UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

Guilherme Ilário Correr

Exploring the role of qubit connectivity in expressibility and entanglement of parameterized quantum circuits architectures

São Carlos

2024

Guilherme Ilário Correr

Exploring the role of qubit connectivity in expressibility and entanglement of parameterized quantum circuits architectures

Dissertation presented to the Graduate Program in Physics at the Instituto de Física de São Carlos da Universidade de São Paulo, to obtain the degree of Master in Science.

Concentration area: Theoretical and Experimental Physics

Advisor: Prof. Dr. Diogo de Oliveira Soares Pinto

Corrected version

(Original version available on the Program Unit)

São Carlos

2024

Correr, Guilherme Ilário
Exploring the role of qubit connectivity in expressibility and entanglement of parameterized quantum circuits architectures / Guilherme Ilário Correr; advisor Diogo de Oliveira Soares Pinto - corrected version -- São Carlos 2024.
135 p.

Dissertation (Master's degree - Graduate Program in Theoretical and Experimental Physics) -- Instituto de Física de São Carlos, Universidade de São Paulo - Brasil , 2024.

1. Quantum computation. 2. NISQ. 3. Parameterized quantum circuits. 4. Variational quantum algorithms. 5. Entanglement. I. Pinto, Diogo de Oliveira Soares, advisor. II. Title.

This work is dedicated to those who are
fighting against the oppression of Palestinian people.
To the lives of all Palestinians.

ACKNOWLEDGEMENTS

Agradeço em português. Primeiramente, porque amei e fui amado em português desde que nasci. Em segundo lugar, porque me recuso a agradecer numa língua forçada na academia pelo imperialismo. Certas escolhas (como escrever o restante do texto em inglês) são mais difíceis de serem tomadas, mas nem sempre podemos fazer como desejamos.

Pensei em estruturar este trecho dividindo em grupos, começando pela família, passando pelas amizades e finalizando com as relações profissionais. Mas entendi que não vejo sentido nessas divisões e meus laços são mais complexos que isso. Assim, gostaria de começar agradecendo a todes funcionáries das universidades por onde passei. Sem o tempo de cada ume, esse trabalho nunca seria concluído. Em especial, agradeço à Neusa e ao Ricardo, essenciais nas etapas decisivas desse processo. A todes que dividiram casa comigo neste período em São Carlos. Às pessoas da República Paracetamal (se eu for nomear todes vou ultrapassar 3 páginas), que me abrigaram e me ajudaram a me entender como pessoa, às pessoas da República Fresnética e aos vizinhos, Igor, Caio, Kin, Laura, Seidy e Danilo, que moravam comigo de certa forma e são relações importantíssimas para sempre. Ao Rafa, Rafinha, Rafuxa, por todos nossos momentos compartilhados seja dividindo quarto ou dividindo casa. Não seria quem sou sem nossa relação e não consigo imaginar melhor companhia para tudo que temos feito e passado juntos, te admiro todos os dias. Tudo que faço é por elas, com elas, para elas e não pretendo mudar. À Fernanda (Pipinha), que esteve comigo em todos os momentos, desde os problemas com a conclusão da graduação até o envio deste texto, pelo amor que compartilhamos, por me deixar aprender e ensinar e por me permitir estar na sua vida e lhe amar de tantos jeitos em tantos momentos. Com você aprendi a amar o verão.

Às pessoas que conheci e compartilhei momentos nas universidades que passei por São Carlos. Aos meus ex-orientadores, Marcio e Du, por me introduzirem à ciência, por serem ótimos amigos e por me proporcionarem as condições exatas para que eu pudesse entender como ser eu mesmo como cientista. Ao Grupo de Óxidos Semicondutores do DF - UFSCar. Ao grupinho do bandejão, Juju sem w, Julha, Clara, Momo, Hot Wheels e Dedé, por todos os pequenos e grandes momentos que se alimentar em conjunto nos proporcionou. Dedé que me acompanha desde a UFSCar, junto com Jão, com quem pude continuar compartilhando momentos e sentimentos durante tantos anos. Ao Guilherme, aluno do Diogo da sala 21 que fez Mec. Estat. e Grupos no primeiro semestre de 2022, por todas as conversas, suportes e empatia com experiências. Aos membros com quem convivi do Grupo de Informação Quântica, em especial Ivan, Elisa, Pedro, Alexandre, Gui, Gustavo, Clara (de certa forma) e Naruna pelas incríveis contribuições para meu amadurecimento e pelas palhaçadas durante os cafezinhos diários (e durante todo o restante

do tempo) e ao meu orientador, Diogo, por confiar em mim quando eu mesmo duvidei e por me fornecer o espaço necessário para me desenvolver dentro da física teórica. Aprecio todas as amizades no nosso grupo e aquelas que conseguimos estender pelo contato com o pessoal do grupo do Prof. Sérgio.

Agradeço aos meus pais, Vanuza e Carlinhos, por todo o suporte, apoio, por acreditarem em mim e por me incentivarem mesmo não entendendo muito bem por que escolhi uma profissão com tantos ônus sem muitos bônus à vista. Quero estar cada dia mais próximo e conhecer vocês cada dia mais. Ao meu irmão, Gadiota, pela companhia, pelo companheirismo e pelos bons momentos que compartilhamos e vamos compartilhar. Aos meus avós, Florzinha, Ilário, Eneida, Ademir e Mercedes, por serem ponto de carinho e de apoio em tantos momentos. Para finalizar, cito a música de Kiko Dinucci: "*Foi batendo o pé na terra que vovó me ensinou a sambar. Foi batendo o pé na terra que vovó me ensinou a sambar. Que vovó me ensinou a sambar, que vovó me ensinou a sambar, que vovó me ensinou a sambar.*"

This study was financed by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

"Modernizar o passado é uma evolução musical.
Cadê as notas que estavam aqui?
Não preciso delas, basta deixar tudo soando bem aos ouvidos.
O medo dá origem ao mal,
O homem coletivo sente a necessidade de lutar,
O orgulho, a arrogância, a glória
Enchem a imaginação de domínio.
São demônios os que destroem o poder bravio da humanidade.
Viva Zapata!
Viva Sandino!
Viva Zumbi!
Antônio Conselheiro,
Todos os panteras negras
Lampião, sua imagem e semelhança
Eu tenho certeza, eles também cantaram um dia."
Chico Science e Nação Zumbi

ABSTRACT

CORRER, G. I. **Exploring the role of qubit connectivity for expressibility and entanglement in parameterized quantum circuits architectures**. 2024. 135p. Dissertation (Master in Science) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2024.

Variational quantum algorithms are one of the promising methods to obtain quantum advantage in the noisy intermediate scale quantum computers era. They rely on a classical optimization procedure, a cost function and a parameterized quantum circuit to build the solution of a particular problem. Most of the work regarding the quantum circuits part is based on heuristic propositions for the circuit structure and reside only within the borders of VQA applications. In this context, the main objective of our work was the characterization of entanglement generation and distribution of generated states for different PQCs structures. Applying the mean entanglement considering the Scott entanglement measures and the expressibility quantifier, we studied the behavior of 5 possible connectivities between qubits that appear in the contemporary quantum computers: No connections, linear, ring, star and all-to-all, for different number of qubits and circuit concatenations (layers). For two circuit architectures with different local parameterizations, we discussed how entanglement and expressibility are connected, showing that the entanglement generation for only 1 layer is influential for the expressibility evolution as a function of the number of layers. Circuits generating mean and standard deviation for entanglement closer to the uniformly distributed states at 1 layer will have a steeper evolution of expressibility. This result is affected by the local parameterization and number of qubits. We then compared the circuits generated entanglement with the entanglement of paradigmatic states EPR_n, GHZ_n and W_n to understand the entanglement characteristics of the different connections. The results showed how the different connectivities will influence parameterized quantum circuits for applications in VQAs and also that these can present the behavior of pseudorandom quantum circuits.

Keywords: Quantum computation. NISQ. Parameterized quantum circuits. Variational quantum algorithms. Entanglement.

RESUMO

CORRER, G. I. **Explorando o papel da conectividade entre qubits para a expressibilidade e emaranhamento em arquiteturas de circuitos quânticos parametrizados**. 2024. 135p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2024.

Algoritmos Quânticos Variacionais são um dos métodos promissores para obter vantagem quântica na era dos computadores quânticos ruidosos de escala intermediária. Eles dependem de um procedimento de otimização clássico, uma função de custo e um circuito quântico parametrizado para construir a solução para um problema em particular. A maior parte dos trabalhos acerca da parcela de circuitos quânticos se baseia em proposições heurísticas para a estrutura do circuito e reside dentro das fronteiras das aplicações em VQA. Neste contexto, o principal objetivo de nosso trabalho foi a caracterização da geração de emaranhamento e distribuição dos estados gerados para diferentes estruturas de PQCs. Aplicando o emaranhamento médio considerando as medidas de emaranhamento de Scott e o quantificador de expressibilidade, nós estudamos o comportamento de 5 possíveis conectividades entre qubits que aparecem em computadores quânticos contemporâneos: sem conexões, linear, anel, estrela e todos com todos, para diferentes números de qubits e concatenações de circuitos (camadas). Para duas arquiteturas de circuitos com diferentes parametrizações locais, nós discutimos como emaranhamento e expressibilidade estão conectados, mostrando que a geração de emaranhamento em apenas 1 camada é influente para a evolução da expressibilidade em função do número de camadas. Circuitos gerando média e desvio padrão de emaranhamento mais próximos dos estados uniformemente distribuídos em 1 camada vão possuir uma evolução da expressibilidade mais íngreme. Esse resultado é afetado pela parametrização local e pelo número de qubits. Nós comparamos o emaranhamento gerado pelos circuitos com o emaranhamento de estados paradigmáticos EPR_n, GHZ_n e W_n para entender as características do emaranhamento das diferentes conexões. Os resultados mostraram como as diferentes conectividades vão influenciar os circuitos quânticos parametrizados para aplicações em VQAs e também que esses podem apresentar o comportamento de circuitos quânticos pseudoaleatórios.

Palavras-chave: Computação quântica. NISQ. Circuitos quânticos parametrizados. Algoritmos quânticos variacionais. Emaranhamento.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

An	Ansatz n
CUE	Circular Unitary Ensemble
ILO	Invertible Local Operator
LOCC	Local Operation and Classical Communication
LU	Local Unitary
MW	Meyer-Wallach
NISQ	Noise Intermediate Scale Quantum
PQC	Parameterized Quantum Circuit
SLOCC	Stochastic Local Operation and Classical Communication
Sm	Scott measure of order m
VQA	Variational Quantum Algorithm

CONTENTS

1 INTRODUCTION

In a documentary about his work, "Sun Ra: A Joyful Noise"[1], Sun Ra said

> "Those of the reality have lost their way, now they must listen to what myth has to say. Those of the reality have been bruised and beaten by the truth. Those of the reality have been slaves of a bad truth, so there is nothing left now but the myth. The myth is neither bad nor good, its potentials are unlimited.
>
> We hold this myth to be potential. They hold their truth to be self-evident. But our myth is not self-evident, because it is a mystery."

In that context, he was talking about myth and mythology as powerful tools to think about possible realities for people who had their past and present stolen by a cruel reality. There, the human created myth reflects and paves a way through reality, reasoned and built without falling into the self-evidence domain. Science and Physics, somehow, have a parallel approach to the comprehension of reality: Applying the human developed collective reasoning, we can understand the natural reality and contemplate different future directions for our society. Here, as in the Sun Ra myth, self-evidence is not enough and mysterious spooky phenomena are important factors of a possible future.

One of those mysterious and insidious phenomena was first recognized in the famous Einstein-Podolsky-Rosen (EPR) paper (1). The effect, at that time, was criticized and said to be a demonstration of the failure of Quantum Theory in being a complete and physical theory. Nowadays, after almost a century of discussions, experiments and scrutiny (2–5), these non-local and non-classical correlations have a very important role in many of the earliest and latest technological and theoretical developments of the Second Quantum Revolution (6). Perhaps the most famous example of the power of computations in the Second Quantum Revolution is the Shor's algorithm (7), which tackles the problem of number factorization and would bring a new paradigm to cryptography. Shor showed an exponential speedup in comparison to classical methods and applied entanglement as an essential ingredient of the computation. Other proposed algorithms solve problems in a broad range of applications, such as the search problem (8) or the solution of linear systems (9), always with rigorous speedup proofs.

Still, the implementation of such algorithms requires a high precision control with low error probabilities of the quantum operations and components of large scale quantum computers. The requirements to achieve such tasks are not available on contemporary

[1] Available on YouTube here.

quantum computers, both on the control and on the number of qubits sides (10). Therefore, to harness the current and short term capabilities of quantum technologies, a different approach shall be taken. This is the nature of many recent proposals that explore the best capabilities available in the Noisy Intermediate Quantum (NISQ) Computers, which have qubits numbers between fifty and a few hundred and can only implement a limited number of operations without destroying the computation by error accumulation (11).

Variational Quantum Algorithms (VQAs) are in the spotlight as a possible approach towards leveraging the best capabilities of available NISQ computers (12). The algorithms apply parameterized quantum circuits (PQCs) and classical optimization methods, which are connected by a problem encoded cost function depending on the parameters of the circuit. These parameters are iteratively classically optimized to accomplish specific tasks. The main goal of this application is to minimize circuit depth and number of necessary qubits in a computation by the use of a hybrid quantum-classical algorithm, providing potential advantages over purely classical methods. This structure of the method is very generic and can solve problems of interest in quantum mechanics, such as ground state search (13) and thermal state preparation (14), or in general contexts, as search algorithms (15) and machine learning classifiers (16).

Recently, literature is giving a lot of attention to discussions about the powers and drawbacks of the ingredients of a VQA: The optimization process (17,18), the cost function (19) and the parameterized quantum circuit (20–23). An important question is whether a specific PQC architecture is reliable for a particular application and if it is reasonable from the real world hardware perspective. In this work, we are going to discuss the properties of possible parameterized quantum circuits and understand how the characteristics of entanglement and states generation is influenced by the possible connectivity between qubits, inspired by contemporary quantum computers. This discussion is important for VQA applications, however, is not restricted to this context. The study of characteristics of PQCs is more general and provides insights about which are the possibilities depending on their structures.

This thesis studies the values of expressibility and entanglement quantifiers of parameterized quantum circuits, building a discussion about their correlations and the role of entangling gates and parameterized gates. The work is structured as follows: Chapter 2 starts with a revision about the properties and principal concepts of VQA, motivating the study of PQCs, which are the central object of the analysis. Chapter 3 will introduce the Haar measure and the concept of $t-$designs, that are going to be important when defining the expressibility quantifier. Chapter 4 presents a review about quantum entanglement and entanglement quantification. In Chapter 5 we define the circuit descriptors, expressibility and mean entanglement, and discuss the methods applied to estimation. Finally, Chapter 6 presents the results and Chapter 7 discuss the prospects and conclusions of

this work.

2 VARIATIONAL QUANTUM ALGORITHMS AND PARAMETERIZED QUANTUM CIRCUITS

This chapter defines the essential ingredients of Variational Quantum Algorithms and the connections with Parameterized Quantum Circuits. The chapter can be seen as a motivation for the work treating the properties of PQCs, even though the application of the results in this work is not restricted to this context. VQAs are very important examples of implementations where PQCs are essential and have been in the spotlight for almost 10 years now (12, 13, 24–26), so the choice of starting this thesis with their definition is justified. We are going to give a brief overview and focus on the principal element of this work: the Parameterized Quantum Circuit.

2.1 General structure of VQAs

Variational Quantum Algorithms are hybrid quantum-classical methods that leverage the best available properties of classical optimization methods and quantum computers (12). The most advantageous part of these kinds of algorithms is that they are generic and provide a possible way to solve a broad range of different problems. The general structure consists of

1. The quantum part: A low depth[1] parameterized quantum circuit, $U(\vec{\theta})$. U is a quantum circuit depending on the vector of parameters $\vec{\theta}$. This PQC is usually called the circuit *ansatz*;

2. The classical part: A classical optimization method that will act on the circuit parameters;

3. The connection between them: A cost function depending on the circuit parameters that encodes a particular problem, $C(\vec{\theta})$. When this function is optimized, the set of parameters $\vec{\theta}^* = \arg\operatorname*{opt}_{\vec{\theta}} C(\vec{\theta})$ solves the problem. Usually, the optimization is done by minimization.

We are going to elaborate on the principal elements after we talk about how the algorithm is executed. Fig. 1 presents a pictorial diagram of the execution of a VQA. The procedure starts with

1. Preparation of the input trial state, ρ_0. This step is done before the optimization loop, so the choice of the input state is independent of the optimization process. Still, it can be done in a way that the final objective is achieved with less iterations.

[1] The depth of a quantum circuit is the longest path of gates executed in the circuit.

2. A first execution of the PQC is done using the first input state

$$\rho_0 \rightarrow U(\vec{\theta_0})\rho_0 U^\dagger(\vec{\theta_0}) \equiv \rho(\vec{\theta_0}). \tag{2.1}$$

We chose $\rho(\vec{\theta})$ to represent the output state after the execution of the circuit $U(\vec{\theta})$. The optimization has not started yet, so the choice of the first parameter vector $\vec{\theta_0}$ for the first execution of the PQC is arbitrary in the sense that there is a lot of freedom of choice, but can be tuned to avoid optimization caveats (12).

3. First calculation of the cost function values using the output state of the circuit. The cost function can be broadly defined as

$$C(\vec{\theta}) := f\left(\{\rho_k\}, \{\mathcal{O}_k\}, U(\vec{\theta})\right), \tag{2.2}$$

where $\{\rho_k\}$ is a set of states, $\{\mathcal{O}_k\}$ is a set of observables and $U(\vec{\theta})$ is the PQC with a set of parameters $\vec{\theta}$. This function can be, e.g., $C(\vec{\theta}) = \mathrm{Tr}\left[\rho(\vec{\theta})H\right]$, where H is the hamiltonian of a system we want to obtain the ground state. The computation of this cost function is performed by making measurements on the output states of the circuit. In this example, a hamiltonian of spin$-1/2$ particles could be decomposed in terms of Pauli strings that will be measured for the computation of the mean values (27). For this step, $\vec{\theta} = \vec{\theta_0}$ and $\{\rho_k\} = \rho_0$.

4. First update of the parameters, taking the initial vector $\vec{\theta_0} \overset{\text{opt}}{\rightarrow} \vec{\theta_i}$.

5. Optimization loop:

   a) Execution of the circuit with parameters $\vec{\theta_i}$ on the trial state, $\rho_0 \rightarrow \rho(\vec{\theta_i})$;

   b) Cost function calculation $C(\vec{\theta_i}) = f\left(\rho_0, \{\mathcal{O}_k\}, U(\vec{\theta_i})\right)$;

   c) Check of a stop criteria, to verify if the value achieved is optimal:

      i. If it is, stop and export $\vec{\theta}^*$;

      ii. If it is not, optimization of parameters using the cost function $\vec{\theta_i} \overset{\text{opt}}{\rightarrow} \vec{\theta}_{i+1}$. Restart the loop by setting $i + 1 \rightarrow i$.

The performance of a VQA is directly connected to these three main elements and it will be better or worse depending on the choice made for the combination. The most usual optimization methods are based on gradients calculation, applying computations by finite differences and analytical expressions for parameter shifts (12, 28). However, there are other proposals including natural gradients, relying on the structure of the state or parameter spaces (17, 18). Either way, the major problems about variational quantum algorithms usually fall over the proper choice of cost function and circuit ansatz. Over the cost function, because it should satisfy the requirements of encoding the solution of a particular problem when it is optimal, being classically hard to compute (so the quantum

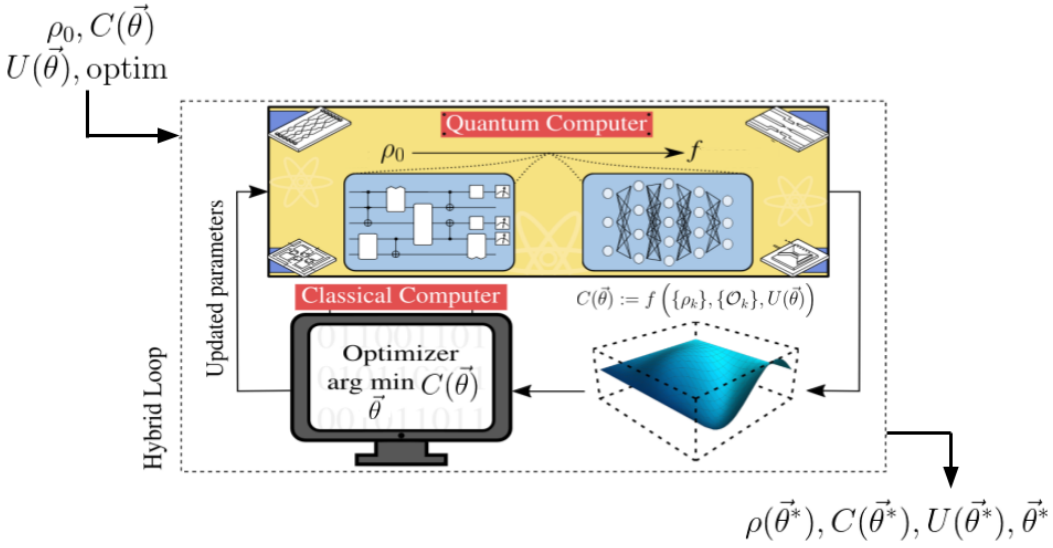

Figure 1 – Schematic representation of the execution of a Variational Quantum Algorithm. A cost function $C(\vec{\theta})$, a trial state ρ_0, an ansatz $U(\vec{\theta})$ and a optimizer are given as inputs. The iterative execution of the hybrid loop is done until a stop criteria is achieved, giving as output the optimal state, $\rho(\vec{\theta}^*)$, optimal value of the cost function, $C(\vec{\theta}^*)$, optimal circuit, $U(\vec{\theta}^*)$, or the optimal parameters, $\vec{\theta}^*$.

Source: Adapted from CEREZO *et al.* (12).

solution is an advantage) and trainable (29). Over the circuit, because it should be capable of constructing the optimal state that gives the optimal value of the cost function, while it has low depth to be implemented in NISQ devices and does not generate difficulties to trainability of the cost function (12, 26).

The versatility of VQAs is mainly connected to the encoding of the problem in the cost function. If the problem can be encoded in a cost function satisfying the above criteria there is at least an expectation that the problem can be solved using the hybrid algorithm. To exemplify, let us talk about two possibilities for the cost function. A first example, in the scope of quantum problems, is the following entanglement witness

$$C(\vec{\theta}) := \mathrm{Tr}\left[(H - E_{\mathrm{sep}}\mathbb{I})\rho(\vec{\theta})\right], \tag{2.3}$$

being H a hamiltonian of a quantum system, E_{sep} the lowest energy eigenvalue of a separable eigenstate of the hamiltonian and $\mathbb{I}$ the identity operator acting on the space of states. This cost function satisfies the criteria for a entanglement witness, being smaller than zero if the state $\rho(\vec{\theta})$ is entangled and greater or equal if the state is separable. This way, when $C(\vec{\theta})$ is minimized below 0, we can say that there are entangled states with energies below the separable eigenstate of lowest energy (30), and the optimal set $\vec{\theta}^*$ allows the circuit to build such state. The task of detecting entangled eigenstates below a particular energy is useful in the context of hamiltonians that present an entanglement gap, i.e., when there is an energy gap between the ground state and the first excited state

that is separable. The separable state energy E_{sep} can be estimated applying a variational quantum eigensolver to obtain the minimum energy eigenstate of the hamiltonian with a circuit comprised only of local gates and an input separable state. This VQA is going to construct only separable states, therefore the minimum energy obtained has to be of a separable eigenstate of the hamiltonian. If this witness is optimized and show entangled states below the first separable eigenstate, it is possible to say that if the energy of the system is below this value, the system is entangled (31).

Another important example is the application of VQAs to data classification. In this case, the cost function considers a set of labelled data, being $\{\rho_k\}$ the set of labelled data and $\{y_k\}$ the labels. This set of states can be inherently quantum, in the case of classification of quantum states, or can encode a classical dataset. In this last case, a encoding procedure should be considered (12). The objective of this method is to find the label of an unlabelled state. Considering the initially unlabelled state as the input of the circuit, the idea is to know the optimal parameters such that the mean value of an observable considering the output state is exactly the specific label of the input state. This way, the cost function is defined as the error when comparing the real label of the labelled state, y_k, with the label provided by the circuit as $\text{Tr}\left[U(\vec{\theta})\rho_k U^\dagger(\vec{\theta})\mathcal{O}\right]$. Therefore, the cost function is

$$C(\vec{\theta}) := \sum_k \left\{ y_k - \text{Tr}\left[U(\vec{\theta})\rho_k U^\dagger(\vec{\theta})\mathcal{O}\right]\right\}^2 \tag{2.4}$$

where $\mathcal{O}$ is an observable arbitrarily defined, chosen to perform the task. When this function is minimized, the optimal parameters are used in the circuit so the input state with unknown label is transformed in a way such that, when obtaining the mean value of $\mathcal{O}$, the correct label of that state is given and the classification is performed.

The trainability of the cost function still remains a concern and can be connected to the occurrence of large regions where the cost function landscape is very flat, turning the calculations of gradients and optimization into a hard procedure (12, 29). This phenomenon is called a *barren plateau* in the literature of VQA and machine learning. The causes of barren plateaus are still a contemporary topic of research and some of the studies connected their appearance to the locality of the cost function, e.g., considering a $k-$local[2] hamiltonian, using the complete hamiltonian as the cost function can lead to barren plateaus, however considering terms with smaller "locality", the landscape is less flat (19). Other reasons can be noise on the quantum circuit (32) or other characteristics of the PQC as high *expressibility* values, which leads us to the discussion about the circuit ansatz.

[2] A $k-$local hamiltonian is an operator composed of a sum of elements that act on at most k qubits, i.e., composed of a sum of tensor products with elements different from the identity in at most k subspaces.

2.2 Parameterized quantum circuits

The parameterized quantum circuit applied in the VQA context is usually called ansatz, so in this section we use the terms interchangeably. The chosen name is due to the nature of the choice of the circuit, which is made usually based on guesses or heuristics. Some proposals try to harness the characteristics of the problem to be solved to define the structure of the circuit (problem inspired ansätze) (21, 22), while others are based on different inspirations (problem agnostic ansätze) (33). The usual structure of a ansatz is presented in Fig. 2, where $U(\vec{\theta})$ is subdivided into the sequential application of L unitaries, such that

$$U(\vec{\theta}) = U_1(\vec{\theta}_1)U_2(\vec{\theta}_2)\cdots U_L(\vec{\theta}_L). \tag{2.5}$$

The number of sequential unitaries, L, is called the total number of *layers* of the circuit. The unitary applied in each of the layers is often the same in structure, however the parameters in each of them can be different. The number of layers needed to solve a particular problem can vary a lot and, sometimes, a large number is required so the optimization can be completed. However, it is worth remarking that this can harm the execution in real quantum computers, as it would imply in high depth circuits.

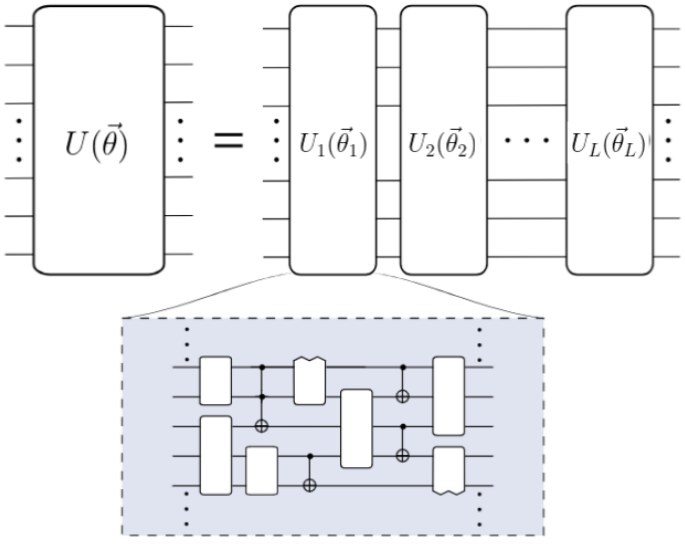

Figure 2 – Decomposition of a parameterized quantum circuit or ansatz, $U(\vec{\theta})$, in terms of the sequential application of L unitaries, $U_i(\vec{\theta}_i)$. The total number L is called number of layers of the circuit.

Source: Adapted from CEREZO *et al.* (12).

Perhaps the most famous example of a problem agnostic ansatz is the Hardware Efficient Ansatz (HEA). The inspiration for the structure is based on quantum computer architectures, with the objective of minimizing the circuit depth in an attempt of reducing error accumulation. So, despite the fact of being problem agnostic, this ansatz has a

strong dependence on the choice of the particular architecture (12,34). Usually, instead of picking a hardware and then defining the possible connections between qubits and gates that can be applied, it is considered that nearest neighbours interactions between qubits are possible (33).

In the case of combinatorial optimization problems, an important ansatz is the Quantum Alternating Operator Ansatz (QAOA) (24). Its structure is based on the application of the circuit digitalization of the evolution generated by two non-commuting hamiltonians. One of them encodes the combinatorial problem of interest, H_P, and is built of Pauli strings of only Pauli-Z operators. The other one, H_X, is based on Pauli strings of Pauli-X operators, which will introduce coherences in the computational eigenstates of the Pauli-Z operator. The hamiltonian that generates the evolution is $H = H_P + H_X$ and, using the Trotter-Suzuki decomposition (35),

$$e^{-it(H_P+H_X)} = \lim_{L\to\infty}(e^{-itH_P/L}e^{-itH_X/L})^L,\ \hbar = 1 \tag{2.6}$$

which gives a way to generate the evolution of the complete hamiltonian considering the individual evolutions of H_P and H_X. Notice that this decomposition would require an infinite number of layers to be reliably implemented. If we consider a finite number of applications, we can obtain an approximation of this evolution in the form

$$U \approx \prod_{k=1}^{L} e^{-iH_P t/L}e^{-iH_X t/L}, \tag{2.7}$$

being L of finite size. We can substitute the time dependencies by parameters, so the digitalization of the evolutions will be given by a parameterized quantum circuit. Considering different parameters for each hamiltonian and for each repetition, we are going to have the structure of the QAOA as

$$U(\vec{\gamma},\vec{\beta}) = \prod_{k=1}^{L} e^{-iH_P\gamma_k}e^{-iH_X\beta_k}, \tag{2.8}$$

so the PQC is based on the alternated application of the evolutions given by the individual hamiltonians.

In fact, the QAOA is based on the quantum annealing problem, where two hamiltonians, one with known and easy to prepare ground state and another whose ground state is sought, are interpolated as a function of time (27). The total hamiltonian is written as $H = s(t)H_P + (1 - s(t))H_X$, where $s(0) = 0$ and, after the complete evolution, $s(\tau) = 1$. The time evolution of this hamiltonian is then decomposed with the time ordered Trotter-Suzuki decomposition and translated to the parameters language. A more detailed discussion of this reasoning is given in Ref. (27).

The characteristics of the PQC have strong influence over the trainability of the cost function and some of them can also lead to barren plateaus. Two of the main contributions of quantum circuits to the appearance of barren plateaus, even in the context of zero noise, happen when they are highly expressible (26) and entangling (36). The expressibility of a quantum circuit is connected to how evenly distributed are the states it generates in the space of states (Chaps. 3 and 5). On the one hand, the more uniformly distributed are the states, the higher the chance of building the objective state that minimizes the cost function for an arbitrary problem. On the other hand, this can lead to a phenomenon called concentration of measure (discussed in Chap. 6), which results in a exponentially close to zero variance of the cost function (26), a characteristic of barren plateaus. Usually, a 2−design (Chap. 3) is already sufficient so the phenomenon occurs. In the case of highly entangling circuits, the argument can be connected to the fact that the reduced states will be close to maximally mixed states and local measurements will harm the information needed to construct the cost landscape (29, 36). Still, PQCs do not have necessarily to be applied in the context of VQAs and can be useful for building random states, states generated by $t-$designs and entangled states. This way, their study is motivated by other research subjects.

Therefore, the study of characteristics of parameterized quantum circuits from the point of view of expressibility and multipartite entanglement generation is very important to understand how and where they can be efficiently applied. In some contexts, highly entangling circuits are needed, so the entangled ground states of systems can be built (37). However, there are situations where too much entanglement will harm the performance and a well-suited amount is necessary to optimization (38, 39). Also, entanglement is a quantum resource and is very important to quantum advantage, so a careful study of its generation is essential. The expressibility identifies the ability of the circuit of generating arbitrary states and can have very deep connections with entanglement (40). This way, in the next chapters we are going to discuss the major concepts in the understanding of expressibility and entanglement, so we can talk through how their quantifiers can be related in the context of PQCs depending on particular connections of qubits in quantum hardware.

3 HAAR MEASURE AND $t-$DESIGNS

Many are the situations in Physics where we have to assign probabilistic descriptions to phenomena. Sometimes this can happen because the best model of the problem lacks complete information, therefore only events with certain probabilities are accessible. Other times, the dynamics is best described by a chaotic model or the problem is too complicated to be completely treated with a unique dynamical description and introducing probabilities is the best approach (41–43). In the specific scenario of Quantum Theory, where the dynamics of closed systems is usually described by a unitary matrix, a "randomness" on the dynamics can be introduced in the form of random matrices and sampled considering a well-suited probability measure. Thinking about this probabilistic approach to unitaries and matrices that are, in fact, representations of groups in the space of states, we are going to introduce the concepts of Haar measure and $t-$designs starting with a description of mean values of functions of group elements.

3.1 Haar measure

Given $\mathfrak{G}$ a finite group and $f : \mathfrak{G} \to \mathbb{C}$ a function that associates a complex number to every element $g \in \mathfrak{G}$, the mean value of f over the group is defined as (44)

$$\mu(f) := \frac{1}{|\mathfrak{G}|} \sum_{g \in \mathfrak{G}} f(g), \tag{3.1}$$

where $|\mathfrak{G}|$ is the number of elements of the group. This mean value possesses some very interesting properties concerning the group action. Firstly, we can say that

$$\mu[f(g)] = \mu[f(hg)] = \mu[f(gh)] = \mu[f(g^{-1})], \ \forall h \in \mathfrak{G}. \tag{3.2}$$

These are called the invariance properties of the means over groups (44). We are going to give arguments for the proofs of the equalities for multiplication from the right by an element of the group and for the inverse. The argument is the same for left multiplication. Starting with the right multiplication, we have

$$\mu[f(hg)] = \frac{1}{|\mathfrak{G}|} \sum_{g \in \mathfrak{G}} f(hg). \tag{3.3}$$

However, by the definition of a group (45,46), if $h, g \in \mathfrak{G} \Rightarrow hg \in \mathfrak{G}$. This way, as the sum is over all the elements of the group and hg is again an element of the group, the sum runs over all the elements of the group for hg analogously to g, and then $\mu[f(hg)] = \mu[f(g)]$. Next, for the inverse case, $f(g^{-1})$, every element in the group has an inverse in the group

and being it the only one, summing over all the inverses is the same as summing over all the elements in the group, therefore $\mu[f(g)] = \mu[f(g^{-1})]$. Notice also that $\mu(f)$ satisfies: if $f(g) = 1\ \forall g \in \mathfrak{G}$, then $\mu(f) = 1$, so the mean is normalized. This mean satisfies positivity, as if $f(g) \geq 0 \Rightarrow \mu(f) \geq 0$.

Here, we defined the mean only for finite groups hitherto, but it can be generalized to infinity groups, compact[1] or non-compact with restrictions (44, 48). Some interesting examples are the following (44): for the $SO(2)$ or $\mathcal{U}(1)$,

$$\mu(f) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta)d\theta. \tag{3.4}$$

As the groups $SO(2)$ and $\mathcal{U}(1)$ are isomorphic, the mean can be written in the same way (θ has an one to one correspondence to the elements of both groups). This formula is valid if the integral is finite. Notice that this integration has an element that was already presented in the finite group case: the "volume" of the group as a normalization of the measure due to compactness, in this case, 2π. If we assign $f(\theta) = 1, \forall\theta$, the mean is normalized.

The next example of the $(\mathbb{R}_+, \cdot)$[2] group show a case where the mean is invariant, but the group does not satisfy compactness

$$\mu(f) := \int_0^{\infty} f(x)\frac{1}{x}dx. \tag{3.5}$$

Due to the condition of compactness not being satisfied, the mean is not normalized.

An invariant integration measure of this kind over the topological space of groups (with some extra peculiarities) is called a Haar measure (44,48). As our great interest are the unitary groups, we are going to devote most of the attention to them. The unitary groups are always compact, as a consequence of the decomposition of the topological space of the group as a subset of an Euclidean real space and due to being a bounded and closed subset (Heine-Borel theorem) (49). This way, we are going to talk about the Haar measure only over compact groups.

The Haar theorem states that given a compact group $\mathfrak{G}$, there is an integration measure $d\mu(g)$ in $\mathfrak{G}$ called Haar measure, satisfying that if the mean value

$$\mathbb{E}(f) = \int_{\mathfrak{G}} f(g)d\mu(g) \tag{3.6}$$

[1] Here, we are not going to be precise about the definition of compactness of topological manifolds. In this text, we use "finite volume" as a notion of compactness, however this is not strictly precise. A proper definition can be found in Szekeres book (47).

[2] The group where the elements are in the positive real set, $\mathbb{R}_+$, and the composition operation is the multiplication, $\cdot$.

is finite, then the measure is invariant under the group action, with the same properties presented in Eq. (3.2). Also, the measure can be normalized, $\int_{\mathfrak{G}} d\mu(g) = 1$, and satisfies positivity as introduced before.

Now, the unitaries in the unitary group can be assigned to matrices/operators acting in the space of states by the representation map that takes elements of the group to the set of automorphisms[3] of the vector space. The connection between the group and the transformations in Quantum Theory is direct. The Haar measure is defined over the group space and assigns a weight to each of the unitaries, which can be assigned to a probability measure over the same space due to its properties (41). This way, connecting both concepts, with the Haar measure we have a tool to sample random operations and calculate statistical moments of functions of operations in the space of states.

Some examples are on point. In the context of quantum reference frames (50), two parties, Alice and Bob, can share reference frames described by Quantum Theory and that are connected by a transformation[4], let's say a unitary transformation $U \in \mathcal{U}(d)$. There will be situations where Alice have no information about how Bob's frame is related to hers. This will introduce an uncertainty over the U transformation, so, if Bob prepares a state ρ and Alice wants to represent it in her reference frame, she would have to weight every possible frame transformation U in the state Bob gives her and take the mean value as

$$\mathcal{G}[\rho] := \int_{\mathcal{U}(d)} d\mu(U) U \rho U^{\dagger}, \tag{3.7}$$

where U is the representation of the element in the space of states and $\mathcal{G}[\rho]$ is the so obtained state in Alice's reference frame. We used the letter U for both the unitary and its representation. Here, the Haar measure is directly introduced as a probability measure over the elements of the unitary group and the function of the unitaries is, in fact, an operator. Sometimes the process $\mathcal{G}$ is called G-twirling (50).

Another interesting application is the No-Free Lunch Theorem for Quantum Machine Learning (51). In a machine learning label task, we want to build a model that relates a state to a label. This can be, in the quantum realm, the classification of states as pure or mixed, as separable or entangled, for example. But this type of model requires some kind of training using labeled states and considering a specific labeling function that connects the set of states to the set of labels that we are trying to achieve. If we consider all possible functions, or, in the quantum case, unitaries, that could relate the sets of states and labels, we could quantify how well on average the model we built would

[3] Automorphisms f are bijective linear transformations defined from a vector space to itself, i.e., $Aut(V) := \{f : V \to V | \text{f invertible and linear}\}$, V vector space.

[4] This can be, e.g., two phase reference frames. In this case, the two reference frames are related by transformations $U \in \mathcal{U}(1)$.

perform in the different situations. This gives some intuition into what will be the machine learning models performance if we change the class of problem we are dealing with. A bona-fide quantifier of how well the machine learning model performs a task is the probability of the unitary model we built, V, gives the wrong answer to the unitary we want to achieve, U. This is called the quantum risk function $R_U(V)$. To obtain the mean over all possible tasks, U, we calculate the average of $R_U(V)$ over all unitary operations, using the equation (51)

$$\mathbb{E}_U[R_U(V)] = \int_{\mathcal{U}(d)} d\mu(U) R_U(V) = \frac{d}{d+1} - \frac{1}{d(d+1)} \int_{\mathcal{U}(d)} d\mu(U) \left|\mathrm{Tr}\left(U^\dagger V\right)\right|^2, \tag{3.8}$$

where the final integral is a function of the entries of the unitary operator U, which is weighted using the Haar measure. The name "no-free lunch" comes from the fact that a particular machine learning algorithm that presents the best performance for a particular labeling function will have a limited performance on average over every possible labeling function (51).

3.2 Unitary and polynomial designs

These examples present the relevance of the Haar measure and its most interesting characteristic: When used as a probability measure, it will be the uniform distribution over the space of unitaries. This way, it provides the possibility of performing the exact calculation of mean values of functions that depend on operators acting on the space of states (48). Now, the Haar measure performs the task of uniform sampling perfectly. A different problem that can be proposed is, given an ensemble of operations in the unitary group, how close the weighted functions estimates calculated with this ensemble can be to those obtained with the real uniform Haar measure? This question motivates the definition of $t-$designs and frame potentials.

Definition 1. (Unitary $t-$design)(52) Let $U(x) \in \mathcal{U}(d)$ be a unitary representative of the equivalence class $U(x) \sim e^{i\phi}U(x)$, where $U(x) \in P\mathcal{U}(d) = \mathcal{U}(d)/U(1)$[5], and μ denote the normalized Haar measure on $P\mathcal{U}(d)$. Call a countable set $\mathcal{I}$ endowed with a weighted function $w : \mathcal{I} \to [0,1]$, such that $\sum_{x\in\mathcal{I}} w(x) = 1$, a weighted set and denote it as $(\mathcal{I}, w)$. Then, a finite weighted set $(\mathcal{I}, w)$, $\mathcal{I} \in P\mathcal{U}(d)$, is called a weighted $t-$design in dimension d if

$$\sum_{x\in\mathcal{I}} w(x)(U(x)^{\otimes t})^\dagger \otimes U(x)^{\otimes t} = \int_{P\mathcal{U}(d)} d\mu(x)(U(x)^{\otimes t})^\dagger \otimes U(x)^{\otimes t}. \tag{3.9}$$

[5] We are considering the projective space due to the form of the integrals and sums: The unitaries are taken in a tensor product with their conjugates. Therefore, if the results are valid for U, then they are valid for $e^{i\phi}U$, $\forall\phi$, justifying the definition over $P\mathcal{U}(d)$ without loss of generality.

The integrals up to the $t-$th power of the products of the unitary operators are called the moments of the distribution. This way, when we say a weighted set is a $t-$design, we are saying that the related ensemble reproduces the moments of the calculated with the Haar measure up to the $t-$th order. Then, every quantity calculated up to this order with the ensemble $(\mathcal{I}, w)$ will share exactly the properties of the one calculated with the Haar measure (52). This definition relaxes the conditions needed to obtain a uniform invariant sampling over the space of unitaries: The Haar integral will not be needed if the objective is to obtain results that are equivalent to their only up to the $t-$th moment.

A different definition of unitary $t-$designs is also given in terms of the degree of polynomials whose means can be calculated using the weighted set. This form is completely equivalent to the former one. However, it is going to be more useful when we talk about the mean values of entanglement quantifiers over the circuits in Chap. 6. For completeness, then, we introduce the definition of polynomial $t-$designs.

Definition 2. (Unitary $t-$design via polynomials)(53, 54) Let $\{U(x)\}_{x=1}^{K} \subset \mathcal{U}(d)$ be a finite set of unitary operators acting on the space $\mathbb{C}^d$. If the set is such that for every polynomial $P_{(t,t)}(U)$ of degree at most t in the elements and complex conjugate elements of the matrix U

$$\frac{1}{K}\sum_{x=1}^{K} P_{(t,t)}(U(x)) = \int_{\mathcal{U}(d)} d\mu(x) P_{(t,t)}(U(x)) \tag{3.10}$$

is satisfied, we say the set is a unitary $t-$design. $d\mu(x)$ is the Haar measure over $\mathcal{U}(d)$.

The unitary $t-$design definition is useful to understand which are the conditions over an ensemble of unitaries needed to obtain a sampling that is equivalent to the one obtained with the Haar measure, but does not provide a specific way to test how close or far a given ensemble is from achieving the objective. To accomplish the task, we define the *frame potential*, which relates a mean quantum channel obtained with the ensemble and the induced via the Haar measure.

Definition 3. (Frame potential)(43) The frame potential is the 2−norm (Hilbert-Schmidt (55)) distance between the $t-$fold channel induced by the Haar ensemble and the one induced by the ensemble of interest. The $t-$fold channel induced by the ensemble ε is defined as

$$\Phi_{\varepsilon}^{(t)}(\,\cdot\,) := \int_{\varepsilon} (U^{\otimes t})^{\dagger}(\,\cdot\,) U^{\otimes t} dU. \tag{3.11}$$

The frame potential will be written in the form

$$F_{\varepsilon}^{(t)} := \frac{1}{|\varepsilon|^2} \sum_{U,V\in\varepsilon} \left|\mathrm{Tr}\left(U^{\dagger}V\right)\right|^{2t}, \tag{3.12}$$

for discrete ensembles and

$$F_{\varepsilon}^{(t)} := \iint_{\varepsilon} \left|\mathrm{Tr}\left(U^{\dagger}V\right)\right|^{2t} dU dV, \tag{3.13}$$

for continuous ensembles.

The connection between the two definitions is made by the following theorem, which provides a useful way to quantify the "Haarness" of an arbitrary ensemble of unitaries.

Theorem 1. (43, 52) For any ensemble ε of unitary operators is valid that

$$F_{\varepsilon}^{(t)} \geq F_{Haar}^{(t)}. \tag{3.14}$$

with equality if and only if ε is a $t-$design.

Proof. We are going to calculate the 2-norm of the difference between the $t-$th moment generated with the ε ensemble and with the Haar measure. This quantity is

$$S := \int_{\varepsilon} (U^{\otimes t})^{\dagger} \otimes U^{\otimes t} dU - \int_{Haar} (U^{\otimes t})^{\dagger} \otimes U^{\otimes t} dU. \tag{3.15}$$

taking the square of the 2-norm

$$0 \leq ||S||_2^2 = \mathrm{Tr}\Big(S^{\dagger} S\Big), \tag{3.16}$$

where we can write

$$\begin{aligned} S^{\dagger}S = &\iint_{U,V\in\varepsilon} (UV^{\dagger})^{\otimes t} \otimes (U^{\dagger}V)^{\otimes t} dU dV - \int_{Haar}\int_{U\in\varepsilon} (UV^{\dagger})^{\otimes t} \otimes (U^{\dagger}V)^{\otimes t} dU dV - \\ &- \int_{V\in\varepsilon}\int_{Haar} (UV^{\dagger})^{\otimes t} \otimes (U^{\dagger}V)^{\otimes t} dU dV + \iint_{Haar} (UV^{\dagger})^{\otimes t} \otimes (U^{\dagger}V)^{\otimes t} dU dV. \end{aligned} \tag{3.17}$$

Now, as $\mathrm{Tr}\Big[(UV^{\dagger})^{\otimes t} \otimes (U^{\dagger}V)^{\otimes t}\Big] = \Big|\mathrm{Tr}\Big(U^{\dagger}V\Big)\Big|^{2t}$, we have

$$\begin{aligned} \mathrm{Tr}\Big(S^{\dagger}S\Big) = &\iint_{U,V\in\varepsilon} \Big|\mathrm{Tr}\Big(U^{\dagger}V\Big)\Big|^{2t} dU dV - 2\int_{Haar}\int_{U\in\varepsilon} \Big|\mathrm{Tr}\Big(U^{\dagger}V\Big)\Big|^{2t} dU dV + \\ &+ \iint_{Haar} \Big|\mathrm{Tr}\Big(U^{\dagger}V\Big)\Big|^{2t} dU dV. \end{aligned} \tag{3.18}$$

The crossed terms are symmetric due to $\Big|\mathrm{Tr}\Big(U^{\dagger}V\Big)\Big|^{2t} = \Big|\mathrm{Tr}\Big(V^{\dagger}U\Big)\Big|^{2t}$, so they are grouped in the second integral. By definition, the first and last terms in Eq. (3.18) are $F_{\varepsilon}^{(t)}$ and $F_{Haar}^{(t)}$, respectively. Also, the crossed term can be rewritten using the properties of the Haar measure. We are going to do this using a finite discrete group Haar measure to have some insight. As the properties are shared by the continuous compact group Haar measure, the results are the same.

First, considering the continuous case, we have the integral

$$\int_{Haar}\int_{U\in\varepsilon}\left|\mathrm{Tr}\left(U^\dagger V\right)\right|^{2t}dUdV. \tag{3.19}$$

Notice that, despite the fact that ε is a continuous non-uniform sampled ensemble, we still have a finite number of $U\in\varepsilon$ and we can always use the Haar measure invariance to take $V\mapsto UV$, $dV\mapsto dV$ by invariance. This way, the term $U^\dagger V$ is taken to V, excluding the ε ensemble part. Therefore

$$\int_{Haar}\int_{U\in\varepsilon}\left|\mathrm{Tr}\left(U^\dagger V\right)\right|^{2t}dUdV=\int_{Haar}\left|\mathrm{Tr}(V)\right|^{2t}dV. \tag{3.20}$$

Now, we are going to show that, in fact

$$\int_{Haar}\int_{U\in\varepsilon}\left|\mathrm{Tr}\left(U^\dagger V\right)\right|^{2t}dUdV=\int_{Haar}\left|\mathrm{Tr}(V)\right|^{2t}dV=\iint_{Haar}\left|\mathrm{Tr}\left(U^\dagger V\right)\right|^{2t}dUdV. \tag{3.21}$$

If we map the problem to the discrete case, we have

$$\iint_{Haar}\left|\mathrm{Tr}\left(U^\dagger V\right)\right|^{2t}dUdV\mapsto\frac{1}{|\mathfrak{G}|^2}\sum_{g\in\mathfrak{G}}\sum_{h\in\mathfrak{G}}\left|\mathrm{Tr}\left(g^{-1}h\right)\right|^{2t}. \tag{3.22}$$

Notice that both g and h are elements of $\mathfrak{G}$. From this, we can say that calculating the sum over all the h of $|\,\mathrm{Tr}(g^{-1}h)|^{2t}$ is the same as summing over all the elements of the group $|\,\mathrm{Tr}(h)|^{2t}$, because if $h,g\in\mathfrak{G}\Rightarrow g^{-1}h\in\mathfrak{G}$. This way, the sum over g will be only a multiplicative factor $|\mathfrak{G}|$ and

$$\frac{1}{|\mathfrak{G}|^2}\sum_{g\in\mathfrak{G}}\sum_{h\in\mathfrak{G}}|\,\mathrm{Tr}\left(g^{-1}h\right)|^{2t}=\frac{1}{|\mathfrak{G}|}\sum_{h\in\mathfrak{G}}|\,\mathrm{Tr}(h)|^{2t}. \tag{3.23}$$

Going back to the original continuous case, we proved the equality

$$\int_{Haar}\int_{U\in\varepsilon}\left|\mathrm{Tr}\left(U^\dagger V\right)\right|^{2t}dUdV=\iint_{Haar}\left|\mathrm{Tr}\left(U^\dagger V\right)\right|^{2t}dUdV=F^{(t)}_{Haar}. \tag{3.24}$$

Plugging this result into Eq. (3.18), we finally have that

$$0\leq||S||_2^2=F^{(t)}_{\varepsilon}+F^{(t)}_{Haar}-2F^{(t)}_{Haar}\Rightarrow F^{(t)}_{\varepsilon}\geq F^{(t)}_{Haar}. \tag{3.25}$$

□

This inequality is saturated only when the 2-norm of S is zero, i.e., when the ε ensemble constitutes a unitary $t-$design and the moments are the same. The frame potential is, therefore, an important tool to test how well a given ensemble will do the job of sampling as the Haar measure and the first estimate we are seeking to analyze

ensembles of unitaries. The notion of frame potentials can be extended also to states and the quantity will be explicitly employed as a descriptor of the expressibility of quantum circuits. A proper introduction of the topic is given in Sec. 5.1.

4 ENTANGLEMENT

The renowned characteristic of Quantum Computing and Quantum Information is the possibility of achieving the "quantum advantage" (56), a limit where tasks performed in quantum systems have performances that surpass the ones performed with classical computation alone. To cross this border, important problems to be solved are the proper control of quantum systems above the limit of decoherence and the exploration of characteristics that are inherently quantum. These only quantum characteristics are reliable resources to rely on, as they are the peculiarities not observed for classical methods, such as coherence and entanglement. In fact, some of the milestones of quantum information applications, namely the protocols of teleportation, superdense coding and Bell theorem based cryptography, use quantum entanglement as the most important element (57).

This way, quantum entanglement is not only a nice spice in quantum circuits, but a very important feature that should be analyzed when studying the performance and possibilities for parameterized quantum circuits. Our approach in this chapter will be directed towards quantification and classification of entanglement. In this sense, the definitions and proofs are driven and justified in terms of how entanglement can be operationally determined. We begin tackling the peculiarities of quantum entanglement defining it for pure states, as this work is not concerned with mixed global states.

Definition 4. (Entangled States)(57) Consider a composite quantum system whose Hilbert space is given by the tensor product of the N Hilbert spaces of the subsystems, $\mathcal{H} = \bigotimes_{i=1}^{N} \mathcal{H}_i$. We say a pure state is entangled if it does not admit a decomposition as a tensor product of local states, i.e., given $\rho := |\psi\rangle\langle\psi|$ entangled state of the composite system,

$$\rho \neq \bigotimes_i \rho_i \quad \text{or} \quad |\psi\rangle \neq \bigotimes_i |\psi_i\rangle \,. \tag{4.1}$$

States that are not entangled are called separable states.

4.1 Bipartite entanglement and entanglement quantifiers

The above definition classifies states either as entangled or separable, but does not provide a way to know whether a state is in one class or in the other. This classification per se will not be very useful if the desire is to use entanglement as a resource for doing tasks, because the "amount of entanglement" would also be a necessary quantity, but doing it is an interesting start. To do so, the Schmidt decomposition theorem will be presented and discussed. The theorem is formulated in the context of pure states and bipartite

entanglement, i.e., when the entanglement is considered between two subdivisions of the composite system.

Theorem 2. (Schmidt Decomposition)(35) Let $|\psi\rangle$ be a pure state of a composite system AB, with $\dim(\mathcal{H}_A) \equiv n_A$ and $\dim(\mathcal{H}_B) \equiv n_B$, supposing $n_A \geq n_B$. Then there are orthonormal states $|i_A\rangle$ for $\mathcal{H}_A$ and $|i_B\rangle$ for $\mathcal{H}_B$, such that every arbitrary pure state in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ can be written

$$|\psi\rangle = \sum_{i=1}^{n_B} \lambda_i |i_A\rangle \otimes |i_B\rangle , \tag{4.2}$$

where λ_i are non-negative real numbers satisfying $\sum_i \lambda_i^2 = 1$. This is called the state Schmidt decomposition and the number of non-zero λ_i is called the state Schmidt number.

Proof. Let $\{|j\rangle\}_{j=1}^{n_A}$ be an orthonormal basis for $\mathcal{H}_A$ and $\{|k\rangle\}_{k=1}^{n_B}$ be an orthonormal basis for system $\mathcal{H}_B$. We can write every arbitrary vector in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ in the decomposed form[1]

$$|\psi\rangle = \sum_{j,k} b_{jk} |j\rangle |k\rangle , \tag{4.3}$$

being $b = [b_{jk}]$ a $n_A \times n_B$ matrix, as the index j runs from 1 to n_A and k from 1 to n_B. Every complex matrix admits a singular value decomposition (35), so we can write the b matrix as

$$b_{n_A \times n_B} = U_{n_A \times n_A} d_{n_A \times n_B} V_{n_B \times n_B}, \tag{4.4}$$

where U and V are unitary matrices of dimension n_A and n_B, respectively, and d is a real positive semidefinite diagonal matrix in the first n_B columns and in the first n_B rows and zero in the rest. Explicitly,

$$d_{n_A \times n_B} = \begin{bmatrix} D_{n_B \times n_B} \\ O_{n_A - n_B \times n_B} \end{bmatrix}. \tag{4.5}$$

Now, if we introduce this decomposition in the form of the arbitrary state $|\psi\rangle$,

$$|\psi\rangle = \sum_{i,k=1}^{n_B} \sum_{j=1}^{n_A} D_{ii} U_{ji} V_{ik} |j\rangle |k\rangle ; \tag{4.6}$$

setting $|i_A\rangle := \sum_{j=1}^{n_A} U_{ji} |j\rangle$ and $|i_B\rangle := \sum_{k=1}^{n_B} V_{ik} |k\rangle$, we can verify that each of the so obtained sets forms an orthonormal set in $\mathcal{H}_A$ and $\mathcal{H}_B$, due to the sets $\{|j\rangle\}$ and $\{|k\rangle\}$

[1] Here, we apply the notation $|j\rangle |k\rangle \equiv |j\rangle \otimes |k\rangle$.

being orthonormal basis[2]. In fact, the set $\{|i_B\rangle\}_{i=1}^{n_B}$ is a basis for $\mathcal{H}_B$; this is not true for $\{|i_A\rangle\}_{i=1}^{n_B}$, because the dimension of $\mathcal{H}_A$ is bigger than $\mathcal{H}_B$. Finally, if we set the elements of the diagonal matrix $D_{ii} \equiv \lambda_i$, then

$$|\psi\rangle = \sum_{i=1}^{n_B} \lambda_i |i_A\rangle \otimes |i_B\rangle . \tag{4.7}$$

□

Now, the connection between the Schmidt decomposition and classification of entanglement can be predicted: the decomposed state is written as a sum of product states, in other words, separable pure states[3]. If the Schmidt number is equal to one, i.e., there is only one term in the sum, the state is certainly separable and is not entangled. This will be precisely the result of the next corollary.

Corollary 1. A pure state $|\psi\rangle$ of a composite system AB is separable (or "not-entangled") if and only if it has Schmidt number equal to one.

Proof. We start by stating the simplest part. If a state has Schmidt number equal to one, then it can be written in the Schmidt decomposition as

$$|\psi\rangle = \sum_i \lambda_i |i_A\rangle |i_B\rangle = \lambda_i |i_A\rangle |i_B\rangle = |i_A\rangle |i_B\rangle , \tag{4.8}$$

which is a separable state, by definition. We used $\lambda_i = 1$ due to the normalization of the state.

To complete the proof, we introduce a discussion about the density matrix of the pure state. We firstly state a simple conclusion about separable pure states. Given a separable pure state of the system AB, $\rho = \rho_A \otimes \rho_B$, the reduced states, i.e., ρ_A and ρ_B, are obviously pure. Now, this will lead to the following result. The density matrix related to $|\psi\rangle$ is given as

$$\rho := |\psi\rangle \langle\psi| = \sum_{i,j} \lambda_i \lambda_j |i_A\rangle |i_B\rangle \langle j_A| \langle j_B| . \tag{4.9}$$

Tracing out one of the subsystems,

[2] Here the reason why the choice is made when building the b matrix with $n_A \geq n_B$ becomes clear: if this was chosen the other way around, the set $\{|i_B\rangle\}$ would have $n_A > n_B$ elements and would not be orthonormal.

[3] Care must be taken. Here, the superposition of pure product states can generate an entangled state. This is completely different from the case of separable mixed states, where a convex combination is considered and represents a statistical mixture related to classical uncertainty about the prepared separable state.

$$\rho_A = \mathrm{Tr}_B(\rho) = \sum_{i,j} \lambda_i \lambda_j \left|i_A\right\rangle \left\langle j_A\right| \delta_{ij} = \sum_i \lambda_i^2 \left|i_A\right\rangle \left\langle i_A\right|, \tag{4.10}$$

which is the reduced density matrix related to system A. The notation Tr_B indicates the partial trace over subsystem B (35). This is the spectral decomposition of the reduced density matrix and, therefore, the number of non-zero λ_i (Schmidt number) is the dimension of the image of ρ_A as an operator or its rank[4]. The result follows directly: if the state ρ is separable, ρ_A has to be pure or a state of rank 1, implying that the Schmidt number is equal to one.

□

One of the consequences of this last part of the proof is that entangled states have reduced states that are not pure, i.e., mixed. Therefore, the reduced states will present classical uncertainties due to the fact that the complete information only can be accessed considering the global state. The Schmidt decomposition is, therefore, an operational way to classify bipartite states: the state will be entangled iff[5] the Schmidt number is different from one. An example of this are the Bell states, which are already in the decomposed form, e.g.,

$$\left|\psi^+\right\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \tag{4.11}$$

whose Schmidt number is equal to two.

Classifying pure quantum states as entangled or separable in bipartitions is an important task and also can lead to other ways of quantifying entanglement. Some proposals generalize the concept of Schmidt number to mixed states (59) and also build ways to quantify bipartite entanglement from the Schmidt decomposition (60). Even more, the Schmidt number was applied to compare the performance of a quantum computation to a classical simulation; Vidal showed that if under any bipartition at all times the Schmidt number is polynomial in the number of qubits, the task can be classically performed (57).

However, by itself, the decomposition does not quantify entanglement in a closed way. To introduce the quantification of entanglement, the axiomatic approach will be considered, where entanglement quantifiers have to satisfy two basic conditions that can be made stronger or added with others, if convenient. This axiomatic approach is in consonance with quantum resource theories (61), which are going to be presented implicitly when talking about entanglement as a resource.

[4] The rank of a linear operator $L : V \to W$, between vector spaces V and W, is defined as $\mathrm{rank}(L) := \dim[\mathrm{Im}(L)]$ (58), where $\mathrm{Im}(L)$ is the image of the linear operator (the proper definition of image is given in App. A).

[5] If and only if.

Entanglement can be thought as a non-local resource in states, a quantity that is used to perform tasks and is shared between parties. As a limited "natural resource", it is not easily created and we should think about a set of operations which cannot increase the amount of entanglement. This set of operations will be the Local Operations and Classical Communication (LOCC), due to the non-local nature of the correlation: We can perform changes locally (local operations, already thought as quantum) and it should not increase the resource. Classical communication is usually considered a task that does not require and does not increase quantum resources, being naturally added to the set. Also, there will be a set of states which do not have the resource and cannot be used to perform tasks requiring it. These will be the not-entangled states: the set of separable states (57, 61).

Now, this specifies the framework where entanglement can be worked out as a resource. The question still not answered is: How entanglement can be quantified? Following the axiomatic approach, we introduce the definition of entanglement quantifiers on the Hilbert space where density operators are defined.

Definition 5. (Entanglement Quantifier)(61) A positive real-valued function from the set of density operators of a Hilbert space $\mathcal{H}$ of the form $f : \mathcal{D}(\mathcal{H}) \to \mathbb{R}_{\geq 0}$ is called an entanglement quantifier if satisfies the minimum requirements:

- **Monoticity under LOCC**: If $\Phi(\cdot)$ is a LOCC, then considering a state $\rho \in \mathcal{D}(\mathcal{H})$,

$$f(\rho) \geq f(\Phi(\rho)). \tag{4.12}$$

- **Vanishing for separable states**: If σ is a state in the set of separable states, then

$$f(\sigma) = 0. \tag{4.13}$$

The conditions in Def. 5 give quantitative interpretation about the concepts discussed in paragraphs above. The first condition makes precise the non-increasing over LOCC condition, while the second one states the existence of states which do not have the entanglement resource, in other words, separable states. Other conditions that could be added to the definition of a "resource quantifier" are convexity, reasonable for mathematical reasons, or subadditivity (61, 62).

One important quantifier in the above sense for bipartite pure states entanglement is the entropy of entanglement. This quantifier is defined as the von Neumann entropy of the reduced density matrix of the subsystems when considering a bipartite system AB, i.e. (55, 62),

$$E(\rho) := S(\mathrm{Tr}_A\, \rho) = S(\mathrm{Tr}_B\, \rho), \quad S(\rho) := -\,\mathrm{Tr}\left[\rho \log_2 \rho\right], \tag{4.14}$$

being $S(\rho)$ the von Neumann entropy of the state ρ.

The entropy of entanglement satisfies the minimum requirements for a quantifier in Def. 5. Its importance is related to the first works about distillation and concentration of entanglement over pure states and, in fact, for pure bipartite states in the asymptotic regime the entropy of entanglement is equal to the distillable entanglement[6] and to the entanglement cost[7](63). This is a condition for reversibility of an entanglement theory under LOCC, as every entanglement quantifier with reasonable restrictions should be bounded above by the entanglement cost and below by the distillable entanglement (57); given that for pure states asymptotically all quantifiers reduces to the entropy of entanglement (62, 64), there is a unique entanglement quantifier and the reversibility is possible (64, 65). In other words, we can establish a constant rate of conversion between two arbitrary pure quantum states (from a number of systems in the state 1 to a set of maximally entangled and then from this set to a number of systems in state 2) similar to the Carnot efficiency of thermal machines (64). This is why reversibility is established: we can go from a number of copies of state 1 to a number of copies of state 2 and conversely from 2 to 1, if we assure that the total entanglement does not increase. This result is valid only in the asymptotic case for pure bipartite states. Entanglement theory in the asymptotic regime under LOCC was already proven to be irreversible if mixed states are included, due to the existence of states that cannot be used to distillate entanglement, however have a nonzero entanglement cost, the bound states (66). In fact, recent results where the allowed operations are the more general set of nonentangling operations, which can include any operation that does not generate entanglement, e.g., LOCC, showed that entanglement theory is irreversible (67).

The quantification of entanglement using the von Neumann entropy is reflected by the relation of this entropy with the purity of the input state. The quantity has two properties: it is minimal and equal to zero when the input state is pure and is maximal and equal to $\log_2(d)$ when the input state is maximally mixed (55). The d stands for the dimension of the subsystem. We saw that for separable states, ρ_{sep}, $\mathrm{Tr}_A(\rho_{sep})$ and $\mathrm{Tr}_B(\rho_{sep})$, namely the reduced density matrices, have to be pure states. Therefore, the entropy of entanglement is zero. Also, a maximally entangled bipartite pure state will have reduced density matrices that are maximally mixed. This can be observed for the case of any of the Bell states. Considering the singlet state,

$$\begin{aligned} \left|\psi^-\right\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle) \\ \Rightarrow \rho &= \left|\psi^-\right\rangle\left\langle\psi^-\right| = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10| - |01\rangle\langle 10| - |10\rangle\langle 01|). \end{aligned} \tag{4.15}$$

Tracing out the subsystem of the first particle, which we are going to call A,

[6] The amount of maximally entangled states that can be obtained from a given state by the means of LOCC.

[7] The amount of maximally entangled states needed to produce the given state via LOCC.

$$\rho_B = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|), \tag{4.16}$$

which is a maximally mixed state for the other subsystem, B. This will give an entropy of entanglement equal to $\log_2(2) = 1$. Therefore, the entropy of entanglement is a normalized entanglement quantifier.

Expanding the standard von Neumann entropy using the Mercator series, we can obtain another interesting entanglement quantifier (68). Considering the expansion

$$\ln(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots, \tag{4.17}$$

we can obtain a linear approximation of the logarithm in the von Neumann entropy by considering only the first term, which is going to give the linear entropy

$$S_L(\rho) = -\operatorname{Tr}\left[\rho(\rho-1)\right] = 1 - \operatorname{Tr}\left(\rho^2\right), \tag{4.18}$$

and we can recognize it as the impurity of the state ρ, i.e., one minus the purity $\operatorname{Tr}(\rho^2)$. Calculating this quantity for the reduced state of one part in a bipartite system, we define the linear entropy of entanglement

$$E_L(\rho) := S_L(\operatorname{Tr}_A \rho). \tag{4.19}$$

This quantity will be central when we obtain the Meyer-Wallach entanglement quantifier in terms of the purity of the reduced states and also for the generalization of the quantifier for systems with different bipartition sizes in the next section.

4.2 Multipartite entanglement

Up to now, we only have discussed bipartite entanglement. When we start thinking about a number of partitions that is greater than two, there will be not only many quantifiers as in the case of bipartite entanglement out of the asymptotic limit, but also different degrees of bipartite entanglement can be shared between different partitions (69). An example of this characteristic is a state of the form

$$|\psi\rangle_{ABC} = \frac{1}{2}(|00\rangle_{AB} + |11\rangle_{AB}) \otimes (|0\rangle_C + |1\rangle_C), \tag{4.20}$$

which is maximally entangled considering the subsystem AB, but is separable when considering the bipartition $AB - C$. We can use the entropy of entanglement to quantify the entanglement between the bipartitions. Consider the bipartition $AB - C$. The state is obviously separable, then the entropy will result in zero. Now, when considering a bipartition $A - BC$, we can calculate the reduced density matrix tracing out the subsystem A. This gives the state

$$\rho_{BC} = \mathrm{Tr}_A(|\psi\rangle\langle\psi|_{ABC}) = \frac{1}{4}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}. \tag{4.21}$$

The state ρ_{BC} is block diagonal in the basis $\{00, 01, 10, 11\}_{BC}$ and it can be diagonalized. The eigenvalues obtained are $\{0, 0, 1/2, 1/2\}$, giving as a result an entropy of entanglement equal to one. When compared with states that present maximum tripartite entanglement, as the GHZ (Greenberger-Horner-Zeilinger) state, $\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$, we can notice that there is a coincidence in the value of entanglement when taking the bipartition $A - BC$: both the example and the GHZ have maximum entropy of entanglement, despite the qubit C being separable in the total state in the example $|\psi\rangle_{ABC}$. This way, bipartite entanglement quantifiers solely cannot characterize entanglement in the multipartite scenario.

Another important characteristic of multipartite entanglement arises when we think about the existence of equivalence classes in the set of entangled states. Considering the case of pure single copies of states of a quantum system, we can ask which are the entangled states that can be obtained from each other under a particular set of operations. These would be equivalence classes of entangled states, in the sense that the states in any of them can be connected to each other using operations in the chosen set. We must choose wisely the set of operations, so there are a finite number of equivalence entanglement classes and we are not lost in a sea of infinitely many classes.

Before, we said that LOCCs are an important set of operations and they appear in many instances when talking about entangled states. This set is a promising choice for the set of operations to build the equivalence classes. We are going to understand what are the effects of this choice by considering the most simple case of a bipartite system. First of all, it was shown in Ref. (70) that, for the case of pure states, the LOCC-equivalence of states is identical to the equivalence under local unitaries (LU). Therefore, we can restrict the study to how states are equivalent under the tensor product of unitaries. Considering the Schmidt decomposition, Eq. (4.2), a pure state of two qubits $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ can be decomposed as

$$|\psi\rangle = \sum_{i=1}^{2} \lambda_i |i_A\rangle |i_B\rangle = \lambda_1 |1_A\rangle |1_B\rangle + \lambda_2 |2_A\rangle |2_B\rangle, \tag{4.22}$$

being the sets $\{|1\rangle, |2\rangle\}_{A,B}$ orthonormal vector basis in $\mathcal{H}_A$ and $\mathcal{H}_B$. Notice that in this case there is only one free parameter for the definition of the state, as the coefficients shall satisfy $\lambda_1^2 + \lambda_2^2 = 1$. We can consider a LU transformation that change from the basis $\{|1\rangle, |2\rangle\}_{A,B}$ to the computational basis (71). We are going to write simply $\{0, 1\}$ without the subindexes A, B for the computational basis. Giving

$$U \otimes V : \begin{cases} U := |0\rangle\langle 1_A| + |1\rangle\langle 2_A| \\ V := |0\rangle\langle 1_B| + |1\rangle\langle 2_B| \end{cases}, \tag{4.23}$$

then

$$U \otimes V \, |\psi\rangle = \lambda_1 \, |00\rangle + \lambda_2 \, |11\rangle \, . \tag{4.24}$$

This state can be written in terms of another parameter $\theta \in [0, \pi/2]$ in an one to one correspondence if we set $\lambda_1 \equiv \cos(\theta)$ and $\lambda_2 \equiv \sin(\theta)$, therefore

$$U \otimes V \, |\psi\rangle = \cos(\theta) \, |00\rangle + \sin(\theta) \, |11\rangle \, . \tag{4.25}$$

Notice that an arbitrary $|\psi\rangle$ state is equivalent to $U \otimes V \, |\psi\rangle$ under a LU and the free parameter does not change in the transformation. Therefore, no LU transformation can take states with different values of θ into each other and there shall be infinitely many equivalence classes of states under LU, or, equivalently, LOCC, which are then continuously parameterized by θ. This way, we have no hope that LOCC will be the wisest choice for defining equivalence classes of entangled states considering single copies. [8]

These equivalence classes can be built using a different set of operations, however very close in spirit to the LOCC. If we consider LOCCs with a related probability of success in achieving the transformation, we obtain the set of stochastic-LOCC (SLOCC). Two equivalent states under SLOCC can be obtained from each other with a certain non-vanishing probability and can be used to perform the same tasks, although with possibly different probabilities of success (72). We said that two states can be obtained from each other by means of LOCC iff they are related by LU (70). If we consider the set of SLOCC operations, this condition is relaxed to the condition that two states are equivalent under SLOCC iff they are related by an invertible local operator (ILO) (72). ILO does not change local ranks of reduced density matrices, therefore we cannot transform states with different local ranks into each other by the means of SLOCC (72). This constraint will be essential in building the equivalence classes of entangled states, as the different classes will have different local ranks.

Considering the two qubits case, we have two possibilities for the local ranks for the states: (i) if the reduced state is pure, and (ii) if the reduced state is mixed. This way, we can already see the two possible equivalence classes in the set of two qubits states. If the state is separable, the reduced states will be pure and the local ranks are equal to 1. If the state is entangled, the reduced states will be mixed and the local ranks are equal to 2;

[8] A disclaimer: in fact, when considering many copies of the state in the asymptotic limit, the LOCC will be enough to classify the equivalence classes of entangled pure bipartite states into two classes: separable or Bell states (65). For the multipartite case, this is not true anymore (72). We chose to exemplify the case of single copies because the reasoning is more direct and will lead straight to the use of stochastic-LOCC in the multipartite entanglement case.

as the local spaces are dimension 2, there is only one possible set, because any state can be converted to a decomposition of two terms in the computational basis. This first set defines the separable states, while the second defines the equivalence class of Bell states, the entangled states.

Now, we can turn the attention to the more complex case of three qubits, whose Hilbert space is the tensor product $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$, being $\mathcal{H}_i$, $i = A, B, C$, one qubit spaces. In this case, considering the SLOCC classification (72), we have 6 equivalence classes:

1. The separable states with all local ranks equal to 1:

$$|\psi\rangle = |A\rangle \otimes |B\rangle \otimes |C\rangle\,; \tag{4.26}$$

2. The three different biseparable states that are separable in one qubit and maximally entangled in the other two, changing only the separable subsystem (one local rank equal to 1 and two local ranks equal to 2):

$$|\psi\rangle = \begin{cases} |A\rangle \otimes (|B\rangle + |C\rangle) \\ (|A\rangle + |B\rangle) \otimes |C\rangle\,; \\ (|A\rangle + |C\rangle) \otimes |B\rangle \end{cases} \tag{4.27}$$

3. The states with all local ranks equal to 2, which decompose into two classes of states:

   a) The W (Wolfgang) states (72) that can be written as a sum of three terms

$$|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle); \tag{4.28}$$

   b) The GHZ (Greenberger-Horne-Zeilinger) (73) states that can be written as a sum of two terms

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \tag{4.29}$$

The GHZ state can be regarded as the maximally entangled state of 3 qubits (72). Also, when we trace out any of the qubits, the reduced GHZ is separable, while the W is highly entangled (considering bipartite mixed state entanglement). This can be understood by the fact that the entanglement of the GHZ state is shared between the three parties and cannot be thought as a composition of three and two parties entanglement. We say that the W state is robust to the loss of information of 1 qubit. This remains true to the generalized case of W states (74).

In the case of a number of qubits greater or equal to 4, there are infinitely many nonequivalent classes of entangled states (72) by the classification of SLOCC and we can see that the multipartite case is not only different from the bipartite case when compared using bipartite entanglement quantifiers, but also has completely different properties for the entanglement the parts can share. This existence of infinitely many classes under SLOCC classification did not limit the entanglement classification for systems with more than 3 qubits. Novel proposals have been made and subdivided the space of states into equivalence classes under other assumptions, each of them with elements inequivalent under ILO, therefore, with infinitely many classes under SLOCC inside them. However, even within these different classifications, the number of equivalence classes gets bigger and bigger with the number of qubits, being, in the case of 4 qubits, nine considering the method in Ref. (74) and thirty four using the method presented in Refs. (75, 76). In the first method, the study is directed towards the 4 qubits case, while in the second one the proposed method is for a general number of qubits.

4.2.1 Meyer-Wallach and Scott quantifiers

The problem introduced now is how to quantify the multipartite entanglement, instead of entanglement between bipartitions solely. This is an important problem for quantum information, as the applications involve many coupled spins and the quantification or analysis of entanglement between them. Two important quantifiers of multipartite entanglement are the $n-$tangle (77) and the $n-$concurrence (78), which can be generalised to the mixed states case. From a perspective of quantifying entanglement and understanding the different values of entanglement depending on the state, with different values for states as W, GHZ and biseparable, another important quantifier is the Meyer-Wallach and its generalisation, the Scott quantifier.

Meyer and Wallach (79) introduced a quantifier for global multipartite entanglement using a clever result from the wedge product between two vectors. The wedge product can be thought as a generalization of the cross product, whose norm squared will have a direct relation with the angle between the vectors involved. We will write the squared norm of the wedge product as (80)

$$D(\vec{a},\vec{b}) \equiv |\vec{a}\wedge\vec{b}|^2 = |\vec{a}|^2|\vec{b}|^2 - |\vec{a}\cdot\vec{b}|^2. \tag{4.30}$$

The definition makes explicit that when $\vec{a} \parallel \vec{b}$, $D(\vec{a},\vec{b}) = 0$. Also, when $\vec{a} \perp \vec{b}$, we have the maximum value of the wedge product squared norm for a given pair $\vec{a},\vec{b}$, $D(\vec{a},\vec{b}) = |\vec{a}|^2|\vec{b}|^2$. To understand how this contributes to the quantification of entanglement between parties, we define a linear map from n to $n-1$ qubits states and prove the following proposition.

Let $(\mathbb{C}^2)^{\otimes n}$ be the space of states of n qubits with a basis labeled by strings $\{b_1,...,b_n\},\ b_j \in \{0,1\}$. For $\lambda \in \{0,1\}$, define the map $i_j : (\mathbb{C}^2)^{\otimes n} \to (\mathbb{C}^2)^{\otimes n-1}$ as

$$i_j(\lambda)\left|b_1...b_j...b_n\right\rangle := \delta_{\lambda b_j}\left|b_1...\hat{b}_j...b_n\right\rangle, \tag{4.31}$$

where the hat symbol, $\hat{b}_j$, means the absence of the $j-$th qubit in the initial state. This linear map does the job of projecting to the 0 or 1 states of the $j-$th qubit and gives as output the state of the rest of the system when projected. For example, if we consider the state $|\varphi\rangle = \frac{1}{\sqrt{2}}(|001\rangle + |100\rangle)$, calling the first qubit 1, the action would be $i_1(0)|\varphi\rangle = \frac{1}{\sqrt{2}}|01\rangle$ and $i_1(1)|\varphi\rangle = \frac{1}{\sqrt{2}}|00\rangle$. We now can dig into how this is interesting to quantify entanglement.

Proposition 1. Given a n qubit state $|\psi\rangle$, the bipartite entanglement between the $k-$th qubit and the rest of the system is maximized with respect to the entropy of entanglement when $D(i_k(0)|\psi\rangle, i_k(1)|\psi\rangle)$ is maximum and minimized when it is minimum. In other words, the entanglement is maximum when the projected states are orthogonal and minimum when they are parallel.

Proof. To start the proof, we introduce a notation for simplification. Let $i_k(0)|\psi\rangle \equiv |\tilde{u}_{\hat{k}}\rangle$ and $i_k(1)|\psi\rangle \equiv |\tilde{v}_{\hat{k}}\rangle$ be the projected not necessarily normalized states. The $\tilde{}$ symbol indicates the state is not normalized, while $\hat{k}$ indicates a state of every qubit, but the $k-$th. This way, we can write the complete state $|\psi\rangle$ as

$$|\psi\rangle = |0_k\rangle \otimes |\tilde{u}_{\hat{k}}\rangle + |1_k\rangle \otimes |\tilde{v}_{\hat{k}}\rangle, \tag{4.32}$$

being $|0_k\rangle$ and $|1_k\rangle$ the basis states of the $k-$th qubit. The associated density operator will be

$$\begin{aligned}\rho = |\psi\rangle\langle\psi| = |0_k, \tilde{u}_{\hat{k}}\rangle\langle 0_k, \tilde{u}_{\hat{k}}| + |0_k, \tilde{u}_{\hat{k}}\rangle\langle 1_k, \tilde{v}_{\hat{k}}| + \\ + |1_k, \tilde{v}_{\hat{k}}\rangle\langle 0_k, \tilde{u}_{\hat{k}}| + |1_k, \tilde{v}_{\hat{k}}\rangle\langle 1_k, \tilde{v}_{\hat{k}}|,\end{aligned} \tag{4.33}$$

and tracing out the $k-$th qubit, we have that the reduced density matrix of all other qubits is given by

$$\rho_{\hat{k}} = \mathrm{Tr}_k(\rho) = |\tilde{u}_{\hat{k}}\rangle\langle\tilde{u}_{\hat{k}}| + |\tilde{v}_{\hat{k}}\rangle\langle\tilde{v}_{\hat{k}}|. \tag{4.34}$$

We can calculate the purity of this state, even though the states $|\tilde{u}_{\hat{k}}\rangle, |\tilde{v}_{\hat{k}}\rangle$ are not normalized. This calculation will allow the analysis of bounds in the purity of the state, that we are going to associate with the orthogonality. Finally, this result can be associated with characteristics of the von Neumann entropy: It is maximum when the state is maximally mixed and minimum when the state is pure. Taking the square of the reduced density matrix and then calculating its trace, we have

$$\mathrm{Tr}\left[(\rho_{\hat{k}})^2\right] = |\langle\tilde{u}_{\hat{k}}|\tilde{u}_{\hat{k}}\rangle|^2 + |\langle\tilde{v}_{\hat{k}}|\tilde{v}_{\hat{k}}\rangle|^2 + 2\,|\langle\tilde{u}_{\hat{k}}|\tilde{v}_{\hat{k}}\rangle|^2\,. \tag{4.35}$$

Now we can finally look at how this equality depends on the inner product between states. The first two terms depend only on the normalization of states $|\tilde{u}_{\hat{k}}\rangle$ and $|\tilde{v}_{\hat{k}}\rangle$. The last one will saturate the purity above if $|\langle\tilde{u}_{\hat{k}}|\tilde{v}_{\hat{k}}\rangle|^2 = 1$ and below if $|\langle\tilde{u}_{\hat{k}}|\tilde{v}_{\hat{k}}\rangle|^2 = 0$. This way, if the states differ only with a phase, being "parallel", the purity of the reduced state is the highest. An equivalent statement is saying that the state $|\psi\rangle$ is separable in the bipartition $k-$th qubit-rest. Now, when the states are orthogonal, we will have the least purity value, giving the highest entanglement possible for the same bipartition.

□

The above proof depends on the normalization of the projected vectors, but it is complete in the sense of minimization or maximization of purity. To explicitly visualize how it works, we propose two extreme examples for 2 qubits: One separable state and one Bell state. For an arbitrary separable state

$$|\psi\rangle = (a\,|0_A\rangle + b\,|1_A\rangle) \otimes |\phi_B\rangle\,, \tag{4.36}$$

such that $|a|^2 + |b|^2 = 1$ and $|\,\langle\phi_B|\phi_B\rangle\,|^2 = 1$. Plugging this into Eq. (4.35)

$$\mathrm{Tr}\left[(\rho_A)^2\right] = |a|^4|\,\langle\phi_B|\phi_B\rangle\,|^2 + |b|^4|\,\langle\phi_B|\phi_B\rangle\,|^2 + 2|a|^2|b|^2|\,\langle\phi_B|\phi_B\rangle\,|^2 = (|a|^2 + |b|^2)^2 = 1, \tag{4.37}$$

implying the reduced state is, indeed, pure and the entropy of entanglement is zero. Now, considering the Φ^+ Bell state, $|\Phi^+\rangle = |0_A\rangle \otimes \frac{1}{\sqrt{2}}\,|0_B\rangle + |1_A\rangle \otimes \frac{1}{\sqrt{2}}\,|1_B\rangle$, the result would be

$$\mathrm{Tr}\left[(\rho_A)^2\right] = \frac{1}{4}|\,\langle 0|0\rangle\,|^2 + \frac{1}{4}|\,\langle 1|1\rangle\,|^2 = \frac{1}{2}, \tag{4.38}$$

implying a maximally mixed state and a maximum entanglement.

Now, the cleverness on the definition of Meyer-Wallach entanglement quantifier can be truly appreciated in the following definition.

Definition 6. (Meyer-Wallach entanglement quantifier)(79) Given the linear map $i_j(\lambda)$ that takes n qubits states to $n-1$ qubits states, projecting over the components of the $j-$th qubit, and the squared norm of the wedge product $D(|\psi\rangle\,, |\phi\rangle)$, the Meyer-Wallach (MW) multipartite entanglement quantifier of a n qubits state $|\psi\rangle$ is defined as

$$Q(|\psi\rangle) := \frac{4}{n}\sum_{k=1}^{n} D(i_k(0)\,|\psi\rangle\,, i_k(1)\,|\psi\rangle). \tag{4.39}$$

The Meyer-Wallach quantifier computes the contribution of every bipartition between one qubit and the rest of the system to the total entanglement using the wedge product. This reflects the total global entanglement and also, with the proper $4/n$ normalization, runs from 0, for a completely separable state, to 1, for some classes of entangled states. We are going to prove the MW satisfies the minimal requirements for entanglement quantifiers in a weaker form (79), but is still interesting for our applications. In the first part of the theorem we consider a different approach from the original reference. The second part is similar, but here we provide a detailed discussion.

Theorem 3. The Meyer-Wallach quantity is zero iff the state is separable and is invariant under local unitary operations, constituting an entanglement quantifier under Def. 5.

Proof. The first part of the proof is vanishing for separable states. This is going to be argued considering directly Eq. (4.35) of Prop. 1 for the purity of the reduced state. Considering the linear operator notation and a $n-$qubit state $|\psi\rangle$, the purity of the reduced state tracing out the $k-$th qubit is

$$\mathrm{Tr}\left[(\rho_{\hat{k}})^2\right] = |\langle\psi| i_k^\dagger(0) i_k(0) |\psi\rangle|^2 + |\langle\psi| i_k^\dagger(1) i_k(1) |\psi\rangle|^2 + 2|\langle\psi| i_k^\dagger(0) i_k(1) |\psi\rangle|^2. \quad (4.40)$$

Now, if the state $|\psi\rangle$ is separable in every bipartition one qubit-rest, the purity in the above equation will be always maximum, therefore the vectors $i_k(0)|\psi\rangle$ and $i_k(1)|\psi\rangle$ differ only by a phase for every k. This way

$$D_k \equiv D(i_k(0)|\psi\rangle, i_k(1)|\psi\rangle) = 0 \ \forall k. \quad (4.41)$$

As $Q(|\psi\rangle)$ is defined as the sum of D_k, $k \in [1, n]$, it is equal to zero for a separable state. Now, if we assume $Q(|\phi\rangle) = 0$, then $D_k = 0$, $\forall k$. The purity of the reduced state is then maximal for every k and the state $|\phi\rangle$ is separable.

The second part of the proof is about the invariance under local unitaries of the MW. Firstly, the Meyer-Wallach quantifier is invariant under unitaries acting only on $(\mathbb{C}^2)^{\otimes n-1}$. This can be seen by the fact that the wedge product $D(|\psi\rangle, |\phi\rangle)$, defined in $(\mathbb{C}^2)^{\otimes n-1}$, is invariant under unitaries acting on the states. Giving $U : (\mathbb{C}^2)^{\otimes n-1} \to (\mathbb{C}^2)^{\otimes n-1}$ unitary,

$$\begin{aligned} D(U|\psi\rangle, U|\phi\rangle) &= |U|\psi\rangle|^2 |U|\phi\rangle|^2 - |\,|\psi\rangle U^\dagger U|\phi\rangle|^2 \\ &= |\,|\psi\rangle|^2 |\,|\phi\rangle|^2 - |\langle\psi|\phi\rangle|^2 = D(|\psi\rangle, |\phi\rangle). \end{aligned} \quad (4.42)$$

Therefore each of the D_k will not be affected by transformations of this kind. Now, we are going to consider the invariance of the wedge product with a general transformation

acting only in one qubit. If we consider an arbitrary linear transformation T^j acting only on the $j-$th qubit of the state $|\psi\rangle \in (\mathbb{C}^2)^{\otimes n}$, the action in the elements of the basis would be

$$T^j |b_1...b_j...b_n\rangle = \sum_{k=0}^{1} T^j_{kb_j} |b_1...k_j...b_n\rangle , \tag{4.43}$$

where the subindexes in $T^j_{kb_j}$ are the elements in the $k-$th row and b_j column of the matrix representation of T^j. For specific basis elements, this action is something like

$$\begin{aligned} T^j |b_1...0_j...b_n\rangle &= T^j_{00} |b_1...0_j...b_n\rangle + T^j_{10} |b_1...1_j...b_n\rangle , \\ T^j |b_1...1_j...b_n\rangle &= T^j_{01} |b_1...0_j...b_n\rangle + T^j_{11} |b_1...1_j...b_n\rangle . \end{aligned} \tag{4.44}$$

A general arbitrary state would then be written as

$$T^j |\psi\rangle = T^j(|0_j\rangle \otimes \left|A_{\hat{j}}\right\rangle + |1\rangle_j \otimes \left|B_{\hat{j}}\right\rangle), \tag{4.45}$$

being $\left|A_{\hat{j}}\right\rangle$ and $\left|B_{\hat{j}}\right\rangle$ arbitrary states of all the qubits except for the $j-$th. Using the results in Eq. (4.44), we get

$$\begin{aligned} T^j |\psi\rangle &= (T^j_{00} |0_j\rangle + T^j_{01} |1_j\rangle) \otimes \left|A_{\hat{j}}\right\rangle + (T^j_{01} |0_j\rangle + T^j_{11} |1_j\rangle) \otimes \left|B_{\hat{j}}\right\rangle \\ &= |0_j\rangle \otimes (T^j_{00} \left|A_{\hat{j}}\right\rangle + T^j_{01} \left|B_{\hat{j}}\right\rangle + |1_j\rangle (T^j_{01} \left|A_{\hat{j}}\right\rangle + T^j_{11} \left|B_{\hat{j}}\right\rangle). \end{aligned} \tag{4.46}$$

This way, it follows that

$$\begin{aligned} i_j(0)T^j |\psi\rangle &= T^j_{00} \left|A_{\hat{j}}\right\rangle + T^j_{01} \left|B_{\hat{j}}\right\rangle ; \\ i_j(1)T^j |\psi\rangle &= T^j_{01} \left|A_{\hat{j}}\right\rangle + T^j_{11} \left|B_{\hat{j}}\right\rangle , \end{aligned} \tag{4.47}$$

and, taking the wedge product, the result will be

$$D(i_j(0)T^j |\psi\rangle , i_j(1)T^j |\psi\rangle) = \left|\det\left(T^j\right)\right|^2 D(i_j(0) |\psi\rangle , i_j(1) |\psi\rangle). \tag{4.48}$$

Therefore, each $j-$th term in the summand that defines the Meyer-Wallach quantifier is invariant also by an operation T^j acting on the $j-$th qubit if the transformation satisfies $|\det(T^j)|^2 = 1$. These transformations constitute the special linear group of dimension 2, $SL(2)$. We can then say that each term in the summand, D_j, is invariant by the group $SL(2) \otimes U(2^{n-1})$. As in the MW there is a superposition of all the D_j, the invariance of Q is reflected by the intersections over all the qubits

$$\bigcap_j SL^j(2) \otimes U^{\tilde{j}}(2^{n-1}) = (U(2))^{\otimes n} , \tag{4.49}$$

which is exactly the group of local unitary operations on each qubit. This way, the MW quantifier is invariant under local unitary operations and is zero iff the state is separable, constituting an entanglement quantifier.

□

In 2003, Brennen (69) obtained an important equivalent form of the Meyer-Wallach entanglement quantifier in terms of the averaged linear entropy of the reduced states considering every possible bipartition one qubit-rest of the system. By considering the Schmidt decomposition of the state of n qubits for the bipartition $k-$th qubit and rest of the system together with the local unitary invariance of the wedge product, we can show that the purity of the reduced state of the $k-$th qubit, $\mathrm{Tr}(\rho_k^2)$, is directly connected to the value of the wedge product calculated in the Meyer-Wallach for the same state. This way, Brennen obtained the decomposition

$$\begin{aligned} Q(|\psi\rangle) &= \frac{4}{n}\sum_{k=1}^{n} D\left(|\tilde{u}_{\hat{k}}\rangle, |\tilde{v}_{\hat{k}}\rangle\right) \\ &= 2\left[1 - \frac{1}{n}\sum_{k=1}^{n}\mathrm{Tr}\left(\rho_k^2\right)\right] = \frac{2}{n}\sum_{k=1}^{n} S_L(\rho_k), \end{aligned} \tag{4.50}$$

which provides an understanding of what is the meaning of the wedge product in the Meyer-Wallach in terms of a well-known entanglement quantifier, the linear entropy. In other words, we can think that the decomposition allows the obtainment of this multipartite entanglement quantifier in terms of the average purity of the reduced states. To compute the multipartite entanglement of a system of n qubits, with state space of dimension $d = 2^n$, in general is necessary the complete knowledge of the state. This would require state tomography and a minimum of d^2 measurements for each sample. In terms of the averaged purity, this requires a number of measurements linear in n, revealing an improvement for multipartite entanglement estimation (69).

We can see that the form in Eq. (4.50) can estimate multipartite entanglement, because it takes into account the different values of entanglement between every possible bipartition in the state. However, the sizes of the bipartitions are fixed: We are only considering divisions one qubit-rest of the system. This is not a problem in the case of three qubits that cannot be addressed considering one bipartite quantifier solely, as the values of entanglement for the states in the different classes Eqs. (4.26), (4.27), (4.28), (4.29) will be different, Table 1, and we can, in fact, use this quantifier to discern the different entanglement classes for 3 qubits. However, if we increase the dimension of the system, the fact that only bipartitions one qubit-rest are being considered will lead to degeneracies for the entanglement values for very different states. Two examples are the four qubits GHZ state and a tensor product of two Bell states $|\phi^+, \phi^+\rangle$,

$$
\begin{aligned}
|GHZ_4\rangle &= \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle), \\
\left|\phi^+, \phi^+\right\rangle &= \frac{1}{2}(|00_{AB}\rangle + |11_{AB}\rangle) \otimes (|00_{CD}\rangle + |11_{CD}\rangle),
\end{aligned} \tag{4.51}
$$

whose values for the MW quantifier are both equal to 1, but considering different bipartitions will have different Schmidt numbers.

Table 1 – Meyer-Wallach entanglement for the 3 qubits entanglement classes

State class	Entanglement value
Separable	0
Biseparable	2/3
W	8/9
GHZ	1

Source: Adapted from MEYER; WALLACH. (79)

Despite this apparent barrier of the Meyer-Wallach being degenerated for very different entangled states, it is possible to generalize this quantifier, now written in a simple form in terms of the linear entropy, for different bipartitions sizes. This is exactly the generalization proposed by Scott (81), that includes even the cases of systems of n qudits. For the special case of qubits, we have the definition below.

Definition 7. Considering a pure state $|\psi\rangle$ of n qubits, $|\psi\rangle\langle\psi| \equiv \rho(n)$, the Scott entanglement quantifier of order m is defined as

$$
Q_m\left[\rho(n)\right] := \frac{2^m}{2^m - 1}\left[1 - \binom{n}{m}^{-1} \sum_{|S|=m} \mathrm{Tr}\left(\rho_S^2(n)\right)\right], \; m \leq \lfloor n/2 \rfloor, \tag{4.52}
$$

being S a subset of size m of $\{1, 2, \cdots, n\}$ and $\rho_S(n)$ the reduced state tracing out the complementary set, $S^{'}$, such that $S^{'} \cup S = \{1, 2, \cdots, n\}$. $\lfloor x \rfloor$ indicates the floor function, whose output is the greatest integer smaller or equal to x.

The value of m will define the sizes of the bipartitions considered and the inverse binomial coefficient appearing in the sum is to consider the average over every possible bipartition m qubits rest in a system of n qubits. When $m = 1$, the original Meyer-Wallach is obtained. Also, the constraint $m \leq \lfloor n/2 \rfloor$ is chosen so the normalization from 0 to 1 will always be valid, independent on the choice of n and m. Another reason for this is that bipartitions of $A-B$, where A has p qubits and B has q qubits will present the same purity values as if A had q qubits and B had p qubits, so there will be redundancy for $m > \lfloor n/2 \rfloor$. This is due to $\mathrm{Tr}\left(\rho_S^2(n)\right) = \mathrm{Tr}\left(\rho_{S'}^2(n)\right)$. A possible example of this is the simple case $n = 3$

and $m = 1$ or $m = 2$. The correlations in both values of m will be the same. Changing the values of m, we can exclude some degeneracies between the entanglement values for the different entanglement classes in different system sizes. The partial ordering of entangled states is not preserved under different values of m. For example, if we consider two states ψ and ϕ, $Q_m(\psi) \leq Q_m(\phi)$ does not imply that the same is true for $Q_{m'}(\psi) \leq Q_{m'}(\phi)$, $m \neq m^{'}$. This can be explained by the fact that the different values of m reflect different aspects of the global entanglement, with increasingly global nature (81). The quantifier presents maximum value 1 when the term inside brackets achieves value $(2^m - 1)/2^m$. This will happen for states whose reduced states are maximally mixed when tracing out $n - m$ qubits.

Another possible characterization of the Scott quantifier was obtained by Rigolin-Oliveira-Oliveira (82,83) and in this work will be used to perform the numerical methods and also analytical calculations. The form they proposed for a system of n qubits and bipartitions with $m = 2$ is written as

$$\begin{aligned} G(2,l) &:= \frac{4}{3}\left[1 - \frac{1}{n-l}\sum_{j=1}^{n-l} \mathrm{Tr}\left(\rho_{j,j+l}^2(n)\right)\right], \\ Q_2\left[\rho(n)\right] &:= \frac{2}{n(n-1)}\sum_{l=1}^{n-1}(n-l)G(2,l), \end{aligned} \tag{4.53}$$

which provides a way to calculate the linear entropies for every possible bipartition and take the mean value. The function G calculates the averaged purity considering the $l-$th possible neighbouring qubits. If $l = 1$, they are first neighbours, $l = 2$, second neighbours and so on. Q_2 will group all the possible G and weight the values. In this work we only considered the Scott quantifier with $m = 1$ (Meyer-Wallach) and $m = 2$, but the work in Ref. (83) proposes a general formula for every possible bipartition size m, which is the same as the quantifier proposed by Scott, but with the bipartitions made explicit.

4.2.2 Paradigmatic states for comparison

In this work, we are going to compare the mean entanglement generated by different circuits with some paradigmatic states, built based on the entanglement classes for 3 qubits and on the tensor product of Bell states if the number of qubits is even, which gives entanglement values equal to the GHZ case for 4 qubits. The GHZ and W states already have a generalisation to an arbitrary number n of qubits, given as

$$|GHZ_n\rangle := \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n}) \tag{4.54}$$

and

$$|W_n\rangle := \frac{1}{\sqrt{n}} \sum_{\pi} |\pi(00\cdots01)\rangle\,, \tag{4.55}$$

being π the n possible cyclic permutations.

The other states we are going to consider are what we are going to call the EPR_n states, which are in fact a subset of the biseparable states (84). For an even number of qubits, this state was already studied in many works (79, 81–83). In this work we propose the possibility for an odd number of qubits, which together with the even number is called EPR_n

$$|EPR_n\rangle = \begin{cases} |\Phi^+\rangle^{\otimes n/2}\,, \text{ if } n \text{ is even} \\ |0_1\rangle \otimes |\Phi^+\rangle^{\otimes (n-1)/2}\,, \text{ if } n \text{ is odd} \end{cases}. \tag{4.56}$$

The choice of which Bell state to consider is arbitrary and $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ could be substituted by any of $\{|\Phi^-\rangle\,, |\Psi^+\rangle\,, |\Psi^-\rangle\}$. Also, the choice of the state in the separable part in the odd n case is arbitrary and we chose $|0\rangle$ for simplicity.

The values of entanglement quantified using the Scott quantifier 1 and 2 on the arbitrary n case for the states EPR_n, n even, GHZ_n and W_n were already obtained in the literature and are presented in Table 2. Together, we included the values calculated in this work for the EPR_n, n odd, and discussed in App. B.

Table 2 – Scott quantifier of order 1 and 2 for the paradigmatic states of n qubits

State class	Scott 1 (MW)	Scott 2
GHZ_n	1	$2/3$
W_n	$4(n-1)/n^2$	$16(n-2)/(3n^2)$
EPR_{odd}	$(n-1)/n$	$(3n-5)/(3n)$
EPR_{even}	1	$(n-2)/(n-1)$

Source: Adapted from OLIVEIRA; RIGOLIN; OLIVEIRA (82) and by the author.

5 CIRCUIT DESCRIPTORS AND METHODS

The main objective of this work is the characterization of how connectivity of parameterized quantum circuits reflect in the generation of arbitrary states, as well as means and standard deviations of the generated multipartite entanglement. Given that there are well-defined and established descriptors that can be applied to parameterized quantum circuits, classifying their capability of generate entanglement or generate arbitrary quantum states (20, 69, 79, 81–83, 85), we can rely on these tools to reveal the characteristics of different quantum circuits architectures considering the possible connectivities. In this sense, this chapter presents the expressibility and entanglement quantification together with the method applied to estimate the sample size. We chose to omit the algorithms and codes, which will be available on GitHub[1]. To obtain the results, the Pennylane Python library was applied. Runtimes and specific details are presented in the codes.

5.1 Expressibility quantifiers

Given a quantum circuit architecture or, in the framework of variational quantum algorithms, an ansatz, natural questions that can come up are: How capable of generating arbitrary states in the space of states is this configuration? Also, how uniformly distributed the generated states are in this same space? These questions are important to the understanding of which architecture shall be applied in which specific context, given that some architectures may not access the portion of the state space where the objective state is, being not useful to the problem we are trying to solve.

We can work this characteristic of exploration of the state space defining a probability measure induced from metrics. Every metric can be used to define a volume element that, in compact spaces, can be normalized to induce a probability measure (41, 86). It will assign probabilistic weights to each of the regions of the state space. With that, we can build ensembles of states that are uniformly sampled over the space of states. In the case of the pure states space, the complex projective space $\mathbb{C}P^{d-1}$, where d is the dimension of the complex space related to the system[2], this can be done using the Fubini-Study metric: We build the volume element from the metric and then normalize it over the entire space of pure states, allowing the obtainment of a probability measure (86). But this can be cumbersome and also provides a new different problem that is how the generated ensemble can be related to the states generated by the particular analyzed circuit.

An interesting intervention is to introduce, instead of the probability measure over the space of states, the probability measure over the space of unitaries that act on the

[1] Author's profile: https://github.com/GICorrer.

[2] For example, n qubits states are related to $\mathbb{C}^{2^n}$.

states. Given a compact Lie group $\mathfrak{G}$, we can say that there is a unique invariant integration measure $d\mu(g)$ in the manifold of $\mathfrak{G}$, the Haar measure (44), discussed in Chap. 3. This measure is not only invariant by the action of the group, but also can be normalized for compact spaces and induces a probability measure (41). Denoting the unitary group of dimension d as $\mathcal{U}(d)$, we can choose a fiducial pure state, $|\psi_0\rangle \in \mathbb{C}P^{d-1}$, and unitaries in $\mathcal{U}(d)$ sampled accordingly to the Haar induced probability measure, $U^i(d), i = 0, 1, ...$ This ensemble of unitaries is uniformly distributed in $\mathcal{U}(d)$, as the probability measure induced by the Haar measure will be the uniform one (41). An induced ensemble of uniformly distributed pure states is then obtained by acting the unitaries over the fiducial state: $U^i(d)\,|\psi_0\rangle\,, i = 0, 1, ...$ In fact, for pure states, the induced volume element via this method coincides with the induced by the Fubini-Study metric, manifesting the equivalence between the approaches (87).

Now, this latter proposed method gives some intuition into how unitaries can induce ensembles of states, but it does not answer the question about how well a given circuit can generate arbitrary uniformly distributed states. To answer these questions, we can, inversely, start with the circuit that would perform this task with perfection and then compare with the circuit available. We already said that the Haar induced measure over the space of states samples uniformly, therefore, the most tailored circuit in this scenario should be one that replicates the behavior of the Haar ensemble. A way to quantify how well a set of unitaries (in other words, our circuit ansatz) can mimetize a Haar ensemble up to the $t-$th moment is defining the estimator for n qubits states (20)

$$\mathcal{A}^t := \int_{Haar} (|\psi\rangle\langle\psi|)^{\otimes t} d\,|\psi\rangle - \int_{\Theta} (|\phi_\theta\rangle\langle\phi_\theta|)^{\otimes t} d\theta, \tag{5.1}$$

where the first integral is obtained by uniformly sampling with the Haar measure over the n qubits state space. The second integral is given by the circuit we desire to characterize, with the sampling performed considering a uniform distribution in the cartesian space of parameters $\Theta = \{\vec{\theta}\}$, where the parameter vectors of the circuit are defined. The choice of uniformly distributed parameters is justified by a lack of bias when analysing the possibilities, so they are equally probable. This way, given an architecture, the space of parameters is uniformly sampled to produce states $|\phi_\theta\rangle$ and an approximation of the state $t-$design. This approximation is then compared to the ideal case obtained via the Haar integral. The Hilbert-Schmidt distance between these states is calculated using the 2-norm (55)

$$||\mathcal{A}^t||_2^2 := \mathrm{Tr}\Big[(\mathcal{A}^t)^\dagger(\mathcal{A}^t)\Big], \tag{5.2}$$

and, the closer this quantity is to zero, the closer the circuit generated $t-$state is to the uniform one. The Haar integral will be decomposed in terms of the projector into the symmetric subspace as (App. A)

$$\int_{Haar} (|\psi\rangle \langle\psi|)^{\otimes t} d|\psi\rangle = \frac{\Pi^t_{symm}}{\dim(\vee^t \mathbb{C}^{2^n})}, \tag{5.3}$$

being Π^t_{symm} the projector into the symmetric subspace of the system of t spaces of n qubits with dimension $\dim(\vee^t \mathbb{C}^{2^n})$ (App. A). If we define the other integral as $\mathcal{I}_t$, then

$$\begin{aligned} ||\mathcal{A}^t||_2^2 &= \mathrm{Tr}\left[\left(\frac{\Pi^t_{symm}}{\dim(\vee^t \mathbb{C}^{2^n})}\right)^2\right] - \frac{2}{\dim(\vee^t \mathbb{C}^{2^n})} \mathrm{Tr}\Big(\Pi^t_{symm} \mathcal{I}_t\Big) + \mathrm{Tr}\Big(\mathcal{I}_t^2\Big) \\ &= -\frac{1}{\dim(\vee^t \mathbb{C}^{2^n})} + \mathrm{Tr}\Big(\mathcal{I}_t^2\Big). \end{aligned} \tag{5.4}$$

The last term can be explicitly obtained in terms of the integral as

$$\begin{aligned} \mathcal{I}_t^2 &= \int_\Theta \int_\Phi \langle\psi_\theta|\psi_\phi\rangle^t (|\psi_\theta\rangle \langle\psi_\phi|)^{\otimes t} d\theta d\phi \\ \Rightarrow \mathrm{Tr}\Big(\mathcal{I}_t^2\Big) &= \int_\Theta \int_\Phi |(\langle\psi_\theta|\psi_\phi\rangle)^{2t}| d\theta d\phi. \end{aligned} \tag{5.5}$$

Now, we can introduce the state frame potential, which generalizes the notion of frame potential introduced in Sec. 3.2 from ensembles of unitaries to ensembles of states. The state frame potential will be defined as (43)

$$F^{(t)}_\varepsilon(\rho) := \iint_\varepsilon \Big\{\mathrm{Tr}\Big[\rho U V^\dagger\Big] \mathrm{Tr}\Big[\rho V U^\dagger\Big]\Big\}^t dU dV, \tag{5.6}$$

where we are considering t copies of the state ρ. This definition preserves the good properties we already had in Chap. 3. In particular, it is minimized by the Haar ensemble (43). Therefore, the state frame potential quantifies how different our ensemble of states obtained by the action of the ensemble ε on ρ is from one generated considering a Haar ensemble of unitaries acting on ρ.

Let us apply this definition for the specific case where the fiducial state ρ is equal to $|0\rangle^{\otimes n}$, i.e., all the n qubits are in the computational basis state $|0\rangle$. This will give

$$\begin{aligned} F^{(t)}_\varepsilon(|0\rangle\langle 0|^{\otimes n}) &= \iint_\varepsilon \{\mathrm{Tr}\Big[|0\rangle \langle 0|^{\otimes n} U V^\dagger\Big] \mathrm{Tr}\Big[|0\rangle \langle 0|^{\otimes n} V U^\dagger\Big]\}^t dU dV \\ &= \iint_\varepsilon dU dV |\langle 0|^{\otimes n} U V^\dagger |0\rangle^{\otimes n}|^{2t}. \end{aligned} \tag{5.7}$$

The states $U|0\rangle$ and $V^\dagger|0\rangle$ are obtained by the action of sampled unitaries considering the ensemble ε on a fiducial state $|0\rangle$. From the discussion above, the Haar induced measure in the space of pure states is the same as the measure over states induced by the Fubini-Study metric (87). This way, we can use this same notion to the ε ensemble and rewrite

$|\psi\rangle \equiv U|0\rangle$, $|\phi\rangle \equiv V^\dagger|0\rangle$ and define $d|\psi\rangle, d|\phi\rangle$ as induced integration measures over the space of states. The frame potential relative to the state $|0\rangle^{\otimes n}$ with ensemble ε is written as

$$F_\varepsilon^{(t)}(|0\rangle\langle 0|^{\otimes n}) = \iint_\varepsilon d|\psi\rangle\, d|\phi\rangle\, |\langle\psi|\phi\rangle|^{2t}. \tag{5.8}$$

If we look closely to the trace of the square of the integral that appeared in the definition of $||\mathcal{A}^t||$, Eq. (5.5), we can see that the quantities are exactly the same when defining the ε ensemble to be the one generated by the circuit when randomly sampling the parameter space.

This frame potential for the state $|0\rangle\langle 0|^{\otimes n}$ is only minimized when the ε ensemble is a $t-$design or, analogously, if the sampling is drawn from the Haar distributed states. Notice that

$$\begin{aligned} \int_{Haar} (|\psi\rangle\langle\psi|)^{\otimes t} d\psi &= \frac{\Pi^t_{symm}}{\dim(\vee^t \mathbb{C}^{2^n})} \\ \Rightarrow \iint_{Haar} d|\psi\rangle\, d|\phi\rangle\, |\langle\psi|\phi\rangle|^{2t} &= \frac{[\mathrm{Tr}\left(\Pi^t_{symm}\right)^2]}{\dim(\vee^t \mathbb{C}^{2^n})^2} = \frac{1}{\dim(\vee^t \mathbb{C}^{2^n})}, \end{aligned} \tag{5.9}$$

where we simply took the trace of the square of the integral. This way, the state frame potential for the Haar ensemble is exactly the inverse of the dimension of the symmetric subspace and the estimator, Eq. (5.4), will be finally written as

$$||\mathcal{A}^t||_2^2 = F_\Theta^{(t)}(|0\rangle\langle 0|^{\otimes n}) - F_{Haar}^{(t)}. \tag{5.10}$$

To simplify the notation, we wrote the Haar frame potential of the $|0\rangle\langle 0|^{\otimes n}$ state simply as $F_{Haar}^{(t)}$, being the other related to the circuit induced ensemble under study. Now, the reason why they were introduced is clear. The quantity $||\mathcal{A}^t||_2^2$ for a $t-$design can be estimated considering the Haar frame potentials and the induced by the circuit considering uniform sampling of the parameter vectors. As this strategy does not provide a unique, well-defined specific estimator and will depend on the design order, we are going to follow it with a different, but analogous, approach[3]. The frame potential can be rewritten as the expected value of the powers of the fidelity comparing two random states. This is visualized in (20)

$$F_\Theta^{(t)}(|0\rangle\langle 0|^{\otimes n}) = \int_\Theta \int_\Phi |\langle\psi_\theta|\psi_\phi\rangle|^{2t} d\theta d\phi = \mathbb{E}_\theta \mathbb{E}_\phi[(|\langle\psi_\theta|\psi_\phi\rangle|^2)^t] = \mathbb{E}[\mathcal{F}(\theta,\phi)^t], \tag{5.11}$$

[3] In Chap. 6 we also test the characteristic of a 2-design for the circuits using Eq. (5.10) to clarify some of the results.

being $\mathcal{F}(\theta,\phi) := |\langle\psi_\theta|\psi_\phi\rangle|^2$ the Uhlmann fidelity for pure states (88). In words, what this says is that the frame potentials of the circuits and the Haar ones are directly connected to the $t-$moments of the distribution of the fidelities. In the circuit case, the fidelities distribution is obtained as a function of parameter vectors, which are uniformly sampled. For the Haar case, the probability density function of the fidelities is obtained considering states that are uniformly sampled using the Fubini-Study measure; in a Hilbert space of dimension d (89),

$$p_{Haar}(\mathcal{F}) = (d-1)(1-\mathcal{F})^{d-2}. \tag{5.12}$$

This way, instead of comparing frame potentials one by one, we can compare directly the fidelities distribution obtained considering quantum states produced by the circuit under investigation and the obtained considering Haar random states. To compare the distributions, the Kullback-Leibler divergence or relative entropy will be used. Despite the fact of not being a distance in the space of probability distributions[4], this is an important quantity to verify how different two probability distributions are from each other and also has an interesting interpretation. We are going to define (20, 55)

$$D_{KL}(P||Q) := \begin{cases} \sum_i P(i)\log\left[\frac{P(i)}{Q(i)}\right], \text{ if supp}(P)\subseteq\text{supp}(Q) \\ +\infty, \text{ else} \end{cases}, \tag{5.13}$$

where $\text{supp}(f)$ is the support of f, i.e., a subset of the domain of f for which the function is different from zero. Therefore, if the distributions have different supports, the relative entropy is saturated[5]. This quantity is measured in nats. The closer D_{KL} calculated between the circuit fidelities distribution and the Haar fidelities distribution, $D_{KL}(p_{circ}(\mathcal{F})||p_{Haar}(\mathcal{F}))$, is to zero, the closer the circuit generated states are to the uniformly distributed ones. When this analysis was introduced (20), the authors said that the closer this is to zero, the higher is the *expressibility* of the circuits. So, the decrease of D_{KL} implies in the increase of the expressibility (20). Some authors chose to introduce an expressibility quantifier (instead of only talking about expressibility in the qualitative sense) that is inversely proportional to the D_{KL}, so there is no confusion (91). Here we will apply the original nomenclature and we are going to look only to the values of D_{KL}.

The relative entropy can be interpreted in terms of a difference between two entropies. Consider that a set of events indexed with i occur with probabilities $P = \{P(i)\}_{i=1}^n$. If we mistakenly assumed that the distribution is given by $Q = \{Q(i)\}_{i=1}^n$,

[4] It does not satisfy the symmetry and triangle inequality requirements (55).

[5] In fact, this is a recurrent phenomenon also in distinguishability quantifiers and distances for the space of states. The trace distance, the Uhlmann fidelity and the Bures distance are examples for which when the support of two operators are orthogonal, the distance is saturated (90).

the predicted averaged uncertainty would be $-\sum_i P(i)\log[Q(i)]$, but the real value, the Shannon entropy, is in fact $-\sum_i P(i)\log[P(i)]$. Calculating the difference between these two quantities we find

$$-\sum_i P(i)\log[Q(i)] - \{-\sum_i P(i)\log[P(i)]\} = \sum_i P(i)\log\left[\frac{P(i)}{Q(i)}\right], \tag{5.14}$$

which is the definition of relative entropy when the support of P is contained in the support of Q. In this line of thought, the quantity measures the differences one would have in protocols that depend on the real predictions of distributions associated to the set of events, in this case, $\{P(i)\}_{i=1}^{n}$.

Within this framework, the numerical estimation of expressibilities of circuits can be practically implemented using classical simulators of quantum circuits: Given the architecture, an ensemble of uniformly distributed parameter vectors is generated and given as an entry to the circuit unitaries. Every pair of circuit output states are then compared using the fidelity, giving the values $\mathcal{F}(\theta,\phi) = |\langle\psi_\theta|\psi_\phi\rangle|^2$. A histogram is built with the values and compared with the histogram obtained with the Haar fidelities, Eq. (5.12), using the relative entropy. The closer the quantity is to zero, the closer we are to the uniform distribution of states and we say the greater is the expressibility of the circuit. A pictorial view of the process of inducing a distribution of states considering the parameter vectors given as an input to circuits is illustrated in Fig. 3. Before exploring the sample size determination and the explicit results, we are going to build a discussion about the entanglement quantifier.

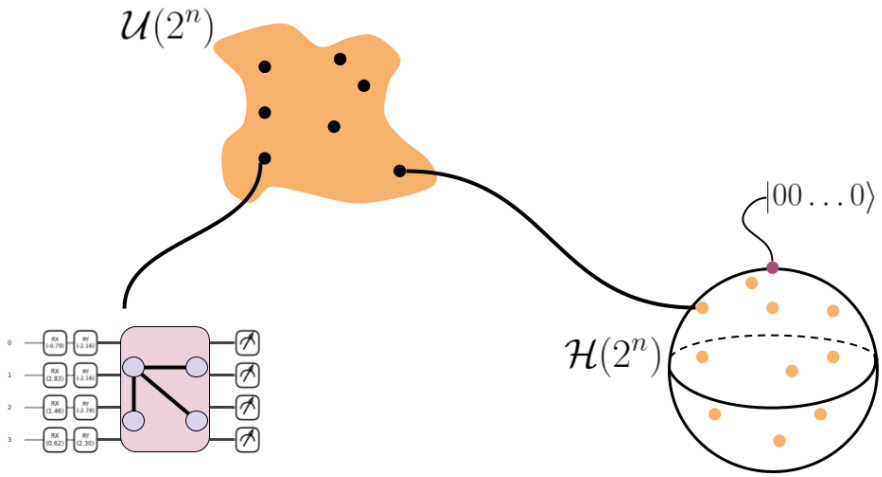

Figure 3 – Pictorial representation of the circuit induced measure over the space of states. Changing the parameters of the circuit will generate different unitaries, inducing a distribution of points on the space of the unitary group in dimension 2^n, $\mathcal{U}(2^n)$ (black dots). These unitaries can be applied to a fiducial state $|0\rangle^{\otimes n}$ (purple dot) in order to induce a distribution of states on $\mathcal{H}(2^n)$ (orange dots).

Source: By the author.

5.2 Entanglement quantifiers

The entanglement generated by parameterized quantum circuits has a strong dependency on the parameters and on the input states. Here, we consider as input state the zero tensor state of n qubits, $|0\rangle^{\otimes n}$, so we can see the effects of creating coherences and entangling with CNOT gates. To understand the role of the structure of the circuit, we average the entanglement quantifier over an ensemble of parameter vectors. Again, as in the expressibility quantifier, the parameters are chosen uniformly at random considering independent distributions for each one of them. This quantifier was proposed in Ref. (20) and is closely connected to the entangling power of quantum evolutions (81, 92) that is, conceptually, the mean entanglement generation of a unitary operation over all the possible input separable states. We will see in Sec. 6.1 that the circuits of our work have an initial parameterized separable state preparation part, followed by connections using CNOT gates. This way, averaging over many possible separable states, we analyse the entangling power of the connections in a restricted separable states set by calculating the mean values.

We are going to consider two entanglement quantifiers to estimate the mean multipartite entanglement of the circuits, the Meyer-Wallach/Scott 1 (S1), Q_1, and the Scott 2 (S2), Q_2, in the analyses. As discussed in Chap. 4, the Meyer-Wallach quantifier presents degeneracies for the entanglement values of very different states, so instead of considering only their values, as done in the work of Ref. (20), we considered the next order Scott quantifier. The quantities we calculated are

$$\langle Q_m \rangle_\Theta = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} Q_m \left(\rho(\vec{\theta}_i) \right), \; m = \{1, 2\}. \tag{5.15}$$

The $\langle \cdot \rangle_\Theta$ symbol represents the average over the parameters space, $\vec{\theta}_i$ is a specific vector of input parameters and $\mathcal{N}$ is the sample size. In the absence of Θ, it should be understood that we are considering the circuit mean entanglement. We also calculated the standard deviations, considering the same procedure, a quantity not well explored in the context of parameterized quantum circuits.

The ensemble of unitary operations sampled considering the uniform Haar measure over the unitary group manifold is often called Circular Unitary Ensemble (CUE) (41). Using this reference sampling, we can calculate average and standard deviation values of entanglement for uniformly distributed pure random states, which are used in our work as the ideal case. The quantities obtained in the literature, for n qubits (81), are

$$\langle Q_m \rangle_{CUE} = \frac{2^n - 2^m}{2^n + 1}, \tag{5.16}$$

and, considering only the Meyer-Wallach (S1), the obtained result is

$$\sigma^2_{CUE}(Q_1) = \frac{6(2^n-4)}{(2^n+3)(2^n+2)(2^n+1)n} + \frac{18\cdot 2^n}{(2^n+3)(2^n+2)(2^n+1)^2}, \tag{5.17}$$

being $\sigma^2_{CUE}(Q_1)$ the variance of S1. In our analysis we applied the standard deviation by taking the square root of the previous equation. Only the standard deviation of the Q_1 was considered, by the fact that the entanglements quantified by Q_1 and Q_2 have very similar properties in the sense we are interested in this comparison with the CUE, as it will be discussed in Chap. 6. The Meyer-Wallach entanglement quantifier is based on the linear entropy (Chap. 4), which is a $(2,2)$ polynomial (40) and requires at least a $2-$design to achieve mean values equivalent to the obtained using the CUE (Def. 2). This way, the circuits are going to replicate the mean entanglement of the CUE when they are $2-$designs and it is possible to test this using frame potentials, as discussed in the previous section. We should be aware that this is a sufficient, but not necessary, condition, so circuits generating mean entanglement close to CUE are not necessarily $2-$designs.

To compare the mean entanglement generated by the circuit with the mean entanglement of the CUE, we propose the normalized quantifier

$$\frac{\langle Q_1\rangle_{CUE} - \langle Q_1\rangle}{\langle Q_1\rangle_{CUE}}. \tag{5.18}$$

This is going to be applied in the discussion of how entanglement and expressibility are connected by the generation of random entanglement, in Chap. 6.

The last discussion performed applying the entanglement quantifiers is the comparison of circuits means and standard deviations with the entanglement of paradigmatic states, as presented in Table 2. An inspection is performed to understand how the characteristics of circuits entanglement can be traced to the paradigmatic states proposed in Subsec. 4.2.2.

5.3 Sample size

To determine the sample size, we considered applying the Chebyshev inequality (93) to the mean estimator. First, the Chebyshev's inequality is a result based on Markov's inequality stating that, if X is a random variable with finite mean μ and variance σ^2, for every real value $k > 0$ is valid that

$$P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}. \tag{5.19}$$

In words, the probability that the modulus of the difference between a realization of the random variable X and its mean μ is greater or equal than k is smaller or equal than the quotient of the variance of X, σ^2, and k^2.

We can choose this random variable to be the sum of N independent and identically distributed random variables, X_i, each with mean μ and variance σ^2, divided by N. Here, N is the size of the sampling and the total random variable is the mean estimator. The expected value and variance of this random variable are

$$\mathbb{E}\left(\frac{\sum_{i=1}^{N} X_i}{N}\right) = \mu \tag{5.20}$$

and

$$\text{var}\left(\frac{\sum_{i=1}^{N} X_i}{N}\right) = \frac{\sigma^2}{N}. \tag{5.21}$$

Using these results in the Chebyshev inequality,

$$P\left(\left|\frac{\sum_{i=1}^{N} X_i}{N} - \mu\right| \geq k\right) \leq \frac{\sigma^2}{Nk^2}. \tag{5.22}$$

Now, using that $P(Y \geq \kappa) + P(Y < \kappa) = 1$ then, if $P(Y \geq \kappa) \leq \xi$, we have that $P(Y < \kappa) \geq 1 - \xi$. Substituting this relation with the variables presented in Eq. (5.22),

$$P\left(\left|\frac{\sum_{i=1}^{N} X_i}{N} - \mu\right| < k\right) \geq 1 - \frac{\sigma^2}{Nk^2}. \tag{5.23}$$

This inequality gives a way to choose the sample size, N, in a manner that the probability that the mean estimator is different from the true mean is bigger than a certain value given in terms of a positive constant, the variance and the sample size. One option for the positive constant k is a positive multiple of the standard deviation, which is always greater than zero for non-deterministic random variables: $k \equiv c\sigma$, implying in

$$P\left(\left|\frac{\sum_{i=1}^{N} X_i}{N} - \mu\right| < c\sigma\right) \geq 1 - \frac{1}{Nc^2}. \tag{5.24}$$

Therefore, choosing the range c and the confidence probability that the difference from the mean is within c of the standard deviation, the sample size is determined. In our work, we chose the confidence probability to be 98% and the range to be 0.1 of the standard deviation. This implies that the sample size should be $N = 10^4$, which we applied for the estimation of the entanglement means and standard deviations. This way, we generate 10^4 input parameter vectors for the circuits, which generate 10^4 entangled states that are used for estimation. In the case of expressibilities, the same sample was applied, but as it is necessary two states for each computation of the fidelity ($\mathcal{F}(\theta, \phi) := |\langle\psi_\theta|\psi_\phi\rangle|^2$), the effective sample was $5 \cdot 10^3$. In this case we did not compute the standard deviation. Discussions in the supplementary material of (20) show that using this sample size is enough so the standard deviations be within 10^{-4}. Finally, for the design estimation, presented in Table 4 of Chap. 6, we considered a sample of the same size as expressibility, $5 \cdot 10^3$, and 20 executions for the mean and standard deviation calculations.

6 RESULTS

In this chapter, we discuss the main contributions of this thesis, including the choices made for the quantum hardware structure considered in this work and how they affect the expressibility and entanglement quantifiers, discussed in Chap. 5, of parameterized quantum circuits. We are going to define the different graphs related to the connections that are going to be performed considering two qubits gates and then study how expressibility, the Scott/Meyer-Wallach quantifier and frame potentials can give important conclusions about the randomness and characteristics of the circuits.

6.1 Circuits choices

6.1.1 Qubits connectivities

As discussed in Chap. 2, VQA circuits are mostly based on heuristics about the problem to be solved. The structure of the connections between qubits and the local gates performed can be chosen inspired by established methods, such as the Quantum Alternating Operator Ansatz and adiabatic quantum computing (24) or Unitary Coupled Cluster Ansatz and chemistry Hartree-Fock methods (25). It can also be based on the optimization of circuit gate structure (ADAPT-VQE) (94) and many other possible heuristics. These are very plausible approaches and are in consonance with most of the methods applied in classical machine learning, where the outputs accomplish the task, even though the reason why the algorithm does it is generally not clear (95).

In our work, we decided to follow a different approach. Instead of considering heuristics and problem inspiration for the circuit structure, we considered the possible topology of connections that can be performed between qubits in currently available quantum platforms and then studied the characteristics of the circuits using the expressibility and mean entanglement figures of merit defined in Chap. 5. This kind of reasoning is similar to the one in Hardware Efficient Ansätze (12), where the gates used are chosen in order to minimize circuit depth depending on the experimental setup (33, 34), usually assumed to support only nearest neighbours interactions due to the feasible implementability (33). Still, our choice is promising in the sense that it takes into account a broad set of quantum platforms connectivities and considers controlled-NOT gates as couplings, comparing the different couplings using quantifiers. This constraint on the connections between qubits has been analysed in different quantum computing topics, including pseudorandom quantum circuits (40), fault tolerant quantum algorithms (96), analysis of superconducting hardware (97) and finally VQA and PQC applications (98–100). Together with the figures of merit, the benchmark of PQCs with different connectivities will provide important data that can be used to choose the architecture based on the necessities one may have

for the problem they are trying to solve.

The different connectivities, also called topologies, are illustrated as graphs in Fig. 4 for 4 qubits, together with their circuit representation of connections using CNOT gates. The way we presented the possible connections as graphs does not rely upon the lattice of the quantum computer, that can be, for instance, square, hexagonal and others (97, 99, 100), even though they are shown as a square lattice in the figure. If these connections can or cannot be implemented for a specific number of qubits in an architecture that has a particular lattice structure, will depend. For example, if only first neighbours connections are considered in a square lattice (vertical and horizontal edges in Fig. 4), the all-to-all topology would not be available without the need of additional gates other than CNOTs. Therefore, we are analysing the role of connection types without the constraint of choosing a specific quantum hardware. We chose to present 4 qubits as it is the minimal number of qubits such that the topologies are not degenerated (linear/star or ring/all-to-all are the same for 3 qubits). When we consider the 3 qubits system (App. C), linear/star will be identified as "linear" and ring/all-to-all will be identified as "ring".

The *no connections* case is the only one that can be implemented in any choice of hardware independent of the number of qubits, as it does not require two qubits gates and only local operations are considered. The other connectivities will be available depending on the particular hardware. *Linear* or nearest neighbour connectivity can be implemented in many quantum platforms, as IBM (Ithaca (98), Vigo (101), Tokyo (100)) or Rigetti (99) (which has two coupled octagon lattices) quantum computers. The other possible connectivities, *ring* and *star*, are available in the same quantum computers depending on the number of qubits desired. Finally, the *all-to-all* is one of the hardest to construct, as it requires connections between every pair of qubits. Still, ion trap quantum computers are able to perform this connectivity (102).

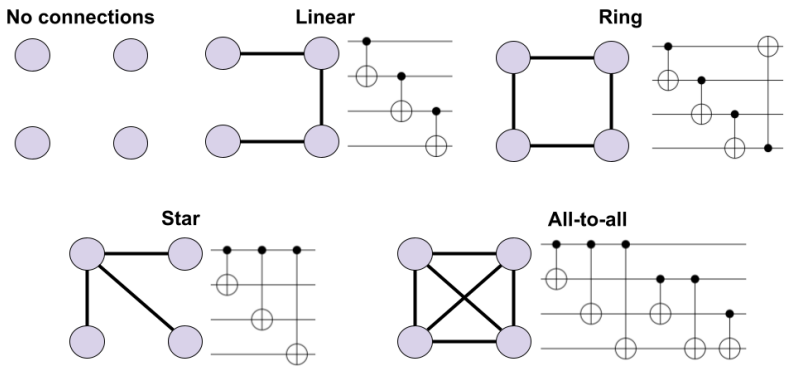

Figure 4 – Graphs of connectivities between qubits and their quantum circuit representation considering CNOT gates.

Source: By the author.

In fact, any of the possible connections could be implemented in different lattices by the use of SWAP gates together with entangling gates. Nevertheless, this would have an additional cost of many gates for each of the couplings and here we are only concerned about how to rely on the architecture to harness the best of the connectivity.

The choice of CNOT for connections is arbitrary, however there is no loss (we can generate multipartite entanglement in both cases) in comparison to other options as CZ gates, even though the results can present differences. We chose CNOT because it is native[1] to many quantum platforms (mainly the IBM quantum computers) and it is not a parameterized gate, so we can isolate contributions related only to connections and the ones related only to parameters, which are going to appear in local operations.

6.1.2 Structure of local parameterized gates

Even though we are considering a general framework for the connections that can be performed between qubits using the different hardware, we have to choose a particular parameterization of the circuits. In this work, the parameters are introduced only in local operations performed on each qubit. To understand how the parameterization can influence the generation of states and the characteristics of the circuits, we divide the PQCs in two classes, called Ansatz 1 and Ansatz 2. For both of them, the sequence of local unitary operations will be RX (rotation in the X direction) followed by RY (rotation in the Y direction) for each qubit. This choice is due to the fact that the sequence of operations will explore the space of one qubit, with a bias around the poles of the Bloch sphere (20), without the need of all the 3 parameters necessary to perform a uniformly distributed unitary (in the group manifold) of one qubit (103).

The ansätze are shown in Fig. 5. Ansatz 1 consists of the sequence of local unitary operations in each qubit, RX-RY, followed by the connections of the specific topology, as illustrated in Fig. 4. These circuits will generate different states considering the possible parameter vectors with each parameter of the local operations uniformly distributed, as discussed in Chap. 5. Ansatz 2 has the same structure as Ansatz 1 up to the two qubit gates part and, after that, the connections are followed by another step of local unitary operations, again with gates RX-RY. The local unitary operations after the connections will not change the entanglement created by the circuit and, therefore, both ansätze will have the same entanglement (considering entanglement quantification or classes of entanglement). Consequently, Ansatz 2 will have same entanglement characteristics, together with more freedom on the parameterization and on building states with different coherences. Comparing both ansätze, it will be possible to isolate what are the effects of entanglement generation and, together with this additional freedom on the local operations, how

[1] The native gates are the quantum gates physically implemented in the quantum hardware. When a gate is not in the native set, it can be decomposed in terms of the native ones.

they affect the task of building different states. These are the structures of the circuits considering only one layer. We must remember that, for each layer, this structure will be concatenated. At first, one may think that the entanglement properties of the circuits are different for more than 1 layer, as more entangling gates will be performed after the additional local operations. In the next sections this is shown to be not true.

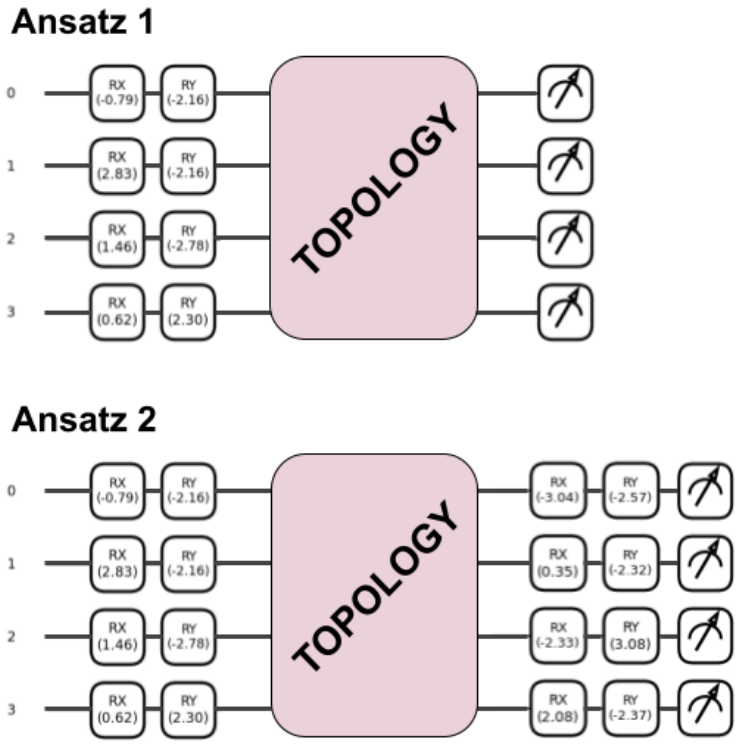

Figure 5 – One layer of the proposed circuits structure, called Ansatz 1 and Ansatz 2. Both ansätze have the same structure up to the connections. In Ansatz 2, the connections are also followed by local parameterized operations. The parameters shown are illustrations of the uniform distribution sampling between $-\pi$ and π.

Source: By the author.

Considering the connections of the graphs and the local gates structure of each ansatz, we can quantify the number of two qubit gates (CNOTs), local gates/parameters and the total number of operations for each ansatz and topology, depending on the number of qubits and layers. Table 3 presents these numbers for Ansatz 1 and Ansatz 2. The number of CNOTs is the same for both ansätze. Also, as each of the ansätze will have the same number of parameters and local operations comparing topologies, we only presented the number of CNOT gates together with the total number of gates. The number of operations will scale linearly as a function of the number of layers.

The table illustrates what we were already expecting: The all-to-all connectivity has an intense scale as a function of the number of qubits, increasing quadratically. The other connected topologies have a very similar increase on the number of CNOT gates, varying linearly with n, even though the linear and star require one less gate for each

layer. For a great number of qubits, therefore, they are very close in terms of cost of gates number. Nevertheless, we still would have to consider how hard it is to perform the connections. The linear has, in general, the simplest structure (only nearest neighbours couplings), while the ring and star can be more difficult to implement. Still, they cannot be strictly compared without choosing a particular lattice of qubits. For example, in a square lattice as presented in Fig. 4 the cost of implementing the linear and ring topologies depends only on the number of gates, as every connection is performed between neighbouring qubits.

Table 3 – Number of CNOT gates and total number of gates comparing topologies for Ansatz 1 and 2, as a function of the number of qubits n and number of layers l.

Topology	Number of CNOTs	Total number Ansatz 1	Total number Ansatz 2
No connections	0	$(2n)l$	$(4n)l$
Linear	$(n-1)l$	$[(n-1)+2n]\,l$	$[(n-1)+4n]\,l$
Ring	nl	$[n+2n]\,l$	$[n+4n]\,l$
Star	$(n-1)l$	$[(n-1)+2n]\,l$	$[(n-1)+4n]\,l$
All-to-all	$\frac{n(n-1)}{2}l$	$\left[\frac{n(n-1)}{2}+2n\right]l$	$\left[\frac{n(n-1)}{2}+4n\right]l$

Source: By the author.

6.2 **Numerical simulations:** the role of local parameterization

In this section we present the numerical results of expressibility and entanglement for both ansätze, changing the number of qubits and number of layers. The discussion will start with expressibility and entanglement, using Meyer-Wallach/Scott 1, for Ansatz 1 and then observe the relations with Ansatz 2. We are going to see that the entanglements for Ansatz 1 and 2 have similar behavior, so this way of presenting the results will be more beneficial for the reasoning. The graphs and tables presented here are a selection of the results and any additional data (for example, relative entropy for a fixed number of qubits not presented in this section) can be found in App. C.

6.2.1 Ansatz 1

6.2.1.1 Expressibility

We begin with the expressibility considering fixed numbers of qubits and changing the number of layers (circuit concatenations). Fig. 6 presents the behavior of Ansatz 1 considering only 4 and 8 qubits for all topologies, with a linear y-axis scale and Fig. 7 presents 4, 6, 7 and 8 qubits for connected topologies, considering a logarithmic y-axis scale. This division is done to understand first the behavior of the no connections case

together with the other topologies (Fig. 6) and then to do a detailed study of the connected topologies (Fig. 7).

From Fig. 6, we can see that the no connections case saturates as a function of the number of layers, an effect that is independent of the dimension of the system for the studied numbers of qubits. This is an expected result for spaces with entangled states (any $n > 1$). The input state is separable and the no connections circuit does not generate entanglement, so only the separable region of state space is explored. As an arbitrary random state sampled from the uniform probability measure will have, on average, nonzero entanglement (81), it will never be possible to achieve values for the relative entropy closer to zero. The saturation values vary and we observed them to be between 0.20 and 0.25 nats. The saturation at this value is not observed for the connected topologies and they evolve to values closer to zero as the number of circuit concatenations increase.

Another interesting characteristic is that, at only 1 layer, there is not much difference between the no connections circuit and the connected topologies. Even though these other circuits generate entanglement, the freedom on building states with different coherences is not enough for them to be closer to the uniform random case than the no connections topology if we apply them only once. This effect is observed at any studied dimension. We already can see that, for the dimension of 4 qubits (Fig. 6 (a)), the connected topologies evolve in a very similar manner to values close to 0. For the higher dimension of 8 qubits (Fig. 6 (b)), a hierarchy between the different topologies starts to build up.

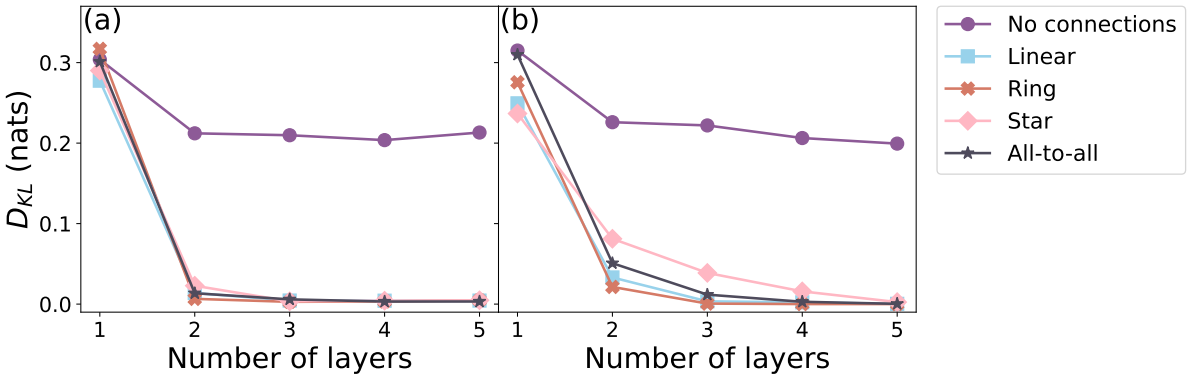

Figure 6 – Ansatz 1: relative entropy as a function of the number of layers for all connectivities, considering the case of (a) 4 qubits and (b) 8 qubits.

Source: By the author.

Now, we focus on the differences between the connected topologies as a function of the number of layers looking at Fig. 7. Increasing the number of layers, the circuits can get closer to the uniform random case and this possibility of evolution is independent on the topology, even though there are differences on how steep is this decrease of D_{KL}. The

behavior implies that the topology of the quantum computer does not matter too much if we take a naive approach and increase the depth of the quantum circuit unrestrictedly, implying in error accumulation and also requiring a greater control of the hardware for real computers due to the excess of operations. This behavior is not observed in the range of parameters studied for the 4 qubits case, Fig. 7 (a), as a saturation is seen to occur at $5 \cdot 10^{-3}$ nats. Similarly, for 3 qubits (App. C) a saturation occurs at the same value.

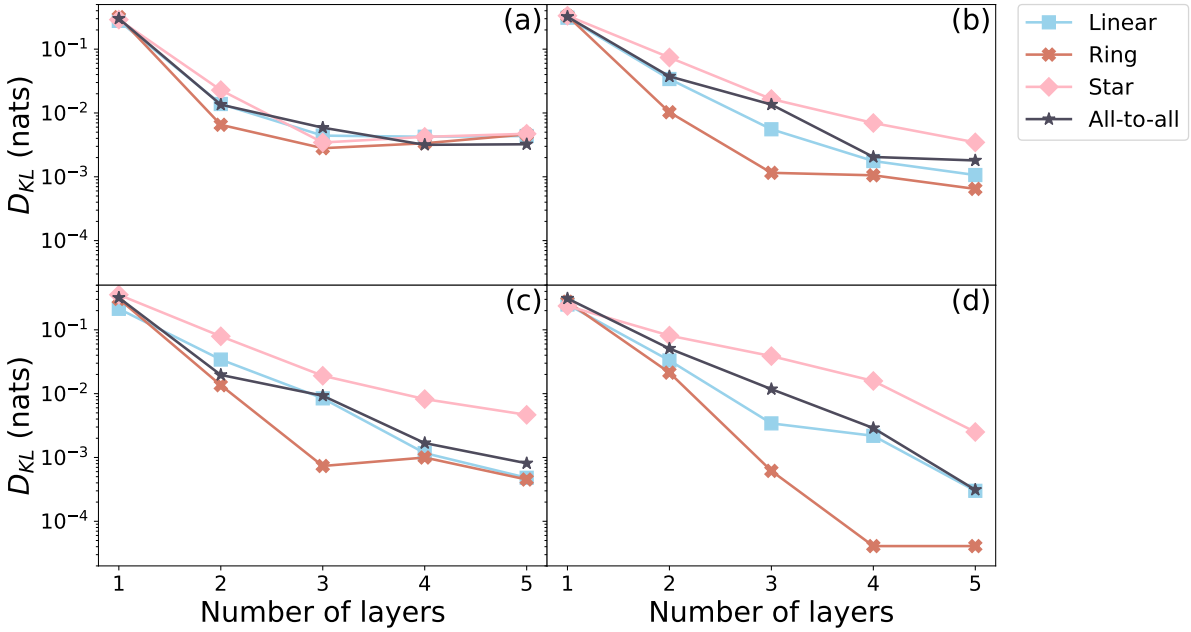

Figure 7 – Ansatz 1: relative entropy in logarithmic scale as a function of the number of layers for all connected connectivities, considering the case of (a) 4 qubits, (b) 6 qubits, (c) 7 qubits and (d) 8 qubits.

Source: By the author.

This result, in fact, is a known characteristic of pseudorandom quantum circuits, which can evolve to a biased replication of the uniform generation of unitaries by increasing its depth (103,104) if the set of gates is universal. The circuit depth needed so the induced measure has nonvanishing support[2] for every $U \in \mathcal{U}(2^n)$, n number of qubits, is of order at least $\mathcal{O}(n^3 2^{2n})$ (103). Notice that this property is different from the analysis made through $t-$designs, as this concept was introduced later in time (52–54) and is based on a relaxation of the conditions imposed to the measure, only requiring replication of statistical moments calculation over a space. A measure presenting nonvanishing support for every point in

[2] Here, support is in the sense of topological spaces. Conceptually, saying that a induced measure, e.g., μ_C, has nonvanishing support for all points in a topological space is the same as saying that for every possible open neighbourhood around every point μ_C will present positive values. In the case of standard probability measures, μ_C will be nonvanishing in the space. Mathematically, if μ_C is the circuit measure on the measurable space of unitary operations of n qubits, $(\mathcal{U}(2^n), \mathcal{B})$, $\mathcal{B}$ Borel sigma-algebra on the space, the support of μ_C is $\text{supp}(\mu_C) := \{U \in \mathcal{U}(2^n) | \forall N_U \in \tau : (U \in N_U \Rightarrow \mu_C(N_U) > 0)\}$. N_U is an open neighbourhood around U and τ is the chosen topology (105).

a space does not necessarily replicate the uniform measure over this same space at any or at a particular moment order. Still, this property is interesting, because it allows for universal quantum computation using the random circuit structure (103). Our circuits do not represent this property in the strict sense, as the set of gates is not universal. A possible universal set of two qubits gates is a universal set of one qubit operations together with the CNOT gate (35). To generate one qubit universal operations using only RX and RY we would need at least three parameters (35). Our circuits only apply two parameters, generating unitaries distributed on the group manifold, however not all of them and not uniformly distributed. The ansätze construction is good enough to provide high convergence rates to the uniform random case with one less operation/parameter.

The structures of standard pseudorandom circuits present some differences when compared to the structure of parameterized quantum circuits. For example, in Ref. (103), the pseudorandom circuits are composed of uniformly sampled unitary operations from $\mathcal{U}(2)$ applied on each qubit, followed by a step of two qubits operations ZZ-Ising[3] with angles $\pi/4$ between every pair of neighbouring qubits. Other possibilities could be to choose a random gate from a universal set of two qubit unitaries, $\mathcal{U}(4)$, and apply it to a pair of qubits randomly chosen (not necessarily neighbours) (106) or, for qudits systems, to draw a unitary from the Haar measure in the set of two qudits operations, $\mathcal{U}(d^2)$, randomly choose two neighbouring qudits and apply the unitary (107).

Distinctly, recent proposals more aligned to VQAs and parameterized quantum circuits involve the application of parameterized local gates and two qubits gates, instead of Haar sampling (38, 39). These circuits are very close to the structures proposed here and replicate the behavior of a random circuit, up to some point. In the context of pseudorandom circuits, the analysis of their characteristics is aligned to the objective of generating Haar-random unitaries. In this more recent approach of characterizing parameterized quantum circuits (20, 38, 39), the final objective is not the generation of random states, but to understand how the randomness affects optimization. Understanding how the structure of the circuits will affect optimization is, therefore, essential. Even more, this understanding can lead to novel proposals outside the scope of VQA, but inspired by the properties of the circuits.

Going back to the data, at only 1 layer there is fluctuation in how close the circuits are to the uniformly distributed case and even the no connections case has a very similar performance when compared to the connected topologies. When we consider, instead of a fixed number of qubits, the evolution as a function of the number of qubits with a fixed number of layers, this feature is highlighted. Fig. 8 shows that there is no hierarchy between circuits at only 1 layer (Fig. 8 (a)), including even the no connections topology, and this is independent on the number of qubits. However, for 3 layers (Fig. 8 (b)), the

[3] A ZZ-Ising gate can be defined as $ZZ(\theta) := \exp\left(-i\frac{\theta}{2}\sigma_z \otimes \sigma_z\right)$.

hierarchy is established and gets more pronounced for increased dimensions. The same occurs for 5 layers (Fig. 7 (d)).

This is also observed in Fig. 7: in the case of 4 qubits (a), there is not much difference between the evolutions and this is also true for 3 qubits (App. C). Increasing the dimension, we can see that the hierarchy of the evolutions as a function of the number of layers gets greater, starting to appear in 6 qubits (b) and has the extremal observed case in 8 qubits (d). It is possible to say that the star case has the slower decrease, followed by the linear and all-to-all almost together and, finally, the ring with the steepest decrease. Therefore, we can place an important relation between the evolution of relative entropies and the connectivities of the circuits. For smaller dimensions, this influence is not so important, however, greater dimensions will imply in more influence of the structure of the connections to the generation of uniformly distributed states. The striking feature is that the connections between qubits are intrinsically connected with the entanglement, as it is generated by the two qubits gates, and we are going to discuss the characteristics of entanglement to understand how it affects the generation of states.

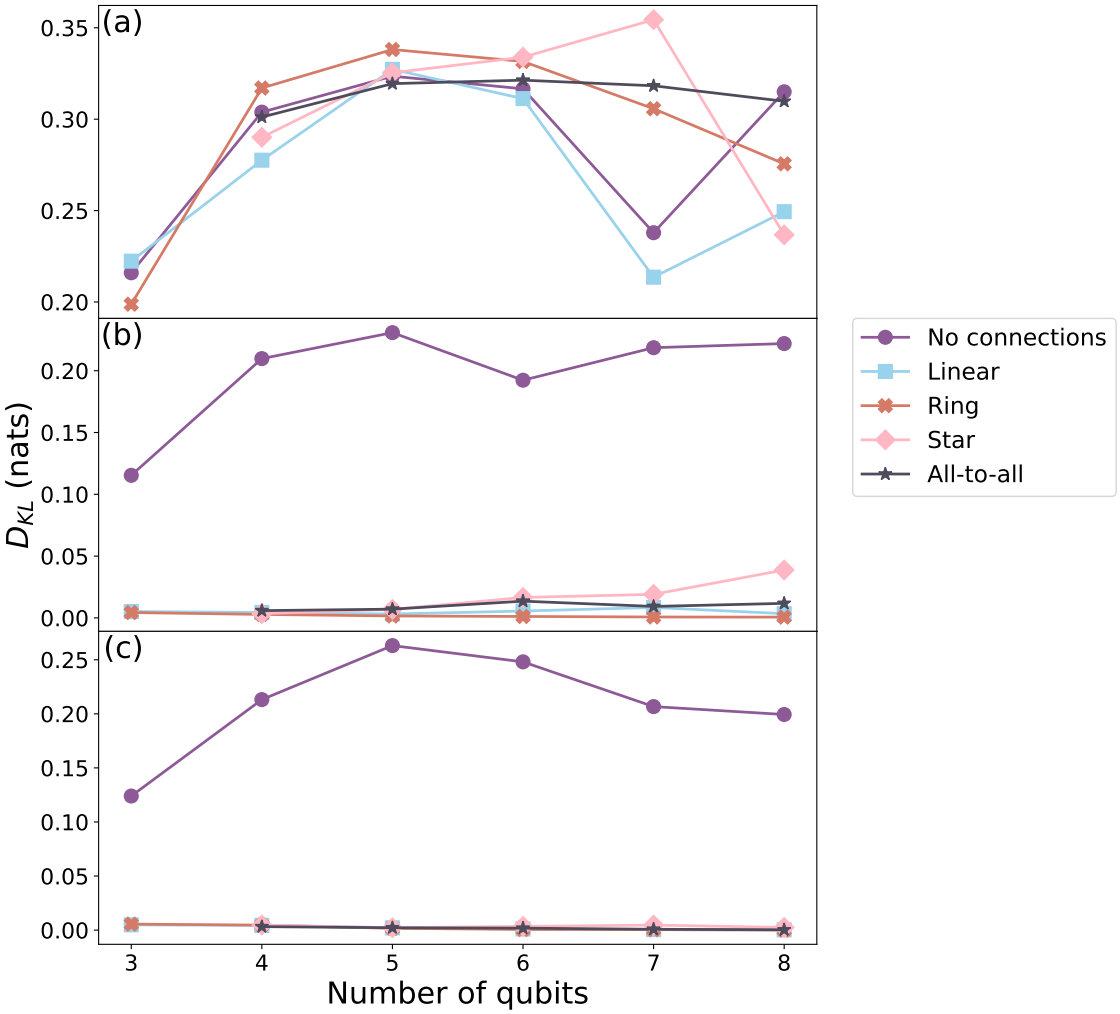

Figure 8 – Ansatz 1: relative entropy as a function of the number of qubits for all connectivities, considering the case of (a) 1 layer, (b) 3 layers and (c) 5 layers.

Source: By the author.

6.2.1.2 Entanglement

Fig. 9 presents the values of entanglement considering the subtraction between the mean for the uniformly distributed case (CUE) and the circuit induced means, for different dimensions. It is important to recall that this CUE mean will depend on the dimension, as depicted in Eq. (5.16), so the plots are normalized with respect to the value of $\langle Q_1 \rangle_{CUE}$[4] for each dimension. As a function of the number of layers, the entanglement comparison evolves to zero, independently on the topology or on the dimension of the system. This is a result of the convergence of the circuits with uniformly sampled parameters to the uniform random unitaries case, as discussed for the D_{KL} in Fig. 7 and 8.

The evolution to the CUE mean is a characteristic of many different systems and also occurs for many different entanglement quantifiers. Important examples are the parameterized quantum circuits presented in Refs. (38–40), whose Meyer-Wallach, von

[4] The notation $\langle \cdot \rangle_{CUE}$ indicates the mean over the Circular Unitary Ensemble (CUE).

Neumann and many orders of the Rényi entropies achieve high entanglement values equal to random mean values with increasing number of layers. In the case of pseudorandom circuits, the Meyer-Wallach convergence can be used as a necessary condition of randomness of the circuit (40). Other established examples are quantized versions of chaotic classical maps, e.g., the kicked rotor (81) or the baker's map (108), which saturate entanglement quantifiers in the parameters regions where the classical map is chaotic.

Usually, this can be interpreted as the generation of arbitrary states due to the convergence of the unitary circuit to the Haar random case. However, this can be understood by connecting the behavior of chaotic and complex systems, such as information scrambling, operatior spreading, the Bohigas-Giannoni-Schmit conjecture[5] (39) and times of relaxation (104) with the behaviors of random circuits. This would lead to important conclusions about the states generated by both kinds of systems and how they can be used in the advancement of physical knowledge. One could apply quantum circuits to generate states analogous to chaotic many-body systems (110), to understand complexity and entanglement generation of unitaries in a more general approach (43, 111) or even to comprehend how information scrambling can harm optimization of parameters in parameterized quantum circuits (39). All these possibilities are the center of many recent research approaches.

How fast is the convergence to zero will depend on the topology of the circuit. Circuits generating entanglement closer to the mean calculated with the uniform measure already at the first layer have a faster convergence. The ring circuit always generates mean entanglement values closer to or bigger than the CUE mean, followed by linear together with all-to-all, and, finally, the star. We can see that the hierarchy observed for the decay of the relative entropy as a function of the number of layers is exactly the same as this hierarchy for entanglement evolution.

Before, we discussed that the relative entropy decay is very similar for the different topologies in smaller dimensions (3 and 4 qubits), however, increasing the dimension, a hierarchy between the circuits is built. In contrast, the data presented in Fig. 9 shows that the entanglement hierarchy exists already at the smaller dimensions (Fig. 9 (a)). This implies that, regarding Ansatz 1, the entanglement generated by the circuits of smaller dimensions is not directly correlated to the decay of the relative entropy. At bigger dimensions, from 5 to 8 qubits, the correlation is observed and values closer to zero in Fig. 9 (b), (c) and (d) imply in a steeper decay of relative entropy.

[5] The Bohigas-Giannoni-Schmit (BGS) conjecture states that the dynamics of a quantum system with chaotic classical analogue will be driven by a hamiltonian with spectrum that follows the Gaussian Unitary Ensemble (GUE), Gaussian Orthogonal Ensemble (GOE) or the Gaussian Sympletic Ensemble (GSE) depending on the symmetries. (109).

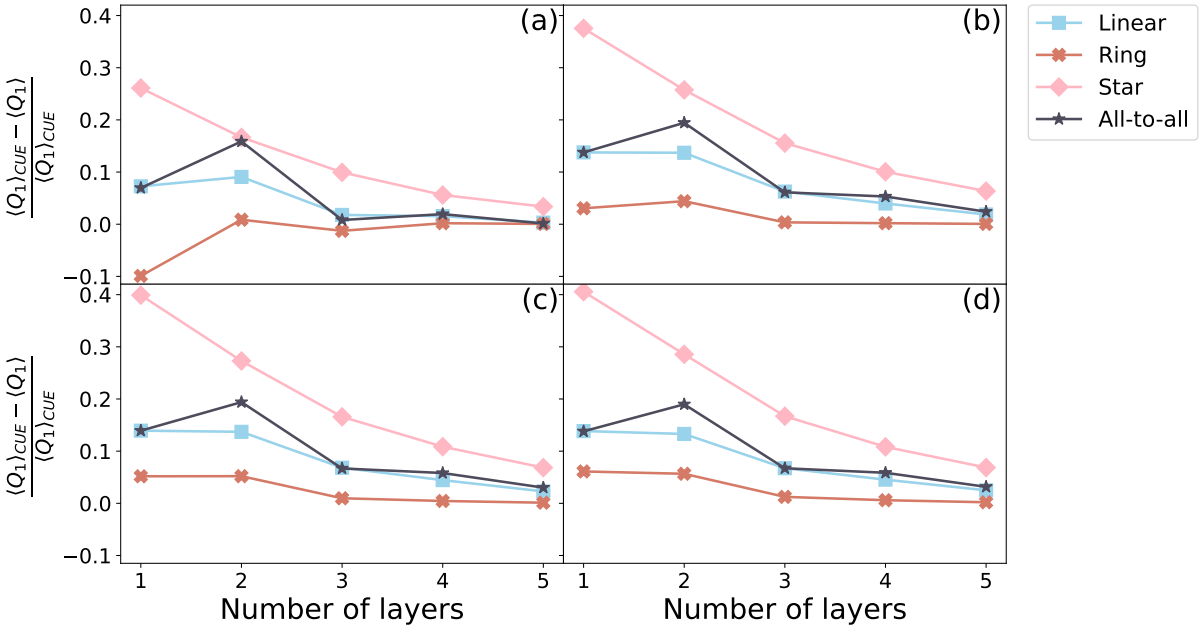

Figure 9 – Ansatz 1: normalized subtraction of the CUE ensemble mean minus the circuit mean of the Meyer-Wallach quantifier for all connected topologies as a function of the number of layers for (a) 4 qubits, (b) 6 qubits, (c) 7 qubits and (d) 8 qubits. The analytical value used for the CUE mean depends on the dimension and is given in Eq. 5.16.

Source: By the author.

The mean entanglement for generic states, Eq. (5.16), increases with the dimension of the system and is very close to maximal entanglement for high dimensional systems (81, 108). Also, comparing the Meyer-Wallach standard deviation of states generated by the circuits with the CUE standard deviation, Fig. 10, we can see that circuits whose mean values are closer to the uniform random distribution also possess standard deviations closer to the uniform case. The decay to the random case is slower for circuits whose initial value (at 1 layer) is already very close.

The reference CUE standard deviation goes exponentially to 0 with the dimension of the system (81, 108), Eq. (5.17)[6], so the generic states are highly entangled and do not deviate a lot from the mean value. This is a possible justification for why the effects of generating entanglement with characteristics closer to the CUE are more influential to higher dimensions, above 4 qubits: There is not much freedom on which are the generic states if we consider high dimensional systems and this will impact the evolution to the uniformly random distributed states. For high dimensions, if the entanglement generated by the circuits does not possess the properties of the entanglement of generic states, the relative entropy decay will be braked.

[6] An effect detailed in a few paragraphs, when the standard deviation as a function of the number of qubits is presented.

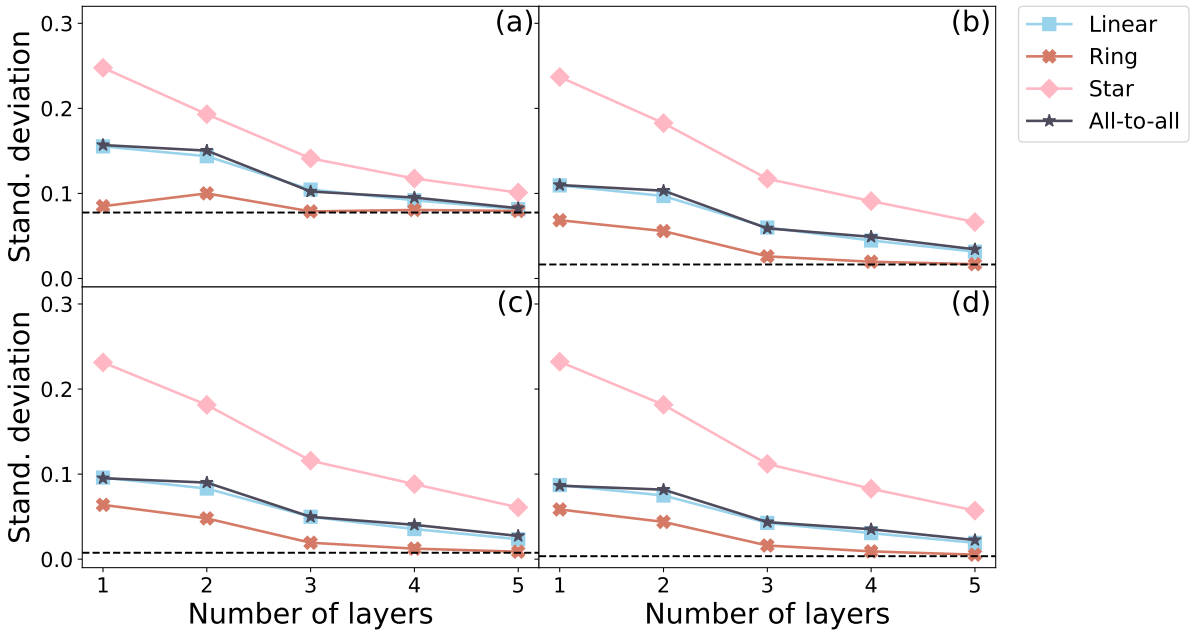

Figure 10 – Ansatz 1: standard deviation of the Meyer-Wallach quantifier for all connected topologies as a function of the number of layers for (a) 4 qubits, (b) 6 qubits, (c) 7 qubits and (d) 8 qubits. The traced line indicates the standard deviation values for the CUE, which depends on the dimension and is given in Eq. 5.16.

Source: By the author.

One could argue that the reason why this correlation exists is the fact that the ring circuit is closer to generating uniformly distributed states at 1 layer when compared to linear/all-to-all or star, therefore its entanglement values are closer to the CUE mean. There are two arguments in opposition to this conclusion. Firstly, in Fig. 8 (a), it is possible to see that the values of the relative entropy comparing each of the circuits for different dimensions and 1 layer do not respect a particular ordering and even the no connections has values closer to the uniform random distribution of states than the connected topologies. Another strong argument is based on the notion of frame potentials. In Chap. 5, we showed that how close a particular ensemble of unitaries is to a $t-$design can be determined using frame potentials, by the means of Eq. (5.10). The closer it is to zero, the closer an ensemble is to a $t-$design. In this same chapter, we discussed that the Meyer-Wallach entanglement quantifier is a $(2, 2)$ polynomial and that a $2-$design is sufficient to replicate the behavior of the Haar obtained mean, as defined in Def. 2. Table 4 shows that there is no ordering between the circuits values for the comparison with a $2-$design. This way, circuits generating entanglement closer to the random uniform mean are doing so because of the structure of the connections and not because of the randomness of the states being closer to a distribution that can replicate the mean values of entanglement. The complete table, including 3 and 5 qubits, is presented in App. C.

Table 4 – Ansatz 1 values for the quantity $||\mathcal{A}^{t=2}||_2^2 = F^{(t=2)}_{|0\rangle\langle 0|^{\otimes n}} - F^{(t=2)}_{Haar}$, that compares the circuit with a 2−design, depending on the topology and on the number of qubits, n. The mean values and standard deviations were obtained over 20 independent calculations.

Topology	$n = 4$	$n = 6$	$n = 7$	$n = 8$
No connections	$(6.5 \pm 0.5) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.6 \pm 0.7) \cdot 10^{-4}$	$(1.6 \pm 0.3) \cdot 10^{-4}$
Linear	$(6.6 \pm 0.7) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.5 \pm 0.6) \cdot 10^{-4}$	$(1.7 \pm 0.3) \cdot 10^{-4}$
Ring	$(6.7 \pm 0.6) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.3 \pm 0.6) \cdot 10^{-4}$	$(1.7 \pm 0.3) \cdot 10^{-4}$
Star	$(6.8 \pm 0.5) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.4 \pm 0.6) \cdot 10^{-4}$	$(1.6 \pm 0.3) \cdot 10^{-4}$
All-to-all	$(6.6 \pm 0.6) \cdot 10^{-3}$	$(1.2 \pm 0.2) \cdot 10^{-3}$	$(4.5 \pm 0.8) \cdot 10^{-4}$	$(1.6 \pm 0.3) \cdot 10^{-4}$

Source: By the author.

In Fig. 11, the evolution of the mean entanglement as a function of the number of qubits is shown for different number of layers. As stated above, the ring circuit always generates mean entanglement higher or closer to the CUE mean, followed by linear/all-to-all almost together and then star, independently on the dimension or number of layers. Increasing the number of layers, the circuits converge to the CUE entanglement values, being the ring the first to achieve a very similar behavior and, at higher dimensions, the entanglement values get very close to 1.

One interesting characteristic is that the generations of entanglement for the linear and all-to-all connectivities are almost the same, up to fluctuations, for all dimensions and number of layers (Fig. 11 (a), (b) and (c)). The number of CNOTs for the linear circuit is $(n-1)l$, while for the all-to-all, $n(n-1)l/2$, n number of qubits and l number of layers. From this, we can say that the entangling gates are not being performed in an optimized manner for the all-to-all circuit, as the number of CNOTs and the complexity of the connections are the greatest among the topologies and a simpler structure is generating equivalent entanglement.

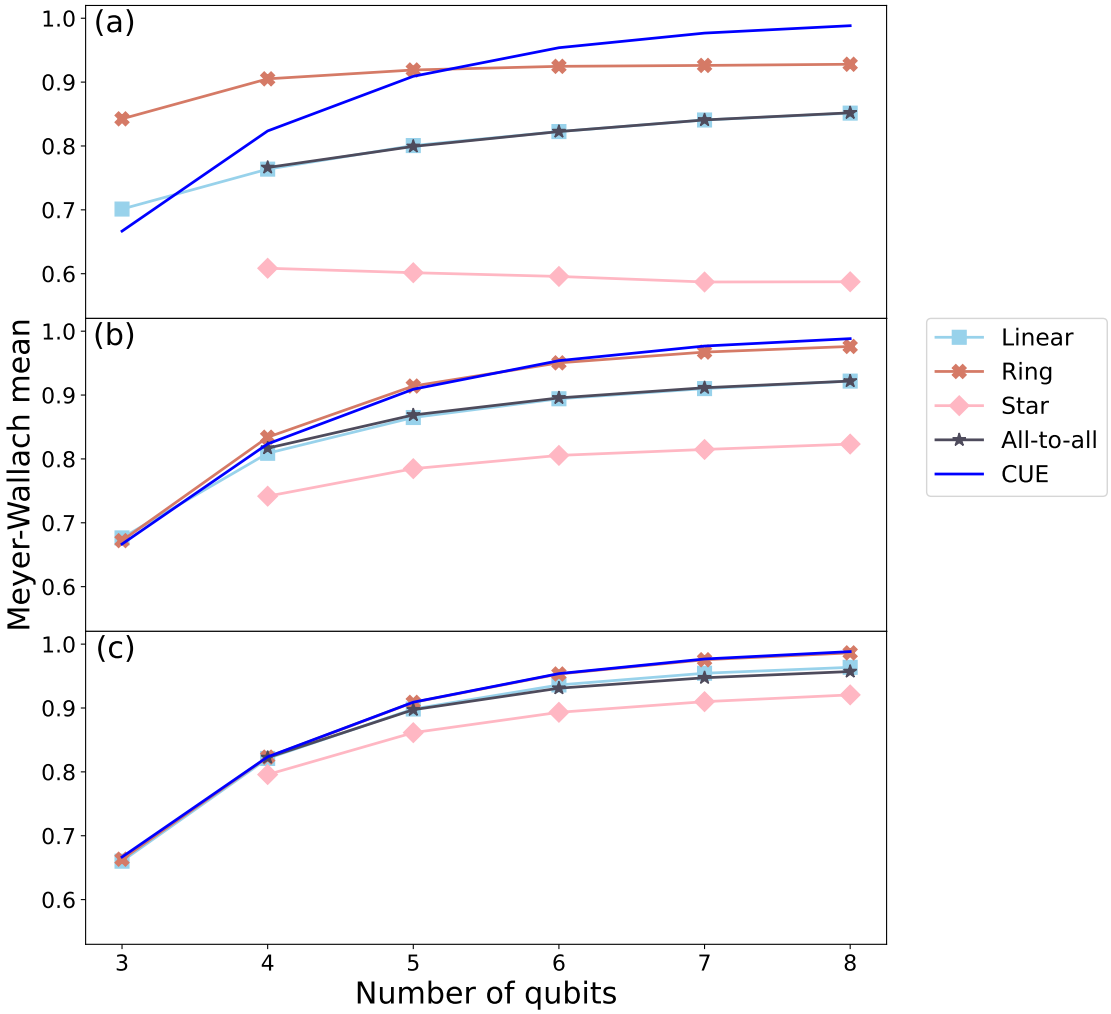

Figure 11 – Ansatz 1: mean of the Meyer-Wallach quantifier for all connected topologies as a function of the number of qubits for (a) 1 layer, (b) 3 layers and (c) 5 layers. The blue line indicates the analytical values of the CUE depending on the dimension of the system and is given in Eq. 5.16.

Source: By the author.

To visualize why, consider an example of 1 layer of Ansatz 1, with 4 qubits and the all-to-all connectivity. The parameterization is going to be fixed as shown in Fig. 12. A Hadamard gate H acting on the input $|0\rangle$ computational basis state is analogous to the sequence RX-RY acting on the same input with parameters

$$H|0\rangle = RX(0)RY\left(\frac{\pi}{2}\right)|0\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \tag{6.1}$$

where the matrices are written in the computational basis $\{|0\rangle, |1\rangle\}$. We are going to choose the parameterization of the first qubit as RX(0) and RY($\pi/2$), so a Hadamard is applied, and on the other qubits nothing is done in the local operations (RX(0) and RY(0)) and then apply the all-to-all connections. We divided the action of the connections

in two steps to clarify what is happening. After step 1, a GHZ_4 state is generated, i.e., $(|0\rangle^{\otimes 4} + |1\rangle^{\otimes 4})/\sqrt{2}$, which is a maximally entangled state according to the Meyer-Wallach quantifier ($Q_1(GHZ_4) = 1$). After step 2, the state will be a tensor product of a Bell state and computational 0 states,

$$|\psi\rangle_{\text{out}} = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \otimes |0\rangle \otimes |0\rangle \,. \tag{6.2}$$

which has entanglement according to the Meyer-Wallach equal to 1/2. This way, even though the all-to-all connectivity is generating a maximally entangled state in the first step, it is destroying the entanglement in step 2, because CNOTs are unitary operations of two qubits and can destroy entanglement. From the results of Fig. 11, we can say that different parameter vectors, on average, will create and destroy entanglement on the process, generating a mean close to the linear case with many more entangling gates.

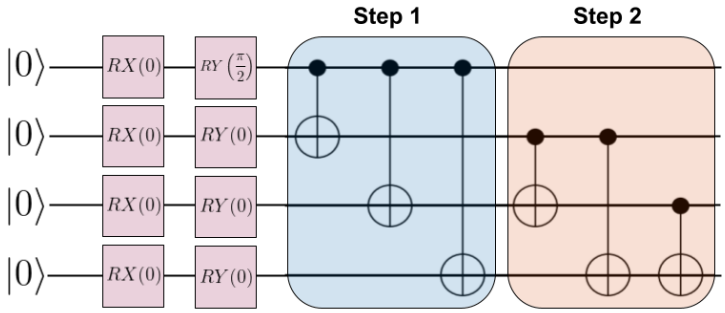

Figure 12 – Circuit representation of Ansatz 1 with 4 qubits with a particular parameterization. In this parameters/input state setup, a Hadamard is applied to the first qubit and nothing is done to all the others on the local unitaries step and then the all-to-all connections with CNOT gates are applied. The connections are divided in two parts, step 1 (blue) and step 2 (red).

Source: By the author.

If the wish is to generate highly entangled states, on average over uniformly distributed parameters, a better choice of ansatz would be to distribute the local operations between entangling gates, so the chance of destroying entanglement by the sequential application of CNOTs is reduced. Nevertheless, the all-to-all circuit can generate highly entangled states if a proper choice of parameters is done. A way to do that is by the optimization of the parameters, searching for entangled states (37). This process will depend on a cost function and has possible complications involving the trainability, as discussed in Chap. 2.

Looking to how entanglement change considering the different dimensions at 1 layer (Fig. 11 (a)), a slight increase is observed for the ring, linear and all-to-all connectivities, while an almost constant value is observed for the star. These differences can be related to the characteristics of the connections: On the one hand, the star circuit only connects one qubit to every other, so the local prepared state of the central qubit has a strong

influence over the entanglement generation. This, in addition to increasing the difficulty of entangling the other qubits between them, will influence the variation of the entanglement depending on the dimension. With more qubits, we are still coupling only one qubit to the others, but the other qubits will still be decoupled, so the entanglement saturates. On the other hand, circuits as ring, linear and all-to-all will have more possibilities of couplings when the number of qubits increases, favoring the generation of multipartite entanglement between different parts of the system. The star will have the possibility of increasing the entanglement with more layers due to the contribution of the local operations on other qubits to the global entanglement generated, so the values will have a weaker dependency on the central qubit.

A similar analysis of the evolution as a function of the number of qubits is presented for the standard deviation in Fig. 13. At only 1 layer (Fig. 13 (a)), the standard deviation of the star and ring circuit do not change a lot, however there is a slight decrease for the linear and all-to-all, again with a very close behavior. Increasing the number of layers, the circuits approach the CUE standard deviation in a very similar manner as discussed for the mean value.

The most interesting feature of the CUE standard deviation is the exponential decay, which the circuits can replicate partially (star, linear, all-to-all) or very closely (ring) at 5 layers (Fig. 13 (c)). This exponential decay of the standard deviation is a consequence of the *concentration of measure* (26, 36) on the space of states, which is a phenomenon characterized by Levy's lemma (112). According to Levy's lemma, if $f : \mathbb{S}^k \to \mathbb{R}$ is a sufficiently slow varying function[7] from a sphere of dimension k, $\mathbb{S}^k$, to the real numbers and $X \in \mathbb{S}^k$ is a point chosen uniformly at random, i.e., sampled according to a uniform probability measure on the sphere, the probability that $f(X)$ differs from the average value of f on the sphere decays exponentially to zero as a function of the dimension k. The space of pure states for a n qubits system without the constraints on the global phase or normalization is $\mathbb{C}^{2^n}$. The elements in this space can be represented as $2 \cdot 2^n$ real vectors. Imposing normalization, the set is represented as $2 \cdot 2^n-$dimensional real vectors with unit norm, i.e., a sphere $\mathbb{S}^{2 \cdot 2^n - 1}$ [8]. This way, the space of pure states is always isomorphic to a sphere and it is possible to consider Levy's lemma. How fast is the exponential decay will depend on how slow is the function variation and the entropy of entanglement is know to be "slow enough" so the effect is observed (112).

From these observations, we can understand why the standard deviation decays exponentially to zero with the dimension and justify the discussions above connecting

[7] Dictated by the characteristic Lipschitz constant of the function, conceptually equivalent to the maximum gradient value.

[8] If we consider an element in $\mathbb{C}^{2^n} = \bigotimes_j^{2^n} \mathbb{C}_j$, it is described by $2 \cdot 2^n$ real numbers, the real, x_j, and the imaginary, y_j, part for every $z_j \in \mathbb{C}_j$, $z_j = x_j + iy_j$. Imposing unit norm is requiring that $\sum_j^{2^n} x_j^2 + \sum_j^{2^n} y_j^2 = 1$, the equation of a $2 \cdot 2^n - 1$ unit sphere.

the concentration of the standard deviation with the decrease of the relative entropy. Any generation of generic states will be affected by this characteristic and this will have strong influences on the training of parameterized quantum circuits when considering possible VQA applications (26, 36). The phenomenon is called concentration of measure due to the fact that functions will concentrate around mean values, so this is analogous to have a measure that is concentrated around a particular region of the space.

In this first part, discussing the relations between how expressible is the circuit and its entanglement, we are only going to employ the MW/S1 quantifier, as the S2 does not provide any additional information in this sense. Fig. 14 shows that the evolutions and saturations of entanglement for both quantifiers are closely related. The same behavior is observed, together with the same hierarchy between the circuits.

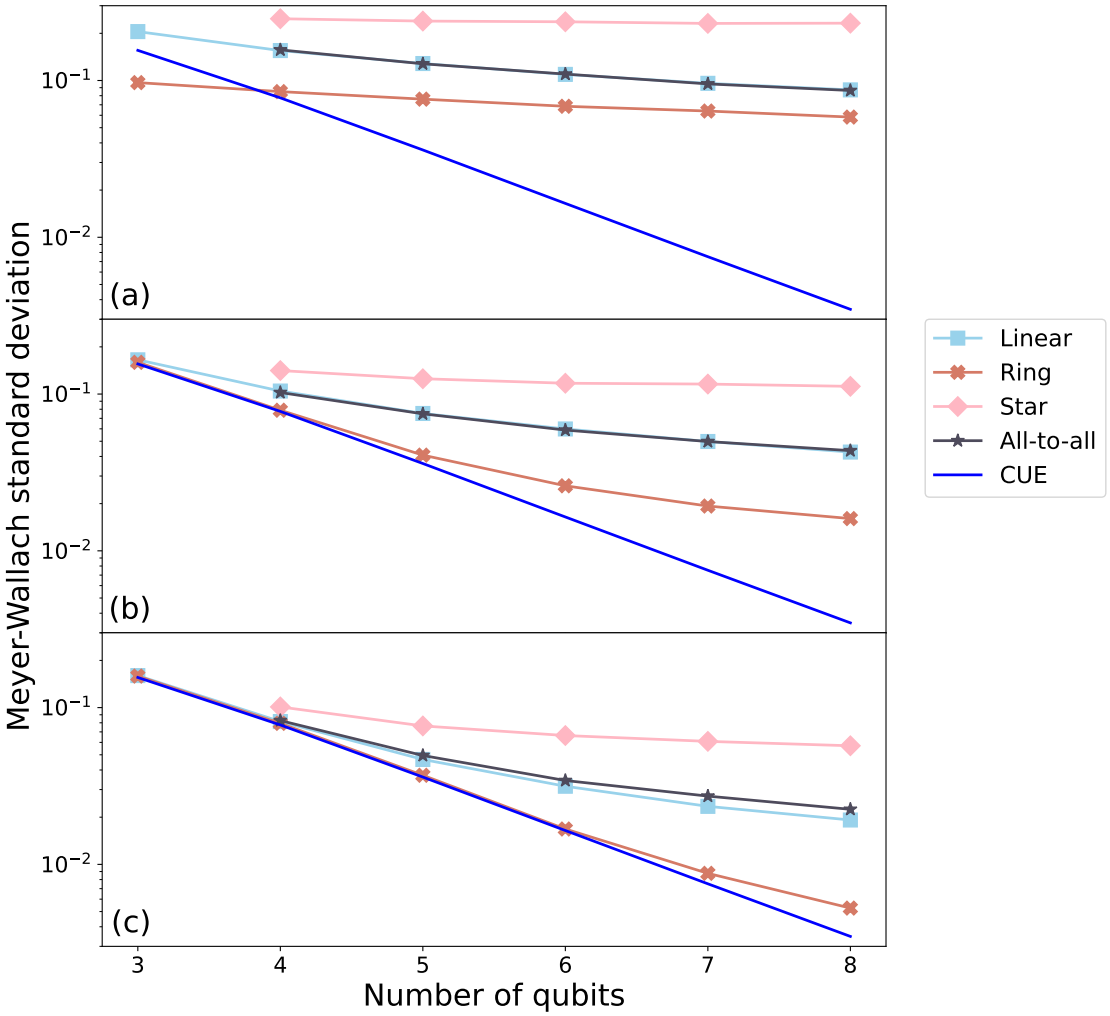

Figure 13 – Ansatz 1: standard deviation of the Meyer-Wallach quantifier for all connected topologies as a function of the number of qubits for (a) 1 layer, (b) 3 layers and (c) 5 layers. The blue line indicates the analytical values of the CUE depending on the dimension of the system and is given in Eq. 5.17.

Source: By the author.

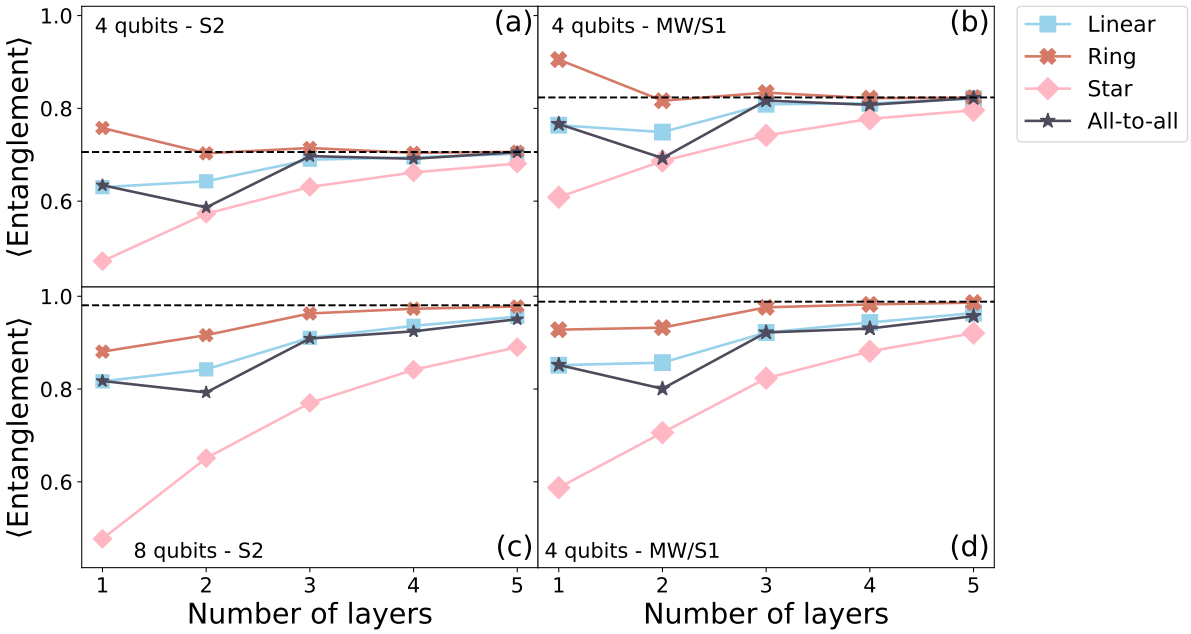

Figure 14 – Ansatz 1: comparison between the entanglement of the different connectivities considering the two different quantifiers presented in this work, Meyer-Wallach/Scott $m = 1$ (S1) and Scott $m = 2$ (S2), where the numbers indicate the sizes of the bipartitions. The graphs present the cases of (a) S2, 4 qubits, (b) S1, 4 qubits, (c) S2, 8 qubits and (d) S1, 8 qubits.

Source: By the author.

6.2.2 Ansatz 2

In this subsection the discussion will be made without separating expressibility and entanglement quantifiers. Fig. 15 shows a comparison between the entanglement generated by Ansatz 1 (A1) and Ansatz 2 (A2) for systems with 4 and 8 qubits. They present little to no differences, independent on the number of qubits (this is also true for other dimensions not presented, as seen in App. C). When defining the ansätze, we discussed that Ansatz 1 and 2 should present the same entanglement considering only 1 layer, but this is not necessarily true for more layers as the effects of the differences due to additional local operations could change the superpositions and entanglement generation. By the data in Fig. 15, the additional local operations between entangling steps do not affect output entangled states. Therefore, the additional randomness introduced will have influences only on the expressibility of the circuits. A slight difference is observed for the all-to-all connectivity at 2 layers, but it disappears for 3 layers. This result is interesting because both ansätze will have equivalent entanglement and we can isolate effects due to the local parameterized gates.

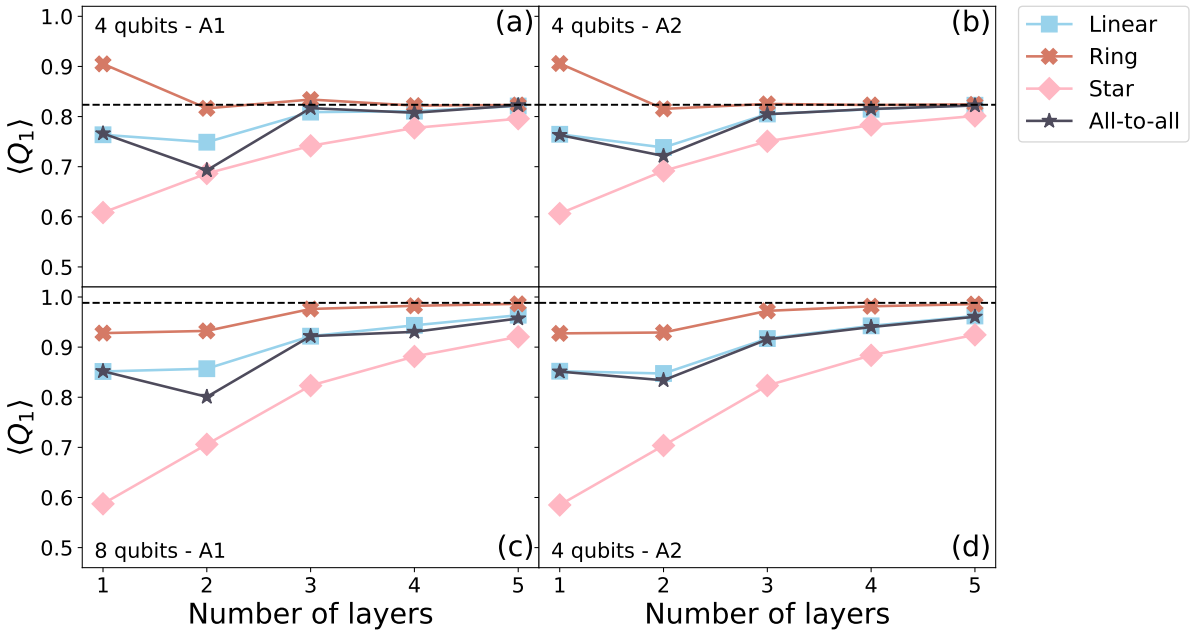

Figure 15 – Comparison between the entanglement of the different connectivities considering the two ansätze, Ansatz 1 (A1) and Ansatz 2 (A2). The graphs present the cases of (a) A1, 4 qubits, (b) A2, 4 qubits, (c) A1, 8 qubits and (d) A2, 8 qubits.

Source: By the author.

As entanglement was already discussed in detail for Ansatz 1 and will have the same properties for Ansatz 2, we turn our attention to the relative entropy and which are the differences between ansätze. In Fig. 16, the peculiarities of Ansatz 2 start to appear: Already at 1 layer, the circuits present a distinction between the no connections and the connected topologies, which have a smaller relative entropy. Ansatz 2 takes advantage of the generated entanglement without additional cost of two qubits gates by changing local coherences with the local unitaries. Introducing more local parameters it is possible to increase the expressibility of the circuits. Another property observed is that the hierarchy between circuits already exists at only one layer, even for the smaller dimension of 4 qubits.

In Fig. 17 the more interesting characteristics of Ansatz 2 can be seen in detail using a logarithmic scale. At smaller dimensions, Fig. 17 (a) and (b), the evolution of the relative entropy for Ansatz 2 is close to the evolution of Ansatz 1, however, the values for 2 layers of Ansatz 1 are equivalent to only 1 layer of Ansatz 2. For example, in (a), the saturation value of $5 \cdot 10^{-3}$ nats is achieved already at only 2 layers, while it takes at least 4 layers in the case of Ansatz 1. In (b), a saturation is achieved at around $1 \cdot 10^{-3}$ nats, while for Ansatz 1 the circuits still have a decreasing relative entropy for 5 layers and above $1 \cdot 10^{-3}$ nats. From this point of view, the structure of Ansatz 2 is optimized to achieve the uniform distribution of states, applying less entangling gates and more local operations.

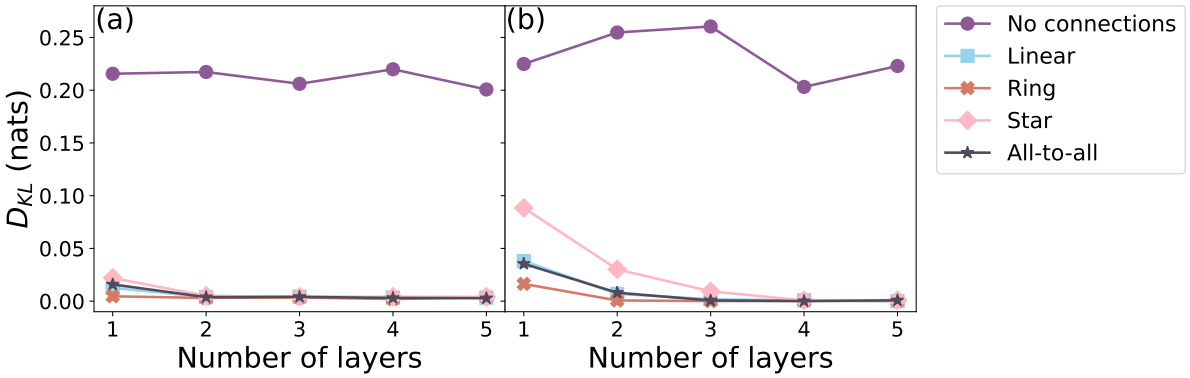

Figure 16 – Ansatz 2: relative entropy as a function of the number of layers for all connectivities, considering the case of (a) 4 qubits and (b) 8 qubits.

Source: By the author.

Increasing the dimension of the system, Ansatz 2 still has a steepest decrease of D_{KL} as a function of the number of layers when compared to Ansatz 1, but this equivalence of 2 layers of A1 to achieve the same values of 1 layer of A2 is not true anymore and more layers are necessary for A2, so the decrease is slower. Up to 3 layers, the effects of local parameters in A2 are very beneficial, decreasing quickly the relative entropy. However, at 4 and 5 layers, we can see that the evolutions start to get erratic, increasing instead of decreasing as in the case of ring topology in Fig. 17 (b), (c) and (d), so the randomness introduced by the local parameters harms the decrease and more layers are not necessarily a condition to obtain circuits closer to the uniform case.

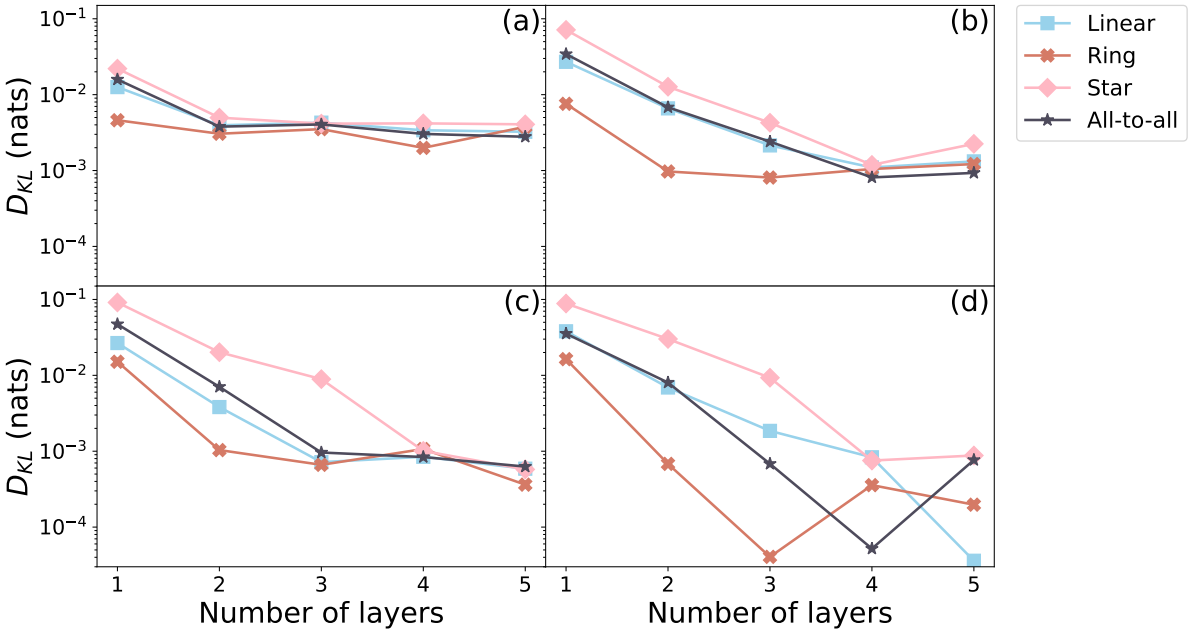

Figure 17 – Ansatz 2: relative entropy in logarithmic scale as a function of the number of layers for all connected connectivities, considering the case of (a) 4 qubits, (b) 6 qubits, (c) 7 qubits and (d) 8 qubits.

Source: By the author.

In the case of Ansatz 1, the entanglement is not so influential at smaller dimensions and the effects of generating entanglement close to the uniformly distributed ensemble are more appreciable for systems with more than 4 qubits. Now, Ansatz 2 presents a hierarchy between topologies already at 1 layer and the effects of generating the entanglement of generic random states are influential from the very beginning. This seems to contradict the discussion made before, but the new elements introduced in Ansatz 2 have to be taken into account: The additional step of local operations will provide more freedom on what are the states that can be built using the circuit. For every entangled state produced in Ansatz 1, Ansatz 2 can create many different states without changing the entanglement. In this sense, generating entanglement closer to the CUE mean is not useful in the smaller dimensions case if one does not have freedom on the local coherences (as observed for A1), but can be valuable if we have access to additional local operations. Still, without this "local freedom", tuned entanglement generation can be very useful in higher dimensions.

These observations are in consonance with results presented in Ref. (20). In this work, there are circuits structures whose mean entanglement generated at 1 layer are very close to the CUE mean, however present relatively lower expressibility values when compared to other circuits with entanglement farther from the CUE mean. It is necessary to highlight that those circuits have very distinct connection structures when compared between them (including the existence or lack of parameterized two qubit gates), which is an analysis very different from the one presented in our work. Going beyond that, remarkably, one of the referred circuits[9] has an expressibility increase as a function of the number of layers that surpass many others which do not possess the same entanglement characteristics. Therefore, even though the mean entanglement possess the characteristics of a uniformly distributed state, the parameterization of this circuit does not provide enough freedom to construct generic states at small layer numbers. By increasing the number of parameters, the circuit can quickly achieve high expressibility values[10]. The results of expressibility evolution as a function of number of number of 2 qubits operations is presented in the supporting information of Ref. (20). This parameter can be seen as equivalent to the number of layers here, as there the number of 2 qubits gates only increases with repeating layers.

It is also important to stress the peculiarities of our results. The use of the Meyer-Wallach and other entanglement quantifiers to characterize pseudorandom circuits is not recent. The work of Ref. (40) studies how the topology of connections and the different gates applied locally to qubits will influence the evolution to a pseudorandom circuit with concatenations. These two parameters are taken into account to verify the convergence

[9] In Ref. (20) called "Circuit 2".

[10] Another example in this same work (Ref. (20)) is "Circuit 9", that is capable of generating the entanglement of the CUE ensemble in only one layer, but has low expressibility and the sole generation of entanglement is not enough.

to pseudorandom circuits and the mean entanglement is used as a second diagnostics of convergence. Despite the apparent degeneracy of our work with those results, here we are looking to a new kind of circuits (PQCs) and entanglement is not only a diagnosis, but an influence to the convergence. Also, the occurrence of convergence is not our final objective and we studied regimes where the circuit is relatively distant from a pseudorandom circuit (1 layer) and how this influences its behavior relatively close to a pseudorandom circuit (5 layers). The results in this section provide a characterization of different topologies available in contemporary quantum hardwares, that can be employed in many different contexts with different requirements depending on the application.

6.3 **Numerical simulations:** comparison with entanglement of paradigmatic states

In this section we discuss how the mean entanglement values can be compared to the entanglement of the paradigmatic states, using the Meyer-Wallach/Scott 1 and Scott 2. The conditions were divided into two groups: even number of qubits, Fig. 18, and odd number of qubits, Fig. 19, due to the differences of the paradigmatic states EPR_{even} and EPR_{odd}.

In Fig. 18, the first noticeable aspect is how the S2 entanglement quantifier is capable of breaking the degeneracy of states GHZ_n and EPR_{even} that appears in the left column (Fig. 18 (a), (c) and (d)). The only situation where this is not possible is in (b), as S2 assumes the same value for both states. It is hard to draw conclusions comparing the entanglement of the circuits with the entanglement of these paradigmatic states, as the two quantifiers have very different values for them. In the case of S1, only circuits of higher dimensions have mean values close to and standard deviations including these states, as seen in Figs. 18 (e) and 19 (d). However, looking at S2, it appears that circuits linear, ring and all-to-all have entanglement values close to the EPR_{even}, however are getting farther from GHZ_n with increased dimension, Fig. 18 (d) and (f).

The W_n state, conversely, has properties very close to the star circuit. Independent of the dimension, including the odd number of qubits in Fig. 19, and of the quantifier, S1 or S2, the star topology at 1 layer always generate entanglement with means close to or standard deviations including W_n. Therefore the characteristics of its entanglement have very similar properties to the W_n states. When the number of layers is increased, the star entanglement gets farther, but this is expected as it is loosing its initial properties and converging to the entanglement of generic states. At 2 layers and considering odd number of qubits, the star circuit is close to the entanglement of EPR_{odd}, as seen in Fig. 19 (b), (c), (d) and (e). The EPR_{odd} states will have entanglement very closely connected to circuits linear and all-to-all, at only one layer. This is seem using both quantifiers and for all odd dimensions.

The ring topology, which has the entanglement properties closer to the entangle-

ment of generic states and presents the steepest evolution to the uniformly distributed generation of states, has entanglement means and standard deviations that usually do not match any of the paradigmatic states. This match only occurs at the 3 qubits dimension. In this sense, we can say that the generic states entanglement is not well described only by the means of specific paradigmatic states. This conclusion was also made in Ref. (108) when quantifying the amount of genuine multipartite entanglement generated by chaotic quantum dynamics using the $n-$tangle quantifier, that is maximal only for GHZ_n states. Generic states present entanglement characteristics that saturate entanglement quantifiers in high dimensions and, in the multipartite case of Scott quantifiers, the ordering of entangled states is not well defined. States more entangled considering a specific m not necessarily are more entangled considering a different $m^{'}$. This is seen comparing the entanglement of states EPR_n and GHZ_n, S1 has values for GHZ_n that are greater than or equal to for EPR_n, but S2 has values the other way around. Still, generic states will always be highly entangled, so it is hard to connect their characteristics with the paradigmatic states studied here.

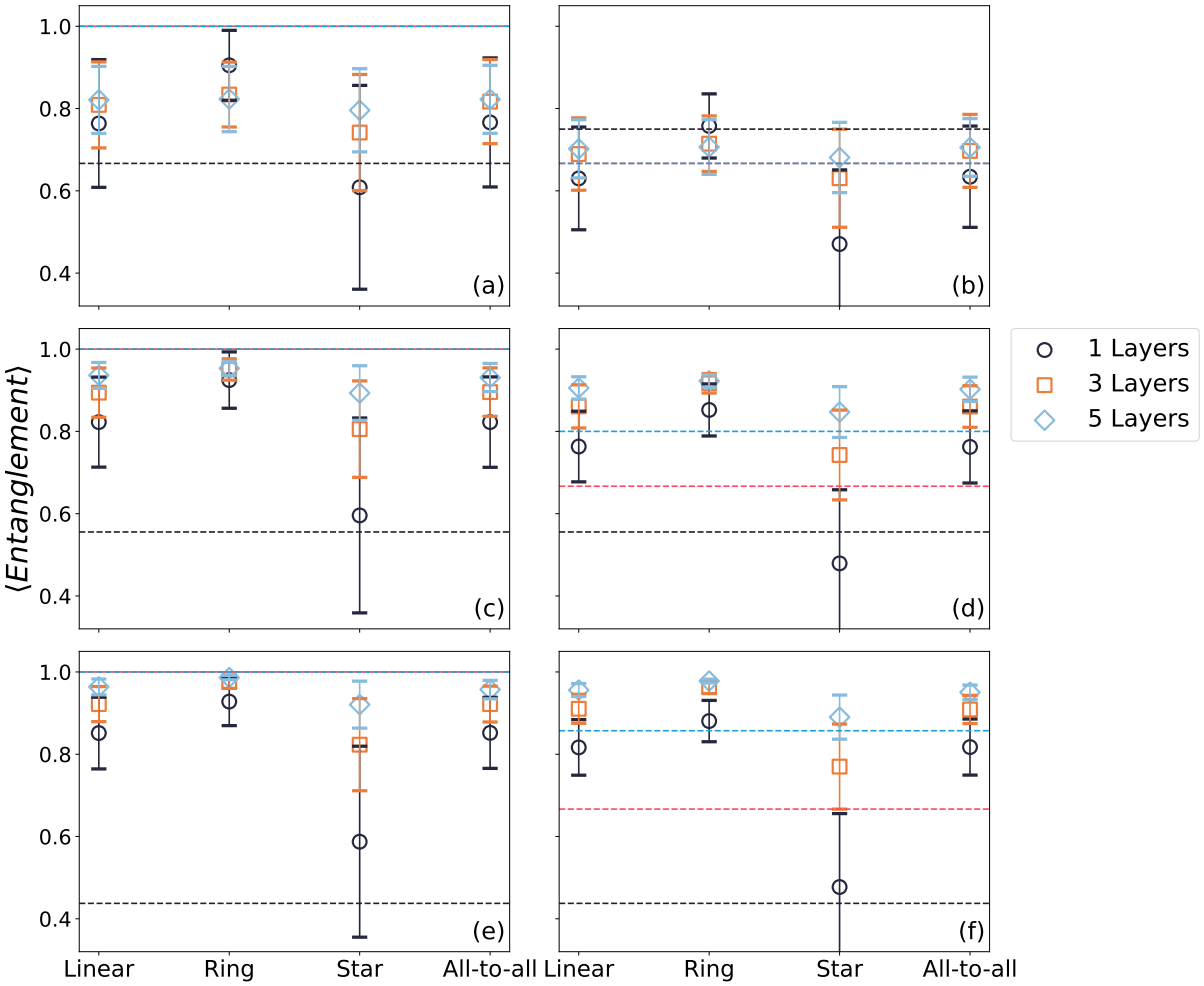

Figure 18 – Mean entanglement values of the quantifiers Meyer-Wallach/Scott 1 (S1) and Scott 2 (S2), for all the topologies considering even number of qubits. The dashed lines are the values for the paradigmatic states: dark purple is the W_n, red the GHZ_n state and blue the EPR_{even} state. The first column presents the values for S1 and the second column for S2. The rows are organized with increasing number of qubits, this way (a) S1, 4 qubits, (b) S2, 4 qubits, (c) S1, 6 qubits, (d) S2, 6 qubits, (e) S1, 8 qubits and (f) S2, 8 qubits.

Source: By the author.

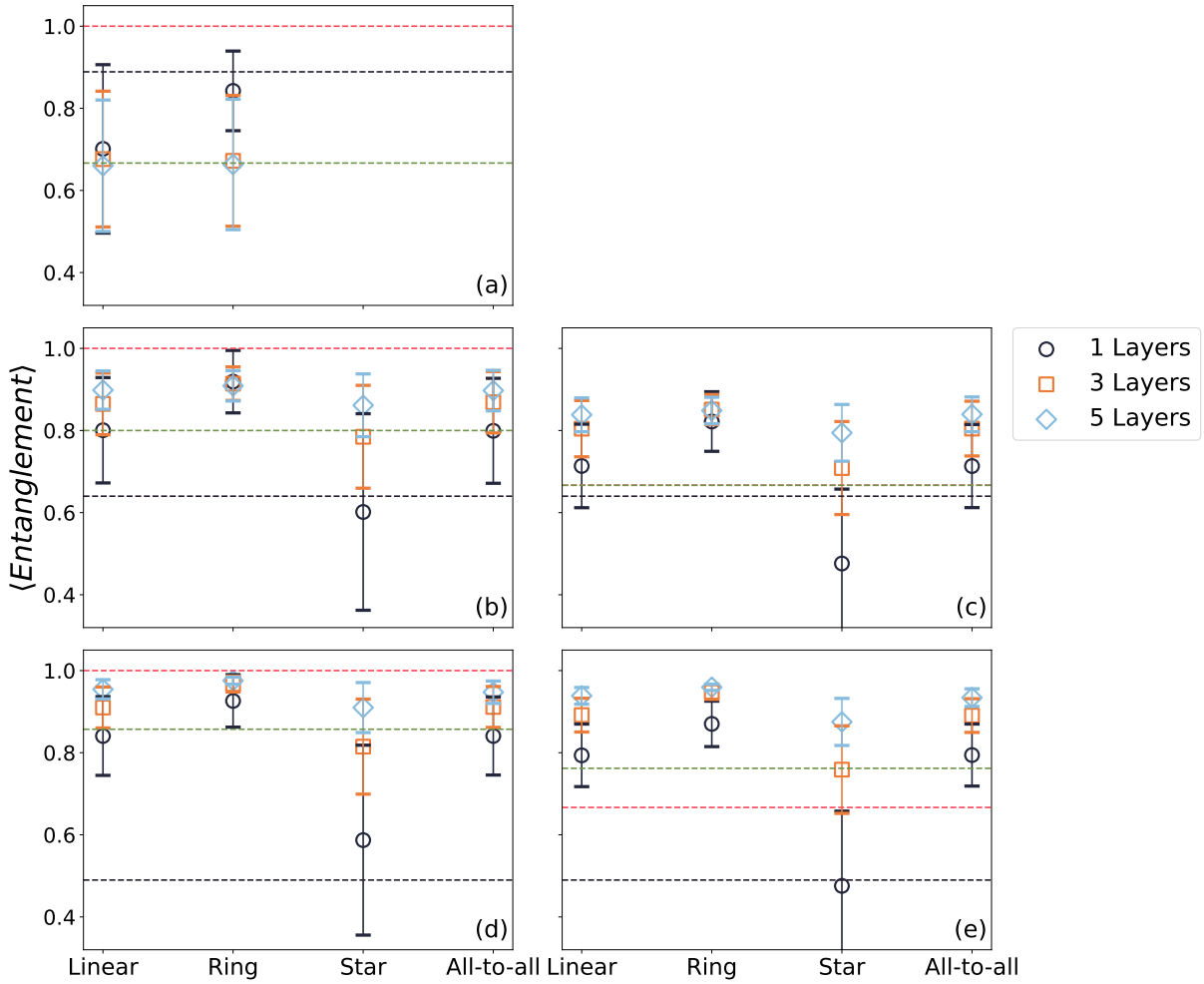

Figure 19 – Mean entanglement values of the quantifiers Meyer-Wallach/Scott 1 (S1) and Scott 2 (S2), for all the topologies considering odd number of qubits. The dashed lines are the values for the paradigmatic states: dark purple is the W_n, red the GHZ_n state and green the EPR_{even} state. The first column presents the values for S1 and the second column for S2. The rows are organized with increasing number of qubits, this way (a) S1, 3 qubits, (b) S1, 5 qubits, (c) S2, 5 qubits, (d) S1, 7 qubits and (e) S2, 7 qubits. The S2 is not defined when $n = 3$ qubits, as $m \leq \lfloor n/2 \rfloor$.

Source: By the author.

7 CONCLUDING REMARKS

The research field of Quantum Technologies is intensely changing and new proposals have been dissected and explored to understand which are the viable future directions. This intensity can be very challenging, as new scientific material is constantly being released, while we observe the rise and fall of methods. The main motivation of our work was the characterization of a building block, the Parameterized Quantum Circuits, in one of these recently proposed methods: Variational Quantum Algorithms. We considered different circuit structures based on the capabilities of available quantum computers in the contemporary NISQ era, encoded in the connectivites between qubits. To compare these connectivities, we chose the expressibility and entanglement capability (mean entanglement) quantifiers, proposed and well known in the VQAs field. It is important to highlight that these quantifiers characterize the static behavior of these circuits. We are not considering the dynamical behavior generated when the circuit is trained, a characteristic of VQA. Still, our characterizations are of great importance for the dynamical behavior, as they can be correlated with the algorithm performance.

We began the thesis by discussing and defining the main properties of uniformly distributed states and entanglement, ingredients needed to understand the quantifiers. These chapters pointed out how the generation of states with these properties is very useful in different contexts and paved the way to the definition of the quantifiers used in the work.

Following, the results chapter started with the definitions of the circuits chosen for this work, Ansatz 1 and 2, which differ only in the local gates and parameterization after entangling operations. We then presented the expressibility and entanglement of Ansatz 1. A saturation of the expressibility as a function of the number of layers for the no connections circuit was observed, due to the inability of generating entanglement. A striking result is that the values observed for the no connections at only 1 layer is comparable to the connected topologies, even without entanglement generation. At this point, there is no hierarchy between the different topologies for every studied dimension. When inspected in detail, the connected topologies evolve to zero very closely in the smaller dimensions of 3 to 5 qubits. However, with increased number of qubits, 6 to 8, a hierarchy between topologies is established. The ring circuits evolves faster to $D_{KL} = 0$ nats, followed by linear/all-to-all almost together and finally star. The expressibility saturated in different values depending on the number of qubits. This is noticed for 3 to 6 qubits, but it seems that it would occur also for 7 and 8 if additional layers were executed.

To argue these differences, we analysed the mean entanglement, quantified using the Meyer-Wallach or Scott 1, generated by these connected topologies. The Scott 2 was

not explicitly applied in this part, as the results showed the observations are very similar as a function of the number of layers. Comparing to the mean entanglement presented by uniformly random generated states (Circular Unitary Ensemble), we noticed that the closer the mean entanglement of a topology is to the CUE value, the steepest is its evolution to maximum expressibility. However, how close these values are is independent of the dimension of the system, contrasting with the evolution of expressibilities that presents a hierarchy only at higher dimensions. We decided to have a closer look at the standard deviation of the mean entanglement generated by the circuits. The circuits generating entanglement closer to the CUE mean also generate standard deviations closer to the CUE ones, following the same hierarchy. At lower dimensions, the entanglement of uniformly distributed states has standard deviations that are not too much concentrated around the mean, so there is some freedom in the entanglement values the states can possibly have. Increasing the dimension, the mean also increases and, due to the concentration of measure phenomenon, the standard deviation exponentially decreases. This way, the states generated by the CUE are highly entangled and highly concentrated around the mean. This, considering higher dimensions, will avoid circuits generating entanglement that is not high and concentrated around the mean to evolve to the real uniformly distributed states, slowing the expressibility evolution as a function of the number of layers. Using the frame potentials, we showed that the entanglement generation is not connected to the circuits being closer or farther to a $2-$design, removing possible doubts about a possible initial biased randomness of the ring circuit. So the hierarchy ring, linear/all-to-all and then star can be justified by entanglement characteristics of the circuits. We need to stress that this correlation between entanglement and expressibility needs to be studied also from the viewpoint of other complexity/comparison to Haar quantifiers, so we can argue that there is no bias coming from the particular quantifier applied in this work. Also, further studies could apply tools from pseudorandom circuits characterization, including analytical rigorous proofs (40, 104, 106, 107), explicitly considering generation of entanglement for each concatenation, a hypothesis not applied yet to the best of our knowledge. The study of other characteristics, as scrambling (39) or magic states generation (113), or different quantum correlations can be a very interesting future direction.

Ansatz 2 was introduced to understand which are the influences of additional local gates and parameters to expressibility when considering the same entanglement as in Ansatz 1. The entanglement plots showed that they indeed generate the same entanglement. We saw that the connected topologies will have lower values of relative entropy already at 1 layer, because the local parameterization allows to build entangled states with different local coherences, when compared to Ansatz 1. Also, the hierarchy between topologies is established already at 1 layer for Ansatz 2, with circuits generating entanglement closer to the CUE values having a steeper increase of expressibility. These results show that generating entanglement closer to CUE is not useful for smaller dimensions if

we do not have additional freedom on building local superpositions, but is very influential if we do.

When comparing the entanglement values of the different circuits with the chosen paradigmatic states $|EPR_n\rangle$, $|GHZ_n\rangle$ and $|W_n\rangle$, we applied both Scott 1 and 2 to break the degeneracies S1 can present. On the one hand, the EPR_{even} and GHZ_n were not very useful to understand the circuits characteristics, as the S1 and S2 quantifiers do not agree with the identification of the states entanglement values with the circuits ones. On the other hand, the W_n state has properties very close to the star circuit entanglement, a conclusion provided by both quantifiers at 1 layer. Increasing to 2 layers and considering odd number of qubits, the star circuit gets closer to the EPR_{odd} entanglement. Also, EPR_{odd} is very closely connected to the linear/all-to-all circuits entanglement at only one layer for every odd dimensions studied. Finally, the ring circuit, whose entanglement is the closest to the CUE, cannot be identified with any of the paradigmatic states.

In the second chapter of this thesis, we discussed how highly expressible quantum circuits can harm parameter optimization in VQAs. One of the contributions of our work is to the choice of possible connectivities between qubits that can generate entanglement and also present limited expressibility, so the optimization is not too damaged. For example, we saw that the star circuit have the slowest convergence to the Haar case as a function of the number of layers. This way, choosing an architecture with star topology, we can increase the number of parameters and the freedom of optimization and, at the same time, alleviate the drawback of high expressibility values. In terms of the capability of generating uniformly distributed states, our results also show that at only 1 layer the no connections circuit will have a performance close to connected topologies, being a less expensive alternative.

Even though the initial motivation of our work was the characterization of PQCs for applications in VQAs, the calculations led us to the realization that these circuits can replicate properties of pseudorandom circuits when we look at the generated unitaries and states by sampling the parameters randomly. This way, the value of our results is not restricted to this family of algorithms, but can be applied in contexts where random circuits are needed. The advantage of PQCs over known pseudorandom circuits is that they do not rely on changing the order of application and qubits involved in the execution of the gate set. This way, using a fixed structure circuit and classically sampling the parameters, we can generated pseudorandom states.

In our next steps, we seek to characterize these circuits when considering different probability distributions for parameters sampling (e.g., exponential, normal, q-exponential (114)), from the point of view of expressibility and average entanglement to understand how they will affect the random states. These results will be compared with a different tool applied for the characterization of pseudorandom circuits: The majorization method

(115). We plan to explore how PQCs will perform in the context of circuit based quantum reservoir computing, that relies on pseudorandom circuits to generate random states whose expectation values for some chosen local observables are given as inputs to a classical neural network. The neural network is then optimized to accomplish a task.

REFERENCES

1 EINSTEIN, A.; PODOLSKY, B.; ROSEN, N. Can quantum-mechanical description of physical reality be considered complete? **Physical Review**, v. 47, p. 777, May 1935. DOI: 10.1103/PhysRev.47.777.

2 BELL, J. S. On the Einstein Podolsky Rosen paradox. **Physics Physique Fizika**, v. 1, p. 195, Nov. 1964. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

3 CLAUSER, J. F. *et al.* Proposed experiment to test local hidden-variable theories. **Physical Review Letters**, v. 23, p. 880, Oct. 1969. DOI: 10.1103/PhysRevLett.23.880.

4 ASPECT, A.; DALIBARD, J.; ROGER, G. Experimental test of Bell's inequalities using time-varying analyzers. **Physical Review Letters**, v. 49, p. 1804, Dec. 1982. DOI: 10.1103/PhysRevLett.49.1804.

5 PAN, J.-W. *et al.* Experimental entanglement swapping: entangling photons that never interacted. **Physical Review Letters**, v. 80, p. 3891, May 1998. DOI: 10.1103/PhysRevLett.80.3891.

6 DOWLING, J. P.; MILBURN, G. J. Quantum technology: the second quantum revolution. **Philosophical Transactions of the Royal Society of London Series A**: mathematical, physical and engineering sciences, v. 361, n. 1809, p. 1655, 2003. DOI: 10.1098/rsta.2003.1227.

7 SHOR, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. **SIAM Review**, v. 41, n. 2, p. 303, 1999. DOI: 10.1137/S0036144598347011.

8 GROVER, L. K. A fast quantum mechanical algorithm for database search. *In*: ANNUAL ACM SYMPOSIUM ON THEORY OF COMPUTING, 28., 1996, Philadelphia. **Proceedings** [...] Philadelphia: ACM, 1996. DOI: 10.1145/237814.237866.

9 HARROW, A. W.; HASSIDIM, A.; LLOYD, S. Quantum algorithm for linear systems of equations. **Physical Review Letters**, v. 103, p. 150502, Oct. 2009. DOI: 10.1103/PhysRevLett.103.150502.

10 FELLOUS-ASIANI, M. *et al.* Limitations in quantum computing from resource constraints. **PRX Quantum**, v. 2, p. 040335, Nov. 2021. DOI: 10.1103/PRXQuantum.2.040335.

11 PRESKILL, J. Quantum computing in the NISQ era and beyond. **Quantum**, v. 2, p. 79, Aug. 2018. DOI: 10.22331/q–2018–08–06–79.

12 CEREZO, M. *et al.* Variational quantum algorithms. **Nature Reviews Physics**, v. 3, n. 9, p. 625–644, Aug. 2021. DOI: 10.1038/s42254–021–00348–9.

13 PERUZZO, A. *et al.* A variational eigenvalue solver on a photonic quantum processor. **Nature Communications**, v. 5, n. 1, p. 4213, 2014. DOI: 10.1038/ncomms5213.

14 VERDON, G. *et al.* **Quantum Hamiltonian-based models and the variational quantum thermalizer algorithm**. 2019. Available at: https://arxiv.org/abs/1910.02071. Accessible at: 15 Nov. 2023.

15 LIU, C.-Y. **Practical quantum search by variational quantum eigensolver on noisy intermediate-scale quantum hardware**. 2023. Available at: https://arxiv.org/abs/2304.03747. Accessible at: 15 Nov. 2023.

16 MAHESHWARI, D.; SIERRA-SOSA, D.; GARCIA-ZAPIRAIN, B. Variational quantum classifier for binary classification: real vs synthetic dataset. **IEEE Access**, v. 10, p. 3705, 2021. DOI: 10.1109/ACCESS.2021.3139323.

17 YAMAMOTO, N. **On the natural gradient for variational quantum eigensolver**. 2019. Available at: https://arxiv.org/abs/1909.05074. Accessible at: 15 Nov. 2023.

18 BERMEJO, P.; AIZPURUA, B.; ORUS, R. **Improving gradient methods via coordinate transformations**: applications to quantum machine learning. Available at: https://arxiv.org/abs/2304.06768. Accessible at: 15 Nov. 2023.

19 CEREZO, M. *et al.* Cost function dependent barren plateaus in shallow parametrized quantum circuits. **Nature Communications**, v. 12, n. 1, p. 1791, 2021. DOI: 10.1038/s41467–021–21728–w.

20 SIM, S.; JOHNSON, P. D.; ASPURU-GUZIK, A. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. **Advanced Quantum Technologies**, v. 2, n. 12, p. 1900070, 2019. DOI: 10.1002/qute.201900070.

21 WIERSEMA, R. *et al.* Exploring entanglement and optimization within the hamiltonian variational ansatz. **PRX Quantum**, v. 1, p. 020319, Dec. 2020. DOI: 10.1103/PRXQuantum.1.020319.

22 LAROCCA, M. *et al.* Group-invariant quantum machine learning. **PRX Quantum**, v. 3, p. 030341, Sept. 2022. DOI: 10.1103/PRXQuantum.3.030341.

23 BALLARIN, M. *et al.* Entanglement entropy production in quantum neural networks. **Quantum**, v. 7, p. 1023, May 2023. DOI: 10.22331/q–2023–05–31–1023.

24 FARHI, E.; GOLDSTONE, J.; GUTMANN, S. **A quantum approximate optimization algorithm**. 2014. Available at: https://arxiv.org/pdf/1411.4028.pdf. Accessible at: 23 Jan. 2023.

25 MCCLEAN, J. R. *et al.* The theory of variational hybrid quantum-classical algorithms. **New Journal of Physics**, v. 18, n. 2, p. 023023, Feb. 2016. DOI: 10.1088/1367–2630/18/2/023023.

26 MCCLEAN, J. R. *et al.* Barren plateaus in quantum neural network training landscapes. **Nature Communications**, v. 9, n. 1, p. 4812, 2018. DOI: 10.1038/s41467–018–07090–4.

27 KOCKUM, A. F. *et al.* **Lecture notes on quantum computing**. 2023. Available at: https://arxiv.org/abs/2311.08445. Accessible at: 22 Nov. 2023.

28 MITARAI, K. *et al.* Quantum circuit learning. **Physical Review A**, v. 98, p. 032309, Sep. 2018. DOI: 10.1103/PhysRevA.98.032309.

29 MANGINI, S. **Variational quantum algorithms for machine learning:** theory and applications. 2023. Thesis (Doctor) – University of Pavia, Pavia, 2023.

30 DRINKO, A. *et al.* **Benchmarking variational quantum eigensolvers for detection of many–body hamiltonian entangled ground state**. In preparation.

31 DOWLING, M. R.; DOHERTY, A. C.; BARTLETT, S. D. Energy as an entanglement witness for quantum many-body systems. **Physical Review A**, v. 70, p. 062113, Dec. 2004. DOI: 10.1103/PhysRevA.70.062113.

32 WANG, S. *et al.* Noise-induced barren plateaus in variational quantum algorithms. **Nature Communications**, v. 12, n. 1, p. 6961, 2021. DOI: 10.1038/s41467–021–27045–6.

33 LEONE, L. *et al.* **On the practical usefulness of the hardware efficient ansatz**. 2022. Available at: https://arxiv.org/pdf/2211.01477.pdf. Accessible at: 23 Jan. 2023.

34 KANDALA, A. *et al.* Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. **Nature**, v. 549, n. 7671, p. 242–246, Sept. 2017. DOI: 10.1038/nature23879.

35 NIELSEN, M. A.; CHUANG, I. L. **Quantum computation and quantum information**. Cambridge: Cambridge University Press, 2000.

36 MARRERO, C. O.; KIEFEROVÁ, M.; WIEBE, N. Entanglement-induced barren plateaus. **PRX Quantum**, v. 2, p. 040316, Oct. 2021. DOI: 10.1103/PRXQuantum.2.040316.

37 MEDINA, I. *et al.* **VQE-inspired optimization for spin chains work extraction**. 2023. Available at: https://arxiv.org/abs/2310.07617. Accessible at: 23 Oct. 2023.

38 KIM, J.; OZ, Y. Entanglement diagnostics for efficient VQA optimization. **Journal of Statistical Mechanics**: theory and experiment, v. 2022, n. 7, p. 073101, Jul. 2022. DOI: 10.1088/1742–5468/ac7791.

39 KIM, J.; OZ, Y.; ROSA, D. Quantum chaos and circuit parameter optimization. **Journal of Statistical Mechanics**: theory and experiment, v. 2023, n. 2, p. 023104, Feb. 2023. DOI: 10.1088/1742–5468/acb52d.

40 WEINSTEIN, Y. S.; BROWN, W. G.; VIOLA, L. Parameters of pseudorandom quantum circuits. **Physical Review A**, v. 78, p. 052332, Nov. 2008. DOI: 10.1103/PhysRevA.78.052332.

41 MEZZADRI, F. How to generate random matrices from the classical compact groups. **Notices of the American Mathematical Society**, v. 54, n. 5, p. 592 – 604, May 2007.

42 LIVAN, G.; NOVAES, M.; VIVO, P. **Introduction to random matrices**: theory and practice. Berlin: Springer International Publishing, 2018. (Springer briefs in mathematical physics). ISBN 9783319708850.

43 ROBERTS, D. A.; YOSHIDA, B. Chaos and complexity by design. **Journal of High Energy Physics**, v. 2017, Apr. 2017. DOI: 10.1007/JHEP04(2017)121.

44 BARATA, J. C. A. **Notas para um curso de física-matemática**. 2023. Available at: http://denebola.if.usp.br/~jbarata/Notas_de_aula/capitulos.html. Accessible at: 3 Feb. 2023.

45 HAMERMESH, M. **Group theory and its application to physical problems**. Mineola: Dover Publications, 1989. (Addison Wesley series in physics). ISBN 9780486661810.

46 FERREIRA, L. A. **Lecture notes on Lie algebras and Lie groups**. 2022. Available at: https://www.ifsc.usp.br/~laf/gauge_2019/notesfinal.pdf. Accessible at: 3 Apr. 2023.

47 SZEKERES, P. **A course in modern mathematical physics**: groups, Hilbert space and differential geometry. Cambridge: Cambridge University Press, 2004. (A course in modern mathematical physics: groups, Hilbert space, and differential geometry). ISBN 9780521829601.

48 MELE, A. A. **Introduction to Haar measure tools in quantum information**: a Beginner's tutorial. 2023. Available at: https://arxiv.org/pdf/2307.08956.pdf. Accessible at: 30 Sept. 2023.

49 MURPHY, T. G. **Lecture notes** - group representations. 2023. Available at: https://www.maths.tcd.ie/pub/coursework/424/GpReps–II.pdf. Accessible at: 3 Feb. 2023.

50 BARTLETT, S. D.; RUDOLPH, T.; SPEKKENS, R. W. Reference frames, superselection rules, and quantum information. **Reviews of Modern Physics**, v. 79, n. 2, 4 2007. ISSN 0034–6861.

51 POLAND, K.; BEER, K.; OSBORNE, T. J. **No free lunch for quantum machine learning**. 2020. Available at: https://arxiv.org/pdf/2003.14103.pdf. Accessible at: 23 Jan. 2023.

52 SCOTT, A. J. Optimizing quantum process tomography with unitary 2-designs. **Journal of Physics A**: mathematical and theoretical, v. 41, n. 5, p. 055308, Jan 2008. DOI: 10.1088/17518113/41/5/055308.

53 GROSS, D.; AUDENAERT, K.; EISERT, J. Evenly distributed unitaries: on the structure of unitary designs. **Journal of Mathematical Physics**, v. 48, n. 5, May 2007. DOI: 10.1063/1.2716992.

54 DANKERT, C. *et al.* Exact and approximate unitary 2-designs and their application to fidelity estimation. **Physical Review A**, v. 80, n. 1, p. 012304, July 2009. DOI: 10.1103/PhysRevA.80.012304.

55 WILDE, M. M. **Quantum information theory**. 2nd ed. Cambridge: Cambridge University Press, 2017.

56 PRESKILL, J. **Quantum computing and the entanglement frontier**. 2012. Available at: https://arxiv.org/pdf/1203.5813.pdf. Accessible at: 23 Jan. 2023.

57 HORODECKI, R. *et al.* Quantum entanglement. **Reviews of Modern Physics**, v. 81, n. 2, p. 865, 2009.

58 ROBERT, A. **Linear algebra**: examples and applications. Singapore: World Scientific, 2005. (Linear algebra: examples and applications). ISBN 9789812564320.

59 TERHAL, B. M.; HORODECKI, P. Schmidt number for density matrices. **Physical Review A**, v. 61, p. 040301, Mar. 2000. DOI: 10.1103/PhysRevA.61.040301.

60 VIDAL, G. Entanglement monotones. **Journal of Modern Optics**, v. 47, n. 2–3, p. 355–376, 2000. DOI: 10.1080/09500340008244048.

61 CHITAMBAR, E.; GOUR, G. Quantum resource theories. **Reviews of Modern Physics**, v. 91, p. 025001, Apr. 2019. DOI: 10.1103/RevModPhys.91.025001.

62 PLENIO, M. B.; VIRMANI, S. An introduction to entanglement measures. **Quantum Information & Computation**, v. 7, n. 1, p. 1–51, Jan. 2007. ISSN 1533–7146.

63 BENNETT, C. H. *et al.* Concentrating partial entanglement by local operations. **Physical Review A**, v. 53, p. 2046–2052, Apr. 1996. DOI: 10.1103/PhysRevA.53.2046.

64 POPESCU, S.; ROHRLICH, D. Thermodynamics and the measure of entanglement. **Physical Review A**, v. 56, p. R3319–R3321, Nov. 1997. DOI: 10.1103/PhysRevA.56.R3319.

65 BENNETT, C. H. *et al.* Concentrating partial entanglement by local operations. **Physical Review A**, v. 53, p. 2046, Apr. 1996. DOI: 10.1103/PhysRevA.53.2046.

66 VIDAL, G.; CIRAC, J. I. Irreversibility in asymptotic manipulations of entanglement. **Physical Review Letters**, v. 86, p. 5803, June 2001. DOI: 10.1103/PhysRevLett.86.5803.

67 LAMI, L.; REGULA, B. No second law of entanglement manipulation after all. **Nature Physics**, v. 19, n. 2, p. 184, 2023. DOI: 10.1038/s41567–022–01873–9.

68 PETERS, N. A.; WEI, T.-C.; KWIAT, P. G. Mixed-state sensitivity of several quantum-information benchmarks. **Physical Review A**, v. 70, p. 052309, Nov. 2004. DOI: 10.1103/PhysRevA.70.052309.

69 BRENNEN, G. K. An observable measure of entanglement for pure states of multi-qubit systems. **Quantum Information & Computation**, v. 3, n. 6, p. 619–626, Nov. 2003. ISSN 1533–7146.

70 BENNETT, C. H. *et al.* Exact and asymptotic measures of multipartite pure-state entanglement. **Physical Review A**, v. 63, p. 012307, Dec. 2000. DOI: 10.1103/PhysRevA.63.012307.

71 QUANTUM COMPUTING. **Prove that there are infinitely many entanglement classes under LU**. Available at: https://quantumcomputing.stackexchange.com/questions/29575/prove–that–there–are–infinitely–many–entanglement–classes–under–lu. Accessible at: 19 Oct. 2023.

72 DÜR, W.; VIDAL, G.; CIRAC, J. I. Three qubits can be entangled in two inequivalent ways. **Physical Review A**, v. 62, p. 062314, Nov. 2000. DOI: 10.1103/PhysRevA.62.062314.

73 GREENBERGER, D. M.; HORNE, M. A.; ZEILINGER, A. Going beyond Bell's theorem. *In*: KAFATOS, M. (ed.) **Bell's theorem, quantum theory, and conceptions of the universe**. Cham: Kluwer Academic Publishers, 1989. p. 69–72. ISBN 978–90–481–4058–9. DOI: 10.1007/978–94–017–0849–4_10.

74 VERSTRAETE, F. *et al.* Four qubits can be entangled in nine different ways. **Physical Review A**, v. 65, p. 052112, Apr. 2002. DOI: 10.1103/PhysRevA.65.052112.

75 LAMATA, L. *et al.* Inductive entanglement classification of four qubits under stochastic local operations and classical communication. **Physical Review A**, v. 75, p. 022318, Feb. 2007. DOI: 10.1103/PhysRevA.75.022318.

76 LAMATA, L. *et al.* Inductive classification of multipartite entanglement under stochastic local operations and classical communication. **Physical Review A**, v. 74, p. 052336, Nov. 2006. DOI: 10.1103/PhysRevA.74.052336.

77 WONG, A.; CHRISTENSEN, N. Potential multiparticle entanglement measure. **Physical Review A**, v. 63, p. 044301, Mar. 2001. DOI: 10.1103/PhysRevA.63.044301.

78 MINTERT, F. *et al.* Measures and dynamics of entangled states. **Physics Reports**, v. 415, n. 4, p. 207–259, 2005. ISSN 0370–1573.

79 MEYER, D. A.; WALLACH, N. R. Global entanglement in multiparticle systems. **Journal of Mathematical Physics**, v. 43, n. 9, p. 4273–4278, 2002. ISSN 0022–2488. DOI: 10.1063/1.1497700.

80 ROY, A. K. *et al.* Geometric quantification of multiparty entanglement through orthogonality of vectors. **European Physical Journal Plus**, v. 136, n. 11, Nov. 2021. DOI: 10.1140/epjp/s13360–021–02127–y.

81 SCOTT, A. J. Multipartite entanglement, quantum-error-correcting codes, and entangling power of quantum evolutions. **Physical Review A**, v. 69, p. 052330, May 2004. DOI: 10.1103/PhysRevA.69.052330.

82 OLIVEIRA, T. R. de; RIGOLIN, G.; OLIVEIRA, M. C. de. Genuine multipartite entanglement in quantum phase transitions. **Physical Review A**, v. 73, p. 010305, Jan. 2006. DOI: 10.1103/PhysRevA.73.010305.

83 RIGOLIN, G.; OLIVEIRA, T. R. de; OLIVEIRA, M. C. de. Operational classification and quantification of multipartite entangled states. **Physical Review A**, v. 74, p. 022314, Aug. 2006. DOI: 10.1103/PhysRevA.74.022314.

84 WYDERKA, N. **Learning from correlations**: what parts of quantum states tell about the whole. 2020. Thesis (Doctor) — University of Siegen, Siegen, 2020.

85 KIM, J.; OZ, Y. Entanglement diagnostics for efficient VQA optimization. **Journal of Statistical Mechanics**: theory and experiment, v. 2022, n. 7, p. 073101, July 2022. DOI: 10.1088/1742–5468/ac7791.

86 HALL, M. J. Random quantum correlations and density operator distributions. **Physics Letters A**, v. 242, n. 3, p. 123–129, 1998.

87 ZYCZKOWSKI, K.; SOMMERS, H.-J. Induced measures in the space of mixed quantum states. **Journal of Physics A**: mathematical and general, v. 34, n. 35, p. 7111, Aug. 2001. DOI: 10.1088/0305–4470/34/35/335.

88 SPEHNER, D. *et al.* Geometric measures of quantum correlations with Bures and Hellinger distances. *In*: FANCHINI, F.; SOARES PINTO, D.; ADESSO, G. (ed.) **Lectures on general quantum correlations and their applications**. Cham: Springer International Publishing, 2017. p. 105–157. ISBN 978–3–319–53412–1. DOI: 10.1007/978–3–319–53412–1_6.

89 ZYCZKOWSKI, K.; SOMMERS, H.-J. Average fidelity between random quantum states. **Physical Review A**, v. 71, p. 032313, Mar. 2005. DOI: 10.1103/PhysRevA.71.032313.

90 PALMA, G. D. *et al.* The quantum Wasserstein distance of order 1. **IEEE Transactions on Information Theory**, v. 67, n. 10, p. 6627–6643, 2021.

91 HUBREGTSEN, T. *et al.* Evaluation of parameterized quantum circuits: on the relation between classification accuracy, expressibility, and entangling capability. **Quantum Machine Intelligence**, v. 3, p. 1, 2021. DOI: 10.1007/s42484–021–00038–w.

92 ZANARDI, P.; ZALKA, C.; FAORO, L. Entangling power of quantum evolutions. **Physical Review A**, v. 62, p. 030301, Aug. 2000. DOI: 10.1103/PhysRevA.62.030301.

93 ROSS, S. **A first course in probability**. Upper Saddle River: Pearson Prentice Hall, 2010. ISBN 9780136033134.

94 GRIMSLEY, H. R. *et al.* An adaptive variational algorithm for exact molecular simulations on a quantum computer. **Nature Communications**, v. 10, n. 1, p. 3007, 2019.

95 ROSCHER, R. *et al.* Explainable machine learning for scientific insights and discoveries. **IEEE Access**, v. 8, p. 42200–42216, 2020.

96 HOLMES, A. *et al.* Impact of qubit connectivity on quantum algorithm performance. **Quantum Science and Technology**, v. 5, n. 2, p. 025009, Mar. 2020. DOI: 10.1088/2058–9565/ab73e0.

97 HU, W. *et al.* Performance of superconducting quantum computing chips under different architecture designs. **Quantum Information Processing**, v. 21, n. 7, July 2022. DOI: 10.1007/s11128–022–03571–0.

98 CHENG, J. *et al.* **TopGen**: topology-aware bottom-up generator for variational quantum circuits. 2022. Availabe at: https://arxiv.org/abs/2210.08190. Accessible at: 30 Sept. 2023.

99 LEYTON-ORTEGA, V.; PERDOMO-ORTIZ, A.; PERDOMO, O. **Robust implementation of generative modeling with parametrized quantum circuits**. 2019. Available at: https://arxiv.org/abs/1901.08047. Accessible at: 30 Sept. 2023.

100 HAMILTON, K. E.; DUMITRESCU, E. F.; POOSER, R. C. Generative model benchmarks for superconducting qubits. **Physical Review A**, v. 99, p. 062323, June 2019. DOI: 10.1103/PhysRevA.99.062323.

101 XIE, L. *et al.* Suppressing ZZ crosstalk of quantum computers through pulse and scheduling co-optimization. *In*: ARCHITECTURAL SUPPORT FOR PROGRAMMING LANGUAGES AND OPERATING SYSTEMS, 27, 2022, Lausanne. **Proceedings** [...] Lausanne: ACM, 2022. DOI: 10.1145/3503222.3507761.

102 BROWN, K. R.; KIM, J.; MONROE, C. Co-designing a scalable quantum computer with trapped atomic ions. **npj Quantum Information**, v. 2, n. 1, Nov. 2016. DOI: 10.1038/npjqi.2016.34.

103 EMERSON, J. *et al.* Pseudo-random unitary operators for quantum information processing. **Science**, v. 302, n. 5653, p. 2098, 2003. DOI: 10.1126/science.1090790.

104 BROWN, W. G. **Random quantum dynamics**: from random quantum circuits to quantum chaos. 2010. Thesis (Doctor) – Dartmouth College, Hanover, 2010.

105 PARTHASARATHY, K. **Probability measures on metric spaces**. New York: Elsevier Science, 2014. (Probability and mathematical statistics). ISBN 9781483225258.

106 HARROW, A. W.; LOW, R. A. Random quantum circuits are approximate 2-designs. **Communications in Mathematical Physics**, v. 291, p. 257, 2009. DOI: 10.1007/s00220–009–0873–6.

107 BRANDAO, F. G.; HARROW, A. W.; HORODECKI, M. Local random quantum circuits are approximate polynomial-designs. **Communications in Mathematical Physics**, v. 346, p. 397, 2016. DOI: 10.1007/s00220–016–2706–8.

108 SCOTT, A. J.; CAVES, C. M. Entangling power of the quantum baker's map. **Journal of Physics A**: mathematical and general, v. 36, n. 36, Aug. 2003. DOI: 10.1088/0305–4470/36/36/308.

109 BOHIGAS, O.; GIANNONI, M. J.; SCHMIT, C. Characterization of chaotic quantum spectra and universality of level fluctuation laws. **Physical Review Letters**, v. 52, p. 1, Jan. 1984. DOI: 10.1103/PhysRevLett.52.1.

110 FISHER, M. P. *et al.* Random quantum circuits. **Annual Review of Condensed Matter Physics**, v. 14, n. 1, p. 335, 2023. DOI: 10.1146/annurev–conmatphys–031720–030658.

111 LIU, Z.-W. *et al.* Entanglement, quantum randomness, and complexity beyond scrambling. **Journal of High Energy Physics**, v. 2018, n. 7, p. 1, 2018. DOI: 10.1007/JHEP07(2018)041.

112 HAYDEN, P.; LEUNG, D. W.; WINTER, A. Aspects of generic entanglement. **Communications in Mathematical Physics**, v. 265, p. 95, 2006. DOI: 10.1007/s00220–006–1535–6.

113 ZHOU, S. *et al.* Single T gate in a Clifford circuit drives transition to universal entanglement spectrum statistics. **SciPost Physics**, v. 9, n. 6, p. 087, 2020. DOI: 10.21468/SciPostPhys.9.6.087.

114 NAVES, C. B. *et al.* Quantum walks in two dimensions: controlling directional spreading with entangling coins and tunable disordered step operator. **Journal of Physics A**: mathematical and general, v. 56, n. 12, Mar. 2023. DOI: 10.1088/1751–8121/acbd25.

115 VALLEJOS, R. O.; MELO, F. de; CARLO, G. G. Principle of majorization: application to random quantum circuits. **Physical Review A**, v. 104, p. 012602, July 2021. DOI: 10.1103/PhysRevA.104.012602.

116 HARROW, A. W. **The church of the symmetric subspace**. 2013. Available at: https://arxiv.org/abs/1308.6595. Accessible at: 30 Sept. 2023.

117 BOHDANOWICZ, T. **Seminar in computer science**: quantum de Finetti theorems. 2014. Available at: http://users.cms.caltech.edu/~vidick/teaching/286_qPCP/lecture13.pdf. Accessible at: 12 Sept. 2023.

118 FRIEDBERG, S.; INSEL, A.; SPENCE, L. **Linear algebra**. Englewood Cliffs: Pearson Education, 2003. (Featured titles for linear algebra advanced series). ISBN 9780130084514.

119 VOROBETS, Y. **Linear algebra II**: operator of orthogonal projection. 2012. Available at: https://www.math.tamu.edu/~yvorobet/MATH423–2012A/Lect3–10web.pdf. Accessible at: 12. Sept. 2023.

APPENDIX

APPENDIX A – SYMMETRIC SUBSPACES AND ORTHOGONAL PROJECTORS

This appendix is devoted to obtain an important relation between states generated by the invariant Haar measure and symmetric subspaces of a particular Hilbert space, applied in Sec. 5.1. Going beyond the importance of this result, the discussion here connects invariance properties of measures with orthogonal projectors. The results presented here are based on the work of Harrow (116) and Bohdanowicz (117) about symmetric subspaces, with extra detail in the calculations. The reference article begins with the particular case of symmetric groups, however, as we already introduced the invariant Haar measure for compact groups and its properties in Sec. 3.1, we are going to start with the general case and then restrict to the particularities of symmetric groups.

We begin by defining and characterizing orthogonal projectors according to the lectures of Vorobets and based on the Friedberg, Insel, Spence linear algebra book (118, 119).

Definition 8. Let W be an inner product space and $V \subset W$ a subspace such that $W = V \oplus V^{\perp}$[1]. We define the orthogonal projector onto V, Π_V, as the operator that satisfies

$$\begin{gathered} \forall \vec{x} \in W \exists \vec{p} \in V,\ \vec{o} \in V^{\perp}, \\ \vec{x} = \vec{p} + \vec{o}; \end{gathered} \tag{A.1}$$

then

$$\Pi_V(\vec{x}) := \vec{p}. \tag{A.2}$$

The defined orthogonal projector operator satisfies linearity conditions. We should be aware that this definition is a restriction of the one posed for general projector operators, for which the space is decomposed as a direct sum of two subspaces not necessarily orthogonal to each other. In the next proposition we are going to proof three characteristics of the orthogonal projector operators, the first two valid for general projectors and the last one due to the orthogonal direct sum decomposition.

Proposition 2. The orthogonal projector operator Π_V satisfies the conditions:

1. The range[2] of Π_V is V and the kernel[3] is $V^{\perp}$;

[1] Direct sum, namely $W = V + V^{\perp}$ and $V \cap V^{\perp} = \varnothing$.

[2] The range of a linear operator is defined as $\text{range}(L) := \text{span}\,[\text{columns}(L)]$, i.e., the space obtained from linear combinations of the column vectors of the operator (58).

[3] Given a linear operator between vector spaces V and W, $L : V \to W$, the kernel is defined as $\ker(L) := \{\vec{v} \in V | L(\vec{v}) = \vec{0}, \vec{0} \in W\}$ (58).

2. Π_V is idempotent, namely $\Pi_V^2 = \Pi_V$;

3. Π_V is self-adjoint, $\Pi_V^\dagger = \Pi_V$.

Proof. By definition, Π_V is zero when restricted to vectors $\in V^\perp$, therefore $\ker(\Pi_V) = V^\perp$. Again, by definition, the image[4] of the projector $\mathrm{Im}(\Pi_V) \subseteq V$. Taking any $\vec{v} \in V$, $\Pi_V(\vec{v}) = \vec{v} \Rightarrow V \subseteq \mathrm{Im}(\Pi_V)$. This way, $\mathrm{Im}(\Pi_V) = \mathrm{range}(\Pi_V) = V$. Property 2 can be proven taking any $\vec{x} = \vec{p} + \vec{o} \in W$, where $\vec{x} \in V$ and $\vec{o} \in V^\perp$, and applying Π_V twice

$$\Pi_V^2(\vec{x}) = \Pi_V(\vec{p}) = \Pi_V(\vec{x}) \Rightarrow \Pi_V^2 = \Pi_V. \tag{A.3}$$

Notice that we only considered up to now that the operator generates vectors inside one subspace of a direct sum decomposition of the space W, so the inner product structure was not necessary. Now, we prove property 3 by assuming that the subspaces are orthogonal. Take any two vectors $\vec{x}, \vec{y} \in W$, decomposed as $\vec{x} = \vec{p}_1 + \vec{0}_1$ and $\vec{x} = \vec{p}_2 + \vec{0}_2$. If $\langle \cdot, \cdot \rangle$ is an inner product on W, we have that

$$\begin{aligned} \langle \Pi_V(\vec{x}), \vec{y} \rangle &= \langle \vec{p}_1, \vec{p}_2 + \vec{o}_2 \rangle = \langle \vec{p}_1, \vec{p}_2 \rangle; \\ \langle \vec{x}, \Pi_V(\vec{y}) \rangle &= \langle \vec{p}_1 + \vec{o}_1, \vec{p}_2 \rangle = \langle \vec{p}_1, \vec{p}_2 \rangle, \end{aligned} \tag{A.4}$$

therefore, $\langle \Pi_V(\vec{x}), \vec{y} \rangle = \langle \vec{x}, \Pi_V(\vec{y}) \rangle, \forall \vec{x}, \vec{y}$, which is the definition of a self-adjoint operator: $\Pi_V^\dagger = \Pi_V$.

□

Now, we can turn our attention to the orthogonal projectors generated using an invariant measure over the group.

Theorem 4. [5] Let $\mathfrak{G}$ be a group with an invariant normalized measure μ and an unitary representation $R : \mathfrak{G} \to L(V)$, where $L(V)$ are the linear operators on the vector space V. Define

$$V^{\mathfrak{G}} := \{ |\psi\rangle \in V | R(g) |\psi\rangle = |\psi\rangle \, \forall g \in \mathfrak{G} \} \tag{A.5}$$

and

$$\Pi := \int_{\mathfrak{G}} d\mu(x) R(x), \tag{A.6}$$

[4] The image of a linear operator L acting between vector spaces V and W is defined as $\mathrm{Im}(L) := \{L(\vec{v}) | \vec{v} \in V\}$ or the set of vectors obtained by the action of L (58).

[5] In the original work by Harrow (116), the normalization of the measure and the unitarity of the representation are not explicitly required, however both conditions are necessary to complete the proof of the theorem.

then Π is an orthogonal projector onto $V^{\mathfrak{G}}$, satisfying the condition $\Pi\Pi^\dagger = \Pi$.

Proof. To prove this theorem, we need first to show that the operator defined satisfies the characteristics of an orthogonal projector and then show that the projected subspace is indeed $V^{\mathfrak{G}}$. To tackle the first part, we use the invariance property of the measure. Let $R(g)$ be a representation on the vector space V of an element $g \in \mathfrak{G}$

$$R(g)\Pi = \int_{\mathfrak{G}} d\mu(x) R(g)R(x) = \int_{\mathfrak{G}} d\mu(x) R(gx) = \int_{\mathfrak{G}} d\mu(gx) R(gx) = \Pi, \tag{A.7}$$

where we just applied the invariance properties of $d\mu$. The same result is obtained multiplying by any $R(g)$ from the right, as the measure is left and right invariant. Now, notice that

$$\Pi^\dagger\Pi = \int_{\mathfrak{G}} d\mu(x) R(x^{-1})\Pi = \int_{\mathfrak{G}} d\mu(x)\Pi = \Pi \Rightarrow \Pi^\dagger\Pi = \Pi, \tag{A.8}$$

the condition of an orthogonal projector. Here, the unitarity of the representation $R(g)^\dagger = R(g^{-1})$ and normalization of the measure $\int_{\mathfrak{G}} d\mu(x) = 1$ were applied to obtain the final result.

Now we are going to prove that the space where Π projects is $V^{\mathfrak{G}}$. This is easy to do using

$$R(g)\Pi\left|\psi\right\rangle = \Pi\left|\psi\right\rangle, \ \forall \left|\psi\right\rangle \in V. \tag{A.9}$$

This implies that $\mathrm{Im}(\Pi) \subseteq V^{\mathfrak{G}}$. Finally, to show that the other side is also true, we consider a vector $\left|\psi\right\rangle \in V^{\mathfrak{G}}$, then

$$\Pi\left|\psi\right\rangle = \int_{\mathfrak{G}} d\mu(x) R(x)\left|\psi\right\rangle = \left|\psi\right\rangle, \tag{A.10}$$

implying $V^{\mathfrak{G}} \subseteq \mathrm{Im}(\Pi)$. Therefore, $V^{\mathfrak{G}} = \mathrm{Im}(\Pi)$.

□

The most important part of this theorem is that it connects invariant measures over group manifolds (which are always available for compact groups, namely the Haar measure) with orthogonal projectors over group-invariant subspaces. The definition of $V^{\mathfrak{G}}$ is of a space that is invariant under the action of the group, an interesting requirement in many physical situations where symmetries are presented. One example of application of these properties was presented in Sec. 3.1 with the so-called G-twirling, $\mathcal{G}[\rho] := \int_{\mathfrak{G}} d\mu(x) U(x)\rho U^\dagger(x) = \Pi\rho\Pi$. This operation symmetrizes the state under the action of the group, turning it into an invariant density matrix (50).

An important symmetry for many purposes is the invariance under permutation of n qudits[6], i.e., invariance under the symmetric group of n elements, S_n. The group invariant subspace is called symmetric subspace in this case. We are going to define a unitary representation, the orthogonal projector onto the symmetric subspace and then deduce some characteristics of these spaces. At the end of the complete discussion, we are going to prove the theorem relating Haar measures and symmetric spaces.

Let $(\mathbb{C}^d)^{\otimes n}$ be the state space of a $n-$qudit system and S_n be the symmetric group of n elements. Considering a permutation $\pi \in S_n$, we define a unitary representation of this group element in the space $(\mathbb{C}^d)^{\otimes n}$ as

$$P_d(\pi) := \sum_{i_1,i_2,\ldots,i_n \in [d]} \left| i_{\pi(1)}, \ldots, i_{\pi(n)} \right\rangle\!\!\left\langle i_1, \ldots, i_n \right|, \tag{A.11}$$

being $[d] = \{1, \ldots, d\}$, $\pi(\cdot)$ the action of the permutation on the index of a qudit and $|i_1, \ldots, i_n\rangle$ an element of the qudit space basis. The orthogonal projector obtained in Theorem 4 will be in this case

$$\Pi_{sym}^{d,n} = \frac{1}{n!} \sum_{\pi \in S_n} P_d(\pi), \tag{A.12}$$

because the invariant Haar measure for finite discrete groups is just $1/|\mathfrak{G}|$, $|\mathfrak{G}|$ number of elements in the group. The subspace invariant under the action of the group, $V^{\mathfrak{G}}$, is the subspace invariant under qudits permutations and will be named symmetric subspace of $(\mathbb{C}^d)^{\otimes n}$, the $\vee^n \mathbb{C}^d$. The symbol $\vee$ denotes the symmetric tensor product, in opposition to the antisymmetric product denoted by $\wedge$.

We are going to state a result about how the symmetric subspace can be generated as a vector space, built from two different yet equivalent approaches. The proof require some work and will not be provided here as it does not have important connections with our main objective. For more details, the interested reader can consult (116). Consider the vector space defined as

$$A := \text{span}\{|\phi\rangle^{\otimes n} \,|\, |\phi\rangle \in \mathbb{C}^d\}, \tag{A.13}$$

where $|\phi\rangle$ is an arbitrary state. The generated space satisfies $A = \vee^n \mathbb{C}^d$. A simple example to clarify the meaning is the 2-qubits case, where

$$A = \text{span}\{|\phi\rangle^{\otimes 2} \,|\, |\phi\rangle \in \mathbb{C}^2\} = \text{span}\{(a\,|0\rangle + b\,|1\rangle)^{\otimes 2}\}. \tag{A.14}$$

[6] Despite our work is completely devoted to qubits, we chose to keep the original approach in (116) and talk about the general case of qudits.

The symmetric subspace in the 2-qubits case is generated by the triplet states $|00\rangle$, $|11\rangle$ and $1/\sqrt{2}(|01\rangle + |10\rangle)$, which are trivially generated considering the set A in this case. If we set $a = 0$ or $b = 0$, we build the first two states. Considering the superposition of $(|0\rangle + i\,|1\rangle)^{\otimes 2}$ and $(i\,|0\rangle + |1\rangle)^{\otimes 2}$ we build the last one, up to a normalization constant. Notice that the space A is constructed already with permutation invariant states.

Now, we are going to provide another characterization for the symmetric subspace. The first one will be valuable in demonstrations and this second will be interesting for calculating the dimension of this subspace. Let $\mathbb{Z}_+$ be the non-negative integers. We define

$$\mathcal{I}_{d,n} := \{\vec{t} = (t_1, \ldots, t_d) | t_i \in \mathbb{Z}_+,\ \sum_{i=1}^{d} t_i = n\} \tag{A.15}$$

as the set of types for a given $n-$qudits space. Every type define a set of states that have a specific number of qudits in each one of the $0, 1, \ldots, d$ states. For example, considering $d = 3$ and $n = 4$, the type $\vec{t} = (3, 0, 1)$ specifies the set of states

$$\{|2000\rangle, |0200\rangle, |0020\rangle, |0002\rangle\} \tag{A.16}$$

where the possible states are $\{|0\rangle, |1\rangle, |2\rangle\}$ for each qutrit. The number of states that can be built from a given type $\vec{t}$ is given by the multinomial coefficient $\binom{n}{\vec{t}} = \frac{n!}{t_1!\ldots t_d!}$, i.e., in how many ways we can arrange n qudits considering that there are t_1 qudits in state 1, t_2 qudits in state 2 and so on. There is a possible resemblance between the types and Fock states, however we have to be aware that types are not vectors defined in a Hilbert space.

If $\mathcal{J}(\vec{t})$ is the set of states with the same type $\vec{t}$, we can build the permutation invariant vector

$$|s_{\vec{t}}\rangle := \sqrt{\binom{n}{\vec{t}}} \sum_{\mathcal{J}(\vec{t})} |i_1, \ldots, i_n\rangle, \tag{A.17}$$

which is invariant because it is an equally weighted superposition of states with the same type, therefore permutations will take one state to the other, and finally the space

$$B := \text{span}\{|s_{\vec{t}}\rangle\, |\vec{t} \in \mathcal{I}_{d,n}\} \tag{A.18}$$

that is again the same as the symmetric subspace, $B = \vee^n \mathbb{C}^d$. This space is less obvious than A, but it is simple to understand why it is a construction of $\vee^n \mathbb{C}^d$: by definition, the vectors in B are symmetric, therefore $B \subseteq \vee^n \mathbb{C}^d$. If we act with the orthogonal projector $\Pi_{sym}^{d,n}$ on an arbitrary state, we are going to generate a superposition of all the different vectors having the same type $\mathcal{J}(\vec{t})$, which is proportional to $\left|s_{\mathcal{J}(\vec{t})}\right\rangle$. This way, $\text{Im}(\Pi_{sym}^{d,n}) \subseteq B \Rightarrow \vee^n \mathbb{C}^d = B$.

By the construction of space B we can determine the dimension of the symmetric subspace. As B is the space spanned by the $\{|s_{\vec{t}}\rangle\}\forall\,\vec{t}\in\mathcal{I}_{d,n}$, we just have to determine the number of possible $\vec{t}$ considering the space $(\mathbb{C}^d)^{\otimes n}$. A type defines the number of qudits in each of the d qudit states considering n qudits, therefore the number of types is the same as the number of ways n indistinguishable balls can be arranged into d boxes. This problem is solved by the famous method of balls and bars, where $d-1$ bars are needed to generate d boxes and we consider the total number of balls: in how many ways is it possible to arrange $n+d-1$ elements into sets of size $d-1$ (the elements that are going to be divisories)? The answer is simply

$$\dim[\vee^n(\mathbb{C}^d)] = \binom{n+d-1}{d-1}. \tag{A.19}$$

Now, the A space construction is going to be applied to prove the following lemma, needed into the proof of the most important result of this section.

Lemma 1. Considering the d-dimensional unitary group acting on $\mathbb{C}^d$, $\mathcal{U}(d)$, the map $U\mapsto U^{\otimes n}$, $U\in\mathcal{U}(d)$, induces an irreducible representation on the space $\vee^n\mathbb{C}^d$.[7]

Proof. Consider vectors $|\psi_1\rangle, |\psi_2\rangle \in \vee^n\mathbb{C}^d$ and $|\varphi_1\rangle, |\varphi_2\rangle \in \mathbb{C}^d$, from which the vectors $|\varphi_1\rangle^{\otimes n}, |\varphi_2\rangle^{\otimes n}$ are built. The vectors are chosen to satisfy

$$\begin{cases} \langle\psi_1|\cdot|\varphi_1\rangle^{\otimes n} \neq 0 \\ \langle\psi_2|\cdot|\varphi_2\rangle^{\otimes n} \neq 0 \end{cases} \tag{A.20}$$

and the conditions are always possible to be satisfied, as these last two vectors are in $\vee^n\mathbb{C}^d$ by the construction of space A. Now, define unitaries V_i drawn uniformly at random from the set of unitaries satisfying $V_i|\varphi_i\rangle = |\varphi_i\rangle$, $i=1,2$. These unitaries can be constructed by choosing a random unitary $U_{d-1}\in\mathcal{U}(d-1)$ and embedding it in the space orthogonal to $|\varphi_i\rangle$. Notice that the vector space generated by $|\varphi_i\rangle$ is invariant under the action of V_i and we can use Theorem 4 cleverly. The unitary $V_i^{\otimes n}$ averaged over the group of unitaries will generate the orthogonal projector over the subspace invariant under $V_i^{\otimes n}$, i.e., generated by $|\varphi_i\rangle^{\otimes n}$ or

$$\mathbb{E}_{V_i}(V_i^{\otimes n}) := \int_{\mathcal{U}(d-1)} d\mu V_i^{\otimes n} = (|\varphi_i\rangle\langle\varphi_i|)^{\otimes n},\ i=1,2. \tag{A.21}$$

If we consider another unitary $W\in\mathcal{U}(d)$ that transforms $|\varphi_2\rangle$ to $|\varphi_1\rangle$ as $W|\varphi_2\rangle = |\varphi_1\rangle$, we can build unitaries U that have the same action, however rely on the invariance properties

[7] This lemma was rephrased from the original version of Harrow (116) for a better understanding. The original version reads: "$\vee^n\mathbb{C}^d$ is and irreducible representation of $\mathcal{U}_d$ under the action $U\mapsto U^{\otimes n}$".

of V_2 and V_1, namely $U \equiv V_1^\dagger W V_2$. We can calculate an average over these U considering the averages over V_1 and V_2, which gives

$$\mathbb{E}_U(U^{\otimes n}) = \iint_{\mathcal{U}(d-1)} d\mu_1 \mu_2 (V_1 U V_2)^{\otimes n} = \left(|\varphi_1\rangle\langle\varphi_2|\right)^{\otimes n}. \tag{A.22}$$

Now, all the necessary tools to achieve the final result are set. To say that $U^{\otimes n}$ constitute an irreducible representation on the space $\vee^n \mathbb{C}^d$ is the same as saying that there are no invariant subspaces in $\vee^n \mathbb{C}^d$ under the action of $U^{\otimes n}$. This condition can be reflected by $\langle\psi_1| U^{\otimes n} |\psi_2\rangle \neq 0$ for arbitrary vectors $|\psi_1\rangle, |\psi_2\rangle \in \vee^n \mathbb{C}^d$, because the existence of an invariant subspace requires an upper triangular block decomposition of the matrix representation in some basis (45,46), so this condition will avoid the reducibility of the representation. In other words, if there is an ensemble of unitaries U such that, on average, this condition is valid, we achieved the result. Using Eq. (A.22),

$$\mathbb{E}_U(\langle\psi_1| U^{\otimes n} |\psi_2\rangle) = \langle\psi_1| \cdot |\varphi_1\rangle^{\otimes n} \langle\varphi_2|^{\otimes n} \cdot |\psi_2\rangle \neq 0, \tag{A.23}$$

due to the initial definition of vectors $|\varphi_1\rangle, |\varphi_2\rangle$ and $|\psi_1\rangle, |\psi_2\rangle$.

□

We now state and prove the final result.

Theorem 5. If $\mathbb{E}_\varphi$ is the average over a randomly chosen unit vector $|\varphi\rangle \in \mathbb{C}^d$ considering an uniform distribution, then

$$\mathbb{E}_\varphi(\varphi^{\otimes n}) := \int_{Haar} |\varphi\rangle\langle\varphi|^{\otimes n} \, d|\psi\rangle = \frac{\Pi_{sym}^{d,n}}{\mathrm{Tr}\left(\Pi_{sym}^{d,n}\right)} = \frac{\Pi_{sym}^{d,n}}{\dim(\vee^n \mathbb{C}^d)} \tag{A.24}$$

Proof. The operator $\rho = \mathbb{E}_\varphi(\varphi^{\otimes n})$ is "symmetrized" under the action of $U^{\otimes n}$, therefore it is in the invariant subspace and commutes with every $U^{\otimes n}$. By the Schur lemma (45) and Lemma 1, ρ must be proportional to the identity in the symmetric subspace, which is exactly $\Pi_{sym}^{d,n}$, therefore $\rho = \lambda \mathbb{I}_{\vee^n \mathbb{C}^d}$. Requiring that ρ is a physical state with trace equal to one, we find that

$$\lambda = \frac{1}{\dim(\vee^n \mathbb{C}^d)}. \tag{A.25}$$

and, finally, joining the constant λ with the fact that the operator is the orthogonal projector on the symmetric subspace,

$$\mathbb{E}_\varphi(\varphi^{\otimes n}) = \frac{\Pi_{sym}^{d,n}}{\dim(\vee^n \mathbb{C}^d)} \tag{A.26}$$

□

APPENDIX B – SCOTT ENTANGLEMENT MEASURE FOR THE PARADIGMATIC ODD EPR

In this section we are going to calculate the Scott measure in the cases $m = 1$ (Meyer-Wallach) and $m = 2$ for the paradigmatic state EPR_n when the number of qubits is odd. These states will have the form

$$|EPR_n\rangle = |0_1\rangle \otimes \left|\Phi^+\right\rangle^{\otimes(n-1)/2}. \tag{B.1}$$

Before we start the calculations, we are going to state a result for the purity of states in the Schmidt decomposition. If a state is decomposed in the Schmidt basis for the bipartition $\mathcal{H}_A \otimes \mathcal{H}_B$, with $m = \min\left[\dim(\mathcal{H}_A), \dim(\mathcal{H}_B)\right]$ as

$$|\psi\rangle = \sum_{i=1}^{m} \lambda_i \left|i_A\right\rangle \otimes \left|i_B\right\rangle, \tag{B.2}$$

the reduced states for system A and B will have the same eigenvalues in a different basis. If $|\psi\rangle\langle\psi| = \rho$, the reduced state for subsystem B is

$$\rho_B = \mathrm{Tr}_A(\rho) = \sum_{i=1}^{m} \lambda_i^2 \left|i_B\right\rangle\left\langle i_B\right|, \tag{B.3}$$

and analogously for A. This way, the Schmidt coefficients provide all the information we need to characterize the subsystems locally, as they provide a kind of spectral decomposition of the subsystems in the local basis. Also, the reduced states will have identical probabilities. If we take the square of the reduced matrix and then calculate the complete trace of this, we are going to achieve the purity, which is given in terms of the Schmidt coefficients as

$$\mathrm{Pur}(\rho_B) \equiv \mathrm{Tr}\left(\rho_B^2\right) = \sum_{i=1}^{m} \lambda_i^4, \tag{B.4}$$

which we are going to apply in the next results.

B.1 Scott 1 or Meyer-Wallach entanglement measure

In the Meyer-Wallach case, using Eq. (4.50) in terms of the purity of the reduced states, we only need to determine the values of the purities when we trace out one qubit. In the state $|EPR_{odd}\rangle = |0_1\rangle \otimes |\Phi^+\rangle^{\otimes(n-1)/2}$ we only have two possible cases: The separable qubit and a qubit inside of a Bell state, which will all have the same purity by symmetry. For the separable qubit case we can argue that the purity of the reduced state has to be one, as the reduced state in a separable state is pure. Now, we are going to consider

qubits inside Bell states. To do so, a subscript on the ket states will be used to index the subsystems, so we can change the position of the vectors in the tensor product to have a better visualization of the Schmidt decomposition, without losing the information about which is the subsystem we are manipulating. To simplify the notation and without loss of generality, we are going to consider the "first" qubit inside a Bell state, namely qubit 2. Reading

$$\begin{aligned} |EPR_n\rangle &= |0_1\rangle \otimes \frac{1}{\sqrt{2}}(|00_{23}\rangle + |11_{23}\rangle) \otimes \left|\Phi^+\right\rangle^{\otimes(n-1)/2-1} \\ &= \frac{1}{\sqrt{2}}|0_2\rangle \otimes \left(|0_1\rangle \otimes \left|\tilde{\Phi}^+\right\rangle\right) + \frac{1}{\sqrt{2}}|1_2\rangle \otimes \left(|0_2\rangle \otimes \left|\tilde{\Phi}^+\right\rangle\right), \end{aligned} \tag{B.5}$$

where $\tilde{\Phi}$ stands for the product of Bell states in the rest of the qubits not explicitly indexed. This state is already in the Schmidt decomposition considering the bipartition qubit 2 and the rest of the system. This way, we can calculate the purity using the Schmidt coefficients as in Eq. (B.4), to obtain $\mathrm{Pur}(\rho_2) = 1/2$. By the symmetry of the state, there will be $n-1$ qubits in this situation (all of them, except for the separable one), we will have a contribution of $1/2(n-1)$ to the purity sum. Substituting in Eq. (4.50), we obtain

$$Q_1(EPR_n) = 2\left[1 - \frac{1}{n}\sum_{k=1}\mathrm{Tr}\left(\rho_k^2\right)\right] = 2\left[1 - \frac{1}{n}\left(1 + \frac{1}{2}(n-1)\right)\right] = \frac{n-1}{n}. \tag{B.6}$$

B.2 Scott 2

Now, to calculate the Scott measure of order 2 for the state, we are going to consider the decomposition proposed by Rigolin-Oliveira-Oliveira, Eq. (4.53). We will have to consider different possible neighbours, l, for the fixed bipartition cut of 2 qubits and rest of the system. We are going to compute $G(2,l)$ for different l values and show that, in fact, there are only two possibilities. The notation for subscripts here will be the same as in the previous section.

First we consider the case $l = 1$, i.e., bipartitions 2 qubits and the rest of the system where these 2 qubits are first neighbours. When the 2 qubits are the qubits inside one of the Bell states in the decomposition, named here j and $j+1$, we have

$$|EPR_n\rangle = \left|\Phi^+_{j,j+1}\right\rangle \otimes \left(|0_1\rangle \otimes \left|\Phi^+\right\rangle^{\otimes k}\right), \tag{B.7}$$

where the k are every other pair of qubits in a Bell state. Using, again, the Schmidt decomposition, we see that the contribution for the purity sum will be $\mathrm{Pur}(\rho_{j,j+1}) = 1$. There will be $(n-1)/2$ such cases, contributing to the purity sum with $(n-1)/2$. We have also the term analogous to the one in the previous section, where we consider one

qubit from a Bell state and the separable $|0\rangle_1$. We already calculated the contribution for the purity and concluded it will be 1/2. Finally, we will have the case including one qubit from each of two neighbouring Bell states and the rest of the system. If we analyse the case of the first two Bell states for simplicity of notation,

$$\begin{aligned} |EPR_n\rangle &= |0_1\rangle \otimes \frac{(|00_{23}\rangle + |11_{23}\rangle)}{\sqrt{2}} \otimes \frac{(|00_{45}\rangle + |11_{45}\rangle)}{\sqrt{2}} \otimes \left|\Phi^+\right\rangle^{\otimes k} \\ &= \frac{1}{2}\{|00_{34}\rangle \otimes |00_{25}\rangle \otimes |\varphi\rangle + |01_{34}\rangle \otimes |01_{25}\rangle \otimes |\varphi\rangle + \quad \text{(B.8)} \\ &+ |10_{34}\rangle \otimes |10_{25}\rangle \otimes |\varphi\rangle + |11_{34}\rangle \otimes |11_{25}\rangle \otimes |\varphi\rangle\}, \quad \text{(B.9)} \end{aligned}$$

where we defined $|\varphi\rangle \equiv |0_1\rangle \otimes |\Phi^+\rangle^{\otimes k}$. Considering the purity calculated using the Schmidt coefficients, we have $\mathrm{Pur}(\rho_{34}) = 4 \cdot 1/2^4 = 1/4$. There will be $(n-1)/2 - 1$ such cases, contributing with $(n-3)/2 \cdot 1/4 = (n-3)/8$. Putting altogether,

$$\sum_{j=1}^{n-1} \mathrm{Tr}\left(\rho_{j,j+1}^2\right) = \frac{n-1}{2} + \frac{1}{2} + \frac{n-3}{8} = \frac{5n-3}{8}, \quad \text{(B.10)}$$

and, finally,

$$G(2,1) = \frac{4}{3}\left[1 - \frac{1}{n-1}\frac{5n-3}{8}\right] = \frac{3n-5}{6(n-1)}. \quad \text{(B.11)}$$

Now, we can start the other possible case, $l \geq 2$. We already know all the possible values the purity can take, given that the state is symmetric. We only have to count the different possibilities and sum them. There will be

- One case involving the separable $|0_1\rangle$, contributing with 1/2;
- $n-1-l$ possible $l-$th neighbours including qubits from different Bell states, contributing with $(n-1-l)/4$. This can be seen by the fact that there all qubit will have a neighbour l qubits ahead, up to the $n-l$ qubit. However, as this case excludes the first qubit, there will be $n-1-l$ such cases.

Resulting in the purity sum

$$\sum_{j=1}^{n-l} \mathrm{Tr}\left(\rho_{j,j+l}^2\right) = \frac{n-1-l}{4} + \frac{1}{2} = \frac{n+1-l}{4}. \quad \text{(B.12)}$$

and

$$G(2, l \geq 2) = \frac{4}{3}\left[1 - \frac{1}{n-l}\frac{n+1-l}{4}\right] = 1 - \frac{1}{3(n-l)}. \quad \text{(B.13)}$$

We can now turn to Q_2 as presented in Eq. (4.53) and sum everything up

$$Q_2(EPR_n) = \frac{2}{n(n-1)}\left[(n-1)G(2,1) + \sum_{l=2}^{n-1}(n-l)G(2, l \geq 2)\right]; \tag{B.14}$$

the term inside bracket has a part already defined in terms of n, proportional to $G(2,1)$ and a part that depends on l. By performing the sums

$$\sum_{l=2}^{n-1}(n-l)\left(1 - \frac{1}{3(n-l)}\right) = \sum_{l=2}^{n-1}\left(n - \frac{1}{3}\right) - \sum_{l=2}^{n-1} l = \left(n - \frac{1}{3}\right)(n-2) - \frac{n(n-1)-2}{2}. \tag{B.15}$$

and the entanglement measure will finally be

$$Q_2(EPR_n) = \frac{2}{n(n-1)}\left[\frac{5 - 8n + 3n^2}{6}\right] = \frac{3n^2 - 8n + 5}{3n(n-1)} = \frac{3n-5}{3n}. \tag{B.16}$$

APPENDIX C – ADDITIONAL DATA

In this section we are going to present the complementary data of conditions discussed in Chap. 6, but not presented at that point to simplify the results presentation and discussion. The data will not be discussed, as all the features were already treated in the main text chapters.

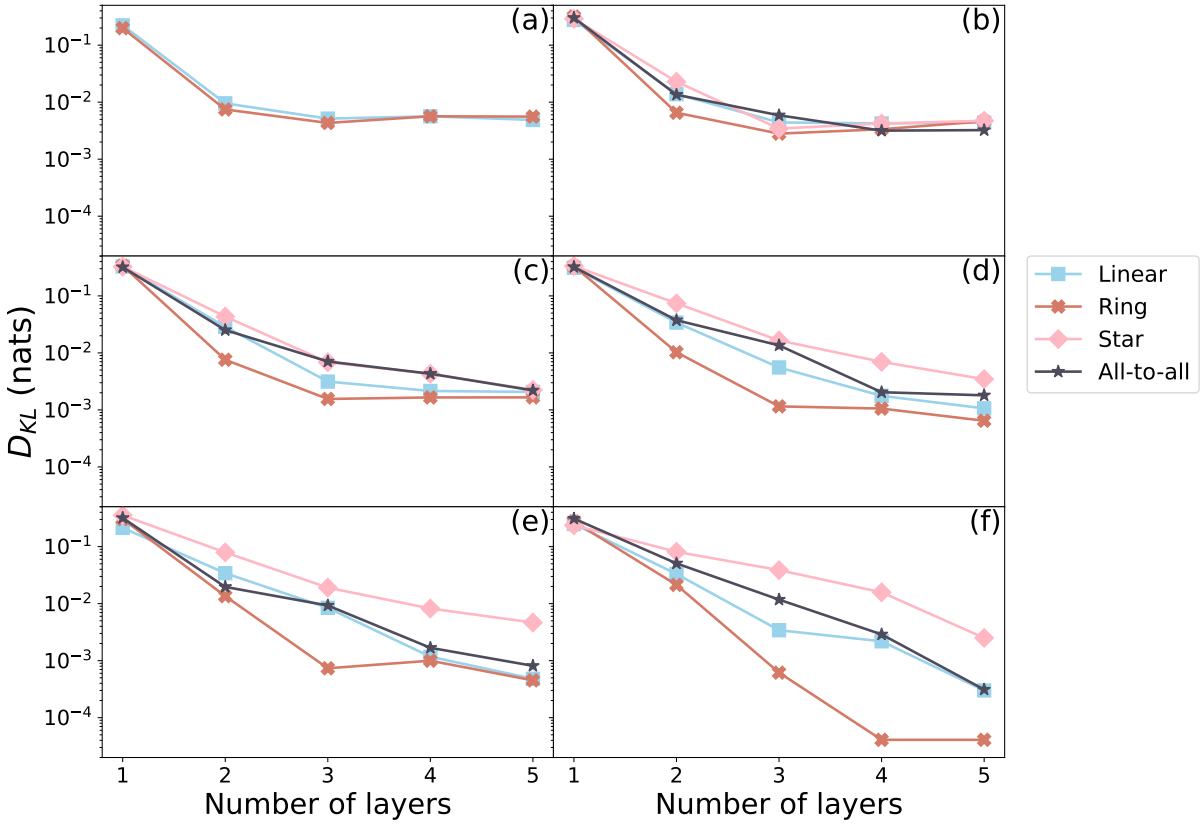

Figure 20 – Ansatz 1: relative entropy in logarithmic scale as a function of the number of layers for all connected connectivities, considering the case of (a) 3 qubits, (b) 4 qubits, (c) 5 qubits, (d) 6 qubits, (e) 7 qubits and (f) 8 qubits.

Source: By the author.

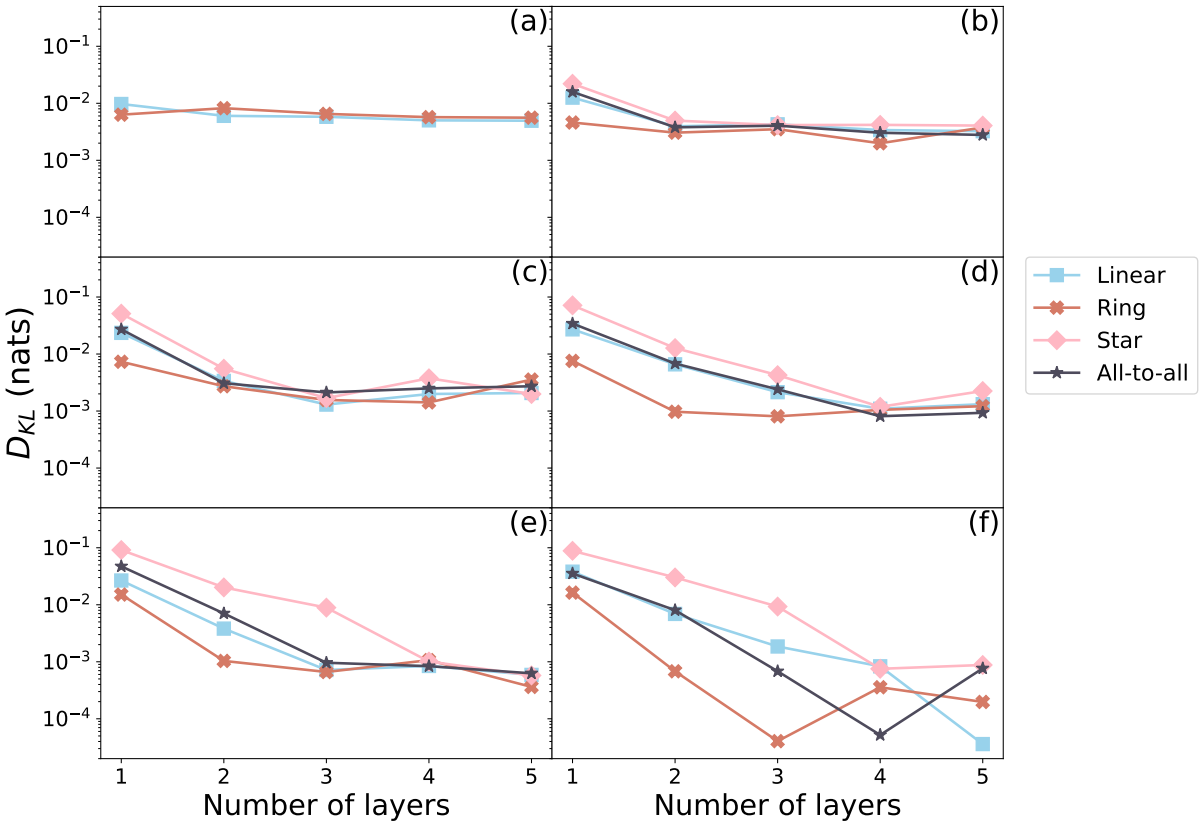

Figure 21 – Ansatz 2: relative entropy in logarithmic scale as a function of the number of layers for all connected connectivities, considering the case of (a) 3 qubits, (b) 4 qubits, (c) 5 qubits, (d) 6 qubits, (e) 7 qubits and (f) 8 qubits.

Source: By the author.

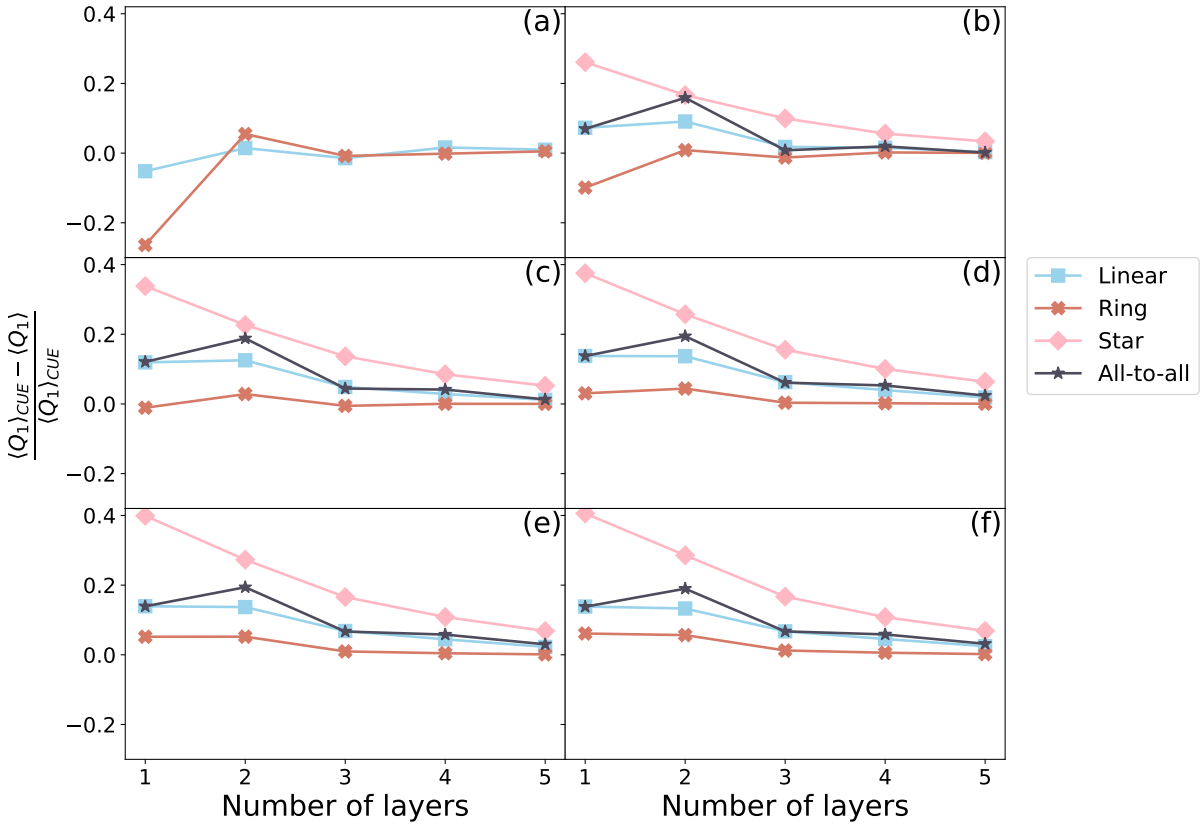

Figure 22 – Ansatz 1: normalized subtraction of the CUE ensemble mean minus the circuit mean of the Meyer-Wallach measure for all connected topologies as a function of the number of layers for (a) 3 qubits, (b) 4 qubits, (c) 5 qubits, (d) 6 qubits, (e) 7 qubits and (f) 8 qubits. The analytical value used for the CUE mean depends on the dimension and is given in Eq. 5.16.

Source: By the author.

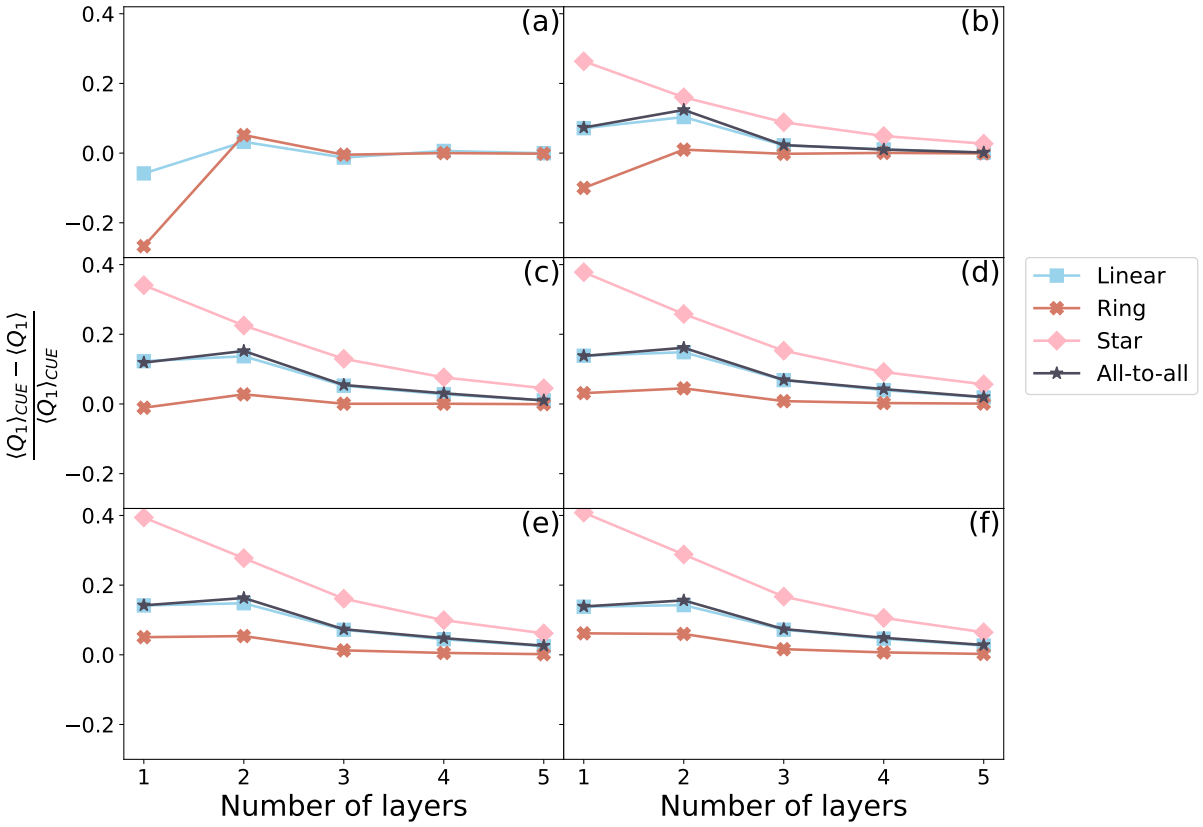

Figure 23 – Ansatz 2: normalized subtraction of the CUE ensemble mean minus the circuit mean of the Meyer-Wallach measure for all connected topologies as a function of the number of layers for (a) 3 qubits, (b) 4 qubits, (c) 5 qubits, (d) 6 qubits, (e) 7 qubits and (f) 8 qubits. The analytical value used for the CUE mean depends on the dimension and is given in Eq. 5.16.

Source: By the author.

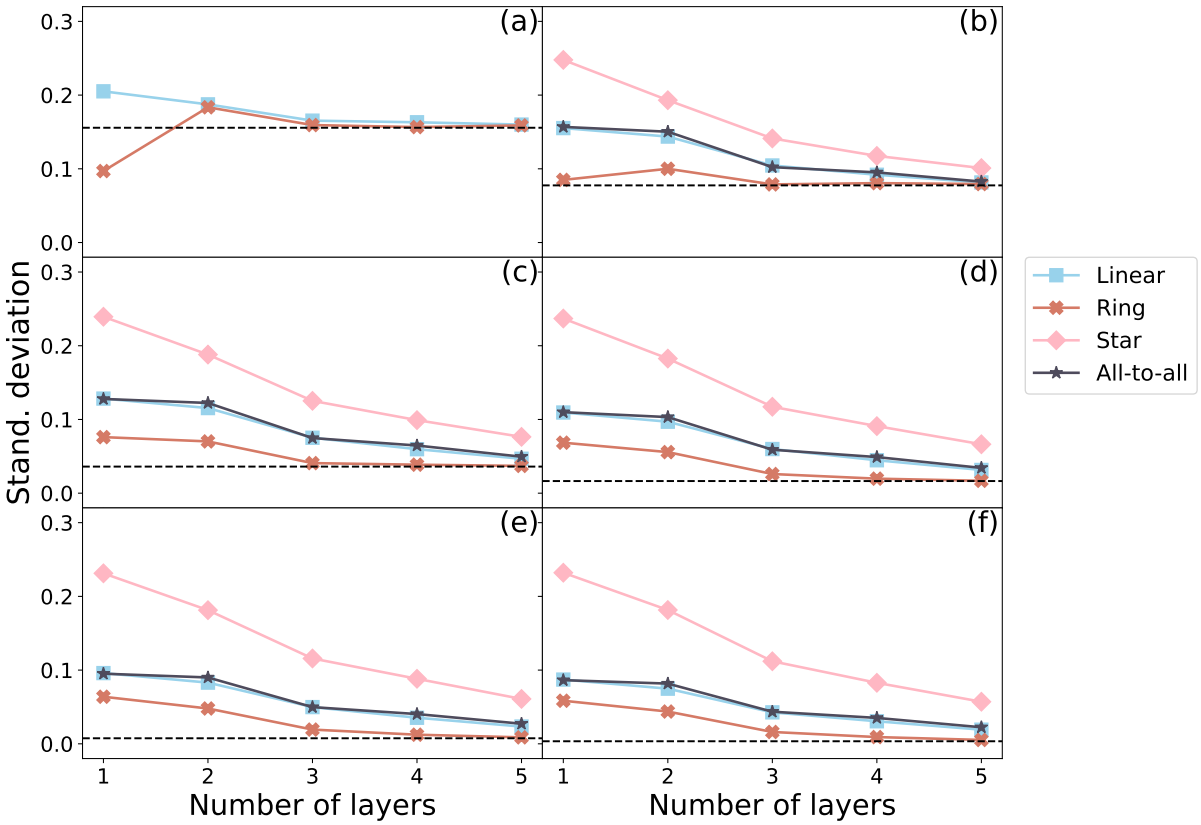

Figure 24 – Ansatz 1: standard deviation of the Meyer-Wallach measure for all connected topologies as a function of the number of layers for (a) 3 qubits, (b) 4 qubits, (c) 5 qubits, (d) 6 qubits, (e) 7 qubits and (f) 8 qubits. The traced line indicates the standard deviation values for the CUE, which depends on the dimension and is given in Eq. 5.17.

Source: By the author.

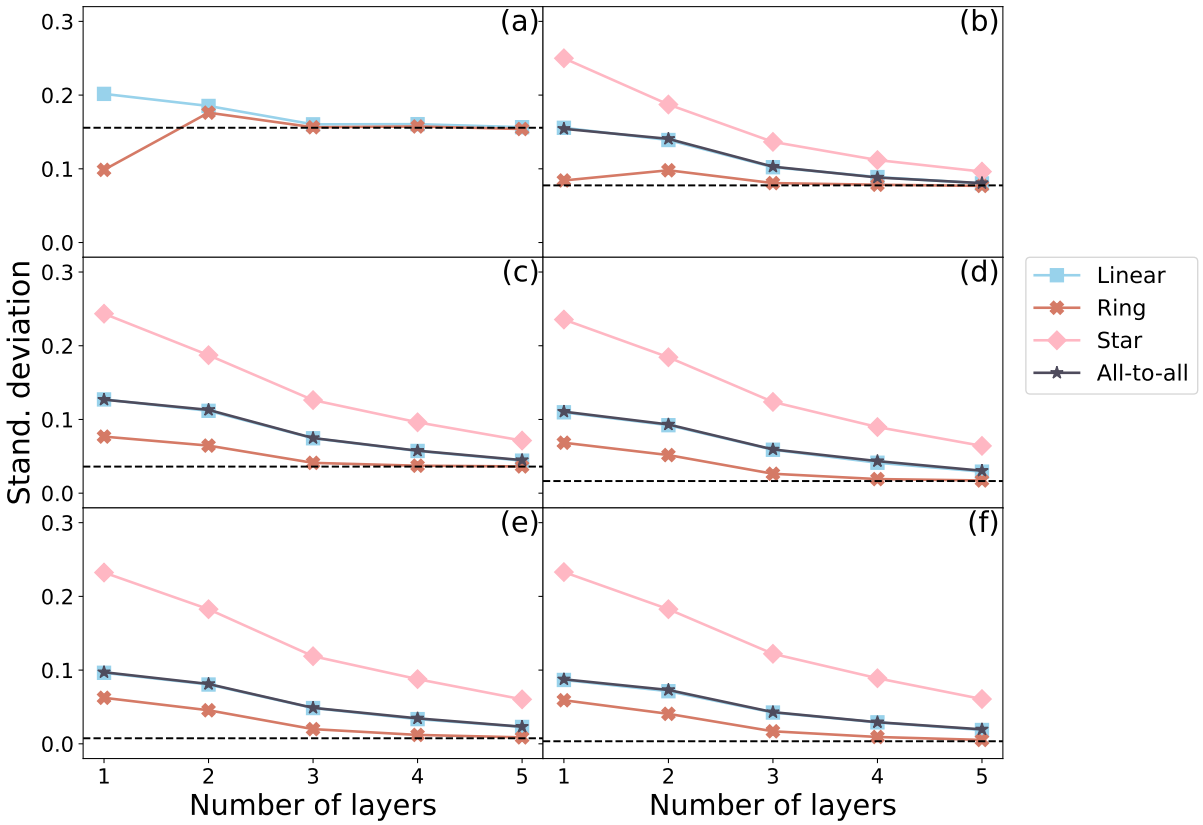

Figure 25 – Ansatz 2: standard deviation of the Meyer-Wallach measure for all connected topologies as a function of the number of layers for (a) 3 qubits, (b) 4 qubits, (c) 5 qubits, (d) 6 qubits, (e) 7 qubits and (f) 8 qubits. The traced line indicates the standard deviation values for the CUE, which depends on the dimension and is given in Eq. 5.17.

Source: By the author.

Table 5 – Ansatz 1 values for the quantity $||\mathcal{A}^{t=2}||_2^2 = F^{(t=2)}_{|0\rangle\langle 0|^{\otimes n}} - F^{(t=2)}_{Haar}$, that compares the circuit with a 2−design, depending on the topology and on the number of qubits, n, for all qubits numbers. The mean values and standard deviations were obtained over 20 independent calculations.

Topology	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$	$n = 8$
No connections	$(1.3 \pm 0.1) \cdot 10^{-2}$	$(6.5 \pm 0.5) \cdot 10^{-3}$	$(3.0 \pm 0.3) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.6 \pm 0.7) \cdot 10^{-4}$	$(1.6 \pm 0.3) \cdot 10^{-4}$
Linear	$(1.3 \pm 0.1) \cdot 10^{-2}$	$(6.6 \pm 0.7) \cdot 10^{-3}$	$(3.0 \pm 0.3) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.5 \pm 0.6) \cdot 10^{-4}$	$(1.7 \pm 0.3) \cdot 10^{-4}$
Ring	$(1.3 \pm 0.1) \cdot 10^{-2}$	$(6.7 \pm 0.6) \cdot 10^{-3}$	$(3.0 \pm 0.3) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.3 \pm 0.6) \cdot 10^{-4}$	$(1.7 \pm 0.3) \cdot 10^{-4}$
Star	–	$(6.8 \pm 0.5) \cdot 10^{-3}$	$(3.0 \pm 0.3) \cdot 10^{-3}$	$(1.2 \pm 0.1) \cdot 10^{-3}$	$(4.4 \pm 0.6) \cdot 10^{-4}$	$(1.6 \pm 0.3) \cdot 10^{-4}$
All-to-all	–	$(6.6 \pm 0.6) \cdot 10^{-3}$	$(3.0 \pm 0.3) \cdot 10^{-3}$	$(1.2 \pm 0.2) \cdot 10^{-3}$	$(4.5 \pm 0.8) \cdot 10^{-4}$	$(1.6 \pm 0.3) \cdot 10^{-4}$

Source: By the author.