UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Corso di Laurea Triennale in Fisica

Tesi di Laurea

Entanglement generation in quantum neural networks

Generazione di entanglement nelle reti neurali quantistiche

Relatore

Dr. Ilaria Siloi

Correlatore

Prof. Simone Montangero

Laureando

Angelica Di Domenico

1224367

Anno Accademico 2022/2023

Abstract

Current quantum devices have limitations in terms of size and noise, then hybrid algorithms are considered the most effective way to reach a quantum advantage in the near term. Quantum versions of Neural Networks (QNNs) can be implemented as variational quantum circuits, namely hybrid quantum-classical algorithms where unitaries typically depend on a set of parameters, that are optimized to minimize a given cost function. In this perspective, we study the role of entanglement as related to the expressibility of a QNN. Starting from previous works, we consider the entanglement properties of random instances of QNN, where parameters are sampled from a Gaussian distribution. By varying the distribution parameters and circuit connectivity, we study the generation of Haar distributed states in terms of entanglement entropy.

Sommario

Gli attuali dispositivi quantistici hanno limitazioni in termini di dimensioni e rumore, pertanto gli algoritmi ibridi sono considerati il modo più efficace per raggiungere un vantaggio quantistico nel breve termine. Le versioni quantistiche delle reti neurali (QNNs) possono essere implementate come circuiti quantistici variazionali, ossia algoritmi ibridi quantistico-classici in cui gli operatori unitari dipendono tipicamente da un insieme di parametri, che vengono ottimizzati per minimizzare una data funzione di costo. In questa prospettiva, si studia il ruolo dell'entanglement in relazione all'espressività di una QNN. Partendo da lavori precedenti, si considerano le proprietà di entanglement delle istanze casuali di una QNN, in cui i parametri sono campionati da una distribuzione Gaussiana. Variando i parametri della distribuzione e la connettività del circuito, studiamo la generazione di stati distribuiti secondo la distribuzione di probabilità uniforme indotta dalla misura di Haar in termini di entropia dell'entanglement.

Contents

Introduction

In the 80s Feynman revolutionized the field of computing by envisioning quantum computers as devices composed of quantum systems that perform operations according to the rules of quantum mechanics [1]. The idea of a quantum simulator arose from the understanding that determining properties of many-body quantum systems is computationally hard with classical resources. Soon after [2, 3], this led to the development of quantum algorithms that exploit quantum phenomena, such as *superposition* of quantum states and *entanglement*, as resources during the computational process, to reach the solution of a given problem [4, 5].

In 1994, Peter Shor [6] proposed a quantum algorithm that efficiently factors prime numbers. This algorithm has an exponential speed-up over any classical algorithms that have been discovered to date, and, if implemented on a large-scale quantum computer, it could potentially compromise the security of the RSA cryptosystem. Soon after Lev Grover [7] also proposed a quantum algorithm for solving the search problem of a marked item in an unstructured database. In this case, the gain with respect to the classical counterpart is only quadratic.

Currently, a universal, fault-tolerant quantum computer has not yet been developed. However, the pursuit of quantum speed-up continues with Noisy Intermediate-Scale Quantum devices (NISQ). These quantum computers have limited capabilities due to noise, and the number of available qubits ranges from 50 to a few hundred [8, 9]. To overcome these limitations Variational Quantum Algorithms (VQAs) are considered a promising alternative [10]. VQAs are hybrid quantum-classical variational algorithms, where the quantum hardware is used to accelerate specific tasks, similar to how GPUs are used in classical computers. Prototype applications of VQAs already exist and have been executed on NISQ devices or quantum computer emulators [11–19]. The optimization procedure in VQAs can be a bottleneck, and the convergence to an optimal solution can be prevented by barren plateaus, namely the emergence of exponentially vanishing gradients in the variational parameters with the number of qubits, the expressibility of the variational Ansatz, or the non-locality of the cost function [20, 21].

Within the domain of VQAs we focus on quantum neural networks (QNNs), a possible quantum alternative for classical supervised learning algorithms called neural networks (NNs) [22, 23]. QNNs are VQAs where the quantum circuit consists of parametrized quantum circuits used to encode the classical inputs into a quantum state, followed by a layered structure of variational circuits, which are trained in order to solve the desired learning task. The entanglement produced by the quantum circuit has an impact on the QNN performances due to its connection with the onset of barren plateaus (BPs) [24]. This motivates the work developed in this thesis. Here, we study the entanglement properties of QNNs where the initial parameters are sampled from a Gaussian distribution, an initialization strategy argued to be efficient to address the emergence of BPs [25]. Building on the entanglement analysis of uniformly initialized QNNs [26], we examine instead how a different choice of the distribution parameters μ and σ in the Gaussian initialization strategy affects the entanglement entropy production of QNNs.

The following work is divided into three chapters. In Chapter 1, we introduce the basic concepts of digital quantum computation and information. In Chapter 2 we give a general description of VQAs, going into more detail on the particular case of the QNNs and their entanglement properties. In Chapter 3, we show the results of our analysis on the entanglement entropy produced by Gaussian initialized QNNs, and we give an outline for further analysis.

Chapter 1

Digital quantum computing

The theory of quantum computation is based on the idea of using quantum mechanics to perform computations [4]. Digital quantum computing using the quantum circuit model is the quantum version of classical digital computation [4, 5]. In the classical regime, we treat each transistor as the physical implementation of a binary variable, which takes values 0 or 1, neglecting the fact that the underlying technology is analog. The same way, in digital quantum computing, we neglect the complex Hamiltonian evolution that underlies the computation, restricting the problem to the one of a two-level system: the qubit. The qubit is a fundamental concept of quantum computation. We are going to describe qubits as mathematical objects, regardless of their realization as actual physical systems such as ion traps, superconducting circuits, Rydberg atoms, and integrated photonic circuits [27–30]. Because we treat qubits as abstract mathematical objects, we can discuss the general hardware-agnostic theory of quantum computation.

In the following sections, we describe the basic concepts of quantum computation. Starting from the basic building block, the qubit, passing through the principles of quantum mechanics, we define the quantum circuit.

1.1 Fundamentals of quantum computation

The elementary unit of classical information, the bit, is a binary variable that can assume values 0 or 1. In contrast, the qubit is a two-level quantum system described by a two-dimensional complex Hilbert space [5]. Once defined a basis of the 2-dimensional Hilbert space $\{|0\rangle, |1\rangle\}$, usually called *computational basis states*, the general state of a qubit can be written as follows:

$$|\psi\rangle = a_0 |0\rangle + a_1 |1\rangle \qquad a_0, a_1 \in \mathbb{C}. \tag{1.1}$$

In contrast with classical computation where we can examine a bit to determine whether it is 0 or 1, when we measure a qubit we get either the result 0, with a probability $|a_0|^2$, or the result 1 with probability $|a_1|^2$. Since $|a_0|^2$ and $|a_1|^2$ represent probabilities, they must sum to one:

$$|a_0|^2 + |a_1|^2 = 1. \tag{1.2}$$

This dichotomy between the unobservable state of a qubit and the observations we can make lies at the heart of quantum computation and quantum information. In most of our abstract models of the world, there is a direct correspondence between elements of abstraction and the real world, but it is different in quantum mechanics. The property of a qubit to be in a

superposition state is counterintuitive. While a classical bit can be only in either 0 or 1, a qubit can exist in a superposition of the two states $|0\rangle$ and $|1\rangle$ until it is observed [4].

In the following work, to better visualize the qubit state, we focus on its geometric representation. Using the Equation (1.2) we can rewrite the qubit state as:

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle \quad \text{with} \quad 0 \leq \theta \leq \pi \quad 0 \leq \varphi \leq \pi, \tag{1.3}$$

where θ and φ define a point on the surface of a unit three-dimensional sphere called the Bloch sphere. In this framework, the state of a single qubit can be visualized as a vector pointing on the surface of such a sphere.

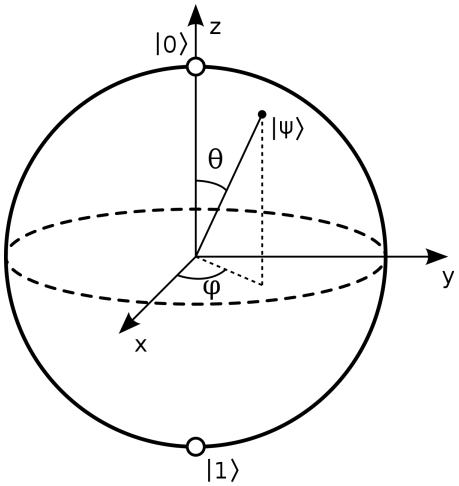

Figure 1.1: Bloch Sphere representation of a qubit state. The angles θ and φ identify the single qubit pure state as shown in Equation (1.3). The $\hat{z}$ unit vector represent the $|0\rangle$ basis state and $-\hat{z}$ the $|1\rangle$ basis state.

The amount of information encoded in a qubit is the one included in two complex numbers a_0 and a_1 defined in Equation (1.1), or likewise in the two angles θ and φ defined in Equation (1.3), so it is potentially infinite. Nevertheless, to access that information we have to measure the quantum state $|\psi\rangle$. The result of the measurement is either $|0\rangle$ or $|1\rangle$: a classical bit. The measurement projects the state of a qubit on $|0\rangle$ or $|1\rangle$, namely on the specific state consistent with the measurement result. Therefore we obtain a single bit of information with a single measurement and only by measuring infinite identically prepared qubits we would be able to determine a_0 and a_1 of the qubit state in Equation (1.1).

If we don't measure the qubit state, we can still exploit the extra information encoded in $|\psi\rangle$ thanks to the quantum phenomenon of *superposition*. The amount of extra information grows exponentially with the number of qubits, thus enabling quantum algorithms to achieve a possible speed up over the classical counterpart. It is the task of quantum algorithms, which are based on quantum logic, to exploit the inherent quantum parallelism of quantum mechanics to achieve the desired output [5].

1.2 Quantum state representation

A state in quantum mechanics is a mathematical object that contains the information of a physical system. If a state contains the maximum amount of information it is called *pure*

state. A pure state is represented by a vector in a Hilbert space over complex numbers with unitary length. A quantum state with partial information on the physical system is defined as a *mixed state*. To describe mixed states we have to introduce a new formalism: *density matrices*. Density matrices are positive semi-definite operators that act on Hilbert spaces. A quantum state, in its more general description, is represented as a density matrix ρ:

$$\rho = \sum_i p_i \left|\psi\right\rangle_i \left\langle\psi\right|_i , \tag{1.4}$$

where p_i is the probability for the system to be in the pure state $\left|\psi\right\rangle_i$. Introducing the trace of an arbitrary matrix A as $Tr(A) = \sum_k^n \left\langle k\right| A \left|k\right\rangle$, where $\{\left|k\right\rangle\}_{k=1,\dots,n}$ is a basis of the Hilbert space, ρ has the following properties:

- ρ is hermitian;
- ρ is a non-negative operator: $\quad \forall \left|\psi\right\rangle \quad \left\langle\psi\right| \hat{A} \left|\psi\right\rangle \geq 0$
- $Tr(\rho) = 1$

We can recognize when ρ represents a pure state because in this case:

$$\rho_p = \left|\psi\right\rangle \left\langle\psi\right| \quad \text{with} \quad p_i = 0 \ \forall i \neq \widetilde{i}, \ p_{\widetilde{i}} = 1, \tag{1.5}$$

such state is a projector: $\rho_p^2 = \rho_p$, and it follows that $Tr(\rho_p^2) = 1$. A mixed state instead presents $Tr(\rho^2) < 1$.

The expectation value of a generic observable $\hat{A}$ relative to a state ρ is written as:

$$\langle\hat{A}\rangle = Tr\Big(\rho\hat{A}\Big). \tag{1.6}$$

We can represent a generic qubit as we did in Equation (1.3) but using the density matrix:

$$\rho(\theta,\varphi) = \left|\psi\right\rangle \left\langle\psi\right| = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2}e^{i\varphi} \end{pmatrix} \cdot \begin{pmatrix} \cos\dfrac{\theta}{2} & \sin\dfrac{\theta}{2}e^{i\varphi} \end{pmatrix} = \begin{pmatrix} \cos^2\frac{\theta}{2} & \sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{-i\varphi} \\ \sin\frac{\theta}{2}\cos\frac{\theta}{2}e^{i\varphi} & \sin^2\frac{\theta}{2} \end{pmatrix}. \tag{1.7}$$

If we consider the change of variables from the spherical coordinates of the Block unitary sphere in the Cartesian ones ($x = \sin\theta\cos\varphi, y = \sin\theta\sin\varphi, z = \cos\theta$), we can express the qubit state on the basis of the Hermitian matrices 2×2 consisting of the identity and Pauli matrices $\{\mathbb{1}, x\hat{\sigma_x}, y\hat{\sigma_y}, z\hat{\sigma_z}\}$:

$$\rho = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix} = \frac{1}{2}(\mathbb{1} + x\hat{\sigma_x} + y\hat{\sigma_y} + z\hat{\sigma_z}). \tag{1.8}$$

1.3 Composite systems and entanglement

Entanglement is the most counter-intuitive phenomenon of quantum mechanics observed in composite quantum systems [5]. In this paragraph, we introduce a formalism for the description of composite systems and we give a general characterization of the *entanglement* and its properties.

1.3.1 Composite systems

A n-body quantum system lives in the tensor product space of the single Hilbert spaces of each component:

$$\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_i = \mathcal{H}_1 \otimes ... \otimes \mathcal{H}_n. \tag{1.9}$$

In the Hilbert space $\mathcal{H}$ the most general state can be written as follows:

$$|\psi\rangle = \sum_{\vec{j}} c_{\vec{j}} \, |j\rangle_1 \otimes ... \otimes |j\rangle_n = \sum_{\vec{j}}^{dim\mathcal{H}} c_{\vec{j}} \, |j_1 ... j_n\rangle \quad with \quad \sum_{\vec{j}} |c_{\vec{j}}| = 1, \tag{1.10}$$

where $\{|j\rangle_i\}_{i=1,...n}$ are the orthonormal basis of each Hilbert space $\{\mathcal{H}_i\}_{i=1,...n}$.
Let us assume that the total system is described by the density matrix ρ:

$$\rho = |\psi\rangle \langle\psi| = \sum_{\vec{j}} \sum_{\vec{k}} c_{\vec{j}} c^*_{\vec{k}} \, |j_1 ... j_n\rangle \langle k_1 ... k_n| \, . \tag{1.11}$$

Taking into consideration a local observable acting only on the subsystem i, we can express it as an operator acting on the entire Hilbert space $\mathcal{H}$:

$$\hat{A} = \mathbb{1}_1 \otimes ... \otimes \hat{A}_i \otimes ... \otimes \mathbb{1}_n. \tag{1.12}$$

To compute the expectation value of such observable in a quantum many-body system we need to define the *reduced density matrix* [4]. The *reduced density matrix* ρ_i is obtained by tracing away all the other degrees of freedom except those relative to subsystem i:

$$\rho_i = Tr_{k \neq i}(\rho). \tag{1.13}$$

From the reduced density matrix the expectation value of the local observable $\hat{A}$, see Equation (1.12), becomes:

$$Tr(\rho \hat{A}) = Tr(\rho_i \hat{A}_i). \tag{1.14}$$

Therefore, it is possible to compute the expectation value of an operator acting only on subsystem i by means of the reduced density matrix ρ_i. We can conclude that ρ_i, obtained after partial tracing over the other subsystems, describes the state of subsystem i [31]. It is important to emphasize that ρ_i does not conserve the purity of the general matrix ρ, namely if ρ is a pure state, there is no guarantee that ρ_i is pure too.

1.3.2 Entanglement

Entanglement is a uniquely quantum mechanical phenomenon that plays a key role in many applications of quantum computation and quantum information [4]. A composite system that has only classical correlations can be described as a separable state, a convex combination of product states. Given a bi-partition of the system in two subsystems A and B, we can define a separable state as:

$$\rho^S_{AB} = \sum_i p_i (\rho^A_i \otimes \rho^B_i) \;\; \text{with} \;\; \sum_i p_i = 1 \tag{1.15}$$

where $\rho^A_i \otimes \rho^B_i$ is a product state. In this case, after the two classical systems A and B have interacted, they are in well-defined individual states. In contrast, after two quantum

systems have interacted, in general, they can no longer be described independently of each other. There are purely quantum correlations between two such systems, independently of their spatial separation [5]. When the state describing a composite system AB cannot be written as separable is said to be *entangled*:

$$\rho \neq \rho_{AB}^{S}. \tag{1.16}$$

An example of a 2-qubits entangled state is the Bell couple $|\phi^+\rangle \langle\phi^+| = \frac{1}{2}(|00\rangle+|11\rangle)(\langle 00|+\langle 11|)$.
To better understand the entanglement phenomenon, let us consider Bell's experiment. We produce an entangled 2-qubit state and then spatially separate the two qubits. If we perform a measurement on one qubit A we can instantly have information on a possible measure result of the state B without actually measuring it. This however does not violate special relativity since the information obtained from system A must travel through a classical channel to reach system B. A possible explanation for the two systems' correlation could be the presence of some additional local variables, called "hidden" because they are not contemplated in quantum mechanics. A theory supported by Einstein, Podolsky, and Rosen that, with the famous paradox, argued that the description of physical reality provided by quantum mechanics was incomplete [32]. The evidence that the quantum theory is complete was achieved through the violation of Bell Inequalities [33]. This experimental evidence brings to the solution of the EPR paradox and to the incompatibility of locality and realism in quantum mechanics.
The Von Newmann *entanglement entropy* quantifies the entanglement in pure quantum states. Taking into consideration a state ρ, we can define the Von Newmann entropy as:

$$S_{vn} = -Tr(\rho \log \rho). \tag{1.17}$$

S_{vn} has three fundamental properties:

- If ρ is a pure state, then $S_{vn}(\rho) = 0$;
- If ρ evolves unitarily in time its entanglement entropy does not change:

$$S_{vn}(U(t)\rho U^{\dagger}(t)) = S_{vn}(\rho); \tag{1.18}$$

- If the Hilbert space $\mathcal{H}$ dimension of the system is N, then:

$$0 \leq S_{vn}(\rho) \leq log(N). \tag{1.19}$$

For example, let us take $\rho = |\psi\rangle \langle\psi|$ as the quantum state of a system composed by n qubits and consider a bi-partition A, B of such system with respectively n_A and $n_B = n - n_A$ number of qubits. The entropy of the subsystem $\rho_A = Tr_B(\rho)$ (1.13), defined as:

$$S_{vn}(\rho_A) = -Tr(\rho_A log \rho_A), \tag{1.20}$$

quantifies the amount of entanglement between subsystem A and its complement B. It is shown that the entanglement entropy of the two subsystems is equal, namely $S_{vn}(\rho_A) = S_{vn}(\rho_B)$.

1.4 Quantum state evolution and projective measurement

The evolution of a qubit in a pure state is described by Schrödinger's equation:

$$i\hbar\frac{\partial\psi}{\partial t} = \hat{H}\psi, \tag{1.21}$$

where $\hat{H}$ is a fixed Hermitian operator known as the Hamiltonian of the closed system. The time evolution postulate requires the system to be closed, i.e. it is not interacting in any way with other systems. We can write the solution of the Equation (1.21) as:

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle\,, \tag{1.22}$$

where we assume that the initial instant is zero $t_0 = 0$, and that $U(t)$ is the state evolution unitary operator [34]. For this reason, we use unitary operators to describe quantum gates in the next paragraph.

The unitary evolution of the density matrix $\rho_0 = |\psi(0)\rangle\langle\psi(0)|$ can be written as:

$$\rho(t) = \sum_i p_i\,|\psi(t)\rangle_i\,\langle\psi(t)|_i = U(t)\rho_0 U^\dagger(t). \tag{1.23}$$

Closed quantum systems evolve according to unitary evolution. However, to extract information from the system we observe it. The interaction needed to observe the system makes it no longer closed, and thus not necessarily subject to unitary evolution [4]. The observation is made by measuring. One kind of measurement is the *projective measurement* based on the Von Neumann projection postulate [35]. We consider a generic quantum state $|\psi(t)\rangle$, we measure an observable $\hat{A}$ at $t = 0$ finding $a \in \sigma(\hat{A})$, where $\sigma(\hat{A})$ is the spectrum of the observable $\hat{A}$. The measurement projects the state $|\psi(0)\rangle$ in an eigenstate $|a\rangle$ of $\hat{A}$, referred to the eigenspace with eigenvalue a. The mathematical object that performs this operation is the *Projector* $P_a^{\hat{A}} = |a\rangle\langle a|$. Immediately after the measurement, we find the system in a state described as:

$$\left|\psi(0^+)\right\rangle = \frac{P_a^{\hat{A}}\,|\psi(0)\rangle}{\sqrt{\langle\psi(0)|\,P_a^{\hat{A}}\,|\psi(0)\rangle}}. \tag{1.24}$$

1.5 Circuit model of quantum computation

A quantum computation can be divided into three main steps:

1. prepare the quantum computer in a well-defined initial state $|\psi_i\rangle$;
2. manipulate the quantum computer state through unitary transformations: $|\psi_f\rangle = U\,|\psi_i\rangle$;
3. performs measurements on a computational basis.

In the following paragraph, we go into detail about the different steps and we focus on the circuit model, one of the fundamental models of quantum computation.

1.5.1 Quantum register of n qubits

A classical computer can be described as a finite register of n bits, likewise, a quantum computer may be thought of as a quantum register of n qubits. The state of an n-qubit quantum computer lives in a 2^n-dimensional Hilbert space, constructed as the tensor product of n 2-dimensional Hilbert spaces, one for each qubit:

$$|\psi\rangle = \sum_{i_{n-1}=0}^{1} ... \sum_{i_0=0}^{1} c_{i_{n-1},...,i_0} |i_{n-1}\rangle \otimes ... \otimes |i_0\rangle . \tag{1.25}$$

Moving from binary to decimal notation, we can represent the same state as :

$$|\psi\rangle = \sum_{i=0}^{2^n-1} c_i |i\rangle \quad with \quad \sum_{i=0}^{2^n-1} |c_i|^2 = 1. \tag{1.26}$$

A set of n classical bits can store only a single integer i. The n-qubit quantum register can be prepared in the corresponding state $|i\rangle$ of the computational basis, but it can also exploit the *superposition* principle. The number of states of the computational basis in a superposition can be as large as 2^n, so it grows exponentially with the number of qubits. When we perform a computation on a classical computer, different inputs require separate runs. In contrast, a quantum computer can perform a computation for exponentially many inputs on a single run [5].

1.5.2 Quantum gates

The time evolution of a quantum register in Equation (1.25), is described by successive applications of unitary operations. In this work, we neglect non-unitary decoherence effects due to the undesired coupling of the quantum computer to the environment. As we emphasized in Section 1.4, the evolution of an n-qubit state is described by a $2^n \times 2^n$ unitary matrix. This unitary matrix can always be decomposed into a product of unitary operations acting only on one or two qubits [5]. In particular, it can be shown that any unitary operation in the Hilbert space of n qubits can be decomposed into the two-qubit CNOT and one-qubit gates. For this reason, this set of gates is said to be *universal for quantum computation* [4, 5]. Therefore, we introduce the fundamental qubit gates that can be part of the universal gate set.

Let us begin with one-qubit gates. Writing the computational basis as vectors in a 2-dimensional Hilbert space:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{1.27}$$

the operations on a qubit are described as 2×2 unitary matrices acting on such Hilbert space. Two quantum gates that have a classical counterpart are the *buffer gate* and the *not gate*. Such gates are respectively defined as:

$$\mathbb{1}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{1.28}$$

The *buffer* acts on the computational basis $\{|0\rangle, |1\rangle\}$ as an identity; while the *not gate* swaps $|0\rangle$ with $|1\rangle$ and vice versa. Two quantum gates that do not have a classical counterpart are:

- the *Hadamard gate*, defined as

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{1.29}$$

it acts on the computational basis as follows

$$H\,|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad H\,|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \tag{1.30}$$

we can visualize its action on the Bloch sphere as a $\frac{\pi}{2}$ rotation of the state vector in the (x, z) plane around the y-axis;

- the *Phase shift gate*, defined as

$$R_z(\delta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{pmatrix}, \tag{1.31}$$

it acts on a generic single qubit state written in Equation (1.3) as follows

$$R_z(\delta)\,|\psi\rangle = \cos\frac{\theta}{2}\,|0\rangle + e^{i(\varphi+\delta)}\sin\frac{\theta}{2}\,|1\rangle \quad with \quad 0 \leq \theta \leq \pi \quad 0 \leq \varphi \leq \pi, \tag{1.32}$$

we can visualize it on the Bloch sphere as a δ rotation of the state around the z-axis.

It is shown that any unitary operation of a single qubit can be constructed using only Hadamard and phase-shift gates [5]. In fact, we can reach a generic state of the quantum Bloch sphere starting from $|0\rangle$ with the following sequence of single-qubit gates:

$$R_z(\frac{\pi}{2} + \varphi)HR_z(\theta)H\,|0\rangle = e^{i\frac{\theta}{2}}(\cos\frac{\theta}{2}\,|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}\,|1\rangle), \tag{1.33}$$

where we can ignore $e^{i\frac{\theta}{2}}$ because is a global phase that has no observable effects [4].

To prepare an entangled state in a generic n-qubit system we need interactions between the qubits, which is why using only single-qubit gates we are unable to generate entanglement. We need to consider a two-qubit gate and the prototypical two-qubit gate able to generate entanglement is the *controlled-NOT gate* [5]. Such gate in matrix representation can be written as a 4×4 matrix:

$$CNOT = \begin{pmatrix} \mathbb{1}_2 & \mathbb{0}_2 \\ \mathbb{0}_2 & \sigma_x \end{pmatrix}, \tag{1.34}$$

in fact, acts on a 4-dimensional Hilbert space $\mathcal{H}$. A computational basis for that Hilbert space is $\{|i_1 i_0\rangle = |00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. A $CNOT$ gate takes as input a generic linear combination of the two-qubit basis states, its action on this latter is:

$$CNOT\,|00\rangle = |00\rangle \;\; CNOT\,|01\rangle = |01\rangle \;\; CNOT\,|10\rangle = |11\rangle \;\; CNOT\,|11\rangle = |10\rangle\,. \tag{1.35}$$

The first qubit of the input state is known as the *control* qubit and the second is the *target* qubit. If the control qubit $|0\rangle$, then the target qubit is unchanged; if the control qubit is $|1\rangle$, then the target qubit is flipped. It is easy to see that a CNOT can generate entanglement. For example, starting from a generic two-qubit separable state:

$$|\psi\rangle_i = (\alpha\,|0\rangle + \beta\,|1\rangle) \otimes |0\rangle \quad \text{with} \quad |\alpha|^2 + |\beta|^2 = 1, \tag{1.36}$$

by applying a CNOT gate we find:

$$CNOT\,|\psi\rangle_i = CNOT(\alpha\,|00\rangle + \beta\,|10\rangle) = \alpha\,|00\rangle + \beta\,|11\rangle\,, \tag{1.37}$$

which is a non-separable state if $\alpha, \beta \neq 0$, this way we create entanglement.

1.5.3 Quantum circuits

The elements described in Subsections 1.5.1 and 1.5.2 are components of what is called a *quantum circuit*. Let us consider the graphic representation of a generic example of a quantum circuit to better describe what it is:

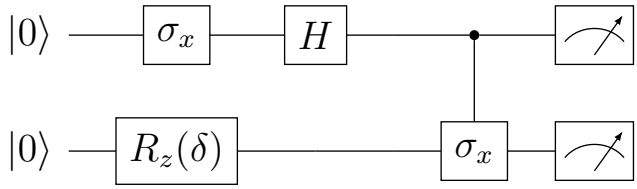

Figure 1.2: Quantum circuit consisting of: a NOT and a Hadamard gate applied to the first qubit; a phase shift gate applied to the second qubit; a CNOT gate using the first qubit as the control qubit and the second as the target; final measurements on both qubits.

Each line in the circuit is associated with a single qubit with time flowing left-to-right [4]. It is conventional to assume that all qubits are initialized in $|0\rangle$. The boxes drawn on a single wire are single qubit gates, and the symbols written inside the box identify the gate. A box drawn on n wires of the circuit represents a gate acting on the n qubits. The CNOT graphical representation is denoted by a point on the control qubit and a NOT gate on the target qubit, sometimes the target qubit is represented with a crossed circle. We choose the first representation to emphasize that we could apply any kind of gate on the qubit target and define a generic control gate. The last symbol we see in Figure 1.2 is the "meter" symbol applied on both qubits. Such a symbol identifies the projective measurement introduced in Equation (1.24).

Chapter 2

Quantum neural networks

A universal fault-tolerant quantum computer that can perform quantum algorithms, like Shor's or Grover's, is not available at the moment. While the experiments advance toward the realization of such devices, the search for quantum speed-up proceeds with the currently available noisy intermediate-scale quantum devices (NISQ). "Intermediate scale" refers to the size of quantum computers with a number of qubits ranging from 50 to a few hundred, while "Noisy" emphasizes that noise places limitations on quantum devices [8, 9]. Variational quantum algorithms (VQAs) are considered a promising route to obtain a quantum advantage in the NISQ era [10]. VQAs are a class of quantum algorithms that use a combination of classical and quantum computation to solve optimization problems. A typical example is finding the ground state of a quantum system, as it is common in quantum chemistry and condensed matter physics [11]. In the machine learning context, quantum neural networks (QNNs) can be used as the underlying architecture, i.e. the quantum circuit, of VQAs. Starting from its classical analog, QNNs are a machine learning model designed to operate and manipulate quantum data, i.e. data encoded into a quantum state. By combining the quantum circuit with classical optimization, the parameters of the QNNs are optimized to minimize a cost function. QNNs seem to represent a promising quantum alternative for classical supervised learning [22, 23].

In the following sections, we introduce variational quantum circuits, then we go into more detail about quantum versions of neural networks (QNNs). Finally, we focus on how to calculate the entanglement entropy produced by QNNs [26] and we introduce Haar distributed random states with their entanglement properties, which are used as a comparison for the entanglement produced in the variational quantum circuit [36, 37].

2.1 Variational quantum algorithms

Variational quantum algorithms are hybrid quantum-classical algorithms. The quantum part is needed to implement the *parametrized quantum circuits (PQCs)*, namely quantum circuits in which some of the unitary operations depend on variational parameters to be optimized. The classical computer instead, deals with the optimization of parameters through the minimization of the so-called *cost function*, whose minimum represents the problem solution. The cost function is optimized using methods based on its gradient or higher-order derivatives, the most common is *gradient descent* [38].

Therefore the two basic elements of VQAs are the cost function and the PQCs *ansatz*. The cost function encodes the problem we want to solve. More specifically, the cost defines a surface

in a multidimensional parameter space, usually called the cost landscape, such that the task of the optimizer is to navigate through the landscape and find the global minimum [10]. A generic cost function can be expressed as:

$$C(\vec{\theta}) = \sum_k f_k(Tr[\hat{O}_k U(\vec{\theta})\rho_k U^\dagger(\vec{\theta})]), \tag{2.1}$$

where $\{f_k\}$ are some set of functions, $U(\vec{\theta})$ is a parametrized unitary, $\vec{\theta}$ is composed of discrete and continuos parameters, $\{\rho_k\}$ are input states from a training set and $\{\hat{O}_k\}$ are a set of observables. $C(\vec{\theta})$ should meet certain criteria to be an exploitable cost function [10]:

- $C(\vec{\theta_{opt}})$ corresponds to the solution of the problem;
- $C(\vec{\theta})$ is efficiently estimated by performing measurements on a quantum computer and possibly performing classical post-processing;
- $C(\vec{\theta})$ is "operationally meaningful", e.g. smaller cost values indicate a better solution;
- $C(\vec{\theta})$ is trainable, e.g. the parameters $\vec{\theta}$ can be efficiently optimized.

An example of an efficient cost function is the expectation value of an observable $\hat{O}$:

$$f(\vec{\theta}) = \langle \hat{O} \rangle = Tr\Big[\hat{O}U(\vec{\theta})\rho U^\dagger(\vec{\theta})\Big], \tag{2.2}$$

where ρ is the initial quantum state generally set to $\rho = |0\rangle^{\otimes n} \langle 0|^{\otimes n}$.

The unitary matrix $U(\vec{\theta})$ describing the PQC can be defined by different gate arrangements. A particular choice of gates is called *ansatz*. Besides the cost function choice, the ansatz selection is an essential constituent of VQAs. If the choice of the ansatz depends on the task to accomplish, such ansatzes are called *problem-inspired ansatzes*, otherwise, if they are independent of the task, they are referred to as *problem agnostic ansatzes* [10]. In figure 2.1 we show a graphic illustration of a hybrid quantum-classical algorithm:

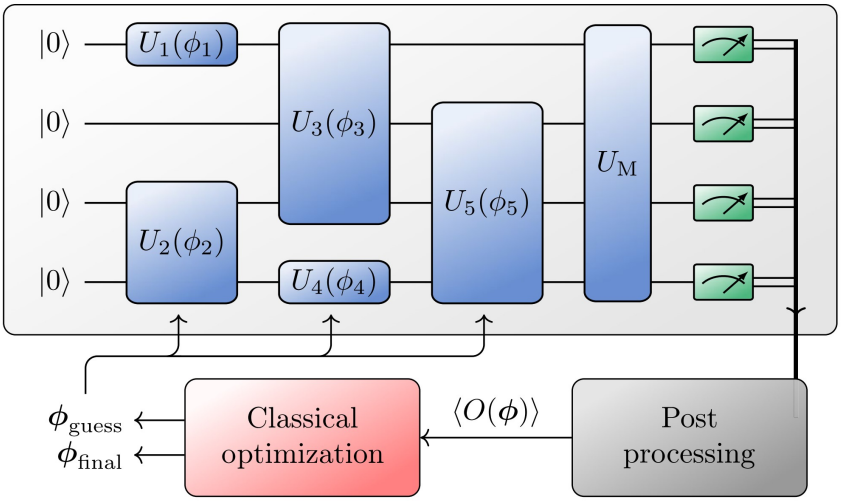

Figure 2.1: A schematic representation of the VQAs hybrid loop consisting of a generic PQC ansatz, from which the cost function $C(\vec{\theta}) = \langle \hat{O}(\psi) \rangle$ is estimated, and the classical optimization process in which the parameters are updated. The process is repeated until convergence is reached. The image is taken from [39].

2.2 Quantum Neural Networks

Quantum neural networks (QNNs) are the quantum version of a subset of classical machine learning called neural networks (NNs) [40]. As the name suggests, NNs are learning algorithms inspired by the functioning of our neurological system. As biological neurons process multiple inputs to generate a single output and then transmit it to other neurons, the fundamental computational unit of a NN follows the same input-output model.

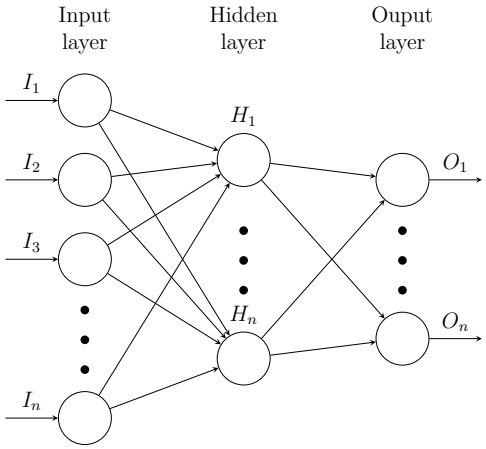

Figure 2.2: Schematic representation of a classical neural network

Figure 2.2 represents a generic NN architecture formed by multiple layers consisting of different computational units, represented by circles. An example of neural networks are those used for *supervised learning*. NNs are "supervised" when both figures and labels are provided to the algorithm to be trained. "Learning" refers to the learning procedure happening during parameter optimization. Therefore a neural network is usually able to generalize and predict well on previously unseen data [40, 41].

Quantum neural networks are a new class of machine learning models, based on classical NN learning techniques, but deployed on quantum computers [22]. A QNN is a variational quantum algorithm made of three main constituents: input data encoding, variational ansatz, and final measurements with a classical update of the parameters [23]. The encoding scheme used to load the input data into the quantum computer is known as *feature map* and consists of a unitary operation, $\mathcal{F}(\vec{x})$, parameterized by the input data $\vec{x}$. There is no fixed form for the feature map, and we thus choose it based on the problem. The other basic element of a QNN is the variational ansatz that we denote as $V(\vec{\theta})$, where $\vec{\theta}$ are the variational parameters. As their classical counterpart, a QNN is composed of several repetitions called *layers*, we define each layer as composed of a feature map $\mathcal{F}(\vec{x})$ and a variational unitary $V(\vec{\theta})$. Therefore a general QNN can be expressed as:

$$U_{QNN}(\vec{x}, \vec{\theta}) = V(\vec{\theta}_L)\mathcal{F}(\vec{x})...V(\vec{\theta}_1)\mathcal{F}(\vec{x}), \tag{2.3}$$

where L is the number of layers and the unitary depends on the input data $\vec{x}$ and the set of trainable parameters $\{\vec{\theta}_i\}$.

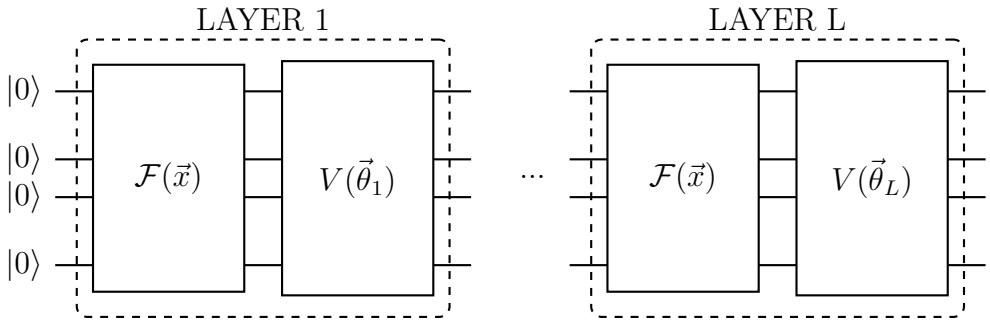

Figure 2.3: Schematic representation of a Quantum Neural Network.

Let us compare Figure 2.3 representing a generic QNN with its classical counterpart in Figure 2.2. We can see that in the latter there is just one input layer, instead in a QNN the input data are uploaded multiple times inside the circuit. Such a procedure is a standard practice called *data re-uploading*, and it is essential for QNN to model higher-order input parameters. [42–44]. After the application of the unitary $U_{QNN}(\vec{x}, \vec{\theta})$, we measure the qubits to infer some relevant information about the system. Given the outcome, the parameters of the circuit are updated through a classical optimizer to minimize the cost function defining the problem [23]. The optimization of a QNN is usually performed by a first-order iterative optimization algorithm called gradient descent [45].

Quantum neural networks are generally affected by a problem known as *barren plateaus* (BPs), the gradient of the cost function vanishes exponentially with the number of qubits hindering the optimization process [20, 21]. The emergence of BP is connected to different phenomena, one of them is the amount of entanglement produced in the variational circuit [24], therefore it is important to keep the entanglement growth under control. Further, different recent works address the BP problem, and, in particular, discuss the relationship between the emergence of BP and the random parameters of the variational quantum circuit. In [25] is shown that using proper Gaussian initialization parameters the norm of the gradient decays at most polynomially with increasing the qubit number and the circuit depth.

2.3 Entanglement entropy production in quantum circuits and Haar-distributed states

Let us assume that we want to uniformly sample, i.e. extract uniformly at random, points in a volume of the three-dimensional Euclidean space. If we consider a unit volume cube, it is sufficient to randomly extract each Cartesian coordinate (x, y, z) of the point in $[0, 1]$. Whereas to perform uniform sampling on the unit volume of a sphere it would be wrong to simply sample each spherical coordinate (r, θ, ϕ) from a uniform distribution over its domain. We must take into account the euclidean measure, which weights differently the distinct portions of the sphere to properly sample points in the sphere volume uniformly at random. For a practical visualization see Figure 2.4.

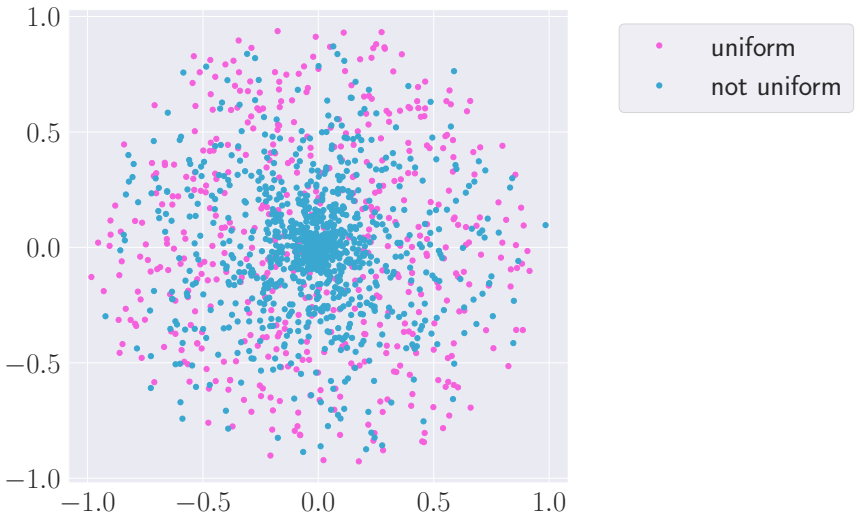

Figure 2.4: Section of a spherical volume showing the difference between sampling points in the spherical volume uniformly at random (uniform) and sampling points from a uniform distribution over the sphere domain (not uniform).

Now, let us turn to the case of uniform sampling from quantum states, which are generated by applying unitaries distributed according to a certain measure. Let us consider the elements U of the group of $n \times n$ unitary matrices $\mathbb{U}(n)$. We can perform operations on the elements of $\mathbb{U}(n)$ such as sampling uniformly over them, just as we can do to points on a sphere. Considering an n-dimensional space, the Haar measure denoted as $\mu(U)$ is the measure that tells us how to weigh the elements of $\mathbb{U}(n)$. To give a more precise definition we can say that the Haar measure is the unique unitarily invariant probability measure $\mu(U)$ defined on the group of $n \times n$ unitary matrices $\mathbb{U}(n)$ [37]. An important property of such a measure is its invariance under unitary transformations:

$$\mu(MU) = \mu(UM) = \mu(U), \tag{2.4}$$

where U and M are unitary matrices that belong to $\mathbb{U}(n)$. $\mathcal{P}(n)$ is the uniform probability induced by the Haar measure $\mu(U)$ on the space of unitary matrices $\mathbb{U}(n)$. We can sample a quantum state according to the Haar measure by generating Haar-random unitaries and applying them to a fixed basis state such as $|0\rangle^{\otimes n}$. We can define such states as *Haar-random distributed states*.

The comparison between the uniform distribution of Haar-random states and the distribution of states obtained from sampling the parameters of a PQC is at the basis of the calculation of *expressibility* [46]. Expressibility is defined as the ability of a PQC to uniformly address the Hilbert space. In the case of the single qubit, this corresponds to a circuit's ability to explore the Bloch sphere. A weak correlation was found between the circuit expressibility and the entanglement generated inside the circuit. Therefore let us consider the entanglement properties of Haar-distributed states. Starting from a Haar-distributed quantum state of n qubits $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$, we create a bi-partition of the n-qubit system into two subsystems A and B, with n_A and $n_B = n - n_A$ qubits respectively. For $n_A \leq n_B$, the expectation value of the

entanglement entropy, calculated as in Equation (1.20), can be written as:

$$\mathbb{E}[S(\psi_A)] = \sum_{j=d_B+1}^{d_A d_B} \frac{1}{j} - \frac{d_A - 1}{2d_B}, \tag{2.5}$$

where $d_A = 2^{n_A}$ and $d_B = 2^{n_B}$ are the dimensions of the two subsystems, and the expectation value is taken over the probability distribution $\mathcal{P}(n)$ [36, 37]. Using this expectation value, it can be shown that Haar-distributed states are generally highly entangled [37]. This is why such states can be used as a comparison for the entanglement entropy production in a circuit, for example in a quantum neural network [26]. We can study the entanglement entropy production of a generic QNN of n qubits by considering every possible linear bi-partition A and B of the system, with respectively n_A and $n_B = n - n_A$ qubits, as shown in Figure 2.5:

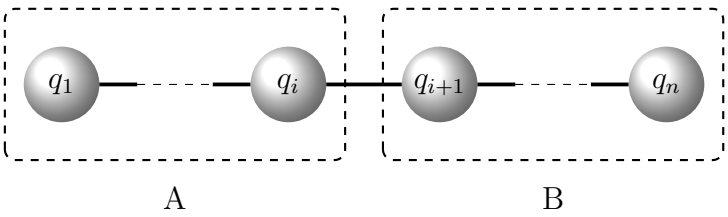

Figure 2.5: Example of a bi-partition of the system. The subsystem A is composed of the qubits q_i with $i \in [1, i]$, while the subsystem B is composed of the qubits q_j with $j \in [i+1, n]$ for $i \in [1, n-1]$

Denoting $\rho_{[1:i]}$ the reduced density matrix of the subsystem A composed of all qubits up to the i-th one, we calculate the entanglement entropy of the considered subsystem division using (1.20) as:

$$S_i = -Tr[\rho_{[1:i]} \log \rho_{[1:i]}], \tag{2.6}$$

where $\rho = U(\vec{x}, \vec{\theta}) \left|0\right\rangle^{\otimes n} \left\langle 0\right|^{\otimes n} U(\vec{x}, \vec{\theta})^\dagger$ is the quantum neural network state obtained by applying the unitary representing the QNN as in (2.3) to the computational basis state $\left|0\right\rangle^{\otimes n} \left\langle 0\right|^{\otimes n}$ [26].

Chapter 3

Results

In this chapter, we study the production of entanglement entropy in a quantum neural network. Our results build on [25, 26] by considering the case where random parameters in the circuit are initialized by sampling from a Gaussian distribution with tunable variance and mean. The considered QNN is composed of several layers, each one consisting of a feature map $\mathcal{F}(\vec{x})$ and a variational unitary $V(\vec{\theta})$ as in Figure 2.3. For both $\mathcal{F}(\vec{x})$ and $V(\vec{\theta})$ composing each layer, we used a particular ansatz introduced in [22] and consisting of single qubit rotations followed by entangling operations as in Figure 3.1:

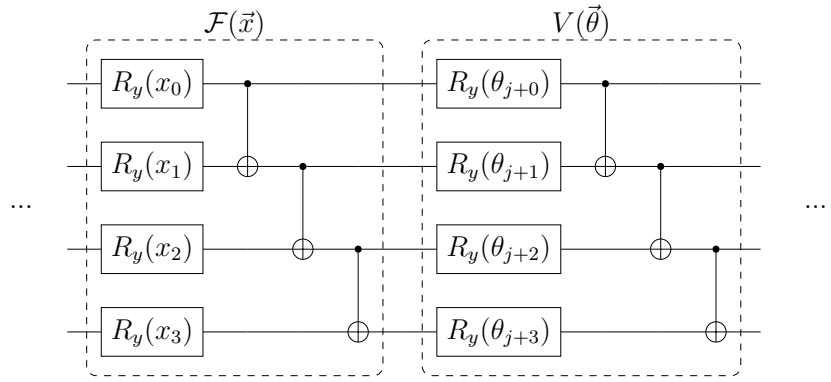

Figure 3.1: Layer $(j+1)$ of the L-layered QNN under consideration, with $j = 0, ..., L-1$. The quantum circuit used as feature map $\mathcal{F}(\vec{x})$ and variational unitary $V(\vec{\theta})$ ansatz is composed of: a series of single qubit rotations around the $\hat{y}$ axis applied to each qubit of the initial state $|0\rangle^{\otimes n}$; a series of CNOT gates applied to each qubit, where the i-th qubit is the control and the $(i+1)$-th qubit is the target. This figure shows an example of the ansatz for a 4-qubit system, but it can be generalized to n qubits.

Taking into consideration $U_L(\vec{x}, \vec{\theta})$, a particular QNN with L layers as in Equation (2.3), we denote the input state as $\vec{x} = (x_1, ..., x_m) \in \mathbb{R}^m$, where m corresponds to the number of qubits, and the variational parameters as $\vec{\theta} = (\theta_1, ..., \theta_p) \in \mathbb{R}^p$, where p is the total number of parameters in the QNN. We sample $\vec{x}$ and $\vec{\theta}$ according to a Gaussian distribution $x_i, \theta_i \sim N(\mu, \sigma)$ and we study the entanglement entropy produced by the QNN by varying the Gaussian parameters μ, σ. It is important to emphasize that the QNN entanglement entropy is calculated as the average among M trials of such quantum neural network and M is set to 100, if not specified otherwise.

3.1 Entanglement entropy production depending on the Gaussian parameters

We compare the entanglement entropy produced in a quantum neural network initialized with Gaussian parameters with that of Haar-random distributed states, in search of the particular value of μ and σ that make the entanglement entropy of the QNN converge to that of Haar-random states.

3.1.1 Gaussian distribution with fixed mean and variable variance

We start by sampling the initial parameters x_i, θ_i from a Gaussian distribution with a fixed $\mu = 0.5\pi$ and variable σ. We compute the entanglement entropy of quantum neural networks $U_L(\vec{x}, \vec{\theta})$ with an increasing number of layers $L = 1, ..., L_{max} = (n-1)$, where n is the number of qubits in the system. In particular, we consider all possible linear bi-partitions A and B of the QNN with respectively n_A and $n_B = n - n_A$ number of qubits as shown in Figure 2.5, and we compute the entanglement entropy as explained in Section 2.3. In this setting, we compare the $U_L(\vec{x}, \vec{\theta})$ entanglement entropy production to that obtained from Haar-random distributed states.

It is shown that sampling initial parameters from a uniform distribution Unif$(0, \pi)$, the entanglement entropy produced by $U_L(\vec{x}, \vec{\theta})$ converges to that of Haar-random distributed states as the number of layers approaches the number of qubits $L \approx n$ [26]. Therefore we vary the Gaussian parameter σ in the search for the same convergence to the entanglement entropy of Haar-random distributed states.

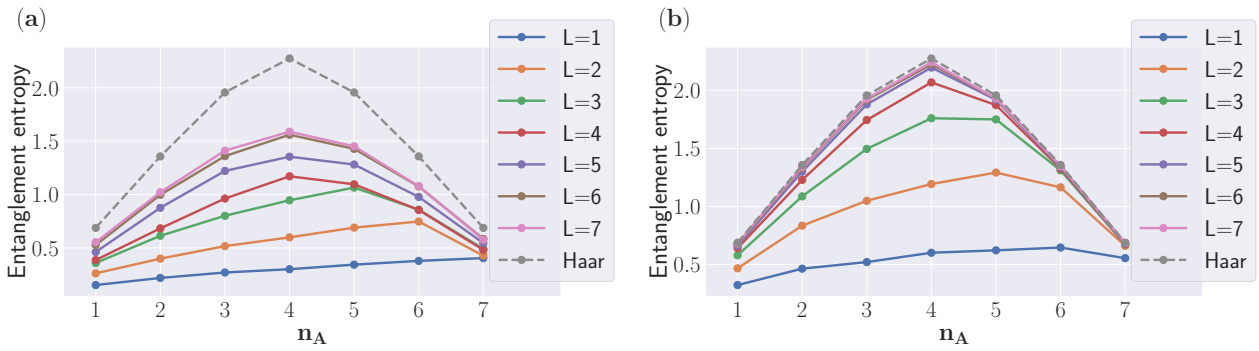

Figure 3.2: Average entanglement entropy depending on the number of qubits n_A in subsystem A for a system composed of $n = 8$ qubits. The curves represent different numbers of layers L in the QNN. (**a**) The initial parameters x_i, θ_i are sampled from a Gaussian distribution $N(\mu = 0.5\pi, \sigma = 0.1\pi)$. (**b**) The initial parameters x_i, θ_i are sampled from a Gaussian distribution $N(\mu = 0.5\pi, \sigma = 0.25\pi)$.

We notice that the findings in Figure 3.2 generally agree with the results in [26]: the entanglement increases with the circuit depth, in particular, when $L = L_{max}$ the entanglement entropy has a maximum in $n_A = \frac{n}{2}$. As panel (**a**) in Figure 3.2 shows, using $\sigma = 0.1\pi$ the entanglement entropy produced by $U_{L_{max}}(\vec{x}, \vec{\theta})$ is still distant from that of Haar-random distributed states. While in panel (**b**), where $\sigma = 0.25\pi$, we can highlight a behavior similar to the one found in [26].

To quantify the comparison between the entanglement entropy produced by a QNN and that of Haar-random states, we compute the normalized difference d_{max} between the maximum values of entanglement entropy produced in the two cases:

$$d_{max} = \frac{(S_{\frac{n}{2}})^{Haar}_{max} - (S_{\frac{n}{2}})^{L_{max}}_{max}}{(S_{\frac{n}{2}})^{Haar}_{max}}, \tag{3.1}$$

where $(S_{\frac{n}{2}})^{Haar}_{max}$ is the maximum entanglement entropy value of Haar random states corresponding to a bi-partition with an equal number of qubits $n_A = \frac{n}{2}$; while $(S_{\frac{n}{2}})^{L_{max}}_{max}$ is the maximum entanglement entropy produced by the maximum-layered QNN, to which corresponds to the same bi-partition.

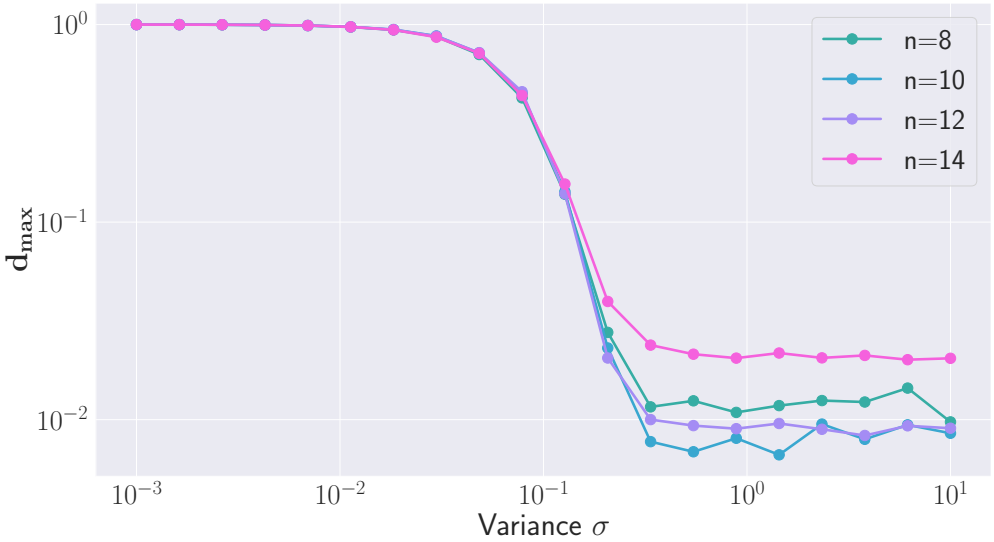

Figure 3.3: Normalized difference between entanglement entropy maximum d_{max} for systems with $n = 8, 10, 12, 14$ qubits. Mean value fixed to $\mu = 0.5\pi$ and $\sigma \in [10^{-3}\pi, 10\pi]$. We highlight that the proper variance value σ must be calculated by multiplying the values reported on the x-axis by π. The x-axis and y-axis are log scaled.

In Figure 3.3 we show d_{max} depending on the Gaussian parameter σ for systems with $n = 8, 10, 12, 14$ qubits. We can outline three different regions for any system size n:

1. for a variance $\sigma \lesssim 10^{-2}$ no entanglement is produced in the system;
2. in the region $10^{-2} \lesssim \sigma \lesssim 1$ there is a growth of entanglement entropy produced in the quantum neural network;
3. for a variance $\sigma \gtrsim 1$ the entanglement entropy stabilizes on a maximum value close to that of Haar-random states.

The results shown in Figure 3.2 are confirmed through this analysis: for a fixed mean value $\mu = 0.5\pi$, the variance value $\sigma \approx 0.25\pi$ seems to be the turning point value after which the entanglement entropy produced by the quantum neural network $U_{L_{max}}(\vec{x}, \vec{\theta})$ is approximately the same of that of Haar-random distributed states.

3.1.2 Gaussian distribution with fixed variance and variable mean

We select one variance value $\sigma = 10^{-3}\pi$ from the first region outlined by Figure 3.3, two values $\sigma = 0.1\pi$, $\sigma = 0.175\pi$ from the central region and one $\sigma = 10\pi$ from the last region. For each one, we compute d_{max} depending on the mean value μ, namely where the Gaussian distribution is centered. We analyze systems with a number of qubits $n = 8, 10, 12, 14$, which present the same trend shown in Figure 3.4.

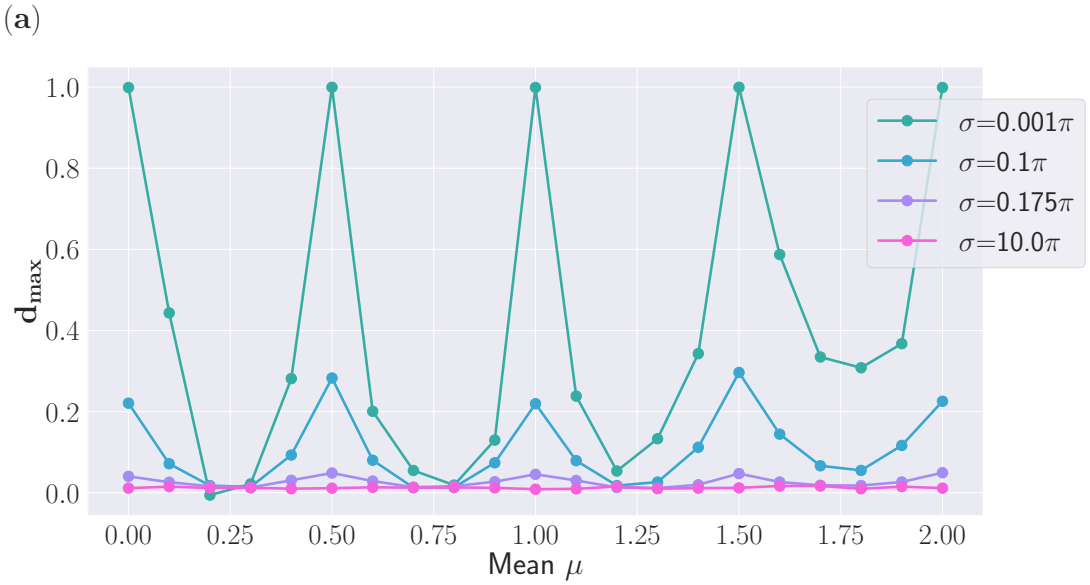

Figure 3.4: Normalized difference between entanglement entropy maximum d_{max} depending on a mean value $\mu \in [0, 2\pi]$. The curves represent different variance values σ. We highlight that the proper mean value μ must be calculated by multiplying the values reported on the x-axis by π. We show plots only for a number of qubits $n = 8$, but we performed simulations for systems with $n = 10, 12, 14$ obtaining similar results.

If the variance is fixed to $\sigma = 10\pi$ the QNN produces an entanglement entropy close to that of Haar-random states regardless of the mean value. We expect such behavior since $\sigma = 10\pi$ lies within the third region outlined in Figure 3.3, namely where the variance σ is large enough that the Gaussian distribution tends to a Uniform one, showing the same behavior of the results in [26].

In Figure 3.4 we notice a periodic pattern: d_{max} oscillates and the amplitude increases as the variance *sigma* decreases. While we show the plots only up to $\sigma = 10^{-3}\pi$, we performed also the simulation for lower σ up to $\sigma = 10^{-8}\pi$, where the same periodic trend was found. Analyzing the case of $\sigma = 10^{-3}\pi$, for $\mu = 0.5\pi$, the entanglement entropy production in the QNN is zero, in agreement with the trend in the low sigma limit shown in Figure 3.3. However, when $\mu = 0.25\pi$ the entanglement entropy produced by the QNN is close to that of Haar random states regardless of the variance value.

By decreasing the value of the variance σ the Gaussian distribution tends to the Dirac delta function. This is equivalent to stating that x_i, θ_i are not random anymore as they all assume the same value μ. By initializing all the initial parameters to a specific value, we find a confirmation of the periodic trend.

To summarize, if we initialize all the parameters x_i, θ_i to the same value μ, the analyzed QNN produces a certain average value of entanglement entropy depending on the chosen value μ. Moreover, there are special initialization values that cause the entanglement entropy produced to be zero or maximum.

3.2 Circuit connectivity and adjacency matrix

Further analyses can be done by investigating how a different choice of ansatz for the feature map $\mathcal{F}(\vec{x})$ or the variational unitary $V(\theta)$ affect the results exposed in Section 3.1. In particular, the choice of ansatz may differ based on circuit *connectivity*, namely how the different qubits of the circuit are connected by multiple qubit gates, such as the CNOT gate. For this reason, we add to the available circuits in the *qucircha library*, used in [26], a new kind of circuit based on a *CNOT gate map*. The circuit CNOT map is represented by a graph, in which the vertices are the qubits and the edges connecting two of them represent the CNOT gates [47]. To mathematically represent the graph we use the adjacency matrix. Specifically, the adjacency matrix of a finite graph G on n vertices is the $n \times n$ matrix where the non-diagonal entry a_{ij} is the number of edges connecting vertex i to vertex j, and the diagonal entry a_{ii} are all zero, since edges from a vertex to itself are not allowed here. In our case, the dimension of the adjacency matrix is given by the number of qubits and the matrix is symmetrical because the graph is undirected [47, 48]. In particular, we implement the possibility of choosing a circuit with linear connectivity, meaning that the qubit i can only be connected to the qubit $i+1$. The circuit added to the *qucircha library* [26] has also $R_y(\theta)$ rotations before and after the CNOT map, with the possibility of disabling the rotations after. For example, if the input adjacency matrix M_{adj} is the following:

$$M_{adj} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{3.2}$$

the correspondent circuit is:

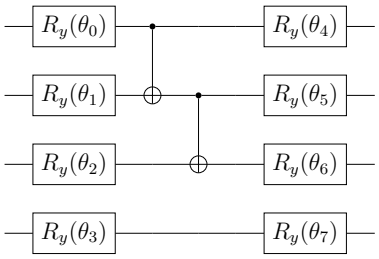

Figure 3.5: The circuit composed of: a series of single qubit rotations around the $\hat{y}$ axis applied to each qubit of the initial state $|0\rangle^{\otimes n}$; a series of CNOT gates based on the input adjacency matrix. This figure is shown an example of the ansatz for a 4-qubit system, but it can be generalized to n-qubits.

Conclusions

In this thesis, we study the entanglement entropy produced by a quantum neural network (QNN) and compare it to the entanglement of Haar-random distributed states. Starting from the analysis of the entanglement generated by QNNs initialized with uniform random parameters [26], we extend the study to the case of Gaussian distributed initial parameters. This is motivated by the effectiveness of Gaussian initialization in preventing the emergence of the barren plateau problem [25]; for this reason, we want to investigate the link between Gaussian initialization and entanglement generation in QNNs, another feature commonly associated with the barren plateau [24].

Once properly defined the QNNs in terms of feature map and variational quantum circuit, we compute the entanglement entropy produced depending on the chosen Gaussian parameters μ and σ. Interestingly, for vanishing variance σ, we find an almost periodic trend in the entanglement entropy depending on the mean value. In particular, there are distinct mean values μ where the entanglement produced is zero and others for which it tends to that of Haar random states. For those mean values that maximize entanglement, the change in sigma has no effect; while for mean values where entanglement is minimal, the entanglement produced depends on the chosen σ value. Specifically, for large sigma values, we recover the same trends we would have by choosing the initial parameters from a uniform random distribution [26].

The entanglement production in various random parametrized QNNs architectures with different feature maps and variational ansatz is discussed in [26], likewise a promising future direction is to investigate also the connection between the entanglement production of Gaussian initialized QNNs to different feature maps and variational ansatzes, for example exploiting the adjacency map circuit developed in this thesis.

Bibliography

[1] Richard P. Feynman. "Simulating physics with computers". en. In: *International Journal of Theoretical Physics* 21.6 (June 1982), pp. 467–488. ISSN: 1572-9575. DOI: 10.1007/BF02650179. URL: https://doi.org/10.1007/BF02650179 (visited on 02/05/2023).

[2] David Deutsch and Roger Penrose. "Quantum theory, the Church–Turing principle and the universal quantum computer". In: *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 400.1818 (Jan. 1997). Publisher: Royal Society, pp. 97–117. DOI: 10.1098/rspa.1985.0070. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1985.0070 (visited on 02/27/2023).

[3] David Deutsch and Richard Jozsa. "Rapid solution of problems by quantum computation". In: *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 439.1907 (Jan. 1997). Publisher: Royal Society, pp. 553–558. DOI: 10.1098/rspa.1992.0167. URL: https://royalsocietypublishing.org/doi/10.1098/rspa.1992.0167 (visited on 02/28/2023).

[4] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. en. 10th anniversary ed. Cambridge ; New York: Cambridge University Press, 2010. ISBN: 978-1-107-00217-3.

[5] Giuliano Benenti, Giulio Casati, and Giuliano Strini. *Principles of Quantum Computation and Information: Volume I: Basic Concepts*. en. WORLD SCIENTIFIC, Apr. 2004. ISBN: 978-981-238-830-8 978-981-279-479-6. DOI: 10.1142/5528. URL: https://www.worldscientific.com/worldscibooks/10.1142/5528 (visited on 02/05/2023).

[6] P.W. Shor. "Algorithms for quantum computation: discrete logarithms and factoring". In: *Proceedings 35th Annual Symposium on Foundations of Computer Science*. Nov. 1994, pp. 124–134. DOI: 10.1109/SFCS.1994.365700.

[7] Lov K. Grover. "A fast quantum mechanical algorithm for database search". en. In: *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC '96*. Philadelphia, Pennsylvania, United States: ACM Press, 1996, pp. 212–219. ISBN: 978-0-89791-785-8. DOI: 10.1145/237814.237866. URL: http://portal.acm.org/citation.cfm?doid=237814.237866 (visited on 01/12/2023).

[8] John Preskill. "Quantum Computing in the NISQ era and beyond". en. In: *Quantum* 2 (Aug. 2018). arXiv:1801.00862 [cond-mat, physics:quant-ph], p. 79. ISSN: 2521-327X. DOI: 10.22331/q-2018-08-06-79. URL: http://arxiv.org/abs/1801.00862 (visited on 02/07/2023).

[9] Kishor Bharti et al. "Noisy intermediate-scale quantum (NISQ) algorithms". In: *Reviews of Modern Physics* 94.1 (Feb. 2022). arXiv:2101.08448 [cond-mat, physics:quant-ph], p. 015004. ISSN: 0034-6861, 1539-0756. DOI: 10.1103/RevModPhys.94.015004. URL: http://arxiv.org/abs/2101.08448 (visited on 02/07/2023).

[10] M. Cerezo et al. "Variational Quantum Algorithms". In: *Nature Reviews Physics* 3.9 (Aug. 2021). arXiv:2012.09265 [quant-ph, stat], pp. 625–644. ISSN: 2522-5820. DOI: 10.1038/s42254-021-00348-9. URL: http://arxiv.org/abs/2012.09265 (visited on 02/07/2023).

[11] Alberto Peruzzo et al. "A variational eigenvalue solver on a photonic quantum processor". en. In: *Nature Communications* 5.1 (July 2014). Number: 1 Publisher: Nature Publishing Group, p. 4213. ISSN: 2041-1723. DOI: 10.1038/ncomms5213. URL: https://www.nature.com/articles/ncomms5213 (visited on 03/05/2023).

[12] Ying Li and Simon C. Benjamin. "Efficient Variational Quantum Simulator Incorporating Active Error Minimization". In: *Physical Review X* 7.2 (June 2017). Publisher: American Physical Society, p. 021050. DOI: 10.1103/PhysRevX.7.021050. URL: https://link.aps.org/doi/10.1103/PhysRevX.7.021050 (visited on 03/05/2023).

[13] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. *A Quantum Approximate Optimization Algorithm*. arXiv:1411.4028 [quant-ph]. Nov. 2014. DOI: 10.48550/arXiv.1411.4028. URL: http://arxiv.org/abs/1411.4028 (visited on 03/05/2023).

[14] Carlos Bravo-Prieto et al. *Variational Quantum Linear Solver*. arXiv:1909.05820 [quant-ph]. June 2020. DOI: 10.48550/arXiv.1909.05820. URL: http://arxiv.org/abs/1909.05820 (visited on 03/05/2023).

[15] Sumeet Khatri et al. "Quantum-assisted quantum compiling". en-GB. In: *Quantum* 3 (May 2019). Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, p. 140. DOI: 10.22331/q-2019-05-13-140. URL: https://quantum-journal.org/papers/q-2019-05-13-140/ (visited on 03/05/2023).

[16] Peter D. Johnson et al. *QVECTOR: an algorithm for device-tailored quantum error correction*. arXiv:1711.02249 [quant-ph]. Nov. 2017. DOI: 10.48550/arXiv.1711.02249. URL: http://arxiv.org/abs/1711.02249 (visited on 03/05/2023).

[17] K. Mitarai et al. "Quantum circuit learning". In: *Physical Review A* 98.3 (Sept. 2018). Publisher: American Physical Society, p. 032309. DOI: 10.1103/PhysRevA.98.032309. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.032309 (visited on 03/05/2023).

[18] Andrew Arrasmith et al. "Variational consistent histories as a hybrid algorithm for quantum foundations". en. In: *Nature Communications* 10.1 (July 2019). Number: 1 Publisher: Nature Publishing Group, p. 3438. ISSN: 2041-1723. DOI: 10.1038/s41467-019-11417-0. URL: https://www.nature.com/articles/s41467-019-11417-0 (visited on 03/05/2023).

[19] Pauline J. Ollitrault, Alexander Miessen, and Ivano Tavernelli. "Molecular Quantum Dynamics: A Quantum Computing Perspective". In: *Accounts of Chemical Research* 54.23 (Dec. 2021). Publisher: American Chemical Society, pp. 4229–4238. ISSN: 0001-4842. DOI: 10.1021/acs.accounts.1c00514. URL: https://doi.org/10.1021/acs.accounts.1c00514 (visited on 03/05/2023).

[20] M. Cerezo et al. "Cost function dependent barren plateaus in shallow parametrized quantum circuits". en. In: *Nature Communications* 12.1 (Mar. 2021). Number: 1 Publisher: Nature Publishing Group, p. 1791. ISSN: 2041-1723. DOI: 10.1038/s41467-021-21728-w. URL: https://www.nature.com/articles/s41467-021-21728-w (visited on 02/08/2023).

[21] Jarrod R. McClean et al. "Barren plateaus in quantum neural network training landscapes". en. In: *Nature Communications* 9.1 (Nov. 2018). Number: 1 Publisher: Nature Publishing Group, p. 4812. ISSN: 2041-1723. DOI: 10.1038/s41467-018-07090-4. URL: https://www.nature.com/articles/s41467-018-07090-4 (visited on 12/11/2022).

[22] Amira Abbas et al. "The power of quantum neural networks". en. In: *Nature Computational Science* 1.6 (June 2021), pp. 403–409. ISSN: 2662-8457. DOI: 10.1038/s43588-021-00084-1. URL: https://www.nature.com/articles/s43588-021-00084-1 (visited on 12/18/2022).

[23] S. Mangini et al. "Quantum computing models for artificial neural networks". en. In: *Europhysics Letters* 134.1 (May 2021). Publisher: EDP Sciences, IOP Publishing and Società Italiana di Fisica, p. 10002. ISSN: 0295-5075. DOI: 10.1209/0295-5075/134/10002. URL: https://dx.doi.org/10.1209/0295-5075/134/10002 (visited on 02/08/2023).

[24] Carlos Ortiz Marrero, Mária Kieferová, and Nathan Wiebe. "Entanglement-Induced Barren Plateaus". en. In: *PRX Quantum* 2.4 (Oct. 2021), p. 040316. ISSN: 2691-3399. DOI: 10.1103/PRXQuantum.2.040316. URL: https://link.aps.org/doi/10.1103/PRXQuantum.2.040316 (visited on 02/08/2023).

[25] Kaining Zhang et al. *Escaping from the Barren Plateau via Gaussian Initializations in Deep Variational Quantum Circuits*. arXiv:2203.09376 [quant-ph]. Dec. 2022. URL: http://arxiv.org/abs/2203.09376 (visited on 12/08/2022).

[26] Marco Ballarin et al. *Entanglement entropy production in Quantum Neural Networks*. arXiv:2206.02474 [quant-ph]. June 2022. URL: http://arxiv.org/abs/2206.02474 (visited on 12/08/2022).

[27] P. K. Ghosh. "Ion traps". English. In: (Dec. 1995). URL: https://www.osti.gov/etdeweb/biblio/252654 (visited on 02/05/2023).

[28] M. H. Devoret, A. Wallraff, and J. M. Martinis. *Superconducting Qubits: A Short Review*. arXiv:cond-mat/0411174. Nov. 2004. URL: http://arxiv.org/abs/cond-mat/0411174 (visited on 02/05/2023).

[29] M. Saffman, T. G. Walker, and K. Mølmer. "Quantum information with Rydberg atoms". In: *Reviews of Modern Physics* 82.3 (Aug. 2010). Publisher: American Physical Society, pp. 2313–2363. DOI: 10.1103/RevModPhys.82.2313. URL: https://link.aps.org/doi/10.1103/RevModPhys.82.2313 (visited on 02/25/2023).

[30] Jianwei Wang et al. "Integrated photonic quantum technologies". en. In: *Nature Photonics* 14.5 (May 2020). Number: 5 Publisher: Nature Publishing Group, pp. 273–284. ISSN: 1749-4893. DOI: 10.1038/s41566-019-0532-1. URL: https://www.nature.com/articles/s41566-019-0532-1 (visited on 02/25/2023).

[31] Giuliano Benenti, Giulio Casati, and Giuliano Strini. *Principles of quantum computation and information*. en. OCLC: ocm60606645. Hackensack, N.J: World Scientific, 2004. ISBN: 978-981-238-830-8 978-981-238-858-2 978-981-256-345-3 978-981-256-528-0.

[32] A. Einstein, B. Podolsky, and N. Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" en. In: *Physical Review* 47.10 (May 1935), pp. 777–780. ISSN: 0031-899X. DOI: 10.1103/PhysRev.47.777. URL: https://link.aps.org/doi/10.1103/PhysRev.47.777 (visited on 01/24/2023).

[33] W. Tittel et al. "Violation of Bell Inequalities by Photons More Than 10 km Apart". en. In: *Physical Review Letters* 81.17 (Oct. 1998), pp. 3563–3566. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.81.3563. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.3563 (visited on 01/24/2023).

[34] Kenichi Konishi and Giampiero Paffuti. *Quantum Mechanics: A New Introduction*. en. Google-Books-ID: 33RdDgAAQBAJ. OUP Oxford, Mar. 2009. ISBN: 978-0-19-156797-1.

[35] John von Neumann. *Mathematical Foundations of Quantum Mechanics: New Edition*. en. Google-Books-ID: B3OYDwAAQBAJ. Princeton University Press, Feb. 2018. ISBN: 978-0-691-17856-1.

[36] Don N. Page. "Average entropy of a subsystem". In: *Physical Review Letters* 71.9 (Aug. 1993). Publisher: American Physical Society, pp. 1291–1294. DOI: 10.1103/PhysRevLett.71.1291. URL: https://link.aps.org/doi/10.1103/PhysRevLett.71.1291 (visited on 02/08/2023).

[37] Patrick Hayden, Debbie W. Leung, and Andreas Winter. "Aspects of Generic Entanglement". en. In: *Communications in Mathematical Physics* 265.1 (July 2006), pp. 95–117. ISSN: 1432-0916. DOI: 10.1007/s00220-006-1535-6. URL: https://doi.org/10.1007/s00220-006-1535-6 (visited on 02/08/2023).

[38] Sebastian Ruder. *An overview of gradient descent optimization algorithms*. arXiv:1609.04747 [cs]. June 2017. URL: http://arxiv.org/abs/1609.04747 (visited on 02/09/2023).

[39] Lennart Bittel and Martin Kliesch. "Training Variational Quantum Algorithms Is NP-Hard". en. In: *Physical Review Letters* 127.12 (Sept. 2021), p. 120502. ISSN: 0031-9007, 1079-7114. DOI: 10.1103/PhysRevLett.127.120502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.120502 (visited on 02/15/2023).

[40] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. en. Google-Books-ID: omivDQAAQBAJ. MIT Press, Nov. 2016. ISBN: 978-0-262-33737-3.

[41] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". en. In: *Nature* 521.7553 (May 2015). Number: 7553 Publisher: Nature Publishing Group, pp. 436–444. ISSN: 1476-4687. DOI: 10.1038/nature14539. URL: https://www.nature.com/articles/nature14539 (visited on 02/07/2023).

[42] Adrián Pérez-Salinas et al. "Data re-uploading for a universal quantum classifier". en-GB. In: *Quantum* 4 (Feb. 2020). Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, p. 226. DOI: 10.22331/q-2020-02-06-226. URL: https://quantum-journal.org/papers/q-2020-02-06-226/ (visited on 02/08/2023).

[43] Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. "Effect of data encoding on the expressive power of variational quantum-machine-learning models". In: *Physical Review A* 103.3 (Mar. 2021). Publisher: American Physical Society, p. 032430. DOI: 10.1103/PhysRevA.103.032430. URL: https://link.aps.org/doi/10.1103/PhysRevA.103.032430 (visited on 02/08/2023).

[44] Francisco Javier Gil Vidal and Dirk Oliver Theis. "Input Redundancy for Parameterized Quantum Circuits". In: *Frontiers in Physics* 8 (2020). ISSN: 2296-424X. URL: https://www.frontiersin.org/articles/10.3389/fphy.2020.00297 (visited on 02/08/2023).

[45] Maria Schuld et al. "Evaluating analytic gradients on quantum hardware". In: *Physical Review A* 99.3 (Mar. 2019). Publisher: American Physical Society, p. 032331. DOI: 10.1103/PhysRevA.99.032331. URL: https://link.aps.org/doi/10.1103/PhysRevA.99.032331 (visited on 02/08/2023).

[46] Sukin Sim, Peter D. Johnson, and Alán Aspuru-Guzik. "Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms". en. In: *Advanced Quantum Technologies* 2.12 (Dec. 2019), p. 1900070. ISSN: 2511-9044, 2511-9044. DOI: 10.1002/qute.201900070. URL: https://onlinelibrary.wiley.com/doi/10.1002/qute.201900070 (visited on 12/20/2022).

[47] John W. Essam and Michael E. Fisher. "Some Basic Definitions in Graph Theory". en. In: *Reviews of Modern Physics* 42.2 (Apr. 1970), pp. 271–288. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.42.271. URL: https://link.aps.org/doi/10.1103/RevModPhys.42.271 (visited on 12/18/2022).

[48] Harmanjit Singh and Richa Sharma. "Role of Adjacency Matrix & Adjacency List in Graph Theory". In: *International Journal of Computers & Technology* 3 (Aug. 2012), pp. 179–183. DOI: 10.24297/ijct.v3i1c.2775.