A heuristic to reduce $\pi/4$-parity-phase circuits

Niel de Beaudrap
Department of Computer Science
University of Oxford
Oxford, UK
niel.debeaudrap@cs.ox.ac.uk

Quanlong Wang
Department of Computer Science
University of Oxford
Oxford, UK
quanlong.wang@cs.ox.ac.uk

Abstract

To approximate arbitrary unitary transformations on one or more qubits, one must perform transformations which are outside of the Clifford group, which is only a discrete set of transformations. The gate most commonly considered for this purpose is the $T = \mathrm{diag}(1, \mathrm{e}^{i\pi/4})$ gate. As T gates are computationally expensive to perform fault-tolerantly in the most promising error-correction technologies, minimising the "T count" (the number of T gates) required to realise a given unitary in a Clifford+T circuit is of great interest. Following Heyfron and Campbell [9], one can transform the circuit into a form involving a CNOT+T circuit, followed by rounds of Pauli measurements and classically controlled Clifford gates. One may further transform the CNOT+T portion into a CNOT circuit, followed by a product of "$\pi/4$-parity-phase" operations: diagonal unitary transformations which induce a relative phase of $\mathrm{e}^{i\pi/4}$ depending on the outcome of a parity computation on the standard basis (which the "phase gadgets" of Kissinger and van de Wetering [10] can easily represent). We present a heuristic technique to simplify circuits of $\pi/4$-parity-phase operations, based on simple identities for such operations. This results in a corresponding reduction of the T-count to realise a unitary circuit. These techniques can be extended to reduce non-Clifford phase gates at various levels of the Clifford hierarchy, by simplifying "$\pi/2^k$-parity-phase" operations for non-negative integers k.

1 Introduction

An important goal of quantum technologies is to realise, as faithfully as possible, an architecture capable of performing approximately universal quantum computation. Ignoring the practical difficulties of imperfectly realised operations and noise in imperfect hardware, such an architecture must be able to approximate an arbitrary unitary transformation with high probability, possibly relative to some embedding of the computational space into the states of its qubits (*e.g.*, to perform error correction).

The above goal requires at minimum that the set of transformations that the architecture can perform do not form a discrete set. This is challenging, as the operations which can be easily performed fault-tolerantly for various error correcting codes form a discrete set, often the Clifford group or a subset of it. As the Clifford group is in any case useful to reason about quantum error correction and very simple procedures on quantum data, this motivates the approach of (a) considering fault-tolerant realisations of the Clifford group, together with a more labour-intensive procedure to realise some unitary transformation outside of the Clifford group, and then (b) minimising the number of non-Clifford gates required to realise or approximate a given unitary. The most popular approach is to consider "Clifford+T" circuits, using a gate-set such as $\{\mathrm{CNOT}, H, S, T\}$, involving CNOT, the Hadamard gate H, and $S = \mathrm{diag}(1, i)$ as generators of the Clifford group, supplemented by the gate $T = \mathrm{diag}(1, \mathrm{e}^{i\pi/4}) = \sqrt{S}$. We then consider the problem of minimising the *T count* of a unitary transformation, *i.e.*, of minimising the number of T gates to realise (or approximate) that unitary.

Submitted to:
QPL 2019

Heyfron and Campbell [9] show that, by simulating H transformations using X observable measurements, controlled-Z operations, and CNOT operations, one can realise a Clifford+T unitary using a circuit consisting of a circuit of CNOT operations, a circuit of diagonal non-Clifford operations, and a sequence of (possibly classically controlled) Clifford operations. This allows them to reduce the problem of T-count reduction to an appropriate analysis of the diagonal non-Clifford portion of this circuit. This can be described as a product of what in this article we call "$\pi/4$-parity-phase" operations, which induce a relative phase of $\mathrm{e}^{i\pi/4}$ on standard basis states depending on the outcome of some parity computation $f(x) = x_{i_1} \oplus x_{i_2} \oplus \cdots \oplus x_{i_m}$, for any string $x \in \{0,1\}^n$, integer $m \geqslant 1$, and $1 \leqslant i_1, i_2, \ldots, i_m \leqslant n$. Each $\pi/4$-parity-phase gate can be realised in principle using a single T or $T^\dagger$ gate, and a circuit of CNOT gates: simplifying the subcircuit of $\pi/4$-parity-phase gates is directly productive to reducing the T count.

The strategy of Heyfron and Campbell [9] represent a strategy of simplification by considering various kinds of *phase polynomials*, and builds on a sequence of results which revolve around such operations [4, 8, 3, 2, 6, 5] presented in various but similar ways. Most recently, Kissinger and van de Wetering [10] consider simplification of circuits by means of analysis in terms of "phase gadgets" representing unitaries generalising $\pi/4$-parity-phase operations (with arbitrary relative phases in place of $\mathrm{e}^{i\pi/4}$). These results all recognise the difficulty of optimising the T count, and related problems, which is NP-hard in general: they approach the problem of systematically minimising T-count by means of approaches to solving other known hard problems to which reducing phase polynomials may be reduced.

In this article, we describe simple identities of $\pi/4$-parity-phase operations. We then make a relatively elementary observation that reducing the T-count amounts to a combinatorial problem of minimising the size of a collection $\{S_1, S_2, \ldots, S_t\}$ of such subsets, subject to representing the same operation. These observations allow us to obtain decompositions of unitiaries with smaller T-count than any previous results, without sophisticated mathematical techniques. They also appear to generalise straightforwardly to the problem of reducing the count of operations $\sqrt[r]{S}$, for $r = 2^k$ for $k > 0$, which would be required to realise a circuit of $\pi/2^{k-1}$-parity-phase operations.

2 Preliminaries

We set out some basic or existing results, using the following notation. Let $[n] := \{1, 2, \ldots, n\}$ and $\mathbb{1}$ be the 2×2 identity matrix. For sets $S, T \subseteq V$ we write $S \Delta T$ for the symmetric difference $(S \cup T) \setminus (S \cap T)$, and $\mathbf{x}^{(S)} \in \{0,1\}^V$ denote the incidence vector of S, where $x_j^{(S)} = 1$ if and only if $j \in S$.

2.1 The Clifford hierarchy

Let $\mathscr{P}^n := \left\{ i^k P_1 \otimes \cdots \otimes P_n \,\middle|\, k \in \mathbb{Z} \;\&\; P_j \in \{\mathbb{1}, X, Y, Z\} \right\}$ denote the n-qubit Pauli group. We define the Clifford hierarchy (on n qubits) by defining $\mathscr{C}_1^n := \mathscr{P}_n$, and

$$\mathscr{C}_k^n = \left\{ U \in \mathrm{U}_n(\mathbb{C}) \,\middle|\, \forall P \in \mathscr{P}^n.\, UPU^\dagger \in \mathscr{C}_{k-1}^n \right\} \tag{1}$$

for $k > 1$. We then define $\mathscr{D}_k^n \subseteq \mathscr{C}_k^n$ to be the subset of diagonal operations. As an abuse of notation, we will identify $\mathscr{C}_k^n$ and $\mathscr{D}_k^n$ with subsets of $\mathscr{C}_k^N$ and $\mathscr{D}_k^N$ (respectively) for $n < N$. As a part of this abuse of notation, we allow ourselves to write $S \in \mathscr{C}_2^n$ and $T \in \mathscr{C}_3^n$ for all $n \geqslant 1$.

2.2 Parity-phase operations

Defining parity-phase operations. It is easy to show that $\mathscr{D}_k^n$ forms an abelian group. In particular, one can show (see Appendix A) that $\mathscr{D}_k^n$ is generated by the operators $\omega \cdot \mathbb{1}^{\otimes n}$ for any global phase ω,

together with all operations of the form $D_{S,k}$ for sets $S = \{s_1, \ldots, s_m\} \subseteq [n]$ for $m \geqslant 1$, defined by

$$D_{S,k} \;=\; \exp\Big(-\frac{i\pi}{2^k}\big(Z_{s_1} \otimes \cdots \otimes Z_{s_m}\big)\Big) \;=\; \exp\Big(-\frac{i\pi}{2^k} Z_S\Big) \;=\; \cos\big(\tfrac{\pi}{2^k}\big)\mathbb{1} - i\sin\big(\tfrac{\pi}{2^k}\big) Z_S\,, \tag{2}$$

where $Z_S = \bigotimes_{j\in S} Z_j$. (We define $D_{S,k}$ for all $k \in \mathbb{Z}$; however, one may show $D_{S,0} = -\mathbb{1}^{\otimes n}$ and $D_{S,-k} = \mathbb{1}^{\otimes n}$ for all $k > 0$ and $S \subseteq [n]$.) Note that $X_j Z_S X_j^\dagger = (-1)^{x_j^{(S)}} Z_S$, and $\mathrm{CNOT}_{h,j}\, Z_S\, \mathrm{CNOT}_{h,j}^\dagger = Z_{S'}$ such that

$$S' \;=\; \begin{cases} S \,\Delta\, \{h\}, & \text{if } j \in S; \\ S, & \text{otherwise.} \end{cases} \tag{3}$$

From this it follows that

$$X_j\, D_{S,k}\, X_j^\dagger \;=\; D_{S,k}^{-1} \;\in\; \mathscr{D}_k^n \tag{4a}$$

if $j \in S$ (and $X_j D_{S,k} X_j^\dagger = D_{S,k}$ otherwise); and

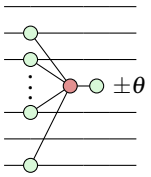
$$\mathrm{CNOT}_{h,j}\, D_{S,k}\, \mathrm{CNOT}_{h,j}^\dagger \;=\; D_{S',k} \;\in\; \mathscr{D}_k^n \tag{4b}$$

so that $\mathscr{D}_k^n$ is preserved under conjugation by CNOT and X operations. Also note that $D_{S,k}^2 = D_{S,k-1}$, from which it follows that $\mathscr{D}_{k-1}^n \subseteq \mathscr{D}_k^n$.

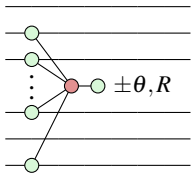
We refer to operations $D_{S,k+1}$, and their inverses, as "$\pi/2^k$-parity-phase" operations. We motivate this terminology as follows. Let $S = \{s_1, s_2, \ldots, s_m\}$ for some $m \geqslant 1$. For a standard basis vector $|z\rangle$, we have $Z_S\,|z\rangle = (-1)^{\mathbf{x}^{(S)}\cdot z}\,|z\rangle$, where we define $\mathbf{x}^{(S)} \cdot z = \sum_i x_i^{(S)} z_i$. From this it follows that

$$D_{S,k+1}\,|z\rangle = \begin{cases} \exp(-i\pi/2^{k+1})\,|z\rangle\,, & \text{if } \mathbf{x}^{(S)}\cdot z = 0; \\ \exp(+i\pi/2^{k+1})\,|z\rangle\,, & \text{if } \mathbf{x}^{(S)}\cdot z = 1. \end{cases} \tag{5}$$

This is equivalent (up to a global phase of $\mathrm{e}^{-i\pi/2^{k+1}}$) to inducing a relative phase of $\pi/2^k$ on $|z\rangle$ for those $z \in \{0,1\}^n$ for which $\mathbf{x}^{(S)}\cdot z = z_{s_1} \oplus z_{s_2} \oplus \cdots \oplus z_{s_m} = 1$; and similarly for $D_{S,k+1}^{-1}$. More generally, we refer to $\exp(\pm\frac{1}{2} i\theta Z_S)$ as a θ-parity-phase operation. We note that θ-phase parity operation, the operators $D_{S,k}$ among them, can be represented by ZX diagrams with the usual denotational semantics (read from left to right in this article), with structure such as the following:



$\Big($or



if conditioned on the parity of a set of bits $R\Big)$ (6)

where the long horizontal wires are the qubits indexed by $[n] = \{1, 2, \ldots, n\}$ and $S \subseteq [n]$ is the subset of those qubits which have (green) degree-3 copy nodes on them. In the right-hand diagram, R denotes a set of boolean variables $s_i \in \{0,1\}$: using the extended annotations of Ref. [7], the diagram denotes that the phase applied is $\pm\theta$ only if $\bigoplus_{s_i \in R} s_i = 1$, and that otherwise the phase is zero. We refer to these as "phase gadgets", adopting the terminology of Ref. [10, Section 4.3]; when $|S| = m$, we may refer to it as a "phase m-gadget". (If θ is an odd multiple of $\pi/4$, we may refer to it as a "T-phase m-gadget"; for θ an integer multiple of $\pi/2$, we refer to it as a "Clifford-phase m-gadget". If $m = 1$, we may also mildly abuse this terminology to refer to a simple green phase node as a "1-gadget".)

Remark. We identified ZX diagrams such as Eqn. (6) as objects of interest early in our work on this topic, independently of Ref. [10]. However, these objects arise from the ZX calculus so naturally that it is reasonable to suppose that any serious investigation of $\mathscr{D}_k^n$ — if it involved the ZX calculus — would quickly identify such diagrams as a structure of central significance.

Parity-phase operations in relation to controlled phases. An important role of $D_{S,3}$ gates for $S \subseteq [n]$ is their relationship to diagonal gates in $\mathscr{D}_3^n$ which are controlled-unitaries of a more straightforward sense, such as CS and CCZ:

$$\begin{aligned} \mathrm{C}S &= |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes S \\ &= \exp\Big(\frac{i\pi}{2}\,|11\rangle\langle 11|\Big), \end{aligned} \qquad \begin{aligned} \mathrm{CC}Z &= \Big(|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|\Big) \otimes \mathbb{1} + |11\rangle\langle 11| \otimes Z \\ &= \exp\Big(i\pi\,|111\rangle\langle 111|\Big); \end{aligned} \tag{7}$$

we may describe how to generate these from $D_{k,3}$ operations by decomposing the projectors $|11\rangle\langle 11|$ or $|111\rangle\langle 111|$ into tensor products of $|1\rangle\langle 1| = \frac{1}{2}\big(\mathbb{1} - Z\big)$, and expanding to obtain a product of $D_{S,3}$ gates (see Eqn. (25) in Appendix A): disregarding the $D_{\emptyset,3}$ factors, which realise global phases, we obtain

$$\mathrm{C}S_{h,j} \propto D_{\{h\},3} D_{\{j\},3} D_{\{h,j\},3}^{-1}\,; \qquad \mathrm{CC}Z_{g,h,j} \propto D_{\{g\},3} D_{\{h\},3} D_{\{j\},3} D_{\{g,h\},3}^{-1} D_{\{g,j\},3}^{-1} D_{\{h,j\},3}^{-1} D_{\{g,h,j\},3}\,. \tag{8}$$

More generally, we may relate $(t-1)$-controlled $\pi/2^k$-phase gates to $\pi/2^{k-t+1}$-phase parity gates, following Eqn. (25):

$$\prod_{\substack{S \in \wp(V) \\ S \neq \emptyset}} D_{S,k}^{(-1)^{|S|}} \propto \exp\Big(\frac{i\pi}{2^{k-|V|+1}}\,|1\rangle\langle 1|^{\otimes V}\Big), \tag{9}$$

where the right-hand operator applies a phase of $\pi/2^{k-|T|-1}$ to those components of a state in which all of the qubits in T are in the state $|1\rangle$. A corollary of this, on which our results depend, is that

$$\prod_{\substack{S \in \wp(V) \\ S \neq \emptyset}} D_{S,k}^{(-1)^{|S|}} \propto \mathbb{1}^{\otimes V}, \quad \text{so that} \quad D_{V,k} \propto \prod_{\substack{S \in \wp(V) \\ S \notin \{\emptyset, V\}}} D_{S,k}^{(-1)^{|V|-|S|+1}}, \qquad \text{for } |V| > k. \tag{10}$$

Connection between parity-phase operations and T-count. From Eqn. (4b), it follows that any operation $D_{S,k}$ can be reduced to an operation $D_{j,k} \propto \mathrm{diag}(1, \mathrm{e}^{2\pi i/2^k})$ acting on a single qubit j, by conjugation with an appropriate CNOT circuit. In particular, it follows that any $D_{S,3}$ circuit has minimal T-count 1. This allows us to approach the question of reducing T count by considering decompositions of unitaries involving few $\pi/4$-parity-phase operations, acting on many qubits.

3 Reduction of T-count through simplification of parity-phase circuits

In this section, we apply the observation of Eqn. (10) to describe simplifications which can reduce the T-count necessary to realise a unitary $\mathscr{D}_3^n$ operation.

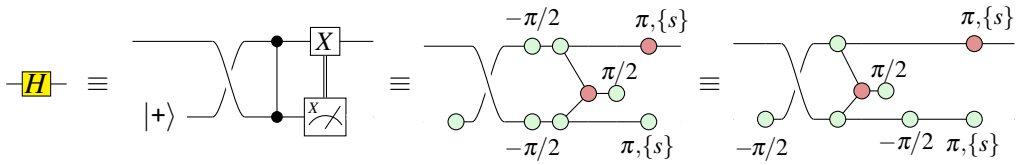
Remark. It will be convenient to us to represent circuits involving parity-phase operations using phase-gadgets (*i.e.*, ZX diagrams) as depicted in Eqn. (6). However, we do not make significant use of the rewriting system of the ZX calculus in our results; apart from our use of phase gadgets, we essentially retain the structure of unitary circuits throughout.

3.1 Realising unitaries with large layers of diagonal operators

We may use a series of circuit transformations, generalising and inspired by that of Heyfron and Campbell [9], to reduce the amount of non-Clifford diagonal operations used to realise a unitary U.

Let $\mathbf{C}$ be a circuit realising U, initially expressed as a Clifford+$\mathscr{D}_k^n$ circuit, *e.g.*, using the gate-set $\big\{\mathrm{CNOT}, X, Z, H, S\big\} \cup \big\{\sqrt[r]{S_j} \;\big|\; r = 2^{t-2},\, t \in \{3,4,\ldots,k\}\big\}$. We transform $\mathbf{C}$ as follows:

1. Rewrite the H gates in the interior of the circuit, using the following circuit / ZX gadgets (using a fresh classical bit label in place of "s" below, each time):



(11)

In the circuit second from the left, the two qubits are subject to a SWAP operation, followed by a $CZ = \exp\big(i\pi\,|11\rangle\langle 11|\big)$ operation. The bottom qubit is measured finally with an X observable measurement (*i.e.*, in the $|\pm\rangle$ basis), and the top operation is acted on finally by an X operation only if the outcome is $|-\rangle$. On the right are annotated ZX diagrams in the style of Ref. [7], in which the measurement is represented as a projection with a random outcome s, which governs the presence of phase operations elsewhere. The leftmost ZX diagram describes the decomposition of the controlled-Z operation, using $CZ_{h,j} \propto D_{\{h,j\},2} D^{-1}_{\{h\},2} D^{-1}_{\{j\},2}$. The final ZX diagram propogates the single-qubit $D^{-1}_{\{*\},2}$ operations towards the preparation and measurement of the second qubit, so that the second qubit is prepared in the $|-\mathrm{y}\rangle \propto |0\rangle - i\,|1\rangle$ state, and (if we absorb the final rotation into the measurement) is measured in the Y eigenbasis, with outcome $s \in \{0,1\}$ on a measured state of $\frac{1}{\sqrt{2}}\big(|0\rangle + i\cdot(-1)^s\,|1\rangle\big)$.

2. Decompose / rewrite operations (possibly with classical controls) as follows, for any $h, j \in [n]$ and for $r = 2^{t-2}$ where $3 \leqslant t \leqslant k$:

$$Z_j \propto D_{\{j\},1} \qquad Y_j \propto X_j D_{\{j\},1} \qquad S_j \propto D_{\{j\},2} \qquad \sqrt[r]{S_j} \propto D_{\{j\},t} \tag{12}$$

3. Commute all X operations — *i.e.*, red π-phase nodes (both unconditioned and classically conditioned) — to the end of the circuit. Note that:
   - X operations immediately before X measurements do not affect the outcome, and may be absorbed into the measurement itself;
   - Commuting X operations past CNOT operations may induce further X operations (which may be classically controlled); we also commute these to the end.
   - Commuting X_j operations (without classical control) past $D_{S,t}$ operations, for $j \in S$, inverts the $D_{S,t}$ operation by Eqn. (4a).
   - Commuting X_j operations (*with* classical control) past $D_{S,t}$ operations, for $j \in S$, induces a classically controlled $D_{S,t-1}$ operation, as from Eqn. (4a) we have

$$D_{S,t} X_j^{s_j} \;=\; X_j^{s_j} D_{S,t}^{(-1)^{s_j}} \;=\; X_j^{s_j} D_{S,t}^{-2s_j} D_{S,t} \;=\; X_j^{s_j} D_{S,t-1}^{-s_j} D_{S,t}\,. \tag{13}$$

4. Commute all SWAP and CNOT operations towards the beginning of the circuit. This may transform various $D_{S,t}$ gates By Eqn. (4b), changing the set S involved without changing the level t of the Clifford hierarchy to which it belongs.

5. Commute any classically-controlled $\mathscr{D}_{S,t}$ gates towards the end of the circuit (but before the X operations), in layers according to the classical control bit involved, and collect together all parity-phase operations acting on a given set of qubits into a single parity-phase operation. Similarly, commute all of the unconditional $\mathscr{D}_{S,t}$ gates towards the beginning of the circuit (but after the CNOT and SWAP operations). Simplify these using $D^2_{S,k} = D_{S,k-1}$ and Eqn. (10). Repeat this step until no further simplifications are possible.

If the original circuit $\mathbf{C}$ had m Hadamard gates, the above procedure realises a transformation

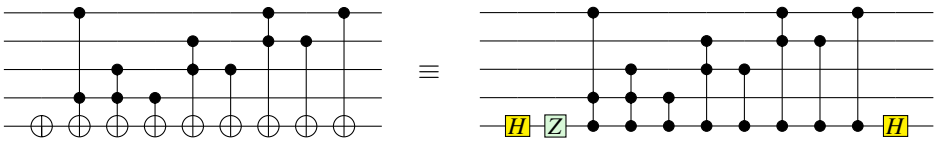
$$\mathbf{C} \longrightarrow \mathbf{H}_1 \mathbf{X}_1 \mathbf{D}_m \cdots \mathbf{D}_2 \mathbf{D}_1 \mathbf{D}_0 \mathbf{X}_0 \mathbf{H}_0 \,, \tag{14}$$

where $\mathbf{H}_0$ prepares a collection of qubits in the $|{-}\mathrm{y}\rangle \propto |0\rangle - i\,|1\rangle$ state (denoted in ZX by a green $-\pi/2$ node) and possibly performs Hadamard gates on the input qubits; $\mathbf{X}_0$ is a circuit of SWAP and CNOT gates; $\mathbf{D}_0$ is a circuit realising a $\mathscr{D}_k^n$ operation; the circuits $\mathbf{D}_j$ (for $1 \leqslant j \leqslant m$) consist of the j^{th} measurement in the $|0\rangle + i \cdot (-1)^{s_j}\,|1\rangle$ basis with outcome s_j (denoted in ZX by a green "$-\pi/2 + \pi, \{s_j\}$" node), followed by $\mathscr{D}_k^n$ operations conditioned on the outcome s_j; $\mathbf{X}_1$ is a circuit consisting of all X operations (classically controlled or not); and $\mathbf{H}_1$ possibly performs some Hadamard gates on the output qubits. We refer to this process as "putting the circuit into HXDXH form". (We recommend pronouncing "HXDXH" as "hexadexah".)

In the above, all of the non-Clifford operations consist of the operations from $\mathscr{D}_t^n$ for $3 \leqslant t \leqslant k$, in the various layers $\mathbf{D}_j$ for $0 \leqslant j \leqslant m$. This motivates the question of how to simply circuit consisting entirely of $D_{S,t}$ operations. In particular, for $k = 3$, reducing the number of $D_{S,3}$ operations involved in such a circuit represents a decrease in the T-count involved in realising the operation.

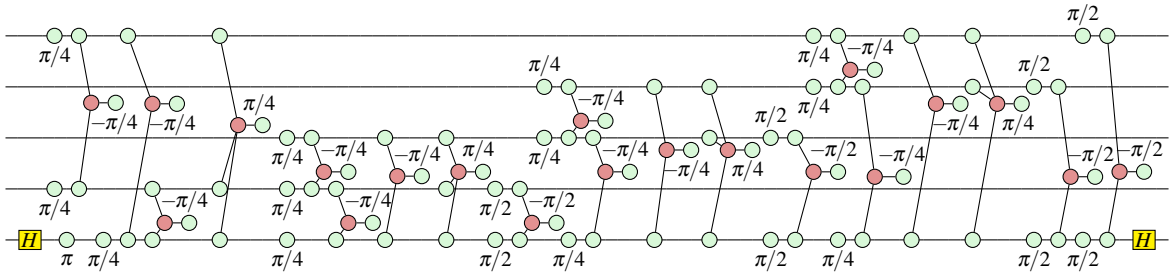
3.2 Example of T-count reduction by conversion to HXDXH form

As an illustrative example, we consider the "Mod5$_4$" circuit on five qubits, which computes $|x\rangle\,|y\rangle \mapsto |x\rangle\,|y \oplus f(x)\rangle$ for $f : \{0,1\}^4 \to \{0,1\}$ such that $f(x) = 1$ if and only if $x \in \{0,1\}^4$ is the binary representation of a multiple of 5. The circuit on the left, below, loosely follows the circuit provided by Ref. [1]. We obtain the right-hand circuit by decomposing each of the (multiply-)controlled NOT gates using CZ and CCZ operations, taking advantage of the fact that all operations use the bottom qubit as a target.



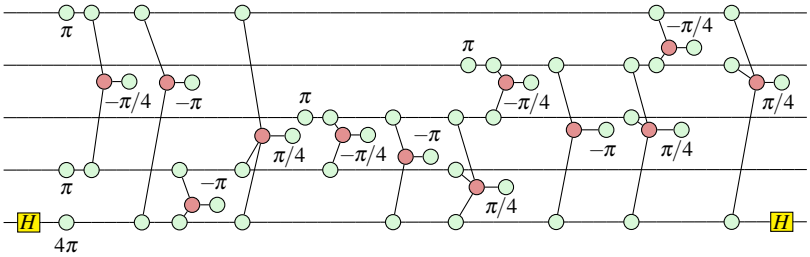
(15)

The right-hand circuit is already nearly in HXDXH form, apart from decomposing CZ and CCZ into $\mathscr{D}_3^n$ gates, which yields the following circuit:



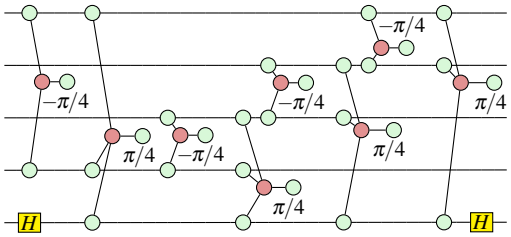
(16a)

We may simplify by collecting together all operations acting on the same set of qubits, *e.g.*, those acting on $\{5\}$, those acting on $\{4,5\}$, and so forth. This yields following circuit:



(16b)

Using the identity $D_{\{a,b\},1} = D_{\{a\},1} D_{\{b\},1}$ and reducing angles mod 2π, we obtain a final simplified circuit:



(16c)

Remark. The circuit above has T-count 8, in contrast to 16 achieved by the algorithm TODD in Ref. [9]. In principle, the same simplification was available through the techniques of Ref. [9], which is also based on parity-phase operations — albeit expressed mostly in terms of "phase polynomial" operators, $U_f := \prod_z \exp\left(\frac{i\pi}{4} f(z) \, |z\rangle\langle z|\right)$, for some polynomial

$$f(z) \;=\; \sum_h \ell_h z_h + 2\sum_{h<j} q_{hj} z_h z_j + 4\sum_{h<j<k} c_{hjk} z_h z_j z_k \tag{17}$$

with integer coefficients ℓ_h, q_{hj}, and c_{hjk}. An advantage of pursuing an analysis of circuits in which all operators of $\mathscr{D}_3^n$ (and more generally $\mathscr{D}_k^n$) are decomposed directly into parity-phase operations, is that simplifications such as the one above come to the fore, and may possible be easier to realise reliably.

3.3 Phase Gadget Elimination tactics

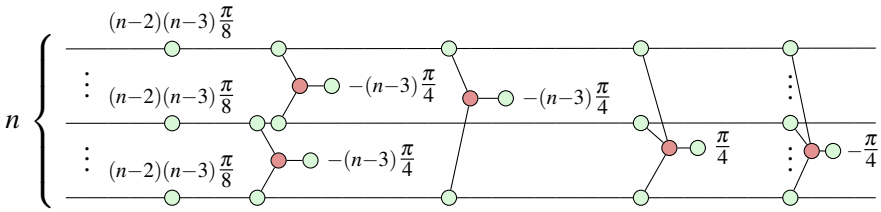
We describe a framework in which to describe reduction of T-count in circuits consisting entirely of $\mathscr{D}_3^n$ operations, and present a specific example which may form the basis of a heuristic to reduce T-count.

Approaches to reduce the T-count while preserving the meaning of a circuit consists of the suitable application of some mathmatical identity, possibly up to scalar factors, and possibly passing through different representations of these circuits. These are often identities of diagonal unitary circuits [3, 5]), though not always [8, 10]. In the special case of reductions by identities on diagonal unitaries, they may in principle be described in terms of a commuting product of operations which are proportional to the identity operator.

Equipped with such an identity on products of diagonal operations, one may attempt to find a collection S of operators, which represents more than half of some set of operators V which are involved in such an identity. One may then substitute the set S with the *inverses* of the operators in $V \setminus S$, simplifying the circuit. When such an identity concerns products of phase gadgets, we describe such a tactic as a Phase Gadget Elimination (or PHAGE) tactic. In principle, the T-optimise subroutine of Ref. [5] and the TOOL and TODD subroutines of Ref. [9] may be interpreted as algorithms to deploy one or more PHAGE tactics. This approach to T-count reduction can be distinguished from that of Ref. [10], in which phases may in principle be aggregated one at a time in circuits which are unitary but not diagonal. While such techniques seem fruitful, we suggest that investigation of identities on parity-phase operations — and the way in which such identities may be deployed as PHAGE tactics — may provide a complementary approach to reduce the T-count.

A spider's nest identity. As evidence in support of this proposal, we now consider a PHAGE tactic which makes use of a specific (and to our knowledge, new) identity involving a single T-phase n-gadget

for $n \geq 4$, and phase k-gadgets with $k \leq 3$:



$$\propto \mathbb{1}^{\otimes n} \quad ; \qquad (18)$$

Let G_n denote the n-qubit circuit on the left-hand side of Eqn. (18). This consists of a 1-gadget with phase angle $(n-2)(n-3)\frac{\pi}{8}$ on each line, a 2-gadget on each pair of lines with phase angle $-(n-3)\frac{\pi}{4}$, and a 3-gadget with phase angle $\frac{\pi}{4}$ on each subset of three lines, and finally an n-gadget with phase angle $-\frac{\pi}{4}$. (We prove this identity in Appendix B.) We refer to this identity (and any other identity involving one large phase-gadget "spider" together with many smaller "spiders") as a *spider nest* identity. For the identity above, we make the following observations:

- If $n = 4$, then it is essentially the same as the rule R_{13} given in [2], and also Eqn. (10).
- If $n \equiv 1 \pmod 4$ or $n \equiv 3 \pmod 4$, all of the 2-gadgets in Eqn. (18) are Clifford-phase gadgets, which do not contribute to the T-count.
- If $n \equiv 3 \pmod 4$ or $n \equiv 2 \pmod 4$, all of the 1-gadgets in Eqn. (18) are Clifford-phase gadgets, which again do not contribute to the T-count.

Let $\mathbf{T}_n$ denote the T-count of the left-hand-side of Eqn. (18), excluding the final n-gadget: then

$$\mathbf{T}_n = \begin{cases} \frac{1}{6}n(n^2+5), & \text{for } n \equiv 0 \pmod 4; \\ \frac{1}{6}n(n^2-3n+8), & \text{for } n \equiv 1 \pmod 4; \\ \frac{1}{6}n(n^2-1), & \text{for } n \equiv 2 \pmod 4; \\ \frac{1}{6}n(n^2-3n+2), & \text{for } n \equiv 3 \pmod 4. \end{cases} \qquad (19)$$

A spider nest PHAGE. For a fixed value of n, and a T-phase gadget on 1 to 3 qubits, we say that the phase gadget is *usable* (in the spider's nest identity above) if the identity involves T-phase gadgets of that size. Thus, T-phase 3-gadgets are usable for all n; T-phase 2-gadgets are only usable if n is even; and T-phase 1-gadgets are only usable if $n \equiv 0, 1 \pmod 4$. These represent the phase gadgets which we may use in a PHAGE tactic (*i.e.*, to reduce the T-count) by means of the spider's nest identity Eqn. (18) for a given value of n. For the sake of definiteness, we describe this tactic below:

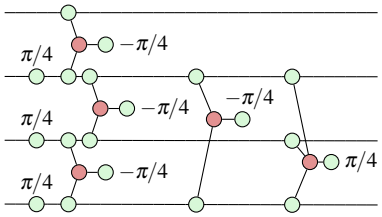
PHAGE TACTIC ("PHAGE A"). *Let V be a set of n qubits, and C be a $\mathscr{D}_3^n$ circuit on V, realised by $\pi/4$-parity-phase operations (represented by phase gadgets) acting on different subsets of qubits. Let S be the number of usable T-phase gadgets acting on V: if $|S| > \frac{1}{2}(\mathbf{T}_n+1)$, then:*

1. *Add the circuit (or the inverse of) G_n described in the left-hand side Eqn.* (18) *to C.*
2. *Simplify the resulting circuit by cancelling phase gadgets on subsets of S, or more generally accumulating them to produce Clifford-phase gadgets.*

In particular: if all of the elements of S are T-phase gadgets which occur in G_n itself, we may substitute the phase-gadgets of S with the inverses of all of the other phase-gadgets in G_n.

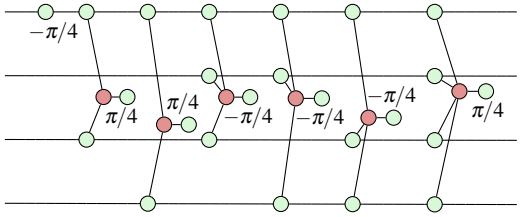
We describe this (and other such substitutions) as a "tactic", in the sense that performing a substitution of this sort for a specific set of phase gadgets is just one circuit transformation that one may apply as part of a higher order algorithm (or "strategy") for T-count reduction. Further below, we consider a heuristic which uses the PHAGE A tactic as a subroutine.

A simple example. Consider the following circuit, with a T-count of 8:

$C \;=$ 

(20a)

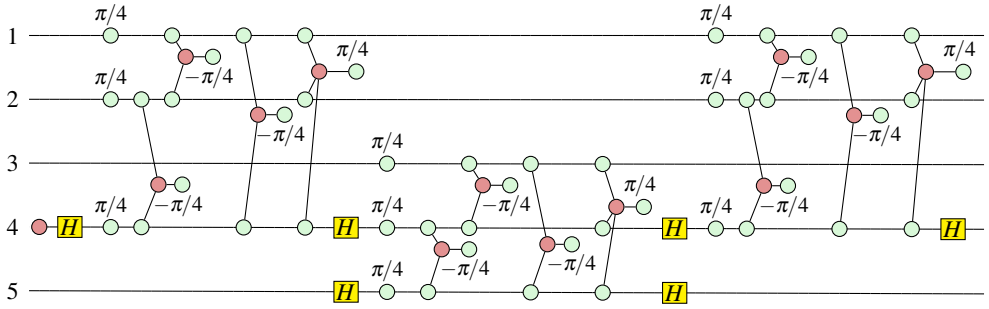
For $n = 4$, all T-phase gadgets on 1–3 qubits are usable for the PHAGE A tactic. As the gadget G_4 has T-count $\mathbf{T}_4 + 1 = 15$, which is less than twice the number of T-phase gadgets in C above, we may apply PHAGE A to this circuit to produce a circuit with a T-count of $\mathbf{T}_4 - 8 = 7$ (shown below, using the inverses of those gadgets of G_4 which are absent from C above):

$C \;\cong$ 

(20b)

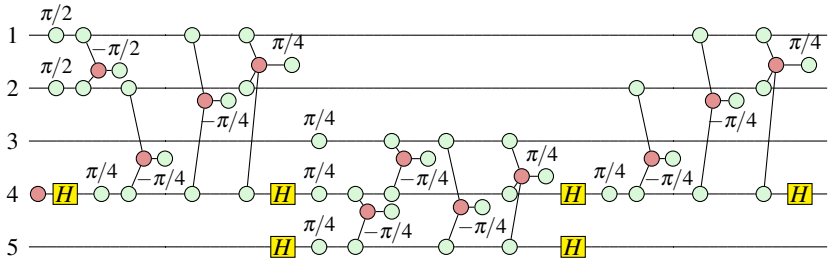
3.4 A practical example of T-count reduction: the triply-controlled NOT

Here we give an example of T-count reduction using the PHAGE A tactic described above. The triply-controlled NOT gate on four qubits (Tof_3) can be constructed using one auxiliary qubit using three Toffoli gates, which leads to the following expression in the ZX calculus:



. (21a)

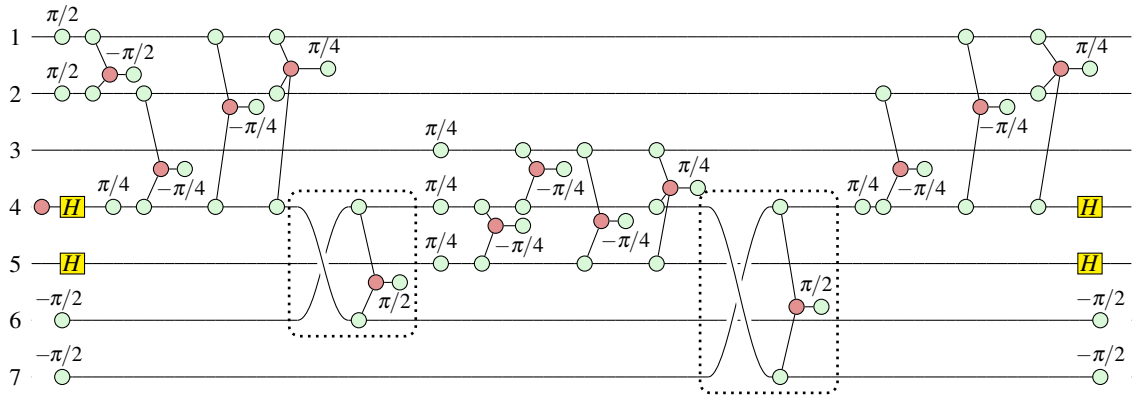
A few of the $\pi/4$-parity-phase operators can be gathered together, allowing us to simplify this to



. (21b)

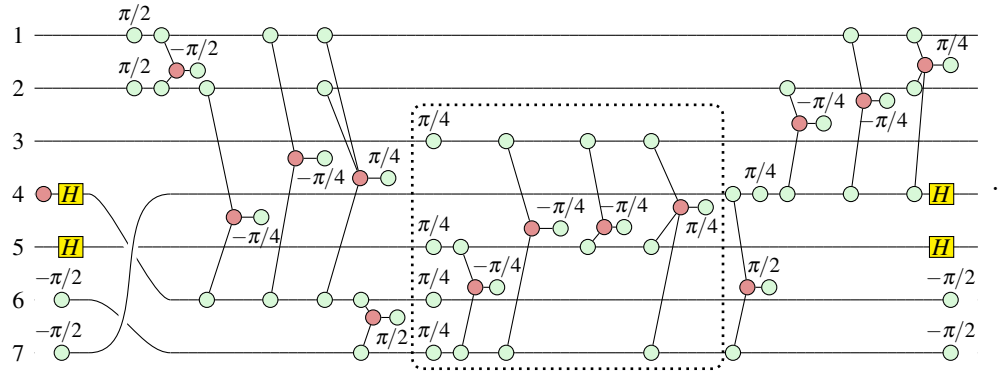
This latter circuit has a T-count of 15, matching the outcome of using the methods presented in Ref. [10] using PyZX and PyZX+TODD.

To reduce the T-count further, we put the circuit into HXDXH form (as outlined in Section 3.1) and look for opportunities to apply the PHAGE A tactic. We first substitute interior H gates on qubit 4 with the gadget of Eqn. (11), as indicated below with dotted lined boxes. (Depending on the outcome of the Y observable measurements — represented by green $-\pi/2$ projection nodes — we may need to perform some classically-controlled $\pi/2$-phase parity operations. As these operations can be easily computed and do not contribute to the T count, we suppose for the sake of brevity that these measurements yield the result $|0\rangle + i|1\rangle$, which will allow us to focus on the task of simplifying $\pi/4$-parity-phase operations.)



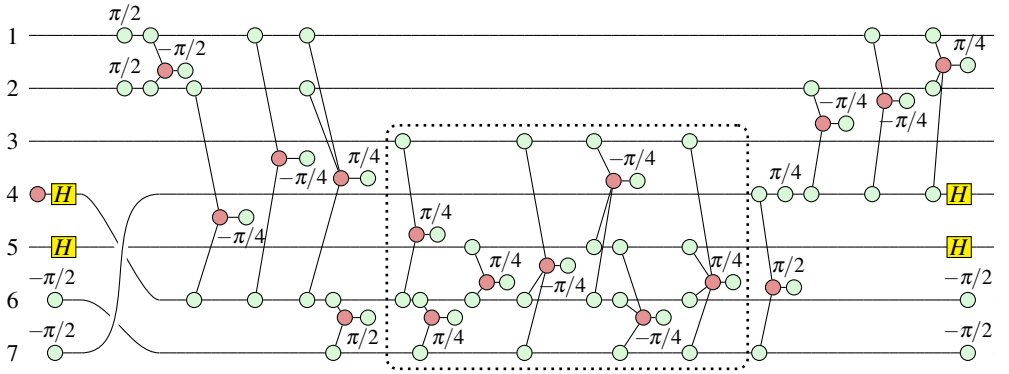
(21c)

We next commute the SWAP gates towards the beginning of the circuit, and collect T-phase operations on common sets of qubits, to put the circuit into HXDXH form:



(21d)

The dotted box contains the operations which act only on the lines $\{3,5,6,7\}$, including eight different T-phase gadgets. As this is greater than $\lfloor(\mathbf{T}_4+1)/2\rfloor = 7$, deploying the PHAGE A tactic (substituting the gates in the dotted box for the above circuit, with those in the circuit below) yields a circuit with a reduced T-count of 14:



(21e)

4 Discussion

4.1 General observations

It seems to us that the ZX calculus not only lends itself to analysis in terms of parity-phase operations, but also leads directly to the idea of analysing T-count in terms of the parity-phase operations and phase gadgets. This is particularly the case when considering circuit transformations such as those of Ref. [9] which isolate a layer of diagonal operators by commuting CNOT gates past them.

It is not difficult to see how the proof that $\mathscr{D}_k^n$ is generated by the operators $D_{S,k}$ for $S \subseteq [n]$ (see Appendix A) may lead one to Fourier analysis as the appropriate tool with which to consider the problem of optimising T count, or the corresponding generalisations to higher levels of the Clifford hierarchy. It may be worthwhile to consider whether identities such as the spider's nest identity of Eqn. (18) have any fruitful connection to Fourier analysis.

Finally, much of our analysis clearly generalises beyond the case of reduction of T-count (as a measure of the complexity of a $\mathscr{D}_3^n$ circuit), to simplifications of $\mathscr{D}_k^n$ circuits. We expect that simple generalisations of Eqn. (18) would provide the opportunity to explore more general simplification of diagonal circuits.

4.2 Ongoing work

This submission represents work in progress. In the results presented above:

- We demonstrate how representing circuits using parity-phase operations can lead directly to simplifications which reduce T-count;
- We describe a programme of T-count reduction by the application of "PHAGE tactics", which are transformations which take advantage of an identity involving parity-phase operations;
- We describe a new identity of this sort, and demonstrate how using this as the basis of a PHAGE tactic can allow us to easily obtain a T-count reduction for a practically important circuit, with results comparable to the state of the art [9, 10].

In future work, we intend to describe an explicit heuristic for T-count reduction, by incorporating the PHAGE A tactic into a procedure which attempts to discover situations in which it can be applied. The following is a sketch of such an approach.

A heuristic to reduce T-count. To quickly locate ways in which we can use the PHAGE A tactic to good effect, we must be able to find "dense" collections of usable T-phase gadgets. There are two obstacles to this task:

- For an initial circuit $\mathbf{C}$ involving many CNOT gates, its HXDXH form may involve relatively few T-gadgets on 3 or fewer qubits (instead having T-phase m-gadgets for moderately large m).
- The subsets on which we could in principle apply the PHAGE A tactic grows exponentially with the number of qubits, in principle requiring exponential time to even determine if any opportunities exist at all to apply PHAGE A.

These concerns motivate using the spider's nest identity of Eqn. (18). Any collection of phase m-gadgets which are not essentially independent of one another must have some significant overlap: this motivates measuring the *density of T-phase gadgets* at each qubit q — which we define by

$$d(q) := \sum_{k \geqslant 1} \frac{\#(T\text{-phase } k\text{-gadgets which act on } q)}{k} \,. \tag{22}$$

We also define $d_3(q)$, the *3-max density (of T-phase gadgets)*, which is the same sum but for $1 \leqslant k \leqslant 3$. It is easy to show that $d_3(q) \leqslant \left(\frac{1}{18} + O(1/n)\right) \cdot n^3$; on any qubit or collection of qubits where $d(q)$ significantly exceeds this bound, there must be several T-phase m-gadgets for $m > 3$, and it may be helpful to apply Eqn. (18) to decompose these into gadgets on at most 3 qubits. Having ensured that the circuit does not have an obvious excess of large T-gadgets, we may then attempt to apply PHAGE A on any large collections of "usable" gadgets that we can find on subsystems of different sizes. Such sets of gadgets will usually consist mostly of 3-gadgets, as $\mathbf{T}_n = \frac{1}{6}n^3 \pm O(n^2)$. The following Lemma follows easily by comparing $\frac{1}{2}(\mathbf{T}_n + 1)$ to the maximum number of usable T-gadgets for each value of n:

Lemma 1. *Let S be a set of T-phase gadgets, each of which acts on 1, 2, or 3 qubits out of a set of $n \geqslant 5$ qubits. If $|S| > \frac{1}{2}(\mathbf{T}_n + 1)$, the number of 3-gadgets in S must be at least $\mathbf{T}'_n$, where*

$$\begin{aligned} \mathbf{T}'_n &:= \tfrac{1}{12}(n^3-6n^2-n+6), \text{ for } n \equiv 0 \pmod 4; & \mathbf{T}'_n &:= \tfrac{1}{12}(n^3-6n^2+5n+6), \text{ for } n \equiv 2 \pmod 4; \\ \mathbf{T}'_n &:= \tfrac{1}{12}(n^3-3n^2-4n+6), \text{ for } n \equiv 1 \pmod 4; & \mathbf{T}'_n &:= \tfrac{1}{12}(n^3-3n^2+2n+6), \text{ for } n \equiv 3 \pmod 4. \end{aligned} \tag{23}$$

This suggests a strategy along the following lines:

1. Compute the 3-density, 2-density, and 1-density of T-phase gadgets acting on each qubit (*i.e.*, the $k \in \{1,2,3\}$ terms of Eqn. (22)). Determine the largest integer $N \geqslant 5$, such that the sum of the N largest 3-densities is at least $\mathbf{T}'_N$. (If no $N \geqslant 5$ satisfies this, then let $N = 4$.)
2. For each $k \in \{4,5,\ldots,N\}$, compute the *score* for each qubit as the sum of the densities of those m-gadgets (for $1 \leqslant m \leqslant 3$) which are useful.
3. Again for each k, rank each qubit in order of descending score, and compute $r(k)$ to be the "lowest" rank such that the sum of the scores of the qubits ranked $\{1, r(k)-k+2, r(k)-k+3, \ldots, r(k)\}$ is at least $\frac{1}{2}(\mathbf{T}_k + 1)$. Then, let $M(k)$ be the sum of the scores of the qubits ranked from 1 to $r(k)$, so that $M(k)$ is proportional to the average total score of a uniformly random subset of these qubits.
4. Let $M := \sum_{1 \leqslant k \leqslant N} M(k)$, and repeat the following poly(n) times:
   a. Randomly select an integer $4 \leqslant k \leqslant N$, with probability distribution governed by $p(k) = M(k)/M$.
   b. Attempt to find one or more as-yet undiscovered subset of size k, from the $r(k)$-highest ranking qubits (according to their scores for k), which is acted on by more than $\mathbf{T}_k + 1$ usable T-gadgets. We may attempt to do so, *e.g.*, by a breadth-first search in the hypergraph of usable T-gadgets from a randomly selected edge.
   c. If successful in finding such a set, compute the *value* of this set, as the difference of the number of usable T-gadgets and $(\mathbf{T}_k + 1)$.
5. If any set with positive value was found, perform the PHASE A tactic on the set with the largest value, and simplify the circuit by collecting together phase gadgets on equal sets of qubits.

This strategy can be made to run in polynomial time, and can itself be repeated a number of times, potentially creating new opportunities to apply the PHASE A tactic when a substitution is performed. We intend to perform numerical experiments on a number of benchmark circuits to test the efficacy of such a procedure, comparing it to the state of the art in T-count reduction [9, 10].

Acknowledgements.

This work was performed in partnership with Cambridge Quantum Computing under the EPSRC Impact Acceleration Award "Compilation and cost-reduction of quantum computations via ZX-calculus". NdB is further supported by the EPSRC National Hub in Networked Quantum Information Technologies (NQIT.org). We would like to thank Earl Campbell and Luke Heyfron for helpful discussions, and XiaoNing Bian for help with quantum circuit diagrams.

References

[1] Matthew Amy: *T-par GitHub*. `https://github.com/meamy/t-par`.

[2] Matthew Amy, Jianxin Chen & Neil J. Ross (2018): *A Finite Presentation of CNOT-Dihedral Operators*. *Electronic Proceedings in Theoretical Computer Science* 266, pp. 84–97. [arXiv:1701.00140].

[3] Matthew Amy, Dmitri Maslov & Michele Mosca (2014): *Polynomial-Time T-Depth Optimization of Clifford+T Circuits Via Matroid Partitioning*. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 33(10), pp. 1476–1489. [arXiv:1303.2042].

[4] Matthew Amy, Dmitri Maslov, Michele Mosca & Martin Roetteler (2013): *A meet-in-the-middle algorithm for fast synthesis of depth-optimal quantum circuits*. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 32(6), pp. 818–830. [arXiv:1206.0758].

[5] Matthew Amy & Michele Mosca (2016): *T-count optimization and Reed–Muller codes*. [arXiv:1601.07363].

[6] Earl T. Campbell & Mark Howard (2017): *A unified framework for magic state distillation and multi-qubit gate-synthesis with reduced resource cost*. *Physical Review A* 95, p. 022316. [arXiv:1606.01904].

[7] Ross Duncan & Simon Perdrix (2010): *Rewriting Measurement-Based Quantum Computations with Generalised Flow*. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide & Paul G. Spirakis, editors: *Automata, Languages and Programming*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 285–296.

[8] David Gosset, Vadym Kliuchnikov, Michele Mosca & Vincent Russo (2014): *An Algorithm for the T-count*. *Quantum Info. Comput.* 14(15-16), pp. 1261–1276. Available at `http://dl.acm.org/citation.cfm?id=2685179.2685180`. [arXiv:1308.4134].

[9] Luke E. Heyfron & Earl T. Campbell (2018): *An efficient quantum compiler that reduces T count*. *Quantum Science and Technology* 4(1), p. 015004. [arXiv:1712.01557].

[10] Aleks Kissinger & John van de Wetering (2019): *Reducing T-count with the ZX-calculus*. [arXiv:1903.10477].

A Parity-phase operators as generators of $\mathscr{D}_k^n$

We show in this Section that, together with arbitrary global phases, the operators $D_{S,k} = \exp(-\frac{i\pi}{2^k} Z_S)$ for $S \subseteq [n]$ generate the group $\mathscr{D}_k^n$.

Let $\mathscr{M}_1^n = \mathscr{P}_n$, and for $k \geqslant 2$, let $\mathscr{M}_k^n$ consist of all products of elements of $\mathscr{D}_k^n$ with products of CNOT and X on various qubits. As $\mathscr{D}_k^n$ is preserved under conjugation by CNOT and X operations, it is easy to show that $\mathscr{M}_k^n$ forms a group for each k, and that in particular that operators in $\mathscr{M}_k^n$ decompose as a product $U_D U_X$ for $U_D \in \mathscr{D}_k^n$ and U_X a circuit of CNOT and X gates.

Lemma 2. *For $k \geqslant 2$ an integer, $\mathscr{D}_k^n \subseteq \mathscr{M}_k^n \subseteq \mathscr{C}_k^n$.*

Proof. For $k = 2$, elements of $\mathscr{M}_k^n$ are Clifford circuits by construction. For $k > 2$, consider $U = U_X U_D \in \mathscr{M}_k^n$, where U_X is a product of CNOT and X gates, and $U_D \in \mathscr{D}_k^n$. Then for any $P \in \mathscr{P}_n$, we have $UPU^\dagger = U_D U_X P U_X^\dagger U_D^\dagger = U_D Q U_D^\dagger$, for $Q = U_X P U_X^\dagger \in \mathscr{P}_n$. As $U_D Q U_D^\dagger \in \mathscr{C}_{k-1}^n$, the Lemma follows. □

Lemma 3. *For integers $n, k \geqslant 1$ and $S \subseteq [n]$, $D_{S,k} \in \mathscr{D}_k^n$.*

Proof. Note that $D_{S,1} = -iZ_{s_1} \otimes \cdots \otimes Z_{s_m} = -iZ_S \in \mathscr{P}_n = \mathscr{C}_1^n$ for any $S \subseteq [n]$. Also, by definition we have $D_{S,k-1} = D_{S,k+1}^2$ for any $k \geqslant 2$. By decomposing any Pauli operator $P \in \mathscr{P}_n$ into a product $P \propto X_A Z_B$ for sets $A, B \subseteq [n]$, it is easy to see that $D_{S,k} \in \mathscr{D}_k^n$: we have

$$\begin{aligned} D_{S,k} P D_{S,k}^{-1} \;=\; \exp\!\big(-\tfrac{i\pi}{2^k} Z_S\big)\, X_A Z_B \exp\!\big(\tfrac{i\pi}{2^k} Z_S\big) \;&=\; \exp\!\big(-\tfrac{i\pi}{2^k} Z_S\big) \exp\!\Big(\tfrac{i\pi}{2^k} X_A Z_S X_A^\dagger\Big) X_A Z_B \\ &=\; \exp\!\big(-\tfrac{i\pi}{2^k} Z_S\big) \exp\!\Big((-1)^{\mathbf{x}^{(S)}\cdot\mathbf{x}^{(A)}} \tfrac{i\pi}{2^k} Z_S\Big) X_A Z_B \\ &=\; \begin{cases} X_A Z_B, & \text{if } \mathbf{x}^{(S)}\cdot\mathbf{x}^{(A)} = 0, \\ \exp\!\big(-\frac{2\pi i}{2^k} Z_S\big) X_A Z_B, & \text{if } \mathbf{x}^{(S)}\cdot\mathbf{x}^{(A)} = 1; \end{cases} \end{aligned} \tag{24}$$

in either case, $D_{S,k} P D_{S,k}^{-1} \in \mathscr{M}_{k-1}^n \subseteq \mathscr{C}_{k-1}^n$. Then $D_{S,k} \in \mathscr{C}_k^n$, and is therefore an element of $\mathscr{D}_k^n$. □

Lemma 4. *For any $n, k \geqslant 1$, any $V \in \mathscr{D}_k^n$ is proportional to a product of operators $D_{S,k}$ for $S \subseteq [n]$.*

Proof. Consider a decomposition of V into a product of operators $V = \prod_z V_z$ for z ranging over $\{0,1\}^n$, where $\langle z|V_z|z\rangle = \langle z|V|z\rangle = \exp(i\theta_z)$ and where $\langle y|V_z|y\rangle = 1$ for all $y \neq z$. We may then express the operator V_z as an exponential of a rank-1 projector on n qubits:

$$\begin{aligned} V_z = \exp\Big(i\,\theta_z\big(|z_1\rangle\langle z_1| \otimes \cdots \otimes |z_n\rangle\langle z_n|\big)\Big) &= \exp\Big(\frac{i\theta_z}{2^n}\big[(\mathbb{1}+(-1)^{z_1}Z)\otimes\cdots\otimes(\mathbb{1}+(-1)^{z_n}Z)\big]\Big) \\ &= \prod_{S\subseteq[n]} \exp\Big(\frac{i\theta_z}{2^n}\bigotimes_{j\in S}(-1)^{z_j}Z_j\Big) = \exp\Big(\sum_{S\subseteq[n]} \frac{i(-1)^{z\cdot\mathbf{x}^{(S)}}\theta_z}{2^n} Z_S\Big). \end{aligned} \tag{25}$$

Taking the product over $z \in \{0,1\}^n$, we then have

$$V = \prod_{z\in\{0,1\}^n} V_z \;=\; \exp\Big(\sum_{S\subseteq[n]} i\hat{\theta}_S Z_S\Big), \tag{26}$$

where $\hat{\theta}_S = \sum_z (-1)^{z\cdot \mathbf{x}^{(S)}} \theta_z / 2^n$ for the sake of brevity. For $j \in [n]$, consider the effect of conjugation of X_j by V: we have

$$\begin{aligned} V X_j V^\dagger &= \exp\Big(\sum_{S\subseteq[n]} i\hat{\theta}_S Z_S\Big) \exp\Big(\sum_{S'\subseteq[n]} i\hat{\theta}_{S'} X_j Z_{S'} X_j^\dagger\Big) X_j \\ &= \exp\Big(\sum_{S\subseteq[n]} i\hat{\theta}_S \big[Z_S + X_j Z_S X_j^\dagger\big]\Big) X_j \;=\; \exp\Big(\sum_{\substack{S\subseteq[n]\\ j\in S}} 2i\hat{\theta}_S Z_S\Big) X_j \;=:\; U_{[j]} X_j \,. \end{aligned} \tag{27}$$

It follows that $U_{[j]} \in \mathscr{D}^n_{k-1}$, and that $U_{[j]}^{2^{k-2}}$ is a Pauli operator. That is, the operator

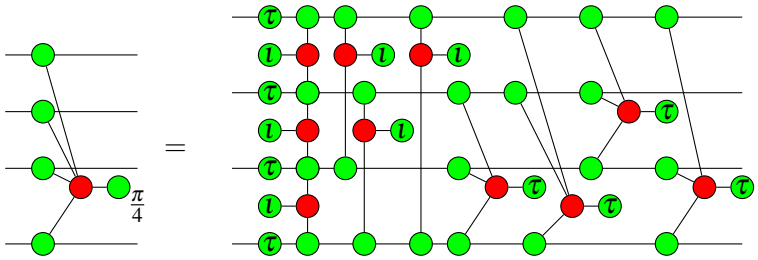
$$U_{[j]}^{2^{k-2}} = \exp\Big(\sum_{\substack{S\subseteq[n]\\ j\in S}} 2^{k-1} i\hat{\theta}_S Z_S\Big) = \prod_{\substack{S\subseteq[n]\\ j\in S}} \Big(\cos\big(2^{k-1}\hat{\theta}_S\big)\mathbb{1} + i\sin\big(2^{k-1}\hat{\theta}_S\big) Z_S\Big) \tag{28}$$

is a tensor product of Z operations. By the linear independence of the operators Z_S, it follows that every factor $\exp(2^{k-1} i\hat{\theta}_S Z_S)$ is either $\mathbb{1}$ or Z_S, for $j \in S$. As this result holds for all j, we obtain the same result for every non-empty set S. This implies that $2^{k-1}\hat{\theta}_S \in \frac{\pi}{2}\mathbb{Z}$ for all $S \neq \varnothing$, or equivalently that $\hat{\theta}_S = m_s \pi / 2^k$ for some $m \in \mathbb{Z}$. It follows that $\exp(i\hat{\theta}_S Z_S) = D_{S,k}^{-m_S}$, so that $V \propto \prod_S D_{S,k}^{-m_S}$ for S ranging over non-empty subsets of $[n]$. □

B Proof of gadget decomposition

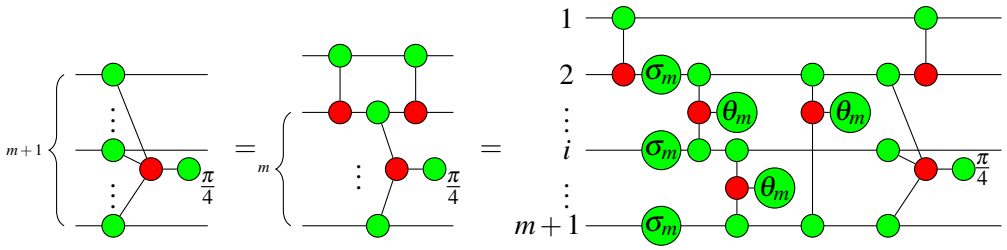
Here we provide a proof of Eqn. (18). We express this as a proof by induction or the proportionality (*i.e.*, the equality of the denotational semantics of ZX-diagrams) of a T-phase n-gadget for $n \geqslant 4$ on one side, and a collection of 3-, 2-, and 1-gadgets as in Eqn. (18) on the other. Below we use the notations $\tau := \frac{\pi}{4}, \iota := -\frac{\pi}{4}, + := \frac{\pi}{2}, - := -\frac{\pi}{2}$.

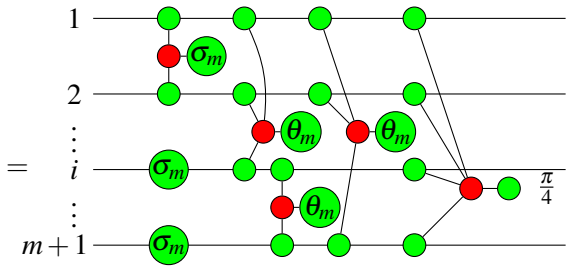
We prove (18) by induction on n, but we need to assume that this equality holds for $n = 4$, the same way as it was taken as a rule in [2]. This means, we have



(29)

where $\tau := \frac{\pi}{4}, \iota := -\frac{\pi}{4}$. Suppose (18) holds for $n = m, m \geq 4$. Let $\sigma_m = \frac{(m-2)(m-3)\pi}{8}, \theta_m = -\frac{(m-3)\pi}{4}$. Then for $n = m+1$, we have





In the last diagram above, we substitute every 4-gadget with the RHS of (29), and fuse together all the phase gadgets that dwell on the same lines. We assert that the resulted diagram after fusion is exactly the decomposition as presented on the RHS of (18) when $n = m+1$. This can be checked by calculating the phase angles of all gadget. For the 1-gadget on line 1, it comes from fusing all the 1-gadgets on line 1 which are obtained from the decomposition of all 4-gadgets connected with line. There are $\binom{m}{2}$ such 4-gadgets, so the angle of the final 1-gadget on line 1 is $\binom{m}{2}\frac{\pi}{4} = \frac{(m-1)(m-2)\pi}{8} = \sigma_{m+1}$. For the final 2-gadget on line 1 and line 2, the phase angle is $\sigma_m + \binom{m-1}{2}\frac{-\pi}{4} = -\frac{(m-2)\pi}{4} = \theta_{m+1}$. Similarly, one can check that the 1-gadget on each line has phase angle σ_{m+1}, 2-gadget on every two lines has phase angle θ_{m+1}, and 3-gadget on every three lines has phase angle $\frac{\pi}{4}$.

Therefore, (18) holds for $n = m+1$. This completes the proof.