Efficient and Correct
Compilation of Quantum Circuits
(Overview Paper)

Robert Wille  Stefan Hillmich  Lukas Burgholzer
Institute for Integrated Circuits, Johannes Kepler University Linz, Austria
{robert.wille, stefan.hillmich, lukas.burgholzer}@jku.at

Abstract—High-level descriptions of quantum algorithms do not take the restrictions of physical hardware into account. Therefore actually executing an algorithm in the form of a quantum circuit on a quantum computer requires compiling it for the desired target architecture first. The compilation of quantum circuits depends on efficient methods to be feasible for all but the trivial instances. To this end, different compiling methods have been introduced in the past, but room for improvement still exists. Moreover, just an efficient compilation process itself is not sufficient—the resulting circuits must be correct as well. In this summary paper, we review how existing compilation approaches can be optimized by utilizing heuristic search algorithms or exact reasoning engines. Furthermore, we review how the correctness of the obtained results can be verified afterwards by clever data structures such as decision diagrams. This illustrates core steps of a compilation flow which can generate minimal or close-to-minimal results for many instances and, additionally, guarantees correctness throughout the process.

I. INTRODUCTION

Quantum computing offers a new model of computation that promises to significantly outperform classical computing at specific tasks. These tasks include integer factorization (Shor’s algorithm [1]), database search (Grover’s search [2]), quantum chemistry [3], Boson sampling [4], and many more. An increasing number of big players such as Google, IBM, and Microsoft as well as start-ups like Rigetti and IonQ are currently investing in the field in order to bring quantum computing closer to reality.

Despite the plethora of promising applications, the capability of physical quantum computers is not quite there yet—the phase we are currently in is often referred to as Noisy Intermediate-Scale Quantum Computing (NISQ [5]). At the moment, only up to around 50 qubits are available, which additionally suffer from comparatively high error rates and short decoherence times. Due to the small number of qubits, quantum error correction methods are not yet feasible.

The shortcomings of the hardware can be mitigated to some extent by carefully selecting algorithms and utilizing an appropriate design flow. Since quantum algorithms are typically formulated on an abstraction level agnostic of the restrictions imposed by physical hardware, one or more compilation steps are required in order to execute them on the actual hardware. This includes decomposing the algorithmic description into gates from the supported quantum operation library (such as Clifford+T) and adhering to possible connectivity constraints imposed by the architecture.

To this end, dedicated compilation flows are required. They have to satisfy two major properties: On the one hand, they have to be efficient, i.e., they should be able to decompose a given algorithm and to address the connectivity constraints using as few elementary operations as possible. On the other hand, they have to be correct, i.e., it should be possible to check whether the resulting quantum circuit indeed realizes the initially given quantum functionality.

In this work, we review existing compilation methods and how they can be optimized using heuristic as well as exact methods. We particularly focus thereby on the so-called mapping problem which addresses the connectivity constraints mentioned above. Afterwards, we review methods that can be used to verify whether the obtained results are correct, i.e., whether they indeed realize the desired functionality. To this end, we focus on methods based on decision diagrams, which are particularly suited for this task. All methods will be illustrated by means of examples. Furthermore, references for further reading and a more in-depth treatment will be provided.

The remainder of this paper is structured as follows: Section II provides a brief overview of quantum computing and architectural constraints of physical devices. Section III discusses the problem of mapping and reviews efficient solutions for that. Afterwards, Section IV covers verification for quantum computing, especially the problem of equivalence checking, which can be used to check whether the results determined from the compilation are correct. Finally, Section V concludes the paper.

II. BACKGROUND

This section provides the background in quantum computations and architectural constraints of quantum computers which are necessary to keep this work self-contained.

A. Quantum Computation

Computations in the quantum realm utilize quantum bits (qubits) [6], which can assume more states than just 0 and 1 known from conventional logic. In fact, while the basis states (denoted by |0⟩ and |1⟩ in Dirac-notation) are the same, a qubit |ψ⟩ can be in any linear combination

|ψ⟩ = α · |0⟩ + β · |1⟩

described through amplitudes α, β ∈ ℂ with |α|^2 + |β|^2 = 1. If both amplitudes α and β are non-zero, the qubits state is also referred to as being in superposition. A second exploitable quantum effect is called entanglement,
where operations on one qubit influence the state of another qubit. On a quantum computer, the values of $\alpha$ and $\beta$ cannot be directly observed. Instead, a measurement of a qubit collapses the state of the qubit back to one of the basis states $|0\rangle$ (with probability $|\alpha|^2$) or $|1\rangle$ (with probability $|\beta|^2$).

In general a quantum system consists of an ensemble of $n$ qubits—spanning a $2^n$-dimensional complex state space, i.e., $|\psi\rangle = \sum_{x \in\{0,1\}^n} \alpha_x |x\rangle$ with $\alpha_x \in \mathbb{C}$ such that $\sum_{x \in\{0,1\}^n} |\alpha_x|^2 = 1$, which is commonly represented as the $2^n$-dimensional complex state vector $|\alpha_x\rangle_{x \in\{0,1\}^n}$. The current state of a quantum system $|\psi\rangle$ can be manipulated using quantum operations, which are defined through unitary matrices $U$. With the exception of measurements, quantum operations are therefore inherently reversible.

**Example 1.** Two common quantum operations with their matrix representations are shown in Fig. 1. The Hadamard transformation $H$ sets a qubit into superposition, e.g., it transforms $|0\rangle$ to $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. The controlled-NOT CNOT operates on two qubits and negates the target qubit, iff the chosen control qubit is in the state $|1\rangle$.

A quantum algorithm is commonly described as a sequence of quantum operations applied to the qubits of a quantum system. Quantum circuit diagrams allow to visualize any such sequence. Here, the qubits are represented by horizontal lines, while quantum operations are placed on the qubits and are applied from left to right. In this context quantum operations are also referred to as quantum gates.

**Example 2.** Fig. 1 illustrates a quantum circuit consisting of a single Hadamard gate and four CNOTs. In case of the CNOTs, the black dot denotes the control qubit, while the $\oplus$-symbol denotes the target qubit. The gates are labeled $g_0, \ldots, g_4$ and are applied to the qubits $q_0, \ldots, q_3$.

### B. Quantum Computer Architectures

The high-level descriptions of quantum algorithms commonly do not satisfy the restrictions imposed by actual physical devices. Hence, the quantum circuit representing the algorithm has to be compiled to a version executable on the hardware. The remainder of this section focuses on the IBM QX architectures [7] as a representative example.

All quantum computers available today only offer a restricted set of quantum gates that can be applied to the devices’ qubits. This does not mean these quantum computers are not universal—in fact, rather “small” sets of libraries such as Clifford+T are sufficient to approximate any quantum state arbitrarily close [6].

**Example 3.** Quantum computers made by IBM share the same set of supported quantum operations. They provide the single-qubit $U$ gate, which allows for arbitrary single-qubit operations. Additionally, IBM architectures support the two-qubit CNOT gate. As these operations form a superset of the Clifford+T gate set, universal quantum computations may theoretically be performed on IBM’s hardware.

In addition to restricting the natively available gates, quantum computers may also impose certain (hardware-dependent) connectivity constraints. These typically limit the possible interactions between qubits in the form of a coupling map.

**Example 4.** Consider the IBM QX4 architecture with five qubits, whose coupling map is shown in Fig. 2. There, the nodes $Q_i$ indicate the physical qubits and an arrow from $Q_c$ to $Q_t$ indicates that a CNOT with control $Q_c$ and target $Q_t$ may be applied. In order to distinguish an algorithm’s logical and the hardware’s physical qubits, logical qubits are denoted with lower-case letters $q_i$ by convention.

Such connectivity constraints only apply to certain types of quantum computers, e.g., those based on superconducting qubits [8]—as opposed to hardware based on trapped ions [9].

### III. Compilation

Compilation of quantum circuits necessitates a multitude of different tasks in order to conduct a given quantum computation on a specific physical device. As already mentioned in the previous section, algorithmic building blocks have to be decomposed into elementary gates provided by the underlying architecture. Then, the resulting circuit’s logical qubits have to be mapped to the device’s physical qubits in a way that all coupling constraints imposed by the architecture are satisfied. Furthermore, several pre- and post-mapping optimizations may be applied in order to reduce gate count and/or increase computation fidelity. In the following, we consider the mapping problem as one of the most important parts of the compilation problem.

---

1 A matrix $U$ is unitary if $UU^\dagger = U^\dagger U = I$, i.e., its inverse $U^{-1}$ is given by the conjugate transpose $U^\dagger$ [6].
Example 5. Consider again the circuit from Fig. 1b and assume that the target architecture is IBM QX4 with the coupling map given in Fig. 2. Choosing the initial mapping $q_1 \rightarrow Q_1$ yields a situation where none of $g_1, \ldots, g_4$ satisfy the coupling constraints. Adding SWAP gates in a naive fashion (i.e., by adding SWAP gates to “move” the qubits so that they eventually satisfy the constraints) may lead to a result as shown in Fig. 3a—and, hence, a circuit $G'$ with substantially more gates.

B. Proposed Solutions

Determining a solution which satisfies the constraints and, at the same time, keeps the cost at a minimum has been proven an NP-complete problem [14], [15]. Accordingly, efficient approaches are key to tackle this problem. In the following paragraphs, we will briefly review two different solutions to the mapping problem, namely a heuristic one based on the $A^*$ search algorithm (introduced in [16]) and a minimal one utilizing reasoning engines (introduced in [17]).

The $A^*$ algorithm [23] is a state-space search algorithm. To this end, (sub-)solutions of the considered problem are represented by state nodes. Nodes that represent a complete solution are called goal nodes (multiple goal nodes may exist).

The main idea is to determine the cheapest path (i.e., the path with the lowest cost) from the root node (representing an empty solution) to a goal node (representing a complete solution). Since the search space is typically exponential, sophisticated mechanisms are employed in order to consider as few paths as possible. This $A^*$-algorithm was one of the first heuristic methods that has been utilized to address the mapping problem [16].

Example 6. Fig. 3b illustrates the mapped circuit $G'$ resulting from the $A^*$-approach. As can be seen, this circuit requires fewer SWAP and $H$ operations and, even if it is not optimal in the number of gates, this result is commonly attained after a short runtime.

A minimal mapped circuit can be obtained by finding the cheapest of all possible mappings — again, an $NP$-complete problem [14], [15]. To this end, the mapping task can be formulated as a problem of Boolean satisfiability (SAT) for which efficient reasoning engines tackling this complexity exists (e.g., [24], [25]). The relationship between logical and physical qubits is encoded in Boolean variables which are then constrained in order to only allow valid assignments conforming with the targeted architecture and ruling out invalid solutions (details on that are provided in [17]). Passing these variables and constraints to a reasoning engine in combination with a cost metric (e.g., number of additionally inserted gates) allows to determine an optimal solution to the mapping problem.

Example 7. Fig. 3c provides the mapped circuit $G'$ determined by the reasoning engine as a result of the SAT encoding. Notably it does not require a single SWAP operation, the only change necessary was the addition of four Hadamard operations to account for the direction in the coupling map.

Both approaches are extremely flexible when it comes to adjusting to different architectures or cost metrics, as was already shown in [26], [27] for the heuristic $A^*$ algorithm, allowing for a broad applicability of these mapping techniques.

\[ q_1 \rightarrow Q_1 \quad q_0 \rightarrow Q_0 \quad q_1 \quad H \quad H \quad H \quad H \]

Fig. 4: Decomposition of a SWAP operation

For the mapping problem, we assume that all building blocks of the specific quantum algorithm are already decomposed into the specific gate set provided by the target architecture (for this purpose, several methods have been proposed in the literature, see, e.g. [10]–[13]). Then, in order to satisfy all connectivity constraints imposed by the device’s coupling map (see Section II-B), the algorithms logical qubits have to be accordingly assigned to the architecture’s physical qubits. Since there are typically no constraints on the application of single qubit gates, these need not be considered during the mapping process. In general, such a mapping cannot be determined in a static fashion, i.e., by fixing an initial assignment, but the mapping has to change dynamically in all but the trivial cases. Typically, this is realized by inserting SWAP operations 2 which allow to exchange the logical qubits the operation acts on (see Fig. 4).

However, tackling this problem in a naive fashion, i.e., by considering gate after gate and resolving arising issues through appropriate SWAP insertions, does not yield a feasible procedure and frequently incurs large overheads.

\[ q_1 \rightarrow Q_1 \quad q_0 \rightarrow Q_0 \quad q_1 \quad H \quad H \quad H \quad H \]

Fig. 3: Solutions $G'$ for the compilation problem

A. The Mapping Problem

2In case a CNOT with control $Q_i$ and target $Q_j$ is to be applied, but there is only an edge from $Q_j$ to $Q_i$ in the coupling map, instead of applying a SWAP operation the CNOT can be surrounded with four Hadamard operations, effectively switching control and target
IV. VERIFICATION

The design task verification comprises multiple levels from the specification of an algorithm down to the actual execution on physical hardware. In this section, we consider equivalence checking as a representative part of verification. The problem of equivalence checking addresses the question whether two given circuits $G$ and $G'$ do realize the same function—an important question during/after the compilation process. While the problem has been intensively studied in the past [28], [29], huge improvements may be achieved by utilizing efficient data structures and exploiting some properties unique to quantum computing.

A. Decision Diagrams

Decision diagrams have been successfully utilized to drastically reduce the required memory to represent state vectors in quantum computing [30]–[33]. Strong simulation approaches based on decision diagrams have recently moved into the spotlight since they can significantly outperform array-based simulators in cases where redundancies can be exploited—in extreme cases leading to an improvement in runtime by from 30 days to 2 minutes [34], [35].

The main idea of decision diagrams is to identify redundancies in the state vector and provide compaction by sharing structures. The vector is split in half into two sub-vectors. This process is repeated until the sub-vectors contain only single elements, i.e., one split for every qubit. If identical sub-vectors occur in the process, this redundancy is exploited by re-using (sharing) the same structure in the resulting decision diagram.

More precisely, consider a quantum system with qubits $q_{n−1}, q_{n−2}, \ldots, q_0$, whereby $q_{n−1}$ represents the most significant qubit. Then, the first $2^n−1$ entries of the corresponding state vector represent the amplitudes for the basis states with $q_{n−1}$ set to $|0\rangle$; the other entries represent the amplitudes for states with $q_{n−1}$ set to $|1\rangle$. This decomposition is represented in a decision diagram structure by a node labeled $q_{n−1}$ and two successors leading to nodes representing the two sub-vectors. By convention, the left (right) edge indicates the 0-successor (1-successor). The sub-vectors are recursively decomposed further until vectors of size 1 (i.e., complex numbers) result. During this decomposition, equivalent sub-vectors can be represented by the same nodes—reducing the complexity of the representation. Then, instead of having a terminal node for every distinct value in the state vector, common factors of the amplitudes are stored in the edge weights. The value can be reconstructed by multiplying the edge weights along the desired path in the decision diagram.

B. DD-based Equivalence Checking

Decision diagrams are ideal for equivalence checking, since the resulting representations are canonic (as proven in [36]), i.e., simply creating decision diagrams for two circuits $G$ and $G'$ and comparing them afterwards solves the problem (as done in [37], [38]). Besides that, further improvements can be achieved.

Fig. 5: DD for circuits $G$ and $G'$ from Fig. 1b and Fig. 3

For example, in case two circuits $G$ and $G'$ are indeed equivalent, it certainly holds that $G·G'^{-1} = \mathbb{I}$, i.e., multiplying $G$ with the inverse of $G'$ results in the identity $\mathbb{I}$.

The full potential of this idea is exploited by utilizing the associativity of the matrix multiplications involved in computing $G · G'^{-1}$, i.e., starting from a decision diagram representing the identity $\mathbb{I}$, gates from $G$ and $G'^{-1}$ are applied successively either from the left ($G$) or from the right ($G'^{-1}$):

$$G · G'^{-1} = (g_0 \cdots g_{m−1}) · (g_m^{−1} \cdots g_0^{−1})$$
$$\equiv (U_{m−1} \cdots U_0) · (U^{−1}_0 \cdots U^{−1}_{m−1})$$
$$\equiv (U_{m−1} \cdots U_0) \rightarrow \mathbb{I} \leftarrow (U^{−1}_0 \cdots U^{−1}_{m−1})$$
$$= G \rightarrow \mathbb{I} \leftarrow G'^{-1}.$$  

However, determining when to apply gates from $G$ and when to apply gates from $G'^{-1}$ is not always obvious. But whenever a “good” strategy for a selection of gates can be employed, equivalence checking of two equivalent quantum circuits can be conducted very efficiently and compactly using decision diagrams.

Example 8. Consider the two circuits $G$ and $G'$ from Fig. 1b and Fig. 3c. Conducting the multiplications in an alternating fashion frequently results in a situation where the application of a gate from $G$ is immediately reverted by the application of gates from $G'^{-1}$, thus effectively keeping the involved decision diagram close to the identity. By this, instead of $13$ nodes (as shown in Fig. 5), no more than $7$ nodes are required during the equivalence check.

V. CONCLUSIONS

In this work, we discussed methods towards an efficient and correct compilation flow for quantum circuits. To this end, we showed how to conduct efficient mapping utilizing reasoning engines such as satisfiability solvers to attain minimal solution as well as an $A^*$-based heuristic for close-to-minimal solutions. However, efficiency by itself is not sufficient, correctness is required, too. In this regard, we showed how the use of clever data structures such as decision diagrams can reduce the runtime significantly in many cases.

ACKNOWLEDGMENT

This work has partially been supported by the LIT Secure and Correct Systems Lab funded by the State of Upper Austria.

*The identity is the best case for decision diagrams due to its recursively equal sub-parts, allowing for a representation linear in the number of qubits.