Appendix for Power and limitations of single-qubit native quantum neural networks

A Detailed Proofs

A.1 Proof of Lemma 1

Lemma 1 *There exist* $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_L) \in \mathbb{R}^{L+1}$ *such that*

$$U^{YZY}_{\boldsymbol{\theta},L}(x) = \begin{bmatrix} P(x) & -Q(x) \\ Q^*(x) & P^*(x) \end{bmatrix} \tag{S1}$$

if and only if real Laurent polynomials $P, Q \in \mathbb{R}[e^{ix/2}, e^{-ix/2}]$ *satisfy*

1. $\deg(P) \leq L$ *and* $\deg(Q) \leq L$*,*
2. P *and* Q *have parity* $L \bmod 2$*,*
3. $\forall x \in \mathbb{R}$*,* $|P(x)|^2 + |Q(x)|^2 = 1$.

Proof First we prove the forward direction, i.e. Eq. (S1) implies the three conditions, by induction on L. For the base case, if $L = 0$, then $R_Y(\theta_0)$ gives $P = \cos(\theta_0/2)$ and $Q = \sin(\theta_0/2)$, which clearly satisfies the conditions.

For the induction step, suppose Eq. (S1) satisfies the three conditions. Then

$$\begin{aligned}
&\begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} R_Z(x) R_Y(\theta_k) && \text{(S2)}\\
&= \begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} \begin{bmatrix} \cos(\theta_k/2)e^{-ix/2} & -\sin(\theta_k/2)e^{-ix/2} \\ \sin(\theta_k/2)e^{ix/2} & \cos(\theta_k/2)e^{ix/2} \end{bmatrix} && \text{(S3)}\\
&= \begin{bmatrix} \cos\frac{\theta_k}{2}e^{-ix/2}P - \sin\frac{\theta_k}{2}e^{ix/2}Q & -\sin\frac{\theta_k}{2}e^{-ix/2}P - \cos\frac{\theta_k}{2}e^{ix/2}Q \\ \cos\frac{\theta_k}{2}e^{-ix/2}Q^* + \sin\frac{\theta_k}{2}e^{ix/2}P^* & \cos\frac{\theta_k}{2}e^{ix/2}P^* - \sin\frac{\theta_k}{2}e^{-ix/2}Q^* \end{bmatrix} && \text{(S4)}\\
&= \begin{bmatrix} \bar{P} & -\bar{Q} \\ \bar{Q}^* & \bar{P}^* \end{bmatrix} && \text{(S5)}
\end{aligned}$$

where

$$\bar{P} = \cos\frac{\theta_k}{2}e^{-ix/2}P - \sin\frac{\theta_k}{2}e^{ix/2}Q, \tag{S6}$$

$$\bar{Q} = \sin\frac{\theta_k}{2}e^{-ix/2}P + \cos\frac{\theta_k}{2}e^{ix/2}Q \tag{S7}$$

clearly satisfy the first condition: $\deg(\bar{P}) \leq L+1$ and $\deg(\bar{Q}) \leq L+1$. They also satisfy the second condition since multiplying by $e^{-ix/2}$ or $e^{ix/2}$ alters the parity. The third condition follows from unitarity.

Next we show the reverse by induction on L that the three conditions suffice to construct the QNN in Eq. (S1). For the base case, we have $L = 0$ and $\deg(P) = \deg(Q) = 0$. Note that P and Q are real polynomials, i.e. the coefficients must be real numbers, so condition 3 implies that $P = \cos(\theta_0/2)$ and $Q = \sin(\theta_0/2)$ for some $\theta_0 \in \mathbb{R}$. Thus the matrix

$$\begin{bmatrix} P & -Q, \\ Q^* & P^* \end{bmatrix} \tag{S8}$$

can be written as a QNN $U^{YZY}_{\boldsymbol{\theta},0}(x) = R_Y(\theta_0)$.

For the induction step, suppose Laurent polynomials P and Q satisfy the three conditions for some $L > 0$. We first observe that condition 3 implies

$$|P|^2 + |Q|^2 = PP^* + QQ^* = 1, \tag{S9}$$

where $PP^* + QQ^*$ is a Laurent polynomial of $e^{ix/2}$ with parity $2n \pmod 2 \equiv 0$. We denote $d := \max(\deg(P), \deg(Q))$, then $d \le L$ by the first condition. Since Eq. (S9) holds for all $x \in \mathbb{R}$, the leading coefficients must satisfy $p_d p_{-d} + q_d q_{-d} = 0$ so they cancel. We choose $\theta_k \in \mathbb{R}$ so that

$$\cos\frac{\theta_k}{2} p_d + \sin\frac{\theta_k}{2} q_d = 0, \tag{S10}$$

$$-\sin\frac{\theta_k}{2} p_{-d} + \cos\frac{\theta_k}{2} q_{-d} = 0. \tag{S11}$$

Now we briefly argue that there always exists a θ_k satisfies equations above. If all p_d, p_{-d}, q_d, q_{-d} are not 0, simply let $\tan\frac{\theta_k}{2} = -\frac{p_d}{q_d}$. Then consider the case of zero coefficients. Since $d = \max(\deg(P), \deg(Q))$, at most two of p_d, p_{-d}, q_d, q_{-d} could be 0. The fact $p_d p_{-d} + q_d q_{-d} = 0$ implies that one of p_d, p_{-d} is 0 if and only if one of q_d, q_{-d} is 0. If $p_d = q_d = 0$ (or $p_{-d} = q_{-d} = 0$), pick θ_k so that $\tan\frac{\theta_k}{2} = \frac{q_{-d}}{p_{-d}}$ (or $\tan\frac{\theta_k}{2} = -\frac{p_d}{q_d}$). If $p_d = q_{-d} = 0$ (or $p_{-d} = q_d = 0$), pick θ_k so that $\sin\frac{\theta_k}{2} = 0$ (or $\cos\frac{\theta_k}{2} = 0$).

Next, for the θ_k that satisfies Eq. (S10) and Eq. (S11) simultaneously, we consider

$$\begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} R_Y^\dagger(\theta_k) R_Z^\dagger(x) \tag{S12}$$

$$= \begin{bmatrix} \cos\frac{\theta_k}{2} e^{ix/2} P + \sin\frac{\theta_k}{2} e^{ix/2} Q & -\cos\frac{\theta_k}{2} e^{-ix/2} Q + \sin\frac{\theta_k}{2} e^{-ix/2} P \\ \cos\frac{\theta_k}{2} e^{ix/2} Q^* - \sin\frac{\theta_k}{2} e^{ix/2} P^* & \cos\frac{\theta_k}{2} e^{-ix/2} P^* + \sin\frac{\theta_k}{2} e^{-ix/2} Q \end{bmatrix} \tag{S13}$$

$$= \begin{bmatrix} \hat{P} & -\hat{Q} \\ \hat{Q}^* & \hat{P}^* \end{bmatrix} \tag{S14}$$

where

$$\hat{P} = \cos\frac{\theta_k}{2} e^{ix/2} P + \sin\frac{\theta_k}{2} e^{ix/2} Q, \tag{S15}$$

$$\hat{Q} = \cos\frac{\theta_k}{2} e^{-ix/2} Q - \sin\frac{\theta_k}{2} e^{-ix/2} P. \tag{S16}$$

Since $P, Q \in \mathbb{R}[e^{ix/2}, e^{-ix/2}]$ are Laurent polynomials with degree at most d, $\hat{P} \in \mathbb{R}[e^{ix/2}, e^{-ix/2}]$ might appear to be a Laurent polynomial with degree $d+1$. In fact it has degree $\deg(\hat{P}) \le d-1$, because the coefficient of $(e^{ix/2})^{d+1}$ term in $\hat{P}$ is $\cos\frac{\theta_k}{2} p_d + \sin\frac{\theta_k}{2} q_d = 0$ by the choice of θ_k, and the coefficient of $(e^{ix/2})^d$ term is 0 by condition 2. Similarly, $\hat{Q}$ has leading coefficient $\cos\frac{\theta_k}{2} q_{-d} - \sin\frac{\theta_k}{2} p_{-d} = 0$ and has degree $\deg(\hat{Q}) \le d-1$. Since $d \le L$, we have $\deg(\hat{P}) \le L-1$ and $\deg(\hat{Q}) \le L-1$, hence condition 1 is satisfied. From the parity of P and Q, it is easy to see that $\hat{P}$ and $\hat{Q}$ have parity $L-1 \bmod 2$, thus condition 2 is satisfied. Condition 3 follows from unitarity. By the induction hypothesis, Eq. (S14) can be written as a QNN in the form of $U^{\mathit{YZY}}_{\boldsymbol{\theta},L-1}(x)$, and therefore the matrix

$$\begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} \tag{S17}$$

can be written as a QNN in the form of $U^{\mathit{YZY}}_{\boldsymbol{\theta},L}(x)$. ■

We note that the above proof is in a similar spirit of the proof of quantum signal processing [33–35].

A.2 Proof of Proposition 2

We first restate the following lemma that was shown in Refs. [32, 34, 40].

Lemma S1 ([32, 34, 40]) *Suppose $A(x) \in \mathbb{R}[e^{ix}, e^{-ix}]$ is a real-valued Laurent polynomial with degree L that satisfies $A(x) \ge 0$ for $x \in \mathbb{R}$, then there exists a Laurent polynomial $P \in \mathbb{R}[e^{ix/2}, e^{-ix/2}]$ with $\deg(P) = L$ and parity $L \bmod 2$ such that $PP^* = A$.*

Note that the Laurent polynomial P in Lemma S1 could be computed via root finding. Then we are ready to prove Proposition 2 as follows.

Proposition 2 *For any even square-integrable function $f : [-\pi, \pi] \to \mathbb{R}$ and for all $\epsilon > 0$, there exists a QNN $U^{\mathit{YZY}}_{\boldsymbol{\theta},L}(x)$ such that $|\psi(x)\rangle = U^{\mathit{YZY}}_{\boldsymbol{\theta},L}(x)\,|0\rangle$ satisfies*

$$\|\, \langle\psi(x)|Z|\psi(x)\rangle - \alpha f(x) \| \leq \epsilon \tag{S18}$$

for some normalizing constant α.

Proof First, according to the Riesz–Fischer theorem [41], the Fourier series of any even square-integrable function $f(x)$ converges to $f(x)$ in the norm of L^2, i.e.

$$\lim_{K\to\infty} \|f_K(x) - f(x)\| = 0, \tag{S19}$$

where $f_K(x)$ is the K-term partial Fourier series of $f(x)$. For any $\epsilon > 0$, there exists a $K \in \mathbb{N}^+$ such that

$$\|f_K(x) - f(x)\| \leq \epsilon, \tag{S20}$$

where $f_K(x)$ is the K-term partial Fourier series of even function $f(x)$. Note that we can always multiply $f(x)$ and $f_K(x)$ by some constant α, without loss of generality, we assume $|f_K(x)| \leq 1$ for all $x \in \mathbb{R}$. Since $f(x)$ is an even function, the partial Fourier series $f_K(x)$ only contains cosine terms, i.e. $f_K(x) \in \mathbb{R}[e^{ix}, e^{-ix}]$. Then $F_K(x) := \frac{1+f_K(x)}{2}$ is a Laurent polynomial in $\mathbb{R}[e^{ix}, e^{-ix}]$ with degree K such that $0 \leq F_K(x) \leq 1$ for $x \in \mathbb{R}$.

By Lemma S1, there exists a Laurent polynomial $P \in \mathbb{R}[e^{ix/2}, e^{-ix/2}]$ with degree K and parity $K \bmod 2$ such that $PP^* = F_K(x)$. Notice that $|P(x)| \leq 1$ for all $x \in \mathbb{R}$ since $|P(x)|^2 = F_K(x) \leq 1$. By Lemma 1, there exists a QNN $U^{\mathit{YZY}}_{\boldsymbol{\theta},K}(x)$ such that

$$U^{\mathit{YZY}}_{\boldsymbol{\theta},K}(x) = \begin{bmatrix} P(x) & -Q(x) \\ Q^*(x) & P^*(x) \end{bmatrix}. \tag{S21}$$

Let $|\psi(x)\rangle = U^{\mathit{YZY}}_{\boldsymbol{\theta},K}(x)\,|0\rangle$, we have

$$\langle\psi(x)|Z|\psi(x)\rangle = PP^* - QQ^* = 2PP^* - 1 = 2F_K(x) - 1 = f_K(x). \tag{S22}$$

It follows from Eq. (S20) that

$$\|\, \langle\psi(x)|Z|\psi(x)\rangle - f(x)\| \leq \epsilon. \tag{S23}$$

■

A.3 Proof of Lemma 3

Lemma 3 *There exist $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_L) \in \mathbb{R}^{L+1}$ and $\boldsymbol{\phi} = (\varphi, \phi_0, \phi_1, \dots, \phi_L) \in \mathbb{R}^{L+2}$ such that*

$$U^{\mathit{WZW}}_{\boldsymbol{\theta},\boldsymbol{\phi},L}(x) = \begin{bmatrix} P(x) & -Q(x) \\ Q^*(x) & P^*(x) \end{bmatrix} \tag{S24}$$

if and only if Laurent polynomials $P, Q \in \mathbb{C}[e^{ix/2}, e^{-ix/2}]$ satisfy

1. $\deg(P) \leq L$ *and* $\deg(Q) \leq L$,
2. P *and* Q *have parity* $L \bmod 2$,
3. $\forall x \in \mathbb{R}$, $|P(x)|^2 + |Q(x)|^2 = 1$.

Proof The entire proof is similar as the proof of Lemma 1, and the only difference is that the coefficients of Laurent polynomials could be complex rather than real. First we prove the forward direction by induction on L. For the base case, if $L = 0$, then $U^{\mathit{WZW}}_{\boldsymbol{\theta},\boldsymbol{\phi},L}(x) = R_Z(\varphi)R_Y(\theta_0)R_Z(\phi_0)$ gives $P = e^{-i(\varphi+\phi_0)/2}\cos(\theta_0/2)$ and $Q = e^{-i(\varphi-\phi_0)/2}\sin(\theta_0/2)$, which clearly satisfies the three conditions.

For the induction step, suppose Eq. (S24) satisfies the three conditions. Then

$$\begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} R_Z(x) R_Y(\theta_k) R_Z(\phi_k) \tag{S25}$$

$$= \begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} \begin{bmatrix} \cos\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} & -\sin\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} \\ \sin\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{i\frac{x}{2}} & \cos\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{i\frac{x}{2}} \end{bmatrix} \tag{S26}$$

$$= \begin{bmatrix} \cos\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} P - \sin\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{i\frac{x}{2}} Q & -\sin\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} P - \cos\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{i\frac{x}{2}} Q \\ \cos\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} Q^* + \sin\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{i\frac{x}{2}} P^* & \cos\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{i\frac{x}{2}} P^* - \sin\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} Q^* \end{bmatrix} \tag{S27}$$

$$= \begin{bmatrix} \bar{P} & -\bar{Q} \\ \bar{Q}^* & \bar{P}^* \end{bmatrix} \tag{S28}$$

where

$$\bar{P} = \cos\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} P - \sin\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} e^{i\frac{x}{2}} Q, \tag{S29}$$

$$\bar{Q} = \sin\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{-i\frac{x}{2}} P + \cos\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} e^{i\frac{x}{2}} Q \tag{S30}$$

clearly satisfy the first condition: $\deg(\bar{P}) \le L+1$ and $\deg(\bar{Q}) \le L+1$. They also satisfy the second condition since multiplying by $e^{-ix/2}$ or $e^{ix/2}$ alters the parity. The third condition is satisfied because of unitarity.

Next we show the backward direction by induction on L that the three conditions suffice to construct the QNN in Eq. (S24). First we consider the base case of $L = 0$, we have $\deg(P) = \deg(Q) = 0$, the above condition implies that $P = e^{-i(\varphi+\phi_0)/2}\cos(\theta_0/2)$ and $Q = e^{-i(\varphi-\phi_0)/2}\sin(\theta_0/2)$ for some $\varphi, \theta_0, \phi_0 \in \mathbb{R}$. Thus the matrix

$$\begin{bmatrix} P & -Q, \\ Q^* & P^* \end{bmatrix} \tag{S31}$$

can be written as a QNN $U^{WZW}_{\boldsymbol{\theta},\boldsymbol{\phi},L}(x) = R_Z(\varphi) R_Y(\theta_0) R_Z(\phi_0)$.

For the induction step, suppose Laurent polynomials P and Q satisfy the three conditions for some $L > 0$. We first observe that condition 3 implies

$$|P|^2 + |Q|^2 = PP^* + QQ^* = 1, \tag{S32}$$

where $PP^* + QQ^*$ is a Laurent polynomial of $e^{ix/2}$ with parity $2n \pmod 2 \equiv 0$. Since Eq. (S32) holds for all $x \in \mathbb{R}$, the leading coefficients must satisfy $p_d p^*_{-d} + q_d q^*_{-d} = 0$ and $p^*_d p_{-d} + q^*_d q_{-d} = 0$, where $d := \max(\deg(P), \deg(Q))$. By the first condition, we have $d \le L$. We choose $\theta_k, \phi_k \in \mathbb{R}$ so that

$$\cos\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} p_d + \sin\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} q_d = 0, \tag{S33}$$

$$-\sin\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} p_{-d} + \cos\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} q_{-d} = 0. \tag{S34}$$

Now we briefly argue that there always exists θ_k and ϕ_k satisfy equations above. If all p_d, p_{-d}, q_d, q_{-d} are not zero, simply let $\tan\left(\frac{\theta_k}{2}\right) e^{-i\phi_k} = -\frac{p_d}{q_d}$. Then consider the case of zero coefficients. Since $\deg(P) = \deg(Q) = d$, at most two of p_d, p_{-d}, q_d, q_{-d} could be 0. The fact $p_d p^*_{-d} + q_d q^*_{-d} = 0$ implies that one of one of p_d, p_{-d} is 0 if and only if one of q_d, q_{-d} is 0. If $p_d = q_d = 0$ (or $p_{-d} = q_{-d} = 0$), pick θ_k and ϕ_k so that $\tan\left(\frac{\theta_k}{2}\right) e^{i\phi_k} = \frac{q_{-d}}{p_{-d}}$. If $p_d = q_{-d} = 0$ or $(p_{-d} = q_d = 0)$, pick θ_k so that $\sin\left(\frac{\theta_k}{2}\right) = 0$ (or $\cos\left(\frac{\theta_k}{2}\right) = 0$).

Next, for the θ_k that we pick, consider

$$\begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} W^\dagger(\theta_k, \phi_k) R_Z^\dagger(x) \tag{S35}$$

$$= \begin{bmatrix} e^{i\frac{\phi_k}{2}} \cos\frac{\theta_k}{2} e^{i\frac{x}{2}} P + e^{-i\frac{\phi_k}{2}} \sin\frac{\theta_k}{2} e^{i\frac{x}{2}} Q & e^{i\frac{\phi_k}{2}} \sin\frac{\theta_k}{2} e^{-i\frac{x}{2}} P - e^{-i\frac{\phi_k}{2}} \cos\frac{\theta_k}{2} e^{-i\frac{x}{2}} Q \\ e^{i\frac{\phi_k}{2}} \cos\frac{\theta_k}{2} e^{i\frac{x}{2}} Q^* - e^{-i\frac{\phi_k}{2}} \sin\frac{\theta_k}{2} e^{i\frac{x}{2}} P^* & e^{-i\frac{\phi_k}{2}} \cos\frac{\theta_k}{2} e^{-i\frac{x}{2}} P^* + e^{i\frac{\phi_k}{2}} \sin\frac{\theta_k}{2} e^{-i\frac{x}{2}} Q^* \end{bmatrix} \tag{S36}$$

$$= \begin{bmatrix} \hat{P} & -\hat{Q} \\ \hat{Q}^* & \hat{P}^* \end{bmatrix} \tag{S37}$$

where

$$\hat{P} = e^{i\frac{\phi_k}{2}} \cos\frac{\theta_k}{2} e^{i\frac{x}{2}} P + e^{-i\frac{\phi_k}{2}} \sin\frac{\theta_k}{2} e^{i\frac{x}{2}} Q, \tag{S38}$$

$$\hat{Q} = e^{-i\frac{\phi_k}{2}} \cos\frac{\theta_k}{2} e^{-i\frac{x}{2}} Q - e^{i\frac{\phi_k}{2}} \sin\frac{\theta_k}{2} e^{-i\frac{x}{2}} P. \tag{S39}$$

Since $P, Q \in \mathbb{C}[e^{ix/2}, e^{-ix/2}]$ are Laurent polynomials with degree at most d, $\hat{P} \in \mathbb{R}[e^{ix/2}, e^{-ix/2}]$ might appear to be a Laurent polynomial with degree at most $d+1$. In fact it has degree $\deg(\hat{P}) \leq d-1$, because the coefficient of $(e^{i\frac{x}{2}})^{d+1}$ term in $\hat{P}$ is $\cos\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} p_d + \sin\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} q_d = 0$ by the selected θ_k and ϕ_k, and the coefficients of $(e^{\frac{ix}{2}})^d$ term is 0 by condition 2. Similarly, $\hat{Q}$ has leading coefficient $-\sin\frac{\theta_k}{2} e^{i\frac{\phi_k}{2}} p_{-d} + \cos\frac{\theta_k}{2} e^{-i\frac{\phi_k}{2}} q_{-d} = 0$ and has degree $\deg(\hat{Q}) \leq d-1$. Since $d \leq L$, we have $\deg(\hat{P}) \leq L-1$ and $\deg(\hat{Q}) \leq L-1$, hence condition 1 is satisfied. From the parity of P and Q, it is easy to see that $\hat{P}$ and $\hat{Q}$ have parity $L-1 \bmod 2$, thus condition 2 is satisfied. Condition 3 follows from unitarity. Thus by the induction hypothesis, Eq. (S37) can be written as a QNN in the form of $U^{WZW}_{\boldsymbol{\theta},\boldsymbol{\phi},L-1}(x)$, and therefore the matrix

$$\begin{bmatrix} P & -Q \\ Q^* & P^* \end{bmatrix} \tag{S40}$$

can be written as a QNN in the form of $U^{WZW}_{\boldsymbol{\theta},\boldsymbol{\phi},L}(x)$. ■

A.4 Proof of Theorem 4

We first restate the following lemma that was proved in Refs. [36, 37].

Lemma S2 ([36, 37]) *Suppose $A(x) \in \mathbb{C}[e^{ix}, e^{-ix}]$ is a real-valued Laurent polynomial with degree L that satisfies $A(x) \geq 0$ for all $x \in \mathbb{R}$, then there exists a Laurent polynomial $P \in \mathbb{C}[e^{ix/2}, e^{-ix/2}]$ with $\deg(P) = L$ and parity $L \bmod 2$ such that $PP^* = A$.*

Lemma S2 is a generalized version of Lemma S1, which extends the coefficients to complex numbers. Similarly, the Laurent polynomial P in Lemma S2 could be computed via root finding. Now we prove Theorem 4 as follows.

Theorem 4 (Univariate approximation properties of single-qubit QNNs.) *For any univariate square-integrable function $f : [-\pi, \pi] \to \mathbb{R}$ and for all $\epsilon > 0$, there exists a QNN $U^{WZW}_{\boldsymbol{\theta},\boldsymbol{\phi},L}(x)$ such that $|\psi(x)\rangle = U^{WZW}_{\boldsymbol{\theta},\boldsymbol{\phi},L}(x)\,|0\rangle$ satisfies*

$$\| \langle\psi(x)|Z|\psi(x)\rangle - \alpha f(x) \| \leq \epsilon \tag{S41}$$

for some normalizing constant α.

Proof Similar as in the proof of Proposition 2, we use the Riesz–Fischer theorem [41] that the Fourier series of any square-integrable function $f(x)$ converges to $f(x)$ in the norm of L^2. For any $\epsilon > 0$, there exists a $K \in \mathbb{N}^+$ such that

$$\|f_K(x) - f(x)\| \leq \epsilon, \tag{S42}$$

where $f_K(x)$ is the K-term truncated Fourier series of $f(x)$. Note that we can always multiply $f(x)$ and $f_K(x)$ by some constant α, without loss of generality, we assume $|f_K(x)| \leq 1$ for all $x \in \mathbb{R}$. Since $f(x)$ is a general square-integrable function, the partial Fourier series $f_K(x)$ is a complex Laurent polynomial with degree K in $\mathbb{C}[e^{ix}, e^{-ix}]$. Then $F_K(x) := \frac{1+f_K(x)}{2}$ is a Laurent polynomial in $\mathbb{C}[e^{ix}, e^{-ix}]$ with degree K such that $0 \leq F_K(x) \leq 1$ for all $x \in \mathbb{R}$.

By Lemma S2, there exists a Laurent polynomial $P \in \mathbb{C}[e^{ix/2}, e^{-ix/2}]$ with degree K and parity $K \bmod 2$ such that $PP^* = F_K(x)$. Notice that $|P(x)| \leq 1$ for all $x \in \mathbb{R}$ since $|P(x)|^2 = F_K(x) \leq 1$. By Lemma 3, there exists a QNN $U^{\textit{WZW}}_{\boldsymbol{\theta},\boldsymbol{\phi},K}(x)$ such that

$$U^{\textit{WZW}}_{\boldsymbol{\theta},\boldsymbol{\phi},K}(x) = \begin{bmatrix} P(x) & -Q(x) \\ Q^*(x) & P^*(x) \end{bmatrix}. \tag{S43}$$

Let $|\psi(x)\rangle = U^{\textit{WZW}}_{\boldsymbol{\theta},\boldsymbol{\phi},K}(x)\,|0\rangle$, we have

$$\langle\psi(x)|Z|\psi(x)\rangle = PP^* - QQ^* = 2PP^* - 1 = 2F_K(x) - 1 = f_K(x). \tag{S44}$$

It follows from Eq. (S42) that

$$\|\,\langle\psi(x)|Z|\psi(x)\rangle - f(x)\| \leq \epsilon. \tag{S45}$$

■

B Limitations on representing multivariate Fourier series

In this section, we show that if the single-qubit native QNN is written in the form of a K-truncated multivariate Fourier series, it has a rich Fourier frequency set Ω, but it cannot meet the requirements of the corresponding Fourier coefficients set C_Ω due to the curse of dimensionality. Specifically, let $\boldsymbol{x} := (x^{(1)}, x^{(2)}, \ldots, x^{(d)}) \in \mathbb{R}^d$ and the single-qubit native QNN is defined as follows:

$$U_{\boldsymbol{\theta},L}(\boldsymbol{x}) = U_3(\theta_0, \phi_0, \lambda_0) \prod_{j=1}^{L} R_Z(x_j) U_3(\theta_j, \phi_j, \lambda_j), \tag{S46}$$

where x_j is a one-dimensional data of $\boldsymbol{x}$, i.e., $x_j \in \{x^{(m)} \mid m = 1, \cdots, d\}$. The quantum circuit is shown in Fig. S1.

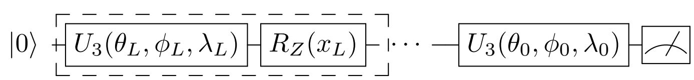


Figure S1: Circuit of $U_{\boldsymbol{\theta},L}(\boldsymbol{x})$, where the trainable block is composed of $U_3(\cdot)$, and the encoding block is $R_Z(\cdot)$. Note that there is no restriction on the encoding order of $x^{(m)}$.

Without loss of generality, assume that each one-dimensional data $x^{(m)}$ is uploaded the same number of times, denoted by K, then we have $Kd = L$. Further, we write $V_j := U_3(\theta_j, \phi_j, \lambda_j)$ for short, that is

$$U_{\boldsymbol{\theta},L}(\boldsymbol{x}) = V_0 \prod_{j=1}^{L} \begin{bmatrix} e^{-ix_j\lambda_0} & 0 \\ 0 & e^{-ix_j\lambda_1} \end{bmatrix} V_j, \tag{S47}$$

where $\lambda_0 = \frac{1}{2}$ and $\lambda_1 = -\frac{1}{2}$. More generally, we have

$$U_{\boldsymbol{\theta},L}(\boldsymbol{x}) = \sum_{j_1,\cdots,j_L=0}^{1} e^{-i(\lambda_{j_1}x_1 + \cdots + \lambda_{j_L}x_L)}\, V_0\, |j_1\rangle\!\langle j_1|\, V_1 \cdots |j_L\rangle\!\langle j_L|\, V_L. \tag{S48}$$

Next, we reformulate the above equation using $x^{(m)}$ instead of x_j. We denote $I_m \subset \{1, \cdots, L\}$ as the index set of encoding block $R_Z(x^{(m)})$ and $I'_m = \{\lambda_{j_k} \mid k \in I_m\}$, then

$$U_{\boldsymbol{\theta},L}(\boldsymbol{x}) = \sum_{\boldsymbol{j}\in\{0,1\}^L} e^{-i(\Lambda^{(1)}_{\boldsymbol{j}} x^{(1)} + \cdots + \Lambda^{(d)}_{\boldsymbol{j}} x^{(d)})}\, V_0\, |j_1\rangle\!\langle j_1|\, V_1 \cdots |j_L\rangle\!\langle j_L|\, V_L, \tag{S49}$$

where $\boldsymbol{j}$ is bit-strings composed of $j_l, l = 1, \cdots, L$ and $\Lambda_{\boldsymbol{j}}^{(m)} = \sum \lambda$, $\lambda \in I'_m$. Further, for the initial state $|0\rangle$ and some observable M, we have

$$f_{\boldsymbol{\theta},L}(\boldsymbol{x}) = \langle 0| U^{\dagger}_{\boldsymbol{\theta},L}(\boldsymbol{x}) M U_{\boldsymbol{\theta},L}(\boldsymbol{x}) |0\rangle = \sum_{\boldsymbol{\omega}\in\Omega} c_{\boldsymbol{\omega}} e^{i\boldsymbol{\omega}\cdot\boldsymbol{x}}, \tag{S50}$$

where $\Omega = \{h_{\boldsymbol{j},\boldsymbol{k}} \coloneqq (\Lambda_{\boldsymbol{k}}^{(1)} - \Lambda_{\boldsymbol{j}}^{(1)}, \cdots, \Lambda_{\boldsymbol{k}}^{(d)} - \Lambda_{\boldsymbol{j}}^{(d)}) \mid \boldsymbol{j}, \boldsymbol{k} \in \{0,1\}^L\}$. We consider the m-th dimension,

$$\Lambda_{\boldsymbol{k}}^{(m)} - \Lambda_{\boldsymbol{j}}^{(m)} = \{(\lambda_{k_1} + \cdots + \lambda_{k_K}) - (\lambda_{j_1} + \cdots + \lambda_{j_K}) \mid \lambda_{k_1} \cdots \lambda_{k_K}, \lambda_{j_1} \cdots \lambda_{j_K} \in I'_m\}. \tag{S51}$$

Since $\lambda = \pm\frac{1}{2}$, for all $\lambda \in I'_m$, we can derive

$$\{\Lambda_{\boldsymbol{k}}^{(m)} - \Lambda_{\boldsymbol{j}}^{(m)} \mid \boldsymbol{j}, \boldsymbol{k} \in \{0,1\}^L\} = \{-K, \cdots, 0, \cdots, K\}. \tag{S52}$$

Then the Fourier frequency spectrum is $\Omega = \{-K, \cdots, 0, \cdots, K\}^d$. For the Fourier coefficient set $C_\Omega = \{c_{\boldsymbol{\omega}} \mid \boldsymbol{\omega} \in \Omega\}$, we have

$$c_{\boldsymbol{\omega}} = \sum_{\boldsymbol{j},\boldsymbol{k}\in\{0,1\}^L, \boldsymbol{\omega}=h_{\boldsymbol{j},\boldsymbol{k}}} C_{\boldsymbol{j},\boldsymbol{k}}, \tag{S53}$$

and

$$C_{\boldsymbol{j},\boldsymbol{k}} = \langle 0| V_L^{\dagger} |k_L\rangle\langle k_L| V_{L-1}^{\dagger} \cdots |k_1\rangle\langle k_1| V_1^{\dagger} M V_1 |j_1\rangle\langle j_1| \cdots V_{L-1} |j_L\rangle\langle j_L| V_L |0\rangle . \tag{S54}$$

If the number of truncation terms K is fixed, as the dimension d increases, the degrees of freedom of the set C_Ω increase with $O(d)$, so it cannot represent any exponential size Fourier coefficients set. Thus for all $\boldsymbol{\theta} \in \mathbb{R}^{3(L+1)}$, the Eq. (S50) cannot express an arbitrary k-truncated multivariate Fourier series.

C Extension to multi-qubit QNNs

We have already shown that single-qubit native QNNs are able to approximate any univariate function but possibly could not approximate arbitrary multivariate functions by Fourier series in Section 3 and Section 4. To address this limitation, research into the extension approach of single-qubit native QNNs is crucial. For classical NNs, a common approach toward overcoming such limitations is to increase the width or depth of the networks. We conjecture that QNNs have similar characteristics. We hereby provide a multi-qubit extension strategy as shown in Fig. S2, called *Parallel-Entanglement*, introducing quantum entanglement by multi-qubit gates such as CNOT gates. Similar to classical NNs in which different neurons are connected by trainable parameters to form deep NNs, QNNs can establish the connection between different qubits through quantum entanglement.

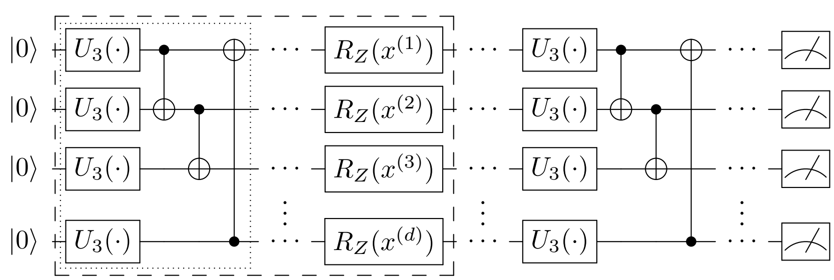
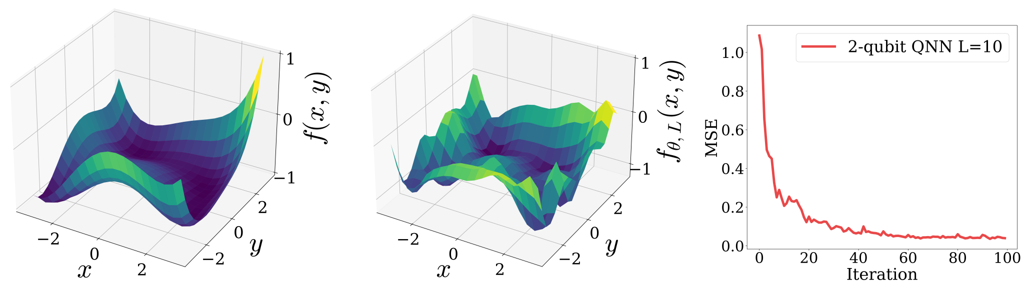


Figure S2: The Parallel-Entanglement model of d qubits and L layers. Each layer consists of a trainable block and an encoding block in the dashed box. Each trainable block is composed of repeated sub-blocks of U_3 rotation gates and CNOT gates, as shown in the dotted box. The number of sub-blocks in each trainable block is denoted by D_{tr}. The encoding block is a d-tensor of $R_Z(x^{(m)})$ gates for a d-dimensional data $\boldsymbol{x} \coloneqq (x^{(1)}, \ldots, x^{(d)})$.

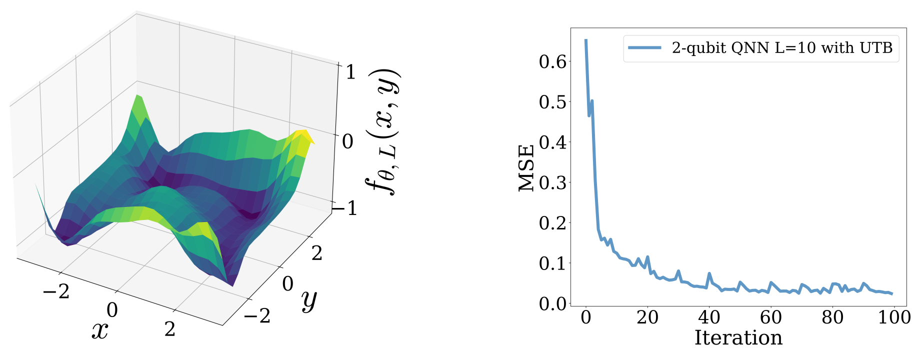
We numerically show that the multi-qubit extension as shown in Fig. S2 could improve the expressivity of single-qubit QNNs. Consider the same bivariate function $f(x, y) = (x^2+y-1.5\pi)^2+(x+y^2-\pi)^2$

used in Section 5.2, we use a two-qubit QNN of $L = 10$ and $D_{tr} = 3$ to approximate the target function $f(x, y)$ with the same training setting. The experimental results are shown in Fig. S3. Compared with the approximation results of single-qubit QNNs in Fig. 6, we can see that the two-qubit QNN has stronger expressive power than single-qubit models. Moreover, we could build the *universal trainable block* (UTB) using a universal two-qubit quantum gate [42, 43] consisting of U_3 and CNOT gates. Specifically, a two-qubit universal trainable block can express any two-qubit unitary matrix. Using a two-qubit QNN with UTBs yields a better approximation result which is shown in Fig. S4. From the numerical results, we could see that the multi-qubit extension could potentially overcome the limitations of single-qubit QNNs on approximating multivariate functions as illustrated in Section 4.



(a) Target function. (b) Approximation result. (c) Training loss.

Figure S3: Panel (a) is the plot of target function $f(x, y)$. Panel (b) shows the approximation result of a two-qubit QNN of $L = 10$ and $D_{tr} = 3$. Panel (c) is the plot of training loss during the optimization.



(a) Approximation result. (b) Training loss.

Figure S4: Panel (a) shows the approximation result of a two-qubit QNN of $L = 10$ with UTB. Panel (b) is the plot of training loss during the optimization.

We further illustrate the ability of Parallel-Entanglement models to address practical problems through experiments on the classification task. The public benchmark data sets are used to demonstrate the capabilities of Parallel-Entanglement models to tackle classification tasks.

The performance of the Parallel-Entanglement model on classifying Iris [44], Breast Cancer [45], and HTRU2 [46] data sets are summarized in Table 1. Specifically, 100 pieces of data are sampled from the dataset, with 80% of them serving as the training set and 20% serving as the test set. We use the Adam optimizer with a learning rate of 0.1 and a batch size of 40 to train the multi-qubit QNNs. In order to reduce the effect of randomness, the results of classification accuracy are averaged over 10 independent training instances.

The Iris data set contains 3 different classes, each class only has four attributes. Obviously, the 4-qubit model easily obtains an average accuracy of over 99% with only 1 layer on Iris data. The HTRU2 data set only has two categories, and each example has 8 attributes. As a result, an 8-qubit QNN is required to complete this task. The average accuracy for binary classification with

Table 1: The performance of the Parallel-Entanglement model.

Dataset	**# of qubits**	L	D_{tr}	**# of parameters**	**Average accuracy**
Iris	4	1	1	16	0.990 ± 0.01
HTRU2	8	1	1	32	0.910 ± 0.09
		3	1	64	0.970 ± 0.03
		3	2	128	0.980 ± 0.02
Breast Cancer	4	1	1	16	0.780 ± 0.06
		3	1	32	0.840 ± 0.02
		3	2	64	0.845 ± 0.04

1 layer achieves above 91%. We can see that adding the number of layers to 3 and the depth of each layer to 2 increases the average accuracy to 98%. Since the Breast Cancer data contains 30 features for binary classification, the principal component analysis (PCA) is used to reduce feature dimension. Here the numerical results of a 4-qubit model are given to illustrate the power of the QNN. Compared to the model using complete information, the QNN model does not perform perfectly. But increasing the number of layers and the depth may improve the test accuracy. Based on the finding of the preceding experiments, it is clear that the Parallel-Entanglement model is capable of handling practical classification problems.