

# One qubit is an universal approximator of any complex function

In these notes we study the Fourier series and the Universal Approximation Theorem in their classical forms. Then, we show that quantum parameterized circuits can make use of quantum versions of both theorems in a natural way. The Fourier series can be recovered term by term with quantum circuits, and the Universal Approximation Theorem has got a quantum analogue.

## I. INTRODUCTION

Several theorems concerning how to represent approximately represent an arbitrary function as a sum series of some other functions. One of the most famous results is the Taylor expansion [1], which has been broadly applied in several fields. In these notes we will focus on two different approaches to demonstrate universality of a series, being the Fourier series and the Universal approximation theorem.

The Fourier series is a constructive method that allows to write any target function defined within an interval as a sum of sines and cosines with fixed frequencies. The final result lies in the fact that sines and cosines of discrete frequencies are orthogonal functions if the inner product is the integral. Following this line of thought, it is possible to map the space of functions into a linear vector space. Thus, any function is a vector of this space, and sines and cosines are the elements of the basis. In conclusion, for any function it is possible to construct any function as a linear combination of basis vectors. The conditions of applicability of this theorem are very broad, suited functions are only required to have a finite number of finite discontinuities and to be integrable.

The Universal Approximation Theorem shows that the set of sums of weighted sigmoidal functions is dense in the space of continuous functions, and that there exist a sum of this kind arbitrarily close to any target function [3]. However, this method is not constructive, we must be able to find the parameters of the expansion by other means. This is the task performed by Neural Networks.

We can recover the results from Fourier series with the UAT theorem. There exists an extension of this theorem that enlarges the possibility. In this approach, sigmoidal functions can be substituted by non-linear, bounded and non-constant functions [2]. In particular, sines and cosines, as in Fourier series. Note also that trigonometric functions can be constructed as a sum of sigmoids, but this would increase the number of terms. Therefore, if the UAT can have the form of a Fourier series, then UAT will be *at least* as good as Fourier. In the worst case situation, the parameters of UAT will be Fourier's, but there may be other parameters that fit the target function better.

In these notes we show that both Fourier series and the UAT can be implemented using Quantum Parameterized circuits. We first tackle Fourier, state the known results and link the quantum circuit to them. Then, we move to the UAT, follow the original proof, and apply a similar method for showing that quantum parameterized circuits can also implement the UAT.

## II. DEMONSTRATION LINKING WITH FOURIER SERIES

The Fourier series is a well-known result of mathematical analysis. It shows that any complex variable can be expressed in terms of a formally infinite sum of sines and cosines with fixed frequencies.

### A. The Fourier series

This is the statement of the Fourier series as explained in Ref. [1].

**Theorem 1** *Let  $f$  be any function  $f : \mathbf{R} \rightarrow \mathbf{R}$  with a finite number of finite discontinuities and a finite number of extreme values integrable within an interval  $[a, b] \in \mathbf{R}$  of length  $P$  with. Then  $f(x)$  can be approximated by*

$$f_N(x) = \frac{a_0}{2} + \sum_{n=1}^N \left( a_n \cos\left(\frac{2\pi nx}{P}\right) + b_n \sin\left(\frac{2\pi nx}{P}\right) \right), \quad (1)$$

where

$$a_n = \frac{2}{P} \int_P f(x) \cos\left(\frac{2\pi nx}{P}\right) dx, \quad (2)$$

$$b_n = \frac{2}{P} \int_P f(x) \sin\left(\frac{2\pi nx}{P}\right) dx. \quad (3)$$

$$(4)$$

Notice that if  $f(x)$  is  $P$ -periodic, this representation is sufficient to approximate the function in all its domain. The series of functions  $\{f_N(x)\}$  satisfy

$$\lim_{N \rightarrow \infty} f_N(x) = f(x) \quad (5)$$

The usage of sines and cosines is very convenient as this functions satisfy orthogonality properties

$$\int_a^b \sin\left(\frac{2\pi n}{P}x\right) \sin\left(\frac{2\pi m}{P}x\right) = \delta_{mn} \quad (6)$$

$$\int_a^b \cos\left(\frac{2\pi n}{P}x\right) \cos\left(\frac{2\pi m}{P}x\right) = \delta_{mn} \quad (7)$$

$$\int_a^b \sin\left(\frac{2\pi n}{P}x\right) \cos\left(\frac{2\pi m}{P}x\right) = 0 \quad (8)$$

$$(9)$$

There is another manner of expressing the formula in Eq. (1) using complex variable:

$$f_N(X) = \sum_{n=-N}^N c_n e^{i \frac{2\pi nx}{P}}, \quad (10)$$

where  $c_n$  is related to the previous coefficients by

$$c_n = \frac{a_n + ib_n}{2} \quad (11)$$

$$c_{-n} = \frac{a_n - ib_n}{2} \quad (12)$$

$$. \quad (13)$$

Once we have the expression Fourier series for a real-valued function, it is easy to understand that it can extended to be a complex-valued function. If  $z(x)$  is complex valued and both real and imaginary components are real-valued and can be represented by a Fourier series, then the coefficients  $c_n$  can be adjusted to yield  $z(x)$

**Theorem 2** *Let  $f$  be any function  $f : \mathbf{R} \rightarrow \mathbf{R}$  with a finite number of finite discontinuities and a finite number of extreme values integrable within an interval  $[a, b] \in \mathbf{R}$  of length  $P$  with. Then  $f(x)$  can be approximated by*

$$z_N(X) = \sum_{n=-N}^N c_n e^{i \frac{2\pi nx}{P}}, \quad (14)$$

where

$$\text{Re}(c_n) = \frac{1}{P} \int_P \text{Re}(z(x)) e^{-i \frac{2\pi nx}{P}} \quad (15)$$

$$\text{Im}(c_n) = \frac{1}{P} \int_P \text{Im}(z(x)) e^{-i \frac{2\pi nx}{P}} \quad (16)$$

$$c_n = \frac{1}{P} \int_P z(x) e^{-i \frac{2\pi nx}{P}} \quad (17)$$

Notice that this result is identical to Eq. (14), but this time  $c_n$  and  $c_{-n}$  are no longer complex conjugates. These tools are sufficient to demonstrate that sines and cosines form a basis of orthogonal functions able to represent any function.

### B. Fourier for quantum circuits (qFourier)

Let  $|\psi(x)\rangle$  be a quantum state encoding some complex function in within. Although a quantum state has got two complex coefficients, only two degrees of freedom define it. As the quantum state must be normalized and a global phase is lost when measuring, there is only one modulus and one complex phase left to define the function. In addition, the modulus of this fuction is bounded by 1. The explicit form of the wavefunction is taken to be

$$|\psi(x)\rangle = e^{i\theta(x)} \left( e^{i\phi(x)} r(x) |0\rangle + \sqrt{1 - r(x)^2} |1\rangle \right), \quad (18)$$

where  $0 \leq r(x) \leq 1$ .

The aim now is to design a unitary gate  $U(x; \omega, \alpha, \beta, \varphi, \lambda)$  such that

$$|\psi_N(x)\rangle = \prod_{i=0}^N U(x; \omega_i, \alpha_i, \beta_i, \varphi_i, \lambda_i) |\psi_0\rangle, \quad (19)$$

where  $|\psi_0\rangle$  may be any arbitrary fixed state, and  $|\psi_N(x)\rangle$  is close to the arbitrary  $|\psi(x)\rangle$ .

The intuition behind this idea is that every gate contributes to the final state with the equivalent of a Fourier term, that is, roughly speaking

$$U(x; \omega, \alpha, \beta, \varphi, \lambda) \sim \tilde{\alpha} e^{i\tilde{\varphi}} e^{i\tilde{\omega}x} + \tilde{\beta} e^{i\tilde{\lambda}} e^{-i\tilde{\omega}x}, \quad (20)$$

and at the end of the day, the more gates we add to the system, the closer we will be to the state  $|\psi(x)\rangle$ . Let us take the gate to be

$$U(x; \omega, \alpha, \beta, \varphi, \lambda) = R_z \left( \frac{\alpha + \beta}{2} \right) R_y(\lambda) R_z \left( \frac{\alpha - \beta}{2} \right) R_z(\omega x) R_y(\varphi) \quad (21)$$

In order to show that this gate is equivalent to one term of the Fourier series, it is convenient to write it in a Fourier form, i. e., with terms of the form  $e^{\pm i\omega x}$ . A simple substitution yields the result

$$U(x; \omega, \alpha, \beta, \varphi, \lambda) = \begin{pmatrix} \cos \lambda \cos \varphi e^{i\alpha} e^{i\omega x} - \sin \lambda \sin \varphi e^{i\beta} e^{-i\omega x} & -\cos \lambda \sin \varphi e^{i\alpha} e^{i\omega x} - \sin \lambda \cos \varphi e^{i\beta} e^{-i\omega x} \\ \sin \lambda \cos \varphi e^{-i\beta} e^{i\omega x} + \cos \lambda \sin \varphi e^{-i\alpha} e^{-i\omega x} & -\sin \lambda \sin \varphi e^{-i\beta} e^{i\omega x} + \cos \lambda \cos \varphi e^{-i\alpha} e^{-i\omega x} \end{pmatrix}, \quad (22)$$

where, for convenience we will define

$$\begin{pmatrix} a_+ & = & \cos \lambda \cos \varphi e^{i\alpha} \\ a_- & = & -\sin \lambda \sin \varphi e^{i\beta} \end{pmatrix} \quad (23)$$

$$\begin{pmatrix} b_+ & = & -\cos \lambda \sin \varphi e^{i\alpha} \\ b_- & = & -\sin \lambda \cos \varphi e^{i\beta} \end{pmatrix} \quad (24)$$

$$\begin{pmatrix} c_+ & = & \sin \lambda \cos \varphi e^{-i\beta} \\ c_- & = & \cos \lambda \sin \varphi e^{-i\alpha} \end{pmatrix} \quad (25)$$

$$\begin{pmatrix} d_+ & = & -\sin \lambda \sin \varphi e^{-i\beta} \\ d_- & = & \cos \lambda \cos \varphi e^{-i\alpha} \end{pmatrix} \quad (26)$$

It is straightforward to notice that all the elements of this matrix are one term of a Fourier series. Let us take, for instance, the terms  $a_{\pm}$ . Let us define  $a_{\pm} = r_{\pm} e^{i\delta_{\pm}}$ . Both numbers must be complex and free, i. e.: we have to be able to adjust them independently. Thus, a total amount of 4 degrees of freedom must appear in the expressions of the complex numbers. The phases are adjusted automatically with  $\alpha$  and  $\beta$ . On the other hand, there are two degrees of freedom shared between two modulus. If we want to solve the values of  $\varphi$  and  $\lambda$  we get the systems of equations

$$\begin{aligned} \cos \lambda \cos \varphi &= r_+ \\ \sin \lambda \sin \varphi &= r_- \end{aligned} \quad (27)$$

which es equivalent to

$$\cos \lambda \cos \varphi \mp \sin \lambda \sin \varphi = \cos(\lambda \pm \varphi) = r_+ \mp r_-, \quad (28)$$

which leads to the solution

$$\begin{aligned} \lambda &= \frac{1}{2} (\arccos(r_+ + r_-) + \arccos(r_+ - r_-)) \\ \varphi &= \frac{1}{2} (-\arccos(r_+ + r_-) + \arccos(r_+ - r_-)) \end{aligned} \quad (29)$$

An analogous derivation holds for all the other elements of the matrix. All of them are related with all others, thus it is only possible to generate one arbitrary function. Once the parameters are fixed, and we need to fix them all for an arbitrary function, the whole matrix is fixed.

With this we showed that a matrix of the form of Eq. (21) have got terms in its entries that are equivalent to one term of the Fourier series. However, this is not enough for showing universality for these gates. It is required that  $N$  gates in a row are somehow equivalent to  $M$  terms of the Fourier series, where the relation between  $N$  and  $M$  is yet to be found. For doing so, we will use an induction method.

Let us suppose now that we have applied  $N$  gates, and the entries of the resulting matrix are a Fourier series in the fashion of Eq. (14). We want to show that if we add another gate, the resulting matrix remains a matrix of Fourier series.

We start with the matrix

$$U_N = \begin{pmatrix} \sum_{n=-N}^N A_n e^{in\omega_0 x} & \sum_{n=-N}^N B_n e^{in\omega_0 x} \\ \sum_{n=-N}^N C_n e^{in\omega_0 x} & \sum_{n=-N}^N D_n e^{in\omega_0 x} \end{pmatrix}, \quad (30)$$

where the elements of this matrix are Fourier series. The new step is computing the Left Multiplication  $U(x; \omega, a, b, \varphi, \lambda)U_N$ . In this derivation we will let the values of the frequencies be free in order to recover the Fourier series later. It is equivalent to perform the derivation with a Right Multiplication.

$$U(x; \omega, a, b, \varphi, \lambda)U_N = \begin{pmatrix} \sum_{n=-N}^N (a_+ A_N + b_+ C_N) e^{i(\Omega_n + \omega)x} + (a_- A_N + b_- C_N) e^{i(\Omega_n - \omega)x} & \sum_{n=-N}^N (a_+ B_N + b_+ D_N) e^{i(\Omega_n + \omega)x} + (a_- B_N + b_- D_N) e^{i(\Omega_n - \omega)x} \\ \sum_{n=-N}^N (c_+ A_N + d_+ C_N) e^{i(\Omega_n + \omega)x} + (c_- A_N + d_- C_N) e^{i(\Omega_n - \omega)x} & \sum_{n=-N}^N (c_+ B_N + d_+ D_N) e^{i(\Omega_n + \omega)x} + (c_- B_N + d_- D_N) e^{i(\Omega_n - \omega)x} \end{pmatrix}, \quad (31)$$

and we want this expression to be identified with

$$U(x; \omega, a, b, \varphi, \lambda)U_N = \begin{pmatrix} \sum_{n=-N-1}^{N+1} \tilde{A}_n e^{in\omega_0 x} & \sum_{n=-N-1}^{N+1} \tilde{B}_n e^{in\omega_0 x} \\ \sum_{n=-N-1}^{N+1} \tilde{C}_n e^{in\omega_0 x} & \sum_{n=-N-1}^{N+1} \tilde{D}_n e^{in\omega_0 x} \end{pmatrix} = U_{N+1} \quad (32)$$

In order to recover the Fourier series we need first to set the frequencies of the given series. The parameters  $\Omega_n$  and  $\omega$  must be adjusted following the table

$$\begin{aligned} \Omega_0 &= \frac{1}{2}\omega_0, \Omega_1 = \frac{3}{2}\omega_0 \\ \Omega_2 &= \frac{5}{2}\omega_0 \\ &\vdots \\ \Omega_n &= \frac{2n+1}{2}\omega_0 \\ \omega &= \frac{1}{2}\omega_0 \end{aligned} \quad (33)$$

and thus

$$\begin{aligned} \Omega_0 \pm \omega &= \frac{1 \pm 1}{2}\omega_0 \\ \Omega_1 \pm \omega &= \frac{3 \pm 1}{2}\omega_0 \\ \Omega_2 \pm \omega &= \frac{5 \pm 1}{2}\omega_0 \\ &\vdots \\ \Omega_n \pm \omega &= \frac{2n+1 \pm 1}{2}\omega_0 \end{aligned} \quad (34)$$

Then it is straightforward to re-arrange terms for the four terms of the matrix.

$$\begin{cases} \tilde{A}_0 &= A_0 a_- + C_0 b_- \\ \tilde{A}_{\pm n} &= A_{\pm n} a_- + C_{\pm n} b_- \\ &+ A_{\pm(n-1)} a_+ + C_{\pm(n-1)} b_+ \\ \tilde{A}_{\pm(N+1)} &= A_{\pm N} a_+ + C_{\pm N} b_+ \end{cases} \quad \begin{cases} \tilde{B}_0 &= B_0 a_- + D_0 b_- \\ \tilde{B}_{\pm n} &= B_{\pm n} a_- + D_{\pm n} b_- \\ &+ B_{\pm(n-1)} a_+ + D_{\pm(n-1)} b_+ \\ \tilde{B}_{\pm(N+1)} &= B_{\pm N} a_+ + D_{\pm N} b_+ \end{cases} \quad (35)$$

$$\begin{cases} \tilde{C}_0 &= A_0 c_- + C_0 d_- \\ \tilde{C}_{\pm n} &= A_{\pm n} c_- + C_{\pm n} d_- \\ &+ A_{\pm(n-1)} c_+ + C_{\pm(n-1)} d_+ \\ \tilde{C}_{\pm(N+1)} &= A_{\pm N} c_+ + C_{\pm N} d_+ \end{cases} \quad \begin{cases} \tilde{D}_0 &= B_0 c_- + D_0 d_- \\ \tilde{D}_{\pm n} &= B_{\pm n} c_- + D_{\pm n} d_- \\ &+ B_{\pm(n-1)} c_+ + D_{\pm(n-1)} d_+ \\ \tilde{D}_{\pm(N+1)} &= B_{\pm N} c_+ + D_{\pm N} d_+ \end{cases} \quad (36)$$

Let us focus in one of them only, for instance the  $\tilde{A}$  terms. In the  $\tilde{A}_0$  term there are 8 degrees of freedom available. Only 4 are needed, thus 4 remain free. For every term until the last one, 4 more degrees of freedom are available, and

4 remain, but only 4 are needed. Then there are 4 degrees of freedom free after every coefficient  $\tilde{A}_n$ . For  $\tilde{A}_{N+1}$  there are no new degrees of freedom, but there are 4 that still remain, and they can be used for fixing this new coefficient.

As a summary, a total amount of 5 degrees of freedom are added with every gate. One of these degrees of freedom, the frequency, is used to free another exponential term in the Fourier series. The other four degrees of freedom are used to allow arbitrary complex values of the coefficients in the Fourier series. ■

### III. DEMONSTRATION LINKING WITH THE UNIVERSAL APPROXIMATION THEOREM

The universal approximation theorem demonstrates that any function can be expressed as a sum of sigmoidal functions with adjustable parameters. This can be understood as a generalization of the Fourier series. If we adjust the terms to be equal to the terms of Fourier, which is valid, we recover the previous result.

#### A. Universal Approximation Theorem

**Theorem 3** *Let  $I_n$  denote the  $n$ -dimensional cube  $[0,1]^n$ . The space of continuous functions on  $I_n$  is denoted by  $C(I_n)$ , and we use  $|\cdot|$  to denote the uniform norm of any function in  $C(I_n)$ . Let  $\sigma$  be a discriminatory and sigmoidal function. Given a function  $f \in C(I_n)$  there exists a function*

$$G(\vec{x}) = \sum_j \alpha_j \sigma(\vec{w}_j \cdot \vec{x} + b_j) \quad (37)$$

such that

$$|G(\vec{x}) - f(\vec{x})| < \epsilon \quad \forall \vec{x} \in I_n \quad (38)$$

for  $w_j \in \mathbf{R}^n$  and  $b_j \in \mathbf{R}$

We also need to define the assumptions of this theorem

**Definition 1** *A function  $\sigma$  is sigmoidal if*

$$\sigma(t) \rightarrow \begin{cases} 1 & \text{as } t \rightarrow +\infty \\ 0 & \text{as } t \rightarrow -\infty \end{cases} \quad (39)$$

**Definition 2** *A function  $\sigma$  is discriminatory if for a measure  $\mu \in M(I_n)$ ,*

$$\int_{I_n} \sigma(\vec{w}_j \cdot \vec{x} + b_j) d\mu(x) = 0 \leftrightarrow \mu = 0, \quad (40)$$

where  $M(I_n)$  is the space of signed regular Borel measures on  $I_n$ .

This theorem is an existence theorem, and thus there is no contribution on how many terms from Eq. (37) are needed for reaching a level precision  $\epsilon$ . This theorem can be proved by means of topology and some previous results from functional analysis. The assumptions of the UAT are given for matching the assumptions of the results supporting it.

The extension made in Ref. [2] extends the UAT to any bounded nonconstant and nonlinear function. This generalization is important from a mathematical and practical point of view, but most common functions can be constructed by means of sigmoidal functions. For instance, a sine (cosine) function can be understood as a sum of sigmoidal functions with the proper shape. Two sigmoids can cancel each other properly and the result would be one cycle of the sine (cosine) function. In principle, this can be repeated periodically. As sine (cosine) functions are a sum of sigmoids, then every function can be represented as a sum of sines (cosines).

#### 1. Proof of the UAT

We will follow the steps of the demonstration of this UAT as in Ref. [3].

Let  $S \subset C(I_n)$  be the set of functions

$$S = \{G(x) : G(x) = \sum_j^N \alpha_j \sigma(\vec{w}_j \cdot \vec{x} + b_j)\}, \quad (41)$$

where  $\alpha_j \in \mathbf{C}$ ,  $\vec{w}_j \in \mathbf{R}^n$ ,  $b_j \in \mathbf{R}$  and  $\vec{x} \in I_n$ . We claim that the closure of  $S$ , namely  $\bar{S}$  is the whole space  $C(I_n)$ .

Let us proceed by contradiction, as in Ref. [3]. Let us suppose that  $\bar{S} \subset C(I_n)$ , but  $\bar{S} \neq C(I_n)$ . By the Hahn-Banach theorem there is a linear functional  $L$  acting on  $C(I_n)$  such that

$$L(S) = L(\bar{S}) = 0, \quad L \neq 0. \quad (42)$$

**Theorem 4 : Hahn-Banach** [4, 5]

Set  $\mathbf{K} = \mathbf{R}$  or  $\mathbf{C}$ . Let  $V$  be a  $\mathbf{K}$ -vector space with a seminorm  $p : V \rightarrow \mathbf{R}$ . If  $\varphi : U \rightarrow \mathbf{K}$  is a  $\mathbf{K}$ -linear functional on a  $\mathbf{K}$ -linear subspace  $U \subset V$  such that

$$|\varphi(x)| \leq p(x) \quad \forall x \in U, \quad (43)$$

then there exists a linear extension  $\psi : V \rightarrow \mathbf{K}$  of  $\varphi$  to the whole space  $V$  such that

$$\psi(x) = \varphi(x) \quad \forall x \in U \quad (44)$$

$$|\psi(x)| \leq p(x) \quad \forall x \in V \quad (45)$$

Using the Riesz Representation theorem we may write  $L$  as

$$L(h) = \int_{I_n} h(x) d\mu(x) \quad (46)$$

for  $\mu \in M(I_n)$  and  $\forall h \in C(I_n)$ .

**Theorem 5 : Riesz Representation** [6]

Let  $X$  be a locally compact Hausdorff space. For any positive linear functional  $\psi$  on  $C(X)$ , there exists a unique regular Borel measure  $\mu$  such that

$$\forall f \in C_c(X) : \quad \psi(f) = \int_X f(x) d\mu(x) \quad (47)$$

In particular,  $\sigma$  is valid as the only imaginary quantities are given by the coefficients  $\alpha_k$ , then

$$\int_{I_n} \sigma(\vec{w}_j \cdot \vec{x} + b_j) d\mu(x) = 0 \quad (48)$$

always.

It is also assumed that  $\sigma$  is discriminatory, and thus  $\mu = 0$ , which contradicts the assumptions. Thus  $\bar{S} = C(I_n)$ , the set  $S$  is dense within  $C(I_n)$  and the theorem is demonstrated. We still have to show that sigmoid functions are discriminatory.

**Lemma 1** Continuous sigmoidal functions are discriminatory.

Let us take  $\sigma(\lambda(\vec{w}_j \cdot \vec{x} + b_j) + c)$ . As  $\lambda \rightarrow \infty$  it is equivalent to the function

$$\gamma(x) = \begin{cases} = 1 & \text{for } \vec{w} \cdot \vec{x} + b > 0 \\ = 0 & \text{for } \vec{w} \cdot \vec{x} + b < 0 \\ = \sigma(c) & \text{for } \vec{w} \cdot \vec{x} + b = 0. \end{cases} \quad (49)$$

If we call the Lebesgue Bounded Convergence Theorem we have

$$0 = \int_{I_n} \sigma_\lambda(x) d\mu(x) = \int_{I_n} \gamma(x) d\mu(x) = \sigma(c)\mu(\Pi_{\vec{w},b}) + \mu(H_{\vec{w},b}) \quad (50)$$

where  $\Pi_{\vec{w},b} = \{x : \vec{w} \cdot \vec{x} + b = 0\}$ , and  $H_{\vec{w},b} = \{x : \vec{w} \cdot \vec{x} + b > 0\}$ .

**Theorem 6 : Lebesgue Bounded Convergence [7]**

Let  $\{f_n\}$  be a sequence of complex-valued measurable functions on a measure space  $(S, \Sigma, \mu)$ . Suppose that  $\{f_n\}$  converges pointwise to a function  $f$  and is dominated by some integrable function  $g(x)$  in the sense

$$|f_n(x)| \leq g(x), \quad \int_S |g| d\mu < \infty \quad (51)$$

then

$$\lim_{n \rightarrow \infty} \int_S f_n d\mu = \int_S f d\mu \quad (52)$$

The measure of all half-planes being 0 implies that  $\mu = 0$ . Let us fix  $\vec{w}$ , and for a bounded measurable function  $h$  we define the linear functional

$$F(h) = \int_{I_n} h(\vec{w} \cdot \vec{x}) d\mu(x), \quad (53)$$

which is bounded on  $L^\infty(\mathbf{R})$  since  $\mu$  is a finite signed measure. Let  $h$  be an indicator of the half planes  $h(u) = 1$  if  $u \geq -b$  and  $h(u) = 0$  otherwise, then

$$F(h) = \int_{I_n} h(\vec{w} \cdot \vec{x}) d\mu(x) = \mu(\Pi_{\vec{w}, b}) + \mu(H_{\vec{w}, b}) = 0. \quad (54)$$

By linearity,  $F(h) = 0$  for any simple function, such as sum of indicator functions of intervals [8].

In particular, for the bounded measurable functions  $s(u) = \sin(\vec{w} \cdot \vec{x})$ ,  $c(u) = \cos(\vec{w} \cdot \vec{x})$  we can write

$$F(c + is) = \int_{I_n} \exp i\vec{w} \cdot \vec{x} d\mu(x) = 0. \quad (55)$$

The Fourier Transform of this  $F$  is null, thus  $\mu = 0$ .

## B. UAT for quantum circuits (qUAT)

Following the strategies from the previous section, we can prove a new universal approximation theorem concerning quantum circuits in the fashion of the Fourier series representation.

**Theorem 7** Let  $I_n$  denote the  $n$ -dimensional cube  $[0, 1]^n$ . The space of quantum single-qubits states encoding continuous functions on  $I_n$  is denoted by  $\Psi_C(I_n)$ , and we use  $\mathcal{F}(\psi, \phi)$  to denote the fidelity between two quantum states in  $\Psi_C(I_n)$ . Let  $\sigma$  be a discriminatory and sigmoidal function. Given a quantum state  $\psi(\vec{x}) \in C(I_n)$  there exists a function

$$\psi(\vec{x})_N = \bigotimes_{j=1}^N R_z(\vec{\omega}_j \cdot \vec{x} + \theta_j) R_y(\varphi_j) |\psi_0\rangle \quad (56)$$

such that

$$1 - \varepsilon \leq \mathcal{F}(\psi_N(\vec{x}), \psi(\vec{x})) \leq 1 \quad \forall \vec{x} \in I_n \quad (57)$$

for  $\omega_j \in \mathbf{R}^n$  and  $\theta_j \in \mathbf{R}$

### 1. Proof of the qUAT

Let us take one circuit to be sequence of gates as in Eq. (56). Every matrix in the sequence is defined to be

$$R_z(\vec{\omega}_j \cdot \vec{x} + \theta_j) R_y(\varphi_j) = \begin{pmatrix} \cos(\varphi_j) e^{i(\vec{\omega}_j \cdot \vec{x} + \theta_j)} & -\sin(\varphi_j) e^{i(\vec{\omega}_j \cdot \vec{x} + \theta_j)} \\ \sin(\varphi_j) e^{-i(\vec{\omega}_j \cdot \vec{x} + \theta_j)} & \cos(\varphi_j) e^{-i(\vec{\omega}_j \cdot \vec{x} + \theta_j)} \end{pmatrix} \quad (58)$$

Thus, if we define the action of the gates recursively

$$\begin{aligned}
\psi_N(\vec{x}) &= \begin{pmatrix} A_N(\vec{x}) \\ B_N(\vec{x}) \end{pmatrix} = R_y(\varphi_N) R_z(\vec{\omega}_N \cdot \vec{x} + \theta_N) \psi_N(\vec{x}) = \\
&= \begin{pmatrix} \cos(\varphi_N) e^{i(\vec{\omega}_N \cdot \vec{x} + \theta_N)} & -\sin(\varphi_N) e^{i(\vec{\omega}_N \cdot \vec{x} + \theta_N)} \\ \sin(\varphi_N) e^{-i(\vec{\omega}_N \cdot \vec{x} + \theta_N)} & \cos(\varphi_N) e^{-i(\vec{\omega}_N \cdot \vec{x} + \theta_N)} \end{pmatrix} \begin{pmatrix} A_{N-1}(\vec{x}) \\ B_{N-1}(\vec{x}) \end{pmatrix} = \\
&= \begin{pmatrix} \cos(\varphi_N) e^{i\theta_N} e^{i\vec{\omega}_N \cdot \vec{x}} A_{N-1}(\vec{x}) - \sin(\varphi_N) e^{i\theta_N} e^{i\vec{\omega}_N \cdot \vec{x}} B_{N-1}(\vec{x}) \\ \sin(\varphi_N) e^{-i\theta_N} e^{-i\vec{\omega}_N \cdot \vec{x}} A_{N-1}(\vec{x}) + \cos(\varphi_N) e^{-i\theta_N} e^{-i\vec{\omega}_N \cdot \vec{x}} B_{N-1}(\vec{x}) \end{pmatrix},
\end{aligned} \tag{59}$$

and we only need to define  $\psi_0(\vec{x})$ , which can be arbitrary. For simplicity, we may take it to be  $|0\rangle$ .

Therefore we may see that, at the end of the day, both  $A_N$  and  $B_N$  are functions of the form

$$A_N(\vec{x}) = \sum_{m=0}^{2^{N-1}} c_m(\varphi_1, \dots, \varphi_N) e^{i\delta_m(\theta_1, \dots, \theta_N)} e^{i\vec{v}_m(\omega_1, \dots, \omega_N) \cdot \vec{x}}, \tag{60}$$

where the inner dependencies of  $c_m$  are products of sines and cosines of  $\varphi_j$ , and those of  $\delta_m$  and  $\vec{v}_m$  are linear combinations.

Let us proceed now as in the proof of the UAT in Ref. [3]. Let us take  $S$  as the set of functions of the form  $A_N(\vec{x})$ , and  $C^{\mathbf{C}}(I_n)$  the set of complex-valued functions in  $I_n$ , defined as in Theorem 3. We assume that  $S \subset C^{\mathbf{C}}(I_n)$ , and  $S \neq C^{\mathbf{C}}(I_n)$ . We can apply the Hahn-Banach theorem and state that there exists a linear functional  $L$  acting on  $C^{\mathbf{C}}(I_n)$  such that

$$L(S) = L(\bar{S}) = 0, \quad L \neq 0. \tag{61}$$

Notice that this theorem is applicable as there are no restriction in working only with real numbers.

Using the Riesz representation theorem we see that we can write the functional  $L$  as

$$L(h) = \int_{I_n} h(x) d\mu(x) \tag{62}$$

for  $\mu \in M(I_n)$  non-null and  $\forall h \in C^{\mathbf{C}}(I_n)$ . In particular,

$$L(h) = A_N(\vec{x}) d\mu(\vec{x}) = 0, \tag{63}$$

and thus

$$\int_{I_n} e^{i\vec{v}_m(\omega_1, \dots, \omega_N) \cdot \vec{x}} d\mu(\vec{x}) = 0. \tag{64}$$

This is the usual Fourier transform of  $\mu$ , and we can conclude that if the  $\mathcal{FT}(\mu) = 0$ , then  $\mu = 0$ , and we come into a contradiction with the only assumption we made. Thus, the sums of the form of Eq. (60)  $S$  are dense in  $C^{\mathbf{C}}(I_n)$  and thus

$$|A - A_N| \leq \epsilon. \tag{65}$$

As the final step we can argue that if  $S$  is dense, then  $A_N$  is close to  $A$ , and  $B_N$  is close to  $B$ . Thus the relative fidelity between two states that are almost equal  $\mathcal{F}(\psi_N(\vec{x}), \psi(\vec{x})) = |AA_N + BB_N|^2$  is nearly one. Thus, we can conclude the results from Eq. 57. ■

## 2. Extension

There is a very simple extension that can be made with no difficult. Let us define now the transformations

$$\psi(\vec{x})_N = \bigotimes_{j=1}^N R_z \left( \left( \frac{\vec{\omega}_j + \vec{v}_j}{2} \right) \cdot \vec{x} + \frac{\beta_j + \alpha_j}{2} \right) R_y \left( \left( \frac{\vec{\omega}_j - \vec{v}_j}{2} \right) \cdot \vec{x} + \frac{\beta_j - \alpha_j}{2} \right) |\psi_0\rangle, \tag{66}$$



where every term turns out to be

$$\begin{aligned}
& R_z \left( \left( \frac{\vec{\nu}_j + \vec{\omega}_j}{2} \right) \cdot \vec{x} + \frac{\beta_j + \alpha_j}{2} \right) R_y \left( \left( \frac{\vec{\nu}_j - \vec{\omega}_j}{2} \right) \cdot \vec{x} + \frac{\beta_j - \alpha_j}{2} \right) = \\
& = \frac{1}{2} \begin{pmatrix} e^{i(\vec{\omega}_j \cdot \vec{x} + \alpha_j)} + e^{i(\vec{\nu}_j \cdot \vec{x} + \beta_j)} & -i(e^{i(\vec{\omega}_j \cdot \vec{x} + \alpha_j)} - e^{i(\vec{\nu}_j \cdot \vec{x} + \beta_j)}) \\ -i(e^{-i(\vec{\omega}_j \cdot \vec{x} + \alpha_j)} - e^{-i(\vec{\nu}_j \cdot \vec{x} + \beta_j)}) & e^{-i(\vec{\omega}_j \cdot \vec{x} + \alpha_j)} + e^{-i(\vec{\nu}_j \cdot \vec{x} + \beta_j)} \end{pmatrix}. \tag{67}
\end{aligned}$$

Following the previous proof it can be shown that this set of functions is also dense in  $C^C(I_n)$ , and thus it is a universal approximator of single-qubits states. In addition, as any arbitrary rotation can be decomposed in three rotations around different axis, in particular around the sequence  $Y - Z - Y$ , this composition shows that there exist universal representations in form of gates that do not depend on the exact form of the gate. In any case the data is introduced several times using the form  $\vec{\omega} \cdot \vec{x} + \theta$ , even setting some of the elements of  $\vec{\omega}$  to 0, a universal approximation can be achieved.

- 
- [1] George B Arfken and Hans J Weber. Mathematical methods for physicists, 1999.
  - [2] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4:251–257, 1991.
  - [3] George Cybenko. Approximations by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2:183–192, 1989.
  - [4] Hans Hahn. Über lineare gleichungssysteme in linearen räumen. *Journal für die reine und angewandte Mathematik*, 157:214–229, 1927.
  - [5] Stefan Banach. Sur les fonctionnelles linéaires ii. *Studia Mathematica*, 1:223–239, 1929.
  - [6] Frédéric Riesz. Démonstration nouvelle d’un théorème concernant les opérations fonctionnelles linéaires. In *Annales scientifiques de l’École Normale Supérieure*, volume 31, pages 9–14, 1914.
  - [7] Alan J Weir, Alan J Weir, et al. *General integration and measure*, volume 2. CUP Archive, 1974.
  - [8] RB Ash. Real analysis and probability academic press. *New York*, 1972.