Efficient Clifford+T approximation of single-qubit operators

Peter Selinger
Department of Mathematics and Statistics
Dalhousie University

Abstract

We give an efficient randomized algorithm for approximating an arbitrary element of $SU(2)$ by a product of Clifford+T operators, up to any given error threshold $\varepsilon > 0$. Under a mild hypothesis on the distribution of primes, the algorithm's expected runtime is polynomial in $\log(1/\varepsilon)$. If the operator to be approximated is a z-rotation, the resulting gate sequence has T-count $K + 4\log_2(1/\varepsilon)$, where K is approximately equal to 10. We also prove a worst-case lower bound of $K + 4\log_2(1/\varepsilon)$, where $K = -9$, so that our algorithm is within an additive constant of optimal for certain z-rotations. For an arbitrary member of $SU(2)$, we achieve approximations with T-count $K + 12\log_2(1/\varepsilon)$. By contrast, the Solovay-Kitaev algorithm achieves T-count $O(\log^c(1/\varepsilon))$, where c is approximately 3.97.

1 Introduction

The decomposition of arbitrary unitary operators into gates from some fixed universal set is a well-known problem in quantum information theory. If the universal gate set is discrete, the decomposition of a general operator can only be done approximately, up to a given accuracy $\varepsilon > 0$. Here, we focus on the problem of approximating single-qubit operators using the Clifford+T universal gate set. The Clifford+T gate set is of particular interest because it is known to be suitable for fault-tolerant quantum computation [1]. Recall that the Clifford group on one qubit is generated by the Hadamard gate H, the phase gate S, and the scalar $\omega = e^{i\pi/4}$. It is well-known that one obtains a universal gate set by adding the non-Clifford operator T.

$$\omega = e^{i\pi/4}, \quad H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}. \tag{1}$$

We present an efficient randomized algorithm for approximating an arbitrary element of $SU(2)$ by a product of Clifford+T operators, up to any given error threshold $\varepsilon > 0$. Under a mild hypothesis on the distribution of primes, the algorithm's expected runtime is polynomial in $\log(1/\varepsilon)$. The algorithm approximates any z-rotation with T-count[1] $K + 4\log_2(1/\varepsilon)$, where K is approximately equal to 10.[2] We also prove a worst-case lower bound of $K + 4\log_2(1/\varepsilon)$, where $K = -9$, so that our algorithm is within an additive constant of optimal for z-rotations. For an arbitrary member of $SU(2)$, we achieve approximations with T-count $K + 12\log_2(1/\varepsilon)$.

By contrast, the Solovay-Kitaev algorithm [3, 4, 5], which was until recently the de facto standard algorithm for this problem, produces circuits of T-count $O(\log^c(1/\varepsilon))$, where c is approximately 3.97. Thus, we have decreased the exponent in the gate complexity from $c = 3.97$ to $c = 1$, which is optimal. Moreover, we have decreased the multiplicative constant to the theoretical optimum, in the case of worst-case z-rotations.

1.1 Prior work

Until recently, the state-of-the-art solution to the approximation problem was the 1995 Solovay-Kitaev algorithm [3, 4, 5]. There are various variants of this algorithm. Perhaps the best-known variant is the one described in [5], which achieves a gate complexity of $O(\log^c(1/\varepsilon))$, for $c \approx 3.97$. Another well-known variant is described in Kitaev et al. [4, Sec. 8.3], and achieves $c = 3 + \delta$, where $\delta > 0$ is any positive real number. Kitaev et al. also gave another algorithm that uses ancillas and achieves gate complexity $O(\log^2(1/\varepsilon)\log\log(1/\varepsilon))$ [4, Sec. 13.7].

[1] We follow [2] in using T-count, rather than the overall gate count, as a convenient measure for the length of a single-qubit Clifford+T circuit. This is justified, on the one hand, because the fault-tolerant implementation of T-gates is far more resource intensive than that of Clifford gates, and on the other hand, because consecutive Clifford gates can always be combined into a single Clifford gate, so that the overall gate count is almost exactly equal to twice the T-count.

[2] Throughout the paper, we use K to denote an additive constant; difference occurrences of K can denote different constants.

At the other end of the spectrum, there is a known information-theoretic lower bound of $c = 1$ for the exponent in the gate complexity. One can make this lower bound more precise: in fact, the decomposition of a typical $SU(2)$ operator into the Clifford+T gate set with accuracy ε requires T-count at least $K + 3\log_2(1/\varepsilon)$. This follows from a result of Matsumoto and Amano [6]. See Section 9 below for details. Heuristically, it appears that, for most operators, this lower bound can in fact be achieved by approximation algorithms based on exhaustive search, such as Fowler's algorithm [7]. However, the runtime of such exhaustive search based algorithms is exponential in $1/\varepsilon$.

Recently, Duclos-Cianci and Svore announced an alternative to the Solovay-Kitaev algorithm that requires ancillas to be prepared in special resource states, using a state distillation procedure [8]. Using this method, and dependent on the particular setting, they reduced the gate complexity exponent c to between 1.12 and 2.27. Moreover, the resource states can be prepared offline, and the expected online gate complexity per single-qubit operation is constant.

Even more recently, in an important milestone, Kliuchnikov, Maslov, and Mosca gave an approximation algorithm for single-qubit operators that has polynomial running time and achieves gate counts of $O(\log(1/\varepsilon))$, thus reducing the gate complexity exponent to $c = 1$ [9]. The Kliuchnikov-Maslov-Mosca approximation algorithm is therefore asymptotically optimal. It uses a fixed, small number of ancilla qubits to approximate a given single-qubit operator. Unlike approaches based on resource states, the ancillas in the Kliuchnikov-Maslov-Mosca algorithm are initialized to $|0\rangle$, and are returned in a state very close (but not exactly equal) to $|0\rangle$; these ancillas do not require any special preparation procedure.

For all practical purposes, this use of ancillas in the Kliuchnikov-Maslov-Mosca algorithm causes no difficulties. However, the question remained open whether there exists an asymptotically optimal, efficient single-qubit approximation algorithm that requires no ancillas, and thus solves exactly the same problem as the Solovay-Kitaev algorithm, for the Clifford+T gate set. The present paper achieves such an algorithm.

1.2 Limitations

Unlike the Solovay-Kitaev algorithm, which can be applied to any universal gate set, the algorithm of this paper is specialized to the Clifford+T gate set. While our number-theoretic approach has already been generalized to specific other universal gate sets (see, e.g., [10]), it is unlikely that it would work for *arbitrary* universal gate sets. While this could be regarded as a limitation of our method, in practice, the tables will likely be turned: any universal gate set that does not permit an $O(\log(1/\varepsilon))$ efficient synthesis algorithm will probably not be considered practical in the future.

Technically, the expected polynomial runtime of our algorithm is contingent on a number-theoretic assumption about the distribution of primes, as stated below in Hypothesis 29. While this hypothesis appears to be heuristically true, it has not been proven to the author's knowledge.

The lower bound of $K + 4\log_2(1/\varepsilon)$ for the T-count of ε-approximations to certain z-rotations only applies to the problem as stated, i.e., for writing operators as a product of single-qubit Clifford+T operators. By the use of other techniques, such as ancillas, resource states, or online measurement, even smaller T-counts can be obtained; see, e.g., [11] for recent results along these lines.

2 Overview of the algorithm

The algorithm can be summarized as follows. Consider the ring

$$\mathbb{D}[\omega] = \mathbb{Z}[\frac{1}{\sqrt{2}}, i] = \{\frac{1}{\sqrt{2}^k}(a\omega^3 + b\omega^2 + c\omega + d) \mid k \in \mathbb{N}; a, b, c, d \in \mathbb{Z}\}. \tag{2}$$

As shown by Kliuchnikov, Maslov, and Mosca [12], a unitary operator $U \in U(2)$ can be exactly synthesized over the Clifford+T gate set (with no ancillas) if and only if all the matrix entries belong to the ring $\mathbb{D}[\omega]$. Moreover, the required number of T-gates is at most $2k$, where $k > 0$ is the minimal exponent required to write all entries of U in the form mentioned in (2).

Suppose we wish to approximate a given z-rotation

$$R_z(\theta) = e^{-i\theta Z/2} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \tag{3}$$

up to a given $\varepsilon > 0$, using Clifford+T gates. We will choose an integer k and a randomized sequence of suitable elements $u \in \mathbb{Z}[\omega]$ such that $\frac{u}{\sqrt{2}^k} \approx e^{-i\theta/2}$. For each u, we attempt to solve the Diophantine equation

$$t^\dagger t + u^\dagger u = 2^k, \tag{4}$$

with $t \in \mathbb{Z}[\omega]$. The parameters are chosen in such a way that this succeeds for a relatively large proportion of the available u. This yields a unitary matrix

$$U = \frac{1}{\sqrt{2}^k} \begin{pmatrix} u & -t^\dagger \\ t & u^\dagger \end{pmatrix} \tag{5}$$

with $\|U - R_z(\theta)\| \leqslant \varepsilon$, and such that the coefficients of U are in the ring $\mathbb{D}[\omega]$. By the Kliuchnikov-Maslov-Mosca exact synthesis algorithm, U can be exactly decomposed into Clifford+T gates with T-count at most $2k$. The remainder of this paper fills in the details of these ideas: in particular, how to choose k, how to find "suitable" u, and how to solve the Diophantine equation (4) with high probability.

3 Some number theory

In this section and the next one, we summarize some well-known facts from algebraic number theory. The following exposition requires almost no prerequisites, and will hopefully be of use to readers who are not experts in number theory, or to those who wish to implement the algorithm of this paper.

3.1 Some rings of algebraic integers

Recall that $\mathbb{N}$ is the set of natural numbers including 0. Let $\omega = e^{i\pi/4} = (1+i)/\sqrt{2}$. Note that ω is an 8th root of unity satisfying $\omega^2 = i$ and $\omega^4 = -1$. We will consider the following rings:

- $\mathbb{Z}$, the ring of integers;
- $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$, the ring of *quadratic integers with radicand 2*;
- $\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$, the ring of *Gaussian integers*;
- $\mathbb{Z}[\omega] = \{a\omega^3 + b\omega^2 + c\omega + d \mid a, b, c, d \in \mathbb{Z}\}$, the ring of *cyclotomic integers of degree 8*.

Remark 1. We have the inclusions $\mathbb{Z} \subseteq \mathbb{Z}[\sqrt{2}] \subseteq \mathbb{Z}[\omega]$ and $\mathbb{Z} \subseteq \mathbb{Z}[i] \subseteq \mathbb{Z}[\omega]$. Of course, all four rings are subrings of the complex numbers.

3.2 Conjugate and norm

Definition 2 (Conjugation)**.** Since ω is a root of the irreducible polynomial $x^4 + 1$, the ring $\mathbb{Z}[\omega]$ has four automorphisms. One of these is the usual *complex conjugation*, which we denote $(-)^\dagger$. It maps i to $-i$, and $\sqrt{2}$ to itself. Equivalently, it is given by $\omega^\dagger = -\omega^3$. It acts trivially on $\mathbb{Z}$ and $\mathbb{Z}[\sqrt{2}]$, and non-trivially on $\mathbb{Z}[\omega]$ and $\mathbb{Z}[i]$, with the following explicit formulas, for real a, b, c, d:

$$(a\omega^3 + b\omega^2 + c\omega + d)^\dagger = -c\omega^3 - b\omega^2 - a\omega + d, \tag{6}$$

$$(a + bi)^\dagger = a - bi. \tag{7}$$

Another automorphism is $\sqrt{2}$-*conjugation*, which we denote $(-)^\bullet$. It maps $\sqrt{2}$ to $-\sqrt{2}$, and i to itself. Equivalently, $\omega^\bullet = -\omega$. It acts trivially on $\mathbb{Z}$ and $\mathbb{Z}[i]$, and non-trivially on $\mathbb{Z}[\omega]$ and $\mathbb{Z}[\sqrt{2}]$, with the following explicit formulas, for rational a, b, c, d:

$$(a\omega^3 + b\omega^2 + c\omega + d)^\bullet = -a\omega^3 + b\omega^2 - c\omega + d, \tag{8}$$

$$(a + b\sqrt{2})^\bullet = a - b\sqrt{2}. \tag{9}$$

The remaining two automorphisms are, of course, the identity function and $(-)^{\dagger\bullet} = (-)^{\bullet\dagger}$.

Remark 3. For $t \in \mathbb{Z}[\omega]$, we have $t \in \mathbb{Z}[\sqrt{2}]$ iff $t = t^\dagger$, $t \in \mathbb{Z}[i]$ iff $t = t^\bullet$, and $t \in \mathbb{Z}$ iff $t = t^\dagger$ and $t = t^\bullet$.

Definition 4 (Norms)**.** We define an integer-valued (number-theoretic) *norm* on each ring:

- For $t = a + bi \in \mathbb{Z}[i]$, let
$$\mathcal{N}_i(t) = t^\dagger t = a^2 + b^2. \tag{10}$$
- For $t = a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$, let
$$\mathcal{N}_{\sqrt{2}}(t) = t^\bullet t = a^2 - 2b^2. \tag{11}$$

- For $t = a\omega^3 + b\omega^2 + c\omega + d \in \mathbb{Z}[\omega]$, let

$$\mathcal{N}_\omega(t) = (t^\dagger t)^\bullet (t^\dagger t) = (a^2+b^2+c^2+d^2)^2 - 2(ab+bc+cd-da)^2. \tag{12}$$

For consistency, we also define $\mathcal{N}_1(t) = t$ for $t \in \mathbb{Z}$.

Remark 5. $\mathcal{N}_i$ and $\mathcal{N}_\omega$ are valued in the non-negative integers, but $\mathcal{N}_1$ and $\mathcal{N}_{\sqrt{2}}$ may take negative values. Each norm is multiplicative, in the sense that $\mathcal{N}(ts) = \mathcal{N}(t)\mathcal{N}(s)$ for all t, s. Moreover, $\mathcal{N}(t) = 0$ if and only if $t = 0$, and $\mathcal{N}(t) = \pm 1$ if and only if t is a unit (i.e., an invertible element) in the ring. The latter property follows easily from multiplicativity and the fact that equations (10–12) define an inverse for t when $\mathcal{N}(t) = \pm 1$.

3.3 Euclidean domains

Remark 6. $\mathbb{Z}$, $\mathbb{Z}[\sqrt{2}]$, $\mathbb{Z}[i]$, and $\mathbb{Z}[\omega]$ are Euclidean domains. Explicitly, for each of these rings, a Euclidean function is given by $|\mathcal{N}(t)|$. For given elements s and $t \neq 0$, the division of s by t with quotient q and remainder r can be defined by first calculating $q' = s/t$ in the corresponding field of fractions (respectively $\mathbb{Q}$, $\mathbb{Q}[\sqrt{2}]$, $\mathbb{Q}[i]$, and $\mathbb{Q}[\omega]$). q' can be expressed with rational coefficients as a, $a + b\sqrt{2}$, $a + bi$, or $\omega^3 a + \omega^2 b + \omega c + d$, respectively. Then q is obtained from q' by rounding each coefficient a, b, c, d to the closest integer, and r is defined to be $qt - s$. One may verify that in each case, $|\mathcal{N}(r)| \leqslant \frac{9}{16}|\mathcal{N}(t)|$.

As usual, we write $t \mid s$ to mean that t is a divisor of s, i.e., that there exists some r such that $rt = s$. We also write $t \sim s$ to indicate that $t \mid s$ and $s \mid t$; in this case, t and s differ only by a unit of the ring, and we say t and s are *associates*.

Remark 7. We note that if R is one of the four rings $\mathbb{Z}$, $\mathbb{Z}[\sqrt{2}]$, $\mathbb{Z}[i]$, and $\mathbb{Z}[\omega]$, and $rt = s$ for some $s, t \in R$ and $r \in \mathbb{Z}[\omega]$ with $t \neq 0$, then $r \in R$. This is easily proved by letting $\overline{R}$ be the corresponding field of fractions (respectively $\mathbb{Q}$, $\mathbb{Q}[\sqrt{2}]$, $\mathbb{Q}[i]$, and $\mathbb{Q}[\omega]$), and noting on the one hand that $r = \frac{s}{t} \in \overline{R}$, and on the other hand that $R = \mathbb{Z}[\omega] \cap \overline{R}$. The latter property follows from Remark 3.

In particular, it follows that $t \mid s$ holds in R if and only if $t \mid s$ holds in $\mathbb{Z}[\omega]$, so that the notion of divisibility is independent of R.

Remark 8. Every Euclidean domain admits greatest common divisors, which can be calculated by repeated divisions with remainder via Euclid's algorithm. Note that greatest common divisors are only unique up to a unit of the ring. Also note that, since each division by t with remainder r satisfies $|\mathcal{N}(r)| \leqslant \frac{9}{16}|\mathcal{N}(t)|$, the computation of the greatest common divisor of two elements of any of the above rings requires at most $O(\log|\mathcal{N}(t)|)$ divisions with remainder.

An element t of a Euclidean domain is called *prime* (or *irreducible*) if t is not a unit, and for all r, s with $rs = t$, either r or s is a unit. We note that the notion of primality is not independent of the ring. For example, 7 is prime in $\mathbb{Z}$, but not in $\mathbb{Z}[\sqrt{2}]$, as $7 = (3+\sqrt{2})(3-\sqrt{2})$.

Remark 9. In each of the above rings, if $\mathcal{N}(t)$ is prime in $\mathbb{Z}$, then t is prime in the ring. Indeed, suppose $t = rs$. Since $\mathcal{N}(t) = \mathcal{N}(r)\mathcal{N}(s)$, either $\mathcal{N}(r)$ or $\mathcal{N}(s)$ is ± 1, hence r or s is a unit by Remark 5.

3.4 Units in $\mathbb{Z}[\sqrt{2}]$

Lemma 10. *The units of $\mathbb{Z}[\sqrt{2}]$ are exactly the elements of the form $(-1)^n(\sqrt{2}-1)^k$, where $n, k \in \mathbb{Z}$. A unit u is a square if and only if $u \geqslant 0$ and $u^\bullet \geqslant 0$.*

Proof. We first note that $(\sqrt{2}-1)(\sqrt{2}+1) = 1$, so $\sqrt{2}-1$ and $\sqrt{2}+1$ are units. Now consider any unit $u = a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$. We prove the first claim by induction on $|b|$. The base case is $b = 0$; in this case, $\mathcal{N}_{\sqrt{2}}(u) = a^2 = \pm 1$ implies $u = a = \pm 1$. For the induction step, note that $\mathcal{N}_{\sqrt{2}}(u) = a^2 - 2b^2 = \pm 1$ implies $a \neq 0$ and $a^2 < 2b^2 + 2 \leqslant 4b^2$, hence $0 < |a| < |2b|$. First consider the case where a, b have the same sign. Then $|a - b| < |b|$, and the claim is proved by applying the induction hypothesis to $\hat{u} = u(\sqrt{2}-1) = 2b - a + (a-b)\sqrt{2}$. The case where a, b have opposite signs is similar, except we use $|a + b| < |b|$ to apply the induction hypothesis to $\hat{u} = u/(\sqrt{2}-1) = a + 2b + (a+b)\sqrt{2}$.

For the second claim, note that $u = (-1)^n(\sqrt{2}-1)^k$ satisfies $u \geqslant 0$ iff n is even, and $u^\bullet \geqslant 0$ iff $n + k$ is even. Therefore u is a square iff n and k are both even, iff $u \geqslant 0$ and $u^\bullet \geqslant 0$. $\square$

3.5 Roots of -1 in $\mathbb{Z}/(p)$

Remark 11. Let $p \in \mathbb{Z}$ be a positive prime satisfying $p \equiv 1 \pmod 4$. It is well-known that there exists $h \in \mathbb{Z}$ such that $h^2 + 1 \equiv 0 \pmod p$. We recall that there is an efficient randomized algorithm for computing h. Consider the field $\mathbb{Z}/(p)$ of integers modulo p. By Fermat's Little Theorem, all non-zero $b \in \mathbb{Z}/(p)$ satisfy $b^{p-1} = 1$, hence $b^{(p-1)/2} = \pm 1$. Because each of the polynomial equations $b^{(p-1)/2} = 1$ and $b^{(p-1)/2} = -1$ has at most $(p-1)/2$ solutions, $b^{(p-1)/2} = -1$ holds for exactly half of the non-zero $b \in \mathbb{Z}/(p)$. Therefore, one can efficiently solve $b^{(p-1)/2} = -1$ by picking b at random until a solution is found; on average, this will require two attempts. Note that, by the method of repeated squaring, the computation of $b^{(p-1)/2}$ only requires $O(\log p)$ multiplications. Finally, we can set $h = b^{(p-1)/4}$.

4 A Diophantine equation

We will be interested in solving equations of the form

$$t^\dagger t = \xi, \tag{13}$$

where $\xi \in \mathbb{Z}[\sqrt{2}]$ is given and $t \in \mathbb{Z}[\omega]$ is unknown. Number theorists will recognize this as a relative norm equation, which can be solved by splitting fully split primes in $\mathbb{Z}[\omega]$ [13]. In the interest of self-containedness, and to aid in the complexity analysis of Section 8, we describe a method for solving equation (13) in detail.

Writing $\xi = x + y\sqrt{2}$ and $t = a\omega^3 + b\omega^2 + c\omega + d$, we can equivalently express (13) as a pair of integer equations:

$$a^2 + b^2 + c^2 + d^2 \;=\; x \tag{14}$$

$$ab + bc + cd - da \;=\; y. \tag{15}$$

Of course, not every $\xi \in \mathbb{Z}[\sqrt{2}]$ can be expressed in the form (13). However, we have the following:

Theorem 12. *Let $\xi = x + y\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$, where x is odd, y is even, $\xi \geqslant 0$, $\xi^\bullet \geqslant 0$, and $p = \xi^\bullet \xi = x^2 - 2y^2$ is prime in $\mathbb{Z}$. Then there exists $t \in \mathbb{Z}[\omega]$ satisfying (13). Moreover, there is an efficient randomized algorithm for computing t.*

Remark 13. Since $t^\dagger t = \xi$ implies $t^{\bullet\dagger} t^\bullet = \xi^\bullet$, the conditions $\xi \geqslant 0$ and $\xi^\bullet \geqslant 0$ are both obviously necessary for (13) to have a solution.

Proof of Theorem 12. Note that p is prime by assumption. Since $\xi \geqslant 0$ and $\xi^\bullet \geqslant 0$, we know that $p \geqslant 0$. Since x is odd and y is even, we have $p \equiv 1 \pmod 4$. Therefore, by Remark 11, we can find $h \in \mathbb{Z}$ with $p \mid h^2 + 1$. Moreover, in $\mathbb{Z}[\sqrt{2}]$, we have $\xi \mid p$ and therefore $\xi \mid h^2 + 1$. Let $s = \gcd(h + i, \xi)$ in the ring $\mathbb{Z}[\omega]$. We claim that $s^\dagger s \sim \xi$.

First, note that ξ is prime in the ring $\mathbb{Z}[\sqrt{2}]$ by Remark 9. By definition of s, we know that $s \mid \xi$. But ξ is real, and therefore also $s^\dagger \mid \xi$. It follows that $s^\dagger s \mid \xi^2$ in $\mathbb{Z}[\omega]$. By Remark 7, $s^\dagger s \mid \xi^2$ in $\mathbb{Z}[\sqrt{2}]$. Since ξ is prime in $\mathbb{Z}[\sqrt{2}]$, it follows that one of three cases holds:

(a) $s^\dagger s \sim 1$. But this is impossible. Indeed, in this case, s is a unit, so $h + i$ and ξ are relatively prime. As ξ is real, it follows that $h - i$ and ξ are also relatively prime, hence $(h+i)(h-i)$ is relatively prime to ξ, contradicting $\xi \mid h^2 + 1$.

(b) $s^\dagger s \sim \xi$. This is what was claimed.

(c) $s^\dagger s \sim \xi^2$. This is also impossible. Indeed, in this case, $\mathcal{N}_\omega(s) = \mathcal{N}_{\sqrt{2}}(s^\dagger s) = \mathcal{N}_{\sqrt{2}}(\xi^2) = \mathcal{N}_\omega(\xi)$. But we also have $s \mid \xi$, so $\xi = us$ for some $u \in \mathbb{Z}[\omega]$. Therefore $\mathcal{N}_\omega(u) = \mathcal{N}_\omega(\xi)/\mathcal{N}_\omega(s) = 1$, hence u is a unit in $\mathbb{Z}[\omega]$, thus $s \sim \xi$. Since, by definition, $s \mid h + i$, we have $\xi \mid h + i$. Since h is an integer, we have $h = h^\bullet$, and thus $\xi^\bullet \mid h + i$. We note that $\xi \not\sim \xi^\bullet$, for otherwise, we would have $p = \xi^\bullet \xi \mid \xi^2 + {\xi^\bullet}^2 = 2x^2 + 4y^2 = 4x^2 - 2p$; since p is an odd prime, this implies $p \mid x$, and from $p = x^2 - 2y^2$, we get $p \mid y$, hence $p^2 \mid x^2 - 2y^2 = p$, a contradiction. Therefore, ξ and $\xi^\bullet$ are non-associate primes in $\mathbb{Z}[\sqrt{2}]$, both dividing $h + i$, which implies $p = \xi^\bullet \xi \mid h + i$, an absurdity since $\frac{1}{p}(h + i) \notin \mathbb{Z}[\omega]$.

We have shown the existence of $s \in \mathbb{Z}[\omega]$ such that $s^\dagger s \sim \xi$ in $\mathbb{Z}[\sqrt{2}]$. Let u be a unit of $\mathbb{Z}[\sqrt{2}]$ such that $us^\dagger s = \xi$. Our next claim is that u is a square in $\mathbb{Z}[\sqrt{2}]$. First note that, by the usual properties of complex numbers, $s^\dagger s \geqslant 0$ and $(s^\dagger s)^\bullet = (s^\bullet)^\dagger (s^\bullet) \geqslant 0$. Also, $\xi \geqslant 0$ and $\xi^\bullet \geqslant 0$ by assumption. It follows that $u \geqslant 0$ and $u^\bullet \geqslant 0$. But then u is a square by Lemma 10. Let $v \in \mathbb{Z}[\sqrt{2}]$ with $v^2 = u$, and let $t = vs$. Noting that $v = v^\dagger$, we have $t^\dagger t = v^2 s^\dagger s = \xi$, as desired. $\square$

Remark 14. We note that the proof of Theorem 12 immediately yields an algorithm for computing t. As a matter of fact, the only randomized aspect is the computation of h; the remainder consists of arithmetic calculations in the various rings, and can be done deterministically. The computation of s requires taking a greatest common divisor in $\mathbb{Z}[\omega]$, which can be done efficiently by Remark 8. The computation of u requires a simple division in $\mathbb{Z}[\sqrt{2}]$, and the computation of v requires taking a square root in $\mathbb{Z}[\sqrt{2}]$, which easily reduces to solving a quadratic equation in the integers. Finally, the computation of t requires a multiplication in $\mathbb{Z}[\omega]$.

Remark 15. Theorem 12 is analogous to a well-known theorem about the integers, stating that every positive prime satisfying $p \equiv 1 \pmod 4$ can be written as a sum of two squares $p = a^2 + b^2$, or equivalently, that the equation $z^\dagger z = p$ has a solution $z = a + bi \in \mathbb{Z}[i]$. A randomized algorithm for computing z was described, for example, by Rabin and Shallit [14]. Our proof and algorithm follows the same general idea, applied to a different pair of Euclidean rings.

5 Approximations in $\mathbb{Z}[\sqrt{2}]$

It is well-known that the set $\mathbb{Z}[\sqrt{2}]$ of integers of the form $\alpha = a + b\sqrt{2}$ is a dense subset of the real numbers. Here, density is of course understood with respect to the Euclidean distance $|\alpha - \beta|$. We note that the Euclidean distance is not at all preserved by $\sqrt{2}$-conjugation; in fact, as we will see in Lemma 16 below, unless $\alpha = \beta$, it is impossible for $|\alpha - \beta|$ and $|\alpha^\bullet - \beta^\bullet|$ to be small at the same time.

The purpose of this section is to find solutions in $\mathbb{Z}[\sqrt{2}]$ to constraints involving α and $\alpha^\bullet$ simultaneously. Specifically, we will be interested in solving problems of the form

$$a + b\sqrt{2} \in [x_0, x_1] \quad \text{and} \quad a - b\sqrt{2} \in [y_0, y_1], \tag{16}$$

where $x_0 < x_1$ and $y_0 < y_1$ are given real numbers, and a, b are unknown integers. We start with a result limiting the number of solutions.

Lemma 16. *Let $[x_0, x_1]$ and $[y_0, y_1]$ be closed intervals of real numbers. Let $\delta = x_1 - x_0$ and $\Delta = y_1 - y_0$, and assume $\delta\Delta < 1$. Then there exists at most one $\alpha = a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$ satisfying (16).*

Proof. Suppose α and β are two solutions. Then

$$|\mathcal{N}_{\sqrt{2}}(\alpha - \beta)| = |\alpha - \beta| \cdot |\alpha^\bullet - \beta^\bullet| \leqslant \delta\Delta < 1. \tag{17}$$

Since $\mathcal{N}_{\sqrt{2}}(\alpha - \beta)$ is an integer, it follows that $\mathcal{N}_{\sqrt{2}}(\alpha - \beta) = 0$, and therefore $\alpha = \beta$. $\square$

The next result establishes the existence of solutions.

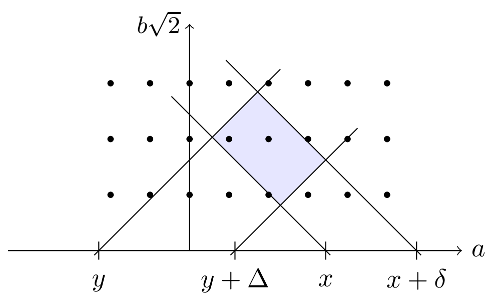
Lemma 17. *Let $[x_0, x_1]$ and $[y_0, y_1]$ be closed intervals of real numbers. Let $\delta = x_1 - x_0$ and $\Delta = y_1 - y_0$, and assume $\delta\Delta \geqslant (1 + \sqrt{2})^2$. Then there exists at least one $\alpha = a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]$ satisfying (16). Moreover, there is an efficient algorithm for computing such a and b.*

Proof. Let us say that a pair of positive real numbers (δ, Δ) has the *coverage property* if for all $x, y \in \mathbb{R}$, there exists $\alpha \in \mathbb{Z}[\sqrt{2}]$ with

$$(\alpha, \alpha^\bullet) \in [x, x + \delta] \times [y, y + \Delta]. \tag{18}$$

The goal, then, is to show that $\delta\Delta \geqslant (1 + \sqrt{2})^2$ implies the coverage property.

Before we continue, it is perhaps helpful to consider the following illustration. Here, a is shown on the horizontal axis, and $b\sqrt{2}$ is shown on the vertical axis. The region defined by $(a + b\sqrt{2}, a - b\sqrt{2}) \in [x, x + \delta] \times [y, y + \Delta]$ is a rectangle oriented at 45 degrees. We are only interested in solutions where a, b are integers; these solutions therefore lie on the grid $\mathbb{Z} \times \sqrt{2}\,\mathbb{Z}$. Note that the horizontal and vertical spacing are not the same.



We prove a sequence of claims leading up to the lemma.

(a) If (δ, Δ) has the coverage property and $\delta' \geqslant \delta$, $\Delta' \geqslant \Delta$, then (δ', Δ') has the coverage property. This is a trivial consequence of the definitions.

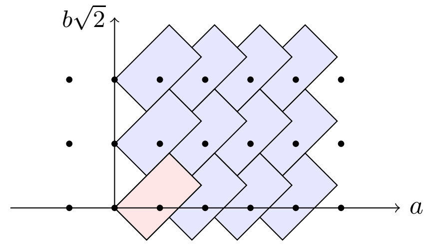
(b) For all $\delta, \Delta > 0$, the pair (δ, Δ) has the coverage property if and only if (Δ, δ) has the coverage property. This is a trivial consequence of the definitions.

(c) Let $\lambda = 1 + \sqrt{2}$. Then for all $\delta, \Delta > 0$, the pair (δ, Δ) has the coverage property if and only if $(\lambda\delta, \lambda^{-1}\Delta)$ has the coverage property. Indeed, recall that $\lambda^\bullet = -\lambda^{-1} = 1 - \sqrt{2}$. Therefore both λ and λ^{-1} are elements of $\mathbb{Z}[\sqrt{2}]$. Suppose (δ, Δ) has the coverage property, and $x, y \in \mathbb{R}$ are given. Then there exists $\alpha \in \mathbb{Z}[\sqrt{2}]$ with $(\alpha, \alpha^\bullet) \in [\lambda^{-1}x, \lambda^{-1}x + \delta] \times [-\lambda y - \Delta, -\lambda y]$. Let $\beta = \lambda\alpha$. It follows that

$$(\beta, \beta^\bullet) = (\lambda\alpha, \lambda^\bullet\alpha^\bullet) = (\lambda\alpha, -\lambda^{-1}\alpha^\bullet) \in [x, x + \lambda\delta] \times [y, y + \lambda^{-1}\Delta], \tag{19}$$

so $(\lambda\delta, \lambda^{-1}\Delta)$ has the coverage property. The converse if proved symmetrically.

(d) $(\delta, \Delta) = (1 + \sqrt{2}, \sqrt{2})$ has the coverage property. Geometrically, this is equivalent to the statement that $\mathbb{R}^2$ is covered by all translations along the grid $\mathbb{Z} \times \sqrt{2}\mathbb{Z}$ of the rectangle R defined by $(a + b\sqrt{2}, a - b\sqrt{2}) \in [0, 1 + \sqrt{2}] \times [0, \sqrt{2}]$.



In order to have an explicit algorithm for finding α, we give an algebraic proof. Consider arbitrary $x, y \in \mathbb{R}$. Let a, b be integers such that

$$a - 1 \leqslant \frac{x + y + \Delta}{2} < a, \tag{20}$$

$$(b - 1)\sqrt{2} \leqslant \frac{x - y - \Delta}{2} < b\sqrt{2}. \tag{21}$$

Let $\alpha = a + b\sqrt{2}$, $\alpha' = a + (b + 1)\sqrt{2}$, and $\alpha'' = (a - 1) + b\sqrt{2}$. We claim that either α, α', or α'' is a solution to (18).

- Case 1: $a - b\sqrt{2} \leqslant y + \Delta$. In this case, we have:

$$\begin{aligned}
\alpha &= a + b\sqrt{2} > \frac{x + y + \Delta}{2} + \frac{x - y - \Delta}{2} = x && \text{by (20) and (21)},\\
\alpha &= a + b\sqrt{2} \leqslant \frac{x + y + \Delta}{2} + 1 + \frac{x - y - \Delta}{2} + \sqrt{2} = x + \delta && \text{by (20) and (21)},\\
\alpha^\bullet &= a - b\sqrt{2} > \frac{x + y + \Delta}{2} - \frac{x - y - \Delta}{2} - \sqrt{2} = y && \text{by (20) and (21)},\\
\alpha^\bullet &= a - b\sqrt{2} \leqslant y + \Delta && \text{by assumption.}
\end{aligned}$$

  Therefore, α is a solution to (18).

- Case 2: $a - b\sqrt{2} > y + \Delta$ and $a + b\sqrt{2} \leqslant x + 1$. In this case, we have:

$$\begin{aligned}
\alpha' &= a + (b + 1)\sqrt{2} > \frac{x + y + \Delta}{2} + \frac{x - y - \Delta}{2} + \sqrt{2} = x + \sqrt{2} > x && \text{by (20) and (21)},\\
\alpha' &= a + (b + 1)\sqrt{2} \leqslant x + 1 + \sqrt{2} = x + \delta && \text{by assumption},\\
\alpha'^\bullet &= a - (b + 1)\sqrt{2} > y + \Delta - \sqrt{2} = y && \text{by assumption},\\
\alpha'^\bullet &= a - (b + 1)\sqrt{2} < \frac{x + y + \Delta}{2} + 1 - \frac{x - y - \Delta}{2} - \sqrt{2} = y + \Delta + 1 - \sqrt{2} < y + \Delta && \text{by (20) and (21).}
\end{aligned}$$

  Therefore, α' is a solution to (18).

- Case 3: $a - b\sqrt{2} > y + \Delta$ and $a + b\sqrt{2} > x + 1$. In this case, we have:

$$\begin{aligned}
\alpha'' &= (a-1) + b\sqrt{2} > x && \text{by assumption,} \\
\alpha'' &= (a-1) + b\sqrt{2} \leqslant \frac{x+y+\Delta}{2} + \frac{x-y-\Delta}{2} + \sqrt{2} = x + \sqrt{2} < x + \delta && \text{by (20) and (21),} \\
\alpha''^{\bullet} &= (a-1) - b\sqrt{2} > y + \Delta - 1 > y && \text{by assumption,} \\
\alpha''^{\bullet} &= (a-1) - b\sqrt{2} \leqslant \frac{x+y+\Delta}{2} - \frac{x-y-\Delta}{2} = y + \Delta && \text{by (20) and (21).}
\end{aligned}$$

  Therefore, α'' is a solution to (18).

(e) $(\delta, \Delta) = (2 + \sqrt{2}, 1)$ has the coverage property. This follows from (b)–(d), by noting that $2 + \sqrt{2} = \lambda\sqrt{2}$ and $1 = \lambda^{-1}(1 + \sqrt{2})$.

(f) Suppose $\delta\Delta \geqslant (1 + \sqrt{2})^2$ and $1 < \delta \leqslant 1 + \sqrt{2}$. Then (δ, Δ) has the coverage property. We consider two cases. Case 1: $\delta > \sqrt{2}$. Since $\delta \leqslant 1 + \sqrt{2}$, we have $\Delta \geqslant 1 + \sqrt{2}$. But $(\sqrt{2}, 1 + \sqrt{2})$ has the coverage property by (d), so the claim follows from (a). Case 2: $\delta \leqslant \sqrt{2}$. Then

$$\Delta \geqslant \frac{(1+\sqrt{2})^2}{\sqrt{2}} > 2 + \sqrt{2}. \tag{22}$$

Also $\delta > 1$. But $(1, 2 + \sqrt{2})$ has the coverage property by (e), so the claim follows from (a).

(g) Suppose $\delta, \Delta > 0$ such that $\delta\Delta \geqslant (1 + \sqrt{2})^2$. Then (δ, Δ) has the coverage property. Indeed, note that there exists some $n \in \mathbb{Z}$ such that $1 < \lambda^n\delta \leqslant \lambda = 1 + \sqrt{2}$. Then $(\lambda^n\delta, \lambda^{-n}\Delta)$ has the coverage property by (f), and (δ, Δ) has the coverage property by (c).

This finishes the proof of the lemma. Note that in each step of the proof, we have shown explicitly how to compute a solution α. Since there are no iterative constructions (apart from the computation of λ^n in (g), which can be done in $O(|\log(\delta)|)$ steps), the computational effort is essentially trivial, and certainly not greater than say $O(\log^2(\delta))$. □

Remark 18. We note that the bounds on $\delta\Delta$ in Lemmas 16 and 17 are sharp. For the sharpness of Lemma 16, consider $(\alpha, \alpha^\bullet) \in [0, 1] \times [0, 1]$, which has two solutions $\alpha = 0$ and $\alpha = 1$. For the sharpness of Lemma 17, consider $(\alpha, \alpha^\bullet) \in [0, 1 + \sqrt{2}] \times [-\sqrt{2}, 1]$, which has exactly four solutions $\alpha = 0$, $\alpha = 1$, $\alpha = \sqrt{2}$, and $\alpha = 1 + \sqrt{2}$. Therefore, for all $\varepsilon > 0$, $(\alpha, \alpha^\bullet) \in [\varepsilon, 1 + \sqrt{2} - \varepsilon] \times [-\sqrt{2} + \varepsilon, 1 - \varepsilon]$ has no solutions at all.

We also give a version of Lemma 17 where a is restricted to be either even or odd:

Corollary 19. *Let $[x_0, x_1]$ and $[y_0, y_1]$ be closed intervals of real numbers. Let $\delta = x_1 - x_0$ and $\Delta = y_1 - y_0$, and assume $\delta\Delta \geqslant 2(1 + \sqrt{2})^2$. Then there exist $a', b' \in \mathbb{Z}$ such that $a' + b'\sqrt{2} \in [x_0, x_1]$ and $a' - b'\sqrt{2} \in [y_0, y_1]$, and a' is even. There also exist $a'', b'' \in \mathbb{Z}$ such that $a'' + b''\sqrt{2} \in [x_0, x_1]$ and $a'' - b''\sqrt{2} \in [y_0, y_1]$, and a'' is odd. Moreover, there is an efficient algorithm for computing such a', b', a'', and b''.*

Proof. This is proved by rescaling. To prove the first claim, use Lemma 17 to find $a, b \in \mathbb{Z}$ with $a + b\sqrt{2} \in [x_0/\sqrt{2}, x_1/\sqrt{2}]$ and $a - b\sqrt{2} \in [-y_1/\sqrt{2}, -y_0/\sqrt{2}]$. Let $a' = 2b$ and $b' = a$. Then we have $a' + b'\sqrt{2} = \sqrt{2}\,(a + b\sqrt{2}) \in [x_0, x_1]$ and $a' - b'\sqrt{2} = -\sqrt{2}\,(a - b\sqrt{2}) \in [y_0, y_1]$, as desired. To prove the second claim, use the first claim to find $a' + b'\sqrt{2} \in [x_0 - 1, x_1 - 1]$ and $a' - b'\sqrt{2} \in [y_0 - 1, y_1 - 1]$, with a' even; then let $a'' = a' + 1$ and $b'' = b'$. □

6 Approximation up to ε

As mentioned in Section 2, given θ and ε, we wish to find $k \geqslant 0$ and $u, t \in \mathbb{Z}[\omega]$ such that

$$U = \frac{1}{\sqrt{2}^k} \begin{pmatrix} u & -t^\dagger \\ t & u^\dagger \end{pmatrix} \tag{23}$$

is unitary and satisfies $\|U - R_z(\theta)\| \leqslant \varepsilon$. We now elaborate how to determine k, u, and t.

6.1 The ε-region

We first note that U is unitary if and only if $u^\dagger u + t^\dagger t = 2^k$. We now examine how to express the error ε as a function of u; this will make explicit the set of available u for a given ε. Let $z = x + yi = e^{-i\theta/2}$. First note that, since both U and $R_z(\theta)$ are 2×2 unitary of determinant 1, the operator norm of $U - R_z(\theta)$ coincides with $1/\sqrt{2}$ of the Hilbert-Schmidt norm. It can therefore be calculated as follows. Let $\hat{u} = \frac{1}{\sqrt{2}^k} u = \alpha + \beta i$ and $\hat{t} = \frac{1}{\sqrt{2}^k} t$.

$$\|U - R_z(\theta)\|^2 = \frac{1}{2}\|U - R_z(\theta)\|_{\mathrm{HS}}^2 = |\hat{u} - z|^2 + |\hat{t}|^2. \tag{24}$$

Using $\hat{u}^\dagger\hat{u} + \hat{t}^\dagger\hat{t} = 1$ and $z^\dagger z = 1$, we can further simplify this to:

$$|\hat{u} - z|^2 + |\hat{t}|^2 = (\hat{u} - z)^\dagger(\hat{u} - z) + \hat{t}^\dagger\hat{t} = \hat{u}^\dagger\hat{u} - \hat{u}^\dagger z - z^\dagger\hat{u} + z^\dagger z + \hat{t}^\dagger\hat{t} = 2 - 2\,\mathrm{Re}(\hat{u}^\dagger z) = 2 - 2\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\cdot\begin{pmatrix}x\\ y\end{pmatrix}. \tag{25}$$

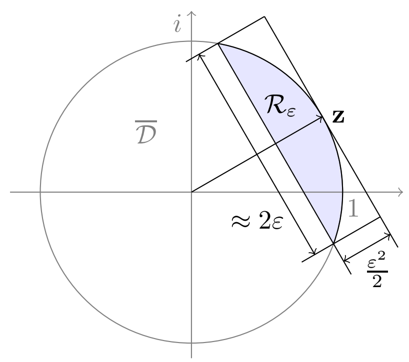
The error is therefore directly related to the dot product of $\hat{u}$ and z, regarded as vectors in $\mathbb{R}^2$. Writing $\mathbf{z} = (x, y)^T$ and $\mathbf{u} = (\alpha, \beta)^T$, we have

$$\|U - R_z(\theta)\|^2 \leqslant \varepsilon^2 \iff 2 - 2\mathbf{u}\cdot\mathbf{z} \leqslant \varepsilon^2 \iff \mathbf{u}\cdot\mathbf{z} \geqslant 1 - \frac{\varepsilon^2}{2}. \tag{26}$$

Let us define the *ε-region* $\mathcal{R}_\varepsilon$ to be the corresponding subset of the unit disk. Let $\overline{\mathcal{D}} = \{\mathbf{u} \mid |\mathbf{u}|^2 \leqslant 1\}$ be the closed unit disk. Then

$$\mathcal{R}_\varepsilon = \{\mathbf{u} \in \overline{\mathcal{D}} \mid \mathbf{u}\cdot\mathbf{z} \geqslant 1 - \frac{\varepsilon^2}{2}\}. \tag{27}$$

The ε-region is shown as a shaded region in the illustration below. Note that the width of this region is $\frac{\varepsilon^2}{2}$ at its widest point, and its length, to second order, is 2ε, and in any case greater than $\sqrt{2}\,\varepsilon$.



In summary, $\|U - R_z(\theta)\| \leqslant \varepsilon$ if and only if $\mathbf{u} \in \mathcal{R}_\varepsilon$.

From now on, it will be convenient to identify the complex plane with $\mathbb{R}^2$, i.e., we will regard $\overline{\mathcal{D}}$ and $\mathcal{R}_\varepsilon$ as subsets of the complex numbers as well as $\mathbb{R}^2$.

6.2 Candidates

Assume, for the moment, that some suitable $k \geqslant 1$ has been fixed. The problem of finding an ε-approximation of $R_z(\theta)$ can now be reduced to the following 2-step problem:

(a) find $u \in \mathbb{Z}[\omega]$ such that $\hat{u} = u/\sqrt{2}^k$ is in the ε-region, and

(b) find $t \in \mathbb{Z}[\omega]$ such that $t^\dagger t + u^\dagger u = 2^k$.

Recall from Section 4 that we have an efficient solution to (b) if $\xi = 2^k - u^\dagger u$ satisfies the hypotheses of Theorem 12. We will quite literally leave one of these hypotheses, namely the primality of $\xi^\bullet\xi$, to chance. However, we will arrange things such that the remaining hypotheses are satisfied by design. To that end, we will say that $\hat{u}$ is a *candidate* if it satisfies all the required conditions, except possibly for the primality of $\xi^\bullet\xi$. This is made precise in the following definition and lemma.

Definition 20. Let $\varepsilon > 0$, $\theta \in \mathbb{R}$, and $k \geqslant 1$ be fixed. Let $\hat{u} = u/\sqrt{2}^k$, where $u \in \mathbb{Z}[\omega]$. Let $\xi = x + y\sqrt{2} = 2^k - u^\dagger u$. Then $\hat{u}$ is called a *candidate* if $\hat{u} \in \mathcal{R}_\varepsilon$, x is odd, y is even, $\xi \geqslant 0$, and $\xi^\bullet \geqslant 0$. We say that $\hat{u}$ is a *prime candidate* if, moreover, $p = \xi^\bullet\xi$ is prime.

We will further restrict attention to candidates $\hat{u} = u/\sqrt{2}^k$ where u is from the ring $\mathbb{Z}[\sqrt{2}, i] \subseteq \mathbb{Z}[\omega]$, i.e., of the form $u = \alpha + \beta i$ with $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$.

Lemma 21. *Let $\hat{u} = u/\sqrt{2}^k$, where $u = \alpha + \beta i$ and $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$. Let us write $\alpha = a + b\sqrt{2}$ and $\beta = c + d\sqrt{2}$. Then $\hat{u}$ is a candidate if and only if the following three conditions hold:*

- *$\hat{u} \in \mathcal{R}_\varepsilon$,*
- *$\hat{u}^\bullet \in \overline{\mathcal{D}}$, and*
- *$a + c$ is odd.*

Proof. With the given choice of notation, we have

$$\begin{aligned} \xi &= 2^k - u^\dagger u \\ &= 2^k - \alpha^2 - \beta^2 \\ &= (2^k - a^2 - 2b^2 - c^2 - 2d^2) - (2ab + 2cd)\sqrt{2}, \end{aligned} \tag{28}$$

hence $x = 2^k - a^2 - 2b^2 - c^2 - 2d^2$ and $y = 2(ab + cd)$. Then x is odd if and only if $a + c$ is odd. The condition that y is even is automatically satisfied. The condition $\xi \geqslant 0$ is equivalent to $u^\dagger u \leqslant 2^k$, or equivalently $\hat{u}^\dagger \hat{u} \leqslant 1$, i.e., $\hat{u} \in \overline{\mathcal{D}}$, which follows from $\hat{u} \in \mathcal{R}_\varepsilon$. Similarly, $\xi^\bullet \geqslant 0$ is equivalent to $\hat{u}^\bullet \in \overline{\mathcal{D}}$. □

6.3 Candidate selection

Theorem 22. *Let $\varepsilon > 0$ and θ be fixed, and let $k \geqslant C + 2\log_2(1/\varepsilon)$, where $C = \frac{5}{2} + 2\log_2(1 + \sqrt{2})$. Then there exists a set of at least $n = \lfloor \frac{4\sqrt{2}}{\varepsilon} \rfloor$ candidates $\hat{u}$ satisfying the conditions of Lemma 21. Moreover, there is an efficient algorithm for computing a random candidate from this set.*

Proof. First note that $k \geqslant C + 2\log_2(1/\varepsilon)$ implies $2^k \geqslant \frac{4\sqrt{2}(1+\sqrt{2})^2}{\varepsilon^2}$. Define

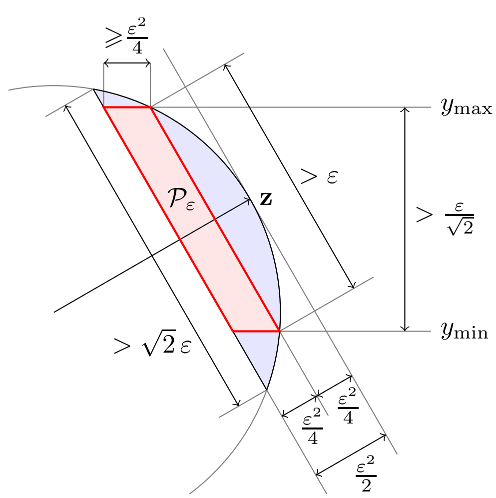
$$\delta = \sqrt{2}^k \, \frac{\varepsilon^2}{8} \quad \text{and} \quad \Delta = \sqrt{2}^{k+1}. \tag{29}$$

We note that δ and Δ satisfy the condition of Lemma 17. Indeed,

$$\delta\Delta = 2^k \, \frac{\sqrt{2}\,\varepsilon^2}{8} \geqslant \frac{8(1+\sqrt{2})^2}{\varepsilon^2} \, \frac{\varepsilon^2}{8} = (1+\sqrt{2})^2. \tag{30}$$

In the following, we will assume, for convenience, that $-\frac{\pi}{2} \leqslant \theta \leqslant \frac{\pi}{2}$. This assumption is without loss of generality, because otherwise, we may simply rotate the ε-region by multiples of 90° without changing the substance of the argument.

Now consider the line $\mathbf{u} \cdot \mathbf{z} = 1 - \varepsilon^2/4$. It intersects the unit circle in two points at y-coordinates $y_{\min}$ and $y_{\max}$, with $y_{\max} - y_{\min} \geqslant \frac{\varepsilon}{\sqrt{2}}$, as shown in this figure:



Consider the parallelogram $\mathcal{P}_\varepsilon$ that is bounded by the lines $\mathbf{u}\cdot\mathbf{z} = 1-\varepsilon^2/4$ and $\mathbf{u}\cdot\mathbf{z} = 1-\varepsilon^2/2$, and by the horizontal lines at $y = y_{\min}$ and $y = y_{\max}$. As illustrated in the above figure, $\mathcal{P}_\varepsilon$ is entirely contained within the ε-region. We will select candidates from within $\mathcal{P}_\varepsilon$.

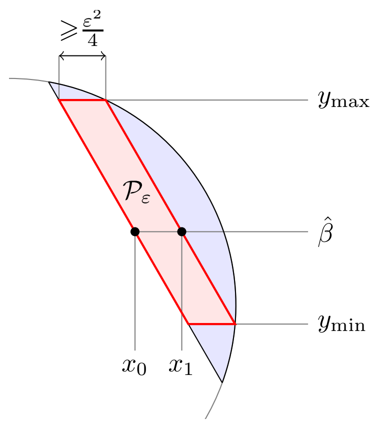
Let $n = \lfloor\frac{4\sqrt{2}}{\varepsilon}\rfloor$, and define $y_j = y_{\min} + \frac{j}{n}(y_{\max} - y_{\min})$ for $j = 0, \ldots, n$, so that $y_{\min} = y_0 < y_1 < \ldots < y_n = y_{\max}$. We note that

$$y_{j+1} - y_j = \frac{y_{\max} - y_{\min}}{n} > \frac{\varepsilon}{\sqrt{2}}\frac{\varepsilon}{4\sqrt{2}} = \frac{\varepsilon^2}{8}. \tag{31}$$

Let $I_j = [y_j, y_j + \frac{\varepsilon^2}{8}]$ for $j = 0, \ldots, n-1$. Then $I_0, \ldots, I_{n-1}$ are non-overlapping closed subintervals of $[y_{\min}, y_{\max}]$ of size $\frac{\varepsilon^2}{8}$. For each $j = 0, \ldots, n-1$, we will find a candidate $\hat{u} \in \mathcal{P}_\varepsilon \cap (\mathbb{R} \times I_j)$ as follows.

First, use Lemma 17 to find $\beta \in \mathbb{Z}[\sqrt{2}]$ such that $\beta \in [\sqrt{2}^k y_j, \sqrt{2}^k(y_j + \frac{\varepsilon^2}{8})]$ and $\beta^\bullet \in [-\sqrt{2}^{k-1}, \sqrt{2}^{k-1}]$. Note that these two intervals are of size δ and Δ, respectively, so the use of Lemma 17 is justified by (30). Let $\hat{\beta} = \beta/\sqrt{2}^k \in I_j$.

Because $\hat{\beta} \in [y_{\min}, y_{\max}]$, the line $y = \hat{\beta}$ intersects the parallelogram $\mathcal{P}_\varepsilon$, as shown in the figure below. Let $x_0 = \min\{x \mid (x, \hat{\beta}) \in \mathcal{P}_\varepsilon\}$ and $x_1 = \max\{x \mid (x, \hat{\beta}) \in \mathcal{P}_\varepsilon\}$, and note that $x_1 - x_0 \geqslant \frac{\varepsilon^2}{4}$.



Now we can use Corollary 19 to find $\alpha \in \mathbb{Z}[\sqrt{2}]$ such that $\alpha \in [\sqrt{2}^k x_0, \sqrt{2}^k(x_0 + \frac{\varepsilon^2}{4})]$ and $\alpha^\bullet \in [-\sqrt{2}^{k-1}, \sqrt{2}^{k-1}]$. Note that these two intervals are of size 2δ and Δ, respectively, so that the use of Corollary 19 is again justified by (30). Furthermore, writing $\alpha = a + b\sqrt{2}$ and $\beta = c + d\sqrt{2}$, Corollary 19 permits us to choose a odd if c is even, or vice versa.

Let $\hat{\alpha} = \alpha/\sqrt{2}^k$, and let $\hat{u} = \hat{\alpha} + \hat{\beta}i$. It is now trivial to show that $\hat{u}$ is a candidate. Indeed, since $\hat{\alpha} \in [x_0, x_1]$, we have that $\hat{u} \in \mathcal{P}_\varepsilon \subseteq \mathcal{R}_\varepsilon$ by construction. Also, note that, by construction, $(\hat{\alpha}^\bullet, \hat{\beta}^\bullet) \in [-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^2 \subseteq \overline{\mathcal{D}}$, so that $\hat{u}^\bullet \in \overline{\mathcal{D}}$. Finally, we already remarked that $a + c$ is odd.

Since we have found a distinct candidate for each $j \in \{0, \ldots, n-1\}$, there exist at least n candidates; moreover, the construction of the jth candidate is clearly algorithmic. $\square$

7 The main algorithm

7.1 Approximating a z-rotation

We are now ready to put together the results of earlier sections to obtain an algorithm for approximating $R_z(\theta)$ up to ε, using only gates from the single-qubit Clifford+T group.

Algorithm 23. Inputs: $0 < \varepsilon \leqslant \frac{1}{2}$ and $\theta \in \mathbb{R}$. Let $k = \lceil C + 2\log_2(1/\varepsilon)\rceil$, where $C = \frac{5}{2} + 2\log_2(1+\sqrt{2}) \approx 5.04$, and let $n = \lfloor\frac{4\sqrt{2}}{\varepsilon}\rfloor$. Use Theorem 22 to generate random candidates $\hat{u} \in \frac{1}{\sqrt{2}^k}\mathbb{Z}[\omega]$. For each candidate $\hat{u} = u/\sqrt{2}^k$, let $\xi = 2^k - u^\dagger u \in \mathbb{Z}[\sqrt{2}]$ and attempt to use the method of Theorem 12 to find $t \in \mathbb{Z}[\omega]$ with $t^\dagger t = \xi$. If this fails, continue with the next candidate. If it succeeds, the operator

$$U = \frac{1}{\sqrt{2}^k}\begin{pmatrix} u & -t^\dagger \\ t & u^\dagger \end{pmatrix}$$

is the desired ε-approximation of $R_z(\theta)$. Use the Kliuchnikov-Maslov-Mosca exact synthesis algorithm [12] to convert U to a sequence of Clifford+T gates; optionally, reduce the sequence of gates to Matsumoto-Amano normal form [6]. Output: the sequence of gates.

Remark 24. It is not necessary to limit the generation of candidates to the particular set of n candidates identified in Theorem 22. Instead, one can simply generate candidates at random until a solution is found.

Remark 25. Although the method of Theorem 12 will usually only work when $p = \xi^\bullet\xi$ is prime, Algorithm 23 requires no explicit primality test. Instead, one can simply perform the method of Theorem 12 under the optimistic assumption that p is prime. In the worst case, this may yield a t that is not a solution, but this can be easily checked after the fact. Some slight care is needed to ensure that the probabilistic algorithm from Remark 11, for finding a square root h of -1 modulo p, does not get into an infinite loop when p is not prime. For this, it is sufficient to limit the number of iterations of this algorithm to some small number, say 1 or 2. If p is indeed prime, this will still yield h with high enough probability; in the worst case, some prime candidates will be unnecessarily rejected.

7.2 Approximating arbitrary gates

To approximate an arbitrary gate $U \in SU(2)$, first decompose it via Euler angles as $U = R_z(\beta)R_x(\gamma)R_z(\delta) = R_z(\beta)\,H\,R_z(\gamma)\,H\,R_z(\delta)$; each rotation can then be approximated separately, say up to $\varepsilon/3$.

Remark 26. If one includes, as we did, only global phases of the form ω^j in the Clifford group, it is of course not possible to approximate unitary operators of arbitrary determinant, using only single-qubit Clifford+T gates. Since the determinant of every Clifford+T operator is a power of ω, an operator $U \in U(2)$ can be approximated to arbitrary ε if and only if $\det U$ is also a power of ω. Otherwise, U can only be approximated up to a global phase. If we include arbitrary global phases in the Clifford group, then of course all operators can be approximated.

8 Complexity analysis

8.1 Gate complexity

Fact 27. *Let U be a single-qubit Clifford+T operator with denominator exponent $k > 0$, and let n be the minimal T-count of U. Then $2k - 3 \leqslant n \leqslant 2k$.*

Proof. By induction on the Matsumoto-Amano normal form of U; see Theorem 7.10 of [15]. □

Since the matrix U constructed by Algorithm 23 has denominator exponent k, it follows by Fact 27 that the approximation of a z-rotation uses T-count at most $2k$, or $K + 4\log_2(1/\varepsilon)$, where $K \approx 10.09$. The approximation of an arbitrary element of $SU(2)$ requires three z-rotations, so the T-count is $K + 12\log_2(1/\varepsilon)$, where $K \approx 30.26$.

Remark 28. Rotations about other "easy" axes, such as x-rotations or y-rotations, can of course also be approximated with the same T-count as a z-rotation, as they differ from a z-rotation only by Clifford operators. More generally, rotations of the form $VR_z(\theta)V^\dagger$, where V is a fixed Clifford+T-operator, can be approximated with T-count $K + 4\log_2(1/\varepsilon)$, where K is a constant depending only on V.

8.2 Time complexity

Most parts of the algorithm are computationally straightforward; for example, the generation of candidates, and the various ring operations required for Theorem 12, can all be done with integers of length $O(k)$. The arithmetic operations, such as addition, multiplication, division, etc., require no more than $O(k^2)$ elementary steps each, and there are $O(k)$ arithmetic operations to perform.

The dominant complexity question is how many candidates must be generated before a solution is found, and indeed, whether a solution will be found at all. Here, we must make an assumption about the distribution of primes:

Hypothesis 29. For a randomly chosen candidate, the probability that $p = \xi^\bullet\xi$ is prime is asymptotically no smaller than for general odd numbers of comparable size.

We note that for a candidate $\hat{u} = u/\sqrt{2}^k$, we have

$$\hat{u}^\dagger\hat{u} \geqslant \left(1 - \frac{\varepsilon^2}{2}\right)^2 \geqslant 1 - \varepsilon^2, \tag{32}$$

and thus

$$\xi \;=\; 2^k - u^\dagger u \;=\; 2^k(1 - \hat{u}^\dagger\hat{u}) \;\leqslant\; 2^k\varepsilon^2 \;\leqslant\; \frac{4\sqrt{2}(1+\sqrt{2})^2}{\varepsilon^2}\,\varepsilon^2 = 4\sqrt{2}(1+\sqrt{2})^2 \leqslant 33. \tag{33}$$

Also, $\xi^\bullet = 2^k - u^{\bullet\dagger}u^\bullet \leqslant 2^k$. It follows that $p = \xi^\bullet\xi \leqslant 33 \cdot 2^k$. By the prime number theorem, a randomly chosen odd p in this range has probability

$$P \approx \frac{2}{\ln(33 \cdot 2^k)} = \Omega(\frac{1}{k}) \tag{34}$$

of being prime; by Hypothesis 29, at least the same probability holds for a randomly chosen candidate. On the other hand, there are $n = \lfloor \frac{4\sqrt{2}}{\varepsilon} \rfloor = O(\sqrt{2}^k)$ candidates available, so asymptotically, a prime will be found with certainty. The expected number of top-level iterations of Algorithm 23 until a solution is found is $O(k)$. Therefore, under Hypothesis 29, the overall runtime of the algorithm is no more than $O(k^4) = O(\log^4(1/\varepsilon))$. A sharper bound could probably be derived by a more sophisticated analysis.

We also note that, due to its randomized nature, and because each candidate is chosen independently, the algorithm can easily be parallelized.

8.3 Seeding

While the above discussion concerns the asymptotic case, one may wonder whether for *particular* values of ε and θ it may happen, by unlucky coincidence, that none of the available candidates are prime. In practice, this never seems to be the case, as primes are always found quite quickly. However, in theory, we may avoid this situation by the method of *seeding*: Instead of approximating $R_z(\theta)$ directly, we may choose a random angle ("seed") ϕ, approximate both $R_z(\phi)$ and $R_z(\theta - \phi)$ to within $\varepsilon/2$, and finally compute $R_z(\theta)$ as $R_z(\phi)R_z(\theta - \phi)$. If a particular seed seems to yield no prime candidates in a given number of iterations, one can simply try a different seed.

The seeding method for z-rotations has the disadvantage that it approximately doubles the T-count, from $K + 4\log_2(1/\varepsilon)$ to $K + 8\log_2(1/\varepsilon)$.

In the approximation of an arbitrary gate $U \in SU(2)$, a different seeding method can be used, which only increases the asymptotic T-count by a small additive constant. Namely, let the seed be some randomly chosen Clifford+T operator V of fixed and relatively small T-count. We can then use the algorithm to approximate UV^{-1}, and multiply the final result by V. The gate complexity, in this case, is unchanged at $K + 12\log_2(1/\varepsilon)$.

It must be emphasized, once again, that seeding is only of theoretical interest, to ensure expected termination of the algorithm under a hypothetical worst-case scenario. In practice, it does not seem to be required.

9 Lower bounds

As mentioned in the introduction, $K + 3\log_2(1/\varepsilon)$ is an easy lower bound for the T-count of a Clifford+T approximation of some arbitrary operator in $SU(2)$ up to ε. Specifically, Matsumoto and Amano showed that there are precisely $192 \cdot (3 \cdot 2^n - 2)$ distinct single-qubit Clifford+T-circuits of T-count at most n [6]. Since $SU(2)$ is a 3-dimensional manifold, it requires $\Omega(1/\varepsilon^3)$ epsilon-balls to cover. The resulting inequality

$$192 \cdot (3 \cdot 2^n - 2) \geqslant \frac{c}{\varepsilon^3}, \tag{35}$$

immediately implies

$$n \geqslant K + 3\log_2(\frac{1}{\varepsilon}). \tag{36}$$

This means that there is some universal constant K such that for every ε, there exists some operator $U \in SU(2)$ that cannot be approximated up to ε with T-count less than $K + 3\log_2(1/\varepsilon)$. In fact, (35) implies that the proportion (in the Haar measure, say) of operators that can be approximated with T-count $K + 3\log_2(1/\varepsilon) - k$ is at most $O(\frac{1}{2^k})$, so that "most" operators require T-counts that are close to or above the lower bound.

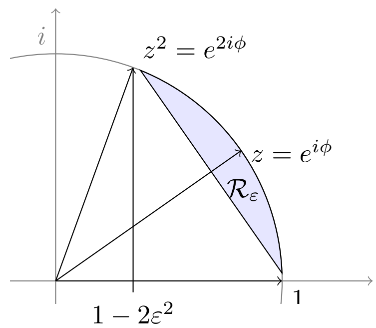
However, like all lower bounds, this must be understood with a grain of salt. For example, the set of z-rotations only forms a 1-dimensional submanifold of $SU(2)$, so it is not a priori clear that the lower bound of $K + 3\log_2(1/\varepsilon)$ also applies to z-rotations.

We now prove a sharper lower bound, which applies to z-rotations in particular.

Theorem 30. *Let $K' = -9$. Then for all $\varepsilon > 0$, there exists an angle θ such that every ε-approximation of $R_z(\theta)$ by a product of Clifford+T operators requires T-count at least $K' + 4\log_2(1/\varepsilon)$.*

Proof. Let ε be given. If $\varepsilon \geqslant 1/2$, then $K' + 4\log_2(1/\varepsilon)$ is negative, so there is nothing to show. Assume, therefore, that $\varepsilon < 1/2$. Let $\phi = \sin^{-1}\varepsilon$, with $0 < \phi < \pi/6$. Let $\theta = -2\phi$, so that $z = e^{-i\theta/2} = e^{i\phi}$. We note the shape of the

ε-region for z:



Identifying complex numbers with vectors in $\mathbb{R}^2$ as before, we estimate the dot product

$$\mathbf{1} \cdot \mathbf{z} = \cos\phi = \sqrt{1-\varepsilon^2} < \sqrt{1-\varepsilon^2+\frac{\varepsilon^4}{4}} = 1 - \frac{\varepsilon^2}{2}. \tag{37}$$

It follows that neither 1 nor z^2 is in the ε-region. We further note that the x-coordinate of z^2 is $\cos 2\phi = 1 - 2\sin^2\phi = 1 - 2\varepsilon^2$. So for all $x + iy \in \mathcal{R}_\varepsilon$, we have

$$1 - 2\varepsilon^2 < x < 1. \tag{38}$$

Now consider some ε-approximation

$$U = \frac{1}{\sqrt{2}^k}\begin{pmatrix} u & s \\ t & v \end{pmatrix} \tag{39}$$

of $R_z(\theta)$ with denominator exponent k, i.e., where $u, t, s, v \in \mathbb{Z}[\omega]$. From $\|U - R_z(\theta)\| \leqslant \varepsilon$, we get $|\hat{u} - z|^2 + |\hat{t}|^2 \leqslant \varepsilon^2$, where $\hat{u} = u/\sqrt{2}^k$ and $\hat{t} = t/\sqrt{2}^k$. By the same reasoning as in Section 6.1, it follows that $\hat{u}$ is in the ε-region. Moreover since $U^\bullet$ is also unitary, we have $\hat{u}^\bullet \in \overline{\mathcal{D}}$. We can write

$$\hat{u} = \frac{1}{\sqrt{2}^k}(a\omega^3 + b\omega^2 + c\omega + d) = \frac{1}{\sqrt{2}^{k+1}}(d\sqrt{2} + c - a) + (b\sqrt{2} + c + a)i = \frac{1}{\sqrt{2}^{k+1}}(\alpha + \beta i),$$

where $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$. Since $\hat{u} \in \mathcal{R}_\varepsilon$, from (38), we have

$$\sqrt{2}^{k+1}(1 - 2\varepsilon^2) < \alpha < \sqrt{2}^{k+1}. \tag{40}$$

Also, from $\hat{u}^\bullet \in \overline{\mathcal{D}}$, we have

$$\alpha^\bullet \in [-\sqrt{2}^{k+1}, \sqrt{2}^{k+1}]. \tag{41}$$

It follows that the constraints

$$(\gamma, \gamma^\bullet) \in [\sqrt{2}^{k+1}(1 - 2\varepsilon^2), \sqrt{2}^{k+1}] \times [-\sqrt{2}^{k+1}, \sqrt{2}^{k+1}] \tag{42}$$

have at least two different solutions in $\mathbb{Z}[\sqrt{2}]$, namely $\gamma = \alpha$ and $\gamma = \sqrt{2}^{k+1}$. Setting $\delta = \sqrt{2}^{k+1}\,2\varepsilon^2$ and $\Delta = 2\sqrt{2}^{k+1}$, we have by Lemma 16:

$$1 \leqslant \delta\Delta = 2^{k+3}\varepsilon^2, \tag{43}$$

or equivalently,

$$k \geqslant \log_2(\frac{1}{8\varepsilon^2}) = -3 + 2\log_2(\frac{1}{\varepsilon}). \tag{44}$$

On the other hand, it follows from Fact 27 that every single-qubit Clifford+T operator U of T-count n can be written with denominator exponent k, where

$$k \leqslant \frac{n+3}{2}. \tag{45}$$

Putting together (44) and (45), we get

$$n \geqslant -9 + 4\log_2(\frac{1}{\varepsilon}), \tag{46}$$

which is the desired result. □

Remark 31. Unlike the lower bound (36), which applies to *typical* operators, the lower bound (46) only applies to carefully chosen worst-case operators. It is plausible that for any fixed ε, approximations of order (36) can be achieved for most angles θ. Moreover, it is also plausible that for any fixed θ, approximations of order (36) can be achieved as $\varepsilon \to 0$. A variation of Algorithm 23 that could potentially achieve this is sketched in Section 10, but the details are left for future work.

ε	k	T-count	Actual error	Runtime	Candidates	Time/Candidate
10^{-10}	72	142	$0.90665 \cdot 10^{-10}$	$0.02s$	37.80	$0.0005s$
10^{-20}	139	278	$0.84346 \cdot 10^{-20}$	$0.05s$	86.60	$0.0005s$
10^{-30}	205	410	$0.82984 \cdot 10^{-30}$	$0.07s$	97.94	$0.0007s$
10^{-40}	272	542	$0.73279 \cdot 10^{-40}$	$0.10s$	114.00	$0.0009s$
10^{-50}	338	676	$0.83841 \cdot 10^{-50}$	$0.13s$	135.66	$0.0010s$
10^{-60}	405	810	$0.73964 \cdot 10^{-60}$	$0.21s$	213.22	$0.0010s$
10^{-70}	471	942	$0.72360 \cdot 10^{-70}$	$0.24s$	201.76	$0.0012s$
10^{-80}	538	1076	$0.96804 \cdot 10^{-80}$	$0.30s$	223.72	$0.0013s$
10^{-90}	604	1208	$0.90793 \cdot 10^{-90}$	$0.42s$	283.36	$0.0015s$
10^{-100}	670	1338	$0.92860 \cdot 10^{-100}$	$0.51s$	327.82	$0.0016s$
10^{-200}	1335	2670	$0.77785 \cdot 10^{-200}$	$2.48s$	526.16	$0.0047s$
10^{-500}	3328	6656	$0.71348 \cdot 10^{-500}$	$47.82s$	1388.42	$0.0344s$
10^{-1000}	6650	13300	$0.80519 \cdot 10^{-1000}$	$504.80s$	2873.80	$0.1757s$

Table 1: Experimental results. The operator approximated is $R_z(\pi/128)$. Errors are measured in the operator norm. The runtime is averaged over 50 independent runs of the algorithm with the same parameters. The runtime is further broken down into average number of candidates tried per run of the algorithm, and time spent per candidate.

10 Overclocking

It is in the nature of Algorithm 23 that k is chosen at the very beginning; the final sequence of gates will have T-count very close to $2k$. This behavior is different from that of search-based algorithms, which would typically try shorter decompositions first, and then gradually move to longer ones.

In practice, many of the approximations and estimates in the above proofs are conservative; it is often possible to choose a smaller k than that prescribed by the algorithm, and still obtain a decomposition. We refer to this technique as "overclocking" the algorithm, by analogy with the practice of running microprocessors at higher clock speeds than they were designed for, and hoping for the best.

Let us first consider the effect of overclocking by a small additive constant, i.e., decreasing the value of C in Algorithm 23 by some fixed amount, independently of ε. This decreases the width of the ε-region by a fixed multiplicative factor, which means that some choices of β (in Theorem 22) will no longer yield a successful solution for α. However, the *proportion* of β for which an α can be found will be asymptotically independent of ε, so that the algorithm will still yield a solution, albeit with a longer running time. In summary, the asymptotic effect of overclocking C by a fixed additive amount is to increase the runtime by a fixed factor, while neither affecting the validity nor the big-O time complexity of the algorithm.

A more interesting question is whether the algorithm can be modified to allow overclocking by a *multiplicative* factor; in the ideal case, one might even hope to achieve approximations with T-count $K+3\log_2(1/\varepsilon)$, for most angles θ. As it is currently stated, the algorithm will not work well for multiplicative overclocking, because the ε-region will then be much too thin to expect to find candidates for a reasonable fraction of y-coordinates β. However, the *area* of the ε-region scales as ε^3, and since the density of grid points scales as $4^k \sim 1/\varepsilon^3$, one still expects, on average, to find a constant number of candidates in the ε-region. This assumes that the angle θ is sufficiently general, so that the ε-region is not aligned with the orientation of the underlying k-grid, so that the lower bound of Theorem 30 can be avoided. It would be an interesting optimization to refine Algorithm 23 to be able to take advantage of such sparsely populated ε-regions, but the details are left to future work.

11 Experimental results

Algorithm 23 has been implemented in the Haskell programming language, and is available from [16]. Table 1 summarizes the results of approximating $R_z(\pi/128)$ up to various ε, using one core of a 3.40GHz Intel i5-3570 CPU.

We may note that the runtime per candidate increases predictably, and polynomially, with decreasing ε, due to the increasing size of the integers and real numbers that must be operated upon. The number of candidates tried per run of the algorithm is approximately proportional to k, which is expected according to (34). This lends some empirical evidence to the validity of Hypothesis 29.

Note that we obtained a decomposition with accuracy $\varepsilon = 10^{-100}$, using only 1338 T-gates. By comparison, the Solovay-Kitaev algorithm with comparable T-count achieves less than $\varepsilon = 10^{-10}$ [12, Table 2].

As an example, here is the decomposition of $R_z(\pi/128)$ up to $\varepsilon = 10^{-10}$:

$$
\begin{aligned}
\hat{u} &= \frac{1}{\sqrt{2}^{72}}(-22067493351\omega^3 - 22078644868\omega^2 + 52098814989\omega + 16270802723) \\
\hat{t} &= \frac{1}{\sqrt{2}^{72}}(18093401340\omega^3 - 18136198811\omega^2 + 7555056984\omega + 7451734762) \\
U &= \texttt{HTHTHTSHTSHTHTSHTSHTHTSHTHTSHTSHTSHTHTSHTHTHTSHTSHTHTHTHTSHTHTHTSHTHTSH} \\
&\ \texttt{THTSHTHTHTSHTHTHTSHTSHTSHTHTSHTHTHTHTSHTHTSHTHTHTSHTSHTHTHTHTSHTSHTHTSH} \\
&\ \texttt{THTHTSHTSHTHTSHTSHTHTSHTHTSHTHTSHTSHTHTHTSHTHTHTSHTHTSHTSHTSHTHTHTHTSHT} \\
&\ \texttt{SHTHTHTHTSHTHTHTSHTHTSHTHTHTHTHTSHTHTSHTHTHTSHTHTSHTSHTHTSHTSHTSHTSHTSH} \\
&\ \texttt{THTHTHTHTSHTHTHTSHTSHTSHTHTHTSHTSHTSHTSHTHTHTHTHTHTHTHTSHTHTHTHTH}\omega^7
\end{aligned}
$$

12 Conclusions

We have given the most efficient algorithm to date for decomposing an element of $SU(2)$ into a product of Clifford+T gates, up to arbitrarily small ε. The algorithm performs well both in theory and in practice, easily achieving decompositions up to $\varepsilon = 10^{-100}$ using T-counts of less than 1400.

Like the recent ancilla-based algorithm by Kliuchnikov, Maslov, and Mosca [9], our algorithm is based on solving a Diophantine equation. Although both algorithms achieve the same optimal asymptotic big-O complexity, the gate decompositions resulting from our algorithm are shorter (by a constant, but non-negligible factor) than those that can be achieved by the Kliuchnikov-Maslov-Mosca approximation algorithm.

Acknowledgements

Thanks to Eric Van Den Berg and Don LeVine for useful comments on an earlier draft of this paper. This research was conducted while the author was a visiting Research Professor at the Mathematical Sciences Research Institute in Berkeley, California. This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). This research was supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D12PC00527. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.

References

[1] X. Zhou, D. W. Leung, and I. L. Chuang. Methodology for quantum logic gate construction. *Phys. Rev. A*, 62:052316 (12 pages), Oct 2000. Also available from `arXiv:quant-ph/0002039`.

[2] M. Amy, D. Maslov, M. Mosca, and M. Roetteler. 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):818–830, 2013. Also available from `arXiv:1206.0758`.

[3] A. Y. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.

[4] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. Graduate Studies in Mathematics 47. American Mathematical Society, 2002.

[5] C. M. Dawson and M. A. Nielsen. The Solovay-Kitaev algorithm. *Quantum Information and Computation*, 6(1):81–95, Jan. 2006. Also available from `arXiv:quant-ph/0505030`.

[6] K. Matsumoto and K. Amano. Representation of quantum circuits with Clifford and $\pi/8$ gates. `arXiv:0806.3834`, June 2008.

[7] A. G. Fowler. Constructing arbitrary Steane code single logical qubit fault-tolerant gates. *Quantum Information and Computation*, 11(9–10):867–873, 2011. Also available from `arXiv:quant-ph/0411206`.

[8] G. Duclos-Cianci and K. M. Svore. Distillation of nonstabilizer states for universal quantum computation. *Phys. Rev. A*, 88:042325 (7 pages), 2013. Also available from `arXiv:1210.1980`.

[9] V. Kliuchnikov, D. Maslov, and M. Mosca. Asymptotically optimal approximation of single qubit unitaries by Clifford and T circuits using a constant number of ancillary qubits. *Phys. Rev. Lett.*, 110:190502 (5 pages), 2013. Also available from `arXiv:1212.0822`.

[10] V. Kliuchnikov, A. Bocharov, and K. M. Svore. Asymptotically optimal topological quantum compiling. `arXiv:1310.4150`, Oct. 2013.

[11] N. Wiebe and V. Kliuchnikov. Floating point representations in quantum circuit synthesis. *New Journal of Physics*, 15:093041 (24 pages), 2013. Also available from `arXiv:1305.5528`.

[12] V. Kliuchnikov, D. Maslov, and M. Mosca. Fast and efficient exact synthesis of single qubit unitaries generated by Clifford and T gates. *Quantum Information and Computation*, 13(7–8):607–630, 2013. Also available from `arXiv:1206.5236`.

[13] H. Cohen. *Advanced Topics in Computational Number Theory*. Graduate Texts in Mathematics 193. Springer, 2000.

[14] M. O. Rabin and J. O. Shallit. Randomized algorithms in number theory. *Communications in Pure and Applied Mathematics*, 39:S239–S256, 1986.

[15] B. Giles and P. Selinger. Remarks on Matsumoto and Amano's normal form for single-qubit Clifford+T operators. `arXiv:1312.6584`, Dec. 2013.

[16] P. Selinger. Newsynth: exact and approximate synthesis of quantum circuits. Software implementation, available from `http://www.mathstat.dal.ca/~selinger/newsynth/`, 2013.