THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

The quantum approximate optimization algorithm: optimization problems and implementations

PONTUS VIKSTÅL

Department of Microtechnology and Nanoscience (MC2)
Applied Quantum Physics Laboratory
Chalmers University of Technology
Göteborg, Sweden, 2023

The quantum approximate optimization algorithm:
optimization problems and implementations
PONTUS VIKSTÅL
ISBN 978-91-7905-893-7

© PONTUS VIKSTÅL, 2023

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5359
ISSN 0346-718X

Applied Quantum Physics Laboratory
Department of Microtechnology and Nanoscience (MC2)
Chalmers University of Technology
SE–412 96 Göteborg
Sweden
Telephone +46 (0)31-772 1000

Cover: Schematic illustration of the compilation steps of a quantum algorithm from computer code to the physical execution on transmon qubits.

Chalmers Digitaltryck
Göteborg, Sweden, 2023

The quantum approximate optimization algorithm:
optimization problems and implementations
PONTUS VIKSTÅL
Department of Microtechnology and Nanoscience (MC2)
Applied Quantum Physics Laboratory
Chalmers University of Technology

Abstract

This thesis explores the Quantum Approximate Optimization Algorithm (QAOA), a hybrid classical-quantum algorithm designed to solve combinatorial optimization problems. The goal of this algorithm is to iteratively optimize a variational state to approximate the ground state of a cost Hamiltonian that encodes a combinatorial optimization problem. The focus of this thesis is the application of QAOA to the Exact Cover problem, an abstraction of the Tail Assignment problem – a problem omnipresent in aviation.

This thesis also includes a demonstration of the practical implementation of QAOA on a superconducting quantum computer, demonstrating empirical proof of QAOA's functionality. It also investigates running QAOA using noise-biased qubits, namely cat qubits, which exhibit resilience to certain types of errors.

This thesis also explores novel multi-qubit gates obtained from the simultaneous application of two controlled-Z gates on current quantum hardware, leading to the efficient creation of large entangled states.

Lastly, the thesis delves into virtual distillation, an error-mitigation protocol, and assesses its performance under various types of errors.

Overall, this thesis validates the promise of QAOA in solving real-world optimization problems while also offering insights into error mitigation.

Keywords: Quantum approximate optimization algorithm, quantum computing, error mitigation, cat qubits

Acknowledgments

Reflecting on my five-year journey, I realize that few experiences offer as much personal growth as pursuing a Ph.D. The journey of becoming an independent researcher is a truly transformative one. I believe and hope others agree that I have managed to achieve this, acquiring the confidence and skills needed to carry out future independent research.

Luck has played a significant part in this journey, and I would be remiss not to acknowledge it. I'm lucky to have had an exceptional high school physics teacher who ignited my interest in this field, and I count myself lucky to have been accepted for this Ph.D. by my supervisor Giulia Ferrini, whose trust and guidance I deeply appreciate. But let's not forget that all of this wouldn't have been possible without a healthy dose of hard work and determination. Hence, as the saying goes, "luck is when preparation meets opportunity".

I'm grateful for the incredible colleagues I've had the privilege of working with, many of whom have become lifelong friends. I owe a massive thanks to Timo Hillmann for proofreading this thesis. Working alongside inspiring individuals such as you and Ingrid has been an absolute pleasure. Finally, I extend my heartfelt gratitude to my friends and family, whose unwavering support and encouragement have been a constant source of strength throughout this journey.

Pontus Vikstål, Göteborg, August 2023

Preface

Pause for a second, and imagine, if you will, all the atoms in the observable Universe – a mind-boggling 10^{80} of them. That is a number so massive that it is too hard to grasp, but now picture a device that could perform computations on an equivalent number of states, all at the same time! Surely, such a device must be as big as the Universe itself? Believe it or not, it would actually not be the size of galaxies, but it would instead fit right in the palm of your hand, and we are only a few years, maybe a decade away from realizing it. What I have in mind is a quantum processor, where a quantum computer with only 300 quantum bits (qubits) could theoretically tackle such an extensive range of states, which speaks to the extraordinary potential of quantum computing. This concept had its origins in 1925 when Erwin Schrödinger wrote down his famous Schrödinger equation that accurately describes the "ticking clock" of any quantum system. This laid the foundation for the first quantum revolution with the development of quantum mechanics, and what followed was the second quantum revolution with remarkable technological inventions such as transistors, MRI, lasers, and superconductors. Today we are at the beginning of the third quantum revolution, which involves quantum computing, quantum communication, quantum sensing, and quantum cryptography.

The power of quantum computing lies in exponential growth: each added qubit effectively doubles the computational power. In simple terms, to double the computational power of a classical computer, you need to double the number of transistors on your chip. In contrast, with a quantum computer, just adding a single qubit doubles the computational power, which makes it already difficult to simulate a quantum computer with 50 qubits using classical supercomputers.

The potential applications of quantum computing are vast and diverse, ranging from cryptography to drug discovery, optimization problems, and beyond. Certainly, while the potential of quantum computing is immense, it does not come without its fair share of skepticism and doubt. There are several key issues that skeptics often highlight. Firstly, there is the technological challenge of building a reliable and scalable quantum computer, and then there is a question about the practical applications of quantum computing. While theoretically,

quantum computers could solve certain problems faster than classical computers, classical computers will likely remain more efficient for many tasks. It is unclear how many problems fall into the former category and whether quantum computers will be able to solve them in practice. But from a more philosophical point of view, it is fascinating to think of the Universe as a giant computer playing out its simulation according to the rules of quantum mechanics. If we can build a computer that can mimic these rules, we would not just be opening up a new frontier in computing but also getting access to the computational inner workings of the Universe itself.

However, reaching the full potential of quantum computing is a bit like balancing a pen on its tip. It is a delicate task, susceptible to the smallest of disturbances. Qubits are no different – any loss of information to the environment makes them decohere and lose their "quantumness". Yet, there exists a silver lining – quantum error correction, but it demands a significant overhead of qubits. Therefore, to fully harness the power of quantum computing, we will likely need millions of qubits.

Nevertheless, on our journey to this goal, we encounter Noisy Intermediate Scale Quantum (NISQ) devices with a few hundred non-error corrected qubits. This poses the question: can we do something useful with these devices? In this thesis, we aim to delve into the current state of quantum algorithms for NISQ devices and try to shed light on the potential of quantum computing to tackle complex computational problems.

Publications

A **Applying the Quantum Approximate Optimization Algorithm to the Tail-Assignment Problem**
Pontus Vikstål, Mattias Grönkvist, Marika Svensson, Martin Andersson, Göran Johansson, and Giulia Ferrini
Phys. Rev. Applied **14**, 034009 (2020)

B **Improved Success Probability with Greater Circuit Depth for the Quantum Approximate Optimization Algorithm**
Andreas Bengtsson, Pontus Vikstål, Christopher Warren, Marika Svensson, Xiu Gu, Anton Frisk Kockum, Philip Krantz, Christian Križan, Daryoush Shiri, Ida-Maria Svensson, Giovanna Tancredi, Göran Johansson, Per Delsing, Giulia Ferrini, and Jonas Bylander
Phys. Rev. Applied **14**, 034010 (2020)

C **Fast Multiqubit Gates through Simultaneous Two-Qubit Gates**
Xiu Gu, Jorge Fernández-Pendás, Pontus Vikstål, Tahereh Abad, Christopher Warren, Andreas Bengtsson, Giovanna Tancredi, Vitaly Shumeiko, Jonas Bylander, Göran Johansson, and Anton Frisk Kockum
PRX Quantum **2**, 040348 (2021)

D **Study of noise in virtual distillation circuits for quantum error mitigation**
Pontus Vikstål, Giulia Ferrini, and Shruti Puri
arXiv:2210.15317 [quant-ph] (2022)
Submitted to Quantum

E **Quantum Approximate Optimization Algorithm with Cat Qubits**
Pontus Vikstål, Laura García-Álvarez, Shruti Puri, and Giulia Ferrini
arXiv:2305.05556 [quant-ph] (2023)
Submitted to Physical Review A

Contents

Abstract **i**

Acknowledgments **iii**

Preface **v**

Publications **vii**

1 Introduction **1**
1.1 Outline of the thesis 2

2 Quantum computing **5**
2.1 Qubits 5
2.1.1 Continuous- versus discrete-variable qubits 6
2.2 Quantum gates 7
2.2.1 Multi-qubit gates 9
2.3 Open quantum systems 13
2.3.1 Closed quantum systems 13
2.3.2 Lindblad master equation 13
2.3.3 Kraus operator formalism 14

3 The quantum approximate optimization algorithm **17**
3.1 From the quantum adiabatic algorithm to the QAOA 18

4 Application of QAOA to a real-world optimization problem **23**
4.1 The tail-assignment problem 23

5 Implementation of QAOA **29**
5.1 Experimental realization of QAOA in a superconducting device . 30
5.2 Study of QAOA implementation with cat qubits 31
5.2.1 Running QAOA with cat qubits 34

6 Error mitigation **37**
6.1 Virtual distillation 39
6.1.1 Noise in the virtual distillation circuit 42

7 Conclusion and outlook **45**
7.1 Summary 45
7.2 Outlook 46

Bibliography **49**

Included papers **56**

Chapter 1

Introduction

The past fifty years have seen an extraordinary reduction in computer hardware size, mainly due to technological advancements that facilitated the miniaturization of transistors within our devices. However, as these transistors approach the size of atoms, they begin experiencing quantum effects that can interfere with their functioning [1, 2]. Researchers are therefore trying to leverage these quantum effects to turn this challenge into an advantage by proposing a new computational paradigm known as "quantum computing".

The origins of quantum computing date back to the early 1980s when the physicists Paul Benioff, Yuri Manin, and Richard Feynman independently and concurrently proposed the idea of a quantum computer [3–5]. They recognized that classical computers require exponentially scaling resources in the number of transistors to accurately simulate quantum systems, which led them to propose the idea of a quantum mechanical system that could simulate other quantum mechanical systems. Building on these ideas, researcher David Deutsch, in 1985, formulated the mathematical model of a quantum Turing machine [6], which paved the way for the quantum circuit model which he later developed in 1989 [7]. The field then took a massive leap forward in 1997 with Peter Shor's discovery of a quantum algorithm capable of solving prime factorization exponentially faster than the best-known classical algorithm [8]. The difficulty of finding the prime factors of large numbers for classical computers is used in public-key cryptography, such as the RSA [9]. As such, with a large enough quantum computer, the public-key cryptosystems could easily be hacked.

Despite these advancements, quantum computing faces several challenges and limitations. One of the most significant issues is the sensitivity to noise, which affects the lifetime of quantum bits or qubits. This noise imposes constraints on the size and complexity of quantum circuits, limiting the scope and applicability of quantum algorithms. Addressing these challenges requires developing sophisticated error correction techniques to achieve fault-tolerant quantum computing architectures [10–13], which typically requires a large overhead

of qubits [14]. For example, estimates on the requirements of running Shor's algorithm for factoring cryptographically hard numbers have shown to require millions of qubits with error correction [15].

Nevertheless, as small-size, Noisy Intermediate Scale Quantum (NISQ) devices become available [16], academics and companies want to explore their usefulness. In fact, experimental demonstrations have already shown the ability to run a quantum algorithm on a NISQ device that would take classical supercomputers thousands of years to replicate the results [17]. However, this experiment has since been disputed with the emergence of clever classical simulation techniques that can, in fact, replicate the results of these quantum computing experiments in only a few hours [18, 19]. Despite these disputes, research continues, with the hope that NISQ devices could demonstrate significant power and usefulness in the near future. Nevertheless, a substantial challenge lies ahead: the quest to discover quantum algorithms capable of solving real-world problems faster than any existing classical algorithm. One promising candidate for this is the Quantum Approximate Optimization Algorithm (QAOA) [20], which aims to solve combinatorial optimization problems. These are types of problems that frequently arise in industry, such as aviation [21], portfolio optimization [22], and vehicle routing [23].

The QAOA, which is the main topic of this thesis, is classified as a heuristic hybrid quantum-classical algorithm. The hybrid character comes from the fact that a quantum computer prepares some N-qubit state that is measured, and the measurement results are then processed by a classical computer that tells the quantum computer how to slightly change how the N-qubit state is prepared, and because no theoretical proof of speed up exists for this algorithm, it is heuristic. Although, for a few problems, such as the MaxCut, lower bounds exist on the performance guarantee [20], which therefore makes it an approximate algorithm. However, in most cases, one has to simply run the algorithm and see what happens.

In this thesis, we set out to explore the QAOA, its application to combinatorial optimization problems, its implementation on a quantum computer, and how error mitigation can reduce noise in the estimation of expectation values.

1.1 Outline of the thesis

In the first chapter of the thesis, Chapter 2, we will give an introduction to quantum computing, the fundamental building block, qubits, and how to operate on them. Next, in Chapter 3, we delve into QAOA, which is the main topic of this thesis. Afterward, in Chapter 4, we show how to use QAOA to solve a real-world optimization problem, thus presenting a concrete example of quantum computing's practical application. Then, in Chapter 5, we show how to implement and run QAOA on an actual superconducting qubit device and discuss the possible role of "cat qubits" in the future for making QAOA more error resilient. In Chapter 6, we study a quantum error-mitigation technique

called “virtual distillation”. This technique can reduce errors in estimating expectation values, and we assess its ability to cope with different types of errors that occur during the protocol. Concluding the thesis in Chapter 7, we present a summary of the central findings, review the appended papers, and give an outlook for potential future research trajectories based on the insights gathered in this thesis.

Chapter 2

Quantum computing

In this chapter, we will be giving an introduction to qubits and how to manipulate them using quantum gates. This is followed by how to model an open quantum system interacting with its environment.

2.1 Qubits

Just as a classical computer uses bits that are either 0 or 1 and that can physically be represented as a low and a high voltage in a wire, a quantum computer uses qubits $|0\rangle$ or $|1\rangle$ (often called computational basis states). A qubit can, for example, be physically manifested as two states of an atom (e.g., ground state $|0\rangle$ and excited state $|1\rangle$) or the number of Cooper pairs on the superconducting island formed by a Josephson junction and a capacitor [24, 25], see Figure 2.1. Qubits are described using "ket" notation, which uses angled brackets and looks like $|0\rangle$. A typical choice to represent the $|0\rangle$ and $|1\rangle$ states are as column vectors:

$$|0\rangle \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.1}$$

Qubits are represented as column vectors instead of single numbers because quantum systems, like atoms, can exist in a linear combination or "superposition" of both states at once. This superposition for a general qubit state can be described as

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle \equiv \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{2.2}$$

In this equation, α and β are two complex numbers with a sum of absolute squares equal to 1 to ensure proper normalization of the qubit, $|\alpha|^2 + |\beta|^2 = 1$. When a qubit is measured in the computational basis, there is a $|\alpha|^2$ probability of finding it in the $|0\rangle$ state, and $|\beta|^2$ probability of finding it in the $|1\rangle$ state.

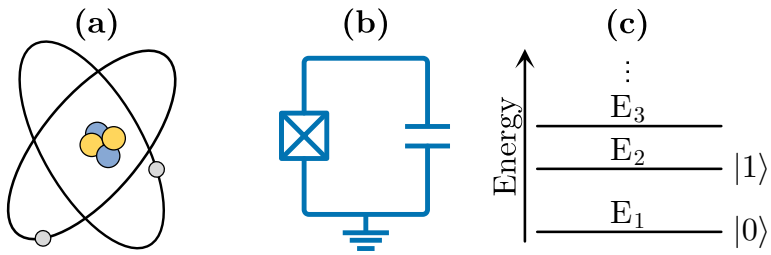

Figure 2.1: **(a)** Representation of an atom. **(b)** Circuit diagram of the transmon qubit, commonly referred to as an "artificial atom". It consists of a Josephson junction (box with a cross) which acts as a non-linear inductor parallel to a capacitor. **(c)** Both systems exhibit an energy spectrum with non-equidistant energy levels. This characteristic can be used to isolate the two lowest energy levels and use them to encode a qubit.

The act of measuring the qubit thus affects its state, causing it to be either $|0\rangle$ or $|1\rangle$ after the measurement.

This idea can be extended to multiple qubits. For example, consider a system of N qubits; its most general state can be expressed as a superposition of all possible binary strings of length N, denoted by $\{0,1\}^N$. Each of these binary strings is associated with a complex amplitude, α_z, resulting in the following representation:

$$|\psi\rangle = \sum_{z\in\{0,1\}^N} \alpha_z |z\rangle, \quad \text{with} \quad \sum_{z\in\{0,1\}^N} |\alpha_z|^2 = 1. \tag{2.3}$$

The critical requirement here is that the sum of the squares of the absolute values of these complex amplitudes must equal 1, ensuring the total probability remains conserved in this larger system. This extension underlines the scalability of quantum computing, demonstrating that the memory available with every added qubit grows exponentially.

Finally, a qubit's quantum state can be visualized as a unit vector inside a unit sphere called the Bloch sphere, as shown in Figure 2.2. The state of a single qubit can be described in terms of polar and azimuthal angles (ϕ, θ) on the Bloch sphere as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle, \quad 0 \le \theta \le \pi, \quad 0 \le \phi \le 2\pi, \tag{2.4}$$

2.1.1 Continuous- versus discrete-variable qubits

Quantum computation is not limited to discrete-variable qubits, meaning that the qubit is encoded in a two-level system. It is also possible to encode information into systems described by continuous variables, like position and momentum or the amplitude of the electromagnetic field [26]. The difference between a discrete- and continuous-variable quantum system is that the discrete-variable quantum system is described by a finite Hilbert space, while

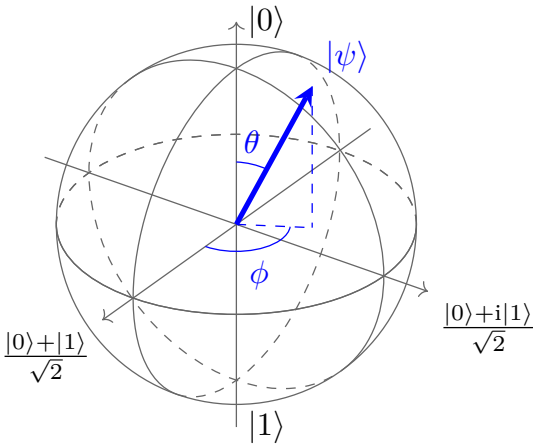

Figure 2.2: Bloch sphere showing the computational basis states $|0\rangle$ and $|1\rangle$, and a general qubit state $|\psi\rangle = \cos\theta/2\,|0\rangle + e^{i\phi}\sin\theta/2\,|1\rangle$.

a continuous-variable quantum system is described by an infinite dimensional Hilbert space [27]. As an example, a discrete-variable qubit is a qubit encoded in a two-dimensional Hilbert space, such as the two lowest energy levels of a transmon, and therefore, the dimension of the qubit is the same as the dimension of the physical state as shown in Figure 2.3a. On the other hand, a continuous-variable qubit, such as the cat qubit, is a qubit encoded in a two-dimensional subspace of an infinite-dimensional Hilbert space, as shown in Figure 2.3b. This redundancy in the physical encoding, as we will explore in Section 5.2, offers protection against certain types of errors, such as bit-flip or dephasing-errors.

To be more precise, a cat state is a superposition of two coherent states with opposite phases. Coherent states $|\alpha\rangle$ are states of the harmonic oscillator that are closest to classical states in that their expectation values correspond to the classical equation of motion of the harmonic oscillator. A *coherent state* is defined as an eigenstate of the annihilation operator a with eigenvalue $\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle$, where α is a complex number. Mathematically the cat state is defined as

$$\left|C_\alpha^\pm\right\rangle = N_\pm(|\alpha\rangle \pm |-\alpha\rangle) \quad \text{with} \quad N_\pm = \sqrt{2(1 \pm e^{-2|\alpha^2|})}, \tag{2.5}$$

where the plus or minus sign corresponds to the even and odd cat states, respectively. When written in the photon number basis, the even cat state is a superposition of even Fock states, while the odd cat state is a superposition of odd number Fock states. The even and odd cat states are thus orthogonal to one another and can therefore be used to encode a qubit.

2.2 Quantum gates

Analogous to classical computers, which employ elementary gates such as AND or NOT to carry out operations on bits, quantum computers use quantum gates

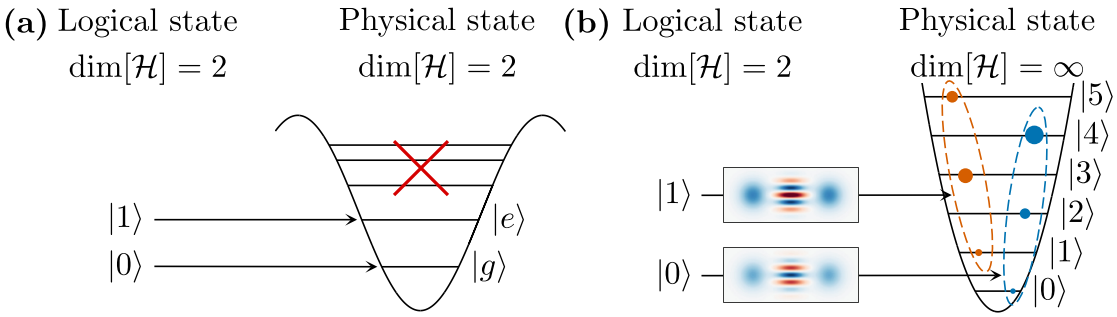

Figure 2.3: **(a)** A discrete-variable qubit. A qubit with a Hilbert space of dimension 2 is encoded in a physical state of the same dimension. **(b)** A continuos variable qubit. Two cat-states seen as their phase space representation (Wigner functions). These cat qubits are orthogonal to one another and can therefore be used to encode a qubit. Each cat qubit lives in an infinite dimensional Hilbert space. The size of the dots of the harmonic potential corresponds to the photon number occupation of each cat qubit.

to change the probability amplitudes of one or several qubits. The Schrödinger equation fundamentally dictates the change of these probability amplitudes in a quantum system over time. This linear differential equation explicates the time evolution of a closed quantum system,

$$\mathrm{i}\hbar \frac{\mathrm{d}\,|\psi(t)\rangle}{\mathrm{d}t} = \hat{H}\,|\psi(t)\rangle\,. \tag{2.6}$$

In this equation, $\hbar$ is the reduced Planck's constant. By opting for suitable units, we can assign $\hbar$ a value of 1, thereby ignoring it in this thesis. The Hamiltonian is represented by $\hat{H}$, whose eigenvalues correspond to the allowed energies of the system. For a N qubit system the Hamiltonian can be represented as a $2^N \times 2^N$ matrix with 2^N eigenstates $|\psi_i\rangle$, and 2^N energy eigenvalues such that $\hat{H}\,|\psi_i\rangle = E_i\,|\psi_i\rangle$. Thus there exists a basis where the Hamiltonian is diagonal

$$\hat{H} = \sum_{i=1}^{2^N} E_i\,|\psi_i\rangle\langle\psi_i|\,. \tag{2.7}$$

When the Hamiltonian is known, we can find a solution to the Schrödinger equation. In the case of a Hamiltonian that is time-independent, the solution is simply given in terms of the evolution operator

$$\hat{U}(t) = \mathrm{e}^{-\mathrm{i}\hat{H}t}. \tag{2.8}$$

If we know the initial state of the system, $|\psi(0)\rangle$, all future states can be computed by applying the time evolution operator $\hat{U}(t)$ on it, represented as $|\psi(t)\rangle = \hat{U}(t)\,|\psi(0)\rangle$. Important types of Hamiltonians are the Pauli matrices:

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.9}$$

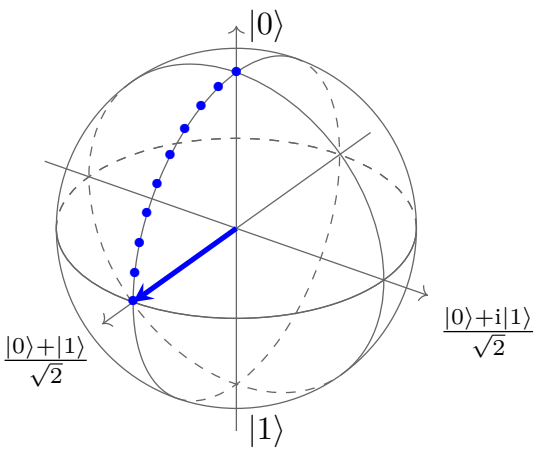

Figure 2.4: Bloch sphere showing the action of the quantum logic operation $\mathrm{R}_y(\pi/2)$ applied to the $|0\rangle$ state.

A quantum state evolving under one of the Pauli matrices, i.e., $\hat{H} = \hat{\sigma}_k$, $k \in \{x, y, z\}$, with $t = \theta/2$, implements the rotational quantum gate

$$\mathrm{R}_k(\theta) \equiv \mathrm{e}^{-\mathrm{i}\frac{\theta}{2}\hat{\sigma}_k}. \tag{2.10}$$

Such gates are labeled as rotational quantum gates since they correspond to rotations around one of the three primary axes of the Bloch sphere. For instance, applying the quantum gate $\mathrm{R}_y(\pi/2)$ corresponds to an anticlockwise rotation of $\pi/2$ around the y-axis, which is depicted for the initial state $|0\rangle$ in Figure 2.4.

The Hadamard gate, represented by H (not to be confused with the Hamiltonian $\hat{H}$), is another significant single-qubit gate. If the input state of the qubit is $|0\rangle$, it will yield the output state $(|0\rangle + |1\rangle)/\sqrt{2}$, and if the input is $|1\rangle$, it will result in $(|0\rangle - |1\rangle)/\sqrt{2}$. The Hadamard gate can be expressed in matrix representation as

$$\mathrm{H}\,|\psi\rangle \Longleftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\alpha+\beta\\ \alpha-\beta\end{pmatrix}, \tag{2.11}$$

or illustrated as a quantum circuit as

$$|0\rangle \;\text{—}\boxed{\mathrm{H}}\text{—}\; \frac{|0\rangle + |1\rangle}{\sqrt{2}} \tag{2.12}$$

The Hadamard gate equates (up to a global phase) to a $\pi/2$ rotation around the y-axis, succeeded by a π rotation around the x-axis, $\mathrm{R}_x(\pi)\mathrm{R}_y(\pi/2) = -\mathrm{i}\mathrm{H}$.

2.2.1 Multi-qubit gates

Apart from single-qubit gates, there exist multi-qubit gates, the simplest among which are two-qubit gates. An example of a common two-qubit gate is the

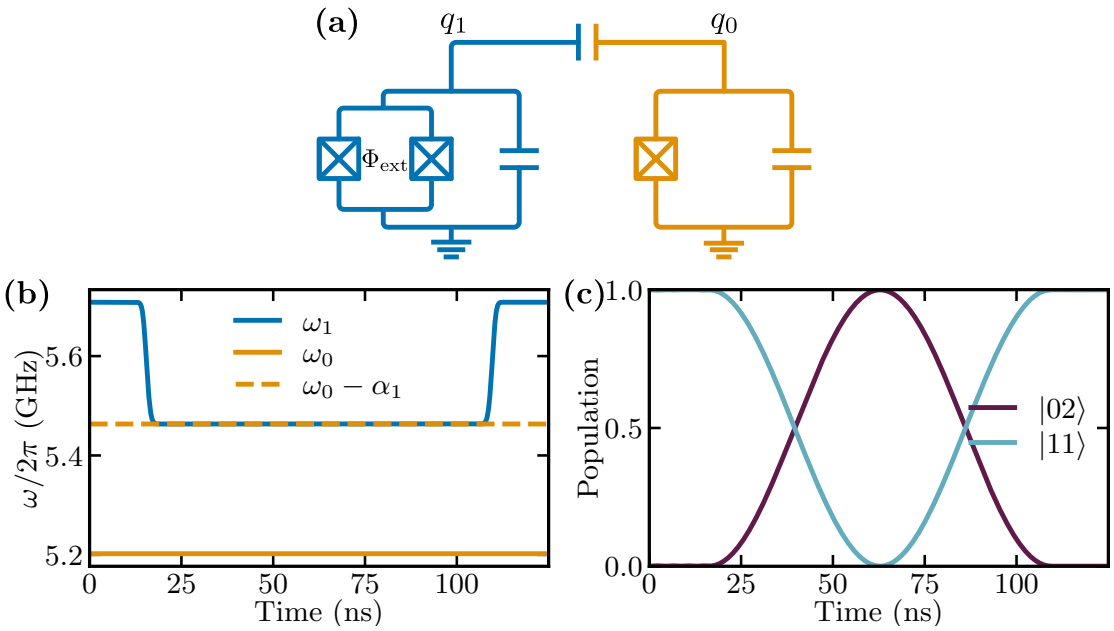

Figure 2.5: **(a)** Two transmon qubits capacitively coupled together. For q_1, the Josephson junction has been replaced with a SQUID which allows the frequency of q_1 to be tuned in time by threading an external flux Φ_{ext} through the loop. **(b)** Frequency of the qubits. **(c)** Population transfer. Here the qubits are ordered as $|q_0 q_1\rangle$.

controlled-Z gate (CZ-gate). This quantum gate adds a minus sign to the $|11\rangle$-state. In mathematical terms $\text{CZ}\,|11\rangle = -\,|11\rangle$, where $|11\rangle$ is compact notation for $|1\rangle \otimes |1\rangle$, and $\otimes$ is the tensor product. In the case of a transmon qubit, such a gate can be implemented using flux-tunable transmon qubits [28–30]. For such a qubit, the Josephson junction of the transmon is replaced by two Josephson junctions in parallel, also known as a Superconducting Quantum Interference Device (SQUID), see leftmost part of Figure 2.5a. The qubit frequency can then be tuned in situ by threading an external magnetic flux Φ_{ext} through the SQUID. From circuit quantization, we can expand the cosine potential of the Josephson junction to the fourth order. With this, the Hamiltonian for the flux tunable transmon can be written in the eigenbases of the quantum harmonic oscillator as [24]

$$\hat{H}_{\text{transmon}} = \omega(\Phi_{\text{ext}})\hat{a}^\dagger \hat{a} + \frac{\alpha}{2}\hat{a}^{\dagger 2}\hat{a}^2, \tag{2.13}$$

where $\omega(\Phi_{\text{ext}})$ is the frequency of the qubit and can be controlled by the external flux, α is the Kerr non-linearity, and $\hat{a}$ and $\hat{a}^\dagger$ are the familiar annihilation and creation operators. The first term in the Hamiltonian corresponds to the standard quantum harmonic oscillator. However, the second term, known as the Kerr term, accounts for the non-linearity introduced by the Josephson junction and provides the necessary anharmonicity to isolate two levels and form a qubit.

The experimental implementation of the CZ-gate involves taking one flux-tunable transmon and one fixed frequency transmon and coupling them to-

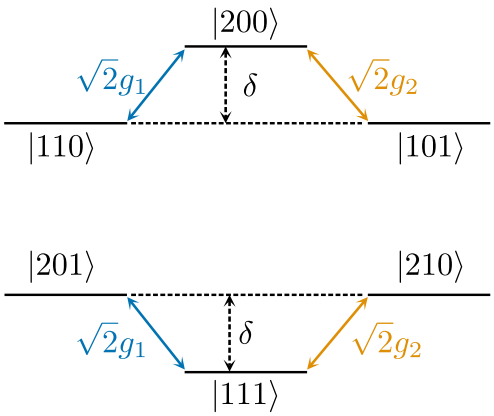

Figure 2.6: The Λ- and V-type systems. g_1 and g_2 are the coupling strengths between q_0, q_1 and q_0, q_2 respectively, and δ is the detuning from resonance.

gether[1], see Figure 2.5a. Next, to achieve the desired -1 phase on the $|11\rangle$ state, the $|11\rangle$ state is brought into resonance with the $|02\rangle$ state, which lies outside of the computational subspace, see Figure 2.5b. When these states become degenerate, oscillations will occur between the state $|11\rangle$ and $|02\rangle$ and vice versa, see Figure 2.5c. During a complete oscillation cycle, the $|11\rangle$ state will acquire a geometrical phase factor of -1, which is just what one wants for the CZ-gate.

Nevertheless, even though single-qubit gates and a two-qubit are enough to facilitate universal quantum computation[2] [1, 31], three-qubit gates are of importance since they can create large-scale entanglement [32] with shorter circuit depth compared to using two-qubit gates, which can facilitate the implementation of more complex quantum algorithms. For example, decomposing a three-qubit gate, like the Fredkin or Toffoli gate, requires at least five two-qubit gates [33, 34]. Therefore it is of practical advantage to have access to three-qubit gates since they can reduce gate synthesis.

In paper Paper C, we show how a three-qubit gate can be synthesized by simultaneously applying two two-qubit gates. Using the setup in Figure 2.7a, the simultaneous application of two CZ-gates, one between qubit q_0 and q_1 and one between q_0 and q_2 will activate the transitions $|110\rangle \leftrightarrow |200\rangle \leftrightarrow |101\rangle$ and $|201\rangle \leftrightarrow |111\rangle \leftrightarrow |210\rangle$, where the states are ordered as $|q_0 q_1 q_2\rangle$. This will form a Λ- and a V-type system, as shown in Figure 2.6. The effect of activating these transitions will result in a non-trivial three-qubit gate. The resulting three-qubit gate, which we named CCZS, was found to be

$$\text{CCZS}(\theta, \phi, \gamma) = |0\rangle\langle 0| \otimes \hat{I} \otimes \hat{I} + |1\rangle\langle 1| \otimes U_{\text{CZS}}(\theta, \phi, \gamma), \tag{2.14}$$

where θ, ϕ, and γ are fixed through the couplings and detunings. For equal

[1] It works if both are flux-tunable as well, but during the implementation of the gate one of them will be fixed in frequency.

[2] Universal quantum computation means that a set of quantum gates can be used to construct any arbitrary quantum computation. An example of a universal gate set is $\{\mathrm{R}_x(\theta), \mathrm{R}_z(\phi), \text{CZ}\}$.

coupling strengths and no detuning, one obtains

$$U_{\text{CZS}}(\pi/2, \pi, 0) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{2.15}$$

This gate has a more similar unitary since one can recognize the effect of a CZ-gate as well as a SWAP-like gate. In particular, when the control qubit is in state 1, it will swap the state of the two target qubits and add an additional minus-phase to the state, see Figure 2.7c. Also the $|111\rangle$ state will also acquire a geometrical phase, see Figure 2.7d. Remarkably, this three-qubit gate proved to be $\sqrt{2}$ times faster than a single CZ-gate. This increase in speed thus enhances the potential of three-qubit gates to facilitate more efficient implementation of quantum algorithms.

From numerical simulations of an ideal quantum computer, we obtained an average gate fidelity[3] of 99.46%. This high fidelity indicates that the gate operation is reliable and effective, producing the intended state of the qubit with high probability.

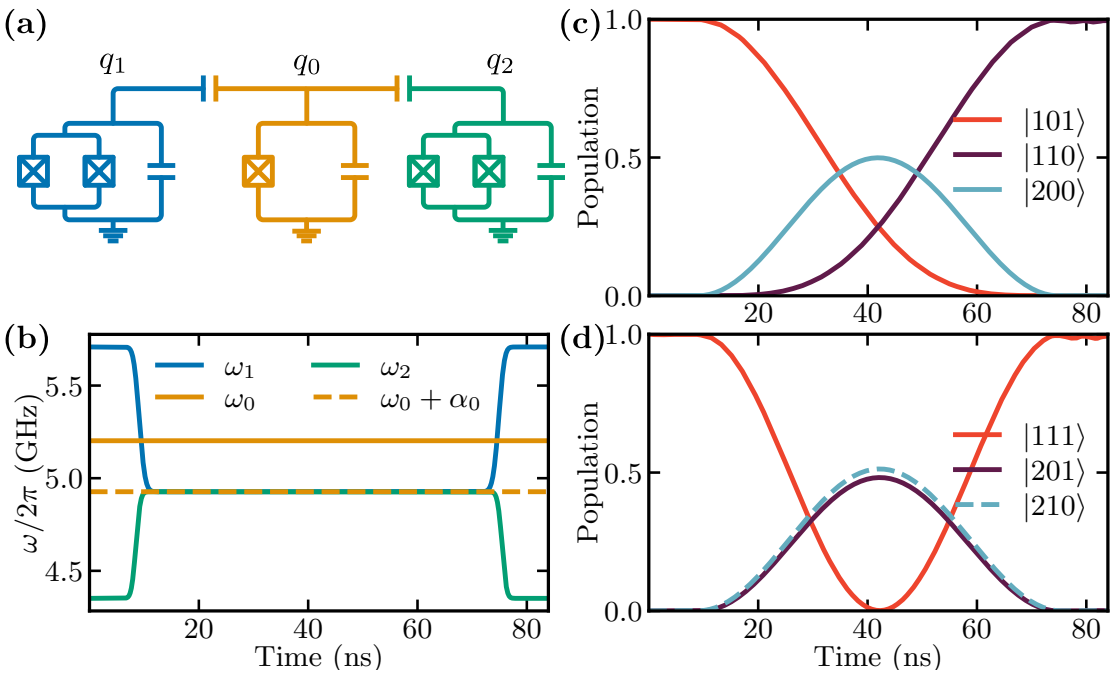

Figure 2.7: The CCZS-gate. **(a)** Illustration of the circuit, q_1 and q_2 are flux-tunable qubits, meaning that their frequency can be tuned by changing the magnetic flux through the SQUID, while q_0 can be a fixed frequency qubit. **(b)** Frequency of the qubits. **(c)-(d)** Population transfer. The qubits are labeled according to $|q_0 q_1 q_2\rangle$.

[3]The average gate fidelity is calculated as the overlap between the output state and the target state averaged over all possible input states [35].

2.3 Open quantum systems

In reality, no quantum system exists in complete isolation. A constant exchange occurs, whether it involves thermal photons energizing the qubit, spurious two-level systems coupled to the qubit causing leakage, or cosmic rays striking it [14, 36]. As such, a comprehensive understanding of the dynamics of a qubit requires a framework that encapsulates its interaction with the surrounding environment. In the upcoming section, we will explore the Lindblad master equation and the Kraus operator formalism. The former equation describes the continuous time-evolution of a quantum system as it interacts with its environment, while the latter can be understood as the integrated version of the Lindblad master equation.

2.3.1 Closed quantum systems

As seen, the dynamics of a closed quantum system are described by the Schrödinger equation (2.6). If a quantum system consists of more than a single state vector $|\psi\rangle$, it is regarded as being in a mixed state. The dynamic behavior of these mixed states is described using the density operator formalism. Here, a density operator $\hat{\rho}$ is formed as a probability-weighted sum of pure states

$$\hat{\rho} = \sum_k p_k \, |\psi_k\rangle\langle\psi_k| \,, \tag{2.16}$$

where p_k is the probability of observing the system in the state $|\psi_k\rangle$ upon measurement. The density operator $\hat{\rho}$ also obeys the normalization condition that the sum of all probabilities p_k should equal 1:

$$\mathrm{Tr}[\hat{\rho}] = \sum_k p_k = 1. \tag{2.17}$$

For a closed quantum system in a mixed state, the density operator's time evolution can be expressed as

$$\frac{\mathrm{d}\hat{\rho}(t)}{\mathrm{d}t} = -\mathrm{i}[\hat{H}, \hat{\rho}(t)] \,, \tag{2.18}$$

which is the well-known, celebrated Liouville-von Neumann equation.

2.3.2 Lindblad master equation

The description of a closed quantum system might be a good approximation for certain scenarios, but it can sometimes fall short. This becomes especially relevant in quantum computing, where it is crucial to understand how the environment influences our qubits and induces decoherence. The environment is typically considered larger than the system, i.e., containing many more degrees

of freedom [37]. For instance, one might imagine the environment as a large number of independent harmonic oscillators, similar to the free modes of the electromagnetic field.

In general, the total Hamiltonian for a quantum system interacting with an environment can be formulated as

$$\hat{H}_{\text{tot}} = \hat{H}_{\text{sys}} + \hat{H}_{\text{env}} + \hat{H}_{\text{int}}, \tag{2.19}$$

where $\hat{H}_{\text{sys}}$ refers to the system's Hamiltonian, $\hat{H}_{\text{env}}$ refers to the environmental Hamiltonian, and $\hat{H}_{\text{int}}$ describes the interaction Hamiltonian coupling the system and the environment together. The whole system's equation of motion is closed, thereby following the Liouville-von Neumann equation

$$\frac{\mathrm{d}\hat{\rho}_{\text{tot}}(t)}{\mathrm{d}t} = -\mathrm{i}[\hat{H}_{\text{tot}}, \hat{\rho}_{\text{tot}}(t)]\,. \tag{2.20}$$

The purpose of a master equation is to determine the system's equation of motion, excluding the combined system and environment. By taking the partial trace of the total density operator over the environmental variables, a reduced density operator for the system can be introduced:

$$\hat{\rho}_{\text{sys}} = \mathrm{Tr}_{\text{env}}[\hat{\rho}_{\text{tot}}]. \tag{2.21}$$

The equation of motion that describes the reduced density operator in the presence of dissipation processes between the system and its environment in the interaction picture is given by the Lindblad master equation

$$\begin{aligned}\frac{\mathrm{d}\hat{\rho}_{\text{sys}}(t)}{\mathrm{d}t} &= -\mathrm{i}[\hat{H}_{\text{sys}}(t), \hat{\rho}_{\text{sys}}(t)] \\ &\quad + \sum_n \frac{\gamma_n}{2}\Big(2\hat{\Gamma}_n\hat{\rho}_{\text{sys}}(t)\hat{\Gamma}_n^\dagger - \hat{\rho}_{\text{sys}}(t)\hat{\Gamma}_n^\dagger\hat{\Gamma}_n - \hat{\Gamma}_n^\dagger\hat{\Gamma}_n\hat{\rho}_{\text{sys}}(t)\Big),\end{aligned} \tag{2.22}$$

where $\hat{\Gamma}_n$ are the jump operators through which the environment couples to the system and γ_n represents the corresponding decay rates[4]. Eq. (2.22) extends the Liouville-von Neumann equation to a system incorporating dissipation processes with its environment, and when all $\gamma_n = 0$ one retains the Loiville-von Neuman equation. The derivation of the Lindblad master equation will not be reproduced in this thesis, but can be found in several other sources [37–39]. However, it is generally assumed that the system and environment are initially uncorrelated, meaning that they are in a product state, and the Born and Markov approximations are then used to determine the jump operators $\hat{\Gamma}_j$.

2.3.3 Kraus operator formalism

The Kraus operator formalism is another powerful method of describing the evolution of open quantum systems [1, 40]. It represents the system's evolution

[4]For a N-qubit system there are at most $2^N - 1$ number of independent jump operators.

through a set of operators $\hat{K}_i$ known as Kraus operators, which encapsulate the system's dynamics and its interaction with the environment. Mathematically it is defined as a completely positive trace preserving map

$$\Lambda(\hat{\rho}) = \sum_{i=1}^{n} \hat{K}_i \hat{\rho} \hat{K}_i^\dagger, \tag{2.23}$$

for some $n \leq d^2$, where $d = 2^N$ is the Hilbert space dimension. Moreover, the Kraus operators fulfill the following completeness relation for the map to be trace preserving

$$\sum_{i=1}^{n} \hat{K}_i^\dagger \hat{K}_i = \hat{I}. \tag{2.24}$$

The Lindblad master equation and the Kraus operator formalism are closely related. The Lindblad equation describes the continuous time-evolution of an open quantum system. In contrast, the Kraus operators describe the changes in the system's state after a finite amount of time.

The Lindblad equation provides the generator for the time-evolution superoperator in the Kraus representation. This is similar to how the Hamiltonian is the generator for the time-evolution operator U for closed quantum systems. In other words, the Kraus operators can be constructed by solving the Lindblad equation for a given time interval. Thus, both provide complementary perspectives on the dynamics of open quantum systems and are invaluable tools in understanding and calculating the evolution of such systems.

In Paper E, we use both the Kraus operator formalism and the Lindblad master equation to model the interaction of a cat qubit with its environment, which is described by single-photon loss, and in Paper D we use the Kraus operator formalism to model dephasing or depolarizing errors that occur during an error-mitigation protocol.

Chapter 3

The quantum approximate optimization algorithm

The QAOA is a versatile and powerful quantum algorithm for solving combinatorial optimization problems on near-term quantum hardware [20, 21, 41–45]. It combines classical optimization and quantum computing as a hybrid algorithm that leverages the capabilities of quantum devices for solving complex problems. QAOA has gained a lot of attention due to its potential applications in a wide range of areas, including aviation [21], finance [22], machine learning [46], and logistics [23], where finding optimal solutions is of paramount importance.

At the core of the QAOA lies a variational approach that aims to approximate the solution to a given optimization problem by preparing a variational state on the quantum computer, parameterized by a set of classical parameters. The classical optimization component of the algorithm is responsible for adjusting these parameters to minimize the expected value of the problem's cost function. The iterative nature of the QAOA allows it to gradually increase the search space in Hilbert space in search of the optimal solution. This approach makes QAOA particularly suitable for near-term quantum devices, as it can be effectively implemented on devices with a limited number of qubits and circuit depth. As quantum hardware continues to advance, QAOA is believed to play a crucial role in unlocking the potential of quantum computing to tackle challenging optimization problems.

In the next section, we will review the formalism of QAOA, and we will make the connection between the quantum adiabatic algorithm and QAOA.

3.1 From the quantum adiabatic algorithm to the QAOA

The QAOA draws inspiration from the Quantum Adiabatic Algorithm (QAA), which was also developed by Farhi *et al.* [47, 48]. The QAA uses adiabatic evolution to slowly transition from the lowest energy eigenstate of an initial Hamiltonian, to the lowest energy eigenstate of a cost Hamiltonian. The QAA Hamiltonian is formulated as the sum of two non-commuting Hamiltonians, which can be represented as

$$\hat{H}(t) = (1 - s(t))\hat{H}_B + s(t)\hat{H}_C, \tag{3.1}$$

where $s(0) = 0$ and $s(T) = 1$, with T denoting the total running time of the algorithm. The initial Hamiltonian ($\hat{H}_B$) has an easily prepared ground state, while the cost Hamiltonian ($\hat{H}_C$) typically has a ground state that is hard to prepare. For a N-qubit system, a simple initial Hamiltonian that is generally used is

$$\hat{H}_B = -\sum_{i=1}^{N} \hat{\sigma}_i^x, \tag{3.2}$$

which has $|+\rangle^{\otimes N}$ as its ground state. For combinatorial optimization problems, $\hat{H}_C$ can often be written in the form of an Ising Hamiltonian

$$\hat{H}_C = \sum_{i=1}^{N} h_i \hat{\sigma}_i^z + \sum_{i<j} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z, \tag{3.3}$$

where $\hat{\sigma}_i^z$ are Pauli-Z operators. The Ising model allows the encoding of a wide variety of problems, where the values of h_i and J_{ij} are problem-specific.

To ensure the highest likelihood of ending up in the ground state of the cost Hamiltonian, the total time T has to be of the order $\mathcal{O}(1/\Delta E_{\min}^2)$, with $\Delta E_{\min}$ signifying the minimum energy gap between the evolution's two lowest instantaneous energy eigenstates. While a linear time-dependence, $s(t) = t/T$, is often assumed, more sophisticated evolution schemes also exist that could potentially speed up the evolution [49, 50].

The QAOA is based on the insight that a possible way to simulate the QAA on a gate-based quantum computer is using trotterization. As discussed by Sun *et al.* [51], the trotterization of a continuous time-evolving operator can be expressed as

$$\hat{U}(T) \equiv \mathcal{T} \exp\left[-\mathrm{i} \int_0^T \hat{H}(t)\,\mathrm{d}t\right] \approx \prod_{k=1}^{p} \exp\Big[-\mathrm{i}\hat{H}(k\Delta t)\Delta t\Big], \tag{3.4}$$

where $\hat{U}(T)$ denotes the evolution operator from 0 to T, $\mathcal{T}$ represents the time-ordering operator, and p is a large integer, which results in $\Delta t = T/p$ being a

small time segment. When considering two non-commuting operators, $\hat{A}$ and $\hat{B}$, and a sufficiently small Δt, the Trotter formula can be used:

$$\mathrm{e}^{\mathrm{i}(\hat{A}+\hat{B})\Delta t} = \mathrm{e}^{\mathrm{i}\hat{A}\Delta t}\mathrm{e}^{\mathrm{i}\hat{B}\Delta t} + \mathcal{O}(\Delta t^2). \tag{3.5}$$

This formula can be applied to the discretized time evolution operator in Eq. (3.4) to yield

$$\hat{U}(T) \approx \prod_{k=1}^{p} \exp\Big[-\mathrm{i}(1 - s(k\Delta t))\hat{H}_B\Delta t\Big] \exp\Big[-\mathrm{i}s(k\Delta t)\hat{H}_C\Delta t\Big]. \tag{3.6}$$

As a result, it becomes feasible to approximate the QAA by alternately applying $\hat{H}_C$ and $\hat{H}_B$ in a sequential manner.

Farhi, Goldstone, and Gutman introduced a groundbreaking yet remarkably simple idea, which involves truncating the product in Eq. (3.6) to an arbitrary positive integer and redefining the time dependence in each exponent as $(1 - s(k\Delta t))\Delta t \to \beta_k$ and $s(k\Delta t)\Delta t \to \gamma_k$. Consequently, the fixed time segments transform into angles or variational parameters subject to optimization:

$$\hat{U} = \prod_{k=1}^{p} \mathrm{e}^{-\mathrm{i}\beta_k\hat{H}_B}\mathrm{e}^{-\mathrm{i}\gamma_k\hat{H}_C}, \quad p \in \mathbb{Z}^+. \tag{3.7}$$

Generally the minus-sign in Eq. (3.2) is absorbed into β_k such that $\hat{H}_B$ is redefined as

$$\hat{H}_B \to \sum_{i=1}^{N} \hat{\sigma}_i^x. \tag{3.8}$$

Hence the final state of the QAOA can be expressed as

$$|\psi_p(\vec{\gamma}, \vec{\beta})\rangle = \prod_{k=1}^{p} \mathrm{e}^{-\mathrm{i}\beta_k\hat{H}_B}\mathrm{e}^{-\mathrm{i}\gamma_k\hat{H}_C} |+\rangle^{\otimes N}, \tag{3.9}$$

where $\vec{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_p)$ and $\vec{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)$ are vectors. Each β_k lies within the range from 0 to π. This can be observed by substituting $\beta_k \to \beta_k \pm \pi$ into Eq. (3.9) and noting that $\mathrm{e}^{\pm\mathrm{i}\pi\hat{H}_B} = \prod_i \mathrm{e}^{\pm\mathrm{i}\pi\hat{\sigma}_i^x} = \prod_i(-\hat{I}_i)$, which is mere a global phase on all qubits. Similarly, if the eigenvalues of the cost Hamiltonian are integers, γ_k can be shown to lie between 0 and 2π for similar reasons. However, the challenge of selecting appropriate angles $(\vec{\gamma}, \vec{\beta})$ remains.

To pick the angles, a cost function must be constructed from which the angles can be optimized, and the cost function must be able to be fed into a classical computer that can optimize the angles and query the quantum computer. The construction of the cost function is an active topic of research [52], but typically a cost function is constructed as an expectation value from the final state and the cost Hamiltonian:

$$F_p(\vec{\gamma}, \vec{\beta}) = \langle\psi_p(\vec{\gamma}, \vec{\beta})|\hat{H}_C|\psi_p(\vec{\gamma}, \vec{\beta})\rangle. \tag{3.10}$$

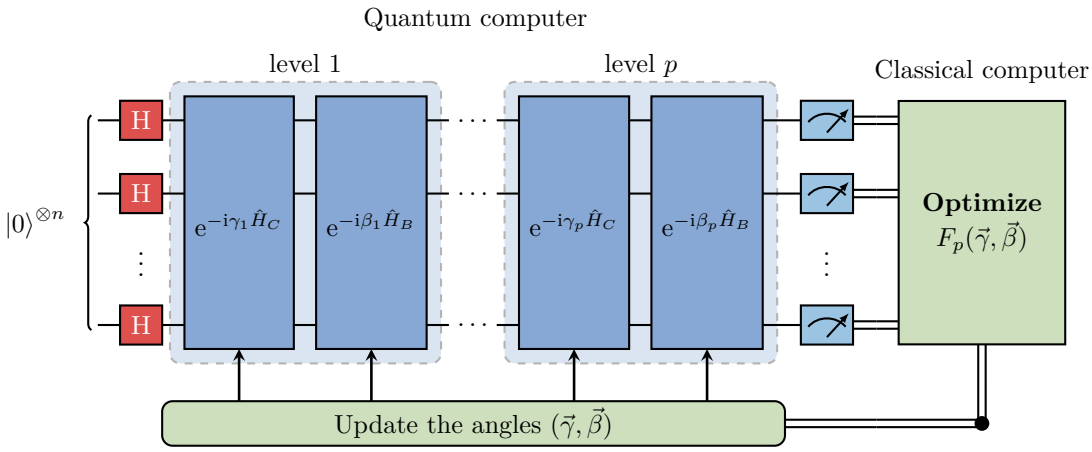

Figure 3.1: Schematic representation of the QAOA. The quantum processor prepares the variational state $|\psi_p(\vec{\gamma}, \vec{\beta})\rangle$. The angles $(\vec{\gamma}, \vec{\beta})$ are optimized in a closed loop using a classical optimizer based on the cost function.

This makes sure that by minimizing the expectation value the variational state occupies more of the low eigenenergy states of the cost Hamiltonian. Typically, this process requires the quantum computer to query a classical optimizer, which informs the quantum computer how to adjust the variational state by slightly modifying the angles to minimize the expectation value, as shown in Figure 3.1.

Significant research has been conducted on the classical optimization part of QAOA [53, 54], as well as to allow for different initial Hamiltonians $\hat{H}_B$ [55–57]. For example, numerous numerical investigations have explored various classical optimizers [58, 59]. Other studies have discovered heuristic methods that improve the classical optimization procedure [44, 60, 61]. Nonetheless, noise and finite sampling errors present unique challenges for optimizers in practice. In a few experiments running QAOA [62–64], Bayesian optimization [65], Nelder-Mead [66], and Model gradient descent [64] are some of the classical optimizers that have been implemented.

Finally, a well-studied problem in the context of QAOA is the MaxCut problem [44, 64, 67–69]. The reason why MaxCut is a well-studied problem is that it is a nontrivial problem that is easy to understand, and it has been extensively studied in classical computing, providing a basis for comparison [70]. It was furthermore the problem that Farhi *et al.* studied in their original QAOA paper, where they found nontrivial performance guarantees for the algorithm [20].

The objective of MaxCut is: given a graph $G = (V, E)$, where V is the set of vertices and E is the set of edges, the goal is to partition the set of vertices of the graph into two subsets, such that the sum of the edge weights going from one partition to the other is maximum. In other words, it seeks to "cut" the graph so that the number of edges between the two partitions (the "cut") is maximized, hence the name MaxCut. Mathematically the problem can be

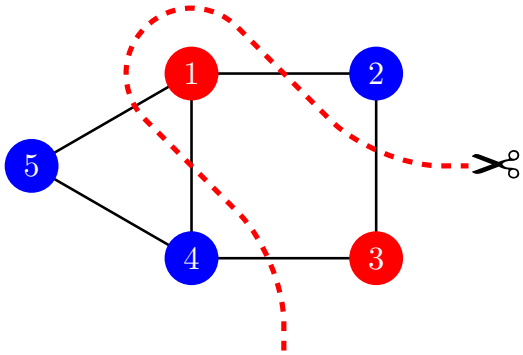

Figure 3.2: A maximum cut of a graph with 5 vertices. The dashed red line corresponds to the cut edges. An edge is cut if two vertices connected by an edge are assigned different colors.

formulated as an optimization problem

$$\text{minimize} \quad -\frac{1}{2}\sum_{\{i,j\}\in E} w_{ij}(1 - s_i s_j), \tag{3.11}$$

$$\text{subject to: } s_i \in \{-1, 1\} \quad i \in V. \tag{3.12}$$

where w_{ij} is the edge weight between vertex i and j, and s_i determines the partition of each vertex i in the graph. An example of a graph and its maximum cut is shown in Figure 3.2.

To map this problem onto a cost Hamiltonian, all we have to do is to replace the classical variables s_i with Pauli-Z matrices. The corresponding MaxCut Hamiltonian then reads

$$\hat{H}_C = -\frac{1}{2}\sum_{\{i,j\}\in E} w_{ij}(\hat{I} - \hat{\sigma}_i^z \hat{\sigma}_j^z), \tag{3.13}$$

where the ground state to this Hamiltonian corresponds to the maximum cut.

Chapter 4

Application of QAOA to a real-world optimization problem

Quantum computing has introduced a new approach for solving complex optimization problems. Notably, the QAOA has emerged as a promising candidate.

In this chapter, we will see how QAOA can be applied to solve a real-world optimization problem, specifically, the tail-assignment problem.

4.1 The tail-assignment problem

The tail-assignment problem is a complex combinatorial optimization problem arising in the airline industry, specifically in aircraft scheduling [71, 72]. The primary objective of the tail-assignment problem is to assign a specific aircraft (identified by its tail number), also called a *route*, to a sequence of flights in such a way that operational constraints are satisfied while minimizing operational costs. The problem is crucial to the efficient utilization of airline resources and has a direct impact on customer satisfaction, punctuality, and overall profitability of an airline.

Several factors and constraints must be considered when solving the tail-assignment problem, including:

1. Maintenance requirements: Each aircraft must undergo periodic maintenance checks, which need to be scheduled within the tail assignment plan.

2. Flight connections: An aircraft assigned to a particular flight must be available at the corresponding departure airport at the correct time.

3. Aircraft compatibility: Some flights may require specific aircraft types due to factors like seating capacity, range, or airport constraints.

4. Crew constraints: Crew schedules and regulations must be taken into account, ensuring that they are assigned to aircraft within legal limits and considering their qualifications.

5. Turnaround times: The minimum time between an aircraft's arrival and its subsequent departure, which includes refueling, cleaning, and passenger boarding, must be respected.

The tail-assignment problem is an NP-hard problem, which means that finding an optimal solution becomes exponentially increasingly difficult as the number of flights and aircraft grows. The typical classical algorithm to tackle this problem is column generation [73].

The tail-assignment problem can formally be defined as follows:

$$\min \quad \sum_{r \in R} c_r z_r, \tag{4.1}$$

$$\text{s.t.} \quad \sum_{r \in R} a_{fr} z_r = 1, \quad \forall f \in F, \tag{4.2}$$

$$z_r \in \{0, 1\}, \quad \forall r \in R, \tag{4.3}$$

where R is the set of all routes, c_r is the cost of route r, F is the set of all activities, i.e., flights or group of sequential flights, a_{fr} is 1 if activity f is included in route r, and z_r are binary decision variables that state if route r is part of the solution or not. Problems with thousands of flights and about 80 aircrafts are common for medium-sized carriers [71].

In Paper A, we study QAOA applied to simplified instances of the tail-assignment problem. We made simulations of instances of the problem derived from real-world data, which was provided by the Jeppesen company. These instances were reduced to fit on a quantum computer with 8, 15, and 25 qubits, with a total of 10 instances for each size. These sizes were chosen so that it is possible to fit the simulation of QAOA on a classical computer. The reduction procedure resulted in only one feasible solution per instance. This means that the cost c_r can be neglected since it is only relevant to include it if there exist multiple solutions to Eq. (4.2). This simplified the problem to only solving Eq. (4.2), which is equivalent to an Exact Cover problem. An example of an Exact Cover problem is depicted in Figure 4.1. The next step was to map this mathematical formulation to a cost Hamiltonian $\hat{H}_C$. To map the Exact Cover problem onto a cost Hamiltonian, one has to first transform Eq. (4.2) into a cost function. This can be done by subtracting 1 from the r.h.s of Eq. (4.2) and squaring the expression:

$$C(z) = \sum_{i=1}^{u} \left(\sum_{j=1}^{v} a_{ij} z_j - 1 \right)^2 . \tag{4.4}$$

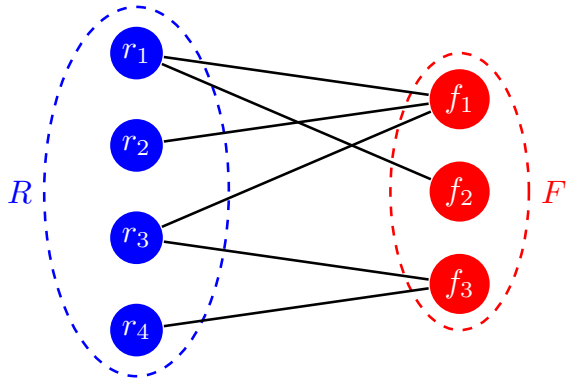

Figure 4.1: Example of an Exact Cover problem. The goal is to pick a subset of vertices from the set R such that one and only one edge from the subset lands on each of the vertices in the set F. In this example, the subset $\{r_1, r_4\}$ corresponds to an exact cover.

Here $u \equiv |F|$ denotes the cardinality of F and $v \equiv |R|$ denotes the cardinality of R. This formulation as a cost function makes sure that there exists a solution if and only if there exists a string z such that $C(z)$ is zero.

Next, we want to write this cost function in the form of an Ising Hamiltonian Eq. (3.3). This is done by substituting the binary variables $z_j \in \{0, 1\}$ with spin variables $s_j \in \{1, -1\}$,

$$z_j = \frac{1 - s_j}{2}.$$

By using this substitution and expanding the square of Eq. (4.4), we obtain the Ising energy function for the Exact Cover problem

$$E(s_1, \ldots, s_v) = \sum_{1 \leq i < j \leq v} J_{ij} s_i s_j + \sum_{i=1}^{v} h_i s_i, \tag{4.5}$$

where

$$J_{ij} \equiv \frac{1}{2} \sum_{k=1}^{u} a_{ki} a_{kj} \tag{4.6}$$

is a penalty term related to having two or more routes covering the same flight, and

$$h_i \equiv \frac{1}{2} \sum_{j=1}^{u} a_{ji} \left(\sum_{k=1}^{v} a_{jk} - 2 \right) \tag{4.7}$$

is a penalty related to over or under-covering the number of flights, i.e., leaving flights unattended will result in a cost. Finally, we quantize Eq. (4.5) by promoting the spin variables s_i to Pauli-Z matrices as $s_i \to \hat{\sigma}_i^z$. The resulting Hamiltonian is

$$\hat{H}_C = \sum_{1 \leq i < j \leq v} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z - \sum_{i=1}^{v} h_i \hat{\sigma}_i^z. \tag{4.8}$$

The ground state of this Hamiltonian corresponds to the solution to the Exact Cover problem.

To assess the performance of QAOA for solving the TAS instances, we look at the success probability, which is defined as the squared magnitude of the overlap between the solution state vector $|z_{\text{sol}}\rangle$ and the variational state $|\psi_p(\vec{\gamma}, \vec{\beta})\rangle$. Mathematically, this is represented as

$$|\langle z_{\text{sol}}|\psi_p(\vec{\gamma}, \vec{\beta})\rangle|^2. \tag{4.9}$$

In Figure 4.2a we plot the mean success probability as a function of the iteration level p. This is done using the optimal variational parameters discovered for the three different problem sizes. As can be observed, the success probability tends to increase with the iteration level across all instances. In Figure 4.2b, we present a simulation of the success probability versus p for a single instance selected from each problem size, extending up to $p = 20$.

An interesting observation from this simulation is that the instances with 25 qubits consistently exhibited a higher success probability than those with 15 qubits at any given iteration level of the QAOA. This might initially seem counter-intuitive, as one would typically expect larger instances (i.e., more qubits) to correspond to more challenging problems. To explain this unexpected finding, we conducted further analysis and proposed that the success probability is not solely determined by the size of the problem (i.e., the number of qubits). Instead, it also depends on the structure of the problem. Specifically, we discovered an empirical relationship between the connectivity of the problem graph, as shown in Figure 4.2c, and the single-shot success probability of the algorithm. From this, we inferred that a more connected graph is associated with a more complex problem. This suggests that the structure of the problem can influence the performance of the QAOA. Finally, Figure 4.2d illustrates the probability of observing a specific eigenvalue of the cost Hamiltonian $\hat{H}_C$ for one of the 25-qubit instances and is demonstrated across the iteration levels $p = 0, 1, 2$. As can be seen, when p increases the variational state occupies more of the eigenstates associated with a low cost.

In summary, these findings suggest that QAOA could potentially be a powerful algorithm for solving the tail-assignment problem. However, it is important to note that these results are obtained for relatively small problem sizes, and further research is needed to understand the potential and limitations of this approach fully.

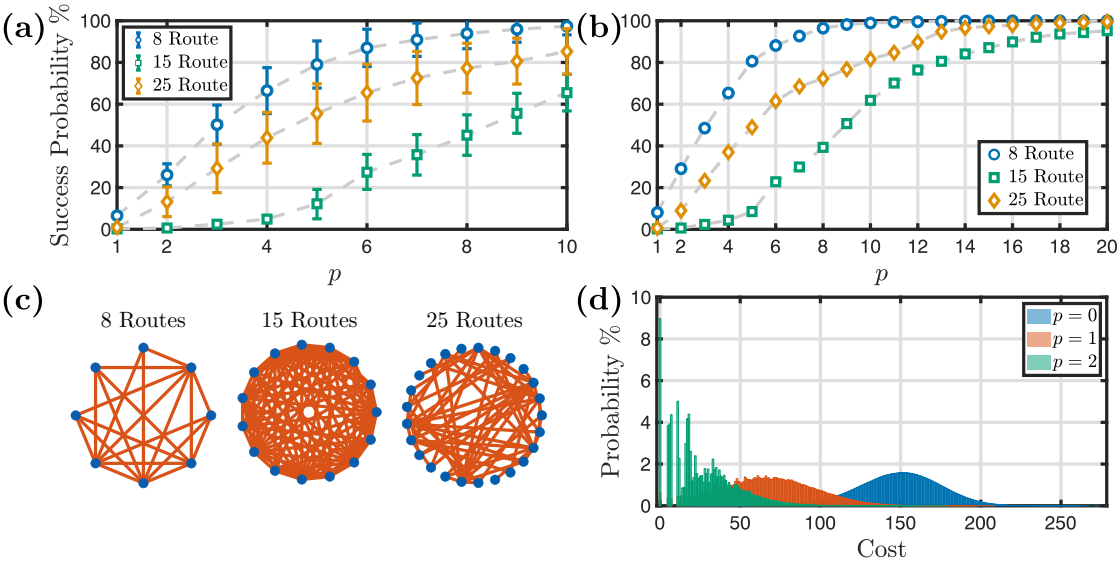

Figure 4.2: **(a)** This figure depicts the mean success probability as a function of the iteration level p, utilizing the optimal found variational parameters for the three distinct problem sizes: 8, 15, and 25 qubits. The error bars in the figure represent the standard deviation of the mean success probability. **(b)** The success probability $|\langle z_{\text{sol}}|\psi_p(\vec{\gamma},\vec{\beta})\rangle|^2$ as a function of p is shown for a single instance selected from each problem size. **(c)** The three instances displayed in **(b)** represented graphically. **(d)** The likelihood of obtaining a specific cost (or equivalently, the likelihood of measuring a specific eigenvalue of the cost Hamiltonian) from a measurement of the state $|\psi_p(\vec{\gamma},\vec{\beta})\rangle$ is shown for the iteration levels $p = 0, 1, 2$ for one of the 25 qubit instances. Here, $p = 0$ represents the initial or "random" state $|+\rangle^{\otimes N}$.

Chapter 5

Implementation of QAOA

The QAOA has been implemented on various hardware platforms such as photonic [74], trapped ion [75], and superconducting quantum processors [62–64, 64, 76]. To prepare the variational state $|\psi_p(\vec{\gamma}, \vec{\beta})\rangle$ on an actual quantum computer, the unitary matrices in Figure 3.1 must be decomposed into single and two-qubit gates. The decomposition depends on the primitive quantum gates available for the specific hardware. In Paper B, we implemented the QAOA on a two-qubit superconducting qubit device. Here we will illustrate how the variational state can be constructed using a universal gate set utilized in Paper B, composed of single-qubit rotations $\mathrm{R}_x(\theta)$, and $\mathrm{R}_z(\phi)$, and the two-qubit CZ-gate.

The initial state for QAOA can be generated by applying the Hadamard gate to each qubit in the all-zero state $|0\rangle^{\otimes N}$. Since the Hadamard gate is not part of the gate set, it must be compiled using a combination of gates, which can be done with three rotations: two around the x-axis and one around the z-axis, $\mathrm{H} = \mathrm{i}\mathrm{R}_x\left(\frac{\pi}{2}\right)\mathrm{R}_z\left(\frac{\pi}{2}\right)\mathrm{R}_x\left(\frac{\pi}{2}\right)$. The unitary involving the sum of Pauli-X matrices

$$\mathrm{e}^{-\mathrm{i}\beta\hat{H}_B} = \mathrm{e}^{-\mathrm{i}\beta\sum_{i=1}^{N}\hat{\sigma}_i^x} = \prod_{i=1}^{N}\mathrm{e}^{-\mathrm{i}\beta\hat{\sigma}_i^x}, \tag{5.1}$$

can be written as a product, as all Hamiltonian terms commute. This unitary can be implemented as parallel single-qubit rotations around the x-axis

$$\mathrm{e}^{-\mathrm{i}\beta\hat{\sigma}_i^x} \equiv \text{—}\boxed{\mathrm{R}_x(2\boldsymbol{\beta})}\text{—} \tag{5.2}$$

The cost Hamiltonian has two components: a two-body Hamiltonian and a single-body Hamiltonian. These terms commute, allowing them to be applied in any order

$$\mathrm{e}^{-\mathrm{i}\gamma\hat{H}_C} = \prod_{1\leq i<j\leq N}\mathrm{e}^{-\mathrm{i}\gamma J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z}\prod_{i=1}^{N}\mathrm{e}^{-\mathrm{i}\gamma h_i\hat{\sigma}_i^z}. \tag{5.3}$$

The single-qubit term can be implemented as rotations around the z-axis

$$\mathrm{e}^{-\mathrm{i}\gamma h_i \hat{\sigma}_i^z} \equiv \quad \mathrm{R}_z(2\gamma h_i) \tag{5.4}$$

while the two-qubit interaction can be implemented using a local single-qubit gate between two controlled-CZ and Hadamard-gates

$$\mathrm{e}^{-\mathrm{i}\gamma J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z} \equiv \tag{5.5}$$

If all elements of the interaction matrix J_{ij} are non-zero, a total number of $N(N-1)/2$ CZ-gates for each level p would be required. However, this assumes that two-qubit gates can be directly applied between any two qubits, which is not always possible due to hardware constraints. Nonetheless, SWAP-gates can be used to move distant qubits close to each other,

$$\equiv \tag{5.6}$$

Furthermore, the number of SWAP gates needed for this depends, of course, on the hardware's connectivity. For instance, in a linear array of qubits, each qubit can be made to interact with all other qubits using a circuit depth of N with $N(N-1)/2$ number of SWAP-gates [77].

5.1 Experimental realization of QAOA in a superconducting device

In Paper B, we explored the application of the QAOA on a hardware platform composed of two superconducting transmon qubits and one parametrically modulated coupler. The primary objective was to conduct a proof of principle experiment aimed at solving small instances of the Exact Cover problem, which was inspired by the TAS problem. The experiment successfully demonstrated the implementation of QAOA with up to $p = 2$ on the Chalmers superconducting quantum processor, with an observed increase in success probability from $p = 1$. This was made possible by achieving sufficiently high gate fidelities.

In the paper, we applied the QAOA to four different Exact Cover problems, as outlined in Table 5.1. For $p = 1$, we applied a simple grid search for finding the optimal parameters $(\gamma_{\mathrm{opt}}, \beta_{\mathrm{opt}})$. From the results in Figure 5.1, it can be seen that our measurements showed excellent agreement with theoretical predictions, indicating low coherent and incoherent error rates. This was a significant finding

Table 5.1: The four distinct Exact Cover problems, each with two subsets, along with their solutions and corresponding sets of coefficients in the cost Hamiltonian $\hat{H}_C = h_1\hat{\sigma}_1^z + h_2\hat{\sigma}_2^z + J\hat{\sigma}_1^z\hat{\sigma}_2^z$, are presented.

Problem	Subsets	h_1	h_2	J	Solution
A	$r_1 = \{f_1, f_2\}$ $r_2 = \{f_1\}$	$-1/2$	0	$1/2$	$\|10\rangle$
B	$r_1 = \{f_1, f_2\}$ $r_2 = \{\}$	-1	0	0	$\|10\rangle$ or $\|11\rangle$
C	$r_1 = \{f_1\}$ $r_2 = \{f_2\}$	$-1/2$	$-1/2$	0	$\|11\rangle$
D	$r_1 = \{f_1, f_2\}$ $r_2 = \{f_1\}$	$-1/2$	0	$1/2$	$\|10\rangle$ or $\|01\rangle$

as it demonstrated the potential of our approach in solving optimization problems using a quantum computer. For $p > 1$, the classical brute force approach for finding the optimal angles scales doubly exponentially in p, necessitating the exploration of more resource-efficient optimization techniques. We investigated three classical optimizers: Bayesian optimization with Gaussian processes (BGP), Nelder-Mead, and covariance matrix adaptation evolution strategy. Ultimately, there was no significant difference in the performance observed among the different classical optimizers, except that BPGs exhibited a slightly higher convergence probability than the other two. At the end of the optimization, a success probability of 96.6% was observed for $p = 2$. This can be compared to the theoretical predicted success probability of 100% that an ideal (noise-free) quantum computer would achieve. We attribute the divergence from 100% to decoherence and imperfect gates.

This research contributed to the ongoing efforts in the field of quantum computing to harness the power of quantum algorithms for solving complex real-world problems and paved the way for future research in the application of quantum algorithms to combinatorial optimization problems.

5.2 Study of QAOA implementation with cat qubits

In our previous work, Paper E, we explored the implementation of QAOA with cat qubits. One of the key characteristics of cat qubits is that the dominant noise mechanism, photon losses, results in Z-biased noise on the logical cat subspace [78, 79], which means that $\hat{\sigma}_z$ errors are more likely to happen than $\hat{\sigma}_x$ and $\hat{\sigma}_y$ errors. This noise bias is a significant aspect as it influences the performance of error correction [80] and error mitigation strategies [81], the latter which we will explore in Chapter 6, and it can yield better performance

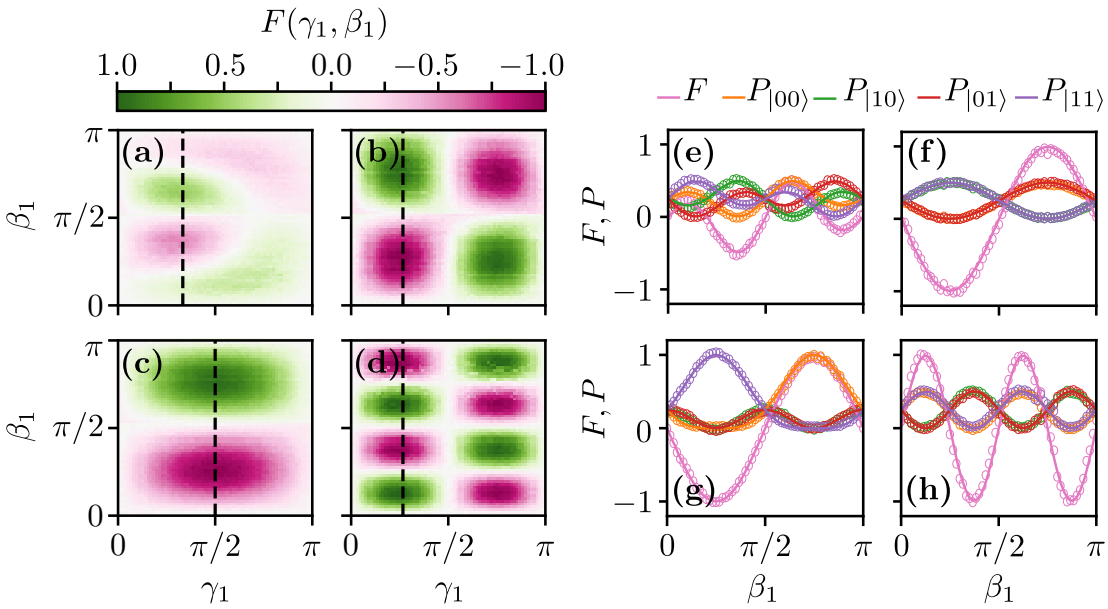

Figure 5.1: The cost function $F(\gamma, \beta)$ for the four different instances of the Exact Cover problems, as given in Table 5.1, is depicted in panels **(a)**–**(d)**. These panels display a 61×61 grid, with each experimental data point being the average of 5000 measurements conducted on the Chalmers quantum processor. The dashed lines represent the positions of the line-cuts in panels **(e)**–**(h)**. These panels provide a comparison between the experimental results (open circles) and theoretical predictions (solid lines) for the four Exact Cover problems. Each color, as indicated at the top, corresponds to either a state probability or the value of the cost function F.

for QAOA in the presence of errors, as we will see in this chapter.

In this section, we will explain how QAOA can be implemented using cat qubits and present the results from benchmarking its performance for small instances of the MaxCut problem. We find that cat qubits are more robust to noise than discrete-variable qubits, showing promise for implementing QAOA using cat qubits on NISQ devices.

Our study focuses on implementing QAOA with cat qubits using two-photon pumped Kerr Nonlinear Resonators (KNRs) [79, 82–84]. Such a system can, for example, be realized by coupling a resonator to a SQUID and adding a varying magnetic flux through the SQUID at twice the resonator frequency to generate pairs of photons in the resonator. The Hamiltonian for the two-photon pumped KNR is given by

$$\hat{H} = -\Delta \hat{a}^\dagger \hat{a} - K \hat{a}^{\dagger 2} \hat{a}^2 + G(\hat{a}^{\dagger 2} + \hat{a}^2), \tag{5.7}$$

where $\Delta = 2\omega_r - \omega_p$ is the detuning of the resonator frequency ω_r and the two-photon pump frequency ω_p, K is the amplitude of the Kerr non-linearity (compared to previous chapters, we have named the Kerr-nonlinearity K instead of α to match with the notation in Paper E, and as not to confuse it with the

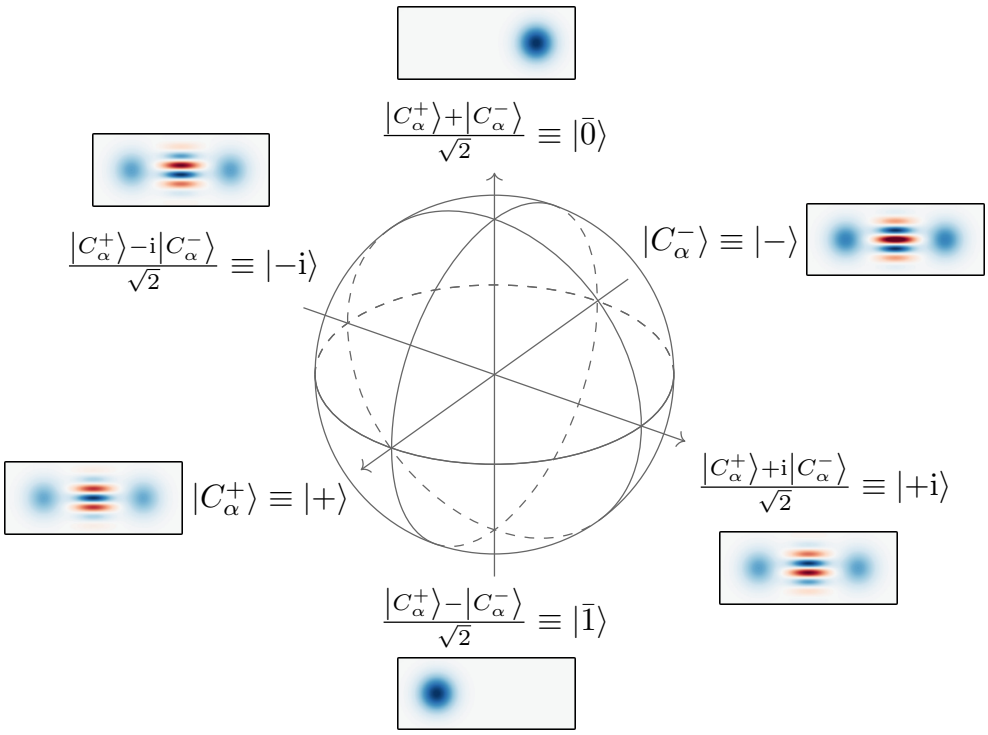

Figure 5.2: The computational states of the cat qubit that lie along one of the three principal axes visualized on the Bloch sphere along with their Wigner function.

displacement amplitude of the coherent state $|\alpha\rangle$) and G is the amplitude of the two-photon pump. When the detuning $\Delta = 0$, the Hamiltonian can be written as

$$\hat{H} = -K\left(\hat{a}^{\dagger 2} - \frac{G}{K}\right)\left(\hat{a}^2 - \frac{G}{K}\right) + \frac{G^2}{K}. \tag{5.8}$$

Since $\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle$, the coherent states $|\pm\alpha\rangle$ with $\alpha = \sqrt{G/K}$ are degenerate eigenstates of the Hamiltonian Eq. (5.8) with eigenenergy G^2/K. Thus, any linear combination of $|\pm\alpha\rangle$ are also degenerate eigenstates, such as the even and odd linear combination of these two coherent states:

$$\left|C_\alpha^\pm\right\rangle = N_\pm(|\alpha\rangle \pm |-\alpha\rangle) \quad \text{with} \quad N_\pm = \sqrt{2(1 \pm \mathrm{e}^{-2|\alpha^2|})}, \tag{5.9}$$

which are known as the even and odd cat states. We can take advantage of this well-defined subspace to encode our computational basis states $|\bar{0}\rangle$, $|\bar{1}\rangle$, defining the qubit (the bar notation is used to distinguish the computational states from the zero and one photon number states). To this aim, one possibility is to directly identify the qubit basis states with $|\alpha\rangle$ and $|-\alpha\rangle$ [85]. However, these states are quasi-orthogonal as $\langle -\alpha|\alpha\rangle = \mathrm{e}^{-2\alpha^2}$. Another possibility consists in choosing the following encoding [86]:

$$|\bar{0}\rangle = \frac{|C_\alpha^+\rangle + |C_\alpha^-\rangle}{\sqrt{2}}, \quad |\bar{1}\rangle = \frac{|C_\alpha^+\rangle - |C_\alpha^-\rangle}{\sqrt{2}}. \tag{5.10}$$

Table 5.2: Average gate fidelity of the $R_{zz}(\Theta)$ and $R_x(\theta)$-gate for the cat qubit and the discrete-variable qubit, respectively. The results are averaged over 20 points evenly spaced between 0 and π.

Avg. gate fid. (%)	$R_{zz}(\Theta)$	$R_x(\theta)$
Cat qubits	99.16	98.60
Discrete-variable qubits	99.16	98.62

In this case, the computational basis states are orthogonal even for small α, while for large α, they are approximately equal to $|\bar{0}\rangle \approx |\alpha\rangle$ and $|\bar{1}\rangle \approx |-\alpha\rangle$. In Figure 5.2, the computational states that lie along one of the principal axes of the Bloch sphere are shown along with their Wigner function. For single-photon losses, the encoding of Eq. (5.10) constitutes a noise-biased qubit where the loss of single-photons results in a phase error plus an exponentially small bit-flip error on the computational states with respect to α. Indeed, by defining the projection operator $\hat{I} = |\bar{0}\rangle\langle\bar{0}| + |\bar{1}\rangle\langle\bar{1}|$, its action on the annihilation operator $\hat{a}$ gives

$$\hat{I}\hat{a}\hat{I} = \frac{\alpha}{2}(\eta + \eta^{-1})\hat{\sigma}_z + \mathrm{i}\frac{\alpha}{2}(\eta - \eta^{-1})\hat{\sigma}_y, \tag{5.11}$$

where $\eta \equiv N_+/N_-$, and $\hat{\sigma}_z$, $\hat{\sigma}_y$ are the two Pauli matrices in the computational subspace of the cat qubit. For large α, $\eta \to 1$ which results in $\hat{I}\hat{a}\hat{I} = \alpha\hat{\sigma}_z$, and we thus see that a single-photon loss event corresponds to a phase-error on the computational basis states.

5.2.1 Running QAOA with cat qubits

In order to run QAOA, one needs to prepare all resonators in the state $|+\rangle$, i.e., in the case of the cat qubit, the cat state $|C_\alpha^+\rangle$. Such a cat state can be generated deterministically in KNRs by starting from the vacuum, which is an eigenstate of Hamiltonian Eq. (5.8) for $G = 0$, and then adiabatically increasing G [82, 84]. Since the Hamiltonian in Eq. (5.8) is symmetric under parity inversion $\hat{a} \to -\hat{a}$, the KNR follows the adiabatic evolution from the vacuum while also conserving the parity, $[\mathrm{e}^{\mathrm{i}\pi\hat{a}^\dagger\hat{a}}, \hat{H}] = 0$, thus ending up in the even parity cat state $|C_\alpha^+\rangle$. Alternatively, a cat state can also be generated using a sequence of SNAP and displacement gates applied to the vacuum state [87].

The other gates needed to run QAOA are implemented according to Refs. [82, 84], where the $\mathrm{R}_z(\phi)$-gate is implemented in KNRs by means of a single-photon drive $E(t)(\hat{a}^\dagger + \hat{a})$. The $\mathrm{R}_x(\theta)$-gate is implemented through a time-dependent detuning $\Delta(t)\hat{a}^\dagger\hat{a}$, and $\mathrm{R}_{zz}(\Theta)$-gate is implemented through a time-dependent beam-splitter interaction between two KNRs $g(t)(\hat{a}_1^\dagger\hat{a}_2 + \hat{a}_1\hat{a}_2^\dagger)$.

In our simulations, we benchmarked MaxCut on 30 randomly generated Erdős–Rényi graphs with 8 vertices and with edge probability 0.5. This means that each vertex of the graph is connected to every other vertex with a probability of 50%.

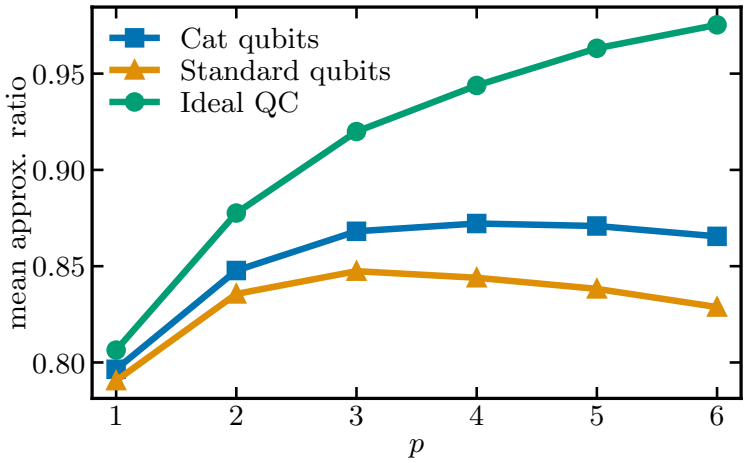

Figure 5.3: The mean approximation ratio averaged over 30 instances of 8-qubit MaxCut graphs. The circle corresponds to the approximation ratio of an ideal (noise-free) quantum computer. The square is the approximation ratio obtained using cat qubits, and the triangle is with discrete-variable qubits, meaning qubits encoded into discrete two-level systems. The average gate fidelity was chosen to be close to identical for the cat qubits and discrete-variable qubits with values reported in Table 5.2.

A typical single-photon loss to Kerr nonlinearity ratio of $1500/K$ was chosen, and in order to make a fair comparison between the performance of the cat qubit and the discrete-variable qubit, we chose the single-photon loss rate for the discrete-variable qubit such that the average gate fidelities are the same between the two systems. By doing so, we can compare which encoding, continuous versus discrete, is the best for QAOA. Using the average gate-fidelities that were numerically calculated for the cat qubits in Table 5.2, the corresponding single-photon loss rate for the discrete-variable qubits that result in the same average gate fidelity as for the cat qubits were obtained, which are also presented in Table 5.2.

As a metric for comparison between the performance of cat qubits and the discrete-variable qubits, we look at the approximation ratio, defined as

$$r \equiv \frac{\mathrm{Tr}(\hat{\rho}\hat{H}_C)}{C_{\max}}, \tag{5.12}$$

where the numerator is the expected cut value with $\hat{\rho}$ the density matrix output from QAOA, and $C_{\max}$ is the value of the maximum cut. The simulation results are presented in Figure 5.3. In both scenarios, the approximation ratio initially increases with p. However, at one point, the noise from the gates implementing the QAOA sequence begins to negatively impact the advantage of using larger depth circuits, leading to a decrease in the approximation ratio. The findings indicate that with identical average gate fidelities, the cat qubit device consistently outperforms the discrete-variable qubit device at all itera-

Table 5.3: Error channel in the computational basis of the cat qubit for the $R_x(\pi/2)$, and $R_{zz}(\pi/2)$-gate. The error channel is shown for coefficients with absolute values greater than $\geq 10^{-3}$.

Gate	Error Channel	Coefficients
$R_x(\pi/2)$	$(1-p_y-p_z)\hat{\rho}+p_z\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z+p_y\hat{\sigma}_y\hat{\rho}\hat{\sigma}_y$ $+p_{yz}\hat{\sigma}_y\hat{\rho}\hat{\sigma}_z+p_{zy}\hat{\sigma}_z\hat{\rho}\hat{\sigma}_y$	$p_y=0.01,\ p_z=0.01,$ $p_{yz}=p_{zy}=-0.002$
$R_{zz}(\pi/2)$	$(1-p_1-p_2)\hat{\rho}+p_1\hat{\sigma}_1^z\hat{\rho}\hat{\sigma}_1^z+p_2\hat{\sigma}_2^z\hat{\rho}\hat{\sigma}_2^z$	$p_1=0.005,\ p_2=0.005$

tion levels p, suggesting an advantage for the cat qubit implementation. In the context of discrete-variable qubits, the peak approximation ratio is reached at $p=3$, whereas for cat qubits, this peak is achieved at $p=4$.

An argument for why the cat qubits perform better can be understood by how the two gates: $\mathrm{R}_x(\theta)$ and $\mathrm{R}_{zz}(\Theta)$ handle errors. For example, consider a $\hat{\sigma}_z$ error that occurs prior to the $\mathrm{R}_x(\theta)$-gate. By commuting the error through the $\mathrm{R}_x(\theta)$-gate we obtain we obtain

$$\mathrm{R}_x(\theta)\hat{\sigma}_z=\mathrm{e}^{-\mathrm{i}\theta\hat{\sigma}_x/2}\hat{\sigma}_z=\cos\left(\frac{\theta}{2}\right)\hat{\sigma}_z-\mathrm{i}\sin\left(\frac{\theta}{2}\right)\hat{\sigma}_x\hat{\sigma}_z=\hat{\sigma}_z\mathrm{R}_x(-\theta), \tag{5.13}$$

where in the last equality, we have used the anti-commutation relation $\hat{\sigma}_x\hat{\sigma}_z=-\hat{\sigma}_z\hat{\sigma}_x$. This shows that a $\hat{\sigma}_z$ error introduces an additional π rotation of the $\mathrm{R}_x(\theta)$-gate. For the $\mathrm{R}_{zz}(\Theta)$-gate, a $\hat{\sigma}_z$ error acting on either of the two qubits commutes through the gate, meaning only the $\mathrm{R}_x(\theta)$-gate will be influenced by dephasing errors, resulting in an additional π-rotation. This is evident from Table 5.3, where we have extracted the error channel for the cat qubit for the two gates. In this scenario, the error channel can be considered as a channel applied post-gate. For the $\mathrm{R}_{zz}(\Theta)$, it is observed that the gate maintains the noise channel structure, implying that the error channel remains as $\hat{\sigma}_z$ errors post-gate, while the $\mathrm{R}_x(\theta)$-gate depolarizes the channel, as indicated by the presence of the $\hat{\sigma}_y$ terms.

For the discrete-variable qubit, in addition to $\hat{\sigma}_z$ errors, both $\hat{\sigma}_x$ and $\hat{\sigma}_y$ errors can occur. This will lead to faulty implementations for both the $\mathrm{R}_x(\theta)$- and $\mathrm{R}_{zz}(\Theta)$-gate. Furthermore, for a typical graph, the number of $\mathrm{R}_{zz}(\Theta)$-gates for each level-p is equivalent to the number of edges in the MaxCut graph. On average, the graphs we examined had a mean edge count of 14.1, while the $\mathrm{R}_x(\theta)$-gate count is equivalent to the number of vertices, or equivalently the number of qubits, which was 8 in this case.

Chapter 6

Error mitigation

The evolution of quantum computing hinges on effectively managing quantum errors. While fault-tolerant quantum error correction is critical for scalable quantum computation, current state-of-the-art quantum processors fail to meet the requisite hardware performance and resource overheads needed for implementing quantum error correction on multiple qubits. As such, alternative techniques for noise mitigation have been developed, including those based on scaling noise [88–91], predictive noise learning [92, 93], and exploiting noise-free circuit symmetries to flag for errors [94–98]. In particular, these techniques are less resource-intensive than quantum error-correcting codes.

Recently, there has been a surge of interest in algorithm- and noise-specific error-mitigation techniques [81, 89, 91, 95, 99–102]. A new technique that has shown considerable promise is virtual distillation, also known as error suppression by derangement [101, 102]. This method has been demonstrated to exponentially suppress errors in the estimation of the expectation value of observables, something that is very relevant for QAOA. The approach involves the use of M redundant copies of a noisy quantum state that are cyclically permuted to calculate the expectation value. Under the assumption that noise merely mixes the ideal state with orthogonal error states, the symmetries inherent in the cyclic-permutation state can exponentially suppress the contributions of the error states to the expectation value.

In practical applications, as we will see in this chapter, virtual distillation is typically implemented using an auxiliary qubit and controlled-SWAP (CSWAP) gates. However, this method is vulnerable to errors itself, as errors can occur during the CSWAP-gates. Therefore, in Paper D, we conducted an in-depth exploration of how different types of noise affects the virtual distillation circuit. We found conditions that make errors less harmful, thereby enhancing the performance of the virtual distillation protocol in practical settings.

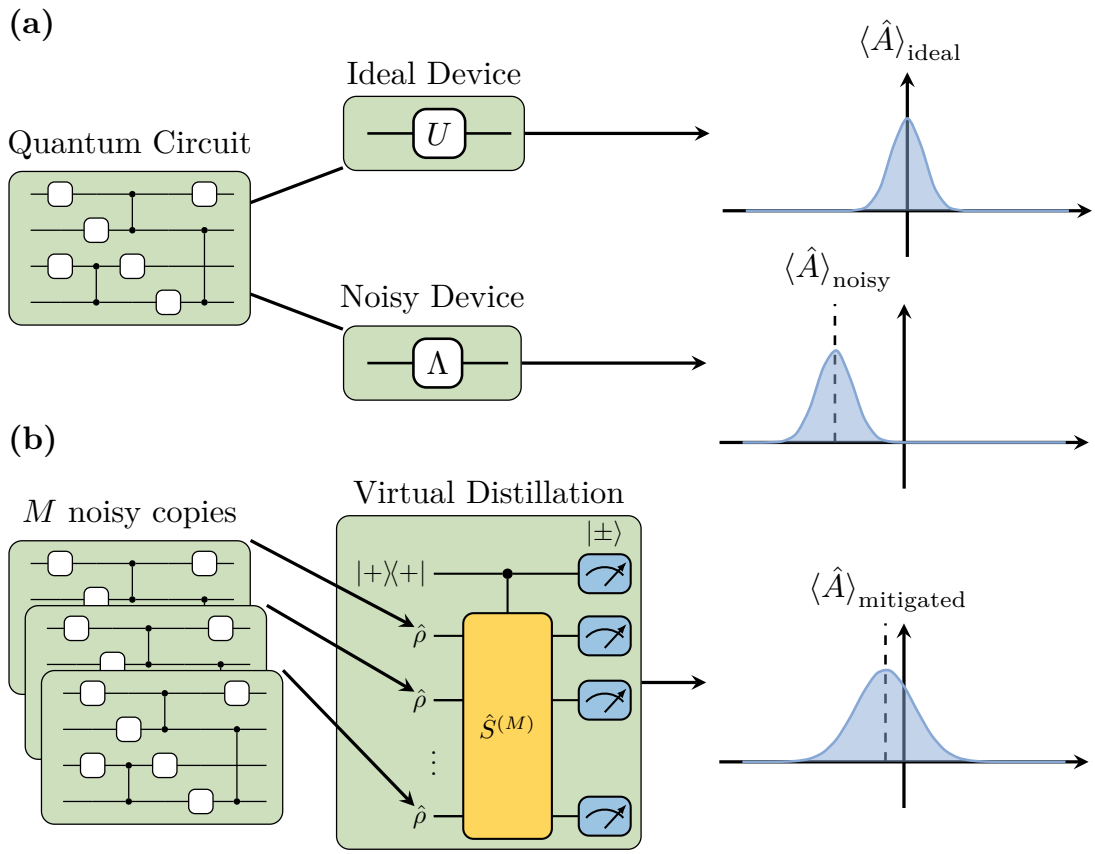

Figure 6.1: **(a)** The expectation value of an observable $\hat{A}$ under ideal and noisy conditions. Noise introduces a bias with respect to the ideal expectation value. **(b)** Quantum computation employing virtual distillation for error mitigation. This technique significantly reduces the bias introduced by noise.

6.1 Virtual distillation

In many quantum computing applications, including QAOA, the estimation of expectation values is of paramount importance. However, due to noise, a quantum computer often prepares a noisy mixed state $\hat{\rho}$, rather than the ideal state $|\psi_{\text{ideal}}\rangle$. Therefore the expectation value of some observable $\hat{A}$ will not be with respect to the ideal state

$$\langle \hat{A} \rangle_{\text{ideal}} = \text{Tr}(\hat{A} \, |\psi_{\text{ideal}}\rangle\langle\psi_{\text{ideal}}|), \tag{6.1}$$

but with respect to the mixed state

$$\langle \hat{A} \rangle_{\text{noisy}} = \text{Tr}(\hat{A}\hat{\rho}). \tag{6.2}$$

This results in a bias in the expectation value of an observable $\hat{A}$, given by $\langle \hat{A} \rangle_{\text{noisy}} - \langle \hat{A} \rangle_{\text{ideal}}$, see Figure 6.1a. The primary objective of quantum error-mitigation is to minimize this bias.

To fix the ideas, consider the output state from a N-qubit quantum circuit. This output state can be decomposed via spectral decomposition as

$$\hat{\rho} = \sum_{k=1}^{d} \lambda_k \, |\psi_k\rangle\langle\psi_k| \, , \tag{6.3}$$

where $d = 2^N$ and λ_k is the probability that the system is found in the state $|\psi_k\rangle$ when measured in the eigenbasis of $\hat{\rho}$. We can assume for convenience that the probabilities λ_k are listed in descending order $\lambda_1 > \lambda_2 \ldots > \lambda_d$. Now, if the dominant eigenvector of the density matrix, i.e., the state with largest eigenvalue, is close to the ideal state, then $\langle\psi_1| \, \hat{A} \, |\psi_1\rangle$ would be a very good approximation of the ideal expectation value[1]. However, erroneous eigenvectors $|\psi_{k\neq 1}\rangle$ contribute to the bias of the estimated expectation value. Virtual distillation aims to suppress these contributions via the following novel idea. Consider the scenario where we generate M instances (copies) of the state $\hat{\rho}$. The most probable outcome during this state generation is that we obtain the dominant eigenvector of the state: with a likelihood of λ^M, the state (right after the state generation) is $|\psi_1, \psi_1, \ldots, \psi_1\rangle$. Calculating the expectation value of an observable $\hat{A}$ on the first subsystem yields the sought-after result

$$\begin{aligned} \langle\psi_1, \ldots, \psi_1, \psi_1| \, \hat{A}_{\mathbf{1}} \, |\psi_1, \ldots, \psi_1, \psi_1\rangle &= \langle\psi_1| \, \hat{A} \, |\psi_1\rangle \, \langle\psi_1|\psi_1\rangle \ldots \langle\psi_1|\psi_1\rangle \\ &= \langle\psi_1| \, \hat{A} \, |\psi_1\rangle \, . \end{aligned} \tag{6.4}$$

[1] This condition is satisfied when the noise in the quantum circuit maps the ideal states to states that are orthogonal to it; otherwise the dominant eigenvector will drift and limit the error suppression efficiency [101, 102]. In general, for a multi-qubit state, both single-qubit depolarizing or dephasing errors can cause the dominant eigenvector to drift. However, it has been shown that the severity of this drift is exponentially smaller than the incoherent decay of fidelity [102].

Here the bold subscript indicates that the operator $\hat{A}$ acts on the first subsystem. The next most probable event is that one of the registers, let us say the first register, is detected in the orthogonal erroneous eigenvector $|\psi_{k\neq 1}\rangle$; a measurement then yields the error term

$$\begin{aligned}\langle\psi_k,\dots,\psi_1,\psi_1|\,\hat{A}_{\mathbf{1}}\,|\psi_k,\dots,\psi_1,\psi_1\rangle &= \langle\psi_k|\,\hat{A}\,|\psi_k\rangle\,\langle\psi_1|\psi_1\rangle\dots\langle\psi_1|\psi_1\rangle \\ &= \langle\psi_k|\,\hat{A}\,|\psi_k\rangle\,. \end{aligned} \tag{6.5}$$

However, if we instead measure the expectation value of the product $\hat{A}\,\textsc{swap}_{1M}$, where $\textsc{swap}_{1M}$ interchanges subsystem 1 and M, we get

$$\langle\psi_k,\dots,\psi_1,\psi_1|\,\hat{A}_{\mathbf{1}}\,|\psi_1,\dots,\psi_1,\psi_k\rangle = \langle\psi_k|\,\hat{A}\,|\psi_1\rangle\,\langle\psi_1|\psi_1\rangle\dots\langle\psi_1|\psi_k\rangle = 0. \tag{6.6}$$

In this case, the SWAP operator altered the sequence of the subsystems, and the outcome is 0 due to the orthogonality of the eigenvectors. This concept can easily be extended to the situation where all subsystems are swapped, allowing only permutation-symmetric states to contribute to the measurement of expectation values.

To illustrate this, consider $\hat{S}^{(M)}$, an operator that acts on M subsystems and cyclically shifts these subsystems and whose operation can be represented as follows

$$\hat{S}^{(M)}\,|\psi_1,\psi_2,\dots,\psi_M\rangle = |\psi_2,\psi_3,\dots,\psi_1\rangle\,. \tag{6.7}$$

By measuring the following expression

$$\begin{aligned}&\mathrm{Tr}[\hat{A}_{\mathbf{1}}\hat{S}^{(M)}\hat{\rho}^{\otimes M}] \\ &= \mathrm{Tr}[\hat{A}_{\mathbf{1}}\sum_{k_1\dots k_M}\lambda_{k_1}\dots\lambda_{k_M}\,|\psi_{k_2},\psi_{k_3},\dots,\psi_{k_1}\rangle\langle\psi_{k_1},\psi_{k_2}\dots,\psi_{k_M}|] \\ &= \sum_{k_1\dots k_M}\lambda_{k_1}\dots\lambda_{k_M}\,\mathrm{Tr}[\hat{A}\,|\psi_{k_2}\rangle\langle\psi_{k_1}|]\,\mathrm{Tr}[|\psi_{k_3}\rangle\langle\psi_{k_2}|]\dots\mathrm{Tr}[|\psi_{k_1}\rangle\langle\psi_{k_M}|] \\ &= \sum_{k_1\dots k_M}\lambda_{k_1}\dots\lambda_{k_M}\,\delta_{k_3,k_2}\dots\delta_{k_1,k_M}\,\mathrm{Tr}[\hat{A}\,|\psi_{k_2}\rangle\langle\psi_{k_1}|] \\ &= \mathrm{Tr}[\hat{A}\sum_{k_1}\lambda_{k_1}^M\,|\psi_{k_1}\rangle\langle\psi_{k_1}|] = \mathrm{Tr}[\hat{A}\hat{\rho}^M], \end{aligned} \tag{6.8}$$

we obtain the expectation value of $\hat{A}$ with respect to $\hat{\rho}^M$, which is not a legitimate state since its trace is not 1. Therefore we have to normalize this expression by dividing it by $\mathrm{Tr}[\hat{S}^{(M)}\hat{\rho}^{\otimes M}] = \mathrm{Tr}[\hat{\rho}^M]$, yielding the final result

$$\langle\hat{A}\rangle_{\text{mitigated}} = \frac{\mathrm{Tr}(\hat{A}_{\mathbf{1}}\hat{S}^{(M)}\hat{\rho}^{\otimes M})}{\mathrm{Tr}(\hat{S}^{(M)}\hat{\rho}^{\otimes M})} = \frac{\mathrm{Tr}(\hat{A}\sum_k\lambda_k^M\,|\psi_k\rangle\langle\psi_k|)}{\sum_k\lambda_k^M}. \tag{6.9}$$

From this, we see that when $|\psi_1\rangle$ corresponds to the output of the ideal (noise-free) quantum circuit, $\langle\hat{A}\rangle_{\text{mitigated}}$ approaches the ideal expectation value at an exponential rate with respect to M. Note that this is all done without actually preparing the state $\hat{\rho}^M/\,\mathrm{Tr}(\hat{\rho}^M)$, hence the name "virtual".

The process of measuring the observable $\hat{S}^{(M)}$ in the denominator of Eq. (6.9) is similar to the Hadamard test [103]. This procedure starts with the preparation of M collective copies of the state $\hat{\rho}$, along with an auxiliary qubit initialized in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. Following this initialization, the cyclic shift operator $\hat{S}^{(M)}$ is applied to the M copies of $\hat{\rho}$. This operation is conditioned on the auxiliary qubit being in the state $|1\rangle$. The final step involves measuring the auxiliary qubit in the X-basis (also known as the plus/minus-basis), as depicted in Figure 6.1b. The measurement of the auxiliary qubit results in a phase-kickback effect, whereby the expectation value obtained from measuring the auxiliary qubit equates to $\mathrm{Tr}(\hat{S}^{(M)}\hat{\rho}^{\otimes M})$. This can quite easily be seen mathematically. The state right before the measurement of the auxiliary qubit is

$$\frac{1}{2}\Big(\hat{I} \otimes \hat{\rho}^{\otimes M} + |0\rangle\langle 1| \otimes \hat{\rho}^{\otimes M}\hat{S}^{(M)} + |1\rangle\langle 0| \otimes \hat{S}^{(M)}\hat{\rho}^{\otimes M}\Big), \tag{6.10}$$

where we have used the fact that $\hat{S}^{(M)}\hat{\rho}^{\otimes M}\hat{S}^{(M)} = \hat{\rho}^{\otimes M}$. By measuring the expectation value of the $\hat{\sigma}_x$ operator of the auxiliary qubit, we get the desired result

$$\begin{aligned}\frac{1}{2}\,&\mathrm{Tr}[\hat{\sigma}_x \otimes \hat{\rho}^{\otimes M} + |1\rangle\langle 1| \otimes \hat{\rho}^{\otimes M}\hat{S}^{(M)} + |0\rangle\langle 0| \otimes \hat{S}^{(M)}\hat{\rho}^{\otimes M}] \\ &= \frac{1}{2}\,\mathrm{Tr}[\hat{I} \otimes \hat{S}^{(M)}\hat{\rho}^{\otimes M}] = \mathrm{Tr}[\hat{S}^{(M)}\hat{\rho}^{\otimes M}]. \end{aligned} \tag{6.11}$$

Moreover, to measure the numerator of Eq. (6.9), it is beneficial to consider the symmetrized version of the operator $\hat{A}$ for simplifying the analysis. The symmetrized operator $\hat{A}$ is defined as follows:

$$\hat{A}^{(M)} = \frac{1}{M}\sum_{\mathrm{i}=1}^{M} \hat{A}_{\mathrm{i}}. \tag{6.12}$$

Given that the symmetrized observable $\hat{A}^{(M)}$ commutes with $\hat{S}^{(M)}$, it is possible to measure the product $\hat{A}^{(M)}\hat{S}^{(M)}$. This is achieved by first measuring $\hat{S}^{(M)}$ using the Hadamard test described above, followed by measuring $\hat{A}^{(M)}$ on the subsystems.

However, a critical question still lingers: how do we decompose $\hat{S}^{(M)}$, and more importantly, could errors that occur during the implementation of $\hat{S}^{(M)}$ potentially undermine the effectiveness of this error-mitigation protocol? These are not trivial concerns, as the protocol's success hinges on accurately implementing $\hat{S}^{(M)}$. Any errors introduced during this process could potentially exacerbate the very issues the protocol aims to mitigate. In the following section, we will delve into these questions, exploring the decomposition of $\hat{S}^{(M)}$ and examining the potential impact of errors on the efficacy of the error-mitigation protocol.

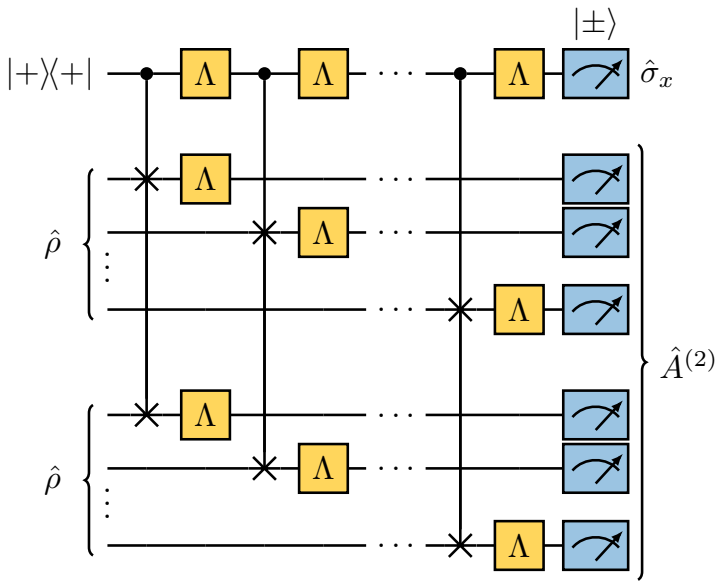

Figure 6.2: The virtual distillation circuit for two copies ($M = 2$). The circuit includes single-qubit errors, denoted by Λ, which are explicitly depicted in the circuit layout.

6.1.1 Noise in the virtual distillation circuit

We now turn our attention to the practical implementation of virtual distillation. Specifically, we will focus on the case where $M = 2$, which will simplify the decomposition of the cyclic shift operator $\hat{S}^{(2)}$.

For $M = 2$, the cyclic shift operator $\hat{S}^{(2)}$ can be factorized into a tensor product of N SWAP-gates, and its controlled version can be factorized into a product of N CSWAP-gates. To model the impact of errors during the protocol, we assume that a single-qubit noise channel is applied after every gate to the qubits involved, as depicted in Figure 6.2.

We consider two types of single-qubit noise channels. The first is the depolarizing channel, which describes a process where information is completely lost with a probability ϵ, and is given by

$$\Lambda_{\text{dep}}(\hat{\rho}) = (1-\epsilon)\hat{\rho} + \frac{\epsilon}{3}(\hat{\sigma}_x\hat{\rho}\hat{\sigma}_x + \hat{\sigma}_y\hat{\rho}\hat{\sigma}_y + \hat{\sigma}_z\hat{\rho}\hat{\sigma}_z), \tag{6.13}$$

where $\{\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}$ are the Pauli operators and ϵ is the error probability. The second error channel is the pure-dephasing channel, a biased noise channel. This channel describes the loss of phase information with a probability ϵ. It describes the loss of phase information with a probability ϵ,

$$\Lambda_{\text{Z}}(\hat{\rho}) = (1-\epsilon)\hat{\rho} + \epsilon\hat{\sigma}_z\hat{\rho}\hat{\sigma}_z. \tag{6.14}$$

Of course, we could have chosen an error channel with biased X- or Y-noise. However, we can always redefine the computational basis states on the Bloch sphere and call all of this Z-biased noise.

In Paper D, we derive an expression for the mitigated expectation value in the presence of errors in the virtual distillation circuit and find that it is given by

$$\langle \hat{A} \rangle^{\Lambda}_{\text{mitigated}} = \frac{\text{Tr}(\bar{\Lambda}(\hat{A})\hat{\rho}^2)}{\text{Tr}(\hat{\rho}^2)}, \tag{6.15}$$

where $\bar{\Lambda} = \Lambda \otimes \ldots \otimes \Lambda$ is a tensor product of N single-qubit error channels $\Lambda \in \{\Lambda_{\text{dep}}, \Lambda_{\text{Z}}\}$. From Eq. (6.15), we see that the influence of errors on the mitigated expectation value depends on the observable $\hat{A}$. Since a general observable on N qubits can be expressed as a sum of N-qubit Pauli strings from the set $\{\hat{I}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}^{\otimes N}$, it is sufficient to consider $\hat{A} \in \{\hat{I}, \hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z\}^{\otimes N}$. In this case, each Pauli matrix transforms as

$$\Lambda_{\text{dep}}(\hat{\sigma}_k) = (1 - 4\epsilon/3)\hat{\sigma}_k, \quad \text{for} \quad k \in \{x, y, z\}, \tag{6.16}$$

$$\Lambda_{\text{Z}}(\hat{\sigma}_k) = \begin{cases} (1 - 2\epsilon)\hat{\sigma}_k, & \text{for } k \in \{x, y\}, \\ \hat{\sigma}_k, & \text{for } k \in \{z\}, \end{cases} \tag{6.17}$$

Therefore the mitigated expectation value for the two types of noise can be expressed as

$$\langle \hat{A} \rangle^{\Lambda_{\text{dep}}}_{\text{mitigated}} = \left(1 - \frac{4}{3}\epsilon\right)^n \frac{\text{Tr}(\hat{A}\hat{\rho}^2)}{\text{Tr}(\hat{\rho}^2)}, \tag{6.18}$$

$$\langle \hat{A} \rangle^{\Lambda_{\text{Z}}}_{\text{mitigated}} = (1 - 2\epsilon)^{n'} \frac{\text{Tr}(\hat{A}\hat{\rho}^2)}{\text{Tr}(\hat{\rho}^2)}, \tag{6.19}$$

where n is the number of Pauli operators in $\hat{A}$ and n' is the number of $\{\hat{\sigma}_x, \hat{\sigma}_y\}$ operators in $\hat{A}$. As an example, if $\hat{A} = \hat{I} \otimes \hat{\sigma}_x \otimes \hat{\sigma}_z$, then $n = 2$ and $n' = 1$. From this, we can see that errors in the virtual distillation circuit attenuate the expectation values, leading to an increase in bias with respect to the ideal expectation value. However, it is important to note that when the mitigated expectation value of an observable $\hat{A}$ is a tensor product of Pauli Z-operators, meaning $n' = 0$, the mitigated expectation value is completely immune to pure dephasing in the distillation circuit, and dephasing noise will therefore not introduce any bias. Moreover, performing local Clifford transformations, e.g., the Hadamard, on the state $\hat{\rho}$ before sending it into the virtual distillation circuit makes it possible to measure any Pauli observable without attenuating the mitigated expectation value given dephasing noise.

In conclusion, our findings suggest that implementing the virtual distillation protocol in a system that is strongly biased towards dephasing noise will not introduce any additional bias in the estimation of the expectation value. Moreover, this also avoids the need to rely on other error mitigation techniques in addition to virtual distillation for mitigating errors.

Chapter 7

Conclusion and outlook

7.1 Summary

This thesis presents a study of the QAOA, a hybrid classical-quantum algorithm that leverages the strengths of both quantum processors and classical optimizers. The primary objective of this algorithm is to iteratively optimize a variational state to approximate the ground state of a cost Hamiltonian that encodes a combinatorial optimization problem. The focus of this thesis has been the application of QAOA to the Exact Cover problem, a simplified model of the tail-assignment problem – a problem that involves assigning aircraft to specific flights. This thesis has also delved into the practical implementation of solving a simple optimization problem on the Chalmers quantum computer, providing empirical evidence of QAOA's functionality. Moreover, this thesis also presented a study of the implementation of QAOA using qubits with favorable noise properties, specifically cat qubits. Finally, this thesis explored virtual distillation, an error-mitigation protocol, and assessed its performance under various types of errors.

In Chapter 2, as further detailed in Paper C, we demonstrate numerical simulations of the simultaneous execution of two CZ-gates on existing quantum hardware resulting in new exotic three-qubit gates.

In Chapter 4, as outlined in Paper A, we investigate the application of QAOA to the Exact Cover problem, an abstraction of the tail-assignment problem. The algorithm was applied to instances corresponding to 8, 15, and 25 qubits, allowing for the calculation of the success probability, a metric indicating the probability of finding the optimal solution in a single-shot measurement. This success probability is shown to increase with the iteration parameter p, which is directly associated with the depth of the quantum circuit.

In Chapter 5, as illustrated in Paper B, we present the implementation and execution of a toy instance of the Exact Cover problem on the Chalmers quantum computer. The experimental results closely matched the theory, providing

the first experimental verification of the impact of the algorithmic depth on the success probability in QAOA. This is followed by a theoretical investigation of the implementation of QAOA with cat qubits using two-photon pumped Kerr non-linear resonators, as detailed in Paper E. Simulation results suggest that cat qubits yield a higher approximation ratio than discrete-variable qubits based on two-level systems when applied to 8-qubit instances of the MaxCut problem in the presence of single-photon losses. This suggests a potential advantage of using cat qubits in QAOA implementations.

Finally, in Chapter 6, as outlined in Paper D, we investigate the effect of noise in a protocol known as virtual distillation, an error mitigation method aimed at reducing noise in the estimation of expectation values. Analytical evidence corroborate the robustness of this protocol against dephasing noise. The findings substantiate that virtual distillation enhances the performance of QAOA amidst dephasing noise, thereby offering a potential avenue for effective error mitigation in such settings.

7.2 Outlook

Looking ahead, there are several promising avenues for future research that build upon the findings of this thesis:

1. **Benchmarking quantum devices with larger optimization problems:** One of the key challenges in quantum computing is scaling up the size of the problems that can be solved. As quantum hardware continues to improve, running larger tasks on quantum devices will be possible. A crucial aspect of this progression will be benchmarking the time to solution with a classical optimizer. This will provide a more comprehensive understanding of the computational advantages offered by quantum devices. It will also allow for a more direct comparison between quantum and classical computing capabilities. This benchmarking process will be instrumental in identifying the areas where quantum computing can provide the most significant benefits and in guiding the development of future quantum algorithms and hardware.

2. **QAOA experiment on cat qubits:** Cat qubits have shown promise regarding their resistance to certain types of noise. A logical next step would be to run a proof-of-principle QAOA experiment using cat qubits. This experiment could provide valuable insights into the practical advantages and challenges of using cat qubits in QAOA and other quantum algorithms. It could also help further to elucidate the behavior of cat qubits under real-world conditions and contribute to the ongoing development of quantum error correction and error-mitigation techniques.

3. **Implementation of virtual distillation using the three-qubit CCZS-gate:** The virtual distillation protocol has shown potential for mitigating the effects of noise in quantum computations. An interesting direction for

future research would be to run virtual distillation using the three-qubit CCZS-gate that we presented in Section 2.2.1. As outlined in Section 6.1.1 the virtual distillation can be implemented using CSWAP-gates. However, compiling this circuit using the CCZS-gate is another possibility that we would like to explore further.

Bibliography

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).

[2] C. P. Williams, *Explorations in quantum computing*, 2nd ed., Texts in computer science (Springer, London, 2011).

[3] P. Benioff, "The computer as a physical system", J. Stat. Phys. **22**, 563–591 (1980).

[4] Y. Manin, *Vychislimoe i nevychislimoe (computable and noncomputable)* (Soviet Radio, 1980) pp. 13–15, in Russian.

[5] R. P. Feynman, "Simulating physics with computers", Int J Theor Phys **21**, 467–488 (1982).

[6] D. Deutsch, "Quantum theory, the church-turing principle and the universal quantum computer", Proc. R. Soc. Lond. A **400**, 97–117 (1985).

[7] D. Deutsch, "Quantum computational networks", Proc. R. Soc. Lond. A **425**, 73–90 (1989).

[8] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer", SIAM J. Comput. **26**, 1484–1509 (1997).

[9] R. L. Rivest, A. Shamir and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems", Commun. ACM **21**, 120–126 (1978).

[10] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory", Phys. Rev. A **52**, R2493–R2496 (1995).

[11] D. Gottesman, "An introduction to quantum error correction and fault-tolerant quantum computation", arXiv:0904.2557 [quant-ph] (2009).

[12] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, 2013).

[13] S. M. Girvin, "Introduction to quantum error correction and fault tolerance", SciPost Phys. Lect. Notes , 70 (2023).

[14] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush, *et al.*, "Suppressing quantum errors by scaling a surface code logical qubit", Nature **614**, 676–681 (2023).

[15] C. Gidney and M. Ekerå, "How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits", Quantum **5**, 433 (2021).

[16] J. Preskill, "Quantum computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[17] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, *et al.*, "Quantum supremacy using a programmable superconducting processor", Nature **574**, 505–510 (2019).

[18] F. Pan and P. Zhang, "Simulation of quantum circuits using the big-batch tensor network method", Phys. Rev. Lett. **128**, 030501 (2022).

[19] F. Pan, K. Chen and P. Zhang, "Solving the sampling problem of the sycamore quantum circuits", Phys. Rev. Lett. **129**, 090502 (2022).

[20] E. Farhi, J. Goldstone and S. Gutmann, "A quantum approximate optimization algorithm", arXiv:1411.4028 [quant-ph] (2014).

[21] P. Vikstål, M. Grönkvist, M. Svensson, M. Andersson, G. Johansson and G. Ferrini, "Applying the Quantum Approximate Optimization Algorithm to the Tail-Assignment Problem", Phys. Rev. Applied **14**, 034009 (2020).

[22] M. Hodson, B. Ruck, H. Ong, D. Garvin and S. Dulman, "Portfolio rebalancing experiments using the quantum alternating operator ansatz", arxiv:1911.05296 [quant-ph] (2019).

[23] D. Fitzek, T. Ghandriz, L. Laine, M. Granath and A. F. Kockum, "Applying quantum approximate optimization to the heterogeneous vehicle routing problem", arxiv:2110.06799 [quant-ph] (2021).

[24] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson and W. D. Oliver, "A quantum engineer's guide to superconducting qubits", Applied Physics Reviews **6**, 021318 (2019).

[25] A. Blais, A. L. Grimsmo, S. M. Girvin and A. Wallraff, "Circuit quantum electrodynamics", Rev. Mod. Phys. **93**, 025005 (2021).

[26] S. Lloyd and S. L. Braunstein, "Quantum computation over continuous variables", Phys. Rev. Lett. **82**, 1784–1787 (1999).

[27] S. L. Braunstein and P. van Loock, "Quantum information with continuous variables", Rev. Mod. Phys. **77**, 513–577 (2005).

[28] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, *et al.*, "Demonstration of two-qubit algorithms with a superconducting quantum processor", Nature **460**, 240–244 (2009).

[29] R. Barends, C. M. Quintana, A. G. Petukhov, Y. Chen, D. Kafri, K. Kechedzhi, R. Collins, O. Naaman, S. Boixo, F. Arute, *et al.*, "Diabatic gates for frequency-tunable superconducting qubits", Phys. Rev. Lett. **123**, 210501 (2019).

[30] S. Li, A. D. Castellano, S. Wang, Y. Wu, M. Gong, Z. Yan, H. Rong, H. Deng, C. Zha, C. Guo, *et al.*, "Realisation of high-fidelity nonadiabatic CZ gates with superconducting qubits", npj Quantum Inf **5**, 84 (2019).

[31] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin and H. Weinfurter, "Elementary gates for quantum computation", Phys. Rev. A **52**, 3457–3467 (1995).

[32] D. M. Greenberger, M. A. Horne and A. Zeilinger, "Going beyond bell's theorem", arXiv:0712.0921 [quant-ph] (2007).

[33] J. A. Smolin and D. P. DiVincenzo, "Five two-bit quantum gates are sufficient to implement the quantum fredkin gate", Phys. Rev. A **53**, 2855 (1996).

[34] N. Yu, R. Duan and M. Ying, "Five two-qubit gates are necessary for implementing the toffoli gate", Phys. Rev. A **88**, 010304 (2013).

[35] M. A. Nielsen, "A simple formula for the average gate fidelity of a quantum dynamical operation", Physics Letters A **303**, 249–252 (2002).

[36] T. Thorbeck, A. Eddins, I. Lauer, D. T. McClure and M. Carroll, "Two-level-system dynamics in a superconducting qubit due to background ionizing radiation", PRX Quantum **4**, 020356 (2023).

[37] H.-P. Breuer and F. Petruccione, *The theory of open quantum systems*, repr ed. (Clarendon Press, 2010).

[38] D. F. Walls and G. J. Milburn, *Quantum optics*, 2nd ed. (Springer, 2008).

[39] G. Ripoll and J. José, *Quantum Information and Quantum Optics with Superconducting Circuits*, 1st ed. (Cambridge University Press, 2022).

[40] D. Greenbaum, "Introduction to quantum gate set tomography", arXiv:1509.02921 [quant-ph] (2015).

[41] E. Farhi, J. Goldstone and S. Gutmann, "A quantum approximate optimization algorithm applied to a bounded occurrence constraint problem", arXiv:1412.6062 [quant-ph] (2015).

[42] S. Lloyd, "Quantum approximate optimization is computationally universal", arXiv:1703.06199 [quant-ph] (2018).

[43] Z. Jiang, E. G. Rieffel and Z. Wang, "Near-optimal quantum circuit for Grover's unstructured search using a transverse field", Phys. Rev. A **95**, 062317 (2017).

[44] L. Zhou, S.-T. Wang, S. Choi, H. Pichler and M. D. Lukin, "Quantum Approximate Optimization Algorithm: Performance, Mechanism, and Implementation on Near-Term Devices", Phys. Rev. X **10**, 021067 (2020).

[45] S. Boulebnane and A. Montanaro, "Solving boolean satisfiability problems with the quantum approximate optimization algorithm", arxiv:2208.06909 [quant-ph] (2022).

[46] P. Torta, G. B. Mbeng, C. Baldassi, R. Zecchina and G. E. Santoro, "Quantum approximate optimization algorithm applied to the binary perceptron", Phys. Rev. B **107**, 094202 (2023).

[47] E. Farhi, J. Goldstone, S. Gutmann and M. Sipser, "Quantum Computation by Adiabatic Evolution", arXiv:quant-ph/0001106 (2000).

[48] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren and D. Preda, "A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem", Science **292**, 472–475 (2001).

[49] E. Farhi, J. Goldstone and S. Gutmann, "Quantum adiabatic evolution algorithms with different paths", arxiv:0208135 [quant-ph] (2002).

[50] E. Crosson, E. Farhi, C. Y.-Y. Lin, H.-H. Lin and P. Shor, "Different strategies for optimization using the quantum adiabatic algorithm", arxiv:1401.7320 [quant-ph] (2014).

[51] Y. Sun, J.-Y. Zhang, M. S. Byrd and L.-A. Wu, "Adiabatic Quantum Simulation Using Trotterization", arXiv:1805.11568 [quant-ph] (2018).

[52] L. Li, M. Fan, M. Coram, P. Riley and S. Leichenauer, "Quantum optimization with a novel gibbs objective function and ansatz architecture search", Phys. Rev. Res. **2**, 023074 (2020).

[53] D. Wecker, M. B. Hastings and M. Troyer, "Training a quantum optimizer", Phys. Rev. A **94**, 022309 (2016).

[54] R. Shaydulin, I. Safro and J. Larson, in *2019 IEEE High Performance Extreme Computing Conference (HPEC)* (2019) pp. 1–8.

[55] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli and R. Biswas, "From the quantum approximate optimization algorithm to a quantum alternating operator ansatz", Algorithms **12**, 34 (2019).

[56] R. Tate, J. Moondra, B. Gard, G. Mohler and S. Gupta, "Warm-started QAOA with custom mixers provably converges and computationally beats goemans-williamson's max-cut at low circuit depths", arxiv:2112.11354 [quant-ph] (2022).

[57] L. Zhu, H. L. Tang, G. S. Barron, F. A. Calderon-Vargas, N. J. Mayhall, E. Barnes and S. E. Economou, "Adaptive quantum approximate optimization algorithm for solving combinatorial problems on a quantum computer", Phys. Rev. Res. **4**, 033029 (2022).

[58] G. G. Guerreschi and M. Smelyanskiy, "Practical optimization for hybrid quantum-classical algorithms", arXiv:1701.01450 [quant-ph] (2017).

[59] W. Lavrijsen, A. Tudor, J. Müller, C. Iancu and W. de Jong, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (2020) pp. 267–277.

[60] K. M. Nakanishi, K. Fujii and S. Todo, "Sequential minimal optimization for quantum-classical hybrid algorithms", Phys. Rev. Research **2**, 043158 (2020).

[61] D. Fitzek, R. S. Jonsson, W. Dobrautz and C. Schäfer, "Optimizing variational quantum algorithms with qBang: Efficiently interweaving metric and momentum to tackle flat energy landscapes", arxiv:2304.13882 [quant-ph] (2023).

[62] J. S. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. S. Fried, S. Hong, *et al.*, "Unsupervised Machine Learning on a Hybrid Quantum Computer", arXiv:1712.05771 [quant-ph] (2017).

[63] A. Bengtsson, P. Vikstål, C. Warren, M. Svensson, X. Gu, A. F. Kockum, P. Krantz, C. Križan, D. Shiri, I.-M. Svensson, *et al.*, "Improved Success Probability with Greater Circuit Depth for the Quantum Approximate Optimization Algorithm", Phys. Rev. Applied **14**, 034010 (2020).

[64] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo, *et al.*, "Quantum approximate optimization of non-planar graph problems on a planar superconducting processor", Nat. Phys. **17**, 332–336 (2021).

[65] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams and N. de Freitas, "Taking the Human Out of the Loop: A Review of Bayesian Optimization", Proceedings of the IEEE **104**, 148–175 (2016).

[66] J. A. Nelder and R. Mead, "A simplex method for function minimization", The Computer Journal **7**, 308–313 (1965).

[67] G. E. Crooks, "Performance of the Quantum Approximate Optimization Algorithm on the Maximum Cut Problem", arXiv:1811.08419 [quant-ph] (2018).

[68] E. Farhi, D. Gamarnik and S. Gutmann, "The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: Worst Case Examples", arXiv:2005.08747 [quant-ph] (2020).

[69] J. Wurtz and P. Love, "Maxcut quantum approximate optimization algorithm performance guarantees for p>1", Phys. Rev. A **103**, 042612 (2021).

[70] M. X. Goemans and D. P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming", Journal of the ACM **42**, 1115–1145 (1995).

[71] M. Grönkvist and J. Kjerrström, in *Operations Research Proceedings 2004*, Vol. 2004, edited by H. Fleuren, D. den Hertog and P. Kort (Springer Berlin Heidelberg, 2005) pp. 166–173.

[72] M. Grönkvist, *The Tail Assignment Problem*, Ph.D. thesis, Chalmers University of Technology and Göteborg University (2005).

[73] G. B. Dantzig and P. Wolfe, "The decomposition algorithm for linear programs", Econometrica **29**, 767–778 (1961).

[74] X. Qiang, X. Zhou, J. Wang, C. M. Wilkes, T. Loke, S. O'Gara, L. Kling, G. D. Marshall, R. Santagati, T. C. Ralph, *et al.*, "Large-scale silicon quantum photonics implementing arbitrary two-qubit processing", Nature Photonics **12**, 534–539 (2018).

[75] G. Pagano, A. Bapat, P. Becker, K. S. Collins, A. De, P. W. Hess, H. B. Kaplan, A. Kyprianidis, W. L. Tan, C. Baldwin, *et al.*, "Quantum approximate optimization of the long-range ising model with a trapped-ion quantum simulator", PNAS **117**, 25396–25401 (2020).

[76] N. Lacroix, C. Hellings, C. K. Andersen, A. Di Paolo, A. Remm, S. Lazar, S. Krinner, G. J. Norris, M. Gabureac, J. Heinsoo, *et al.*, "Improving the performance of deep quantum optimization algorithms with continuous gate sets", PRX Quantum **1**, 110304 (2020).

[77] B. O'Gorman, W. J. Huggins, E. G. Rieffel and K. B. Whaley, "Generalized swap networks for near-term quantum computing", arxiv:1905.05118 [quant-ph] (2019).

[78] A. Joshi, K. Noh and Y. Y. Gao, "Quantum information processing with bosonic qubits in circuit QED", arXiv:2008.13471 [quant-ph] (2020).

[79] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. E. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais, *et al.*, "Bias-preserving gates with stabilized cat qubits", Sci Adv **6**, eaay5901 (2020).

[80] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, *et al.*, "Real-time quantum error correction beyond break-even", Nature **616**, 50–55 (2023-04-01).

[81] P. Vikstål, G. Ferrini and S. Puri, "Study of noise in virtual distillation circuits for quantum error mitigation", arxiv:2210.15317 [quant-ph] (2022).

[82] H. Goto, "Universal quantum computation with a nonlinear oscillator network", Phys. Rev. A **93**, 050301 (2016).

[83] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar and M. H. Devoret, "Stabilization and operation of a Kerr-cat qubit", Nature **584**, 205–209 (2020).

[84] S. Puri, S. Boutin and A. Blais, "Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving", npj Quantum Inf **3**, 18 (2017).

[85] S. Puri, C. K. Andersen, A. L. Grimsmo and A. Blais, "Quantum annealing with all-to-all connected nonlinear oscillators", Nat Commun **8**, 15785 (2017).

[86] S. Puri, A. Grimm, P. Campagne-Ibarcq, A. Eickbusch, K. Noh, G. Roberts, L. Jiang, M. Mirrahimi, M. H. Devoret and S. M. Girvin, "Stabilized cat in driven nonlinear cavity", Phys. Rev. X **9**, 041009 (2019).

[87] M. Kudra, M. Kervinen, I. Strandberg, S. Ahmed, M. Scigliuzzo, A. Osman, D. P. Lozano, M. O. Tholén, R. Borgani, D. B. Haviland, *et al.*, "Robust preparation of wigner-negative states with optimized snap-displacement sequences", PRX Quantum **3**, 030301 (2022).

[88] Y. Li and S. C. Benjamin, "Efficient variational quantum simulator incorporating active error minimization", Phys. Rev. X **7**, 021050 (2017).

[89] K. Temme, S. Bravyi and J. M. Gambetta, "Error mitigation for short-depth quantum circuits", Phys. Rev. Lett. **119**, 180509 (2017).

[90] S. Endo, S. C. Benjamin and Y. Li, "Practical quantum error mitigation for near-future applications", Phys. Rev. X **8**, 031027 (2018).

[91] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow and J. M. Gambetta, "Error mitigation extends the computational reach of a noisy quantum processor", Nature **567**, 491–495 (2019).

[92] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin and Y. Li, "Learning-based quantum error mitigation", PRX Quantum **2**, 040330 (2021).

[93] P. Czarnik, A. Arrasmith, P. J. Coles and L. Cincio, "Error mitigation with clifford quantum-circuit data", Quantum **5**, 592 (2021).

[94] X. Bonet-Monroig, R. Sagastizabal, M. Singh and T. E. O'Brien, "Low-cost error mitigation by symmetry verification", Phys. Rev. A **98**, 062339 (2018).

[95] S. McArdle, X. Yuan and S. Benjamin, "Error-mitigated digital quantum simulation", Phys. Rev. Lett. **122**, 180501 (2019).

[96] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, *et al.*, "Experimental error mitigation via symmetry verification in a variational quantum eigensolver", Phys. Rev. A **100**, 010302(R) (2019).

[97] Google AI Quantum and Collaborators, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, *et al.*, "Hartree-fock on a superconducting qubit quantum computer", Science **369**, 1084–1089 (2020).

[98] W. J. Huggins, J. R. McClean, N. C. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley and R. Babbush, "Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers", npj Quantum Inf **7**, 1–9 (2021).

[99] S. Wang, P. Czarnik, A. Arrasmith, M. Cerezo, L. Cincio and P. J. Coles, "Can error mitigation improve trainability of noisy variational quantum algorithms?", arXiv:2109.01051 [quant-ph] (2021).

[100] T. E. O'Brien, S. Polla, N. C. Rubin, W. J. Huggins, S. McArdle, S. Boixo, J. R. McClean and R. Babbush, "Error mitigation via verified phase estimation", PRX Quantum **2**, 020317 (2021).

[101] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush and J. R. McClean, "Virtual distillation for quantum error mitigation", Phys. Rev. X **11**, 041036 (2021).

[102] B. Koczor, "Exponential error suppression for near-term quantum devices", Phys. Rev. X **11**, 031057 (2021).

[103] D. Aharonov, V. Jones and Z. Landau, "A polynomial quantum algorithm for approximating the jones polynomial", Algorithmica **55**, 395–421 (2009).