation,
>
> with non-trivial constant error parameter due to concatenated coding and
>
> with increasing error parameter $1-\frac{k}{n_w}$ due to QPCP.

Note that, if the QPCP conjecture holds, a $\mathrm{QMA}_{\epsilon}^{\mathrm{eraa}}$ class cannot be QMA-intermediate unless the error parameter ϵ increases faster than $1 - \frac{k}{n_w}$. We know by theorem 7.31 that the constant error parameter of erasure noise up to which QMA remains invariant due to concatenated coding equals at least 18% since the erasure channel transforms into the depolarizing channel when discarding the flag qubits.

The reason why researchers pretty much agree on the correct formulation of the QPCP conjecture [64, 65] is its equivalence to the important statement that a k-Local Hamiltonian problem with a constant relative energy gap $\frac{a-b}{m}$ is still QMA-complete. This stands in strong analogy to the classical PCP conjecture, which is equivalent to the NP-completeness of the simplified 3-SAT problem with the promise that either all or at most a constant fraction (e.g. 90%) of the clauses is satisfiable. We call this statement the constraint satisfaction variant of the PCP theorem.

There also exists a third equivalent formulation of the classical PCP theorem in terms of multiprover protocols, which has not been proven for the quantum PCP conjecture yet. Fitzsimons and Vidick [3] made the first approach in this direction and provided a multiprover protocol for the k-Local Hamiltonian problem. Yet, this protocol does not prove the implication of a reasonable multiprover QPCP statement due to an additional inverse polynomial in the probability gap.

In the first sections of this chapter we will present and prove the equivalences of the different formulations of the PCP theorem and the QPCP conjecture, including the noisy QMA version. Afterwards, we will revisit the multiprover protocol for the Local Hamiltonian problem by Fitzsimons and Vidick [3] for two purposes: Firstly, by adapting the underlying quantum code we reduce the original number of 5 provers to the minimum possible number of 3 for this protocol structure. Secondly, we manage to decrease the polynomial that determines the probability gap, although a remaining term still prevents that the QPCP conjecture implies a reasonable multiprover variant.

8.2 The classical PCP theorem

Definition 8.1. *A problem is in the complexity class* $\mathrm{PCP}(r, q, c, s)$ *("probabilistically checkable proof") iff it has a probabilistic polynomial time verifier of completeness* c *and soundness* s *that* *uses at most $\mathcal{O}\left(r(n)\right)$ coin tosses and reads at most $\mathcal{O}\left(q(n)\right)$ witness bits (precisely, this means the verifier can use $\mathcal{O}\left(q(n)\right)$ queries to an oracle tape which returns the value of the i-th witness bit for query i).*

Note that in the above definition the verifier can choose *adaptively* which witness bits he wants to read, i.e. the choice of the witness bit i is allowed to depend on the values of the previously read witness bits $1, \dots, i-1$. Of course, it is also conceivable to demand a non-adaptive choice of witness bits, i.e. the verifier has to decide on all witness bits he wants to read before reading the value of the first. Most literature uses the above definition for $\mathrm{PCP}(r, q, c, s)$, but some literature, e.g. [13], demands non-adaptive choices. Notice that a protocol reading q witness bits adaptively can easily be changed into a protocol reading

$$q' = \sum_{i \in [q]} 2^{i-1} = 2^q - 1$$

witness bits non-adaptively. Hence, the question of allowing adaptive or demanding non-adaptive choices of the witness bits in the PCP definition is actually irrelevant for the formulation of the PCP theorem which only demands that q is a constant.

Theorem 8.2 (PCP theorem). *$\exists$ constant $s \in]0,1[$ such that $\mathrm{NP} \subseteq \mathrm{PCP}(\log, 1, 1, s)$.*

For the other two equivalent formulations of the PCP theorem we need the following definitions:

Definition 8.3. *We define the problem 3-SAT(t) as 3-SAT with the additional promise that either all or at most a fraction t of the clauses is satisfiable.*

Definition 8.4. *An $\mathrm{MIP}(l, r)$ verifier ("multiprover interactive protocol") is a probabilistic polynomial time algorithm that has l rounds of communication with each of r non-communicating provers. The i-th round of communication with prover j starts with the verifier writing a question q_i^j from a question set Q_i^j onto a prover specific message tape upon which the prover overwrites it with the answer $f_i^j(q_1^j, q_2^j, \dots, q_i^j)$ according to an answer function $f_i^j : Q_1^j \times \dots \times Q_i^j \to A_i^j$, which depends implicitly on the input x. A specific collection of answer functions $(f_i^j)_{i \in [l], j \in [r]}$ is called a* strategy *of the provers.*

A problem $A = (A_{\text{yes}}, A_{\text{no}})$ *lies in the complexity class* $\mathrm{MIP}(l, r, c, s)$ *iff there exists an* $\mathrm{MIP}(l, r)$ *verifier* V *such that*

- $\forall x \in A_{\text{yes}}$ *there exists a strategy of the provers such that* V *accepts with probability at least* c *(completeness),*

- $\forall x \in A_{\text{no}}$ *and for all provers' strategies* V *accepts with probability at most* s *(soundness).*

A strategy is the best *or* optimal *one for an input* x *iff it leads to the highest possible acceptance probability.*

Definition 8.5. $\mathrm{MIP}(l, r, c, s)_{\text{log,const}}$ *equals the complexity class* $\mathrm{MIP}(l, r, c, s)$ *with the restriction that at most logarithmically many coins are tossed and that the answers of the provers have constant size.*

Proposition 8.6. *The following statements are equivalent:*

1. *(PCP version).* $\exists$ *constant* $s \in]0, 1[$ *such that* $\mathrm{NP} = \mathrm{PCP}(\log, 1, 1, s)$.

2. *(CSP version).* $\exists$ *constant* $t \in]0, 1[$ *such that 3-*$\mathrm{SAT}(t)$ *is NP-complete.*

3. *(Multiprover version).* $\exists$ *constant* $s \in]0, 1[$ *such that* $\mathrm{NP} \subseteq \mathrm{MIP}(2, 1, 1, s)_{\text{log,const}}$.

Proof. (1 ⇒ 2). Assume there exists a constant $s \in]0, 1[$ such that $\mathrm{NP} = \mathrm{PCP}(\log, 1, 1, s)$. Let V be the $\mathrm{PCP}(\log, 1, 1, s)$ verifier for 3-SAT that, for input x of length n, tosses $r(n) \in \mathcal{O}(\log n)$ random coins with outcome R and reads $q \in \mathcal{O}(1)$ many bits $Q_{x,R}$ from the witness. Since the verifier can thus access maximally $q2^{r(n)}$ different witness bits, we can assume w.l.o.g. that the witness is of polynomial length.

Let V_R denote the NP verifier that equals V up to the fact that the random coin outcomes are replaced by the fixed assignment R. The efficient Cook-Levin construction provides us with a Boolean formula $\phi_{x,R}$ (w.l.o.g. in 3-CNF) in the variables $\{y_i\}_{i \in Q_{x,R}}$ which is satisfiable iff there exists a witness such that V_R accepts input x.

Consequently, if V accepts input x with certainty, the conjunction formula $\phi_x := \bigwedge_{R \in \Sigma^{r(|x|)}} \phi_{x,R}$ is satisfiable.

If, on the other hand, V accepts input x only with probability at most s, then any witness maximally satisfies a fraction s of the $\phi_{x,R}$. Each $\phi_{x,R}$ can consist of several clauses, but the maximal number of clauses l is a constant, since $\phi_{x,R}$ only contains constantly many variables. Consequently, if at most a fraction s of the $\phi_{x,R}$ is satisfied, this implies that at most a constant fraction $t := s + (1-s)(1-\frac{1}{l}) \in]0,1[$ of the clauses of ϕ_x is satisfied.

Let $A \in \mathrm{NP}$ be arbitrary and f its Karp reduction onto 3-SAT. Then $x \to \phi_{f(x)}$ is a valid Karp reduction from A to 3-SAT(t).

(2 ⇒ 3). Assume $\exists$ constant $\epsilon \in]0,1[$ such that 3-SAT$(1, 1-\epsilon)$ is NP-complete. We will show that the following is a valid $\mathrm{MIP}(2,1,1,1-\frac{\epsilon}{3})_{\text{log,const}}$ protocol for 3-SAT$(1,1-\epsilon)$:

- Given a Boolean formula in 3-CNF, choose a clause at random and afterwards a variable of that clause at random.
- Ask prover 1 for an assignment of the clause and prover 2 for an assignment of the variable.
- Accept iff the two answers are consistent and the clause is satisfied by the assignment.

Clearly, if a Boolean formula is satisfiable, then the provers can simply supply answers according to the satisfying assignment and the above protocol accepts with certainty.

Assume, on the other hand, that at most $(1-\epsilon)$ of the clauses of a Boolean formula in 3-CNF can be satisfied. Then also the assignment composed of the second prover's best strategy answers can satisfy at most $(1-\epsilon)$ of the clauses. The acceptance probability of the above protocol is consequently at most $(1-\epsilon) + \epsilon\frac{2}{3} = 1 - \frac{\epsilon}{3}$.

(3 ⇒ 1). Consider an $\mathrm{MIP}(2,1,1,s)_{\text{log,const}}$ protocol. Let Q^1, Q^2 be the polynomially sized sets of questions that prover 1 and 2 can be asked and A^1, A^2 be the respective constant sized answer sets. The protocol can easily be transformed into a $\mathrm{PCP}(\log, 1, 1, s)$ protocol by expecting the witness to consist of all the answers according to the best strategy for questions from Q^1 and Q^2, tossing the logarithmically many coins and checking the respective constantly many witness bits. □

We did not specify the constants for which the statements of proposition 8.6 hold. In fact, the PCP and multiprover version hold for any constant soundness $0 < s < 1$, since it can be decreased to another arbitrary constant soundness $0 < s' < s$ by carrying out the PCP or MIP protocol with perfect completeness constantly many times and accepting iff all repetitions accept. Clearly, for i repetitions the soundness of a PCP$(\log, 1, 1, s)$ protocols goes down to $s' = s^i$.

For MIP$(2, 1, 1, s)$ the analogous assumption is wrong as the counter example of Feige's NA game ("Noninteractive Agreement") shows [66]. The critical issue is that by keeping the protocol a single round game the provers can exploit their knowledge of all questions before answering. Still, Raz [67] proved by his parallel repetition theorem that the soundness of any MIP$(2, 1, 1, s)$ goes down exponentially with the number of parallel repetitions, just the basis of the scaling might be different. For the NA game, which has soundness $\frac{1}{2}$, the soundness decreases for example to $\left(\frac{1}{\sqrt{2}}\right)^{2i}$ for $2i$ protocol rounds instead of the naively expected $\left(\frac{1}{2}\right)^{2i}$ [66].

The CSP formulation of the PCP conjecture holds for any constant $t > \frac{7}{8}$ as Håstad proved [68]. Interestingly, $\frac{7}{8}$ is exactly the fraction of clauses of a 3-SAT formula that any assignement at least satisfies. Consequently, 3-SAT (t) is already trivially in P for $t < \frac{7}{8}$.

8.3 The quantum PCP conjecture

In analogy to the classical complexity class we can define a class for quantumly probabilistically checkable proofs [64, 65]:

Definition 8.7. *A problem* $(A_{\text{yes}}, A_{\text{no}})$ *lies in the complexity class* QPCP(q, c, s) *iff it has a quantum polynomial time verifier of completeness* c *and soundness* s *that has only uniformly random access to* $\mathcal{O}(q)$ *out of polynomially many witness qubits (precisely, this means that the generating Turing machine has as additional input a set* S *of* $\mathcal{O}(q)$ *uniformly random witness indices and outputs a quantum circuit* V_x^S *that only acts on the witness qubits from* S*).*

In contrast to a classical PCP verifier, a quantum PCP verifier cannot decide himself which witness qubits to access upon some random coin tosses. Instead, he is restricted to a uniformly random access. Any other definition usually results in some kind of unnatural classical-quantum mixture regarding the witness access. A disadvantage of the random witness access is that standard amplification by parallel repetition gets more complicated.

An advantage on the other hand is that an equivalence between a probabilistically checkable proof version and a constraint satisfaction version of a natural quantum PCP conjecture holds. The following equivalence proof relies upon the sketch of [65] and the more detailed lecture notes by [69].

Conjecture 8.8 (Quantum PCP conjecture)**.**

1. *(PCP version).* $\exists c$, $\exists s$ *with* $c - s > 0$ *constant such that* $\mathrm{QMA} = \mathrm{QPCP}(1, c, s)$*.*

2. *(CSP version).* $\exists k$ *constant,* $\exists a$ *and* $\exists t > 0$ *constant such that the problem of deciding if a* k*-local Hamiltonian with* m *interaction terms has ground state energy at most* a *(yes instance) or at least* $a + tm$ *(no instance) is* QMA*-compete under quantum polynomial time reductions.*

Proposition 8.9. *The two versions of the QPCP conjecture 8.8 are equivalent.*

Proof. (1 ⇒ 2). Assume $\exists c$, $\exists s$ with $c - s > 0$ constant such that $\mathrm{QMA} = \mathrm{QPCP}(1, c, s)$. Let $(V_x^S)_{x \in \Sigma^*}^{S \subseteq [n_w], |S| = q}$ be the $\mathrm{QPCP}(1, c, s)$ verifier for the QMA-complete problem k-LH working on z ancilla qubits and $q \in \mathcal{O}(1)$ qubits from the witness of polynomial length p. We want to map a k-LH instance x onto the q-local Hamiltonian H that consists of the positive semi-definite $m = \binom{p}{q}$ interaction terms

$$H_S \coloneqq \langle 0^z | \, (V_x^S)^\dagger \Pi_0 V_x^S \, | 0^z \rangle$$

of norm $\|H_S\| \leqslant 1$. Clearly, H has ground state energy at most sm for a yes-instance and at least cm for a no-instance. Hence, to prove the CSP version with $k \coloneqq q$, $a \coloneqq s$ and $t \coloneqq c - s$ it only remains to show that H_S can be computed in quantum polynomial time.

The approach is a variant of quantum process tomography [4, §8.4.2]: Simulating V_x^S and the output measurement on the fixed zero ancilla and q witness qubits corresponds exactly to measuring the observable H_S with regard to the witness state reduced to the constantly sized register S. Hence, polynomially many such simulations for a reduced witness state $|n\rangle$, $n \in \Sigma^{2^q}$, lead to an approximation of the diagonal matrix element $\langle n| H_S |n\rangle$ within arbitrary polynomial accuracy according to the central limit theorem.

Off-diagonal matrix elements $\langle n| H_S |m\rangle$ can simply be computed via four diagonal elements:

$$\begin{aligned} \langle n| H_S |m\rangle = & \frac{\langle n| + \langle m|}{\sqrt{2}} H_S \frac{|n\rangle + |m\rangle}{\sqrt{2}} + i \frac{\langle n| + i \langle m|}{\sqrt{2}} H_S \frac{|n\rangle - i |m\rangle}{\sqrt{2}} \\ & - \frac{1+i}{2} \langle n| H_S |n\rangle - \frac{1+i}{2} \langle m| H_S |m\rangle . \end{aligned}$$

(2 ⇒ 1). Assume there exists a constant k, an a and a constant $t > 0$ such that the problem of deciding whether a k-local Hamiltonian with m interaction terms has ground state energy at most a (yes-instance) or at least $a + tm$ (no-instance) is QMA-complete under quantum polynomial time reductions.

The QMA protocol with completeness $c := 1 - \frac{a}{m}$ and soundness $s := 1 - \frac{a}{m} - t$ for this k-LH problem described in section 3.5 is basically equivalent to a QPCP$(1, c, s)$ protocol; the verifier's uniformly random choice of the witness qubits is simply replaced by the QPCP protocol's inherent choice. □

It appears unusual, but not necessarily wrong, that the construction of the above q-Local Hamiltonian from the QPCP$(1, c, s)$ protocol for the LH problem requires a quantum instead of a classical polynomial time reduction. The reason is that a quantum algorithm can simulate the circuit V_x^S efficiently whereas a classical algorithm cannot apply the exponentially sized matrix efficiently.

Although the above equivalence shows a strong analogy to the classical PCP theorem the goal of proving or disproving it seems far off. Works on the QPCP conjecture so far mainly succeeded in restricting the form of the Hamiltonian family for which the CSP version can hold. The following is one result of this kind to which we will return in sections 8.6 and 8.7:

Theorem 8.10. *Assuming* NP $\neq$ QMA, *the CSP version of the QPCP conjecture 8.8 can be restricted to* 2*-Local Hamiltonians of constant degree, i.e. each qubit is involved non-trivially in only constantly many interaction terms.*

Proof. The theorem is a conclusion of two works: [70] provide for every $\epsilon > 0$ an efficient mapping from any k-local Hamiltonian on n qubits to a 2-local Hamiltonian by introducing an error of $\mathcal{O}(\epsilon n)$ to the ground state energy. Hence, a constant relative energy gap of the Hamiltonian persists. Moreover, by studying product state approximations of ground states, [71, corollary 5] proved that the problem of deciding whether a 2-local Hamiltonian of degree D has at most energy a or at least energy $a + tm$ with m the number of interaction terms is in NP for any constant $t > 0$ and D above some constant. □

8.4 Noisy QMA version of the quantum PCP conjecture

A QPCP$(1, c, s)$ protocol can easily be understood as a noisy QMA protocol that destroys and indicates all but $k = \mathcal{O}(1)$ witness qubits uniformly at random. For the equivalence of the two classes we hence need to consider the channel

$$T_{n_w} = \binom{n_w}{k}^{-1} \sum_{\substack{s \in \{0,1\}^{n_w},\, i \in [n_w] \\ w(s) = k}} \bigotimes \left((T_0^{\mathrm{eras}})^{s_i} + (T_1^{\mathrm{eras}})^{1-s_i} \right)_i$$

which is the equal mixture of all channels that are composed of the non-erasing channel on exactly k qubits and the maximal erasing channel on the remaining n_w qubits.

Note that the equivalence would not hold for the suspicious class $\mathrm{QMA}_{\mathcal{T}}^{s}$, since its malicious prover could send a state with specific $n - k$ erased witness qubits leading to an acceptance probability above the soundness value of the QPCP$(1, c, s)$ protocol which only has to be attained for randomly erased states.

Unfortunately, the channel T_{n_w} is not of i.i.d. form. A reasonable alternative for the QPCP conjecture is to demand the equivalence of QMA to the noisy QMA class with i.i.d. erasure noise $T_\epsilon^{\mathrm{eras}}$ such that the expected number of undisturbed qubits equals a constant k. Since the probability that exactly i witness qubits are not erased is given by the binomial distribution $\binom{n_w}{i}(1-\epsilon)^i \epsilon^{n_w - i}$ with with mean value $(1-\epsilon) n_w$, this is accomplished by the channel $T_\epsilon^{\mathrm{eras}}$ with $\epsilon = 1 - \frac{k}{n_w}$.

Requiring a constant expectation number of accessible qubits is in a certain way weaker than requiring exactly k accessible qubits. Still we can show in the proposition below that both conjectures are equivalent by applying Chernoff bound. Hence, the robustness of QMA against $\epsilon = 1 - \frac{k}{n_w}$ increasing i.i.d. erasure noise constitutes a third physically interesting, equivalent formulation of the prevailing QPCP conjecture.

A stronger conjecture for that an equivalence is not obvious is obtained by eliminating the flag qubits and conjecture $\mathrm{QMA} = \mathrm{QMA}_{T_\epsilon^{\mathrm{depol}}}(c, s)$ for the depolarizing channel $T_\epsilon^{\mathrm{depol}}$ with $\epsilon = 1 - \frac{k}{n_w}$. One might favor this conjecture over the prevailing QPCP conjecture, since the depolarizing channel is a thoroughly quantum channel which is liberated from the classical feature of marking the accessible witness qubits. We propose $\mathrm{QMA} = \mathrm{QMA}_{T_p^{\mathrm{depol}}}(c, s)$ with $\epsilon = 1 - \frac{k}{n_w}$ for a constant k and c and s such that $c - s > 0$ constant as a reasonable and thouroughly quantum alternative for a QPCP conjecture. Unfortunately, we are missing evidences for expressing a profound intuition about its possible validity or invalidity.

Conjecture 8.11 (Noisy QMA Quantum PCP conjectures)**.**

1. *(Non-i.i.d. erasure noise version). $\exists k$ constant, $\exists c$, $\exists s$ with $c - s > 0$ constant such that $\mathrm{QMA} = \mathrm{QMA}_{\mathcal{T}}(c, s)$ with $\mathcal{T} = (T_{n_w})_{n_w \in \mathbb{N}_0}$ the channel family defined for $n_w \geqslant k$ by*

$$T_{n_w} = \binom{n_w}{k}^{-1} \sum_{\substack{s \in \{0,1\}^{n_w},\\ w(s)=k}} \bigotimes_{i \in [n_w]} \left((T_0^{\mathrm{eras}})^{s_i} + (T_1^{\mathrm{eras}})^{1-s_i} \right)_i .$$

2. *(I.i.d. erasure noise version). $\exists k$ constant, $\exists c$, $\exists s$ with $c - s > 0$ constant such that $\mathrm{QMA} = \mathrm{QMA}_{T_\epsilon^{\mathrm{eras}}}(c, s)$ with $\epsilon = 1 - \frac{k}{n_w}$.*
3. *(I.i.d. depolarizing noise version). $\exists k$ constant, $\exists c$, $\exists s$ with $c - s > 0$ constant such that $\mathrm{QMA} = \mathrm{QMA}_{T_\epsilon^{\mathrm{depol}}}(c, s)$ with $\epsilon = 1 - \frac{k}{n_w}$.*

Proposition 8.12. *The non-i.i.d. erasure noise version of the QPCP conjecture 8.11 is equivalent to the QPCP conjecture 8.8.*

Proof. *"$\Longrightarrow$".* Since the output of a channel T_{n_w} equals a mixture of states with flag qubits in the computational basis, we can assume without changing the acceptance probability that each $\mathrm{QMA}_{\mathcal{T}}(c,s)$ protocol initially measures the flag qubits of the witness in the computational basis and that the following circuit, denoted by V_x^S, just accesses the input witness qubits measured with flag $|0\rangle$ defining the set S. Instead of acting on the witness qubits measured with flag $|1\rangle$ the circuit can act on auxillary qubits initialized in the completely mixed state. Because each k-tuple S is measured with equal probability, $(V_x^S)_{x\in\Sigma^*}^{S\subseteq[n],|S|=k}$ is clearly a proper $\mathrm{QPCP}(1,c,s)$ verifier with k accessible qubits.

"$\Longleftarrow$". If $\mathrm{QMA} = \mathrm{QPCP}(1,c,s)$, any $\mathrm{QPCP}(1,c,s)$ protocol can be simulated by a $\mathrm{QPCP}(1,c,s)$ protocol with the number of accessible qubits restricted to the number k of accessible qubits needed for verifying the Local Hamiltonian problem. A $\mathrm{QPCP}(1,c,s)$ protocol V_x^S with k accessible qubits can clearly be transformed into a $\mathrm{QMA}_{\mathcal{T}}(c,s)$ protocol by first measuring the flag qubits and then applying V_x^S with S the set of the k input qubits measured with flag $|0\rangle$. □

Proposition 8.13. *The i.i.d. erasure noise version of the QPCP conjecture 8.11 is equivalent to the QPCP conjecture 8.8.*

Proof. "$\Longrightarrow$". Assume $\mathrm{QMA} = \mathrm{QMA}_{T_\epsilon^{\mathrm{eras}}}(c,s)$ with $\epsilon = 1 - \frac{k}{n_w}$ for a constant k and a c and a s such that $c - s \geqslant 0$ constant. Choose $\delta > 0$ small enough such that $c - \delta - s$ is still lower bounded by positive constant. Define $k' := \alpha k$ with a constant $\alpha > 1$ chosen large enough such that $e^{-\frac{(\alpha-1)^2 k}{2}} \leqslant \delta$.

Let $(V_x)_{x\in\Sigma^*}$ be the $\mathrm{QMA}_{T_\epsilon^{\mathrm{eras}}}(c,s)$ verifier for an arbitrary QMA problem A. Consider the following QPCP protocol $V = (V_x^{S'})_{x\in\Sigma^*}^{S'\subseteq[n],|S'|=k'}$ with k' accessible qubits for the problem A:

- ○ Toss n_w biased coins (value 0 with probability ϵ) with outcome string c.
- ○ If $w(c) > k'$, reject. Otherwise, choose a set S of $w(c)$ witness qubits out of the set S' of k' accessible qubits uniformly at random.
- ○ Construct a "noisy witness" by tensoring each witness qubit from S by a flag qubit $|0\rangle$, interleaved with qubit pairs $\frac{\mathbb{I}}{2} \otimes |1\rangle\langle 1|$ at the $n_w - w(c)$ positions $[n_w]\backslash S$.
- ○ Simulate V_x on this noisy witness.

The above protocol constructs a specific noisy witness with $s \leqslant k'$ undisturbed qubits with the same probability as the erasure channel family leaves these qubits undisturbed:

$$\begin{aligned}
&\mathbb{P}[V_x^{S'} \text{ chooses a specific set } A \text{ of } s \leqslant k' \text{ qubits}]\\
&\qquad = \mathbb{P}[w(c) = s] \cdot \mathbb{P}[A \subseteq S'] \cdot \mathbb{P}[s \text{ specific qubits are chosen from } S']\\
&\qquad = \binom{n_w}{s}(1-\epsilon)^s \epsilon^{n_w - s} \cdot \frac{\binom{n_w - s}{k'-s}}{\binom{n_w}{k'}} \cdot \frac{1}{\binom{k'}{s}}\\
&\qquad = (1-\epsilon)^s \epsilon^{n_w - s}.
\end{aligned}$$

Hence, while the soundness of V is at most s, its completeness is lower bounded by $c - \mathbb{P}[w(c) > k']$. Using Chernoff bound 3.19 and $q := \frac{k}{n_w}$ we can bound

$$\begin{aligned}
\mathbb{P}[w(c) > k'] &= \sum_{i=k'+1}^{n_w} \binom{n_w}{i} q^i (1-q)^{n_w - i}\\
&\leqslant e^{-\frac{(k'-k)^2}{2k}}\\
&\leqslant e^{-\frac{(\alpha-1)^2 k}{2}}\\
&\leqslant \delta.
\end{aligned}$$

Hence, V is a valid $\mathrm{QPCP}(1, c-\delta, s)$ verifier for A and $\mathrm{QMA} = \mathrm{QPCP}(1, c-\delta, s)$.

"$\Longleftarrow$". Assume $\mathrm{QMA} = \mathrm{QPCP}(1, c, s)$ for a c and a s such that $c - s \geqslant 0$ constant. Let $(V_x^S)_{x \in \Sigma^*}^{S \subseteq [n], |S| = k}$ be the $\mathrm{QPCP}(1, c, s)$ verifier with k accessible witness qubits for the Local Hamiltonian problem. Choose $\delta > 0$ small enough such that $c - \delta - s$ is still larger than a positive constant. Define $k' := \alpha k$ with a constant $\alpha > 2$ chosen large enough such that $e^{-\frac{(\alpha-1)k}{4}} \leqslant \delta$.

Consider the following $\mathrm{QMA}_{T_\epsilon^{\text{eras}}}$ protocol $V = (V_x)_{x \in \Sigma^*}$ with $\epsilon = 1 - \frac{k'}{n_w}$ for the Local Hamiltonian problem:

- Measure the flag qubits to determine the set S' of non-erased witness qubits.
- If $|S'| < k$, reject. Otherwise, choose a set S of k witness qubits out of the set S' uniformly at random.

∘ Simulate V_x^S on the witness after discarding the flag qubits.

Clearly, the probability that the above protocol determines a specific set S is the same for all k-tuple sets. Apart from the case $|S'| < k$, V_x simulates hence V_x^S with uniformly random S. Consequently, the soundness of V is at most s while its completeness is at least $c - \mathbb{P}[|S'| < k]$. Using Chernoff bound 3.19 and $q = \frac{k'}{n_w}$ we can bound again

$$\begin{aligned}\mathbb{P}[|S'| < k] &= \sum_{i=0}^{k-1} \binom{n_w}{i} q^i (1-q)^{n_w - i} \\ &\leqslant e^{-\frac{(k'-k)^2}{2k'}} \\ &\leqslant e^{-\frac{(\alpha-1)^2 k}{2\alpha}} \\ &\leqslant e^{-\frac{(\alpha-1)k}{4}} \\ &\leqslant \delta.\end{aligned}$$

Hence, V is a valid $\mathrm{QMA}_{T_\epsilon^{\mathrm{eras}}}(c-\delta, s)$ verifier for the Local Hamiltonian problem and $\mathrm{QMA} = \mathrm{QMA}_{T_\epsilon^{\mathrm{eras}}}(c-\delta, s)$. □

Proposition 8.14. *The i.i.d. depolarizing noise version of the QPCP conjecture 8.11 implies the i.i.d. erasure noise version.*

Proof. This is trivial, since the erasure channel equals the depolarizing channel when ignoring the flag qubits. □

8.5 A multiprover version of the quantum PCP conjecture

In section 8.3 we showed that the prevailing QPCP conjecture possesses a PCP and a CSP variant, but an equivalent multiprover variant like in the classical case has not been found until today. Already formulating a one round multiprover protocol with constantly sized answers for the Local Hamiltonian problem obeying an inverse polynomial gap for the acceptance probability is highly non-trivial. Such a protocol was not found until 2014 by [3]. Before presenting this protocol, we introduce the quantum classes of multiprover interactive protocols:

Definition 8.15. *The complexity class* $\mathrm{QMIP}(l, r, c, s)$ *equals the complexity class* $\mathrm{MIP}(l, r, c, s)$ *with the difference that verifier and provers are quantum, i.e. each prover possesses a private register initialized in the all zero state and a prover's action corresponds to a unitary operator applied to his private and his personal message register. A strategy hence equals a set of unitary operators* $(U_{l'}^{r'})_{l' \in [l], r' \in [r]}$*. (Of course, the operators also depend on the input* x*, but for readability we neglect this index from now on.)*

$\mathrm{QMIP}_{\mathrm{le}}(l, r, c, s)$ *equals* $\mathrm{QMIP}(l, r, c, s)$ *with polynomially sized parts of the provers' private registers initialized in an entangled state* $|\psi\rangle$ *("limited entanglement"). A strategy is hence the tuple* $(U_{l'}^{r'}, |\psi\rangle)_{l' \in [l], r' \in [r]}$*.*

$\mathrm{QMIP}^*(l, r, c, s)$ *equals* $\mathrm{QMIP}_{\mathrm{le}}(l, r, c, s)$ *without the limitation on the size of the entangled state shared by the provers.*

$\mathrm{MIP}^*(l, r, c, s)$ *equals* $\mathrm{MIP}(l, r, c, s)$ *with quantum provers sharing unlimited entanglement but returning classical answers.*

We use the short forms QMIP, $\mathrm{QMIP}_{\mathrm{le}}$, QMIP^*, MIP *and* MIP^* *for the union of the respective multiprover classes with a polynomial number of provers and rounds, completeness* $\frac{2}{3}$ *and soundness* $\frac{1}{3}$*.*

Theorem 8.16. $\mathrm{QMIP}_{\mathrm{le}} \subseteq \mathrm{QMIP} = \mathrm{MIP} = \mathrm{NEXP} \subseteq \mathrm{QMIP}^* = \mathrm{MIP}^*$.

Proof. The containments are a result of the work in [72] proving $\mathrm{QMIP}_{\mathrm{le}} \subseteq \mathrm{QMIP} = \mathrm{NEXP}$, [73] proving $\mathrm{MIP} = \mathrm{NEXP}$, [74] proving $\mathrm{NEXP} \subseteq \mathrm{MIP}^*$ and [75] proving $\mathrm{MIP}^* = \mathrm{QMIP}^*$. □

Interestingly, the above containment results show that allowing limited entanglement to the provers weakens the class of quantum multiprover protocols while unlimited entanglement seems to strengthen it. Limited entanglement gives provers an additional possibility to cheat, while unlimited entanglement can be exploited by the verifier for his benefit. It is worth mentioning that no upper bound on the complexity classes $\mathrm{QMIP}^* = \mathrm{MIP}^*$ is known. Up to current knowledge, they might even comprise undecidable problems.

The work in [72, 73] also shows that MIP, QMIP and $\mathrm{QMIP}_{\mathrm{le}}$ protocols can be reduced and amplified to 2 provers, 1 round, perfect completeness and exponentially small soundness (by the cost of increasing the question and answer sizes). For the classes MIP^* and QMIP^* of unlimited entanglement, on the contrary, no amplification results are known yet.

Let us study how to build a multiprover protocol for the k-Local Hamiltonian problem with one round and constantly sized answers. In the classical protocol for 3-SAT presented in the proof of proposition 8.6 one prover is asked for a random clause and the other one for a bit from that clause, followed by a consistency check and the evaluation of the clause. Clauses correspond to interaction terms of the LH instance, but unfortunately, the states for one interaction term and a single qubit will usually be mixed states for which we cannot check equality. The well-known SWAP test [53] for comparing quantum states just works for pure states.

A workaround for this problem could be to replace the quantum answers of constantly many qubits by classical descriptions which are comparable. Unfortunately, a remaining problem is that the provers could send answers of reduced density matrices and single qubits that pass a comparision test perfectly but which do not form a global quantum state. In fact, checking if several reduced states are consistent with a global state is itself a QMA-complete problem under randomized Turing reduction [76]. This is a thoroughly quantum obstacle that does not exist in the classical case of 3-SAT.

We realize that a straightforward quantum adaption of the multiprover protocol for 3-SAT is not possible for k-LH. In 2014 Fitzsimons and Vidick [3] presented a different approach for a multiprover protocol based on the idea that the provers share an encoding of the ground state and are asked probabilistically for their share of a logical qubit or interaction term. Each prover possesses one physical qubit for each logical qubit and since the applied code can correct any single qubit error the verifier is able to detect any single cheating prover. Similarly to the classical 3-SAT protocol, the verifier either evaluates an interaction term, i.e. measures the energy with respect to that interaction term, or carries out a consistency check in terms of a code space check.

Input: A k-local Hamiltonian $H = \sum_{S \in \mathcal{C}} H_S$ with H_S positive semi-definite, $\|H_S\| \leqslant 1$ and $S \subseteq [n]$ indicating the k qubits on that H_S acts non-trivially.

Expect r provers to share the ground state of H encoded by a code with the properties of table 8.2 such that prover t possesses for each logical qubit the physical qubits indicated by the set Λ^t (the prover's *share*).

Apply with probability $\frac{1}{2}$ each:

Test A Choose an interaction term H_S uniformly at random and ask each prover for their share of the qubits in S. Error correct and decode the received state. Measure the resulting state with regard to the obversable H_S like in the QMA membership protocol of proposition 3.33. Accept iff $|1\rangle$ is measured, otherwise reject.

Test B Choose a qubit $i \in [n]$ uniformly at random and afterwards a set $S \in \mathcal{Q}$ containing i uniformly at random with $\mathcal{Q} := \mathcal{C} \cup \{\{i\} \mid i \in [n]\}$. Ask a random prover for his share of the qubits in S and the remaining provers for their share of qubit i. Reject iff a code space check on the logical qubit i fails.

Accept.

Table 8.1: Multiprover protocol for the k-Local Hamiltonian problem based on the protocol by Fitzsimons and Vidick.

The original protocol equals the protocol of table 8.1 with an addtional limitation on the code ($\Lambda^t = \{t\}$, $\rho^t = \frac{\mathbb{I}}{2}$ and detectibility is strengthened to correctibility) and the difference that the question for a single qubit $\{i\}$ is posed with probability 50% in test B, while in test B of table 8.1 the question $\{i\}$ is chosen with the same probability as any interaction term involving i. This difference is irrelevant for the performance of the protocol, but avoids a case differentiation in the soundness proof.

Studying the protocol it becomes clear that the obfuscation for the provers which test is applied is achieved by asking one of the provers for a whole interaction term in test

B, although the verifier only needs the value of a single qubit i (see figure 8.1). Thus, if a prover is asked for an interaction term, he cannot adapt his answer to minimize the energy of that interaction term to achieve a higher acceptance probability in test A, since he has to consider the possibility that test B is running in which the other provers are not aware of the term information and cannot adapt their answers consistently.

Since the smallest code that fulfills the requirements of the original paper [3] is the 5-qubit code, the paper establishes a 5-prover protocol:

Theorem 8.17. *There exists a* $\mathrm{QMIP}_{\mathrm{le}}\left(1, 5, 1 - \frac{a}{2m}, 1 - \frac{Cb}{mn^2(n^k)^c}\right)$ *protocol with logarithmically sized classical questions, constantly sized answers and* C *and* c *constants just depending on* k *that decides if a* k*-local Hamiltonian with* m *interaction terms has ground state energy at most* a *(yes-instance) or at least* b *(no-instance).*

Conjecture 8.18 (Multiprover QPCP conjecture)**.** *Any* QMA *problem can be solved by a* $\mathrm{QMIP}^*(1, r, c, s)$ *protocol with a constant number of provers* r*, a constant probability gap* $c - s > 0$*, logarithmically sized questions and constant sized answers.*

Note that no implication between the above conjecture 8.18 and the original QPCP conjecture 8.8 is known. A $\mathrm{QMIP}^*(1, r, c, s)$ protocol cannot be transformed as easily into a $\mathrm{QPCP}(1, c, s)$ protocol as it works for the classical analogue in proposition 8.6. Regarding the classical multiprover protocol for 3-SAT one can imagine that each prover possesses a string representing all his possible answers of which he only sends certain bits depending on the question they receive. These provers' strings together can be considered as witness for a $\mathrm{PCP}(\log, 1, c, s)$ protocol. For the quantum case this concept does not work since the question one prover is asked might influence the answer of the other provers due to the operation on the shared entangled state.

Proving the opposite implication direction (CSP variant of QPCP conjecture 8.8 $\Longrightarrow$ multiprover QPCP conjecture 8.18) via the multiprover protocol of table 8.1 is prevented by the high soundness term in theorem 8.17. The additional polynomial $n^2(n^k)^c$ causes an inverse polynomial probability gap even for Hamiltonians with a constant relative energy gap (note, moreover, that in the current situation, the low energy value a has to be smaller than $\frac{2Cb}{n^2(n^k)^c}$ for completeness and soundness being separated at all). If the soundness in theorem 8.17 could be upper bounded by $1 - \frac{b}{2m}$, then indeed, we knew that conjecture 8.18 was weaker than the original QPCP conjecture 8.8.

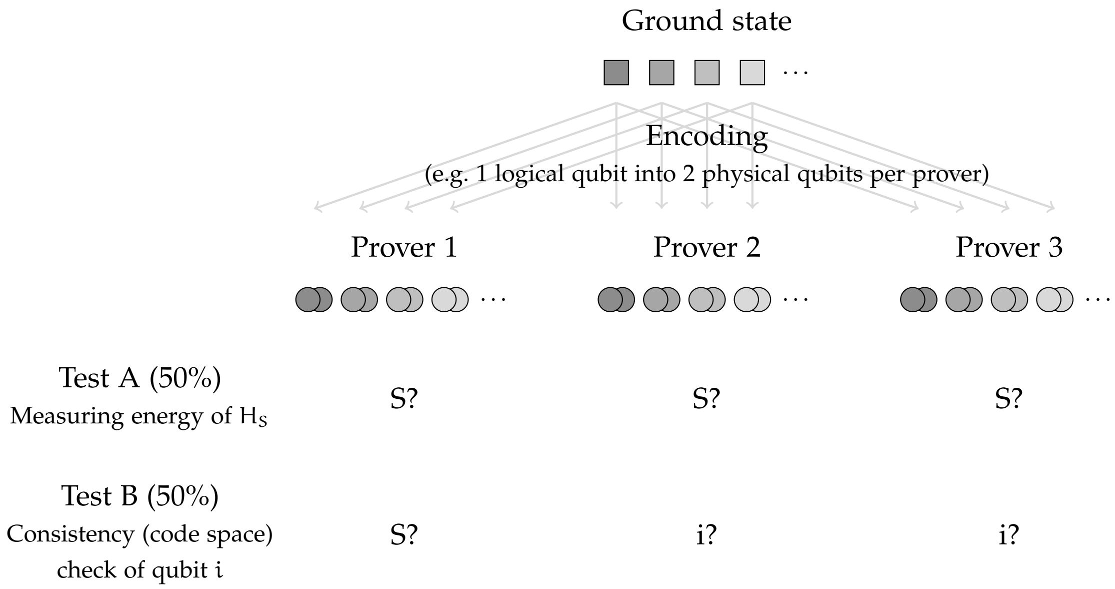


Figure 8.1: Questions posed to the (randomly labelled) provers by protocol 8.1.

8.6 Improving the multiprover protocol

The only improvement of the multiprover protocol by [3] was accomplished in 2016 by [77]. The author succeeds in restricting the protocol to four entangled provers with classical, constantly sized answers by keeping the protocol of 1 round and logarithmically sized questions. The answers can be made classical by playing a so-called stabilizer game with the provers which originates from the method of delegated quantum computation [75].

That the protocol of [3] can be restriced to classical answers in such a way does not follow directly from $\text{QMIP}^* = \text{MIP}^*$ by theorem 8.16, because the proof of this equality [75] introduces an inverse polynomial error into the acceptance probability that would close the gap between completeness and soundness. This could be avoided by initial amplification of the completeness and soundness parameter – which is possible since the protocol is actually of QMIP_{le} type – but this would blow up the question and answer sizes to a polynomial.

1. The logical qubits of the code are partitioned into r blocks Λ^t: $[N] = \bigcup_{t=1}^{r} \Lambda^t$.
2. Every error restricted to one block is detectable.
3. The partial trace over all but one block is independent of the code state, i.e. $\forall t \in [r]\ \exists \rho^t \in \mathcal{D}(\mathbb{C}^{2^{|\Lambda^t|}})\ \forall |\psi\rangle \in V_C : \mathrm{tr}_{[N]\setminus\Lambda^t}(|\psi\rangle\langle\psi|) = \rho^t$.

Table 8.2: Conditions on the $(N, 2)$ code required by the multiprover protocol.

The reduction from 5 to 4 provers by [77] is accomplished by simply substituting the 5-qubit code by the $[4, 1, 2]$ code of [78], since the protocol by [3] just makes use of the error detection property of the code and does not actually need the tool of error correction. Studying the protocol in detail, we can relax the code properties even further: Firstly, a prover can possess more than one qubit as long as every error on his share is detectable. Secondly, the reduced density matrix onto a prover's share does not necessarily have to equal the completely mixed state, it only has to be independent of the code state. In summary, we require the code properties listed in table 8.2.

Can we reduce the number of provers needed for the protocol even further with these relaxations of the code properties? The next two lemmata give an answer to this.

Lemma 8.19. *A code with properties 1 and 2 of table 8.2 is partitioned into at least* 3 *blocks.*

Proof. Assume a code with properties 1 and 2 of table 8.2 and a partition into m blocks such that block i comprises s_i qubits. Clearly, the blocks can be padded by additional qubits in the state $|0\rangle$ until they contain the same number of qubits $s := \max_{i\in[m]} s_i$ while keeping the ability to detect any block error (for detection first carry out a projection of the additional qubits onto the all zero state and then the original error detection procedure).

Hence we can restrict our proof to the case that each of the r blocks comprises exactly s qubits, which we consider as a qudit of dimension $d = 2^s$. The detectibility of every block error is equivalent to the detectibility of every single qudit error. Hence in the qudit formulation, the code is of distance $\delta \geqslant 2$ encoding $K = 2$ orthonormal states on r qudits. The quantum Singleton bound from the end of section 4.4 tells us

$$r \geqslant 2(\delta - 1) + \log_d K > 2. \qquad \square$$

Lemma 8.20. *The code with code states*

$$|\bar{0}\rangle := \frac{1}{2}(|00\,00\,00\rangle + |01\,01\,01\rangle + |10\,10\,10\rangle + |11\,11\,11\rangle)$$
$$|\bar{1}\rangle := \frac{1}{2}(|00\,01\,10\rangle + |01\,10\,11\rangle + |10\,11\,00\rangle + |11\,00\,01\rangle)$$

fulfills the conditions of table 8.2 with the partition $\Lambda^1 = \{1, 2\}$, $\Lambda^2 = \{3, 4\}$ and $\Lambda^3 = \{5, 6\}$.

Proof. This code can detect any error restricted to a single block Λ^t of qubits, since $\langle c| E_i |c'\rangle = 0$ for any any $c, c' \in \{\bar{0}, \bar{1}\}$ and any non-identity Pauli operator E_i that acts non-trivially only on the qubits of a set Λ^t.

The third property is fulfilled, since for every prover $t \in \{1, 2, 3\}$

$$\mathrm{tr}_{[6]\setminus\Lambda^t}(|\bar{0}\rangle\langle\bar{0}|) = \mathrm{tr}_{[6]\setminus\Lambda^t}(|\bar{1}\rangle\langle\bar{1}|) = \frac{\mathbb{I}\,\mathbb{I}}{4}. \qquad \square$$

The above code allows us to reduce the multiprover protocol to 3 provers which is the minimum possible by code adaption (the adapted proof is contained in the next section). But it is unclear if these provers can also be restricted to classical answers, since the code is not a stabilizer code which is necessary for the techniques of [77]. But note that the slightly changed code

$$|\bar{0}\rangle := \frac{1}{\sqrt{3}}(|00\,00\,00\rangle + |01\,01\,01\rangle + |10\,10\,10\rangle)$$
$$|\bar{1}\rangle := \frac{1}{\sqrt{3}}(|00\,01\,10\rangle + |01\,10\,00\rangle + |10\,00\,01\rangle)$$

fulfills the condition of table 8.2 as well and equals a qubit implemention of the qutrit stabilizer code of [79, A2] ($|00\rangle \to |0\rangle$, $|01\rangle \to |1\rangle$ and $|10\rangle \to |2\rangle$). Hence, it is worth checking if the techniques of [77] nevertheless work for this code and even prove the existence of a 3-prover protocol with classical answers.

Besides a reduction of the provers we would like to improve the soundness of the protocol. As we saw at the end of the last section, the additional polynomial $n^2(n^k)^c$ in the soundness term of theorem 8.17 prevents to prove that the standard QPCP conjecture 8.8 implies the multiprover QPCP conjecture 8.18 (note that of course the constant C also plays a role depending on the values of a and b). A weaker conjecture could be a good starting point to make progress on the standard QPCP conjecture.

For a general k-local Hamiltonian we find one improvable point in the original proof, which reduces the factor n^2 in the polynomial to n. We wrote the polynomial in the denominator of the soundness term on purpose in the form $n^2(n^k)^c$, since the term n^k has a different origin and is related to the number of possible questions and hence to the number of interaction terms. It is therefore not surprising that for Hamiltonians of degree d this term can be reduced from n^k to nd (the constant c in theorems 8.17 and 8.21 is indeed the same). This is interesting since theorem 8.10 states that the QPCP conjecture can be restricted to 2-LH with constant degree.

Unfortunately, the remaining polynomial $n(nd)^c$ still prevents that the multiprover QPCP conjecture 8.18 is implied by the standard QPCP conjecture 8.8. The first factor n is accumulated in proposition 8.33 by changing the order of extracted qubits that allow the construction of a low energy state in case of high acceptance and seems inevitable. The term nd originates from the argument in lemma 8.28 that a failure with probability ϵ of test B with its $\mathcal{O}(nd)$ different probabilistic branches covers the case that a single probabilistic branch can fail with probability up to $\mathcal{O}(nd\epsilon)$ while most other branches rarely fail. If we could ensure instead that each branch fails with a similar probability, i.e. of order $\mathcal{O}(\epsilon)$, the soundness of the protocol got rid of this problematic polynomial. This seems like a difficult if not impossible task; most conceivable is a proof in the style of [71] that argues that the ground state energy problem for Hamiltonian families without this property is simple (e.g. contained in NP).

Note that classically a single consistency check either fails with probability 1 or 0. Hence, that each consistency check fails with a probability close to the average can only occur in the quantum protocol where a single consistency check is itself probabilistic.

The reason why the provers should apply a strategy that makes a consistency check fail can only be an energetic advantage in test A. The current proof of the relevant lemma 8.28 works without any energy argument; a proof for a more limited failure probability of each B branch clearly needs to connect the performances of test A and B. A rather even distribution of the failure probability over the B branches requires from the Hamiltonian family that small changes across many interaction terms decrease the ground state energy more than large changes of few interaction terms. A difficult point is that quantifying the amount of the "change" involves the combination of Hamiltonian and chosen code, since it refers to the extent to that an encoded low energy state is pushed outside the code space by a partial change of the physical qubits.

We can summarize the currently possible improvements of theorem 8.17 to:

Theorem 8.21. *There exists a* $\mathrm{QMIP}_{\mathrm{le}}\left(1, 3, 1-\frac{a}{2m}, 1-\frac{Cb}{mnp^c}\right)$ *protocol with logarithmically sized classical questions, constant sized answers,* $p = nd$ *according to lemma 8.28 and* C *and* c *constants just depending on* k *that decides if a* k*-local Hamiltonian of degree* d *either has ground state energy at most* a *(yes-instance) or at least* b *(no-instance).*

The improved theorem still bases on the protocol of table 8.1 but with a different code, some different bounding techniques and by considering the degree d of the Hamiltonian. The completeness of the protocol of table 8.1 is clearly $1 - \frac{a}{2m}$, since for a yes-instance and provers answering honestly according to the ground state, the protocol accepts with probability $1 - \frac{a}{m}$ for test A and with certainty for test B. In the next section we will discuss the adaptions of the original soundness proof by [3] that are needed due to the new code and our improved bounds.

Note that we keep stating the protocol performance for general k and d. This has two reasons. First, as long as the QPCP conjecture is not proven, one might want to consider the protocol for k-local Hamiltonians whose relative energy gap scales inverse polynomially. Although these Hamiltonians can be transformed into 2-local Hamiltonians of constant degree, a formulation of the protocol for the original Hamiltonians is useful, since the transformation [80] changes the scaling of the energy gap. Secondly, even if the QPCP was proven and although we know that it can be restricted to 2-local Hamiltonians of constant degree, we do not know if it can be restricted to energy values at most a or at least b obeying $a < \frac{2Cb}{n(nd)^c}$. But this is necessary to separate completeness and soundness of the multiprover protocol as long as the additional polynomial $n(nd)^c$ and the small constant C remain in the soundness term.

8.7 Adaption of the soundness proof

8.7.1 Sketch of the extraction protocol for a low energy state σ

In this section we adapt the soundness proof of [3] (note that the full proof is only contained in the long arXiv version of the paper and not in the journal version) to prove theorem 8.21. We assume $H = \sum_{S \in \mathcal{C}} H_S$ to be a k-local Hamiltonian of degree d with S indicating the qubits on that the interaction term H_S acts non-trivially. Moreover, we assume C to be a code with the properties and definitions of table 8.2.

Definition 8.22. *The superoperator* DEC : $\mathcal{L}(\mathbb{C}^{2^n}) \to \mathcal{L}(\mathbb{C}^2)$,

$$\mathrm{DEC}(\rho) := \mathcal{D}(P_C \rho P_C) + \mathrm{tr}\left((\mathbb{I} - P_C)\rho\right) |0\rangle \langle 0| ,$$

maps code states via the decoding map $\mathcal{D}$ to their logical qubit and states orthogonal to the code space to the default state $|0\rangle$.

The soundness proof for theorem 8.21 works via contraposition, i.e. one proves the existence of a low energy state σ for the case that the protocol of table 8.1 accepts with high probability.

Since the questions are classical we can avoid considering a question register and denote the strategy of the provers by $\left(U_S^t, |\psi\rangle \langle\psi|\right)$ with $t \in [r]$ indicating the prover and $S \in \mathcal{Q} := \mathcal{C} \cup \{\{i\} \mid i \in [n]\}$ indicating the question which equals a set of 1 or k logical qubits. We can assume w.l.o.g. that the register P^t on that a prover $t \in [r]$ acts is composed of a private register S^t and the answer registers Q_i^t, each comprising $|\Lambda^t|$ qubits for the prover's share of the logical qubit $i \in [n]$. Note that throughout this section an upper index t always indicates the prover and a lower index i the logical qubit.

The idea for the construction algorithm of the low energy state σ is to apply an operation $U_{\{i\}}^t$ to the strategy state $|\psi\rangle$ for all t, swap the answers into an extraction register by substituting ρ^t and undoing the operation $U_{\{i\}}^t$. Since a single prover cannot differentiate his share of a logical qubit from ρ^t, this action rarely influences further answers of the provers and we can distill the encoding of the low energy state σ into the extraction register by repeating the procedure for all $i \in [n]$.

Picture 8.2 depicts the registers that the extraction algorithm needs for each prover. The collection of registers S^t and Q_i^t, $t \in [r]$, $i \in [n]$, is initialized in $|\psi\rangle$ (a prover's operation U_S^t is later always applied to these registers without explicitely marking this by an index). Each extraction register R_i^t has the same size as Q_i^t and the size of $\bar{R}_i^t$ is chosen such that $R_i^t \cup \bar{R}_i^t$ can be initialized in a pure state $|\eta^t\rangle$ with $\mathrm{tr}_{\bar{R}_i^t}(|\eta^t\rangle) = \rho^t$. The purification registers $\bar{R}_i^t$ are irrelevant for the protocol, but they will save some writing by avoiding mixed states.

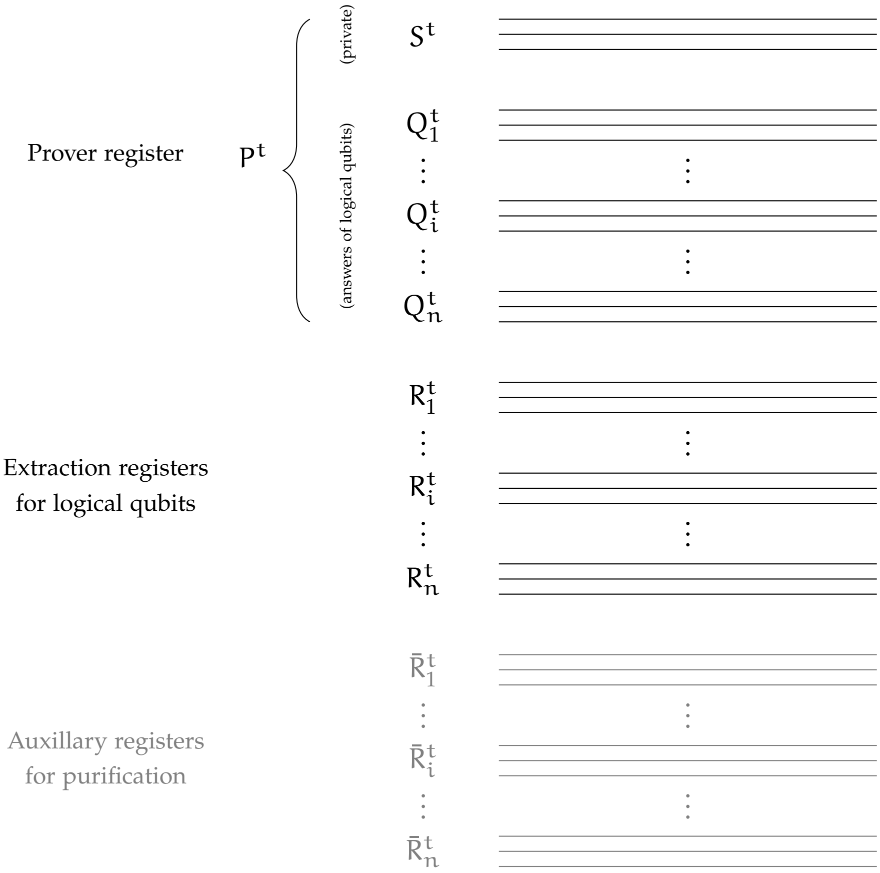


Figure 8.2: Registers of the extraction algorithm for each prover $t \in [r]$.

Definition 8.23. *If we write a (super)operator in brackets with a lower index, the lower index denotes the register the (super)operator acts on within a larger system.*

Replacing an index of a register by a set of indices denotes the union over this set of indices. Neglecting the index denotes the union over the whole range of this index.

According to the previous definition we write for example $Q^t_U := \bigcup_{i \in U} Q^t_i$, $P := \bigcup_{t \in [r]} P^t$ and $Q := \bigcup_{t \in [r]} \bigcup_{i \in [n]} Q^t_i$.

Definition 8.24. *Given a strategy* $(U^t_S, |\psi\rangle\langle\psi|)$ *we define for all* $i \in S \in \mathcal{Q}$

$$\begin{aligned} D^t_{i,S} &:= (U^t_S)^\dagger \, \mathrm{SWAP}_{Q^t_i, R^t_i} \, U^t_S \\ \mathcal{D}^t_{i,S}(\rho) &:= D^t_{i,S} \rho (D^t_{i,S})^\dagger \end{aligned}$$

with $\mathrm{SWAP}_{Q^t_i, R^t_i}$ *the operator swapping the qubits of the register* Q^t_i *with those in* R^t_i.

Moreover, for every set $I \subseteq [n]$ *of* l *elements* $i_1 < i_2 < \cdots < i_l$ *and any choice of* $S_i \in \mathcal{Q}$ *containing* $i \in I$ *we use the convention*

$$\bigcirc_{i \in I} \mathcal{D}^t_{i,S_i} := \mathcal{D}^t_{i_l,S_l} \circ \cdots \circ \mathcal{D}^t_{i_2,S_2} \circ \mathcal{D}^t_{i_1,S_1}$$

Definition 8.25. *For the strategy* $(U^t_S, |\psi\rangle\langle\psi|)$ *we define the states*

$$\begin{aligned} |\tilde{\psi}\rangle &:= (|\psi\rangle)_P \otimes \bigotimes_{t \in [r]} \bigotimes_{i \in [n]} (|\eta^t\rangle)_{R^t_i \cup \bar{R}^t_i} \\ \tilde{\rho} &:= |\tilde{\psi}\rangle\langle\tilde{\psi}| \\ \tau &:= \left(\bigotimes_{t \in [r]} \bigcirc_{i \in [n]} \mathcal{D}^t_{i,\{i\}} \right) (\tilde{\rho}) \\ \sigma &:= \mathrm{DEC}_R \left(\mathrm{tr}_{\bar{R} \cup P} (\tau) \right). \end{aligned}$$

If the protocol of table 8.1 accepts with probability $1 - \epsilon$, it will turn out by theorem 8.34 that there exists a constant C such that σ has energy at most $\frac{mnp^c\epsilon}{C}$ with $p = nd$. Equivalently stated, there exists a constant C such that, if the protocol accepts with

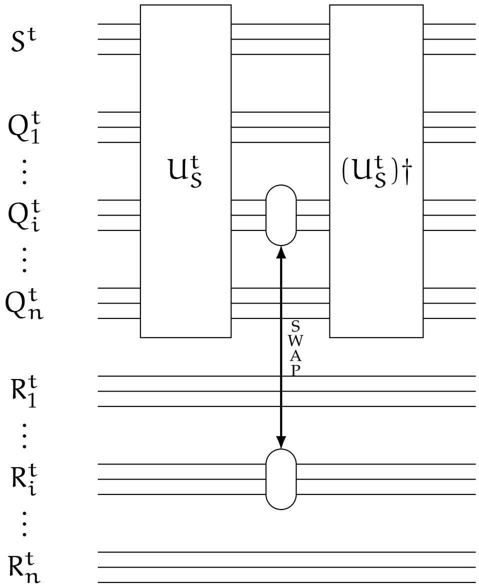


Figure 8.3: The operation $D_{i,S}^t$.

probability greater than $1 - \frac{Cb}{mnp^c}$, σ has energy less than b and hence at most a due to the promise. This establishes the soundness proof.

8.7.2 Closeness of the extraction operators $D_{i,S}^t$ and $D_{i,\{i\}}^t$

The first step in the soundness proof is to show that the extraction operators $D_{i,S}^t$ and $D_{i,\{i\}}^t$ for $i \in S \in \Omega$ almost act the same on any state $|\phi\rangle$ if $\rho = \mathrm{tr}_{R \cup \bar{R}}(|\phi\rangle\langle\phi|)$ is a strategy state that leads to high acceptance. In this subsection we revisit this first step (claim 8 in the original paper [3]) in detail for several reasons:

- lemma 8.27 makes use of the error detection abilities of our adapted code,
- the necessity for the additional lemma 8.31 occured due to a small bug in the original work,
- we study the origin of the additional polynomial in the soundness term and express it in terms of the Hamiltonian degree.

Definition 8.26. *Let $W^{t,\sigma}_{i,S,S'}$, $\sigma \in \mathcal{P}^{\otimes|\Lambda^t|}$, $i \in S \cap S' \in \mathcal{Q}$, be the operators on register P^t such that*

$$D^t_{i,S}\left(D^t_{i,S'}\right)^\dagger = \sum_{\sigma \in \mathcal{P}^{\otimes|\Lambda^t|}} (\sigma)_{R^t_i} \otimes W^{t,\sigma}_{i,S,S'}.$$

Moreover, we define for $i \in S \in \mathcal{Q}$ and a state $|\phi\rangle$ on all registers of the extraction protocol:

$$\begin{aligned} |\phi^t_{i,S}\rangle &:= D^t_{i,S} \bigotimes_{t' \in [r] \setminus \{t\}} D^{t'}_{i,\{i\}} |\phi\rangle \\ |\phi^{t,\text{suc}}_{i,S}\rangle &:= (P_C)_{R_i} |\phi^t_{i,S}\rangle \\ |\phi^{t,\text{fail}}_{i,S}\rangle &:= \left(\mathbb{I} - (P_C)_{R_i}\right) |\phi^t_{i,S}\rangle . \end{aligned}$$

The indices of $|\phi^t_{i,S}\rangle$ indicate that prover t is asked question S while all other provers are only asked for qubit i and the actions are reversed after the logical qubit i is extracted. We want to show closeness of the states $|\phi^t_{i,S}\rangle$ and $|\phi^t_{i,S'}\rangle$. The next lemma provides a first step to express $|\phi^t_{i,S}\rangle$ as $|\phi^t_{i,S'}\rangle$ plus small failure terms:

Lemma 8.27. *For any $i \in S \cap S'$ with $S, S' \in \mathcal{Q}$ we have that*

$$|\phi^t_{i,S}\rangle = |\phi^{t,\text{fail}}_{i,S}\rangle + (P_C)_{R_i}\left(D^t_{i,S}\left(D^t_{i,S'}\right)^\dagger\right)|\phi^{t,\text{fail}}_{i,S'}\rangle + (P_C)_{R_i} W^{t,\mathbb{I}}_{i,S,S'} |\phi^t_{i,S'}\rangle .$$

Proof.

$$\begin{aligned} |\phi^t_{i,S}\rangle &= |\phi^{t,\text{fail}}_{i,S}\rangle + (P_C)_{R_i}\left(D^t_{i,S}\left(D^t_{i,S'}\right)^\dagger\right)|\phi^t_{i,S'}\rangle \\ &= |\phi^{t,\text{fail}}_{i,S}\rangle + (P_C)_{R_i}\left(D^t_{i,S}\left(D^t_{i,S'}\right)^\dagger\right)|\phi^{t,\text{fail}}_{i,S'}\rangle \\ &\quad + (P_C)_{R_i} W^{t,\mathbb{I}}_{i,S,S'} |\phi^t_{i,S'}\rangle + (P_C)_{R_i} \sum_{\sigma \in \mathcal{P}^{\otimes|\Lambda^t|} \setminus \{\mathbb{I}\}} \left((\sigma)_{R^t_i} \otimes W^{t,\sigma}_{i,S,S'}\right) |\phi^{t,\text{suc}}_{i,S'}\rangle \end{aligned}$$

with the last summand vanishing since the non-identy Pauli operator σ on register R^t_i turns the code state $|\phi^{t,\text{suc}}_{i,S'}\rangle$ orthogonal to the code space according to the error detection abilities of the code. □

For the closeness of $|\phi^t_{i,S}\rangle$ and $|\phi^t_{i,S'}\rangle$ we have to show that the failure states have a small norm and that $W^{t,\mathrm{Id}}_{i,S,S'}$ almost acts like the identity on $|\phi^t_{i,S'}\rangle$. The next lemma proves the first fact:

Lemma 8.28. *If a strategy* (U^t_S, ρ) *with* $\rho = \mathrm{tr}_{R\cup\bar{R}}(|\phi\rangle\langle\phi|)$ *succeeds in test B with probability at least* $1-\epsilon$, *then*

$$\| \, |\phi^{t,\mathrm{fail}}_{i,S}\rangle \, \|^2 = \mathcal{O}(p\epsilon)$$

for all $i \in S$ *with* $S \in \mathcal{Q}$ *and* $p = nd$.

Proof. Considering $(P_C)_{Q_i} = \mathrm{SWAP}_{Q_i,R_i}(P_C)_{R_i}\,\mathrm{SWAP}_{Q_i,R_i}$ and that any prover operation U^t_S commutes with the projection $(P_C)_{R_i}$ of the register R_i, we can express the probability that test B is passed as

$$\mathbb{P}[\text{test B passed}] = \frac{1}{r}\sum_{t\in[r]}\frac{1}{n}\sum_{i\in[n]}\frac{1}{|\{S \mid i\in S\in\mathcal{Q}\}|}\sum_{\substack{S\in\mathcal{Q}\\ i\in S}}\|(P_C)_{R_i}|\phi^t_{i,S}\rangle\,\|^2 \geqslant 1-\epsilon.$$

Hence, the average failure state can be bounded by

$$\frac{1}{r}\sum_{t\in[r]}\frac{1}{n}\sum_{i\in[n]}\frac{1}{|\{S \mid i\in S\in\mathcal{Q}\}|}\sum_{\substack{S\in\mathcal{Q}\\ i\in S}}\| \, |\phi^{t,\mathrm{fail}}_{i,S}\rangle \, \|^2 \leqslant \epsilon.$$

Since the probability that a specific tuple (i,S) is picked by the protocol is lower bounded by $\frac{1}{rn(d+1)}$, it holds for all $i\in S\in\mathcal{Q}$ that

$$\| \, |\phi^{t,\mathrm{fail}}_{i,S}\rangle \, \|^2 = \mathcal{O}\left(nd\epsilon\right). \qquad \square$$

Note that at this point we acquired an additional n dependence of the soundness value that prevents the probability gap of the multiprover protocol from scaling like the relative energy gap of the Hamiltonian.

Corollary 8.29. *If a strategy* (U^t_S, ρ) *with* $\rho = \mathrm{tr}_{R\cup\bar{R}}(|\phi\rangle\langle\phi|)$ *succeeds in test B with proba-*

bility at least $1 - \epsilon$, *then*

$$\left\| W^{t,\mathbb{I}}_{i,S,S'} |\phi^{t}_{i,S'}\rangle \right\|^2 = 1 - \mathcal{O}(p\epsilon)$$

for all $i \in S \cap S'$ *with* $S, S' \in \mathcal{Q}$ *and* $p = nd$ *according to lemma 8.28.*

Proof. Rearranging the equation of lemma 8.27 allows to bound $(P_C)_{R_T} W^{t,\mathbb{I}}_{i,S,S'} |\phi^{t,\text{suc}}_{i,S'}\rangle$ by triangle equality. Inserting the bound of the failure states by lemma 8.28 and noting that the projection $(P_C)_{R_i}$ does not increase the norm leads to the desired statement. □

The above corollary already states that $W^{t,\mathbb{I}}_{i,S,S'}$ rarely changes the norm of $|\phi^{t}_{i,S'}\rangle$. We want to get one step further and show that it even acts almost as the idenity on this state. The authors of [3] claim that $W^{t,\mathbb{I}}_{T,S,S'} = \mathbb{I}$ with the reason that $\text{tr}_{R^t_i}(U^t_S) = \text{tr}_{R^t_i}(U^t_{S'}) = \text{Id}$. The counter example $W^{t,\mathbb{I}}_{\{i\},S,S'} = 0$ for $U^t_S = (X)_{Q^t_i}$ and $U^t_{S'} = (Z)_{Q^t_i}$ for any $i \in S \cap S'$ reveals this as a misconception. In email discussions with the authors the workaround of lemma 8.31 was found, proving at least that $W^{t,\mathbb{I}}_{T,S,S'}$ acts almost as the identity on $|\phi^{t}_{i,S'}\rangle$.

Lemma 8.30. *The swap operator can be written in the form*

$$\text{SWAP} = \frac{1}{2}(\mathbb{I}\,\mathbb{I} + XX + YY + ZZ).$$

Proof. Writing SWAP as a linear combination of Pauli operators $\mu \otimes \nu \in \{\text{Id}, X, Y, Z\}^{\otimes 2}$ and computing their prefactors via $\frac{1}{4}\,\text{tr}\left(\text{SWAP}\,\mu^{\dagger} \otimes \nu^{\dagger}\right)$ shows that the only non-zero prefactors equal $\frac{1}{2}$ for $\mu = \nu$. □

Lemma 8.31. *If a strategy* (U^t_S, ρ) *with* $\rho = \text{tr}_{R \cup \bar{R}}(|\phi\rangle\langle\phi|)$ *succeeds in test B with probability at least* $1 - \epsilon$, *then*

$$W^{t,\mathbb{I}}_{i,S,S'} |\phi_{i,S'}\rangle = |\phi^{t}_{i,S'}\rangle + |\phi^{t,\text{err}}_{i,S,S'}\rangle,$$

$$\left\| |\phi^{t,\text{err}}_{T,S,S'}\rangle \right\|^2 = \mathcal{O}(p\epsilon)$$

for all $i \subseteq S \cap S'$ *with* $S, S' \in \mathcal{Q}$ *and* $p = nd$ *according to lemma 8.28.*

Proof. With the SWAP expression from lemma 8.30 and the definition

$$D^t_{i,S,\sigma} \coloneqq (U^t_S)^\dagger (\sigma)_{Q^t_i} U^t_S$$

for all $i \in S \in \mathcal{Q}$ and all $\sigma \in \mathcal{P}^{\otimes|\Lambda^t|}$ we can rewrite

$$\begin{aligned} D^t_{i,S} &= (U^t_S)^\dagger \,\mathrm{SWAP}_{Q^t_i,R^t_i}\, U^t_S \\ &= \frac{1}{2^{|\Lambda^t|}} \sum_{\sigma\in\mathcal{P}^{\otimes|\Lambda^t|}} D^t_{i,S,\sigma} \otimes (\sigma)_{R^t_i} \\ W^{t,\mathbb{I}}_{i,S,S'} &= \frac{1}{2^{|\Lambda^t|}} \operatorname{tr}_{R^t_i}\left(D^t_{i,S}(D^t_{i,S'})^\dagger\right) \\ &= \frac{1}{2^{2|\Lambda^t|}} \sum_{\sigma\in\mathcal{P}^{\otimes|\Lambda^t|}} D^t_{i,S,\sigma}(D^t_{i,S',\sigma})^\dagger. \end{aligned}$$

This allows us to interpret $W^{t,\mathbb{I}}_{i,S,S'} |\phi^t_{i,S'}\rangle$ with $i \in S \cap S'$ and $S, S' \in \mathcal{Q}$ as the average of the $c \coloneqq 2^{2|\Lambda^t|}$ many vectors

$$|\phi^t_{i,S,S',\sigma}\rangle \coloneqq D^t_{i,S,\sigma}(D^t_{i,S',\sigma})^\dagger |\phi^t_{i,S'}\rangle .$$

Since the operators $D^t_{i,S,\sigma}$ are unitary, the states $|\phi^t_{i,S,S',\sigma}\rangle$ are normalized and have to be almost the same in order to average to a state of norm almost one (recall corollary 8.29). Since $|\phi^t_{i,S,S',\mathrm{Id}}\rangle = |\phi^t_{i,S'}\rangle$, we prove that each $|\phi^t_{i,S,S',\sigma}\rangle \approx |\phi^t_{i,S'}\rangle$ by splitting

$$|\phi^t_{i,S,S',\sigma}\rangle = \sqrt{1-\delta_\sigma}\,|\phi^t_{i,S'}\rangle + \sqrt{\delta_\sigma}\,|\phi^{t,\perp}_{i,S,S',\sigma}\rangle$$

with the normalized state $|\phi^{t,\perp}_{i,S,S',\sigma}\rangle$ orthogonal to $|\phi^t_{i,S'}\rangle$. We can derive $\delta_\sigma \in \mathcal{O}(nd\epsilon)$ via the norm result of corollary 8.29 and the inequality $\sqrt{1-\delta_\sigma} \leqslant 1 - \frac{\delta_\sigma}{2}$ by assuming the worst case that all but one $|\phi^t_{i,S,S',\sigma}\rangle$ are equal to $|\phi^t_{i,S'}\rangle$:

$$\begin{aligned} \mathcal{O}(p\epsilon) &= 1 - \left\| W^{t,\mathbb{I}}_{i,S,S'} |\phi_{i,S'}\rangle \right\|^2 \\ &\geqslant 1 - \left(\left(1 - \frac{1}{c} + \frac{\sqrt{1-\delta_\sigma}}{c}\right) + \left(\frac{\sqrt{\delta_\sigma}}{c}\right)^2 \right) \\ &\geqslant \left(\frac{1}{2c} - \frac{1}{c^2}\right)\delta_\sigma. \end{aligned}$$

Consequently,

$$W^{t,\mathbb{I}}_{i,S,S'} \left|\phi^t_{i,S'}\right\rangle = \left|\phi^t_{i,S'}\right\rangle + \left|\phi^{t,\mathrm{err}}_{i,S,S'}\right\rangle$$

with

$$\begin{aligned}\left|\phi^{t,\mathrm{err}}_{i,S,S'}\right\rangle &= \frac{1}{c} \sum_{\sigma \in \mathcal{P}^{\otimes|\Lambda^t|}} \left(\sqrt{1-\delta_\sigma} - 1\right) \left|\phi^t_{i,S'}\right\rangle + \sqrt{\delta_\sigma} \left|\phi^{t,\perp}_{i,S,S',\sigma}\right\rangle \\ \left\|\left|\phi^{t,\mathrm{err}}_{T,S,S',R}\right\rangle\right\|^2 &= \mathcal{O}(p\epsilon).\end{aligned}$$ □

We can finally prove the desired result of this subsection:

Proposition 8.32. *If a strategy* (U^t_S, ρ) *with* $\rho = \mathrm{tr}_{R \cup \bar{R}}(|\phi\rangle\langle\phi|)$ *succeeds in test B with probability at least* $1-\epsilon$*, then*

$$\|(D^t_{i,S} - D^t_{i,\{i\}})\,|\phi\rangle\,\|^2 = \mathcal{O}(p\epsilon)$$

for all $i \subseteq S$ *with* $S \in \mathcal{Q}$ *and* $p = nd$ *according to lemma 8.28.*

Proof. The statement is proven if we can show that $\|\,|\phi^t_{i,S}\rangle - |\phi^t_{T,S'}\rangle\,\|^2 = \mathcal{O}(p\epsilon)$, since the unitary $\bigotimes_{t' \in [r]\setminus\{t\}} D^{t'}_{i,\{i\}}$ leaves the norm invariant. One can see that this is indeed the case by taking the norm of the expression in lemma 8.27 and using the inequality $(a+b+c)^2 \leqslant 3(a^2+b^2+c^2)$:

$$\begin{aligned}&\|\,|\phi^t_{i,S}\rangle - |\phi^t_{i,S'}\rangle\,\|^2 \\ &\qquad \leqslant \left(\|\,|\phi^{t,\mathrm{fail}}_{i,S}\rangle\,\| + \|\,|\phi^{t,\mathrm{fail}}_{i,S'}\rangle\,\| + \|(P_C)_{R_T} W^{t,\mathbb{I}}_{i,S,S'}\,|\phi^t_{i,S'}\rangle - |\phi^t_{i,S'}\rangle\,\|\right)^2 \\ &\qquad \leqslant \left(\|\,|\phi^{t,\mathrm{fail}}_{i,S}\rangle\,\| + 2\|\,|\phi^{t,\mathrm{fail}}_{i,S'}\rangle\,\| + \|\,|\phi^{t,\mathrm{err}}_{i,S,S'}\rangle\,\|\right)^2 \\ &\qquad = \mathcal{O}(p\epsilon).\end{aligned}$$ □

8.7.3 Energy of the extraction state σ

In the first step of the proof it was shown that the extraction operators $D^t_{i,\{i\}}$ and $D^t_{i,S}$ almost act the same on a high acceptance state. By several induction arguments it is possible to extent this statement to a series of extraction operators with one series asking for each single qubit successively leading to $\tau = \left(\bigotimes_{t\in[r]} \bigcirc_{i\in[n]} \mathcal{D}^t_{i,\{i\}}\right)(\tilde{\rho})$ and the other series asking for an interaction term S and then for all other single qubits successively:

Proposition 8.33. *For a provers' strategy* $(U^t_S, |\phi\rangle\langle\phi|)$ *that passes test B with probability* $1-\epsilon$ *there exists a constant* c *depending on* k *only such that*

$$\left\| \mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}}[\tau] - \mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}} \left[\left(\bigotimes_{t\in[r]} \bigcirc_{i\in\bar{S}} \mathcal{D}^t_{i,\{i\}} \right) \circ \left(\bigotimes_{t\in[r]} \mathcal{D}^t_{S,S} \right) (\tilde{\rho}) \right] \right\|_1 = \mathcal{O}(np^c\epsilon)$$

for all $S \in \mathcal{C}$ *with* $\bar{S} := [n] \backslash S$ *and* $p = nd$ *according to lemma 8.28.*

Proof. This statement equals equation (25) of the original paper [3].

The step from equation (21) to equation (22) in the original proof [3] is the only one that makes use of the fact that each code state reduced to a single prover register t is equal to the state ρ^t from the extraction register R^t_i. Since the actual value of the state is irrelevant for the argument, this step and the remaining proof remain valid. □

The additional factor n in the above proposition accumulates by pulling the operators $\mathcal{D}^t_{i,\{i\}}$ for $i \in S$ past all operators with lower index i to express τ.

In the next theorem we can finally formulate the main statement about the low energy of σ. This statement ressembles claim (12) in [3]. By slightly different bounding tools we can avoid an additional factor of n in comparision to the original paper [3].

Theorem 8.34. *Let* $(U^t_S, |\psi\rangle\langle\psi|)$ *be a provers' strategy that passes test A with probability* $1-\delta$ *and test B with probability* $1-\epsilon$*. Then there exists a constant* c *depending on* k *only such that*

$$\frac{1}{m}\,\mathrm{tr}\left(H\sigma\right) = \mathcal{O}(\delta + np^c\epsilon)$$

with $p = nd$ *according to lemma 8.28.*

Proof. For any $S \in \mathcal{C}$ and for $(\mathrm{DEC})_{R_S} := \bigotimes_{i \in S} (\mathrm{DEC})_{R_i}$ it holds that

$$\begin{aligned}
\mathcal{O}(np^c\epsilon) \geqslant{} & \|H_S\| \left\| (\mathrm{DEC})_{R_S} \left(\mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}} [\tau] \right) \right. \\
& \left. - (\mathrm{DEC})_{R_S} \left(\mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}} \left[\left(\bigotimes_{t \in [r]} \bigcirc_{i \in \bar{S}} \mathcal{D}^t_{i,\{i\}} \right) \circ \left(\bigotimes_{t \in [r]} \mathcal{D}^t_{S,S} \right) (\tilde{\rho}) \right] \right) \right\|_1 \\
\geqslant{} & \left\| \mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}} [(H_S)_R (\mathrm{DEC})_{R_S} (\tau)] \right. \\
& \left. - \mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}} \left[(H_S)_R (\mathrm{DEC})_{R_S} \circ \left(\bigotimes_{t \in [r]} \bigcirc_{i \in \bar{S}} \mathcal{D}^t_{i,\{i\}} \right) \circ \left(\bigotimes_{t \in [r]} \mathcal{D}^t_{S,S} \right) (\tilde{\rho}) \right] \right\|_1 \\
\geqslant{} & \left| \left\| \mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}} [(H_S)_R (\mathrm{DEC})_{R_S} (\tau)] \right\|_1 \right. \\
& \left. - \left\| \mathrm{tr}_{Q_{\bar{S}} \cup \bar{R}_{\bar{S}}} \left[(H_S)_R (\mathrm{DEC})_{R_S} \circ \left(\bigotimes_{t \in [r]} \bigcirc_{i \in \bar{S}} \mathcal{D}^t_{i,\{i\}} \right) \circ \left(\bigotimes_{t \in [r]} \mathcal{D}^t_{S,S} \right) (\tilde{\rho}) \right] \right\|_1 \right|
\end{aligned}$$

with

1. the first inequality sign based on the expression of proposition 8.33, $\|H_S\| \leqslant 1$ and the contractivity of the trace norm under the cpt map $(\mathrm{DEC})_{R_S}$ by lemma 2.15,

2. the second inequality sign based on the property $\|H_j\| \|B\|_1 \geqslant \|H_j B\|_1$ by lemma 2.15 and the fact that R_S and $Q_{\bar{S}} R_{\bar{S}}$ are distinct registers and

3. the third inequality sign based on triangle inequality.

Since the trace norm of a positive semidefinite operator equals the trace and stays invariant under the unitary map $\bigotimes_{t \in [r]} \bigcirc_{i \in \bar{S}} \mathcal{D}^t_{i,\{i\}}$ which does not act on registers R_S, we obtain

$$\mathcal{O}(np^c\epsilon) \geqslant \left| \mathrm{tr}[H_S \sigma] - \mathrm{tr} \left[(H_S)_R (\mathrm{DEC})_{R_S} \circ \left(\bigotimes_{t \in [r]} \mathcal{D}^t_{S,S} \right) (\tilde{\rho}) \right] \right| .$$

The second term equals the probability that test A fails for clause S. Averaging over all clauses, using triangle inequality and the fact that test A fails on average with probability δ leads to the desired result

$$\frac{1}{m}\operatorname{tr}[H\sigma] \leqslant \mathcal{O}(\delta + np^{c}\epsilon). \qquad \square$$

Chapter 9

Conclusion

The first research focus of this thesis was an extension of the uniform diagonalization theorem in chapter 5 to make it applicable to quantum complexity classes such as QMA and BQP. Assuming BQP $\neq$ QMA the construction proves the existence of QMA-intermediate problems and the undecidability of BQP membership for QMA problems. The former motivated us in the following chapters to study hierarchies of QMA-intermediate classes by restricting the witness of the class QMA.

The necessity for the extension of the uniform diagonalization theorem arose from the fact that quantum complexity classes such as QMA and BQP consist of promise problems, while the original formulation of the theorem only covered decision problems. The extended uniform diagonalization theorem now applies to a whole variety of standard complexity classes including quantum and classically randomized classes. Still, there exist classes, such as QMIP* or the decision problem classes MA and BPP, that are not known to fulfill the prerequisites of the theorem. Broadening the classical randomized classes MA and BPP to promise problems, the theorem applies again. This and the lack of complete decision problems for MA and BPP are an indication for us that classical randomized classes should – like their quantum counterparts – be considered as sets of promise problems.

Intermediate problems are not just interesting from a theoretical but also from a practical perspective. In the current era of *NISQ* ("noisy intermediate scale quantum computation") the first quantum computing devices have been delevoped, but are still far off from offering fault-tolerant, universal quantum computing. Intermediate problems of the complexity classes BQP and NP (both with regard towards P) are therefore attractive candidates to demonstrate the advantages and usefulness of NISQ devices over classical computers. Unfortunately, the strictly intermediate problems constructed by the uniform diagonalization theorem lack practical relevance and do not owe their intermediateness to an obviously simplified circuit structure. They are therefore less suitable candidates for applications of NISQ devices and more serve as theoretical motivation for a further structural study of QMA.

This further study was pursued in chapters 6 and 7 by introducing noisy QMA classes which differ from QMA by restricting the witness to outputs of quantum channels. For specific non-uniform channels the noisy QMA classes are strictly QMA-intermediate under quantum polynomial time reductions, since they consist of all problems quantum polynomial-time reducible onto the previously constructed, strictly QMA-intermediate problems. For more physical channels, such as the partly depolarizing and partly dephasing channel, the noisy QMA classes loose the proof of strict intermediateness but allow a simple interpolation between the complexity classes QMA – BQP and QMA – QCMA, respectively.

After showing that QMA remains invariant for quantum channels whose noise decreases with the witness length due to amplification, we were able to show the same result for small constant i.i.d. noise by the tool of concatenated coding. Besides a bound for general i.i.d. noise we derived improved bounds for the error parameter of the partly depolarizing and partly dephasing channel, which tell us that QMA keeps its power even if each witness qubit is affected by 18% depolarizing or 27% dephasing noise. Better bounds may be achieved by optimizing the applied codes.

For the partly erasing channel, which also interpolates between QMA and BQP, QMA might even stay invariant for all constant error parameters below 1. This would at least be implied by the QPCP conjecture as we saw in chapter 8. Quantum complexity theorists believe in this conjecture as an analogue of the classical PCP theorem. According to the prevailing QPCP formulation QMA remains unchanged even if the verifier with constant acceptance probability gap can access only constantly many random witness qubits. We proved the equivalence to the scenario in which only the expected number of accessible qubits is constant due to equally partly erased witness qubits. As a thoroughly quantum alternative for the QPCP conjecture we propose the same statement with partly depolarizing instead of erasing channels. This statement is clearly stronger, but we find it hard to take up a position for its validity or invalidity.

Assuming at least the validity of the prevailing QPCP conjecture we can summarize the noisy QMA results of this thesis as follows: The complexity class QMA stays invariant if each of its n_w witnes qubits is disturbed by a noise channel

with decreasing error parameter $\frac{k}{n_w}$ due to amplifcation,

with constant error parameter k due to concatenated coding

and, in case of erasure, with increasing error parameter $1 - \frac{k}{n_w}$ due to QPCP.

Last but not least robustness of QMA against maximal erasure would imply the collapse QMA = BQP and presumably reward the prover with 1 million dollars for solving the quantum analogue of the open millenmium problem NP $\stackrel{?}{=}$ P.

The second half of chapter 8 dealt with a possible multiprover formulation of the QPCP conjecture, which had not been found before now despite its classical analogue. After an overview of the different classes of classical and quantum multiprover interactive proof systems, we improved a multiprover protocol for the local Hamiltonian problem by Fitzsimons and Vidick [3]. We slightly expanded the acceptance probability gap that prevents the establishment of a reasonable multiprover QPCP conjecture and reduced the number of provers to the minimum possible by preserving the characteristic protocol structure of energy and code space check.

The following interesting open tasks emerge from the studied topics in this thesis:

- Find a physically relevant problem that is QMA-intermediate under reasonable complexity theoretic assumptions, like Graph Isomorphism is for NP (recall the discussion in section 5.1).
- Find non-trivial, interesting complete problems for noisy QMA classes.

- Improve the upper bounds on the witness noise up to which QMA remains invariant by optimizing the applied concatenated codes.
- Find an upper robustness bound on the witness noise by a Shannon theoretic argument similar to the discussion at the end of section 6.3.
- Find a lower bound on the witness noise for which QMA collapses to the classes QCMA or BQP.

- Either disprove the possibly stronger QPCP conjecture in terms of i.i.d. depolarizing noise or show its equivalence to the prevailing QPCP conjecture.
- Find a multiprover protocol for the Local Hamiltonian problem allowing a reasonable equivalent multiprover QPCP conjecture.
- Prove that the three necessary provers for the presented multiprover protocol can be assumed to be classical by adapting the techniques of [77].

And, of course, prove BQP $\neq$ QMA and the QPCP conjecture!

Publications

Dziemba, F. A., “Uniform Diagonalization Theorem for Complexity Classes of Promise Problems including Randomized and Quantum Classes.” December 2017. arXiv:1712.07276.

Dziemba, F. A., “Robustness of QMA against Witness Noise,” *Quantum Information and Computation*, vol. 17, November 2017. arXiv:1611.07332.

Dziemba, F. A., “Adiabatic Quantum Computation.” Master’s Thesis. October 2016. arXiv:1610.04708.

Beer, K. and Dziemba, F. A., “Phase Context Decomposition of Diagonal Unitaries for Higher-dimensional Systems,” *Physical Review A*, vol. 93, May 2016. arXiv:1511.05758.

CV

Friederike Anna Dziemba

2014 - 2018	PhD studies in Physics *Quantum Information Group & RTG "Quantum mechanical noise in complex systems" at the Leibniz Universität Hannover*
2012 - 2014	Master of Science in Physics *Master's thesis "Adiabatic quantum computation", Leibniz Universität Hannover*
2012	Semester abroad *Technion Israel Institute of Technology*
2010 - 2014	Tutor "Complexity of Algorithms" and "Theoretical Computer Science"
2009 - 2010	Student assistant for the project "Quantus 2" (Quantum Optics Group)
2009 - 2012	Bachelor of Science in Physics *Bachelor's thesis "Simulation of rotations of rigid bodies", Leibniz Universität Hannover*
2007 - 2015	Bachelor of Science in Mathematics *Bachelor's thesis "Frobenius pseudo primes", Fernuniversität Hagen*
2008 - 2009	Tutor "Technical Mechanics 3"
2007 - 2010	Bachelor of Engineering *Bachelor's thesis "De-echoing of audio signals", Beuth Hochschule für Technik Berlin*
1997 - 2006	German Abitur *Städtisches Gymnasium Porta Westfalica*
2003 - 2004	High School graduation *Akron, Ohio*

Bibliography

[1] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM Journal on Computing*, vol. 26, October 1997. arXiv:quant-ph/9508027v2.

[2] J. Rosenberger, "P vs. NP Poll Results," *Communications of the ACM*, vol. 55, May 2012.

[3] J. Fitzsimons and T. Vidick, "A Multiprover Interactive Proof System for the Local Hamiltonian Problem," in *Proceedings of the 6th Innovations in Theoretical Computer Science*, ACM, 2015. arXiv:1409.0260.

[4] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 10 ed., 2010.

[5] C. M. Dawson and M. A. Nielsen, "The Solovay-Kitaev Algorithm," *Quantum Information and Computation*, vol. 6, January 2006. arXiv:quant-ph/0505030.

[6] C. King and M. B. Ruskai, "Minimal Entropy of States Emerging from Noisy Quantum Channels," *IEEE Transactions on Information Theory*, vol. 47, January 2001. arXiv:quant-ph/9911079.

[7] M. Wolf, "Quantum Channels & Operations: Guided Tour." `https://www-m5.ma.tum.de/foswiki/pub/M5/Allgemeines/MichaelWolf/QChannelLecture.pdf`, 2012. Lecture notes.

[8] F. Hiai and M. B. Ruskai, "Contraction Coefficients for Noisy Quantum Channels," *Journal of Mathematical Physics*, vol. 57, December 2016. arXiv:1508.03551.

[9] D. Aharonov, A. Kitaev, and N. Nisan, “Quantum Circuits with Mixed States,” in *Proceedings of the 30th ACM Symposium on Theory of Computing*, ACM, 1998. arXiv:quant-ph/9806029.

[10] M. M. Wilde, *Quantum Information Theory*. Cambridge University Press, 2 ed., 2017. arXiv:1106.1445.

[11] J. Watrous, “Theory of Quantum Information.” `https://cs.uwaterloo.ca/~watrous/LectureNotes/CS766.Fall2011/all.pdf`, 2011. Lecture notes.

[12] C. H. Papadimitriou, *Computational Complexity*. Addison Wesley Longman, 1994.

[13] S. Arora and B. Barak, *Computational Complexity – A Modern Approach*. Cambridge University Press, 2009.

[14] O. Goldreich, *Computational Complexity – A Conceptual Perspective*. Cambridge University Press, 2008.

[15] A. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation (Graduate Studies in Mathematics)*, vol. 47. AMS, 2002.

[16] J. Watrous, “Quantum Computational Complexity,” in *Encyclopedia of Complexity and Systems Science* (R. A. Meyers, ed.), Springer New York, 2009. arXiv:0804.3401.

[17] A. W. Strzeboński, “Computing in the Field of Complex Algebraic Numbers,” *Journal of Symbolic Computation*, vol. 24, December 1997.

[18] C. Marriott and J. Watrous, “Quantum Arthur-Merlin Games,” *Computational Complexity*, vol. 14, June 2005. arXiv:cs/0506068.

[19] S. A. Cook, “The Complexity of Theorem-Proving Procedures,” in *Proceedings of the 3rd ACM Symposium on Theory of Computing*, ACM, 1971.

[20] O. Goldreich, “On Promise Problems: A Survey,” in *Theoretical Computer Science* (O. Goldreich, A. L. Rosenberg, and A. L. Selman, eds.), Springer Berlin Heidelberg, 2006.

[21] S. Bravyi, D. P. DiVincenzo, R. I. Oliveira, and B. M. Terhal, “The Complexity of Stoquastic Local Hamiltonian Problems,” *Quantum Information and Computation*, vol. 8, May 2008. arXiv:quant-ph/0606140.

[22] E. Knill and R. Laflamme, “Quantum Computing and Quadratically Signed Weight Enumerators,” *Information Processing Letters*, vol. 79, May 2001. arXiv:quant-ph/9909094.

[23] S. Gharibian and J. Sikora, “Ground State Connectivity of Local Hamiltonians,” in *Proceedings of the 42nd International Colloquium of Automata, Languages, and Programming*, Springer Berlin Heidelberg, 2015. arXiv:1409.3182.

[24] D. Aharonov and T. Naveh, “Quantum NP - A Survey.” October 2002. arXiv:quant-ph/0210077.

[25] J. Kempe, A. Kitaev, and O. Regev, “The Complexity of the Local Hamiltonian Problem,” *SIAM Journal on Computing*, vol. 35, May 2006. arXiv:quant-ph/0406180.

[26] T. Cubitt and A. Montanaro, “Complexity Classification of Local Hamiltonian Problems,” *SIAM Journal on Computing*, vol. 45, June 2016. arXiv:1311.3161.

[27] B. A. Trakhtenbrot, “A Survey of Russian Approaches to Perebor (Brute-Force Searches) Algorithms,” *IEEE Annals of the History of Computing*, vol. 6, October 1984.

[28] E. Knill and R. Laflamme, “Theory of quantum error-correcting codes,” *Physical Review A*, vol. 55, February 1997.

[29] A. Cross, G. Smith, J. A. Smolin, and B. Zeng, “Codeword Stabilized Quantum Codes,” *IEEE Transactions on Information Theory*, vol. 55, January 2009. arXiv:0708.1021.

[30] A. Ekert and C. Macchiavello, “Quantum Error Correction for Communication,” *Physical Review Letters*, vol. 77, September 1996. arXiv:quant-ph/9602022.

[31] A. R. Calderbank, E. M. Rains, P. M. Shor, and N. J. A. Sloane, “Quantum Error Correction via Codes over GF(4),” *IEEE Transactions on Information Theory*, vol. 44, July 1998. arXiv:quant-ph/9608006.

[32] E. Rains, "Nonbinary Quantum Codes." March 1997. arXiv:quant-ph/9703048.

[33] D. Gottesman, *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, Pasadena, California, 1997. arXiv:quant-ph/9705052.

[34] J. Katz, "Complexity Theory: Graph Non-Isomorphism is in AM (Lecture 17)." `http://www.cs.umd.edu/~jkatz/complexity/f11/lecture17.pdf`, 2011. Lecture notes.

[35] L. Babai and S. Moran, "Arthur-Merlin Games: A Randomized Proof System, and a Hierarchy of Complexity Classes," *Journal of Computer and System Sciences*, vol. 36, April 1988.

[36] P. Beame, "Computational Complexity Essentials: Arthur-Merlin Games (Lecture 10)." `https://courses.cs.washington.edu/courses/cse532/04sp/lect10.pdf`, 2004. Lecture notes.

[37] S. Gharibian, M. Santha, J. Sikora, A. Sundaram, and J. Yirka, "Quantum generalizations of the polynomial hierarchy with applications to QMA(2)." May 2018. arXiv:1805.11139.

[38] R. E. Ladner, "On the Structure of Polynomial Time Reducibility," *Journal of the ACM*, vol. 22, January 1975.

[39] T. Schäfer, "The Complexity of Satisfiability Problems," in *Proceedings of the 10th ACM Symposium on Theory of Computing*, ACM, 1978.

[40] F. A. Dziemba, "Uniform Diagonalization Theorem for Complexity Classes of Promise Problems including Randomized and Quantum Classes." December 2017. arXiv:1712.07276.

[41] D. Schmidt, "The Recursion-Theoretic Structure of Complexity Classes," *Theoretical Computer Science*, vol. 38, 1985.

[42] H. Vollmer, "The Gap-Language-Technique Revisited," in *Proceedings of the 4th Workshop on Computer Science Logic*, Springer Berlin Heidelberg, 1991.

[43] K. W. Regan and H. Vollmer, "Gap-Languages and log-Time Complexity Classes," *Theoretical Computer Science*, vol. 188, November 1997.

[44] U. Schöning, "A Uniform Approach to Obtain Diagonal Sets in Complexity Classes," *Theoretical Computer Science*, vol. 18, April 1982.

[45] J. L. Balczar, J. Diaz, and J. Gabarró, *Structural Complexity I*. Springer Berlin Heidelberg, 1995.

[46] S. Bravyi, "Efficient Algorithm for a Quantum Analogue of 2-SAT." February 2006. arXiv:quant-ph/0602108.

[47] R. Downey and L. Fortnow, "Uniformly Hard Languages," *Theoretical Computer Science*, vol. 298, April 2003.

[48] S. Even, A. L. Selman, and Y. Yacobi, "The Complexity of Promise Problems with Applications to Public-Key Cryptography," *Information and Control*, vol. 61, May 1984.

[49] T. Morimae, K. Fujii, and H. Nishimura, "Quantum Merlin-Arthur with Noisy Channel." August 2016. arXiv:1608.04829.

[50] F. A. Dziemba, "Robustness of QMA against Witness Noise," *Quantum Information and Computation*, vol. 17, November 2017. arXiv:1611.07332.

[51] D. Janzing, P. Wocjan, and T. Beth, "Non-Identity Check is QMA-Complete," *International Journal of Quantum Information*, vol. 3, September 2005. arXiv:quant-ph/0305050.

[52] P. Wocjan, D. Janzing, and T. Beth, "Two QCMA-complete Problems," *Quantum Information and Computation*, vol. 3, June 2003. arXiv:quant-ph/0305090.

[53] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, "Quantum Fingerprinting," *Phyical Review Letters*, vol. 87, September 2001. arxiv:quant-ph/0102001.

[54] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, "Capacities of Quantum Erasure Channels," *Physical Review Letters*, vol. 78, April 1997. arXiv:quant-ph/9701015.

[55] S. Beigi, N. Datta, and F. Leditzky, "Decoding Quantum Information via the Petz Recovery Map," *Journal of Mathematical Physics*, vol. 57, August 2016. arXiv:1504.04449.

[56] C. Morgan and A. J. Winter, ""Pretty strong" Converse for the Quantum Capacity of Degradable Channels," *IEEE Transactions on Information Theory*, vol. 60, January 2014. arXiv:1301.4927.

[57] D. Aharonov and M. Ben-Or, "Fault-tolerant Quantum Computation with Constant Error," in *Proceedings of the 29th ACM Symposium on Theory of Computing*, ACM, 1997. arXiv:quant-ph/9906129.

[58] B. Rahn, A. C. Doherty, and H. Mabuchi, "Exact Performance of Concatenated Quantum Codes," *Physical Review A*, vol. 66, Jun 2002. arXiv:quant-ph/0206061.

[59] J. Fern, J. Kempe, S. N. Simić, and S. Sastry, "Generalized Performance of Concatenated Quantum Codes – A Dynamical Systems Approach," *IEEE Transactions on Automatic Control*, vol. 51, March 2006. arXiv:quant-ph/0409084.

[60] J. Watrous, "Semidefinite Programs for Completely Bounded Norms," *Theory of Computing*, vol. 5, November 2009. arXiv:0901.4709.

[61] S. Arora and S. Safra, "Probabilistic Checking of Proofs: A New Characterization of NP," *Journal of the ACM*, vol. 45, January 1998.

[62] I. Dinur, "The PCP Theorem by Gap Amplification," *Journal of the ACM*, vol. 54, July 2007.

[63] S. Aaronson, "The QPCP Manifesto." `https://www.scottaaronson.com/blog/?p=139`, 2006. Blog entry.

[64] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, "The Detectibility Lemma and Quantum Gap Amplification," in *Proceedings of the 41st ACM Symposium on Theory of Computing*, ACM, 2009.

[65] D. Aharonov, I. Arad, and T. Vidick, "The Quantum PCP Conjecture," *SIGACT News*, vol. 44, June 2013. arXiv:1309.7495.

[66] U. Feige, "On the Success Probability of the Two Provers in One-Round Proof Systems," in *Proceedings of the 6th Annual Structure in Complexity Theory Conference*, IEEE, 1991.

[67] R. Raz, "A Parallel Repetition Theorem," *SIAM Journal on Computing*, vol. 27, June 1998. arXiv:quant-ph/0406180.

[68] J. Håstad, "Some Optimal Inapproximability Results," *Journal of the ACM*, vol. 48, July 2001.

[69] T. Vidick, "Quantum PCP Conjectures." `http://users.cms.caltech.edu/~vidick/teaching/286_qPCP/lecture7.pdf`, 2014. Lecture notes.

[70] S. Bravyi, S. DiVincenzo, D. Loss, and B. Terhal, "Quantum Simulation of Many-Body Hamiltonians Using Perturbation Theory with Bounded-Strength Interactions," *Physical Review Letters*, vol. 101, August 2008. arXiv:0803.2686.

[71] F. G. S. L. Brandao and A. Harrow, "Product-State Approximations to Quantum States," *Communications in Mathematical Physics*, vol. 342, February 2016. arXiv:1310.0017.

[72] H. Kobayashi and K. Matsumoto, "Quantum Multi-Prover Interactive Proof Systems with Limited Prior Entanglement," *Journal of Computer and System Sciences*, vol. 66, May 2003. arXiv:cs/0102013.

[73] L. Babai, L. Fortnow, and C. Lund, "Non-Deterministic Exponential Time has Two-Prover Interactive Protocols," *Computational Complexity*, vol. 1, March 1991.

[74] T. Ito and T. Vidick, "A Multi-Prover Interactive Proof for NEXP Sound againt Entangled Provers," in *Proceedings of the 53rd IEEE Symposium on Foundations of Computer Science*, IEEE, 2012. arXiv:1207.0550.

[75] B. Reichardt, F. Unger, and U. Vazirani, "A Classical Leash for a Quantum System: Command of Quantum Systems via Rigidity of CHSH Games," in *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, ACM, 2013. arXiv:1209.0448.

[76] Y.-K. Liu, "Consistency of Local Density Matrices is QMA-Complete," in *Approximation, Randomization and Combinatorial Optimization. Algorithms and Techniques.* (J. Díaz, K. Jansen, J. D. P. Rolim, and U. Zwick, eds.), Springer Berlin Heidelberg, 2006. arXiv:quant-ph/0604166.

[77] Z. Ji, "Classical Verifications of Quantum Proofs.," in *Proceedings of the 48th ACM Symposium on Theory of Computing*, ACM, 2016. arXiv:1505.07432.

[78] M. Grassl, T. Beth, and T. Pellizzari, "Codes for the Quantum Erasure Channel," *Physical Review A*, vol. 56, July 1997. arXiv:quant-ph/9610042.

[79] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, "Holographic Quantum Error-Correcting Codes: Toy Models for the Bulk/Boundary Correspondence," *Journal of High Energy Physics*, vol. 2015, Jun 2015. arXiv:1503.06237.

[80] R. Oliveira and B. Terhal, "The Complexity of Quantum Spin Systems on a Two-dimensional Square Lattice," *Quantum Information and Computation*, vol. 8, November 2008. arXiv:quant-ph/0504050.