

Quantum Circuit Optimization in the ZX-Calculus



Ryan Krueger
Christ Church College
University of Oxford

A thesis submitted for the degree of
Master of Science in Mathematical Sciences

Trinity 2021

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod
tempor incididunt ut labore et dolore magna aliqua. Ut enim ad
minim veniam, quis nostrud exercitation ullamco laboris nisi ut
aliquip ex ea commodo consequat.

Acknowledgements

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Contents

1	Introduction	1
2	Background	3
2.1	Qubits	3
2.2	Quantum Circuits	4
2.3	ZX-Calculus	6
2.4	Circuit Optimization in the ZX-Calculus	8
3	Methods	12
3.1	Congruences	12
3.2	Search Procedures	12
3.2.1	Simulated Annealing	12
3.2.2	Genetic Algorithms	12
3.2.3	Target Functions	13
3.3	Circuit Optimization Strategies	13
4	Results	14
5	Application: Quantum Chemistry	15
6	Conclusion	16
	Bibliography	17

Chapter 1

Introduction

Quantum computing arose from the idea that the “bugs” of quantum mechanics (e.g., superposition) may serve as useful features for information processing. Since this framing was posed in the late 1970’s, a rich theory of quantum information has developed that treats a *qubit*, a superposition of two classical bits, as the basic unit of information. Quantum information science is an exciting field of work as many quantum algorithms have been discovered that offer significant speedups from their classical counterparts (e.g., integer factorization). However, like the formulation of the Turing machine decades before the invention of the transistor, efforts to engineer a real-world quantum computer that can implement these theories are in their infancy.

The current state-of-the-art are so-called noisy intermediate-scale quantum (NISQ) computers [7]. These are devices without enough qubits to spare for error correction (i.e., noisy) and with a relatively small qubit number (i.e., intermediate-scale). With ~ 50 qubits, NISQ devices cannot run particularly resource-intensive quantum algorithms (e.g., Shor’s algorithm at a meaningful scale) but do afford the occasional speedup over classical computers. To stress these capabilities, there is significant effort towards minimizing *quantum circuits*, the standard model for a quantum computation, to mitigate noise and decoherence. More formally, this field is known as *quantum circuit optimization*.

The ZX-calculus, a graphical formalism for quantum circuits rooted in category theory, provides a unique setting for circuit optimization. In this language, quantum circuits are represented as *ZX-diagrams* that are manipulated via a set of rewrite rules. ZX-diagrams provide a lower-level representation for quantum computations and can be deformed arbitrarily, enriching the space of possible simplifications. While there is no known procedure for recovering a quantum circuit from a general ZX-diagram, subsets of transformations have been identified that preserve circuit extractability.

These transformations have enabled a suite of circuit optimization procedures in the ZX-calculus (e.g., for reducing T-count) with various tradeoffs.

Still, circuit extraction from a ZX-diagram remains a bottleneck for optimization; the best known extraction procedures can introduce an unwieldy degree of complexity in the resulting circuit. For example, circuit extraction can introduce many CNOT gates which are particularly problematic on NISQ devices. Interestingly, however, closely related ZX-diagrams can induce drastically different circuits. Current optimization procedures employing the ZX-calculus do not exploit this fact (e.g., by searching local variants) but rather return a ZX-diagram (or its associated circuit) once no more graph simplifications can be applied.

In this thesis, we explore the role that such a search over local variants of ZX-diagrams could play in circuit optimization. More specifically, we apply various search procedures (i.e., simulated annealing and genetic algorithms) over the application of rewrite rules that do not necessarily simplify the graph itself but may reduce the complexity of the associated circuit. By default, we apply this search over an already simplified diagram (via off-the-shelf methods) but also explore more general strategies for combining these search procedures and/or rewrite rules with existing optimization procedures. We find that (FIXME: this combination is effective for reducing the 2-qubit count on random circuits when the number of qubits is less than 10). FIXME: Also present application.

Chapter 2

Background

To understand quantum circuit optimization in the ZX-calculus, one first must understand quantum circuits and the ZX-calculus. In this section we provide the requisite background in both of these, preceded by a brief discussion on qubits. We conclude this section with an overview of quantum circuit optimization in the ZX-calculus.

2.1 Qubits

A classical bit has a value of 0 or 1. A quantum bit, or *qubit*, encodes a quantum superposition of these two values and therefore more information than a classical bit. In the traditional bra-ket notation of quantum mechanics, a qubit is represented by a vector in \mathbb{C}^2 ,

$$|\psi\rangle = \lambda_1|0\rangle + \lambda_2|1\rangle = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

where

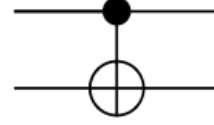
$$\begin{aligned} |0\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |1\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

and λ_1 and λ_2 are the probability amplitudes of observing $|0\rangle$ and $|1\rangle$ upon measurement in this basis, respectively. Notably, qubits can be written with respect to any orthonormal basis. The set $\{|0\rangle, |1\rangle\}$ is known as the *computational* basis. Another common basis is the $+/-$ basis, where

$$\begin{aligned} |+\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ |-\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}} \end{aligned}$$



(a) The Hadamard gate.



(b) The CNOT gate. The top and bottom qubits are the control and target, respectively.

Figure 2.1: The circuit representation of two common quantum logic gates. Wires represent qubits and boxes represent quantum logic gates. Quantum circuits are read from left to right.

The computational and $+/-$ bases are also known as the Z and X bases, respectively.

More generally, a system of n qubits is represented by the tensor product of the individual states. So, an n -qubit state corresponds to a 2^n -dimensional vector. For example, the two-qubit state with both qubits in $|0\rangle$ is

$$|0\rangle \otimes |0\rangle = |00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

In the ZX-calculus and this work, qubits are abstracted as wires in string diagrams. However, this definition is included both for completeness and to aid understanding of quantum circuits.

2.2 Quantum Circuits

A quantum circuit models a quantum computation as a sequence of discrete gates. In the traditional notation with qubits as vectors in \mathbb{C}^2 , quantum gates correspond to unitary matrices. For example, the Hadamard gate, a quantum gate that acts on a single qubit and maps

$$\begin{aligned} |0\rangle &\mapsto \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ |1\rangle &\mapsto \frac{|0\rangle - |1\rangle}{\sqrt{2}} \end{aligned}$$

has the matrix form

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

Figure 2.1a depicts the circuit representation of the Hadamard gate. Note that $|+\rangle = H|0\rangle$ and $|-\rangle = H|1\rangle$. An important family of single-qubit gates are the *phase shift*

gates that modify the phase of the quantum state by some α . A phase shift gate maps

$$\begin{aligned} |0\rangle &\mapsto |0\rangle \\ |1\rangle &\mapsto e^{i\alpha}|1\rangle \end{aligned}$$

and has the matrix form

$$R_\alpha = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{bmatrix}$$

Examples of phase gates are the T-gate ($\alpha = \frac{\pi}{4}$) and the S-gate ($\alpha = \frac{\pi}{2}$). Phase gates where α is a multiple of $\pi/2$ are called *Clifford* gates.

Quantum gates are not restricted to acting on single qubits. For example, the controlled-NOT (CNOT) gate acts on two qubits and flips the second (target) qubit only when the first (control) qubit is $|1\rangle$. The CNOT gate is typically depicted as shown in Figure 2.1b and has the following matrix form:

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

In general, an n -qubit gate is represented by a 2^n -dimensional unitary. However, attention is typically restricted to single- and 2-qubit gates as it is well known that the Clifford + T gate set, consisting of the H, T, S, and CNOT gates, is *universal* – any other operation can be represented by a finite sequence from this set. The *Clifford circuits* are those that can be generated by $\{H, S, CNOT\}$. As the T and S gates are instances of phase shift gates, $\{H, R_\alpha, CNOT\}$ is also universal though in practice any non-Clifford phase gate is implemented with T gates.

Quantum gates can be composed sequentially or in parallel. Consider two gates A and B . The effect of B applied in series after A can be described by a single gate (i.e., linear map) via matrix multiplication (i.e., $B \cdot A$). Alternatively, the tensor product is used to describe A and B in parallel (i.e., $B \otimes A$). For example, the simple circuit shown in Figure 2.3a can be described by the following linear map:

$$(\mathbb{I} \otimes CNOT) \cdot (\mathbb{I} \otimes H \otimes \mathbb{I}) \cdot (CNOT \otimes \mathbb{I}) \cdot (S \otimes \mathbb{I} \otimes \mathbb{I})$$

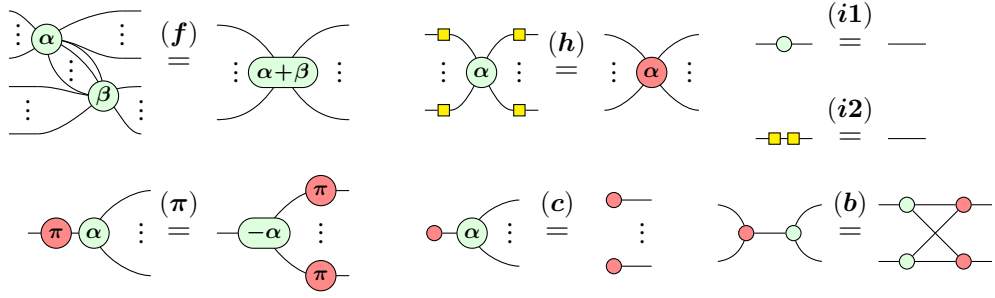


Figure 2.2: The rules of the ZX-calculus. All rules hold with the colors interchanged and for $\alpha, \beta \in [0, 2\pi)$.

2.3 ZX-Calculus

The ZX-calculus is a graphical language for representing and reasoning about quantum processes. Quantum processes are represented by *ZX-diagrams* which consist of *wires* and *spiders*. Spiders can have an arbitrary number of inputs and outputs and come in two flavors: Z spiders (shown in green) and X spiders (shown in red). As with quantum circuits, a ZX-diagram is interpreted from left to right.

ZX-diagrams abstract away the complexities of linear algebra and tensor products from traditional formalisms of quantum processes. Wires still represent qubits while spiders provide a general form for linear maps, where

$$\begin{aligned} \text{Z spider with } \alpha &:= |0\dots 0\rangle\langle 0\dots 0| + e^{i\alpha} |1\dots 1\rangle\langle 1\dots 1| \\ \text{X spider with } \alpha &:= |+\dots +\rangle\langle +\dots +| + e^{i\alpha} |-\dots -\rangle\langle -\dots -| \end{aligned}$$

Two diagrams can be composed in serial by joining the outputs of the first to the inputs of the second, or in parallel by stacking the two diagrams.

From this definition, we can identify the diagrammatic representations for several common components of quantum computation:

$$\begin{aligned} \text{green circle} &= |0\rangle + |1\rangle \propto |+\rangle & \text{red circle} &= |+\rangle + |-\rangle \propto |0\rangle \\ \text{green circle with } \pi &= |0\rangle\langle 0| - |1\rangle\langle 1| = Z & \text{red circle with } \pi &= |+\rangle\langle +| - |-\rangle\langle -| = X \end{aligned}$$

where Z and X are the corresponding Pauli matrices. Note that spiders without any inputs can be regarded as qubit state preparations. The Hadamard gate is so pervasive that it merits shorthand notation; we use a yellow square to represent the Hadamard gate as shown below

$$\text{yellow square} = \text{green circle with } \frac{\pi}{2} \text{ --- red circle with } \frac{\pi}{2} \text{ --- green circle with } \frac{\pi}{2}$$

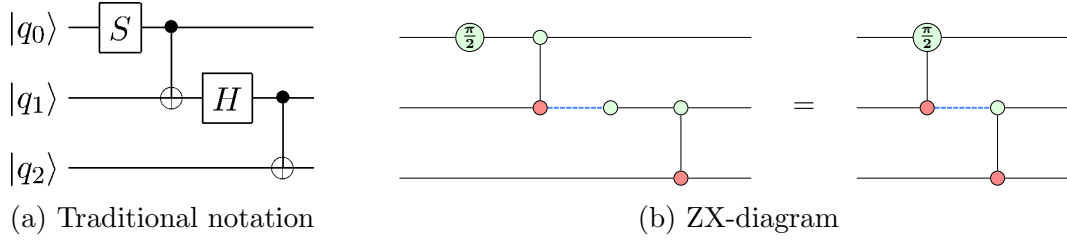


Figure 2.3: A simple 3-qubit quantum circuit shown in both the traditional notation and as a ZX-diagram.

and replace a Hadamard gate between two spiders with a blue dashed edge:

$$\begin{array}{c} \diagup \quad \diagdown \\ \vdots \end{array} \text{---} \text{---} \begin{array}{c} \diagdown \quad \diagup \\ \vdots \end{array} := \begin{array}{c} \diagup \quad \diagdown \\ \vdots \end{array} \text{---} \text{---} \begin{array}{c} \diagdown \quad \diagup \\ \vdots \end{array}$$

We can also reason about ZX-diagrams. The first rule for determining equality is that *only connectivity matters* (OCM). In other words, two ZX-diagrams are equal when one can be deformed into the other (e.g., via moving vertices around in the plane or bending wires) while maintaining the connectivity and the order of the inputs and outputs. In addition to this OCM principle, the ZX-calculus includes a primary set of rewrite rules as shown in Figure 2.2. These rules only hold up to non-zero scalar factors, though these scalars are typically ignored as they correspond to negligible differences in global phase. Additional rules can be derived from this set, such as the antipode rule

$$\begin{array}{c} \diagup \quad \diagdown \\ \vdots \end{array} \text{---} \text{---} \begin{array}{c} \diagdown \quad \diagup \\ \vdots \end{array} = \begin{array}{c} \diagup \quad \diagdown \\ \vdots \end{array} \text{---} \begin{array}{c} \diagdown \quad \diagup \\ \vdots \end{array} \quad (a)$$

and the π -copy rule

$$\begin{array}{c} \diagup \quad \diagdown \\ \vdots \end{array} \text{---} \begin{array}{c} \diagdown \quad \diagup \\ \vdots \end{array} = \begin{array}{c} \pi \\ \pi \\ \vdots \\ \pi \end{array} \quad (\pi c)$$

All quantum circuits can be translated to ZX-diagrams. This is a consequence of the fact that the following universal gate set can be easily represented in the ZX-calculus:

$$CNOT = \begin{array}{c} \text{---} \text{---} \text{---} \\ \text{---} \text{---} \end{array} \quad R_\alpha = \text{---} \text{---} \quad H = \text{---} \text{---}$$

In the CNOT diagram, the green and red spiders are the control and target qubits, respectively. As an example, Figure 2.3 depicts both the traditional and diagrammatic representations of a simple quantum circuit.

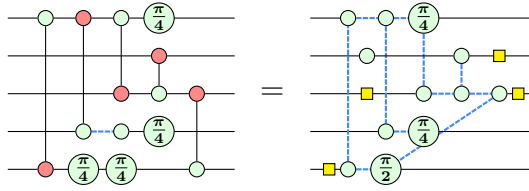


Figure 2.4: An example of a ZX-diagram and an equivalent, graph-like ZX-diagram.

2.4 Circuit Optimization in the ZX-Calculus

The goal of quantum circuit optimization is to simplify quantum circuits (e.g., via reducing the number of gates) to reduce noise and decoherence when run on modern quantum computers. More specifically, the T-count and 2-qubit-count are typically targeted; fault tolerant implementations of T gates generally require an order of magnitude more resources than Clifford gates [2], and the fidelity of single-qubit gates is demonstrably higher than that of 2-qubit gates [1]. It is standard to perform optimizations at the circuit-level. Examples of circuit-level optimizations are 3-qubit gate decomposition and cancelling back-to-back CNOT gates.

The ZX-calculus enables circuit optimization at the graph-level, a more flexible setting that is well-studied in its own right. Such an optimization involves converting a circuit to a ZX-diagram, simplifying the diagram, and converting the diagram back to a circuit. However, there is no known general-purpose procedure for recovering a quantum circuit from a generic ZX-diagram. This restricts the space of permissible simplifications to those that preserve whichever diagrammatic properties are required by the extraction procedure.

In 2020, Duncan et al. reported an extraction procedure that enabled current best practices [3]. Firstly, this procedure operates on *graph-like* ZX-diagrams. A ZX-diagram is graph-like when:

1. All spiders are Z-spiders.
2. Z-spiders are only connected via Hadamard edges
3. There are no parallel Hadamard edges or self-loops
4. Every input or output is connected to a Z-spider and every Z-spider is connected to at most one input or output

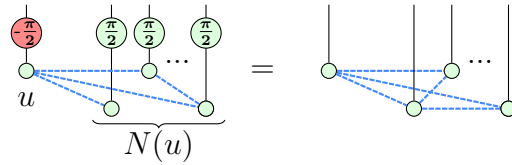
Every ZX-diagram is equal to a graph-like diagram (see Figure 2.4 for an example), and this form admits an underlying graph structure that permits graph-theoretic

analyses. Secondly, this procedure requires that the underlying graph of the input ZX-diagram satisfies a graph-theoretic invariant called *focused generalized flow* (focused gFlow). Given such an input ZX-diagram, extraction proceeds in such a way that each non-zero phase corresponds to one phase gate in the resulting circuit but each edge can correspond to multiple CNOTs. Importantly, this extraction procedure is not optimal; if a given circuit is converted to a ZX-diagram and the extraction procedure is immediately applied, the extracted circuit is often more complex than the original circuit. In particular, this procedure can drastically increase the number of 2-qubit gates. For more details on the extraction procedure or focused gFlow, see [3].

Alongside this extraction procedure, Duncan et al. also introduced an optimization procedure that relies on focused gFlow-preserving graph simplifications. More specifically, this procedure relies on two graph-theoretic transformations that each correspond to rewrite rule in the ZX-calculus: *local complementation* and *pivoting*. Given a graph G and some vertex u of G , the *local complementation* of G according to u , denoted $G \star u$, is the graph in which the connectivity of all pairs of neighbors of u is inverted. For example,



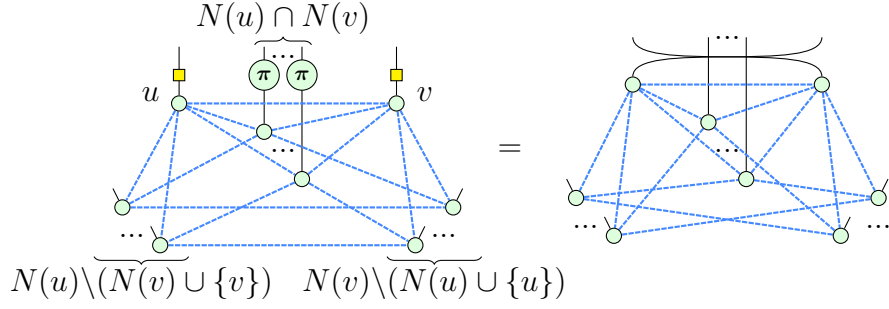
Given a ZX-diagram, local complementation can be induced in the underlying graph structure while maintaining equality by applying an $X_{-\pi/2}$ gate on (the spider corresponding to) u and a $Z_{\pi/2}$ gate to its neighbors [4]:



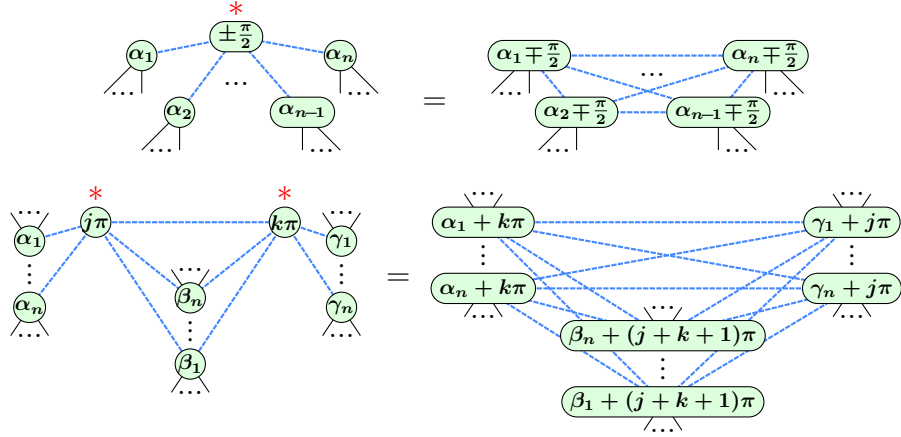
Relatedly, given two connected vertices u and v in G , the *pivot* of G along the edge uv is the graph $G \wedge uv := G \star u \star v \star u$. In practice, this consists of complementing the edges between three subsets of vertices: (1) the common neighborhood of u and v , (2) the exclusive neighborhood of u , and (3) the exclusive neighborhood of v :



Again, given a ZX-diagram, the rewrite rule reported in [5] introduces a pivot by applying Hadamard gates on u and v and Z_π gates on their common neighborhood:



Each rewrite rule can be extended to a simplification by performing local complementation (resp. pivoting), removing the vertex (resp. the pair of vertices), and updating the phases:



Importantly, the existence of a focused gFlow is preserved in both cases. At a high-level, optimization can then be performed by applying these simplifications to fixpoint and extracting a circuit from the resulting ZX-diagram. Details on this optimization procedure and proofs of the simplification rules or their focused gFlow-preservation can be found in [3]. A variant of this procedure is implemented in the `full_reduce` method of the PyZX library.

One issue with `full_reduce` is its reliance on circuit extraction; it is not uncommon that the circuit extracted from the final, simplified ZX-diagram has more gates than the input circuit. Kissinger and van de Wetering introduced an alternative optimization method that uses *phase teleportation* to sidestep the issue of circuit extraction altogether [6]. Crucially, simplification of the ZX-diagram is performed symbolically. This symbolic simplification leverages the fact that several rewrite rules involving *phase-gadgets*, a particular ZX-diagrammatic motif, correspond to changes in the original phases. Therefore, non-Clifford phases can potentially cancel or combine with each other while leaving the original graphical structure of the ZX-diagram intact. So, as simplification proceeds, application of these rules is simply tabulated to enable reconstruction of a final circuit with the same connectivity but a possible

reduction in non-Clifford gates (and therefore T-count). Importantly, using phase teleportation rather than circuit extraction ensures that the number of 2-qubit and Hadamard gates is unchanged. By changing the angles of many non-Clifford phase gates to either 0 or multiples of $\pi/2$, this procedure can also make a circuit-level optimization procedure much more effective. This procedure is implemented in the `teleport_reduce` method of PyZX.

Chapter 3

Methods

TODO: Previous approaches failed to do FIXME. We want to do that. Here we describe methods for doing so. TODO: First, we want rewrite rules that don't necessarily simplify but preserve focused gFlow and may change circuit extraction. Pivoting and LC are good for this because they change connectivity. Lets show that. TODO: We then need a way to find the best application of these congruences to simplify circuit. We describe some general purpose search procedures that we use. TODO: We then describe how we are going to use all of this for circuit optimization.

3.1 Congruences

FIXME: Congruences. We desire rewrite rules that won't make the graph about as complex but could lead to different extraction.

TODO: Proof in appendix

3.2 Search Procedures

TODO: Here we describe some search procedures to apply the congruences

3.2.1 Simulated Annealing

TODO

3.2.2 Genetic Algorithms

TODO

3.2.3 Target Functions

TODO: weighted 2-qubit gates, number of edges, etc

3.3 Circuit Optimization Strategies

TODO: Here we describe how we use conruences and these search procedures for quantum circuit optimization. First describe primary method: searching local space after simplification. Then describe several other variants.

FIXME: Other, more general combinations (e.g., GA with all simps, or search with other “safe” procedures)

Chapter 4

Results

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Chapter 5

Application: Quantum Chemistry

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Chapter 6

Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

Bibliography

- [1] CJ Ballance, TP Harty, NM Linke, MA Sepiol, and DM Lucas. High-fidelity quantum logic gates using trapped-ion hyperfine qubits. *Physical review letters*, 117(6):060504, 2016.
- [2] Earl T Campbell, Barbara M Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, 2017.
- [3] Ross Duncan, Aleks Kissinger, Simon Perdrix, and John Van De Wetering. Graph-theoretic simplification of quantum circuits with the zx-calculus. *Quantum*, 4:279, 2020.
- [4] Ross Duncan and Simon Perdrix. Graph states and the necessity of euler decomposition. In *Conference on Computability in Europe*, pages 167–177. Springer, 2009.
- [5] Ross Duncan and Simon Perdrix. Pivoting makes the zx-calculus complete for real stabilizers. *arXiv preprint arXiv:1307.7048*, 2013.
- [6] Aleks Kissinger and John van de Wetering. Reducing t-count with the zx-calculus. *arXiv preprint arXiv:1903.10477*, 2019.
- [7] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.