Notes on the Heisenberg Model

These are personal notes on the theory behind the implementation of a variational quantum eigensolver (VQE) applied to the Heisenberg model, with an emphasis on symmetry exploitation. This is not a formal publication, but rather an informal summary of the conceptual and mathematical background supporting my work. It is intended for educational and reference purposes only.

1 Statement of the Heisenberg model

1.1 Spin chain

The model consists on a one-dimensional array of L atoms. We impose periodic boundary conditions (close chain). That means that the element $L+1$ in the chain will be the first one.

The microscopic degrees of freedom are spins. That means that every element of the chain comes with a vector space V_l and spin operator $\boldsymbol{S}_i = (S_i^x, S_i^y, S_i^z)$, where each $S_i^\alpha \in \mathrm{End}(\mathbb{C}^2)$. Where $\mathrm{End}(V_l)$ is the space of all linear operators $V_l \to V_l$. These operators are elements of the Lie Algebra $\mathfrak{su}(2)$ since they are the generators of the $SU(2)$ group and fulfill the usual commutation relations $[S_l^i, S_l^j] = i\hbar\epsilon^{ijk}S_l^k$. For spin 1/2 V_l is $\mathbb{C}^2$ and the representation of the spin operators are the Pauli matrices.

The full Hilbert space is the tensor product of each $\mathbb{C}^2$,

$$\mathcal{H} = \bigotimes_{i=1}^{L} \mathbb{C}^2 = (\mathbb{C}^2)^{\otimes L} \tag{1}$$

In this case, each $\mathbb{C}^2$ carries a 2-dimensional representation of $\mathfrak{su}(2)$. The whole Hilbert space can be represented by a 'global' $\mathfrak{su}(2)$-representation given by the total spin operator $\boldsymbol{S} = (S^x, S^y, S^z) \in \mathrm{End}(\mathcal{H})$ defined as

$$S^\alpha := \sum_{i=1}^{L} S_i^\alpha \tag{2}$$

But this representation is reducible in the usual SU(2) multiplet decomposition. In other words, the whole Hilbert space can be written as a tensor product of each individual spin, but also like:

$$\mathcal{H} = \bigoplus_j (V_j \otimes \mathbb{C}^{m_j}) \tag{3}$$

This decomposition shows that the Hilbert space is a direct sum of irreducible SU(2) spaces V_j with spin j and dimension $2j+1$. The $\mathbb{C}^{m_j}$ is the multiplicity space, which count how many equal dimension V_j appear.

For example for $L = 3$ the Hilbert space can be written as

$$\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2 = V_{3/2} \oplus V_{1/2} \oplus V_{1/2} = V_{3/2} \oplus (V_{1/2} \otimes \mathbb{C}^2) \tag{4}$$

This decomposition is particularly useful when we work with Hamiltonians invariant under SU(2) transformations, because the energy eigenstates can be classified by total spin quantum number j and the Hamiltonian becomes block-diagonal in this decomposition as we will see.

1.2 Hamiltonian

In our model, we have a Hamiltonian $H \in \mathrm{End}(\mathcal{H})$ describing the exchange interaction between the spins. We impose that these interactions satisfy a few conditions. First, interactions are only between the nearest neighbors, they must also be homogeneous (translationally invariant) and partially isotropic ($[S^z, H] = 0$)

The most general Hamiltonian that satisfies these conditions is

$$H = \sum_{i=1}^{L} J_x S_i^x S_{i+1}^x + J_y S_i^y S_{i+1}^y + J_z S_i^z S_{i+1}^z \tag{5}$$

where J_i are the exchange coupling parameters.

2 Variational quantum eigensolvers (VQE)

Our goal is to use variational quantum eigensolvers to determine a prediction for the ground state energy of our model. The VQE is a quantum algorithm based on the variational principle. This principle establishes that the expectation value of the Hamiltonian for a given state $|\psi\rangle$ is always bigger or equal to the ground state energy

$$\langle \hat{H} \rangle_\psi \geq E_0 \tag{6}$$

There is a method to give an estimation of the ground state energy using this principle. It consists of taking a parametrized state $|\psi(\theta)\rangle$ and finding an optimal θ value that minimizes the expectation value. The minimum energy will give a prediction of the ground state energy.

The VQE is just an algorithm implementation of this method. We start by preparing an initial quantum state $|\psi_0\rangle$. Then we use a variational circuit with parametrized gates to obtain $U(\theta)|\psi_0\rangle = |\psi(\theta)\rangle$ and we calculate the expectation value with the Hamiltonian. We finally use a classical optimizer to adjust the parameter θ till we find a θ^* that minimizes the energy.

$$E_0 \approx \langle \psi(\theta^*)|\hat{H}|\psi(\theta^*)\rangle \tag{7}$$

2.1 Gate parametrization with QAOA

To parametrize a state for the VQE, we need a parametrized quantum circuit, sometimes called ansatz. Usually, this quantum circuit is built with parametrized rotation gates. However, there are many instances where it is useful to think of a quantum circuit in terms of a Hamiltonian. The QAOA is an algorithm with p cost and mixing layers $\hat{U}_C(\beta_m)$, $\hat{U}_M(\alpha_m)$ each of them generated by Hamiltonians $\hat{H}_C$ and $\hat{H}_M$. Then the ansatz given an initial state is

$$|\psi_p(\boldsymbol{\alpha},\boldsymbol{\beta})\rangle = \prod_{m=1}^{p} e^{-i\alpha_m \hat{H}_M} e^{-i\beta_m \hat{H}_C} |\psi_0\rangle \qquad \begin{matrix} \boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_m, \cdots, \alpha_p) \\ \boldsymbol{\beta} = (\beta_1, \cdots, \beta_m, \cdots, \beta_p) \end{matrix} \tag{8}$$

The Hamiltonian $\hat{H}_C$ represents the cost function we want to optimize. In our case, plays the role of the Hamiltonian of our model, the one we want to know the ground state energy. The $\hat{H}_M$ term is a mixing term used to gain expressivity and explore more regions of the Hilbert space. Usually, this Hamiltonian is made out of Pauli matrices like, for instance, $\hat{Y}_i$ on each site.

$$\hat{H}_M = \sum_{i=1}^{L} \hat{Y}_i \Longrightarrow \hat{U}_M(\alpha_m) = e^{-i\alpha_m \hat{H}_M} = \prod_{i=1}^{L} e^{-i\alpha_m Y_i} \tag{9}$$

And the final ansatz is

$$|\psi_p(\boldsymbol{\alpha},\boldsymbol{\beta})\rangle = \prod_{m=1}^{p} \prod_{i=1}^{L} e^{-i\alpha_m Y_i} e^{-i\beta_m \hat{H}_C} |\psi_0\rangle \tag{10}$$

2.2 Ansatz symmetrization

These ansatz are in general non-equivariant, meaning that they do not preserve the symmetry of the problem. However, there is a procedure to transform any generic ansatz into an equivariant one, *i.e*, symmetry preserving. This procedure is known as ansatz symmetrization.

The ansatz symmetrization process begins by considering the generators G of the parametrized circuit

$$U_G(\boldsymbol{\theta}) = e^{-i\boldsymbol{\theta} G} \tag{11}$$

A generator (and therefore the corresponding unitary gate) is said to be equivariant under a given symmetry if it satisfies:

$$[U_G(\theta), U_s] = 0 \iff [G, U_s] = 0 \tag{12}$$

Where U_s is the unitary representation of the element $s \in S$, with S being the symmetry group.

The ansatz symmetrization procedure consists of modifying the generators so that they commute with the symmetry representation. To achieve this, we use the twirling formula, which ensures that a given generator transforms into another that commutes with U_s

$$\mathcal{T}_U[X] = \int d\mu_s \, U_s X U_s^\dagger \quad \Longrightarrow \quad [\mathcal{T}_U[X], U_s] = 0 \;\; \forall X \;\; \text{and} \;\; s \in S \tag{13}$$

Applying this transformation allows us to systematically construct an equivariant ansatz.

While symmetrizing the ansatz can be advantageous in some cases, it is not always the best choice. There is always a trade-off between expressivity and efficiency. By enforcing equivariance, we restrict our model to the relevant symmetry subspace, and thus we lose the ability to explore other sectors. In other words, we sacrifice expressivity. This can be problematic if the symmetry sector of the ground state is unknown.

On the other hand, an equivariant ansatz is more computationally efficient because the symmetry restriction significantly reduces the number of parameters to optimize. This, in turn, lowers the computational cost and can lead to faster convergence. See Appendix 5.2 for more context about the twirling formula and how to use it.

3 xxx Heisenberg model

3.1 Hamiltonian and symmetries

The simplest spin chain is the Heisenberg xxx model. Here $J_x = J_y = J_z \equiv J$. Then the Hamiltonian is

$$H_{xxx} = -J \sum_{i=1}^{L} \boldsymbol{S}_i \cdot \boldsymbol{S}_{i+1} \tag{14}$$

This model is a particular case where the interaction is completely isotropic; let's see why. One can prove that [1]

$$[\boldsymbol{S}^2, H_{xxx}] = 0 \quad [S^z, H_{xxx}] = 0 \tag{15}$$

This property has 2 direct implications. The first and more intuitive implication is that we can find a common basis that diagonalizes simultaneously all three operators. We will start considering the spin multiplet basis $|j, m\rangle \in \mathcal{H}$ such that

$$\boldsymbol{S}^2 |j, m\rangle = j(j+1) |j, m\rangle \quad S^z |j, m\rangle = m |j, m\rangle \tag{16}$$

Since $\boldsymbol{S}^2$ and S^z have to be conserved quantities among the time evolution generated by the Hamiltonian ($\frac{d\hat{A}}{dt} = [\hat{H}, \hat{A}] + \frac{\partial \hat{A}}{\partial t}$), the quantum numbers s and m of a multiplet state will also be conserved along the time evolution.

The other implication is that the Hamiltonian is invariant under $SU(2)$ transformations. A SU(2) transformation among the Hilbert space can be described by

$$U_s = e^{i\theta \hat{n} \cdot \boldsymbol{S}} \tag{17}$$

Where $\boldsymbol{S} = (S^x, S^y, S^z)$ are the group generators.

We say that one operator is invariant under a SU(2) transformation if

$$U_s \hat{A} U_s^\dagger = \hat{A} \iff [\hat{A}, U_s] = 0 \iff [\hat{A}, S^\alpha] = 0 \quad \forall \alpha = x, y, z \tag{18}$$

So we can say that our Hamiltonian is invariant under SU(2) transformations iff it commutes with all three generators. To prove that, we use a general result in Lie Algebras which says that if an operator commutes with the Casimir operator $\boldsymbol{S^2}$, then it will also commute with all generators iff the system is irreducible. Then, since we've proved that our Hilbert space is irreducible in SU(2) multiplets, we can ensure that and H commutes all generators.

$$[H, S^\alpha] = 0 \quad \forall \alpha = x, y, z \tag{19}$$

And therefore the Hamiltonian is completely invariant under SU(2) transformation.

This property is extremely important. As we said in the introduction, if the Hamiltonian is invariant under a SU(2) transformation, it gets block diagonal in the SU(2) multiplet decomposition. To prove that we will use that H is a rank 0 tensor under SU(2) transformation because its invariance. This allows us to use the Wigner-Eckart theorem (see Appendix 5.1) to prove that the Hamiltonian cannot connect different $\boldsymbol{|s, m\rangle}$ states if they have different quantum number s.

$$\langle s, m | H | s', m' \rangle \neq 0 \iff s' = s \tag{20}$$

And thus the Hamiltonian gets decomposed into irreducible blocks, each of them characterized by a SU(2) multiplet, just as the Hilbert space itself does.

[1] https://github.com/DominguezIsaac/Heisenberg-model

This will allow us to simplify our problem. To efficiently find the ground state, we only need to search within the symmetry sector corresponding to the correct multiplet, significantly reducing the computational cost of a VQE algorithm.

3.2 VQE on the antiferromagnetic (J<0) xxx model with even L

We use the Hamiltonian

$$H_{xxx} = \sum_{i=1}^{L} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_{i+1} = \sum_{i=1}^{L} X_i X_{i+1} + Y_i Y_{i+1} + Z_i Z_{i+1} \tag{21}$$

Which corresponds to the one in Eq. (14) with $J = -1 < 0$ and with Pauli matrices instead of spin matrices. For $J < 0$ the system is considered to be in the anti-ferromagnetic regime. In this regime, the lower energy is associated with a configuration where the spins are anti-aligned. For odd L the minimum spin configuration corresponds to a total $j = 1/2$ and for even L will be $j = 0$. Then, using an even L spin $1/2$ chain we can ensure that the ground state will be in the multiplet subspace where $j = 0$ *i.e* a SU(2) singlet.

Just to see how efficient it is to restrict the VQE to this subspace, let's see a comparison between the dimension of the whole Hilbert space and the singlet dimension. The total dimension for a L spin chain is $\dim(\mathcal{H}) = 2^L$, this space gets decomposed in multiplets of V_j each of them with dimension $\dim(V_s) = 2j + 1$. In the case of the singlet the dimension is $\dim(V_0) = 1$ so we just need to know how many singlet subspaces are in the decomposition, how many degeneracy we have. In general, the number of subspaces with spin j in the decomposition of L spin $1/2$ is

$$N_s = \frac{(2j+1)L!}{\left(\frac{L}{2}+j+1\right)!\left(\frac{L}{2}-j\right)!} \longrightarrow \dim(V_{j=0}) = \frac{L!}{\left(\frac{L}{2}+1\right)!\frac{L}{2}!} \tag{22}$$

In the next table we can see a comparison between the whole Hilbert space and the singlet for different even L values.

L	$\dim(\mathcal{H})$	$\dim(V_{j=0})$
2	4	1
4	16	2
6	64	5
8	256	14
10	1024	42

Now we can start to build the VQE of our system. To initialize the circuit, we need a state with zero total spin to ensure that we are in the right symmetry sector. The easiest state with total $s = 0$ is made by taking pairs of spins and building a singlet for each pair

$$|\psi_0\rangle = \underbrace{\frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle)}_{\text{singlet between 2 spins}} \otimes \cdots \otimes \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle) = \bigotimes_{k=1}^{L/2} \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \tag{23}$$

This state is expressed in the local basis for convenience, but we know that this state is an eigenvector of S^2 with total $s = 0$, so it is in the right symmetry sector.

The next step is to build the ansatz with the QAOA method we have explained in Section 2.1. In this case, the initial ansatz is

$$|\psi(\boldsymbol{\theta})\rangle = \prod_{m=1}^{p} \prod_{i=1}^{L} e^{-i\alpha_m Y_i} e^{-iH_{\text{odd}}(\boldsymbol{\beta}_m)} e^{-iH_{\text{even}}(\boldsymbol{\gamma_m})} |\psi_0\rangle \tag{24}$$

There are some key differences with the QAOA ansatz we have given above. The most evident is that the cost Hamiltonian is not exactly the Heisenberg, but it is divided into two terms defined as

$$\begin{aligned} H_{\text{odd}}(\boldsymbol{\beta}_m) &= \sum_{\substack{i=1 \\ \text{odd}}}^{L} \beta_m^x X_i X_{i+1} + \beta_m^y Y_i Y_{i+1} + \beta_m^z Z_i Z_{i+1} \qquad \boldsymbol{\beta}_m = (\beta_m^x, \beta_m^y, \beta_m^z) \\ H_{\text{even}}(\boldsymbol{\gamma}_m) &= \sum_{\substack{i=2 \\ \text{even}}}^{L} \gamma_m^x X_i X_{i+1} + \gamma_m^y Y_i Y_{i+1} + \gamma_m^z Z_i Z_{i+1} \qquad \boldsymbol{\gamma}_m = (\gamma_m^x, \gamma_m^y, \gamma_m^z) \end{aligned} \tag{25}$$

These Hamiltonians are built in a way that they split the spin operators in pairs. The odd term groups the sites $\{i, i+1\} = \{1,2\},\{3,4\},\{5,6\}\cdots$ while the even groups $\{i, i+1\} = \{2,3\},\{4,5\},\{6,7\}\cdots$. Doing this allows us to work with two Hamiltonians that all their terms commute with each other, since for example $[X_1X_2, Y_3Y_4] = 0$ but $[X_1X_2, Y_2Y_3] \neq 0$. So any term of H_{odd} will commute with any other term present in this Hamiltonian and the same for H_{even}, while cross terms won't commute in general.

If we look at the number of parameters, we see that for a fixed layer (fixed m) we have 3 values $\boldsymbol{\beta}_m$ for H_{odd}, 3 values $\boldsymbol{\gamma}_m$ for H_{even} and one extra value α_m for the mixing term. This gives us in total 7 parameters per layer. With p layers we obtain a total of $7p$ parameters.

3.2.1 Equivariant ansatz

This ansatz does not take into account the SU(2) symmetry underlying the Heisenberg model. Then we need to use the ansatz symmetrization we have explained above if we want to make it equivariant. In our specific case, since we know a priori that the ground state belongs to the same symmetry sector as the initial state, an equivariant ansatz is expected to be a great choice. We restrict the optimizing task only to explore the singlet subspace, gaining efficiency and reducing the number of parameters.

To symmetrize the ansatz, we should apply the twirling formula to the ansatz generators. This is to the sigma-y Pauli matrix and to the product of two consecutive Pauli matrices X_iX_{i+1}. Here X has to be taken as any of the Pauli matrices, it could be the X matrix but also would be valid for Y and Z matrices.

We start with the Y symmetrization

$$\mathcal{T}_U[Y] = \int d_\mu(U) UYU^\dagger \equiv \mathcal{M}^{(1)}(Y) \tag{26}$$

There is a subtle mark here. We should notice that in principle $Y_i = \mathbb{I} \otimes \cdots \otimes Y_i \otimes \cdots \otimes \mathbb{I} \in \text{End}((\mathbb{C}^2)^{\otimes L})$ but since the relevant part of the operator is restricted only to $\mathbb{C}^2$ we consider that the operator is equivalent to the usual Pauli matrix acting on $\mathbb{C}^2$.

We see that the twirling formula is equivalent to the first moment operator applied to the sigma-y matrix. The explicit computation of this operation has been done in the Appendix 5.2. Using Eq. (5.2.7) we obtain

$$\mathcal{T}_U[Y] = \frac{Tr(Y)}{2}\mathbb{I} = 0 \tag{27}$$

Since the Pauli matrices are traceless, the symmetrization vanishes. This result implies that in our equivariant ansatz the term $\prod_{i=1}^{L} e^{-i\alpha_m Y_i}$ must vanish.

Concerning the other generators:

$$\mathcal{T}_U[X_iX_{i+1}] = \mathcal{M}^{(2)}(X_iX_{i+1}) = C_1\mathbb{I} + C_2\text{SWAP} \tag{28}$$

Here we see that the symmetrization is equivalent to the second moment, which following the Appendix procedure we saw that is equal to a combination of the identity and SWAP operators. The coefficients can be found using Eq. (5.2.7). In this particular case we obtain.

$$\begin{aligned} C_1 &= \frac{2Tr(X_iX_{i+1}) - Tr(SWAP \cdot X_iX_{i+1})}{3} = \frac{\cancelto{0}{2Tr(X)Tr(X)} - Tr(X^2)}{3} = \frac{-2}{3} \\ C_2 &= \frac{2Tr(SWAP \cdot X_iX_{i+1}) - Tr(X_iX_{i+1})}{3} = \frac{2Tr(X^2) - \cancelto{0}{Tr(X)Tr(X)}}{3} = \frac{4}{3} \end{aligned} \tag{29}$$

We then conclude that the symmetrization of a product of Pauli matrices is

$$\begin{aligned} &\mathcal{T}_U[X_iX_{i+1}] = -\frac{2}{3}\mathbb{I} + \frac{4}{3}\text{SWAP} = -\frac{2}{3}\mathbb{I} + \frac{4}{3}(\mathbb{I} + X\otimes X + Y\otimes Y + Z\otimes Z) \\ &\mathcal{T}_U[X_iX_{i+1}] \sim X_iX_{i+1} + Y_iY_{i+1} + Z_iZ_{i+1} \end{aligned} \tag{30}$$

This ansatz symmetrization yields the equivariant ansatz

$$|\psi(\boldsymbol{\theta})\rangle = \prod_{m=1}^{p} e^{-i\theta_m H_{\text{odd}}} e^{-i\phi_m H_{\text{even}}} |\psi_0\rangle \tag{31}$$

Where now H_{odd} and H_{even} are the isotropic versions. This means that their parameters are $\beta_m = (1,1,1)$ and $\gamma_m = (1,1,1)$. Note that the equivariant ansatz has $2p$ parameters in total compared to the $7p$ of the non-symmetric ansatz.

3.2.2 Numerical implementation

We compare the performance of the two ansatz using the Heisenberg model with $L = 8$ spins and different values of p, from $p = 1$ up to $p = 5$. We use the L-BFGS optimizer to find the optimal value iteratively. The results are shown in Figure 1.

The first thing we should notice is that the maximum number of iterations to find the optimal value was restricted to 50 due to computation limitations. As we can see, the non-equivariant ansatz reached the maximum iteration in almost all depths. This could imply that the optimizer didn't reach the optimal value and thus the results could be inconclusive. This fact is particularly evident with $p = 5$ layers, where the ground state value clearly diverges from the expected value.

However, we can draw conclusions for the equivariant ansatz. The first thing to notice is that the computational cost is indeed lower, since in all depths except the last one we didn't reach the maximum iterations to find the optimal value. Also, since the maximum was not reached, the results must be more conclusive. Comparing with the actual ground state energy marked with strip lines, we can see how the optimal value converges to the actual ground state, giving better results than the non-equivariant ansatz.

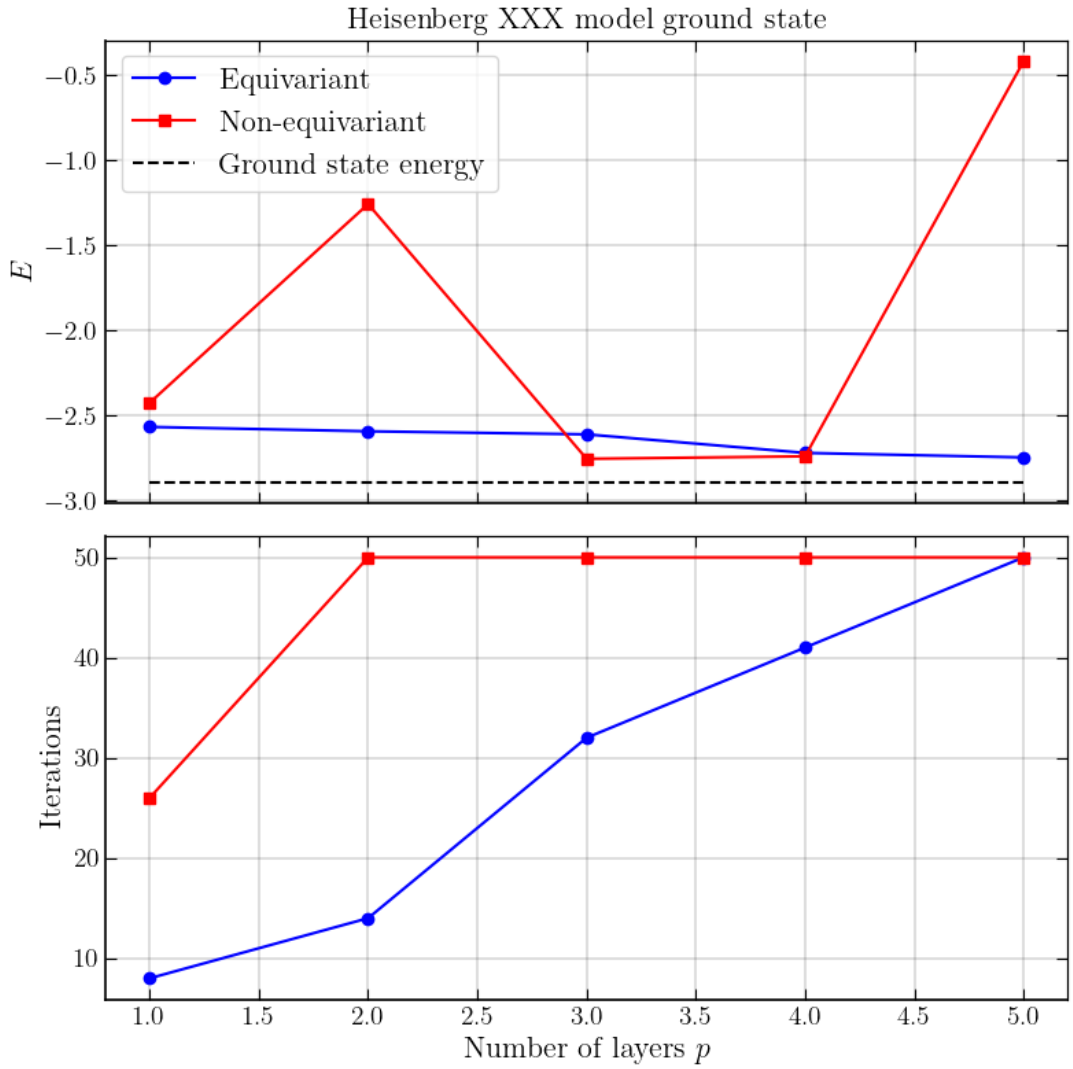

Figure 1: Comparison of the equivariant and the non-equivariant ansatz for the Heisenberg model on $L = 8$ spins. For every number of layers p, we plot the achieved energy (top) and the number of iterations necessary to reach it (bottom).

4 xxz Heisenberg model

4.1 Hamiltonian and symmetries

In this case, we have a more general chain whose Hamiltonian is

$$H_{xxz} = -J \sum_{i=1}^{L} \left(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z \right) \tag{32}$$

The parameter Δ controls the anisotropy between the $x-y$ plane and the z-direction. For $\Delta = 1$ the model is isotropic (XXX model), while $\Delta \neq 1$ breaks the full rotational symmetry.

Using the ladder operators ($S_i^{\pm} := S_i^x \pm i S_i^y$) the Hamiltonian becomes

$$H_{xxz} = -\frac{J}{2} \sum_{i=1}^{L} \left(S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+ + 2\Delta S_i^z S_{i+1}^z \right) \tag{33}$$

Now, due to the anisotropy term, we don't have full rotational symmetry. The Hamiltonian doesn't commute with the S^2 anymore. However, it still commutes with S^z so we can ensure that the Hamiltonian is still invariant under rotations around the $z-$axis.

$$[H, S^z] = 0 \Longrightarrow [H, R_z(\theta)] = 0 \Longrightarrow R_z(\theta)\, H\, R_z^\dagger(\theta) = H \tag{34}$$

Where $R_z(\theta) := e^{-i\theta S^z}$ is a global rotation around the z-axis.

We can also check it explicitly on the Hamiltonian. The S^z term of the Hamiltonian is obviously invariant ($R_z(\theta) S^z R_z^\dagger(\theta) = S^z$), and the ladder operators transform in the following way

$$\begin{aligned} R_z(\theta) S^x R_z^\dagger(\theta) &= S^x \cos\theta - S^y \sin\theta \\ R_z(\theta) S^y R_z^\dagger(\theta) &= S^x \sin\theta + S^y \cos\theta \end{aligned} \Longrightarrow R_z(\theta) S^{\pm} R_z^\dagger(\theta) = e^{\mp i\theta} S^{\pm} \tag{35}$$

So we see that the ladder operators only get a global phase from the rotation around the z-axis that cancels out when in the Hamiltonian we multiply a raising with a lowering operator.

This kind of symmetry transformation will preserve the quantum number m associated with the operator S^z. It also has the underlying group $U(1)_z \subseteq SU(2)$ behind the symmetry.

$$U(1) = \{ e^{i\theta} \,|\, \theta \in \mathbb{R} \} \tag{36}$$

Since the model lacks a global SU(2) symmetry, the decomposition of the Hilbert space into irreducible SU(2) multiplets is no longer helpful for diagonalizing the Hamiltonian. However, we can still exploit the partial isotropy. Since H and S^z commutes they can be simultaneously diagonalized. This implies that the eigenstates of S^z form a basis in the Hilbert space in which the Hamiltonian is block diagonal.

The local computational basis

$$\{ |\uparrow\uparrow \cdots \uparrow\rangle, |\uparrow\uparrow \cdots \downarrow\rangle, \cdots |\downarrow\downarrow \cdots \downarrow\rangle \} \tag{37}$$

consists of eigenstates of S^z, since each product state has a well-defined total magnetization m, given by the sum over individual spin projections $m_i = \pm\frac{1}{2}$:

$$m = \sum_{i=1}^{L} m_i \tag{38}$$

For example:

$$S^z \left|\uparrow\uparrow\cdots\uparrow\right\rangle = \sum_{i=1}^{L} \frac{1}{2} \left|\uparrow\uparrow\cdots\uparrow\right\rangle = \frac{L}{2} \left|\uparrow\uparrow\cdots\uparrow\right\rangle \tag{39}$$

Therefore, this basis naturally decomposes the Hilbert space into sectors of fixed magnetization m, where the Hamiltonian acts block-diagonally. Explicitly,

$$\mathcal{H} = \bigoplus_m \mathcal{H}_m \tag{40}$$

Where $\mathcal{H}_m = \{|\psi\rangle \in \mathcal{H} \,|\, S^z |\psi\rangle = m |\psi\rangle\}$

The dimension of each magnetization sector is given by the number of spin configurations with total m, which corresponds to the binomial coefficient

$$\dim(\mathcal{H}_0) = \binom{L}{L/2} = \frac{L!}{(L/2)!^2} \tag{41}$$

Then, although the symmetry structure is less restrictive than in the isotropic XXX model, the residual U(1) symmetry still allows for a decomposition into symmetry sectors. This can be exploited to significantly reduce the computational cost of the VQE algorithm.

4.2 VQE on the antiferromagnetic xxz model with even L

In this case, the Hamiltonian we consider is

$$H_{xxz} = -\sum_{i=1}^{L} X_i X_{i+1} + Y_i Y_{i+1} - 2 Z_i Z_{i+1} \tag{42}$$

This Hamiltonian corresponds to the xxz one with $J = 1 > 0$ and $\Delta = -2 < 0$. This particular configuration corresponds to an antiferromagnet along the z direction (since $J\Delta < 0$) but a ferromagnetic order in the $x - y$ plane (since $J > 0$). Also, since $\Delta < -1$ we expect to have a phase dominated by the Ising antiferromagnetic model. At the limit $\Delta \to -\infty$ the ground state is the degenerated Néel ground state $|\uparrow\downarrow\uparrow\downarrow\cdots\rangle$ and $|\downarrow\uparrow\downarrow\uparrow\cdots\rangle$, so we expect our ground state to belong to the $m = 0$ sector. As we did in the xxx model, we can compare the dimension of the whole Hilbert space with the symmetry sector we want to explore:

L	$\dim(\mathcal{H})$	$\dim(\mathcal{H}_{m=0})$
2	4	2
4	16	6
6	64	20
8	256	70
10	1024	252

As we can see, we can decrease significantly the dimension of the space we explore, but not as much as with the SU(2) symmetry.

To initialize the circuit we use one of the Néel states

$$|\psi_0\rangle = |\uparrow\downarrow\uparrow\downarrow\cdots\rangle \tag{43}$$

And we use the same Ansatz (Eq. (24)) to start the optimization. However, we are interested in studying the use of symmetries in this model, so we should do a symmetrization process as with the xxx model to find an ansatz that commutes with the elements of the U(1) group, or

in particular with S^z.

In this case, the symmetrization procedure is easier than the one with SU(2) symmetry because we don't need the Haar measure. To integrate along the U(1) group we just need to integrate the parameter θ and the probability measure is just $\rho(\theta) = \frac{1}{2\pi}$.

Then, for instance, the symmetrization applied to the Y generator will be:

$$\mathcal{T}_{U(1)}[Y] = \int_0^{2\pi} \frac{d\theta}{2\pi} R_z(\theta) Y R_z^\dagger(\theta) = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{-i\theta S^z} Y e^{i\theta S^z} \tag{44}$$

We now use that a rotation along the z-axis applied to the Y matrix is equivalent to:

$$e^{-i\theta S^z} Y e^{i\theta S^z} = \cos\theta\, Y + \sin\theta\, X \tag{45}$$

But both functions are 0 if they are integrated in an interval between 0 and 2π

$$\mathcal{T}_{U(1)}[Y] = Y \underbrace{\int_0^{2\pi} \frac{d\theta}{2\pi} \cos\theta}_{0} + X \underbrace{\int_0^{2\pi} \frac{d\theta}{2\pi} \sin\theta}_{0} = 0 \tag{46}$$

So, similarly to the xxx case, the mixing term doesn't commute with the elements of the symmetry group, so it must vanish in the equivariant ansatz.

For the terms in the odd and even Hamiltonians, clearly the symmetrization process applied to Z_iZ_{i+1} is left invariant. On the other hand, X_iX_{i+1} and Y_iY_{i+1} are not invariant under z rotations by themselves, but the combination of both is, as we will see

$$\mathcal{T}_{U(1)}[X_iX_{i+1}] = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{-i\theta S^z} X_iX_{i+1} e^{i\theta S^z} = \int_0^{2\pi} \frac{d\theta}{2\pi} e^{-i\theta S^z} X_i e^{i\theta S^z} e^{-i\theta S^z} X_{i+1} e^{i\theta S^z} = \tag{47}$$

Now we use the following:

$$e^{-i\theta S^z} X e^{i\theta S^z} = \cos\theta\, X - \sin\theta\, Y \tag{48}$$

Then

$$\begin{aligned} e^{-i\theta S^z} X_i e^{i\theta S^z} e^{-i\theta S^z} X_{i+1} e^{i\theta S^z} &= (X_i\cos\theta - Y_i\sin\theta)(X_{i+1}\cos\theta - Y_{i+1}\sin\theta) \\ &= X_iX_{i+1}\cos^2\theta - (X_iY_{i+1} + Y_iX_{i+1})\sin\theta\cos\theta + Y_iY_{i+1}\sin^2\theta \end{aligned} \tag{49}$$

So the twirling is

$$\begin{aligned} \mathcal{T}_{U(1)}[X_iX_{i+1}] &= X_iX_{i+1}\int_0^{2\pi} \cos^2\theta \frac{d\theta}{2\pi} - (X_iY_{i+1} + Y_iX_{i+1}) \cancel{\int_0^{2\pi} \cos^2\theta \frac{d\theta}{2\pi} \sin\theta\cos\theta} + \\ &+ Y_iY_{i+1}\int_0^{2\pi} \cos^2\theta \frac{d\theta}{2\pi} \sin^2\theta \\ &= \frac{1}{2}(X_iX_{i+1} + Y_iY_{i+1}) \end{aligned} \tag{50}$$

Similarly, one can prove that

$$\mathcal{T}_{U(1)}[Y_iY_{i+1}] = \frac{1}{2}(X_iX_{i+1} + Y_iY_{i+1}) \tag{51}$$

These results actually make sense, since the whole Hamiltonian was invariant under S^z because the $x-y$ term was isotropic. Then the symmetrization yields the equivariant ansatz:

$$|\psi(\boldsymbol{\theta})\rangle = \prod_{m=1}^{p} e^{-iH_{\text{odd}}(\boldsymbol{\beta}_m)} e^{-iH_{\text{even}}(\boldsymbol{\gamma_m})} |\psi_0\rangle \tag{52}$$

Where the parameters must be isotropic in the $x-y$ plane

$$\boldsymbol{\beta}_m = (\beta_m^{xy}, \beta_m^{xy}, \beta_m^z) \quad \boldsymbol{\gamma}_m = (\gamma_m^{xy}, \gamma_m^{xy}, \gamma_m^z) \tag{53}$$

This way the odd and even Hamiltonians are

$$\begin{aligned} H_{\text{odd}}(\boldsymbol{\beta}_m) &= \sum_{\substack{i=1 \\ \text{odd}}}^{L} \beta_m^{xy}(X_iX_{i+1} + Y_iY_{i+1}) + \beta_m^z Z_iZ_{i+1} \\ H_{\text{even}}(\boldsymbol{\gamma}_m) &= \sum_{\substack{i=2 \\ \text{even}}}^{L} \gamma_m^{xy}(X_iX_{i+1} + Y_iY_{i+1}) + \gamma_m^z Z_iZ_{i+1} \end{aligned} \tag{54}$$

Therefore, with this symmetrization we see that for a fixed layer we have 4 different parameters. This gives us a total of $4p$ parameters. Is slightly less in comparison with the $7p$ of the non equivariant ansatz, but more than the one we obtained with SU(2) symmetry which was $2p$.

5 Appendix

5.1 Wigner-Eckart theorem

The Wigner-Eckart theorem states that matrix elements of tensor operators with respect to angular-momentum eigenstates satisfy:

$$\langle \alpha', j'm' | \, T_q^{(k)} \, | \alpha, jm \rangle = \langle jk; mq | jk; j'm' \rangle \, \frac{\langle \alpha' j' | \, T^{(k)} \, | \alpha j \rangle}{\sqrt{2j+1}}. \tag{55}$$

In order to prove the Eq. (55) we first need to give a definition for a spherical tensor operator. One usual definition of a tensor operator is given by the commutation relation between the angular momentum operators and the tensor operator. In general, a tensor operator will follow the following relations

$$\left[J_z, T_q^{(k)} \right] = \hbar q T_q^{(k)} \tag{56}$$

$$\left[J_\pm, T_q^{(k)} \right] = \hbar \sqrt{(k \mp q)(k \pm q + 1)} T_{q\pm 1}^{(k)} \tag{57}$$

Using the second one we can compute this operation

$$\langle \alpha', j'm' | \, [J_\pm, T_q^{(k)}] \, | \alpha, jm \rangle = \hbar \sqrt{(k \mp q)(k \pm q + 1)} \, \langle \alpha', j'm' | \, T_{q\pm 1}^{(k)} \, | \alpha, jm \rangle \tag{58}$$

But also, using the definition of the commutator

$$\begin{aligned} &\langle \alpha', j'm' | \, [J_\pm, T_q^{(k)}] \, | \alpha, jm \rangle = \langle \alpha', j'm' | \, J_\pm T_q^{(k)} \, | \alpha, jm \rangle - \langle \alpha', j'm' | \, T_q^{(k)} J_\pm \, | \alpha, jm \rangle \\ &= \hbar \sqrt{(j' \pm m')(j' \mp m' + 1)} \, \langle \alpha', j'm' \mp 1 | \, T_q^{(k)} \, | \alpha, jm \rangle \\ &- \hbar \sqrt{(j \mp m)(j \pm m + 1)} \, \langle \alpha', j'm' | \, T_q^{(k)} \, | \alpha, jm \pm 1 \rangle \end{aligned} \tag{59}$$

Where we've used the usual property of the ladder operator $J_\pm$
$J_\pm \, |j, m\rangle = \hbar \sqrt{(j \mp m)(j \pm m + 1)} \, |j, m \pm 1\rangle$, and the fact that $(J_\pm)^\dagger = J_\mp$

So if we equalize the Eq. (58) with Eq. (59) we obtain

$$\begin{aligned} &\sqrt{(j' \pm m')(j' \mp m' + 1)} \, \langle \alpha', j'm' \mp 1 | \, T_q^{(k)} \, | \alpha, jm \rangle = \\ &\sqrt{(j \mp m)(j \pm m + 1)} \, \langle \alpha', j'm' | \, T_q^{(k)} \, | \alpha, jm \pm 1 \rangle \\ &+ \sqrt{(k \mp q)(k \pm q + 1)} \, \langle \alpha', j'm' | \, T_{q\pm 1}^{(k)} \, | \alpha, jm \rangle \end{aligned} \tag{60}$$

The next step is to realize that this equation is similar to the recursion relation of the Clebsch-Gordan coefficient.

Given some j_1, j_2 and j, one can show that there exists a recursion relation between the Clebsch-Gordan coefficients of the different m_1 and m_2. That relation is the following, and we note that it is very similar to the one we've obtained before.

$$\begin{aligned} &\sqrt{(j \mp m)(j \pm m + 1)} \, \langle j_1 j_2; m_1 m_2 | j_1 j_2; j, m \pm 1 \rangle = \\ &\sqrt{(j_1 \pm m_1)(j_1 \mp m_1 + 1)} \langle j_1 j_2; m_1 \pm 1, m_2 | j_1 j_2; j, m \rangle \\ &+ \sqrt{(j_2 \pm m_2)(j_2 \mp m_2 + 1)} \langle j_1 j_2; m_1, m_2 \pm 1 | j_1 j_2; j, m \rangle. \end{aligned} \tag{61}$$

Both of them are a sum of coefficients (the square roots) times some unknown quantity (the matrix elements in Eq.(60) and the Clebsch-Gordan coefficients in Eq.(61)). In fact, both coefficients are the same realizing that, in Eq.(60) j' and m' play the role of j and m, j and m

play the role of j_1 and j_2 and finally k and q are j_2 and m_2. So, in general, we can say that we therefore have two sets of linear homogeneous equations:

$$\sum_j a_{ij}x_j = 0 \qquad \sum_j a_{ij}y_j = 0 \tag{62}$$

Where a_{ij} are the square root coefficients and x_j and y_j are the matrix elements and the Clebsch-Gordan coefficients. This type of equations can not be solved exactly for x_j (our goal), but we can say that x_j and y_j are proportional. $x_j = Cy_j$. Then, noting that the Clebsch-Gordan coefficient $\langle j_1 j_2; m_1, m_2 \pm 1 | j_1 j_2; j, m\rangle$ corresponds to the matrix element $\langle \alpha', j'm' | \, T^{(k)}_{q\pm1} \, | \alpha, jm\rangle$, what we obtain is (after changing $j' \to j$ etc.)

$$\boxed{\langle \alpha', j'm' | \, T^{(k)}_{q\pm1} \, | \alpha, jm\rangle = C\langle jk; m, q \pm 1 | jk; j', m'\rangle} \tag{63}$$

Where C is a constant that has to be independent of the coefficients m, q and m'.

Now if we look to the particular case where the tensor is a scalar tensor $T^{(0)}_0 = S$.

$$\langle \alpha', j'm' | \, S \, | \alpha, jm\rangle = \langle j0; m0 | j0; j'm'\rangle \, \frac{\langle \alpha' j' | \, S \, | \alpha j\rangle}{\sqrt{2j+1}} \tag{64}$$

This Clebsch-Gordan coefficient corresponds to adding an initial $j_1 = j$ with $j_2 = 0$ and $m_1 = m$ $m_2 = 0$ to get j' and m', *i.e.*

$$j' = j \otimes 0 = j \qquad m' = m + 0 = m \tag{65}$$

So there is only one possibility. If $j' = j$ and $m' = m$ the coefficient is 1 and 0 otherwise. Or in other words $\langle j0; m0 | j0; j'm'\rangle = \delta_{jj'}\delta_{mm'}$

$$\langle \alpha', j'm' | \, S \, | \alpha, jm\rangle = \delta_{jj'}\delta_{mm'} \, \frac{\langle \alpha' j' | \, S \, | \alpha j\rangle}{\sqrt{2j+1}} \tag{66}$$

The physical interpretation is that a scalar operator cannot connect elements with different angular momentum. The only possibility for non-vanishing matrix elements of a scalar operator in an angular momentum basis is that the basis elements have $j' = j$ and $m' = m$.

5.2 Regarding the twirling formula and the Haar measure

We have stated a general process to construct an equivariant ansatz using the twirling formula. Now we need to understand how to compute this calculation.

$$\mathcal{T}_U[X] = \int_{U(d)} UXU^\dagger \, d\mu(U) \tag{67}$$

This is an integral over the continuous unitary group $U(d)$, where we integrate the function $f(U) = UXU^\dagger$ with respect to the Haar measure $d\mu(U)$. Here U is treated as a variable running over all the elements of the group. The natural question that arises is: what does it mean to integrate over a group? To answer this, we must understand the notion of the Haar measure.

The Haar measure is a probability measure defined on groups such as $U(d)$. It assigns real values to measurable subsets of the group, which allows us to define integrals over these sets. The key feature of the Haar measure is that it is both left and right invariant, meaning that, given a function f and a group element V:

$$\int_{U(d)} f(U)\, d\mu(U) = \int_{U(d)} f(VU)\, d\mu(U) = \int_{U(d)} f(UV)\, d\mu(U) \quad \forall\, V \in U(d) \tag{68}$$

Moreover, since the Haar measure is a probability measure, it satisfies the following properties:

$$\int_S 1\, d\mu(U) \geq 0 \quad \forall S \subseteq U(d) \quad \text{and} \quad \int_{U(d)} 1\, d\mu(U) = 1 \tag{69}$$

These properties allow us to interpret the integral of any function $f(U)$ over the Haar measure as the expected value of $f(U)$ with respect to μ, denoted as

$$\mathop{\mathbb{E}}_{U\sim\mu}[f(U)] := \int_{U(d)} f(U)\, d\mu(U) \tag{70}$$

5.2.1 Space of operators

We denote $\mathrm{End}((\mathbb{C}^d)^{\otimes k})$ the set of linear operators acting on $(\mathbb{C}^d)^{\otimes k} = \underbrace{\mathbb{C}^d \otimes \cdots \otimes \mathbb{C}^d}_{k \text{ times}}$

$$\mathrm{End}((\mathbb{C}^d)^{\otimes k}) : (\mathbb{C}^d)^{\otimes k} \to (\mathbb{C}^d)^{\otimes k} \tag{71}$$

This set forms a vector space itself and it can be equipped with an inner product, the so-called Hilbert-Schmidt inner product

$$\langle A, B\rangle_{H,S} := Tr(A^\dagger B) \quad A, B \in \mathrm{End}((\mathbb{C}^d)^{\otimes k}) \tag{72}$$

This vector space equipped with the Hilbert-Schmidt inner product is a Hilbert space, which means that it can have projectors, self-adjoints, and more linear operations between operators.

$$\mathrm{End}((\mathbb{C}^d)^{\otimes k}) \to \mathrm{End}((\mathbb{C}^d)^{\otimes k}) \tag{73}$$

5.2.2 k-th Moment operator

One example of a linear operation between operators is the k-th moment operator, defined as

$$\mathcal{M}^{(k)}_\mu(O) := \mathop{\mathbb{E}}_{U\sim\mu}[U^{\otimes k} O U^{\dagger\otimes k}] = \int_{U(d)} d\mu(U)\, U^{\otimes k} O U^{\dagger\otimes k} \tag{74}$$

This linear operation takes an operator $O \in \mathrm{End}((\mathbb{C}^d)^{\otimes k})$ and gives another operator $\mathcal{M}(O) \in \mathrm{End}((\mathbb{C}^d)^{\otimes k})$
The moment operator is linear, trace preserving and self-adjoint with respect to the Hilbert-Schmidt inner product, i.e

$$\langle \mathcal{M}(A), B\rangle_{H,S} = \langle A, \mathcal{M}(B)\rangle_{H,S} \tag{75}$$

5.2.3 k-th order commutant

Given $S \subseteq \text{End}(\mathbb{C}^d)$ a subset of the linear operators vector space acting on $\mathbb{C}^d$, we define the commutant as:

$$\text{Comm}(S,k) := \{A \in \text{End}((\mathbb{C}^d)^{\otimes k}) \,|\, [A, B^{\otimes k}] = 0 \,\forall B \in S\} \tag{76}$$

This is the set of all the operators A that commute with some symmetry group and form a vector subspace of the whole set of operators. For example, the commutant with respect to the unitary group $U(d)$ is:

$$\text{Comm}(U(d),k) := \{A \in \text{End}((\mathbb{C}^d)^{\otimes k}) \,|\, [A, V^{\otimes k}] = 0 \,\forall V \in U(d)\} \tag{77}$$

5.2.4 Moment operator as the orthogonal projector onto the commutant

One useful property of the moment operator is the following:

For all $A \in \text{End}((\mathbb{C}^d)^{\otimes k})$ its moment operator $\mathcal{M}^{(k)}(A) \in \text{Comm}(U(d),k)$.

That means that given an arbitrary operator A, its moment operator will commute with any element of the unitary group.

$$\left[\mathcal{M}^{(k)}(A), V^{\otimes k}\right] = 0 \quad \forall V \in U(d) \tag{78}$$

Proof:

$$\begin{aligned} V^{\otimes k}\mathcal{M}^{(k)}(A) = \mathop{\mathbb{E}}_{U\sim\mu}[(VU)^{\otimes k} O U^{\dagger\otimes k}] &\underset{\text{left invariance}}{=} \mathop{\mathbb{E}}_{U\sim\mu}[U^{\otimes k} O U^{\dagger\otimes k}] V^{\otimes k} = \mathcal{M}^{(k)}(A) V^{\otimes k} \\ &\implies \left[\mathcal{M}^{(k)}(A), V^{\otimes k}\right] = 0 \quad \square \end{aligned} \tag{79}$$

This was exactly what we wanted to symmetrize our ansatz and make it equivariant. Now we can look back to the twirling formula and notice that it's nothing but the moment operator applied to the generators of the ansatz. Therefore, we just need a practical way of computing the k-th moment operator to a given generator.

5.2.5 Permutation operators

Given $\pi \in S_k$, an element of the symmetric group S_k, we define the permutation matrix or operator $V_d(\pi) \in \text{End}((\mathbb{C}^d)^{\otimes k})$ to be the unitary matrix such that

$$V_d(\pi) \,|\psi_1\rangle \otimes \cdots \otimes |\psi_k\rangle = |\psi_{\pi^{-1}(1)}\rangle \otimes \cdots \otimes |\psi_{\pi^{-1}(k)}\rangle \tag{80}$$

This means that, given an initial state, the permutation operator rearranges the indices $j \in \{1, \cdots, k\}$ according to the symmetric group action $\pi(j)$.

Example:

With $k = 2$, we work with the group $S_2 = \{Id, \leftrightarrow\}$ meaning we have only two possibilities, do nothing or invert the elements. The associated permutation operators are

$$V_d(\pi) = \{\mathbb{I}_{2d}, \text{SWAP}\} \in \text{End}(\mathbb{C}^d \otimes \mathbb{C}^d) \tag{81}$$

For the element Id it is obvious why the associated matrix is the identity matrix of dimension $2d$. For the second element we notice the following

$$\pi(1) = 2 \quad \pi(2) = 1 \implies \pi^{-1}(1) = 2 \quad \pi^{-1}(2) = 1 \tag{82}$$

Then, following our definition, the permutation operator should perform the following operation

$$|\psi_1\rangle \otimes |\psi_2\rangle \underset{V_d(\pi)}{\longrightarrow} |\psi_2\rangle \otimes |\psi_1\rangle \tag{83}$$

And this action is accomplished by the SWAP operator

5.2.6 Schur-Weyl duality

This theorem is vital to compute the moment operator in a practical way. The theorem establishes the following:

The k-th order commutant of the unitary group is the span of the permutation operators associated to S_k

$$\text{Comm}(U(d), k) = \text{Span}\left(V_d(\pi) \,:\, \pi \in S_k\right) \tag{84}$$

This means that if you consider all possible permutations $\pi \in S_k$ and construct all the $V_d(\pi)$ associated, then any operator $A \in \text{Comm}(U(d), k)$ (*i.e* that commutes with any unitary group element) will be a linear combination of $V_d(\pi)$. The $V_d(\pi)$ form a basis in the Comm(U(d),k) vector subspace.

This fact is so powerful because, given an arbitrary operator $O \in \text{End}((\mathbb{C}^d)^{\otimes k})$, the moment operator belongs to the commutant subset as we proved. Therefore, it can be expanded as a combination of the permutation operators

$$\mathcal{M}^{(k)}(O) = \int_{U(d)} d\mu(U)\, U^{\otimes k} O U^{\dagger \otimes k} = \sum_{\pi \in S_k} C_\pi(O) V_d(\pi) \tag{85}$$

Then, we've found a way of computing the moment operator of a given operator and we just need to find the coefficients $C_\pi(O)$. The way to find them is to compute the product of the operator with each permutation operator and compute the trace. In general we obtain:

$$\underbrace{Tr(V_d^\dagger(\sigma))}_{b_\sigma} = \sum_{\pi \in S_k} \underbrace{C_\pi(O)}_{c_\pi} \underbrace{Tr(V_d^\dagger(\sigma) V_d(\pi))}_{G_{\sigma\pi}} \tag{86}$$

Then we can solve for C_π solving the system of linear equations $\vec{b} = G\vec{c} \iff \vec{c} = G^+\vec{b}$ or in components

$$C_\pi(O) = \sum_{\sigma \in S_k} (G^+)_{\pi\sigma} Tr(V_d^\dagger(\sigma) O) \tag{87}$$

Here, G is the so-called Gram matrix and, since it doesn't necessarily have an inverse, we define G^+ as the pseudo-inverse Defining the Weingarten coefficients as $Wg(\pi^{-1}\sigma, d) := (G^+)_{\pi\sigma}$ we find a closed result for the moment operator calculation

$$\boxed{\mathcal{M}^{(k)}(O) = \int_{U(d)} d\mu(U)\, U^{\otimes k} O U^{\dagger \otimes k} = \sum_{\pi,\sigma \in S_k} Wg(\pi^{-1}\sigma, d)\, Tr(V_d^\dagger(\sigma) O)\, V_d(\pi)} \tag{88}$$

With $Wg(\pi^{-1}\sigma, d) := (G^+)_{\pi\sigma}$ and $G_{\pi\sigma} = Tr(V_d^\dagger(\pi) V_d(\sigma)) = Tr(V_d(\pi^{-1}\sigma))$

5.2.7 Explicit computation of the first and second moment

For the ansatz symetrization we need to compute the twirling formula to operators that end up being a first and second moment operator computations, so let's see the explicit form of these operators using the Schur-Weyl duality.

First moment (k=1)
Given an operator $O \in \text{End}(\mathbb{C}^d)$ the first moment is

$$\mathcal{M}^{(1)}(O) = \int_{U(d)} d\mu(U)\, UOU^\dagger \tag{89}$$

For $k = 1$ the symmetric group only have one element $S_1 = \{Id\}$ whose permutation operator is $V_d(Id) = \mathbb{I}_d$ the idenity operator acting on $\mathbb{C}^d$. Then the moment becomes

$$\mathcal{M}^{(1)}(O) = C_{\mathbb{I}}(O)\mathbb{I}_d \tag{90}$$

To obtain $C_{\mathbb{I}}(O)$ we just need to perform the trace of the product $V_d O$ for each V_d, in our case, since we only have the identity we just writte:

$$Tr(O) \underset{\mathcal{M}\text{ is trace preserving}}{=} Tr(\mathcal{M}^{(1)}(O)) = Tr(C_{\mathbb{I}}(O)\mathbb{I}_d) = dC_{\mathbb{I}}(O) \iff C_{\mathbb{I}}(O) = \frac{Tr(O)}{d} \tag{91}$$

So

$$\boxed{\mathcal{M}^{(1)}(O) = \int_{U(d)} d\mu(U)\, UOU^\dagger = \frac{Tr(O)}{d}\mathbb{I}_d} \tag{92}$$

Second moment (k=2)
Given an operator $O \in \text{End}(\mathbb{C}^d \otimes \mathbb{C}^d)$ its second moment is

$$\mathcal{M}^{(2)}(O) = \int_{U(d)} d\mu(U)\, U^{\otimes 2} O U^{\dagger \otimes 2} \tag{93}$$

For $k = 2$ the symmetric group have elements $S_2 = \{Id, \leftrightarrow\}$ whose permutation operators are $V_d(Id) = \mathbb{I}_{2d}$ and $V_d(\leftrightarrow)$ =SWAP as we already see in an example. Then the second moment is

$$\mathcal{M}^{(2)}(O) = C_1(O)\mathbb{I}_{2d} + C_2(O)\text{SWAP} \tag{94}$$

To obtain the coefficients we multiply by each V_d element and perform the trace

$$\begin{aligned} Tr(O) &= Tr(\mathcal{M}^{(2)}(O)) = C_1 \underbrace{Tr(\mathbb{I}_{2d})}_{d^2} + C_2 \underbrace{Tr(\text{SWAP})}_{d} \Rightarrow d^2 C_1 + dC_2 = Tr(O) \\ Tr(\text{SWAP} \cdot O) &= C_1 \underbrace{Tr(\text{SWAP})}_{d} + C_2 \underbrace{Tr(\text{SWAP}^2)}_{d^2} \Rightarrow dC_1 + d^2 C_2 = Tr(\text{SWAP} \cdot O) \end{aligned} \tag{95}$$

Solving this set of linear equations gives us the coefficients:

$$\boxed{\begin{aligned} C_1(O) &= \frac{dTr(O) - Tr(\text{SWAP} \cdot O)}{d^2 - 1} \\ C_2(O) &= \frac{dTr(\text{SWAP} \cdot O) - Tr(O)}{d^2 - 1} \end{aligned}} \tag{96}$$