A Review of Unitary Approximation with Clifford+T Gates
Thesis for the Honors Program in Computer Science

Junyi "Bob" Zou*

1 Introduction

In this thesis, I will review the state-of-art Ross-Selinger algorithm for single-qubit unitary approximation with Clifford and T gates, give a more detailed and self-contained description of it and provide an alternative, lattice-based approach for simplifying part of the algorithm.

Single-bit universal quantum computation requires the ability to approximate any 2×2 unitary matrix with a circuit composed of gates from a universal gate set. A universal gate set is a fixed and finite set consisting of topological generators of $PU(2)$ (not $U(2)$ because global phase isn't relevant), which means it generates a topologically dense set in $PU(2)$ (the metric used is often the l_2 operator norm). There are many possible universal gate sets, and in this thesis, I will focus on the Clifford+T gate set, which consists of the Clifford gates and the T gate. The Clifford gates are a group of order 24 generated by the Hadamard gate (H), the Phase gate (S) and a global phase $\omega = e^{i\pi/4}$. There are two reasons for choosing this particular set. First, they are suitable for fault-tolerant implementations, which is crucial for practical purposes [ZLC00]. Second, the Clifford+T gate set is one of the few "Super-Golden-Gate" sets defined in [PS18], which means not only Clifford+T can approximate any member of $PU(2)$ with a reasonably short circuit, they also require minimal number of T gates. The second consideration mainly arises because of the fact that the cost of the T gate is considerably higher than the Clifford gates and therefore T-count has been the measure of circuit size complexity in this context.

The classical approach for finding circuits of a given universal gate set to approximate members of $PU(2)$ is the Solovay-Kitaev algorithm. Given error tolerance ε, it runs in $O(\log^{2.71}(1/\varepsilon))$ time and gives circuits of size $O(\log^{3.97}(1/\varepsilon))$[DN06]. This is not ideal, considering there is no guarantee for the T-count of the circuit found to be low. In recent years, there have been attempts to optimize T-count. For example, [KMM16] proposed an exponential run time algorithm that guarantees optimal T-count with the use of ancillas, which can already deal with $\varepsilon \approx 10^{-15}$. Their most significant contribution, however, is the incorporation of the use of a Diophantine's equation (norm equation) in their algorithm, which then inspired an ancilla-free version of the algorithm by [Sel15]. Later, based on Selinger's earlier work, Ross and Selinger proposed this new state-of-art algorithm [RS16] that runs in polylog time and guarantees optimal circuit T-count with a factoring oracle. Even in the absence of a factoring oracle, the typical case T-count is $3\log_2(1/\varepsilon) + O(\log(\log(1/\varepsilon))$, where $3\log_2(1/\varepsilon)$ is the information theoretical lower bound.

*Courant Institute of Mathematical Sciences, New York University. Email: jz2819@nyu.edu.

1.1 A High Level Overview

Both [KMM16] and [RS16] actually only considered the problem of approximating arbitrary z-rotations because of the well-known decomposition of a unitary into 2 Hadamards and 3 z-rotations. Another key observation, made by [KMM13], is that the group generated by Clifford+T is exactly $D[\omega]^{2\times 2} \cap U(2)$, and they also provided a synthesis algorithm for conversion. For approximation purposes one only needs $G = D[\omega]^{2\times 2} \cap SU(2)$ [RS16].

Furthermore, it is noted in [GS19] that the T-count of a member of G is given by the SDE (smallest denominator exponent) of its entries. Simply put, the lower the SDE, the smaller the T-count.

Given a target z-rotation R_z and an error tolerance ε, the goal of the algorithm is thus simply to search for the member of G that is within ε distance from R_z and has the smallest SDE. This is done in two steps in Ross-Selinger. First, it searches for a $u \in D[\omega]$, such that any member of $D[\omega]^{2\times 2} \cap U(2)$ whose first entry is u will be close enough to the target, in increasing order of SDE. Second, for each candidate found, it enforces the condition that it leads to a matrix belonging to $SU(2)$ by solving a norm equation which simply equates the determinant to 1. Since the search part returns candidates in increasing order of SDE, as long as the norm equation can be solved efficiently (which requires a factoring oracle), we are guaranteed to have minimal T-count. In the absence of a factoring oracle, certain candidates found may lead to norm equations that take too long to solve and thus must be skipped. But it can be shown that typically, this does not happen very much so the T-count is still reasonably low.

1.2 Content Roadmap

This thesis, including this introduction section and the last acknowledgement section, has 10 sections.
In Section 2, some basic mathematical definitions are given. Readers familiar with these definitions can skip this section.
In Section 3, concepts and definitions key to the understanding of Ross-Selinger are introduced.
In Section 4, an overview and discussions about the two major components of the algorithm are given.
In Section 5, the grid-operator approach to the search component, adopted in [RS16], is summarized and discussed.
In Section 6, the novel lattice-region-search approach to the search component is presented.
In Section 7, a detailed explanation of the norm equation solver used in [RS16] is given.
In Section 8, the full algorithm is given.
In Section 9, analysis of T-count and time complexities for various cases is given.

For readers only interested in understanding the algorithm at a high level, Section 5, 6, 7, 9 can be skipped. For readers who are primarily interested in novelties in this thesis, Section 5 can be skipped and more attention should be paid to Section 4 and 6.

Overall, much of this thesis is written based on existing results, except for Section 4 and 6. However, even for existing results, the proofs have been rewritten and reorganized, with the addition of my own comments and remarks. The hope is that this thesis provides a more comprehensive and easy-to-follow introduction to the unitary approximation problem than the references cited.

1.3 Future Work

Several interesting problems can be further explored.

First, it is still unknown if decomposing an arbitrary unitary matrix into z-rotations and approximating the z-rotations only would yield the most cost-efficient circuit. There might be more efficient ways for the Clifford+T set to navigate through $PU(2)$, and optimal/near-optimal algorithms would be highly meaningful.

Second, it should be noted that Clifford+T is not the only "Super-Golden-Gate" set available. Do similar algorithms for other gate sets exist? If so, how easy would it be to generalize Ross-Selinger to other "Super-Golden-Gate" sets?

Third, it is plausible that for other gate sets, one would need to solve norm equations in different cyclotomic rings. A systemic approach to solve such equations using the Gentry-Szydlo algorithm was proposed in [HGS04]. Nevertheless, it only works for cyclotomic extensions of prime order. Can we make the approach from [HGS04] work for non-prime order cyclotomic rings?

2 Preliminaries

The Clifford Group is generated by the following three generators:

1. The Hadamard Gate:
$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$
2. The Phase Gate:
$$S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$$
3. A global phase:
$$\begin{bmatrix} \omega & 0 \\ 0 & \omega \end{bmatrix} \cong \omega = e^{i\pi/4}$$

The T gate is the square root of the Phase gate:

$$T = \sqrt{S} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$$

The ring D represents the ring of all the dyadic integer:

$$D = \{a/2^k \mid a \in \mathbb{Z}, k \in \mathbb{N}\}$$

The ring $\mathbb{Z}[\sqrt{2}]$ is obtained by adjoining $\sqrt{2}$ to $\mathbb{Z}$:

$$\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$$

Similarly, the ring $D[\sqrt{2}]$ is obtained by adjoining $\sqrt{2}$ to D:

$$D[\sqrt{2}] = \{a + b\sqrt{2} \mid a, b \in D\}$$

The ring $\mathbb{Z}[\omega]$ is obtained by adjoining ω to $\mathbb{Z}$:

$$\mathbb{Z}[w] = \{a + bw + cw^2 + dw^3 \mid a, b, c, d \in \mathbb{Z}, w = e^{i\pi/4}\}$$

Similarly, the ring $D[\omega]$ is obtained by adjoining ω to D:

$$D[w] = \{a + bw + cw^2 + dw^3 \mid a, b, c, d \in D, w = e^{i\pi/4}\}$$

The l_2 operator norm of a 2×2 matrix U is given by:

$$\|U\|_2 = \sup_{x \in \mathbb{R}^2, x \neq 0} \frac{\|Ux\|}{\|x\|} = \sigma_{\max}(U)$$

where $\sigma_{\max}(U)$ is the largest singular value. When U is normal, this is also equal to $|\lambda|_{\max}(U)$, the largest absolute eigenvalue of U.

3 Problem Set-up

3.1 Decomposition of 2 by 2 unitary matrices

In this subsection, I will show that any 2 by 2 unitary matrix can be decomposed into three z-rotations and two Hadamards. Therefore, the task of approximating any 2 by 2 unitary matrix can be reduced to approximating any z-rotation.

Definition 3.1. *The z-rotation matrix is defined as as the exponential of the Pauli-Z gate and Hadamard gate is the standard H gates*

$$R_z(\theta) = e^{-i\theta Z/2} = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{1}$$

Theorem 3.2. *Any 2 by 2 unitary matrix can be decomposed into exactly 3 z-rotations and 2 Hadamards, up to a global phase.*

Proof. First note we have the following general form for any 2 by 2 unitary matrix:

$$U = \begin{bmatrix} a & b \\ -e^{i\phi} b^\dagger & e^{i\phi} a^\dagger \end{bmatrix} \quad |a|^2 + |b|^2 = 1 \tag{2}$$

Since the norm square of a, b must add up to 1, without loss of generality we can assume a, b have norm $\sin\theta$ and $\cos\theta$. Also note that a, b can always be replaced by $e^{i\phi/2}c, e^{i\phi/2}d$, with c, d having the same norm. Further, suppose c has phase ψ_1 and d has phase ψ_2. We can rewrite 2 as:

$$U = \begin{bmatrix} e^{i\phi/2}c & e^{i\phi/2}d \\ -e^{i\phi/2}c^\dagger & e^{i\phi/2}d^\dagger \end{bmatrix} = e^{i\phi/2} \begin{bmatrix} e^{i\psi_1}\cos\theta & e^{i\psi_2}\sin\theta \\ -e^{-i\psi_2}\sin\theta & e^{-i\psi_1}\cos\theta \end{bmatrix} \tag{3}$$

Ignoring the global phase, we see that U has 3 degrees of freedom, which are controlled exactly by the 3 z-rotations. To see this, use the surrogate variables $\alpha = (\psi_1 + \psi_2)/2$ and $\beta = (\psi_1 - \psi_2)/2$, we can rewrite 3 as:

$$U = e^{i\phi/2} \begin{bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{i\beta} & 0 \\ 0 & e^{-i\beta} \end{bmatrix} \tag{4}$$

The first and third matrix in the product are already z-rotations. We can obtain plane in the middle rotation by the following:

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{5}$$

Putting everything together we obtain:

$$U = e^{\phi/2} R_Z(-2\alpha) H R_Z(-2\theta) H R_Z(-2\beta) \tag{6}$$

□

Therefore, in this thesis, we will focus on how to approximate z-rotations with a circuit of Clifford gates and T gates.

3.2 Global Phase

So far we have always been ignoring the global phase because the global phase has no effect on the actual rotation that the operator (gate) represents. This is true, but it is also dangerous to completely omit the discussion of the global because the metric used, the l_2 operator norm, is not invariant with respect to the global phase. Specifically, let $R_z(\theta)$ be our target z-rotation and U be the circuit we synthesized with Clifford and T gates, then $\|R_z(\theta) - \lambda U\|_2$ is not invariant with respect to λ. An extreme example would be to approximate $R_z(-\pi/4) = e^{-i\pi/8}T$. Setting $\lambda = e^{i\pi/8}$ leads to the exact solution $U = \omega^{-1}T$, while setting $\lambda = 1$ would obviously lead to longer circuits. In this section, I will discuss how [RS16] showed that it actually suffices to only consider two cases $\lambda = 1$ and $\lambda = \sqrt{\omega} = e^{i\pi/8}$.

Lemma 3.3. *[RS16] Suppose $W \in SU(2)$ and $tr(W) > 0$, then we have:*

$$\forall \lambda \in \{z \in \mathbb{C} \mid \|z\| = 1\}, \quad \|I - W\|_2 \leq \|I - \lambda W\|_2 \tag{7}$$

Proof. With loss of generality, one can assume W takes the following form:

$$W = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix}, \quad 0 \leq \phi \leq \pi/2$$

This is because: first, when computing the distance between W and another diagonal matrix, only diagonal entries of W matter; second, determinant being 1 suggests the diagonal entries are complex conjugates; third, having trace being greater than 1 suggests the entries are in the positive half plane; forth by symmetry we have $\|I - W\|_2 = \|I - W^\dagger\|_2$.

Then suppose $\lambda = e^{i\psi}$, no matter what ψ is, it will make either $|1 - e^{i(\phi+\psi)}|$ or $|1 - e^{i(-\phi+\psi)}|$ greater than $|1 - e^{i\phi}|$ (this only requires a simple geometric visualization) and since

$$\|I - \lambda W\|_2 = \max(|1 - e^{i(\phi+\psi)}|, |1 - e^{i(-\phi+\psi)}|), \quad \|I - W\|_2 = |1 - e^{i\phi}|$$

we completes the proof. □

Lemma 3.4. *[RS16] Given error ε, $R \in SU(2)$ and U synthesized by Clifford+T, we have the following:*

$$\exists \lambda \in \{z \in \mathbb{C} \mid \|z\| = 1\},\ \|R - \lambda U\|_2 \leq \varepsilon \iff \exists n \in \mathbb{Z},\ \|R - \omega^{n/2} U\| \leq \varepsilon \tag{8}$$

Proof. Only the only if direction requires a proof. First, since U is synthesized by Clifford+T, its determinant is ω^k for some $k \in \mathbb{Z}$ (see 9) and thus $\det(R^{-1}U) = \omega^k$. Then choose $n = -k/-k+8$, $\lambda' = \omega^{-n/2}\lambda$, $W = \omega^{n/2}R^{-1}U$ such that:

$$\det(W) = 1, \quad \text{tr}(W) > 0$$

then by Lemma 3.3, we have:

$$\begin{aligned} & \|I - W\|_2 \leq \|I - \lambda' W\|_2 \\ \implies & \|I - \omega^{n/2}R^{-1}U\|_2 \leq \|I - \lambda'\omega^{n/2}R^{-1}U\|_2 \\ \implies & \|R - \omega^{n/2}U\|_2 \leq \|R - \lambda'\omega^{n/2}U\|_2 \\ \implies & \|R - \omega^{n/2}U\|_2 \leq \|R - \lambda U\|_2 \end{aligned}$$

□

Theorem 3.5. *[RS16] Given an error ε and a target z-rotation $R_z(\theta)$, to find a U synthesized by Clifford+T and a unit scalar λ such that:*

$$\|R_z(\theta) - \lambda U\| \leq \varepsilon$$

while minimizing the T-count of U, it suffices to consider just two cases: $\lambda = 1$ and $\lambda = \sqrt{w}$.

Proof. Suppose there is a pair of U, λ that is within the error tolerance, then by Lemma 3.3, there also exists $\lambda = \omega^{n/2}$ such that λU is also within the error tolerance. Next choose some $k \in \mathbb{Z}$ such that $\lambda' = \omega^k\lambda \in \{1, w^{1/2}\}$, and let $U' = \omega^{-k}U$, we have:

$$\lambda U = \lambda' U'$$

Furthermore, note ω is a Clifford gate, so U' can be synthesized with the same T-count as U. □

Therefore, to include global phase in the approximation algorithm, we simply need to run the algorithm twice, once with $\lambda = 1$ and another with $\lambda = \sqrt{\omega}$. There are only minor differences between the two runs and therefore, for the rest of the thesis I will focus on just the $\lambda = 1$ case.

3.3 The role of $D[\omega]$ and Best Approximation

In this section, we briefly look at how the ring $D[\omega]$ is related to Clifford+T as well as the general matrix form for the best approximation that can be synthesized by Clifford+T

Theorem 3.6. *[KMM13] The set of unitary matrices that can be generated by a circuit consisting of gates from the Clifford+T set is exactly the set $D[w]^{2\times 2} \cap U(2)$. Furthermore, given $U \in D[w]^{2\times 2} \cap U(2)$, there is an efficient algorithm to convert it into a circuit consisting of H, T gates only.*

Proof. The only if direction is obvious because all elements of Clifford+T have their entries inside $D[\omega]$ and therefore, any gate synthesized must have entries inside $D[\omega]$ as well. The if direction can be shown with their synthesis algorithm. Here I would just briefly talk about the ideas behind the algorithm. The KMM synthesis algorithm relies on the following observations:

1. We can define a "potential" for each element of $D[w]^{2\times 2} \cap U(2)$, which is the SDE of the top left entry. This makes sense because it is not hard to show that for any element of $D[w]^{2\times 2} \cap U(2)$, the SDEs of all the entries are the same. For the definition of SDE, see definition 3.9.

2. Given $U \in D[w]^{2\times 2} \cap U(2)$ where $sde(U) \geq 4$, there is some $k \in 0, 1, 2, 3$ such that:
$$sde(HT^{-k}U) = sde(U) - 1$$

3. There are finitely many elements of $D[w]^{2\times 2} \cap U(2)$ with SDE no more than 3, and an exhaustive approach shows that all of them can be synthesized with H and T only.

With these observations, it should be pretty clear how the synthesis algorithm works. It basically recursively reduces a given U with SDE greater than 3 to a U' whose SDE is no more than 3, look up for the synthesis formula of U' and then append all the $T^k H$ used in the reduction step to the front. It should also be obvious that this algorithm runs in O(sde(U)). □

Corollary 3.7. *The general form of a unitary that can be generated by Clifford+T set is:*
$$\begin{bmatrix} u & -t^\dagger \omega^k \\ t & u^\dagger \omega^k \end{bmatrix}, \quad u, t \in D[w], \quad \|u\|^2 + \|t\|^2 = 1 \tag{9}$$

Theorem 3.8. *[RS16] Given a z-rotation $R_z(\theta)$, defined as:*
$$R_z(\theta) = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}$$
For small $\varepsilon < |1 - \sqrt{\omega}|$, the best ε-approximation is always given by a unitary of the form:
$$\begin{bmatrix} u & -t^\dagger \\ t & u^\dagger \end{bmatrix}, \quad u, t \in D[w], \quad \|u\|^2 + \|t\|^2 = 1 \tag{10}$$

Proof. This proof is very similar to that of Lemma 3.4. Suppose we have U of form in equation 9 that is within ε distance from $R_z(\theta)$, let $e^{i\phi_1}, e^{i\phi_2}$ be the eigenvalues of $UR_z(\theta)^{-1}$, then we have:
$$\|I - UR_z(\theta)^{-1}\|_2 \leq \varepsilon < |1 - \sqrt{\omega}|$$
On the other hand, we also have:
$$\|I - UR_z(\theta)^{-1}\|_2 = \max(|1 - e^{i\phi_1}|, |1 - e^{i\phi_2}|$$
It follows that both ϕ_1, ϕ_2 are in $(-\pi/8, \pi/8)$, and
$$-\pi/4 < \phi_1 + \phi_2 < \pi/4$$
$$|1 - e^{i(\phi_1+\phi_2)}| < |1 - \omega|$$
Note also we have:
$$\det(UR_z(\theta)^{-1}) = e^{i(\phi_1+\phi_2)} = \omega^k$$
Therefore:
$$|1 - e^{i(\phi_1+\phi_2)}| < |1 - \omega| \implies |1 - \omega^k| < |1 - \omega|$$
$$\implies k = 0$$

□

From this point, unless otherwise specified, U will always refer to a unitary of the form in 10.

3.4 T-count and SDE

Definition 3.9. *For any element $x \in D[w]$, it can be written as:*

$$x = \frac{a + bw + cw^2 + dw^3}{\sqrt{2}^k}, \quad a, b, c, d \in \mathbb{Z} \tag{11}$$

Abstractly, this means:

$$D[w] = \bigcup_{k=0}^{\infty} \frac{1}{\sqrt{2}^k}\mathbb{Z}[w] \tag{12}$$

This representation is not unique in general, it is unique for fixed k.
The smallest denominator exponent (SDE) of x is define as:

$$sde(x) = \min\{k \in \mathbb{N} | \sqrt{2}^k x \in \mathbb{Z}[w]\}$$

Theorem 3.10. *[RS16] Given a unitary U of the form in (10), and let k be the SDE of u, then, for the purpose z-rotation approximation, its T-count is always $2k - 2$.*

Proof. The first part of the proof involves close inspection of figure 2 in [GS19], which shows that for a U of the form in 10, the T-count is always $2k$ or $2k-2$. In case the T-count is $2k$, then $TUT^\dagger$ would always have T-count $2k - 2$. Since a detailed explanation of the reasoning behind involves understanding the normal form of Clifford+T circuits, which lies beyond of the scope of this thesis, I will omit the details.

The second part is simply to note that U and $TUT^\dagger$ have the same distance from $R_z(\theta)$ because z-rotations commute with T:

$$\|R_z(\theta) - U\|_2 = \|TR_z(\theta)T^\dagger - TUT^\dagger\|_2 = \|R_z(\theta)TT^\dagger - TUT^\dagger\|_2 = \|R_z(\theta) - TUT^\dagger\|_2$$

□

4 Ross-Selinger Algorithm Overview

Now we are ready to discuss the algorithm itself. Note that the distance from U to a z rotation only depends on u and θ:

$$\|R_z(\theta) - U\|_2^2 = 2 - 2Re(e^{i\theta/2}u)$$

Thus, considering Theorem 3.10, the circuit-size and precision of an approximation candidate U only depends on u, and we can reformulate our approximation problem as follows:

Given an angle θ and a precision ε, the goal is to efficiently solve the optimization problem:

$$\text{argmin}_{u \in D[w]}\ sde(u)$$

Under the constraints:

$$2 - 2Re(e^{i\theta/2}u) \leq \varepsilon^2 \tag{13}$$

$$\exists t \in D[w], \quad \|t\|^2 + \|u\|^2 = 1 \tag{14}$$

The idea behind Ross and Selinger's algorithm is to first find a set of u candidates that satisfy constraint 13, enumerate these candidates in increasing order of SDE and for each of them, try to solve the norm equation until success. It is clear that this approach is accurate, in the same that it always finds the optimal solution (assuming that the norm equation solver is perfect). We will first consider the search part of the algorithm.

4.1 The search part

The searching is the mostly challenging part of the algorithm and, in fact, is the most significant novelty of Ross-Selinger. Therefore, it deserves some special attention.

4.1.1 Reduction from $D[\omega]$ to $\mathbb{Z}[\omega]$

Since we eventually want to list all the possible candidates in the order of increasing T-count, Theorem [3.10] suggests that it suffice to list the candidates in the order of increasing SDE. And then (3.9) motivates us to break the search problem for $u \in D[\omega]$ into a sequence of search problems for $u \in 2^{-k/2}\mathbb{Z}[\omega]$ in increasing order of $k = 0, 1, 2, \ldots$.

Remarks: Note that the sets $\{2^{-k/2}\mathbb{Z}[\omega]\}_{k=1}^{\infty}$ are not disjoint (e.g. $\frac{1}{\sqrt{2}} = \frac{\omega - \omega^3}{2}$), so if we carry out the search problem in the order $k = 0, 1, 2 \ldots$, some candidates for $k = 0$ will reappear in the problem for $k = 1$ (in fact it is not hard to see that if x has SDE n, then x can be written in the form of (11) for any $k \geq n$). So there is a bit of inefficiency, but this does not affect the optimality of the solution in the sense that, suppose the optimal candidate u has sde k^*, then our search algorithm is guaranteed to find it in the search problem for $2^{-k^*/2}\mathbb{Z}[\omega]$. It is also not hard to address the issue by noting that for an x in the form of (11), its denomenator exponent can be reduced if and only if $b-d, a-c$ are both even, so our algorithm can check this and skip candidates of this form.

Hence, we will just focus on the search problem in the space of $\mathbb{Z}[\omega]$

4.1.2 Search region

Consider the distance constraint in equation (13), it is easy to get:

$$Re(e^{i\theta/2}u) \geq 1 - \frac{\varepsilon^2}{2}$$

Let $z = e^{-i\theta/2}$, and let $\vec{z}, \vec{u}$ be the vectors representing the two complex numbers on the complex plane. Then we can actually represent the constraint as:

$$\vec{u} \cdot \vec{z} \geq 1 - \frac{\varepsilon^2}{2}$$

This can be thought of as the projection of $\vec{u}$ onto $\vec{z}$ is positive and greater than $1 - \varepsilon^2/2$. Denoting the unit ball as $\overline{\mathcal{D}}$, Figure 1, taken from [RS16], shows the search region graphically.

4.1.3 Is the problem well-posed?

So intuitively, we want to find all the elements $u \in \mathbb{Z}[\omega]$ such that $u \in \mathcal{R}_\varepsilon$, but there is a problem. The ring $\mathbb{Z}[\omega]$ is dense in $\mathbb{C}$ (its density follows from the density of $\mathbb{Z}[\sqrt{2}]$ in $\mathbb{R}$), so there are

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

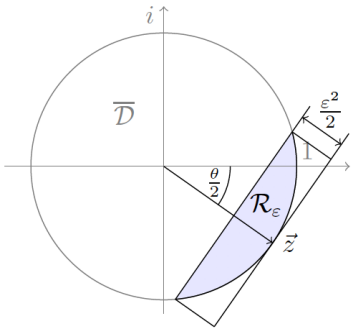

Figure 1: A graphical illustration of the search region in the complex plane

infinitely many elements of $\mathbb{Z}[\omega]$ inside the search region and clearly an exhaustive search would not be feasible. Also, it seems too good to be true that there will be infinitely many candidates just for the case $k = 0$. One is thus tempted to ask if it is possible to impose further constraints on the choice of u such that it reduces the number of candidates u inside $\mathcal{R}_\varepsilon$ to a finite number.

Lemma 4.1. *The automorphisms of $\mathbb{Z}[\omega]$ is exactly the Galois group of the 8-th cyclotomic extension of $\mathbb{Q}$:*

$$Gal(\mathbb{Q}[\omega]/\mathbb{Q}) = \{\sigma_i\}_{i=0}^{3}, \quad \sigma_k(\omega) = \omega^{2k+1}$$
$$\sigma_0 = id, \quad \sigma_1(\omega) = \omega^\circ, \quad \sigma_3(\omega) = \omega^\dagger, \quad \sigma_2(\omega) = (\omega^\dagger)^\circ$$

We will call σ_1 as the square-2 conjugate as it maps $\sqrt{2}$ to $-\sqrt{2}$ and σ_3 is the normal complex conjugate.
And suppose x has form $a + b\omega + c\omega^2 + d\omega^3, a, b, c, d \in \mathbb{Z}$, then:

$$\sigma_1(x) = a + d\omega - c\omega^2 + b\omega^3$$
$$\sigma_2(x) = a - b\omega + c\omega^2 - d\omega^3$$
$$\sigma_3(x) = a - d\omega - c\omega^2 - b\omega^3$$

Lemma 4.2. *Consider the canonical embedding of $\mathbb{Z}[\omega]$ and the definition of the norm:*

$$\rho : \mathbb{Z}[\omega] \to \mathbb{C}^4, \quad \rho(x) = (x, x^\circ, x^\dagger, (x^\circ)^\dagger), \quad norm(x) = xx^\circ x^\dagger (x^\circ)^\dagger$$

For any two elements $x, y \in \mathbb{Z}[\omega]$, the norm of $x - y$ is a non-negative integer:

$$norm(x - y) \in \mathbb{Z}$$

Proof. Note that the norm of any element in $\mathbb{Z}[\omega]$ is invariant under all automorphisms of $\mathbb{Z}[\omega]$:

$$\forall \sigma_i \in G, x \in \mathbb{Z}[\omega] \quad \sigma_i(\text{norm}(x)) = \prod_{k=0}^{3} \sigma_i\sigma_k(x) = \prod_{j=0}^{3} \sigma_j(x) = \text{norm}(x)$$

because the automorphisms form a group and multiplication by a group element simply permutes the group elements. So simply checking observation [4.1], we deduce that $b, c, d = 0$ and $\text{norm}(x) \in \mathbb{Z}$. Finally, notice that:

$$\text{norm}(x) = |x|^2 |x^\circ|^2 \geq 0$$

□

Remarks: A similar result holds for any field extension, that the norm (as well as trace) of any element is in the fixed field of the Galois group, which is equal to the base field when the extension is normal.

Theorem 4.3. *For any distinct $x, y \in \mathbb{Z}[\omega]$, the following inequality holds:*

$$|x - y||x^\circ - y^\circ| \geq 1 \tag{15}$$

Proof. This follows from Lemma [4.2] since:

$$\begin{aligned} x \neq y &\implies \text{norm}(x - y) \geq 1 \\ &\implies |(x - y)|^2 |(x - y)^\circ|^2 \geq 1 \\ &\implies |x - y||x^\circ - y^\circ| \geq 1 \end{aligned}$$

□

The above theorem suggests that it is possible to obtain a discrete grid if we can have a extra constraint that controls how large $|x^\circ - y^\circ|$ can be. If $|x^\circ - y^\circ|$ is bounded above, that is:

$$x^\circ, y^\circ \in B, \quad \text{diam}(B) = \sup_{a,b \in B} \|a - b\| < \infty$$

Then $|x - y|$ is bounded below by $1/\text{diam}(B)$ and we get a discrete grid. Fortunately, there is a very natural B to consider for candidates u.

Lemma 4.4. *Using constraint (14), we have:*

$$uu^\dagger = 1 - tt^\dagger \implies (uu^\dagger)^\circ = 1^\circ - (tt^\dagger)^\circ = |u^\circ|^2 = 1 - |t^\circ|^2$$

$$|u^\circ|^2, |t^\circ|^2 > 0 \implies 0 \leq |u^\circ|^2 \leq 1$$

Thus, we see that the determinant constraint (14) naturally limits u° to the unit disk and therefore, if we require u° to be inside the unit disk, then the candidates $u \in \mathbb{Z}[\omega]$ will form a discrete grid over the complex the plane and within a given convex region $\mathcal{R}_\varepsilon$, there will only be finitely candidates and thus an exhaustive search will be possible.

Definition 4.5. *To summarize, the problem formulation we eventually want to work with is: given an angle θ and a precision ε, the goal is to efficiently solve the optimization problem:*

$$argmin_{u \in D[w]} \; sde(u)$$

Under the constraints:

$$u \in \mathcal{R}_\varepsilon, \quad u^\circ \in \overline{\mathcal{D}} \tag{16}$$

$$\exists t \in D[w], \quad \|t\|^2 + \|u\|^2 = 1 \tag{17}$$

4.2 The norm equation

Note that to satisfy the second constraint in Definition 4.5, one needs a solver that decides if Equation 17 has a solution, and if it does, finds the solution. It is commonly believed that solving such norm equations in cyclotomic rings is at least as hard as factoring [KMM16]. Indeed, in solvers proposed by [RS16],[KMM16] and [HGS04], a factoring oracle is always needed. Fortunately, this is not a big bottle-neck for the Ross-Selinger algorithm itself, as the absence of a factoring oracle only impacts the T-count complexity to a reasonable extent (see Section 9). For a detailed description of the solver to the norm equation over $D[\omega]$, see Section 7.

5 The Grid Operator Approach to the Search Problem

In this section, I will review the grid-operator approach for the searching problem introduced in [RS16]. Note that in this section, many lemmas and theorems will be stated without proofs mainly because the proofs are extremely long and tedious, and there is little point in reproducing such work. It should also give the readers a hint why we wanted an alternative to the grid operator approach.

5.1 1-d grid problems

Definition 5.1. *[RS16] Let $B \subset \mathbb{R}$, the grid of B is the set:*

$$G(B) = \{\alpha \in \mathbb{Z}[\sqrt{2}] | \alpha^\circ \in B\}$$

Definition 5.2. *[RS16] The 1-d grid problem is the follow: given $A, B \subset \mathbb{R}$, find $\alpha \in \mathbb{Z}[\sqrt{2}]$ (we will see why $\mathbb{Z}[\sqrt{2}]$ is a good choice later)*

$$\alpha \in A, \quad \alpha^\circ \in B$$

equivalently, this is to find:

$$A \cap G(B)$$

For the sake of solving problem defined in [5.7], we will only consider closed convex sets in $\mathbb{R}^2$ (because $\mathcal{R}_\varepsilon, \overline{\mathcal{D}}$ are indeed closed convex sets), and that corresponds to closed intervals in $\mathbb{R}$, so we will assume $A = [a_0, a_1], B = [b_0, b_1]$

Theorem 5.3. *[RS16] Let $A = [a_0, a_1], B = [b_0, b_1]$ be closed real intervals, then there is an algorithm that takes $O(|A||B|)$ number of arithmetic operations to enumerate all the solution.*

Proof. Let $\lambda = 1 + \sqrt{2}$, then $(\lambda^{-1})^\circ = -\lambda$, and they are both in $\mathbb{Z}[\sqrt{2}]$. The 1-d grid problem for A, B is equivalent to the grid problem for $\lambda^{-1}A, -\lambda B$ in the sense that:

$$\alpha \in A \iff \lambda^{-1}\alpha \in \lambda^{-1}A$$

$$\alpha^\circ \in B \iff (\lambda^{-1}\alpha)^\circ \in -\lambda B$$

This means α is a solution to grid problem A, B if and only if $\lambda^{-1}\alpha$ is a solution to grid problem $\lambda^{-1}A, -\lambda B$.

This means that if $\lambda^k \leq |A| < \lambda^{k+1}$, we can re-scale it to a problem with $\lambda^{-1} \leq |A| < 1$ and we just need to multiply the solution by λ^{k+1} to get back solution to the original problem.

Thus, we assume $\lambda^{-1} \leq |A| < 1$, and then consider any solution $x \in \mathbb{Z}[\sqrt{2}]$

$$x = a + b\sqrt{2}, \quad x \in [a_0, a_1], \quad x^\circ \in [b_0, b_1]$$

We can deduce the following relations:

1. For each $b \in \mathbb{Z}$, there is at most one $a \in \mathbb{Z}$ that leads to a solution x:

$$a_0 \leq a + b\sqrt{2} \leq a_1 \implies a \in [a_0 - b\sqrt{2}, a_1 - b\sqrt{2}]$$

since we assume $|A| = a_1 - a_0 < 1$,this means for a fixed $b \in \mathbb{Z}$, $|[a_0 - b\sqrt{2}, a_1 - b\sqrt{2}]| < 1$.

2. $b \in [(a_0 - b_1)/\sqrt{2}^3, (a_1 - b_0)/\sqrt{2}^3]$

$$b = (x - x^\circ)/\sqrt{2}^3$$

$$\implies \max b = (\max x - \min x^\circ)/\sqrt{2}^3$$

And the result follows

□

So the algorithm is quite straight-forward, given intervals $[a_0, a_1], [b_0, b_1]$, first construct the interval in which b lies and enumerate all the integer points inside that interval. Then, for each b, compute the unique a (if any) and finally check if $a + b\sqrt{2}$ is a solution. The time complexity is dominated by the number of possible b candidates so it is bounded by $O(b_1 - b_0) = O(|A||B|)$ (we need to take re-scaling into account).

Remarks: One might ask if the algorithm is efficient, in the sense that whether there will be as many as $\Omega(|A||B|)$ solutions so that we will not have to wait too long to get a solution in case $|A|, |B|$ are large. [Sel15] showed that this is true, but the whole proof is a bit tedious so here is an improved version of his argument:
we can view $\mathbb{Z}[\sqrt{2}]$ as a copy of the 2-d lattice $\mathbb{Z} \times \sqrt{2}\mathbb{Z}$ embedded in the 2-d plane, where the first coordinate indicates the value of a and the second indicates the value of $b\sqrt{2}$. Then, each real number x can be represented by a line in the plane that connects $(x, 0)$ and $(0, x)$ on which every point represents one way of writing $x = a + b\sqrt{2}$ with $a, b \in \mathbb{R}$ and if a lattice point is on the line, then $x \in \mathbb{Z}[\sqrt{2}]$. Therefore, we can represent the constraint $a + b\sqrt{2} \in [x, x + \delta]$ as the region between the two lines representing $x, x + \delta$. Similarly, the line connecting $(y, 0)$ and $(0, -y)$ represents all the possible ways of writing $y = a - b\sqrt{2}$ and the constraint $(a + b\sqrt{2})^\circ$ is equivalent to the region between the two lines representing $y, y + \Delta$. This means the constraint $u \in [x, x + \delta], u^\circ \in [y, y + \Delta]$ is simply a rectangle tilted at 45 degrees, see Figure 2, taken from [Sel16], for a graphical illustration.

The question to ask is for what values of δ, Δ, we can be sure that the rectangle formed by the four lines contains at least one lattice point for any $x, y \in R$ and Selinger used rather tedious (but elementary) tools to show this holds for $\Delta\delta \geq (1+\sqrt{2})^2$, but actually we can just apply Minkowski's

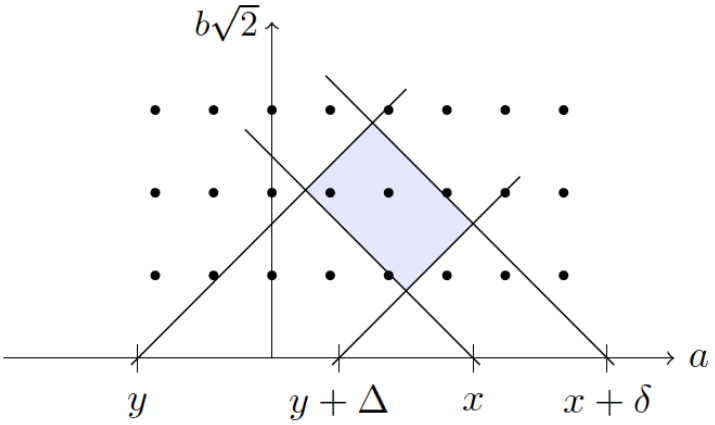

Figure 2: A graphical illustration of the 1-d grid search

theorem to conclude any convex set of area $2^2\sqrt{2}$ would contain a lattice point in $\mathbb{Z} \times \sqrt{2}\mathbb{Z}$. As the area of the rectangle is obtained by $\Delta\delta/2$, this gives a rougher bound $\Delta\delta \geq 8\sqrt{2}$. But this is totally fine to show that there are $\Omega(|A||B|)$ many solutions because given any A, B, simply split the problem into:

$$A_i = [a_0 + i\sqrt{8\sqrt{2}}, a_0 + (i+1)\sqrt{8\sqrt{2}}], B_j = [b_0 + j\sqrt{8\sqrt{2}}, b_0 + (j+1)\sqrt{8\sqrt{2}}]$$

Each problem A_i, B_j is guaranteed to have one distinct solution so in total we get $\Omega(|A||B|)$ many solutions. And this actually implies the algorithm takes $O(1)$ time to enumerate one solution.

5.2 2-d grid problems

We have spent quite some effort in solving the 1 dimensional case because as we will see, the ultimate goal of the 2-d algorithm is to reduce the problem into 2 1-d problems. For the rest of the discussions, we will consider the 2-d plane as the complex plane or $\mathbb{R}^2$ interchangeably and the meaning should be clear from the context.

Definition 5.4. *[RS16] Let $B \subset \mathbb{R}^2$ be a closed convex set, then the grid for B is defined as:*

$$G(B) = \{u \in \mathbb{Z}[w] | u^\circ \in B\}$$

And as in the 1-d case we have the definition:

Definition 5.5. *[RS16] The 2-d grid problem is: given $A, B \subset \mathbb{R}^2$ are two closed convex sets, find $\alpha \in \mathbb{Z}[w]$ such that*

$$\alpha \in A, \quad \alpha^\circ \in B$$

equivalently, this is to find:

$$A \cap G(B)$$

Lemma 5.6. *[RS16] The scaled 2-d grid search problem, given $A, B \in \mathbb{R}^2, k \geq 0$ is to find $u \in \frac{1}{\sqrt{2}^k}\mathbb{Z}[\omega]$ such that:*

$$u \in A, \quad u^\circ \in B$$

And it is computationally equivalent to the un-scaled 2-d grid problem for $\sqrt{2}^k A, (-\sqrt{2})^k B$

Definition 5.7. *[RS16] The 2-d grid problem in the context of Ross-Selinger is to enumerate all the $u \in \mathbb{Z}[\omega]$, in increasing order of SDE, such that:*

$$u \in \mathcal{R}_\varepsilon, \quad u^\circ \in \overline{\mathcal{D}}$$

Alternatively, this means to enumerate, in increasing order of SDE, all the elements of the set $\mathcal{R}_\varepsilon \cap G(\overline{\mathcal{D}})$, where:

$$G(\overline{\mathcal{D}}) = \{u \in \mathbb{Z}[\omega] \,|\, u^\circ \in \overline{\mathcal{D}}\}$$
$$|G(\overline{\mathcal{D}})| = \inf_{a,b \in G(\overline{\mathcal{D}})} |a - b| \geq \frac{1}{2}$$

So, we just need en efficient algorithm for the general 2-d grid search problem. Before, we get to the complicated case, it is helpful to consider simple cases.

5.3 Upright rectangles

Definition 5.8. *A convex set A is called an upright rectangle if it is the direct product of closed intervals $[ax_0, ax_1] \times [ay_0, ay_1]$.*

If both A, B in the 2-d grid problem are upright rectangle, then it is easy to reduce it to two 1-d grid problems.

Lemma 5.9. *[RS16] Let $u \in \mathbb{Z}[\omega]$, then we have:*

$$u = \alpha + \beta i \vee \alpha + \beta i + \omega, \quad \alpha, \beta \in \mathbb{Z}[\sqrt{2}]$$

Proof. This follows directly from writing $u = a + b\omega + c\omega^2 + d\omega^3$ explicitly in terms of a, b, c, d only. □

Theorem 5.10. *[RS16] The 2-d grid problem with A, B being upright rectangles can be solved efficiently with an algorithm that enumerates each solution in $O(1)$ time.*

Proof. Let $A = A_x \times A_y, B = B_x \times B_y$, by Lemma [5.9], there are two possible forms for a potential solution u:

1. $u = \alpha + \beta i$, then:
$$u \in A \implies \alpha \in A_x, \quad \beta \in A_y$$
$$u^\circ \in B \implies \alpha^\circ \in B_x, \quad \beta^\circ \in B_y$$

2. $u = \alpha + \beta i + \omega$, then:
$$u \in A \implies \alpha + \frac{1}{\sqrt{2}} \in A_x, \quad \beta + \frac{1}{\sqrt{2}} \in A_y$$
$$u^\circ \in B \implies \alpha^\circ - \frac{1}{\sqrt{2}} \in B_x, \quad \beta^\circ - \frac{1}{\sqrt{2}} \in B_y$$

□

So the 2-d problem reduces to solve the 1-d problem for (A_x, B_x), (A_y, B_y) or $(A_x - 1/\sqrt{2}, B_x + 1/\sqrt{2})$, $(A_y - 1/\sqrt{2}, B_y + 1/\sqrt{2})$, and by Theorem [5.3], enumeration of each solution takes constant time.

5.4 Smallest Bounding Box of a convex set

The most natural extension of the upright rectangle case to more general sets is to consider the smallest bounding box of the set:

Definition 5.11. *[RS16] The smallest bounding box of a set A, denoted as $BBox(A)$, is the smallest upright rectangle that contains A, and we defined the uprightness of a convex set A to be:*

$$up(A) = \frac{area(A)}{area(BBox(A))}$$

And we say A is M-upright if $up(A) \geq M$.

Lemma 5.12. *[RS16] Each solutions to the 2-d grid problem for two M-upright sets can be enumerated in $O(1/M^2)$ time. And if M is fixed, then it is constant.*

Ross and Selinger did not provide a rigorous proof, but this proposition is reasonable enough for people to believe it with some intuition. First, we solve the problem for $BBox(A), BBox(B)$, and consider the case where the solution is of the form $\alpha + \beta i, \alpha, \beta \in \mathbb{Z}[\sqrt{2}]$. Then, our solutions can be put in a grid form:

$$u = \{\alpha_1, \alpha_2, \dots\} \times \{\beta_1, \beta_2, \dots\}$$

where α_i, β_j are the solutions to the 1-d problems $(BBox(A)_x, BBox(B)_x)$, $(BBox(A)_y, BBox(B)_y)$ respectively arranged in increasing order.

Next, we simply need to check if $\alpha + \beta i$ is in A. Since the area of A is about M of the area of $BBox(A)$, and the fact that u can be arranged in a grid form means they distribute relatively evenly in $BBox(A)$ (note we still have the lower bound $|\alpha - \alpha'|$ determined by the diameter of $BBox(B)$), it makes sense to say that when A, B are large enough to contain a large number of solutions, then the number of u inside A is about M of the solutions to $BBox(A), BBox(B)$. And of course we need to apply the same argument to B so the total number of solutions that satisfy $u \in A, u^\circ \in B$ is about M^2 of the solutions to the problem for $BBox(A), BBox(B)$.

Now let us take a look back at the regions $A = \mathcal{R}_\varepsilon, B = \overline{\mathcal{D}}$, we see that $up(B)$ is a constant, so M is solely determined by ε. As a rough bound, we can choose the bounding box to be the unit box and by observing $\mathcal{R}_\varepsilon$ contains a circle of radius $\varepsilon^2/4$, we can get a lower bound $up(A) = \Omega(\varepsilon^4)$. So if we are happy with an algorithm that enumerates each solution in $O(1/\varepsilon^4)$ time, we can simply stop here. But this is not quite what we want, as the goal is to have an algorithm that runs in $O(\log(1/\varepsilon))$, so we need to do a bit more work.

5.5 Grid operators

Definition 5.13. *[RS16] A grid operator is a linear operator $L : \mathbb{R}^2 \to \mathbb{R}^2$ such that:*

$$L(\mathbb{Z}[\omega]) \subseteq \mathbb{Z}[\omega]$$

and a grid operator is called special if it has determinant ± 1

Lemma 5.14. *[RS16] Using the standard basis, a matrix $L \in \mathbb{R}^{2\times 2}$ is a grid operator if and only it is of the form:*

$$L = \begin{bmatrix} a + a'/\sqrt{2} & b + b'/\sqrt{2} \\ c + c'/\sqrt{2} & d + d'/\sqrt{2} \end{bmatrix}$$

where $a + b + c + d \equiv 0 \mod 2$, $a' \equiv b' \equiv c' \equiv d' \mod 2$

Proof. The proof for this lemma is straight-forward computation so I will just mention the idea. The only if direction can be shown by expressing:

$$u = \begin{bmatrix} x_1 + x_2/\sqrt{2} \\ y_1 + y_2/\sqrt{2} \end{bmatrix}, \quad x_1, x_2, y_1, y_2 \in \mathbb{Z}, \quad x_2 \equiv y_2 \mod 2$$

The if direction simply comes from applying L to standard basis and Hadamard basis vectors.

□

Lemma 5.15. *[RS16] The set of special grid operators form a multiplicative group and is closed under square-2 conjugation, which is defined entry-wise. Further we have $(Lu)^\circ = L^\circ u^\circ$*

Theorem 5.16. *[RS16] Let L be a special grid operator and $A, B \subset \mathbb{R}^2$, then $u \in \mathbb{Z}[\omega]$ is a solution to the 2-d grid problem for (A, B) if and only if Lu is a solution to the 2-d grid problem for $L(A), G^\circ(B)$.*
This means that the 2-d grid problem is computationally invariant under L-transformation.

Proof.

$$u \in A \iff Lu \in L(A), \quad u^\circ \in B \iff L^\circ u^\circ = (Lu)^\circ \in L^\circ(B)$$

□

5.6 Ellipses and the enclosing ellipse of a bounded convex set

The power of grid operator is that for certain convex sets with low uprightness, there exists a grid operator that transforms them into convex sets with high uprightness and we can use this to improve the efficiency of our algorithm.

Definition 5.17. *[RS16] Let D be a positive-definite 2-by-2 matrix and p a point in $\mathbb{R}^2$, the ellipse defined by (D, p) is the set:*

$$E(D, p) = \{u \in \mathbb{R}^2 \mid (u - p)^\dagger D(u - p) \leq 1$$

Next, we have the most important theorem of the grid-search problem

Theorem 5.18. *[RS16] Suppose $A, B \subseteq \mathbb{R}^2$ are M-upright ellipses, then there exists a grid operator L such that $L(A), L^\circ(B)$ are $1/6$-upright, and L can be found in $O(\log(1/M))$ time.*

Lemma 5.19. *[RS16] For any convex set $A \in \mathbb{R}^2$ with non-empty interior, there exists an ellipse E such that:*

$$area(E) \leq \frac{4\pi}{3\sqrt{3}} area(A)$$

The upper bound is sharp and is attained when A is a equilateral triangle and E is the circle inscribing A.

5.7 General 2-d grid search outline

It should be clear from previous discussion how the algorithm works, given closed convex sets $A, B \in \mathbb{R}^2$ represented by their enclosing rational polygon (polygon with rational vertices) such that the area of the polygon is bounded by a fixed constant times the area of A, B, our algorithm does the following:

1. Find ellipses A', B' enclosing A, B such that $\text{area}(A') = O(\text{area}(A))$ and same for B'. The existence is ensure by proposition [5.19] and there is an efficient algorithm (a.k.a. smallest enclosing ellipse algorithm) to compute these ellipses (some software even have it in their own library) provided that we have the enclosing rational polygon of A, B mentioned above. We can assume that this step takes constant time as the run-time is independent of ε.

2. Use Theorem [5.18] to find a grid operator L that transforms A', B' into 1/6-upright sets, and it follows that $L(A)$ and $L^\circ(B)$ are N/6-upright for some constant N and by proposition [5.12], it takes constant time to enumerate each solution for 2-d grid problem of $L(A), L^\circ(B)$.

So overall the improved grid-search algorithm takes $O(\log(1/M))$ plus constant time for each solution enumerated.

6 The Lattice Approach to the Search Problem

As I and my mentor went through the problem set-up, we realized that there is a much simpler way to do the searching, which is through lattice region search. Selinger also confirmed with us via email that the lattice-based approach works but they did not realize this possibility while writing the paper. He himself had actually mentioned about this possibility later in a public lecture [Sel16].

6.1 Embedding $\mathbb{Z}[\omega]$ in $\mathbb{R}^4$

Recall that the goal of the candidate search problem is to find all the elements $u \in \mathbb{C}$ of $\frac{1}{\sqrt{2}^k}\mathbb{Z}[\omega]$ such that $u \in R_\varepsilon$ and $u^\circ \in \overline{D}$ in the complex plane. It suffices to consider the case where $k = 0$ because other cases simply involves re-scaling the lattice basis and there is nothing different.

We have previously shown that for any $u \in \mathbb{Z}[\omega]$, it can be written as:

$$u = a + b\omega + c\omega^2 + d\omega^3, \quad a, b, c, d \in \mathbb{Z}$$

To apply lattice-based search algorithms, one would naturally consider embedding $\mathbb{Z}[\omega]$ as a free $\mathbb{Z}$-module in the corresponding $\mathbb{R}$-vector space spanned by the same basis. While there are many ways to choose a basis-space pair, it is important to make sure the search constraints correspond to

a compact convex set in the coordinate space. Note that there are two complex constraints, one on u and another on u°, so the most natural way is to choose a basis from $\mathbb{C}^2$, which we will treat as $\mathbb{R}^4$.

Specifically, we will apply a partial canonical embedding $\sigma(u) = (u, u^\circ)$ (we do not need the full embedding as the other two conjugates are simply the standard complex conjugates of the first and constraints are on them are trivial and unnecessary) on the cyclotomic extension basis $\{1, \omega, \omega^2, \omega^3\}$ which maps:

$$\begin{aligned}\forall u \in \mathbb{Z}[\omega], \quad \sigma(u) &= \sigma(a + b\omega + c\omega^2 + d\omega^3) \\ &= a\sigma(1) + b\sigma(\omega) + c\sigma(\omega^2) + d\sigma(\omega^3) \\ &= a(1,1) + b(\omega, \omega^3) + c(\omega^2, -\omega^2) + d(\omega^3, \omega^1)\end{aligned}$$

Therefore, the basis for the embedding $\mathbb{Z}[\omega]$ in $\mathbb{C}^2$ is exactly:

$$\{(1,1), (\omega, \omega^3), (\omega^2, -\omega^2), (\omega^3, \omega)\}$$

It is however hard to directly work with $\mathbb{C}^2$, or to visualize the constraints as a convex body. Fortunately, things become simple linear algebra if we instead consider $\mathbb{C}^2$ as $\mathbb{R}^4$ (the usual way of splitting real and imaginary parts), in $\mathbb{R}^4$, the basis written in the form of a 4 by 4 matrix is:

$$B = \begin{bmatrix} 1 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ 1 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -1 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Therefore, given any $u \in \mathbb{Z}[\omega]$ specified by the coordinates $\vec{u} = (a, b, c, d) \in \mathbb{Z}^4$, we have:

$$B\vec{u} = \vec{v}, \quad \vec{v} = (v_1, v_2, v_3, v_4) = (Re(u), Im(u), Re(u^\circ), Im(u^\circ))$$

Note that this gives us a very convenient way to specify the constraints as quadratic and linear inequalities:

$$u \in R_\varepsilon \iff v_1^2 + v_2^2 \le 1, \quad (v_1, v_2) \cdot \vec{z} \ge 1 - \varepsilon^2/2, \quad \vec{z} = e^{-\theta/2}$$
$$u^\circ \in D \iff v_3^2 + v_4^2 \le 1$$

Thus, we have reduced the search problem into finding all the integral vectors in the coordinate space ($\mathbb{R}^4$) that are inside the set $\mathcal{K}$:

$$\mathcal{K} = \{\vec{x} \in \mathbb{R}^4 \mid \vec{y} = B\vec{x}, \quad y_1^2 + y_2^2, y_3^2 + y_4^2 \le 1, \quad (y_1, y_2) \cdot \vec{z} \ge 1 - \varepsilon^2/2\}$$

It it easy to show $\mathcal{K}$ is a convex body:

Lemma 6.1. *$\mathcal{K}$ is closed, bounded and convex.*

Proof. Note that invertible linear operators in $\mathbb{R}^n$ preserve closedness, boundedness and convexity. It should be easy to check from the definition that $B(\mathcal{K})$ is closed, bounded and convex and B is invertible. Therefore $B^{-1}(B(\mathcal{K})) = \mathcal{K}$ is also closed bounded and convex. □

Therefore, our problem has been reduced to finding integral lattice points in a convex body in $\mathbb{R}^4$, for which there is an efficient algorithm to find one solution when the dimension of the lattice space is fixed. The algorithm is efficient in the sense that, for fixed dimension, its run-time is a polynomial function only dependent on the bit-length of the input. For the sake of being self-contained, below are the details of the algorithm.

Definition 6.2. *To specify the convex set $\mathcal{K}$, we use a membership oracle that can solve the separation problem.This oracle takes an $x \in V$ and outputs whether $x \in \mathcal{K}$. In the case that $x \notin \mathcal{K}$, the oracle also finds us a hyperplane that separates x and $\mathcal{K}$. For the rest of the discussions, a membership oracle is assumed to be able to solve the separation problem.*

Lemma 6.3. *A membership oracle for $\mathcal{K}$ can be constructed efficiently.*

Proof. Given $\vec{y} \in \mathbb{R}^n$, it is straight-forward to test membership. So suppose $B\vec{y} \notin B(\mathcal{K})$, and we want to solve the separation problem. The idea is to first find the hyperplane that separates $B\vec{y}$ from $B(\mathcal{K})$. Let $\vec{u} = B\vec{y}$, since we already know that y fails to be inside $\mathcal{K}$, there are two cases:

1. u is outside R_ε.
2. u° is outside D.

In the first case, we again compute the projection of $\vec{u}$ onto $\vec{z}$. Let u_z be the scalar projection, if it is smaller than $1 - \varepsilon^2/2$. Take $\nu = [1 - \varepsilon^2/2 + u_z]/2$, and the line perpendicular to $\vec{z}$ that intersects z at νz is a line that separates u from R_ε. If u_z is larger or equal to $1 - \varepsilon^2 2$, then we can deduce that u is outside the unit circle, so the tangent line that cuts the unit circle at $\vec{u}/\|\vec{u}\|$ will separate R_ε from u.

In the second case, again we know u° is outside the unit circle, then the tangent line that cuts the unit circle at $u^\circ/\|u^\circ\|$ separates D and u°.

The above procedure involves minimal computations as most of the calculations can be done analytically with basic 2D Euclidean geometry. Therefore suppose we have found a hyperplane , specified be $\vec{c} \in \mathbb{R}^n$ and $D \in \mathbb{R}$ such that:

$$c^T B y > D$$

$$\forall x \in \mathcal{K}, \quad c^T B x < D$$

Next, we simply define $c' = (c^T B)^T$, and we get a hyperplane specified by $c' \in \mathbb{R}^n$ and $D \in \mathbb{R}$ such that:

$$c'^T y > D$$

$$\forall x \in \mathcal{K}, \quad c'^T x < D$$

□

6.2 LLL-based Integer Programming in a Fixed Dimension

Definition 6.4. *[Lov86] For a given convex set $\mathcal{K} \subseteq \mathbb{R}^n$, a pair of ellipsoids (E', E) such that $E' \subseteq \mathcal{K} \subseteq E$, E', E are concentric and E' is generated by shrinking E by a factor of $1/n$ is called a Lówner-John pair for $\mathcal{K}$.*
If the factor is relaxed to be $1/(n+1)\sqrt{n}$, then such a pair is called a weak Löwner -John pair for $\mathcal{K}$.

Theorem 6.5. *[Lov86]* ***Löwner and John*** *For any convex set $\mathcal{K} \subseteq \mathbb{R}^n$, there exists a Löwner-John pair for it.*

Theorem 6.6. *[Lov86]* ***Lovász****: For any convex set $\mathcal{K} \in \mathbb{R}^n$ specified by a membership oracle, a weak Lówner-John pair can be computed in polynomial time.*

Proof. Use the shallow cut ellipsoid method, we can construct a sequence of ellipsoids $E_0, E_1, \ldots, E_k$ such that $\mathcal{K}$ is contained in E_k for all k and we have the shrinking factor:

$$\mathrm{vol}(E_{k+1}) < \exp\left(-\frac{3}{2(n+1)(2n+1)^2}\right)\mathrm{vol}(E_k)$$

The claim is that suppose the sequence terminates at E_k, then E_k determines a weak Lówner-John pair (E'_k, E_k). The proof is by contradiction, if E_k fails to define a pair, then we can construct E_{k+1}. This construction uses the shallow-cut ellipsoid method as well. (For details of the shallow-cut ellipsoid method, refer to Lovász's text) □

Now we are ready to prove the major theorem that is relevant to solving the search problem:

Theorem 6.7. *[Lov86] Given a convex body $\mathcal{K} \in \mathbb{R}^n$ specified by a membership oracle, we can achieve in polynomial time one of the following two (but not simultaneously):*

1. *find an integral vector $x \in \mathbb{Z}^n$ that is in $\mathcal{K}$.*
2. *find an integral vector $c \in \mathbb{Z}^n$ such that:*

$$\max_{x \in \mathcal{K}} x \cdot c - \min_{x \in \mathcal{K}} x \cdot c \leq 2n^2 9^n$$

Proof. First, we find a weak Löwner-John pair (E', E) for $\mathcal{K}$, let o be their common center. Next, let T be the linear transformation on $\mathbb{R}^n$ that maps E', E onto two concentric balls with radius $1, (n+1)\sqrt{n}$ respectively. Let $y = T(o)$ be the common center. And let $T\mathbb{Z}^n$ be the image of the lattice under the linear transformation. (The construction of T can be found in detail in [Len83]).

Apply LLL to $T\mathbb{Z}^n$ (for the LLL step, the lattice basis of $T\mathbb{Z}^n$ is given by matrix form of T with standard basis). Let $b_1 \ldots b_n$ be the reduced basis. Then find a lattice point $x \in T\mathbb{Z}^n$ such that:

$$\|x - y\| \leq n\|b_k\|, \quad b_k = \mathrm{argmax}_{1 \leq i \leq n}\|b_i\|$$

We can quickly check if $T^{-1}(x) \in \mathcal{K}$, and if it is not, since TE' is contained in $T\mathcal{K}$, this implies $x \notin TE'$ and thus:

$$1 < \|x - y\| \leq n\|b_k\|, \quad \|b_k\| \geq 1/n$$

Partition $T\mathbb{Z}^n$ into hyperplanes of $n-1$ dimension by taking out the basis b_k:

$$T\mathbb{Z}^n = \bigcup_{i\in\mathbb{Z}} \mathcal{L}(\{b_i\}_{i=1}^n \setminus b_k) + ib_k$$

Let L' be the sublattice with $i = 0$, using the property of LLL reduced basis, we have:

$$d(b_k, L') \geq f(n)\|b_k\| \geq g(n)$$

where f, g are functions of n. This gives a lower bound on the distance between successive hyperplanes and therefore the number of i such that $L' + ib_k$ intersects with TE' is given by:

$$(n+1)\sqrt{n}\frac{1}{d(b_k, L')} = h(n)$$

Thus, the number of recursion cases solely depends on the dimension and since n is fixed, this is essentially a constant and we could repeat this process for the hyperplanes.

□

Theorem 6.8. *Using the algorithm mentioned in Theorem 6.7 as an oracle, there is an efficient algorithm that enumerates all the solutions of the grid search problem for fixed k. Moreover, it can be efficiently extended to the case of increasing k in the sense that it takes on average constant time to enumerate each solution for any k.*

Proof. The idea is simple: split the search region into sub-regions in which there can be at most one candidate, and apply integer programming to each of the sub-region. First, one can bound $\mathcal{K}$ within the bounding box $\mathcal{B}$ of a ball with radius $\|B^{-1}\|_2$. This is because, suppose $\vec{x} \in \mathcal{K}$, then from the definition of $\mathcal{K}$ we have:

$$\|B\vec{x}\| = \|\vec{y}\| \leq 1$$

$$\|\vec{x}\| = \|B^{-1}\vec{y}\| \leq \frac{\|B^{-1}\vec{y}\|}{\|\vec{y}\|} \leq \sup_{\vec{y}\in\mathbb{R}^4} \frac{\|B^{-1}\vec{y}\|}{\|\vec{y}\|} = \|B^{-1}\|_2$$

Next, one can evenly split $\mathcal{B}$ into small boxes such that each small box contains at most one lattice point. This can be achieved by choosing small boxes of length $< 1/2$ for the case $k = 0$. Note for the case $k = 0$, it does not matter how many small boxes there are compared to the number of solutions. What matters is how fast they grow with respect to k asymptotically. Since the dimension is 4, when k increases, the number of small boxes will increase in $O(\sqrt{2}^{4k}) = O(4^k)$. Fortunately, as noted in remark 5.25 of [RS16], the number of solutions also grow as $O(4^k)$. Therefore, the average time for enumerating one solution is always $O(1)$.

□

7 Solving the Norm Equation

The goal is to solve the equation:

$$t^\dagger t = \xi = 1 - u^\dagger u$$
$$t, u \in D[w], \quad \xi \in D[\sqrt{2}]$$

One easy to check necessary condition is that:

$$\xi = \|t\| > 0, \quad \xi^\circ = \|t^\circ\| > 0$$
$$(a + b\sqrt{2})^\circ = a - b\sqrt{2}$$

We call such ξ, doubly positive. The following lemma allows us to consider a relaxation of the problem:

Lemma 7.1. *$t^\dagger t = \xi$ has a solution if and only if $t^\dagger t \sim \xi$ has a solution and ξ is doubly positive, $\sim$ denotes differ by a unit in $\mathbb{Z}[\sqrt{2}]$.*

Proof. The only if direction is obvious. For the if direction, suppose $t^\dagger t u = \xi$, then since both t^t and ξ are doubly positive, u must be doubly positive. A unit in $\mathbb{Z}[\sqrt{2}]$ is doubly positive if and only if it is a square because all units are of the form $(-1)^n(1+\sqrt{2})^m$, and first positive condition forces n to be even and second positive condition forces $n+m$ to be even. Note that v is always real, so:

$$t^\dagger t v^2 = (tv)^\dagger tv = \xi$$

□

Corollary 7.2. *Once we have checked that ξ is doubly positive (otherwise the algorithm tries the next ξ candidate), it suffices to find:*

$$t^\dagger t \sim \xi$$

If a solution exists, we call ξ decomposable.

The intuition is that things become easier when we have a prime factorization of ξ since all the rings we are working with are Euclidean domains. But to do this, we first consider a special case.

7.1 The case $\xi \in \mathbb{Z}[\sqrt{2}]$

Lemma 7.3. *[RS16] Given $\xi = \alpha\beta$, where $\alpha, \beta \in \mathbb{Z}[\sqrt{2}]$, $\gcd(\alpha, \beta) \sim 1$, then ξ is decomposable if and only if both α, β are decomposable.*

Proof. The if direction is obvious. For the only if direction, suppose $\xi = t^\dagger t$, then $t \in \mathbb{Z}[w]$. This is because in $\mathbb{Z}[w]$, $\sqrt{2}^k | t$ if and only if $2^k | t^\dagger t$, we can choose $t' = t\sqrt{2}^k \in \mathbb{Z}[w]$ and we have:

$$t'^\dagger t' = 2^k \xi \implies 2^k | t'^\dagger t' \implies \sqrt{2}^k | t' \implies t \in \mathbb{Z}[w]$$

Another more straight-forward way to see this is to write out t explicitly in terms of a, b, c, d and conclude that they must be integers if the norm is in $\mathbb{Z}[\sqrt{2}]$.
Let $s = \gcd(t, \alpha)$ in $\mathbb{Z}[\sqrt{2}]$, then we easily have

$$s^\dagger s | \alpha^2, \quad s^\dagger s | t^\dagger t \sim \alpha\beta, \quad \gcd(\alpha, \beta) \sim 1 \implies s^\dagger s | \alpha$$

On the other hand, we have:

$$\alpha = \gcd(t^\dagger t, \alpha), \quad \gcd(t^\dagger t, \alpha) | \gcd(t, \alpha) \cdot \gcd(t^\dagger, \alpha) = s^\dagger s \implies \alpha | s^\dagger s$$

Thus $\alpha \sim s^\dagger s$, and same argument for β. □

Corollary 7.4. *[RS16] Given a prime factorization of $\xi = \xi_1^{m_1} \dots \xi_k^{m_k}$ in $\mathbb{Z}[\sqrt{2}]$, then ξ is decomposable if and only if each $\xi_i^{m_i}$ is decomposable. And we can easily find solutions for ξ once we have the solutions for each $\xi_i^{m_i}$*

Thus, only two questions remain, how to find a prime factorization of ξ, and for each prime power factor, how to decompose them.

7.2 Decomposability of $\xi_i^{m_i}$

First, we note some special property of primes in $\mathbb{Z}[\sqrt{2}]$

Lemma 7.5. *[RS16] For $\xi \in \mathbb{Z}[\sqrt{2}]$ a prime element, there is one and only one integer prime p such that $\xi | p$ in $\mathbb{Z}[w]$, and consequently the prime factorization of this particular p in $\mathbb{Z}[\sqrt{2}]$ is either ξ or $\xi^\circ \xi$*

Proof. Existence is easy by considering prime factorization of the norm of ξ. Uniqueness follows from the co-primality of integers primes. That is , suppose there are 2 distinct prime integers p, q such that:

$$p = a\xi, q = b\xi, \quad a, b \in \mathbb{Z}[\sqrt{2}]$$
$$p^\circ p = a^\circ a \xi^\circ \xi = p^2, \quad a^\circ a, \xi^\circ \xi \in \mathbb{Z}$$
$$q^\circ q = b^\circ b \xi^\circ \xi = q^2, \quad b^\circ b, \xi^\circ \xi \in \mathbb{Z}$$

Clearly, we see $\xi^\circ \xi$ is an integer and is a common divisor of p^2, q^2, so it has to be 1 and consequently ξ is a unit, not a prime.
The second part comes from considering $\xi^\circ \xi | p^2$ in $\mathbb{Z}$. There are only three possibilities:

$$A : \xi^\circ \xi \sim 1, \quad B : \xi^\circ \xi \sim p, \quad C : \xi^\circ \xi \sim p^2$$

A is impossible, B is what we claim, for C, since $\xi | p$, there must be some $u \in \mathbb{Z}[\sqrt{2}]$ (we know u must be real and all real elements of $\mathbb{Z}[w]$ are in $\mathbb{Z}[\sqrt{2}]$) such that $\xi u = p$, then:

$$\xi^\circ \xi u^\circ u = p^2, \quad \xi^\circ \xi \sim p^2 \implies u \text{ is a unit}, \quad \xi \sim p$$

□

Lemma 7.6. *(**Quadratic Reciprocity**) Let the legendre symbol be defined as:*

$$\left(\frac{q}{p}\right) = \begin{cases} +1 & \text{if } q \text{ is a square modulo } p \\ -1 & \text{else} \end{cases}$$

Then the following is true:

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{(p-1)(q-1)}{4}}$$

Lemma 7.7. *[RS16] Let p be a prime in $\mathbb{Z}$, then its prime factorization in $\mathbb{Z}[\sqrt{2}]$ is either p or $\alpha^\circ\alpha$, and can be computed efficiently.*

Proof. There are 3 cases:

1. If $p = 2$, then $p = \sqrt{2}^2$, this is a trivial case.
2. If $p \equiv 3, 5 \mod 8$, then p is also a prime in $\mathbb{Z}[\sqrt{2}]$. This is because, suppose α is a prime factor of p in $\mathbb{Z}[\sqrt{2}]$, then by lemma 7.5, we have either $\alpha \sim p$ or $\alpha^\circ\alpha \sim p$. The first case if what we claim, for the second case, let $\alpha = a + b\sqrt{2}$, then:
$$\xi^\circ\xi = a^2 - 2b^2 = \pm p \equiv \pm 3 \mod 8$$
But this is impossible as squares are $0, 1, 4$ modulo 8.
3. If $p \equiv 1, 7 \mod 8$, then $p = \alpha^\circ\alpha$ in $\mathbb{Z}[\sqrt{2}]$. By quadratic reciprocity, 2 is a square in $\mathbb{F}_p$, and let x be the solution to $x^2 - 2 \equiv 0 \mod p$ and let $\alpha = \gcd(p, x + \sqrt{2})$.
Note that $p|(x + \sqrt{2})^\circ(x + \sqrt{2})$ but p does not divide $x + \sqrt{2}$ or $x - \sqrt{2}$ so p is not a prime element in $\mathbb{Z}[\sqrt{2}]$. And therefore by lemma 7.5, there is some prime element $\xi \in \mathbb{Z}[\sqrt{2}]$ such that $\xi^\circ\xi \sim p$. And thus $\xi^\circ\xi|(x + \sqrt{2})^\circ(x + \sqrt{2})$, so $(x + \sqrt{2})$ will contain ξ as a factor as well therefore $\alpha \sim \xi$.

□

Lemma 7.8. *[RS16] Let $\xi \in \mathbb{Z}[\sqrt{2}]$ be a prime element and p be the unique integer prime such that $\xi|p$, then ξ is decomposable if and only if $p = 2$ or $p \equiv 1, 3, 5 \mod 8$*

Proof. we split into different cases

1. if $p = 2$, then $\xi \sim \sqrt{2}$, let $t = 1 + w$, we see $(1 + w)^\dagger(1 + w) = (1 + \sqrt{2})\sqrt{2} \sim \sqrt{2} \sim \xi$
2. if $p \equiv 1, 5 \mod 8$, then we know $p \equiv 1 \mod 4$, by fact 6, there exists integer u such that $u^2 \equiv -1 \mod p$, thus we have:
$$\xi \mid p \mid u^2 + 1 = (u + i)(u - i)$$
Let $t = \gcd(\xi, u + i)$ in $\mathbb{Z}[w]$, then we have $t^\dagger t|\xi^2$. Since ξ is prime in $\mathbb{Z}[\sqrt{2}]$, we again have three cases:
$$A : t^\dagger t \sim 1, \quad B : t^\dagger t \sim \xi, \quad C : t^\dagger t \sim \xi^2$$
A is impossible as it implies ξ is coprime with $u \pm i$ but we have $\xi|u^2 + 1$. B is what we claim. For C, by the same argument in lemma 5, we would have $t \sim \xi$, this implies $\xi|u \pm i$, but then we have:
$$\xi|(u + i) - (u - i) \sim 2 \neq p$$
contradicting that p is unique. Thus we have found $t^\dagger t \sim \xi$
3. if $p \equiv 3 \mod 8$, then by fact 6, there exists integer u such that $u^2 \equiv -2 \mod p$, thus we have:
$$\xi \mid p \mid u^2 + 2 = (u + i\sqrt{2})(u - i\sqrt{2})$$

Let $t = \gcd(\xi, u + i\sqrt{2})$ in $\mathbb{Z}[w]$, then we have $t^\dagger t | \xi^2$, we have the exact same three cases as part 2. A and B are exactly analogous. For C, follow the argument in part 2 we get:

$$\xi | (u + i\sqrt{2}) - (u - i\sqrt{2}) \sim 2\sqrt{2}$$

$$\xi | 2\sqrt{2}, \quad \xi | p \implies \xi | \gcd(p, 2\sqrt{2}) = 1$$

where the co-primality of $p, 2\sqrt{2}$ can be shown by considering their norm and ξ cannot be a unit.

4. if $p \equiv 7 \mod 8$, then we know the prime factorization of p in $\mathbb{Z}[\sqrt{2}]$ is $p \sim \xi^\circ \xi$. This is because by fact 6, 2 is a square modulo p and define $u^2 \equiv 2 \mod p$, we have $(u + \sqrt{2})^\circ (u + \sqrt{2}) \equiv 0 \mod p$, then in $\mathbb{Z}[\sqrt{2}]$, we have

$$p | (u + \sqrt{2})^\circ (u + \sqrt{2}), \quad p \nmid (u + \sqrt{2}), \quad p \nmid (u + \sqrt{2})^\circ \implies p \text{ not a prime in } \mathbb{Z}[\sqrt{2}]$$

This means $\xi \sim p$ is impossible thus by lemma 5. Thus suppose we also have $\xi \sim t^\dagger t$, then:

$$p \sim (t^\dagger t)^\circ (t^\dagger t) = (t^\circ t)^\dagger (t^\circ t)$$

Note by definition, the $\sqrt{2}$ conjugation acts on elements of $\mathbb{Z}[w]$ in the following way:

$$(aw^3 + bw^2 + cw + d)^\circ = -aw^3 + bw^2 - cw + d$$

Thus one can easily verify that $t^\circ t$ is a Gaussian integer $\mathbb{Z}[i]$. Then we have:

$$p \sim (a + bi)(a - bi) = a^2 + b^2$$

This is impossible when $p \equiv 7 \mod 8$ as perfect squares are either 0,1 or 4 modulo 8.

□

Now we can extend the discussion to prime powers of the form ξ^m, we have a similar result:

Lemma 7.9. *[RS16] Let $\xi \in \mathbb{Z}[\sqrt{2}]$ be a prime element, and p be the unique integer such that $\xi | p$ in $\mathbb{Z}[w]$. Then ξ^m is decomposable if and only if m is even or $p \equiv 1, 2, 3, 5 \mod 8$*

Proof. The case m is even is simple, just take the square root.
If $p \equiv 1, 2, 3, 5 \mod 8$, (here $p \equiv 2$ is really $p = 2$), then by lemma 7:

$$\exists s \in \mathbb{Z}[w], \quad s^\dagger s \sim \xi \implies (s^m)^\dagger (s^m) \sim \xi^m$$

The case $m \equiv 1 \mod 2 \wedge p \equiv 7 \mod 8$ is impossible for the same reason as in lemma 7 (that $p^m = a^2 + b^2$).

With lemma 7.8 and lemma 7.9, we actually have an algorithm to both check and solve the decomposition problem of $\xi_i^{m_i}$. The only thing we need is to be able to solve:

$$u^2 \equiv -1 \mod p$$

$$u^2 \equiv -2 \mod p$$

And there is a well-know algorithm by [Rab80] that solves these two equations in probabilistic polynomial time. □

7.3 Generalization to the case $\xi \in D[\sqrt{2}]$

In section 3 we have solve the relaxed decomposition problem when $\xi \in \mathbb{Z}[\sqrt{2}]$. Thus, we would want to reduce the general case problem to the special case problem. This is motivated by the following lemma:

Lemma 7.10. *[RS16] Given $\xi \in D[\sqrt{2}]$, then ξ is decomposable if and only if $\sqrt{2}\xi$ is decomposable.*

Proof. First notice that:

$$(1+w)^\dagger(1+w) = 2+\sqrt{2} = \sqrt{2}(1+\sqrt{2}) \sim \sqrt{2}$$

As we know $1+\sqrt{2}$ is a unit in $\mathbb{Z}[\sqrt{2}]$. Also $1+w$ is invertable in $D[w]$, with:

$$(1+w)\cdot\frac{1+w}{(2+\sqrt{2})w} = \frac{(1+w)^2}{(1+w)^\dagger(1+w)w} = \frac{1+w}{(1+w^\dagger)w} = 1$$

$$\frac{1+w}{(2+\sqrt{2})w} = \frac{1+w}{\sqrt{2}}\cdot w^{-1}\cdot(1+\sqrt{2})^{-1} \in D[w]$$

Thus if we have $t^\dagger t \sim \xi$, then let $s = (1+w)t$, we have:

$$s^\dagger s = (1+w)^\dagger(1+w)t^\dagger t \sim \sqrt{2}\xi$$

If we have $t^\dagger t \sim \sqrt{2}\xi$, then let $s = (1+w)^{-1}t$. □

Theorem 7.11. *[RS16] Given an integer factoring oracle, we can solve the equation:*

$$t^\dagger t \sim \xi, \quad \xi \in D[\sqrt{2}], \quad t \in D[w]$$

in probabilistic polynomial time

Proof. Without loss of generality, we assume ξ to be doubly positive, thus:

$$\xi^\circ\xi = \frac{n}{2^k} \in D \text{ for some } n \in \mathbb{Z}^+, k \in \mathbb{N}$$

Let $\xi' = \sqrt{2}^k\xi$, then we have (note k must be even):

$$\xi' \in \mathbb{Z}[\sqrt{2}], \quad \xi'^\circ\xi' = n$$

With the factoring oracle, let the integer prime factorization be $n = \prod_{i=1}^N p_i^{m_i}$. Then each p_i can be efficiently factored into primes in $\mathbb{Z}[\sqrt{2}]$ by lemma 5, which yields a prime factorization of $\xi'^\circ\xi'$ in $\mathbb{Z}[\sqrt{2}]$. This allows us to find the prime factorization of ξ' easily, (because each prime factor of n has only two possible factorizations, both are easy to handle).Thus:

$$\exists s \in D[w], \quad s^\dagger s \sim \sqrt{2}^k\xi$$

By Lemma 7.10, this implies that:

$$[(1+w)^{-k}s]^\dagger[(1+w)^{-k}s] \sim \xi$$

And we are done. □

7.4 A Step-by-step Summary

1. First, we assume that our search algorithm returns a valid candidate $u \in D[\omega]$ such that:

$$\xi = 1 - u^\dagger u$$

is doubly positive. This is automatically ensured if one requires $u \in \mathcal{R}_\varepsilon, u^\circ \in \overline{\mathcal{D}}$.
Easy calculation would show that $\xi \in D[\sqrt{2}]$ and therefore:

$$\xi^\circ \xi \in D$$

Note that the search algorithm also produces the sde of u, denote it as k, then we can write:

$$\xi^\circ \xi = \frac{n}{2^{2k}}, n \in \mathbb{Z}^+$$

This is because, since u has sde k, we have:

$$u = \frac{a}{\sqrt{2}^k}, \quad a \in \mathbb{Z}[\omega]$$

$$u^\dagger u = \frac{a^\dagger a}{2^k} = \frac{|a|^2}{2^k} = \frac{b}{2^k}, \quad b \in \mathbb{Z}[\sqrt{2}]$$

$$\xi = \frac{2^k - b}{2^k} = \frac{c}{2^k}, \quad c \in \mathbb{Z}[\sqrt{2}]$$

$$\xi^\circ \xi = \frac{c^\circ c}{2^{2k}} = \frac{n}{2^{2k}}, \quad n \in \mathbb{Z}$$

And n is not negative follows from the fact that ξ is doubly positive.

2. Define $\xi' = \sqrt{2}^{2k}\xi = 2^k\xi$, then our previous calculation shows:

$$\xi' \in \mathbb{Z}[\sqrt{2}], \quad \xi'^\circ \xi' = n$$

Lemma 7.10 suggests that the decomposition of ξ is related to the decomposition of ξ':

$$t'^\dagger t' \sim \xi' = \sqrt{2}^{2k}\xi \implies [(1+\omega)^{-2k}t]^\dagger[(1+\omega)^{-2k}t] \sim \xi$$

And we are also guaranteed that $(1+\omega)^{-2k}t$ is in $D[\omega]$ if $t \in D[\omega]$. because $(1-\omega)^{-1} \in D[\omega]$. Therefore, our algorithm proceeds by computing the decomposition of ξ'.

3. With a factoring oracle, we first find a prime factorization of n:

$$n = \prod_{i=1}^{N} p_i^{m_i}$$

For each p_i, we can use Lemma 7.7 to easily compute its prime decomposition in $\mathbb{Z}[\sqrt{2}]$. Note that if p_i is a prime in $\mathbb{Z}[\sqrt{2}]$, then m_i will have to be even and otherwise the prime factorization of p_i will be of form $\alpha_i^\circ \alpha_i$, and therefore we can also deduce the prime factorization of ξ' in $\mathbb{Z}[\sqrt{2}]$:

$$\xi' = \prod_{i=1}^{N} q_i^{n_i}, \quad q_i^{n_i} = \begin{cases} p_i^{m_i/2} & \text{if } p_i \text{ is prime in } \mathbb{Z}[\sqrt{2}] \\ \alpha_i^{m_i} & \text{if } p_i = \alpha_i^\circ \alpha_i \end{cases}$$

4. For each $q_i^{n_i}$, (note q_i are prime elements in $\mathbb{Z}[\sqrt{2}]$, apply Lemma 7.9 to obtain its decomposition, if there is one. Note that the unique prime integer such that q_i divides it is exactly p_i, so the test can be carried out efficiently. Also, note Lemma 7.9 is constructive, so if the decomposition exists, we can also compute it efficiently.

5. Lemma 7.3 suggests that the decomposition of ξ' exists if and only the decomposition of each $q_i^{n_i}$ exists. So if we managed to find $t_i^\dagger t_i \sim q_i^{n_i}$ for all i, then the decomposition of ξ' is easily given by:

$$\xi' = t'^\dagger t'. \quad t' = \prod_{i=1}^{N} t_i$$

   And then we also get $\xi = t^\dagger t, \quad t = (1+\omega)^{-2k} t'$

6. Once we have $t^\dagger t \sim \xi$, simply compute $\beta^2 = \xi / t^\dagger t$, Lemma 7.1 suggests that it will be a unit and a square in $\mathbb{Z}[\sqrt{2}]$. Then we just need to divide it repeatedly by $(1+\sqrt{2})$ to obtain its factorization and thus to obtain β. And finally, we get $(\beta t)^\dagger (\beta t) = \xi$

8 Putting Everything Together

This section outlines the Ross-Selinger Algorithm

1. Step 1: Use either the approach introduced in Section 5 or Section 6, search for a candidate u satisfying the constraints introduced in Definition 4.5.
   This step takes $O(1)$ time (ignoring bit-length of inputs/outputs).

2. Step 2: For the candidate u found, attempt to solve the norm equation:

$$\|t\|^2 = 1 - \|u\|^2, \quad t \in D[\omega]$$

   with the algorithm summarized in Section 7.4. On success, proceed to Step 3, otherwise go back to Step 1. In the absence of a factoring oracle, if the factoring heuristic runs longer than a threshold, skip this u and go back to Step 1.
   This step takes $O(\text{polylog}(n))$ time, where n is defined the same way as in Section 7.4.

3. Step 3: Use the algorithm introduced in Theorem 3.6 to convert (u, t) into a circuit.
   This step takes $O(sde(u))$ time.

9 Complexity Analysis

First, it is natural to ask if it makes any difference, for the sake of complexity analysis, that we adopt the grid operator approach (Section 5) or the lattice region search approach (Section 6)to do the searching. The answer is no. This is because both approaches solves the same search problem, in the sense that they both enumerate candidate u in increasing order of SDE under the same set of constraints. So without loss of generality, one may assume the outputs of the two approaches are identical, as permutations of candidates with same SDE does not matter in complexity analysis (no particular order is assumed to exist within candidates of the same SDE anyway). Furthermore, while there is a slight difference in the run time complexity, both approaches on average enumerates

a candidate in constant time (ignoring bit-length of the solution), which is all one needs for time complexity analysis.

Note, by the construction of Ross-Selinger, both run time and T-count would depend on how many candidates u we have to try before we manage to solve the norm equation. Therefore, it is helpful to first show the complexities of some key quantities capturing the notion of how "larg" the SDE of u has to be in order to achieve ε-approximation and how "larger" the SDE has to be to ensure the norm equation is also solvable.

Looking at the procedure outlined in section 7.4, we note that if n happens to be a prime number, then the analysis is much simplified and if n is a prime congruent to $1, 2, 3, 5$ modulo 8, then Lemma 7.9 guarantees the existence of the solution. Therefore, we would argue that we will need to try at most as many u candidates as needed to get such a prime n congruent to $1, 2, 3, 5$ modulo 8. The following lemma suggests we can further simplify our analysis.

Lemma 9.1. *[RS16] The integer n in step 1 of 7.4 is either 0 or an odd number congruent to 1 modulo 8. Further, we have $n \leq 4^k$, where k is the sde of u.*

Proof. Note that this lemma is important because the procedure in 7.4 does not assume or require n to be odd, but since the denominator is a power of 2, we can always reduce the factor to its simplest form where n is odd, and we just modify the subsequent procedure slightly (replace $\xi' = \sqrt{2}^{2k}$ with $\xi' = \sqrt{2}^{l}$ where l is the most reduced denominator exponent).

Motivated by the algorithm's design, we note that any $u \in D[\omega]$ also has the form:

$$u = \frac{a\omega^3 + b\omega^2 + c\omega + d}{(1+\omega)^{k'}}$$

And some explicit calculation shows that k' would be minimal if and only if $a + b + c + d$ is odd. Let k' be the minimal denominator $(1 + \omega)$ power.

1. Case 1: $k' \leq 1$, one can show through direct calculation that the search part only has 9 solutions and one has $n = 1$ while the rest have $n = 0$.

2. Case 2: $k' \geq 2$, Therefore, we can write (this again is some direct calculation):

$$u^\dagger u = \frac{1}{[(1+\omega)^\dagger(1+\omega)]^{k'}}(A + B\sqrt{2})$$

$$A = a^2 + b^2 + c^2 + d^2, \quad B = cd + bc + ab - da$$

Further, we notice that $[(1+\omega)^\dagger(1+\omega)]^{k'} = \sqrt{2}^{k'}(1+\sqrt{2})^{k'}$ is an element of $\mathbb{Z}[\sqrt{2}]$ and is divisible by 2, so we can write $[(1+\omega)^\dagger(1+\omega)]^{k'} = C + D\sqrt{2}$ for some even integer $C, D.$, so we have:

$$\xi = 1 - u^\dagger u = \frac{1}{[(1+\omega)^\dagger(1+\omega)]^{k'}}(C + D\sqrt{2} - A - B\sqrt{2}) = \frac{1}{[(1+\omega)^\dagger(1+\omega)]^{k'}}(x + y\sqrt{2})$$

$x = C - A$ is odd because A is odd (k' is minimal implies $a+b+c+d$ is odd) and $y = D - B$ is even because B has same parity as $(a+c)(b+d)$ which must be even. Finally we have:

$$\xi^\circ\xi = \frac{1}{\{[(1+\omega)^\dagger(1+\omega)]^\circ[(1+\omega)^\dagger(1+\omega)]\}^{k'}}(x^2 - 2y^2)$$

The denominator is fact just $2^{k'}$ and $x^2 - 2y^2 \equiv 1 \mod 8$ follows from parity of x, y.

The last part simply follows from the fact that $\xi^\circ\xi \leq 1$ as ξ, ξ° are in $[0,1]$ by construction. □

Thus, we only need to try as many u as needed to get n being a prime number. So what is the probability of n being a prime number? This requires a number theoretical hypothesis:

Hypothesis 9.2. *[RS16] The number n produced in step 1 of the procedure, being an odd number 1 modulo 8, is asymptotically at least as likely to be a prime as a randomly chosen odd number of comparable size (i.e. $\leq 4^k$). That is to say, we are assuming that:*

$$\pi(x; 8, 1) \sim \frac{1}{4}\frac{x}{\log(x)}$$

which says that approximately a quarter of the primes are 1 modulo 8. And one can model the primality of each n as an independent random variable.

Let u_i be the i-th candidates produced by the search algorithm and k_i, n_i be the corresponding SDE and integer calculated in step 1 of 7.4. We have that $n_i \leq 4^{k_i}$ by Lemma 9.1. And using the prime number theorem, we have:

$$P(n_j \text{ is a prime }) = p_j \geq \frac{2}{\log(4^{k_j})} = \frac{1}{k_j \log(2)}$$

Moreover, one could show that if $j \leq 2^l + 1$, then $k_j \leq k_2 + 2l$, this comes from a lemme for the grid search that says that if the grid search problem for k has at least 2 solutions and the problem for $k + 2l$ as at least $2^l + 1$ solutions. Setting $l = 1 + \log_2(j)$, we get:

$$k_j \leq k_2 + 2(1 + \log_2 j) \tag{18}$$

Therefore:

$$p_j \geq \frac{1}{(k_2 + 2(1 + \log_2 j))\log(2)} = \frac{1}{(k_2+2)\log(2) + 2\log(j)}$$

Let n_{j_0} be the first prime n, assuming independence, it is easy to get:

$$P(j_0 > j) = \prod_{i=1}^{j}(1 - p_i) \leq \left(1 - \frac{1}{(k_2+2)\log(2) + 2\log(j)}\right)^j$$

$$E(j_0) = \sum_{j=0}^{\infty} P(j_0 > j) \leq 1 + \sum_{j=1}^{\infty}\left(1 - \frac{1}{(k_2+2)\log(2) + 2\log(j)}\right)^j = O(k_2)$$

It is not obvious that the last expression is $O(k_2)$, but one could show this with elementary analysis techniques (truncating series into 2 parts, with the tail being dominated by $1/j^2$ and the head being $O(k_2)$).

Therefore, it suffices to show that $k_2 = O(\log(1/\varepsilon))$. This is to ask, how large must k be in order for the search problem with $\mathcal{R}_\varepsilon$ and $\overline{\mathcal{D}}$ to have at least 2 solutions. There are many ways to show this, depending on what kind of algorithm we are using for the searching problem, but essentially one just needs to consider the un-scaled problem for $\mathbb{Z}[\omega]$, and it is not to hard to come up with some bound that the problem will have at least 2 solutions if:

$$p(\varepsilon) \geq C$$

for some constant C and polynomial $p(\varepsilon)$ (note the region $\mathcal{R}_\varepsilon$ always has a circle of radius $\varepsilon^2/4$ in it), and then the scaled version is simply:

$$p(\varepsilon) \geq \frac{C}{2^k}$$

which leads to

$$k_2 = O(\log(1/\varepsilon)) \tag{19}$$

and then it also follows that:

$$E(j_0) = O(\log(1/\varepsilon)) \tag{20}$$

9.1 T-Count Complexity

9.1.1 The information theoretical lower bound

Before we use the above results to show the T-count complexity of Ross-Selinger in the absence of a factoring oracle, it is interesting to have a lower bound for the number of T gates needed to be sufficient to ε-approximate any member of $SU(2)$ (note not just z-rotations). The following information theoretical bound follows from a simple counting argument.

Theorem 9.3. *[Sel15] Let $T(U;\varepsilon)$ represent the optimal T-count for approximating $U \in SU(2)$ up to ε and define:*

$$T_{\max}(\varepsilon) = \sup_{U \in SU(2)} T(U;\varepsilon)$$

Then the information theoretical lower bound for $T_{\max}$ is:

$$K + 3\log_2(1/\varepsilon) \leq T_{\max}(\varepsilon)$$

where K is some Constant

Proof. Using the Matsumoto and Amano normal form for Clifford+T circuits, [MA08] showed that there are exactly $192(3 \cdot 2^n - 2)$ distinct Clifford+T circuits of T-counts $\leq n$. Since $SU(2)$ is 3-dimensional (this can be seen from Theorem 3.2 as well), it takes $O(1/\varepsilon^3)$ ε-balls to fully cover the space, therefore to make sure there is at least one distinct circuit per ball, we get the inequality:

$$192(3 \cdot 2^n - 2) \geq \frac{K'}{\varepsilon^3}$$

$$\implies n \geq K + 3\log_2(1/\varepsilon)$$

□

9.1.2 T-Count Complexity for Ross-Selinger

Note that Ross-Selinger, with a factoring oracle, is guaranteed to produce the least T-count possible. Therefore, it is only meaningful to analyze T-count complexity in the absence of a factoring oracle.

Theorem 9.4. *[RS16] Let m be the T-count of the circuit found by Ross-Selinger in the absence of a factoring oracle, and let m' and m'' be the optimal and second-to-optimal T-count of the same problem, then we have:*

$$E[m] = m'' + O(\log\log(1/\varepsilon))$$

Proof. Let k be the SDE of the solution produced by the algorithm. Let k' and k'' be the corresponding SDEs for the optimal and second-optimal solutions, recall that k_{j_0} is the SDE for u_{j_0} (the first u found by the search algorithm that is guaranteed to have a quickly solvable norm equation) and k_2 is the SDE of u_2 (the second u found by the search algorithm). By definition of k_{j_0} and by Equation **??**, we have:

$$k \leq k_{j_0} \leq k_2 + 2(1 + \log_2 j_0) \leq k'' + 2(1 + \log_2 j_0)$$

Using Lemma 3.10, we have $m \leq 2k$ and $2k'' - 2 \leq m''$, hence:

$$m \leq 2k \leq 2k'' + 4(1 + \log_2 j_0) \leq m'' + 6 + 4\log_2 j_0$$

$$E[m] \leq m'' + 6 + 4E[\log_2 j_0] \leq m'' = 6 + 4\log_2 E[j_0]$$

Using result **??**

$$= m'' + O(\log\log(1/\varepsilon)$$

□

9.1.3 Some Comparisons

Given the z-rotation approximation problem $(R_z(\theta), \varepsilon)$, we define the function $T(\theta;\varepsilon)$ to be the optimal T-count for this problem. Let $A(\theta;\varepsilon)$ be the T-count of the solution produced by Ross-Selinger with a factoring oracle, then we have shown by construction:

$$\forall \varepsilon, \theta, \quad A(\theta;\varepsilon) = T(\theta;\varepsilon)$$

While in the absence of a factoring oracle (indicated by A'), by Theorem 9.4, we have:

$$E[A'(\theta;\varepsilon)] = T(\theta;\varepsilon) + O(\log(\log(1/\varepsilon)))$$

[Sel15] also showed that for an old version of the algorithm (indicated by A_s):

$$10 + 4\log(1/\varepsilon) \geq \sup_\theta A_s \geq \sup_\theta T \geq -9 + 4\log(1/\varepsilon)$$

In [RS16], they have argued that it is also true that

$$\sup_\theta A_s \geq \sup_\theta A'$$

$$E_\theta[A'] = K + 3\log(1/\varepsilon)$$

So overall we have:

$$10 + 4\log(1/\varepsilon) \geq \sup_{\theta} A_s \geq \sup_{\theta} A' \geq \sup_{\theta} T \geq -9 + 4\log(1/\varepsilon)$$

On the other hand, we have the information theoretical lower bound for T-count needed for universal ε-approximation of $SU(2)$:

$$\sup_{U \in SU(2)} T(U;\varepsilon) \geq K + 3\log(1/\varepsilon)$$

So if we use Ross and Selinger and do approximation by decomposing arbitrary unitary into 3 z-rotations, we get:

$$30 + 12\log(1/\varepsilon) \geq \sup_{U} T(U;\varepsilon) \geq K + 3\log(1/\varepsilon)$$

9.2 Run Time Complexity

Theorem 9.5. *[RS16] Ross-Selinger runs in expexted time $O(polylog(1/\varepsilon))$. This is true regardless of the use of a factoring oracle and which approach one adopts for the searching problem.*

Proof. For simplicity, suppose one uses the lattice region search approach, the story is quite simple. First it takes on average $O(1)$ time to enumerate a candidate u. And then the norm equation solver tries to solve for t in expected time $O(\text{poly}(\text{sizeof}(n))) = O(\text{polylog}(n))$ where n is defined the same way as in section 7.4. We have argued that this process takes on average j_0 candidates for the norm equation solver to be guaranteed to succeed in short time (this is true in the absence of factoring). And the last step is to convert u_{j_0} to a circuit using the algorithm described in Theorem 3.6, which runs in $O(k_{j_0})$. Therefore, in total we have:

$$j_0 \cdot O(\text{polylog}(n)) + O(k_{j_0})$$

Equation 18, combined with Equation 19,20, suggests:

$$k_{j_0} \leq k_2 + 2 + 2\log_2(j_0) = O(\log(1/\varepsilon)) + O(\log\log(1/\varepsilon)) = O(\log(1/\varepsilon))$$

Lemma 9.1 suggests $n \leq 4^{k_{j_0}}$ and using the above result:

$$O(\text{polylog}(n)) = O(\text{poly}(k_{j_0})) = O(\text{polylog}(1/\varepsilon))$$

To show the same is true for grid operator approach, one needs to also take into account the time needed to construct the grid operators, which is not necessary in the lattice approach. Nevertheless, [RS16] showed that this extra time is also $O(\log(1/\varepsilon))$ □

10 Acknowledgement

I would like to thank my mentor, Prof. Oded Regev, for his kind and patient mentorship throughout my undergraduate career. This thesis would not be complete without his generous assistance.

References

[DN06] Christopher M. Dawson and Michael A. Nielsen. The Solovay-Kitaev algorithm. *Quantum Inf. Comput.*, 6(1):81–95, 2006.

[GS19] Brett Giles and Peter Selinger. Remarks on Matsumoto and Amano's normal form for single-qubit Clifford+T operators. 2019.

[HGS04] Nick Howgrave-Graham and Mike Szydlo. A method to solve cyclotomic norm equations $f * \overline{f}$. In *Algorithmic number theory*, volume 3076 of *Lecture Notes in Comput. Sci.*, pages 272–279. Springer, Berlin, 2004.

[KMM13] Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca. Fast and efficient exact synthesis of single-qubit unitaries generated by Clifford and T gates. *Quantum Inf. Comput.*, 13(7-8):607–630, 2013.

[KMM16] Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca. Practical approximation of single-qubit unitaries by single-qubit quantum Clifford and T circuits. *IEEE Trans. Comput.*, 65(1):161–172, 2016.

[Len83] H. W. Lenstra, Jr. Integer programming with a fixed number of variables. *Math. Oper. Res.*, 8(4):538–548, 1983.

[Lov86] László Lovász. *An algorithmic theory of numbers, graphs and convexity*, volume 50 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1986.

[MA08] Ken Matsumoto and Kazuyuki Amano. Representation of quantum circuits with Clifford and $\pi/8$ gates, 2008.

[PS18] Ori Parzanchevski and Peter Sarnak. Super-golden-gates for $PU(2)$. *Adv. Math.*, 327:869–901, 2018.

[Rab80] Michael O. Rabin. Probabilistic algorithms in finite fields. *SIAM J. Comput.*, 9(2):273–280, 1980.

[RS16] Neil J. Ross and Peter Selinger. Optimal ancilla-free Clifford + T approximation of z-rotations. *Quantum Inf. Comput.*, 16(11-12):901–953, 2016.

[Sel15] Peter Selinger. Efficient Clifford+T approximation of single-qubit operators. *Quantum Inf. Comput.*, 15(1-2):159–180, 2015.

[Sel16] Peter Selinger. Number-theoretic methods in quantum computing, Sep 2016.

[ZLC00] Xinlan Zhou, Debbie W. Leung, and Isaac L. Chuang. Methodology for quantum logic gate construction. *Physical Review A*, 62(5), Oct 2000.