Proof that VQC is a Linear Classifier

consider

this works for arbitrary f.

that $\text{pred} = \langle \Phi(x) | W_\theta^\dagger f W_\theta | \Phi(x) \rangle$

$$= \mathrm{Tr}\left[\underset{\in \mathbb{R}^{2^n,2^n}}{|\Phi(x)\rangle\langle\Phi(x)|} \cdot \underset{\in \mathbb{R}^{2^n,2^n}}{W_\theta^\dagger Z W_\theta} \right]$$

← used: $\mathrm{Tr}[AB] = \mathrm{Tr}[BA]$

even if A, B are not square!

$A \in \mathbb{C}^{n,k}$ $B \in \mathbb{C}^{k,n}$

$n, k \in \mathbb{N}$

Now: decompose each of these in the operator basis (Pauli basis)

$$\Phi(x) = |\Phi(x)\rangle\langle\Phi(x)| = \frac{1}{2^n} \sum_{\alpha \in 4^n} \mathrm{Tr}\left[|\Phi_x\rangle\langle\Phi_x| \cdot P_\alpha \right] \cdot P_\alpha$$

$$\Phi(x) = \frac{1}{2^n} \sum_{\alpha \in 4^n} \Phi_\alpha(x) \cdot P_\alpha$$

$$H_\theta = \underbrace{W_\theta^\dagger Z W_\theta}_{} = \frac{1}{2^n} \sum_{\alpha \in 4^n} \mathrm{Tr}\left[W_\theta^\dagger Z W_\theta \cdot P_\alpha \right] \cdot P_\alpha$$

$$H_\theta = \frac{1}{2^n} \sum_{\alpha \in 4^n} w_\alpha(\theta) \cdot P_\alpha$$

← used: for a basis $\mathcal{B}$ and an operator O

we write O in the basis $\mathcal{B}$ as

$$O = \sum_{\beta_i \in \mathcal{B}} \langle \beta_i | O \rangle \beta_i$$

inner product between basis element β_i and O.

Which inner product depends on the space $\mathcal{B}, O$ live in

In that case, it's the Hilbert-Schmidt inner product.

$\mathrm{Tr}\left[|\Phi(x)\rangle\langle\Phi(x)| \cdot W_\theta^\dagger Z W_\theta \right]$

$\mathrm{Tr}\left[\Phi(x) \; H(x) \right]$

$$= \mathrm{Tr}\left[\left(\frac{1}{2^n} \sum_\alpha \Phi_\alpha(x) \cdot P_\alpha \right) \left(\frac{1}{2^n} \sum_\alpha w_\alpha(\theta) \; P_\alpha \right) \right]$$

$$= \mathrm{Tr}\left[\frac{1}{4^n} \sum_{\alpha,\beta} \Phi_\alpha(x) \, w_\beta(\theta) \cdot P_\alpha P_\beta \right]$$

$$= \frac{1}{4^n} \sum_{\alpha,\beta} \Phi_\alpha(x) \, w_\beta(\theta) \cdot \mathrm{Tr}\left[P_\alpha P_\beta \right]$$

$$= \frac{1}{4^n} \sum_\alpha \Phi_\alpha(x) \, W_\alpha(\theta) \cdot 2^n$$

$$= \frac{1}{2^n} \sum_\alpha \Phi_\alpha(x) \, W_\alpha(\theta)$$

$\in [-1, 1]$

$$\mathrm{Tr}\left[P_\alpha P_\beta \right] = \begin{cases} 2^n & \text{if } \alpha = \beta \\ 0 & \text{if } \alpha \,!\!= 0 \end{cases}$$

The obvious decision rule would be to threshold the output on 0,

However, say we pick an arbitrary threshold $b \in [-1, 1]$

then our decision rule is:

$$\text{label}(x) = \text{sign}\left(\sum_\alpha W_\alpha(\theta) \, \Phi_\alpha(x) + b \right)$$

Which is the exact same equation as our linear classifier. $\text{Sign}\left(\sum_i w_i x_i + b \right)$