Accuracy and Compactness in Decision Diagrams for Quantum Computation

Alwin Zulehner*, Philipp Niemann†, Rolf Drechsler‡, and Robert Wille*

*Institute for Integrated Circuits, Johannes Kepler University Linz, Austria
†Cyber-Physical Systems, DFKI GmbH, Bremen, Germany
‡Department of Computer Science, University of Bremen, Bremen, Germany

alwin.zulehner@jku.at, philipp.niemann@dfki.de, drechsle@informatik.uni-bremen.de, robert.wille@jku.at

Abstract—Quantum computation is a promising research field since it allows to conduct certain tasks exponentially faster than on conventional machines. As in the conventional domain, decision diagrams are heavily used in different design tasks for quantum computation like synthesis, verification, or simulation. However, unlike decision diagrams for the conventional domain, decision diagrams for quantum computation as of now suffer from a trade-off between accuracy and compactness that requires parameter fine-tuning on a case-by-case basis. In this work, we—for the first time—describe and evaluate the effects of this trade-off. Moreover, we propose an alternative approach that utilizes an algebraic representation of the occurring irrational numbers and outline how this can be incorporated in a decision diagram in order to overcome this trade-off.

I. INTRODUCTION

Quantum computation [1] is a promising computation paradigm that—by exploiting quantum phenomena like superposition, entanglement, and phase shifts—allows for substantial speedups compared to conventional computers for certain problems. While the basic idea of quantum computation as well as corresponding algorithms with remarkable speed-ups are around for several decades [2], [3], physical realizations recently gained a new momentum with frequent “breakthroughs”, e.g. in increasing the number of available qubits and their rapidly improving fidelity [4]. In order to be prepared for future quantum devices, also research on automated design methods for quantum computations is underway. Since quantum computations are usually described in terms of exponentially large state vectors and unitary matrices, this often leads to rather intractable solutions when using straight-forward representations like, e.g., 1- and 2-dimensional arrays [5], [6], [7].

Motivated by that, alternative representations are currently investigated. Inspired by the conventional domain, decision diagrams are considered a promising approach for the efficient representation of quantum computations [8], [9], [10]. The key idea of decision diagrams is to exploit redundancies to gain a more compact representation that is non-exponential in many practically relevant cases. In combination with manipulation algorithms whose complexity grows polynomially with respect to the size of the decision diagram, this allows to efficiently conduct certain design tasks. In this regard, especially Quantum Multiple-valued Decision Diagrams (QMDDs) [10] and its variations serve as promising representative—leading to a broad variety of efficient approaches e.g. for synthesis [11], [12], [13], verification [14], [15], and simulation [16].

However, current decision diagrams for the quantum domain suffer from a trade-off between accuracy and compactness, since (1) small errors that are inevitably introduced by the limited precision of floating-point arithmetic can harm the compactness (i.e. the size of the decision diagram) significantly, and (2) overcompensating these errors (to increase compactness) may lead to an information loss and introduces numerical instabilities (cf. Section III). This motivates the need for an alternative solution that inherently achieves accuracy and compactness at once and, thus, allows to overcome the application-specific trade-off present in current solutions.

In this work, we propose such an alternative approach that utilizes an algebraic number field (within the complex numbers) that is strongly connected with the well established Clifford+T gate library. This allows to represent the considered quantum functionality algebraically rather than numerically—thereby completely avoiding the trade-off between accuracy and compactness. In fact, the proposed decision diagram can fully exploit redundancies for a compact representation and, at the same time, guarantees a perfectly accurate result.

This paper is structured as follows: In Section II, we review the basics of quantum computation as well as corresponding decision diagrams with a particular focus on Quantum Multiple-valued Decision Diagrams. In Section III, we discuss the trade-off that is necessary when using decision diagrams in the quantum domain—motivating the need for a decision diagram that represents complex numbers algebraically. In Section IV, we illustrate how to achieve this algebraic representation and how these findings can be exploited in decision diagrams in order to achieve both, perfect accuracy and compactness. Section V concludes the paper.

II. BACKGROUND

This section briefly reviews the basics of quantum computation and Quantum Multiple-valued Decision Diagrams (QMDDs). For a more detailed introduction, we refer to [1] and [10], respectively.

A. Quantum Computation

Quantum systems are composed of qubits that can be in one of the basis states $|0\rangle$ and $|1\rangle$, or in a superposition of both, i.e. $\alpha|0\rangle + \beta|1\rangle$ where $\alpha$ and $\beta$ are complex values with $|\alpha|^2 + |\beta|^2 = 1$. Accordingly, an $n$-qubit quantum system can be in one of the $2^n$ basis states $(|0\ldots0\rangle, |0\ldots1\rangle, \ldots, |1\ldots1\rangle)$ or a superposition of these states. The state of such a quantum system is represented by a state vector of dimension $2^n$. By the postulates of quantum mechanics, the evolution of a quantum state due to a quantum operation can be described by a unitary transformation matrix of dimension $2^n \times 2^n$.

Example 1. Commonly used quantum operations include the Hadamard operation $H$ (setting a qubit into a superposition), the NOT operation $X$ (flipping the basis states $|0\rangle$ and $|1\rangle$), as well as the phase shift operations $T$ ($\pi/4$ gate), $S = T^2$.
(Phase gate) and $Z = S^2$. The corresponding unitary matrices are defined as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & \omega \end{pmatrix},$$

where $\omega = \frac{1 + \sqrt{5}}{2} = e^{i\pi/4}$. Besides these operations that are applied to a single target qubit, there are also controlled operations on multiple qubits. The state of the additional control qubits determines whether the operation is performed on the target qubit.

Complex, high-level quantum operations like e.g. quantum computations are realized in terms of a sequence of elementary quantum operations (so-called quantum gates). The unitary matrix of the entire high-level operation can then be computed as the matrix product of the individual gate matrices (in reversed order).

B. Decision Diagrams for Quantum Computation

As state vectors and unitary matrices corresponding to quantum states and operations grow exponentially with the size of the quantum system (i.e., the number of qubits), straight-forward representations like 1- or 2-dimensional arrays quickly become infeasible (e.g. those proposed in [5], [6], [7]). One dedicated data-structure that allows for a more compact representation and efficient computation by exploiting redundancies in the vectors/matrices is provided by Quantum Multiple-Valued Decision Diagrams (QMDDs, [10]). The general idea of QMDDs is to represent a (unitary) matrix in terms of a directed acyclic graph such that sub-matrices which occur multiple times are represented by a shared graph structure. While there are several data-structures that follow a similar approach (e.g. those proposed in [8], [9]), only QMDDs additionally make use of weighted edges. This unique property allows them to share structures also for submatrices that differ by a scalar factor—a frequently occurring case.

Example 2. Fig. 1a shows the transformation matrix of the quantum operation $U = H \otimes I_2$, i.e. a Hadamard operation is performed on one qubit of a 2-qubit quantum system. Fig. 1b shows a decision diagram representation of this matrix. Here, the single root node (labeled $q_0$) represents the whole matrix and has four outgoing edges to nodes representing the top-left, top-right, bottom-left, and bottom-right sub-matrix (from left to right). Likewise the $2 \times 2$ sub-matrices (represented by nodes labeled $q_1$) are decomposed until the terminal nodes are reached—each of which represents a distinct number.

Apparently, the top-left, top-right and bottom-left sub-matrices of the original matrix are identical and can be represented by a shared graph structure (the left-most node labeled $q_1$ in Fig. 1b). However, the bottom-right sub-matrix is represented by a separate graph structure, although it has the same structure and differs only by a scalar factor of $-1$. If this similarity is taken into account (as it is done in QMDDs), an even more compact representation can be achieved by extracting such scalar factors and annotating them to the corresponding edges. Then, a single node at the $q_1$ level is sufficient as shown in Fig. 1c. Here, the common factor $\frac{1}{\sqrt{2}}$ is extracted and annotated to the root edge (an additional edge pointing to the root node). For simplicity, edge weights equal to 1 are usually suppressed and edges with a weight of 0 are indicated by stubs.

```
(a) Unitary matrix
(b) DD w/o edge weights
(c) QMDD
```

To obtain the value of a particular matrix entry, one has to follow the corresponding path from the root to the terminal node and multiply all edge weights on this path. For example, the matrix entry $\frac{1}{\sqrt{2}}$ from the bottom-left sub-matrix of Fig. 1a (highlighted bold) can be determined as the product of the weights on the highlighted path of the QMDD in Fig. 1c.

In order to determine the common scalar factors that allow for the enhanced use of shared structures, the nodes of a QMDD are normalized. To this end, a normalization factor is determined (usually the left-most non-zero weight of the outgoing edges). This factor is then applied to all outgoing edges and also propagated to all incoming edges. By doing so, QMDD even become canonical, i.e. unique, representations of (unitary) matrices [10].

III. ACCURACY VS. COMPACTNESS IN QMDDS

QMDDs have successfully been applied for design tasks like synthesis [11], [12], [13], verification [14], [15], and simulation [16] since they allow for a compact representation and, thus, an efficient processing of the respective matrices and vectors. This compactness is achieved by exploiting redundancies such that the entire information of the matrix/vector is preserved.

However, while gaining a compact representation without information loss is not complicated as long as the set of possible values is finite (e.g. 0 and 1 for Boolean functions) or discrete (e.g. only integer numbers occur), this is different for the quantum domain, where we have to deal with arbitrary complex numbers and, thus, irrational coefficients. As a consequence, decision diagrams that represent conventional computations like e.g. Binary Decision Diagrams (BDDs, [17]), Kronecker Functional Decision Diagrams (KFDDs, [18]), or Binary Moment Diagrams (BMDs, [19]) do not face problems with the accuracy of the representation, while accuracy can be an important issue for decision diagrams representing quantum computations. In this section, we discuss the resulting challenges and issues that lead to a trade-off between accuracy and compactness—motivating the need for an algebraic representation of quantum computations overcoming this.

To this end, first note that, in the area of quantum computation, most design automation tasks require hundreds or even thousands of multiplications of unitary matrices (e.g. to compute the unitary matrix for an entire quantum circuit from the gate matrices) or multiplications of a vector and a unitary matrix (e.g. to simulate the evolution of a quantum state). From a numerical perspective, these tasks do not constitute an issue per se, since the multiplication with a unitary matrix

1In this regard, state vectors are interpreted as matrices of dimension $2^n \times 1$.

2Using the left-most edge weight with the largest absolute value as normalization factor increases the numerical stability of the representation while preserving canonicity.
is a well-conditioned problem. In fact, the resulting error, i.e. the deviation from the exact result, can be expected to be in the order of the input error. Furthermore, applying several multiplications successively will only lead to an error that grows linearly with the number of matrix multiplications. Consequently, using a numerical, i.e. approximated, representation of the complex numbers with a high resolution can yield numerically stable computations.

However, this approximation can have a significant impact on the decision diagram representation. To this end, recall that the key idea of decision diagrams is to exploit redundancies in order to gain a compact representation. This compact representation is indeed a key factor for their efficiency, since the complexity of the manipulation algorithms (e.g. matrix multiplication) grows with the size of the decision diagram. By using approximate representations of the complex numbers, the detection of these redundancies can become a tough challenge. An example demonstrates the problem.

**Example 3.** Recall Example 2 where the matrix shown in Fig. 1a can compactly be represented by the QMDD shown in Fig. 1c. This compact representation is possible since several redundancies can be exploited. However, representing the irrational entries with floating point numbers on a machine with limited accuracy, may break these redundancies e.g. when using rounding towards $\infty$ or when the matrix is constructed as the product of several other matrices. Then, two occurrences of $\pm \frac{\sqrt{2}}{2}$ might be represented by slightly different floating point numbers (e.g. differing in the last significant bit of the mantissa) and no redundancy can be detected anymore.

In general, this will likely lead to a decision diagram where none of the existing redundancies are detected at all—leading to an exponentially large representation. A solution to this issue (due to tiny errors caused by machine accuracy) is to identify numbers that do not differ by more than a so-called **tolerance value** (denoted as $\epsilon$ in the following).

**Example 4.** Assume that two floating point numbers that shall represent $\frac{1}{\sqrt{2}}$ differ only in the last three bits of the mantissa (assuming an IEEE 754 single precision floating point number with 23 mantissa bits). Then, setting e.g. $\epsilon = 10^{-5}$ allows to detect that the two entries are equal.

However, choosing a proper value for $\epsilon$ is crucial. If $\epsilon$ is chosen too small, it might not be able to compensate the limited machine accuracy and, thus, to determine more redundancies. If $\epsilon$ is chosen too large, additional redundancies might be “detected” that are not actually present—leading to an undesired approximation (information loss) and numerical instabilities of the multiplication algorithm. In the worst case, this may falsify the result such that an invalid quantum state (e.g. a vector composed of zeros only) or a non-unitary matrix results. Nevertheless, in many cases there exist proper configurations for $\epsilon$, but this heavily depends on the considered application and determining an adequate tolerance value may require time-consuming fine-tuning of parameters on a case-by-case basis.

**Example 5.** Fig. 2 shows the size of the state vector (by means of QMDD nodes) while simulating the Ground State Estimation (GSE, [20]) quantum algorithm. As can be seen, the number of QMDD nodes is highly affected by $\epsilon$. Choosing $\epsilon = 0$, i.e. (almost) no two different numbers are considered to be equal, yields the highest precision that is possible using floating point numbers, but results in a rather large representation. Instead, choosing $\epsilon = 10^{-3}$ yields a vector composed of zeros only—a perfectly compact but obviously wrong result. As a trade-off, choosing $\epsilon = 10^{-15}$ leads to almost the same numerical result as $\epsilon = 0$, but yields a better compactness and, thus, a smaller run-time.

Overall, determining a “perfect” $\epsilon$, i.e. finding the best trade-off between accuracy and compactness (which heavily influences the run-time), is a non-trivial task. So far, it has to be evaluated on a case-by-case level for each application.

**IV. PROPOSED SOLUTION**

In this section, we propose to overcome the trade-off discussed and evaluated above by using an algebraic representation of the complex numbers that occur in QMDDs as solution—allowing for both, a perfect accuracy together with a perfect exploitation of redundancies.

**A. Utilizing the Ring $\mathbb{D}[\omega]$**

In order to obtain an algebraic representation of the complex numbers, the most obvious choice would be to extend the well-known Gaussian numbers $\mathbb{Z}[i]$ to the ring $\mathbb{Z}[i, \sqrt{2}]$. By doing so, all complex numbers of the form $a + bi + 2i(c + d\sqrt{2})$ can be represented in an exact fashion. This ring is already a dense subset of the complex numbers such that any complex number can be approximated by an element from $\mathbb{Z}[i, \sqrt{2}]$ up to an arbitrary precision (this density is a known property of $\mathbb{Z}[\sqrt{2}]$ in the real numbers and can easily be lifted to the complex numbers). However, the irrational number $\frac{1}{\sqrt{2}}$ that plays a vital role in quantum computation, is not contained in this ring. Thus, it seems more promising to study the ring $\mathbb{Z}[i, \frac{i}{\sqrt{2}}]$ which trivially contains $\mathbb{Z}[i, \sqrt{2}]$ (since $\sqrt{2} = 2 \cdot \frac{1}{\sqrt{2}}$), but allows to represent $\frac{1}{\sqrt{2}}$ and all its potencies exactly.

In the following, we will make use of a different interpretation of this ring that is more convenient from an algebraic perspective. More precisely, we will use the interpretation as an extension of the so-called dyadic fractions $\mathbb{D} = \{ \frac{a}{2^k} \mid a, k \in \mathbb{Z}, k \geq 0 \}$, namely $\mathbb{D}[\omega]$ (where $\omega = \frac{1 + i}{\sqrt{2}} = e^{i\pi/4}$ as in Example 1).

Using the latter representation, all complex numbers that can be represented exactly can be written as $\alpha = \frac{1}{2^k} (a\omega^k + bw^k + cw + d)$ for coefficients $a, b, c, d, k \in \mathbb{Z}$, i.e. using five integers (cf. [5]).

Note that the ring $\mathbb{D}[\omega]$ is also strongly related to the well established Clifford+T gate library [21]. This library is very popular in quantum computation due to its **universality** (any quantum operation, i.e. any unitary transformation matrix,

3Note that this is a statement about the matrix multiplication problem itself and not about a certain implementation.
can be realized up to an arbitrarily small error) as well as fault-tolerance. The most elementary gates in this library are the Clifford group gates (H, CNOT, S) and the T gate as discussed in Example 1. The relation between the ring $\mathbb{D}[\omega]$ and the Clifford+T gate library is that the quantum operations which can be realized exactly by Clifford+T gates (i.e. without any rounding error) are precisely given by those matrices whose entries are from the ring $\mathbb{D}[\omega] = \mathbb{D}[\sqrt{2}, i]$ (as shown in [5]). As a consequence, all such quantum operations can be represented with perfect accuracy using our approach. Hence, $\mathbb{D}[\omega]$ provides the ideal basis for a decision diagram that employs an accurate, algebraic representation of complex numbers.

B. Incorporating $\mathbb{D}[\omega]$ into QMDDS

In order to use the algebraic representation of complex numbers presented above within QMDDS, there are two aspects that have to be taken into account:

1) In order to determine common factors and structural similarities (that are required to find redundancies), a unique representation of $\mathbb{D}[\omega]$ numbers is required. However, there are in general infinitely many possibilities to represent a $\mathbb{D}[\omega]$ number.

2) The extracted (normalization factors) have to be applied to the edge weights. More precisely, the weights have to be divided by these factors. However, as division means multiplication by the (multiplicative) inverse, this division can only be conducted properly for $\mathbb{D}[\omega]$ numbers that indeed have a multiplicative inverse in $\mathbb{D}[\omega]$, but not for $\mathbb{D}[\omega]$ numbers in general (e.g. all odd integers greater than or equal to 3 do not have an inverse in $\mathbb{D}[\omega]$ and the result of a division by such a number cannot be represented as a $\mathbb{D}[\omega]$ number).

Recall that each number $\alpha \in \mathbb{D}[\omega]$ can be written as

$$\alpha = \frac{1}{\sqrt{2}} (a\omega^3 + b\omega^2 + c\omega + d)$$

for coefficients $a, b, c, d, k \in \mathbb{Z}$. If the exponent $k$ is fixed, the representation is clearly unique since two different representations would yield a non-trivial representation of 0 in $\mathbb{Z}[\omega]$. Thus, a unique representation can be achieved when using the smallest denominator exponent $k_{\text{min}}$ such that there is no representation with an exponent $k < k_{\text{min}}$. The existence of such an exponent has already been discussed in [5], but no constructive criterion for minimality has been derived. To this end, we note that $\sqrt{2} = -\omega^3 + \omega$, such that

$$\alpha = \frac{1}{\sqrt{2}} \left( (b-d)\omega^3 + (c+a)\omega^2 + (b+d)\omega + (c-a) \right)$$

where $a', b', c', d' \in \mathbb{Z}$ if, and only if, $a = c \mod 2$ and $b = d \mod 2$. Thus, we know that the exponent is minimal if, and only if, $a \neq c \mod 2$ or $b \neq d \mod 2$—yielding a constructive algorithm to obtain unique representations of $\mathbb{D}[\omega]$ numbers.

Regarding the division by normalization factors, we exploit the fact that a similar argumentation as above can be performed for the field $\mathbb{Q}[\omega]$ as well. In fact, each $\mathbb{Q}[\omega]$ number has a unique representation as $\frac{1}{\sqrt{2}}$ where $\alpha \in \mathbb{D}[\omega]$ and $\epsilon$ is an odd integer ($\epsilon \in 2\mathbb{Z} + 1$) that is co-prime to the integer coefficients of $\alpha$. Having this, all computations can be made in the field $\mathbb{Q}[\omega]$ where all non-zero numbers have a multiplicative inverse, in summary, by spending one additional integer and switching to the algebraic number field $\mathbb{Q}[\omega]$, a division/normalization becomes possible.

Overall, the proposed solution allows for an algebraic and, thus, perfectly accurate representation of complex numbers within the QMDD data-structure. Since this allows to detect all existing redundancies, our solution allows to overcome the currently existing trade-off between accuracy and compactness.

V. Conclusions

In this work, we have—for the first time—discussed and evaluated the trade-off between accuracy and compactness of decision diagrams for quantum computation. Since this requires fine-tuning of parameters on a case-by-case basis and might still yield useless results, we propose to overcome this issue by an algebraic decision diagram. The proposed algebraic representation guarantees perfect accuracy while remaining compact (all redundancies that are actually present are detected). Future work covers the evaluation of the proposed solution regarding computational and memory overhead—particularly with respect to the bitwidth of the integers required for an algebraic representation.

References


