tweedledum: A Compiler Companion for Quantum Computing

Bruno Schmitt, Giovanni De Micheli
Integrated Systems Laboratory (LSI), EPFL, Switzerland

Abstract—This work presents tweedledum—an extensible open-source library aiming at narrowing the gap between high-level algorithms and physical devices by enhancing the expressive power of existing frameworks. For example, it allows designers to insert classical logic (defined at a high abstraction level, e.g., a Python function) directly into quantum circuits. We describe its design principles, concrete implementation, and, in particular, the library’s core: An intuitive and flexible intermediate representation (IR) that supports different abstraction levels across the same circuit structure.

Index Terms—quantum, design automation, compilation

I. INTRODUCTION

The emergence of quantum computers with an increasingly higher number of qubits and longer coherence times marks the beginning of an exciting era in quantum technology, as it empowers us to solve problems that are out of reach for any of the best classical supercomputers [1], [2]. However, finding applications and algorithms for which a quantum algorithm yields a significant scaling advantage in time-to-solution, known as a quantum speed-up, over classical algorithms has been a significant concern in the quantum computing research community. The “Quantum Algorithm Zoo” website [3] holds a comprehensive list of quantum algorithms with their respective speed-up factors. Finding more of these algorithms is crucial to quantum computing progress.

The quantum computing community hopes that the availability of quantum hardware and the development of quantum programming languages will stimulate and aid the discovery of new quantum algorithms [4]. However, most programming systems available for quantum computing are intertwined with the quantum circuit model, which means that the developer must describe the algorithm in terms of basic unitary operators. Not surprisingly, the implementation of quantum algorithms on such a low level of abstraction is very time-consuming, error-prone, and results in non-portable implementations—given the diversity of quantum devices.

To explore new algorithms and programs for quantum computing as well as to enhance the productivity of programmers, we have seen many quantum frameworks striving to support higher-level abstractions, e.g., Q# [5], Qiskit [6], PyQuil/Forest [7], PennyLane [8]. These frameworks allow developers to create increasingly more complex programs by combining and adapting a small set of known quantum algorithms that, often, use arbitrary technology-independent operations. Nevertheless, given the stringent resource constraints in near-term quantum hardware, the use of higher levels of abstraction is only possible when allied with sophisticated compilation algorithms capable of generating highly optimized low-level circuits. Since circuit compilation and optimization are essential for quantum computing [9], there are numerous competing compilers and toolkits, e.g., qiskit-terra [6], quilc [10], ScaffCC [11], staq [12], and t|ket| [13]. We refer to [14] for a survey of quantum software stacks.

The embodiment of our research contribution is a compiler companion library for the synthesis and compilation of quantum circuits called tweedledum. In contrast to most solutions, we designed it to enhance other compilers and frameworks, and some of these other tools indeed use it already, e.g., qiskit-terra, quilc and staq. We implement tweedledum as an open-source library \(^1\) in C++-17 and provide Python bindings for easy integration into existing compilers/frameworks. The library integrates state-of-the-art algorithms used for quantum compilation, targeting most of the pipeline of a quantum software stack: from the abstract higher algorithmic layers to the physical mapping layer.

II. OVERVIEW

The goal of compilation is to bridge the gap between high-level quantum programs and technology-dependent implementations. Internally, compilation breaks down into a sequence of tasks: lexing the source code into a sequence of tokens, parsing these tokens into an abstract syntax tree (or AST), validating this AST, and finally, translating it into technology-dependent representation.

At this point, tweedledum mainly aims at enhancing a compiler’s capabilities of doing the last translation step. Hence its compiler companion denomination. The library provides various algorithms for synthesizing, optimizing, and manipulating quantum circuits. Fig. I shows a bird’s eye view of a compilation process. A compiler can use tweedledum to build a high-level quantum circuit and then use its passes to progressively lower it into a technology-dependent circuit using the fewest resources possible. For example, tweedledum empowers qiskit-terra to accept classical logic functions, written in Python, as an oracle definition and handles their translation into quantum circuits. Also, it adds the ability to synthesize permutation to Rigetti’s quilc.

III. REPRESENTATION OF QUANTUM FUNCTIONALITY

The basic mathematical objects to be dealt with when representing quantum functionality are Hamiltonians of a quantum system. These are linear, unitary mappings \(\mathbb{C}^n \rightarrow \mathbb{C}^n\) that describe the system’s evolution. There are several ways for representing quantum operators, and tweedledum supports the most common ones. Each way has its strengths and

---

\(^1\)https://github.com/boschmitt/tweedledum
weaknesses. We evaluate the efficiency of a representation conversion is essential during compilation, where various representation is universally suitable for all applications, of supporting transformations and manipulations. Since no by its succinctness in describing operators and its capability of supporting transformations and manipulations. Since no representation is universally suitable for all applications, conversion is essential during compilation, where various synthesis and transformation techniques are applied.

Fig. 2 shows some of the different ways to define quantum functionality. The most natural way to represent a Hamiltonian is to choose a basis of the Hilbert space and then consider the corresponding transformation matrix, a $2^n \times 2^n$ complex-valued unitary matrix. Unitary matrices are canonical: two quantum operators are equivalent if and only if they have the same unitary matrix. Note that canonicity is an essential property for many applications of synthesis and verification. However, they are impractical to represent operators acting on many qubits since their size grows exponentially with the number of qubits.

On the other hand, the quantum circuit model is a simple and convenient tool for representing quantum programs that do not suffer from the same exponential growth. One of the disadvantages is its lack of canonicity, i.e., there are many different ways of representing a given computation using quantum circuits. We can simplistic view many parts of the compilation process as a search for a circuit that better optimizes a cost function of our interest. We refer to [16] and [17] for discussions on quantum decision diagrams and phase polynomials, respectively.

IV. INTERMEDIATE REPRESENTATION

In tweedledum, the standard IR is a quantum circuit. A circuit has a set of wires (qubits and bits) and a sequence of operators applied to those wires, the so-called instructions. An operator is an abstract effect that may modify the state of a subset of wires. An instruction is an operator applied to a specific subset of qubits and bits. In other words, to represent a quantum computation, we first create an empty circuit to which we add qubits and bits. Then we create instructions by applying operators to these qubits and bits.

A. Fundamental concepts

1) Wires: Qubits and bits: We represent qubits and bits using memory semantics, which means instructions act on references to qubits (and bits) and do not consume their value; we say they affect their state via side effects. A benefit of using memory semantics is that the IR inherently prevents a program from violating the no-cloning theorem [18]. Indeed, no IR mechanism allows the copy of a quantum state.

2) Operators: Conceptually, an operator is an effect that we can apply to a subset of qubits and bits. Most often, this effect is unitary evolution. To ensure a high level of customizability, the library does not have a fixed set of operators nor imposes restrictions on how their effects are defined. On the contrary, it encourages and facilitates the implementation of user-defined ones. Indeed, one of the main strengths of tweedledum lies in its use of a uniform concept, known as Operator, to enable
describing different levels of abstractions and computations. For example, we can define high-level operators such as the truth table operator, the permutation operator, or low-level operators like the Hadamard operator, e.g., Fig. 3.

The library imposes two minimal requirements for the Operator concept. Namely, every Operator must be identifiable and must target at least one qubit. (Note that the number of targets is intrinsic to an operator; the number of controls is not.)

An exciting feature of *tweedledum* is the ability to insert classical logic directly into quantum circuits and to synthesize it during compilation. Moreover, the library provides means for a user to define such classical logic as a Python function. By combining most of the EPFL logic synthesis libraries [19], we can create compilation flows that handle the translation of this Python function into a sequence of elementary quantum operators.

The library also accepts Boolean functions provided in other forms, e.g., a logic network. Sometimes the input is already a reversible Boolean function. In such cases, we can use specialized methods to synthesize a circuit [20], [21]. More often, however, the Boolean function will be irreversible. Given an irreversible function $f$, it is known that there must exist a reversible Boolean function $f'$ such that $f'(x, y) = (x, y \oplus f(x))$.

where $x = x_0, \ldots, x_{n-1}$ and ‘⊕’ refers to the XOR operation. (For the sake of clarity, we limit the discussion to single-output Boolean functions, but the technique can be extended to accommodate multiple-output functions.) Such an embedding is also referred to as Bennett embedding [22], and implies the existence of the following quantum operation:

$$B_f : |x\rangle|y\rangle \mapsto |x\rangle|y \oplus f(x)\rangle$$

The operation $B_f$ is also known as a single-target operator. Single-target operators describe complex operations that cannot generally be implemented natively on a quantum computer. *tweedledum* provides three techniques to synthesize single-target operators directly. Starting from a functional representation of $f$, i.e., a truth table, both *pkrm_synth* and *prrm_synth* techniques synthesize a particular case of an exclusive sum-of-products (ESOP) expression for $f$. We can easily translate such ESOP into a cascade of multiple-control $X$ operators. These techniques, however, are only applicable to small Boolean functions as they can be both very time-consuming and generate a quantum circuit with a prohibitive number of instructions [23], [24]. *spectrum_synth* [25] uses the Rademacher-Walsh spectrum of a truth table to generate a circuit over the operator set Clifford+$R_z$ directly.

For a more scalable solution, we combine these direct methods with the hierarchical synthesis approaches (Fig. 4).
The latter allows us to achieve scalability by decomposing the initial function into small parts suitable for functional synthesis. The synthesis method generates a reversible circuit for each part and combines them following the structural representation of the function. The combination of subcircuits might require additional qubits, which store intermediate computation steps. Given an irreversible Boolean function \( f \) they find an \((n+1+a)\)-qubit quantum circuit that realizes the unitary

\[
B_f : |x\rangle|y\rangle|0\rangle^a \rightarrow |x\rangle|y \oplus f(x)\rangle|0\rangle^a
\]

where \( a \geq 0 \), which means that the synthesis algorithm can use the \( a \) additional qubits to store intermediate computations. \texttt{lhrs\_synth} \cite{26} and \texttt{xag\_synth} \cite{27} are examples of hierarchical synthesis.

Linear reversible circuits form a subclass of quantum circuits in which implementation requires only \textsc{cnot} operators. Synthesis methods aiming to reduce the size of these circuits play an essential role in \texttt{tweedledum} since other algorithms depend on them. Given a binary matrix describing the classical function, the library has two techniques for synthesizing such circuits. Both methods rely on a modified implementation of Gaussian elimination, which yields asymptotically optimal circuits. They differ in that the \texttt{steiner\_gauss\_synth} \cite{12}, \cite{28} algorithm can synthesize circuits that respect the connectivity constraints of device architectures. We also provide two techniques, \texttt{a\_star\_swap\_synth} and \texttt{sat\_swap\_synth} \cite{29}, to synthesize an even more constrained class of linear reversible circuits: those composed entirely of \textsc{swap} operators.

VI. COMPILATION PASSES

A. Utility

Passes in this category provide simple utilities that do not fit any other category, which more complex passes might either require or significantly benefit from using it. For example, since \texttt{tweedledum} does not modify circuits in place, all passes that modify a circuit must do a shallow duplication, i.e., create a new circuit with the same wires. There are also passes to reverse and invert circuits. The \texttt{reverse} pass creates a new circuit with the instruction applied in the reverse topological order. Inversion is similar, but with the addition of applying the adjoint instruction.

The library also provides an instruction canonicalization pass. The goal of canonicalization is to make optimizations more effective. Very often, we can write instructions in multiple forms. For example, we can write them with equivalent operators \( T = P(\frac{\pi}{8}) = P(-\frac{\pi}{4}) \); or apply an operator to a permutation of the same wires \( \text{SWAP}(q_0, q_1) = \text{SWAP}(q_1, q_0) \). Canonicalization means selecting one of these forms to be canonical and then going through a circuit and rewriting all instructions into the canonical form. Thus, canonicalization allows optimization passes that look for specific patterns to focus only on the canonical forms rather than all forms.

B. Decomposition

We define decomposition as the process of systematically breaking down high-level instructions into a series of lower-level ones. We emphasize “systematically” because it is the characteristic that differentiates decomposition from synthesis. A decomposition technique builds a lower-level implementation through the application of some construction rule(s). Compared to synthesis, decomposition techniques are faster and produce predictable results, i.e., we know in advance the resulting number of instructions and qubits. The cost to decompose a multiple-controlled \textsc{pauli} instruction depends on whether the circuit has clean ancillae available. (Here, we measure cost by the number of instructions.) For example, the circuit might have enough clean ancillae to allow a “\textsc{v-chain}” decomposition, as illustrated in Fig. 5. If that is not the case, we need to use the more costly “dirty-ancilla” decomposition (based on Lemmas 7.1 and 7.2
A unidirected graph \( G = (V, E) \) where \( V \) is the set of virtual qubits \( V = \{v_0, v_1, \ldots, v_{n-1}\} \) and \( E \subseteq \binom{V}{2} \) is a set of qubit pairs \( \{v_i, v_j\} \) used by the instructions in the circuit. (Note that one-qubit instructions can be safely ignored since the coupling constraints do not affect their mapping.) We model the coupling constraints in a similar way. An undirected graph \( (P, E_p) \) where \( P = \{p_0, p_1, \ldots, p_{m-1}\} \) is a set of physical qubits and an edge \( \{p_u, p_w\} \in E_p \subseteq \binom{P}{2} \) means that a instruction can be executed using the two physical qubits \( p_u \) and \( p_w \). The library divides mapping into two subtasks: placement and routing.

1) **Placement:** The goal is to find a subgraph isomorphism \( \pi : V \mapsto P \) which respects \( (v_i,v_j) \in E_v \implies (\pi(v_i),\pi(v_j)) \in E_p \). If it exists, then we call \( \pi \) a perfect placement and the mapping task requires only a relabeling of the virtual qubits to physical qubits, i.e., replace each virtual qubit \( v_i \) in the unmapped circuit by a physical qubit \( \pi(v_i) \). Otherwise, we must resort to routing.

2) **Routing:** Given an initial placement, a router transforms a circuit so that all two-qubit instructions operate on physically adjacent qubits. Thus, our goal becomes finding a way of mapping a given circuit on a given device architecture with low overhead, whether in the number of additional instructions or depth of the resulting circuit. All routers implemented in *tweedledum* guarantee the compilation of any quantum circuit to any architecture represented as a simple connected graph. They are, therefore, completely hardware agnostic. These routing algorithms iteratively construct a new circuit that conforms to the desired architecture constraints.

**D. Optimization**

With the limited space and time resources available on current quantum devices, aggressive circuit optimization techniques are essential to extract all performance out of the machines. They often play a crucial role in whether a circuit can or cannot execute in a device or a simulator. Furthermore, they provide more accurate resource estimates, which might guide quantum algorithms and hardware development.

*tweedledum* provides circuit optimizations as a set of orthogonal compilation passes that can be composed into compilation flows. Optimization techniques range from purely structural, relying only on the relationship between instructions on a quantum circuit representation, to purely functional, e.g., when they rely on resynthesizing a circuit from a unitary matrix.

---

**Fig. 5:** Different ways of decomposing a multiple controlled X instruction (a) depending on the number of ancillae available.
There exists a significant trade-off between the quality of results and scalability. On the one hand, structural transformations offer better scalability at the cost of inferior quality of results. On the other hand, functional optimizations offer the contrary: better quality of results and poor scalability.

a) Instruction cancellation: As the name implies, this pass performs basic instruction cancellation: it traverses a circuit and removes pairs of adjacent joint instructions. This optimization is purely structural, and its effectiveness is highly dependent on the canonicalization of the instructions.

b) Phase folding: This optimization pass extends [17]'s $T$-count optimization algorithm to enable merging parameterized phase operators and handling arbitrary operators. The implementation handles operators not belonging to set $\{X, CNOT, SWAP, P(\theta)\}$ conservatively. It ignores their phase contribution: this pass only keeps track of phase polynomial terms that are trivially mergeable.

c) Linear resynthesis: The instruction cancellation pass suffers from structural bias: the input/output relationship between instructions (the structure) strongly influences the quality of results. Resynthesis techniques circumvent this problem by traversing the circuit searching subcircuits that they know how to represent functionally and synthesize in the linear resynthesis pass, we first identify linear subcircuits and represent their functionality as a binary matrix. Then we try to find a less (or equally) costly implementation by resynthesizing a new subcircuit using linear_synthesis algorithm.

VII. CONCLUSION

It is widely believed that a language’s expressive power influences the depth at which people can think. Existing quantum computing programming languages and frameworks limit the kinds of control structures, data structures, and abstractions developers can use; thus, the forms of algorithms they can construct are likewise limited. This paper introduced tweedledum—a library that augments the expressive power of current frameworks by providing methods for synthesis, compilation, and optimization of quantum circuits. The library also seeks to create an open-source environment where state-of-the-art quantum compilation techniques can be developed and compared.

Finally, we see that the field of quantum software is evolving rapidly, driven by various factors, ranging from progress in quantum hardware to improved compilation techniques. While we cannot foresee what the future will hold, the flexible design of tweedledum presents many possibilities for future improvements.

ACKNOWLEDGMENTS

We thank Mathias Soeken for valuable discussions regarding the library implementation and the development of its various algorithms. This research was partially supported by the ERC project H2020-ERC-2014-ADG 669354 CyberCare.

REFERENCES


