Computational Power of Hamiltonians in Quantum Computing

Zur Erlangung des akademischen Grades eines Doktors
der Naturwissenschaften der Fakultät für Informatik
der Universität Karlsruhe (Technische Hochschule)

genehmigte

Dissertation

von

Pawel Wocjan

aus Warschau

Tag der mündlichen Prüfung:	17.11.2003
Erster Gutachter:	Prof. Dr. Thomas Beth
Zweiter Gutachter:	Prof. Dr. Mario Rasetti

Danksagung

Mein herzlichster Dank gilt meinem Doktorvater Professor Dr. Beth, ohne den ich den Weg zu dem faszinierenden Gebiet des Quantencomputing nicht gefunden hätte. Schon vor dem Studium weckte er mein Interesse für mathematische Aspekte der Informatik. Die anregende Atmosphäre am Institut für Algorithmen und Kognitive Systeme (IAKS) und die von ihm großzügig gewährten Freiheiten haben entscheidend zu dieser Arbeit beigetragen.

Für die freundliche Übernahme des Korreferats möchte ich Professor Dr. Rasetti danken. Seine Vorträge über graphentheoretische Methoden in der Festkörperphysik haben mich zu neuen interessanten Denkansätzen inspiriert.

Meinem Freund und Zimmerkollegen Dominik Janzing verdanke ich eine angenehme und ideenreiche Zusammenarbeit. Durch ausführliche Diskussionen wurde mein Verständnis der Quantenmechanik entscheidend gefördert. Zusammen mit ihm und meinem Kollegen Martin Rötteler konnten wichtige Fortschritte im Arbeitsgebiet erreicht werden, die zu gemeinsamen Veröffentlichungen führten.

Weiterhin möchte ich Professor Dr. Lazic für Diskussionen über Codierungstheorie danken, die mir einen Zugang zur Quanteninformationstheorie eröffnet haben.

Die Quantencomputinggruppe mit Markus Grassl, Dominik Janzing, Jörn Müller-Quade, Martin Rötteler und Robert Zeier stand immer für Kommunikation und Austausch wissenschaftlicher Ideen zur Verfügung.

Meine Kollegen Ingo Boesnach, Markus Grassl, Matthias Hahn, Martin Haimerl, Jörg Moldenhauer waren mir eine unentbehrliche Hilfe bei technischen Problemen.

Thomas Decker und Dominik Janzing danke ich für eine überaus sorgfältige und gründliche Durchsicht dieser Arbeit. Nicole Ising möchte ich für die sprachlichen Korrekturen bei der Einleitung danken.

Contents

Zusammenfassung der Dissertation

Dem Gebiet des Quantencomputings und der Quanteninformationstheorie wird seit einigen Jahren von einer interdisziplinären Gruppe von Forschern wie Informatikern, Mathematikern und Physikern große Aufmerksamkeit gewidmet. Motiviert wird diese fächerübergreifende Forschung durch das ehrgeizige Ziel, einen Quantencomputer zu bauen. Unter einem Quantencomputer versteht man eine Berechnungsmachine, die sich der Quantenmechanik eigene Effekte zu Nutze macht. Damit soll eine neue Art der Informationsverarbeitung realisiert werden mit dem Ziel Berechnungsprobleme effizienter lösen zu können, als es mit klassischen Computern möglich wäre.

Das Bestreben der Experimentalphysiker richtet sich darauf, Techniken zu entwickeln, mit denen Quantensysteme derart kontrolliert werden können, dass sie als Hardware für künftige Quantencomputer dienen können. Vom Standpunkt eines Informatikers aus gibt es verschiedene Herausforderungen. Die Ziele der Quanteninformatik sind analog zu den Zielen der theoretischen Informatik. Diese bestehen hauptsächlich darin, die Komplexität von Berechnungsproblemen zu klassifizieren und, falls möglich, effiziente Algorithmen zu deren Lösung anzugeben. Die Definition der Berechnungskomplexität kann dabei nur im Kontext eines Berechnungsmodells formuliert werden. Dieses spezifiziert die Menge der elementaren Operationen, die dem Computer zur Verfügung stehen. Ein Algorithmus beschreibt eine Sequenz solcher elementarer Operationen, die benötigt werden, um den Input (die Spezifikation des Problems) zu dem gewünschten Output (der Lösung des Problems) zu transformieren. Die Zeitkomplexität eines Problems ist die minimale Anzahl elementarer Operationen, die notwendig sind, um es zu lösen. Die wichtigsten Forschungsthemen der Quanteninformatik sind:

1. die Entwicklung abstrakter Berechnungsmodelle, die die elementaren Operationen des Quantencomputers festlegen (diese Festlegung sollte sich dabei an den physikalischen Gesetzen des betrachteten Systems orientieren),

2. die Klassifikationen von Berechnungsproblemen hinsichtlich ihrer Komplexität innerhalb dieser Modelle,

3. die Entwicklung effizienter Quantenalgorithmen, um diese verschiedenen Arten von Berechnungsproblemen zu behandeln.

Die Entwicklung der *Berechnungsmodelle* ist zentral, da diese die Komplexität der Probleme und die Effizienz der Algorithmen bestimmen. Deswegen konzentriert sich die vorliegende Arbeit auf Berechnungsmodelle und ihre physikalische Rechtfertigung. Das gängigste Modell des Quantencomputers ist das Quantenschaltkreismodell, das als eine quantenmechanische Erweiterung des Modells der booleschen Schaltkreise verstanden werden kann. Die elementaren Operationen im Quantenschaltkreismodell werden durch Gatter beschrieben, die nur auf ein oder zwei Quantenbits zugreifen. Die gewünschten Transformationen werden implementiert, indem man eine geeignete Sequenz von elementaren Gattern anwendet. Die Komplexität einer Transformation entspricht dabei der benötigten Anzahl elementarer Gatter.

In dieser Arbeit hingegen bauen wir auf einem Berechnungsmodell auf, das sich mehr an den Eigenschaften quantenmechanischer Systeme orientiert. Unsere Vorgehensweise begründet sich in einem kontrolltheoretischen Modell des Quantencomputers. Innerhalb dieses Modells werden die gewünschten Transformationen implementiert, indem man die folgenden zwei Schritte mehrmals anwendet: Ausführung einer Kontrolloperation und die natürliche Entwicklung des Systems für eine bestimmte Zeit. Die natürliche Zeitentwicklung des Systems wird durch seinen Hamiltonoperator charakterisiert. Genauer gesagt definiert der Hamiltonoperator die lineare Differentialgleichung, die die Zeitentwicklung beschreibt (die sogenannte Schrödingergleichung). Somit ist das kontrolltheoretische Modell der Physik näher als das Quantenschaltkreismodell, da es berücksichtigt, dass die quantenmechanischen Zeitentwicklungen kontinuierliche Prozesse sind, die durch Hamiltonoperatoren beschrieben werden. Die Komplexität einer Transformation im kontrolltheoretischen Modell wird durch die benötigte Zeit und die Anzahl der Kontrolloperationen angegeben. Diese Betrachtungen führen zu einer erweiterten Definition von Quantenalgorithmen und einer anderen Sichtweise von Komplexität. In diesem Zusammenhang betrachten wir zeitkontinuierliche Algorithmen (im Gegensatz zu den diskreten Algorithmen, die auf dem Quantenschaltkreismodell beruhen).

Das kontrolltheoretische Modell kann jedoch nicht nur dazu verwendet werden, Transformationen auszuführen, die klassische Berechnungsprobleme lösen sollen, sondern auch Transformationen zu implementieren, die die Zeitentwicklung eines anderen quantenmechanischen Systems repräsentieren. Die Betrachtung des Aufwands, die Zeitentwicklung eines Quantensystems durch ein gegebenes System zu simulieren, definiert sozusagen einen Vergleich der "Berechnungsmächtigkeit" der verschiedenen Systeme. Die Bestimmung des Aufwands stellt im allgemeinen ein schwieriges Problem dar. Lösen läßt es sich weitgehend, wenn man es nur für infinitesimale Zeitintervalle betrachtet. Dabei werden die Kontrolloperationen so gewählt, dass die Zeitentwicklung des einen Systems die des anderen bis auf einen

Fehler widerspiegelt, der von zweiter Ordnung der Länge des betrachteten Intervalls ist. Diese Vorgehensweise basiert auf der sogenannten Average-Hamiltonian-Theorie, die eine wichtige Rolle in der Kernspinresonanz-Spektroskopie (Nuclear Magnetic Resonance Spectroscopy) spielt.

Die gegenseitige Simulation von Zeitentwicklungen auf *infinitesimalen* Zeitintervallen wurde von uns formal definiert und gegenseitige Simulation von *Hamiltonoperatoren* genannt. Das Hauptresultat der vorliegenden Arbeit ist eine Theorie, die die für die gegenseitige Simulation von Hamiltonoperatoren benötigten Berechnungsressourcen untersucht. Wir geben untere Schranken für die Simulationskomplexität an. Eine wichtige Methode zur Herleitung dieser Schranken ist die Charakterisierung der Hamiltonoperatoren durch Graphen; dabei repräsentieren die Graphen die Kopplungstopologie der Hamiltonoperatoren, d.h. sie kennzeichnen, welche Paare von Subsystemen interagieren. Unsere Schranken zeigen, dass die Simulationskomplexität stark von den Eigenschaften der Graphen abhängt. Wir entwickeln effiziente Simulationsalgorithmen, die auf Methoden der Graphentheorie, der Darstellungstheorie endlicher Gruppen, der Designtheorie und der konvexen Optimierung beruhen. Optimalität der Algorithmen läßt sich in für das Quantencomputing relevanten Fällen mit Hilfe der unteren Schranken beweisen. Des Weiteren zeigen wir, wie elementare Gatter des Quantenschaltkreismodells innerhalb des kontrolltheoretischen Modells implementiert werden können. Es zeigt sich, dass Gatter, die auf disjunkten Qubitpaaren operieren, parallel ausgeführt werden. Dies zeigt eine interessante Verbindung zwischen dem Schaltkreismodell und dem kontrolltheoretischen Modell auf.

Mit dieser Fundierung der elementaren Quantengatter innerhalb des kontrolltheoretischen Modells wird der Notwendigkeit Rechnung getragen, Berechnungsmodelle auf Naturgesetzen aufzubauen. Dies steht im Einklang zu Deutschs Gedanken in seinem Buch "Fabric of Reality" [Deu97]. Darin kritisiert er, dass die Berechnungstheorie traditionell als ein Gegenstand der reinen Mathematik untersucht wurde. Der Autor plädiert für eine auch auf der Physik aufbauende Komplexitätstheorie. Schließlich stellen Berechnungsmodelle lediglich Idealisierungen bestehender physikalischer Systeme dar. Von einem fundamentalen Standpunkt aus gesehen sind es die physikalischen Gesetze, die festlegen, welche Probleme effizient mit den jeweiligen Computern gelöst werden können.

Alle gängigen Berechnungsmodelle (wie Turingmaschinen, Familien boolescher Schaltkreise und zelluläre Automaten) basieren auf klassischer Physik. Aber die fundamentale Theorie der Natur ist die Quantentheorie. Sie ist es, die Berechnungsmodelle und Komplexität bestimmt. Abgesehen von diesen theoretischen Gründen gibt es auch praktische Gründe, eine Theorie der Informationsverarbeitung auf der Basis der Quantentheorie zu entwickeln. Eine solche Theorie wird für das klassische Rechnen und die neu entstehenden Nanotechnologien von Bedeutung werden. Man muss hierbei in Betracht ziehen, dass in künftigen Computern mit zunehmender Integrationsdichte jeder Berechnungsprozess mehr und mehr als

ein mikroskopischer Prozess zu beschreiben ist, der quantenmechanische Effekte einbezieht. Deswegen wird diese erweiterte Theorie wichtig sein, wenn klassische Rechner derart miniaturisiert werden, dass die logischen Zustände des Bits durch einfache quantenmechanische Freiheitsgrade repräsentiert werden. Beispielsweise kann Information im Zustand eines Atoms oder eines Photons codiert werden. (Dies ist bereits bei den ersten experimentellen Realisierungen des Quantencomputings mit wenigen Qubits und der Quantenkryptographie der Fall.)

Die Motivation für unsere Theorie der gegenseitigen Simulierbarkeit von Hamiltonoperatoren geht auf Feynmans Vorschlag zur Simulation von Quantensystemen zurück. Er bemerkte, dass alle Ansätze, Dynamiken (d.h. die zeitlichen Entwicklungen) allgemeiner Vielteilchenquantensysteme mit klassischen Computern zu simulieren, mit einer exponentiellen Verlangsamung (aufgrund der exponentiell in der Anzahl der wechselwirkenden Komponenten anwachsenden Dimension des System-Hilbert-Raums) verbunden sind. Um diesen Zeitaufwand zu umgehen, schlug Feynman vor, Quantensysteme anstatt klassischer Systeme für die Simulation von Quantendynamiken zu benutzen. Schon Feynman vermutete die Existenz eines universellen Quantensimulators. Unter einem Quantensimulator versteht man eine Berechnungsmaschine, die auf der quantenmechanischen Ebene operiert und die eine effiziente Simulation von Dynamiken verschiedener Vielteilchenquantensyteme ermöglicht. Während erste Arbeiten zu diesem Thema Simulationen angaben, die auf dem Quantenschaltkreismodell beruhen, gehen wir von einem fundamentaleren Standpunkt aus. Wie bereits oben erwähnt, schlagen wir ein Modell vor, bei dem die gewünschten Dynamiken simuliert werden, indem die natürliche Hamiltonsche Zeitentwicklung des Systems durch externe Kontrolloperationen unterbrochen wird.

Eine interessante Anwendung unserer Methoden der gegenseitigen Simulation von Hamiltonoperatoren bietet sich im adiabatischen Quantencomputing an. Hier wird zur Lösung eines harten Berechnungsproblem das System im Energieminimum eines einfach gewählten Hamiltonoperators initialisiert. Der Hamiltonoperator wird langsam variiert und geht am Ende des Algorithmus in einen Hamiltonoperator über, dessen Energieminimum die Lösung des betrachten Problems codiert. Während der Berechnung befindet sich das System in einem Zustand minimaler Energie bezüglich des aktuell vorliegenden Hamiltonoperators. Unsere Simulationsmethoden erlauben eine effiziente Realisierung der adiabatischen Algorithmen innerhalb des kontrolltheoretischen Modells. Dabei legen wir Wert darauf, als Berechnungsressource nur Hamiltonoperatoren zu verwenden, die *physikalisch realistische* Kopplungstopologien aufweisen. Wir konstruieren solche Hamiltonoperatoren, deren Energieminima die Lösungen des NP-vollständige Problem “independent set” codieren.

Zusammenfassend beinhalten die wichtigsten Resultate der vorliegenden Arbeit:

- die formale Definition der gegenseitigen Simulation von Hamiltonoperatoren,
- die Herleitung unterer und oberer Schranken für die Komplexität der gegenseitigen Simulation von Hamiltonoperatoren,
- die Entwicklung optimaler Simulationsalgorithmen, die Methoden der Darstellungstheorie endlicher Gruppen (irreduzible projektive Darstellungen), der Designtheorie (orthogonale Arrays, Differenzschemata, Hadamardmatrizen und Spreads) und der algebraischen Graphentheorie benutzen,
- die Anwendung der gegenseitigen Simulation von Hamiltonoperatoren für das adiabatische Quantencomputing zur Lösung NP-vollständiger Berechnungsprobleme und die Beschreibung neuer Berechnungsprobleme und ihrer Quantenkomplexitätsklassen.

Gliederung der Dissertation

Kapitel 1 Quantencomputing und Quantenschaltkreismodell

Hier wird eine kurze Einführung in die wichtigsten Ideen des Quantencomputings und der Quanteninformationstheorie gegeben. Das gängigste Modell des Quantencomputers, das Quantenschaltkreismodell, wird eingeführt. Es ist ein diskretes Modell, das in enger Analogie zu dem klassischen Modell der booleschen Schaltkreise steht. Hierauf aufbauend wird die Quantenkomplexitätsklasse BQP eingeführt. Diese Klasse umfasst diejenigen Probleme, die effizient (d.h. mit einer polynomialen Anzahl von elementaren Gattern) auf einem Quantenrechner gelöst werden können.

Kapitel 2 Kontrolltheoretisches Modell und Simulation von Hamiltonoperatoren

Dieses Kapitel führt ein kontrolltheoretisches Modell des Quantencomputers ein, das sich enger an die physikalische Natur der Prozesse anlehnt als das Quantenschaltkreismodell. Die Zeitentwicklung in diesem Modell ist durch den Hamiltonoperator des Systems und durch die externen Kontrolloperationen bestimmt. Insbesondere berücksichtigt dieses Modell, dass alle physikalischen Dynamiken kontinuierliche Prozesse sind.

Basierend auf dieser kontrolltheoretischen Beschreibung definieren wir das mathematische Modell der gegenseitigen Simulation von Hamiltonoperatoren. Die

Komplexität der Simulation von Hamiltonoperatoren durch einen gegebenen Hamiltonoperator wird durch den Zeitaufwand und die Anzahl der Zeitschritte definiert. Der Zeitaufwand gibt das Verhältnis von der simulierten Laufzeit zur dafür benötigten Laufzeit des simulierenden Hamiltonoperators an. Die Anzahl der Zeitschritte ist die Anzahl der Kontrolloperationen, die notwendig sind, um den gewünschten Hamiltonoperator zu simulieren.

Kapitel 3 Decoupling- und Zeitumkehralgorithmen

Wir betrachten in diesem Kapitel Decoupling und Zeitumkehr (Refokussierung). Decoupling ist ein Spezialfall eines Simulationsalgorithmus, bei dem die zu simulierende Zeitentwicklung die triviale Zeitentwicklung (der Stillstand des Quantensystems) ist. Hingegen bewirkt Zeitumkehr, dass sich das Quantensystem so verhält, als ob seine Zeitentwicklung rückwärts laufen würde. Beide Aufgaben spielen eine wichtige Rolle in der Kernspinresonanz-Spektroskopie (Nuclear Magnetic Resonance Spectroscopy).

Als Verallgemeinerung bestehender Verfahren für das Decoupling von gekoppelten Qubits, konstruieren wir effiziente Algorithmen für gekoppelte Qudits, d.h. zusammengesetzte Systeme, die aus mehreren Subsystemen einer beliebigen Dimension bestehen. Diese Verallgemeinerung basiert auf dem Konzept eines *Annihilators*, das eine minimale Menge an Kontrolloperationen angibt, die ausreicht, um die Zeitentwicklung abzuschalten. Des Weiteren zeigen wir, wie man aus Decouplingalgorithmen effiziente Algorithmen für die Zeitumkehr erhalten kann.

Unsere Algorithmen können bei einer breiten Klasse von Hamiltonoperatoren eingesetzt werden (sogar wenn die genaue Struktur der Hamiltonoperatoren unbekannt ist), während frühere Resultate in der Kernspinsresonanz-Spektroskopie für spezielle Hamiltonoperatoren zugeschnitten waren. Unsere Konstruktionen der Algorithmen beruhen auf Methoden der Darstellungstheorie endlicher Gruppen und der Designtheorie (orthogonale Arrays, Differenzschemata, Hadamardmatrizen, und Spreads). Die Decouplingalgorithmen stellen ein wichtiges Werkzeug dar, um die Universalität der gegenseitigen Simulation von Hamiltonoperatoren in Kapitel 4 herzuleiten.

Kapitel 4 Universelle Simulation von Hamiltonoperatoren

Wir untersuchen die Bedingungen, die die Menge der externen Kontrolloperationen erfüllen muss, um eine universelle Simulation von Hamiltonoperatoren zu erlauben. Dies führt zu dem Begriff der *Transformergruppen*. Wir charakterisieren die endlichen Transformergruppen mit einer notwendigen und hinreichenden Bedingung, die mit Hilfe von Gruppencharakteren formuliert wird. Mittels dieser Beschreibung konnten wir verschiedene Transformergruppen finden, indem alle Gruppen bis Größe 255 mit Computeralgebrasystemen durchsucht wurden.

Kapitel 5 Komplexitätsschranken für Simulation von Hamiltonoperatoren

Wir konzentrieren uns in diesem Kapitel auf die Komplexität, verschiedene Zeitentwicklungen mittels eines gegebenen Hamiltonoperators zu simulieren. Dabei werden Zeitaufwand und Anzahl der Zeitschritte als Komplexitätsmaße benutzt. Basierend auf der Tatsache, das normalerweise die vorkommenden Hamiltonoperatoren aus Paarwechselwirkungen bestehen, entwickeln wir eine nützliche Notation (sogenannte Kopplungsmatrizen), die eine Beschreibung der Hamiltonoperatoren und der Kontrolloperationen mit Methoden der algebraischen Graphentheorie ermöglicht. Wir zeigen, dass die Simulationskomplexität stark von den Invarianten der Graphen abhängt, die die Kopplungstopologie der Hamiltonoperatoren beschreiben. Diese Invarianten sind beispielsweise die chromatische Zahl, der chromatische Index, der "clique coloring"-Index, und die Eigenwerte der Adjazenzmatrizen. Mit Hilfe dieser Methoden leiten wir untere und obere Schranken für die Simulationskomplexität allgemeiner Hamiltonoperatoren her. Dieses Vorgehen erlaubt auch, neue Schranken für Grapheninvarianten abzuleiten.

Kapitel 6 Spezielle Simulationsaufgaben und ihre Komplexität

Wir betrachten spezielle Simulationsaufgaben und analysieren ihre Komplexität, indem wir die in Kapitel 5 beschriebenen Komplexitätsschranken anwenden. Zuerst beweisen wir, dass unsere Algorithmen für Decoupling und Zeitumkehr aus Kapitel 3 optimal sind. Dann wenden wir unsere Komplexitätsschranken auf einige konkrete physikalische Systeme an, die für die Realisierung eines Quantencomputers relevant werden könnten. Beispielsweise betrachten wir spezielle Arten von Ising-Wechselwirkungen, die geeignet sind, um den sogenannten "one-way computer" zu implementieren, der einen interessanten Vorschlag für eine skalierbare Hardwarearchitektur darstellt. Weiterhin können mit diesen Wechselwirkungen Graphencodes realisiert werden, mit deren Hilfe Quanteninformation gegen Fehler geschützt werden kann. Die Komplexitätsschranken folgen bei beiden Systemen direkt aus den Eigenwerten der zugehörigen Graphen.

Eine wichtige Simulationsaufgabe ist die Implementierung der Gatter des Quantenschaltkreismodells innerhalb des kontrolltheoretischen Modells. Daher geben wir Simulationstechniken an, um Quantengatter mit Hilfe von Hamiltonoperatoren zu implementieren. Insbesondere zeigen wir, dass Quantengatter, die auf disjunkten Paaren von Qubits wirken, parallel ausgeführt werden können. Basierend auf diesem Resultat schlagen wir ein physikalisch motiviertes Komplexitätsmaß für Quantenschaltkreise vor. Dieses Maß stellt eine Verbindung zwischen dem Quantenschaltkreismodell und dem kontrolltheoretischen Modell her. Die Berechnungsmächtigkeit des Schaltkreismodells und des kontrolltheoretischen Modells wird dabei im Detail verglichen. Für spezielle Wechselwirkungstypen zeigen wir,

dass das kontrolltheoretische Modell mächtiger ist als das Quantenschaltkreismodell, da es einen höheren Grad an Parallelität erlaubt.

Kapitel 7 Quantenkomplexitätsklassen

Dieses Kapitel geht kurz auf einige Aspekte der Quantenkomplexitätstheorie ein. Damit verfolgen wir zweierlei Ziele. Einerseits bildet die Tatsache, dass es schwierig ist, Energieminima von Hamiltonoperatoren zu bestimmen, die Grundlage für das Verständnis des adiabatischen Quantencomputings, das in Kapitel 8 behandelt wird. Auf der anderen Seite lässt sich nach einem kurzen Einblick in die Komplexitätstheorie besser abschätzen, wie komplex die kontrolltheoretischen Probleme der optimalen Implementierung von Transformationen sind. Dadurch erhält unser Zugang, der das Problem auf nur infinitesimalen Zeiten betrachtet, im Nachhinein eine Rechtfertigung.

Ausgangspunkt für unsere komplexitätstheoretischen Betrachtungen ist die Tatsache, dass die Lösungen der NP-vollständigen Probleme "max cut" und "max independent set" in den Grundzuständen einfacher Hamiltonoperatoren mit Paarwechselwirkungen codiert werden können. Anschließend betrachten wir die Quantenkomplexitätsklassen QMA und QCMA, die zwei möglichen quantenmechanischen Verallgemeinerungen von NP, und beschreiben das Problem "local Hamiltonian". Dieses Problem war bis vor kurzem das einzige, von dem bewiesen wurde, dass es QMA-vollständig ist; QCMA-vollständige Probleme waren hingegen keine bekannt. Unsere wichtigsten Beiträge sind neue Probleme, von denen wir beweisen, dass sie für beide Klassen vollständig sind. Wir zeigen, dass das Problem "low-energy and low-complexity states of local Hamiltonians" QCMA-vollständig ist. Dies ist eine Erweiterung des Problems "local Hamiltonian". Beide Probleme hängen mit den Fragen "hat ein gegebener Hamiltonoperator Zustände mit Energie kleiner als eine gegebene Schranke?" bzw. "falls ja, können diese Zustände effizient präpariert werden?" zusammen. Die Hamiltonoperatoren, die bei diesen Problemen benutzt werden, sind komplizierter als diejenigen bei NP. Des Weiteren zeigen wir, dass die Probleme "identity check" und "identity check on basis states" vollständig für QMA bzw. QCMA sind. Diese Probleme bestehen darin, zu entscheiden, ob ein Quantenschaltkreis wie der Identitätoperator auf dem ganzen Raum bzw. auf den Basiszuständen operiert. Ein wichtiges Problem, das auf "identity check" reduziert werden kann, ist "equivalence check". Dieses besteht darin, zu entscheiden, ob zwei Quantenschaltkreise dieselbe Transformation implementieren. Die Tatsache, dass "equivalence check" QMA- bzw. QCMA-vollständig ist, deutet darauf hin, dass es schwierig ist, zu entscheiden, ob die Zeitentwicklungen bzgl. zweier gegebener Hamiltonoperatoren nach bestimmten Zeiten zu derselben Transformation führen. In gleicher Weise legt das Resultat nahe, dass es ein schwieriges Problem ist, zu entscheiden, ob ein Quantenschaltkreis durch eine hamiltonsche Zeitentwicklung mittels eines gegebenen Hamiltonoperators implementiert werden kann.

Deswegen scheint es, dass es schwierig ist, die Berechnungmächtigkeit verschiedener Hamiltonoperatoren innerhalb des kontrolltheoretischen Modells zu bewerten. Zum Beispiel dürfte es ein schwieriges Problem sein, zu entscheiden, mit welchem Hamiltonoperator aus einer gegebenen Menge von möglichen Operatoren eine bestimmte Transformation am schnellsten implementierbar ist. Aufgrund der Ähnlichkeit dieser Probleme mit den oben beschriebenen QMA- und QCMA-Problemen, ist es angebracht, sich auf infinitesimale Zeitschritte in unserer Theorie der gegenseitigen Simulation von Hamiltonoperatoren zu beschränken.

Kapitel 8 Anwendung der Simulation von Hamiltonoperatoren im adiabatischen Quantencomputing

Wir untersuchen hier, ob die Zeitentwicklungen, die durch die Hamiltonoperatoren aus dem vorhergehenden Kapitel gegeben sind, ausgenutzt werden können, um NP-Probleme zu lösen. Unser Vorgehen basiert auf dem adiabatischen Quantencomputing. Dieses stellt einen Vorschlag für eine Klasse von kontinuierlichen Quantenalgorithmen dar, mit denen Grundzustände (d.h. Zustände minimaler Energie) eines gewünschten Hamiltonoperators bestimmt werden können. Dazu wird der Hamiltonoperator des Systems derart variiert, dass er langsam von einem besonders einfachen zu dem gewünschten Hamiltonoperator übergeht, dessen Grundzustände die Lösungen eines schwierigen Berechnungsproblems codieren. Das Resultat dieses Kapitels ist, dass es möglich ist, physikalisch realistische Hamiltonoperatoren zu konstruieren, deren Grundzustände die Lösungen des Problems “max independent set” codieren. Unsere Konstruktion basiert auf der Tatsache, dass das Problem “max independent set” auch für kubische planare Graphen NP-vollständig ist und dass jeder solche Graph effizient in ein zweidimensionales Gitter eingebettet werden kann. Aufgrund seiner speziellen Kopplungsstruktur kann der Hamiltonoperator effizient mit Hilfe unserer Algorithmen simuliert werden.

Introduction

The field of quantum computation and quantum information theory has been attracting a lot of attention from an interdisciplinary group of people like computer scientists, mathematicians, and physicists. This cross-disciplinary research is motivated by the ambitious aim to build a quantum computer. A quantum computer is a computational device that harnesses physical phenomena unique to quantum mechanics to realize a fundamentally new mode of information processing. The aim is to solve computational problems more efficiently than any classical computer can do.

The experimentalist's efforts consist in devising techniques allowing control of quantum systems in such a way that they could serve as hardware for future quantum computers. From a computer scientist's point of view the challenges are different. The aims of quantum computer science are analog to the aims of theoretical computer science. These consist mainly in classifying the complexity of computational problems and, if possible, to give efficient algorithms for solving them. The definition of computational complexity can only be formulated in the context of a computational model. A computational model specifies the set of elementary operations available in a computer. An algorithm describes a sequence of such elementary operations needed to transform the input (the specification of the problem) to the desired output (the solution of the problem). The time complexity of a problem is the minimal number of elementary operations required to solve it. The most important research topics of quantum computer science are

1. development of abstract computational models that determine the elementary transformations of a quantum computer (this determination should be based on physical laws governing the considered quantum system),

2. classification of computational problems with respect to their complexity within these models, and

3. development of efficient quantum algorithms to handle these various types of computational problems.

The development of *computational models* is central because they determine the complexity of problems and the efficiency of algorithms. Therefore, the focus of

this thesis is on computational models. The most common model for the quantum computer is the quantum circuit model that may be considered as the quantum extension of the Boolean circuit model (a classical computation device). The elementary operations of the quantum circuit model are described by elementary gates operating on only one or two qubits. The desired transformations are implemented by applying a sequence of elementary gates. The complexity of a transformation corresponds to the number of elementary gates.

In contrast, we work in this thesis with a computational model that better respects the properties of real quantum systems. Our approach builds upon a control-theoretic model of quantum computing. In this model the desired transformations are implemented by performing (1) an external control operation and (2) letting the system evolve for a certain time, and then repeating both steps several times. The natural time evolution of the system is characterized by its Hamiltonian. More precisely, the Hamiltonian defines the linear differential equation describing the time evolution (the so-called Schrödinger equation). For these reasons, the control-theoretic model is closer to physics than the quantum circuit model because it takes into account that quantum time evolutions are continuous processes described by Hamiltonians. The complexity of a transformation in the control-theoretic model is the required time and the number of control operations. These considerations lead to an extended definition of quantum algorithms and another view on complexity. In this context, we also consider continuous time algorithms (in contrast to discrete algorithms based on the quantum circuit model).

Not only can the control-theoretic model be used to implement transformations solving classical computation problems, but also transformations representing the time evolution of another quantum system. Considering the cost of simulating the time evolution of a quantum system by a given one defines in a certain sense the comparison of the "computational power" of different systems. The determination of the cost, in general, is a computationally hard problem. It can be solved to a large extent if we consider infinitesimal time intervals only. The control operations are chosen such that the simulated time evolution is reproduced with an error in the order of the square of the length of the considered time interval. This approach is based on the so-called Average Hamiltonian Theory that plays an important role in Nuclear Magnetic Resonance Spectroscopy.

We defined formally the mutual simulation of time evolutions for *infinitesimal* time intervals and called it mutual simulation of *Hamiltonians*. The main result of this thesis is a theory investigating the computational resources required for mutual simulation of Hamiltonians. We give lower bounds on the simulation complexity. An important tool for obtaining these bounds is graphs which are used to characterize Hamiltonians; the graphs represent the coupling topology of the Hamiltonians, that is, which pairs of subsystems interact. Our bounds show that the simulation complexity depends strongly on the properties of these graphs.

We develop efficient simulation algorithms based on methods of graph theory, representation theory of finite groups, design theory, and convex optimization theory. Optimality of the algorithms follows in many relevant cases from our lower bounds. Furthermore, we show how the elementary quantum gates of the quantum circuit model can be implemented within the control-theoretic model. It follows that quantum gates operating on disjoint qubits can be executed in parallel. This shows an interesting connection between the quantum circuit model and the control-theoretic model.

With this foundation of the elementary quantum gates within the control-theoretic model we account for the necessity to develop computational models based on physical laws. This is in line with Deutsch's thoughts in his book "Fabric of Reality" [Deu97]. He criticizes that the theory of computation has been studied traditionally as a subject of pure mathematics and asks for a complexity theory that is also based on physics. After all computational models are only idealizations of existing physical systems. From a fundamental point of view it is the laws of physics that determine which computational problems can be solved efficiently by computers.

All current computational models (like Turing machines, families of boolean circuits and cellular automata) are based on classical physics. But the fundamental theory is quantum mechanics. It is quantum theory that determines computational models and complexity. Besides these theoretical reasons, there are also practical reasons for developing a theory of information processing in terms of quantum mechanics. Such a theory will be important for classical computing and emerging nanotechnologies. We must take into consideration that with increasing integration density in future computers every computational process will have to be described more and more as a microscopic process involving quantum effects. Therefore, this refined description will become necessary as classical computers are miniaturized in such a way that the logical states of bits are represented by simple quantum mechanical degrees of freedom. For example, information can be encoded in the state of an atom or of a photon. (This is already the case in experimental implementations of quantum computing with a few qubits and quantum cryptography.)

The motivation for our theory of mutual simulation of Hamiltonians goes back to Feynman's proposal for simulating quantum systems. He observed that all attempts to simulate dynamics (that is, time evolution) of general many-particle quantum systems by any classical computer involve an exponential slowdown in running time (due to the exponentially increasing dimension of the system's Hilbert-space with the number of components). To avoid this time overhead, Feynman proposed the use of quantum systems instead of classical systems to simulate quantum dynamics. Feynman already conjectured the existence of a universal quantum simulator. A universal quantum simulator is a controlled device operating at the quantum level, which allows the efficient simulation of dynamics

of different many-particle quantum system. Whereas the first works on this topic gave simulations based on the quantum circuit model only, we adopt a more fundamental approach. As mentioned above, we propose a model where the desired Hamiltonian time evolutions are simulated by interspersing the Hamiltonian time evolution inherent to the quantum system by external control.

Our methods of mutual simulation of Hamiltonians offer an interesting application in adiabatic quantum computing. In this approach for solving computationally hard problems the system is first initialized in the energy minimum of a simple Hamiltonian. The Hamiltonian is varied slowly and corresponds at the end of the algorithm to a Hamiltonian whose energy minimum encodes the solution of the consider problem. During the computation the system is in a state of minimal energy with respect to the actually present Hamiltonian. Our simulation techniques allow an efficient realization in the control-theoretic model. It is important that only *physically realistic* Hamiltonians (that is, with a local coupling topology) are used as computational resources. We construct such Hamiltonians whose energy minima encode the solutions of the NP-complete problem "max independent set".

In summary, the main results of this thesis include:

- the formal definition of mutual simulation of Hamiltonians,
- the derivation of lower and upper bounds on the complexity of mutual simulation of Hamiltonians,
- the development of optimal simulation algorithms using methods of representation theory of finite groups (irreducible projective representations), design theory (orthogonal arrays, difference schemes, Hadamard matrices, and spreads) and algebraic graph theory, and
- application of mutual simulation of Hamiltonians to adiabatic quantum computing for solving computationally hard problems and the description of new computational problems and their quantum complexity classes.

Structure of the thesis

Chapter 1 Quantum computing and quantum circuit model

Here we give a short introduction to the main ideas of quantum computing and quantum information theory. The most common model of the quantum computer, namely the quantum circuit model, is introduced. It is a discrete model that is in close analogy to the classical model of boolean circuits. Based on this model we introduce the quantum complexity class BQP. This class consists of all problems that can be solved efficiently (that is, with polynomially many elementary gates of the quantum circuit model) on a quantum computer.

Chapter 2 Control-theoretic model and simulation of Hamiltonians

This chapter introduces a control-theoretic model of the quantum computer that better respects the physical nature of processes than the quantum circuit model. The time evolution in this model is determined by the Hamiltonian inherent to the system and by external control operations. Especially, this model takes into account that all physical evolutions are continuous processes.

Based on the control-theoretic description we define the mathematical model of mutual simulation of Hamiltonians. The complexity of simulating Hamiltonians by a given Hamiltonian is measured by the time overhead and the number of time steps. The time overhead gives the ratio of the running time of the desired Hamiltonian and the running time of the simulating Hamiltonian that is necessary to achieve the simulation. The number of time steps is the number of control operations that are necessary to simulate the desired Hamiltonians.

Chapter 3 Decoupling and time-reversal algorithms

We consider in this chapter decoupling and time-reversal (refocusing). Decoupling is a special case of a simulation algorithm, the desired time evolution is the trivial time evolution (the standstill of the quantum system). Time reversal makes the system behave as if its time evolution moved backwards. Both tasks play an important role in Nuclear Magnetic Resonance Spectroscopy.

We generalize known decoupling schemes for coupled qubits to coupled qudits, that is, quantum systems consisting of several subsystems of arbitrary dimension. This generalization is based on the concept of an *annihilator* describing the minimal set of control operations that is sufficient for switching off general time-evolutions. Furthermore, we show how to efficiently obtain algorithms for time-reversal from decoupling algorithms.

Our algorithms can be applied to a broad class of Hamiltonians (even if the exact structure of the Hamiltonians is unknown), whereas earlier results in Nuclear Magnetic Resonance Spectroscopy were designed for special Hamiltonians only. Our constructions make use of methods of representation theory of finite groups (irreducible projective representations) and design theory (orthogonal arrays, dif-

ference schemes, Hadamard matrices, and spreads). The decoupling schemes provide an important tool in establishing the universality of simulating Hamiltonians in the next chapter.

Chapter 4 Universal simulation of Hamiltonians

We study the requirements on the set of external control operations that allows universal simulation of Hamiltonians. This leads to the notion of transformer groups. We characterize the finite transformer groups with the help of a necessary and sufficient condition that is formulated in terms of group characters. Using this condition we found different transformer groups by performing a search over all groups of size up to 255 with computer algebra systems.

Chapter 5 Bounds on complexity of simulating Hamiltonians

We focus on the complexity of simulating Hamiltonians by a given Hamiltonian. Both time overhead and number of time steps are used as complexity measures. Based on the fact that Hamiltonians occurring in nature are usually pair-interaction Hamiltonians, we use a convenient notation (called coupling matrices) that permits a description of the Hamiltonians and the control operations by methods of algebraic graph theory. We show that the simulation complexity depends strongly on the invariants of the graphs that describe the coupling topology of the Hamiltonians. These invariants are for example chromatic number, clique coloring index, chromatic index and eigenvalues of the adjacency matrices. Based on these methods, we derive lower and upper bounds on the simulation complexity for general Hamiltonians. Surprisingly, this method also leads to the derivation of new bounds on graph invariants.

Chapter 6 Special simulation tasks and their complexity

We consider special simulation tasks and analyze their complexity by applying the complexity bounds described in Chapter 5. First, we prove that our algorithms for decoupling and time-reversal presented in Chapter 3 are optimal. Then we apply the complexity bounds to some concrete physical systems that could be relevant for the realization of future quantum computers. For instance, we consider special types of Ising-interactions that are suitable for implementing the so-called "one-way quantum computer" that is a proposal for a scalable hardware architecture. Furthermore, these interactions can be used to realize graph codes that protect quantum information against errors. The complexity bounds follow directly from the graph spectra for both systems.

An important simulation task is the implementation of the elementary gates of the quantum circuit model within the control-theoretic model. Therefore, we give simulation techniques allowing to simulate quantum gates using Hamiltonians. Especially, we show that quantum gates operating on disjoint qubits can be executed in parallel. Based on this observation we propose a physically motivated

complexity measure for quantum circuits. This complexity measure establishes an interesting connection between the quantum circuit model and the control-theoretic model. The computational power of the quantum circuit model and the control-theoretic model are compared in detail. We show for special interactions that the control-theoretic model is more powerful because it allows a higher degree of parallelization.

Chapter 7 Quantum complexity classes

In this chapter we address some questions of quantum complexity theory. There are two reasons for doing this. On the one hand, the computational complexity of determining energy minima of Hamiltonians is the basis for understanding adiabatic quantum computing that is treated in Chapter 8. On the other hand, we obtain an impression of how difficult it is to decide if an implementation of a transformation in the control-theoretic model is optimal. This provides a justification for restricting the problem to infinitesimal time intervals in our approach.

The starting point for our considerations is the fact that the solutions of the NP-complete problems "max cut" and "max independent set" may be encoded in the ground states of simple pair-interaction Hamiltonians. Then we study the quantum complexity classes *QMA* and *QCMA* that are two natural extensions of NP to the quantum model and describe the "local Hamiltonian problem". This problem was the only problem known to be QMA-complete; in contrast, no QCMA-complete problems were known.

Our main contribution in this area is the discovery of new problems which we prove to be complete for both classes. We show that the problem of "low-energy and low-complexity states of local Hamiltonians" is QCMA-complete. This is an extension of the "local Hamiltonian" problem. Both problems are related to the questions "does a given Hamiltonians have states with energy lower than a given bound?" and "if yes, can the low-energy states be prepared efficiently?", respectively. The Hamiltonians occurring in these problems are more complicated than in the case of NP. Furthermore, we show that the problems "identity check" and "identity check on basis states" are complete for QMA and QCMA, respectively. These problems consist in checking whether a unitary operator specified by a quantum circuit acts as the identity operator on the whole space or on the basis states, respectively. An important problem that can be reduced to identity check is "equivalence check". It consists in deciding whether two quantum circuits implement the same unitary operator. The fact that "equivalence check" is complete for QMA and QCMA indicates that it is difficult to decide if two autonomous time evolutions according to different Hamiltonians implement the same unitary operator after certain times. In the same way, it suggests that it is difficult to decide if a unitary operator specified by a quantum circuit can be implemented by an autonomous Hamiltonian time evolution according to a given Hamiltonian.

Therefore, it seems that it is difficult to evaluate the computational power of

Hamiltonians within the control-theoretic model. For example, the problem of deciding which Hamiltonian in a given set of operators allows to implement within the shortest time a desired transformation should be very hard. Because of the resemblance to the questions on QMA and QCMA it seems to be appropriate to restrict oneself to infinitesimal time intervals in our theory of mutual simulation of Hamiltonians.

Chapter 8 Application of Hamiltonian simulation within adiabatic quantum computing

We study in this chapter whether the time evolutions according to the Hamiltonians of Chapter 7 could be used to solve the computational problems. Our approach is based on adiabatic quantum computing. Adiabatic quantum computing is a proposal for a class of continuous time algorithms that find ground states (that is, states having minimal energy) of a desired Hamiltonian. This is done by slowly varying the Hamiltonian from a simple Hamiltonian to the desired Hamiltonian whose ground states encode the solution to a computationally hard problem. The result of this chapter is that it is possible to construct *physically realistic* Hamiltonians whose ground states encode the solution to problem "max independent set". Our construction relies on the facts that "max independent set" problem remains NP-complete for planar cubic graphs and that any such graph can be embedded efficiently in a two dimensional rectangular lattice. Due to its special structure the Hamiltonian can be simulated efficiently using our simulation algorithms.

Chapter 1

Quantum computing and quantum circuit model

The aim of this chapter is to give a general overview of quantum computing and quantum information theory. We present the *quantum circuit model* that is the most common abstract model. It allows to construct a general theory of quantum computation and quantum complexity that does not depend upon a specific physical system for its realization.

1.1 Quantum circuit model

In any computer, we need to encode the problem we want to solve (input), to extract the answer (output), and to manipulate the state of the computer to transform the input into the desired output (computation). In the following we discuss how these tasks are done within the quantum circuit model. We give a brief introduction of the quantum circuit model. The reader is referred to the books Nielsen and Chuang [NC00] and Gruska [Gru99] for more details.

The quantum circuit model is a computational model that is based on quantum mechanics. Historically, Planck's proposal to associate discrete units of energy with black radiation body radiation at the end of the 19th century, followed by Einstein's explanation of the photo-electric effect by assuming a corpuscular nature of light and Bohr's atom model with discrete energy levels, led to the birth of quantum mechanics. The first formulations of quantum mechanics are Heisenberg-Born-Jordan's Matrizenmechanik (matrix mechanics) and Schrödinger's Wellenmechanik (wave mechanics) (cf. von Neumann [vNJ32]). Dirac (cf. Dirac [Dir58]) and Jordan (cf. Born and Jordan [BJ30]) unified these both equivalent theories to a theory called transformation theory. This formulation allows an especially simple treatment of physical questions (cf. von Neumann [vNJ32]) and is based on the mathematical notion of Hilbert spaces.

Quantum systems	Qubits	References
Ion traps	Energy levels of the ions	[CZ95, MMK+95, NLR+99]
Nuclear spin resonance	Magnetic moment	[HSTC00, KCL98, MFM+00]
QED with optical cavities	Photon states	[DRBH95, THL+95]
Josephson junctions	Cooper pairs	[MSS00]
Quantum dots	Spin state	[LD98, ZR98]
Solid state NMR	Magnetic moment	[Kan98]

Figure 1.1: Quantum systems proposed for quantum computing

While the origins of quantum mechanics were about 100 years ago, the idea of information processing with the help of quantum mechanical systems was stated not until the early eighties. Feynman proposed to use a quantum computer to simulate quantum dynamics (cf. Feynman [Fey82]).[1] Nowadays, the most common model for quantum computation is the quantum circuit model that was introduced by Yao (cf. Yao [Yao93]) as a generalization of boolean circuits. It extends the general model of Hilbert spaces by specifying the elementary operations of a quantum computer. This allows to develop quantum complexity theory.

1.1.1 Qubits

The *bit* is the fundamental concept of classical computation and classical information. At the heart of quantum computation and quantum information lies the *quantum bit*, or *qubit* for short, as the natural extension of a bit.

We describe qubits as mathematical objects with certain properties. The reason for treating qubits as abstract entities is that it allows us to construct a general theory of quantum computation and quantum information that does not depend upon a specific physical system for its realization. Similarly, when we speak of a bit we do not have to think each time whether it is realized physically as a magnetization state on a hard disk or an electric charge in a memory cell. In the literature, several physical systems are discussed as possible realizations of the standard model of the quantum computer. They are summarized in Table 1.1.

Just as a classical bit has a *state* – either 0 or 1 – a qubit also has a state. Two possible states for a qubit are the states $|0\rangle$ and $|1\rangle$ that correspond to the states 0 and 1 for a classical bit. The standard notation used in quantum computing is *Dirac's bra-ket* notation. The difference between bits and qubits is that a qubit can be in a state *other* than $|0\rangle$ or $|1\rangle$. It is also possible to form *linear combinations* of states, called *superpositions*

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle\,. \tag{1.1}$$

[1] This issue will be discussed in Chapter 2.

The numbers α and β are complex numbers. Mathematically, the state of a qubit is a unit vector in the two-dimensional complex Hilbert space $\mathbb{C}^2$. The special states $|0\rangle$ and $|1\rangle$ are known as *computational basis states*, and form an orthonormal basis for this vector space:

$$|0\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |1\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{1.2}$$

In Dirac's bra-ket notation a column vector (representing the state of a qubit) is called a *ket*

$$|\Psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{1.3}$$

The corresponding *bra* $\langle\Psi|$ is the row vector $(\bar{\alpha}, \bar{\beta})$, where $\bar{\alpha}$ and $\bar{\beta}$ are the complex conjugates of α and β, respectively. The inner product of the kets $|\Phi\rangle$ and $|\Psi\rangle$ is given by the *bra-ket* $\langle\Phi|\Psi\rangle$.

Let us now turn to the problem of reading out the state of a qubit. Classically, we can determine whether a bit is in the state 0 or 1. For example, this happens in computers all the time when the contents of memory cells are read out. Rather remarkably, we cannot examine a qubit to determine its state, that is the values of α and β. Instead, quantum mechanics tells us that we can only acquire much more restricted information about the quantum state. When we measure a qubit we get either the result 0, with probability $|\alpha|^2$, or the result 1, with probability $|\beta|^2$. Naturally, we must have $|\alpha|^2 + |\beta|^2 = 1$ in eq. (1.1), since the probabilities must sum to one. Geometrically, we can interpret this as the condition that the qubit's state is normalized to one. The numbers α and β are called *probability amplitudes* since they determine the probabilities of the measurement outcomes. Besides the probabilistic nature of the measurement process quantum mechanics tells us that measuring a qubit alters its state. The state of a qubit after the measurement (post-measurement state), is $|0\rangle$ or $|1\rangle$ (depending on the outcome), and not $\alpha|0\rangle + \beta|1\rangle$. This means that although a qubit can be prepared in an infinite number of different quantum states by choosing different probability amplitudes α and β in eq. (1.1) it cannot be used to store more than one (classical) bit of information. As we will explain later this is because no measurement process can reliably differentiate between nonorthogonal states. However, we will see that the superposed states are essential for quantum computing.

Multiple qubits and quantum registers

Having introduced the qubit – the fundamental concept of quantum information – we turn to consider composite quantum systems.

One of the fundamental principles of quantum mechanics is that the joint quantum state space of two systems is the tensor product of their individual quantum

state spaces. Thus, the quantum state space of n qubits is the Hilbert space

$$\mathcal{H} := \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 \,. \tag{1.4}$$

We will call such collection of qubits a *quantum register*, or simply a *register*. As in the classical case, it can be used to store more information. For instance, the binary representation of 6 is 110 and encoding this value in a quantum register is done by preparing the three qubits in the state $|1\rangle \otimes |1\rangle \otimes |0\rangle$. The *computational basis states* of $\mathcal{H}$ are parameterized by binary strings of length n. The computational basis state corresponding to the binary string $i_n \cdots i_2 i_1$ is given by the tensor product

$$|i_n\rangle \otimes \cdots \otimes |i_2\rangle \otimes |i_1\rangle$$

and denotes a register prepared with the value $i = 2^{n-1} i_n + \ldots + 2^1 i_2 + 2^0 i_1$. Two states $|i\rangle$ and $|i'\rangle$ are orthogonal as soon as $i \neq i'$:

$$\langle i'|i\rangle = \langle i'_1|i_1\rangle \cdots \langle i'_n|i_n\rangle \tag{1.5}$$

and if $i' \neq i$ then at least one of the inner products of the r.h.s. of the above expression is zero so that $\langle i'|i\rangle = 0$.

The most general state in $\mathcal{H}$ can be written as a ket

$$|\Psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle \tag{1.6}$$

satisfying $\sum_{i=0}^{2^n-1} |\alpha_i|^2 = 1$. The coefficients α_i are called probability amplitudes of the states $|i\rangle$, and we say that $|\Psi\rangle$ is a *superposition* of $|i\rangle$. When we measure each qubit of $\mathcal{H}$ we get i with probability $|\alpha_i|^2$.

Note that the state in eq. (1.6) describes the situation in which several different values of the register are present *simultaneously*; just as in the case of a single qubit, there is no classical counterpart to this situation. Furthermore, in a composite quantum system we may have so-called *entangled* states. *Entanglement* is one of the most interesting and puzzling ideas associated with composite quantum systems. Consider the Bell state

$$|\Psi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \,. \tag{1.7}$$

This state has the remarkable property that there are no single qubit states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ such that $|\Psi\rangle = |\Psi_1\rangle \otimes |\Psi_2\rangle$. We say that a state of a composite system that cannot be written as a product of states of its component systems is an *entangled* state. States that are not entangled are called *product states*. Entangled states play a crucial role in quantum computing and quantum information theory.

As we have seen quantum computation takes place in the composite quantum system $\mathcal{H}$, and we obtain extra computational power from the exponential size of

the state space and the quantum phenomena like superposition and entanglement. However, there is no direct way to use the exponential dimension to speed up computations. This is due to the probabilistic nature of measurement in quantum mechanics.

Measurement

A fundamental postulate of quantum mechanics is that a measurement is described by an orthogonal decomposition of $\mathcal{H}$. Let $\mathcal{P} = (P_1, P_2, \ldots, P_m)$ be an orthogonal decomposition of $\mathcal{H}$, that is, $\mathcal{P}$ is a collection of projections such that

$$\sum_{i=1}^{m} P_i = \mathbf{1}\,,$$

where $\mathbf{1}$ denotes the identity matrix. It follows from this condition that the projections P_i are mutually orthogonal. Therefore, any vector $|\Psi\rangle$ in $\mathcal{H}$ can be expressed as a linear superposition of its components along each subspace given by the image of P_i:

$$|\Psi\rangle = \mathbf{1}|\Psi\rangle = \sum_{i=1}^{m} P_i|\Psi_i\rangle = \sum_{i=1}^{m} \beta_i|\Phi_i\rangle\,, \tag{1.8}$$

where $\beta_i := \langle\Psi|P_i|\Psi\rangle$ and

$$|\Phi_i\rangle := \frac{P_i|\Psi\rangle}{\langle\Psi|P_i|\Psi\rangle}$$

is a unit vector in the image of P_i for $i = 1, \ldots, m$. Measuring the state $|\Psi\rangle$ with respect to $\mathcal{P}$ will cause the following:

1. The measurement outcome i is obtained with probability $|\beta_i|^2$.
2. The state $|\Psi\rangle$ changes ("collapses") to $|\Phi_i\rangle$ if the measurement outcome is i.
3. The only classical information given by the measurement is which i was obtained. All information about the state $|\Psi\rangle$ that is not contained in the image of P_i is lost.

We see that there is no way to gain complete knowledge of the state of a register through a single measurement.

In the quantum circuit model the elementary read out consists of measuring a subset of the qubits with respect to $\{|0\rangle\langle 0|, |1\rangle\langle 1|\}$. This measurement is considered as elementary since it can (often) be realized directly in physical systems. More general measurements have be accomplished by first carrying out some computation process and then performing the above elementary measurement.

1.1.2 Quantum gates

We describe how changes occur to states of a quantum computer. In order to motivate the rules for state manipulation within the quantum circuit model, we briefly describe the classical circuit model. In a classical computer, the processing of information is done by logic gates. The output of a gate is the value of a boolean function evaluated on its input. For example, the simplest non-trivial classical gate is the NOT gate. Its output is given by flipping the input bit, i.e., 0 becomes 1 and vice-versa. Recall that any (classical) logic function can be implement by a circuit using only the gates AND, OR, and NOT. These three gates are thus said to form a *universal* set of gates for (classical) computing.

Classical gates usually correspond to physical devices (e.g. a collection of transistors in a processor). Quantum gates are fundamentally different. They do not correspond to physical devices, but to processes acting on quantum registers. A quantum gate may be realized by switching on a laser for a certain period of time.

Besides AND, OR, and NOT gates, there are also elements that copy bits (fan-out) in classical circuits. It is arguable that these elements should also be considered as gates. The essential difference between quantum and classical information processing is that input and output cannot exist simultaneously in the quantum domain. The reason is the so-called no-cloning theorem. It states that quantum information cannot be copied.

Single qubit gates

What does an analogous quantum NOT look like? Imagine that we had some quantum process that interchanged the states $|0\rangle$ and $|1\rangle$. Such a process would obviously be a good candidate for a quantum analog to the NOT gate. However, specifying the action of the gate on the states $|0\rangle$ and $|1\rangle$ does not tell us what happens to superpositions of the states $|0\rangle$ and $|1\rangle$, without further knowledge of quantum gates. In fact, the quantum NOT gate acts *linearly*, that is, it takes the superposition $\alpha|0\rangle + \beta|1\rangle$ to the state $\alpha|1\rangle + \beta|0\rangle$. There is a convenient way of representing the quantum NOT gate in matrix form following from the linearity of quantum gates. It is described by a complex 2×2 matrix

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{C}^{2\times 2} \tag{1.9}$$

with respect to the basis $\{|0\rangle, |1\rangle\}$. In general, quantum gates on single qubits are described by complex 2×2 matrices. Are there any constraints on what matrices may be used as quantum gates? It turns out that there are. Recall that the normalization condition requires $|\alpha|^2 + |\beta|^2 = 1$ for a general quantum state $|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle$. This must also be true for the quantum state $|\Psi'\rangle = \alpha'|0\rangle + \beta'|1\rangle$ after the gate has acted. It turns out that the appropriate condition on the matrix

U representing the gate is that U must be unitary. Recall that a matrix U is unitary if $U^\dagger U = \mathbf{1}$, where $U^\dagger$ is the *adjoint* of U (obtained by transposing and then complex conjugating U), and $\mathbf{1}$ is the 2×2 identity matrix. For the NOT gate it is easy to see that $X^\dagger X = \mathbf{1}$.

One of the fundamental postulates of quantum mechanics is that this *unitarity* constraint is the *only* constraint on quantum gates. Any unitary matrix specifies a valid quantum gate. The interesting implication is that in contrast to the classical case there are many non-trivial single qubit gates. Therefore, we have gates that have no classical counterpart. For instance, the *Hadamard transformation* H is the unitary transformation

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} . \tag{1.10}$$

Note that it evolves "classical" states $|0\rangle$ and $|1\rangle$ into superpositions

$$\begin{aligned} H|0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) , \\ H|1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) . \end{aligned} \tag{1.11}$$

Therefore it cannot be regarded as classical. This gate is of great utility: take an k-bit quantum register initially in the state $|0 \ldots 00\rangle$ and apply to every single qubit of the register the gate H. The resulting state is

$$\begin{aligned} |\Psi\rangle &= H \otimes H \otimes \cdots \otimes H \, |0 \ldots 00\rangle \\ &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \cdots \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ &= \frac{1}{\sqrt{2^k}} \sum_{i=0}^{2^k-1} |i\rangle . \end{aligned} \tag{1.12}$$

Thus, with a linear number of operations we have created a register state that contains an exponential number of distinct terms. It is quite remarkable that using quantum registers, k elementary operations can create a state containing all 2^k possible values of the register. In contrast, in classical registers k elementary operations can only prepare one state of the register representing one specific number. It is this ability of creating quantum superpositions that makes "quantum parallelism" possible. If after preparing the register in a superposition of several numbers all subsequent computational operations are unitary then with each computational step the computation is performed simultaneously on all the numbers present in the superposition. However, once again we must emphasize that this quantum parallelism cannot be used directly to speed up computations because of the probabilistic nature of the measurement process. It is not possible to obtain all information about the state.

Multiple qubit gates

Let us now turn to the issue of universality. It is clear that starting from the state $|00\ldots0\rangle$ and using only single qubit gates, i.e., gates of the form

$$\mathbf{1}\otimes\cdots\otimes\mathbf{1}\otimes U\otimes\mathbf{1}\otimes\cdots\otimes\mathbf{1}\,,$$

we can only create product states. Therefore, it is necessary to have (at least) two-qubit operations in order to be able to create entangled states. Of all possible unitary transformations acting on a pair of qubits, an interesting subset is the one that contains gates that can be written as

$$|0\rangle\langle 0|\otimes\mathbf{1}+|1\rangle\langle 1|\otimes U\,, \tag{1.13}$$

where $\mathbf{1}$ is the single-qubit identity gate, U is some other single-qubit gate, and $|i\rangle\langle i|$ is the orthogonal projection onto the subspace spanned by $|i\rangle$ $(i=0,1)$. Such a two-qubit gate is called a "controlled gate", since which of the transformations $\mathbf{1}$ and U is performed on the second qubit is controlled by whether the first qubit is in the state $|0\rangle$ or $|1\rangle$. For example, the effect of controlled-NOT (CNOT) on the basis states is

$$|00\rangle\mapsto|00\rangle\,;\quad |01\rangle\mapsto|01\rangle\,;\quad |10\rangle\mapsto|01\rangle\,;\quad |11\rangle\mapsto|10\rangle\,. \tag{1.14}$$

The matrix representation of the CNOT is given by

$$U_{CN}=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{pmatrix}\in\mathbb{C}^{4\times4}\,. \tag{1.15}$$

Another way of describing the action of the CNOT is as a generalization of the classical XOR gate, since the action of this gate may be summarized as the linear extension of the map $|x\rangle|y\rangle\mapsto|x\rangle|x\oplus y\rangle$, where $\oplus$ is the addition modulo two.

As for the single qubit case, the requirement that probability is conserved is expressed by the fact that U_{CN} is a *unitary matrix*, that is $U_{CN}^{\dagger}U_{CN}=\mathbf{1}$, where $\mathbf{1}$ denotes here the 4×4 identity matrix. Two-qubit gates are described by 4×4 unitary matrices.

More generally, one of the fundamental postulates of quantum mechanics is that all reversible transformations of states are described by unitary matrices. In particular, the transformations on $\mathcal{H}$ correspond to elements of $\mathcal{U}(2^n)$, the group of unitary matrices of size $2^n\times2^n$. However, as in the classical case not every quantum transformation is available directly but has to be realized as a sequence of elementary transformations; these are specified by the underlying computational model. Quantum circuits describe how transformations are built of elementary gates.

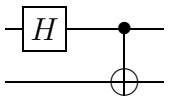


Figure 1.2: Quantum circuit preparing a Bell state

1.1.3 Quantum circuits

A quantum circuit describes a quantum process that is realized as a sequence of elementary processes. It consists of gates and quantum wires. Quantum wires denote qubits in a quantum register. A qubit may correspond to a physical particle such as an ion inside an ion trap or a photon moving from one location to another through space. Whereas a classical gate corresponds to a physical device (e.g. transistors realizing some logic transformation), a quantum gate describes an elementary physical process (quantum dynamic) implemented by some external control. This issue will be discussed in more detail in Section 2. We assume that the state input to the circuit is the computational basis state $|0\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle$.

The first example of a simple quantum circuit is shown in Figure 1.2. It describes a unitary transformation on $\mathbb{C}^2 \otimes \mathbb{C}^2$ since there are two wires. The upper wire and the lower wire correspond to the first and the second qubit, respectively. The state of the system is $|0\rangle \otimes |0\rangle$ is at the beginning. The first gate corresponds to the Hadamard gate on the first qubit. It describes the unitary transformation $H \otimes \mathbf{1}_2$, where $\mathbf{1}_2$ is the two by two identity matrix. The second gate corresponds to the controlled NOT gate, where the first qubit is the control and the second qubit is the target. It corresponds to the unitary transformation U_{CN} defined in eq. (1.15). The dot means that the NOT gate (denoted by $\oplus$) is carried out on the second qubit if the first qubit is in the state $|1\rangle$. The state of the system changes as follows:

$$|0\rangle \otimes |0\rangle \mapsto \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) \otimes |0\rangle \mapsto \frac{1}{\sqrt{2}}\big(|0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle\big)$$

We introduce notation that allows us to embed a unitary transformation on a single qubit or on two qubits in the joint Hilbert space of all qubits. Let U be an arbitrary single qubit transformation. If it is applied to the kth qubit, then the resulting transformation on $\mathcal{H}$ is given by the tensor product

$$U^{(k)} := \underbrace{\mathbf{1} \otimes \cdots \otimes \mathbf{1}}_{k-1} \otimes U \otimes \underbrace{\mathbf{1} \otimes \cdots \otimes \mathbf{1}}_{n-k} .$$

Consider a quantum register consisting of three qubits. The quantum circuit corresponding to $U^{(2)}$ is shown in Figure 1.3.

We need an analogous definition for 4×4 matrices. Let V be an arbitrary 4×4 matrix and $1 \leq k < l \leq n$. We denote by $V^{(k,l)}$ the transformation on $\mathcal{H}$ that acts as V on the subsystem consisting of the qubits k and l and as identity on

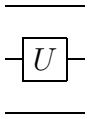


Figure 1.3: A single qubit gate acting on qubit 2

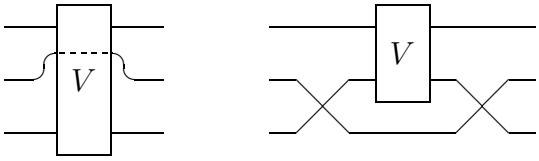


Figure 1.4: Representations of a two-qubit gate acting on qubits 1 and 3

the other qubits. For $l = k + 1$ this transformation may be written conveniently as the tensor product

$$V^{(k,k+1)} := \underbrace{\mathbf{1} \otimes \cdots \otimes \mathbf{1}}_{k-1} \otimes V \otimes \underbrace{\mathbf{1} \otimes \cdots \otimes \mathbf{1}}_{n-k-1} .$$

Consider again a quantum register consisting of three qubits. The quantum circuit corresponding to $V^{(1,3)}$ in shown in Figure 1.4. We denote by $\mathrm{CNOT}^{(k,l)}$ (with $1 \leq k < l \leq n$) the controlled-NOT operation with qubit k as control and qubit l as target. For a quantum register consisting of three qubits, the quantum circuit for $\mathrm{CNOT}^{(1,3)}$ is shown in Figure 1.5.

To define formally quantum circuits we need a notation that allows us to specify on which qubits the gates acts on. U will be either a 2×2 or 4×4 unitary matrix specifying a quantum gate. If U is a 2×2 matrix, then L will always be a list containing one number, that is $L := (k)$, and we define $U^L := U^{(k)}$. Analogously, if U is a 4×4 matrix, than L will always be an ordered list containing two numbers, that is, $L := (k, l)$, and we define $U^L := U^{(k,l)}$.

Definition 1.1 (Quantum circuit)
Let $\mathcal{G}$ be a set of gates. A quantum circuit $\mathcal{C}$ on n qubits of size N is given by a tuple $(U_1, L_1; U_2, L_2; \ldots; U_N, L_N)$, where $U_j \in \mathcal{G}$ and either $L_j := (k_j)$ or $L_j := (k_j, l_j)$ depending on whether U_j is a single qubit or a two qubit gate. We say that $\mathcal{C}$ realizes a unitary transformation $U \in \mathcal{U}(2^n)$ if

$$U = U_N^{L_N} \cdots U_2^{L_2} U_1^{L_1} .$$

As for classical circuits, there are universal sets of gates for quantum circuits; such a universal set of gates is sufficient to realize any quantum transformation. One particularly useful universal set of gates is the set of all single qubit gates and the controlled NOT gate, that is, any unitary operation on $\mathcal{H}$ can be generated by a sequence of controlled NOT and single qubit gates (cf. Barenco et

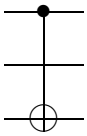


Figure 1.5: controlled NOT with qubit 1 as control and qubit 3 as target

AL. [BBC+95]). Therefore, there is no (mathematical) reason to consider more complicated gates (acting on more than two qubits) as elementary ones.

Once we have fixed a universal set of gates we can define complexity of quantum transformations.

Definition 1.2 (Circuit complexity)
Let $U \in \mathcal{U}(2^n)$ be a unitary transformation on a quantum register of size n. The circuit complexity of U with respect to the universal set of gates $\mathcal{G}$ is the minimal number N such that there is a quantum circuit of size N realizing U.

This definition of complexity is justified by the mathematical fact that single and two qubit gates are the simplest ones that are universal. More complicated gates should not be considered as elementary ones. In the next section we define a continuous complexity measure based on physically founded assumptions and relate it to the discrete circuit complexity. This gives further justification for the quantum circuit model.

It is often reasonable to assume that gates acting on disjoint qubits can be performed in parallel. This leads to a slightly modified definition of quantum circuits and complexity.

Definition 1.3 (Parallelized quantum circuits)
A parallelized quantum circuit on n qubits of depth N is a sequence of N steps S_j. Each step S_j is a collection of m_j gates $S_j := \{U_{j,1}, L_{j,1}; \ldots, U_{j,m_j}, L_{j,m_j}\}$ acting on disjoint qubits, that is $L_{j,m} \cap L_{j,m'} = \emptyset$ for all $1 \leq m < m' \leq m_j$. The circuit implements the unitary transformation

$$U = \prod_{s=1}^{N} \prod_{j=1}^{m_j} U_{s,j}^{L_{s,j}} \,. \tag{1.16}$$

The depth of a unitary transformation U is the minimal number N such that there is a parallelized quantum circuit implementing U.

Note that the depth of unitary transformation differs by at most a factor n form the circuit complexity. Figure 1.6 shows a quantum circuit of size 7 and the corresponding parallelized quantum circuit of depth 3.

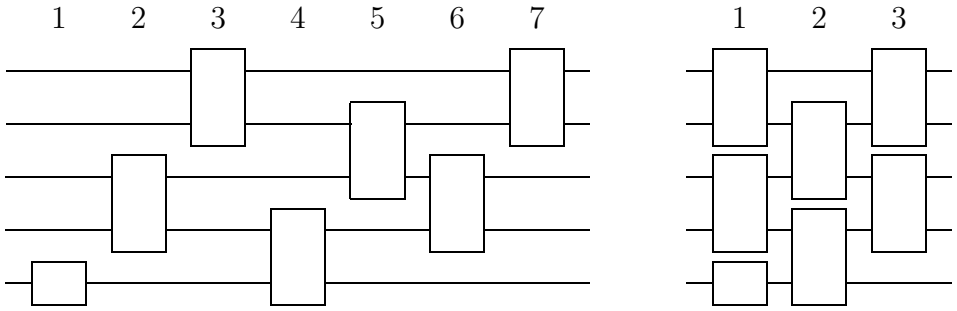


Figure 1.6: Discrete complexity measure of quantum circuits

1.1.4 Quantum complexity class BQP

The complexity of quantum algorithms is measured by the number of gates required to implement them. The major theoretical success for the field of quantum computing came when Peter Shor discovered an efficient factorization algorithm . This discovery has attracted a lot of attention because the security of public key cryptography based on RSA-encryption relies the computational difficulty of the factorization problem of large numbers. The best classical algorithms (number field sieve) have a running time that is subexponential in the size of the number we want to factorize.

We refer the reader to Nielsen and Chuang [NC00] for a complete analysis of Shor's algorithm. Let us just note that at the heart of Shor's algorithm lies the discrete Fourier transform . It is given by

$$DFT|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} \exp(2\pi i \frac{xy}{2^n})|y\rangle \,, \tag{1.17}$$

where n is the size of the register. The efficiency of Shor's algorithm relies on the fact the DFT of size 2^n can be realized on a quantum computer using only $O(n^2)$ gates, whereas fast classical algorithms require $O(n2^n)$ operations (cf. Rötteler [Röt01]).

Let us define efficiency more formally. We define BQP to be the class of all computational problems that can be solved efficiently on a quantum computer, where a bounded probability of error is allowed. Here efficiency means that the problem can be solved with bounded probability of error using a *polynomial size* quantum circuit. More formally, we say a language L is in BQP if there is a family of polynomial size quantum circuits that decides the language, accepting strings in the language with probability at least 2/3, and rejecting strings that are not

in the language with probability at least 2/3. In practice, what it means is that the quantum circuit takes as input binary strings, and tries to determine whether they are elements of the language or not. At the end of the quantum circuit one qubit is measured, with 1 indicating that the circuit has been accepted, and 0 indicating that the string has been rejected. By testing the string a few times to determine whether it is in L, we can determine with very high probability whether a given string is in L. Note that BQP is the quantum analog of BPP; the corresponding circuits consist of probabilistic classical gates [Pap94].

The problem of integer factorization is in BQP because of Shor's algorithm (to be consistent the above definition of BQP as a class of decision problems we note that Shor's algorithms allows to solve the the decision problem corresponding to integer factorization. It is as follows: given positive integers N and k, decide if N has a factor satisfying $2 \leq M \leq k$ [Kob99]).

Based on the quantum circuit model we introduce in Chapter 7 two quantum complexity classes QMA and QCMA. They contain BQP and may be considered as two natural extensions of the classical complexity class NP to the quantum model. We establish new problems that are complete for these classes. Some of these problems are closely related to questions concerning spectral properties of Hamiltonians.

1.2 Mixed states and density operators

We have introduced quantum mechanics using the language of state vectors. An alternate formulation is possible using a description based on *density operators* or *density matrices*. This formulation is much more general than the state description since it includes the situation where the state vector in not known and allows a statistical description of this situation.

This happens, for instance, when we measure a qubit but discard the measurement outcome. In this situation, all we know is that the qubit is in the state $|0\rangle$ with probability $|\alpha|^2$ and in the state $|1\rangle$ with probability $|\beta|^2$. We say that the qubit is in a probabilistic mixture of the pure states $|0\rangle$ and $|1\rangle$. Such mixtures are generalizations of pure states. The naive way to think about a mixture is that we have a probability distribution over pure quantum states. However, this point of view is not correct since different probability distributions on the set of pure states may lead to the same mixed state.

Suppose a quantum system is in one of a number of states $|\Psi_i\rangle$, where i is a index, with respective probability p_i. We call $\{(p_i, |\Psi_i\rangle)\}$ an *ensemble of pure states*. The density operator describing the incomplete knowledge of the system's state is defined by the equation

$$\rho := \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \,. \tag{1.18}$$

All postulates of quantum mechanics can be stated in terms of the density operators. Suppose, for example, the evolution of the quantum system is described by the unitary operator U. If the system was initially in the state $|\Psi_i\rangle$ with probability p_i, then after the evolution the system will be in the state $|\Psi_i'\rangle = U|\Psi_i\rangle$ with probability p_i. Therefore, the evolution of the density operator is described by the equation

$$\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i| \longrightarrow \sum_i p_i U|\Psi_i\rangle\langle\Psi_i|U^\dagger = U\rho U^\dagger \,. \tag{1.19}$$

Measurements can also be described easily in the density operator language. We use $\mathrm{tr}(A)$ to denote the trace of the operator A. Suppose we perform a measurement described by the projections P_m. If the initial state was $|\Psi_i\rangle$, then the probability of getting the measurement result m is

$$p(m|i) = \langle\Psi_i|P_m|\Psi_i\rangle = \mathrm{tr}(P_m|\Psi_i\rangle\langle\Psi_i|)\,,$$

where we have used the equality $\langle\Psi|A|\Psi\rangle = \mathrm{tr}(A|\Psi\rangle\langle\Psi|)$ that is valid for all kets $|\Psi\rangle$ and all operators A. Therefore, the probability of obtaining result m is

$$\begin{aligned} p(m) &= \sum_i p_i p(m|i) \\ &= \sum_i p_i \mathrm{tr}(P_m|\Psi_i\rangle\langle\Psi_i|) \\ &= \mathrm{tr}(P_m\rho)\,. \end{aligned}$$

If the initial state was $|\Psi_i\rangle$, then the state after obtaining the measurement outcome m is

$$|\Psi_{im}\rangle = \frac{P_m|\Psi_i\rangle}{\langle\Psi_i|P_m|\Psi_i\rangle}\,.$$

By some elementary calculations we see that the density operator ρ changes to

$$\rho \mapsto \rho_m = \frac{P_m\rho P_m}{\mathrm{tr}(P_m\rho P_m)}$$

if we obtain the measurement result m.

The following theorem is needed to establish a characterization of density operators. Furthermore, we will also use it when considering Hamiltonians. The proof can be found in Nielsen and Chuang [NC00].

Theorem 1.4 (Spectral decomposition)
Any Hermitian operator A ($A^\dagger = A$) on $\mathbb{C}^d$ is diagonal with respect to some orthonormal basis $\{|\Psi_1\rangle, \ldots, |\Psi_d\rangle\}$ for $\mathbb{C}^d$ and has real eigenvalues $\{\lambda_1, \ldots, \lambda_d\}$. We have

$$A = \sum_{i=1}^{d} \lambda_i |\Psi_i\rangle\langle\Psi_i|\,, \tag{1.20}$$

where λ_i are the eigenvalues corresponding to the eigenvectors $|\Psi_i\rangle$.
Conversely, any diagonalizable operator with real eigenvalues is Hermitian.

The class of operators that are density operators are characterized by the following useful theorem:

Theorem 1.5 (Characterization of density operators)
An operator ρ is the density operator associated to some ensemble $\{(p_i, |\Psi_i\rangle)\}$ if and only if it satisfies the conditions:

1. *ρ has trace equal to one.*
2. *ρ is a* positive *operator, that is, all its eigenvalues are non-negative.*

Proof. Suppose $\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|$ is a density operator. Then we have

$$\mathrm{tr}(\rho) = \sum_i p_i \mathrm{tr}(|\Psi_i\rangle\langle\Psi_i|) = \sum_i p_i = 1\,,$$

so the trace condition $\mathrm{tr}(\rho) = 1$ is satisfied. Suppose $|\Phi\rangle$ is an arbitrary ket in the state space. Then

$$\begin{aligned}\langle\Phi|\rho|\Phi\rangle &= \sum_i p_i \langle\Phi|\Psi_i\rangle\langle\Psi_i|\Phi\rangle \\ &= \sum_i p_i |\langle\Phi||\Psi_i\rangle|^2 \\ &\geq 0\,,\end{aligned}$$

so the positivity condition is satisfied.
Conversely, suppose that ρ is any operator satisfying the trace and positivity conditions. Since ρ is positive, it must have a spectral decomposition

$$\rho = \sum_j \lambda_j |\Psi_j\rangle\langle\Psi_j|\,,$$

where the eigenvectors $|\Psi\rangle$ are mutually orthogonal, and λ_j are real, non-negative eigenvalues of ρ. From the trace condition we see that $\sum_j \lambda_j = 1$. Therefore, a system in state $|\Psi_j\rangle$ with probability λ_j will have density operator ρ. □

The Bloch sphere representation is a geometric representation of qubits' states. We will use a similar representation for Hamiltonians. Let us start with the pure states. Because of $|\alpha|^2 + |\beta|^2 = 1$, we may rewrite eq. (1.1) as

$$|\Psi\rangle = e^{i\gamma}\left(\cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle\right)\,, \tag{1.21}$$

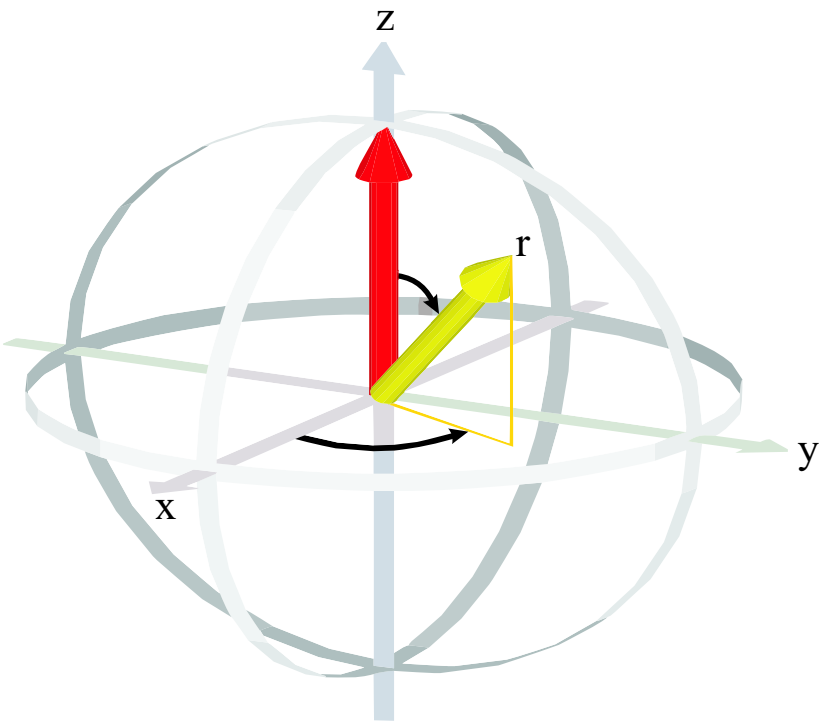


Figure 1.7: Bloch sphere

where θ, ϕ and γ are real numbers. The global phase factor $e^{i\gamma}$ can be ignored since it has no observable effects in the measurement. For that reason we effectively write

$$|\Psi\rangle = \cos(\theta/2)|0\rangle + e^{i\varphi}\sin(\theta/2)|1\rangle\,. \tag{1.22}$$

The numbers θ and φ define a point on the unit three-dimensional sphere, as shown in Figure 1.2. This sphere is called the *Bloch sphere*. It provides a useful means of visualizing the state of a single qubit. Many operations on a single qubit are conveniently described in the Bloch sphere picture.

For the generalization to mixed states we need a basis for Hermitian operators on $\mathbb{C}^2$. An extremely useful basis consists of the the identity matrix $\mathbf{1}$ and the Pauli matrices.

Definition 1.6 (Pauli matrices) *The Pauli-matrices are given by*

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.23}$$

An arbitrary density matrix of a single qubit may be written as

$$\rho = \frac{\mathbf{1} + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z}{2}, \tag{1.24}$$

where $r = (r_x, r_y, r_z)$ is a real three-dimensional vector such that $\|r\| \leq 1$. This vector is known as the *Bloch vector* for the state ρ. A state is pure if and only if $\|r\| = 1$. All this is seen as follows. The density operator corresponding to the pure state $|\Psi\rangle$ in eq. (1.22) is

$$|\Psi\rangle\langle\Psi| = \begin{pmatrix} \cos^2\frac{\theta}{2} & \frac{1}{2}\,e^{i\varphi}\sin\theta \\ \\ \frac{1}{2}\,e^{-i\varphi}\sin\theta & \sin^2\frac{\theta}{2} \end{pmatrix}, \tag{1.25}$$

where we have used the equality $\sin\frac{\theta}{2}\cos\frac{\theta}{2} = \frac{1}{2}\sin\theta$. By elementary calculations one sees that $|\Psi\rangle\langle\Psi| = \frac{1}{2}(\mathbf{1} + r_x\sigma_x + r_y\sigma_y + r_z\sigma_z)$, where $r_x = \sin\phi\sin\theta$, $r_y = \cos\phi\sin\theta$, and $r_z = \cos\frac{\theta}{2} - \sin\frac{\theta}{2}$.

Chapter 2

Control-theoretic model and simulation of Hamiltonians

In this chapter we consider a *control-theoretic model* that is closer to physics than the quantum circuit model. Our theory of mutual simulation of Hamiltonians is based on this model.

For the implementation of a quantum computer it is necessary to control the time evolution of the underlying physical system in a universal way. We formulate the task to implement unitary transformations on quantum registers as a control-theoretic problem following the approach of Khaneja [Kha00] and Khaneja et al. [KBG01]. In particular, we take into account that the desired unitary transformation have to be implemented by the continuous time evolution of the quantum system rather than by a discrete sequence of gates.

So far we have described the evolution of a *closed* quantum system by *unitary transformations*. That is, the state $|\Psi\rangle$ of the system at time t_1 is related to the state $|\Psi'\rangle$ of the system at time t_2 by a unitary operator $U(t_1, t_2)$ which depends only on the time t_1 and t_2, $|\Psi'\rangle = U(t_1, t_2)|\Psi\rangle$. Just as the abstract framework of quantum mechanics does not tell us the state space or a quantum state of a *particular* quantum system, it does not tell us which unitary operators $U(t_1, t_2)$ describe real-world quantum dynamics. Since every unitary has to be implemented by using physical interactions and there is only a restricted set of interactions in nature it seems clear that not every unitary transformation can be obtained directly but has to be realized as a sequence of elementary or natural unitary transformations. Therefore, an important question to ask is: "What unitary transformations are natural to consider?" Recall that within the quantum circuit model we assume that all single qubit and all two-qubit gates are elementary.

The usefulness and attraction of the quantum circuit model may be partially explained by the following reasons. From a mathematical point of view, it is quite natural to look for a subset of the Lie group of unitary transformations on

the Hilbert space that generate the whole group. Obviously, the set of all single qubit operations does not generate the whole Lie group. However, the set of all single qubit gates augmented by a nontrivial two-qubit gate (e.g. the CNOT) is already sufficient to generate the whole Lie group. There is thus no reason to consider more complicated transformations like three-qubit unitary operators as basic ones. Furthermore, the model of two-qubit gates might be considered as the attempt to develop quantum computation in strong analogy to the theory of classical devices. Building complex logical networks from two-bit gates is a successful concept of classical computation.

Despite these reasons in favor of the quantum circuit model, it is necessary to keep in mind possible "criticism", modifications, and extensions. The aim of this thesis is to carry out an examination and to give a physical explanation of the the quantum circuit model based on the control-theoretic model. By posing these questions we emphasize the fundamental point that every computation process is carried out by physical processes. Therefore, in our attempts to formulate models for information processing we should always attempt to go back to fundamental physical laws. This is also important if we want to understand the fundamental limits of computation.

To define a physically founded complexity measure of unitary transformations we have to work with a more refined description of quantum dynamics taking into account that the evolution of a quantum system is a *continuous time* process. This is explained in the following section.

2.1 Hamiltonian time evolutions

The time evolution of a closed quantum system $\mathcal{H}$ (of dimension d) is described by the time-independent *Schrödinger equation*,

$$\frac{d}{dt}U(t) = -iH\,U(t)\,, \quad U(0) = \mathbf{1}\,, \tag{2.1}$$

where H is a fixed Hermitian operator known as the *Hamiltonian* of the closed system $\mathcal{H}$.

If we know the Hamiltonian of the system, then we understand its dynamics completely, at least in principle. In general figuring out the Hamiltonian needed to describe a particular physical system is a very difficult problem that relies on knowledge from experiment or on a theory about the relevant forces (interactions).

Because the Hamiltonian is a Hermitian operator it has a spectral decomposition

$$H = \sum_E E|E\rangle\langle E|\,, \tag{2.2}$$

with eigenvalues E and corresponding normalized eigenvectors $|E\rangle$ that are mutually orthogonal. The states $|E\rangle$ are called *energy eigenstates* , or *stationary states* , and E is called the energy of the state $|E\rangle$. The minimal eigenvalue is called *ground state energy* and the corresponding eigenvectors *ground states* . The reason why the states $|E\rangle$ are called stationary states is because their only change after time t is to acquire a phase factor,

$$|E\rangle \mapsto \exp(-iEt)|E\rangle\,. \tag{2.3}$$

As an example, suppose a single qubit has Hamiltonian $H = \omega\sigma_z$. Here ω is a parameter that, in practice, needs to be experimentally determined, or can be computed from background knowledge. The energy eigenstates of this Hamiltonian are obviously the same as the eigenstates of σ_z, namely $|0\rangle$ and $|1\rangle$, with corresponding energies $-\omega$ and ω.

Let us now explain the connection between the Hamiltonian picture of dynamics and the corresponding unitary evolution. The answer is given by the solution to Schrödinger's equation

$$U(t_1, t_2) = \exp\left(-iH(t_2 - t_1)\right)\,. \tag{2.4}$$

This operator is unitary since we have

$$U(t_1, t_2) = \sum_E \exp\left(-iE(t_2 - t_1)\right)|E\rangle\langle E|\,.$$

More generally, any unitary operator U can be realized in the form $U = \exp(-i\tilde{H})$ for some Hermitian operator $\tilde{H}$. Therefore, we have a one-to-one correspondence between Hamiltonians H and one-parameter groups $(U_t)_{t\in\mathbb{R}}$ of unitaries, i.e. unitary representations of the additive group $\mathbb{R}$. The unitary transformation U_t is the solution of the Schrödinger equation.

We have already seen that a global phase factor $e^{i\gamma}$ has no physical meaning; therefore it is sufficient to consider $SU(d)$, i.e. the group of special unitary matrices of size $d \times d$, instead of $\mathcal{U}(d)$. Note that $\exp(-i\gamma\mathbf{1}t) = e^{-i\gamma t}\mathbf{1}$ gives only a global phase factor; therefore we may assume w.l.o.g. that the Hamiltonians are elements of $su(d)$, i.e., the Lie algebra of traceless Hermitian matrices of size $d \times d$.

In quantum computation and quantum information theory we often speak of *applying* a unitary operator to a particular quantum system. For example, within the quantum circuit model we may speak of applying the NOT gate X to a single qubit. This seems to contradict that unitary operators describe the time evolution of *closed* quantum systems. After all, applying a unitary operation may require some sort of external system interacting with the quantum system, and the quantum system is not closed. For instance, this occurs when a laser is focused on an atom in an ion trap. The effect of the laser field on the atom is

described by a time-dependent atomic Hamiltonian. It is not necessary to include the laser explicitly in the description. This resolves the (seeming) contradiction, and the atomic Hamiltonian contains terms depending on laser intensity and some other parameters of the laser, that can be controlled at will. Therefore, the time evolution of the atom may be described by a Hamiltonian that we can vary, despite the atom is not a closed system.

2.2 Definition of the control-theoretic model

More generally, for many systems it is possible to describe their time evolution by a time-dependent Hamiltonian that may be changed according to some parameters under the experimentalist's control. Within this picture the time evolution of a closed quantum system is described by the time-dependent *Schrödinger equation*

$$\frac{d}{dt}U(t) = -iH(t)\,U(t)\,, \quad U(0) = \mathbf{1}\,, \tag{2.5}$$

where $H(t)$ is the *time-dependent Hamiltonian* and $U(t)$ the resulting transformation at time t. Recall that the transformations on $\mathcal{H}$ are elements of the Lie group $SU(d)$ of special unitary matrices of size d. The Hamiltonians are elements of the Lie algebra $su(d)$ of traceless Hermitian matrices of size d.

The set of external *controls* is characterized by a subset $\mathcal{C}$ of $su(d)$ of control Hamiltonians that can be switched on. By definition we can split the Hamiltonian

$$H = H_0 + H_c(t)\,, \tag{2.6}$$

where H_0 is the part of the Hamiltonian that is internal to the system, called the *drift* or *free Hamiltonian* and H_c is the part of the Hamiltonian that can be externally changed, called the *control Hamiltonian*. Now we are able to formalize complexity.

Definition 2.1 (Control-theoretic complexity)
Let $U \in SU(d)$ be a unitary transformation. The complexity of U is the minimal time T required to steer the system evolving according to the differential equation

$$\frac{d}{dt}U(t) = -i\big(H_0 + H_c(t)\big)U(t)\,, \tag{2.7}$$

from the identity, $U(0) = \mathbf{1}$, to the desired unitary $U = U(T)$.

Note that the design of time-optimal control sequences in quantum systems is a special case of the optimal control problem on compact Lie groups (cf. Jurdjevic and Sussmann [JS72]). This theory has many applications in engineering (cf. Jurdjevic [Jur97]) ranging from steering space shuttles to determining optimal movements of robots.

2.2.1 Fast control limit

This model may simplified further since in many physically relevant situations the so-called *fast control limit* may be used. In this approximation, we assume that the strength of the control Hamiltonians $H_c(t)$ can be made arbitrary large compared to the natural Hamiltonian H. Equivalently, we may assume that the directly accessible control possibilities are unitaries in the control group[1] $\mathcal{K}$ and that they can be performed arbitrarily fast. This approximation is justified e.g. in NMR because the coupled and local evolutions act on significantly different time scales.

In the following we denote the system Hamiltonian by H instead of H_0. To implement the desired unitary we alternate the natural time evolution $\exp(-iHt)$ with control operations $V_j \in \mathcal{K}$. A general control sequence has the form

> perform the control operation V_1, wait for the time t_1, perform V_2, wait t_2,..., perform V_N, wait for the time t_N.

The resulting unitary is given by

$$U = \exp(-iHt_N)V_N \cdots \exp(-iHt_2)V_2 \exp(-iHt_1)V_1 \,. \tag{2.8}$$

This may be rewritten as

$$U = U_N U_N^\dagger \exp(-iHt_j)U_N \cdots U_2^\dagger \exp(-iHt_j)U_2 U_1^\dagger \exp(-iHt_j)U_1 \,, \tag{2.9}$$

where $U_j = \prod_{k=1}^{j} V_k$. Using the identity $U^\dagger \exp(A)U = \exp(U^\dagger A U)$ we get

$$U = U_N \exp(-iH_N t_N) \cdots \exp(-iH_2 t_2) \exp(-iH_1 t_1),, \tag{2.10}$$

where $H_j = U_j^\dagger H U_j$. The operators H_j are the Hamiltonians in the so-called "toggling frame". Note that U and $U_N^\dagger U$ can be implemented in the same time since $U_N \in \mathcal{K}$. The implementation time of a unitary depends only on the coset of $\mathcal{K}$ in G (cf. Khaneja [KBG01]). Therefore we may drop the unitary U_N on the right side of the product in eq. (2.10).

Let $Ad_{\mathcal{K}}(H)$ denote the set of conjugates of H under the action of $\mathcal{K}$ via conjugation

$$Ad_{\mathcal{K}}(H) = \{U^\dagger H U \mid U \in \mathcal{K}\} \,. \tag{2.11}$$

By applying the control operations we effectively change the Hamiltonian into a piecewise constant time-dependent Hamiltonian. Then the unitary U is the solution of a time-dependent Schrödinger equation with piecewise constant Hamiltonians in $Ad_{\mathcal{K}}(H)$.

[1] This control group is the Lie group corresponding to the Lie algebra generated by the control Hamiltonians (cf. Khaneja et al. [KBG01]).

2.2.2 Pair-interaction Hamiltonians

So far we have not considered the internal structure of the quantum system. We assume that the natural Hamiltonian H to act on an n-fold tensor product Hilbert space

$$\mathcal{H} := \mathbb{C}^d \otimes \mathbb{C}^d \otimes \cdots \otimes \mathbb{C}^d \,,$$

where each $\mathbb{C}^d$ denotes the Hilbert space of a qudit, i.e. a d-dimensional system. Recall that the Lie algebra $su(d)$ of traceless Hermitian operators on $\mathcal{H}$ is a $m := (d^2 - 1)$-dimensional real vector space. Let $B := \{\sigma_1, \ldots, \sigma_m\}$ be a basis of $su(d)$. For two-dimensional systems, a useful basis for $su(2)$ is given by the Pauli matrices (see Definition 1.6).

Let $\sigma_\alpha \in B$. By $\sigma_\alpha^{(k)}$ we denote the operator

$$\mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes \sigma_\alpha \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} \,. \tag{2.12}$$

that acts as σ_α on the kth qudit.

A general feature of the most important Hamiltonians available in nature is that particles interact with other particles in such a form that the total Hamiltonian is a sum of pair-interactions. Mathematically, this is described by the following definition.

Definition 2.2 (Pair-interaction Hamiltonian)
A pair-interaction Hamiltonian of a quantum network of n interacting qudits is a traceless Hermitian operator H that can be decomposed as follows:

$$H = \sum_{k<l} \sum_{\alpha\beta} J_{kl;\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} + \sum_k \sum_\alpha r_{k;\alpha} \sigma_\alpha^{(k)} \,, \tag{2.13}$$

where $J_{kl;\alpha\beta} \in \mathbb{R}$ and $r_{k;\alpha} \in \mathbb{R}$. The first part of eq. (2.13) defines the couplings between the qudits and the second part the free Hamiltonians of the qudits. We denote the couplings by H_{kl} and the free Hamiltonians by H_k.

Now we define the interaction graphs that describe the coupling topology of Hamiltonians. Let us first introduce some basic graph-theoretical notions. A *graph* is an ordered pair $G = (V, E)$ with $V \subseteq \{1, 2, \ldots, n\}$ and $E = \{e_1, e_2, \ldots, e_m\} \subseteq V \times V$. Elements of V are called *vertices* and label the qudits. Elements of E are called *edges* and label the pair-interactions between the qudits. An edge $e = (k, l)$ is an ordered pair of vertices k and l called the ends of e. We consider only *undirected* graphs with no *loops*, i.e., edges of the form (k, k). To have a unique representation we require that $k < l$. Two distinct edges are called *adjacent* if and only if they have a common vertex. A graph G is called *complete* if every pair of distinct vertices of G are adjacent in G; such a graph is denoted by K.

Definition 2.3 (Interaction graph)
Let H be a pair-interaction Hamiltonian of n coupled qudits

$$H := \sum_{k<l} H_{kl} + \sum_k H_k \,.$$

The corresponding interaction graph is defined to be the graph $G := (V, E)$ with $V := \{1, \ldots, n\}$ and $(k, l) \in E$ if $H_{kl} \neq \mathbf{0}$, i.e., two vertices k and l are adjacent if the nodes k and l are coupled with each other. We say that H has a complete interaction graph if G is the complete graph on n vertices.

Having described "physical" Hamiltonians let us consider what kind of control operations are possible. In the setting discussed here and in most other articles on simulation of Hamiltonians the only possibilities of external control are given by local unitaries on each qudit. It is assumed that they can be implemented independently on the qudits and arbitrarily fast compared to the natural time evolution (fast control limit or bang-bang control). This is the usual model used to describe experiments in Nuclear Magnetic Resonance (cf. Ernst et al. [EBW87], Haeberlen [Hae76], and Slichter [Sli90]). These assumptions are summarized in the following definition:

Definition 2.4 (Control group)
The control operations are elements of the control group

$$\mathcal{K} := SU(d) \otimes SU(d) \otimes \cdots \otimes SU(d) \,. \tag{2.14}$$

Every $U \in \mathcal{K}$ can be written as

$$U := U^{(1)} \otimes U^{(2)} \otimes \cdots \otimes U^{(n)} \tag{2.15}$$

to denote the corresponding local operations $U^{(1)}, \ldots, U^{(n)}$ on the qudits.

This model requires a refined definition of complexity of unitaries.

Definition 2.5 (Continuous complexity measure)
Let H be a pair-interaction Hamiltonian of n qudits, $G := SU(d^n)$ the group of unitary transformations on the joint Hilbert space, and $\mathcal{K} := SU(d) \otimes \cdots \otimes SU(d)$ be the control group. The complexity of $U \in G$ is the minimal time t such that

$$U = V_N \exp(-i\, H\, t_j) \cdots V_2 \exp(-i\, H\, t_2) V_1 \exp(-i\, H\, t_1) \,, \tag{2.16}$$

where $V_j \in \mathcal{K}$ and $t_1, \ldots, t_N$ are real positive numbers summing up to t.

The problem to compute the minimal implementation time in the general case is a very difficult task. The case of two coupled qubits has been solved by Khaneja et al. [KBG01]:

Theorem 2.6 (Two qubit case)
Let $H = \sigma_z \otimes \sigma_z$ and $\mathcal{K} = SU(2) \otimes SU(2)$. The minimum time required to produce a unitary $U \in SU(4)$ is the smallest value of $\sum_{i=1}^{3} |\alpha_i|$, such that we can solve

$$U = W_1 \exp(-i\alpha_1 \sigma_x \otimes \sigma_x + \alpha_2 \sigma_y \otimes \sigma_y + \alpha_3 \sigma_z \otimes \sigma_z) W_2 \,, \tag{2.17}$$

where $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$, W_1 and W_2 belong to $SU(2) \otimes SU(2)$. The number of switches is at most 3.

Optimal solutions of more general cases are not known. Approximate solutions may be obtained from average Hamiltonian theory.

2.3 Average Hamiltonian theory

The formalism of average Hamiltonian theory allows to determine the resulting time evolution at a time t by writing the evolution of a time independent average Hamiltonian $\bar{H}$. We briefly sketch the key element of average Hamiltonian theory. We refer reader to the books [EBW87, Hae76, Sli90] for more details.

The overall dynamic after time t of evolution is given by

$$U(t) = \mathcal{T} \exp\left(-i \int_0^t \mathrm{d}\tau H(\tau)\right) = \exp(-i\bar{H}t) \,,$$

where $\mathcal{T}$ denotes the Dyson time ordering operator. We look for a Hamiltonian $\bar{H}$ such that $\exp(-i\bar{H}t) = U(t)$. A solution of this equation is a time-independent Hamiltonian that would result in the same unitary if it were applied over the same period. If the Hamiltonians at different times commute, i.e., $[H(\tau), H(\tau')] = \mathbf{0}$ for all τ, τ', the average Hamiltonian is given by $\bar{H} = \int_0^t \mathrm{d}\tau H(\tau)$. However, this is rarely the case.

For sufficiently small t, the *Magnus expansion* provides a formal means of calculating the average Hamiltonian:

$$\bar{H} = \bar{H}^{(0)} + \bar{H}^{(1)} + \bar{H}^{(2)} + \dots \,, \tag{2.18}$$

where the operators $\bar{H}^{(0)}, \bar{H}^{(1)}, \bar{H}^{(2)}, \dots$ are the average Hamiltonians of increasing order

$$\bar{H}^{(0)} = \frac{1}{t} \int_0^t \mathrm{d}\tau H(\tau) \,, \tag{2.19}$$

$$\bar{H}^{(1)} = \frac{-i}{2t} \int_0^t \mathrm{d}\tau' \int_0^{\tau'} \mathrm{d}\tau'' [H(\tau'), H(\tau'')] \,. \tag{2.20}$$

The operator $\bar{H}^{(0)}$ will be called the *average Hamiltonian* . The norms of the average Hamiltonians can be bounded as follows: $\|H^{(0)}\| \leq 1$ and $\|\bar{H}^{(1)}\| \leq t/2$

since $\|[H(\tau'), H(\tau'')]\| \leq 2\|H(\tau')\|\|H(\tau'')\| = 2$ and the integration is taken over the simplex of area $t^2/2$. The norm of the higher order terms is bounded by higher orders of t. Therefore, for sufficiently small time t the resulting unitary $U(t)$ is essentially determined by $\bar{H}^{(0)}$.

Note that the approximation used in eq. (2.19) is closely related to the Trotter formula . Let $A_i \in \mathbb{C}^{d\times d}$ be any N matrices. The Trotter formula states that

$$\exp\left(\sum_{j=1}^{N} A_i\right) = \lim_{m\to\infty}\left(\prod_{j=1}^{N}\exp(A_i/m)\right)^m . \tag{2.21}$$

This observation leads directly toward the notion of simulating Hamiltonians that we introduce in the next section.

2.4 Simulation of Hamiltonians

Simulation of Hamiltonian time evolutions of general quantum systems is an interesting application for future quantum computers. Historically, this idea was the first motivation to study quantum computers:

> Can physics be simulated by a universal computer? [...] the physical world is quantum mechanical, and therefore the proper problem is the simulation of quantum physics [...] the full description of quantum mechanics for a large system with R particles [...] has to many variables, it **cannot be simulated** with a normal computer with a number of elements proportional to R [... but it can be simulated with] quantum computer elements. [...] Can a quantum system be probabilistically simulated by a classical (probabilistic, I'd assume) universal computer? [...] If you take the computer to be the classical kind I've described so far [...] the answer is certainly, No!
>
> Feynman [Fey82], as quoted in [NC00]

Let H be the Hamiltonian of a quantum system. Simulation of H means to mimic the dynamical behavior of the quantum system, described by the Schrödinger's equation

$$\frac{d}{dt}|\Psi(t)\rangle = -iH|\Psi(t)\rangle\,. \tag{2.22}$$

Although "physical" Hamiltonians consist of pair-interactions only (see Definition 2.2), the resulting quantum dynamics may be highly non-trivial. Simulation of quantum systems by classical computers is possible, but generally there are no known efficient algorithms. This fact has consequences in the study of solid state systems since our understanding of a large spectrum of collective quantum phenomena is hindered by the intractability of tracking quantum dynamics.

In Chapter 7 we consider some computational problems in statistical physics to illustrate the complexity of quantum dynamics.

2.4.1 Simulation with quantum circuits

In early works on simulation of Hamiltonian evolution the desired Hamiltonians are simulated by quantum circuits (cf. LLOYD [Llo96] and SOMMA ET AL. [SOG+01]). The idea behind this is as follows.

Recall that n particles mostly interact in the form of pair-interactions, that is, the system Hamiltonian H can be decomposed as

$$H = \sum_{k<l} H_{kl} + \sum_{k} H_k$$

where H_{kl} is a Hermitian operator acting on the joint Hilbert space of particles k and l and H_k is the free Hamiltonian of particle k. Note that pair-interactions H_{kl} are infinitesimal versions of two qubit gates: every unitary of the form $\exp(-iH_{kl}t)$ for $t \geq 0$ is a two-qubit gate. Analogously, free Hamiltonians are infinitesimal versions of single qubit gates. This may be considered as an important justification for the quantum circuit model, since it refers to the form of the fundamental interactions in nature. But this argument is not really correct. In general, there is no obvious correspondence between the time evolution $\exp(-iHt)$ and a sequence of single qubit and two-qubit gates, where H is a pair-interaction Hamiltonian. Nevertheless, there is a straightforward simulation by quantum circuits in an *approximative* sense given by the *Trotter formula*:

$$\lim_{m\to\infty} \left(\prod_{k<l} \exp\big(-iH_{kl}\, t/m\big) \right)^m = \exp\Big(-i \sum_{k<l} H_{kl}\, t\Big). \tag{2.23}$$

This example shows, that the simulation of the time evolution caused by a time-independent pair-interaction Hamiltonian might require an *infinite* number of qubit gates. However, it has been shown that the number m of gates required to simulate the time evolution $\exp(-iHt)$ up to an error ϵ is only growing with t^2/ϵ (cf. LLOYD [Llo96]). Despite the fact that infinite accuracy requires an infinite number of gates, the *time* it takes to implement the growing number of gates does not tend to infinity if one assumes that the implementation of $\exp(-iH_k\, t/m)$ and $\exp(-iH_{kl}\, t/m)$ requires the time $O(t/m)$ [Llo96]. Based on this assumption, it was shown by JANZING AND BETH [JB01] that the running time of a quantum circuit simulating the dynamics using the Trotter formula is determined by a graph invariant (*chromatic index*) of the interaction graph.

2.4.2 Control-theoretic simulation

A more control-theoretic formulation of simulating Hamiltonians has become popular [DNBT01, WJB02b, NBD$^+$01, WRJB02b, SM01, BCL$^+$02]. In this formulation we assume that the dynamics of the quantum computer is determined by its Hamiltonian together with external control possibilities as in the control-theoretic model. Simulation of Hamiltonians is based on the average Hamiltonian.

We use the notation of the control-theoretic model. Let U_j be control operations and τ_j positive real numbers. The τ_j specify the fraction of time between the control operations. If t is small compared to the "time scale" of the H, then according to average Hamiltonian theory the evolution

$$U(t) = \prod_{j=1}^{N} \exp(-i\, U_j^\dagger H U_j\, \tau_j t)\,. \tag{2.24}$$

is approximatively given by

$$\exp(-i\bar{H}t)\,,$$

where $\bar{H}$ is the *average Hamiltonian*

$$\bar{H} := \sum_{j=1}^{N} \tau_j U_j^\dagger H U_j\,.$$

The sum $\tau := \sum_{j=1}^{N} \tau_j$ gives the slow-down of the resulting time evolution with respect to the original one.

The discussion above motivates the notion of simulating a Hamiltonian $\tilde{H}$ by the system Hamiltonian H. This problem has been studied for the first time in Wocjan et al. [WJB02b]. The complexity of the simulation task is measured by the time overhead and the number of time steps.

Definition 2.7 (Simulation of Hamiltonians)
Let $\tilde{H}$ be an arbitrary Hamiltonian. We say $\tilde{H}$ can be simulated by H with time overhead τ and N time steps if and only if there are N control operations $U_j \in \mathcal{K}$ and N positive real numbers τ_j summing to τ such that

$$\tilde{H} = \sum_{j=1}^{N} \tau_j U_j^\dagger H U_j\,. \tag{2.25}$$

We write $\tilde{H} \leq \tau H$ to denote that H can simulate $\tilde{H}$ with time overhead τ.

The time-overhead gives the slow-down of the simulated Hamiltonian evolution with respect to the original one. The problem of time-optimal simulation of a Hamiltonian is reduced to a convex optimization problem (this has been noted in [WJB02b, BCL$^+$02]).

Remark 2.8 These convex problems are closely related to the method for obtaining pseudo-pure states by averaging over random unitary transformations (cf. KNILL ET AL. [KCL98]). Pseudo-pure states are states that can be written as a convex combination of the maximally mixed state (with density matrix $\mathbf{1}/d$) and a pure state $|\psi\rangle\langle\psi|$. Writing a general state as $\rho = \mathbf{1}/d + A$, where A is the traceless part, we have that ρ can be transformed into the pseudo-pure state $(1-\lambda)\mathbf{1}/d + \lambda|\psi\rangle\langle\psi|$ by averaging over unitary transformations if and only if A can be transformed (without time overhead) into the traceless operator $-\lambda\mathbf{1}/d + \lambda|\psi\rangle\langle\psi|$. Determining the optimal *signal-to-noise ratio* of the attainable pseudo-pure state is hence directly related to determining minimal time overhead of simulation schemes.

We introduce the notion of control sequences to describe conveniently simulation schemes.

Definition 2.9 (Control sequence)
A control sequence C is a tuple

$$C := (\tau_1, U_1, \tau_2, U_2, \ldots, \tau_N, U_N) \tag{2.26}$$

where $U_j \in \mathcal{K}$ and $\tau_j > 0$ positive real numbers. We call $\tau := \sum_{j=1}^{N} \tau_j$ the time overhead *and N the* number of time steps *of the control sequence. The control sequence C acts on a Hamiltonian H as*

$$C(H) := \sum_{j=1}^{N} \tau_j U_j^\dagger H U_j \,.$$

It is useful to introduce the concatenation of control sequences. This allows to build complex control sequences from basic ones.

Definition 2.10 (Concatenation of control sequences)
Let $C_1 := (r_1, U_1, \ldots, s_M, U_M)$ and $C_2 := (s_1, V_1, \ldots, s_N, V_N)$ be two control sequences. The concatenation $C := C_2 \circ C_1$ is defined to be the control sequence

$$C = (s_j r_j, V_j U_i)_{i=1,\ldots,M;j=1,\ldots,N} \,. \tag{2.27}$$

The time overhead of C is the sum of time overheads, and the number of time steps is product of the time overheads.

It is easily verified that with this definition of concatenation we have

$$(C_2 \circ C_1)(H) = C_2(C_1(H)) \tag{2.28}$$

for all Hamiltonians H. Note that the concatenation of control sequences is not a commutative operation as shows the following example. Let $C_1 := (\frac{1}{2}, \mathbf{1}, \frac{1}{2}, i\sigma_y)$ and $C_2 := (1, R_z)$, where is a unitary with $R_z^\dagger \sigma_x R_z = \sigma_y$. Then we have for the concatenations $C_2(C_1(\sigma_x)) = \mathbf{0}$, but $C_1(C_2(\sigma_x)) = \sigma_y$.

Chapter 3

Decoupling and time-reversal algorithms

We have seen in the first chapter that quantum computing based on the quantum circuit model requires the ability to perform gates on the quantum register. The couplings created by the gates can only originate from the natural couplings in the quantum systems involved. However, naturally available Hamiltonians do not couple specific pairs of qubits as desired in most applications of quantum computation. The fundamental task to turn off unwanted couplings is so difficult that, coercing a complex system to do nothing – stopping all evolution – can be just as difficult as making it do something computationally useful. In the literature this task is usually referred to as *decoupling*. A closely related task is the so-called *time-reversal*, i.e., to simulate $-H$ if H is the system Hamiltonian.

We construct decoupling schemes for general qudit systems. The constructions use irreducible projective representations of finite groups and the design-theoretic concepts of orthogonal arrays and difference schemes. Then we show how to convert decoupling schemes to time-reversal schemes. These schemes have been derived in our paper Wocjan et al. [WRJB02a]. Their optimality is proved in Chapter 5.

Lastly, we discuss the case of *partially* coupled quantum networks. This is formally described by the graph-theoretical notion of *chromatic number*.

3.1 Annihilators

Before solving the problem of decoupling in quantum networks with many nodes we consider the case of a single node. We introduce the concept of an annihilator characterizing control sequences for switching off the possibly unknown dynamics of the node. We prove some optimality properties of annihilators and show how a minimal annihilator can be explicitly constructed using nice error bases. The

problem how to distribute the conjugations of an annihilator among the nodes of a quantum network can be formulated combinatorially. An efficient solution of this problem is given by orthogonal arrays.

Definition 3.1 (Annihilator)
A control sequence $C = (\tau_1, U_1, \tau_2, U_2, \ldots, \tau_N, U_N)$ is called an annihilator (of dimension d) with N times steps if

$$C(a) = \mathbf{0} \tag{3.1}$$

for all $a \in su(d)$. An annihilator is called minimal if there is no annihilator with less time steps.

We establish a correspondence between minimal annihilators and *error bases*. Error bases are used in quantum error correction to track the evolution of a state in the presence of noise and to allow a simple recovery procedure after measuring the syndrome (cf. Knill [Kni96a, Kni96b]). To define error bases we need the *trace inner product*.

Definition 3.2 (Trace inner product)
The trace inner product (also called Hilbert-Schmidt inner product) on $\mathbb{C}^{d\times d}$ is defined by

$$\langle A|B\rangle_{\mathrm{tr}} := \mathrm{tr}(A^\dagger B)/d \tag{3.2}$$

for all $A, B \in \mathbb{C}^{d\times d}$.

Definition 3.3 (Error basis)
An error basis $\mathcal{E} = \{U_1, \ldots, U_{d^2}\}$ is an orthogonal basis of $\mathbb{C}^{d\times d}$ with respect to the trace inner product consisting of unitary matrices.[1]

We derive an equivalent definition of error bases in the next lemma. Especially, this correspondence shows that error bases define annihilators.

Lemma 3.4 (Characterization of error bases)
Let $\mathcal{E} = \{U_1, \ldots, U_{d^2}\}$ be a set of unitary matrices in $\mathbb{C}^{d\times d}$. The following statements are equivalent:

1. *$\mathcal{E}$ is an error basis*
2. *$\frac{1}{d^2}\sum_{j=1}^{d^2} U_j^\dagger C U_j = \mathrm{tr}(C)\,\mathbf{1}/d$ for all $C \in \mathbb{C}^{d\times d}$*

[1] We will usually assume that $U_1 = \mathbf{1}$. Note that if $\{U_1, \ldots, U_{d^2}\}$ is an orthogonal basis then $\{UU_1, \ldots, UU_{d^2}\}$ is also an orthogonal basis for any unitary $U \in \mathbb{C}^{d\times d}$. Therefore, by setting $U := U_1^\dagger$ we can always achieve $U_1 = \mathbf{1}$.

Proof. Since operators of rank one span the whole vector space $\mathbb{C}^{d\times d}$ it is sufficient to prove the equivalence for all operators of the form $C := |\Psi_1\rangle\langle\Phi_1|$, where $|\Psi_1\rangle$ and $|\Phi_1\rangle$ are arbitrary ket in $\mathbb{C}^d$.

The first statement is (by definition of error basis) equivalent to the statement that $\mathcal{E}$ forms an orthogonal basis of $\mathbb{C}^{d\times d}$ with respect to the trace inner product. This again is equivalent to the completeness relation saying that

$$\langle A|B\rangle_{\mathrm{tr}} = \sum_{j=1}^{d^2} \langle A|U_j\rangle_{\mathrm{tr}}\langle U_j|B\rangle_{\mathrm{tr}}\,. \tag{3.3}$$

for all $A, B \in \mathbb{C}^{d\times d}$. We prove the lemma by showing that the completeness relation is equivalent to condition 2.

Let $A = |\Phi_1\rangle\langle\Phi_2|$ and $B = |\Psi_1\rangle\langle\Psi_2|$ be two arbitrary operators of rank one. We evaluate the completeness relation in eq. (3.3) with these operators and obtain for the l.h.s.

$$\langle\Psi_2|\Phi_2\rangle\langle\Phi_1|\Psi_1\rangle/d$$

and for the r.h.s.

$$\langle\Psi_2|M|\Phi_2\rangle\,,$$

where $M := \frac{1}{d^2}\sum_{j=1}^{d^2} U_j^\dagger C U_j$.

Since both sides are equal for all $|\Psi_2\rangle, |\Phi_2\rangle$ it follows that

$$M = \langle\Phi_1|\Psi_1\rangle\mathbf{1}/d = \mathrm{tr}(C)\mathbf{1}/d\,. \tag{3.4}$$

as desired. □

Before we can prove that *minimal* annihilators correspond to error bases, some basic definitions and results of quantum information theory are necessary. They can be found in Nielsen and Chuang [NC00]. Let $\mathcal{P} = (p_1, \ldots, p_d)$ be a probability distribution. The *Shannon entropy* is defined by the equation

$$H(\mathcal{P}) = -\sum_{i=1}^{d} p_i \log_2 p_i\,.$$

Shannon entropy measures the disorder of probability distributions. If $p_i = 1$ for some i, then the entropy is zero. The entropy takes its maximum value $\log_2 d$ for the uniform distribution. The notion of entropy extends to density operators, and is usually called *von Neumann entropy* . Let ρ be an arbitrary density operator on $\mathbb{C}^d$. Due to the properties of density operators and the spectral decomposition (see Theorem 1.4) we have

$$\rho = \sum_{i=1}^{d} \lambda_j |\Psi_i\rangle\langle\Psi_i|$$

such that the eigenvalues $\lambda_1, \ldots, \lambda_d$ form a probability distribution and the eigenvectors $|\Psi_1\rangle, \ldots, |\Psi_d\rangle$ form an orthogonal basis of $\mathbb{C}^d$. The von Neumann entropy $S(\rho)$ of ρ is defined by the equation

$$S(\rho) = -\sum_{i=1}^{d} \lambda_i \log_2 \lambda_i \,.$$

The von Neumann entropy takes its minimal value 0 on pure states, i.e. for $\rho = |\Psi\rangle\langle\Psi|$, and its maximal value $\log_2 d$ for the maximally mixed state $\rho = \mathbf{1}/d$. Let $U_1, \ldots, U_N \in \mathbb{C}^{d\times d}$ be arbitrary unitary matrices, $p_1, \ldots, p_N$ a probability distribution, and $|\Psi\rangle$ a state of $\mathbb{C}^d$. We have the following inequality (cf. Nielsen and Chuang [NC00], p. 518)

$$S\left(\sum_{j=1}^{N} p_j U_j^\dagger |\Psi\rangle\langle\Psi| U_j\right) \leq H(p_1, \ldots, p_N) \leq \log_2 N \,. \tag{3.5}$$

Now we have all the necessary tools to prove the next lemma characterizing the conditions on of minimal annihilators.

Lemma 3.5 (Conditions on minimal annihilators)
Minimal annihilators of dimension d have necessarily d^2 time steps. Furthermore, all times τ_j are necessarily equal.

Proof. Let $(\tau_1, U_1, \ldots, \tau_N, U_N)$ be an arbitrary annihilator of dimension d. We may assume w.l.o.g. that $\tau_1, \ldots, \tau_N$ form a probability distribution.
Let $|\Psi_1\rangle, \ldots, |\Psi_d\rangle$ be an orthonormal basis of $\mathbb{C}^d$. We define a special state in the composite system $\mathbb{C}^d \otimes \mathbb{C}^d$ together with its corresponding density operator

$$|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |\Psi_k\rangle \otimes |\Psi_k\rangle \,, \quad |\Psi\rangle\langle\Psi| = \frac{1}{d} \sum_{k,l=1}^{d} |\Psi_k\rangle\langle\Psi_l| \otimes |\Psi_k\rangle\langle\Psi_l| \,.$$

We use the annihilator to define a unitary mixing in order to show that $N \geq d^2$:

$$\begin{aligned}
\sum_{j=1}^{N} \tau_j (\mathbf{1} \otimes U_j^\dagger) |\Psi\rangle\langle\Psi| (\mathbf{1} \otimes U_j) &= \frac{1}{d} \sum_{k,l=1}^{d} |\Psi_k\rangle\langle\Psi_l| \otimes \sum_{j=1}^{N} \tau_j U_j^\dagger |\Psi_k\rangle\langle\Psi_l| U_j \\
&= \frac{1}{d} \sum_{k=1}^{d} |\Psi_k\rangle\langle\Psi_k| \otimes \mathbf{1}_d/d \\
&= \mathbf{1}_d/d \otimes \mathbf{1}_d/d \\
&= \mathbf{1}_{d^2}/d^2 \,.
\end{aligned}$$

It follows from the above equation that we need at least d^2 unitaries since the rank of each pure state $(\mathbf{1} \otimes U_j^\dagger)|\Psi\rangle\langle\Psi|(\mathbf{1} \otimes U_j)$ is 1 and since they must sum to d^2 (the rank of the maximally mixed state).

For $N = d^2$ all τ_j must be equal due to the inequality (3.5). □

Now we have all the necessary prerequisites needed to establish a one-to-one correspondence between minimal annihilators and error bases.

Theorem 3.6 (Minimal annihilators and error bases)
Let $\mathcal{E} = \{U_1, \ldots, U_{d^2}\}$ be a set unitaries in $\mathbb{C}^{d\times d}$. The following statements are equivalent:

1. *$\mathcal{E}$ is an error basis.*
2. *$A := (p, U_1, \ldots, p, U_{d^2})$ is a minimal annihilator (with $p := 1/d^2$).*

Proof. If $\mathcal{E}$ is an error basis then it follows from the correspondence established in Lemma 3.4 that C is an annihilator. Minimality follows from Lemma 3.5.
Let A be a minimal annihilator. Its actions extends to all matrices in $\mathbb{C}^{d\times d}$. Let $sl(d)$ be the vector space of all traceless matrices in $\mathbb{C}^{d\times d}$. Note that we have $sl(d) = su(d) + isu(d)$. Therefore, A maps every matrix in $sl(d)$ onto the zero matrix $\mathbf{0}$. The identity matrix $\mathbf{1}$ is mapped by A onto itself. By writing an arbitrary matrix $C \in \mathbb{C}^{d\times d}$ as a sum of a multiple of $\mathbf{1}$ and a traceless matrix in $sl(d)$ we see that

$$\frac{1}{d^2}\sum_{j=1}^{d^2} U_j^\dagger C U_j = \mathrm{tr}(C)\mathbf{1}/d\,. \tag{3.6}$$

Lemma 3.4 shows then that $\{U_1, \ldots, U_{d^2}\}$ forms an error basis. □

The connection between annihilators and error bases has been already hinted at in Viola et al. [VKL99] without giving a proof.
We now deal with the problem to construct minimal annihilators. As we have seen this is equivalent to constructing error bases. One way of constructing such bases relies on the concept of *nice error bases*. We refer to Knill [Kni96a] and Klappenecker and Rötteler [KR02] for an overview of this method and mention that nice error bases are used in the construction of quantum error control codes (cf. [CRSS98, Got96, Kni96b, Ste96]). They are also of interest in the theory of noiseless subsystems (cf. [KLV00, Zan01]).

Definition 3.7 (Nice error basis)
Let G be a group of order d^2 with identity element e. A nice error basis on $\mathbb{C}^d$ is a set $\mathcal{E} = \{U_g \in \mathcal{U}(d) \mid g \in G\}$ of unitary matrices such that

(i) U_e is the identity matrix,

(ii) $\mathrm{tr}\, U_g = d\,\delta_{g,e}$ for all $g \in G$,

(iii) $U_g U_h = \alpha(g,h) U_{gh}$ *for all* $g, h \in G$,

where $\alpha(g,h)$ *is a function from* $G \times G$ *to the group* $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$.

Following Isaacs [Isa76] we summarize some definitions of representation theory of finite groups so that we can understand better nice error bases. Let G be a group. $\mathrm{GL}(n,\mathbb{C})$ denotes the group of invertible $n \times n$ matrices with entries in $\mathbb{C}$.

A *representation* of G over $\mathbb{C}$ is a homomorphism φ from G to $\mathrm{GL}(n,\mathbb{C})$, for some n. The *degree* of ρ is the integer n. The representation φ is called *irreducible* if there are no invariant subspaces of $\mathbb{C}^n$ under the action of the matrices $\{\varphi(g)\}_{g\in G}$ apart from the trivial subspaces $\{0\}$ and $\mathbb{C}^n$. It is called *unitary* if all matrices $\varphi(g)$ are unitary.

A *projective representation* is a map $\varphi : G \to \mathrm{GL}(n,\mathbb{C})$ such that for every $g, h \in G$ there exists a scalar $\alpha(g,h) \in \mathbb{C}$ such that

$$\varphi(g)\varphi(h) = \alpha(g,h)\varphi(gh)\,.$$

The *degree* of φ is n and the function $\alpha : G \times G \to \mathbb{C}$ is the associated *factor system* of φ. The values $\alpha(g,h)$ for $g, h \in G$ form the factor set of φ. The factor set α has nonzero values and is uniquely determined by the matrices $\varphi(g)$. Both observations follow from the fact that the matrices $\varphi(g)$ are nonsingular. The projective representation is called *irreducible* there are no invariant subspaces of $\mathbb{C}^n$ under the action of the matrices $\{\varphi(g)\}_{g\in G}$ apart from the trivial subspaces $\{0\}$ and $\mathbb{C}^n$. It is called *unitary* if all $\varphi(g)$ are unitary. If $\alpha(g,h) = 1$ for all $g, h \in G$, then φ is an *ordinary* representation .

Let $\mathcal{E} = \{U_g \in \mathcal{U}(d) \mid g \in G\}$ be a nice error basis. In Klappenecker and Rötteler [KR02] it is shown that the map $g \mapsto U_g$ defines a projective representation of G; this is a consequence of conditions (i) and (iii). Condition (ii) shows that the matrices U_g are pairwise orthogonal with respect to the trace inner product. Hence, a nice error basis is an irreducible unitary projective representation of the finite group G. The group G itself is also called *index group* since its elements index the matrices of the nice error basis $\mathcal{E}$.

Note that in general the group generated by the matrices U_g for $g \in G$ will be larger than G, since these matrices are not closed under multiplication. A well-known theorem from projective representation theory (cf. Huppert [Hup83, Theorem V.24.6], Isaacs [Isa76, Theorem 11.15]) states that it is always possible to switch to an equivalent projective representation such that the images U_g generate a finite group $\hat{G}$. This group is called the *abstract error group* corresponding to $\mathcal{E}$. Whereas $g \mapsto U_g$ is an irreducible projective representation of G, this yields an irreducible *ordinary* representation of $\hat{G}$. It is a well-known fact that $\hat{G}$ is a central extension of G (cf. [Isa76]): denoting the center of $\hat{G}$ by $\zeta(\hat{G})$ this means that $\hat{G}/\zeta(\hat{G}) \cong G$.

Given a nice error basis $\{U_g \,|\, g \in G\}$, the abstract error group is isomorphic to the group generated by the matrices U_g. The assumption that the factor system α is of finite order ensures that the abstract error group is finite.

The following example shows the existence of nice error bases for any dimension $d \in \mathbb{N}$.

Example 3.8 (Heisenberg group)
The discrete Fourier transform of length $d \in \mathbb{N}$ is the unitary transformation defined by

$$\mathrm{DFT}_d := \frac{1}{\sqrt{d}}(\omega^{k\cdot l})_{k,l=0,\ldots,d-1}\,,$$

where ω denotes the primitive d-th root of unity $e^{2\pi i/d}$. Define $\mathcal{E}_d := \{S^i T^j : i,j = 0,\ldots,d-1\}$, where

$$S := \begin{pmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & 1 & \\ & & & 0 & \\ 1 & & & & \end{pmatrix}, \quad T := \mathrm{DFT}_d^{-1} \cdot S \cdot \mathrm{DFT}_d = \begin{pmatrix} 1 & & & \\ & \omega & & \\ & & \ddots & \\ & & & \omega^{d-1} \end{pmatrix}.$$

$\mathcal{E}_d$ is a nice error basis on $\mathbb{C}^d$ showing the existence of nice error bases for any dimension $d \in \mathbb{N}$. The index group in this case is the abelian group $G = Z_d \times Z_d$ whereas the corresponding abstract error group is a non-abelian group isomorphic to a semi-direct product $\hat{G} \cong (Z_d \times Z_d) \rtimes Z_d$ (the so-called Heisenberg group). The projective representation of G leading to the error basis $\mathcal{E}_d$ is defined by mapping the generators of G as follows: $(1,0) \mapsto S$ and $(0,1) \mapsto T$. The identity $ST = \omega TS$ is readily verified which shows that the commutator subgroup of $\hat{G}$ is contained in the center $\zeta(\hat{G})$. This also shows that the factor system α corresponding to the projective representation of G defined $\mathcal{E}_d$ is given by

$$\alpha((i,j),(k,l)) = \omega^{-jk},$$

for all $(i,j),(k,l) \in G$.

We give a brief account of some general properties of nice error bases (cf. [KR02]). A complete classification of abstract error groups on $\mathbb{C}^d$ for $1 \leq d \leq 11$ is given in [KR02]. Index groups of abstract error groups are in general not abelian: in [KR02] a family of groups having non-abelian index groups is constructed. It is known that all abstract error groups are solvable. Moreover, it is known that all solvable groups can occur as *subgroups* of index groups of nice error bases. On the other hand, it is known that not all solvable groups can occur as index groups.

Using the concept of nice error bases we describe the idea of switching off an interaction by averaging over a group. Whereas usual techniques are based on ordinary irreducible representations (cf. Zanardi [Zan99] and Viola et al.

[VKL99]), the following lemma shows that averaging over a projective irreducible representation also projects onto the set of scalar matrices.

Lemma 3.9 *Let $M \in \mathbb{C}^{d\times d}$, G be a finite group, and $R : g \mapsto U_g \in \mathcal{U}(d)$ an irreducible projective representation of G. Then the following equation holds:*

$$\frac{1}{|G|}\sum_{g\in G} U_g^\dagger M U_g = \frac{\mathrm{tr}(M)}{d}\,\mathbf{1}$$

Proof. We have seen that each projective representation of an index group G with associated factor system α gives rise to an ordinary representation of the corresponding abstract error group $\hat{G}$ and that $\hat{G}$ is a central extension of G. It follows that $\{U_g : g \in G\}$ is a set of coset representatives for $\zeta(\hat{G})$ in $\hat{G}$, i. e.,

$$\hat{G} = \bigcup_{g\in G} \zeta(\hat{G})U_g.$$

Each element $\sigma \in \hat{G}$ has a unique factorization of the form $\sigma = zg$ where $z \in \zeta(\hat{G})$ and $g \in G$. From Schur's Lemma (cf. Serre [Ser77, Section 2.2]) follows that for $M \in \mathbb{C}^{d\times d}$, $\hat{G}$ a finite group, and $R : \sigma \mapsto U_\sigma \in \mathcal{U}(d)$ an irreducible (ordinary) representation of $\hat{G}$ the following identity holds:

$$\frac{1}{|\hat{G}|}\sum_{\sigma\in\hat{G}} U_\sigma^\dagger M U_\sigma = \frac{\mathrm{tr}(M)}{d}\,\mathbf{1}. \tag{3.7}$$

Using this we obtain

$$\begin{aligned}
\frac{1}{|G|}\sum_{g\in G} U_g^\dagger M U_g &= \frac{1}{|G|}\frac{1}{|\zeta(\hat{G})|}\sum_{g\in G}\sum_{z\in\zeta(\hat{G})} U_g^\dagger U_z^\dagger M U_z U_g \\
&= \frac{1}{|\hat{G}|}\sum_{\sigma\in\hat{G}} U_\sigma^\dagger M U_\sigma \\
&= \frac{\mathrm{tr}(M)}{d}\,\mathbf{1},
\end{aligned}$$

where the last line follows from Schur's Lemma (3.7) for ordinary representations. □

3.2 Decoupling schemes

We now turn to quantum networks of n coupled qudits. We derive sufficient conditions on how to distribute control operations of annihilators on the nodes

in order to decouple general qudit systems. This problem is solved efficiently by using combinatorial concepts orthogonal arrays and difference schemes.

The idea to use orthogonal arrays for decoupling qubit networks was made by Stollsteimer and Mahler (cf. [SM01]). We showed in Wocjan et al. [WRJB02a] that this idea extends to qudit systems by using minimal annihilators. In Chapter 5 we prove that these schemes are optimal with respect to the number of steps.

Let us first introduce the general definition of decoupling schemes:

Definition 3.10 (General decoupling scheme)
A decoupling scheme for a system consisting of n qudits is a control sequence $C := (\tau_1, U_1, \tau_2, U_2, \ldots, \tau_N, U_N)$ such that

$$C(H) = \mathbf{0} \tag{3.8}$$

for all pair-interaction Hamiltonians H of n qudits. We call N the number of time steps.

In the following we construct decoupling schemes with equally long time steps. We first consider the case of two coupled qudits. The general case can be reduced to this case since the Hamiltonians contain only pair-interactions (see Definition 2.2).

The general Hamiltonian of a bipartite system has the form

$$H = \sum_{\alpha\beta} J_{\alpha\beta}\, \sigma_\alpha \otimes \sigma_\beta + \sum_{\gamma} r_\gamma\, \mathbf{1} \otimes \sigma_\gamma + \sum_{\delta} s_\delta\, \sigma_\delta \otimes \mathbf{1}\,. \tag{3.9}$$

Let $\mathcal{E}_1 = \{U_i\}_i$ and $\mathcal{E}_2 = \{V_i\}_i$ be error bases of the respective systems and let $\mathcal{A}$ denote the alphabet $\{1, \ldots, d^2\}$. By applying the annihilators defined by $\mathcal{E}_1$ and $\mathcal{E}_2$ independently on the nodes we switch off the Hamiltonian, i.e.

$$\frac{1}{|\mathcal{A}^2|} \sum_{(i,j)\in\mathcal{A}^2} (U_i^\dagger \otimes V_j^\dagger)\, H\, (U_i \otimes V_j) = \mathbf{0}\,. \tag{3.10}$$

We now define decoupling schemes based on minimal annihilators.

Definition 3.11 (Decoupling scheme)
Let $\mathcal{E} := \{U_1, \ldots, U_{d^2}\}$ be an arbitrary error basis. Denote by $\mathcal{A}$ the alphabet $\{1, \ldots, d^2\}$. A decoupling scheme is characterized by $n \times N$-matrix $M = (m_{ij})$ over $\mathcal{A}$ such that

$$\frac{1}{N} \sum_{j=1}^{N} (U_{m_{1j}}^\dagger \otimes \cdots \otimes U_{m_{nj}}^\dagger)\, H\, (U_{m_{1j}} \otimes \cdots \otimes U_{m_{nj}}) = \mathbf{0} \tag{3.11}$$

for all pair-interaction Hamiltonians H.

The entries of the matrix specify the conjugations of the nodes in the time intervals. The N time intervals correspond to the columns and the n nodes correspond to different rows. For instance, the decoupling scheme corresponding to eq. (3.10) is given by the matrix

$$\left(\begin{array}{cccc|cccc|c|cccc} 1 & 1 & \dots & 1 & 2 & 2 & \dots & 2 & \dots & N & N & \dots & N \\ 1 & 2 & \dots & N & 1 & 2 & \dots & N & \dots & 1 & 2 & \dots & N \end{array}\right). \tag{3.12}$$

We now turn to the problem of constructing decoupling schemes. The simplest approach for decoupling is to choose the columns of M as all tuples of $\mathcal{A}^n$. The reason is that if $\mathcal{E} := \{U_1, \dots, U_{d^2}\}$ is an error basis of $\mathbb{C}^{d\times d}$ then

$$\mathcal{E}^{\otimes n} := \{U_{i_1} \otimes \cdots \otimes U_{i_{d^2}} \mid (i_1, \dots, i_{d^2}) \in \mathcal{A} \times \cdots \times \mathcal{A}\} \tag{3.13}$$

is an error basis of $\mathbb{C}^{d^n\times d^n}$. This scheme is not efficient in terms of the number of time steps. The number of time steps scales exponentially as d^{2n} because the sequence has to be repeated d^2 times for each added node.

It is clear that more efficient schemes can be constructed if we take into account that the Hamiltonians contain only pair-interactions and are not arbitrary elements in $su(d^n)$. These constructions rely on orthogonal arrays.

3.2.1 Orthogonal arrays

We refer the reader to Beth et al. [BJL99], Colbourn and Dinitz [CD96] and Hedayat et al. [HSS99] for the general theory of orthogonal arrays. Orthogonal arrays have numerous applications e.g. in the design of experiments. There are also connections between orthogonal arrays and mutually orthogonal Latin squares and transversal designs (cf. [BJL99, Section VIII]). The following definition takes account of the fact that for purposes of decoupling we need a special type of orthogonal arrays, namely those of strength $t = 2$. Also the notation used is in this thesis is adapted to this situation.

Definition 3.12 (Orthogonal arrays of strength 2)
Let $\mathcal{A}$ be a finite alphabet and let $n, N \in \mathbb{N}$. An $n \times N$ array M with entries from $\mathcal{A}$ is an orthogonal array with $|\mathcal{A}|$ levels, strength $t = 2$, and index λ if and only if each pair of elements of $\mathcal{A}$ occurs λ times in the list $((m_{kj}, m_{lj}) \mid j = 1, \dots N)$ for $1 \le k < l \le n$. We use the notation $OA_\lambda(n, N)$ to denote a corresponding orthogonal array.[2]

The following theorem shows that decoupling in networks of arbitrary dimensions can be achieved using control sequences obtained from orthogonal arrays.

[2]Note that in [BJL99] the notation $OA_\lambda(n, s)$ is used for an orthogonal array with $N = \lambda s^2$ in our notation, where $s := |\mathcal{A}|$.

Theorem 3.13 (Decoupling with OAs)
Let $\mathcal{A}$ be the finite alphabet $\{1,\ldots,d^2\}$. Then any orthogonal array with parameters $OA_\lambda(n,N)$ over $\mathcal{A}$ can be used to decouple a quantum network consisting of n nodes of dimension d. The number of local operations used in this scheme is given by N.

Proof. Let H be the (unknown) pair-interaction Hamiltonian

$$H = \sum_{k<l} H_{kl} + \sum_k H_k\,.$$

Let $M = (m_{kj})$ be the $n \times N$-matrix over $\mathcal{A}$ of the orthogonal array $OA(n,N)$. Choose an arbitrary error basis $\mathcal{E} := \{U_1,\ldots,U_{d^2}\}$. By $U_i^{(k)}$ we denote the ith matrix of $\mathcal{E}$ acting on the kth node.
Let H_k be the free Hamiltonian of the kth node. Then we have

$$\begin{aligned}
&\frac{1}{N}\sum_{j=1}^{N}(U_{m_{1j}}^\dagger \otimes\cdots\otimes U_{m_{nj}}^\dagger)\,H_k\,(U_{m_{1j}}\otimes\cdots\otimes U_{m_{nj}})\\
=\;& \frac{1}{N}\sum_{j=1}^{N} U_{m_{kj}}^{(k)\dagger}\,H_k\,U_{m_{kj}}^{(k)}\\
=\;& \frac{1}{|\mathcal{A}|}\sum_{i\in\mathcal{A}} U_i^{(k)\dagger}\,H_k\,U_i^{(k)} = \mathbf{0}\,.
\end{aligned}$$

The equality of the last two lines follows from the fact that the every row of an orthogonal array contains all elements of $\mathcal{A}$ exactly λ times. The sum in the last equality is zero since the control sequence is an annihilator and thus cancels H_k.
Now we turn to the couplings. Let H_{kl} be the interaction between the nodes k and l. Then we have

$$\begin{aligned}
&\frac{1}{N}\sum_{j=1}^{N}(U_{m_{1j}}^\dagger \otimes\cdots\otimes U_{m_{nj}}^\dagger)\,H_{kl}\,(U_{m_{1j}}\otimes\cdots\otimes U_{m_{nj}})\\
=\;& \frac{1}{N}\sum_{j=1}^{N} U_{m_{kj}}^{(k)\dagger}U_{m_{lj}}^{(l)\dagger}\,H_{kl}\,U_{m_{kj}}^{(k)}U_{m_{lj}}^{(l)}\\
=\;& \frac{1}{|\mathcal{A}|^2}\sum_{(i,i')\in\mathcal{A}^2} U_i^{(k)\dagger}U_{i'}^{(l)\dagger}\,H_{kl}\,U_i^{(k)}U_{i'}^{(l)} = \mathbf{0}\,.
\end{aligned}$$

The equality between the last to lines follows from the property that each pair of elements of $\mathcal{A}$ occurs precisely λ times in the list $((m_{kj},m_{lj}) \mid j=1,\ldots,N)$. The last sum is equal to zero since both (resulting) annihilators are applied independently on both nodes. □

For any given number $n \in \mathbb{N}$ of nodes there are parameters λ, N such that an orthogonal array $OA(n, N)$ exists. However, since we are interested in efficient schemes, N has to be a polynomial in the number n of nodes. Also it is of interest to give explicit constructions of such schemes, i.e. of orthogonal arrays. Whereas little is known about the existence of efficient schemes for general n and alphabet sizes s the situation is much better in the case when s is a prime power. We have the following two theorems that give explicit constructions.

Theorem 3.14 (Construction of orthogonal arrays)
Rao-Hamming: If s is a prime power then an orthogonal array

$$OA((s^i - 1)/(s - 1), s^i)$$

exists whenever $i \geq 2$.
Addelman and Kempthorne: If s is an odd prime power then an orthogonal array

$$OA(2(s^i - 1)/(s - 1) - 1, 2s^i)$$

exists whenever $i \geq 2$.

Proof. cf. [HSS99, Theorem 3.20 and 3.16]. □

Corollary 3.15 *Consider a quantum network of n coupled nodes. Assume that the dimension d of each node is a prime power. Then there exists a decoupling scheme for this quantum network using N local operations, where $N \leq cn$ and c is a constant depending only on d.*

Proof: Let $s := d^2$ be the size of a minimal annihilator for a d-dimensional system. In view of Theorem 3.13 we have to show that there exists an orthogonal array with parameters $OA_\lambda(\tilde{n}, N)$ with $n \leq \tilde{n}$ and $N \leq cn$ as above. The Rao-Hamming construction gives an explicit method for constructing for an $OA((s^i - 1)/(s-1), s^i)$ for any $i \geq 2$. Hence, if for the number of nodes $n = (s^i - 1)/(s - 1)$ holds, we have found a decoupling scheme with $N = s^i = (s-1)n+1$ operations, i. e., $N = O(n)$. For general n we embed into an OA of this form. Switching to the next number of the form $(s^i - 1)/(s - 1)$ with suitable $i \geq 1$ can be achieved by multiplying n with a number less or equal s, i. e., $\tilde{n} \leq sn$. □

Remark 3.16 There are tables of orthogonal arrays covering the small instances (cf. [BJL99, CD96, HSS99]). The Addelman and Kempthornes Construction gives a family of OAs with parameters $OA(2\frac{s^i-1}{s-1} - 1, 2s^i)$. This shows that the constant c in Corollary 3.15 can be improved to $c/2$.

The following example illustrates that OAs give more efficient schemes already for small systems.

Example 3.17 We consider the case of four three-level systems. Using the exponential scheme we need $s^4 = 6561$ local operations to decouple all interactions. Following Corollary 3.15 we obtain a decoupling scheme with the same property that uses only $s^2 = 81$ local operations. This scheme we can decouple even up to ten three-level systems.

The selective decoupling scheme presented above generalizes straightforwardly to the case that the dimensions of the n subsystems do not agree. Then one has to use so-called *mixed orthogonal arrays* , i.e., one has different alphabets for different nodes. Although little is known about constructions of efficient mixed orthogonal arrays, it is known that exponential ones exist (cf. [HSS99, Section 9.3]).

3.2.2 Difference schemes

We turn to a special situation for quantum networks consisting of qubits and having only *diagonal* couplings.

Definition 3.18 (Diagonal coupling)
We call a coupling diagonal if it is of the form

$$J_x \, \sigma_x \otimes \sigma_x + J_y \, \sigma_y \otimes \sigma_y + J_z \, \sigma_z \otimes \sigma_z \, , \tag{3.14}$$

where $\sigma_x, \sigma_y, \sigma_z$ are Pauli matrices and $J_x, J_y, J_z \in \mathbb{R}$.

Diagonal couplings encompass important couplings that occur in spin systems. Examples are the *strong scalar coupling* with $J_x = J_y = J_z := 1$, the *weak scalar coupling* $J_x = J_y = 0$ and $J_z = 1$, and the *dipolar coupling* with $J_x = J_y = -1$ and $J_z = 2$ (cf. Luy and Glaser [LG01]).

The fact that all couplings are diagonal allows us to construct more efficient schemes than with orthogonal arrays. This has been shown in Stollsteimer and Mahler [SM01]. In Chapter 5 we will show the optimality of these schemes.

The nice error basis $\mathcal{E} = \{\mathbf{1}, \sigma_x, \sigma_y, \sigma_z\}$ has as index group the abelian group $Z_2 \times Z_2 = \{(0,0), (0,1), (1,0), (1,1)\}$ as index group. The components are added modulo 2. This allows us to derive a sufficient condition for decoupling.

Lemma 3.19 (Sufficient condition for decoupling)
Let D be an $n \times N$ matrix with entries in $Z_2 \times Z_2$. If the vector of the element-wise differences of any two rows contains each element of $Z_2 \times Z_2$ equally often, then D can be used to decouple networks consisting of n qubits with diagonal couplings within N time steps.

Proof. Note that we have

$$\sigma_h \sigma_g \sigma_h = \begin{cases} \sigma_g & \text{if } g = (0,0) \text{ or } h = g \\ -\sigma_g & \text{if } g \neq (0,0) \text{ and } h \neq g \, . \end{cases}$$

Furthermore, we have

$$\begin{aligned}(\sigma_h \otimes \sigma_{h'})\,\sigma_g \otimes \sigma_g\,(\sigma_h \otimes \sigma_{h'}) &= (\sigma_{h+h'} \otimes \mathbf{1})\,\sigma_g \otimes \sigma_g\,(\sigma_{h+h'} \otimes \mathbf{1}) \\ &= (\mathbf{1} \otimes \sigma_{h+h'})\,\sigma_g \otimes \sigma_g\,(\mathbf{1} \otimes \sigma_{h+h'})\,.\end{aligned}$$

Now if we apply a scheme specified by D then the terms of every diagonal pair-interaction in eq. (3.14) acquire in exactly half of the time intervals a minus sign. □

The sequences based on orthogonal arrays fulfill automatically this condition, but there exist suitable matrices with a smaller number of columns, leading to shorter decoupling sequences. Such matrices are a special case of difference schemes. Difference schemes play an important role in constructing orthogonal arrays (cf. [BJL99, CD96, HSS99]).

Definition 3.20 (Difference scheme)
An $n \times N$ array $D_{\mathcal{A}}(n, N)$ with entries from $\mathcal{A}$ is called a difference scheme based on $(\mathcal{A}, +)$ if for all k, l with $1 \leq k < l \leq n$ the vector difference between the kth and lth rows contains every element of $\mathcal{A}$ equally often.

There are a lot of constructions known for difference schemes over additive groups of finite fields (cf. [HSS99]). As we have seen in Lemma 3.19 the relevant group is the the group $Z_2 \times Z_2$ that is the additive group of the finite field $GF(4)$. We obtain the following theorem.

Theorem 3.21 (Decoupling with difference schemes)
A difference scheme $D(n, N)$ over the finite field $(GF(4), +)$ defines a decoupling scheme for an n-qubit network with diagonal couplings within N time steps.

For difference schemes $D(n, N)$ one has $n \leq N$, so that at most N qubits can be decoupled with N time steps.

Corollary 3.22 *Let an n-qubit network with diagonal couplings be given. Then there exists a decoupling scheme based on difference schemes using $N = cn$ local operations, where $1 \leq c < 2$.*

Proof. A difference scheme $D(p^s, p^s)$ exists over the finite field $GF(p^r)$ if $s \geq r \geq 1$ [HSS99, Theorem 6.6]. For $p = 2$ and $r = 2$ we obtain the required difference scheme. □

The advantage of decoupling schemes based on difference schemes is that the factor c is smaller than the factor that occurs for decoupling schemes based orthogonal arrays (see Corollary 3.15 and Remark 3.16).

3.2.3 Hadamard matrices

We have seen that restricting to diagonal couplings permits to use more efficient decoupling schemes. The decoupling schemes can be improved even further if we consider a special case of diagonal coupling, namely

$$\sum_{k<l} J_{kl}\sigma_z^{(k)}\sigma_z^{(l)}\,. \tag{3.15}$$

In this case it is sufficient to use Hadamard matrices instead of difference schemes over $GF(4)$. The Hadamard matrices are difference schemes over $GF(2)$. The construction of decoupling schemes based on Hadamard matrices has been presented in Leung et al. [LCYY00].

We will need these schemes for showing that the lower bounds on time overhead and number of time steps derived in Chapter 5 are tight.

Definition 3.23 (Hadamard matrix)
A Hadamard matrix of order N is an $N \times N$ matrix H_N of $+1$'s and -1's whose rows are orthogonal, i.e. which satisfies

$$H_N H_N^T = N\mathbf{1}_N\,.$$

When do Hadamard matrices exist? The answer is perhaps surprising: it is not known. It is easy to show that if an H_N exists then N is $1, 2$ or a multiple of 4 (cf. [HSS99], Corollary 7.2), and it is almost certainly true that if N is a multiple of 4 then an H_N exists. However, this assertion, known as the *Hadamard conjecture*, is a basic unsolved problem in discrete mathematics [HSS99].

Theorem 3.24 (Decoupling with Hadamard matrices)
Let $n \leq N$. A Hadamard matrix H_N gives a decoupling scheme for an n-qubit network with only zz-couplings with N time steps.

Proof. Note that we have

$$\begin{aligned} -\sigma_z \otimes \sigma_z &= (\sigma_x \otimes \mathbf{1})\,\sigma_z \otimes \sigma_z\,(\sigma_x \otimes \mathbf{1}) \\ &= (\mathbf{1} \otimes \sigma_x)\,\sigma_z \otimes \sigma_z\,(\mathbf{1} \otimes \sigma_x) \end{aligned}$$

and

$$\sigma_z \otimes \sigma_z = (\sigma_x \otimes \sigma_x)\,\sigma_z \otimes \sigma_z\,(\sigma_x \otimes \sigma_x)\,.$$

The value -1 corresponds to conjugation with σ_x and $+1$ correspond to doing nothing. The entry at position (k, j) of H_N tells whether or not to perform a σ_x-conjugation on the kth qubit in the jth time step. Since the kth and lth columns are orthogonal, the coupling $\sigma_z^{(k)} \otimes \sigma_z^{(l)}$ acquires in exactly half of the time steps the $-$ sign. Consequently, the coupling between k and l is cancelled. □

Due to the existence of a special type Hadamard matrices we have the following corollary.

Corollary 3.25 *Let an n-qubit network with zz-couplings be given. Then there exists a decoupling scheme based on Hadamard matrices with N time steps, where $n \leq N < 2n$.*

Proof. There exists for every power $N := 2^i$ a Hadamard matrix. Let

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

be a Hadamard matrix of size 2. Then the tensor products

$$H_{2^i} := \underbrace{H_2 \otimes \cdots \otimes H_2}_{i \text{ times}}$$

are Hadamard matrices of size 2^i. □

There are other types of constructions for Hadamard matrices (cf. HEDAYAT ET AL. [HSS99]).

3.3 Equivalence of decoupling schemes

We have seen that orthogonal arrays may be used to construct decoupling schemes for quantum networks with arbitrary pair-interactions. There is a different approach for constructing decoupling schemes for qubit networks with general couplings (cf. LEUNG [Leu02]). We extend Leung's approach to coupled qudits with the help nice error basis with abelian index groups. Then we show that both approaches lead to the same class of decoupling schemes. The presentation is based on our work WOCJAN AND BETH [WB02]. Furthermore, we show that one of the constructions in [Leu02] can also be explained using the combinatorial notion of *maximal spreads*. This result is based on our work WOCJAN ET AL. [WRJB02a].
In the following we explain briefly the approach of [Leu02] for decoupling general qubit networks. The system Hamiltonian has the form

$$H = \sum_{kl;\alpha\beta} J_{kl;\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} + \sum_{k;\alpha} r_{k;\alpha} \sigma_\alpha^{(k)} , \tag{3.16}$$

where σ_α ($\alpha = x, y, z$) are the Pauli matrices. The control operations are assumed to be the identity or one of the Pauli matrices. Then, in each time step, each $\sigma_\alpha^{(k)}$ acquires either a $+$ or $-$ sign, which is controlled by the applied control operation. Therefore, the coupling $J_{kl;\alpha\beta}\sigma_\alpha^{(k)}\sigma_\beta^{(l)}$ is unchanged (negated) if the

signs of $\sigma_\alpha^{(k)}$ and $\sigma_\beta^{(l)}$ agree (disagree). Note that the signs of the three Pauli matrices σ_α^k acting on the same qubit k are not independent; they must multiply to $+$. Therefore, the only possible signs for $\sigma_x^{(k)}, \sigma_y^{(k)}, \sigma_z^{(k)}$ are given by the triples $(+++), (+--), (+-+), (--+)$, and are realized by applying $\mathbf{1}^{(k)}, \sigma_x^{(k)}, \sigma_y^{(k)}, \sigma_z^{(k)}$, respectively, before and after the time step.

Based on these observations a decoupling scheme for n qubits with N time steps is specified in [Leu02] by three $n \times N$ sign matrices S_x, S_y, S_z, related by the entry-wise product $S_x * S_y = S_z$. We say that these three matrices satisfy the Schur condition since this entry-wise product is usually called the Schur product (and also Hadamard product). The (k, j) entry of S_α is the sign of $\sigma_\alpha^{(k)}$ in the jth time step.

Within this framework sufficient and necessary conditions for decoupling are:

$$\sum_{j=1}^{N} S_{\alpha;kj} = 0 \tag{3.17}$$

for all α and all k, and

$$\sum_{j=1}^{N} S_{\alpha;kj} S_{\beta;lj} = 0 \tag{3.18}$$

for all α, β and all $k < l$. The first condition ensures that the local terms are removed. The second condition ensures that the coupling terms are removed. Both conditions may also be expressed as follows: any two rows taken from S_x, S_y, S_z are orthogonal and all row sums are zero.

3.3.1 Generalization to qudit networks

Now we generalize the above decoupling conditions to coupled qudits. To do this we use nice error basis with abelian index group.

Let $\mathcal{E} = \{U_g | g \in G\}$ be a nice error basis on $\mathbb{C}^d$ with abelian index group G. Since the matrices U_g form a basis of $\mathbb{C}^{d\times d}$ every pair-interaction Hamiltonian on n coupled qudits may be written as

$$H := \sum_{k<l} \sum_{h,h' \neq e} J_{kl;hh'} U_h^{(k)} U_{h'}^{(l)} + \sum_{k} \sum_{h \neq e} r_{k;h} U_h^{(k)}\,, \tag{3.19}$$

where the coefficients $J_{kl;hh'} \in \mathbb{C}$ and $r_{k;h} \in \mathbb{C}$ are chosen such that H is a traceless Hermitian matrix.

> **Decoupling scheme:** We consider decoupling schemes with time steps of equal length and with elements of $\mathcal{E}$ (a nice error basis with

abelian index group G) as control operations. In this case, a decoupling scheme may be described by a matrix $M = (g_{kj})$ of size $n \times N$ matrix over G as follows:

$$\sum_{j=1}^{N} (U_{g_{1j}}^{\dagger} \otimes U_{g_{2j}}^{\dagger} \otimes \cdots \otimes U_{g_{nj}}^{\dagger}) H (U_{g_{1j}} \otimes U_{g_{2j}} \otimes \cdots \otimes U_{g_{nj}}) \qquad (3.20)$$

The columns correspond to the time steps and the rows to the nodes.

Analogously to the qubit case, in each time step, each U_h in Eq. (3.19) acquires a phase factor that is controlled by the applied local unitaries of the nice error basis. We define $\chi(g, h)$ to be the phase factor that U_h acquires when it is conjugated by U_g, i.e. $\chi(g, h)$ is defined via the relation

$$U_g^{\dagger} U_h U_g = \chi(g, h) U_h \,. \qquad (3.21)$$

The $d^2 \times d^2$ matrix with entries $\chi(g, h)$ is denoted by X. We will call it the *phase matrix*. The coupling term $U_h \otimes U_{h'}$ acquires the phase factor $\chi(g, h)\chi(g', h')$ if it is conjugated by $U_g \otimes U_{g'}$.

Now we can state the necessary and sufficient criteria for decoupling within this framework.

Criteria for decoupling: Let G be an abelian index group and $M = (g_{kj})_{k=1,\ldots,n;\, j=1\ldots,N}$ be a $n \times N$ matrix over G. Decoupling is achieved if and only if

$$\sum_{j=1}^{N} \chi(g_{kj}, h) = 0 \qquad (3.22)$$

for all k and for all $h \neq e$, and

$$\sum_{j=1}^{N} \chi(g_{kj}, h)\chi(g_{lj}, h') = 0 \qquad (3.23)$$

for all $k < l$ and for all (h, h') with $h, h' \neq e$.

Condition (3.22) ensures that all local terms are removed and Condition (3.23) that all coupling terms are removed.

Remark 3.26 *The starting point for the generalization from the qubit case with Pauli matrices to the qudit case with nice error basis with abelian index was the following simple observation. The triples that give the possible sign assignments*

to the Pauli matrices are a part (the last three columns) of the character table of $Z_2 \times Z_2$

$$\left|\begin{array}{cccc} + & + & + & + \\ + & + & - & - \\ + & - & + & - \\ + & - & - & + \end{array}\right| .$$

This is because $Z_2 \times Z_2$ *is the abelian index group for the error basis* $\{\mathbf{1}, \sigma_x, \sigma_y, \sigma_z\}$.

3.3.2 Proof of equivalence

Now we show that the decoupling conditions (3.22) and (3.23) are equivalent to the condition that the decoupling matrix is an orthogonal array.

To prove this we need the following theorem for abelian groups [Hup83, Chap. V. §6].

Theorem 3.27 (Characters of abelian groups)
Let G *be an abelian group.*

1. *Every irreducible representation* ρ *of* G *has degree* 1*, i.e. we have* $\rho : G \to \mathbb{C}^\times$*. Furthermore, the number of different irreducible representations (irreducible characters) of* G *is* $|G|$.

2. *The characters form a group* $\hat{G} = \mathrm{Hom}(G, \mathbb{C}^\times)$ *under pointwise multiplication,*

$$\chi\tilde{\chi}(h) = \chi(h)\tilde{\chi}(h)$$

for all irreducible characters $\chi, \tilde{\chi}$ *and* $h \in G$*. The character group* $\hat{G}$ *is isomorphic to* G.

Now we show that the phase matrix is a character table of the index group G.

Lemma 3.28
Let $\mathcal{E} := \{U_g \mid g \in G\}$ *be a nice error basis with an abelian index group* G*. Then the corresponding phase matrix* X *is a character table of the group* G.

Proof. Let α be the factor system corresponding to the nice error basis $\mathcal{E}$ with abelian index group G. We prove that X is a character table by showing that the rows of X form a group under pointwise multiplication that is isomorphic to G (see the theorem above).

We first show that

$$\chi(g, h) = \frac{\alpha(h, g)}{\alpha(g, h)} . \tag{3.24}$$

We have

$$U_g U_h = \alpha(g,h) U_{gh} \tag{3.25}$$
$$U_h U_g = \alpha(h,g) U_{hg} = \alpha(h,g) U_{gh}\,. \tag{3.26}$$

By multiplying Eq. (3.26) by $U_g^\dagger$ from the left and using Eq. (3.25) we obtain

$$\begin{aligned} U_g^\dagger U_h U_g &= \alpha(h,g) U_g^\dagger U_{gh} \\ &= \frac{\alpha(h,g)}{\alpha(g,h)} U_g^\dagger U_g U_h \\ &= \frac{\alpha(h,g)}{\alpha(g,h)} U_h\,. \end{aligned}$$

We now prove that the rows of X form a group under pointwise multiplication that is isomorphic to G. Let $g, \tilde{g}$ be arbitrary elements of G. Note that we have $\overline{\alpha(\tilde{g}^{-1},g)}\alpha(\tilde{g}^{-1},g) = 1$ (otherwise the matrix $U_{\tilde{g}^{-1}} U_g = \alpha(\tilde{g}^{-1},g) U_{\tilde{g}^{-1}g}$ would not be unitary). The group property is verified by

$$\begin{aligned} \chi(g,h)\chi(\tilde{g}^{-1},h) U_h &= U_g^\dagger U_{\tilde{g}^{-1}}^\dagger U_h U_{\tilde{g}^{-1}} U_g \\ &= \overline{\alpha(\tilde{g}^{-1},g)}\alpha(\tilde{g}^{-1},g)\, U_{g\tilde{g}^{-1}}^\dagger U_h U_{g\tilde{g}^{-1}} \\ &= U_{g\tilde{g}^{-1}}^\dagger U_h U_{g\tilde{g}^{-1}} \\ &= \chi(g\tilde{g}^{-1},h) U_h \end{aligned}$$

for all $h \in G$.

The rows of X form a group that is isomorphic to G (and not only to a proper subgroup of G) since there is a bijection between the rows of X and the elements of G. This is seen as follows. Assume that there are $g \neq \tilde{g}$ such that $\chi(g,h) = \chi(\tilde{g},h)$ for all $h \in G$. This is equivalent to $U_g^\dagger U_h U_g = U_{\tilde{g}}^\dagger U_h U_{\tilde{g}}$. Set $U = U_{\tilde{g}} U_g^\dagger$. Then we have $UM = MU$ for all $M \in \mathbb{C}^{d\times d}$ since the matrices U_h for a basis of $\mathbb{C}^{d\times d}$. Therefore U must be a multiple of the identity matrix. Due to the properties of a nice error basis this is only possible for $g = \tilde{g}$. This proves that there is a bijection between the group elements of G and the rows of X. □

The next lemma provides a criterion in terms of group characters that shows whether all group elements appear equally often. This allows us to check whether a matrix is an orthogonal array.

Lemma 3.29
Let G be an abelian group. Denote by $\chi_1, \chi_2, \ldots, \chi_k$ all irreducible characters of G, where χ_1 is the trivial character (i.e. $\chi_1(h) = 1$ for all $h \in G$). Let v be an arbitrary element of the group ring $\mathbb{C}[G]$, i.e. v is a formal sum of (weighted) group elements

$$v := \sum_{g\in G} \mu_g g\,, \quad \mu_g \in \mathbb{C}\,. \tag{3.27}$$

If $\chi_i(v) = 0$ for all $i = 2, \ldots, k$ then we have

$$v = \frac{\mu}{|G|} \sum_{g \in G} g \,, \tag{3.28}$$

where $\mu := \chi_1(v) = \sum_{g \in G} \mu_g$.

Proof. Let $G := \{g_1, \ldots, g_k\}$ be an arbitrary ordering of the group elements, where g_1 is the identity element of G. Denote by X the (normalized) character table of G, i.e.

$$X_{ij} := |G|^{-1/2} \chi_i(g_j) \tag{3.29}$$

for $i, j = 1, \ldots, k$. Recall that the (normalized) character table X is a unitary matrix and has the following form

$$X = \frac{1}{|G|^{1/2}} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & * & \\ 1 & & \end{pmatrix} \tag{3.30}$$

The conditions above can be expressed as

$$|G|^{1/2} X \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

Multiplying by the inverse X^{-1} we obtain

$$(\mu_1, \mu_2, \ldots, \mu_k)^T = \frac{\mu}{|G|} (1, 1, \ldots, 1)^T \,.$$

due to the special form in Eq. (3.30). This show that all coefficients μ_g in Eq. (3.27) are equal to $\mu/|G|$. □

Theorem 3.30 (Equivalence of decoupling schemes)
Let $\mathcal{E}$ be a nice error basis with an abelian index group G. An $n \times N$ matrix M over G defines a decoupling scheme (using the matrices of $\mathcal{E}$ as control operations) if and only if M is an orthogonal array $OA(n, N)$ over G of strength 2.

Proof. Pick any two rows (g_{kj}) and (g_{lj}) of M. We define an element of the group ring $\mathbb{C}[G \times G]$ as the formal sum

$$v_{kl} := \sum_{j=1}^{N} (g_{kj}, g_{lj}) \,.$$

To abbreviate the notation we denote by χ_g the map $\chi(g,\cdot)$ and by $\chi_{g,g'}$ the map $\chi(g,\cdot)\chi(g',\cdot)$. Lemma 3.28 shows that χ_g are all irreducible characters of G. Consequently, $\chi_{g,g'}$ are all irreducible characters of $G \times G$.

The decoupling conditions (3.22) and (3.23) are equivalent to

$$\chi_{g,g'}(v_{kl}) = 0$$

for all $(g,g') \neq (e,e)$. By Lemma 3.29 this is equivalent to the case that all elements of $G \times G$ appear equally often in the sum v_{kl}. This shows that M is an orthogonal array $OA(n,N)$ of strength 2 over G. □

Note that our reasoning can be generalized in a straightforward way for strengths greater than 2.

3.3.3 Construction of Schur sets based on spreads

Sign matrices S_x, S_y, S_z satisfying the decoupling criteria can be constructed from special Hadamard matrices endowed with certain extra structures. Suppose we want to decouple n qubits using a Hadamard matrix H_M. The orthogonality condition is automatically satisfied if the rows of S_x, S_y, S_z are taken to be distinct rows of H_M. It remains to ensure that $S_x * S_y = S_z$. A set of three vectors of equal length and with entries ± 1 is called a Schur-set if they entry-wise multiply to $+ + \cdots +$. For example, $\{[+--],[-+-],[--+]\}$ is a Schur-set. If H_M has at least $3n$ rows that partition into n Schur-subsets, one can ensure $S_x * S_y = S_z$ by choosing the ith rows of S_x, S_y, S_z to be the rows of the ith Schur-subset. This poses the first extra property on H_M – its rows partition into many Schur-subsets. The immediate lower bound on the size of H_M is $\lfloor (M-1)/3 \rfloor \geq n$ under this construction.

By considering Hadamard matrices of Sylvester type (i.e. of size 2^r), Leung [Leu02] obtained two families of decoupling schemes:

Lemma 3.31 (Decoupling schemes with Sylvester matrices)
The rows of the Sylvester matrix H_{2^r} can be partitioned into $(2^r-1)/3$ and $(2^r-5)/3$ Schur-subsets when r is even and odd, respectively.

We give an alternative proof for the existence of a decoupling scheme for $n = (2^{2m}-1)/3$ qubits using $N = 2^{2m}$ time intervals. A decoupling scheme with these parameters can be constructed using orthogonal arrays [SM01] and Hadamard matrices [Leu02].

Let V be the vector space $\mathbb{F}_4^m$, where $m \geq 1$ and let $\mathbb{F}_4 = \{0, 1, \omega, \omega^2 = 1+\omega\}$, where $\omega^3 = 1$, be the Galois field with 4 elements. Recall that the number of

d-dimensional subspaces of an m-dimensional vector space over $\mathbb{F}_q$ is given by

$$\left[\begin{array}{c} m \\ d \end{array} \right]_q := \frac{(q^m-1)(q^{m-1}-1)\cdots(q^{m-d+1}-1)}{(q^d-1)(q^{d-1}-1)\cdots(q-1)} \tag{3.31}$$

(cf. [BJL99, Lemma 2.14, Section I]). For the special case $q = 4$ and $d = 1$ formula (3.31) shows that there are $(4^m-1)/(4-1) = (2^{2m}-1)/3$ lines in $\mathbb{F}_4^m$. Note that different lines intersect in the point $\{0\}$ only. Hence by taking the set of all one-dimensional subspaces of $\mathbb{F}_4^m$ we obtain a maximal spread in $\mathbb{F}_4^m$, i. e. a collection of subspaces U_i partitioning $\mathbb{F}_4^m$ with the additional property that

$$U_i \cap U_j = \{0\}\,.$$

We define a map φ from $\mathbb{F}_4$ onto $\{-1,+1\}^4$ as follows:

$$\begin{aligned} \varphi(0) &= (+1,+1,+1,+1) \\ \varphi(\omega) &= (+1,-1,+1,-1) \\ \varphi(\omega^2) &= (+1,+1,-1,-1) \\ \varphi(1) &= (+1,-1,-1,+1) \end{aligned}$$

These row vectors form the Hadamard matrix $H_4 = H_2 \otimes H_2$, where H_2 is the usual Hadamard matrix of size 2. Therefore, all rows are orthogonal.[3] Note that the last three rows satisfy the Schur condition.

We extend the map φ to vectors $\vec{v} = (v_1, \ldots, v_m) \in \mathbb{F}_4^m$ by defining the map

$$\phi(\vec{v}) := \varphi(v_1) \otimes \cdots \otimes \varphi(v_m) \in \{-1,1\}^{4m}\,. \tag{3.32}$$

The image of ϕ is the set of all rows of the Hadamard matrix $H_2^{\otimes 2m}$. Let $U_k = \langle \vec{v}_k \rangle$ be a maximal spread of $\mathbb{F}_4^m$. By evaluating ϕ on the three elements of U_k (except for the zero vector $\vec{0}$) we get three orthogonal vectors satisfying the Schur condition. We can take them as rows of S_x, S_y, S_z:

$$\begin{aligned} S_{x;k} &= \phi(\omega \cdot \vec{v}_k) \\ S_{y;k} &= \phi(\omega^2 \cdot \vec{v}_k) \\ S_{z;k} &= \phi(1 \cdot \vec{v}_k) \end{aligned}$$

This shows that the second through the last rows of the Hadamard matrix $H_2^{\otimes 2m}$ can be divided into $(2^{2m}-1)/3$ disjoint 3-subsets, each with rows that satisfy the Schur condition. The rows in a 3-subset can be chosen as rows of S_x,S_y and S_z, respectively.

[3] The fact that this matrix is indeed the Hadamard matrix can also be derived with the help of group characters. More precisely, we consider $\mathbb{F}_4$ as a two-dimensional vector space over $\mathbb{F}_2$ and let tr denote the trace map of this field extension [Jac74, Section 4.15]. For all $z \in \mathbb{F}_4$ the map $\varphi(z) : x \mapsto (-1)^{tr(zx)}$ is an irreducible character of the additive group $(\mathbb{F}_4, +)$ (which is isomorphic to $Z_2 \times Z_2$). Hence, orthogonality of the rows follows from the orthogonality of the characters.

3.4 Time-reversal schemes

We now consider the problem to invert an arbitrary, possibly unknown Hamiltonian in a quantum network, i.e. to simulate $-H$ when H is present. This question is closely related to the construction of decoupling schemes.

Definition 3.32 (General time-reversal scheme)
A time-reversal scheme for a network consisting of n qudits is a control sequence $C := (\tau_1, U_1, \tau_2, U_2, \ldots, \tau_N, U_N)$ such that

$$C(H) = -H \tag{3.33}$$

for all pair-interaction Hamiltonians H of n qudits. We call $\tau = \sum_{j=1}^{N} \tau_j$ the time overhead and N the number of time steps.

We consider time-reversal schemes where all τ_j are equal. In the case of a single qudit we can invert the time evolution by summing over all elements of an error basis $\{U_1, U_2, \ldots, U_{d^2}\}$ but the identity (we assume that $U_1 = \mathbf{1}$):

$$\sum_{i=2}^{d^2} U_i^\dagger H U_i = -H \tag{3.34}$$

The resulting time overhead is $|G| - 1 = d^2 - 1$ and the number of time steps is $d^2 - 1$. This observation can be generalized to the case of multiple nodes. For that we introduce the notion of *normal form* for orthogonal arrays.

Definition 3.33 (Normal form)
Let M be an $OA_\lambda(n, N)$ over an alphabet $\mathcal{A}$. We say that M is in normal form if each entry in the first column of M is the first element of $\mathcal{A}$.

The next lemma shows that it is always possible possible to bring an orthogonal array into normal form without changing its parameters:

Lemma 3.34 *Let M be an $OA_\lambda(n, N)$. Then there is an orthogonal array with the same parameters that is in normal form.*

Proof. We identify the underlying alphabet $\mathcal{A}$ with an arbitrary finite group G of order $|\mathcal{A}|$. Consider two rows $(g_1, \ldots, g_N)$ and $(h_1, \ldots, h_N)$ of M. Multiplying the elements of the rows by g_1^{-1} and h_1^{-1}, respectively, preserves the property that all pairs occur with frequency λ since $G \times G$ is closed under multiplication by fixed elements, that is, $(g_1^{-1}, h_1^{-1})G \times G = G \times G$. □

Based on the normal form of OAs we now give a time-reversal schemes for a general, possibly unknown pair-interaction Hamiltonian.

Theorem 3.35 (Time-reversal with OAs)
Let M be an orthogonal array $OA(n, N)$ over an alphabet of size d^2. It can be used to reverse the time evolution of a quantum network consisting of n qudits. The number of time steps is $N-1$ and the time overhead is $N-1$.

Proof. We choose $\mathcal{A} := \{1, \ldots, d^2\}$ as alphabet. Let $\mathcal{E} = \{1, \ldots, d^2\}$ be an arbitrary error basis of $\mathbb{C}^{d\times d}$. The elements of $\mathcal{A}$ enumerate the elements of an error basis $\mathcal{E} = \{1, \ldots, d^2\}$.
By applying the transformation of Lemma 3.34 we may assume that the orthogonal array M is in normal form, that is, all entries of the first columns are 1. The entries of M are denoted by m_{kj}. Then we have

$$-H = \sum_{j=2}^{N} (U^\dagger_{m_{1j}} \otimes \cdots \otimes U^\dagger_{m_{nj}}) H (U_{m_{1j}} \otimes \cdots \otimes U_{m_{nj}}) .$$

□

It is obvious that this method also works for difference schemes and Hadamard matrices.

3.5 Decoupling and time-reversal schemes for partially coupled systems

The assumption that every node is coupled to all the other nodes is too strong in many physical systems since many coupling terms might be neglected. This reduces the overhead for decoupling and inverting the time evolution.
The interaction graph of a partially coupled network is a non-complete graph; this property allows to construct more efficient decoupling and time-reversal schemes. To see how this can be done we need some basic results of graph theory (cf. Bollobás [Bol98]).

Definition 3.36 (Chromatic number)
A graph can be colored, by assigning each vertex one of a number of different colors. The coloring scheme is called a proper coloring if no two connected vertices have the same color. The chromatic number χ is the smallest number of colors required to properly color the graph.

To illustrate this definition let us consider a scheduling problem. We wish to arrange talks in a conference in such way that no participant will be forced to miss a talk they would like to hear: there are no undesirable clashes. Assuming a good supply of lecture rooms enabling us to hold as many parallel talks as we like, how long the conference has to last? What is the smallest number of time

slots required? Let G be the graph whose vertices are the talks and in which two talks are joined if and only if there is a participant wishing to attend both. What is the minimal value k for which V can be partitioned into k classes such that no edge joins two vertices of the same class? It is given by the chromatic number of G.

In a complete graph (a fully coupled network) $\chi = n$, but in a partially coupled network the chromatic number can be much smaller. This observation permits to derive more efficient decoupling schemes (cf. JONES AND KNILL [JK99]) since if the network is represented by a properly colored graph, then there are no constraints on the control operations between nodes with the same color. It is sufficient to create a decoupling scheme of a completely coupled χ-node network, and apply identical sequences to all nodes of the same color.

So the chromatic number gives an upper bound on the complexity of time-reversal. Lower bounds on the complexity of time-reversal are derived in Chapter 5. It follows that these lower bounds on the complexity of time-reversal are also lower bounds on the chromatic number.

Chapter 4

Universal simulation of Hamiltonians

4.1 Transformers

In Chapter 3 we have given a necessary and sufficient condition on the minimal set of available control operations in order to enable decoupling and time-reversal of Hamiltonians. We have established a one-to-one correspondence between error bases and minimal sets of control operations for decoupling and time-reversal. However, if we want to simulate an arbitrary Hamiltonian by any other this condition is not sufficient. This can be seen by the following example. Assume that the only control operations on $\mathbb{C}^2$ are given by the Pauli-matrices (in their role as unitary operators). If the one-qubit Hamiltonian $H = \sigma_z$ is given, conjugation of H by a Pauli-matrix can only lead to either H or $-H$. All the Hamiltonians which can be obtained as average Hamiltonians are scalar multiples of H. Hence one cannot simulate e.g. σ_x using only identity and Pauli matrices (that form together an error basis of $\mathbb{C}^{2\times 2}$) as control operations. The concept of *transformers* introduced in Wocjan et al. [WRJB02b] is useful in order to characterize and to find control groups that enable universal simulation of Hamiltonians. This section is based on [WRJB02b].

Definition 4.1 (Transformer)
A subgroup $\mathcal{T}$ of $SU(d)$ is called a universal transformer of Hamiltonians if every $\mathbb{R}$-linear map L on $su(d)$ can be written as

$$L(A) = \sum_{j=1}^{N} \tau_j U_j^\dagger A U_j$$

with $N \geq 1$, positive real numbers τ_j and $U_j \in \mathcal{T}$.

The physical meaning of this is that a transformer allows to simulate the Hamiltonian $L(H)$ for an arbitrary map L if the *unknown* Hamiltonian H is present. JANZING AND BETH showed in [JB02] without using this terminology that $SU(d)$ is a transformer for every dimension d.

Remark 4.2 In particular, a transformer is able to simulate an arbitrary Hamiltonian $\tilde{H} \in su(d)$ by an arbitrary (non-zero) Hamiltonian $H \in su(d)$ by realizing a map L such that $L(H) = \tilde{H}$.

Remarkably, the condition for a *finite* group to be a transformer can be characterized in terms of irreducibility of certain representations. In contrast to the condition for an annihilator, it refers to the adjoint action on the set of operators instead of the underlying Hilbert space:

Definition 4.3 (Adjoint action)
Let G be a finite group and φ a unitary representation of degree d, i.e., φ acts on $V = \mathbb{C}^d$. We define a representation φ_{ad} on $V \otimes V$ by $\varphi_{\mathrm{ad}}(g) := \overline{\varphi(g)} \otimes \varphi(g)$ for all $g \in G$, where $\overline{U}$ denotes complex conjugation of a matrix U. We call φ_{ad} the adjoint action *of φ. Note that this action can be identified with the action of G on matrices in $\mathbb{C}^{d\times d}$ via conjugation $g \mapsto (M \mapsto \varphi(g)^\dagger M \varphi(g))$.*

In the following we make use of the fact that the algebra generated by the images of an irreducible m-dimensional representation ϑ of a finite group G over the complex numbers is equal to the full matrix algebra $\mathbb{C}^{m\times m}$. We cite the corresponding theorem from ISAACS [Isa76, Theorem 9.2]. A representation ϑ defined over a field F is called *absolutely irreducible* if it remains irreducible when considered over an extension field E/F.

Theorem 4.4 *Let ϑ be an absolutely irreducible representation of a finite group G which has degree m and is defined over the field F. Then*

$$\left\{ \sum_{g\in G} \alpha_g \vartheta(g) : \alpha_g \in F \right\} = F^{m\times m}.$$

In particular for any m-dimensional irreducible representation over the field $\mathbb{C}$ of complex numbers the vector space generated by the images equals $\mathbb{C}^{m\times m}$.

We now have the necessary prerequisites to characterize finite transformers.

Theorem 4.5 (Characterization of finite transformers)
A finite group $\mathcal{T} \leq SU(d)$ is a transformer if and only if the adjoint representation ϑ given by

$$\vartheta(U) := (A \mapsto U^\dagger A U)$$

with $U \in \mathcal{T}$ acts irreducibly on $sl(d) = su(d) + i\ su(d)$, i.e., the space of traceless operators.

Proof. ($\Longleftarrow$) Let L be a given $\mathbb{R}$-linear map on $su(d)$. Assume that the adjoint action of $\mathcal{T}$ is irreducible on $sl(d)$ and denote this representation by ϑ. From Theorem 4.4 follows that the complex linear span of the images of ϑ is the full matrix algebra acting on $sl(d)$. Hence, the mapping L can be written as a complex linear combination of the form

$$L : A \mapsto \sum_j c_j U_j^\dagger A U_j \,, \quad c_j \in \mathbb{C} \,.$$

We now show that the coefficients can be chosen to be real for $A \in su(d)$: since $U_j^\dagger A U_j$ is Hermitian for all j we have $L(A) = L(A)^\dagger = \sum_j \bar{c}_j U_j^\dagger A U_j$. Therefore we can write L in the form

$$L : A \mapsto \sum_j r_j U_j^\dagger A U_j \,, \quad r_i = \frac{1}{2}(c_i + \bar{c}_i) \in \mathbb{R} \,.$$

Since the adjoint representation acts irreducibly on $sl(d)$, we can realize a time-reversal scheme by conjugating A with all elements of $\mathcal{T}$ but the identity. This shows that it is sufficient to consider positive real numbers as coefficients c_j; the minus signs can be achieved by applying the time-reversal scheme.

($\Longrightarrow$) Assume that every $\mathbb{R}$-linear map on $su(d)$ can be implemented in the sense of Definition 4.1 using $\mathcal{T}$. Let $\mathcal{M}$ be the complex linear span of the maps $\vartheta(U)$ with $U \in \mathcal{T}$. The idea is to show that any $F \in sl(d)$, $F \neq 0$ can be mapped to any other $\tilde{F} \in sl(d)$ by a map $T \in \mathcal{M}$. This in turn shows that the adjoint action is irreducible since there cannot be a nontrivial invariant subspace.

Let $F = H_1 + iH_2$ and $\tilde{F} = \tilde{H}_1 + i\tilde{H}_2$ with $H_1, H_2, \tilde{H}_1, \tilde{H}_2 \in su(d)$. To construct T we consider two cases. If H_1 and H_2 are linearly dependent, i.e. $H_1 = \lambda H_2$ for some $\lambda \in \mathbb{R}$, we may assume w.l.o.g. that $F \in su(d)$ (otherwise multiply F with a suitable complex number). There are maps L_1 and L_2 in $\mathcal{M}$ with $L_1(F) = \tilde{H}_1$ and $L_2(F) = \tilde{H}_2$. Then $T := L_1 + i\, L_2$ is the desired map.

If H_1 and H_2 are linearly independent, there are maps L_1 and L_2 in $\mathcal{M}$ such that

$$\begin{array}{rclcrcl} L_1(H_1) & = & \tilde{H}_1 & \qquad & L_2(H_1) & = & \tilde{H}_2 \\ L_1(H_2) & = & \mathbf{0} & & L_2(H_2) & = & \mathbf{0} \end{array}$$

Then $T := L_1 + i\, L_2$ is the desired map. □

We derive a necessary and sufficient condition for a finite group to be a transformer group in the sense of Definition 4.1. Theorem 4.5 shows that the problem to construct a finite transformer group is to find for given dimension $d > 1$ a finite group G and an irreducible (unitary) representation φ of G such that the adjoint action becomes irreducible if we split off the trivial representation $\mathbf{1}$ of G. The trivial representation is always contained in φ_{ad} since the one-dimensional space corresponding to the linear span of the identity matrix remains invariant,

i.e. $\varphi_{\mathrm{ad}} = \mathbf{1} \oplus \pi$ for some representation π of G. Abusing the notation we will write $\varphi_{\mathrm{ad}} - \mathbf{1}$ to denote the summand π in this decomposition.

Once we have found a suitable pair (G, φ) with $\deg(\varphi) = d$ this yields a transformer group as in Definition 4.1. For basic results concerning representation theory of finite groups we refer the reader to Isaacs [Isa76].

Example 4.6 (Two-dimensional transformer)
We examine the case of a two-dimensional system, i. e., $d = 2$. Starting from the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

we first note that the group $\langle i \cdot \sigma_x, i \cdot \sigma_y, i \cdot \sigma_z \rangle$ is isomorphic to the quaternion group Q_8 of order 8. This group has an (outer) automorphism of order 3 which permutes the Pauli matrices cyclically. This automorphism is given by the matrix

$$R := \frac{i-1}{2} \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}.$$

Setting $s_k := i\sigma_k$ for $k \in \{x, y, z\}$ the automorphism is given by $R^{-1} s_x R = s_y$, $R^{-1} s_y R = s_z$, and $R^{-1} s_z R = s_x$. The group generated by the s_k and R is isomorphic to $\mathrm{SL}(2, \mathbb{F}_3)$, i.e. the group of 2×2 matrices over the finite field $\mathbb{F}_3$ which have determinant 1. Let φ be the (natural) representation of the matrix group given by $\langle s_x, s_y, s_z, R \rangle$. Then the 24 matrices in the image of φ form a faithful irreducible representation of $\mathrm{SL}(2, \mathbb{F}_3)$. Choosing the basis $\{s_x, s_y, s_z\}$ of $sl(2)$ we see that the images of $\varphi_{\mathrm{ad}} - \mathbf{1}$ are given explicitly by $s_x \mapsto \mathrm{diag}(1, -1 - 1)$, $s_y \mapsto \mathrm{diag}(-1, 1, -1)$, $s_z \mapsto \mathrm{diag}(-1, -1, 1)$, and R maps to the permutation matrix corresponding to the 3-cycle $(1, 2, 3)$. It is readily verified that this is an irreducible representation.

Let G be a finite group having an irreducible representation φ such that the images of φ are a transformer in the sense of Definition 4.1. Then necessarily φ must be non-monomial[1] for otherwise the set of diagonal matrices would be an invariant subspace under the action of $\varphi_{\mathrm{ad}} - \mathbf{1}$. This give a necessary condition for transformers. Note that in fact the group $\mathrm{SL}(2, \mathbb{F}_3)$ is the smallest group which is not an $\mathcal{M}$-group, i.e., $\mathrm{SL}(2, \mathbb{F}_3)$ has representations which are not equivalent to monomial ones.

There is a necessary and sufficient characterization of transformer groups which can be verified from the character table alone. The *character* χ of a representation φ is defined by $\chi(g) := \mathrm{tr}(\varphi(g))$ and is called irreducible if and only if the corresponding representation is irreducible (cf. Isaacs [Isa76]).

[1] A representation is called *monomial* if all representing matrices have the property to contain precisely one non-vanishing entry in each row and each column.

Theorem 4.7 (Characterization of finite transformers)
Let G be a finite group and χ be an irreducible character of G with corresponding representation φ. Then χ corresponds to a universal transformer if and only if the following identity holds:

$$\sum_{g\in G} |\chi(g)|^4 = 2|G|\,.$$

Proof. The representation $\varphi_{\mathrm{ad}} - \mathbf{1}$ has character values $\overline{|\chi(g)|^2} - 1$ for all $g \in G$ since $\mathrm{tr}(\varphi_{\mathrm{ad}}(g)) = \overline{\chi(g)}\chi(g)$; this follows from $\varphi_{\mathrm{ad}}(g) = \varphi(g) \otimes \varphi(g)$. Recall that the vector space of class functions on G (i.e., functions that are constant on the conjugacy classes of G) has a normalized scalar product given by

$$\langle \chi_1 | \chi_2 \rangle = \frac{1}{|G|} \sum_{g\in G} \chi_1(g)\chi_2(g^{-1})$$

for characters χ_1, χ_2 of G. A character χ is irreducible if and only if $\langle \chi|\chi\rangle = 1$. By computing the scalar product of the character corresponding to $\varphi_{\mathrm{ad}} - \mathbf{1}$ we obtain

$$\begin{aligned} \frac{1}{|G|} \sum_{g\in G} (|\chi(g)|^2 - 1) \cdot (\overline{|\chi(g)|^2 - 1}) &= \frac{1}{|G|} \sum_{g\in G} |\chi(g)|^4 - \frac{2}{|G|} \sum_{g\in G} |\chi(g)|^2 + 1 \\ &= \frac{1}{|G|} \sum_{g\in G} |\chi(g)|^4 - 1 \end{aligned}$$

On the other hand this scalar product is equal to 1 due to the irreducibility of $\varphi_{\mathrm{ad}} - \mathbf{1}$. Rearranging terms and clearing denominators yields the claimed statement. □

In the following we present a transformer for a three dimensional system.

Example 4.8 (Three-dimensional transformer)
The minimal group having a representation φ for which $\varphi_{\mathrm{ad}} - \mathbf{1}$ is irreducible is the linear group $\mathrm{GL}(3, \mathbb{F}_2)$ of invertible 3×3 matrices over the field $\mathbb{F}_2$. This is a simple group of order 168. As generators of this group we choose the matrices

$$x := \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad y := \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

where x is an element of order 2 and the order of y is 7. The group $\mathrm{GL}(3, \mathbb{F}_2)$ has a three-dimensional irreducible representation φ over the complex numbers which on the generators x and y is given by

$$\varphi(x) := \frac{2}{\sqrt{7}} \begin{pmatrix} \cos(\frac{5\pi}{14}) & -\zeta_7^3 \cos(\frac{\pi}{14}) & -\zeta_7^2 \cos(\frac{3\pi}{14}) \\ -\zeta_7^4 \cos(\frac{\pi}{14}) & -\cos(\frac{3\pi}{14}) & \zeta_7^6 \cos(\frac{5\pi}{14}) \\ -\zeta_7^5 \cos(\frac{3\pi}{14}) & \zeta_7 \cos(\frac{5\pi}{14}) & -\cos(\frac{\pi}{14}) \end{pmatrix}$$

and

$$\varphi(y) := \begin{pmatrix} \zeta_7 & \cdot & \cdot \\ \cdot & \zeta_7^2 & \cdot \\ \cdot & \cdot & \zeta_7^4 \end{pmatrix}.$$

Here $\cdot$ is an abbreviation for 0 and ζ_7 denotes the primitive 7-th root of unity $e^{2\pi i/7}$. The character of the representation φ takes the values

$$3,\ -1,\ 1,\ 0,\ \zeta_7 + \zeta_7^2 + \zeta_7^4,\ \zeta_7^3 + \zeta_7^5 + \zeta_7^6$$

on the conjugacy classes of $\mathrm{GL}(3, \mathbb{F}_2)$. Consulting the character table of $\mathrm{GL}(3, \mathbb{F}_2)$ in the Atlas (cf. Conway et al. [CCNW85, p. 3]) we find that φ is irreducible. The representation $\varphi_{\mathrm{ad}} - \mathbf{1}$ has character values

$$8,\ 0,\ 0,\ -1,\ 1,\ 1$$

from which follows that it is also irreducible, again by checking the character table of $\mathrm{GL}(3, \mathbb{F}_2)$. Overall we obtain that the representation φ of $\mathrm{GL}(3, \mathbb{F}_2)$ yields a transformer of size 168.

Based on the Neubüser catalogue used in MAGMA (cf. Bosma et al. [BCP97]) and GAP (cf. [GAP97]) we performed an exhaustive search over all groups of smaller sizes which has shown that this indeed is the minimal possible group size. In Table 4.1 we summarize the results of this search. Groups of sizes up to 255 have been considered. The number in the Neubüser catalogue is given such that for instance the first row of this table corresponds to the group (in GAP syntax) `SmallGroup(24,3)` which has been studied in Example 4.6. Note that we only give transformer groups which act faithfully.

4.2 Universal simulation

By putting together the results on selective decoupling and transformers we show that all pair-interaction Hamiltonians can be simulated by any pair-interaction Hamiltonian (with non-zero coupling and non-zero local terms) provided that the available control operations on each subsystem form a transformer.

We first prove a lemma showing when universal simulation is possible in bipartite quantum systems. The general case reduces to this since we deal with pair-interactions.

Lemma 4.9 (Bipartite quantum system)
Let H be the system Hamiltonian of a bipartite system

$$H := \sum_{\alpha\beta} J_{\alpha\beta} \sigma_\alpha \otimes \sigma_\beta + a \otimes \mathbf{1} + \mathbf{1} \otimes b$$

Group size	Numbers in library	Dimension
24	3	2
48	28, 29, 33	2
72	3, 25	2
96	67, 74, 192	2
120	5	2
144	36, 121, 122, 157	2
168	22	2
168	42	3
192	187, 204, 963	2
216	3, 38	2
216	88	3
240	93, 102, 103, 154	2

Table 4.1: Transformer groups of small sizes

with a non-trivial coupling between the nodes and with non-trivial local terms $a, b \in su(d)$. Assume that it is possible to implement all unitary transformations of the form $U \otimes V \in \mathcal{T}_1 \otimes \mathcal{T}_2$, where $\mathcal{T}_1$ and $\mathcal{T}_2$ are the transformers of the left and the right subsystem, respectively.
Then H can simulate every Hamiltonian

$$\tilde{H} := \sum_{\alpha\beta} \tilde{J}_{\alpha\beta}\sigma_\alpha \otimes \sigma_\beta + \tilde{a} \otimes \mathbf{1} + \mathbf{1} \otimes \tilde{b}\,, \tag{4.1}$$

where $\tilde{a}, tildeb \in su(d)$.

Proof. We first do not consider the local terms. Write H in the form

$$H = \sum_\gamma A_\gamma \otimes B_\gamma\,,$$

where $A_j \in su(d)$ are all nonzero and $B_j \in su(d)$ are all linearly independent. Then H can be transformed into any interaction of the form $C \otimes D$ with arbitrary $C, D \in su(d)$. We construct a map L such that $L(H) = C \otimes D$. This can be done by choosing $\mathbb{R}$-linear maps L_1 and L_2 on $su(d)$ with $L_1(A_1) = C$, $L_2(B_1) = D$, and $L_2(B_\gamma) = 0$ for $\gamma \neq 1$. Since $\mathcal{T}_1$ and $\mathcal{T}_2$ are universal transformers we can find positive real numbers r_i and s_j and unitary transformations $U_i \in \mathcal{T}_1$ and $V_j \in \mathcal{T}_2$ such that

$$L_1(A) = \sum_{i=1}^{N} r_i U_i^\dagger A U_i \quad \text{and} \quad L_2(A) = \sum_{j=1}^{M} s_j V_j^\dagger A V_j$$

for all $A \in su(d)$. Concatenation of the corresponding control sequences $(r_1, U_1 \otimes \mathbf{1}; \ldots; r_N, U_N \otimes \mathbf{1})$ and $(s_1, \mathbf{1} \otimes V_1; \ldots; s_M, \mathbf{1} \otimes V_M)$ implements the desired linear

map L

$$\sum_{ij} r_i s_j (U_i \otimes V_j)^\dagger H (U_i \otimes V_j) = C \otimes D \,.$$

This proves that we can simulate each tensor product operator $C \otimes D$. By setting $C := \tilde{J}_{\alpha\beta}\sigma_\alpha$ and $D := \sigma_\beta$ we can simulate all bilinear terms in eq. (4.1). This shows that we can simulate the bilinear part of $\tilde{H}$.

Let us now consider the local terms. As we have seen we can simulate the Hamiltonian

$$H' = \sum_{\alpha\beta} \tilde{J}_{\alpha\beta}\sigma_\alpha \otimes \sigma_\beta + a' \otimes \mathbf{1} + \mathbf{1} \otimes b' \,,$$

that coincides with the desired Hamiltonian $\tilde{H}$ except for the local terms a' and b'. The local terms a and b are during the simulation of the desired coupling and become a' and b', respectively.

By concatenating an annihilator on the right and a suitable control sequence on the left we can simulate the Hamiltonian $(\tilde{a} - a') \otimes \mathbf{1}$. Using a similar scheme we can simulate $\mathbf{1} \otimes (\tilde{b} - b')$. Finally, the sum of these Hamiltonians gives the desired Hamiltonian $\tilde{H}$. □

Note that this proof also applies if the dimensions of the two subsystems are different. We have shown in that any bipartite Hamiltonian can simulate any other provided that it consists of non-trivial local Hamiltonians on both nodes and a non-trivial coupling. If this criterion is met by all pairs of an n-partite Hamiltonian then universal simulation of all pair-interaction Hamiltonians[2] is possible by applying the selective decoupling techniques. By canceling all couplings but the coupling between nodes k and l we end up with the Hamiltonian $H_{kl} + H_k + H_l$.

Theorem 4.10 (Universal simulation)
Let H be a pair-interaction Hamiltonian such that all bilinear terms and all local terms are non-trivial. Then H can simulate any other pair-interaction Hamiltonian provided that transformers can be implemented on all nodes.

Proof. We apply a selective decoupling scheme to cancel all couplings but the coupling between nodes k and l. After this we end up the Hamiltonian $H_{kl} + H_k + H_l$. By applying a suitable control sequence of transformer operations we can simulate the Hamiltonian $\tilde{H}_{kl} + H'_k + H'_l$, where $\tilde{H}_{kl}$ is the desired coupling between nodes k and l. The local Hamiltonians are changed by this control sequence to H'_k and H'_l.

[2]The condition that *all* the couplings have to be non-trivial is only necessary in the average Hamiltonian approach. If higher order terms in the time interval are considered interactions between nodes k and l and k and m can be used for simulating arbitrary couplings between l and m as noted in Dodd et al. [DNBT01] and Nielsen et al. [NBD+01].

By repeating this procedure for all pairs (k, l) we can simulate a Hamiltonian that coincides with the desired one except for the local terms. The local Hamiltonians are adjusted analogously to the bipartite case. □

The time overhead and number of time steps are of order $O(n^2)$ since there are $n(n-1)/2$ pairs of nodes.

Chapter 5

Bounds on complexity of simulating Hamiltonians

Having shown in Chapter 4 under which conditions universal simulation of Hamiltonians is possible we turn to investigate the complexity, i.e., the computational resources required for simulating Hamiltonians. The main result is consists in a complexity theory that permits to compare quantitatively the computational power of different interactions. Whereas other works examine the question of optimality for the two-qubit case we obtain lower and upper bounds on the time overhead and the number of time steps for general quantum networks. This chapter presents the bounds derived in Wocjan et al. [WJB02b], Wocjan et al. [WRJB02a], and Janzing, Wocjan and Beth [JWB02a].

In Section 5.1 we derive lower bounds on the time overhead and the number of time steps. The lower bounds on the time overhead are derived with the help of majorization theory. Majorization induces a partial order on Hermitian matrices by comparing their spectra (i.e. their eigenvalues). We will show that if a Hamiltonian H can simulate a Hamiltonian $\tilde{H}$ with time overhead 1, then necessarily H is greater or equal to $\tilde{H}$ with respect to majorization. This observation leads to lower bounds on the time overhead.

The difficulty in working with the Hamiltonians is that it is hard to compute their spectra since their size grows exponentially with the number of subsystems. Therefore, we represent Hamiltonians by coupling matrices that form the fundamental data structure of our theory. It is much easier to compute the spectra of coupling matrices since their size grows only linearly with the number of subsystems. Furthermore, lower bounds on the number of time steps can be derived by comparing the ranks of the coupling matrices.

The most important advantage of the representation by coupling matrices is that homogeneous Hamiltonians can be described by weighted interaction graphs. A homogeneous Hamiltonian is a pair-interaction Hamiltonian where all subsystems interact via the same type of coupling and only the strengths and signs of

the interactions vary. This situation often arises in physics; the Ising model is one example. In this case the computation of the spectra of the coupling matrices is facilitated substantially since they are directly linked to spectra of the adjacency matrices of corresponding interaction graphs. This connection allows to use the results of *algebraic graph theory* that investigates precisely how, or whether, properties of graphs are reflected by the spectra of their adjacency matrices (cf. Chung [Chu97], Cvetković et al. [CDS95], and Godsil and Royle [GR01]).

The lower bounds are tight bounds in many cases that are relevant for quantum computing. Especially, the derived bounds prove the optimality of the decoupling and time-reversal schemes presented in Chapter 3.

In Section 5.2 Carathéodory's theorem provides a general upper bound on the number of time steps.

5.1 Lower bounds

5.1.1 Basic results of majorization theory

Majorization was developed to answer the following question: what does it mean to say that one probability distribution is more mixed than another? In the quantum mechanical context, this question becomes: given two density operators, what does it mean that one is more mixed than the other? Recent results in entanglement theory make extensive use of majorization theory (cf. Nielsen and Vidal [NV01]). The aim of this section is to show that majorization theory can also be applied with great success to the problem of optimal simulation of Hamiltonians.

We state and explain a series of theorems, mainly without including complete proofs. Any reader seriously interested in majorization is referred to Marshall and Olkin's book [MO79], the Chapter 2 and 3 of Bhatia's book [Bha96], Andos's survey articles [And89, And94], and Alberti and Uhlmann's monograph [AU82].

The basic intuition underlying majorization may be understood from the following definition:

Definition 5.1 (Vector majorization)
Let $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ be two d-dimensional real vectors. We say that x is majorized by y, written $x \preceq y$, if there exists a set $\{\Pi_1, \ldots, \Pi_N\}$ of d-dimensional permutation matrices and a probability distribution $\{p_1, \ldots, p_N\}$ such that

$$x = \sum_{j=1}^{N} p_j \Pi_j y \,. \tag{5.1}$$

That is, x is majorized by y precisely when x can be obtained from y by randomly permuting the components of y, and then averaging over the permutations with respect to a probability distribution.

Suppose the vectors x and y are probability distributions. Then, at least naively, this definition appears to be a natural and appealing approach to capturing the notation that one probability distribution is more disordered or mixed than the other.

As a simple example of majorization, suppose the vector x is an arbitrary probability distribution on d outcomes. Then it is easy to see that

$$\left(\frac{1}{d}, \ldots, \frac{1}{d}\right) \preceq x\,,$$

since the the uniform distribution $(1/d, \ldots, 1/d)$ may be obtained by averaging over the d permutations that cyclically shift the components of x with respect to the uniform distribution on d elements. This simple example agrees with our intuition that the uniform distribution on d elements is at least as disordered as any other probability distribution over d elements.

The definition for the majorization relation $x \preceq y$ in eqs. (5.1) and (5.5) in terms of random permutations is satisfying from an intuitive point of view, but is rather inconvenient for actual calculations. Given two vectors x and y, is there some simple procedure to determine whether $x \preceq y$? Rather remarkably, such a procedure does exit. First, we rearrange the components of x and y into decreasing order, writing for example $x^{\downarrow} = (x_1^{\downarrow}, \ldots, x_d^{\downarrow})$ for the vector whose components are the same as those of x, but ordered so that

$$x_1^{\downarrow} \geq x_2^{\downarrow} \geq \cdots \geq x_d^{\downarrow}\,. \tag{5.2}$$

It turns out that $x \preceq y$ if and only if

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow} \tag{5.3}$$

for $k = 1, \ldots, d-1$ and

$$\sum_{j=1}^{d} x_j^{\downarrow} = \sum_{j=1}^{d} y_j^{\downarrow}\,. \tag{5.4}$$

Note that we could have actually taken this *calculational* definition from the very beginning, and obtain eq. (5.1) as a result. However, this set of inequalities is probably less suggestive than eq. (5.1) as far as the intuitive meaning of the majorization relation is concerned. The equivalence of both definitions is discussed in full detail in Nielsen and Vidal [NV01].

Let us now turn to the connection between majorization and quantum mechanics. The quantum mechanical analogue of a probability distribution is the density operator, so the first step is to define an operator notation of majorization.

Definition 5.2 (Operator majorization)
Let A and B be two arbitrary d-dimensional Hermitian operators. We say A is majorized by B, written $A \preceq B$, if

$$\lambda(A) \preceq \lambda(B)\,, \tag{5.5}$$

where $\lambda(A)$ and $\lambda(B)$ denote the spectra of A and B, i.e., the vectors whose components are the eigenvalues of A and B, respectively.

Uhlmann's theorem extends eq. (5.1) to operator majorization.

Theorem 5.3 (Uhlmann's theorem)
Let A and B be two arbitrary Hermitian operators of size d. We have $A \preceq B$ if and only if there exists a set $\{U_1, \ldots, U_N\}$ of unitary matrices of size d and a probability distribution $\{p_1, \ldots, p_N\}$ such that

$$A = \sum_{j=1}^{N} p_j U_j^\dagger B U_j\,. \tag{5.6}$$

Uhlmann's theorem clearly illustrates the idea that the Hermitian operator A is more random than B, since A can be obtained by independently applying to B unitary operations $\{U_1, \ldots, U_N\}$, and mixing the resulting operators $U_j^\dagger B U_j$ according to the probability distribution $\{p_1, \ldots, p_N\}$. This transformation is known as *unitary mixing*.

We will need the following result when working with coupling matrices.

Remark 5.4 (Real symmetric matrices)
If the Hermitian matrices in Uhlmann's theorem are replaced by real symmetric matrices, then it is sufficient to use only orthogonal instead of unitary matrices.

5.1.2 Bounds from majorization

In order to derive lower bounds on the time overhead and the number of time steps of a simulation scheme we neglect the free evolutions of the nodes and consider the weaker problem to simulate the desired Hamiltonian up to local terms of each node. Note that the local terms become irrelevant when allowing arbitrary unitary operations on the nodes.

We consider first the time overhead. Let H and $\tilde{H}$ be two Hamiltonians. Recall that the relation "$\tilde{H}$ can be simulated by H with time overhead 1" defines a quasi-order of Hamiltonians. A partial characterization of this quasi-order gives the following majorization criterion:

Theorem 5.5 (Lower bound on time overhead)
Let H and $\tilde{H}$ be arbitrary pair-interaction Hamiltonians of n coupled qudits:

$$H = \sum_{k<l} H_{kl} \quad \text{and} \quad \tilde{H} = \sum_{k<l} \tilde{H}_{kl}$$

A necessary condition that $\tilde{H}$ can be simulated by H with time overhead τ is that $\tilde{H}$ is majorized by τH. Furthermore, it is necessary that this majorization criterion is still satisfied after rescaling the bilinear terms

$$H'_{kl} := s_{kl} H_{kl}\,, \quad \tilde{H}'_{kl} := s_{kl} \tilde{H}_{kl}$$

with arbitrary real weights (s_{kl}).

Proof. By Definition 2.7 the Hamiltonian $\tilde{H}$ can be simulated by H with time overhead τ if there is a control sequence $C := (\tau_1, U_1; \ldots; \tau_N, U_N)$ with time overhead τ such that $\tilde{H} = C(H)$. This means that there are unitary matrices $U_1, \ldots, U_N \in K$ and real positive numbers $\tau_1, \ldots, \tau_N$ summing to τ such that

$$\tilde{H} = \sum_{j=1}^{N} \tau_j U_j^\dagger H U_j\,.$$

Dividing both sides by τ yields

$$\tilde{H}/\tau = \sum_{j=1}^{N} p_j U_j^\dagger H U_j\,,$$

where $p_1, \ldots, p_N$ is a probability distribution. Therefore, Uhlmann's theorem implies that $\tilde{H}/\tau \preceq H$. This proves the first statement.

Note that the same control sequence C can be used for the rescaled problem. This is due to the special form of the control operations and the fact that both Hamiltonians are pair-interaction Hamiltonians. Therefore, the second statement is a direct consequence from the first one. □

For the case of two coupled qubits the majorization criterion is a necessary and sufficient condition (cf. Bennett et al. [BCL$^+$02] and Vidal and Cirac [VC01]). However, the situation is different in the case of two coupled subsystems of dimension $d > 2$. There are examples of Hamiltonians of two coupled three level systems that have the same spectrum but cannot simulate each other with time overhead $\tau = 1$ (cf. Chen [Che02]). This shows that Theorem 5.5 gives in general only a necessary but not a sufficient condition.

The difficulty in applying Theorem 5.5 is that it is hard to compute the eigenvalues of the Hamiltonians. This is because the dimension of the underlying Hilbert

space grows exponentially as d^n with the number of nodes, whereas the number of parameters grows quadratically as $(d^2-1)^2(n-1)n/2$; the number of couplings is $(n-1)n/2$ and the number of parameters needed to describe a coupling between two qudits is $(d^2-1)^2$.

5.1.3 Coupling matrix

In the following, $\{\sigma_\alpha \mid \alpha = 1, \ldots, m\}$ denotes an orthogonal basis of $su(d)$ with respect to the trace inner product, where $m = d^2 - 1$ is the dimension of $su(d)$ viewed as a real vector space.

Definition 5.6 (Coupling matrix)
Let H be a general pair-interaction Hamiltonian on n coupled qudits (without local terms). Its coupling matrix J is a real symmetric $mn \times mn$-matrix such that

$$H := \sum_{k<l} \sum_{\alpha\beta} J_{kl;\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} \,. \tag{5.7}$$

This condition determines all entries $J_{kl;\alpha\beta}$ for $k < l$. From symmetry of the coupling matrix it follows that $J_{lk;\beta\alpha} = J_{kl;\beta\alpha}$ for all $k < l$. We set $J_{kk;\alpha\beta} := 0$ for all k.

The coupling matrix J consists of $m \times m$-blocks. The $m \times m$-matrix J_{kl} given by the block at position (k, l) describes the coupling between the qudits k and l. We have $J_{lk} = J_{kl}^T$, i. e. the matrix describing the coupling between the qudits l and k is just the transpose of the matrix describing the coupling between k and l. The blocks on the diagonal are zero matrices. Note that the symmetry of the coupling matrix J does not imply any physical symmetry of the interaction. It is a consequence of our redundant representation that makes the coupling matrix symmetric. This is necessary for applying majorization theory to coupling matrices.

Before we can apply majorization theory to coupling matrices, we have to work out the effect of the control operations on the coupling matrices. To do that we first introduce the notion of adjoint action of the Lie group $SU(d)$ on the Lie algebra $su(d)$ and then show that is defines an orthogonal action on $\mathbb{R}^m$.

Definition 5.7 (Adjoint action)
The adjoint action of the Lie group $SU(d)$ on the Lie algebra $su(d)$ is defined by

$$A \mapsto U^\dagger A U \tag{5.8}$$

for all $A \in su(d)$ and all $U \in SU(d)$.

Lemma 5.8 *The adjoint action of $U \in SU(d)$*

$$A \mapsto U^\dagger A U$$

defines an orthogonal operation $O \in \mathbb{R}^{m\times m}$ via the relation

$$\vec{a} \mapsto O\vec{a}\,,$$

where $\vec{a} = (a_1, \ldots, a_m)^T \in \mathbb{R}^m$ is the row column vector whose coefficients are given by $a_i := \langle \sigma_i | a \rangle_{\rm tr}$ (i.e. the coefficients of the representation of the matrix $A \in su(d)$ with respect to the ONB $\{\sigma_1, \ldots, \sigma_m\}$).

Proof. We show that O is orthogonal with respect to the inner product on $\mathbb{R}^n$ defined by

$$\langle \vec{a}, \vec{b} \rangle := \langle A | B \rangle_{\rm tr}\,, \tag{5.9}$$

where $\vec{a}, \vec{b} \in \mathbb{R}^m$ and A, B are the corresponding operators in $su(d)$. It is sufficient to check that the scalar product on $\mathbb{R}^m$ defined by eq. (5.9) is left invariant under the operation of O. By the cyclic property of the trace we obtain

$$\begin{aligned} \langle O\vec{a}, O\vec{b} \rangle &= {\rm tr}((U^\dagger A U)^\dagger U^\dagger B U))/d \\ &= {\rm tr}(A^\dagger B)/d \\ &= \langle \vec{a}, \vec{b} \rangle\,. \end{aligned}$$

□

Note that for $d = 2$ the representation in Lemma 5.8 corresponds to the representation of density matrices by Bloch vectors and unitary operations to rotations on the Bloch sphere. However, for $d > 2$ the adjoint action of $SU(d)$ on $su(d)$ defines a *proper* subgroup of $SO(d^2 - 1)$, i.e., not every orthogonal matrix in $O \in SO(d^2 - 1)$ corresponds to conjugation by a unitary in $U \in SU(d)$. This is seen by counting the degrees of freedom.

In the following lemma we work out the effect of control operations on the blocks of a coupling matrix:

Lemma 5.9 *Let H be a general bipartite Hamiltonian*

$$H = \sum_{\alpha\beta} C_{\alpha\beta} \sigma_\alpha \otimes \sigma_\beta\,, \tag{5.10}$$

where $C \in \mathbb{R}^{m\times m}$ describes the coupling. Then the adjoint action of $U_1 \otimes U_2$ on H, i.e.,

$$H \mapsto (U_1 \otimes U_2)^\dagger H (U_1 \otimes U_2)$$

translates into

$$C \mapsto O_1 C O_2^T\,, \tag{5.11}$$

where O_1 and O_2 are the orthogonal matrices corresponding to U_1 and U_2, respectively.

Proof. The matrix C can be written as

$$C = \sum_{\alpha,\beta=1}^{m} C_{\alpha\beta} |\alpha\rangle\langle\beta| \,,$$

where $|\alpha\rangle$ and $|\beta\rangle$ are the standard basis vectors of $\mathbb{R}^m$. By Lemma 5.8 we know that conjugation of σ_α by U_1 corresponds to $|\alpha\rangle \mapsto O_1|\alpha\rangle$ and conjugation of σ_β by U_2 to $|\beta\rangle \mapsto O_1|\beta\rangle$ or equivalently to $\langle\beta| \mapsto \langle\beta|O_2^T$. This proves the correspondence in eq. (5.11). □

In the case of two qubits we have a normal form for the coupling:

Lemma 5.10 (Normal form for coupling between qubits)
Let H be a general Hamiltonian of two coupled qubits (without local terms) and $C \in \mathbb{R}^{3\times 3}$ the matrix describing the coupling as in Lemma 5.9. Then by rotating the reference frame we may assume that the coupling is diagonal

$$H = c_x \sigma_x \otimes \sigma_x + c_y \sigma_y \otimes \sigma_y + c_z \sigma_z \otimes \sigma_z \,.$$

The coefficients c_x, c_y, c_z are the singular values of C.

Proof. Lemma 5.9 shows that conjugation of H by a control operation $U_1 \otimes U_2 \in SU(2) \otimes SU(2)$ corresponds to multiplication of C by O_1 from the left and by O_2^T from the right, where O_1, O_2 are the corresponding operations in $SO(3)$. By the *singular value decomposition* (cf. [Bha96]) there are $O_1, O_2 \in SO(3)$ such that $O_1 C O_2^T = \mathrm{diag}(c_x, c_y, c_z)$, where c_x, c_y, c_z are the singular values of C. Equivalently, there is control operation $U \in SU(2) \otimes SU(2)$ such that $UHU^\dagger = H_{c_x,c_y,c_z}$ where $H_{c_x,c_y,c_z} = c_x \sigma_x \otimes \sigma_x + c_y \sigma_y \otimes \sigma_y + c_z \sigma_z \otimes \sigma_z$. □

Note that such a simple normal form can only be derived for qubits because for $d > 2$ the adjoint action of $SU(d)$ on $su(d)$ does not induce every possible transformation in $SO(d^2-1)$. This is needed for the singular value decomposition.

The action of control operations on coupling matrices is now obtained straightforwardly from Lemma 5.9.

Theorem 5.11 (Action on coupling matrices)
Let H be an arbitrary Hamiltonian with coupling matrix J. Conjugation of H by $U := U_1 \otimes U_2 \otimes \cdots \otimes U_n \in K$ translates to conjugation of J by a block diagonal matrix of the form

$$O := O_1 \oplus O_2 \oplus \cdots \oplus O_n \in \bigoplus_{k=1}^{n} SO(m) \,, \tag{5.12}$$

where the orthogonal matrices O_k correspond to U_k as in Lemma 5.8.

Let $\tilde{H}$ be the Hamiltonian with coupling matrix $\tilde{J}$ that we want to simulate. The condition for simulation in Definition 2.7 translates into

$$\tilde{J} = \sum_{j=1}^{N} \tau_j O_j J O_j^T \,, \tag{5.13}$$

where the orthogonal block diagonal matrices $O_j = O_{j1} \oplus O_{j2} \oplus \cdots \oplus O_{jn}$ correspond to the control operations $U_j = U_{j1} \otimes U_{j2} \otimes \cdots \otimes U_{jn} \in K$ as in eq. (5.12).

5.1.4 Bounds from spectra of coupling matrices

Now it becomes evident why we have chosen a redundant data structure encoding the Hamiltonians. Since the coupling matrices are real symmetric we can apply once majorization theory to eq. (5.13) to derive lower bounds on the time overhead.

Theorem 5.12 (Lower bound on time overhead)
Let H and $\tilde{H}$ be arbitrary pair-interaction Hamiltonians. A necessary condition that $\tilde{H}$ can be simulated with overhead τ by H is that the spectrum of $\tilde{J}$ is majorized by the spectrum of τJ. Furthermore, it is necessary that this majorization criterion is still satisfied after rescaling the couplings as follows: $J'_{kl} := s_{kl} J_{kl}$ and $\tilde{J}' := s_{kl} \tilde{J}$, where $S = (s_{kl})$ is an arbitrary real symmetric $n \times n$-matrix.

Proof. This is proved by applying Uhlmann's theorem to eq. (5.13). □

As a corollary we obtain a lower bound on the time overhead of time-reversal.

Corollary 5.13 (Lower bound on time overhead of time-reversal)
Let r be the greatest eigenvalue and q the smallest eigenvalue of the coupling matrix J representing H. Then $\tau \geq -r/q$ is a lower bound on the overhead of time-reversal of H.

Proof. Let J be the coupling matrix of H. Then $\tilde{J} := -J$ is the coupling matrix of $-H$. Rescale all blocks of both coupling matrices by -1, i.e. we have $J' = -J$ and $\tilde{J}' := J$. The bound follows now from the fact that the greatest eigenvalues of $\tilde{J}'$ and J' are r and $-q$, respectively. □

The advantage of the representation with coupling matrices is that their size grows only linearly with n. Therefore, this representation is more useful for calculations. Most importantly, it is much easier to compute the spectrum of the coupling matrices than the spectrum of the Hamiltonians themselves. This makes it possible to apply Theorem 5.12 successfully in order to derive lower bounds on the time overhead.

The usefulness of coupling matrices becomes especially evident if we consider interactions with an additional symmetry that is characterized by a weighted graph. It is defined as follows.

Definition 5.14 (Weighted graph)
Let $W = (w_{kl})$ be a real symmetric matrix of size n with zeros on the diagonal. The matrix W defines a so-called weighted graph[1] *$G = (V, E)$ as follows:*

1. *$V := \{1, \ldots, n\}$,*
2. *$(k, l) \in E$ if $w_{kl} \neq 0$,*
3. *the edge (k, l) has the weight w_{kl}.*

W is called the adjacency matrix of the weighted graph G. Conversely, a weighted graph defines a real symmetric matrix with zeros on the diagonal. An unweighted graph can be considered as a weighted one whose edges have the weight 1.
The spectrum of a weighted graph G is the spectrum of its adjacency matrix W, the vector of eigenvalues of W. We assume that the eigenvalues are listed in non-increasing order.

Definition 5.15 (Homogeneous Hamiltonian)
A homogeneous Hamiltonian H is a pair-interaction Hamiltonian of the following form

$$H := \sum_{k<l} w_{kl} \sum_{\alpha\beta} c_{\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} , \tag{5.14}$$

where $W := (w_{kl})$ is a real symmetric $n \times n$-matrix with zeros on the diagonal and $C = (c_{\alpha\beta})$ is a real symmetric $m \times m$-matrix. The matrix W describes the coupling strengths and the signs of the interactions between all qudits. The matrix C characterizes the type of the coupling.

The essence of this definition is that all qudits interact with each other via the same interaction and that only the coupling strengths and the signs vary. It is important that in this special case the coupling matrix J can be written as a tensor product of W and C

$$J = W \otimes C . \tag{5.15}$$

Therefore, the eigenvalues of J are the products of eigenvalues of W and C.
To derive a general lower bound for simulating arbitrary Hamiltonians by a homogeneous Hamiltonian we make use of the rescaling property: each $m \times m$-block at position (k, l) of J and $\tilde{J}$ may be multiplied with the same factor $s_{kl} \in \mathbb{R}$. We

[1] In graph theory one usually considers only non-negative weights. We include also negative weights to describe the signs of the interactions.

may assume that W is a matrix with only 1 as non-diagonal entries; W is the adjacency matrix of the interaction graph of H. Rescaling is described conveniently with the help of the following definition.

Definition 5.16 (Hadamard quotient)
Let $A = (a_{kl})$ be an arbitrary square matrix of size n. We define the support of A, denoted by $\mathrm{supp}(A)$*, to be the set of index pairs such that the corresponding entries of A are not zero, i.e.*

$$\mathrm{supp}(A) = \{(k,l) \mid a_{kl} \neq 0\}\,. \tag{5.16}$$

Let A and B be arbitrary square matrices of size n with $\mathrm{supp}(A) \subseteq \mathrm{supp}(B)$*. We define the Hadamard quotient of A and B, denoted by A/B, to be the matrix $C = (c_{kl})$ with entries given by*

$$c_{kl} = \begin{cases} a_{kl}/b_{kl} & \text{if } (k,l) \in \mathrm{supp}(A) \\ 0 & \text{if } (k,l) \notin \mathrm{supp}(A) \end{cases} \tag{5.17}$$

We denote by I the all-one-matrix of size m, i.e. the matrix whose all entries are 1.

Remark 5.17 Let $J = W \otimes C$ be the coupling matrix of a homogeneous Hamiltonian. Then we may consider the simulation of $\tilde{J}/(W \otimes I)$ by $A \otimes C$, where A is a matrix having only 0 and 1 as entries ($A := W/W$). The matrix A is the adjacency matrix of the interaction graph.

In the following theorem we derive a general lower bound on the number of time steps. To do that we have to define positive (positive semidefinite) and to state some basis inequalities. A Hermitian matrix A is called positive (positive semidefinite) if

$$\langle\Psi|A|\Psi\rangle > 0 \qquad (\,\langle\Psi|A|\Psi\rangle \geq 0\,) \quad \text{for all } |\Psi\rangle\,,$$

or equivalently if all its all its eigenvalues are positive (non-negative). Furthermore, we have

$$\lambda_{\min}(A) \leq \langle\Psi|A|\Psi\rangle \leq \lambda_{\max}(A) \tag{5.18}$$

for all unit vectors $|\Psi\rangle$.
Let A and B be two arbitrary Hermitian matrices (of the same size). We have the following inequalities for the maximal and minimal eigenvalues:

$$\lambda_{\max}(A+B) \leq \lambda_{\max}(A) + \lambda_{\max}(B) \tag{5.19}$$
$$\lambda_{\min}(A+B) \geq \lambda_{\min}(A) + \lambda_{\min}(B)\,. \tag{5.20}$$

These inequalities are special cases of the inequalities in Bhatia [Bha96, Theorem III.2].

Lemma 5.18 *Let A and B be two arbitrary Hermitian matrices and $\mathcal{I}$ be the interval $[-r,-q]$, where q and r are the minimal and maximal eigenvalues of B, respectively. Then the rank of the sum $A+B$ is at least the number of eigenvalues of A outside $\mathcal{I}$.*

Proof. Let P and Q be the projections onto the sums of all eigenspaces of A with eigenvalues smaller than $-r$ and greater than $-q$, respectively, and s be the number of eigenvalues of A outside $\mathcal{I}$. Clearly, s is equal to the dimension of the image of $P \oplus Q$. Denote by $\tilde{A}$ and $\tilde{B}$ the $s \times s$-submatrices of A and B defined by $(P \oplus Q)A(P \oplus Q)$ and $(P \oplus Q)B(P \oplus Q)$, respectively.

Due to the choice of P and Q the spectrum of $\tilde{A}$ is contained in the interval $(-\infty,-r) \cup (-q,\infty)$. The spectrum of $\tilde{B}$ is contained in the interval $[q,r]$ since the minimal (maximal) eigenvalues of a matrix cannot decrease (increase) when projecting the matrix.

We prove the lemma by showing that $\tilde{A}+\tilde{B}$ has full rank. Set $\lambda := \frac{1}{2}(q+r)$. From the triangle inequality it follows for every unit vector $|\Psi\rangle \in \mathbb{R}^s$ that

$$\begin{aligned} \|(\tilde{A}+\tilde{B})|\Psi\rangle\| &= \|(\tilde{A}+\lambda\mathbf{1}+\tilde{B}-\lambda\mathbf{1})|\Psi\rangle\| \\ &\geq \|(\tilde{A}+\lambda\mathbf{1})|\Psi\rangle\| - \|(\tilde{B}-\lambda\mathbf{1})|\Psi\rangle\|\,. \end{aligned}$$

The eigenvalues of the shifted operators $\tilde{A}+\lambda\mathbf{1}$ and $\tilde{B}-\lambda\mathbf{1}$ are contained in the intervals $\big(-\infty,-\frac{1}{2}(r-q)\big) \cup \big(\frac{1}{2}(r-q),\infty\big)$ and $[-\frac{1}{2}(r-q),\frac{1}{2}(r-q)]$, respectively. This implies that

$$\|(\tilde{A}+\lambda\mathbf{1})|\Psi\rangle\| - \|(\tilde{B}-\lambda\mathbf{1})|\Psi\rangle\| > 0$$

because the norms can be bounded by the absolute values of the eigenvalues: $\|(\tilde{A}+\lambda\mathbf{1})|\Psi\rangle\| > |\frac{1}{2}(r-q)|$ and $\|(\tilde{B}-\lambda\mathbf{1})|\Psi\rangle\| \leq |\frac{1}{2}(r-q)|$. □

By applying this lemma to the coupling matrices we obtain the following lower bound on the number of time steps.

Theorem 5.19 (Lower bound on number of time steps)
Let $J := W \otimes C$ be the coupling matrix of a homogeneous Hamiltonian H with complete interaction graph, $\tilde{J}$ an arbitrary coupling matrix of the Hamiltonian $\tilde{H}$ that we want to simulate. We denote by

1. *$\lambda_{\min}$ and $\lambda_{\max}$ the minimal and maximal eigenvalues of C, respectively,*
2. *r the rank of C,*
3. *s the number of eigenvalues of $\tilde{J}/(W \otimes I)$ that are not contained in the interval*

$$\mathcal{I} := [-\tau\lambda_{\max}, -\tau\lambda_{\min}]\,,$$

where τ is a positive real number.

If there is a control sequence that simulates $\tilde{H}$ by H with time overhead τ then it has at least s/r time steps.

Proof. The condition for a control sequence $(\tau_1, O_1; \tau_2, O_2; \ldots; \tau_N, O_N)$ is

$$\sum_{j=1}^{N} \tau_j O_j (W \otimes C) O_j^T = \tilde{J}\,. \tag{5.21}$$

Since the matrices O_j are block-diagonal we can rescale each $m \times m$-block in eq. (5.21) such that we obtain

$$\sum_{j=1}^{N} \tau_j O_j (K_n \otimes C) O_j^T = \tilde{J}/(W \otimes I_m)\,, \tag{5.22}$$

where K_n is the adjacency matrix of the complete graph with n vertices. We denote the rescaled coupling matrix $\tilde{J}/(W \otimes I_m)$ by J'. Set $R := \sum_{j=1}^{N} \tau_j O_j(\mathbf{1}_n \otimes C) O_j^T$. Adding R on both sides of eq. (5.22) gives

$$\sum_{j=1}^{N} \tau_j O_j \big(I_n \otimes C\big) O_j^T = J' + R\,. \tag{5.23}$$

The rank of the matrix I_n is 1 since all its entries are 1. Consequently, the rank of the l.h.s. of eq. (5.23) is at most $N\,r$.

It follows from the inequalities (5.19) and (5.20) that the eigenvalues of R are contained in the interval $[\tau\lambda_{\min}, \tau\lambda_{\max}]$. We now apply Lemma 5.18 to the sum $J' + R$. It follows that the rank of $J' + R$, r.h.s. of eq. (5.23), is at least s, the number of eigenvalues of J' outside the interval $\mathcal{I}$.

By combining the bounds on the ranks of both sides of eq. (5.23), we obtain the lower bound on the number of time steps $N \geq s/r$. □

There are interesting cases where the bound of Theorem 5.19 can be tightened. Assume $\tilde{J} = \tilde{W} \otimes C$ with the same matrix C as the interaction that is used for the simulation. In other words, only the strengths and the signs of some interactions should be changed. *Selective decoupling* is a special case where we want to cancel certain interactions without changing the others. An example is shown in Figure 5.1. Starting from the complete interaction graph (1) we must cancel certain interactions to obtain the interactions graphs (2) and (3).

Theorem 5.20 (Lower bound on number of time steps)
Let $W \otimes C$ be the coupling matrix of the system Hamiltonian and $\tilde{W} \otimes C$ the coupling matrix of the Hamiltonian that we want to simulate. Assume all non-diagonal entries of W to be non-zero.

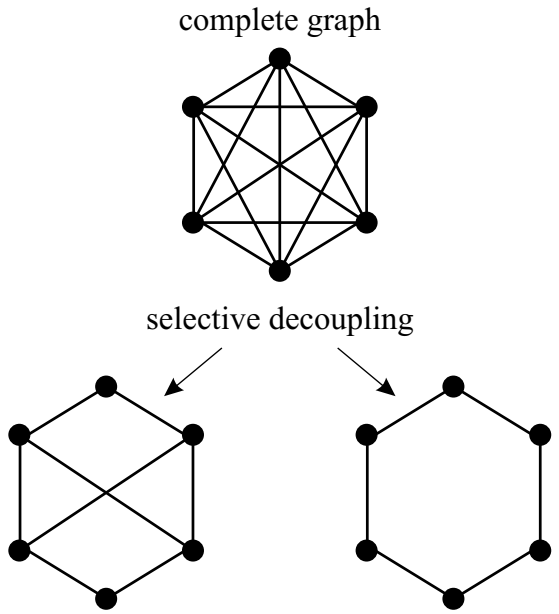


Figure 5.1: Selective decoupling

1. *If C is a positive semidefinite matrix, then the number of time steps is at least the number of positive eigenvalues of $A := \tilde{W}/W$.*

2. *If C is the identity matrix, then the number of time steps is at least $n - k$, where k is the multiplicity of the smallest eigenvalue $\mu_{\min}$ of $A := \tilde{W}/W$.*

Proof. To prove the statements we consider the rescaled problem to simulate $A \otimes C$ by $K \otimes C$, where $A := \tilde{W}/W$.

If C is positive semidefinite, then the interval $\mathcal{I} = [-\tau\lambda_{\max}, -\tau\lambda_{\min}]$ does not contain positive numbers (since $\lambda_{\max}, \lambda_{\min} \geq 0$ and the time overhead τ is always positive). The first statement readily follows from Theorem 5.19 since the number of positive eigenvalues of $A \otimes C$ is r times the number of positive eigenvalues of A, where r denotes the rank of C.

Assume that C is the identity matrix $\mathbf{1}_m$. In this case the r.h.s. of eq. (5.23) reduces to

$$A \otimes \mathbf{1}_m + \tau \mathbf{1}_n \otimes \mathbf{1}_m \,, \tag{5.24}$$

where $\tau := \sum_{j=1}^{N} \tau_j$ is the time overhead. Note that the time overhead τ is at least $-\mu_{\min}$, where $\mu_{\min}$ is the minimal eigenvalue of A. The reason is that the matrix in eq. (5.24) is necessarily positive semidefinite since it is a sum of positive semidefinite matrices given by the l.h.s. of eq. (5.23). Therefore its rank is at least $m(n-k)$. Since the l.h.s. of eq. (5.23) has at most the rank $m\,N$ the number of time steps is at least $n - k$. □

5.2 Upper bounds

The following upper bound can be easily derived from Carathéodory's theorem (cf. ROCKAFELLER [Roc70]).

Theorem 5.21 (Upper bound on the number of time steps)
Every simulation that is possible can be achieved within at most

$$\frac{n(n-1)}{2}m^2+1$$

time steps.

Proof. If $\tilde{J}$ can be simulated by J with time overhead τ then $\tilde{J}/\tau$ is in the convex span of the matrices $O_j J O_j^T$ with notation as above. The dimension of this convex set is at most $m^2 n(n-1)/2$ since the diagonal blocks are empty and each matrix $O_j J O_j^T$ is symmetric. Carathéodory's theorem states that each point in an M dimensional convex set can be written as a convex sum of at most $M+1$ extreme points. □

In the next chapter we derive upper bounds on the time overhead and the number of time steps for special Hamiltonians in terms of invariants of the interaction graphs.

5.3 Spin-off: bounds for graph properties

We have seen that chromatic number and clique coloring index of graphs associated to Hamiltonians give upper bounds on the complexity of decoupling, time-reversal and simulating Hamiltonians. Lower bounds on the complexity are given in terms of the eigenvalues of the adjacency matrices. By abstracting from the physical problem of simulating Hamiltonians we derive in this chapter new lower bounds on chromatic number and clique coloring index. The results presented here are based on our work WOCJAN ET AL. [WJB02a].

Let G be a graph on n vertices. Recall that the chromatic number $\chi(G)$ of a graph G is the smallest integer k such that the vertices of G can be k-colored so that any two adjacent vertices have different colors (see Definition 3.36). The problem to determine the chromatic number is NP-complete (cf. GAREY AND JOHNSON [GJ79]).

Bounds are known for $\chi(G)$ in terms of the eigenvalues of the adjacency matrix $A=(a_{kl})$ of G. Besides adjacency matrices there are various matrices that are naturally associated with a graph, such as the Seidel matrix and the Laplacian. One of the main problems of algebraic (or spectral) graph theory is to determine

precisely how, or whether, properties of graphs are reflected in the algebraic properties of such matrices (cf. CVETKOVIĆ [CDS95], CHUNG [Chu97], and GODSIL AND ROYLE [GR01]).

Let us recall two well-known bounds in terms of adjacency matrices. Denote by $\lambda_{\min}$ and $\lambda_{\max}$ the smallest and largest eigenvalues of A, respectively. An upper bound of $\chi(G)$ was first given by WILF [Wil67] who showed that

$$\chi(G) \leq 1 + \lambda_{\max}\,.$$

For the lower bound, HOFFMAN [Hof70] proved (cf. HAEMERS [Hae95])

$$\chi(G) \geq 1 - \frac{\lambda_{\max}}{\lambda_{\min}}\,. \tag{5.25}$$

The following theorem proved by BARNES [Bar01] gives a generalization of the lower bound in eq. (5.25).

Theorem 5.22 *Let A denote the adjacency matrix for a connected graph G on n vertices, and let $D = \mathrm{diag}(d_1, \ldots, d_n)$ be a diagonal matrix such that $A + D$ is positive semidefinite. Then each d_i is positive, and the largest eigenvalue of the matrix $D^{-\frac{1}{2}}AD^{-\frac{1}{2}} + \mathbf{1}$ is a lower bound for χ, where $\mathbf{1}$ is the identity matrix of size n.*

Note that the lower bound in eq. (5.25) can be obtained from Theorem 5.22 by taking $D = |\lambda_{\min}|\mathbf{1}$.

The second graph invariant we consider in this chapter is the clique coloring index (cf. WALLIS AND ZHANG [WZ93]). We introduce the necessary notations. A clique of G is a complete subgraph of G, i.e., all vertices are adjacent in the subgraph. A clique of G is called a maximal clique of G if it is not properly contained in another clique of G. A clique partition P of G is a partition of $E(G)$ (the edges) such that its classes induce maximal cliques. Given a set C of k colors, an k-coloring of P in G is a mapping from P to C, such that cliques sharing a vertex have different colors. Let the clique coloring index $\kappa(G)$ be the smallest k such that there is a partition P permitting an k-coloring. We say the graph G consists of independent cliques if $\kappa(G) = 1$.

Upper and lower bounds on the clique coloring index are given for several graphs in [WZ93]; in some cases the exact value is determined.

5.3.1 Majorization of graph spectra

To derive lower bounds on the chromatic number χ and the clique coloring index κ we use basic results of majorization theory that are summarized in Chapter 5. Motivated by Uhlmann's theorem (Theorem 5.3) we define conversion maps for Hermitian matrices:

Definition 5.23 *Let A and B be two traceless Hermitian matrices. We say that A can be converted to B with cost τ and N steps if there are N positive real numbers τ_j summing to τ and N unitary matrices U_j such that*

$$\sum_{j=1}^{N} \tau_j U_j^* A U_j = B\,. \tag{5.26}$$

The following lemma gives the minimal cost and a lower bound on the number of steps.

Lemma 5.24 *Let A and B be two traceless Hermitian matrices. The optimal cost to convert A to B is the minimal non-negative real number τ such that*

$$B \prec \tau A\,. \tag{5.27}$$

It is gives explicitly by

$$\tau := \max_{m=1,\ldots,n-1} \left\{ \frac{\sum_{i=1}^{m} \lambda_i(A)}{\sum_{i=1}^{m} \lambda_i(B)} \right\} \tag{5.28}$$

where $\lambda_1(A), \ldots, \lambda_n(A)$ and $\lambda_1(B), \ldots, \lambda_n(B)$ denote the eigenvalues sorted in non-increasing order of A and B, respectively.
Let $r(A)$ and $r(B)$ denote the rank of A and B, respectively. Then

$$\frac{r(B)}{r(A)} \leq N \tag{5.29}$$

is a lower bound on the number of steps required to convert A to B.

Proof. The minimal cost follows directly from Uhlmann's Theorem. Note that the rank of the sum in eq. (5.26) is at most $Nr(A)$. This observation gives the lower bound. □

The above lemma refers to arbitrary traceless Hermitian matrices. To derive lower bounds on the chromatic number and the clique coloring index we apply this lemma to special matrices constructed from adjacency matrices. For this construction we need the definition of the Hadamard product (cf. Bhatia [Bha96]). If $A = (a_{ij}) \in \mathbb{C}^{n\times n}$ and $B = (b_{ij}) \in \mathbb{C}^{n\times n}$ are given, then the Hadamard product of A and B is the matrix $A * B = (a_{ij} b_{ij})$.

Let G be a graph with adjacency matrix A. Weighted adjacency matrices of G are Hadamard products of the form $W * A$, where W are arbitrary Hermitian matrices.

5.3.2 Lower bounds on the chromatic number

We consider the problem of reversing the sign of a weighted adjacency matrix of G and show that $\chi(G)-1$ is an upper bound on the cost. Combining this upper bound together with the optimal cost (derived in Lemma 5.24) yields a lower bound on the chromatic number.

The following lemma shows that the chromatic number gives an upper bound on the cost of conversion for two special cases.

Lemma 5.25 *Let G be an arbitrary graph with chromatic number χ. Denote by $A=(a_{kl})$ the adjacency matrix of G. We can always convert $W*A$ to the zero matrix $\mathbf{0}$ with χ time steps and $W*A$ to $-W*A$ with cost $\chi-1$ for all Hermitian matrices W.*

Proof. Let $V=\{1,\ldots,n\}$ be the set of vertices of G. Choose a partition of the vertices $V=V_1\cup\ldots\cup V_\chi$ corresponding to a (minimal) coloring. The set V_c contains all vertices of color c ($c\in\{1,\ldots,\chi\}$). We assume that the vertices are ordered according to their colors, i.e., first come the vertices of color 1, then of color 2, etc.

Set $M:=W*A$. Let $\omega\in\mathbb{C}$ be a primitive χ-th root of unity, i.e., $\chi^j\neq 1$ for $j=1,\ldots,\chi-1$ and $\omega^\chi=1$. Define the diagonal matrix

$$D:=\mathrm{diag}(\omega^{c_1},\ldots,\omega^{c_n})\,,$$

where c_k is the color of the vertex k, i.e., $k\in V_{c_k}$. We show that

$$\bar{M}:=\sum_{j=1}^{\chi}D^{-j}MD^j\,. \tag{5.30}$$

is the zero matrix $\mathbf{0}$. Let $M=(m_{kl})_{k,l=0,\ldots n-1}$ and $\bar{M}=(\bar{m}_{kl})_{k,l=0,\ldots n-1}$. The entries of $\bar{M}$ are given by

$$\bar{m}_{kl}=\sum_{j=1}^{\chi}(\omega^{-c_k})^j m_{kl}(\omega^{c_l})^j\,.$$

Let $k,l\in V_c$ with $k\neq l$. The vertices k and l have the the same color and consequently they cannot be adjacent, i.e., $a_{kl}=0$. Therefore we have $\bar{m}_{kl}=0$ since $m_{kl}=w_{kl}a_{kl}=0$. Note that $\bar{m}_{kk}=0$ since $a_{kk}=0$ (the diagonal entries of the adjacency matrix are all zero).

Now let $k\in V_c$ and $l\in V_{\tilde{c}}$ with $c\neq\tilde{c}$. We have

$$\bar{m}_{kl}=\sum_{j=1}^{\chi}\omega^{-cj}m_{kl}\omega^{\tilde{c}j}=0$$

since the vectors $(\omega^{c\,1}, \omega^{c\,2}, \ldots, \omega^{c\,\chi})$ and $(\omega^{\tilde{c}\,1}, \omega^{\tilde{c}\,2}, \ldots, \omega^{\tilde{c}\,\chi)})$ are orthogonal. Note that they are rows of the matrix of the discrete Fourier transform of size χ.

By letting the sum in eq. (5.30) run to $j = \chi - 1$, we see that $W * A$ can be converted to $-W * A$ with cost $\chi - 1$ since D^χ is the identity matrix (take $\tau_j = 1$ for $j = 1, \ldots, \chi - 1$). □

By combining this lemma with Lemma 5.24 we obtain the following lower bound on the chromatic number.

Theorem 5.26 *Let G be a graph on n vertices with chromatic number χ. Denote by A its adjacency matrix. Then we have*

$$\chi \geq 1 + \max_{W} \max_{m=1,\ldots,n-1} \left\{ \frac{\sum_{i=1}^{m} \lambda_i(W * A)}{-\sum_{i=1}^{m} \lambda_{n+1-i}(W * A)} \right\}, \tag{5.31}$$

where W ranges over all Hermitian matrices and $\lambda_i(W * A)$ are the eigenvalues of the weighted adjacency matrix $W * A$ sorted in non-increasing order.

Proof. We consider the problem to convert $W * A$ to $-W * A$ for all Hermitian matrices W. Lemma 5.25 shows that this can be done with cost $\chi - 1$. Lemma 5.24 gives

$$\max_{m=1,\ldots,n-1} \left\{ \frac{\sum_{i=1}^{m} \lambda_i(W * A)}{-\sum_{i=1}^{m} \lambda_{n+1-i}(W * A)} \right\}$$

as optimal cost since $\lambda_{n+1-i}(W*A)$ are the eigenvalues of $-W*A$ in non-increasing order. Both results imply the lower bound in eq. (5.31). □

Note that we obtain as a special case the well-known lower bound $\chi \geq \frac{\lambda_1}{|\lambda_n|} + 1$ in eq. (5.25). Set W to be the matrix all of whose entries are equal to 1. Then we have $A = W * A$. Instead of taking the maximum over $m = 1, \ldots, n-1$ consider only $m := 1$ in eq. (5.31). This corresponds to the maximal and minimal eigenvalues of A.

Furthermore, Theorem 5.22 can also be understood as a special case of Theorem 5.26. Let $D := \mathrm{diag}(d_1, \ldots, d_n)$ be the diagonal matrix defined as in Theorem 5.22. Take W to be the matrix with entries $w_{kl} := d_k^{-\frac{1}{2}} d_l^{-\frac{1}{2}}$. Then we have $W * A = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$. This shows that this modified adjacency matrix can also be expressed with the Hadamard product.

Theorem 5.26 permits to consider a larger class of modified adjacency matrices and to take into account all eigenvalues (Theorem 5.22 considers only the maximal and minimal eigenvalues). The advantage of the method presented in [Bar01] is that the matrix D is the solution of a semidefinite programming problem that can be computed by an algorithm given there. It remains to be shown whether the larger class of modified adjacency matrices permits to obtain better lower bounds

and whether there are efficient algorithms to compute them. Nevertheless, the derivation in our approach is intuitive and gives a generalization of Theorem 5.22.

The next theorem gives another lower bound on the chromatic number; it also follows from Lemma 5.24 and Lemma 5.25.

Theorem 5.27 *Let G be a graph on n vertices with chromatic number χ. Denote by A its adjacency matrix. Then we have*

$$\chi \geq \max_{W} \max_{\lambda \in \mathrm{Spec}(W*A)} \left\{ \frac{n}{n - m_\lambda} \right\}, \tag{5.32}$$

where W ranges over all Hermitian matrices and m_λ denotes the multiplicity of the eigenvalue λ of the weighted adjacency matrix $W * A$.

Proof. We consider the problem to convert $W * A$ to $\mathbf{0}$. By Lemma 5.25 we can do this with χ steps. Now we derive a lower bound on the number of steps. By Uhlmann's theorem we know that there is a probability distribution $\{p_j\}$ on N outcomes and N unitary matrices U_j such that

$$\sum_{j=1}^{N} p_j U_j^* (W * A) U_j = \mathbf{0}\,. \tag{5.33}$$

To derive a lower bound on N we add a scalar multiple of the identity matrix $\lambda\mathbf{1}$ (where λ is an eigenvalue of $W * A$) to $W * A$ and obtain:

$$\sum_{j=1}^{N} p_j U_j^* (W * A + \lambda\mathbf{1}) U_j = \lambda\mathbf{1}\,. \tag{5.34}$$

The rank of the l.h.s. is at most $N(n - m_\lambda)$ while the rank of the r.h.s. is n. Therefore $\frac{n}{n-m_\lambda}$ is a lower bound on the number of steps. This implies the lower bound on the chromatic number. □

5.3.3 Lower bound on the clique coloring index

The ideas of the previous section can also be used to derive a lower bound on the clique coloring index. Denote by K the adjacency matrix of the complete graph.

Lemma 5.28 *Let G be a graph with clique coloring index κ. Then we can convert $W * K$ to $W * A$ with cost κ for all Hermitian matrices.*

Proof. Let $E = E_1 \cup \cdots \cup E_\kappa$ be the corresponding partition of the edges such that all subgraphs $G_c := (V, E_c)$ consist of independent cliques for $c = 1, \ldots, \kappa$.

To prove the upper bound it suffices to show that $W * A_c$ can be obtained from $W * K$ with cost 1, where A_c is the adjacency matrix of the subgraph G_c. We consider one fixed G_c.

Let m be the number of cliques of G_c and $V = V_1 \cup \cdots \cup V_m$ be the corresponding partition of the vertices. Let $\omega \in \mathbb{C}$ be a mth-root of unity. Define a diagonal matrix $D := \operatorname{diag}(d_1, \ldots, d_n)$, where $d_k := \omega^h$ if $k \in V_h$. Then we have

$$\frac{1}{m} \sum_{j=1}^{m} D^{-j}(W * K)D^j = W * A_c \,.$$

Repeating this procedure for all $c = 1, \ldots, \kappa$ yields the upper bound. □

By combining the above lemma with Lemma 5.24 we obtain the following lower bound on the clique coloring index.

Theorem 5.29 *Let G be a graph with clique coloring index κ. Denote by A its adjacency matrix. Then we have*

$$\kappa \geq \max_{W} \max_{m=1,\ldots,n-1} \left\{ \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{m} \mu_i} \right\} , \tag{5.35}$$

where W ranges over all Hermitian matrices and $\mu_1, \ldots, \mu_n$ and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues sorted in non-increasing order of $W * K$ and $W * A$, respectively.

Proof. Lemma 5.28 show that we can convert $W * K$ to $W * A$ with cost κ. Lemma 5.24 gives

$$\max_{m=1,\ldots,n-1} \left\{ \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{m} \mu_i} \right\}$$

as a lower bound on the cost of this conversion. Combining both results gives the lower bound on the clique coloring index. □

Note that by choosing $W := -K$ and considering only $m = 1$, we obtain

$$\kappa \geq -\lambda_{\min}(A) \,.$$

Chapter 6

Special simulation tasks and their complexity

6.1 Time-reversal and decoupling

Due to wide applications of nuclear magnetic resonance (NMR) in medicine, chemistry, and physics much effort has been undertaken to understand the dynamics of nuclear spins in solids and liquids (cf. Maciel [Mac94] and Vlaardingbroeck and den Boer [VdB96]). Nowadays, NMR plays an important role in the first experimental realizations of quantum computation (cf. Nielsen and Chuang [NC00]).

Phenomenologically, the dynamical evolution of a single spin seems to be a relatively simple relaxation and dephasing process of an open quantum system, formally described by the Bloch equations (cf. Ernst et al. [EBW87] and Slichter [Sli90]). However, the fact that not all dephasing and relaxation processes are irreversible is important for practical purposes since refocusing techniques are applied with great success (cf. Rhim et al. [RPW70], Haeberlen [Hae76], [EBW87], [Sli90]).

The simplest version of refocusing is possible when dephasing of the (uncoupled) spins is caused by spatial inhomogeneities of the strength of the magnetic field. In this case the time evolution can be reversed by sandwiching the natural dynamics by 180°–rotations of all spins around an axis orthogonal to the magnetic field. More sophisticated versions of refocusing are necessary if dephasing and relaxation processes are caused by spin-spin interactions. An important example is the so-called dipole-dipole coupling (cf. [RPW70]).

The so-called truncated dipole-dipole Hamiltonian H_d is given by

$$H_d := \sum_{k<l} w_{kl} \left(\sum_{\alpha} \sigma_\alpha^{(k)} \sigma_\alpha^{(l)} - 3\sigma_z^{(k)} \sigma_z^{(l)} \right),$$

where w_{kl} is the strength of the interaction between the spins k and l.

Let $V_y \in K$ a control operation that realizes a rotation of each spin around the y-axis by 90°, i.e., if the spin is in z-direction it is in x-direction after the rotation. Formally, we have a unitary transformation of the form

$$V_y := U_y \otimes U_y \otimes \cdots \otimes U_y \tag{6.1}$$

with $U_y^\dagger \sigma_z U_y = \sigma_x$.

We choose a second rotation U_x around the x-axis such that $U_x^\dagger \sigma_z U_x = \sigma_y$ and set $V_x := U_x \otimes U_x \otimes \cdots \otimes U_x$. It is decisive to note that

$$-H_d = V_y^\dagger H_d V_y + V_x^\dagger H_d V_x \,.$$

We have thus constructed a time-reversal scheme for the dipole-dipole Hamiltonian that is simple with respect to three aspects: (a) it does not use selective pulses, i.e., in each time step the same unitary transformation is applied to all spins. (b) The number of steps of this time-reversal scheme does not increase with n since it is always 2, and (c) the time overhead does not increase with n since it is also 2.

In this section we show that there are interactions that are considerably more complex to reverse with respect to all three criteria and give a characterization of the complexities of decoupling and time-reversal in qubit networks. These results were derived in Janzing, Wocjan and Beth [JWB02b].

Note that it is sufficient to restrict our attention to time-reversal schemes in order to study the complexity of both decoupling and time-reversal. A time-reversal scheme with N' time steps defines a decoupling scheme with $N'+1$ steps simply by appending the identity (recall that the notion of time overhead does not make sense for decoupling). Conversely, each decoupling scheme can be converted to a time-reversal scheme as follows. The equation

$$\sum_{j=1}^{N} p_j U_j^\dagger H U_j = 0 \,, \quad U_j \in K$$

implies

$$\sum_{j=2}^{N} (p_j/p_1)(U_j U_0^\dagger)^\dagger H (U_j U_0^\dagger) = -H$$

by elementary calculation.

Time-reversal and decoupling schemes that apply to general Hamiltonians have been presented in Chapter 3. These schemes are universal since they can be used even if the system Hamiltonian is unknown. Here we focus on optimality criteria for schemes for given Hamiltonians. We will see that there are Hamiltonians that cannot be reversed significantly more efficiently than unknown Hamiltonians.

We restrict our attention to homogeneous Hamiltonians

$$H = \sum_{k<l} w_{kl} \sum_{\alpha\beta} c_{\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} \,. \tag{6.2}$$

Recall that the matrix $W := (w_{kl})$ is a real symmetric matrix $n \times n$-matrix (with zeros on the diagonal) describing the coupling strengths and the signs of the interactions between all spins. The matrix $C = (c_{\alpha\beta})$ is a real symmetric 3×3-matrix characterizing the type of the coupling. This means that all spins interact with each other via the same coupling and that only the coupling strengths and the signs vary. The coupling matrix of H can be written as a tensor product of W and C, i.e., $J = W \otimes C$.

Theorem 6.1 (Complexity of time-reversal)
Let $J := W \otimes C$ be the coupling matrix of a complete homogeneous Hamiltonian of n coupled qubits. To discuss the complexity of time reversal we distinguish between the following three cases:

1. *C is traceless.*

   All spins can be subjected to the same transformations in each time step, the number of time steps and the time overhead are at most 2.

2. *C has negative and positive eigenvalues but $\mathrm{tr}(C) \neq 0$.*

   The spins have to be addressed separately, the number of time steps is at least $n/3 - 1$. But the time overhead does not depend on n. It depends only on the eigenvalues of C.

3. *C is either positive or negative semidefinite, i.e., the non-zero eigenvalues have the same sign.*

   Then the spins have to be addressed separately, the number of time steps is at least $n - 1$, and the time overhead is also at least $n - 1$.

Proof. To prove these statements we assume w.l.o.g. that the interaction between all pairs is of the form

$$c_x \sigma_x \otimes \sigma_x + c_y \sigma_y \otimes \sigma_y + c_z \sigma_z \otimes \sigma_z \,,$$

where c_x, c_y, c_z are the eigenvalues of C. This can always be achieved by rotating the reference frame as shown in Lemma 5.10.

Case 1. Assume C to be traceless. Let S be a rotation on the Bloch sphere that realizes the cyclic permutation of the axis according to $x \mapsto y \mapsto z \mapsto x$. Then we have

$$\sum_{j=0}^{2} O^j J {O^j}^T = \mathbf{0} \,.$$

where $O := S \oplus S \oplus \cdots \oplus S$ and O^j is the jth power of O. Hence the Hamiltonian is inverted by a sequence of length 2. Due to the equation

$$-J = OJO^T + O^2JO^{2T}$$

the time overhead of this inversion scheme is 2.

Case 2. The fact that the spins have to be addressed separately has been also noted in Masanes et al. [MVL02]. This can be seen as follows. If all spins are subjected to the same transformation then all 3×3–blocks of J are conjugated by the same element of $SO(3)$. This conjugation preserves the trace of each block. Positive linear combination preserves the sign of the trace of each block. Therefore, one can never obtain the result $-C$ in any block.

To prove the minimal number of time steps for time-reversal and decoupling schemes it is useful to recall that such schemes for a given Hamiltonian can also be applied to a rescaled Hamiltonian obtained by changing the strength and the sign of an arbitrary spin pair-interaction. If all the coefficients w_{kl} are non-vanishing for $k \neq l$ then we can equivalently describe a time-reversal or decoupling scheme for the coupling matrix $J := K \otimes C$ where all non-diagonal entries of K are 1 and all diagonal ones are 0 (i.e. the adjacency matrix of the complete graph). Hence the fact that interactions decrease with distance of the spins is irrelevant for considerations of the complexity of time-reversal (as long as the interaction cannot be neglected). Therefore we assume w.l.o.g. that all non-diagonal entries of W to be 1.

Now we show that the number of time steps for a decoupling scheme is at least $n/3$. Let $\tau_1, \tau_2, \ldots, \tau_N$ be the (relative) times and $O_1, O_2, \ldots, O_N$ with $O_j \in \bigoplus_{k=1}^{n} SO(3)$ be the operations of a decoupling scheme, i.e.,

$$\sum_{j=1}^{N} \tau_j O_j (K \otimes C) O_j^T = \mathbf{0}\,. \tag{6.3}$$

Set $R := \sum_{j=1}^{N} \tau_j O_j (\mathbf{1} \otimes C) O_j^T$. We add R to both sides of eq. (6.3) and obtain

$$\sum_{j=1}^{N} \tau_j O_j \big((I \otimes C)\big) O_j^T = R\,. \tag{6.4}$$

The matrix I has only 1 as entries and has therefore rank 1. The l.h.s. of eq. (6.4) has therefore rank Nr at most, where $r \leq 3$ is the rank of C. The rank of the r.h.s is at least n. This can be seen as follows. Each matrix O_j as well as $\mathbf{1} \otimes A$ are block-diagonal. Conjugating $\mathbf{1} \otimes C$ by O_j corresponds to a conjugation of each block by orthogonal 3×3–matrices. Each 3×3–block of R is hence a positive sum of matrices that are orthogonally equivalent to C and their traces have the same sign as the trace of C. Hence they cannot add up to the zero matrix in any

of the n blocks. We conclude that the number N of time steps is at least $n/3$ and $N \geq n/3 - 1$ for time-reversal.

We show that the time overhead does not depend on n. We assume w.l.o.g. $c_x > 0 > c_z$ for the eigenvalues of C. First we describe a partial decoupling scheme selecting for instance the $\sigma_z \otimes \sigma_z$ coupling terms while switching off the xx and yy terms. Such a partial decoupling can be achieved by certain sequences of local conjugations by the unitary σ_z. Each time step of this scheme is described by a column of a Hadamard matrix (see Chapter 3). The entries determine which spins are conjugated by σ_z transformations. Recall that the idea is that the xx and yy terms acquire in exactly half of the time steps a minus sign. Note that this scheme does not weaken the zz-terms. In other words, although this scheme requires a number of *time steps* of the order n the *time overhead* for the simulation of the interaction $\sum_{k<l} w_{kl} c_z \sigma_z^{(k)} \sigma_z^{(l)}$ by the system Hamiltonian H is 1.

There are two cases depending on whether $c_y \geq 0$ or $c_y < 0$. Consider w.l.o.g. the case $c_y \geq 0$. First switch off the xx and yy interactions. Apply on each spin a conjugation by V_y as defined in eq. (6.1). This simulates the interaction

$$\sum_{k<l} w_{kl}\, c_z\, \sigma_x^{(k)} \sigma_x^{(l)} .$$

By applying this interaction for the (relative) time $-c_x/c_z$ we have reversed the xx-components of the original Hamiltonian. Since c_y is not negative we can reverse similarly the yy-components with time overhead $-c_y/c_z$. To invert the zz-components we use the xx-components. The time overhead is $-c_z/c_x$. In summary we see that the time overhead is independent of n. It only depends on the eigenvalues of C.

Case 3. If C is positive or negative semidefinite a similar argument as in Case 2 can be used to tighten the bound on the number of time steps. In this case the rank of each 3×3 block on the r.h.s. of eq. (6.4) is at least r. Hence the number of time steps for decoupling is at least n and for time-reversal at least $n - 1$.

To prove the minimal time overhead we make use of Lemma 5.13. Let r and q be the maximal and minimal eigenvalues of J, respectively. Since the maximal and minimal eigenvalues of K are $(n-1)$ and -1, respectively, the maximal and minimal eigenvalues of J are $(n-1)c_{\max}$ and $-c_{\max}$, where $c_{\max}$ is the maximal eigenvalue of C. Hence the minimal time overhead is at least $n - 1$. □

The truncated dipole-dipole coupling is an example for Case 1. The corresponding matrix C is given by $C := \mathrm{diag}\,(1, 1, -2)$. An important example for Case 3 is the strong scalar coupling where A is the identity matrix. By combining both types one can easily find examples for Case 2 (cf. Luy and Glaser [LG01]).

Besides of their practical application complexity bounds on time-inversion are also interesting from the fundamental point of view. Consider a Hamiltonian evolution that creates large-scale entanglement between a large number n of spins.

If all algorithms that disentangle the system again could be proven to be rather complex, than the created (phenomenological) entropy might be interpreted as *algorithmic entropy*. Note that the relevance of complexity theory for the definition of physical entropy has been advocated for by several authors (cf. ZUREK [Zur89, Zur90] and LI AND VITANYI [LV93]).

Remark 6.2 For n coupled *qudits* we do not have derived such a hierarchy of complexity since the symmetric matrix C cannot be diagonalized as in the case of qubits. However, if C is semi-positive one shows by similar arguments that $n-1$ is a lower bound on both the time overhead and the number of time steps of time-reversal.

6.2 Simulation of Hamiltonians with complete Ising-Hamiltonians

In this section we consider the problem to simulate arbitrary qubit Hamiltonians $\tilde{H}$ by an Ising-Hamiltonian H with complete interaction graph, i.e.,

$$H = \sum_{k<l} w_{kl}\, \sigma_z^{(k)} \sigma_z^{(l)} \tag{6.5}$$

and $w_{kl} \neq 0$ for all $k < l$. The control operations are elements of the control group $\mathcal{K} = SU(2) \otimes SU(2) \otimes \cdots \otimes SU(2)$.

The interesting aspect of this problem is that concepts of graph theory helps substantially to to devise efficient control schemes for simulating Hamiltonians. We show how properties of the interaction graphs determine the simulation complexity. Furthermore, we establish a direct connection to separability problems. This section relies on our articles WOCJAN ET. AL [WJB02b] and JANZING, WOCJAN, AND BETH [JWB02a].

In the rest of the section we make use of the observation that by rescaling the system Hamiltonians H in eq. (6.5) may be assumed to be of the form

$$H = \sum_{k<l} \sigma_z^{(k)} \sigma_z^{(l)}\,. \tag{6.6}$$

The corresponding coupling matrix is given by

$$J = K \otimes C\,, \tag{6.7}$$

where K is the adjacency matrix of the complete graph and

$$C := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We call the Hamiltonian in eq. (6.6) *the complete Ising-Hamiltonian.*

6.2.1 Simulation of Ising-Hamiltonians

Before we treat the problem to simulate *general* Hamiltonians by the complete Ising-Hamiltonian, we consider only Ising-Hamiltonians

$$\tilde{H} = \sum_{k<l} w_{kl} \sigma_z^{(k)} \sigma_z^{(l)} , \quad w_{kl} \in \mathbb{R} . \tag{6.8}$$

The specific form of the complete Ising-Hamiltonian H allows to represent the toggling frame Hamiltonians, i.e., the elements of $Ad_{\mathcal{K}}(H) := \{U^\dagger H U \mid U \in \mathcal{K}\}$ in a convenient way. This representation will be used to establish a connection with separability problems (see Theorem 6.19). It is also used in Lemma 6.4 that simplifies control schemes for simulating Ising-Hamiltonians.

Lemma 6.3 (Representation of toggling frame Hamiltonians)
Let H be the complete Ising-Hamiltonian. Any toggling frame Hamiltonian $\hat{H} \in Ad_{\mathcal{K}}(H)$ can be represented by n unit vectors $|J_k\rangle = (J_{k;x}, J_{k;y}, J_{k;z})^T \in \mathbb{R}^3$ such that the blocks J_{kl} of its coupling matrix are given by $|J_k\rangle\langle J_l|$ for $k \neq l$.

Proof. Let $\hat{H} = U^\dagger H U$ where $U := U_1 \otimes \cdots \otimes U_n \in \mathcal{K}$. The coupling matrix J can be represented by n three dimensional real unit vectors: to each spin we associate the vector $|J_k\rangle = (J_{k;x}, J_{k;y}, J_{k;z})^T \in \mathbb{R}^3$, where

$$U_k^\dagger \sigma_z U_k = J_{k;x}\sigma_x + J_{k;y}\sigma_y + J_{k;z}\sigma_z .$$

Note that this is the usual Bloch sphere representation for density operators. It is straightforward to verify that the blocks of the coupling matrix are given by $J_{kl} = |J_k\rangle\langle J_l|$. □

Lemma 6.4 (Simplification of control schemes)
Assume that there is a control sequence that simulates the Ising-Hamiltonian $\tilde{H}$ by the complete Ising-Hamiltonian H with time overhead τ. Then this control sequence can be modified to a new one that has the same time overhead τ but uses control operation only from $\mathcal{X} := \{\mathbf{1}, \sigma_x\} \otimes \cdots \otimes \{\mathbf{1}, \sigma_x\}$.

Proof. Let

$$\hat{H} := \sum_{kl;\alpha\beta} \hat{J}_{kl;\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} \in Ad_{\mathcal{K}}(H)$$

be a toggling frame Hamiltonian that arises during the simulation. To prove the lemma it suffices to show that the Hamiltonian consisting of only the Ising terms of the toggling frame Hamiltonian $\hat{H}$, i.e.,

$$\sum_{kl} \hat{J}_{kl;zz} \sigma_z^{(k)} \sigma_z^{(l)} \tag{6.9}$$

can be obtained from H by a control scheme that uses only elements of $\mathcal{X}$ and has time-overhead 1.

Lemma 6.3 implies that the entries of the coupling matrix $\hat{J}$ of $\hat{H}$ can be factorized as $\hat{J}_{kl;\alpha\beta} = \hat{J}_{k;\alpha}\hat{J}_{l;\beta}$, where $|\hat{J}_k\rangle = (\hat{J}_{k;x}, \hat{J}_{k;y}, \hat{J}_{k;z})$ are the corresponding n three-dimensional vectors. For the proof we use the factorization $\hat{J}_{kl;zz} = \hat{J}_{k;z}\hat{J}_{l;z}$ for the Ising-terms of $\hat{H}$.

We express each $\hat{J}_{k;z} = c_k^+ - c_k^-$ with $0 \leq c_k^+, c_k^- \leq 1$ and $c_k^+ + c_k^- = 1$. In the new control sequence the complete Ising-Hamiltonian H is conjugated by

$$X = X_1 \otimes X_2 \otimes \ldots \otimes X_n \in \mathcal{X}\,.$$

The resulting Hamiltonian $X^\dagger H X$ acts for the time

$$\tau_X = \prod_{k=1}^{n} c_k(X)$$

where $c_k(X) = c_k^+$ if $X_k = \mathbf{1}$ and $c_k(X) = c_k^-$ if $X_k = \sigma_x$. Note that we have $(\sigma_x \otimes \mathbf{1})\sigma_z \otimes \sigma_z(\sigma_x \otimes \mathbf{1}) = -\sigma_z \otimes \sigma_z$. Therefore, we obtain

$$\sum_{X \in \mathcal{X}} \tau_X X^\dagger H X = \sum_{k<l} \hat{J}_{kl;zz}\sigma_z^{(k)}\sigma_z^{(l)} \tag{6.10}$$

since the strengths and signs of the Ising terms between the spins k and l are given by

$$c_k^+c_l^+ - c_k^-c_l^+ - c_k^+c_l^- + c_k^-c_l^- = (c_k^+ - c_k^-)(c_l^+ - c_l^-) = \hat{J}_{k;z}\hat{J}_{l;z} = \hat{J}_{kl;zz}\,.$$

This completes the proof that the elements of $Ad_{\mathcal{K}}(H)$ that arise during the simulation can be chosen to be elements of $Ad_{\mathcal{X}}(H)$, i.e., Hamiltonians containing only Ising terms, without increasing the time overhead. □

This case is of special interest since it deals with simulation procedures that do not rely on any first order approximation. All Hamiltonians that arise during the simulation procedure in eq. (6.10) commute. Therefore the unitary transformation implemented by the simulation scheme coincides exactly with the exponent of the average Hamiltonian.

To prove a statement on time optimal simulation of Ising-Hamiltonians we need to introduce some graph-theoretical notions. An (unweighted) graph is *bipartite* if its vertex set can be partitioned into two nonempty subsets V_1 and V_2 such that each edge of G has one end in V_1 and the other in V_2. The pair (V_1, V_2) is called a cut or bipartition of the graph. The *complete bipartite* graph with bipartition (V_1, V_2) is the bipartite graph with the additional property that every vertex of V_1 is connected with every vertex in V_2. It is denoted by $G(V_1, V_2)$.

A *Seidel matrix* defines a modified adjacency matrix $S := (s_{kl})$ for (unweighted) graphs in the following way (cf. CVETKOVIĆ ET AL. [CDS95]):

$$s_{kl} = \begin{cases} 0 & \text{if } k = l \\ -1 & \text{if } k \text{ and } l \text{ are adjacent and } k \neq l \\ 1 & \text{if } k \text{ and } l \text{ are non-adjacent and } k \neq l \end{cases}$$

Obviously, $S = K - 2A$, where K is the adjacency matrix of the complete graph and A the adjacency matrix of G.

Theorem 6.5 (Time optimal simulation)
Let G be a weighted graph with adjacency matrix $W = (w_{kl})$. Then the corresponding Ising-Hamiltonian

$$\tilde{H} = \sum_{kl} w_{kl} \sigma_z^{(k)} \sigma_z^{(l)}$$

can be simulated with overhead 1 if and only if the adjacency matrix W can be expressed as a convex combination

$$W = \sum_j p_j S_j \tag{6.11}$$

where the sum runs over a subset of the set of Seidel adjacency matrices of all complete bipartite graphs, i.e., over 2^{n-1} possible matrices.

Proof. The arguments in Lemma 6.4 show that it is sufficient to apply σ_x-rotations only and that each simulation step can be characterized by an n-dimensional vector $|x\rangle$ with entries from $\{+, -\}$. The entry kl (for $k \neq l$) of the matrix $|x\rangle\langle x|$ indicates the sign of the pair-interaction between the spins k and l in this step. The matrix $|x\rangle\langle x| - \mathbf{1}$ is a Seidel matrix of a complete bipartite graph. This can be seen as follows: By assigning to each vertex either "+" or "−" we have a bipartition of the vertex set: V_1 contains all vertices with "+" and V_2 all vertices with "−". The sign of the edge (k, l) is "−" if and only if the edge has one end in V_1 and the other end in V_2 and "+" otherwise. The edges with "−" define the complete bipartite graph $G(V_1, V_2)$. We also include the case $V_1 = \emptyset$ and $V_2 = V$ to cover the case when + is assigned to all nodes. Therefore, all we can achieve in a single step is $K - 2A(V_1, V_2)$, where K is the adjacency matrix of the complete graph and $A(V_1, V_2)$ is the adjacency matrix of $G(V_1, V_2)$. Note that $K - 2A(V_1, V_2)$ is the Seidel matrix of $G(V_1, V_2)$.

Conversely, to each Seidel matrix of a complete bipartite graph one can find a vector $|x\rangle$ such that the matrix $|x\rangle\langle x|$ coincides with the Seidel matrix (up to the diagonal). □

This proof shows that each $S_j + \mathbf{1}$ is a positive matrix since it is a projection up to a scalar factor. Hence, we obtain the following corollary.

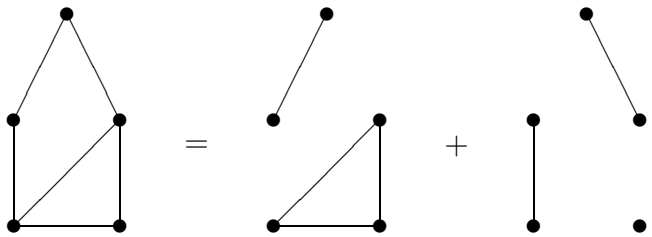


Figure 6.1: Clique coloring index

Corollary 6.6 (Lower bound on time-overhead)
The absolute value of the minimal eigenvalue of the weighted adjacency matrix W is a lower bound on the time overhead.

We present now some upper bounds on the time overhead. A graph $G' = (V', E')$ is called a subgraph of G if $V' \subseteq V$ and $E' \subseteq E$. A *clique* of G is a complete subgraph of G. A clique of G is called a *maximal clique* of G if it is not properly contained in another clique of G. A *clique partition* P of G is a partition of $E(G)$ such that its classes induce maximal cliques of G. Given a set C of h colors, an h-coloring of P in G is a mapping from P to C, such that cliques sharing a vertex have different colors. Let the *clique coloring index* $\kappa(G)$ be the smallest h such that there is a partition P permitting an h-coloring [WZ93]. We say the graph G consists of *independent cliques* if $\kappa(G) = 1$.

Figure 6.1 shows a graph with clique coloring index 2 and the corresponding partition.

Lemma 6.7 (Clique decoupling)
Let $G = (V, E)$ be an unweighted graph with clique coloring index 1. The corresponding Ising-Hamiltonian

$$\tilde{H} = \sum_{(k,l)\in E} \sigma_z^{(k)} \sigma_z^{(l)} ,$$

can be simulated with time overhead 1. This simulation is time optimal.

Proof. Let $\omega \geq 2$ be the number of maximal cliques. We choose a Hadamard matrix H of size $N := c\omega$ with $c < 2$. Let $s_k = (s_{k1}, \ldots, s_{kN})$ be the first ω rows of H. We partition the time interval into N time steps of length of equal length $\frac{1}{N}$. At the beginning and the end of the jth time interval we apply σ_x on all qubits of the k clique if $s_{kj} = -$, otherwise we do nothing.

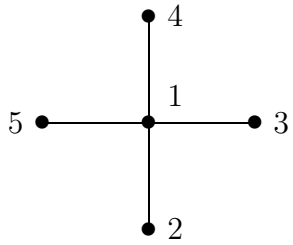
This scheme cancels all interactions among different cliques since the rows are orthogonal. It is optimal since $q \leq -1$ where q is the minimal eigenvalue of G (cf. Cvetković [CDS95, Theorem 0.13]). The fact that the scheme used in the proof is time optimal is also obvious for physical reasons since an overhead of less than

1 is only possible if the interactions to be simulated are all weaker than those of the system Hamiltonian. □

By partitioning a graph G into $\kappa(G)$ subgraphs consisting of independent cliques we obtain:

Corollary 6.8 (Upper bound on the time overhead)
Let G be an unweighted graph with clique coloring index κ. The corresponding Ising-Hamiltonian $\tilde{H}$ can be simulated with the time overhead κ.

Since the optimal simulation of an (unweighted) graph consisting of independent cliques has overhead 1 one might think that the clique index is the minimal overhead for general (unweighted) graphs. However, the following example shows that this is not true. Consider the *star* $G = (V, E)$ with $V = \{1, \ldots, 5\}$ and $E = \{(1,2), (1,3), (1,4), (1,5)\}$:



Due to the special form of the star it is easily seen that the clique coloring index of G is 4. However, the optimal time overhead is 2 only. The vectors can be chosen as $s_1 = (++++)$, $s_2 = (-+++)$, $s_3 = (+-++)$, $s_4 = (++-+)$, $s_5 = (+++-)$ and each of the four intervals has length $1/2$ since $\frac{1}{2}\langle s_1|s_i\rangle = 1$ for $i = 2, \ldots, 5$ and $\langle s_i|s_j\rangle = 0$ for $2 \leq i < j \leq 5$. This is optimal since the smallest eigenvalue of the adjacency matrix of G is -2.

To obtain a upper bound on the time overhead for weighted graphs we generalize the notion of clique coloring index:

Definition 6.9 (Weighted clique coloring index)
Let G be a weighted graph with non-negative weights w_{kl}. For every non-negative real number s we define the (unweighted) graph G_s as follows: the vertices k and l are adjacent in G_s if $w_{kl} > s$. The weighted clique coloring index of G is

$$\kappa := \int_0^\infty \kappa_s \, ds \,, \tag{6.12}$$

where κ_s denotes the clique coloring index of G_s.

Figure 6.2 shows a graph with weighted clique coloring index and the corresponding partition.

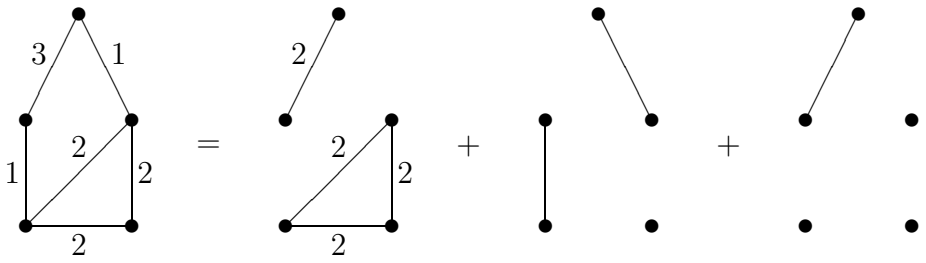


Figure 6.2: Weighted clique coloring index

Why do we consider graphs with non-negative weights when working with the clique coloring index? Consider the problem of time-reversal of the complete Ising-Hamiltonian on n qubits. Then $n-1$ is a lower bound on the time overhead due to Corollary 5.13. This shows that the minus signs decrease here the possible degree of parallelization. Therefore, we will introduce in the next section the *chromatic index* to study the time overhead of graph with arbitrary weights.

Lemma 6.10 (Upper bound on time overhead)
Let G be a weighted graph with non-negative weights w_{kl} and weighted clique coloring index κ. The corresponding Ising-Hamiltonian

$$\tilde{H} = \sum_{k<l} w_{kl}\, \sigma_z^{(k)} \sigma_z^{(l)}$$

can be simulated with time overhead κ.

Proof. Denote by $s_1, \ldots, s_m$ the different positive numbers $\{w_{kl}\}$ sorted in increasing order. Set $s_0 := 0$. Obviously, the function $s \mapsto \kappa(G_s)$ takes constant values on every interval $[s_{j-1}, s_j)$ for $1 \leq j \leq m$.
For $j = 1, \ldots, m$ we simulate the simulate the Hamiltonian with the graph $G_{s_{j-1}}$ for the relative time $s_j - s_{j-1}$. This scheme simulates the desired Hamiltonian. The time overhead is given by the weighted clique coloring index. □

Now we use spectral methods to derive lower bounds on the number of time steps. An upper bound is given with the help of Carathéodory's theorem.

Theorem 6.11 (Bounds on number of time steps)
Let A be the weighted adjacency matrix of the Hamiltonian we want to simulate. Then one requires at least $n - k$ time steps, where k is the multiplicity of the minimal eigenvalue $\mu_{\min}$ of A. Furthermore, if $\mu_{\min}$ is irrational then at least n steps are necessary. In any case, $n(n-1)/2 + 1$ time steps are always sufficient.

Proof. We characterize the each step a diagonal matrix X_j of size $n \times n$. The diagonal entries are ± 1 and indicate which spins are subjected to conjugation in

this time step. The simulation condition is then

$$\sum_{j=1}^{N} \tau_j X_j K X_j = A\,.$$

We add the identity matrix on both sides. Due to $X_j \mathbf{1} X_j = \mathbf{1}$ we obtain

$$\sum_{j=1}^{N} \tau_j X_j I X_j = A + \sum_{j} \tau_j \mathbf{1}\,.$$

The rank of the l.h.s. is at most the number N of time steps. To estimate the rank of the r.h.s. note that the time overhead $\tau := \sum_j \tau_j$ is at least $-\mu_{\min}$. Hence the dimension of the kernel of $A + \tau\mathbf{1}$ can be at most the multiplicity of the eigenvalue $\mu_{\min}$. This proves that the number of time steps is at least $n - k$.

Assume $\mu_{\min}$ to be irrational. Then the time overhead τ is necessarily greater than $-\mu_{\min}$. This can be seen as follows. The optimization with respect to the time overhead reduces to the following convex problem. Consider the matrix $X_j K X_j$ for an arbitrary time step j. Its non-diagonal entries are ± 1 and indicate which interactions acquire a minus sign in the jth step. In graph-theoretical language, the set of matrices that can occur as $X_j K X_j$ are exactly the Seidel matrices of complete bipartite graphs. Then the optimal τ is the minimal positive number such that A/τ is in the convex span of the set of Seidel matrices of complete bipartite graphs. Geometrically, the convex span is a polytope having the Seidel matrices as its extreme points. It is embedded in the $n(n-1)/2$ dimensional vector space of real symmetric matrices with zeros on the diagonal. Let O be the origin. Consider the semi-line νA for $\nu \geq 0$. Then the optimal simulation is the unique intersection point P of the semi-line with the boundary of the polytope. The quotient of the distance between O and A and between O and P is the optimal time overhead. This quotient can never be irrational. The reason is that P has rational entries since it is the solution of a linear equation over the field $\mathbb{Q}$ of rational numbers. Hence τ is greater than $-\mu_{\min}$ and $A + \tau\mathbf{1}$ has necessarily full rank. This proves that we need at least n time steps.

To see that $n(n-1)/2 + 1$ time steps are always sufficient we can argue as in the proof of Theorem 5.21 with Carathéodory's theorem. The dimension of the convex span of the matrices $X_j K X_j$ is at most $n(n-1)/2$. □

It is surprising that it is relevant for our lower bound on the number of time steps whether the smallest eigenvalue of the (weighted) adjacency matrix is irrational. It is not clear if this is only a feature of our proof or if there is a true connection to the irrationality of graph spectra. The question of whether graphs have a rational spectrum is studied in algebraic graph theory (cf. Cvetković [CDS95]).

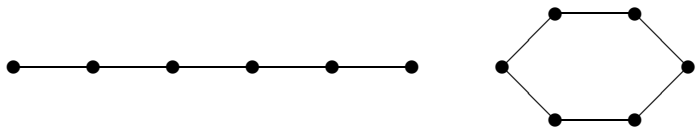


Figure 6.3: Open chain (path) and cyclic chain (circuit) on 6 vertices

6.2.2 Simulation of graph Hamiltonians

In this section we consider the simulation of special interaction graphs that appeared in the literature in various applications. Some interesting models in quantum information theory refer to quite idealized types of Hamiltonians like nearest neighbor interactions. If the natural Hamiltonian contains long range interactions between all nodes one may try to simulate the idealized interaction. Then the problem is to cancel the unwanted terms without destroying the desired interactions. The examples below show that it may cause a large number of time steps to cancel unwanted long-range interactions no matter how fast they are decreasing with the distance. As long as they are not neglected, the control sequences that cancel them may be rather long. Here we restrict our attention to Ising-interactions between n qubits.

Since we consider the task to cancel some interactions and keep others the rescale matrix $A := \tilde{W}/W$ has only 1 and 0 as entries. In graph-theoretical language, it is the adjacency matrix of the desired interaction graph. Therefore, we can use graph spectra for deriving lower bounds based on the above theorems. We show that some of them are almost tight by sketching simulation schemes based on well-known results on selective decoupling.

Open and cyclic spin chains

We consider the problem to simulate *open* and *cyclic chains* on n spins with only nearest neighbor interactions, i.e., we have to cancel the interactions between all non-adjacent pairs. The corresponding interaction graphs are shown in Figure 6.3 for 6 spins. In graph theory, these graphs are called paths and circuits. More precisely, a *path* P_n is a graph on n vertices with $E := \{(1,2),(2,3),\ldots,(n-1,n)\}$. A *circuit* C_n is a graph on vertices n with $E := \{(1,2),(2,3),\ldots,(n-1,n),(n,1)\}$.

To determine the characteristic polynomial $P_{C_n}(\lambda)$ of a circuit C_n we have to introduce special polynomials. The *Chebyshev polynomials of the first kind* form a set of orthogonal polynomials defined as the solutions to the Chebyshev differential equation (cf. Abramowitz and Stegun [AS72]). They can be expressed as

$$T_n(x) = \cos\big(n \arccos(x)\big) \tag{6.13}$$

or as the product

$$T_n(x) = 2^{n-1} \prod_{k=1}^{n} \left\{ x - \cos\left[\frac{(2k-1)\pi}{2n} \right] \right\} \tag{6.14}$$

(cf. Zwillinger [Zwi95]). Zeros occur when

$$x = \cos\left[\frac{\pi(k-\frac{1}{2})}{n} \right],$$

where $k = 1, 2, \ldots, n$. Extrema occur for

$$x = \cos\left(\frac{\pi k}{n} \right), \tag{6.15}$$

where $k = 1, 2, \ldots, 2n$. At maximum, $T_n(x) = 1$ for $k = 2, 4, \ldots, 2n$, and at minimum, $T_n(x) = -1$ for $k = 1, 3, \ldots, 2n-1$.

Lemma 6.12 (Spectrum of the circuit C_n)
The spectrum of a circuit C_n consists of all the numbers of the form

$$2\cos\left(\frac{2\pi}{n}\, i \right), \quad i = 1, \ldots, n\,. \tag{6.16}$$

The characteristic polynomial is

$$P_{C_n}(\lambda) = 2\, T_n\Big(\frac{\lambda}{2} \Big) - 2\,. \tag{6.17}$$

Proof. The adjacency matrix of the circuit C_n is $S + S^T$, where S denotes the cyclic shift in n dimensions. The eigenvalues of S are the nth roots of unity and the corresponding eigenvalues of S^T are their complex conjugates. Therefore, the eigenvalues of $S + S^T$ are given by twice the real parts of the nth roots of unity. Due to the properties of Chebyshev polynomials of the first kind it is easily verified that the characteristic polynomial is given by the polynomial in eq. (6.17). □

Now we can study the complexity of simulating circuits. For odd n the lower bound on the number of time steps is n since the minimal eigenvalue is irrational for all $n > 3$. For n even the minimal eigenvalue occurs for occurs for $i = n/2$ in eq. (6.16) and is -2; it multiplicity is 1. Hence $n - 1$ is a lower bound on the number of time steps and 2 is a lower bound on time-overhead. There are numbers n, where this lower bound on the number of time steps is almost tight. This is shown by the following example.

Example 6.13 (Simulating circuits C_n with n even)
Let n be an even number with the property that a Hadamard matrix of dimension $n/2$ exists. This is for instance the case for each power of 2 (cf. Hedayat

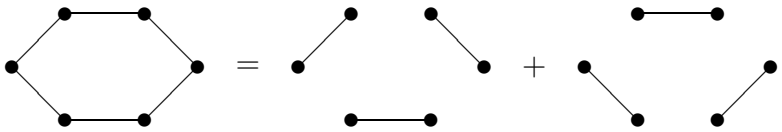


Figure 6.4: Decomposition of the circuit C_6 into independent cliques

et al. [HSS99]). We construct a simulation scheme that consists of two subroutines. The first subroutine simulates the interactions between the pairs $\{1,2\}$, $\{3,4\},\ldots,\{n-1,n\}$. The second subroutine the interactions between the pairs $\{n,1,\},\{2,3\},\ldots,\{n-2,n-1\}$. This is shown in Figure 6.4. Note that all the pairs in the same subroutine are disjoint. The problem to simulate the interactions between disjoint pairs is a special instance of clique decoupling (see Lemma 6.7), where the interaction between independent cliques are cancelled and the interactions within the same clique remains. It can be achieved using Hadamard matrices having at least the number of cliques as dimension. Using this method, we need $n/2$ steps in each subroutine. Therefore we have given a simulation with n steps, whereas our lower bound is $n-1$. In general, the number of steps for simulating the circle grows only linearly in n. This shows that the lower bound is quite good even for general n. This selective decoupling scheme is time-optimal since it saturates the lower bound on time overhead.

To determine the spectrum and the characteristic polynomial of a path P_n we need the following lemma and the definition of Chebyshev polynomials of second kind.

Lemma 6.14 (Differentiation of the characteristic polynomial)
Let G be a graph with vertex set $V=\{x_1,x_2,\ldots,x_n\}$. Let G_i be the subgraph induced by the vertices $V\setminus\{x_i\}$. If all the subgraphs G_i are isomorphic with some graph H, then the characteristic polynomial of H can be computed as

$$p_H(\lambda)=\frac{1}{n}\frac{d}{d\lambda}p_G(\lambda)\,. \tag{6.18}$$

Proof. Let $A\in\mathbb{C}^{n\times n}$ be an arbitrary matrix, and let A_i the principal sub-matrix resulting from deleting the ith row and the ith column, $i=1,\ldots,n$. Then the derivative of the characteristic polynomial of A can be expressed as sum of the characteristic polynomials of A_i (cf. Horn and Johnson [HJ85])

$$\frac{d}{d\lambda}p_A(\lambda)=\sum_{i=1}^{n}p_{A_i}(\lambda)\,.$$

The proof follows by applying this identity to the adjacency matrices and the fact that all subgraphs G_i have the same spectrum as H. □

If we delete any vertex of a circuit with $n+1$ we obtain a path on n vertices. This observation put together with the lemma above allows us to compute the characteristic polynomials of paths.

Chebyshev polynomials of second kind, denoted by $U_n(x)$, can be written as

$$U_n(x) = \frac{\sin\left[(n+1)\arccos(x)\right]}{\sqrt{1-x^2}} \tag{6.19}$$

and as the product

$$U_n(x) = 2^n \prod_{k=1}^{n} \Big[x - \cos\Big(\frac{\pi k}{n+1}\Big)\Big] \tag{6.20}$$

(cf. Zwillinger [Zwi95]). Note that the Chebyshev polynomials of the first and second kind are related via differentiation:

$$\frac{1}{n+1} T'_{n+1}(x) = U_n(x)\,. \tag{6.21}$$

Now we have everything to derive the spectrum and the characteristic polynomial of the path P_n.

Lemma 6.15 (Spectrum of the path P_n)
The spectrum of the path P_n consists of the numbers

$$2\cos\big(\frac{\pi}{n+1}\, i\big) \quad i = 1, \ldots, n\,. \tag{6.22}$$

The characteristic polynomial is

$$P_{P_n}(\lambda) = U_n\Big(\frac{\lambda}{2}\Big)\,, \tag{6.23}$$

where $U_n(x)$ denotes the Chebyshev polynomial of the second kind.

Proof. By applying Corollary 6.14 we can deduce from the previous result on circuits the spectrum and the characteristic polynomial of the *path* P_n with n vertices. All subgraphs of the circuit C_{n+1} induced by n vertices are isomorphic with the path P_n. Therefore, we obtain

$$P_{P_n}(\lambda) = \frac{1}{n+1} P'_{C_{n+1}}(\lambda) = U_n\Big(\frac{\lambda}{2}\Big)\,.$$

It then easily follows from the product form in eq. (6.20) that the spectrum of the path P_n consists of the numbers given in eq. (6.22). □

The minimal eigenvalue of the path P_n is irrational for all $n > 2$. By Theorem 6.11 we conclude that the number of time steps is at least n.

Rectangular lattice

We consider a quantum system of $n = l^2$ spins located on a two-dimensional rectangular lattice. For simplicity assume l to be even. We want to simulate a lattice with only nearest neighbor interactions.

To compute the spectrum of the rectangular lattice we need to introduce the notion of a sum of graphs. The rectangular lattice can be expressed as the sum of two paths. The advantage is that this observation allows to compute easily the spectrum of the rectangular lattice in terms of the spectra of the paths.

Definition 6.16 (Sum of graphs)
Let X and Y be two graphs. The vertices (x, y) and (x', y') are adjacent in the sum $X + Y$ if and only if

1. *either $x = x'$ and y, y' are adjacent in Y*
2. *or x, x' are adjacent in X and $y = y'$.*

Theorem 6.17 (Spectrum of the sum of two graphs)
Let X, Y be arbitrary graphs with eigenvalues $\lambda_1, \ldots, \lambda_m$ and $\mu_1, \ldots, \mu_n$, respectively. Then the spectrum of the sum $X + Y$ consists of numbers of the form

$$\lambda_i + \mu_j \tag{6.24}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, n$.

Proof. The adjacency matrix of the sum $X + Y$ is given by

$$A(X + Y) = \mathbf{1}_n \otimes A(Y) + A(X) \otimes \mathbf{1}_m \,.$$

This follows directly from Definition 6.16, since the first case corresponds to $\mathbf{1}_n \otimes A(Y)$ and the second case corresponds to $A(X) \otimes \mathbf{1}_m$.

Let $|\lambda_1\rangle, \ldots, |\lambda_m\rangle$ and $|\mu_1\rangle, \ldots, |\mu_n\rangle$ be the eigenvectors of the adjacency matrices $A(X)$ and $A(Y)$, respectively. The matrices $\mathbf{1}_n \otimes A(Y)$ and $A(X) \otimes \mathbf{1}_m$ commute. They can be simultaneously diagonalized in the basis $|\lambda_i\rangle \otimes |\mu_j\rangle$. Then the corresponding eigenvalues are given by eq. (6.24). □

Corollary 6.18 (Spectrum of the square lattice)
The sum of two paths having m and n vertices, respectively is the graph of an $m \times n$ rectangular lattice. The spectrum of this graph consists of all the numbers of the form

$$2\cos\left(\frac{\pi}{m+1}\, i\right) + 2\cos\left(\frac{\pi}{n+1}\, j\right), \quad i = 1, \ldots, m\,;\, j = 1, \ldots, n\,. \tag{6.25}$$

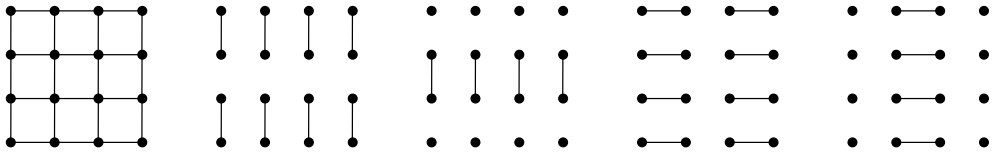


Figure 6.5: Simulation of the rectangular lattice interaction with 4 subroutines

The desired interaction graph is shown on the left of Fig. 6.5. This kind of interaction can for instance be used for preparing the initial state in the 'One-Way Quantum Computer' proposed by Raussendorf and Briegel [RB00].

We first consider the time overhead. An upper bound is given by 4 since this is the clique coloring index. It is easy to see that the minimal eigenvalue of A is given by

$$\lambda_{\min} = 2\cos\left(\frac{\pi}{l+1}\,l\right) + 2\cos\left(\frac{\pi}{l+1}\,l\right). \tag{6.26}$$

By Theorem 6.11 the lower bound on the number of time steps is n since the smallest eigenvalue is irrational. Note that this example shows that the complexity measures *time overhead* and *number of time steps* may differ significantly.

An upper bound on the number of time steps can be obtained as follows. The graph has $2(l-1)l$ edges. We can partition the edges into 4 sets of edges such that each set contains only disjoint interacting pairs. These 4 partitions are shown in Figure 6.5. The simulation consists of 4 subroutines simulating one the interactions in one of the 4 classes. For each subroutine we choose Hadamard matrices with a dimension that is at least the number of cliques. The numbers of cliques are $l^2/2$ or $l^2/2 + l$ in each subroutine. Since there exist Hadamard matrices for every power of 2 the square lattice graph can always be simulated in $O(l^2) = O(n)$ time steps.

Graph codes

The computational power of different n qubit interactions is not well understood yet. It would be interesting to know which n qubit transformations can easily be implemented when a certain Hamiltonian is given. However, one of the few examples where the power of specific Hamiltonians is *directly* used (without using them to implement two qubit gates) is the preparation of states of graph codes (cf. Schlingemann and Werner [SW00], Schlingemann [Sch01, Sch02] and Grassl et al. [GKR02]). Here the codes states are obtained by the free time evolution according to a Ising-Hamiltonian. The graph representing a code is the interaction graph that can be used for preparing the states. We assume the natural interaction to be equal Ising interactions between all 6 spins and want to simulate the interaction graph in Fig 6.6. This interaction graph is required for the preparing the states of a graph code of length 5.

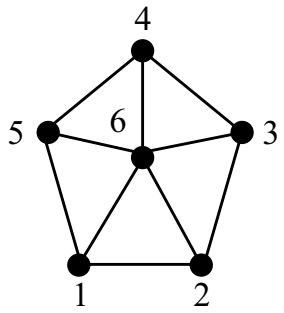


Figure 6.6: Wheel on 6 vertices

The eigenvalues of the wheel in Figure 6.6 can easily be computed by any computer algebra system. They are given by

$$1+\sqrt{6}\,,\ \ \frac{1}{2}\sqrt{5}-\frac{1}{2}\,,\ \ \frac{1}{2}\sqrt{5}-\frac{1}{2}\,,\ \ 1-\sqrt{6}\,,\ \ -\frac{1}{2}\sqrt{5}-\frac{1}{2}\,,\ \ -\frac{1}{2}\sqrt{5}-\frac{1}{2}\,.$$

The minimal eigenvalue $-1/2-\sqrt{5}/2$ is irrational. Therefore, by Theorem 6.11 the minimal number of time steps is 6. An implementation with 12 time steps is given as follows. The scheme consists of 3 subroutines each consisting of 4 time steps. As above each subroutine simulates the interaction between disjoint cliques and cancels the interaction between different cliques. Subroutine 1 has the cliques $\{1,2,6\},\{4,5\},\{3\}$. The clique partitions in subroutine 2 and 3 are $\{3,4,6\},\{1,5\},\{2\}$ and $\{1\},\{4\},\{2,3\},\{5,6\}$, respectively. In each subroutine decoupling the different cliques can be achieved by Hadamard matrices of dimension 4 since no subroutine has more than 4 cliques. Hence we have 4 time steps in each subroutine.

6.2.3 Simulation of general Hamiltonians

Remarkably, the problem of specifying the set of Hamiltonians that can be simulated with time overhead 1 is related to the problem of characterizing separable states of n-qubit systems. More specifically, the convex problem can be reduced to the question "what kind of pair-correlations can occur in a separable n-qubit quantum state?"

Theorem 6.19 (Time optimal simulation)
An arbitrary Hamiltonian $\tilde{H}$ can be simulated with overhead τ if and only if there is a separable quantum state ρ in $(\mathbb{C}^2)^{\otimes n}$ such that

$$\frac{1}{\tau}J+\mathbf{1}=\mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(l)}))_{kl;\alpha\beta}\,,$$

where J denotes the coupling matrix of $\tilde{H}$ and $\mathbf{1}$ the $3n\times 3n$ identity matrix.

Proof. Assume that $\tilde{H}$ can be simulated with overhead τ. Then the Hamiltonian $\tilde{H}/\tau$ can be simulated with overhead 1 and consequently can be written as a convex combination $\tilde{H}/\tau := \sum_j p_j H_j$ of elements $H_j \in Ad_{\mathcal{K}}(H)$.

Now we consider the coupling matrix J of one H_j. We show how to construct a separable state ρ such that $J + \mathbf{1} = (\mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(l)}))_{kl;\alpha\beta}$. Let $H_j = U^\dagger H U$. Use the representation of H_j by n three dimensional real vectors $|J_k\rangle$ as explained in Remark 6.3. The Bloch sphere gives the correspondence between the unit vectors $|J_k\rangle$ and the projections ρ_k in $\mathbb{C}^2$ defined by $J_{k;\alpha} = \mathrm{tr}(\rho_k\sigma_\alpha)$. Let ρ be the product state $\rho := \rho_1 \otimes \ldots \otimes \rho_n$. Then we have $J_{kl;\alpha\beta} = \mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(l)})$ for all $k \neq l$. This is almost the desired state corresponding to H_j. The only problem that remains is that we may have $\mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(k)}) \neq 0$ for $\alpha \neq \beta$ since the product of two different Pauli matrices (acting on the same spin) is the third Pauli matrix multiplied by a scalar. Therefore we substitute ρ by a state $\bar{\rho}$ in such a way that the expectation values of all traceless 1-qubit observables vanish and the expectation values of all considered 2-qubit observables remain unchanged. This is done as follows. For every $|J_k\rangle$ we can find $O'_k \in SO(3)$ such that $O'_k|J_k\rangle = -|J_k\rangle$. This rotation corresponds to conjugation of the qubit k by a unitary u'_k. The vector $-|J_k\rangle$ corresponds to the projection $\rho'_k := I_2 - \rho_k$. Let $\bar{\rho}$ be the state

$$\bar{\rho} := \frac{1}{2}(\rho_1 \otimes \ldots \otimes \rho_n + \rho'_1 \otimes \ldots \otimes \rho'_n)\,.$$

hen we have $\mathrm{tr}(\bar{\rho}\sigma_\alpha^{(k)}\sigma_\beta^{(l)}) = \mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(l)})$ for all $k \neq l$ (all vectors $|J_k\rangle$ are multiplied by -1 and therefore there is no effect on the pairs) and $\mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(k)})$ is the 3×3 identity matrix.

Once we have constructed the separable states ρ_j corresponding to each H_j (as explained above) then $\rho := \sum_j p_j\rho_j$ is the desired separable state corresponding to H/μ.

Assume conversely we have a separable state ρ of the desired form. Let

$$\rho = \sum_j p_j\rho_j$$

be a decomposition into pure product states. Each ρ_j can be characterized by n real unit vectors via the Bloch sphere representation. We consider one ρ_j. Let $|J_k\rangle = (J_{k;x}, J_{k;y}, J_{k;z})$ be the real unit vector of the kth qubit. Choose unitaries U_k such that $U_k\sigma_z U_k^\dagger = J_{k;x}\sigma_x + J_{k;y}\sigma_y + J_{k;z}\sigma_z$. Set $U := U_1 \otimes \cdots \otimes U_n$ and $H_j := U^\dagger H U$ (note that we omitted the index j for U and U_k). It remains to show that $H/\tau = \sum_j p_j H_j$. The coupling matrix of H/τ is given by J/τ. For $k \neq l$ the block kl of the latter matrix is given (by assumption) by the 3×3-matrix

$$\mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(l)})_{\alpha\beta} = \sum_j p_j(\mathrm{tr}(\rho_j\sigma_\alpha^{(k)}\sigma_\beta^{(l)})_{\alpha\beta})\,.$$

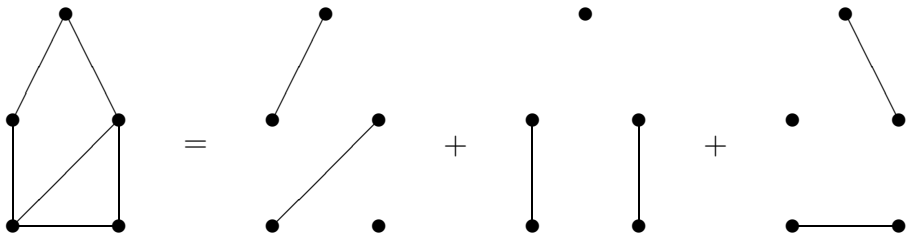


Figure 6.7: Chromatic index

Since the 3×3-matrices in the convex sum on the right-hand side coincide with the blocks kl of the coupling matrices of H_j we have shown that $H/\tau = \sum_j p_j H_j$ □

A simple lower bound on the simulation time overhead can be derived from Theorem 6.19. Note that the matrix

$$(\mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(l)}))_{kl;\alpha\beta}$$

is positive since for every vector $|d\rangle = (d_{k;\alpha})$ of length $3n$ we have

$$\sum_{kl;\alpha\beta} d_{k;\alpha}\mathrm{tr}(\rho\sigma_\alpha^{(k)}\sigma_\beta^{(l)})d_{l;\beta} = \mathrm{tr}(\rho AA^\dagger) \geq 0\,,$$

where $A = \sum_{k;\alpha} d_{k;\alpha}\sigma_\alpha^{(k)}$. Hence $J+\tau\mathbf{1}$ is a positive matrix. Due to the fact that J is traceless its minimal eigenvalue is negative. With these observations we obtain the following corollary.

Corollary 6.20 (Lower bound on time overhead)
The absolute value of the minimal eigenvalue of the coupling matrix of $\tilde{H}$ is a lower bound on the simulation overhead of $\tilde{H}$ by the complete Ising-Hamiltonian.

The Hamiltonians we want to simulate contain interactions of different strengths. It will be convenient to encode this information by a *family* of graphs. Therefore we define:

Definition 6.21 *Let H be an arbitrary pair-interaction Hamiltonian. For every non-negative real number r we define the* interaction graph *G_r as follows: Let the qubits $\{1,\ldots,n\}$ label the vertices and let the edges be all the pairs (k,l) with the property $\|H_{kl}\| > r$.*

A decisive property of the interaction graphs G_r is their chromatic index.

Definition 6.22 (Chromatic index)
Recall that the chromatic index χ' of a graph $G = (V,E)$ is the minimum number

of colors permitting an edge-coloring such that no two adjacent edges receive the same color or equivalently a partition $E = M_1 \cup M_2 \cup \ldots \cup M_{\chi'}$ into independent subsets of E.

To illustrate this definition let us consider a scheduling problem. Each of n businessmen wishes to hold confidential meetings with some others. Assuming that each meeting lasts a day and at each meeting exactly two businessmen are present, in how many days can the meeting be over? In this case we consider the graph G whose vertices correspond to the n businessmen and where two vertices are adjacent if and only if the two businessmen wish to hold a meeting. Then the problem above asks for the minimal number of colors in an edge coloring of G in such a way that no two adjacent edges have the same color.

An example of a graph with chromatic number 3 is shown in Figure 6.7. Its clique coloring index is only 2 (see Figure 6.1). We extend the notion of chromatic index for weighted graphs as we have done for the clique coloring index.

Definition 6.23 (Weighted chromatic index)
We define the weighted chromatic index *of a Hamiltonian H as*

$$\chi' := \int_{r=0}^{\infty} \chi'_r dr \tag{6.27}$$

where χ'_r denotes the chromatic index of the interaction graph G_r of H.

The idea to consider the chromatic index as a complexity measure for the interaction is intuitive: in general, it should be easy to control interactions on disjoint qubit pairs, whereas one should expect that its unlikely that one can *control* simultaneously the interaction between qubit 1 and 2 and the interaction 1 and 3 at the same moment.

Corollary 6.24 (Chromatic index one)
Let M be a set of qubit pairs, such that no two pairs contain a common qubit. Then we can simulate the Hamiltonian

$$\tilde{H} = \sum_{(k,l)\in M} \sigma_z^{(k)} \sigma_z^{(l)} \tag{6.28}$$

with time overhead 1 by using the complete Ising-Hamiltonian.

Proof. This is a special case of clique decoupling (see Lemma 6.7) since a graph with chromatic index 1 consists of independent cliques. □

Lemma 6.25 (Simulation of two qubit Hamiltonians)
Let $H = \sigma_z \otimes \sigma_z$ be the drift Hamiltonian of a 2-spin system. All Hamiltonians $\tilde{H} \in su(4)$ can be simulated with time overhead less than $\|\tilde{H}\|$ using the drift Hamiltonian H.

Proof. We first assume that $\tilde{H}$ contains no local terms, i.e., we have $\tilde{H} = \sum_{\alpha\beta} J_{12;\alpha\beta}\, \sigma_\alpha \otimes \sigma_\beta$. In the following we omit the indices $1, 2$ and describe the coupling by the 3×3 matrix J in abuse of notation.

Recall that conjugation of $\tilde{H}$ by $k = u \otimes v \in SU(2) \otimes SU(2)$ corresponds to multiplication of J by U from the left and by V^T from the right, where $U, V \in SO(3)$. By the singular value decomposition (cf. Horn and Johnson [HJ85]) there are $U, V \in SO(3)$ such that $UJV^T = \mathrm{diag}(s_x, s_y, s_z)$ where s_x, s_y, s_z are the singular values of J. Equivalently, there is $k \in SU(2) \otimes SU(2)$ such that $kHk^\dagger = H_{s_x,s_y,s_z}$ where $H_{s_x,s_y,s_z} = s_x\sigma_x \otimes \sigma_x + s_y\sigma_y \otimes \sigma_y + s_z\sigma_z \otimes \sigma_z$. By computing the eigenvalues we see that $\|H_{s_x,s_y,s_z}\| = \sum_\alpha |s_\alpha|$. The simulation time overhead cannot be more than the right hand side since each term $s_\alpha\sigma_\alpha \otimes \sigma_\alpha$ can be simulated with overhead $|s_\alpha|$ by $\sigma_z \otimes \sigma_z$.

Let $\tilde{H}$ contain local terms. By applying the singular value decomposition to the non-local part we may assume that it has the following form $\tilde{H} = \sum_\alpha s_\alpha\sigma_\alpha \otimes \sigma_\alpha + 1 \otimes a + b \otimes 1$ with $a, b \in su(2)$. We can split $\tilde{H} = H' + H''$ where H' is the non-local part and H'' the local one. By the Trotter formula

$$\exp(A + B) = \lim_{m\to\infty} (\exp(A/m) \exp(B/m))^m$$

we can simulate the parts independently with arbitrary accuracy. The simulation of H'' takes no time by assumption. It remains to show that $\|H'\| \leq \|\tilde{H}\|$. We may assume that $\tilde{H}$ is invariant with respect to qubit permutation since $\|\frac{1}{2}\tilde{H} + \frac{1}{2}\tilde{H}_{ex}\| \leq \|\tilde{H}\|$ where $\tilde{H}_{ex}$ is the Hamiltonian obtained from $\tilde{H}$ by exchanging the qubits. Let $\tilde{H} = H_{s_x,s_y,s_z} + 1 \otimes c + c \otimes 1$ with $c \in su(2)$. The eigenvectors of H_{s_x,s_y,s_z} are the Bell states. We have $\langle\Psi|\tilde{H}|\Psi\rangle = \langle\Psi|H_{s_x,s_y,s_z}|\Psi\rangle$ since $\langle\Psi|1 \otimes c + c \otimes 1|\Psi\rangle = 0$ for all Bell states. Therefore the norm of $\tilde{H}$ cannot be smaller than the norm of H_{s_x,s_y,s_z} □

By putting together Corollary 6.24 and Lemma 6.25 we obtain the following theorem.

Theorem 6.26 (Upper bound on time overhead)
The time overhead for simulating an arbitrary pair-interactions Hamiltonian $\tilde{H}$ by the complete Ising-Hamiltonian is at most the weighted chromatic index of $\tilde{H}$.

6.3 Connecting the control theoretic model and the circuit model

If one assumes that the implementation time of a two qubit gate is proportional to its distance from the identity then we will show that the complete Ising-Hamiltonian can implement quantum circuits with no time overhead. We present an example showing that the complete Ising-Hamiltonian is even more powerful.

Clearly the assumption that the interactions between all spins have the same magnitude is unphysical since in real physical systems the interaction strength always decreases with the distance between the interacting particles. However, many aspects of our theory can be developed in strong analogy for more general drift Hamiltonians. Furthermore, the results of our idealized model indicate that long-range interactions imply strong computational power.

Now we show that our computational model is at least as powerful as the usual model with two qubit gates, even if we care about constant overhead. We describe here briefly the quantum circuit model and introduce the *weighted depth* following Janzing and Beth [JB01]. It is a complexity measure for unitary transformations extending the notion of depth in Definition 1.3.

The following quantity measures the deviation of a unitary operator from the identity:

Definition 6.27 (Angle of unitary operators)
The angle *of an arbitrary unitary operator $U \in SU(4)$ is the smallest possible norm*[1] *$\|H\|$ of a Hermitian operator $H \in su(4)$ (i.e. an element of the Lie algebra of traceless Hermitian 4×4-matrices) that satisfies* $\exp(-iH) = U$.

It coincides with the time required for the implementation of U if the norm of the used Hamiltonian is 1. We consider only the angle of two-qubit gates, i.e., we do not include the angle of local gates in the definition of the weighted depth. The notion of angle allows us to formulate a modification of the term 'depth' that will later turn out to be decisive in connecting complexity measures of discrete and continuous algorithms:

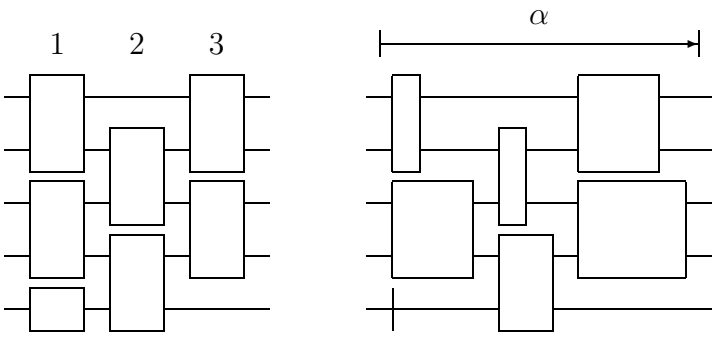
Definition 6.28 (Weighted depth)
Let α_i be the maximum of the angles of the two-qubit unitaries performed in step i of A (see Definition 1.3). Then the weighted depth *of A is defined to be the sum $\alpha = \sum_i \alpha_i$.*

Figure 6.8 illustrates weighted depth compared to depth. The complexity measure weighted depth takes into account that the two qubit gates have to be implemented by the couplings between the qubits.

Corollary 6.29 (Lower bound on implementation time)
A quantum circuit of weighted depth α can be implemented by using local operations in $\mathcal{K}$ and the complete Ising-Hamiltonian with time α.

Proof. Let $M = \{(k, l)\}$ be the set of the pairs that the two-qubit gates act on. No two pairs in M contain a common vertex and therefore we can simulate the

[1] Here $\|.\|$ denotes the operator norm given by $\|H\| := \max_{|\Psi\rangle} \|H|\Psi\rangle\|$ where $|\Psi\rangle$ runs over the unit vectors of the corresponding Hilbert space.



a. Depth b. Weighted depth

Figure 6.8: Discrete and continuous complexity measures of quantum circuits

Hamiltonian

$$\tilde{H} = \sum_{(k,l)\in M} \sigma_z^{(k)} \sigma_z^{(l)}$$

with time overhead 1 (see Lemma 6.24). Therefore the disjoint pairs can be considered independently and the gates U_{kl} can be implemented in parallel. To implement the gate U_{kl} we choose a Hamiltonian H_{kl} of minimal norm such that $U_{kl} = \exp(-iH_{kl})$. Due to Lemma 6.25 we can simulate H_{kl} with overhead less than $\|H_{kl}\|$. By definition $\|H_{kl}\|$ is the angle of U_{kl}. □

Now we present an example showing that the complete Ising-Hamiltonian is able to implement some unitary transformations on n qubits faster than every circuit consisting of one and two qubit gates.

Theorem 6.30 (Higher parallelization)
The control theoretic model with the complete Ising-Hamiltonian allows to implement unitaries faster than within the quantum circuit model if we assume that the implementation time is given by the weighted depth.

Proof. Let H be the complete Ising-Hamiltonian acting on n qubits. Assume that for all $t > 0$ the transformation $\exp(-iHt)$ could be performed by a circuit with weighted depth t. Then its inverse $\exp(iHt)$ could also be implemented with weighted depth t (run this circuit backwards and substitute each gate by its inverse). Due to Corollary 6.29 this would imply that $\exp(iHt)$ could be implemented by H with running time t. For $t \to 0$ this implies that H can simulate $-H$ without time overhead. But this is a contradiction since Lemma 5.13 gives $n-1$ as a lower bound on the time overhead of time-reversal of H. □

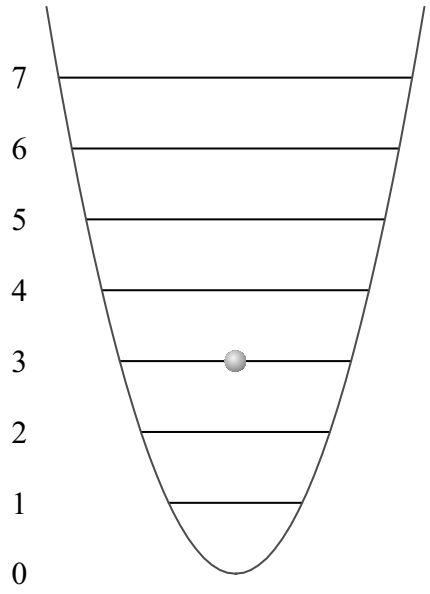


Figure 6.9: Discrete energy levels of a harmonic oscillator

6.4 Coupled harmonic oscillators

We consider the problem to simulate Hamiltonians of quantum networks consisting of n coupled harmonic oscillators.

The harmonic oscillator is one of the most important systems in quantum mechanics (cf. Sakurai [Sak94]). From a practical point of view it has applications in a variety of branches of modern physics – molecular spectroscopy, quantum optics, solid state physics, quantum statistical mechanics, and so forth. From a historical point of view it was M. Planck's proposal to associate discrete units of energy with radiation oscillators that led to the birth of quantum mechanics.

The state space of a harmonic oscillator is the Hilbert space $l^2(\mathbb{N})$ of square sumable sequences $|\Psi\rangle = \sum_{E=0}^{\infty} a_E |E\rangle$, i.e., $\sum_{E=0}^{\infty} |a_E|^2 < \infty$. The energy values of a harmonic oscillator with angular frequency ω are given by

$$\hbar\omega \left(n + \frac{1}{2} \right) , \quad \text{where } n = 0, 1, 2, \dots , \tag{6.29}$$

and $\hbar$ is Planck's constant. After rescaling and shifting the energy levels we may assume that the Hamiltonian is given by the (infinite) diagonal matrix

$$D = \operatorname{diag}(0, 1, 2, 3, \dots) ,$$

The discrete energy levels of this Hamiltonian are shown in Figure 6.9.

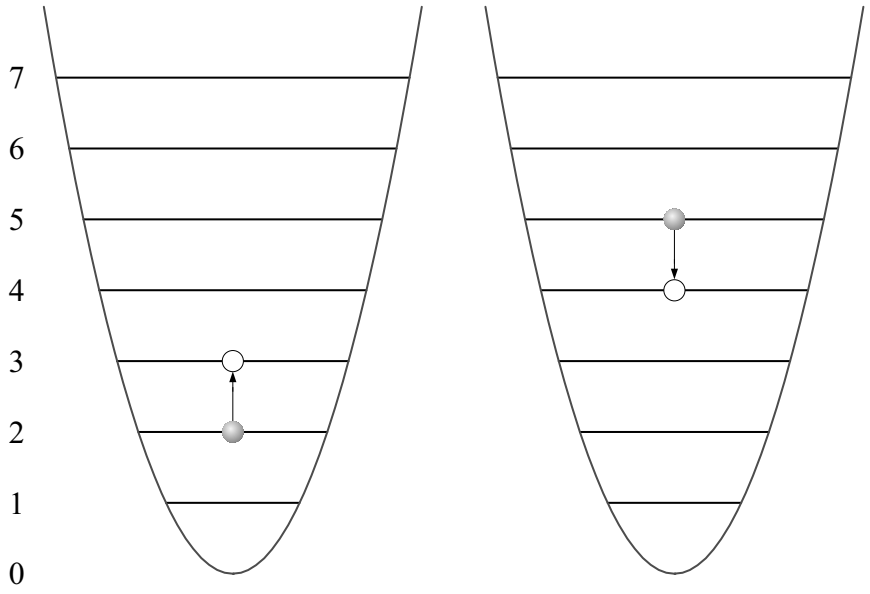


Figure 6.10: Coupling between two oscillators

There are two important operators associated with an oscillator. The *annihilation operator* a is defined by

$$a|0\rangle = 0\,, \quad a|E\rangle = \sqrt{E}|E-1\rangle \text{ for } E = 1, 2, \ldots\,.$$

The *creation operator* $a^\dagger$ is the adjoint of a. It is given by

$$a^\dagger|E\rangle = \sqrt{E+1}|E+1\rangle\,.$$

We consider a coupling between two oscillators of the form

$$a \otimes a^\dagger + a^\dagger \otimes a\,. \tag{6.30}$$

Intuitively, the effect of this coupling is that if one oscillator is excited to the next higher level then the other falls down to the next lower level. Therefore, the energy of the joint system (consisting of both oscillators) is conserved. The effect of this Hamiltonian is depicted in Figure 6.10.

Let us now consider a quantum network consisting of n oscillators. The Hamiltonian of the uncoupled system is given by

$$H_f := \sum_{k=1}^{n} D^{(k)}\,. \tag{6.31}$$

We assume that the couplings between the n oscillators are given by the homogeneous Hamiltonian

$$H := \sum_{k<l} w_{kl}\big(a^{(k)}a^{\dagger(l)} + a^{\dagger(k)}a^{(l)}\big)\,. \tag{6.32}$$

The matrix $W = (w_{kl})$ is a real symmetric $n \times n$-matrix with zeros on the diagonal determining the sign and strength of the couplings. We often write H_W to express the dependence of the Hamiltonian on W.

Interactions of the form in eq. (6.32) often appear if higher order terms in creation and annihilation operators are neglected and only the part of the total interaction term is considered that commutes with the uncoupled evolution corresponding to H_f.

Here we restrict our attention to the case of energy values less than d and obtain as an approximation a d-dimensional Hilbert space for each oscillator. The free Hamiltonian of an (uncoupled) oscillator is the diagonal matrix $D = \mathrm{diag}(0, 1, 2, \ldots, d-1)$ and the annihilation operators can be written as a finite sum

$$a = \sum_{E=1}^{d-1} \sqrt{E} |E-1\rangle\langle E| \,. \tag{6.33}$$

The creation operator is $a^\dagger = \sum_{E=0}^{d-2} \sqrt{E+1} |E+1\rangle\langle E|$.

6.4.1 Decoupling and time-reversal

The decoupling and time-reversal schemes based on orthogonal arrays in Section 3.2 can be used to decouple general pair-interaction Hamiltonians. It turns out that more efficient decoupling and time-reversal schemes can be constructed for interactions of the form in eq. (6.32). These schemes are based on generalized Hadamard matrices.

The coupling H_W between the harmonic oscillators can be removed by control operations in the one-parameter group $\{\exp(iDr) \mid r \in \mathbb{R}\}$. The control operation $\exp(iDr)$ is denoted by U_r.

Conjugation of the annihilation operator a by U_r gives a global phase factor e^{ir}, i.e., we have $U_r^\dagger a U_r = e^{ir} a$. This is seen as follows:

$$\begin{aligned} U_r^\dagger a U_r |E\rangle &= U_r^\dagger \, a \, e^{iEr} |E\rangle \\ &= e^{iEr} U_r^\dagger \sqrt{E} |E-1\rangle \\ &= \sqrt{E} e^{iEr} e^{-(E-1)r} |E-1\rangle \\ &= e^{ir} a |E\rangle \,. \end{aligned}$$

For the creation operator we have $U_s^\dagger a^\dagger U_s = e^{-is} a^\dagger$.

This shows that if the time evolution according to H_W is conjugated by transformations $U_r^{(k)}$ and $U_s^{(l)}$ on oscillator k and l, respectively, the coupling term $a^{(k)} a^{\dagger(l)}$ is multiplied with the phase factor $\exp(i(r-s))$, i.e.,

$$U_r^{\dagger\,(k)} U_s^{\dagger(l)} \, a^{(k)} a^{\dagger\,(l)} \, U_r^{(k)} U_s^{(l)} = \exp(i(r-s)) a^{(k)} a^{\dagger\,(l)} \,. \tag{6.34}$$

The adjoint term $a^{\dagger\,(k)}a^{(l)}$ obtains the phase factor $\exp(i(s-r))$.

Based on these observations we construct decoupling schemes with N time step all of equal length. We describe the decoupling schemes by $n \times N$ matrices M with complex numbers of modulus one as entries such that $MM^\dagger = N\mathbf{1}_n$. Such a matrix M defines a decoupling schemes for the class of Hamiltonians H_W as follows. If the vector

$$m_k = (e^{ir_{1j}}, e^{ir_{2j}}, \ldots, e^{ir_{nj}})$$

denotes the jth column of M, then this means that during the jth time step the natural time evolution of the n oscillators is conjugated by the local transformation

$$\prod_{k=1}^{n} U^{(k)}_{r_{kj}} \,. \tag{6.35}$$

The total effect of this scheme is that the term $a^{(k)}a^{\dagger\,(l)}$ obtains the factor

$$\langle m_k | m_l \rangle := \frac{1}{N} \sum_{j=1}^{N} e^{i(r_{kj} - r_{lj})} \,,$$

where $\langle \cdot | \cdot \rangle$ is the usual inner product in $\mathbb{C}^N$. This gives the following *decoupling criterion*: all couplings are removed if and only if the vectors m_k are orthogonal. This condition is satisfied if and only if $MM^\dagger = N\mathbf{1}_n$.

There is a canonical way of finding n vectors having this property by taking the Fourier transform of the standard basis of $\mathbb{C}^n$. However, the rotations $\exp(2\pi i\, kj/n)$ required to be implemented are very close to the identity for large n. This could be a severe disadvantage when applying the schemes to real physical systems.

An alternative way of constructing such matrices is given by difference schemes (cf. Beth et al. [BJL99] and Hedayat et al. [HSS99]). Recall that difference schemes have already been used in Subsection 3.2.2 to construct decoupling schemes for qubit networks with diagonal couplings.

Let $D(n, N)$ be a difference scheme based on the cyclic group $Z_u := \mathbb{Z}/u\mathbb{Z}$. We construct an $n \times N$ complex matrix M from D by replacing each entry d_{kl} of D by $e^{2\pi i\, d_{kl}/u}$. The rows of M are mutually orthogonal vectors in $\mathbb{C}^N$. Set $\omega := e^{2\pi i\, u}$. Let $d_k = (d_{k1}, \ldots, d_{kN})$ and $d_l = (d_{l1}, \ldots, d_{lN})$ be two different row of M. Then we have

$$\begin{aligned} \langle m_k | m_l \rangle &= \frac{1}{N} \sum_{j=1}^{N} e^{2\pi i\,(d_{kj} - d_{lj})/u} \\ &= \frac{\lambda}{N} \sum_{s=0}^{u-1} \omega^u \\ &= 0 \end{aligned}$$

because the difference vector of any two rows of D has the property that each element of Z_u occurs equally often (λ times) and all uth roots of unity sum to 0. This shows that M satisfies the decoupling criterion $MM^\dagger = N\mathbf{1}_N$.

More generally, matrices M of size $n \times n$ with the properties that all entries have modulus one and $MM^\dagger = n\mathbf{1}$ are called *generalized Hadamard matrices* . As we have seen such matrices define decoupling schemes. The construction based on difference schemes is one method to obtain generalized Hadamard matrices. We refer the read to HAAGERUP [Haa97] for other methods to construct generalized Hadamard matrices.

6.4.2 Optimality of decoupling and time-reversal

We now prove that the decoupling and time-reversal schemes based on generalized Hadamard matrices are optimal with respect to time overhead and number of time steps.

Theorem 6.31 (Optimality of schemes)
Generalized Hadamard matrices of size n define optimal[2] *decoupling and time-reversal schemes for a quantum network consisting of n coupled harmonic oscillators with the homogeneous system Hamiltonian*

$$H := \sum_{k<l} w_{kl}\big(a^{(k)}a^{\dagger(l)} + a^{\dagger(k)}a^{(l)}\big)$$

having a complete interaction graph.

Proof. The decoupling scheme based on generalized Hadamard matrices has n time steps. The time overhead and number of time steps of the corresponding time-reversal scheme is $n-1$.

To prove optimality we have to work out the coupling matrix of this Hamiltonian. By rescaling the Hamiltonian we may assume that $W = (w_{kl})$ is K, the adjacency matrix of the complete graph on n vertices.

Define the following linearly independent elements of $su(d)$.

$$X_E := |E\rangle\langle E-1| + |E-1\rangle\langle E|$$

and

$$Y_E := i|E\rangle\langle E-1| - i|E-1\rangle\langle E|$$

for $E = 0, 1, \ldots, d-1$. These matrices are orthogonal with respect to the inner product $\langle V|W\rangle := \mathrm{tr}(V^\dagger W)/d$. One may supplement these $2d-2$ vectors to an

[2] within the approximation that we consider only the first d energy levels of the oscillators

orthogonal basis of $su(d)$, but the completion is not important since the interaction among each pair of oscillators can be written as an expression in X_E and Y_E only:

$$a \otimes a^\dagger + a^\dagger \otimes a = \sum_{E,F} \sqrt{EF}(X_E \otimes X_F + Y_E \otimes Y_F)\,.$$

The coupling matrix J of the rescaled Hamiltonian can be constructed as follows. Define a $(2d-2) \times (2d-2)$-matrix C' by

$$C' := |\phi\rangle\langle\phi| + |\psi\rangle\langle\psi|\,,$$

with

$$|\phi\rangle := (\sqrt{1}, \sqrt{2}, \ldots, \sqrt{d-1}, \underbrace{0, 0, \ldots, 0}_{d-1})^T$$

and

$$|\psi\rangle := (\underbrace{0, 0, \ldots, 0}_{d-1}, \sqrt{1}, \sqrt{2}, \ldots, \sqrt{d-1})^T\,.$$

With respect to the basis described above, for all pairs (k, l) of oscillators the coupling matrices J_{kl} are the same and given by filling C' with zeros to the size $(d^2-1) \times (d^2-1)$-matrix. Denote this matrix by C. Since the vectors $|\phi\rangle$ and $|\psi\rangle$ are orthogonal, the spectrum of J_{kl} contains twice the value c, where $c = \langle\phi|\phi\rangle = \langle\psi|\psi\rangle > 0$. The other eigenvalues are all zero.

The spectrum of J contains twice the value $(n-1)c$ and $n-1$ times the value $-c$. The other eigenvalues are zero. This can be seen by writing J as a tensor product $K \otimes C$. Lemma 5.13 shows that $n-1$ is a lower bound on time overhead. This proves optimality with respect to time overhead.

To prove that n is a lower bound on the number of time steps for decoupling we use similar arguments to those in the proof of Theorem 6.1 (Case 3). If $(\tau_1, O_1; \ldots, \tau_N, O_N)$ is a decoupling scheme then we have

$$\sum_{j=1}^{N} \tau_j O_j (I_n \otimes C) O_j^T = \mathbf{1}_n \otimes C\,,$$

where $I_n := K_n + \mathbf{1}_n$. The rank of the l.h.s. is at most $2N$ and the rank of the r.h.s. is $2n$ since C has rank 2. It follows that $N \geq n$. This proves the optimality with respect to number of time steps. □

Note that a partially coupled system[3] consisting of harmonic oscillators can be decoupled with χ time steps, where χ is the chromatic number of the interaction graph. Time-reversal can be achieved with time overhead $\chi - 1$ and $\chi - 1$ time steps.

[3]We refer the reader to Section 3.5 for definitions of partially coupled systems and chromatic number.

6.4.3 Simulation of different coupling strengths and signs

The schemes based on generalized Hadamard matrices do not only allow to remove all interactions but also to achieve the following selective decoupling without time overhead. Partition the set of nodes into n' disjoint cliques and remove only the couplings between nodes in different cliques. This can be achieved by applying the same sequences of transformations on all nodes in the same clique, since this does not affect the interactions among them. Then it is sufficient to construct a difference scheme with only n' rows since each row refers to one of the cliques.

Note that an analogous way of clique decoupling is also possible for the following kind of n-qubit interaction. Assume that all qubits are coupled by the interaction $\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z$. Then the interacting is invariant with respect to simultaneous unitary rotations on both qubits. Hence decoupling schemes for n' qubits define a "clique decoupling" scheme for n' cliques.

Selective decoupling is a special instance of the general problem to simulate a coupling Hamiltonian $H_{\tilde{W}}$ using H_W. We define the matrix A as the Hadamard quotient $\tilde{W}/W$. Recall that we must assume that if an entry of W is equal to zero than the corresponding entry of $\tilde{W}$ must vanish. This is due to the fact that one cannot simulate a coupling between nodes which are not coupled within average Hamiltonian theory.

Let us first consider the case that the matrix A has only non-negative values, i.e., the signs of the interactions are not changed. An *upper* bound on the time overhead for this simulation problem can be derived in strong analogy to the problem to simulate Ising-Hamiltonians by the complete Ising-Hamiltonian. It is given by the *weighted clique coloring index* κ of the matrix A (see Definition 6.9). Analogously, if W contains negative values, than the *weighted chromatic index* of W gives an upper bound on the time overhead (see Definition 6.23).

Lower bounds on the time overhead and number of time steps can be derived from the theorems in Chapter 5. A lower bound on the time overhead is the smallest τ such that

$$\mathrm{Spec}(\tilde{W}) \preceq \mathrm{Spec}(\tau W)\,.$$

It follows from Theorem 5.12 since both Hamiltonians $J = W \otimes C$ and $\tilde{J} = \tilde{W} \otimes C$ are homogeneous and C has only two different eigenvalues $c > 0$ and 0. Theorem 5.20 (Case 1) gives the number of positive eigenvalues of A as a lower bound on the number of time steps.

Chapter 7

Quantum complexity classes

The field of complexity theory has long been studied in terms of classical physics. One of the most important concepts of (classical) complexity theory is NP-completeness. NP-completeness captures the combinatorial difficulty of a number of central problems which resisted the efficient solution and provides a method for proving that a combinatorial problem is as intractable as any NP problem.

In this chapter we study the quantum complexity classes *QMA* and *QCMA* that are two natural extensions of NP to the quantum model. While many problems are known to be NP-complete, there was only one problem known to be QMA-complete, namely the "local Hamiltonian" problem. This problem consists in deciding whether Hamiltonians have eigenstates with low energy. For QCMA there was no complete problem known.

In Janzing, Wocjan and Beth [JB03] we have defined the problem "identity check". It consists in deciding if a quantum circuit acts as the identity operator on the underlying Hilbert space. We have proved that it is QMA-complete. In Wocjan et al. [WJB03b] we have introduced the problem "identity check on basis states". It consists in checking if a quantum circuits acts as the identity operator on the basis states of the underlying Hilbert space (and not on the whole space). We have proved that it is QCMA-complete. Furthermore, we have defined in Wocjan et at [WJB03b] the problem of low-energy and low-complexity eigenstates and proved that it is QCMA-complete. This problem is related to the question "do Hamiltonians have eigenstates with low energy that can be prepared efficiently?".

7.1 NP-complete problems and Hamiltonians

Before introducing the quantum analogs of NP, we show that determining the ground state energy for some classes of pair-interaction Hamiltonians is NP-complete. The intractability is shown by reduction to the *NP-complete problems*

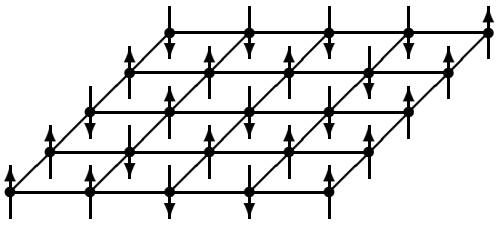


Figure 7.1: Two dimensional Ising model with spin a configuration represented by $\downarrow$ and $\uparrow$

max cut and *max independent set* in graph theory. This result is the basis for applying simulation of Hamiltonians within *adiabatic quantum computing* in Chapter 8. There we show how to simulate efficiently a *physically realistic* Hamiltonian whose ground states encode the solution to the max independent set.

We consider the Ising model that is of great interest in statistical physics. The Ising model can be formulated on any graph as follows. Consider a graph $G = (V, E)$ having n vertices $V = (1, \ldots, n)$, and a set E of edges. Each edge $(k, l) \in E$ has an associated constant *interaction energy* or *coupling constant* J_{kl}. The model is usually defined as the graph of a crystal lattice where the vertices represent lattice sites, and the edges represent near-neighbor interactions. The Ising-Hamiltonian is given by

$$H = \sum_{(k,l)\in E} J_{kl}\sigma_z^{(k)}\sigma_z^{(l)} + F\sum_k \sigma_z^{(k)}\,, \tag{7.1}$$

where F is the strength of an exterior magnetic field. Since all terms of the Hamiltonian commute, its energy eigenstates are the computational basis states.

To each vertex $k \in V$ we associate a *magnetic spin variable* m_k; it takes values $m_k = \pm 1$, where $+1$ represents the "up spin" and "-1" the "down spin". A *state* or a *spin configuration* $\mathbf{m}$ is an assignment of ± 1 to the variables m_k. An example of a two-dimensional Ising model with a spin configuration ($\uparrow$ corresponds to $+1$ and $\downarrow$ to -1) is shown in Figure 7.1. The energy of a spin configuration $\mathbf{m}$ is given by

$$H(\mathbf{m}) = \sum_{(k,l)\in E} J_{kl}m_k m_l + F\sum_{k\in V} m_k\,. \tag{7.2}$$

In this model two mathematical problems arise. The first is the study of the minimum energy configurations called *ground states*, and the second is the calculation of the partition function

$$\mathcal{Z}(\beta) = \sum_{\mathbf{m}\in\mathcal{M}} e^{-\beta H(\mathbf{m})}\,, \tag{7.3}$$

where $\mathcal{M}$ is the set of all spin configurations, $\beta = \frac{1}{\kappa T}$ is the inverse temperature, κ is the Boltzmann constant, and T is the temperature. Many physical properties

may be derived from the magnetic partition function. Let us mention that the free energy from the magnetic degree of freedom is $\kappa T \log \mathcal{Z}(T)$, and the equilibrium magnetic properties, magnetization, entropy, magnetic energy, specific heat and susceptibility, can all be obtained by differentiating the partition function with respect to the temperature.

The study of finite lattices belongs to the field of algorithmic combinatorics. For n spins, finding a ground state consists of searching for a spin configuration among 2^n that minimizes the energy. The partition function is a sum with 2^n terms.

For problems of this type there is a general agreement that if a problem cannot be solved in less than a number of calculations that grows exponentially with the size of the problem, then the problem should be considered inherently intractable. On the other hand, it is accepted that an efficient algorithm is an algorithm that requires a number of calculations bounded by a polynomial function of the size of the problem.

In the case of spin glasses the size of the problem is the number of spins. To understand the computational complexity of computing the ground state and the partition function of spin glasses, we must have a rough idea of the theory of *NP-completeness*. We refer the interested reader to the books by Ado et al. [AHU74] and Garey and Johnson [GJ79]. This theory provides straightforward techniques for proving that a given problem is "just as hard" as a large number of other problems that are widely recognized as being difficult and that have been confounding experts for decades. The class of these problems is called the class of non-deterministic polynomial-time complete (NP-complete) problems. It includes many "classical problems" in combinatorics, such as the traveling salesman problem, the Hamiltonian circuit problem, colorability of graphs and integer linear programming. All problems in the class have been shown to be equivalent, in the sense that if one problem is tractable, then all are.

The NP-problems are stated as decision problems. For example, in the traveling salesman problem the question is whether there is a tour between all cities having length B or less.

Now the mechanism for proving intractability is the same in statistical mechanics as in computational complexity: polynomial reduction.

7.1.1 Max cut

We reduce the problem to determine the energy of ground states in the Ising model with no exterior magnetic field to the "max cut" problem (cf. Karp [Kar72] and Garey and Johnson [GJ79]).

- INSTANCE: Weighted graph $G = (V, E)$, weight $w(e) \in \mathbb{N}$ for each $e \in E$, positive integer w.

- QUESTION: Is there a partition or *cut* of V into disjoint V_1 and V_2 such that the sum of the weights of the edges that have one end point in V_1 and one endpoint in V_2 is at least w?

This problem remains NP-complete if $w(e) = 1$ for all $e \in E$ (the "simple max cut" problem) (cf. GAREY ET AL. [GJS76]), and if, in addition, no vertex has degree exceeding 3 (cf. YANNAKAKIS [Yan78]). It can be solved in polynomial time if G is *planar* (cf. HADLOCK [Had75] and ORLOVA AND DORFMANN [OD72]). A graph is called planar if it is possible to represent it by a drawing in the plane in which the vertices correspond to distinct points and the edges to simple curves connecting its end vertices such that every two curves are either disjoint or meet only at a common endpoint.

The corresponding search problem is to find a maximal cut. Now we show that the problem to determine the ground state can be reduced to the "max cut" problem. Let us consider an Ising model on a graph $G = (V, E)$ with its edges weighted by $J > 0$ and with no exterior magnetic field. The energy of a state $\mathbf{m} = (m_1, \ldots, m_n)$ is given in this case by

$$H(\mathbf{m}) = J \sum_{(k,l)\in E} m_k m_l \,. \tag{7.4}$$

It is easy to see that the energies H can be defined in terms of *cuts* in G. Indeed, for a state $\mathbf{m} = (m_1, \ldots, m_n)$, let us denote by $V^+ = \{k \mid m_k = +1\}$, and by $V^- = \{k \mid m_k = -1\}$. This defines the cut $C = (V^-, V^+)$ of G. Let us also define E^+, and E^- as the set of edges with both end points in V^+, and respectively V^-. We divide the set of vertices in two parts. The cut refers to the set of edges that cross between the "up spins" vertices to the "down spin" vertices. Let E^{+-} be the set of edges in the cut, i.e., all edges with one endpoint in the V^+ and the other in the V^-. The weight of the cut is $w(C) = J|E^{+-}|$.

Clearly as $\mathbf{m}$ varies over all spin configurations, the corresponding cut C varies over all cuts of G. Observing that we actually have a one-to-one correspondence between spin configurations and cuts we can write the corresponding energies as follows:

$$\begin{aligned} H(C) &= J\left(|E^+| + |E^-| - |E^{+-}|\right) \\ &= J|E| - 2w(C)\,. \end{aligned} \tag{7.5}$$

If C is the cut defined by $\mathbf{m}$, $H(\mathbf{m})$ is the same as $H(C)$. For a given cut C, the energy is a constant term $J|E|$ (for the graph) minus twice the weight $w(C)$ of the cut. Minimizing the energy, i.e., finding the ground state, is therefore equivalent to solving the max cut problem in our graph. This shows the intractability for general interactions graphs. As a polynomial time algorithm to compute the partition function would permit us, in this case, to know the energy of the ground state, we conclude that computing the partition function is NP-hard.

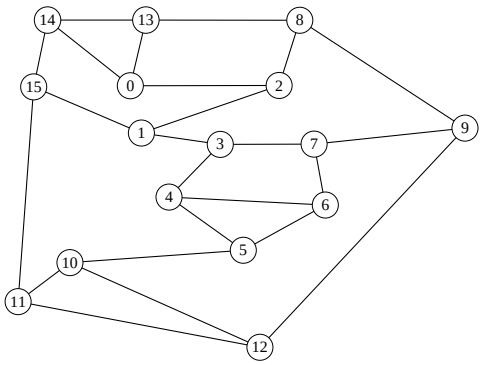


Figure 7.2: An example of a planar cubic graph

The situation does not improve if we restrict the graphs to more "physical" crystal lattices (cf. ISTRAIL [Ist00]).

7.1.2 Max Independent Set

We show that the "max independent set" problem may also be encoded in a suitably chosen Ising model. This observation is the basis for applying simulation of Hamiltonians in Chapter 8. The independent set problem [GJ79] is defined as follows:

- INSTANCE: Graph $G = (V, E)$, positive integer $v \leq |V|$.
- QUESTION: Does G contain an *independent set* whose cardinality is at least v, i.e., a subset $V' \subseteq V$ such that $|V'| \geq v$ and such that no two vertices in V' are joined by an edge in E?

The "independent set" problem problem remains NP-complete for cubic planar graphs [GJS76]. A graph is called *cubic* if all vertices have degree 3, i.e., all vertices are connected to exactly three vertices. Recall that a graph is called *planar* if it can be drawn in the plane such that the edges do not intersect. An example of a planar cubic graph is shown in Figure 7.2.

We now show that "max independent set" problem can be encoded in the ground states of a pair-interaction Hamiltonian.

Theorem 7.1 (Planar spin glass within a magnetic field)
Let $G = (V, E)$ be a cubic planar. Determining the energy of the ground states of the corresponding Hamiltonian

$$H = \sum_{(k,l)\in E} \sigma_z^{(k)}\sigma_z^{(l)} + \sum_{k\in V} \sigma_z^{(k)} \tag{7.6}$$

is equivalent to determining the maximum cardinality of independent sets of G.

Proof. This has been shown in [Bar82] (see also [WB03]). We include the proof here for completeness. We associate a variable $X_k \in \{0,1\}$ to each vertex $k \in V$. There is an independent set whose cardinality is at least v if and only if there is an assignment to the variables $\{X_k \mid k \in V\}$ such that

$$L = \sum_{k \in V} X_k - \sum_{(k,l) \in E} X_k X_l \geq v \,. \tag{7.7}$$

This is seen as follows. If V' is an independent set whose cardinality is at least v, then the assignment $X_k = 1$ for $k \in V'$ and $X_k = 0$ for $k \in V \setminus V'$ fulfills inequality (7.7).

Now let $X_1, \ldots, X_n$ be an assignment that fulfills inequality (7.7). If $V' = \{k \mid X_k = 1\}$ is not an independent set, then we must have $|V'| \geq v + p$, where $p := \sum_{(k,l) \in E} X_k X_l > 0$ is the "penalty" for V' not being an independent set. Let $(\tilde{k}, \tilde{l}) \in E$ with $X_{\tilde{k}} = X_{\tilde{l}} = 1$. By removing $\tilde{k}$ from V' (i.e. setting $X_{\tilde{k}} := 0$) the cardinality of V' drops by 1, while p drops by at least 1. After repeating this several times, we end up with an independent set whose cardinality is at least v. Setting $S_k = 2X_k - 1$ for all $k \in V$ and observing that $|E| = \frac{3}{2}|V|$ for all cubic graphs, we obtain

$$L = -\frac{1}{4} \sum_{k \in V} S_k - \frac{1}{4} \sum_{(k,l) \in E} S_k S_l + \frac{1}{8} |V| \,. \tag{7.8}$$

For $E = -4L + \frac{1}{2}|V|$ we see that there exists an independent set whose cardinality is at least k if and only if there is an assignment of values to the variables $S_k \in \{-1, 1\}$ (corresponding to the eigenvalues of σ_z) such that

$$E = \sum_{k \in V} S_k + \sum_{(k,l) \in E} S_k S_l \leq \frac{1}{2} |V| - 4v \,. \tag{7.9}$$

Now it is clear that determining the minimal energy E is equivalent to determining the maximal cardinality v of independent sets of G. □

7.2 QCMA and QMA - quantum analogues of NP

We have seen that the problem of determining the ground state energy for certain classes of Hamiltonians is NP-complete. We now introduce the complexity classes QCMA and QMA that are two possible extensions of NP to the quantum model. In the next section we show that the problem of determining the ground state

energy for specially constructed classes of Hamiltonians is complete for QCMA and QMA.

To define the quantum analogies of NP we adopt the common viewpoint that NP is the class of languages of those strings for which there exist polynomial-length proofs of membership that can be checked in polynomial time [Pap94].

This is usually explained in the following setting [KSV02]. Imagine two persons: King Arthur, whose mental capabilities are polynomially bounded, and a wizard Merlin, who is intellectually omnipotent. A is interested whether $x \in L$ (for example, to be sure if some Boolean formula x is satisfiable). M wants to convince A that indeed $x \in L$. But A does not trust M ("he is too clever to be loyal") and wants to make sure $x \in L$, not just believe M. So Arthur arranges that, after both see the input string, M writes a note to A where he "proves" that $x \in L$. Then A verifies this proof by some polynomial proof-checking procedure. In the case of satisfiability, the proof y is just a truth assignment such that the formula x evaluated on y gives true. If $x \notin L$ then it there is no proof y that convinces A wrongly that $x \in L$.

One may extend this verification viewpoint on NP to the quantum setting in several ways. For instance, we may consider *quantum proofs* (or *quantum certificates*), or we may consider ordinary (classical) certificates that are checked by polynomial-time quantum computers (a quantum circuit consisting of a polynomial number of gates). The certificates are then general quantum states or classical states (i.e. computational basis states), respectively.

We consider here both cases. The certificates may be classical or quantum and the polynomial time quantum verification procedure may operate with (two-sided) bounded error. The corresponding complexity classes are called QCMA and QMA. The "C" in QCMA means that the certificates are only allowed to be classical. These versions of "quantum NP" represent the quantum generalizations of the class MA (based on the Arthur-Merlin games of Babai [BM88]) that is the "probabilistic version" of NP.

The version with quantum certificates was apparently first discussed by Knill [Kni96c], and was later studied by Kitaev [KSV02] (who instead referred to the class we call QMA as BQNP). The class QCMA was mentioned in [Wat00] and defined explicitly in [AN03].

Now we define formally both complexity classes. Our definition of QMA is based on the presentation in [Wat00]. Let us begin by stating the assumptions regarding the uniformity of quantum circuits. A family $\{U_x\}$ of quantum circuits is said to be *polynomial-time uniformly generated* if there exists a deterministic procedure that, on input x, outputs a description of U_x and runs in polynomial time. For simplicity, we may assume that all input string are over the alphabet $\Sigma = \{0, 1\}$. It is assumed that the circuits in such a family are composed of a finite number of quantum gates that are universal (for instance, of gates in the Shor basis

[BMP+99]). Furthermore, it is assumed that the size of any circuit in such a family is not more than the length of that circuit's description (i.e., no compact descriptions of large circuits are allowed), so that U_x must have size polynomial in $|x|$.

For each circuit U_x, some number n_x of the qubits upon which U_x acts are specified as *input qubits*, and all other m_x qubits as *ancilla qubits*. The input qubits are assumed to be initialized in some specified input state $|\Psi\rangle$, while all ancilla qubits are initialized to the $|0\rangle$ state. One of the qubits is also specified as the output qubit and is assumed to be observed after the circuit has been applied. The probability that U_x accepts $|\Psi\rangle$ is defined to be the probability that an observation of the output qubit (in the $\{|0\rangle, |1\rangle\}$ yields 1, given that the input qubits are initially set to $|\Psi\rangle$.

Let us denote by $\mathcal{B}$ the Hilbert space of one qubit, by $\mathcal{B}^{\otimes n}$ the Hilbert space of n qubits, and by $\mathcal{S}(\mathcal{H})$ the set of density matrices on the Hilbert space $\mathcal{H}$. We now define the class as follows.

Definition 7.2 (QMA)
A language $L \subseteq \Sigma^*$ is in QMA if there exists a polynomial-time uniformly generated family of quantum circuits $\{U_x\}_{x\in\Sigma^*}$ such that

1. *if $x \in L$ then there exists a quantum state ρ that is accepted by U_x with probability greater than $2/3$, i.e.,*

$$\exists \rho \in \mathcal{S}(\mathcal{B}^{\otimes n_x}) \,:\, \mathrm{tr}(U_x(\rho \otimes |0\cdots 0\rangle\langle 0\cdots 0|)U_x^\dagger)P_1) > \frac{2}{3}\,,$$

   where P_1 is the projection onto $|1\rangle$ of the output qubit.

2. *if $x \notin L$ then all quantum states ρ are accepted by U_x with probability less than $1/3$, i.e.,*

$$\forall \rho \in \mathcal{S}(\mathcal{B}^{\otimes n_x}) \,:\, \mathrm{tr}(U_x(\rho \otimes |0\cdots 0\rangle\langle 0\cdots 0|)U_x^\dagger)P_1) < \frac{1}{3}\,,$$

Note that in our definition the quantum circuit U_x does not take x as input; but rather the procedure that produces the description of U_x takes x as input – the input ρ to a given circuit U_x corresponds to a quantum certificate that proves the property that $x \in L$. Information regarding x may of course be hard-coded into U_x; this eliminates the need for inputing x. It should be noted that the class QMA would not change if the definition was such that there were just one circuit for each input length (rather than for each input), with each circuit taking ρ and x as input (as would be the case for more standard notion of uniformity).

Note that our certificates are mixed states in contrast to the definitions in [KSV02, KR03]. Due to linearity arguments this modification does not change the complexity class.

Remark 7.3 (Amplification of probabilities)
Similar to classical bounded error classes, the bounds of $1/3$ *and* $2/3$ *in the definition of QMA may be replaced by* $2^{-e(|x|)}$ *and* $1-2^{-e(|x|)}$*, respectively, for any polynomial* e*. In the other direction, the bounds* $1/3$ *and* $2/3$ *may be replaced by functions* $b(|x|)$ *and* $a(|x|)$*, respectively, for* $a, b : \mathbb{N} \to [0,1]$ *such that (i)* a *and* b *are computable in polynomial time and (ii)* $a(|x|) - b(|x|) \geq 1/p(|x|)$ *for some polynomial* p*. In both cases, this follows from the fact that for any polynomial* q *we may run* $q(|x|)$ *independent copies of given verification procedure on a compound certificate consisting of* $q(|x|)$ *certificates of independent copies, and make a decision to accept or reject depending on the proportion of the individual copies that accept appropriately (Chernoff bound). A simple analysis reveals that entanglement among the individual certificates can yield no increase in the probability of acceptance as compared to the situation in which the certificates are not entangled, and that the probability of error is bounded by the tail of a binomial series as expected. This "amplification of probabilities" is described in [KSV02] in full detail.*

In the following we often assume that the probabilities are $1-\epsilon$ *and* ϵ*, where* ϵ *is a number that is exponentially small in* $|x|$*.*

We now allow only classical certificates, i.e., computational basis states. This class is called quantum classical MA (QCMA).

Let $\mathbb{B}^n$ denote the set of binary strings of length n. QCMA is defined as follows:

Definition 7.4 (QCMA)
A language $L \subseteq \Sigma^*$ *is in QCMA if there exists a polynomial-time uniformly generated family of quantum circuits* $\{U_x\}_{x\in\Sigma^*}$ *such that*

1. *if* $x \in L$ *then there exists a classical state* $|y\rangle$ *that is accepted by* U_x *with probability greater than* $2/3$*, i.e.,*

$$\exists |y\rangle \in \mathbb{B}^{n_x} \,:\, \mathrm{tr}(U_x(|y\rangle\langle y| \otimes |0\cdots 0\rangle\langle 0\cdots 0|U_x^\dagger)P_1) > \frac{2}{3}\,,$$

   where P_1 *is the projection onto* $|1\rangle$ *of the output qubit.*

2. *if* $x \notin L$ *then all classical states* $|y\rangle$ *are accepted by* U_x *with probability less than* $1/3$*, i.e.,*

$$\forall |y\rangle \in \mathbb{B}^{n_x} \,:\, \mathrm{tr}(U_x(|y\rangle \otimes |0\cdots 0\rangle\langle\Psi| \otimes \langle 0\cdots 0|U_x^\dagger)P_1) < \frac{1}{3}\,,$$

It is clear that $MA \subseteq QCMA \subseteq QMA$. The left inclusion is trivial. The right inclusion follows from the fact that the quantum verifier can force Merlin to send him a classical witness by measuring the witness before applying the quantum circuit on it.

Before we continue, let us summarize what is known about these classes in terms of complexity (for more detail see [AN03]). The most important class in quantum complexity theory is the class BQP (see Section 1.1.4). We have $BQP \subseteq QMA$. But how powerful is the class QMA? Can we upper-bound it? It was proved that BQP is contained in a large class, called PP . A language is in PP if there is a polynomial-time uniformly generated family of quantum circuits $\{U_x\}$ such that if $x \in L$ the circuit U_x outputs 1 with probability larger than 1/2, and if $x \notin L$ it outputs 0 with probability larger that 1/2. Note that the difference between the output probability and 1/2 can be exponentially small. This makes the class possibly much stronger than the class BPP; in particular, PP contains NP. It turns out that the above upper bound on BQP can be generalized to prove the same inclusions for the class QMA, i.e. $QMA \subseteq PP$.

Theorem 7.5 $BPP \subseteq BQP \subseteq QCMA \subseteq QMA \subseteq PP$.

This is almost all that is known regarding the relation of BQP and QMA to classical complexity classes. To give intuation about what this upper bound means regarding the quantum complexity power, we note that the class PP is known to be contained in perhaps a more natural class, PSPACE , which is the class of languages that can be recognized by a quantum circuits that use polynomial space (but can consist of exponentially many gates).

7.3 Local Hamiltonian problem

Having defined QMA and QCMA we show that the 3-local Hamiltonian problem is complete for both classes. We have alredy proved in Section 7.1 that pair-interaction qubit (2-local) Hamiltonians are sufficient to encompass NP. It is still an open question whether 2-local Hamiltonians are sufficient to achieve QMA-completeness.

One way to describe locality is a follows. Let $\mathbf{L}(\mathcal{B}^{\otimes s})$ denotes the set of linear operators from $\mathcal{B}^{\otimes s}$ to $\mathcal{B}^{\otimes s}$. Let $A \in \mathbf{L}(\mathcal{B}^{\otimes s})$ be an arbitrary operator and $S \subseteq \{1, \ldots, n\}$ with $|S| = s$. We denote by $A[S] \in \mathbf{L}(\mathcal{B}^{\otimes n})$ the embedding of the operator A into the Hilbert space $\mathcal{B}^{\otimes n}$, i.e., the operator that acts as A on the qubits specified by S and on the other as the identity.

Definition 7.6 (Local Hamiltonian)
A Hamiltonian $H : \mathcal{B}^{\otimes n} \to \mathcal{B}^{\otimes n}$ is called an s-local Hamiltonian if it is expressible in the form

$$H = \sum_j H_j[S_j]\,, \tag{7.10}$$

where each term $H_j \in \mathbf{L}(\mathcal{B}^{\otimes |S_j|})$ is a Hermitian operator acting on a set S_j, $|S_j| \leq s$, i.e., if it can be expressed as a sum of terms, where each term acts on

a bounded number of qubits.

Now we define the local Hamiltonian problem.

Definition 7.7 (Local Hamiltonian problem)
Let $\{H_x\}_{x\in\Sigma^*}$ be a family of s-local Hamiltonians. It required for all $x \in \Sigma^*$ that all summands $H_{x,j}$ of the decomposition $H_x = \sum_{j=1}^{r_x} H_{x,j}$ have bounded operator norm $\|H_{x,j}\| \leq poly(|x|)$ and that their entries are computable in polynomial time and are specified by only $poly(|x|)$ bits. Let $\{a_x\}_{x\in\Sigma^*}$ and $\{b_x\}_{x\in\Sigma^*}$ be two families of numbers computable in polynomial time such that $b_x - a_x > 1/poly(|x|)$. We are promised that the minimal eigenvalue of H_x is either at most a_x or greater than b_x. The local Hamiltonian problem consists in deciding which case is true.

The original definition of Kitaev required that $0 \leq H_j \leq 1$ (i.e., that both H_j and $\mathbf{1} - H_j$ have nonnegative eigenvalues). However, it is easy to see that the two definitions are equivalent. Given H_j such that $\|H_j\| \leq poly(|x|)$ for each j, normalize a, b, and all the H_j by a factor such that $\|H_j\| \leq 1/2$. Then, add $\frac{1}{2}\mathbf{1}$ to each H_j (such that $0 \leq H_j \leq 1$) and $r/2$ to a and b, where r is the number of terms in H.

Remark 7.8 (Local Hamiltonian problem as generalization of 3SAT)
Kitaev defined this problem as a quantum analogue of the 3SAT. The analogy is seen as follows. Let $\phi = C_1 \wedge C_2 \wedge \cdots \wedge C_r$ be a 3SAT formula on n variables, where each C_i is a clause, i.e., an OR over three variables or their negations. For each clause C_i we define an 8×8 matrix operating on three qubits. H_i is the projection on the unsatisfying assignment of C_i. For example, for the clause $C_1 = (X_1 \vee X_2 \vee \neg X_4)$ we get the matrix

$$H_1 = |001\rangle\langle 001|$$

since 001 is the only unsatisfying assignment for C_1.
We consider the operation of H_i on all qubits by taking the tensor product of H_i with identity $\mathbf{1}$ on the rest of the qubits. The new matrix will be denoted by again H_i in slight abuse of notation. For example if there are 5 variables, then we have

$$H_1 = |0\rangle\langle 0| \otimes |0\rangle\langle 0| \otimes \mathbf{1} \otimes |1\rangle\langle 1| \otimes \mathbf{1}\,.$$

If z is an assignment to the n variables satisfying a clause C_i, i.e. $C_i(z) = 1$, then we have $H_i|z\rangle = 0$, where $|z\rangle$ is the computational basis state $|z_1\rangle\otimes|z_2\rangle\otimes\cdots\otimes|z_n\rangle$. Otherwise, we have $H|z_i\rangle = |z\rangle$. The matrix H_i "penalizes" assignments that that do not satisfy C_i by giving them one unit of energy.
Let $H = \sum_{i=1}^{r} H_i$. With the observation above we see that

$$H|z\rangle = q|z\rangle\,,$$

where q is the number of clauses unsatisfied by z. All eigenvalues of H are non-negative numbers, and zero is an eigenvalue of H if and only if H corresponds to a satisfiable formula. Otherwise, the smallest eigenvalue of H is at least 1. Thus, 3SAT is equivalent to the problem: "Is the smallest eigenvalue of H 0 or is it at least 1?".

7.3.1 QMA-completeness

In Wocjan and Beth [WB03] we have shown that 2-locality is sufficient to encompass NP. This done by formulating the NP-complete problems "max cut" and "max independent set" in the setting of the 2-local Hamiltonian problem.

The results of Kitaev [KSV02] and Kempe and Regev [KR03] show that the 3-local problem is QMA-complete:

Theorem 7.9 *The 3-local Hamiltonian is QMA-complete.*

The proof that the 3-local Hamiltonian problem is in QMA can be found in [KSV02].

We only outline the idea of the proof that QMA may be reduced to the 3-local Hamiltonian problem. The task is to construct a 3-local Hamiltonian H corresponding to the quantum circuit U such that H has a small eigenvalue if and only if there is a quantum certificate that is accepted with high probability by U; otherwise H should have only large eigenvalues.

Let $U = U_L U_{L-1} \dots U_1$ consist of L two-qubit gates. The circuit U acts on $\mathcal{B}^l = \mathcal{B}^n \otimes \mathcal{B}^m$ qubits, where n and m are the size of the input and the ancilla register, respectively. Furthermore, we need an additional register called "clock" to construct the Hamiltonian. The clock consists of L qubits.

The total Hamiltonian consists of four parts:

$$H := H_{in} + H_{out} + H_{prop} + H_{clock} \,.$$

The intuition behind the construction is that low energy states of H represent in some sense the whole history of the quantum circuit U. The correlations between the clock register and the Hilbert space the circuit acts on contain the information at which time step which gate has been applied. This can be achieved as follows.

H_{clock} is defined as

$$H_{clock} := L^{12} \sum_{1 \le i < j \le L} |01\rangle_{ij}\langle 01|_{ij} \,.$$

The subscripts i, j indicate the considered qubits. H_{clock} acts only on the L clock qubits. It penalizes all states in the clock register that are not of the form

$$|\underbrace{11\dots1}_{t}\underbrace{0\dots0}_{L-t}\rangle \,,$$

called "unary representation" of the numbers $1, \ldots, L$. States of this form are denoted by $|\hat{t}\rangle$. They are the allowed states of the clock.

H_{in} is defined as

$$H_{in} := \sum_{i=m+1}^{N} |1\rangle_i\langle 1|_i \otimes |0\rangle_1\langle 0|_1 \,,$$

where $N = n + m$. The first component of the tensor product acts on the ancilla register and the second on the first qubit of the clock. It penalizes all states where the ancilla register is not initialized whenever the clock is in its starting position.

H_{prop} ensures that the correlations between the clock and the remaining registers are according to the history of the quantum circuit. It is defined as

$$H_{prop} := \sum_{t=1}^{L} H_{prop,t}$$

where

$$\begin{aligned} H_{prop,t} \;:=\; & \frac{1}{2}\big(\mathbf{1} \otimes |10\rangle_{t,t+1}\langle 10|_{t,t+1} + \mathbf{1} \otimes |10\rangle_{t-1,t}\langle 10|_{t-1,t} \\ & -U_t \otimes |1\rangle_t\langle 0|_t + U_t^\dagger \otimes |0\rangle_t\langle 1|_t\big) \end{aligned}$$

for $2 \leq t \leq L-1$ and

$$\begin{aligned} H_{prop,1} \;:=\; & \frac{1}{2}\big(\mathbf{1} \otimes |10\rangle_{1,2}\langle 10|_{1,2} + \mathbf{1} \otimes |0\rangle_1\langle 0|_1 \\ & -U_1 \otimes |1\rangle_1\langle 0|_1 + U_1^\dagger \otimes |0\rangle_1\langle 1|_1\big) \\ H_{prop,L} \;:=\; & \frac{1}{2}\big(\mathbf{1} \otimes |1\rangle_L\langle 1|_L + \mathbf{1} \otimes |10\rangle_{L-1,L}\langle 10|_{L-1,L} \\ & -U_L \otimes |1\rangle_L\langle 0|_L + U_L^\dagger \otimes |0\rangle_L\langle 1|_L\big) \,. \end{aligned}$$

H_{out} penalizes all states where the output is not in the state $|0\rangle$ whenever the clock is in its end position. It is defined as

$$H_{out} := |0\rangle_1\langle 0|_1 \otimes |1\rangle_L\langle 1|_L \,.$$

Assume there is a quantum certificate $|\Psi\rangle$ that is accepted by U with probability greater than $1 - \epsilon$. Define

$$|\eta\rangle = \sum_{t=0}^{L} U_j \cdots U_1\Big(|\Psi\rangle_{\text{input}} \otimes |00\cdots 0\rangle_{\text{ancillas}}\Big) \otimes |\tilde{t}\rangle_{\text{clock}} \,.$$

For this state representing the history of the quantum circuit U we have

$$\langle \eta|H|\eta\rangle < \frac{\epsilon}{L+1} \,.$$

This shows that H has an eigenvalue smaller than $\epsilon/(L+1)$ if there is a quantum certificate that is accepted with probability greater than $1-\epsilon$.

Now assume that all quantum states $|\Psi\rangle$ are accepted with probability smaller than b. Then one can show that all eigenvalues of H are greater than c/L^3 for some constant c. In case, ϵ is to large such that $\epsilon/(L+1) \geq c/L^3$ or the gap between both values is too small, we may use probability amplification to achieve a smaller ϵ.

7.3.2 QCMA-completeness

We have seen that the problem of determining the ground state energy of local Hamiltonians is QMA-complete. But there is a slightly different problem that is also interesting: determine if there are low energy states that can be prepared *efficiently*. Here efficiency is defined by the number of required elementary gates.

Note that it is not clear that it should be easier to decide whether there exist low energy states that can be prepared efficiently than to decide whether low energy states do exist at all. In special instances one may have arguments proving that small eigenvalues exist although one has no idea how to prepare the corresponding low energy states. However, we show that the *problem class* of deciding whether a general 3-local Hamiltonian has low-energy states includes the problem of deciding whether there are low-complexity and low-energy states. This follows from the fact that the first problem class is QMA and the latter one is QCMA.

Now we state the considered problem class formally.

Definition 7.10 (Problem of low-complexity and low-energy states) *Let $\{H_x\}_{x\in\Sigma^*}$ be a family of s-local Hamiltonians. It required for all $x \in \Sigma^*$ that all summands $H_{x,j}$ of the decomposition $H_x = \sum_{j=1}^{r_x} H_{x,j}$ have bounded operator norm $\|H_{x,j}\| \leq poly(|x|)$ and that their entries are computable in polynomial time and are specified by only $poly(|x|)$ bits. Let $\{a_x\}_{x\in\Sigma^*}$ and $\{b_x\}_{x\in\Sigma^*}$ be two families of numbers computable in polynomial time such that $b_x - a_x > 1/poly(|x|)$. We are promised that either*

1. *there is a quantum circuit V_x consisting of κ_x elementary gates such that*

$$|\Psi_x\rangle := V_x|00\cdots0\rangle$$

   is a state with energy less than a_x, i.e.,

$$\langle\Psi_x|H_x|\Psi_x\rangle < a_x\,,$$

2. *or all quantum circuits V that consist of at most κ_x gates can only prepare states $|\Psi\rangle$ with energy at least b, i.e.,*

$$\langle\Psi_x|H_x|\Psi_x\rangle > b_x\,,$$

where κ_x is some polynomial in $|x|$.
The problem of low-complexity and low-energy states is to decide which case is true.

Now we establish QCMA-completeness of this problem.

Theorem 7.11 *The problem of low-complexity and low-energy states is QCMA-complete.*

Proof. It is easy to see that the problem is in QCMA: the certificate is a classical string describing the preparation procedure V_x. The fact that $|\Psi_x\rangle$ is indeed a state with energy smaller than a_x can be checked as in the local Hamiltonian problem.

Now we show that the problem encompasses QCMA. We consider a quantum circuit U and the task is to decide whether there is a classical certificate (i.e. a computational basis state) that is accepted with high probability. Let n be the number of qubits of the input register and m be the number of ancilla qubits. We construct a circuit $\tilde{U}$ with n input qubits and $m+n$ ancillas as follows: n controlled-NOT gates copy the input to the n additional ancillas and then the circuit U is performed on the $n+m$ qubits. This copying operation renders the input classical. Therefore, $\tilde{U}$ has a quantum certificate that is accepted with high probability if and only if U has a classical certificate that is accepted with high probability.

Now we use the construction of [KR03] and obtain the 3-local Hamiltonian H associated with $\tilde{U}$. Let $\tilde{L}$ be the number of gates of $\tilde{U}$.

Consider the case that there is a classical certificate $|y\rangle$ that is accepted by U with probability at least $1-\epsilon$. It is accepted by $\tilde{U}$ with the same probability. Then the state

$$|\eta\rangle = \sum_{t=0}^{\tilde{L}} \tilde{U}_j \cdots \tilde{U}_1 \Big(|y\rangle_{\text{input}} \otimes |00\cdots 0\rangle_{\text{ancillas}} \Big) \otimes |\tilde{t}\rangle_{\text{clock}} \,.$$

is a low energy states (as for the local Hamiltonian problem). It remains to show that it can be prepared efficiently. Note that we have $|2^t - 1\rangle = |\tilde{t}\rangle$ on the clock register (that has size $L+1$). We omit technical details but it is obvious that the superposition

$$|\text{clock}\rangle = \frac{1}{\sqrt{L+1}} \sum_{t=0}^{L} |2^t - 1\rangle$$

can be prepared efficiently. The transformation

$$W = \sum_{t=0}^{\tilde{L}} \tilde{U}_j \cdots \tilde{U}_1 \otimes |1\rangle_t\langle 1|_t$$

can also be implemented efficiently. Now we obtain $|\eta\rangle$ by applying W to the state $|y\rangle \otimes |0 \cdots 0\rangle \otimes |\text{clock}\rangle$.

Therefore, we have shown that the question whether there is a low energy state that can be prepared with at most κ_x elementary gates is equivalent to the question whether there is a computation basis state that is accepted by U. □

7.4 Identity check

There was only one QMA-complete problem known so far, namely the 3-local Hamiltonian problem [KSV02, KR03]. In our work Janzing, Wocjan and Beth [JB03] we have constructed a new QMA-complete problem called *identity check*. This problem occurs naturally when constructing quantum networks from elementary gates. Given a classical description of a quantum circuit, determine whether it is almost equivalent to the identity. Explicitly, the task is to decide whether the corresponding unitary is close to a complex multiple of the identity matrix with respect to the operator norm.

A generalization of this problem is *equivalence check*: given two descriptions of quantum circuits and a description of a common invariant subspace, decide whether the restrictions of the circuits to this subspace almost coincide. We show that equivalence check is also in QMA and hence QMA-complete.

Furthermore, we show that the problem *identity check on basis states* is QCMA-complete.

7.4.1 Equivalence check

Now we introduce the problem equivalence check. Let U be a quantum circuit acting on n qubits that consists of two-qubit gates

$$U = U_k \cdots U_2 U_1 \,.$$

Imagine someone claims that U could also be implemented by another (perhaps much simpler) sequence of elementary gates

$$V_l \cdots V_2 V_1 \,.$$

Assume that he did not tell us why he thinks that this sequence also implements U. How difficult is it to determine whether it really does? Also the following slight modification of the above problem is natural. Usually, we are not interested in the whole physical state space but rather in a computational subspace. This subspace may, for instance, be defined by a quantum error correcting code [Ste96] or a decoherence free subspace [ZR97, VKL99]. In this case it is not relevant whether

the alternative circuit coincides with the original one on the whole space but only on the code space.

Assume that we already know (for example by construction) that the alternative circuit leaves the computational subspace invariant. Does the alternative circuit agree with the original one when it is restricted to this common invariant subspace? This is obviously equivalent to the question whether

$$V_1^\dagger V_2^\dagger \cdots V_l^\dagger U_k \cdots U_2 U_1$$

acts as the identity on the invariant subspace.

Now we define formally the problem equivalence check.

Definition 7.12 (Equivalence check)
Let $\{U_x\}_{x\in\Sigma^*}$ and $\{V_x\}_{x\in\Sigma^*}$ be two polynomial-time uniformly generated family of quantum circuits acting on n_x qubits. Furthermore, let $\{\mathcal{S}_x\}_{x\in\Sigma^*}$ be the family of the common invariant subspaces of U_x and V_x. It is assumed that the $\mathcal{S}_x$ are specified by a polynomial-time uniformly generated family $\{S_x\}_{x\in\Sigma^*}$ of quantum circuits such that $S_x\mathcal{S}_x = \mathcal{B}_1$ where $\mathcal{B}_1$ is the space of all states of $\mathcal{B}^{\otimes(n_x+1)}$ where the last qubit is in the state $|1\rangle$.

The problem equivalence check is to decide whether the restrictions of U_x and V_x to $\mathcal{S}_x$ coincide approximatively. More precisely, we assume that it is promised that

1. *either there is a vector $|\Psi\rangle \in \mathcal{V}$ such that*

$$\|(U_x V_x^\dagger - e^{i\phi}\mathbf{1})|\Psi_x\rangle\| \geq \delta$$

   for all $\phi \in [0, 2\pi)$

2. *or there exists an angle $\phi \in [0, 2\pi)$ such that for all vectors $|\Psi\rangle \in \mathcal{S}_x$*

$$\|(U_x V_x^\dagger - e^{i\phi}\mathbf{1})|\Psi\rangle\| \leq \mu\,,$$

where $\delta - \mu \geq 1/poly(|x|)$.

We first show that equivalence check is in QMA. Then we show that a special instance of equivalence check, namely to decide whether a circuit is almost equivalent to the identity, encompasses QMA. Hence equivalence check and identity check are both QMA-complete.

To prove that equivalence check is in QMA we have to describe how to give a certificate proving that U_x and V_x do not coincide. For an arbitrary unitary operator W the difference from multiples of the identity is a normal operator. Hence its operator norm is given by the greatest modulus of the eigenvalues. Therefore the operator norm distance between W and the set of trivial transformations (global phases) can be determined as follows.

Let W be an operator that has $\exp(i\alpha)$ and $\exp(i\beta)$ has as eigenvalues. Then the norm distance of W to $\exp(i\phi)\mathbf{1}$ is at least

$$\max\{|e^{i\alpha} - e^{i\phi}|, |e^{i\beta} - e^{i\phi}|\} \tag{7.11}$$

If $|\alpha - \beta| \leq \pi$ then the minimum of (7.11) is attained for $\phi := (\alpha + \beta)/2$ and the norm distance to the trivial transformations implementing global phases is consequently at least

$$|1 - e^{i(\alpha+\beta)/2}| = \sqrt{2(1 - \cos((\alpha - \beta)/2))}\,. \tag{7.12}$$

Let U'_x, V'_x be the restrictions of U_x and V_x to $\mathcal{S}_x$.

If the first case in the definition of equivalence check is true then there exist eigenvectors[1] $|\Psi_a\rangle$ and $|\Psi_b\rangle$ of $U'V'^{\dagger}$ with eigenvalues $e^{i\alpha}$ and $e^{i\beta}$, respectively, such that

$$\delta \leq \sqrt{2(1 - \cos((\alpha - \beta)/2))}\,.$$

In order to check that the eigenvalues corresponding to the given eigenvectors satisfy this criterion we can use the phase estimation algorithm [CEMM98].

If the second case in the definition of equivalence check is true then we have $\sqrt{2(1 - \cos((\alpha - \beta)/2)} \leq \mu$. To distinguish between the two cases the accuracy of the phase estimation has to be chosen such that $\cos((\alpha-\beta)/2)$ can be determined up to an error of $(\delta^2 - \mu^2)/4$. It remains to check whether $|\Psi_a\rangle$ and $|\Psi_b\rangle$ are elements of $\mathcal{S}$. This can be done using the circuit S.

Actually the setting of QMA problems (see Definition 7.2) requires that the certificate is one quantum state instead of two.

Formulated as an Arthur-Merlin game [KSV02] Merlin proves Arthur that U_x and V_x are not equivalent on $\mathcal{S}_x$ by proving that $U'V'^{\dagger}$ has eigenvalues of non-negligible distance. The quantum certificate is the state $|\Psi_a\rangle \otimes |\Psi_b\rangle$. The verification procedure can be done with the quantum circuit in Figure 7.3. It verifies that $U_x V_x^{\dagger}$ is not close to the identity on the subspace $\mathcal{S}$. The two copies of S check that the certificates are really elements of $\mathcal{S}$. The results of this check are copied to additional ancilla qubits by controlled-NOT gates. The main part of the circuit (A^k and F) is a usual phase estimation algorithm. The ancilla registers are initialized into the superposition state $(1/\sqrt{m}) \sum_{k \leq m} |k\rangle$ and control the implementation of $A^k := (U_x V_x^{\dagger})^k$. The state $|k\rangle$ obtains a phase according to the eigenvalues of A^k. After applying the Fourier transformations F the phases can be read out from the ancilla registers. A circuit D computes the phase difference and C checks whether the difference is sufficiently large and the certificates are elements of the subspace $\mathcal{S}$.

A priori it is not clear that Merlin cannot cheat by sending entangled (wrong) certificates. However, it is easily checked that the circuit in Figure 7.3 treats any

[1] We drop the subscript x

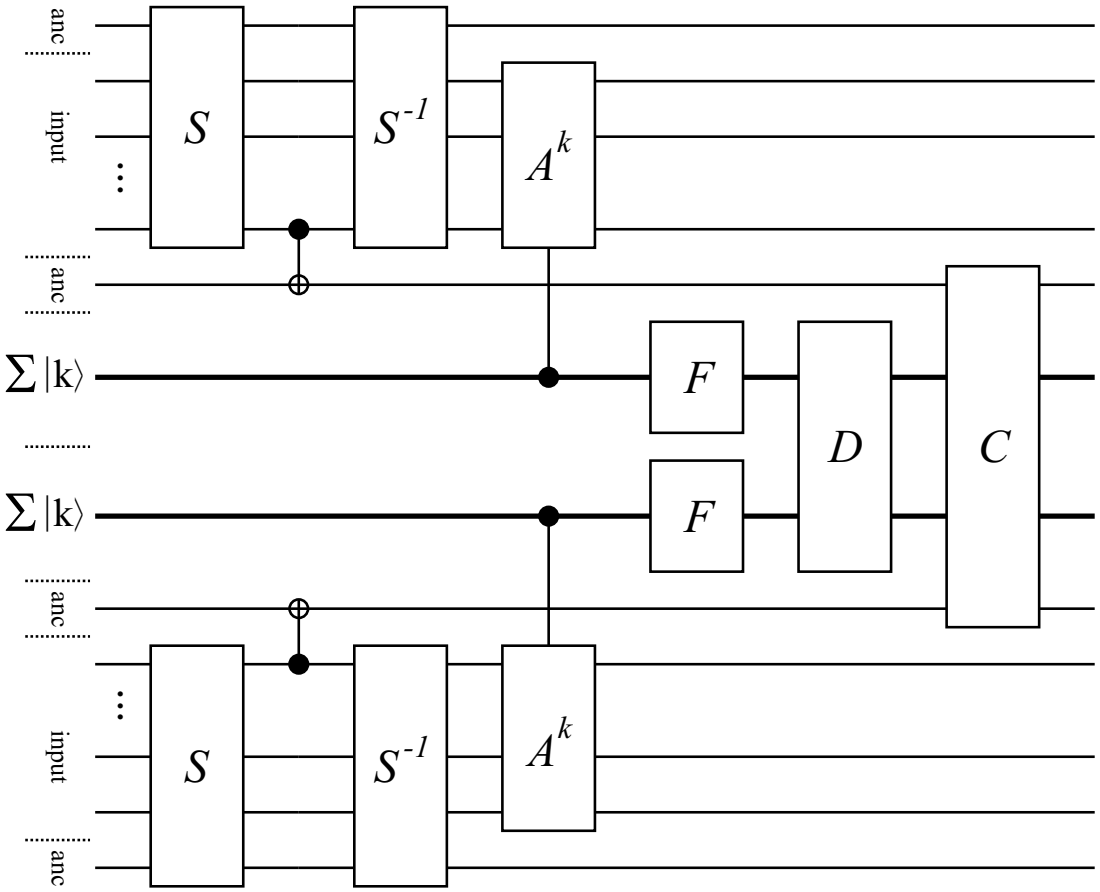


Figure 7.3: Quantum circuit used to verify that $U_x V_x^\dagger$ is not close to the identity on the subspace $\mathcal{S}$.

state of the form

$$\sum_j c_j |\Psi_a^j\rangle \otimes |\Psi_b^j\rangle$$

as an incoherent mixture of product states $|\Psi_a^j\rangle \otimes |\Psi_b^j\rangle$ with weights $|c_j|^2$. Note that it is also irrelevant whether the witness states $|\Psi_a\rangle$ and $|\Psi_b\rangle$ are really eigenstates of $UV^\dagger$. The phase estimation algorithm can only produce outputs that really exists as eigenvalues (up to the accuracy that is determined by the size of the used ancilla register).

7.4.2 QMA-completeness of identity check

Now we define identity check.

Definition 7.13 (Identity Check)
Let $\{U_x\}_{x\in\Sigma^*}$ be a polynomial-time uniformly generated family of quantum circuits

acting on n_x. It is promised that either

1. *for all $\phi \in [0, 2\pi)$*
$$\|U_x - e^{i\phi}\mathbf{1}\| \geq \delta_x$$
or
2. *or there exists an angle $\phi \in [0, 2\pi)$ such that*
$$\|U_x - e^{i\phi}\mathbf{1}\| \leq \mu_x \,.$$

is true, where δ_x and μ_x are computable in polynomial time and $\delta_x - \mu_x \geq 1/poly(|x|)$. The identity check problem is to decide which case is true.

Since it is a special instance of equivalence check we already know that it is in QMA. Now we prove that it is QMA-complete.

Recall that in the general QMA setting there is a quantum circuit U given and the problem consists in deciding whether there is a certificate $|\Psi\rangle$ such that the state
$$U|\psi\rangle \otimes |00\cdots0\rangle$$
has the property that the first qubit is with high probability in the state $|1\rangle$ (i.e. is accepted with high probability). In order to show that identity check encompasses QMA we construct a circuit Z that implements a unitary close to the identity whenever there is no state that is accepted by U and less close to the identity if there is a certificate.

The construction is as follows (see Figure 7.4). The register of the circuit U of identity check is extended by one qubit. The circuit Z is the transformation
$$Z := U^\dagger R_o U R_a \,.$$

The transformation R_a is a phase shift controlled by the states of the ancillas. Whenever the ancilla part of the register is initialized in the state $|00\cdots0\rangle$ the additional qubit gets a phase $\exp(i\varphi)$. The gate R_o is a phase shift controlled by the output qubit of U. The additional qubit gets a phase $\exp(i\varphi)$ whenever the circuit has accepted.

Theorem 7.14 (Identity check in QMA-complete)
Let $\{U_x\}_{x\in\Sigma^*}$ be a polynomial-time uniformly generated family of quantum circuits as in the definition of identity check. Then for the corresponding circuits Z_x in Figure 7.4 the following statements hold:

If the first case in the definition of identity check is true then we have
$$\|Z_x - e^{i\phi}\mathbf{1}\| \geq \sqrt{2(1-\cos\varphi)} - 2\sqrt{\epsilon}$$
for all $\phi \in \mathbb{R}$.

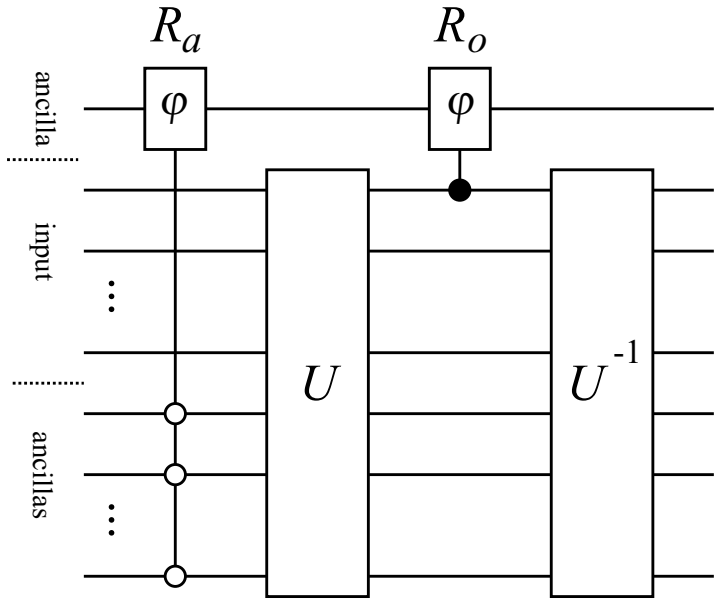


Figure 7.4: Quantum circuit Z consisting of $U, U^\dagger$ and two controlled φ-phase shifts R_a and R_o. If U rejects all states with high probability then the corresponding circuit Z is closer to the identity than in the case that there is a certificate that is accepted with high probability. The upper ancilla obtains a phase shift 2φ if and only if the ancillas of U have been initialized correctly and the input has been accepted by U.

If the second case is true then we have

$$\|Z_x - e^{i\varphi/2}\mathbf{1}\| \leq 2\sqrt{1-\cos(\varphi/2)} + 2\sqrt{2\epsilon}\,.$$

Proof. The effect of Z on a general state $|\Psi\rangle$ can be understood if we express $|\Psi\rangle$ as

$$|\Psi\rangle = |\Psi_1\rangle \oplus |\Psi_2\rangle\,,$$

where $|\Psi\rangle$ is a state with ancillas all set to 0 and $|\Psi_2\rangle$ is a state with ancilla register in states different from $|00\cdots0\rangle$. We have

$$Z|\Psi\rangle = U^\dagger R_o U R_a|\Psi_1\rangle \oplus U^\dagger R_o U R_a|\Psi_2\rangle\,.$$

Case 2. We first consider the second case (i.e. all quantum states are accepted by U with probability less than ϵ). The effect of Z on the summand $|\Psi_1\rangle$ is

$$U^\dagger R_o U R_a|\Psi_1\rangle = U^\dagger R_o P_1 U|\Psi_1\rangle \oplus U^\dagger R_o(\mathbf{1}-P_1)U R_a|\Psi_1\rangle\,,$$

where P_1 is the projection onto the state $|1\rangle$ of the output qubit. By the definition of the controlled phase shift R_o we have

$$R_o(\mathbf{1}-P_1) = (\mathbf{1}-P_1)\,.$$

This gives

$$\begin{aligned} Z|\Psi_1\rangle &= U^\dagger R_o P_1 U R_a|\Psi_1\rangle \oplus U^\dagger(\mathbf{1}-P_1)UR_a|\Psi_1\rangle \\ &= U^\dagger R_o\big(P_1UR_a|\Psi_1\rangle\big) + R_a|\Psi_1\rangle - U^\dagger\big(P_1UR_a|\Psi_1\rangle\big)\,. \end{aligned}$$

Since the probability of acceptance is at most ϵ the length of the vector $P_1UR_a|\Psi_1\rangle$ is at most $\sqrt{\epsilon}\|\,|\Psi_1\rangle\|$. We conclude that

$$\|Z|\Psi_1\rangle - R_a|\Psi_1\rangle\| \leq 2\sqrt{\epsilon}\|\,|\Psi_1\rangle|\,. \tag{7.13}$$

Note that

$$\|R_a|\Psi_1\rangle - e^{i\varphi/2}|\Psi_1\rangle\| \leq |1-\exp(i\varphi/2)|\,\|\,|\Psi_1\rangle\| \tag{7.14}$$

because $\|R_a - \exp(i\varphi/2)\mathbf{1}\| = |1-\exp(i\varphi/2)\|$ (due to arguments as in (7.11) and (7.12)). By combining the inequalities (7.13) and (7.14) we obtain

$$\|Z|\Psi_1\rangle - e^{i\varphi/2}|\Psi_1\rangle\| \leq \big(2\sqrt{\epsilon}+|1-\exp(i\varphi/2)|\big)\,\|\,|\Psi_1\rangle\|\,. \tag{7.15}$$

Now we consider the effect of Z on the second summand $|\Psi_2\rangle$. We have

$$\begin{aligned} \|Z|\Psi_2\rangle - e^{i\varphi/2}|\Psi_2\rangle\| &= \|U^\dagger R_o U R_a|\Psi_2\rangle - e^{i\varphi/2}|\Psi_2\rangle\| \\ &= \|U^\dagger(R_o - e^{i\varphi/2}\mathbf{1})U|\Psi_2\rangle \\ &\leq \|R_o - e^{i\varphi}\mathbf{1}\|\,\|\,|\Psi_2\rangle\| \\ &\leq |1-\exp(i\varphi/2)|\,\||\Psi_2\rangle\|\,. \end{aligned}$$

Together with inequality (7.15) we obtain

$$\begin{aligned} \|Z|\Psi\rangle - e^{i\varphi/2}|\Psi\rangle\| &\leq \big(|1-\exp(i\varphi/2)|+2\sqrt{\epsilon}\big)\,\big(\|\,|\Psi_1\rangle\| + \|\,|\Psi_2\rangle\|\big) \\ &\leq \sqrt{2}(|1-\exp(i\varphi/2)|+2\sqrt{\epsilon})\,. \end{aligned}$$

With $|1-\exp(i\varphi/2)| = \sqrt{2(1-\cos\varphi/2)}$ we obtain the desired inequality

$$\|Z - e^{i\varphi/2}\mathbf{1}\| \leq 2\sqrt{1-\cos(\varphi/2)} + 2\sqrt{2\epsilon}\,.$$

Case 1. We consider now the first case. Let $|\psi\rangle$ be a quantum certificate that is accepted by U with probability at least $1-\epsilon$. Define $P_0 := \mathbf{1} - P_1$. To prove the desired inequality we take the state vector

$$|\Psi\rangle := \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |\psi\rangle \otimes |00\cdots 0\rangle\,.$$

We have

$$\begin{aligned}
Z|\Psi\rangle &= U^\dagger R_o U R_a \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle \\
&= U^\dagger R_o U \frac{1}{\sqrt{2}}(|0\rangle + e^{i\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle \\
&= U^\dagger R_o (\mathbf{1} - P_0) U \frac{1}{\sqrt{2}}(|0\rangle + e^{i\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle + \\
&\quad U^\dagger R_o P_0 U \frac{1}{\sqrt{2}}(|0\rangle + e^{i\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle \\
&= U^\dagger (\mathbf{1} - P_0) U \frac{1}{\sqrt{2}}(|0\rangle + e^{i2\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle + \\
&\quad U^\dagger R_o P_0 U \frac{1}{\sqrt{2}}(|0\rangle + e^{i\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle \\
&= \frac{1}{\sqrt{2}}(|0\rangle + e^{i2\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle - \\
&\quad U^\dagger P_0 U \frac{1}{\sqrt{2}}(|0\rangle + e^{i2\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle + \\
&\quad U^\dagger P_0 U \frac{1}{\sqrt{2}}(|0\rangle + e^{i\varphi}|1\rangle) \otimes |\psi\rangle \otimes |0\ldots 0\rangle \\
&=: |\hat{\Psi}\rangle - |\varphi_1\rangle + |\varphi_2\rangle \,.
\end{aligned}$$

Note that the vectors $|\varphi_1\rangle$ and $|\varphi_2\rangle$ have at most norm $\sqrt{\epsilon}$ due to the high probability of acceptance. One checks easily that

$$\min_{\gamma\in\mathbb{R}} \| \, |\hat{\Psi}\rangle - e^{i\gamma}|\Psi\rangle\| = \| \, |\hat{\Psi}\rangle - e^{i\varphi}|\Psi\rangle\| = |1 - \exp(i\varphi)| \,.$$

We conclude

$$\min_{\gamma\in\mathbb{R}} \|Z|\Psi\rangle - e^{i\gamma}|\Psi\rangle\| \geq |1 - \exp(i\varphi)| - 2\sqrt{\epsilon} \,.$$

With $|1-\exp(i\varphi)| = \sqrt{2(1-\cos\varphi)}$ we conclude that the minimal norm difference between Z and $e^{i\gamma}\mathbf{1}$ is at least

$$\sqrt{2(1-\cos\varphi)} - 2\sqrt{\epsilon} \,.$$

□

As mentioned in Remark 7.3 the value of ϵ can be made arbitrarily small. For small φ the lower and upper bounds on the norm distances between U and the trivial transformations are approximatively given by

$$\varphi + 2\sqrt{2\epsilon}$$

and

$$\sqrt{2}\varphi - 2\sqrt{\epsilon}\,,$$

respectively. This shows that for sufficiently small ϵ there is a polynomial gap between the lower and upper bound. This shows that identity check is QMA-complete.

7.4.3 QCMA-completeness of identity check on basis states

In the previous section we stated the problem identity check. The task is to decide whether a (classically described) quantum circuit U is almost equivalent to the identity in the sense that there is a global phase ϕ such that the operator norm $\|U - \exp(i\phi)\mathbf{1}\|$ is close to zero. But also a weaker definition of equivalence is natural. Usually, quantum algorithms start with classical input (basis states as input) and end with measurements in the computational basis to obtain the classical output. In this context we do not care whether the circuits agree on all states; rather, we are only interested whether they agree on the computational basis states. In the following we show that this problem is QCMA-complete.

But what makes the difference between the original requirement and the weaker formulation? It is clear that a unitary operator that maps every basis state $|x\rangle\langle x|$ on itself may give different phases to different basis states. But one can see easily that this *does not* make the difference between QCMA and QMA (in case these classes are indeed different): The statement that a quantum circuit gives different phases to different basis vectors has still a classical proof. It is given by two numbers of basis states with non-negligible phase difference. The verifier can check the phase difference efficiently with the help of quantum phase estimation [CEMM98] (see also the quantum circuit in Figure 7.3).

So what can possibly make the difference between QMA and QCMA? It is the fact that there exist unitary transformations U that have large norm distance to all trivial transformations $\exp(i\phi)\mathbf{1}$ even though the distance between $U|x\rangle$ and $|x\rangle$ is exponentially small on all basis states $|x\rangle$. Let $U = HDH$, where H is the Hadamard transformation on n qubits and $D = \mathrm{diag}(-1, 1, 1, \ldots, 1)$ is a controlled phase shift on the first qubit. The minimal norm distance is attained for $\phi = 0$ and is in this case $\|\mathbf{1} - D\| = 2$. But for all computational basis states $|y\rangle$ we have

$$\|(\mathbf{1} - HDH)|y\rangle\| = \|H(\mathbf{1} - D)H|y\rangle\| = \|(2/2^n)\sum_{\tilde{y}} |\tilde{y}\rangle\| = 2/2^{n/2}$$

since $H\,\mathrm{diag}(1, 0, 0, \ldots, 0)\,H$ is the all-one-matrix.

Let us define the problem identity check on basis states.

Definition 7.15 (Identity check on basis states)
Let $\{Z_x\}_{x\in\Sigma^*}$ be a polynomial-time uniformly generated family of quantum circuits acting on n_x qubits. It is promised that

1. *either there is a binary string $z \in \mathbb{B}^{n_x}$ such that*

$$|\langle z|Z_x|z\rangle|^2 \leq 1 - \mu_x \,,$$

   i.e., Z_x does not act as the identity on the basis states,

2. *or for all binary strings $z \in \mathbb{B}^{n_x}$*

$$|\langle z|Z_x|z\rangle|^2 \geq 1 - \delta_x \,,$$

   i.e., Z_x acts "almost" as the identity on all computational basis states,

where $\mu_x - \delta_x \geq 1/poly(|x|)$. The problem of identity check on basis states is to decide which case is true.

It is easily seen that this problem is contained in QCMA since the proof for case 1 is given by a string that describes the basis state $|z\rangle$. Then we perform the quantum circuit Z_x. An n-fold controlled-NOT can be used to flip an additional ancilla qubit if and only if the output is $|y\rangle$. The additional ancilla is the output qubit of the verifier.

QCMA-completeness of this problem can be proved in strong analogy to the proof for QMA-completeness of identity check. Let U be a quantum circuit as in Definition 7.4. We construct the quantum circuit Z that uses U and $U^\dagger$ as subroutines.

Let R be the rotation

$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix},$$

with $0 < \varphi < \pi/2$ and R_a be the rotation R controlled by the m ancilla qubits corresponding to U. R_a is implemented if and only if the ancillas are correctly initialized in the state $|0\ldots0\rangle$. Let R_o be the same rotation R controlled by the output qubit of U. (To prove QMA-completeness of identity check we have used controlled phase shifts that are diagonal in the computational basis.) The whole circuit $Z := U^\dagger R_o U R_a$ is shown in Figure 7.5.

The following theorem shows that the problem of deciding whether there are basis states that are likely to be accepted by U can be reduced to identity check on basis states.

Theorem 7.16 (QCMA-completeness)
Let $\{U_x\}_{x\in\Sigma^*}$ be a polynomial-time uniformly generated family of quantum circuits

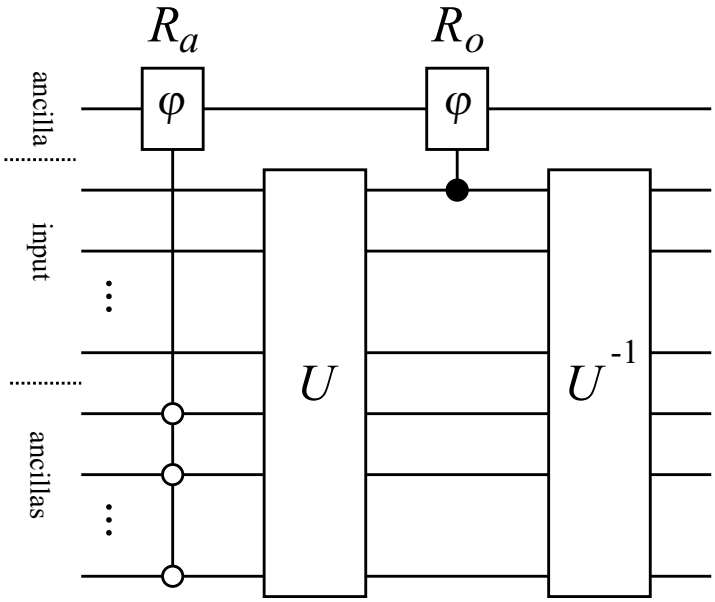


Figure 7.5: The circuit Z acts on all basis states almost as the identity if and only if no basis state is likely to be accepted by U. Note that Z has the same structure as the circuit in Figure 7.4; the difference is that the controlled operations are rotations and not phase shifts.

as in the definition of QCMA (Definition 7.4). Then the following statement holds for the corresponding circuit Z_x:
If case the first case of Definition 7.4 is true then there is a binary string z such that

$$|\langle z|Z|z\rangle|^2 \leq (\cos(2\varphi) + \sqrt{\epsilon})^2 \,,$$

where $|z\rangle = |0\rangle \otimes |y\rangle \otimes |00\cdots 0\rangle$ and y is the classical certificate for the the circuit U.
If the second case is true then for all binary strings z we have

$$|\langle z|Z|z\rangle|^2 \geq (\cos(\varphi) - 2\sqrt{\epsilon})^2 \,.$$

Proof. The proof is in strong analogy to the proof of Theorem 7.14. The important difference is that no superpositions between states with correctly and wrongly initialized ancillas have to be considered. Therefore the bounds are easier to derive.

Case 1. Let $|y\rangle$ be a binary string that is accepted by U with high probability (we drop the subscripts for fixed x). We consider the binary string $|z\rangle := |0\rangle \otimes |y\rangle \otimes |00\cdots 0\rangle$ to show that Z is "far" from the identity on the basis states.

$$\begin{aligned} Z|z\rangle &= U^\dagger R_o U R_a |z\rangle \\ &= U^\dagger R_o U (\cos\varphi)|0\rangle + \sin\varphi|1\rangle) \otimes |y\rangle \otimes |0\ldots 0\rangle \\ &= U^\dagger R_o (\cos\varphi)|0\rangle + \sin\varphi|1\rangle) \otimes (c_1|1\rangle \otimes |\psi_1\rangle + c_0|0\rangle \otimes |\psi_0\rangle) \end{aligned}$$

Due to the high probability of acceptance we have $|c_0| \leq \sqrt{\epsilon}$. Now we consider only the term with c_1 and obtain

$$\begin{aligned} & U^\dagger R_o (\cos\varphi)|0\rangle + \sin\varphi|1\rangle) \otimes c_1|1\rangle \otimes |\psi_1\rangle \\ = \; & (\cos(2\varphi))|0\rangle + \sin(2\varphi)|1\rangle) \otimes c_1 U(|1\rangle \otimes |\psi_1\rangle)\,. \end{aligned} \tag{7.16}$$

The first component is the single ancilla on which the rotation R is performed, the second component is the output of U and the third tensor component is the remaining part of the register where U acts on.

The overlap between the initial vector $|z\rangle$ and the vector of eq. (7.16) is at most $|c_1|\cos(2\varphi)$. Taking into account the length of the neglected vector we obtain

$$|\langle z|Z|z\rangle| \leq \cos(2\varphi) + \sqrt{\epsilon}\,.$$

Case 2. Let z be a string such that the bits corresponding to ancillas of U are all set to 0. Let P_1 as in Definition 7.4 be the projection onto the state $|1\rangle$ of the output qubit corresponding to U. Note that $R_o(\mathbf{1} - P_1) = \mathbf{1} - P_1$. Therefore, we have

$$\begin{aligned} |\langle z|Z|z\rangle| &= |\langle z|U^\dagger R_o U R_a|z\rangle| \\ &= |\langle z|U^\dagger R_o (P_1 + \mathbf{1} - P_1) U R_a|z\rangle| \\ &= |\langle z|U^\dagger R_o P_1 U R_a + R_a - U^\dagger P_1 U R_a|z\rangle| \\ &\geq |\langle z|R_a|z\rangle| - |\langle z|U^\dagger R_o P_1 U R_a|z\rangle| - |\langle z|U^\dagger P_1 U R_a|z\rangle| \\ &\geq \cos\varphi - 2\sqrt{\epsilon}\,. \end{aligned}$$

The latter inequality follows from the fact that the length of the vector $P_1 U R_a|z\rangle$ is at most $\sqrt{\epsilon}$ due to the small probability of acceptance.

Let z be a string such that the bits corresponding to the ancillas of U are not all set to 0. Then we have

$$\begin{aligned} |\langle z|U^\dagger R_o U R_a|z\rangle| &= |\langle z|U^\dagger R_o U|z\rangle| \\ &\geq \cos(\varphi)\,. \end{aligned}$$

This can be seen by writing $|z\rangle$ as $|\Psi_{-\varphi}\rangle \oplus |\Psi_0\rangle \oplus |\Psi_\varphi\rangle$, where $|\Psi_{-\varphi}\rangle, |\Psi_0\rangle$ and $|\Psi_\varphi\rangle$ are vectors in the eigenspaces of $U^\dagger R_o U$ corresponding to the eigenvalues $e^{-i\varphi}, 1$ and $e^{i\varphi}$, respectively. Therefore, we have

$$|\langle z|U^\dagger R_o U|z\rangle| = |pe^{-i\varphi} + qe^{i\varphi} + r|$$

with $p := \| |\Psi_{-\varphi}\rangle\|^2$, $q := \| |\Psi_0\rangle\|^2$ and $r := \| |\Psi_\varphi\rangle\|^2$. By elementary geometry the shortest vector in the convex span of the complex values $e^{-i\varphi}, 1, e^{i\varphi}$ has length $\cos(\varphi)$. This completes the proof. □

7.5 Remark on some problems in QCMA

We have shown that the problems "identity check" and "low-complexity and low-energy states" are QCMA-complete. In this section we discuss some problems that are contained in QCMA.

It is an interesting question of quantum information theory to find simple procedures for preparing certain entangled multi-particle states from unentangled initial states. Having found a procedure that prepares a desired state $|\psi\rangle$ from the state $|00\cdots0\rangle$, for instance, we may want to know whether there is also a simpler way to prepare $|\psi\rangle$.

Hence the following type of problems seems natural: Given a classical description of a quantum circuit U, decide whether there is also a simpler preparation procedure for $|\psi\rangle := U|00\cdots0\rangle$ in the following sense:

1. Given an elementary set of universal quantum gates. Decide whether there exists a quantum circuit V consisting of at most k gates preparing almost the same state (norm difference at most $1-\delta$) or all states prepared using at most k gates have at least the norm distance $1-\mu$ from $|\psi\rangle$

2. Let l and T be given, decide whether there is a l-local Hamiltonian preparing $|\psi\rangle$ approximatively by its autonomous evolution within the time T, i.e.,

   $$\exp(-iHt)|00\cdots0\rangle$$

   is almost the same state as $|\psi\rangle$ for an appropriate value $t \leq T$.

3. Consider the control-theoretic setting as, for instance, the one appearing in NMR-experiments (as described in Definition 2.5): given a pair-interaction Hamiltonian H and a maximal running time T. Let a state $|\Psi\rangle$ be specified by a quantum circuit as above, decide if it is possible to intersperse the natural time-evolution by at most k fast local operations (i.e. one-qubit rotations) such that the resulting unitary prepares the desired state without exceeding the maximally allowable running time T.

These types of problems are clearly in QCMA when the desired accuracies are defined as in Definition 7.4 since the proof consists of a classical description of the preparation procedure (i.e. the gate sequence, the Hamiltonian or the control sequence). The verifier can check that the procedure does indeed prepare the

Class	locality	deviation	number of queries	queried eigenvalues
NP	2	$O(1)$	1	1
QMA/QCMA	3	$O(1/p(\|x\|))$	1	1
PP	3	$O(1)$	1	$2^{n-1}+1$
#P	3	$O(1)$	n	binary search

Figure 7.6: Classifying complexity classes with respect to sorting eigenvalues of local Hamiltonians

desired state by simulating the described preparation procedure on a quantum computer and applying $U^\dagger$. Then he measures the obtained state in the computational basis.

We do not know whether these problems are contained in any lower complexity class.

7.6 Sorting eigenvalues of local Hamiltonians

We consider the problem of sorting the eigenvalues of local Hamiltonians. More precisely, we consider a hypothetical quantum machine that given i outputs the ith smallest eigenvalue of a local Hamiltonian with a certain precision. We show that such a machine would very powerful since some complexity classes could be solved efficiently.

NP: This follows e.g. from the fact that the NP-complete problems max cut and independent set may be encoded in pair-interaction Hamiltonian (see Section 7.1), i.e., determining the ground state energy is equivalent to solving these two problems.

QMA/QCMA: This follows from the fact that the local Hamiltonian problem is QCMA- and QMA-complete (see Section 7.3), i.e., determining the ground state energy with polynomial precision is complete for both classes.

PP: We have shown in Remark 7.8 that 3SAT may be encoded in a 3-local Hamiltonian. There is a satisfying assignment if and only if the minimal eigenvalue is 0. Otherwise, the minimal eigenvalue is at least 1. Furthermore, it is easily see that the multiplicity of the eigenvalue 0 is the number of satisfying assignments.

Consider the problem MAJSAT (majority SAT) that is PP-complete [Pap94]. MAJSAT is the problem to decide if is it true that the majority of the 2^n truth assignments to its variables (that is, at least $2^{n-1}+1$ of them) satisfy it? This can be decided if we know if the $(2^{n-1}+1)$th largest eigenvalue of the 3-local Hamiltonian corresponding to the Boolean expression. The majority of the truth assignments is true if and only if $\lambda_{2^{n-1}+1}$ is 0, where λ_i denotes the ith largest eigenvalue.

#P: Consider the problem #3SAT that is #P-complete [Pap94]. #3SAT is the problem to count the satisfying assignments. It is clear that the number of satisfying assignments equals the multiplicity of the eigenvalue 0 of the 3-local Hamiltonian corresponding to the Boolean expression. Therefore, #3SAT is equivalent to the problem of determining the dimension of the kernel of H. This dimension can be determined by invoking "sorted diag" at most n times. The method corresponds to binary search. We use "sorted diag" to determine the 2^{n-1}th eigenvalue of H. If it is 0 we continue the search in the upper half and ask what is the $(2^{n-1} + 2^{n-2})$-th eigenvalue. Otherwise, we continue in the lower half and ask what is the 2^{n-2}th eigenvalue. After at most $n = \log_2(2^n)$ queries we can determine the multiplicity of the eigenvalue 0.

Chapter 8

Application of Hamiltonian simulation within adiabatic quantum computing

In this chapter we construct a nearest-neighbor Hamiltonian whose ground states encode the solutions to the NP-complete problem "independent set" in cubic planar graphs. The Hamiltonian can be easily simulated by Ising interactions between adjacent particles on a 2D rectangular lattice. We describe the required pulse sequences. Our methods could help to implement adiabatic quantum computing by "physically reasonable" Hamiltonians like short-range interactions. The presentation is based on our work Wocjan et al. [WJB03a].

8.1 Adiabatic quantum computing

Adiabatic quantum computation has been proposed as a general way of solving computationally hard problems on a quantum computer [Fea01]. Adiabatic quantum algorithms proposed so far work by applying a time-dependent Hamiltonian

$$H(t) = (1 - \frac{t}{T})H_B + \frac{t}{T}H_P \tag{8.1}$$

that interpolates linearly from an initial Hamiltonian H_B to the final Hamiltonian H_P. The Hamiltonians are chosen such that the ground states of H_B are easily prepared and the ground states of the final Hamiltonian H_P encode the solutions to the problem [Fea01].

The running time of the algorithm is denoted by T. If $H(t)$ varies sufficiently slowly, i.e., T is sufficiently high, then one hopes that the final state of the quantum computer will be close to the ground state of the final Hamiltonian H_P, so a measurement will yield a solution to the problem with high probability. The adiabatic theorem is the justification for this hope. However, it is not clear whether

all necessary conditions for adiabatic evolution are satisfied. For instance, it is not clear whether the gap between the ground states and first excited states of $H(t)$ is sufficiently high for all t.

The adiabatic method can only succeed if the Hamiltonian $H(t)$ changes slowly. But how slow is slow enough? Unfortunately, this question has proved difficult to analyze in general. Some numerical evidence suggests the possibility that the adiabatic method might efficiently solve computationally interesting instances of hard combinatorial problems, outperforming classical algorithms [Fea01]. Whether adiabatic quantum computing provides a definite speedup over classical methods for certain problems remains an interesting open question.

Our objective is not to explore the computational power of the adiabatic quantum computing, but rather to investigate how to implement the adiabatic time evolution starting from "physically reasonable" Hamiltonians like short-range interactions.

A Hamiltonian can be considered as physically reasonable only if it is at least "local" (see Definition 7.6). Recall that a Hamiltonian is local if it can be expressed as a sum of terms, where each term acts on a bounded number of qubits. Indeed, in this case, the corresponding time evolution can be approximately simulated by a universal quantum computer [NC00].

For a *direct physical implementation* of the continuously varying Hamiltonian $H(t)$ we require a stronger *locality condition*. Physical interactions are usually pair-interactions (see Definition 2.2), unless one considers effective Hamiltonians. The system Hamiltonian can be thus decomposed as

$$H = \sum_{k<l} H_{kl} + \sum_{k} H_k \,, \tag{8.2}$$

H_{kl} is a Hermitian operator acting on the joint Hilbert space of particle k and l and H_k is the free Hamiltonian of particle k. Furthermore, the interaction strength is decreasing with the distance. Therefore, we do not want to propose a scheme that relies on "weak" interaction terms among distant particles. We thus require that each particle is coupled to only a few other particles in its direct neighborhood.

One of the most simple nontrivial examples are the Ising interactions on a 2D lattice. Our resource is the Ising Hamiltonian on an $r \times s$ rectangular lattices, i.e.,

$$H_{\text{Ising}} = \sum_{(k,l)\in L} \sigma_z^{(k)} \sigma_z^{(l)} \,, \tag{8.3}$$

where L are the edges of a rectangular lattice, i.e, a graph of order rs obtained by placing vertices at the coordinates $\{(i,j) \mid 0 \leq i < r, 0 \leq s < J\}$ with edges joining just the pairs at unit distance.

8.2 "Planar orthogonal" Hamiltonians

Due to the lattice structure of our resource Hamiltonian we need to simulate a pair-interaction Hamiltonian $\hat{H}_P$ whose interaction graph is a subgraph of the rectangular lattice; this necessity is because one cannot simulate within average Hamiltonian theory a coupling between nodes that coupled by th resource Hamiltonian.

Let L' be a subgraph of the rectangular lattice L. We construct a final Hamiltonian

$$\hat{H}_P = \sum_{(k,l)\in L'} w_{kl}\, \sigma_z^{(k)}\sigma_z^{(l)} + \sum_k w_k \sigma_z^{(k)} \quad \text{with } w_{kl}, w_k \in \mathbb{Z}\,, \tag{8.4}$$

such that its ground states encode the solution to the NP-complete problem "independent set". Clearly, such Hamiltonians satisfy the locality condition.

The idea behind this construction is as follows. Recall that determining the ground state energy of the pair-interaction Hamiltonian

$$H_P = \sum_{(k,l)\in E} \sigma_z^{(k)}\sigma_z^{(l)} + \sum_{k\in V} \sigma_z^{(k)}\,,$$

where $G = (V, E)$ is a planar cubic graph, is equivalent to solving the NP-complete problem "max independent set" (see Section 7.1.2). The problem with the Hamiltonian H_P is that does not fit into the structure of the rectangular lattice. Therefore, we need to embed our graph into this structure. The embedding can be done using known result on *planar orthogonal embedding* of graphs [KW01]. The idea to use planar orthogonal embedding was inspired by [KL02, RAS02].

In the following we show how the Hamiltonians $\hat{H}_P$ is constructed from the planar orthogonal embedding of G and explain why its the ground state energy encodes the solution to the "max independent set" problem for G. Furthermore, we show how $\hat{H}_P$ can be obtained efficiently from the 2D Ising model Hamiltonian H_{Ising}.

Together with the choice of a local initial Hamiltonian

$$\hat{H}_B = \sum_k \sigma_x^{(k)} \tag{8.5}$$

our results allow to simulate efficiently the adiabatic quantum evolution according to

$$\hat{H}(t) = (1 - \frac{t}{T})\hat{H}_B + \frac{t}{T}\hat{H}_P\,.$$

Now we define formally planar orthogonal embedding of graphs.

Definition 8.1 (Planar orthogonal embedding)
A planar orthogonal embedding Γ of a graph $G = (V, E)$ is a mapping that

- *maps vertices $k \in V$ to lattice points $\Gamma(k)$ and*

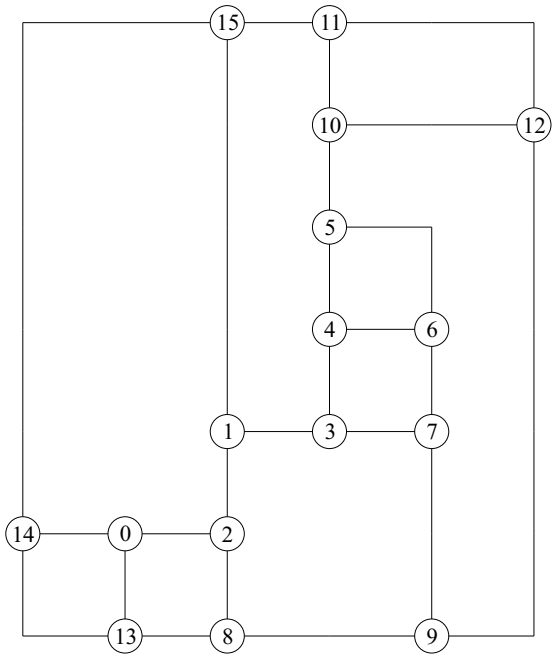


Figure 8.1: Planar orthogonal embedding of the graph in Figure 7.2

- *edges* $(k, l) \in E$ *to paths in the lattice such that the images of their endpoints* $\Gamma(k)$ *and* $\Gamma(l)$ *are connected and such that the paths do not share any vertices (besides the endpoints).*

Note that the map inserts "dummy vertices" are necessary to create the paths connecting the vertices Γ_k and Γ_l. A planar orthogonal embedding is shown in Figure 8.1.

Every planar graph with maximum degree 3 admits a planar orthogonal embedding on an $\lfloor n/2 \rfloor \times \lfloor n/2 \rfloor$. The algorithm presented in [Kan96] computes efficiently such planar orthogonal embedding of graphs. We used AGD (Library of Algorithms for Graph Drawing) to compute the embedding [AGD02].

In the proposal of [KL02] the Hamiltonian H_P is considered. The planar orthogonal embedding gives a regular wiring among the qubits. This means that the couplings are not spatially local. In contrast, we need a Hamiltonian $\hat{H}_P$ that contains only nearest-neighbor interactions. This is necessary that it can be simulated by H_{Ising}. The idea is to use the dummy vertices as wires that propagate the state of a (real) vertex spin to the neighborhood of another vertex. This can be achieved by constructing a path of adjacent dummy vertices, each interacting with its neighbor by a strong ferromagnetic coupling. Furthermore, the first dummy at one end of this "dummy path" is strongly ferromagnetically coupled to a vertex and the last dummy at the other end is in the neighborhood of another real vertex, coupled to it via a usual anti-ferromagnetic interaction. The interaction strength is chosen in such a way that it is always energetically better when

all dummies have the same state as the real vertex to which they are connected to than to have a mismatch along the "ferromagnetic path".

Formally, this construction is as follows:

- The dummy vertices have no local σ_z term.
- The vertices $\Gamma(k)$ have σ_z as local Hamiltonians.
- Let $(k,l) \in E$ be an edge of G.

  If Γ_k and Γ_l are adjacent, then the coupling between Γ_k and Γ_l is chosen to be anti-ferromagnetic, i.e., $\sigma_z \otimes \sigma_z$.

  If Γ_k and Γ_l are not adjacent, then there are m dummy vertices $v_1, \ldots, v_m$ such that the path $(\Gamma_k, v_1, \ldots, v_m, \Gamma_l)$ connects the vertices Γ_k and Γ_l. The couplings between Γ_k and v_1 and v_i and v_{i+1} for $i = 1, \ldots, m-1$ are chosen to be ferromagnetic with coupling strength c, i.e., $-c\,\sigma_z \otimes \sigma_z$. The coupling between v_m and Γ_l is chosen to be anti-ferromagnetic, i.e., $\sigma_z \otimes \sigma_z$.

The corresponding "planar orthogonal Hamiltonian" is shown in Figure 3. The filled circles correspond to dummy vertices that do not have any local Hamiltonian. The circles with indices correspond to the original vertices of G. They have σ_z as local Hamiltonians. The thin lines correspond anti-ferromagnetic interactions and the thick lines to ferromagnetic interactions.

The idea behind this construction is that there is a direct one-to-one correspondence between the ground states of H_P and $\hat{H}_P$. The same is true for the first excited states. This can be seen as follows:

Let $(k,l) \in E$ be an edge of G and $(\Gamma_1, v_1, \ldots, v_m, \Gamma_l)$ be the path on the lattice connecting Γ_k and Γ_l. The variables $S_{\Gamma_k}, S_1, \ldots, S_m \in \{0,1\}$ indicate whether the corresponding qubit is spin up or spin down.

A ground state satisfies the condition that $S_1, \ldots, S_m$ are all equal to S_{Γ_k}. To show this we define the number of mismatches along the path to be the number of occurrences of $S_{\Gamma_k} \neq S_1$, $S_i \neq S_{i+1}$ for $i = 1, \ldots, m-1$. This number is denoted by δ.

Then the minimal possible energy (due to the couplings along the path) is

$$c(-m + \delta) - 1\,. \tag{8.6}$$

If we remove the mismatches (by setting $S_i := S_{\Gamma_k}$ for $i = 1, \ldots, m$) then the maximal possible energy is

$$-cm + 1\,. \tag{8.7}$$

By choosing $c = 3$ minimal energy can be achieved only if the states of all dummy vertices are equal to the state of the qubit corresponding to Γ_k.

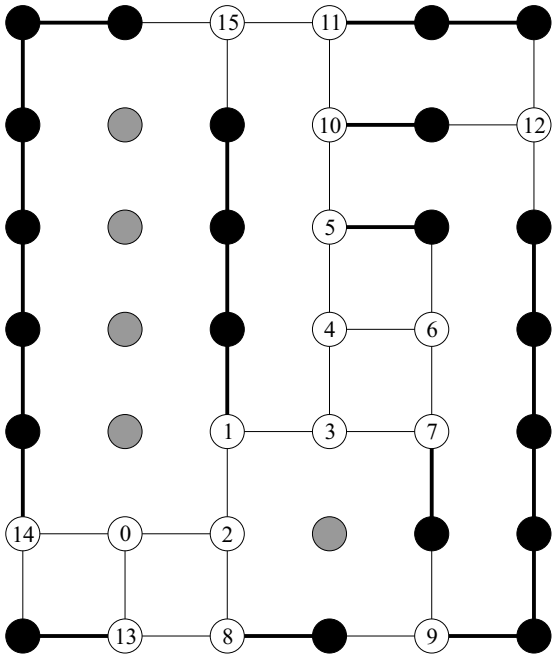


Figure 8.2: Hamiltonian corresponding to the planar orthogonal embedding in Figure 8.1

For adiabatic quantum computing it is important that the gap between the ground and first excited states of the Hamiltonian at all times is sufficiently large. We show that the modification of H_P to $\hat{H}_P$ does not decrease this gap.

The gap between the ground and the first excited states of H_P is smaller or equal to 8. This is seen as follows. Let $S_1, \ldots, S_n \in \{-1, +1\}$ be an assignment corresponding to a ground state of H_P. Pick any vertex k and let l_1, l_2, l_3 be at the three vertices connected to k. By flipping S_k the energy can increase by at most 8 because the relevant Hamiltonian is

$$\sigma_z^{(k)} + \sum_{i=1}^{3} \sigma_z^{(k)} \sigma_z^{(l_i)} .$$

By choosing $c = 9$ it is seen that the first excited states of $\hat{H}_P$ satisfy the condition that the states all of dummy vertices are equal to the vertex of Γ_k.

8.3 Simulating "planar orthogonal" Hamiltonians

Starting from the Ising Hamiltonian H_{Ising}, we can implement the Hamiltonian $\hat{H}_P$ with time overhead (slow-down) $2c + 1$ and 16 time steps by interspersing

the time evolution according to H_{Ising} by local operations in $X \otimes X \otimes \cdots \otimes X$, where $X = \{\mathbf{1}, \sigma_x\}$.

Based on the results of [LCYY00, JWB02b] (see also Section 3.2.3) we construct a selective decoupling scheme based on Hadamard matrices. Due to the special form of H_{Ising} it is sufficient to use the Hadamard matrix of size 4 only.

Our scheme consists of 4 subroutines that implement the following couplings of $\hat{H}_P$:

1. horizontal $\sigma_z \otimes \sigma_z$,
2. vertical $\sigma_z \otimes \sigma_z$,
3. horizontal $-c\,\sigma_z \otimes \sigma_z$, and
4. vertical $-c\,\sigma_z \otimes \sigma_z$

The indices i, j enumerate the rows and the columns of the lattice. We denote the columns of the Hadamard matrix of size 4

$$W := \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

by $W(0,0), W(0,1), W(1,0)$ and $W(1,1)$.

Let $v = (v_1, v_2, v_3, v_4) \in \{-1, 1\}^4$ be a column vector. We use the abbreviation

"apply v at (i, j)"

to denote the following control sequence with 4 equally long time steps: at the beginning and the end of the sth time step we apply σ_x on the qubit at position (i, j) if $v_s = -1$ and do nothing if $v_s = 1$, where time step s runs from 1 to 4.

Let $v, v' \in \{-1, 1\}^4$. One easily checks that applying v and v' at adjacent lattice points changes $\sigma_z \otimes \sigma_z$ to $\langle v, v' \rangle\, \sigma_z \otimes \sigma_z$, where $\langle v, v' \rangle$ denotes the inner product of v and v'. This is the key observation for constructing the selective decoupling scheme.

In the first and second subroutines the length of the 4 time steps is chosen to $1/4$. Let us consider the first subroutine. The vertical couplings are automatically removed if we apply in rows with even indices only $W(0,0)$ and $W(1,0)$ and in rows with odd indices $W(1,0)$ and $W(1,1)$. The choice between $W(a,0)$ and $W(a,1)$ depends on which horizontal interactions should remain or be switched off. Explicitly, this choice is as follows. Choose $W(a,0)$ for the leftmost spin. If the interaction between the spins $(j-1)$ and j should remain, then apply the same $W(a,b)$ to j as to $(j-1)$. Otherwise (i.e. the coupling should be switched off) apply the second possible $W(a,b')$ to j.

The second subroutine is obtained from the first subroutine by exchanging the roles of rows and columns of the lattice.

In the third and fourth subroutines the length of the 4 time steps is chosen to $c/4$. The third subroutine is obtained from the first subroutine by apply $(-1)^j v$ instead of v to the spin j. Finally, the fourth subroutine is obtained from the third subroutine by exchanging the roles rows and columns of the lattice.

Bibliography

[AGD02] Algorithms for Graph Drawing. *User manual*, 2002. `www.ads.tuwien.ac.at /AGD/MANUAL/MANUAL.html`.

[AHU74] A. Aho, J. E. Hopcroft, and J. D. Ullman. *The design and analysis of computer algorithms*. Addison-Wesley, 1974.

[AN03] D. Aharonov and T. Naveh. Quantum NP - A Survey. LANL e-print quant-ph/0210077, 2003.

[And89] T. Ando. Majorization, doubly stochastic matrices, and comparison of eigenvalues. *Linear Algebra and Its Applications*, 118:163–248, 1989.

[And94] T. Ando. Majorizations and inequalities in matrix theory. *Linear Algebra and Its Applications*, 199:17–64, 1994.

[AS72] M. Abramowitz and C. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,* 9*th printing*. Dover, New York, 1972.

[AU82] P. M. Alberti and A. Uhlmann. *Stochasticity and partial order: doubly stochastic maps and unitary mixing*. Dordrecht, Boston, 1982.

[Bar82] F. Barahona. On the computational complexity of Ising spin models. *J. Phys. A: Math. Gen.*, 15:3241–3253, 1982.

[Bar01] E. R. Barnes. A lower bound for the chromatic number of a graph. *Contemporary Mathematics*, 275:3–12, 2001.

[BBC+95] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, J. A. Smolin, and H. Weinfurter. Elementary gates for quantum computation. *Physical Review A*, 52(5):3457–3467, 1995.

[BCL+02] C. H. Bennett, J. I. Cirac, M. S. Leifer, D. W. Leung, N. Linden, S. Popescu, and G. Vidal. Optimal simulation of two-qubit Hamiltonians using general local operations. *Physical Review A*, 66:012305, 2002.

[BCP97] W. Bosma, J.J. Cannon, and C. Playoust. The Magma algebra system I: The user language. *J. Symb. Comp.*, 24:235–266, 1997.

[Bha96] R. Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer, 1996.

[BJ30] M. Born and P. Jordan. *Elementare Quantenmechanik*. Springer, 1930.

[BJL99] Th. Beth, D. Jungnickel, and H. Lenz. *Design Theory*, volume I of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, 2nd edition, 1999.

[BM88] L. Babai and S. Moran. Arthur-merlin games: a randomized proof system, and a hierchy of complexity classes. *Journal of Computer and System Sciences*, 36(2):254–276, 1988.

[BMP+99] P. Boykin, T. Mor, M. Pulver, V. Roychowdhury, and F. Vatan. On universal and fault-tolerant quantum computing: a novel basis and a new constructive proof of universality for shor's basis. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, pages 486–494, 1999.

[Bol98] B. Bollobás. *Modern graph theory*, volume 184 of *Graduate Texts in Mathematics*. Springer, 1998.

[CCNW85] J. H. Conway, R. T. Curtis, S. P. Norton, and R. A. Wilson. *Atlas of Finite Groups*. Clarendon Press, Oxford, 1985.

[CD96] C. J. Colbourn and J. H. Dinitz. *Handbook of Combinatorial Designs*. CRC Press, 1996.

[CDS95] D. M. Cvetković, M. Doob, and H. Sachs. *Spectra of Graphs*. Johann Ambrosius Barth, 3rd edition, 1995.

[CEMM98] R. Cleve, A. Ekert, C. Macchiavello, and M Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London A*, 454(1969):339–354, 1998.

[Che02] H. Chen. Necessary conditions for the efficient simulation of Hamiltonians using local unitary operations. LANL e-print quant-ph/0109115, 2002.

[Chu97] F.R.K. Chung. *Spectral Graph Theory*, volume 92 of *CBMS*. American Mathematical Society, 1997.

[CRSS98] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane. Quantum Error Correction Via Codes over GF(4). *IEEE Trans. Inform. Theory*, 44:1369–1387, 1998.

[CZ95] I. Cirac and P. Zoller. Quantum computation with cold trapped ions. *Physical Review Letters*, 74:4091–4094, 1995.

[Deu97] D. Deutsch. *The Fabric of Reality*. The Penguin Press, 1997.

[Dir58] P. M. A. Dirac. *The Principles of Quantum Mechanics*. Clarendon Press, Oxford, 4th edition, 1958.

[DNBT01] J. L. Dodd, M. A. Nielsen, M. J. Bremner, and R. T. Thew. Universal quantum computation and simulation using any entangling Hamiltonian and local unitaries. LANL e-print quant-ph/0106064, 2001.

[DRBH95] P. Domokos, J. Raimond, M. Brune, and S. Haroche. Simple cavity QED two-bit universal logic gate. *Physical Review A*, 52(5):3554–3559, 1995.

[EBW87] R. R. Ernst, G. Bodenhausen, and A. Wokaun. *Principles of nuclear magnetic resonance in one and two dimensions*. Clarendon Press, 1987.

[Fea01] E. Farhi et al. A Quantum Adibatic Evolution Algorithm Applied to Random Instances of an NP-complete Problem. *Science*, 292:472, 2001.

[Fey82] R. P. Feynman. Simulating physics with computers. *Int. J. Theor. Phys.*, 21:467, 1982.

[GAP97] The GAP Team, Lehrstuhl D für Mathematik, RWTH Aachen, Germany and School of Mathematical and Computational Sciences, U. St. Andrews, Scotland. GAP – *Groups, Algorithms, and Programming, Version 4*, 1997.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A guide to the Theory of NP-Completness*. W. H. Freeman and Company, 1979.

[GJS76] M. R. Garey, D. S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1(2):237–267, 1976.

[GKR02] M. Grassl, A. Klappenecker, and M. Rötteler. Graphs, quadratic forms, and quantum codes. In *Proc. 2002 IEEE International Symposium on Information Theory*, page 45, 2002.

[Got96] D. Gottesman. Class of quantum error-correcting codes saturating the quantum Hamming bound. *Physical Review A*, 54:1862–1868, 1996.

[GR01] C. Godsil and G. Royle. *Algebraic Graph Theory*, volume 207 of *Graduate Texts in Mathematics*. Springer, 2001.

[Gru99] J. Gruska. *Quantum Computing*. McGraw-Hill, London, 1999.

[Haa97] U. Haagerup. Operator algebras and quantum field theory. In S. Doplichter, editor, *Orthogonal maximal abelian *-subalgebras of the n x n matrices and cyclic n-roots*, pages pp. 296–322. International Press, 1997.

[Had75] F. O. Hadlock. Finding a maximum cut of a planar graph in polynomial time. *SIAM Journal on Computing*, 4(3):221–225, 1975.

[Hae76] U. Haeberlen. *High resolution NMR in solids: selective averaging*. Academic Press, 1976.

[Hae95] W. H. Haemers. Interlacing eigenvalues and graphs. *Linear Algebra and its Applications*, pages 593–616, 1995.

[HJ85] R. A. Horn and C. R. Johnson. *Matrix anaysis*. Cambridge University Press, 1985.

[Hof70] A. J. Hoffman. On eigenvalues and colorings of graphs. In B. Harris, editor, *Graph Theory and its Applications*, pages 78–91. Academic Press, 1970.

[HSS99] A. S. Hedayat, N. J. A. Sloane, and J. Stufken. *Orthogonal Arrays*. Series in Statistics. Springer, 1999.

[HSTC00] T. Havel, S. Somaroo, C.-H. Tseng, and D. Cory. Principles and demonstration of quantum information processing by NMR spectroscopy. *Applicable Algebra in Engineering, Communication and Computing (AAECC)*, 10(4/5):339–374, 2000.

[Hup83] B. Huppert. *Endliche Gruppen*, volume I. Springer, 1983.

[Isa76] I. M. Isaacs. *Character theory of finite groups*. Pure and Applied Mathematics. Academic Press, 1976.

[Ist00] S. Istrail. Statistical Mechanics, Three-Dimensionality and NP-completeness: Universality of Intractability for the Partition Function of the Ising Model Across Non-Planar Lattices. to appear in *Proceeding of the 31st ACM Annual Symposium on the Theory of Computing* (STOC 2000), May 21-23, 2000, Portland, Oregon. ACM Press 2000, 2000.

[Jac74] N. Jacobson. *Basic Algebra I.* Freeman and Company, 1974.

[JB01] D. Janzing and Th. Beth. Complexity measure for continuous time quantum algorithms. *Physical Review A*, 64:022301, 2001. see also LANL e-print quant-ph/0009094.

[JB02] D. Janzing and Th. Beth. Distinguishing n Hamiltonians on $\mathbb{C}^n$ by a single measurement. *Physical Review A*, 65:022303, 2002. see also LANL e-print quant-ph/0103021.

[JB03] P. Janzing, D. Wocjan and Th. Beth. Identity check is QMA-complete. LANL e-print quant-ph/0305050, 2003.

[JK99] J. A. Jones and E. Knill. Efficient refocusing of one spin and two spin interactions for NMR quantum computation. *J. Magn. Resonance*, 141:323–325, 1999.

[JS72] V. Jurdjevic and H. J. Sussmann. Control systems on Lie Groups. *Journal of Differential Equations*, 12:313, 1972.

[Jur97] V. Jurdjevic. *Geometric control theory*, volume 52 of *Cambridge studies in advanced mathematics*. Cambridge University Press, Cambridge, 1997.

[JWB02a] D. Janzing, P. Wocjan, and Th. Beth. Bounds on the number of time steps for simulating arbitrary interaction graphs. LANL e-print quant-ph/0203061, submitted to Int. J. Found. Comp. Sci., 2002.

[JWB02b] D. Janzing, P. Wocjan, and Th. Beth. Complexity of decoupling and time-reversal for n spins with pair-interactions: Arrow of time in quantum control. *Physical Review A*, 66:042311, 2002. see also LANL e-print quant–ph/0109088.

[Kan96] G. Kant. Drawing Planar Graphs Using the Canonical Ordering. *Algorithmica*, 16:4–32, 1996.

[Kan98] B. Kane. A silicon-based nuclear spin quantum computer. *Nature*, 393:133–137, 1998.

[Kar72] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computation*, pages 85–103. Plenum Press, New York, 1972.

[KBG01] N. Khaneja, R. Brockett, and S. Glaser. Time Optimal Control in Spin Systems. *Physical Review A.*, 63:032308, 2001.

[KCL98] E. Knill, I. Chuang, and R. Laflamme. Efficient pure states for bulk quantum computation. *Physical Review A*, 57(3):3348–3363, 1998.

[Kha00] N. Khaneja. *Geometric Control in Classical and Quantum Systems.* PhD thesis, Harvard University, Cambridge, Massachusetts, 2000.

[KL02] W. M. Kaminsky and S. Lloyd. Scalable Architecture for Adiabatic Quantum Computing of NP-Hard Problems. Quantum Computing & Quantum Bits in Mesoscopic Systems (Kluwer Academic 2003), see also LANL e-print quant-ph/0211152, 2002.

[KLV00] E. Knill, R. Laflamme, and L. Viola. Theory of Quantum Error Correction for General Noise. *Physical Review Lett*, 84(11):2525–2528, 2000.

[Kni96a] E. Knill. Group representations, error basis and quantum codes. LANL e-print quant-ph/9608049, 1996.

[Kni96b] E. Knill. Non-binary Unitary Error Bases and Quantum Codes. LANL e-print quant-ph/9608048, 1996.

[Kni96c] E. Knill. Quantum randomness and nondeterminism. Technical Report LAUR-96-2186, Los Alamos National Laboratory, 1996.

[Kob99] N. Koblitz. *Algebraic Aspects of Cryptography*, volume 3 of *Algorithms and Computation in Mathematics.* Springer, 1999. 2nd printing.

[KR02] A. Klappenecker and M. Rötteler. Beyond Stabilizer Codes I: Nice Error Bases. *IEEE Trans. Inf. Th.*, 48(8):2392–2395, 2002. see also LANL e-print quant-ph/0010082.

[KR03] J. Kempe and O. Regev. 3-local hamiltonian is qma-complete. *Quantum Computation and Information*, 3(3):258–264, 2003.

[KSV02] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*, volume 27 of *Graduate Studies in Mathematics.* American Mathematical Society, 2002.

[KW01] M. Kaufmann and D. Wagner, editors. *Drawing Graphs: Method and Models*, volume 2025 of *Lecture Notes on Computer Science*. Springer, 2001.

[LCYY00] D. W. Leung, I. L. Chuang, F. Yamaguchi, and Y. Yamamoto. Efficient implementation of coupled logic gates for quantum computation. *Physical Review A*, 61(4):042310–1–7, 2000.

[LD98] D. Loss and D. DiVincenzo. Quantum computation with quantum dots. *Physical Review A*, 57:120–126, 1998.

[Leu02] D. W. Leung. Simulation and reversal of n-qubit Hamiltonians using Hadamard matrices. *J. Modern Optics*, 49(8):1199–1217, 2002.

[LG01] B. Luy and S. Glaser. Superposition of scalar and residual dipolar couplings: Analytical transfer functions for three spins 1/2 under cylindrical mixing conditions. *J. Magn. Reson.*, 148:169–181, 2001.

[Llo96] S. Lloyd. Universal quantum simulators. *Science*, 273:1073–1078, 1996.

[LV93] M. Li and P. Vitanyi, editors. *An introduction to Kolmogorov complexity and its applications*. Springer, 1993.

[Mac94] G. Maciel, editor. *Nuclear magnetic resonance in modern technology*, volume 447 of *NATO ASI XV*. Kluwer, 1994.

[MFM+00] R. Marx, A. Fahmy, J. Myers, W. Bermel, and S. Glaser. Approaching five-bit NMR quantum computing. *Physical Review A*, 62, 2000.

[MMK+95] C. Monroe, D. Meekhof, B. King, W. Itano, and D. Winlend. Demonstration of fundamental quantum logic gate. *Physical Review Letters*, 75(25):4714–4717, 1995.

[MO79] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[MSS00] Y. Makhlin, G. Schön, and A. Shnirman. Josephson junction based quantum computing. *Applicable Algebra in Engineering, Communication and Computing (AAECC)*, 10(4/5):375–382, 2000.

[MVL02] Ll. Masanes, G. Vidal, and J. I. Latorre. Time-optimal Hamiltonian simulation and gate synthesis using homogeneous local unitaries. *Quant. Inform. & Comp.*, 2(4):285–296, 2002.

[NBD+01] M. A. Nielsen, M. J. Bremner, J. L. Dodd, A. Childs, and C. Dawson. Universal simulation of Hamiltonian dynamics for qudits. LANL e-print quant-ph/0109064, 2001.

[NC00] M. A. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[NLR+99] H. Nägler, D. Leibfried, H. Rohde, J. Thalhammer, G. Eschner, F. Schmidt-Kaler, and R. Blatt. Laser addressing of individual ions in a linear ion trap. *Physical Review A*, 60(1):145–148, 1999.

[NV01] M. A. Nielsen and G. Vidal. Majorization and the interconversion of bipartite states. *Quantum Inform. & Comp.*, 1(1):76–93, 2001.

[OD72] G. I. Orlova and Y. G. Dorfman. Finding the maximum cut in a graph. *Engineering Cybernetics*, 10:502–506, 1972.

[Pap94] C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.

[RAS02] H. Rosé, T. Asselmeyer, and A. Schramm. Private communication (meeting of the BMBF project MARQUIS). 2002.

[RB00] R. Raussendorf and H. J. Briegel. Quantum computing via measurements only. LANL e-print quant-ph/0010033, 2000.

[Roc70] R. T. Rockafeller. *Convex Analysis*. Princeton University Press, Princeton, 1970.

[Röt01] M. Rötteler. Schnelle Signaltransformationen für Quantenrechnern. Dissertation, Universität Karlsruhe, 2001.

[RPW70] W. K. Rhim, A. Pines, and J. S. Waugh. Violation of the spin-temperature hypothesis. *Physical Review Lett.*, 25(4):218–220, 1970.

[Sak94] J. J. Sakurai. *Modern Quantum Mechanics*. Addison Wesley, 1994.

[Sch01] D. Schlingemann. Stabilizer codes can be realized as graph codes. LANL e-print quant-ph/0111080, 2001.

[Sch02] D. Schlingemann. Logical network implementation for cluster states and graph codes. LANL e-print quant-ph/0202007, 2002.

[Ser77] J. P. Serre. *Linear representations of finite groups*. Springer, 1977.

[Sli90] C. P. Slichter. *Principle of magnetic resonance*. Springer, 3rd edition, 1990.

[SM01] M. Stollsteimer and G. Mahler. Suppression of arbitrary internal coupling in a quantum register. *Physical Review A*, 64:052301, 2001.

[SOG+01] R. Somma, G. Oritz, J. Gubernatis, E. Knill, and R. Laflamme. Simulating physical phenomena by quantum networks. LANL e-print quant-ph/0108146, 2001.

[Ste96] A. Steane. Simple quantum error correcting codes. *Physical Review A*, 54:4741, 1996.

[SW00] D. Schlingemann and R. F. Werner. Quantum error-correcting codes associated with graphs. LANL e-print quant-ph/0012111, 2000.

[THL+95] Q. Turchette, C. Hood, W. Lange, H. Mabuchi, and H. Kimble. Measurement of conditional phase shifts for quantum logic. *Physical Review Lett.*, 75(25):4710–4713, 1995.

[VC01] G. Vidal and J. I. Cirac. Optimal simulation of nonlocal Hamiltonians using local operations and classical communication. LANL e-print quant-ph/0108076, 2001.

[VdB96] M. Vlaardingbroeck and J. den Boer. *Magnetic Resonance Imaging.* Springer, 1996.

[VKL99] L. Viola, E. Knill, and S. Lloyd. Dynamical decoupling of open quantum systems. *Physical Review Lett.*, 82:2417–2421, 1999.

[vNJ32] von Neumann. J. *Mathematische Grundlagen der Quantenmechanik*, volume XXXVIII of *Grundlehren der mathematischen Wissenschaften.* Springer, 1932.

[Wat00] J. Watrous. Succinct quantum proofs for properties of finite groups. In *Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, pages 537–546, 2000.

[WB02] P. Wocjan and Th. Beth. Equivalence of decoupling schmes. unpublished, 2002.

[WB03] P. Wocjan and Th. Beth. The 2-local Hamiltonian problem encompasses NP. LANL e-print quant-ph/0301087, 2003.

[Wil67] H. S. Wilf. The eigenvalues of a graph and its chromatic number. *J. London Math. Society*, 42:300–332, 1967.

[WJB02a] P. Wocjan, D. Janzing, and Th. Beth. Lower bound on the chromatic number by spectra of weighted adjacency matrices. LANL e-print cs.DM/0112023, submitted to Journal of Graph Theory, 2002.

[WJB02b] P. Wocjan, D. Janzing, and Th. Beth. Simulating arbitrary pair-interactions by a given Hamiltonian: graph-theoretical bounds on the time complexity. *Quantum Information & Computation*, 2(2):117, 2002. see also LANL e-print quant–ph/1010677.

[WJB03a] P. Wocjan, D. Janzing, and Th. Beth. Treating the Independent Set Problem by 2D Ising Interactions with Adiabatic Quantum Computing. LANL e-print quant-ph/0302027, 2003.

[WJB03b] P. Wocjan, D. Janzing, and Th. Beth. Two QCMA-complete problems. LANL e-print quant-ph/0305090, 2003.

[WRJB02a] P. Wocjan, M. Rötteler, D. Janzing, and Th. Beth. Simulating Hamiltonians in Quantum Networks: Efficient Schemes and Complexity Bounds. *Physical Review A*, 65:042309, 2002. see also LANL e-print quant-ph/0109088.

[WRJB02b] P. Wocjan, M. Rötteler, D. Janzing, and Th. Beth. Universal simulation of Hamiltonians using a finite set of control operations. *Quantum Information & Computation*, 2(2):133, 2002. see also LANL e-print quant-ph/0109063.

[WZ93] W. D. Wallis and Guo-Hui Zhang. On the partition and coloring of a graph by cliques. *Discrete Math.*, 120(1–3):191–203, 1993.

[Yan78] M. Yannakakis. Node- and edge-deletion NP-complete problems. In 10*th Ann. ACM Symp. on Theory of Computing*, pages 253–264, New York, 1978. Association for Computing Machinery.

[Yao93] A. C.-C. Yao. Quantum circuit complexity. In 34*th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 352–361. IEEE Computer Society Press, 1993.

[Zan99] P. Zanardi. Symmetrizing evolutions. *Physical Letters A*, 258:77, 1999.

[Zan01] P. Zanardi. Stabilizing quantum information. *Physical Review A*, 63:012301, 2001.

[ZR97] P. Zanardi and M. Rasetti. Noiseless quantum codes. *Physical Review Letters*, 79:3306–3309, 1997.

[ZR98] P. Zanardi and F. Rossi. Quantum information in semiconductors: noiseless encoding in a quantum dot array. *Physical Review Lett*, 81:4752–4755, 1998.

[Zur89] W. Zurek. Algorithmic randomness and physical entropy. *Physical Review A*, 40:4731–4751, 1989.

[Zur90] W. Zurek, editor. *Complexity, entropy and the physics of information*. Addison–Wesley, 1990.

[Zwi95] D. Zwillinger, editor. *CRC Standard Mathematical Tables and Formulae*. CRC Press, Boca Raton, FL, 1995.

Index

Lebenslauf

19.05.1974	geboren in Warschau
1981–1984	Grundschule in Warschau
1984–1987	Max-Planck-Gymnasium in München
1987–1994	Gymnasium in Karlsbad (Abitur Juni 1994)
1994–1999	Studium der Informatik an der Universität Karlsruhe (TH)
	– 1996 Vordiplom
	– 1999 Diplomarbeit am “Institut National de Recherche en Informatique et en Automatique”, Nancy
	– 1999 Diplom in Informatik
seit März 99	beschäftig als wissenschaftlicher Mitarbeiter am Institut für Algorithmen und Kognitive Systeme, Universität Karlsruhe (TH)

Publications

Articles in journals:

1. D. Janzing, P. Wocjan, and Th. Beth, "Bounds on the number of time steps for simulating arbitrary interaction graphs", 2002, LANL e-print quant-ph/0203061, to appear in *International Journal of Foundations of Computer Science*.

2. P. Wocjan, D. Janzing, and Th. Beth, "Simulating arbitrary pair-interactions by a given Hamiltonian: graph-theoretical bounds on the time complexity", *Quantum Information & Computation*, 2(2):117, 2002. see also LANL e-print quant–ph/0106077.

3. P. Wocjan, M. Rötteler, D. Janzing, and Th. Beth, "Universal simulation of Hamiltonians using a finite set of control operations", *Quantum Information & Computation*, 2(2):133, 2002. see also LANL e-print quant-ph/0109063.

4. P. Wocjan, M. Rötteler, D. Janzing, and Th. Beth, "Simulating Hamiltonians in Quantum Networks: Efficient Schemes and Complexity Bounds", *Physical Review A*, 65:042309, 2002. selected for the April 8, 2002 issue of the Virtual Journal of Nanoscale Science & Technology, selected for the April 2002 issue of the Virtual Journal of Quantum Information, see also LANL e-print quant-ph/0109088.

5. D. Janzing, P. Wocjan, and Th. Beth, "Complexity of decoupling and time-reversal for n spins with pair-interactions: Arrow of time in quantum control", *Physical Review A*, 66:042311, 2002. selected for the October 28, 2002 issue of the Virtual Journal of Nanoscale Science & Technology, selected for the November, 2002 issus of the Virtual Journal of Quantum Information, see also LANL e-print quant–ph/0109088v2.

6. D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth, "Thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the Second Law", *International Journal of Theoretical Physics*, 39(12):2217–2753, 2000. see also LANL e-print quant-ph/0002048.

Contributions to books:

7. Th. Beth, M. Grassl, D. Janzing, M. Rötteler, P. Wocjan, and R. Zeier, "Algorithms for quantum systems - quantum algorithms", in G. Leuchs and Th. Beth (eds.), Quantum Information Processing, Wiley-VCH, Berlin, 2003.

Articles in conference proceedings:

8. P. Wocjan, D. E. Lazic, and Th. Beth, "Performances of binary block codes used on classical-quantum channels", Proceedings of the Fifth International Conference on "Quantum Communication, Measurement & Computing", Capri, Italy, 2000, p. 43–46.

Preprints/submitted articles:

9. P. Wocjan, D. Janzing, Th. Decker, and Th. Beth, "Measuring 4-local n-qubit observables could probabilistically solve PSPACE", LANL e-print quant-ph/0308011.

10. P. Wocjan, D. Janzing, and Th. Beth, "Two QCMA-complete problems", LANL e-print quant-ph/0305090.

11. D. Janzing, P. Wocjan, and Th. Beth, "Identity Check is QMA-complete", LANL e-print quant-ph/0305050.

12. D. Janzing, P. Wocjan, and Th. Beth, "Cooling and Low Energy State Preparation for 3-local Hamiltonians are FQMA-complete", LANL e-print quant-ph/0303186.

13. P. Wocjan, D. Janzing, and Th. Beth, "Treating the Independent Set Problem by 2D Ising Interactions with Adiabatic Quantum Computing", LANL e-print quant-ph/0302027.

14. P. Wocjan and Th. Beth, "The 2-local Hamiltonian problem encompasses NP", LANL e-print quant-ph/0301087.

15. P. Wocjan, D. Janzing, and Th. Beth, "Required sample size for learning sparse Bayesian networks with many variables", LANL e-print cs.LG/0204052, 2002.

16. P. Wocjan, D. Janzing, and Th. Beth, "Lower Bound on the Chromatic Number by Spectra of Weighted Adjacency Matrices", LANL e-print cs.DM/0112023, 2001.

Technical reports:

17. M. Eck, P. Wocjan, and Th. Beth, "Entwurf und Implementierung eines Simulators für Quantenschaltkreise", E.I.S.S.-Report 03/2000, European Institute for System Security, University of Karlsruhe, 2001.

Diplomarbeit (diploma thesis) and Studienarbeit:

- P. Wocjan, "Entwurf diffraktiver Strahlteiler mit der Methode der finiten Elemente", Studienarbeit, University Karlsruhe, 1997.
- P. Wocjan, "Der Brill-Noether-Algorithmus: Konstruktion geometrischer Goppa-Codes und absolute Faktorisierung", Diploma thesis, University Karlsruhe, 1999.