CLASSICAL CONSTRUCTION OF QUANTUM CODES

RUOCHUAN XU

Abstract. This paper explores a generalization of quantum error detection and correction in representations of Lie algebras, specifically in a class of irreducible representations of $\mathfrak{sl}_n\mathbb{C}$. The central method introduced is a two-step construction that transforms the problem of finding quantum codes into a classical problem of convex geometry. The purpose of this paper is also two-fold. First, we wish to build up background for this rarely studied topic and explain essential aspects of classical and quantum codes that motivate our study. Second, we discuss the interplay of irreducible representations of $\mathfrak{sl}_n\mathbb{C}$ with the so-called Lie type graph metric and present code constructions in representations of $\mathfrak{sl}_n\mathbb{C}$ for $n \leq 4$.

Contents

1. Introduction

The essential question to ask in the theory of classical linear codes is the following. Given the dimension of a Hamming space n and the intended code distance d, find a distance d code subspace C in $(\mathbb{Z}/2)^n$ with dimension as large as possible. If such a code C of dimension k can be found, we say C is an $[n, k, d]$ code. This is essentially a problem of sphere packing in a discrete metric, and it invites upper and lower bounds for the packing density.

Date: January 2023.

In the usual theory of quantum codes, people consider finding a distance d code subspace $\mathcal{C}$ in the Hilbert space $(\mathbb{C}^2)^{\otimes n}$ representing n qubits. It is often the case that we want $\mathcal{C}$ to be isomorphic to $(\mathbb{C}^2)^{\otimes k}$, so that $\mathcal{C}$ encodes k qubits, and in this case we say $\mathcal{C}$ is an $[[n, k, d]]$ code. Of course, given n and d, we want k to be as large as possible. Comparing these two central problems in classical and quantum codes, we already see a strong analogy.

Some aspects of this analogy arise from practical concerns. For instance, classically we compute with combinations of two bits 0 and 1, and quantum-ly people wish to compute with a pair of basis states $|0\rangle$ and $|1\rangle$ of $\mathbb{C}^2$. Thus, limiting to Hilbert spaces of the form $(\mathbb{C}^2)^{\otimes n}$ is negotiable from a mathematical point of view. The central aspect of the classical-quantum analogy we here explore is the role of distance in the classical and quantum theories of codes. People have long realized that finding classical codes in any metric space including the Hamming space is just a game of sphere packing. Likewise, once we have the error correction and detection criteria for quantum codes, we can generalize the notion of distance for codes in $(\mathbb{C}^2)^{\otimes n}$ and consider quantum codes in other so-called quantum metric spaces. We follow the definition of a quantum metric proposed in [9].

Just like it is natural to study upper and lower bounds for sphere packing, we want to construct quantum codes and try to see if in any sense they are optimal. We primarily do this in the so-called $\mathfrak{sl}_n\mathbb{C}$ graph metric spaces, as first studied in [1]. That is, we investigate code subspaces in certain irreducible representations V of $\mathfrak{sl}_n\mathbb{C}$. We think of the action of $\mathfrak{sl}_n\mathbb{C}$ as quantum errors and based on this assign a natural quantum metric *i.e.* a filtration on $\mathcal{L}(V)$ that in turn puts conditions on code subspaces $\mathcal{C}$ of a desired distance d.

We start with brief introductions to classical and quantum codes. Then, we explain a central two-step construction for quantum codes: in the first step, we find a classical code, but this must be in service for the second step, as we need to satisfy a convex geometry criterion. The general form of this two-step construction was proposed in [7]. The remaining half of the paper discusses quantum metric spaces arising from certain representations of $\mathfrak{sl}_n\mathbb{C}$ and applications of the two-step construction in these spaces, as well as a different code construction strategy in the case of $\mathfrak{sl}_2\mathbb{C}$.

2. Classical Linear Codes

We briefly review the theory of classical linear codes, a simple story that turns out to be a rich source of ideas for quantum codes. The mathematical model for classical codes is as follows. Assume a finite space of messages X subject to errors. That is, a message $x \in X$ has a probability $p(x, y)$ to be altered to some $y \in X$. A code is a subset $C \subset X$, where elements in C are called *codewords*, to be interpreted as the only valid messages. To avoid working with transition probabilities, we further assume that X is equipped with a metric d such that if $d(x', y') > d(x, y)$, then $p(x', y') \ll p(x, y)$. In other words, an erroneous message is most likely close to the original message in the metric space (X, d). As a result of this simplification, a **code** C is a minimum-distance set, with a minimum distance or simply **distance** $d(C) \coloneqq \min_{x,y\in C, x\neq y} d(x, y)$. Similarly, we define the distance of an **error** $e : X \to X$ to be $d(e) \coloneqq \max_{x\in C} d(x, e(x))$. Since the metric d is usually understood in context, we make a slight abuse of notation and consider a code C with distance $d(C) = d$, then

(i) errors up to distance $< d$ can be detected;
(ii) errors up to distance $< d/2$ can be corrected.

Here, an error with distance less than d either results in no change to a codeword or maps it to some element in X outside C, in which case the erroneous message is declared invalid. As for error correction, notice that metric balls of radius $t < d/2$ centered at each point in C are disjoint. If all errors we wish to correct have distance $\leq t$, we can interpret an erroneous message lying in a radius t closed metric ball to have as original value the center of that ball.

Definition 2.1. Let X be a metric space, and $\mathrm{Isom}(X)$ the group of all distance-preserving bijections of X. Then X is **homogeneous** if $\mathrm{Isom}(X)$ acts on X transitively.

In particular, if X is homogeneous, then balls of radius t around any point $x \in X$ has a volume independent of x. This further assumption allows us to come up with elementary bounds for the

size of a distance-d code. We shall henceforth also assume that the metric is integer-valued, as all metrics we are concerned with are.

The **Hamming bound** is the volume upper bound that for a code C of distance d in a homogeneous metric space X,

$$|C| \leq \frac{|X|}{|B_t(x)|},$$

where $t = \lfloor \frac{d-1}{2} \rfloor$ and $|B_t(x)|$ is the volume of the closed ball of radius t centered at any point x in X. The inequality holds because **error balls** of radius t centered at each codeword in C must be disjoint by the triangular inequality, as required for error correction. When equality is achieved, such error balls **tile** X and we say the code C is **perfect**.

Corresponding to the Hamming upper bound, the most elementary lower bound for classical codes is the Gilbert-Varshamov bound. Given a metric space X and desired distance d, a **lower bound** for a code is an upper bound on the size of the code, below which a code $C \subset X$ of distance d must exist. It is a lower bound for the maximal size of a code in X with distance d. There are two similar versions of the Gilbert-Varshamov bound, attributed respectively to Gilbert and Varshamov.

Proposition 2.2 (Gilbert Bound for General Codes)**.** *Let X be a homogeneous metric space. There exists a code C in X of distance d if*

$$|B_{d-1}(0)||C| \leq |X|.$$

Here, 0 is an arbitrary point in X that we choose to be in C.

Proof. We greedily construct the code as follows. If $B_{d-1}(0)$ does not cover X, choose an arbitrary point $x \in X \backslash B_{d-1}(0)$ as a codeword. If $B_{d-1}(0) \cup B_{d-1}(x)$ does not cover X, then we may choose another point in the complement as a codeword, such that the pairwise distance between all codewords chosen so far is at least d. In this way, k codewords can be chosen as long as we do not exhaust X, i.e., as long as $|B_{d-1}(0)| \cdot k \leq |X|$. □

The Varshamov bound considers a smaller class of codes in a smaller class of metric spaces.

Definitions 2.3. A **classical Hamming space** X is a product set A^n where A is a finite set, equipped with the **classical Hamming metric** defined by

$$d((a_1, \ldots, a_n), (b_1, \ldots, b_n)) = |\{a_i \mid a_i \neq b_i\}|.$$

Denoting $|A|$ by m, we can endow A with the group structure of $\mathbb{Z}/m$, so that $X = (\mathbb{Z}/m)^n$. A code C that is a submodule of the free $\mathbb{Z}$-module $(\mathbb{Z}/m)^n$ is called a **linear code**.

The following proof probabilistically exhibits the existence of linear codes.

Proposition 2.4 (Varshamov Bound for Linear Codes)**.** *For a subset $Y \subset A^n$ containing 0, denote $Y \backslash \{0\}$ by Y°. There exists a k-dimensional linear code C in A^n of distance d if*

$$\left|B^\circ_{d-1}(0)\right||C^\circ| < |A^{n\circ}|.$$

Proof. Choose a k-dimensional submodule C of A^n uniformly at random. Then, for any $x \in A^{n\circ}$, $\mathbb{P}[x \in C] = |C^\circ|/|A^{n\circ}|$, although the events $\{x \in C\}$ are not independent. By the union bound,

$$\begin{aligned}\mathbb{P}[d(C) < d] &= \mathbb{P}[x \in C \text{ for some } x \in B^\circ_{d-1}(0)] \\ &\leq \sum_{x \in B^\circ_{d-1}(0)} \mathbb{P}[x \in C] \;=\; \left|B^\circ_{d-1}(0)\right||C^\circ|/|A^{n\circ}|\end{aligned}$$

There must exists a k-dimensional linear code C of distance $\geq d$ if the above probability is less than 1, from which the bound follows. □

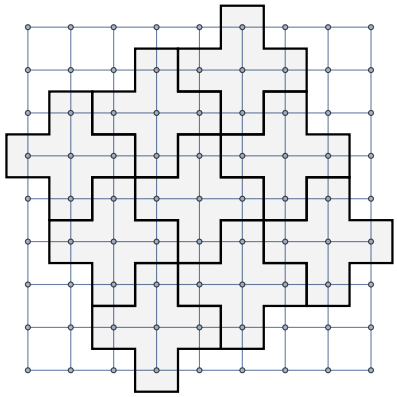

FIGURE 1. A tiling of the $\mathbb{Z}^2$ lattice, or a perfect packing of spheres of radius 1. The centers of the spheres form a sublattice of distance 3. In fact, this figure also exhibits a distance 3 perfect code in $(\mathbb{Z}/5)^2$.

The above theory can be applied to the prototypical class of classical Hamming spaces $X_n = (\mathbb{Z}/2)^n$ consisting of bit strings of length n. Here the **Hamming distance** between two bit strings is the number of bits where they differ. Since the Hamming metric is translation invariant, *i.e.* satisfies $d(x+z, y+z) = d(x, y)$ for all $x, y, z \in X_n$, there is an associated norm called the **Hamming weight**, given by $\mathrm{wt}(x) \coloneqq d(x, 0)$; it simply counts the number of 1's in the bit string x. The errors on X_n are described by bit flips that act on X_n by $x \mapsto x + e$ for all $x \in X_n$ and some $e \in X_n$; the 1's in e indicate bits that flip. The distance of a bit flip error e is simply its weight.

For a linear code $C \subset X_n$, minimum distance reduces to minimum weight, *i.e.*, $d(C) = \min_{c \in C, c \neq 0} \mathrm{wt}(c)$. A k-dimensional linear code in an n-dimensional Hamming space can be compactly described as the image of an $n \times k$ **generator matrix** G, or as the kernel of an $(n-k) \times n$ **parity check matrix** H. These matrices have nice interpretations in terms of encoding and decoding. G encodes k bits of information as a bit string of length n. To decode an erroneous message $x + e$, one applies H to get the **syndrome** $H(x + e) = He$ of the error e; a key point is that the syndrome does not depend on the message x. If we restrict to a sufficiently small subset $\mathcal{E}$ of errors such that the syndromes He are distinct for distinct e, then we are able to identify e from its syndrome, and perform error correction by $x + e \mapsto (x + e) + e = x$.

From the parity check matrix formalism, one obtains a concrete instance of the error correction criterion. Recall from (ii) above that a linear code of distance $d = 2t+1$ can correct errors of weight at most t. More explicitly, for distinct errors e_1 and e_2 each of weight at most t, $e_1 + e_2$ has weight more than 0 and less than d, hence it lies outside C. As a result, $H(e_1 + e_2) \neq 0$ and $He_1 \neq He_2$, so that the distinct syndromes guarantee error correction.

The distance of a linear code can be described in terms of its parity check matrix.

Proposition 2.5. *Let $C \subset X_n$ be a linear code with parity check matrix H. Then $d(C)$ is the minimum number of linearly dependent columns of H.*

Proof. We know $d(C)$ is the minimum weight of nonzero codewords in C, or the minimum number of nonzero entries achieved by nonzero bit strings in $\ker H$. The statement follows from the fact that Hx is a linear combination of columns of H with coefficients the entries of x, for any $x \in X_n$. □

Since perfect codes geometrically correspond to tilings, they are natural objects of investigation. We provide two simple examples in Figure 1 and the following. As the example in Figure 1 suggests, the notion of perfect codes remains well-defined in an infinite metric space. We say that a distance d code C in an infinite homogeneous metric space X is **perfect** if X is covered by the balls $\{B_t(y) \mid y \in C\}$, where $t = \lfloor \frac{d-1}{2} \rfloor$ as before.

Example 2.6. Let H_r be the $r \times (2^r - 1)$ matrix where column j is the binary representation of j. The **Hamming code** C_r is defined to be the code with H_r as the parity check matrix. Since H_r has rank r and the minimum number of linearly dependent columns is 3, by Prop. 2.5 we see that C_r are $[2^r - 1, 2^r - 1 - r, 3]$ codes. The Hamming code is perfect because

$$|C_r||B_1(0)| = 2^{2^r-1-r} \cdot \left(\binom{2^r-1}{0} + \binom{2^r-1}{1}\right) = 2^{2^r-1} = |X_{2^r-1}|.$$

3. Quantum Error Correction and Quantum Metric

3.1. **Quantum Errors.** A quantum vector state $|\psi\rangle$ in a complex Hilbert space $\mathcal{H}$ has a much more subtle nature compared to a classical bit string. What's more, quantum vector states are susceptible to a continuum of errors modelled by linear operators $E \in \mathcal{L}(\mathcal{H})$ that maps $|\psi\rangle$ to $E|\psi\rangle$, where $\mathcal{L}(\mathcal{H})$ is the space of bounded linear operators on $\mathcal{H}$. We are only interested in finite dimensional Hilbert spaces, so $\mathcal{H} = \mathbb{C}^n$ and $\mathcal{L}(\mathcal{H}) = M_n(\mathbb{C})$. It is perhaps surprising that to detect or correct a subspace $\mathcal{E}$ of quantum errors, it suffices to consider a finite collection of errors that is a basis for $\mathcal{E}$. This **discretization of errors** is a crucial fact that makes quantum error correction possible.

The physical description of quantum error correction is complicated by entanglement. For two quantum systems A and B with corresponding Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the joint system AB is described by the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. Physically, we may have access only to the subsystem A, for instance when B represents the environment. In this case, vector states are no longer sufficient to describe the state of a subsystem. If $|\psi\rangle$ is a vector state in $\mathcal{H}_A \otimes \mathcal{H}_B$, then the corresponding state of the subsystem A is described by the **density operator** $\rho_A = \mathrm{tr}_B(|\psi\rangle\langle\psi|)$, where tr_B denotes partial trace over $\mathcal{H}_B$. It is always possible to choose an appropriate basis and write ρ_A in diagonal form as

$$\rho_A = \sum_i p_i |i\rangle\langle i| .$$

This allows the realization of ρ_A as an *ensemble* of the vector states $|i\rangle$, each prepared with probability p_i. For a density matrix ρ, a quantum error is a **quantum operation** described by a collection of matrices $\{E_a\}$ that satisfies $\sum_a E_a^* E_a = I$ and acts by

$$\rho \mapsto \sum_a E_a \rho E_a^*. \tag{3.1}$$

In terms of the ensemble interpretation of density matrices, this general form of quantum error acts on each $|i\rangle$ in the ensemble by $|i\rangle \mapsto E_a|i\rangle$ with probability $|E_a|i\rangle|^2$.

Warning 3.2. There is ambiguity in the choice of the ensemble $\{|i\rangle\}$ to represent a density matrix, and the operators $\{E_a\}$ to represent a quantum operation.

Note that the above general description of quantum errors, just like density matrices, can be derived from the usual vector state formalism of quantum mechanics. That is, when a quantum system A is part of a product system AB with B interpreted as the environment, errors are caused by unitary evolution of the joint system. If we denote the initial state of A as $|\psi\rangle_A$ and the environment as $|0\rangle_B$, then in terms of an orthonormal basis $|a\rangle_B$ of $\mathcal{H}_B$, a unitary transformation on $\mathcal{H}_A \otimes \mathcal{H}_B$ can be written as

$$|\psi\rangle_A \otimes |0\rangle_B \mapsto \sum_a E_a |\psi\rangle_A \otimes |a\rangle_B \tag{3.3}$$

for some $E_a \in \mathcal{L}(\mathcal{H}_A)$, where unitarity is equivalent to the completeness relation $\sum_a E_a^* E_a = I$. After taking outer product and tracing out B, we recover the previous description 3.1 of quantum operation. Conversely, any quantum operation in the form of 3.1 can be realized as the result of looking only at subsystem A when the joint system evolves as in 3.3. These two descriptions of quantum operations match the physical requirement for a quantum operation to be *completely-positive* and trace-preserving, in a result called Stinespring's dilation theorem.

The operator formalism of quantum states and quantum errors incorporates the simple error model as a linear transformation $|\psi\rangle \mapsto E\,|\psi\rangle$ — if we let $\rho = |\psi\rangle\langle\psi|$ and $\{E_a\}$ to consist of a single operator E. Since a quantum state can be entangled in the sense discussed above, the simple error model is mathematically and physically less general.

3.2. **Error Correction and Detection.** Similar to classical linear codes, a **quantum code** is a subspace $\mathcal{C}$ of the Hilbert space in question. Recall that to correct or detect a classical error, we apply the parity check matrix H to $x + e$. This syndrome measurement of the erroneous state does not change the state. On the other hand, we do not have full access to a quantum state, and quantum measurement necessarily *collapses* the measured state. But this complication is also an asset that allows us to discretize the errors.

Consider first **nondegenerate** quantum codes, whose correction procedure captures the spirit of the general scheme. In this case, it is possible to choose an orthonormal basis $\{|i\rangle\}$ of $\mathcal{C}$ and a collection of operators $\{E_a\}$ such that the error in question has the form 3.1 and

$$\langle i|E_a^* E_b|j\rangle = \delta_{ab}\delta_{ij}.$$

The analog of disjoint error balls in classical error correction for a nondegenerate quantum code $\mathcal{C}$ is the mutually orthogonal *error subspaces* $\mathcal{H}_a = E_a\mathcal{C}$. If any $|\psi\rangle \in \mathcal{C}$ is subject to the error $|\psi\rangle\langle\psi| \mapsto \sum_a E_a\,|\psi\rangle\langle\psi|\,E_a^*$, the orthogonality of $\mathcal{H}_a$ allows us to make a measurement that yields one of the density matrices $E_a\,|\psi\rangle\langle\psi|\,E_a^*$, or equivalently one of the vector states $E_a\,|\psi\rangle$. The measurement result discloses E_a, and error correction is completed by applying E_a^*.

The general criterion for quantum error correction is more subtle. Under the assumption that a quantum error $\mathcal{E}$ and a quantum error recovery procedure $\mathcal{R}$ are both quantum operations, A code $\mathcal{C}$ is said to correct $\mathcal{E}$ if there exists $\mathcal{R}$ such that

$$(\mathcal{R}\circ\mathcal{E})(\rho) = \rho$$

for any ρ that can be formed as an ensemble of states in $\mathcal{C}$.

Remark 3.4. If we consider more general quantum operations and weaken $\sum_a E_a^* E_a = I$ to $\sum_a E_a^* E_a \le I$, then the above characterization of error recovery may also be weakened to $(\mathcal{R}\circ\mathcal{E})(\rho) \propto \rho$.

We have the following key result.

Theorem 3.5 (Quantum error correction criterion)**.** *Let $P_{\mathcal{C}}$ be the projector associated with a quantum code $\mathcal{C}$. Then a quantum operation represented by operators $\{E_a\}$ can be corrected if and only if for some complex numbers ε_{ab},*

$$P_{\mathcal{C}} E_a^* E_b P_{\mathcal{C}} = \varepsilon_{ab}\, P_{\mathcal{C}};$$

equivalently, if and only if for an orthonormal basis $\{|i\rangle\}$ for $\mathcal{C}$,

$$\langle i|E_a^* E_b|j\rangle = \varepsilon_{ab}\,\delta_{ij}. \tag{3.6}$$

Here, ε_{ab} form the entries of a Hermitian matrix.

Proof. The original paper is [6]. See also section 10.3 of [10] and section 7.2 of [11]. □

The criterion for quantum error detection is analogous to the correction criterion, while the detection protocol requires only a measurement.

Theorem 3.7 (Quantum error detection criterion)**.** *Suppose a quantum error is described by a collection of operators $\{E_a\}$. Then a code $\mathcal{C}$ can detect the error if for some $\varepsilon_a \in \mathbb{C}$*

$$P_{\mathcal{C}} E_a P_{\mathcal{C}} = \varepsilon_a P_{\mathcal{C}};$$

or equivalently, for an orthonormal basis $\{|i\rangle\}$ of $\mathcal{C}$,

$$\langle i|E_a|j\rangle = \varepsilon_a \delta_{ij}. \tag{3.8}$$

*The constant ε_a is called the **slope** of the error E_a.*

Proof. Suppose the above criterion is satisfied. For any $|\psi\rangle \in \mathcal{C}$, $E_a|\psi\rangle = \varepsilon_a|\psi\rangle + |\psi_a^\perp\rangle$, where $|\psi_a^\perp\rangle$ is some vector perpendicular to $\mathcal{C}$. Then the error operation is

$$|\psi\rangle \otimes |0\rangle_B \mapsto \sum_a E_a|\psi\rangle \otimes |a\rangle_B = |\psi\rangle \otimes \big(\sum_a \varepsilon_a |a\rangle_B\big) + \sum_a \big(|\psi_a^\perp\rangle \otimes |a\rangle\big). \tag{3.9}$$

Denote by ρ_A the contaminated state of the subsystem A obtained by taking outer product and partial trace of 3.9. Then we may perform a measurement that asks whether ρ_A is supported in $\mathcal{C}$ using the projectors $\{P_\mathcal{C}, I - P_\mathcal{C}\}$. There are two outcomes: either the answer is yes and the post-measurement state is proportional to $P_\mathcal{C}\rho_A P_\mathcal{C} \propto |\psi\rangle\langle\psi|$, equivalently just the vector state $|\psi\rangle$; or the answer is no. In the first case we actually recovered the original state, and in the second case we know an error occurs and discard the erroneous state. □

The similarity of the correction and detection criteria is to our great advantage. For one thing, the discretization of errors mentioned at the beginning of this section follows from the linearity of the above criteria.

Theorem 3.10. *Suppose a quantum error $\mathcal{E}$ described by operators $\{E_a\}$ on a code $\mathcal{C}$ can be corrected (resp. detected), and $\mathcal{F}$ is a quantum error with operation elements $\{F_b\}$, where each F_b is a linear combination of E_a. Then $\mathcal{C}$ can also correct (resp. detect) $\mathcal{F}$.*

Proof. One easily checks that 3.6 and 3.8 both hold for the collection $\{F_b\}$ if they hold for $\{E_a\}$, with different constants ε. □

Remark 3.11. In fact, by Theorem 10.2 of [10], the same recovery operation that corrects $\mathcal{E}$ also corrects $\mathcal{F}$. The detection protocol by definition does not depend on the error.

If a quantum operation $\mathcal{E}$ is described by operators $\{E_a\}$, denote the subspace of matrices spanned by $\{E_a\}$ also as $\mathcal{E}$. As a consequence of the above theorem, to correct (*resp.* detect) any error whose operation elements are supported in $\mathcal{E}$, it suffices to check the error correction (*resp.* detection) criteria for $\{E_a\}$ or any spanning set $\{F_b\}$ of $\mathcal{E}$. In particular, errors $E \in \mathcal{E}$ acting simply by $|\psi\rangle \mapsto E|\psi\rangle$ are special cases of quantum operations that have the form $|\psi\rangle\langle\psi| \mapsto E|\psi\rangle\langle\psi|E^*$, hence can also be corrected (*resp.* detected).

For an n-dimensional Hilbert space $\mathcal{H}_n$, the quantum errors $\mathcal{E}$ we consider are subspaces of $M_n(\mathbb{C})$, with the understanding that we mean quantum operations with operation elements supported in $\mathcal{E}$. Looking back at the criteria 3.6 and 3.8, we see that the slope is best defined with respect to an error space $\mathcal{E}$ and a code $\mathcal{C}$ as a linear functional $\mathcal{E} \to \mathbb{C}$. But linearity implies that we only need to check the criteria against a convenient basis for $\mathcal{E}$ and $\mathcal{C}$, with a finite number of constants ε_a involved.

3.3. **Quantum Metrics and Code Distances.** We are now ready to introduce a notion of distance that plays a similar role as the metric used in the theory of classical codes. The physically most relevant Hilbert space to consider is the n-qubit space $\mathcal{H}_2^{\otimes n}$. It is natural to assume that each qubit experiences a small error $I + \epsilon E_i$, so that the overall error is

$$\bigotimes_{i=1}^n (I + \epsilon E_i) = I + \epsilon\big(E_1 \otimes I^{\otimes(n-1)} + I \otimes E_2 \otimes I^{\otimes(n-2)} + \cdots + I^{\otimes(n-1)} \otimes E_n\big) + O(\epsilon^2).$$

When ϵ is small, we see that the number of non-identity entries in an error operator measures its severity. Since tensor products of the identity and the Pauli matrices $\{I, X, Y, Z\}^{\otimes n}$ span $\mathcal{L}(\mathcal{H}_2^{\otimes n})$, it is customary to restrict attention to these errors, called the **multi-Pauli operators**. Define the **weight** of a Pauli operator as its number of non-identity entries. This definition shares form and motivation with the weight of a classical bit flip error. Thanks to the discretization of errors we saw above, it is natural to observe that by the weight of multi-Pauli operators we actually mean a filtration on the space of errors $\mathcal{L}(\mathcal{H}_2^{\otimes n})$, as noted in [7] and generalized in [9]. Again, we are only concerned with finite dimensions.

Definitions 3.12. Let $\mathcal{H}$ be a finite-dimensional Hilbert space. A **quantum metric** on $\mathcal{L}(\mathcal{H})$ is a filtration *i.e.* a one parameter family of subspaces $\{\mathcal{V}_t\}_{t\geq 0}$ of $\mathcal{L}(\mathcal{H})$ satisfying

(i) $\mathcal{V}_s \subset \mathcal{V}_t$ for $s \leq t$;
(ii) $\mathcal{V}_s\mathcal{V}_t \subset \mathcal{V}_{s+t}$;
(iii) $\mathcal{V}_t^* = \mathcal{V}_t$ for all t;
(iv) $I \in \mathcal{V}_0$.

When we only consider the parameters $t \in \mathbb{N}$ and $\mathcal{V}_t = \mathcal{V}^t$ is the linear span of t-fold products of elements in some subspace $\mathcal{V} \subset \mathcal{L}(\mathcal{H})$, we say $\{\mathcal{V}_t\}$ is a **quantum graph metric**.

The name "graph" is used due to analogy with graph distance. We see that the metric we want to define with multi-Pauli operators is the **quantum Hamming metric** where

$$\begin{aligned}\mathcal{V}_t &= \operatorname{span}\{E_1 \otimes \cdots \otimes E_n \mid E_i \in M_2(\mathbb{C}) \text{ and } E_i \not\propto I \text{ for at most } t \text{ values of } i\}\\ &= \operatorname{span}\{\text{multi-Pauli operators of weight at most } t\}.\end{aligned}$$

Equivalently, the quantum Hamming metric is the quantum graph metric with

$$\mathcal{V} = \operatorname{span}\{\text{multi-Pauli operators of weight at most } 1\}.$$

Code distance is defined in most of the quantum computing literature with respect to the quantum Hamming metric, but it can be easily generalized.

Definition 3.13. Given a Hilbert space $\mathcal{H}$ and an integer-indexed quantum metric $\{\mathcal{V}_t\}_{t\in\mathbb{Z}_{\geq 0}}$ on $\mathcal{L}(\mathcal{H})$, a code $\mathcal{C} \subset \mathcal{H}$ has distance d if for all $E \in \mathcal{V}_{d-1}$, the detection criterion 3.8 is satisfied.

That is, a distance d code can detect errors up to distance $d-1$. Since $\mathcal{V}_t\mathcal{V}_t \subset \mathcal{V}_{d-1}$ for $t = \lfloor \frac{d-1}{2} \rfloor$, from the correction criterion 3.6 we see that a distance d code can correct errors up to distance t. Thus, given a quantum metric, we may talk only about error detection with correction implied.

This language of distance suggests an alternative definition of quantum metric as a function $D : \mathcal{L}(\mathcal{H}) \to [0, \infty]$ satisfying

(i) $D(A + B) \leq \max(D(x), D(y))$;
(ii) $D(\lambda A) \leq D(A)$ for $\lambda \in \mathbb{C}$;
(iii) $D(AB) \leq D(A) + D(B)$;
(iv) $D(A^*) = D(A)$;
(v) $D(I) = 0$.

The two definitions are bijectively corresponded via $\mathcal{V}_t = \{A \in \mathcal{L}(\mathcal{H}) \mid D(A) \leq t\}$ and $D(A) = \min\{t \geq 0 \mid A \in \mathcal{V}_t\}$, but in the quantum case, the metric is on the space of operators.

This difference highlights the fact that in contrast to the classical case where codewords reside in a metric space, in the quantum case code distance can only be defined with respect to error correction and detection. Nevertheless, we have the same statement as in the classical case that a distance d code can detect errors up to distance $d - 1$ and correct errors up to distance $\lfloor \frac{d-1}{2} \rfloor$. What's more, the classical Hamming bound translates to the **quantum Hamming bound** that says that for a *nondegenerate* quantum code $\mathcal{C} \subset \mathcal{H}$, if $\mathcal{C}$ can *correct* $\mathcal{E}$ then

$$\dim \mathcal{E} \cdot \dim \mathcal{C} \leq \dim \mathcal{H}.$$

Here, $\dim \mathcal{E}$ is analogous to the size of an error ball. The bound follows from the definition that for a nondegenerate code $\mathcal{C}$, the error subspaces $E_a\mathcal{C}$ are orthogonal where $\{E_a\}$ is a basis of $\mathcal{E}$. We similarly say that a nondegenerate quantum code is **perfect** if it satisfies the quantum Hamming bound with equality. As for code lower bounds, we will see a quantum code construction motivated by Gilbert's classical lower bound in the following section.

4. A Two-Step Construction for Quantum Codes

4.1. **Detecting One Error.** The formulation of the error detection condition 3.8 suggests that it would be convenient to first consider a basis consisting of Hermitian operators of the error space we wish to correct, and find basis elements for $\mathcal{C}$ in terms of the simultaneous eigenbasis for these Hermitian operators. As an example, consider the simple case of detecting one error on $\mathcal{H} = \mathbb{C}^n$, or more precisely detecting a two-dimensional error space $\mathcal{E}$. Given that $\mathcal{E} = \mathcal{E}^*$, we can consider a basis $\{I, E\}$ of $\mathcal{E}$ where E is Hermitian, and it suffices to check the error detection condition for E.

Label the real eigenvalues of E in order as $\lambda_1 \leq \cdots \leq \lambda_n$, and the corresponding eigenstates as $|1\rangle, \ldots, |n\rangle$. If we pair up the eigenstates and define $|k, l\rangle = \alpha |k\rangle + \beta |l\rangle$ for some coefficients α, β satisfying the normalization condition $|\alpha|^2 + |\beta|^2 = 1$, then the states $|k, l\rangle$ and $|k', l'\rangle$ for disjoint pairs $\{k, l\}$ and $\{k', l'\}$ are orthogonal. Moreover, $E|k, l\rangle$ and $|k', l'\rangle$ are orthogonal, so that if we construct a subspace $\mathcal{C}$ of $\mathbb{C}^n$ with $\{|k, l\rangle\}$ as an orthonormal basis, to satisfy 3.8 it remains only to ensure that $\langle k, l|E|k, l\rangle = \varepsilon$ for some slope ε. That this slope cannot depend on the basis element $|k, l\rangle$ is the difficulty.

In this case it is not difficult after all. One easily calculates that $\langle k, l|E|k, l\rangle = |\alpha|^2 \lambda_k + |\beta|^2 \lambda_l$ lies in the interval $[\lambda_k, \lambda_l]$ if $k < l$. Pick ε as the median of the eigenvalues $\lambda_1, \ldots, \lambda_n$. By choosing symmetric elements $\{k, l\}$ as a pair with $k + l = n + 1$ and appropriate coefficients α and β for each pair, we can ensure that $\langle k, l|E|k, l\rangle = \varepsilon$. When n is odd we may also include $|\lceil n/2 \rceil\rangle$ as a basis element. The dimension of the code $\mathcal{C}$ we find is the number of orthonormal basis elements, which is $\lceil n/2 \rceil$.

Remark 4.1. If the median eigenvalue is degenerate when n is odd, or if at least one of the two middle eigenvalues is degenerate when n is even, it is easy to see that $\mathcal{C}$ can be constructed with dimension larger than $\lceil n/2 \rceil$. The following proposition also applies to these cases after slight modification.

It turns out that this strategy is optimal.

Proposition 4.2. *Suppose all eigenvalues $\lambda_1, \ldots, \lambda_n$ of E are distinct in the above setting. Then a code that detects E can have dimension at most $\lceil n/2 \rceil$.*

Proof. First, we observe that the Hermiticity of E ensures that the associated slope ϵ must be real. The error detection criterion can be equivalently formulated using $F_\varepsilon = E - \varepsilon I$. That is, with respect to the bilinear form defined by F_ε, $\mathcal{C}$ detects E if it is isotropic. Note that F_ε is a real quadratic form diagonal in the eigenbasis $\{|1\rangle, \ldots, |n\rangle\}$, and recall that the signature of a real quadratic form is a triple (n_0, n_+, n_-), where n_0, n_+, n_- respectively denotes the number of zero, positive, negative entries on the diagonal. It is a straightforward linear algebra fact that the dimension of a maximal isotropic subspace is $n_0 + \min\{n_+, n_-\}$.

If we choose $\varepsilon \in (\lambda_k, \lambda_{k+1})$, then the signature of F_ε is $(0, n - k, k)$. If we choose $\varepsilon = \lambda_k$, then the signature is $(1, n - k, k - 1)$. Thus we see that the choice of ε as the median of the eigenvalues is optimal and the dimension of an isotropic subspace is at most $\lceil n/2 \rceil$. □

4.2. **Detecting D Errors.** To find a code that detects a larger error space, a two-step construction is proposed in Theorems 3 and 4 of [7]. First, find an intermediate subspace $\mathcal{B} \subset \mathcal{H}$ such that the error space restricted to $\mathcal{L}(\mathcal{B})$, $\mathcal{F} = P_\mathcal{B} \mathcal{E} P_\mathcal{B}$, generates a commutative algebra. Suppose $\dim \mathcal{B} = m$ and $\dim \mathcal{F} = D + 1$; we need to detect D linearly independent errors in addition to the identity. Since $\mathcal{E} = \mathcal{E}^*$ by assumption, we have $\mathcal{F} = \mathcal{F}^*$ and it is possible to find a simultaneous eigenbasis of $\mathcal{B}$ with respect to which $\mathcal{F}$ acts diagonally. Label the simultaneous eigenstates as $|1\rangle, \ldots, |m\rangle$, and write $\mathcal{F} = \mathrm{span}\{F_0 = I, F_1, \ldots, F_D\}$ where each F_j is Hermitian. Let

$$\vec{F} = (F_1, \ldots, F_D),$$

so that $\vec{F}|k\rangle = \vec{\lambda}_k |k\rangle$ for some $\vec{\lambda}_k \in \mathbb{R}^D$.

Remark 4.3. The use of the letter D here differs by one from the same letter in [7]. The result of Theorem 4 in [7] can be slightly improved because errors proportional to the identity automatically satisfy the detection criterion.

The second step in this construction is a generalization of the previous example. Consider a partition of $\{1,\dots,m\}$ into disjoint subsets Y_i. For each Y_i define

$$|\psi_i\rangle = \sum_{k\in Y_i} \sqrt{\beta_i^k}\,|k\rangle\,,$$

where the nonnegative coefficients β_i^k satisfying the normalization condition $\sum_{k\in Y_i}\beta_i^k = 1$ are to be chosen later. The states $|\psi_i\rangle$ are orthonormal, and $\langle\psi_i|\vec{F}|\psi_j\rangle = 0$ if $i \neq j$. The remaining challenge to satisfy the detection criterion is to find a generalized slope $\vec{\varepsilon}\in\mathbb{R}^D$ such that

$$\langle\psi_i|\vec{F}|\psi_i\rangle = \vec{\varepsilon} \quad \forall i. \tag{4.4}$$

Note that $\langle\psi_i|\vec{F}|\psi_i\rangle = \sum_{k\in Y_i}\beta_i^k\vec{\lambda}_k$ is a convex sum and can take any value in the convex hull spanned by $\{\vec{\lambda}_k\}_{k\in Y_i}$ with appropriate choice of the coefficients β_i^k. Thus 4.4 is possible if and only if the convex hulls spanned by $\{\vec{\lambda}_k\}_{k\in Y_i}$ for all i have a common intersection. To maximize the dimension of a code $\mathcal{C}$ with an orthonormal basis of the form $|\psi_i\rangle$ is to maximize the size of the partition $\{Y_i\}$ of $\{1,\dots,m\}$ such that this convex geometrical property holds. This problem can in general be solved thanks to the result in [12].

Theorem 4.5 (Tverberg's theorem)**.** *Given m points in $\mathbb{R}^D$, a partition of the m points into r disjoint subsets with intersecting convex hulls is possible if*

$$m > (D+1)(r-1)$$

Remark 4.6. For a general configuration of points, Tverberg's theorem is optimal just by counting the number of variables and constraints. Here we relax the condition $\beta_i^k \geq 0$ to $\beta_i^k \in \mathbb{R}$, and all other variables and equations are over the reals. Let the configuration of points $\{\vec{\lambda}_k\}$ be given. We have D degrees of freedom in choosing a point $\vec{\varepsilon}$ in the common intersection of the convex hulls, and we need to make another m choices of the coefficients β_i^k. The linear constraints include r equations $\sum_{k\in Y_i}\beta_i^k = 1$ and rD equations $\sum_{k\in Y_i}\beta_i^k\vec{\lambda}_k = 1$. Assuming that the linear system has full rank for a general point configuration $\{\vec{\lambda}_k\}$, the system is over determined if $m + D < r + rD$; equivalently if $m \leq (D+1)(r-1)$.

As a result of Tverberg's theorem, if the intermediate space $\mathcal{B}$ in this two-step construction has dimension m and the intermediate error space $\mathcal{F}$ has dimension $D+1$, then there exists a code $\mathcal{C}$ detecting $\mathcal{E}$ such that

$$\dim\mathcal{C} = \lceil\frac{m}{D+1}\rceil$$

Since we will use this construction extensively, three caveats are in order. First, unlike in the special case of detecting one error, when the error space has higher dimension, other code constructions might beat this two-step strategy in maximizing code dimension.

Second, maximizing the dimension of $\mathcal{C}$ depends on maximizing the dimension of the intermediate space $\mathcal{B}$. In [7], the authors are concerned with a general error space $\mathcal{E}$. They use a greedy construction for $\mathcal{B}$ in the spirit of the constructive Gilbert bound for not necessarily linear classical codes (2.2). In more detail, this construction crudely takes $\mathcal{F} = \mathcal{E}$ and $D+1 = \dim\mathcal{E}$. To ensure that $\mathcal{E}$ acting on $\mathcal{B}$ is commutative, we want a basis $\{|k\rangle\}$ of $\mathcal{B}$ with respect to which $E_0 = I, E_1, \dots, E_D$ are diagonal. Suppose we have already chosen $|1\rangle,\dots,|k\rangle$ acted diagonally by $E_0,\dots,E_D$. As long as $k(D+1) < \dim\mathcal{H}$, we may choose another basis element $|k+1\rangle$ in the orthogonal complement of $E_a|l\rangle$, $a = 0,\dots,D$, $l = 1,\dots,k$. Thus the resulting dimension of $\mathcal{B}$ is at least $\lceil\frac{\dim\mathcal{H}}{\dim\mathcal{E}}\rceil$. We will use the idea of this greedy construction in finding the intermediate space $\mathcal{B}$, but we can often do better than $\lceil\frac{\dim\mathcal{H}}{\dim\mathcal{E}}\rceil$ because the error space $\mathcal{F}$ restricted to $\mathcal{B}$ can be smaller than $\mathcal{E}$.

Third, when the generalized eigenvalues $\{\vec{\lambda}_k\}$ are placed in some special positions in $\mathbb{R}^D$, it is possible to partition them into more disjoint subsets with intersecting convex hulls than promised by Tverberg's theorem. When such a partition into r disjoint subsets exists for $r > \lceil \frac{m}{D+1} \rceil$, we have a **super-Tverberg** behavior, and a point in the common intersection of the convex hulls is called a **super-Tverberg point**.

4.3. **The Two-Step Construction of CSS Codes.** Although this point is only implicit in most discussions of the CSS Codes, the two-step construction of a CSS code is exactly an example of the two-step construction we discussed above. For an introduction to CSS codes and stabilizer codes, see Chapter 10 in [10] and Chapter 7 of [11], or any other relevant physical text on quantum codes.

Example 4.7. Consider the subspace of the 4-qubit Hilbert space stabilized by $Z^{\otimes 4}$ and $X^{\otimes 4}$. The intermediate subspace $\mathcal{B}$ stabilized by $Z^{\otimes 4}$ has basis elements associated to length 4 bit strings of even weight. Explicitly, they are

$$|0000\rangle, |1111\rangle, |1100\rangle, |0011\rangle, |1010\rangle, |0101\rangle, |1001\rangle, |0110\rangle.$$

Since $Y = iZX$ is a composition of a bit flip and a phase flip, we see that any weight one multi-Pauli that involves X or Y maps a basis element above to $|x\rangle$ where x is a bit string of odd weight, hence $|x\rangle$ is orthogonal to $\mathcal{B}$. That is, if the error space $\mathcal{E}$ we wish to detect is $\mathcal{V}_1$ in the quantum Hamming metric, then the weight one multi-Pauli elements involving X and Y vanish in the restricted error space $\mathcal{F} = P_{\mathcal{B}}\mathcal{E}P_{\mathcal{B}}$. Therefore, $\mathcal{F}$ is spanned by the identity and weight one multi-Pauli elements involving only Z and I, such as $Z \otimes I \otimes I \otimes I$, which all commute.

In the second step where we take $\mathcal{C}$ to be the subspace of $\mathcal{B}$ stabilized by $X^{\otimes 4}$, we cut the dimension further by half, as usual when taking simultaneous eigenstates of stabilizers; this is a super-Tverberg behavior. Here $m = 8$ and $D = \dim \mathcal{F} - 1 = 4$, so Tverberg's theorem guarantees a code of dimension $\lceil \frac{8}{4+1} \rceil = 2$. But we actually have a 4-dimensional code $\mathcal{C}$ with basis elements

$$|0000\rangle + |1111\rangle, |1100\rangle + |0011\rangle, |1010\rangle + |0101\rangle, |1001\rangle + |0110\rangle.$$

Since each basis element of $\mathcal{C}$ is mapped outside $\mathcal{C}$ by each basis element of $\mathcal{F}$, we see that the corresponding super-Tverberg point is $\vec{\varepsilon} = \vec{0}$.

This example illustrates how general CSS codes align with our scheme. The CSS codes are constructed from classical codes. Given the n-qubit Hilbert space $\mathcal{H}_2^{\otimes n}$, suppose we have two classical codes C_Z and C_X that are subspaces of $(\mathbb{Z}/2)^n$ satisfying $C_X^\perp \subset C_Z$ and $d(C_Z), d(C_X) \geq d$. The CSS code constructed from C_Z and C_X is a quantum Hamming distance d quantum code that associates a codeword to every coset in $C_Z/C_X^\perp$. Equivalently, a CSS code associated with C_Z and C_X is a special subclass of the stabilizer codes where the stabilizers consist of elements in the n-qubit Pauli group that is either a tensor product of Z's and I's or a tensor product of X's and I's. For a length n bit string e, let $\bar{Z}_e$ denote the multi-Pauli whose k-th component is Z if the k-th bit of e is 1, and I if the k-th bit of e is 0; define $\bar{X}_e$ similarly. If H_Z and H_X are respectively the parity check matrices of C_Z and C_X, then the stabilizers of the associated CSS code are $\bar{Z}_e$ where e are the rows of H_Z and $\bar{X}_e$ where e are the rows of H_X.

In the first step, consider the subspace stabilized by the $\bar{Z}_e$'s. Since each Z acts by flipping phase, $|x\rangle$ where $x \in (\mathbb{Z}/2)^n$ is stabilized by $\bar{Z}_e$ if and only if $x \cdot e = 0$. As a result, the intermediate subspace is $\mathcal{B} = \mathbb{C}[C_Z]$, the subspace spanned by $|x\rangle$ for $x \in C_Z$. Suppose a multi-Pauli $\bar{E}$ consists only of I, X, Y and has weight less than $d = d(C_Z)$. Then $\bar{E} \propto \bar{X}_e$ where e has weight less than d, and $\bar{E}|x\rangle \propto |x + e\rangle$ for $x \in C_Z$. Since C_Z has distance d, $x + e \notin C_Z$, hence $|x + e\rangle \notin \mathbb{C}[C_Z]$. This shows that the restricted error space $\mathcal{F}$ is spanned by multi-Pauli elements involving only Z and I with weight less than d, which all commute.

In the second step, we need to detect errors of the form $\bar{Z}_e$ where e has weight less than d. Since $d(C_X) \geq d$, there exists a row e' of H_X such that $e \cdot e' \neq 0$; otherwise e would lie inside C_X. Since Z and X anticommute, this implies that $\bar{Z}_e$ and $\bar{X}_{e'}$ anticommute, so one knows that $\bar{Z}_e$ can be detected by the general formalism of stabilizer codes. In fact, the slope or the super-Tverberg point

in the second step of CSS code construction is always the zero vector: since any $|\psi\rangle$ in the resulting CSS code is stabilized by $\bar{X}_{e'}$,

$$\langle\psi|\bar{Z}_e|\psi\rangle = \langle\psi|\bar{Z}_e\bar{X}_{e'}|\psi\rangle = -\langle\psi|\bar{X}_{e'}\bar{Z}_e|\psi\rangle = -\langle\psi|\bar{Z}_e|\psi\rangle,$$

so $\langle\psi|\bar{Z}_e|\psi\rangle = 0$. We have shown that the two-step construction is a generalization of a well-known quantum code construction, and we proceed to apply this generalized method in geometries arising from Lie algebra representations.

5. Quantum Error Detection in Irreducible Representations of $\mathfrak{sl}_n\mathbb{C}$

5.1. **Lie Type Graph Metric Spaces.** The preceding discussions already suggested that from the mathematical point of view, finding quantum codes that can correct or detect errors only requires a Hilbert space equipped with a quantum metric, and while the n-qubit quantum Hamming space is both relatively simple and extremely rich, there is no need to restrict attention to these quantum metric spaces.

Here we are interested in the cases where the Hilbert space in question is a representation of a matrix Lie algebra, equipped with a **Lie type graph metric**. Assume that $\mathfrak{g} = \mathfrak{g}_{\mathbb{C}}$ is a complex Lie algebra of traceless matrices; then the Lie type graph metric for a representation $d\rho$ of $\mathfrak{g}$ is the quantum graph metric defined by $\mathcal{V}_1 = d\rho(\mathfrak{g}) \oplus \mathbb{C}I$ and $\mathcal{V}_t = \mathrm{span}(\mathcal{V}_1)^t$. In fact, we only consider $\mathfrak{g} = \mathfrak{sl}_n\mathbb{C}$, and we call the associated Lie type graph metric spaces **$\mathfrak{sl}_n\mathbb{C}$-metric spaces**. This setting sounds abstract, but in practice, methods of code construction we find remain elementary. Thus we will use many Lie theoretical facts without proof, with details to be found in [3] and [5].

A Lie algebra representation only specifies a module structure on a vector space, and we need to impose an inner product. This is done as follows. Given an n-dimensional Lie algebra $\mathfrak{g}$, $\mathfrak{g}_{\mathbb{R}} \coloneqq \mathfrak{g} \cap \mathfrak{u}_n$ is also a Lie algebra. Here $\mathfrak{g}_{\mathbb{R}}$ is called a **real form** of $\mathfrak{g}$, for $\mathfrak{g} = \mathbb{C} \otimes_{\mathbb{R}} \mathfrak{g}_{\mathbb{R}}$. This identity holds because $\mathfrak{u}_n$ consists of anti-Hermitian $n \times n$ matrices and any matrix $E \in M_n(\mathbb{C})$ can be written as $E = X + iY$ for a pair $X, Y \in \mathfrak{u}_n$:

$$E = \frac{E - E^*}{2} + i\frac{E + E^*}{2i}.$$

Observe that there is a one-to-one correspondence between representations of $\mathfrak{g}$ and $\mathfrak{g}_{\mathbb{R}}$. Although real forms in general are not unique, this particular real form $\mathfrak{g}_{\mathbb{R}}$ is canonical in that as $U(n)$ is compact, $G \coloneqq e^{\mathfrak{g}_{\mathbb{R}}} \subset U(n)$ is a compact Lie group. Assume that the Lie algebra representation $d\rho$ of $\mathfrak{g}_{\mathbb{R}}$ one gets by restricting the representation of $\mathfrak{g}$ is the differential of a Lie group representation ρ of G. Then by the compactness of G, we choose the inner product such that the action of G is unitary. Under such an inner product, unitarity of ρ implies that for all $X \in \mathfrak{g}_{\mathbb{R}}$,

$$(e^{td\rho(X)})^* = \rho(e^{tX})^* = \rho(e^{tX})^{-1} = e^{-td\rho(X)} \qquad \forall t \in \mathbb{R},$$

so that after differentiation at $t = 0$, $d\rho(X)^* = -d\rho(X)$. Since $X = -X^*$ by the definition of $\mathfrak{g}_{\mathbb{R}}$, we get the requirement

$$d\rho(X)^* = d\rho(X^*). \tag{5.1}$$

Since here $\mathfrak{g} = \mathfrak{sl}_n\mathbb{C}$, we have $\mathfrak{g}_{\mathbb{R}} = \mathfrak{su}_n$. Given a representation V of $\mathfrak{sl}_n\mathbb{C}$, we shall subsequently choose an inner product satisfying 5.1 such that the action of $SU(n)$ on V is unitary.

5.2. **Irreducible Representations of $\mathfrak{sl}_n\mathbb{C}$.** We will be interested in the following class of irreducible representations of $\mathfrak{sl}_n\mathbb{C}$. We start with $\mathfrak{sl}_2\mathbb{C}$, which has a canonical basis

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad E = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad F = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Let $\mathcal{H}_N = \mathbb{C}[x, y]_N$ be the space of degree N homogeneous polynomials in two variables, where the action of H, E, F are given by

$$h \coloneqq d\rho(H) = x\partial_x - y\partial_y \quad e \coloneqq d\rho(E) = x\partial_y \quad f \coloneqq d\rho(F) = y\partial_x.$$

Up to an isomorphism switching $(x, y) \mapsto (y, -x)$, this action restricted to $\mathfrak{su}_2$ can be checked to be the differential of the action of $SU(2)$ on $\mathbb{C}[x, y]_N$ by

$$(\rho(g)\phi)(x, y) = \phi(g^{-1}\begin{pmatrix} x \\ y \end{pmatrix}) \qquad \text{for } g \in SU(2), \phi \in \mathbb{C}[x, y]_N,$$

where the action of g^{-1} is simply by 2×2 matrix multiplication. In honor of the inspiration from physics and simply for convenience, we keep the Dirac notation and write $x^k y^{N-k}$ as $\left|x^k y^{N-k}\right\rangle$. The unitarity condition 5.1 is satisfied if we define an inner product such that $\langle x^k y^{N-k} | x^{N-l} y^l \rangle = \delta_{kl} / \binom{N}{k}$, as will be shown more generally below.

It is well-known that the representations above indexed by N are exactly all the irreducible representations of $\mathfrak{sl}_2\mathbb{C}$. For $\mathfrak{sl}_n\mathbb{C}$ where $n \geq 3$, there is also a similar class of irreducible representations on the space of degree N homogeneous polynomials in n variables, denoted $\mathbb{C}[x_1, \ldots, x_n]_N$; but there exist other irreducible representations. Let 1_{ij} denote the $n \times n$ matrix with 1 in the ij entry and 0 elsewhere, and let H_{ij} be the $n \times n$ traceless diagonal matrix $1_{ii} - 1_{jj}$. Then the action of $\mathfrak{sl}_n\mathbb{C}$ is given by

$$h_{ij} \coloneqq d\rho(H_{ij}) = x_i\partial_i - x_j\partial_j \quad e_{ij} \coloneqq d\rho(1_{ij}) = x_i\partial_j \quad f_{ji} \coloneqq d\rho(1_{ji}) = x_j\partial_i \quad i < j.$$

The inner product we choose is such that distinct monomials are orthogonal and

$$\langle x_1^{k_1} \cdots x_n^{k_n} | x_1^{k_1} \cdots x_n^{k_n} \rangle = 1 \Big/ \binom{N}{k_1, \ldots, k_n}.$$

It is easily checked that the unitarity condition is satisfied for the canonical basis elements; we only need to observe that h_{ij} acts diagonally and

$$\begin{aligned} &\left\langle x_1^{k_1} \cdots x_i^{k_i+1} \cdots x_j^{k_j-1} \cdots x_n^{k_n} \right| \left(e_{ij} \left| x_1^{k_1} \cdots x_i^{k_i} \cdots x_j^{k_j} \cdots x_n^{k_n} \right\rangle \right) \\ =&\left(f_{ji} \left| x_1^{k_1} \cdots x_i^{k_i+1} \cdots x_j^{k_j-1} \cdots x_n^{k_n} \right\rangle \right)^* \left| x_1^{k_1} \cdots x_i^{k_i} \cdots x_j^{k_j} \cdots x_n^{k_n} \right\rangle = \frac{k_1! \cdots (k_i+1)! \cdots k_j! \cdots k_n!}{N!}. \end{aligned}$$

This class of irreducible representations can be understood geometrically. Define

$$|k_1 \cdots k_n\rangle \coloneqq \binom{N}{k_1, \ldots, k_n}^{1/2} \left| x_1^{k_1} \cdots x_n^{k_n} \right\rangle, \tag{5.2}$$

so that $|k_1 \cdots k_n\rangle$ form an orthonormal basis of $\mathbb{C}[x_1, \ldots, x_n]_N$. In this notation, Figure 2 completely describes the action of $\mathfrak{sl}_3, \mathbb{C}$ on $\mathbb{C}[x_1, x_2, x_3]_3$. For instance, the edge connecting $|111\rangle$ and $|210\rangle$ indicates that $e_{13}|111\rangle = x_1\partial_3\sqrt{3}|x_1x_2x_3\rangle = \sqrt{3}\left|x_1^2x_2\right\rangle = \sqrt{2}|210\rangle$, and similarly $f_{31}|210\rangle = \sqrt{2}|111\rangle$. To better understand the meaning of Figure 2, we need a more abstract description of these representations on homogeneous polynomials.

Let $\mathfrak{sl}_n\mathbb{C}$ act on the n-dimensional vector space V simply by matrix multiplication; V is called the **defining representation** of $\mathfrak{sl}_n\mathbb{C}$. It is straightforward to observe that the representation $\mathbb{C}[x_1, \ldots, x_n]_N$ is isomorphic to the representation $\mathrm{Sym}^N(V)$, where each x_i corresponds to the standard basis element 1_i of V, and the coefficient of differentiation corresponds to the coefficient derived from the product rule when $\mathfrak{sl}_n\mathbb{C}$ acts on $\mathrm{Sym}^N(V)$. In either description, it is clear that the subspace of traceless diagonal elements in $\mathfrak{sl}_n\mathbb{C}$, denoted $\mathfrak{h}$, remains diagonal in the representations. The monomials in $\mathbb{C}[x_1, \ldots, x_n]_N$, equivalently the basis of $\mathrm{Sym}^N(V)$ obtained by equivalence classes of tensor products of the standard basis elements of V, form a simultaneous eigenbasis of $\mathfrak{h}$.

The diagonal elements in $\mathfrak{h}$ play a crucial role in subsequent analysis. First, consider the case of $\mathfrak{sl}_2\mathbb{C}$, where the commutation relations satisfied by H, E, F are also satisfied by h, e, f that act on a given representation $\mathcal{H}$, as Lie algebra representations preserve the Lie bracket:

$$[h, e] = 2e \quad [h, f] = -2f \quad [e, f] = h.$$

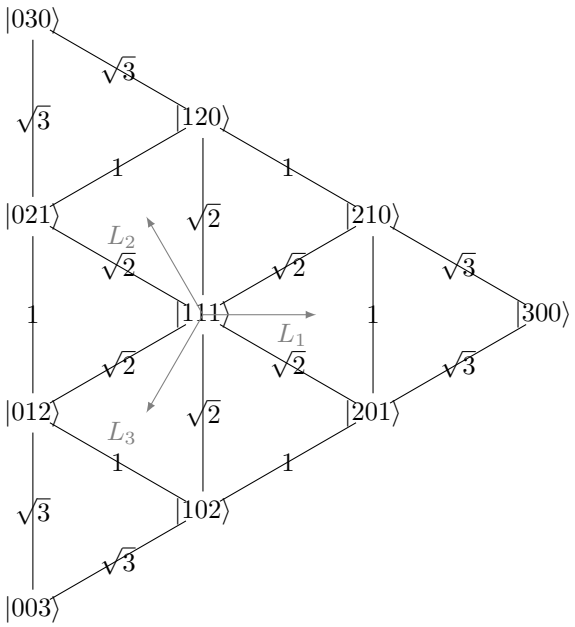

FIGURE 2. Weight diagram for $\mathbb{C}[x_1, x_2, x_3]_3$ as an irreducible representation of $\mathfrak{sl}_3\mathbb{C}$; equivalently for $\mathrm{Sym}^3(V)$ where V is the defining representation of $\mathfrak{sl}_3\mathbb{C}$. Indicated on the edges are the coefficients of the translating actions of $1_{ij},\ i \neq j$.

Using these relations we obtain the key fact that e and f permute the eigenspaces of h and respectively raises or lowers the eigenvalue. If $hv = \alpha \cdot v$ for some $v \in V$ and $\alpha \in \mathbb{C}$, then

$$h(e(v)) = (eh + [h, e])v = (\alpha + 2) \cdot ev \qquad h(f(v)) = (fh + [h, f])v = (\alpha - 2) \cdot fv. \tag{5.3}$$

At some point the action of e will annihilate a **highest weight** eigenvector v, for instance when $e = x\partial_y$ acts on $x^N \in \mathbb{C}[x, y]_N$; similarly $f = y\partial_x$ annihilates y^N.

The general case works as follows. The space of traceless diagonal elements $\mathfrak{h}$ act on $\mathfrak{g} = \mathfrak{sl}_n\mathbb{C}$ by restricting the adjoint representation of $\mathfrak{g}$ on itself. That is, $h \in \mathfrak{h}$ acts by $[h, -]$. Since elements in $\mathfrak{h}$ commute, this action splits $\mathfrak{g}$ into eigenspaces $\mathfrak{g} = \mathfrak{h} \oplus \left(\bigoplus_\alpha \mathfrak{g}_\alpha\right)$, where each α is a generalized eigenvalue *i.e.* an element in $\mathfrak{h}^*$ such that $[h, v] = \alpha(h) \cdot v$ for $v \in \mathfrak{g}_\alpha$. The elements α are called **roots** of $\mathfrak{g}$ and the eigenspaces $\mathfrak{g}_\alpha$ **root spaces**.

Given a representation V of $\mathfrak{g}$, the action of $\mathfrak{h}$ also splits V into eigenspaces $V = \bigoplus_\beta V_\beta$. Here the linear functionals β are called the **weights** of the representation V and V_β the **weight spaces**. If $e \in \mathfrak{g}_\alpha$ and $v \in V_\beta$ for some $\alpha, \beta \in \mathfrak{h}^*$, then for $h \in \mathfrak{h}$ we have the key computation

$$h(e(v)) = ([h, e] + eh)v = (\alpha(h) + \beta(h)) \cdot e(v).$$

That is, elements in the root space $\mathfrak{g}_\alpha$ map elements in the weight space V_β to the weight space $V_{\alpha+\beta}$; this action is often visualized in $\mathfrak{h}^*$ as a translation of α by β.

For $\mathfrak{g} = \mathfrak{sl}_n\mathbb{C}$, observe that if $D = (d_i)$ is a diagonal matrix and $M = (m_{ij})$ is another matrix, then the ij entry of $[D, M]$ is $(d_i - d_j)m_{ij}$. From this we find that the decomposition of $\mathfrak{sl}_n\mathbb{C}$ into root spaces is

$$\mathfrak{sl}_n\mathbb{C} = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathfrak{g}_{L_i - L_j},$$

where $\{L_i\}$ is the restriction to $\mathfrak{h}$ of the dual to the standard diagonal basis, with $L_i(1_{jj}) = \delta_{ij}$. Each $\mathfrak{g}_{L_i - L_j}$ is the linear span of 1_{ij}.

Now we restrict attention to the representations $\mathrm{Sym}^N(V)$ of $\mathfrak{sl}_n\mathbb{C}$ discussed above. Note that the weights of the defining representation V are just $\{L_i\}_{i=1,\dots,n}$ with weight vectors the standard basis of V, so that by the product rule, the weights of $\mathrm{Sym}^N(V)$ are the N-th symmetric powers of $\{L_i\}_{i=1,\dots,n}$. Since $\mathfrak{sl}_n\mathbb{C}$ is traceless, $\{L_i\}_{i=1,\dots n}$ satisfy $\sum_{i=1}^n L_i = 0$, so we can represent $\{L_i\}_{i=1,\dots n}$ in $\mathfrak{h}^*$ as the position vectors of a regular $(n-1)$-simplex centered at the origin. Figure 2 illustrates the case $n = 3$. In this **weight diagram**, each node indicates a weight space. The action of a root vector proportional to 1_{ij} in $\mathfrak{g}_{L_i - L_j}$ translates a weight vector by $L_i - L_j$ or annihilates it. For instance, 1_{13} act by $\mathrm{span}\{|012\rangle\} \mapsto \mathrm{span}\{|111\rangle\} \mapsto \mathrm{span}\{|210\rangle\} \mapsto 0$.

The weight diagram can be made into a graph, where there is an edge between two vertices representing weight spaces if the action of some element in $\mathfrak{g}\backslash\mathfrak{h}$ translates one weight space to the other. From this description of the errors, we see that if we use the two-step construction described in the previous section, then the first step can be done readily. That is, since each basis error translates a weight space to another within graph distance ≤ 1, the image of a weight vector under errors in $\mathcal{V}_t = \mathfrak{g}^t$ cannot have support on a weight space of graph distance more than t away. The problem of finding an intermediate subspace reduces to the classical problem of finding a minimum distance set in the weight diagram equipped with the graph metric. We state this result for $\mathfrak{g} = \mathfrak{sl}_n\mathbb{C}$, although the above argument also works for other Lie type graph metric spaces; see Lemma 3.2.1 in [1].

Proposition 5.4. *In the $\mathfrak{sl}_n\mathbb{C}$ metric spaces $\mathrm{Sym}^N(V)$ where V is the defining representation of $\mathfrak{sl}_n\mathbb{C}$, let $\mathcal{B}$ be the subspace spanned by weight vectors associated to a minimum distance d set in the weight diagram. Then the error set $\mathcal{F} = P_{\mathcal{B}}\mathcal{V}_{d-1}P_{\mathcal{B}}$ is a commutative algebra.*

6. Quantum Codes in Symmetric Power Representations of $\mathfrak{sl}_n\mathbb{C}$

We apply the recipe introduced in the preceding sections to construct codes in symmetric power representations of $\mathfrak{sl}_n\mathbb{C}$. In the case of $\mathfrak{sl}_2\mathbb{C}$, we also discuss another possible, in fact in this case better code construction that plays on the canonical inner product we choose to make the representation unitary. It is possible to consider generalizations of this strategy to higher n, but we do not work in this direction.

As before, we equip a relevant representation with the canonical inner product and call the resulting Hilbert space $\mathcal{H}$. We are mostly interested in the asymptotic behavior of code dimension as $\dim\mathcal{H} \to \infty$, but we also consider some interesting special cases when $\dim\mathcal{H}$ is small.

6.1. Codes in String Representations of $\mathfrak{sl}_2\mathbb{C}$.

Consider the case $n = 2$. Note that the orthonormal eigenstates we found in 5.2 are also eigenstates of h, and we can instead label them by eigenvalues as $|-N\rangle, |-N+2\rangle, \dots, |N-2\rangle, |N\rangle$. This representation on $\mathbb{C}[x,y]_N$ is nicknamed a string representation because nodes in the weight diagram form an arithmetic progression, and e and f map nodes to their neighbors as observed in 5.3. This new labeling is related to the old one by

$$|2k - N\rangle = \binom{N}{k}^{1/2} \left|x^k y^{N-k}\right\rangle$$

In the new labelling, we find that if $e\,|l\rangle = A\,|l+2\rangle$ for some coefficient A, then $e\,|-l-2\rangle = A\,|-l\rangle$, $f\,|l+2\rangle = A\,|l\rangle$, and $f\,|-l\rangle = A\,|-l-2\rangle$; see Figure 3. More explicitly, we can calculate that

$$\begin{aligned} e\,|l\rangle &= \frac{1}{2}\sqrt{(N-l)(N+l+2)}\,|l+2\rangle \qquad l \neq N; \\ f\,|l\rangle &= \frac{1}{2}\sqrt{(N+l)(N-l+2)}\,|l-2\rangle \qquad l \neq -N. \end{aligned} \tag{6.1}$$

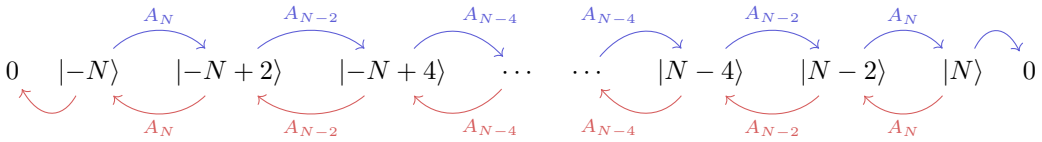

FIGURE 3. The string representation of $\mathfrak{sl}_2\mathbb{C}$ on $\mathbb{C}[x,y]_N$. The action of e is shown in blue and f in red.

FIGURE 4. A distance 2 code construction. The filled shapes indicate a basis for the intermediate space $\mathcal{B}$; each combination of filled shape and color indicates the support of an orthogonal basis for this code.

6.1.1. *$d = 2$.* We wish to detect errors in $\mathcal{E} = \mathcal{V}_1$. Using the two-step construction, the intermediate subspace $\mathcal{B}$ can be chosen to be spanned by eigenstates that are a graph distance 2 apart in Figure 3, so $\dim \mathcal{B} = \frac{\dim \mathcal{H}}{2} + O(1)$. The remaining error space $\mathcal{F} = P_{\mathcal{B}}\mathcal{E}P_{\mathcal{B}}$ is spanned by I and h, so using the strategy in section 4.1, we can achieve $\dim \mathcal{C} = \frac{\dim \mathcal{B}}{2} + O(1) = \frac{\dim \mathcal{H}}{4} + O(1)$.

The two-step strategy is not optimal in this case. Instead of finding an intermediate subspace $\mathcal{B}$ where the remaining error space $\mathcal{F} = P_{\mathcal{B}}\mathcal{E}P_{\mathcal{B}}$ is diagonal, the following strategy seeks to find $\mathcal{B}$ such that $\mathcal{F}$ is block diagonal — in this case in two by two blocks. This construction was found by Rui Okada at UC Davis.

The idea is to use the symmetry of coefficients indicated in Figure 3. Let

$$|\psi_1\rangle = \sqrt{N-2}\,|-N\rangle + \sqrt{N}\,|N-2\rangle \qquad |\phi_1\rangle = \sqrt{N}\,|-N+2\rangle - \sqrt{N-2}\,|N\rangle\,.$$

Note that e takes the support of $|\psi_1\rangle$ to the support of $|\phi_1\rangle$, and we easily calculate that $\langle\phi_1|e|\psi_1\rangle = A_N\sqrt{N(N-2)} - A_N\sqrt{N(N-2)} = 0$. Dually, f takes the support of $|\phi_1\rangle$ to the support of $|\psi_1\rangle$, and taking the adjoint of $\langle\phi_1|e|\psi_1\rangle$ we get $\langle\psi_1|f|\phi_1\rangle = 0$. Further, we see that $\langle\psi_1|h|\psi_1\rangle = \langle\phi_1|h|\phi_1\rangle = 0$.

Going on with the code construction, we do not use $|-N+4\rangle$ or $|N-4\rangle$ to construct the code, and we imitate the construction of $|\psi_1\rangle$ and $|\phi_1\rangle$ by taking

$$|\psi_2\rangle = \sqrt{N-8}\,|-N+6\rangle + \sqrt{N-6}\,|N-8\rangle \qquad |\phi_2\rangle = \sqrt{N-6}\,|-N+8\rangle - \sqrt{N-8}\,|N-6\rangle\,,$$

so that we still have $\langle\phi_2|e|\psi_2\rangle = \langle\psi_2|f|\phi_2\rangle = \langle\psi_2|h|\psi_2\rangle = \langle\phi_2|h|\phi_2\rangle = 0$. This strategy is schematically depicted in Figure 4. Let $\mathcal{C}$ be the subspace spanned by the orthogonal basis elements $\{|\psi_i\rangle, |\phi_i\rangle\}$. Then the error detection criterion 3.8 is satisfied with slope 0 for e, h, f, hence for $\mathcal{E}$. It is clear from Figure 4 that this construction yields a code $\mathcal{C}$ of dimension $\dim \mathcal{C} = \frac{\dim \mathcal{H}}{3} + O(1)$, beating the usual two-step strategy.

Using Proposition 4.2 and taking $E = h$, we see that an upper bound for code dimension is $\frac{\dim \mathcal{H}}{2}$. It is not clear if it is possible to close the gap between the lower bound $\frac{\dim \mathcal{H}}{3}$ and the upper bound $\frac{\dim \mathcal{H}}{2}$. The upper bound obtained by Okada using more sophisticated linear programming methods is also $\frac{\dim \mathcal{H}}{2}$. Here we show that if a code can do better than Okada's construction described above, it must have a somewhat complicated description in terms of the eigenbasis for h.

Proposition 6.2. *The construction above yielding $\dim \mathcal{C} = \frac{\dim \mathcal{H}}{3}$ is asymptotically optimal among distance 2 codes $\mathcal{C}$ equipped with an orthonormal basis where the basis elements have disjoint support on the eigenbasis for h.*

Proof. First, we show that it is not possible for the support of three basis elements of $\mathcal{C}$ to have the pattern shown in Figure 5. Assuming the contrary, we can adjust the phases and write the three

basis elements as

$$|\psi_-\rangle = a_- |k-2\rangle + e^{i\theta_-} b_- |l-2\rangle \quad |\psi_0\rangle = a_0 |k\rangle + e^{i\theta_0} b_0 |l\rangle \quad |\psi_+\rangle = a_+ |k+2\rangle + e^{i\theta_+} b_+ |l+2\rangle$$

for some relative phases θ_j and $a_j, b_j \in \mathbb{R}_{>0}$ satisfying the normalization conditions $a_j^2 + b_j^2 = 1$ for $j = -, 0, +$. Then, the slope of h with respect to $|\psi_j\rangle$ is

$$\langle\psi_0|h|\psi_0\rangle = a_0^2 k + (1-a_0^2) l \qquad \langle\psi_\pm|h|\psi_\pm\rangle = a_\pm^2 (k \pm 2) + (1 - a_\pm^2)(l \pm 2). \tag{6.3}$$

For this to equal a prescribed slope ε_h of h, ε_h must lie in $(k+2, l-2)$, and a_j and b_j are already determined if $\mathcal{C}$ detects h. On the other hand, to satisfy the detection condition for e, we must have

$$\langle\psi_0|e|\psi_-\rangle = A_- a_- a_0 + e^{i(\theta_- + \theta_0)} B_- b_- b_0 \equiv 0$$
$$\langle\psi_+|e|\psi_0\rangle = A_+ a_+ a_0 + e^{i(\theta_+ + \theta_0)} B_+ b_+ b_0 \equiv 0$$

for some positive coefficients $A_\pm, B_\pm$. Thus we see that $e^{i(\theta_- + \theta_0)} = e^{i(\theta_+ + \theta_0)} = -1$, and we need

$$A_- B_+ a_- b_+ = A_+ B_- a_+ b_-. \tag{6.4}$$

However, from the explicit form of the coefficients 6.1 we observe that as a function of l, the coefficient of the action of e is a concave down function symmetric with respect to $l = -1$. This means that the coefficient of e *i.e.* $\frac{1}{2}\sqrt{(N-l)(N+l+2)}$ in 6.1 achieves maximum at $l = -1$ (N odd) or $l = -2$ and $l = 0$ (N even); the coefficient monotonically decreases as l increases to more than 0 or decreases to less than -2. Therefore, if $l \geq 2$ and $k \leq -2$ in the context of this proof, we must have $A_- < A_+$ and $B_- > B_+$, so that $A_- B_+ < A_+ B_-$. In the other case, if $l \leq -1$ and $k \leq -7$, since the second derivative of the coefficient of e is negative, we can again deduce that $A_- B_+ < A_+ B_-$. The inequality still holds if $k \geq 1$ and $l \geq 7$, by symmetry.

Moreover, since the slopes in 6.3 all equal to ε_h, we see that $a_- < a_+$ hence $b_+ < b_-$, so that $a_- b_+ < a_+ b_-$. Therefore, we get $A_- B_+ a_- b_+ < A_+ B_- a_+ b_-$, contradicting 6.4. This finishes the first part of the proof, that is, the pattern in Figure 5 is impossible.

Now, note that with the disjoint support hypothesis, it suffices to argue locally that the code density cannot exceed 1/3. To this end, consider an element $|\psi\rangle$ of the given orthonormal basis for $\mathcal{C}$. To detect h, $|\psi\rangle$ must be supported on at least two eigenvectors of h, at least one below ε_h and at least one above. If the support of other basis elements is disjoint from the graph distance 1 neighbors of the support of $|\psi\rangle$, then at best we get a local code density of 1/4.

Otherwise, we have another basis element $|\phi\rangle$ whose support has distance 1 to the support of $|\psi\rangle$. Without loss of generality, assume that $\operatorname{supp}|\psi\rangle$ has larger or equal cardinality compared to $\operatorname{supp}|\phi\rangle$.

First, consider $|\psi\rangle$ to be supported on two eigenvectors of h, say $|k\rangle, |l\rangle$ where $k < l$. Since we are concerned with asymptotic behavior, let $k+2 < l$. Assume without loss of generality that $|k+2\rangle$ is in $\operatorname{supp}|\phi\rangle$; then to satisfy $\langle\phi|e|\psi\rangle = 0$, we see that $|l+2\rangle$ must also be in $\operatorname{supp}|\phi\rangle$. Now, since the pattern in Figure 5 is not possible, we see that other basis elements of $\mathcal{C}$ must have graph distance at least 1 from $\operatorname{supp}|\psi\rangle \cup \operatorname{supp}|\phi\rangle$. Thus the local density is 1/3, leveling the known construction.

Second, suppose $|\psi\rangle$ is supported on three eigenvectors of h, say two above ε_h and one below. If those two above are consecutive, then the slope ε_e of e must be nonzero. For this to be possible, all but at most one of the given basis for $\mathcal{C}$ must have support of size at least three. That is, the code density is asymptotically at most 1/3. Suppose those two above are not consecutive. Say the h-eigenvectors in the support are $|k\rangle, |l_1\rangle, |l_2\rangle$. Again, we can assume that a neighboring basis element $|\phi\rangle$ is supported on at most three eigenvectors of h. If they are $|k+2\rangle, |l_1+2\rangle, |l_2+2\rangle$, then we get a density of at most 1/3. If they are $|k+2\rangle, |l_1+2\rangle$, then $|k+4\rangle$ must be orthogonal to $\mathcal{C}$ by the same analysis as in the proof that Figure 5 is impossible, so the code density is again at most 1/3.

Third and last, if $|\psi\rangle$ has support of size at least 4, then to detect e, another basis element $|\phi\rangle$ of $\mathcal{C}$ must have support of size at least 2, so that the code density cannot exceed 1/3. □

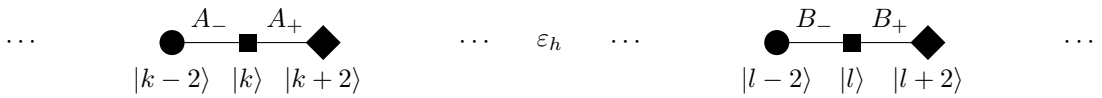

FIGURE 5. If each filled shape indicates the support of a basis element of a subspace $\mathcal{C}$, then $\mathcal{C}$ fails to detect all distance 1 errors.

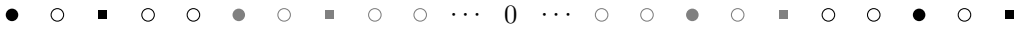

FIGURE 6. A distance 3 code construction. Each combination of filled shape and color indicates the support of an orthogonal basis for a second intermediate space $\mathcal{B}'$; a basis for the code is formed by pairing up the intermediate basis elements.

6.1.2. *General d.* Using the two-step construction, the intermediate subspace $\mathcal{B}$ for a distance d code need to have graph distance d, so $\dim \mathcal{B} = \frac{\dim \mathcal{H}}{d} + O(1)$. The remaining error space to be detected is $\mathcal{F} = P_{\mathcal{B}} \mathcal{V}_{d-1} P_{\mathcal{B}}$. By exploiting the commutation relations of e, f, h, we can deduce that $\mathcal{F} = \operatorname{span}\{I, h, h^2, \ldots, h^{d-1}\}$. For instance, ef and fe can be written as degree 2 polynomials in h. In the notation of section 4.2, we have $D = d - 1$, so in the second step using Tverberg's theorem, we get a code with dimension

$$\dim \mathcal{C} = \frac{\dim \mathcal{B}}{d} + O(1) = \frac{\dim \mathcal{H}}{d^2} + O(1).$$

For $d \geq 3$, Okada's idea can be generalized, but no known code can beat the two-step construction asymptotically. For instance, consider the code sketched in Figure 6. The black circles indicate the support of a state $|\psi_1\rangle$, and the black squares indicate the support of a state $|\phi_1\rangle$; similarly, the gray filled shapes indicate the support of $|\psi_2\rangle$ and $|\phi_2\rangle$, and so on we get pairs of states $|\psi_i\rangle$ and $|\phi_i\rangle$. If $\mathcal{B}$ is the intermediate space spanned by the h-eigenvectors corresponding to the filled shapes, then the errors remaining on $\mathcal{B}$ are spanned by h, h^2, e^2, f^2. By wisely choosing coefficients as we did for the distance 2 construction, we can ensure that $\langle \phi_i | e^2 | \psi_i \rangle = \langle \psi_i | f^2 | \phi_i \rangle = \langle \psi_i | h | \psi_i \rangle = \langle \phi_i | h | \phi_i \rangle = 0$. Denote the space spanned by $\{|\psi_i\rangle, |\phi_i\rangle\}$ as $\mathcal{B}'$, then $\dim \mathcal{B}' = \frac{\dim \mathcal{H}}{5} + O(1)$. The unfortunate fact is that we still need to detect the remaining error h^2 on $\mathcal{B}$, so that we need to pair up elements in $\{|\psi_i\rangle, |\phi_i\rangle\}$ appropriately and get a code $\mathcal{C}$ of dimension

$$\dim \mathcal{C} = \frac{\dim \mathcal{B}'}{2} + O(1) = \frac{\dim \mathcal{H}}{10} + O(1),$$

which is less efficient than $\frac{\dim \mathcal{H}}{9} + O(1)$ promised by the usual two-step construction.

Although the construction in Figure 6 is not optimal asymptotically, it is interesting to note that when $N = 7$, it yields a code of dimension 2, whereas the usual two-step construction can only give a code of dimension 1. In this small case, we do not need to pair up $|\psi\rangle$ and $|\phi\rangle$ to satisfy the detection condition for h^2.

6.2. **Codes in Triangular Lattice Representations of $\mathfrak{sl}_3\mathbb{C}$.** In this subsection and the next concerning quantum metric spaces associated to $\mathfrak{sl}_3\mathbb{C}$ and $\mathfrak{sl}_4\mathbb{C}$, we drop Okada's idea due to the complexity of the coefficients associated to the actions of off-diagonal elements in these higher dimensional cases. Instead, we pursue the usual two-step construction and utilize symmetries of the weight diagrams.

Recall that the symmetric power representations for $\mathfrak{sl}_3\mathbb{C}$ can be visualized as in Figure 2. That is, the weight lattice for the N-th symmetric power representation

$$\Lambda = \{a_1 L_1 + a_2 L_2 + a_3 L_3 \mid L_1 + L_2 + L_3 = 0;\ a_1, a_2, a_3 \in \mathbb{Z}_{\geq 0};\ a_1 + a_2 + a_3 = N\}$$

is the part of a translation of a triangular lattice called the A_2 lattice that is inside the triangle with vertices NL_1, NL_2, NL_3, with no degeneracy. By Proposition 5.4, the first step in the two-step

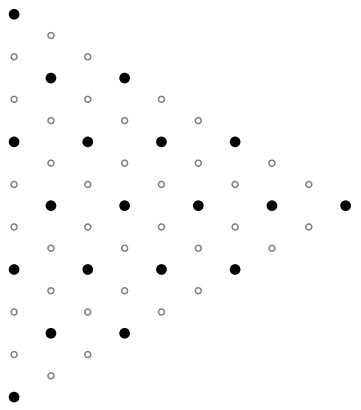

FIGURE 7. Weight diagram for $\mathrm{Sym}^9(V)$ where V is the defining representation of $\mathfrak{sl}_3\mathbb{C}$. Black circles indicate a basis for a distance 2 intermediate code subspace.

construction for a distance d code in this Lie type quantum metric space is to find a graph distance d subset of the weight diagram seen as a graph, where two vertices in the weight diagram are connected by an edge if the action of some $L_i - L_j$ translates one corresponding weight space to the other.

Remark 6.5. More generally, the weight diagram for the representation $\mathrm{Sym}^N(V)$ where V is the n-dimensional defining representation of of $\mathfrak{sl}_n\mathbb{C}$ is the part of an $(n-1)$-dimensional lattice inside an $(n-1)$-dimensional regular simplex with vertices $\{NL_i\}_{i=1,\dots,n}$, described by

$$\Lambda_n^N = \Big\{ \sum_{i=1}^n a_i L_i \mid \sum_{i=1}^n L_i = 0;\, a_i \in \mathbb{Z}_{\geq 0};\, \sum_{i=1}^n a_i = N \Big\}. \tag{6.6}$$

The corresponding lattice structure called A_{n-1} is usually defined by

$$A_{n-1} = \{(x_1, \dots, x_n) \in \mathbb{Z}^n \mid \sum_{i=1}^n x_i = 0\}.$$

We can identify L_i with the n-th standard basis vector of $\mathbb{Z}^n$, so that A_{n-1} is the lattice generated by $L_i - L_j$, which are closest to the origin and called the *roots* of the lattice. The origin is inside Λ_n^N if and only if n divides N. We shall restrict attention to such cases, so that the weight diagram is centered and Λ_n^N is part of the A_{n-1} lattice in the strict sense. No generality is lost if we only care about codes in the limit $N \to \infty$.

6.2.1. *$d = 2$.* Under the identification of the A_2 lattice with $\mathbb{Z}[\omega]$ where ω is a primitive third root of unity, we may choose the intermediate subspace $\mathcal{B}$ to be spanned by weight vectors corresponding to the intersection of the ideal generated by $1-\omega$ with the triangle having vertices NL_1, NL_2, NL_3, as indicated in Figure 7. This choice of $\mathcal{B}$ is asymptotically optimal as it is associated to a tiling of the plane by regular hexagons.

The following code construction was proposed by my REU partner Ian Shors. Note that $\mathcal{B}$ corresponds to part of a sublattice of the weight lattice, with $\dim \mathcal{B} = \frac{\dim \mathcal{H}}{3} + O(1)$. Recall from previous general discussions that the restricted error space $\mathcal{F}$ on $\mathcal{B}$ consists only of the diagonal elements in $\mathfrak{h}$. In the notation of section 4.2, we may choose $\vec{\varepsilon}$ to be the origin and partition the eigenbasis of $\mathcal{B}$ into disjoint subsets such that the convex hull spanned by each subset contains the origin. In the hexagonal center of the triangle with vertices NL_1, NL_2, NL_3, we may pair up weight vectors in opposite weight spaces with respect to the origin. From the remaining basis elements in $\mathcal{B}$ whose weights lie in the smaller triangles at the corners, we may choose one basis element in each of

the three triangle to form a subset, so that the triangle spanned by the three corresponding weights contain the origin. Since the hexagonal center occupies $\frac{2}{3}$ in area of the larger triangle, we see that

$$\dim \mathcal{C} = \big(\frac{2}{3}\cdot\frac{1}{2}+\frac{1}{3}\cdot\frac{1}{3}\big)\dim \mathcal{B} + O(1) = \frac{4}{27}\dim \mathcal{H} + O(1).$$

Since we used the symmetry of the weight lattice inside the hexagonal region, this is a super-Tverberg behavior. Tverberg's theorem guarantees only a code of dimension $\frac{\dim \mathcal{B}}{3}$ in the second step of the construction.

6.2.2. *General d.* As observed in [8], a general distance d sublattice of the A_2 lattice may be chosen to be the ideal generated by $\lfloor\frac{d}{2}\rfloor - \lfloor\frac{d+1}{2}\rfloor\omega$, so that one needs to travel $\lfloor\frac{d}{2}\rfloor + \lfloor\frac{d+1}{2}\rfloor = d$ edges in the weight diagram seen as a graph to get from the origin to the nearest point in the sublattice. The resulting choice of the intermediate code $\mathcal{B}$ is again asymptotically optimal and corresponds to a tiling of the plane. When $d = 2t$, the sublattice corresponds to a tiling by regular hexagons of radius t touching along edges in the lattice seen as a nearest neighbor graph. The number of vertices in the A_2 lattice on the boundary of a hexagon of radius t is $6t$; each vertex that is a vertex of the hexagon is shared by two other hexagons, and the remaining vertices on the boundary are shared by one other hexagon in the tiling. Thus, the volume of *i.e.* the number of vertices in each hexagon is

$$1 + \big(\sum_{k=1}^{t-1} 6k\big) + 6\cdot\frac{1}{3} + 6(t-1)\cdot\frac{1}{2} = 3t^2.$$

In the case that $d = 2t+1$, the sublattice corresponds to a tiling by hexagonal polyhex, as illustrated in Figure 5b of [8]. Each such polyhex contains in its interior a hexagon of radius t, so the number of vertices in each hexagonal polyhex is

$$1 + \sum_{k=1}^{t} 6k = 3t^2 + 3t + 1.$$

Accounting for boundary effects along the triangle bounding the weight diagram of $\mathrm{Sym}^N(V)$ in the A_2 lattice, we get $\dim \mathcal{B} = \frac{\dim \mathcal{H}}{3t^2} + O(N)$ or $\frac{\dim \mathcal{H}}{3t^2+3t+1} + O(N)$ respectively in these two cases, where $\dim(\mathcal{H}) = \Theta(N^2)$.

Now, since the error space we wish to detect is $\mathcal{V}_{d-1} = \mathrm{span}\,(\mathfrak{sl}_3\mathbb{C}\oplus\mathbb{C}I)^{d-1}$, when $d > 2$ it is not clear if the symmetry of the weight diagram can be exploited to construct more efficient codes. Instead, we rely on Tverberg's theorem and count the dimension of the error space $\mathcal{F} = P_{\mathcal{B}}\mathcal{V}_{d-1}P_{\mathcal{B}}$. Note that $\mathcal{F}$ is spanned by monomials in 1_{ij} and h_{ij} for $i \neq j$ of degree $\leq d-1$, such that the composed translations by $L_i - L_j$ in the weight diagram corresponding to the 1_{ij}'s is the identity. Returning to the concrete description of the symmetric power representations in terms of homogeneous polynomials, recall that the action of $\mathfrak{sl}_3\mathbb{C}$ on $\mathbb{C}[x_1,x_2,x_3]_N$ is defined by

$$d\rho(1_{ij}) = x_i\partial_j\ (i\neq j)\quad d\rho(1_{11}-1_{22}) = h_{12} = x_1\partial_1 - x_2\partial_2\quad d\rho(1_{11}-1_{33}) = h_{13} = x_1\partial_1 - x_3\partial_3.$$

Viewing k_1 and k_2 as variables with $k_3 = N - k_1 - k_2$, we see that the action of each 1_{ij} or h_{ij} on $x_1^{k_1}x_2^{k_2}x_3^{k_3}$ is linear in k_1 and k_2. Consequently, each monomial in 1_{ij} and h_{ij} of degree $\leq d-1$ satisfying the aforementioned condition acts diagonally such that the diagonal entry for $x_1^{k_1}x_2^{k_2}x_3^{k_3}$ is a polynomial in k_1 and k_2 of degree $\leq d-1$.

On the other hand, observe that the eigenvalues of the action of h_{12} and h_{13} on $x_1^{k_1}x_2^{k_2}x_3^{k_3}$ are respectively $k_1 - k_2$ and $k_1 - k_3 = 2k_1 + k_2 - N$. Therefore, polynomials in h_{12} and h_{13} of degree $\leq d-1$ are the collection of diagonal matrices acting on $\mathbb{C}[x_1,x_2,x_3]_N$ such that the diagonal entry for $x_1^{k_1}x_2^{k_2}x_3^{k_3}$ is a polynomial in k_1 and k_2 of degree $\leq d-1$. We have shown that $\mathcal{F}\subset\mathrm{span}(\mathfrak{h}\oplus\mathbb{C}I)^{d-1}$, and that for the latter, degree $\leq d-1$ polynomials in h_{12} and h_{13} is a basis. Thus, the desired dimension upper bound for $\mathcal{F}$ is $\binom{2+d-1}{2} = \frac{d^2+d}{2}$. The same argument shows that more generally,

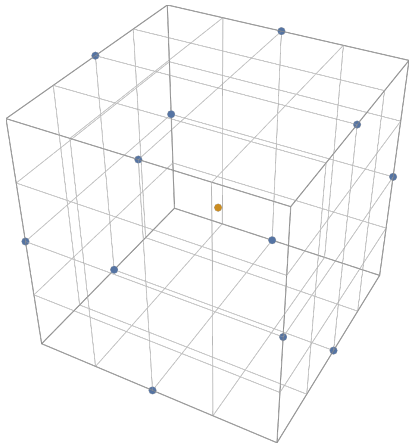

FIGURE 8. Neighbors of a point in the A_3 lattice

Proposition 6.7. *In a symmetric power representation for $\mathfrak{sl}_n\mathbb{C}$, let $\mathcal{V}_{d-1} = \mathrm{span}(\mathfrak{sl}_n\mathbb{C} \oplus \mathbb{C}I)^{d-1}$, $\mathcal{B}$ be a subspace spanned by weight vectors associated to a minimum distance $\geq d$ subset of the weight diagram, and $\mathcal{F} = P_{\mathcal{B}}\mathcal{V}_{d-1}P_{\mathcal{B}}$. Then $\mathcal{F}$ is contained in the space spanned by polynomials in $h_{12}, h_{13}, \ldots, h_{1n}$ of degree $\leq d-1$, hence $\dim \mathcal{F} \leq \binom{n+d-2}{n-1}$.*

For $\mathfrak{sl}_3\mathbb{C}$, with $\dim \mathcal{F} \leq \frac{d^2+d}{2}$, the resulting lower bound for code dimension is $\lceil \frac{2 \dim \mathcal{B}}{d^2+d+2} \rceil$.

6.3. **Codes in Tetrahedral Lattice Representations of $\mathfrak{sl}_4\mathbb{C}$.** The N-th symmetric power representations of $\mathfrak{sl}_4\mathbb{C}$ are understood using 6.6 with $N = 4$. To visualize, $L_1, \ldots, L_4$ can be represented as vertices of a regular tetrahedron that coincides with four of the vertices of a cube centered at the origin. Explicitly, set

$$L_1 = (1,1,1) \qquad L_2 = (1,-1,-1) \qquad L_3 = (-1,1,-1) \qquad L_4 = (-1,-1,1). \tag{6.8}$$

As before, Λ_n^N is part of the A_3 lattice generated by $L_i - L_j$ inside the tetrahedron with vertices NL_1, NL_2, NL_3, NL_4.

In this subsection, we only look at distance 2 codes. A general construction for a distance 2 subset of Λ_n^N can be described as the kernel of the map $\Lambda_n^N \to \mathbb{Z}/n$ given by

$$a_1L_1 + \cdots + a_nL_n \mapsto a_1 + 2a_2 + \cdots + na_n, \tag{6.9}$$

where the last term na_n in the sum can of course be omitted. If some $a_1L_1 + \cdots + a_nL_n$ is in the kernel, then the translation of this point by $L_i - L_j$ for any $i \neq j$ is not in the kernel, so the distance between any two points in the kernel is at least 2. When $n = 3$ and 3 divides N, this construction coincides with the construction discussed in the previous subsection, where the distance 2 subset is part of the sublattice generated by $3L_1$ and $3L_2$. As discussed above and illustrated in Figure 7, this distance 2 subset is optimal.

When $n = 4$, the distance 2 subset chosen above, denoting now as $\Lambda_{\mathcal{B}}^N$, is again optimal. Suppose $p = a_1L_1 + \cdots + a_4L_4$ in Λ_4^N is not in the kernel of 6.9. If $a_1 + 2a_2 + 3a_3 \equiv 1 \mod 4$, then p has exactly four neighbors in the kernel after translation by $L_4 - L_1, L_1 - L_2, L_2 - L_3, L_3 - L_4$; when $a_1 + 2a_2 + 3a_3 \equiv 3$, the neighbors are obtained by the inverse of the above translations; when $a_1 + 2a_2 + 3a_3 \equiv 2$, we have four neighbors after translations by $L_4 - L_2, L_2 - L_4, L_3 - L_1, L_1 - L_3$.

On the other hand, the neighbors of a point p in the A_3 lattice are given as midpoints of the twelve edges of a cube centered at p, as shown in Figure 8. It is straightforward to observe that we can choose at most 4 points out of these twelve points such that the nearest neighbor graph distance between any two chosen points is at least 2. This shows the optimality of $\Lambda_{\mathcal{B}}^N$. When $N = 8k$ for

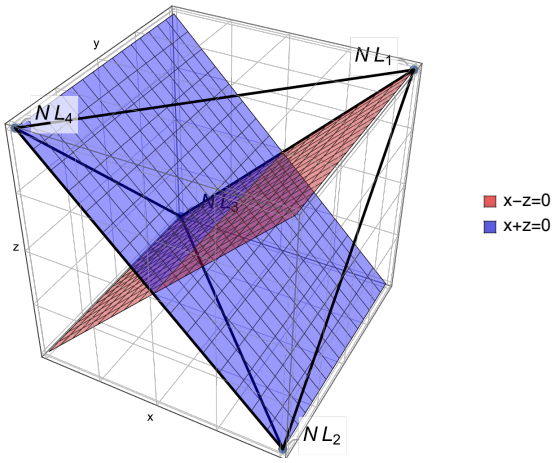

FIGURE 9. A schematic picture of the weight lattice for the N-th symmetric power representation of $\mathfrak{sl}_4\mathbb{C}$. The planes indicated are exploited to form triangles containing the origin.

some integer k, the origin is contained in Λ_4^N and $\Lambda_{\mathcal{B}}^N$ is part of a sublattice of the A_3 lattice. We restrict attention to these cases.

In the second step of the two-step construction, the most intuitive choice for $\vec{\varepsilon}$ is the center of the tetrahedron, or the origin in our coordinates. In the convex octahedral region formed by midpoints of the edges of the tetrahedron, we may pair up points opposite with respect to the origin. The opposite of a point in the remaining region is no longer inside the tetrahedron, so we consider the next best possibility; that is, to find collections of three points forming a triangle that contains the origin.

Inspired by numerical simulations, we utilize symmetries of $\Lambda_{\mathcal{B}}^N$ to construct a class of triangles. Note that if $p = a_1L_1 + a_2L_2 + a_3L_3 + a_4L_4$ is in $\Lambda_{\mathcal{B}}^N$, from $a_1 + 2a_2 + 3a_3 \equiv 0 \mod 4$ and $a_1 + a_2 + a_3 + a_4 = N \equiv 0 \mod 2$, we deduce that $a_1 - a_3 \equiv a_2 - a_4 \equiv 0 \mod 2$. In Euclidean coordinates specified by 6.8,

$$p = a_1L_1 + a_2L_2 + a_3L_3 + a_4L_4 = (a_1 + a_2 - a_3 - a_4, a_1 - a_2 + a_3 - a_4, a_1 - a_2 - a_3 + a_4).$$

We easily calculates that the distance from p to the plane $x + z = 0$ is $\sqrt{2}\,|a_1 - a_3|$, so the difference between p and its reflection with respect to $x + z = 0$ is $\pm(a_1 - a_3)(L_1 - L_3)$, which is in the kernel of the map 6.9. Therefore, $\Lambda_{\mathcal{B}}^N$ is symmetric with respect to $x + z = 0$. Similarly, the distance from p to the plane $x - z = 0$ is $\sqrt{2}\,|a_2 - a_4|$, so the difference between p and its reflection with respect to $x - z = 0$ is $\pm(a_2 - a_4)(L_2 - L_4)$, which is again in the kernel of 6.9. Thus, $\Lambda_{\mathcal{B}}^N$ is also symmetric with respect to $x - z = 0$.

With p as above and assuming p has $x - z > 0$, the reflection of p across $x - z = 0$ is given by

$$p' = p - (a_2 - a_4)(L_2 - L_4) = a_1L_1 + a_4L_2 + a_3L_3 + a_2L_4.$$

Then we calculate that

$$\begin{aligned} -(p + p') &= -(2a_1L_1 + (a_2 + a_4)L_2 + 2a_3L_3 + (a_2 + a_4)L_4) \\ &= \big(\frac{3N}{4} - 2a_1\big)L_1 + \big(\frac{3N}{4} - a_2 - a_4\big)L_2 + \big(\frac{3N}{4} - 2a_3\big)L_3 + \big(\frac{3N}{4} - a_2 - a_4\big)L_4. \end{aligned}$$

We have written $-(p + p')$ in the form where the sum of coefficients of L_i is N. Since $-(p + p')$ is always in the A_3 lattice, if it is also inside the tetrahedron with vertices NL_i, it must be a member

of Λ_4^N, and in this case the above coefficients of L_i are nonnegative. With the assumption that N is divisible by 8, we easily see that

$$\big(\frac{3N}{4} - 2a_1\big) + 2\big(\frac{3N}{4} - a_2 - a_4\big) + 3\big(\frac{3N}{4} - 2a_3\big) \equiv 0 \mod 4.$$

Thus, if $-(p+p')$ is inside the tetrahedron, then it must be inside $\Lambda_{\mathcal{B}}^N$. Similarly, if p'' is the reflection of p across $x+z=0$ and $-(p+p'')$ is inside the tetrahedron, then $-(p+p'')$ is inside $\Lambda_{\mathcal{B}}^N$.

The recipe to find triangles through the origin is then to pair up p_2 relatively close to NL_2 and p_4 relatively close to NL_4 where p_2 and p_4 are symmetric with respect to $x-z=0$, and let the third vertex be $-(p_2+p_4)$, as long as it is inside the tetrahedron. Similarly, we form triangles by picking suitable p_1 relatively close to NL_1 and p_3 relatively close to NL_3 and find the third vertex on the plane $x+z=0$.

Unfortunately, this strategy does not improve code efficiency in the limit $N \to \infty$, as the number of triangles is limited by the number of lattice points on the planes $x+z=0$ and $x-z=0$, which is a negligible fraction of $\Lambda_{\mathcal{B}}^N$ when N is large. Still, it is an interesting fact that the intermediate lattice $\Lambda_{\mathcal{B}}^N$ possesses symmetries that make possible the construction of these triangles, and for small N this strategy improves code size.

Example 6.10. In the case $N=8$, shown in Figure 10 is the intermediate lattice $\Lambda_{\mathcal{B}}^8$ and the result of our two-step construction with the origin as $\vec{\varepsilon}$, with convex hull spanned by two, three, or four points in $\Lambda_{\mathcal{B}}^8$. Although we cannot prove the optimality of our two-step strategy among all possible code constructions, in this specific example the resulting code is optimal at each step in our framework of two-step constructions.

In more detail, this framework incorporates both the general recipe we described in section 4.2 and our choice to work in the eigenbasis of the diagonal $\mathfrak{h}$ and choose basis elements for the intermediate subspace $\mathcal{B}$ to have disjoint support on this eigenbasis. We have previously shown that our construction of $\Lambda_{\mathcal{B}}$ is the densest graph distance ≥ 2 subset of the lattice Λ_4^N, hence the resulting $\mathcal{B}$ is optimal in the first step.

Now, we examine four regions R_i in the tetrahedron T spanned by $8L_i$. Each R_i is defined to be the corner of T bounded by the plane through the center perpendicular to L_i. These four planes respectively contain the four triangles in Figure 10, and in our coordinates the four regions we refer to are the intersection of the tetrahedron with $x+y+z>0; x-y-z>0; -x+y-z>0; -x-y+z>0$. Note in Figure 10 that for each $i=1,2,3,4$, we have associated a convex hull to every point in R_i, and we maximized the number of convex sets spanned by points on the boundary plane of R_i. It is not possible to partition $\Lambda_{\mathcal{B}}^8$ into more disjoint subsets such that the convex hull spanned by each subset contains the origin, because once we have associated a subset to every point in R_i and its boundary, we cannot obtain more subsets – the convex hull spanned by any collection of points in $T \backslash \overline{R_i}$ cannot contain the origin. This kind of "centroid argument" in general requires the fact that $T = \cup_i \overline{R_i}$ because each R_i potentially contains a different number of points, but in this case each R_i contains the same number of points due to symmetries of $\Lambda_{\mathcal{B}}^4$. Beware that we have shown that both the first step and the second step are optimal in our construction, but it is not clear if the combined two steps are optimal, although intuition says that it is in this case.

The above argument fails for larger N because lattice points fill up the tetrahedral region and it is no longer certain if we can associate a convex hull to each point in R_i. However, the exactly same "centroid argument" applied to the two-dimensional case shows that for any N, the codes we constructed in the previous subsection for the N-th symmetric power representations of $\mathfrak{sl}_3\mathbb{C}$ is optimal in the second step. Recall that the first step as indicated in Figure 7 is also optimal. Again, it is highly likely that the combined two steps are optimal among two-step constructions, but more sophisticated analysis is required.

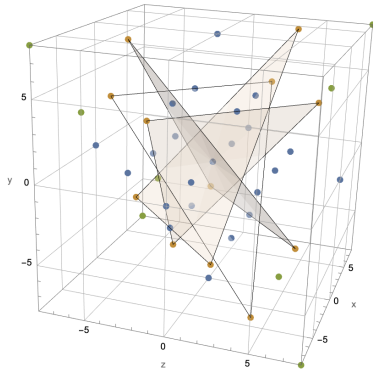

FIGURE 10. Distance 2 intermediate lattice for the 8-th symmetric power representation of $\mathfrak{sl}_4\mathbb{C}$. The blue points form pairs opposite with respect to the origin; yellow points form four triangles passing though the origin; green points span two tetrahedrons containing the origin.

7. Discussion

There are a lot of remaining questions that stem from materials discussed in this paper. For a concrete instance, it is clear from inspection that we can construct more triangles in the tetrahedral lattice for large N, but is it possible to construct enough that matters asymptotically? Besides optimizing the two-step construction, one wonders if Okada's code construction for the $\mathfrak{sl}_2\mathbb{C}$ case can be generalized to other cases we considered. My REU group member Ian Shors has made some progress in this direction. Since Okada's idea relies on the inner product we choose to make the representation in question unitary, it seems to me that more explanation for the relevance of this choice is also needed.

Another line of investigation is to look at more quantum metric spaces. This includes symmetric power representations of $\mathfrak{sl}_n\mathbb{C}$ for higher values of n, other irreducible representations of $\mathfrak{sl}_n\mathbb{C}$, tensor product of these representations, Lie type graph metric spaces with $\mathfrak{g}$ chosen to be other Lie algebras, etc. To find a focus among these many potential topics, it would be good if we can develop more theory. For instance, what is the implication of a degenerate weight space? What is the relation between quantum codes in a reducible representation to quantum codes in its irreducible components?

One may also try to elevate the objects of study from elementary convex geometrical symmetries to more algebraic symmetries. It is possible to define and study the isometry group of a Lie type graph metric space, which is a Lie group with some relation to the Lie algebra in question. It is even possible to use subgroups of the isometry group to define quantum codes. Unfortunately, I did not have time to study these topics in more detail. I hope that by pursuing these questions, we can find a more natural explanation for the introduction of Lie algebra to quantum codes and go beyond the basis-dependent two-step strategy to a more algebraic recipe of code construction.

Acknowledgement

It is a pleasure to work with Professor Greg Kuperberg over the summer, and I am grateful for the organization of the UC Davis REU program by Professor Kuperberg and Professor Javier Arsuaga. I also wish to express gratitude for graduate students Sanchayan Dutta and Rui Okada for their guidance, and Jennifer Brown and Rohit Thomas for their generous assitance throughout the REU program. Finally, I wish to thank my research group members Bella Finkel, Yuanyuan Shen, Ian Shors, and Jonathan Webb.

References

[1] Chris Bumgardner. *Codes in W^*-metric spaces: theory and examples*. PhD thesis, University of California, Davis, 2012. arXiv:1205.4517.

[2] John H. Conway and Neil J. A. Sloane. *Sphere packings, lattices and groups*, volume 290 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, 3rd edition, 1993.

[3] William Fulton and Joseph Harris. *Representation theory: a first course*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 3rd edition, 1998.

[4] Daniel Gottesman. An introduction to quantum error correction and fault-tolerant quantum computation. 2009. arXiv:0904.2557.

[5] Brian Hall. *Lie groups, Lie algebra, and representations: an elementary introduction*. Graduate Texts in Mathematics. Springer, 2nd edition, 2015.

[6] Emanuel Knill and Raymond Laflamme. A theory of quantum error-correcting codes. 1995. arXiv:quant-ph/9604034.

[7] Emanuel Knill, Raymond Laflamme, and Lorenza Viola. Theory of quantum error correction for general noise. *Phys. Rev. Lett.*, 84(11):2525–2528, 2000. arXiv:quant-ph/9908066.

[8] Greg Kuperberg. Numerical cubature from Archimedes' hat-box theorem. *SIAM J. Numer. Anal.*, 44(3):908–935, 2006. arXiv:math/0405366.

[9] Greg Kuperberg and Nik Weaver. A von Neumann algebra approach to quantum metrics. *Mem. Amer. Math. Soc.*, 215(1010):1–80, 2012. arXiv:1005.0353.

[10] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge, 2000.

[11] John Preskill. Lecture notes on quantum information. `http://theory.caltech.edu/~preskill/ph229/`.

[12] Helge Tverberg. A generalization of radon's theorem. *Journal of the London Mathematical Society*, s1-41:123–128, 1996.