

A SURVEY OF TWO APPLICATIONS OF REPRESENTATION THEORY IN QUANTUM COMPUTATION

YUHAD OLARTE IBRAHIM

1. A SELF-CONTAINED INTRODUCTION TO QUANTUM COMPUTING

1.1. Postulates of quantum mechanics. In what follows, we introduce the standard postulates of quantum mechanics. Throughout these notes, we will see that the main idea of a quantum state is a superposition of classical states, described by a vector of complex amplitudes. Such a state can be manipulated through either measurements or unitary transformations. This section is mainly based on [8, 6, 5].

Postulate 1. The set of all quantum states of a given quantum system forms a complex vector space equipped with an inner product, i.e it is a Hilbert Space denoted as \mathcal{H} , called the state space.

In our setting, we take $\mathcal{H} = \mathbb{C}^N$ with $N = 2^n$, where n is a non-negative integer called the number of *quantum bits* (or *qubits*). Is common to find in the literature the name of *qudit* whereas $n := d$. A quantum state $|\psi\rangle \in \mathbb{C}^N$ can be expressed as a column vector, and its Hermitian conjugate corresponds to a row vector, as shown below

$$|\psi\rangle = \begin{pmatrix} \psi_0 \\ \vdots \\ \psi_{N-1} \end{pmatrix}, \quad \langle\psi| := |\psi\rangle^\dagger = (\bar{\psi}_0, \dots, \bar{\psi}_{N-1}).$$

Here, \bar{c} denotes the complex conjugate of a complex number $c \in \mathbb{C}$. The notation $|\cdot\rangle$ and $\langle\cdot|$, called ket and bra respectively, used to represent quantum states is known as *Dirac notation*. The *inner product between two states* corresponds to the dot product between a bra and a ket vector

$$\langle\psi|\varphi\rangle := \langle\psi, \varphi\rangle = \sum_{i \in \{0, \dots, N-1\}} \bar{\psi}_i \varphi_i.$$

The *outer product* of two states, denoted by $|\psi\rangle\langle\varphi|$, corresponds to the $N \times N$ matrix whose (i, j) -th entry is given by the product $\psi_i \bar{\varphi}_j$. We may assume that quantum states are normalized, meaning $|\psi\rangle$ is a unit vector satisfying $\langle\psi|\psi\rangle = 1$.

Example 1.1. A single qubit corresponds to the state space $\mathcal{H} = \mathbb{C}^2$. We define the computational basis vectors as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Since these vectors form an orthonormal basis of \mathcal{H} , any qubit state $|\psi\rangle \in \mathcal{H}$ can be expressed as a linear combination $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ ensures that $|\psi\rangle$ is a unit vector. We refer to $\{|0\rangle, |1\rangle\}$ as the *computational basis*.

In the above example, we say that $|\psi\rangle$ is in a *superposition* of the basis states $|0\rangle$ and $|1\rangle$, with complex amplitudes α and β . The concept of superposition depends on the chosen basis: a state that is a superposition in one basis might be a basis vector in another.

Postulate 2. The time evolution of a closed quantum system is described by a unitary transformation on the Hilbert space. That is, if a quantum system is in the state $|\psi\rangle$ at some initial time, then at a later time its state is given by

$$|\psi'\rangle = U |\psi\rangle.$$

where U is a unitary operator on the state space \mathcal{H} , meaning $U^\dagger U = U U^\dagger = I$.

Example 1.2. The following are popular quantum operations

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

These are called the *Pauli matrices*. Another important quantum operation is the *Hadamard matrix* defined as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Postulate 3. When we perform a measurement on a quantum system, the outcome is random and described by a set of measurement operators $\{M_m\}$ acting on the state space \mathcal{H} . These operators tell us both the possible outcomes and how the state changes when we observe it. The index m refers to the measurement outcomes that may occur in the experiment. These operators must satisfy

$$(1) \quad \sum_m M_m^\dagger M_m = I.$$

If the system is in state $|\psi\rangle$ before the measurement, then the probability of obtaining outcome m is given by

$$p(m) = \langle \psi | M_m^\dagger M_m | \psi \rangle,$$

then the state collapses after measurement to

$$|\psi'\rangle = \frac{M_m |\psi\rangle}{\sqrt{\langle \psi | M_m^\dagger M_m | \psi \rangle}}.$$

Note that the equation (1) implies that the probabilities, indeed sum to one as

$$1 = \sum_m p(m) = \sum_m \langle \psi | M_m^\dagger M_m | \psi \rangle.$$

Before presenting our first example, let us make the following remark. Consider the state $|\bar{\psi}\rangle = e^{i\theta} |\psi\rangle$, where $|\psi\rangle$ is a state vector and $\theta \in \mathbb{R}$ is a real number. The state $|\bar{\psi}\rangle$ differs from $|\psi\rangle$ only by a *global phase factor*. Importantly, this global phase has no physical consequences: it does not affect the outcomes of any measurement. To see this, let M_m be a measurement operator corresponding to outcome m . Then, the probability of obtaining m when measuring $|\bar{\psi}\rangle$ is

$$\langle \bar{\psi} | M_m^\dagger M_m | \bar{\psi} \rangle = \langle \psi | e^{-i\theta} M_m^\dagger M_m e^{i\theta} | \psi \rangle = \langle \psi | M_m^\dagger M_m | \psi \rangle.$$

Therefore, from an observational point of view, $|\psi\rangle$ and $e^{i\theta} |\psi\rangle$ are indistinguishable. This is why global phase factors are considered physically irrelevant and are often ignored.

Example 1.3. In this example we explicit the measurement of a qubit in the computational basis. Define $M_0 = |0\rangle\langle 0|$ and $M_1 = |1\rangle\langle 1|$. Since this two matrices are Hermitian and $M_0^2 = M_0$, $M_1^2 = M_1$ thus $I = M_0^\dagger M_0 + M_1^\dagger M_1 = M_0 + M_1$. Now suppose the state being measured is $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with easy calculation we see that the probability of obtaining measurement outcome 0 is

$$p(0) = \langle \psi | M_0^\dagger M_0 | \psi \rangle = \langle \psi | M_0 | \psi \rangle = (\bar{\alpha}\langle 0| + \bar{\beta}\langle 1|) M_0 (\alpha|0\rangle + \beta|1\rangle) = |\alpha|^2.$$

Analogous, the probability of obtaining measurement outcome 1 is $|\beta|^2$. The state after measurement in the two cases is therefore $|0\rangle$ and $|1\rangle$.

Postulate 4. The state space of a composite quantum system is the tensor product of the state spaces of its subsystems. Explicitly, if two systems have state spaces \mathcal{H}_A and \mathcal{H}_B then the joint system has state space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$.

Note that this postulate aligns with the notion of a qudit introduced earlier. Specifically, taking the state space of a single qubit to be \mathbb{C}^2 , the composite system of d qubits has the state space

$$\underbrace{\mathbb{C}^2 \otimes \dots \otimes \mathbb{C}^2}_{d \text{ qubits}} = (\mathbb{C}^2)^d = \mathbb{C}^N.$$

Here we make several important remarks. Following the previous postulate, we can generalize the calculation of measurement outcomes to systems of multiple qubits. For instance, consider a two-qubit state expressed in the computational basis as $|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$. We define the measurement operators corresponding to each computational basis state as

$$M_{00} = |00\rangle\langle 00|, \quad M_{01} = |01\rangle\langle 01|, \quad M_{10} = |10\rangle\langle 10|, \quad M_{11} = |11\rangle\langle 11|.$$

Then the probability of each outcome is given by $p(00) = |a|^2$, $p(01) = |b|^2$, $p(10) = |c|^2$ and $p(11) = |d|^2$ as can be verified by direct computation, analogously to Example 1.3. More generally, for an n -qubit system in a state of the form $|\psi\rangle = \sum_{j=0}^{N-1} \alpha_j |j\rangle$ where $\{|j\rangle\}$ is the computational basis for \mathbb{C}^N and $N = 2^n$, the probability of obtaining outcome j under measurement is $p(j) = |\alpha_j|^2$. This formula is known as *Born's Rule*.


Remark 1.4 (Partial measurement). In quantum mechanics, we often perform partial measurements as follows. Consider the two-qubit state expressed in the computation basis $|\psi\rangle = \sum_{i,j} \alpha_{i,j} |i\rangle|j\rangle$ and $i, j \in \{0, 1\}$. To measure only the first qubit, without disturbing the second, we define the measurement operators $M_i^A = |i\rangle\langle i| \otimes I$. Note that $(M_i^A)^2 = M_i^A$ and thus $(M_i^A)^\dagger = M_i^A$. Then, with straightforward calculations, we see that the probability of measure i in the first qubit is given by

$$p(\text{measure } 0 \text{ in the first qubit}) = \langle \psi | M_0^A | \psi \rangle = |\alpha_{00}|^2 + |\alpha_{01}|^2.$$

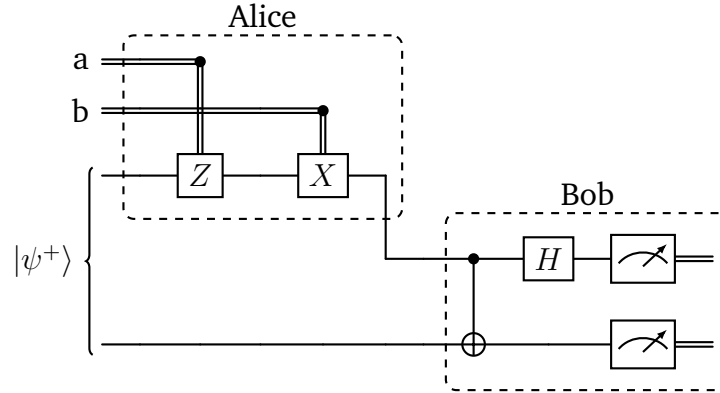
Analogously, to measure the second we define the operators $M_j^B = I \otimes |j\rangle\langle j|$ and then the probability of measure of measure j in the second qubit is given by $|\alpha_{10}|^2 + |\alpha_{11}|^2$. From this and Postulate 3, is easy to calculate the state after the measurement.

1.2. Quantum Circuits. A *quantum circuit* is a model for quantum computation that describes how quantum information (qubits) is initialized, transformed, and measured. The general process consists of four main steps:

- (1) **Initialization:** Each quantum circuit begins with a well-defined quantum state, which are represented individually by lines, or cables. The times goes from left to right.
- (2) **Quantum Gates (Unitary Evolution):** Apply quantum operations, like the one we introduce in the Example 1.2 this gates can act on one or more qubits. This operations are denoted by boxed indicating their corresponding name. In the Table 4 we describe the most common quantum gates in detail.
- (3) **Measurement:** We measure the states result by the Quantum gates. This will be denoted

as 

Example 1.5 (Superdense coding protocol). The following example is called *Superdense coding protocol* which allows sending 2 classical bits using 1 qubit, provided that the sender (traditionally named Alice) and receiver (Bob) share an entangled Bell pair beforehand. The superdense coding protocol is defined via the following quantum circuits



where $|\psi^+\rangle = \frac{1}{\sqrt{2}} |01\rangle + \frac{1}{\sqrt{2}} |10\rangle$.

1.3. Density matrices. So far we have described quantum states using state vectors $|\psi\rangle$. This corresponds to the ideal situation in which we have complete knowledge of the system. However, in practice, a quantum system might be in one of several possible states with some probability distribution. To handle this situation, we use the *density operators* (also called density matrices). A *density operator* is a positive semidefinite, Hermitian operator ρ on the Hilbert space \mathcal{H} with unit trace

$$\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|, \quad p_i \geq 0, \quad \sum_i p_i = 1.$$

Here, the system is in the pure state $|\psi_i\rangle$ with probability p_i . If one of the probabilities $p_i = 1$, the system is said to be in a *pure state*, and we have $\rho = |\psi\rangle \langle \psi|$, if more than one probability is nonzero, the system is in *mixed state*. A density operator ρ satisfies:

- (1) Hermitian: $\rho^\dagger = \rho$.
- (2) Positive semidefinite: $\langle \phi | \rho | \phi \rangle$ for all $|\phi\rangle \in \mathcal{H}$.
- (3) Trace one: $\text{Tr}(\rho) = 1$.

An important remark is that Postulate 3 can be restated using density operators. If $\{M_m\}$ is a collection of measurement operators satisfying $\sum_m M_m^\dagger M_m = I$, then the probability of obtaining outcome m when the system is in state ρ is

$$p(m) = \text{Tr}(M_m^\dagger M_m \rho).$$

If the outcome m occurs, the post-measurement state is given by

$$\rho' = \frac{M_m \rho M_m^\dagger}{\text{Tr}(M_m^\dagger M_m \rho)}.$$

Example 1.6. Consider the pure qubit state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. The corresponding density matrix is

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} |\alpha|^2 & \alpha\bar{\beta} \\ \bar{\alpha}\beta & |\beta|^2 \end{pmatrix}.$$

If instead the system is in $|0\rangle$ with probability p and in $|1\rangle$ with probability $1 - p$, the density matrix is

$$\rho = p|0\rangle\langle 0| + (1 - p)|1\rangle\langle 1| = \begin{pmatrix} p & 0 \\ 0 & 1 - p \end{pmatrix}.$$

This shows how the density operator formalism generalizes the description of quantum states and naturally incorporates both classical randomness and quantum superposition.

2. BASIC FACTS ABOUT REPRESENTATION THEORY OF S_n AND GL

In this section, we introduce basic ideas and facts from representation theory, focusing on the representations of S_n and GL. For the references, we will mainly follow [3, 4].

A *representation* of a finite group G on a finite-dimensional complex vector space V is a group homomorphism $\rho: G \rightarrow \text{GL}(V)$. When the context is clear, we often abuse notation by referring to V itself as the representation of G and write gv for $\rho(g)(v)$ if it does not lead to confusion. The *degree* of V , as a representation of G is $\dim V$. A subspace W of V is a *subrepresentation* if W is G -invariant, this is, $gw \in W$ for all $w \in W$ and $g \in G$. Note that 0 and V itself are subrepresentations of V . We say that V is *irreducible* if it contains no proper nonzero invariant subspaces, this is, 0 and V are the unique subrepresentations of V .

Example 2.1. The following are basics examples of representations.

- (1) Let G be a finite group. The *trivial representation* \mathbb{C} defined as $gv = v$ for all $g \in G$ and $v \in \mathbb{C}$. Note that this is a degree 1 representation.
- (2) Let $V = \mathbb{C}^n$ with standard basis $\{e_1, \dots, e_n\}$. Define a representation of the symmetric group S_n on V called *permutation representation* by permuting the basis vectors $ge_i = e_{g(i)}$. This representation is not irreducible: the vector $v_0 = e_1 + \dots + e_n$ is fixed by all of S_n , so the subspace $\langle v_0 \rangle$ is invariant and gives the trivial representation. The orthogonal complement

$$V_{\text{std}} = \{(x_1, \dots, x_n) \in \mathbb{C}^n : x_1 + \dots + x_n = 0\}$$

is also S_n -invariant, and is called the *standard representation*. For $n \geq 2$, this representation is irreducible and has degree $n - 1$.

As a concrete example, for $n = 3$, the standard representation $\rho: S_3 \rightarrow \text{GL}_2(\mathbb{C})$ can be specified on the generators by

$$\rho((12)) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \rho((123)) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

- (3) Let $V = \mathbb{C}$. Define a representation of the symmetric group S_n on V called *sign representation* as $gv = \text{sgn}(g)v$ for all $g \in S_n$ and $v \in \mathbb{C}$. Note that this is a degree 1 representation.

Note that given a representation $\rho: G \rightarrow \text{GL}(V)$ of a finite group G , we can define a G -action on V by setting $g \cdot v = \rho(g)(v)$. Conversely, a G -action on a complex vector space V defines a representation via $\rho(g)(v) = g \cdot v$.

Let V and W be representations of a finite group G . A G -map f is a linear map $f: V \rightarrow W$ that commute with the action of G , this is, $gf(v) = f(gv)$ for all $g \in G$ and $v \in V$. Additionally we say that a G -map f is a *isomorphism of representation* if it is an isomorphism of vector spaces, in this case we say that V and W are *equivalent representations* and write $V \simeq W$. The set of all G -maps is denoted by $\text{Hom}_G(V, W)$. We write $\text{End}_G(V) := \text{Hom}_G(V, V)$. Note that $\text{Hom}_G(V, W) \subseteq \text{Hom}_{\mathbb{C}}(V, W)$.

Proposition 2.2 (Schur's Lemma). *Let $f: V \rightarrow W$ be a non-zero G -map. Then*

- (1) *If V is irreducible, then f is injective.*
- (2) *If W is irreducible, then f is surjective.*
- (3) *If $V = W$, then $f = tI$ for some nonzero $t \in \mathbb{C}$.*

Example 2.3. Let G be an abelian group and V a irreducible representation of G . Consider $h \in G$ and define $f_h: V \rightarrow V$ as $f_h(v) := hv$. Note that $gf_h(v) = g(hv) = (gh)v = h(gv) = f_h(gv)$ and thus $f_h \in \text{End}_G(V)$ then by Schur's lemma $f_h = c_h I$ for some nonzero $c_h \in \mathbb{C}$. Fix a nonzero $v \in V$ then $f_h(v) \in \mathbb{C}v$ showing that the span of any nonzero v is invariant subspace of V and hence $V = \mathbb{C}v$. Therefore, all irreducible representations of an abelian group are 1-dimensional.

Now, recall the definition of *complex group algebra* $\mathbb{C}[G]$ which corresponds to a $|G|$ -dimensional complex vector space with basis G where multiplication is inherited from group multiplication, this is

$$\left(\sum_{g \in G} c_g g \right) \left(\sum_{h \in G} c_h h \right) := \sum_{q \in G} c_q q \quad \text{where } c_g c_h = c_q \text{ and } gh = q$$

which is well-defined since G is finite. Note that this algebra is associative, in general a representation V of an associative algebra A over \mathbb{C} is an algebra homomorphism $\rho: A \rightarrow \text{End}(V)$. An important remark is that any representation V of a finite-group G can be extended, by linearity, to a representation of the group algebra $\mathbb{C}[G]$. And conversely, we can restrict any representation of $\mathbb{C}[G]$ to G and get a representation of G . So we get a bijection

$$\text{Hom}_{\mathbb{C}\text{-alg}}(\mathbb{C}[G], \text{End}(V)) \longleftrightarrow \text{Hom}_{\text{Grp}}(G, \text{GL}(V)).$$

This establishes an equivalence between representations V of G and $\mathbb{C}[G]$ -modules, where irreducible representations correspond to simple modules and subrepresentations to submodules.

As we can build new vector spaces from old using basic operations, the same methodology applies in representation theory. Let V, W be representations of a group G we can define the representations $V \oplus W, V \otimes W, \text{Sym}^n V, \wedge^n V, \text{Hom}(V, W)$ and V^* where each is endowed with its usual underlying vector space structure, and the G -actions are obviously defined.

2.1. Isotypic decomposition. A natural question involving these operations and irreducible representations then arises: when can a general finite-dimensional representation be expressed as a direct sum of irreducible subrepresentations? This is answered by *Maschke's Theorem*, which states: *If W is a subrepresentation of V of a finite group G , then there exists a complementary*

subrepresentation W' of V so that $V = W \oplus W'$. From this result we obtain the following useful corollaries.

Corollary 2.4. *Every representation of a finite group can be decomposed as a direct sum of irreducible representations.*

Corollary 2.5 (Isotypic decomposition). *For any representation V of a finite group, we can write*

$$V \simeq V_1^{\oplus n_1} \oplus \cdots \oplus V_k^{\oplus n_k}$$

where V_i are distinct irreducible representations. The number of factors k and the V_i that occur (up to isomorphism) along with their multiplicities n_i are all unique.

In this context, a decomposition of the form $V \simeq W_1 \oplus \cdots \oplus W_k$, where $W_i \simeq V_i^{\oplus n_i}$, is called the *isotypic decomposition* of V . The subspace W_i is called the *V_i -isotypic component*, and the integer n_i is called the *multiplicity* of V_i in V .

Remark 2.6. The V_i -isotypic component of V can be characterized as the image of the evaluation map $\Phi: V_i \otimes \text{Hom}(V_i, V) \rightarrow V$, $(v, f) \mapsto f(v)$. Consequently, the isotypic decomposition can be written as $V \simeq \bigoplus_{i=1}^k V_i \otimes \text{Hom}(V_i, V)$. In this case, $V_i \otimes \text{Hom}(V_i, V)$ correspond to the *i -th isotypic decomposition*.

Let V be a representation of a finite group G and consider the isotypic decomposition of V made in the previous remark it can be proven that the projection of V onto the i -th isotypic decomposition is given by the formula

$$(2) \quad P_i = \frac{\dim V_i}{|G|} \sum \overline{\chi_i(g)} g \in \mathbb{C}[G]$$

where $\chi_i(g) = \text{Tr}(\rho_i(g): V_i \rightarrow V_i)$ correspond to the character of V_i . The characters of S_n are extremely well-studied, as for example in [2].

2.2. Irreducible representations of S_n . Lets start by recalling that two permutation $g, h \in S_n$ are conjugate if there exists a permutation $\pi \in S_n$ such that $g = \pi h \pi^{-1}$. This conjugacy classes are completely determined by cycle type which can be encoded as a partition. A *partition* λ of n is an integer sequence $\lambda = (\lambda_1, \dots, \lambda_k)$ such that $\lambda_1 \geq \dots \geq \lambda_k \geq 1$ and $\sum_i \lambda_i = n$. Let $P(n)$ denote the number of partitions of n We introduce the notation $|\lambda| = k$ for a given $\lambda = (\lambda_1, \dots, \lambda_k)$ and $\ell(\lambda)$ denote the number of nonzero λ_i . We introduce an order between partitions as follows, given two partitions $\lambda = (\lambda_1, \dots, \lambda_n), \mu = (\mu_1, \dots, \mu_n)$ we say that μ is *dominated by* λ writing $\mu \preceq \lambda$ if for every $k \geq 1, \mu_1 + \mu_2 + \dots + \mu_k \leq \lambda_1 + \lambda_2 + \dots + \lambda_k$. It is a well-known fact that the number of distinct irreducible representations of G is equal to the number of its conjugacy classes, so $P(n)$ correspond to the number of irreducible representations of S_n .

Given any partition λ of n may be represented as *Young diagram* this is collection of n boxes, or cells, arranged in left-justified rows with λ_i in the i -th row. A *Young Tableau* is a labeling of a Young diagram with the numbers $\{1, \dots, n\}$. For example, the following is a Young tableau for the partition $\lambda = (3, 3, 2, 1)$ with the *canonical labeling* corresponds to

1	2	3
4	5	
6	7	
8		

From now on, all Young tableaux will be considered with the canonical labeling. Given a Young tableau corresponding to a partition λ of n , we define two subgroups of S_n as

$$P_\lambda := \{g \in S_n : g \text{ preserves each row}\}, \quad Q_\lambda := \{g \in S_n : g \text{ preserves each column}\}.$$

Now, in the group algebra $\mathbb{C}[S_n]$, define the elements

$$a_\lambda = \sum_{g \in P_\lambda} g, \quad b_\lambda = \sum_{g \in Q_\lambda} \text{sgn}(g) g.$$

The *Young symmetrizer* is defined as the product $c_\lambda = a_\lambda b_\lambda$. We denote by $V_\lambda := \mathbb{C}[S_n]c_\lambda$ the image obtained by left multiplication of $\mathbb{C}[S_n]$ on c_λ . The following theorem characterizes the irreducible representations of S_n via the modules V_λ .

Theorem 2.7. *For any given partition λ of n the subspace $\mathbb{C}[S_n]c_\lambda$ by left multiplication on $\mathbb{C}[S_n]$ is an irreducible representation of S_n . Furthermore, every irreducible representation of S_n is isomorphic V_λ for some partition λ of n .*

Example 2.8. We will compute the irreducible representation of S_3 . The following are the canonical Young tableau for $n = 3$

$$\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline \end{array} \quad \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & \\ \hline \end{array} \quad \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array}$$

For $\lambda = (3)$, we have $P_{(3)} = S_3$ and $Q_{(3)} = \{1\}$. Hence, the corresponding Young symmetrizer is $c_{(3)} = \sum_{g \in S_3} g$. One can verify that for all $h \in S_3$, we have $hc_{(3)} = c_{(3)}$, so the corresponding module $V_{(3)}$ is the trivial representation.

For $\lambda = (1, 1, 1)$, we have $P_{(1,1,1)} = \{1\}$ and $Q_{(1,1,1)} = S_3$. Thus, $c_{(1,1,1)} = \sum_{g \in S_3} \text{sgn}(g)g$ and for any $h \in S_3$, we find $hc_{(1,1,1)} = \text{sgn}(h)c_{(1,1,1)}$. Therefore, $V_{(1,1,1)}$ is the sign representation.

The case $\lambda = (2, 1)$ is slightly more involved. We have $P_{(2,1)} = \{1, (12)\} \simeq S_2$ and $Q_{(2,1)} = \{1, (13)\} \simeq S_2$. Then, $c_{(2,1)} = (1 + (12))(1 - (13)) = 1 + (12) - (13) - (132)$. It can be shown that $V_{(2,1)}$ is spanned by $c_{(2,1)}$ and $(13)c_{(2,1)}$, so it is a 2-dimensional irreducible representation. Since S_3 has exactly three irreducible representations (as seen from its character table), the representation $V_{(2,1)}$ must be the standard representation.

The dimension of the irreducible representations V_λ can be calculated with the formula

$$(3) \quad \dim V_\lambda = \frac{n!}{\prod h(i, j)}$$

where the hook length $h(i, j)$ corresponds to number of cells directly to the right of the cell (i, j) + number of cells directly below of the cell (i, j) + 1. For example for the partition $\lambda = (4, 3, 1, 1)$ of $n = 9$ we have that $\dim V_\lambda = 216$.

2.3. Schur-Weyl duality. Let V be a complex vector space and denote as $V^{\otimes n}$ the n -th tensor power. Over this space we have two canonical actions defined as follows. For $v_1 \otimes \cdots \otimes v_n \in V^{\otimes n}$ and $g \in S_n$ define the (right) action of S_n as $(v_1 \otimes \cdots \otimes v_n)g = v_{g(1)} \otimes \cdots \otimes v_{g(n)}$ and for $f \in \text{GL}(V)$ define the (left) action as $f(v_1 \otimes \cdots \otimes v_n) = f(v_1) \otimes \cdots \otimes f(v_n)$, extending by linearity we induce a structure of $(\text{GL}(V)\text{-}S_n)$ -bimodule over $V^{\otimes n}$.

For the following let λ be a partition of n and consider the operators $a_\lambda: V^{\otimes n} \rightarrow V^{\otimes n}$ and $b_\lambda: V^{\otimes n} \rightarrow V^{\otimes n}$ given by (right) multiplication. The following lemma characterize the image of this operators.

Lemma 2.9. *Let V be a finite dimensional complex vector space and $\lambda = (\lambda_1, \dots, \lambda_k)$ be a partition of n . Then $a_\lambda(V^{\otimes n}) = \text{Sym}^{\lambda_1} V \otimes \cdots \otimes \text{Sym}^{\lambda_k} V$ and $b_\lambda(V^{\otimes n}) = \bigwedge^{\lambda'_1} V \otimes \cdots \otimes \bigwedge^{\lambda'_k} V$, where $\lambda' = (\lambda'_1, \dots, \lambda'_n)$ correspond to conjugate partition obtained by swapping the rows and the columns of λ .*

As we do for a_λ and b_λ let λ be a partition of n . Define the *Young symmetrizer* as the product $c_\lambda = a_\lambda b_\lambda$ and consider the operator $c_\lambda: V^{\otimes n} \rightarrow V^{\otimes n}$ and denote their image as $\mathbb{S}_\lambda(V) = c_\lambda(V^{\otimes n})$. This assignment define a functor $\mathbb{S}_\lambda: \mathbf{Vect}_{\mathbb{C}} \rightarrow \mathbf{Vect}_{\mathbb{C}}$ to the category of complex vector spaces to itself. Moreover, it turns out that $\mathbb{S}_\lambda(V)$ is irreducible as $\text{GL}(V)$ -module, although not every irreducible representation of $\text{GL}(V)$ can be decomposed into the images \mathbb{S}_λ of V under Schur functors. We can prove that the actions provided above for $V^{\otimes n}$ commute each other and their span in $\text{End}(V^{\otimes n})$ are centralizers of each other, then the Double Commutant theorem applies together with Corollary 2.5 we get a decomposition of $V^{\otimes n}$ into irreducible representations of $S_n \times \text{GL}(V)$ as follows.

Theorem 2.10 (Schur-Weyl duality). *Let V be a finite dimensional complex vector space. Then the n -fold tensor power $V^{\otimes n}$ can be decomposed as*

$$V^{\otimes n} \simeq \bigoplus_{\substack{|\lambda|=n \\ \ell(\lambda) \leq \dim V}} V_\lambda \otimes \mathbb{S}_\lambda(V)$$

as a representation of $S_n \times \text{GL}(V)$ where V_λ runs through all the irreducible representations of S_n and each $\mathbb{S}_\lambda(V)$ is an irreducible representation of $\text{GL}(V)$ or is zero.

As in Equation (3), we can provide an explicit formula for the irreducible representations \mathbb{S}_λ of $\text{GL}(V)$, known as the *hook-content formula*:

$$(4) \quad \dim \mathbb{S}_\lambda(V) = \frac{\dim V + c(i, j)}{\prod h(i, j)}$$

where $h(i, j)$ denotes the hook length of the box (i, j) in the Young diagram of λ , and the *content* $c(i, j)$ is given by $c(i, j) = j - i$. In particular, $c(i, j) = 0$ for boxes on the main diagonal, it equals $+k$ for boxes on the k -th diagonal above the main diagonal, and $-k$ for boxes on the k -th diagonal below the main diagonal.

In order to present a basic example we first note the following basic fact: Note that if U is a complex vector space of dimension m (without any group action) and W is a representation of a group G , then $U \otimes W \simeq W^{\oplus m}$ as representations of G . Applying this to $V^{\otimes n}$ with only the action

of $GL(V)$, we obtain the decomposition

$$(5) \quad V^{\otimes n} \simeq \bigoplus_{|\lambda|=n} (\mathbb{S}_\lambda(V))^{\oplus f_\lambda}$$

where $f_\lambda = \dim V_\lambda$. For further details in the following examples we refer [2].

Example 2.11. We compute the decomposition of $V^{\otimes n}$ given by Schur-Weyl duality for a specific values of n .

- (1) For $n = 2$, the partitions of 2 are (2) and (1, 1). Analogous to Example 2.8 the irreducible representations $V_{(2)}$ and $V_{(1,1)}$ of S_2 correspond to the trivial and sign representations, both of dimension 1. On the other hand, we can compute the corresponding Schur functor images $\mathbb{S}_{(2)}(V)$ and $\mathbb{S}_{(1,1)}(V)$, note that in this case $c_{(2)} = a_{(2)}$ and $c_{(1,1)} = b_{(1,1)}$, now Lemma 2.9 applies and thus following the equation (5) we got $V^{\otimes 2} \simeq \text{Sym}^2 V \oplus \bigwedge^2(V)$.
- (2) For $n = 3$, the partitions of 3 are (3), (2, 1) and (1, 1, 1). By Example 2.8 the irreducible representations $V_{(3)}$ and $V_{(1,1,1)}$ of S_3 correspond to the trivial and sign representations, both of dimension 1. As we already mentioned, $V_{(2,1)}$ has dimension 2 so $f_{(2,1)} = 2$. Now we claim that $\mathbb{S}_{(2,1)}(V) = \text{span} \{(v_1 \wedge v_3) \otimes v_2 + (v_2 \wedge v_3) \otimes v_1\}$, for that consider the isomorphism $\varphi: V^{\otimes 3} \rightarrow V^{\otimes 2} \otimes V$ and the natural embedding $i: \bigwedge^2 V \otimes V \rightarrow V^2 \otimes V$ defined as $(v_1 \wedge v_3) \otimes v_2 \mapsto v_1 \otimes v_2 \otimes v_3 - v_3 \otimes v_2 \otimes v_1$, under the composition $\varphi^{-1} \circ i$ we can see

$$\varphi^{-1} \circ i((v_1 \wedge v_3) \otimes v_2 + (v_2 \wedge v_3) \otimes v_1) = c_{(2,1)}(v_1 \otimes v_2 \otimes v_3),$$

which proves the claim. Furthermore, this vector also spans the kernel of the canonical map $\bigwedge^2 V \otimes V \rightarrow \bigwedge^3 V$, so we have $\mathbb{S}_{(2,3)}(V) \simeq \ker(\bigwedge^2 V \otimes V \rightarrow \bigwedge^3 V)$. Finally, by Lemma 2.9 and following the equation (5) we got $V^{\otimes 3} \simeq \text{Sym}^2 V \oplus \bigwedge^3 V \oplus \ker(\bigwedge^2 V \otimes V \rightarrow \bigwedge^3 V)^{\oplus 2}$.

3. SCHUR TRANSFORMATION AND SPECTRUM ESTIMATION THEOREM

In this section we explore the *Schur basis* arising from Schur–Weyl duality on the system of n d -dimensional quantum systems, i.e., n -qudits $(\mathbb{C}^d)^{\otimes n}$. In this case restrict the action from $GL_d(\mathbb{C})$ to $U(d)$ where denote the group of $d \times d$ unitary operators, because in quantum mechanics the state space of a qudit is a Hilbert space with an inner product, and the physically allowed evolutions must preserve that inner product.

In the following we will restrict the irreducible representations of $U(d)$ to the group of orthogonal matrices with determinant 1, $SU(d)$. In representation theory, the group $SU(2)$ has irreducible representations indexed by a half-integer J , each spin- J representation having dimension $2J + 1$, corresponding to the familiar addition of angular momentum in quantum mechanics, where the fundamental representation (spin- $\frac{1}{2}$) is 2-dimensional. Following Example 2.11 and taking $(n = 2, d = 2)$, the two-qubit space decomposes as

$$(\mathbb{C}^2)^{\otimes 2} = (Q_1 \otimes P_{\text{triv}}) \oplus (Q_0 \otimes P_{\text{sign}})$$

where P_{triv} and P_{sign} denote the 1-dimensional symmetric and antisymmetric irreducible representations of S_2 , and Q_1 and Q_0 are the spin-1 (triplet) and spin-0 (singlet) subspaces, respectively. Similarly, the Hilbert space of three qubits $(n = 3, d = 2)$ decomposes as

$$(\mathbb{C}^2)^{\otimes 3} = (Q_{3/2} \otimes P_{\text{triv}}) \oplus (Q_{1/2} \otimes P_{2,1}),$$

where $Q_{3/2}$ and $Q_{1/2}$ are the spin-3/2 and spin-1/2 subspaces, and $P_{2,1}$ denotes the 2-dimensional mixed-symmetry irrep of S_3 . This illustrates how the total spin of multiple qubits arises from combining the fundamental spin- $\frac{1}{2}$ representations while keeping track of permutation symmetries.

This previous decomposition, obtained from the Schur-Weyl duality is natural to introduce a basis that respects both the $SU(d)$ and S_n symmetries. This is the *Schur basis*, whose elements are labeled as $|\lambda, p_\lambda, q_\lambda\rangle$, where λ indexes the irreducible representations of $SU(d)$, p labels a basis within the S_n irrep corresponding to the symmetry type, and q labels a basis within the $SU(d)$ irrep. Now, we introduce some notation. We fix the standard computational basis $\{|i\rangle\}_{i=1}^d$ for the state space \mathbb{C}^d of each qudit. From this, the basis for the system $(\mathbb{C}^d)^{\otimes n}$ is then $|i_1\rangle \otimes \cdots \otimes |i_n\rangle = |i_1 \dots i_n\rangle$ where $i_k = 0, \dots, d$.

We define a circuit called schur transformation, denoted as U_{Sch} which is constructed recursively using the Clebsch-Gordan (CG) transformation for $U(d)$. The core mathematical step is the decomposition of the tensor product of an irrep of $U(d)$ with the fundamental representation. Let V_λ be the irrep of $U(d)$ associated with partition λ . The tensor product with the standard module \mathbb{C}^d decomposes as

$$V_\lambda \otimes \mathbb{C}^d \cong \bigoplus_{\mu} V_\mu$$

where the direct sum is over all partitions μ obtained from λ by removing a single box in accordance with the branching rules for $U(d) \supseteq U(d-1)$. This decomposition is multiplicity free. The Clebsch-Gordan transform, U_{CG} , is the unitary matrix whose matrix elements are the Clebsch-Gordan coefficients. It performs the following basis change:

$$U_{\text{CG}} : \bigoplus_{\lambda} V_\lambda \otimes \mathbb{C}^d \rightarrow \bigoplus_{\mu} V_\mu.$$

In other words it maps the input state $|\lambda, q_\lambda\rangle \otimes |i\rangle$ (where $|i\rangle$ is a computational basis state for the new qudit) to an output state which is a superposition over the valid μ

$$U_{\text{CG}} |\lambda, q_\lambda\rangle |i\rangle = \sum_{\mu} \sum_{q_\mu} C_{\lambda, q_\lambda; i}^{\mu, q_\mu} |\mu, q_\mu\rangle.$$

The coefficients $C_{\lambda, q_\lambda; i}^{\mu, q_\mu}$ are the Clebsch-Gordan coefficients, which are computed classically, and well studied, as for example [1]. In the full Schur transform circuit, this U_{CG} step is applied sequentially. One starts with the first two qudits, applies U_{CG} , then takes the resulting state and the third qudit and applies U_{CG} again, and so forth for all n qudits.

For the following theorem we start by defining the *relative entropy* which is a measure of how different two probability distributions are. For two discrete probability distributions $p = (p_i)_i$ and $q = (q_i)_i$ over the same set, it is defined as

$$H(p||q) = \sum_i p_i \log \frac{p_i}{q_i}.$$

Equivalently, one may write

$$H(p||q) = - \sum_i p_i \log \frac{q_i}{p_i} = -H(p) - \sum_i p_i \log(q_i),$$

where $H(p) = -\sum_i p_i \log p_i$ is the Shannon entropy of p . Relative entropy is always nonnegative, and it vanishes if and only if $p = q$. It can be interpreted as some sort of distance between p and q , quantifying how inefficient it is to assume that the distribution is q when the true distribution is p .

Now we introduce the *Spectrum Estimation Theorem* which provides bounds on how a tensor power $\rho^{\otimes n}$ of a density operator ρ is distributed across the isotypic components of the Schur-Weyl decomposition. In this setting the projector operator defined in (2) measures the weight of $\rho^{\otimes n}$ in the component associated to λ . Formally the theorem states:

Theorem 3.1 (Spectrum Estimation Theorem). *Let $\dim V = n$, and let $\rho: V \rightarrow V$ be a density operator. Let $|\lambda| = d$, assume $\ell(\lambda) \leq n$ and in what follows, if $\ell(\lambda) < n$ we add zero to it so we may write $\lambda = (\lambda_1, \dots, \lambda_n)$. Let $P_\lambda: V^{\otimes d} \rightarrow V_\lambda \otimes \mathbb{S}_\lambda(V)$ be the isotypic projection defined in (2). Write $\text{spec}(\rho) = (r_1, \dots, r_n)$, then*

$$\begin{aligned} \text{Tr}(P_\lambda(\rho^{\otimes d})) &\leq (d+1) \binom{n}{2} 2^{-dH(\bar{\lambda} \parallel \text{spec}(\rho))} \\ &= (d+1) \binom{n}{2} 2^{dH(\lambda)} 2^{\sum_i \lambda_i \log(r_{e_i})}. \end{aligned}$$

Thus, the theorem shows that $\rho^{\otimes n}$ is concentrated in those isotypic components $V_\lambda \otimes \mathbb{S}_\lambda(V)$ for which the partition $\bar{\pi}$ is close to the spectrum of ρ . In other words, the Schur-Weyl decomposition acts as a tool to estimate the eigenvalues of ρ by examining which components π carry significant weight.

Before providing a proof, we introduce some notation. Consider $\{e_1, \dots, e_n\}$ as a basis for V and thus we denote $e_I = e_{i_1} \otimes \dots \otimes e_{i_d}$ the induced basis of $V^{\otimes d}$. The *weight* of e_I is defined as $\text{wt}(e_I) = (w_1, \dots, w_n)$ where w_j is the number of i_i 's equal to j . As a toy example consider $n = 3$ and $d = 4$, choose $e_I = e_1 \otimes e_3 \otimes e_1 \otimes e_2$ so $\text{wt}(e_I) = (2, 1, 1)$.

Lemma 3.2. *Let $\{e_1, \dots, e_n\}$ and let $e_I = e_{i_1} \otimes \dots \otimes e_{i_d}$ the induced basis for $V^{\otimes d}$. Let $w = (w_1, \dots, w_n)$ the weight of e_I and fix a partition λ of d , then if $w \preceq \lambda$ then $P_\lambda(e_I) \neq 0$.*

Sketch of the proof. To show that the projection is nonzero, it suffices to find one operator mapping into the λ -isotypic component that gives a nonzero result. Recall that the Young symmetrizer $c_\lambda = a_\lambda b_\lambda$ projects onto the corresponding Schur module. Thus, if there exists a tableau T of shape λ such that $c_T(e_I) \neq 0$, then e_I has a nonzero component in the λ -isotypic subspace, and hence $P_\lambda(e_I) \neq 0$. It is known that $c_T(e_I) \neq 0$ precisely when any two entries k, l in the same column of T correspond to distinct indices in e_I . Now, when the weight $w = \text{wt}(e_I)$ satisfies $w \preceq \lambda$, there exists a *semistandard Young tableau* of shape λ and content w , that is, a filling of λ whose rows are weakly increasing and columns strictly increasing. Such a tableau guarantees that the column indices in e_I are distinct, so $c_T(e_I) \neq 0$. Therefore e_I has a nonzero projection onto the λ -isotypic component and so $P_\lambda(e_I) \neq 0$. \square

Proof Theorem 2.10. Write $r = \text{spec}(\rho) = (r_1, \dots, r_n)$ with $r_j \geq r_{j+1}$. Since ρ is a density operator, by the Spectral Decomposition Theorem [6, Theorem 2.1] we can write $\rho = \sum_{i=1}^n r_i |e_i\rangle \langle e_i|$ and hence

$$\rho^{\otimes d} = \sum_{I=(i_1, \dots, i_d)} r_{i_1} \cdots r_{i_d} |e_{i_1} \otimes \dots \otimes e_{i_d}\rangle \langle e_{i_1} \otimes \dots \otimes e_{i_d}| = \sum_I r_I |e_I\rangle \langle e_I|$$

where $r_I = r_{i_1} \cdots r_{i_d}$ and $|e_I\rangle = |e_{i_1}\rangle \otimes \dots \otimes |e_{i_d}\rangle$. By the previous lemma, the eigenvectors that do not project to zero in $V_\pi \otimes \mathbb{S}_\pi(V)$ are those satisfying $\text{wt}(e_I) \preceq \pi$. If $\text{wt}(e_I) = \pi$ then

$r_I = r_1^{\pi_1} \cdots r_n^{\pi_n}$, and generally $\text{wt}(e_I) \preceq \pi$ implies $r_I \leq r_1^{\pi_1} \cdots r_n^{\pi_n}$. Thus

$$P_\pi(\rho^{\otimes d}) = \sum_{\text{wt}(e_I) \preceq \pi} r_I P_\pi(|e_I\rangle \langle e_I|) \leq r_1^{\pi_1} \cdots r_n^{\pi_n} \sum_{\text{wt}(e_I) \preceq \pi} P_\pi(|e_I\rangle \langle e_I|).$$

Taking the trace and using positivity and monotonicity of trace we got

$$\text{Tr}(P_\pi \rho^{\otimes d}) \leq r^\pi \text{Tr}\left(P_\pi \sum_{\text{wt}(e_I) \preceq \pi} |e_I\rangle \langle e_I|\right) \leq r^\pi \text{Tr}(P_\pi)$$

where $r^\pi = \prod_{j=1}^n r_j^{\pi_j}$. Since P_π projects onto the π -isotypic component $V_\pi \otimes \mathbb{S}_\pi(V)$, we have

$$\text{Tr}(P_\pi) = \dim(V_\pi \otimes \mathbb{S}_\pi(V)) = \dim V_\pi \cdot \dim \mathbb{S}_\pi(V)$$

From [5, Exercise 7.1.21] and [5, Equation (7.1.7)] we have $\dim \mathbb{S}_\pi(V) \leq (d+1)^{\binom{n}{2}}$ and $\dim V_\pi \leq 2^{dH(\bar{\pi})}$ so that

$$\text{Tr}(P_\pi \rho^{\otimes d}) \leq (d+1)^{\binom{n}{2}} 2^{dH(\bar{\pi})} r^\pi.$$

Let $p = \bar{\pi} = (\pi_1/d, \dots, \pi_n/d)$. Then $\log_2 r^\pi = \sum_j \pi_j \log_2 r_j = d \sum_j p_j \log_2 r_j = -dH(p) - dD(p||r)$ hence $r^\pi = 2^{-dH(p)} 2^{-dD(p||r)}$. Now substituting this into the previous bound cancels the entropy factors and gives

$$\text{Tr}(P_\pi \rho^{\otimes d}) \leq (d+1)^{\binom{n}{2}} 2^{-dD(\bar{\pi}||r)}.$$

□

4. QUANTUM MACHINE LEARNING

In this section we introduce an application with all the representation theory machinery previously developed for quantum computing. As a reference we will present the basic ideas introduced in [7]. Before that we define some needed concepts about Lie algebra and lie groups representations.

We start by recalling some basic facts. A *Lie algebra* is a complex vector space \mathfrak{g} with a *lie bracket* $[\cdot, \cdot]: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, which satisfies the following axioms: For $X_1, X_2, X_3 \in \mathfrak{g}$ and $a, b \in \mathbb{C}$ we have

- (1) *Antisymmetry*: $[X_1, X_2] = -[X_2, X_1]$. A direct consequence of this is that $[X, X] = 0$ for any $X \in \mathfrak{g}$.
- (2) *Bilinearity*: $[aX_1 + bX_2, X_3] = a[X_1, X_3] + b[X_2, X_3]$ and $[X_1, aX_2 + bX_3] = a[X_1, X_2] + b[X_1, X_3]$.
- (3) *Jacobi Identity*: $[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0$.

We say that two elements $X_1, X_2 \in \mathfrak{g}$ *commute* if their Lie bracket is zero. A *Lie subalgebra* is a linear subspace $\mathfrak{h} \subseteq \mathfrak{g}$ which is closed under the Lie bracket. The following are basic examples of Lie algebras.

Example 4.1. (1) We define *Trivial Lie algebra* as any complex vector space V by setting $[x, y] = 0$ for all $x, y \in V$.

- (2) We define *Matrix Lie algebras*. We start first by defining *General linear algebra* $\mathfrak{gl}_n(\mathbb{C})$, the space of $n \times n$ matrices with the bracket $[A, B] = AB - BA$. As a Lie subalgebras we got *Special linear algebra*, $\mathfrak{sl}_n = \{A \in \mathfrak{gl}_n(\mathbb{C}) : \text{Tr}(A) = 0\}$ and *Orthogonal algebra*, $\mathfrak{so}_n(\mathbb{C}) = \{A \in \mathfrak{gl}_n(\mathbb{C}) : A^T = -A\}$.

Next we define *(Matrix) Lie groups* as a closed subgroup of $GL_d(\mathbb{C})$ where by closed we mean that if $\{A_m\} \subseteq G$ is a sequence of matrices with $\lim_{m \rightarrow \infty} A_m = A \in GL_d(\mathbb{C})$, then $A \in G$. The following are basic examples of Matrix Lie groups.

- Example 4.2.** (1) *Special linear group*, $SL_d(\mathbb{C}) = \{A \in GL_d(\mathbb{C}) : \det(A) = 1\}$.
 (2) *Orthogonal group*, $O_d(\mathbb{C}) = \{A \in GL_d(\mathbb{C}) : A^T A = I\}$.
 (3) *Special orthogonal group*, $SO_d(\mathbb{C}) = O_d(\mathbb{C}) \cap SL_d(\mathbb{C})$.
 (4) *Unitary group*, $U_d(\mathbb{C}) = \{A \in GL_d(\mathbb{C}) : A^* A = I\}$.

Given any (matrix) lie group G we define the *associated lie algebra* algebra as

$$(6) \quad \mathfrak{g} = \{X \in M_n(\mathbb{C}) : e^{tX} \in G \text{ for all } t \in \mathbb{R}\}$$

with bracket $[X, Y] = XY - YX$. In the above definition, we note that we can recover any $X \in \mathfrak{g}$ by starting with a one-parameter subgroup e^{tX} and taking derivatives at the identity e^0 , this is, $X = \left. \frac{d}{dt}(e^{tX}) \right|_{t=0}$. The following are examples of associated lie algebras for given (matrix) lie groups.

Example 4.3. Given,

- (1) For $G = GL_n(\mathbb{C})$, the Lie algebra is $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{C}) = M_n(\mathbb{C})$.
- (2) For $G = SL_n(\mathbb{C})$, the Lie algebra is $\mathfrak{g} = \mathfrak{sl}_n(\mathbb{C}) = \{X \in M_n(\mathbb{C}) : \text{Tr}(X) = 0\}$.
- (3) For $G = SO_d(\mathbb{C})$, the Lie algebra is $\mathfrak{g} = \mathfrak{so}_d(\mathbb{C}) = \{X \in M_d(\mathbb{C}) : X^T + X = 0\}$.
- (4) For $G = SU_d$, the Lie algebra is $\mathfrak{g} = \mathfrak{su}(d) = \{X \in M_d(\mathbb{C}) : X^* + X = 0 \text{ and } \text{Tr}(X) = 0\}$.

Let \mathfrak{g} be a Lie algebra and V a finite dimensional complex vector space. A *representation of \mathfrak{g}* action on V is a linear map $r : \mathfrak{g} \rightarrow \mathfrak{gl}(V)$ such that $r([X, Y]) = [r(X), r(Y)]$ for all $X, Y \in \mathfrak{g}$. There are strategies to relate representations of Lie groups and their Lie algebras.

Theorem 4.4. *Let R be a matrix Lie group representation of G . One can always induce a representation r of the associated Lie algebra defined in (6) by setting $r(X) = \left. \frac{d}{dt} R(e^{tX}) \right|_{t=0}$ for $X \in \mathfrak{g}$. Conversely, if G is simply connected, then every representation r of the Lie algebra \mathfrak{g} uniquely determines a representation R of G , satisfying $R(e^X) = e^{r(X)}$ for all $X \in \mathfrak{g}$.*

In the last case, for a general matrix Lie group G , even if it is not simply connected, any Lie algebra representation r can still be lifted *locally* to a representation of G near the identity element, via the formula $R(g) = e^{r(X)}$ whenever g^{e^X} close to the identity. So, if G is simply connected, there is a one-to-one correspondence between their representations.

Example 4.5. Recall that $SU(2) = \{U \in GL(\mathbb{C}^2) : U^\dagger U = 1 \text{ and } \det(U) = 1\}$. Taking $V = \mathbb{C}^2$ we define the *fundamental representation* $U : G \rightarrow GL(V)$ is given by $U_g = g$. We illustrate the Theorem 4.4 by inducing a representation of $\mathfrak{su}(2)$. Following the fact that $\mathfrak{su}(2)$ is a 3 dimensional real vector space. It thus suffices to find 3 linearly independent tangent vectors, which we can do by taking derivatives of parameterized paths and evaluating at the identity. We choose

$$\left. \frac{d}{d\theta} e^{-i\theta X/2} \right|_{\theta=0} = \frac{-i}{2} X, \quad \left. \frac{d}{d\theta} e^{-i\theta Y/2} \right|_{\theta=0} = \frac{-i}{2} Y, \quad \left. \frac{d}{d\theta} e^{-i\theta Z/2} \right|_{\theta=0} = \frac{-i}{2} Z.$$

Since the Pauli matrices are linearly independent, these three resulting matrices form a basis for the representation $r : \mathfrak{su}(2) \rightarrow \mathfrak{gl}(V)$.

Example 4.6 (Adjoint representation of $SU(2)$). Let $V = M_2(\mathbb{C})$ be the space of 2×2 complex matrices. Consider the operator given by conjugation $A \mapsto U_g A U_g^\dagger$ where $U_g = g \in SU(2)$ is in $GL(V)$. The *adjoint representation* of $SU(2)$ is defined as $gA = U_g A U_g^\dagger$. It can be proven that $\{X, Y, Z, 1\} \subseteq V$ forms an orthonormal basis of V with respect to the Hilbert Schmidt inner product $\langle A, B \rangle_{HS} = \frac{1}{2} \text{Tr}[A^\dagger B]$. This representation can be decomposed under two invariant subspaces $V = \text{span}\{X, Y, Z\} \oplus \text{span}\{1\}$. Since, $\mathfrak{su}(2) = \text{span}\{X, Y, Z\}$ we got $V = \mathfrak{su}(2) \oplus \mathbb{C}1$. Then we can do the same as the previous example, to lift this representation to a lie algebra representation of $\mathfrak{su}(2)$ by taking derivatives.

With the representation theory developed so far, we now briefly introduce how these ideas can be applied in the context of quantum machine learning. *Quantum Machine Learning* (QML) is an emerging field at the intersection of quantum computing and machine learning (ML). The idea is to use the principles of quantum mechanics, such that superposition, entanglement, and quantum circuits, to design algorithms that can process and analyze data in ways that classical computers cannot do efficiently. In the following we will describe *supervised QML* problem of classifying labeled quantum data.

In a supervised QML classification task, one is given a dataset of the form $\{\rho_i, y_i\}$ with $\rho_i = |\rho_i\rangle\langle\rho_i|$ and where $|\rho_i\rangle \in \mathcal{R} \subseteq \mathcal{H}$ are pure n -qubit quantum states in a set \mathcal{R} belonging to the d -dimension Hilbert space \mathcal{H} (with $d = 2^n$) and $y_i \in \mathcal{Y}$ are labels associated with each state ρ_i according to some unknown function $f: \mathcal{R} \rightarrow \mathcal{Y}$. That is, $f(\rho_i) = y_i$. We start by taking a data set $\mathcal{S} = \{\rho_i, y_i\}_{i=1}^N$ in which we have repeated access to, sampling states from the dataset according to a given i.i.d probability distribution over $\mathcal{R} \times \mathcal{S}$. Then, one trains a model $h_\theta: \mathcal{R} \rightarrow \mathcal{S}$ where θ are trainable parameters where the objective is for the model's prediction $h_\theta(\rho_i)$, to match the true label y_i as closely as possible, this is, with high probability.

Geometric Quantum Machine Learning (GQML) is a newer subfield of quantum machine learning (QML) that focuses on exploiting the geometric and symmetry structures of data and quantum states, this is, the goal of the GQML is to embed about the symmetries of the data in \mathcal{S} into the model. In the following we will focus in the QML model

$$(7) \quad h_\theta(\rho_i) = \text{Tr} [\mathcal{W}_\theta(\rho_i^{\otimes k}) M_i].$$

where $\mathcal{W}_\theta: B(\mathcal{H}^{\otimes k}) \rightarrow B(\mathcal{H}^{\otimes k'})$ is a trainable parameterized quantum channel and M_i is a Hermitian measurement operator. In general *parameterized quantum channel* is a family of completely positive, trace-preserving of maps $\mathcal{W}_\theta: \mathcal{B}(\mathcal{H}_{int}) \rightarrow \mathcal{B}(\mathcal{H}_{out})$ that depends smoothly on a set of trainable parameters, where each $\mathcal{W}_\theta(\rho) = \sum_j K_j(\theta) \rho K_j^\dagger(\theta)$ with $\sum_j K_j(\theta) K_j^\dagger(\theta) = I$

Now, a *symmetry* refers to a property of the data or of the target function which remains invariant under a given transformation. In the quantum setting, such transformations are naturally modeled by unitary operators acting on the state space, by maps of the form $\rho \mapsto U \rho U^\dagger$ for some unitary U . In order to study those symmetries we note that the set of all unitary symmetry transformations has a structure where representation theory can help with; a group.

Proposition 4.7. *Let G be the set of all unitary symmetry transformations, such that for any $U \in G$, the map $\rho \mapsto U \rho U^\dagger$ leaves some property of ρ unchanged. Then, G forms a group.*

Furthermore, in the above definition, the action of the group G is said to leave the data labels y_i invariant if $f(U \rho_i U^\dagger) = f(\rho_i) = y_i$ for all ρ_i with label y_i and for all $U \in G$. Note that if the states

themselves are invariant $U\rho_i U^\dagger = \rho_i$ then label invariance is vacuously satisfied. Now we need to ensure the compatibility of G in the QML h_θ as follows:

- (1) Equivariance under G : $\mathcal{W}_\theta(U^{\otimes k} \rho_i^{\otimes k} (U^\dagger)^{\otimes k}) = U^{\otimes k'} \mathcal{W}_\theta((\rho_i)^{\otimes k}) (U^\dagger)^{\otimes k'}$ for all $U \in G$ and,
- (2) Equivariance of the measurement operator: $[M, U^{\otimes k}] = 0$ for all $U \in G$.

With these conditions now one can prove that the model h_θ is invariant under the action of G :

$$\begin{aligned} h_\theta(U\rho_i U^\dagger) &= \text{Tr} [\mathcal{W}_\theta((U\rho_i U^\dagger)^{\otimes k}) M_i] \\ &= \text{Tr} [U^{\otimes k'} \mathcal{W}_\theta((\rho_i)^{\otimes k}) (U^\dagger)^{\otimes k'} M_i] \\ &= \text{Tr} [\mathcal{W}_\theta((\rho_i)^{\otimes k}) M_i] = h_\theta(\rho_i). \end{aligned}$$

Lets introduce some basic examples of QML tasks.

Example 4.8 (Purity classification). The goal is to classify single-qubit states according to their *purity* (pure vs. mixed). The label depends only on the eigenvalues of the density matrix ρ , which remain invariant under any unitary conjugation $\rho \mapsto U\rho U^\dagger$. Hence, the symmetry group is $G = SU(2)$.

For a single copy, in a conventional experiment, we may classify the data by computing $h(\rho_i) = \text{Tr}[\rho_i M]$. Here, $SU(2)$ acts on \mathbb{C}^2 through its fundamental representation (Example 4.5) which is irreducible. Consequently, $SU(2)$ admits no non-trivial block-diagonal structure on \mathbb{C}^2 . By Schur's Lemma, any measurement operator commuting with all $U \in SU(2)$ satisfy $[M, U] = 0$ for all $U \in SU(2)$ then $M = cI$. Therefore, the model $h(\rho_i) = \text{Tr}[\rho_i M]$ cannot distinguish between pure and mixed states, since every $SU(2)$ -invariant measurement acts trivially.

For two copies, in a quantum-enhanced setting, we classify the data using $h(\rho_i) = \text{Tr}[\rho_i^{\otimes 2} M]$. Now $SU(2)$ acts through the tensor representation $U \otimes U$, which is reducible and admits a block-diagonal structure in the symmetric-antisymmetric basis (Example 2.11)

$$\mathbb{C}^2 \otimes \mathbb{C}^2 \simeq \underbrace{\text{Sym}^2(\mathbb{C}^2)}_{\text{triplet, spin-1}} \oplus \underbrace{\wedge^2(\mathbb{C}^2)}_{\text{singlet, spin-0}}.$$

The operators M that commute with this $SU(2)$ action have the form $M = c_0 \mathbb{1}_3 \oplus c_1 \mathbb{1}_1$, acting independently on the triplet and singlet components. This explicit block-structure, absent in the fundamental representation, allows the model to access information about purity.

Example 4.9 (Distinguish ferromagnetic from antiferromagnetic two-qubit states). The goal is to distinguish between *ferromagnetic* (aligned) and *antiferromagnetic* (anti-aligned) two-qubit states. The dataset consists of density operators ρ that are invariant under global $SU(2)$ rotations, that is

$$\rho \mapsto (U \otimes U) \rho (U^\dagger \otimes U^\dagger)$$

so the symmetry group is $G = SU(2)$ acting via the tensor-product representation $U \otimes U$. The tensor product of two fundamental representations of $SU(2)$ decomposes as $U \otimes U \simeq U_{\text{triplet}} \oplus U_{\text{singlet}}$ corresponding respectively to the spin-1 (symmetric) and spin-0 (antisymmetric) subspaces. According to Schur-Weyl duality, any operator commuting with the global action of $SU(2)$ must be block-diagonal across these invariant subspaces.

Now we note that projectors onto the symmetric and antisymmetric subspaces are given by

$$M_{\text{ferro}} = \frac{1}{2}(I + \text{SWAP}), \quad M_{\text{antiferro}} = \frac{1}{2}(I - \text{SWAP}).$$

The corresponding measurement probabilities $\text{Tr}[\rho M_{\text{ferro}}]$ for ferromagnetic and $\text{Tr}[\rho M_{\text{antiferro}}]$ for antiferromagnetic provides the likelihood that a two-qubit state lies in either subspace.

In the following we relate all the representation theory of Lie groups/Lie algebras introduced before to help in the tasks of QML.

Consider the task of classifying two-dimensional classical data points $x = (x^1, x^2)$ into two classes. Each coordinate is encoded into the state of a qubit via a single-qubit rotation $R(x^j)$, and the two-qubit system is then processed by a parameterized quantum circuit (quantum neural network) $W(\theta)$. The problem exhibits a natural permutation symmetry: exchanging x^1 and x^2 leaves the label unchanged. In this case we restrict the generators of unitaries to those that commute with SWAP at the Lie algebra level $[\text{SWAP}, H] = 0$. Instead of solving this directly at the group level, we work at the Lie algebra level, which is often easier. Consider a single-layer QNN $W(\theta) = e^{-i\theta H}$. As we would like to have $[W(\theta), U_g] = 0$ for all $g \in G$. Using Taylor expansion around $\theta = 0$ we got

$$e^{-i\theta H} = \mathbb{1} + \theta H + \mathcal{O}(\theta^2)$$

Then we can substitute into the commutator $[W(\theta), U_g] = \theta[H, U_g] + \mathcal{O}(\theta^2)$ which is zero so we end up requiring $[H, U_g] = 0 \quad \forall g \in G$. Note that finding generators H that commute with U_g is easier because

- H is Hermitian
- The condition is linear: $[H, U_g] = 0$
- For Lie groups, we only need to check on a basis of the Lie algebra

Going back to the the original QML task we see that after encoding into two qubits, the symmetry group is $G_{\text{SWAP}} = \{\mathbb{1}, \text{SWAP}\}$. By the previous discussion we need generators H such that $[H, \text{SWAP}] = 0$, these generators must be symmetric under qubit exchange and thus commute with SWAP

$$H \in \text{span} \{X \otimes \mathbb{1} + \mathbb{1} \otimes X, Y \otimes \mathbb{1} + \mathbb{1} \otimes Y, Z \otimes \mathbb{1} + \mathbb{1} \otimes Z, X \otimes X, Y \otimes Y, Z \otimes Z\}$$

Finally we build the quantum neural network as $\mathcal{W}(\theta) = \prod_l e^{-i\theta_l H_l}$. This illustrates that: *when we have a problem with Lie group symmetry, pass to the Lie algebra, analyze it, and return to the Lie group.*

Gate	Matrix	Description
X (Pauli- X)	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	Bit-flip gate. Equivalent to a classical NOT gate.
Y (Pauli- Y)	$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$	Bit and phase flip.
Z (Pauli- Z)	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$	Phase-flip gate.
H (Hadamard)	$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	Creates superposition; maps basis states to equal superpositions.
S (Phase)	$\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$	Applies a $\pi/2$ phase. Square root of Z .
T ($\pi/8$ gate)	$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$	Applies a $\pi/4$ phase.
$RX(\theta)$	$\begin{bmatrix} \cos(\theta/2) & -i\sin(\theta/2) \\ -i\sin(\theta/2) & \cos(\theta/2) \end{bmatrix}$	Rotation around the X -axis by angle θ .
$RY(\theta)$	$\begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix}$	Rotation around the Y -axis by angle θ .
$RZ(\theta)$	$\begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}$	Rotation around the Z -axis by angle θ .
CNOT	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$	Controlled-NOT gate. Flips the target qubit if the control is 1.

TABLE 1. Common quantum gates, their matrix representations, and brief descriptions.

REFERENCES

- [1] B. Diu and F. Laloe. *Quantum Mechanics: Volume II: Angular Momentum, Spin, and Approximation Methods*. Wiley-VCH Verlag GmbH & Company KGaA, 2020.
- [2] W. Fulton and J. Harris. *Representation theory: a first course*, volume 129. Springer Science & Business Media, 2013.
- [3] E. Kowalski. *An introduction to the representation theory of groups*, volume 155. American Mathematical Society, 2014.
- [4] V. Lakshmibai and J. Brown. Representation theory of finite groups. In *Flag Varieties: An Interplay of Geometry, Combinatorics, and Representation Theory*, pages 47–58. Springer, 2009.
- [5] J. M. Landsberg. *Quantum Computation and Quantum Information: A Mathematical Perspective*, volume 243. American Mathematical Society, 2024.
- [6] M. A. Nielsen and I. L. Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.
- [7] M. Ragone, P. Braccia, Q. T. Nguyen, L. Schatzki, P. J. Coles, F. Sauvage, M. Larocca, and M. Cerezo. Representation theory for geometric quantum machine learning. *arXiv preprint arXiv:2210.07980*, 2022.
- [8] J. Watrous. Understanding quantum information and computation. *arXiv preprint arXiv:2507.11536*, 2025.