A MASTER THESIS FOR THE

MASTER IN QUANTUM SCIENCE AND TECHNOLOGY

# Quantum Benchmarking: entanglement measures in quantum computers

**Author:**
Juan Francisco Martín Bravo

**Supervisor:**
Dra. Alba Cervera-Lierta

- University of Barcelona, July 2023 -

# Quantum Benchmarking: entanglement measures in quantum computers

Juan Francisco Martín Bravo

Supervised by: Dra. Alba Cervera-Lierta

Barcelona Supercomputing Center, 08034 Barcelona
10 July 2023

Quantum computation has emerged as a promising paradigm shift in the field of computing, and with the advent of new quantum computers, it has become crucial to assess and quantify their performance. Benchmarking, a well-established practice in the field, plays a vital role in this regard. One effective way to evaluate a quantum computer's capabilities is by measuring the amount of entanglement it exhibits, as entanglement is a fundamental characteristic of quantum systems. In this thesis, we provide a comprehensive overview of the current landscape of quantum benchmarking and propose several protocols for estimating the Rényi entropy of quantum states, which offers valuable insights into the entanglement structure of these states. We present a protocol based on the renowned Swap test, specifically designed for future fault-tolerant devices, as well as another protocol based on randomized measurements to address the limitations of current NISQ devices. We have implemented these protocols on the quantum simulation framework of Qibo, ensuring an efficient and reliable execution on any quantum computer, in particular the one at the Barcelona Supercomputing Center (BSC). Through this work, we aim to contribute to the advancement of quantum benchmarking and facilitate the assessment of entanglement in quantum computing systems.

*Keywords*: Quantum computing, Entanglement measures, Benchmarking, Rényi entropy, Random matrices.

Juan Francisco Martín Bravo: jmartibr88@alumnes.ub.edu

## Contents

## 1 Introduction

Over the past decade, we have witnessed an astonishing development of quantum technologies promising a paradigm shift in the way classical computers have conventionally addressed our computational problems. Quantum computation has risen up to exploit the principles of quantum physics in order to store and process information. While classical computers rely on bits representing 0s or 1s, quantum devices employ qubits, which manifest two interesting properties: superposition and entanglement.

On the one hand, superposition states that when a physical system has the potential to exist in multiple configurations or states, it can be described by a combination of all these possibilities. In other words, rather than being restricted to a single state like in classical physics, a quantum system can simultaneously occupy different states with specific probabilities assigned to each configuration. On the other hand, entanglement involves creating pairs of qubits that are interconnected in such a way that their states become correlated beyond what classical correlation can describe. When qubits are entangled, the state of one qubit is intrinsically linked to the state of the other, persisting regardless of space. By leveraging the principles of superposition and entanglement, quantum computers hold the potential to solve complex computational problems more efficiently than classical computers. Quantum algorithms, such as Shor's algorithm [1] for integer factorization and Grover's algorithm [2] for database searching, demonstrate the power of qubits in solving problems substantially faster than classical algorithms, providing a remarkable advantage over classical computation.

Building the perfect quantum computer is hard and several challenges must be overcome. Those challenges stem from the inherent nature of the quantum world, as observing a quantum system disrupts its behavior. To exemplify this concept, we can take a look at decoherence. Qubits lose their quantum properties due to environmental interactions, therefore needing protective measures and error correction techniques. Thus, it is imperative to achieve near-perfect isolation from the external environment while enabling strong interactions between qubits. But despite these hurdles, quantum computers hold promise for transforming industries and advancing scientific research, offering unprecedented computational power that could revolutionize scientific discovery and unlock breakthroughs that were once thought to be out of reach.

A significant breakthrough has been the recent appearance of noisy intermediate-scale quantum (NISQ) devices [3], acting as a bridge between the classical realm and the eventual realization of fault-tolerant quantum computers. Rather than being a world-changing technology on its own, NISQ devices should be seen as a step towards more powerful quantum technologies that will be developed in the future. The term "noisy" emphasizes that we will have imperfect control over those qubits, arising noise that will place a serious limitation on what quantum devices can achieve in the near term. Several different quantum platforms fall under the NISQ category, each with its own unique architecture and technology. Some of the commonly known quantum platforms on NISQ devices include: superconducting qubits [4], trapped ion qubits [5], topological qubits [6], photonic quantum computing [7] and quantum dot qubits [8].

The characterization and performance study of NISQ devices has emerged as an important challenge within the field. Fortunately, quantum researchers find solace in the practice of benchmarking, a longstanding tradition within the realm of computation. Similar to traditional classical benchmarking, quantum benchmarking involves running specific tests and algorithms on quantum hardware to evaluate their efficiency, accuracy, and reliability. As quantum technology advances, it becomes increasingly important to have standardized benchmarks that can provide meaningful comparisons across different quantum platforms. Benchmarking helps researchers and developers identify the strengths and weaknesses of quantum systems, make improvements, and optimize their performance. It also enables the evaluation of different quantum algorithms and protocols, contributing to the overall progress and development of quantum computing. By establishing benchmarks, the quantum community can ensure reliable and objective measurements, foster competition, and drive innovation in the field.

This project can be distinctly divided into two parts. First, we will embark on a comprehensive exploration of a variety of certification protocols that hold great significance in the quantum domain, serving as powerful tools for benchmarking the performance of quantum devices. The primary aim of this initial segment is to serve as an introductory encounter, inviting all those intrigued by the recently unfolding and captivating realm of benchmarking and quantum computer characterization. Regardless, for the second part of the work we will adopt a more practical approach. Our attention will be centered on presenting two benchmark protocols that assess the quantum device's capability to generate entanglement in qubits. In particular, we will focus on the estimation of the so-called Rényi entropy (Section 3). The first protocol is relatively modest and intended for future fault-tolerant quantum computers (Section 4). Nevertheless, the second protocol is more complex and it is specifically designed to be effectively applied on current NISQ devices (Section 5). Both protocols will be programmed in *Qibo* [9], an open-source full-stack application programming interface (API) for quantum simulation and quantum hardware control, with the final objective of establishing a library to test quantum entanglement in the future quantum computer located at the Barcelona Supercomputing Center (BSC).

## 2 Quantum Benchmarking

This section will serve as a concise overview of the most commonly used and well-known benchmarks in the field of NISQ computers. We will classify them into three categories according to their level of application in the circuit: gate-level benchmarks, circuit-level benchmarks, and application benchmarks. To elucidate, Figure 1 shows a schematic view of the benchmarks that we will discuss below.

Although all the protocols we are addressing provide some insight into the quantum computer's performance, relying on a single application may be insufficient for testing overall system performance. This is why benchmark suites have been introduced. They allow us to address a variety of different problems in a more simple and easy-access way. A few noteworthy examples include: the Application-Oriented Performance Benchmarks [10, 11]; QASMBench [12, 13]; and SupermarQ [14, 15].

A fundamental and straightforward way to assess the performance of a Quantum Processor Unit (QPU) is by examining its inherent physical features, serving as indicators of its most general capabilities. Important physical metrics for evaluating quantum computers include gate fidelity and readout fidelity, the number of qubits and their interaction, and T1 and T2 times. By considering these metrics together, users can gain insights into the performance of a quantum processor and assess its capabilities based on several parameters.

To begin with, it is important to define the concept of quantum fidelity, as it is a key technique used in the vast majority of quantum benchmarking. The fundamental principle of fidelity estimation is to achieve quantum state verification from only a few measurements, without the necessity of applying full tomography [16] over the desired computed state $\rho$. Many methods have already been presented in order to accomplish this in an efficient and rapid way, without using an exponentially increasing number of measurements [17]. One of

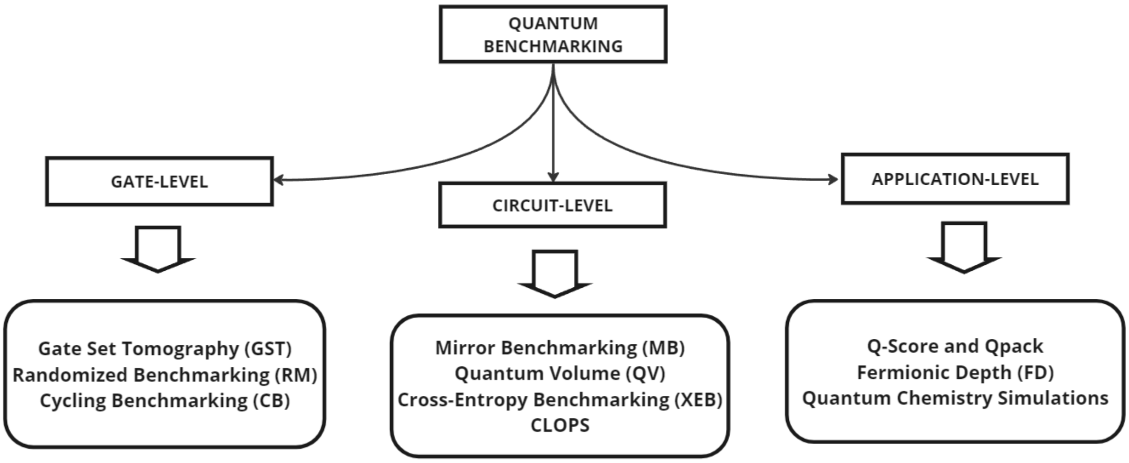

Figure 1: Diagram of the quantum benchmarks we are addressing in this introduction.

its majors drawbacks is its focus on a concrete variety of states, such as Bell states, cluster states, graph states, and Dick states. That being said, many benchmarking protocols use this technique to extract some knowledge about the current quantum state of the device.

Next, the number of qubits plays a crucial role in the complexity and scale of quantum algorithms, thus impacting the computational power and capabilities of the system. Certain quantum algorithms, such as the already mentioned Shor's algorithm [1] and Grover's algorithm [2], exhibit significant speedup over their classical counterparts but require a sufficient number of qubits to be effective. As the number of qubits increases, more information could potentially be processed and manipulated.

Besides, direct connections between all qubits may not always be present within the QPU. A simple way to comprehend and address the interaction between qubits is by treating the problem as a graph. In this representation, qubits are situated at the vertices, while the edges symbolize the existing coupling between adjacent/successive qubits. Consequently, when qubits are not directly coupled, executing two-qubit gate operations becomes unfeasible. In essence, the quality of qubit connectivity directly impacts the capabilities of the QPU: the better the connectivity among qubits, the better the capabilities of the QPU.

Finally, we need also to address the issue of NISQ devices regarding the decoherence time. This way, longer coherence time is essential to the device's performance, as a more significant number of operations can be accomplished before getting an erroneous output beyond a tolerance limit [18]. The period for a qubit's natural decay from the excited state $|1\rangle$ to the ground state $|0\rangle$ is called T1 coherence time (referred to as amplitude damping or energy decay). This temporal parameter characterizes the longitudinal relaxation rate. However there is another important metric to benchmark the quality of qubits, known as T2 coherence time (also referred to as phase dumping), focused on dephasing time. T1 can be estimated using the Rabi experiment [19] while T2 can be computed with the Ramsey experiment [20]. Together, these coherence time measurements serve as vital benchmarks to assess the quality and reliability of qubits in quantum computing systems.

## 2.1 Gate-Level Benchmarks

The manipulation and transformation of qubits are performed by quantum gates. Unlike classical logic gates, which operate on bits, quantum gates leverage the properties of qubits, such as superposition and entanglement, to enable powerful computations. They can flip qubit states, create superposition states, and enable interactions between qubits. By combining sequences of quantum gates, quantum algorithms can be constructed to solve complex problems.

Therefore, various protocols have been developed over the past few years for assessing the performance of these fundamental building blocks of QPUs, quantum gates. Among these protocols to evaluate the performance of quantum gates there are the so-called gate-level protocols, designed to assess the quality of individual quantum gates and their ability to perform specific tasks. By providing meaningful metrics, gate-level protocols enable the comparison of different implementations of quantum technologies while offering insights into the suitability of QPUs for running certain types of quantum circuits. Here, we will discuss some of the most widely used protocols for characterizing the gate-level quality of any QPU.

**Gate Set Tomography**. The initial motivation behind the creation of quantum benchmarks was the need to comprehend the specific processes implemented by the quantum hardware in the presence of noise and imperfect control gates. Thus, quantum process tomography [16] is a widely recognized technique that enables the complete characterization of any quantum process. When this technique is used in a QPU to test its performance by measuring its ability to implement a set of pre-determined quantum gates, the procedure gets the name of Gate Set Tomography (GST) [21–23]. The GST method yields the specific error and noise models of each quantum operation performed by the QPU (including gates, state preparation, and measurement) in order to accurately determine those underlying operations.

To carry out GST, a set of carefully chosen quantum gates is applied to the QPU, and the resulting output states are measured. By analyzing the data obtained from these measurements, GST aims to reconstruct a model of the QPU's underlying gate set and quantify the errors associated with each gate operation. Thus, GST has become an important tool for benchmarking and diagnosing errors, as it provides a comprehensive and detailed characterization of their performance at the gate level.

However, the ability to account for higher-order errors beyond one-qubit and two-qubit errors is challenging in terms of both classical processing and data collection. Its cost scales exponentially with the number of qubits involved, making it a non-trivial task to extend GST to larger QPUs. As a result, while GST has been useful for diagnosing and characterizing the performance of small-scale QPUs, its applicability to larger ones in real-world applications is limited by these scalability issues. In response to this problem, randomized benchmarking (RB) [24] was introduced.

**Randomized Benchmarking**. RB refers to a collection of methods used to reliably estimate the magnitude of average fidelity $f$ or average error rate $\epsilon = 1 - f$ of a given quantum gate set, requiring only polynomial classical resources. Its execution is simple: apply sequences of randomly selected gates with varying lengths. Besides, this approach is robust against state preparation and measurement (SPAM) errors. SPAM errors are often considered together as they can be challenging to distinguish. For instance, when preparing the state $|0\rangle$ and measuring it, if the outcome is $|1\rangle$, it is difficult to determine whether the inconsistency is primarily due to errors in the state preparation or errors in the measurement process.

However, just as GST, RB gives little information about neither the performance of circuits nor their applications, thus being difficult to predict the performance of a concrete algorithm given only the obtained RB metrics of its independent gate sets [25]. This is because specific-designed circuits are more sensitive to error than randomized ones. Their arrangement and order of the gates are carefully chosen to perform a specific computational task or implement a particular algorithm. This constrained structure can make the circuit more susceptible to errors because any deviation or perturbation in the individual gates can propagate and affect the overall computation.

Another RB's major drawback is its limited ability to detect crosstalk errors [26], which can significantly impact QPU's performance at the circuit level. In the context of quantum computing, crosstalk errors refer to undesired effects resulting from the interaction of various components or subsystems within a QPU. When two or more subsystems inadvertently interact, crosstalk emerges. Thus, unwanted interactions can cause deviations from the expected behavior of quantum gates and circuits, which are crucial for performing accurate quantum computations. Although recent research has proposed ways to extend RB to include crosstalk estimation [27, 28], the need for alternative benchmarking techniques that can better address crosstalk errors remains a priority in the field of quantum computing.

RB has received significant attention in the literature, leading to the development of numerous variations and types of RB protocols. As a result, the classification of RB protocols has become complex and difficult to navigate. We recommend Ref. [29] to get a more general and efficient approach to all kinds of RB protocols.

**Cycle Benchmarking**. Finally, it is worth mentioning that Cycle Benchmarking [30] (CB) has been recently introduced to go beyond the limitations of RB and GST. CB allows us to estimate the fidelity of a global noise process affecting a QPU that occurs when a cycle of operations is applied to a quantum register. We define each cycle as a parallel set of gates (in analogy with a digital clock cycle), typically cycles of single-qubit gates and cycles of multi-qubit gates. The CB operates under the assumption that the noise occurring during each cycle of independent single-qubit gates is independent of the specific gates being executed, thus following the Markovian assumption. Besides, CB is designed to be resilient against SPAM errors and can effectively characterize processes in larger quantum registers, as the number of required measurements to accurately estimate the process fidelity remains relatively unaffected by the number of qubits.

## 2.2 Circuit-Level Benchmarks

We have previously discussed the performance at the gate level, reflecting aspects of a few qubit's performance. Now, we are going to focus on some aggregated metrics proposed to directly characterize the performance of the overall quantum computer system.

**Mirroring Benchmarking**. While many benchmarks provide valuable insights into a processor's performance, some fall short in directly measuring the processor's ability to run real-world programs (e.g. RB), particularly on devices large enough to potentially demonstrate quantum advantage. To address this, Mirroring Benchmarking (MB) [25, 31, 32] has been introduced. MB involves the use of "mirror circuits" that execute a set of calculations and then reverse them, providing predictable outcomes compared to complex quantum programs. Researchers created two types of benchmark programs using circuit mirroring: one with random sequences of operations and another with highly structured procedures. The method offers a more accurate and flexible assessment of quantum processor capabilities compared to existing benchmarks, which often lack scalability and fail to capture the relationship between circuit structure and performance.

**Quantum Volume**. IBM's proposal to circuit-level benchmarks was Quantum Volume (QV) [33]. QV is a metric used to evaluate the performance of near-term quantum computers with limited size, considering both the number of qubits and the quality of gate operations and measurements. It is a single-number measurement that also takes into account the error rates of the system and it is affected by uncontrolled interactions that may occur within the system. QV assesses the largest random circuit of equal width and depth that a quantum computer can successfully execute, based on the performance of random circuits with a fixed and generic structure. Its core metric is based on the heavy output generation probability (HOG) [34], so it assures that the correct output is generated with a probability equal to or greater than $2/3$. Thus, it requires an exponentially costly computation of probability amplitudes, and these approaches are not scalable.

Its main drawback is that QV considers only square circuits (meaning circuits with equal width and depth), failing to capture algorithms that do not present this specific shape (e.g. the Shor algorithm [1], with width $n$ but depth $n^3$). Therefore, QV is hardly suitable for giving an accurate performance of a QPU on a real application. However, Volumetric Benchmarking (VB) [35] has recently been presented to generalize QV to non-square circuits. Furthermore, a new metric called Algorithmic Qubits (AQ) [36, 37] has also been introduced to address the limited scalability of QV.

**Cross-Entropy Benchmarking**. In October 2019, Google claimed to have reached quantum supremacy [38], where a series of operations was made in 200 seconds that would take a supercomputer about 10 thousand years to complete. In order to verify that their QPU was working properly they used Cross-Entropy Benchmarking (XEB) [39], which is a quantum benchmarking method that quantifies the similarity between the output distribution of a quantum device and the ideal one. The XEB technique compares the frequency of each bitstring observed experimentally with its corresponding ideal probability computed through simulation on a classical computer. For a given circuit, bitstring measurements are collected, and the linear cross-entropy fidelity $\mathcal{F}_{XEB}$ is computed as the average of the simulated probabilities of the measured bitstrings [38],

$$\mathcal{F}_{XEB} = 2^n \langle P(x_i) \rangle_k - 1 = \frac{2^n}{k} \left( \sum_{i=1}^{k} | \langle 0^n | \, C \, | x_i \rangle |^2 \right) - 1, \tag{1}$$

where $k$ represents the number of samples, $n$ is the number of qubits and $P(x_i)$ is the probability of a bitstring $x_i$ for a quantum circuit $C$.

It can be intuitively understood as a measure of how frequently we sample high-probability bitstrings in a quantum circuit. In the absence of errors, the probability distribution follows an exponential pattern, leading to a $\mathcal{F}_{XEB} = 1$. Conversely, sampling from a uniform distribution results in an average probability $\langle P(x_i) \rangle = 1/2^n$, yielding an $\mathcal{F}_{XEB} = 0$. Thus, $\mathcal{F}_{XEB}$ values between 0 and 1 indicate the likelihood of no errors occurring during the circuit execution. The XEB approach aims to demonstrate that the samples obtained from the quantum device achieve high fidelity values, indicating a strong correlation with the ideal distribution. While obtaining the probabilities $P(x_i)$ requires classical simulation of the quantum circuit, computing $\mathcal{F}_{XEB}$ becomes computationally intractable in the quantum supremacy regime. However, by employing specific circuit simplifications, quantitative fidelity estimates for a fully operational processor executing complex quantum circuits with large dimensions can be derived. Overall, XEB is a promising tool for evaluating the performance of NISQ devices and assessing their capabilities for practical quantum applications, such as quantum chemistry simulations, optimization problems, machine learning tasks, cryptography, and quantum simulation.

Unfortunately, one of the main challenges with the XEB protocol is its reliance on classical resources. This means it scales exponentially with the number of qubits, limiting its scalability. However, a new method has been proposed to address this issue: Clifford XEB [40], as Clifford circuits can be simulated in polynomial time [41], enabling the scaling to much larger systems.

**Circuit Layer Operations per Second**. Up until now, we focused on the quality and scale of the QPUs, but there is one feature we have not taken into account yet: speed.

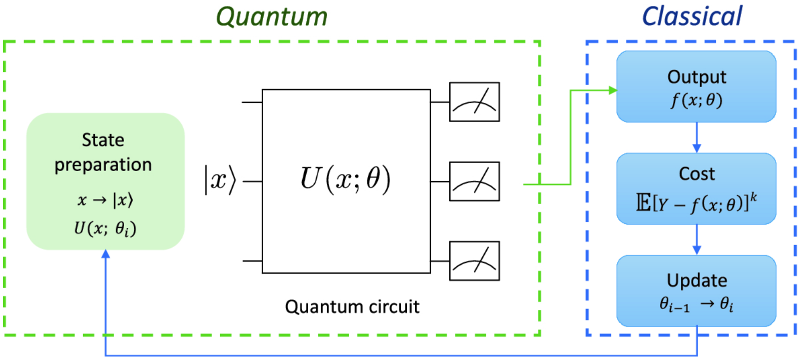

Figure 2: Scheme of the VQA execution [43].

The speed of a quantum computer can be estimated by the Circuit Layer Operations per Second (CLOPS) method [42]. It is a metric that considers the interaction between classical and quantum computing, as real-world applications may include both classical and quantum processing. Some examples of hybrid algorithms are called Variational Quantum Algorithm (VQA) and they allow to combine a short-depth quantum circuit with classical optimization to evaluate a cost function that depends on the parameters of a quantum gate sequence. Figure 2 show a schematic representation of how VQAs work. By minimizing this cost function using well-established classical optimizers, VQA strikes a balance between quantum circuit depth and classical processing, making them efficient for solving complex problems. Thus, they provide a promising approach to exploiting the capabilities of NISQ devices by utilizing both quantum and classical resources in a synergistic manner.

CLOPS protocol is formally defined as the number of QV layers executed per second, and it includes hundreds of parameterized predefined circuits that may be used, similar to the ones of quantum volume, except that each SU(4) random unitary is left fully parameterized. The main drawback of CLOPS is its primary focus on the quantum aspect while considering classical computations as merely auxiliary. As a result, any enhancements in the performance of the classical component would have minimal impact on improving the CLOPS metric.

## 2.3 Application-Level Benchmarks

It is challenging to ascertain whether one quantum computer can outperform another solely based on its physical characteristics. Consequently, some performance metrics have emerged, focusing on the evaluation of quantum computers' real-world applications. Using application-level metrics enables straightforward cross-platform comparisons between different quantum architectures and classical approaches.

**Q-score and QPack**. The first application-level benchmark we are addressing is Q-score [44]. Q-score [44] is a single-numeric metric (just like QV) that measures the performance of a quantum computer in solving combinatorial optimization problems. It allows us to measure the maximum number of quantum bits that a quantum computer can effectively use to solve combinatorial optimization problems, specifically MaxCut [45]. Besides, Q-score is perfectly scalable, so it is not restricted to NISQ devices. The Quantum Approximate Optimization Algorithm (QAOA) [46] is commonly used for Q-score, but any other suitable quantum algorithm tackling the combinatorial optimization problem can be considered as well. An open-source implementation of Q-score regarding the MaxCut problem with QAOA is available for anyone who wants to test it on their QPU [47].

Despite its widespread use, the Q-score benchmark is limited in its ability to evaluate the overall performance of a QPU. Its use of a single problem, MaxCut, raises concerns about its applicability to other problems, as a QPU could be optimized for specific tasks but perform poorly on others. To address this limitation, QPack [48, 49] was recently introduced as a universal benchmark that includes several combinatorial optimization problems, such as MaxCut, the dominating set problem (DSP) [50] and the traveling salesman problem (TSP) [51].

QPack is designed as a comprehensive evaluation protocol for NISQ computers using the QAOA algorithm. It provides a universal benchmark for quantum computers that focuses on the maximum problem size a device can handle, its required runtime, and the accuracy achieved. This protocol comprises three main parts: the problem library, which contains a set of problems to be evaluated; the QAOA, which is used to find approximate solutions to the problems; and the performance algorithm, which evaluates runtime, accuracy, and scaling of the QPU's performance. QPack's current implementation on Qskit is available on the GitHub repository [52], while another implementation using the XACC library can be found at [53].

**Quantum many-body problems and chemistry**. One of the most promising applications of QPUs is their capability to solve quantum many-body problems. Thus, the Fermionic Depth (FD) [54] benchmark has been presented as a performance metric to assess this issue. FD employs the one-dimensional Fermi-Hubbard model [55], a well-established model in solid-state physics useful in simulating strongly correlated fermionic systems. This model is particularly helpful as an application benchmark for variational quantum simulations as it can be solved exactly using the Bethe ansatz [56] on a classical computer for both finite and infinite chains.

The protocol involves calculating the approximate ground state energy of the 1D Hubbard model using the variational quantum eigensolver (VQE) [57, 58] for different system sizes and comparing the results with the exact energy at the infinite size. Due to the effects of decoherence, the deviation curve will display a minimum at a particular length, referred to as the *fermionic length* of the QPU. This fermionic length is an indication of the maximum size of a fermionic problem that the QPU can handle. Although the fermionic length is specific to the 1D Fermi-Hubbard model, it provides valuable insight into a QPU's ability to tackle quantum many-body problems, which is a crucial factor in evaluating QPU's performance.

A quantum chemistry simulation benchmark has also been proposed in the literature [59], with a specific focus on small molecules, such as alkali metal hydrides. This benchmark allows for the evaluation of a QPU by defining a series of electronic structure calculation examples that can be executed on the hardware to determine the molecule's ground state. This provides a performance metric by comparing the results with the theoretically exact solution. An open-source implementation is available on [60].

## 3 Entanglement Benchmarking

An intriguing aspect of benchmarking in quantum computing involves exploring the level of entanglement that can be attained during the execution of quantum circuits. We categorize entanglement benchmarks as part of the circuit-level benchmarking approach, as they enable to assess entanglement performance of the entire QPU without focusing on a unique quantum application.

Measurement and certification of entanglement have long been subjects of study due to their fundamental significance in quantum systems. As a result, numerous approaches have been proposed to experimentally detect, certify, and quantify entanglement. For comprehensive insights into prominent entanglement detectors and quantifiers, we highly recommend consulting Ref. [61], which provides valuable perspectives on the state-of-the-art methods in the field of entanglement assessment. Moreover, a recent noteworthy contribution in this field is presented in Ref. [62], where a volumetric benchmark is introduced based on the generalization and verification of entanglement across multiple qubits using graph states.

The importance and relevance of entanglement in quantum computation has motivated our thesis to focus on advancing this particular area of study, presenting a couple of recent and hardly-known protocols that address the quantification of entanglement on the quantum device after a circuit execution (Sections 4 and 5). We would also like to comment that other entanglement measures where also considered, like the ones presented in Ref. [63, 64], where a variational quantum singular value decomposition (QSVD) algorithm is studied with the final objective of computing the Von Neumann entropy. Unfortunately, due to a lack of time, it was not possible to include the QSVD approach into this thesis.

Let's focus now on how entanglement could potentially be estimated. Given a density matrix $\rho_{AB}$ that represents a bipartite quantum state, there are several ways to quantify entanglement between two subsystems. The best-known is the so-called Von Neumann entropy, which is defined over one of the subsystems as $\mathcal{S}(\rho_A) = -\operatorname{Tr}\{\rho_A \log \rho_A\} = -\operatorname{Tr}\{\rho_B \log \rho_B\} = \mathcal{S}(\rho_B)$, with $\rho_A = \operatorname{Tr}_B\{\rho_{AB}\}$ being the reduced density matrix of subsystem $A$. However, during this work we will be dealing with the so-called Rényi entropies $S_n$, which are also defined in terms of the reduced density matrices and the order $n$. Thus, the $n$-order Rényi entropy $S_n$ is given by the expression

$$S_n = \frac{1}{1-n} \log R_n, \qquad R_n = \operatorname{Tr}\{(\rho_A^n)\}, \tag{2}$$

where we make this explicit distinction of $R_n$ because the algorithms we work with do not estimate directly the entropy but this concrete magnitude $R_n$. In the limit when $n \to 1$, we recover the Von Neumann entropy.

It is important to distinguish between entanglement estimation and entanglement certification. Thus, the protocols we will be discussing work only as entanglement estimators. However, they can also certificate entanglement under the assumption that the initial quantum state is pure, but not if it is mixed. Therefore, we can always run these methods to quantify the entropy, but that does not mean that actual entanglement is present in the system if the initial state is mixed. In these mixed-state cases, once entanglement on the system is proven, we can accept the entropy value given by the algorithm. A major discussion regarding entanglement certification and quantification is performed in Appendix C, where we also present a method to assess entanglement certification through the randomized protocol we introduce in Section 5.

## 4 n-order Rényi entropy estimation by the Swap Test

In this section, we will present a straightforward and effective protocol for estimating the $n$-order Rényi entropy. This method builds upon previous works by Johri et al. [65] and Linke et al. [66]. Our main objective is to provide a general framework for computing the Rényi entropy of any order $n$ using *Qibo*.

This protocol allows us to estimate the entanglement spectrum via the Swap Test [67], a well-known algorithm used to determine the overlap of two quantum states. It provides a way to compare both states without directly measuring them, involving an extra auxiliary qubit. Figure 3 shows a better perspective on how the algorithm is implemented.

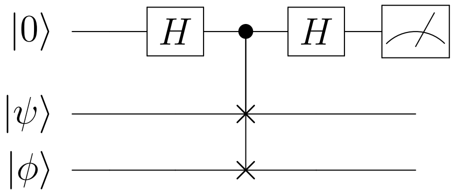

Figure 3: General circuit for the Swap Test, estimating the overlap $|\langle\psi|\phi\rangle|^2$.

One of its main challenges is the requirement of a large number of qubits in a quantum computer. To compute the $n$-order Rényi entropy, we need $n$ identical copies of the initial quantum state $|\psi\rangle$. For example, estimating the 2-order Rényi entropy necessitates two copies, while estimating the 5-order Rényi entropy requires five identical copies. Specifically, the number of qubits needed is given by $\#_{qubits} = n \times m + 1$, where $n$ is the order of the Rényi entropy, $m$ is the number of qubits in the initial pure state $|\psi\rangle$, and the additional "$+1$" accounts for the auxiliary qubit. Consequently, the algorithm becomes impractical when the available quantum computer does not have a sufficient number of qubits.

Furthermore, implementing this method on NISQ devices presents additional challenges, as NISQ computers suffer from noise and decoherence issues, leading to errors in state preparation, measurement, and gate operations. These errors can introduce significant inaccuracies in the estimation of the Rényi entropy, making it challenging to obtain reliable results. That is why other alternatives have been studied, like the one we present and address in Section 5. However, it is still a functional protocol that will efficiently work on future fault-tolerant quantum computers.

Before we delve into the intricate steps of the methodology, it is important to clarify a fundamental limitation of this protocol: it can only be applied to pure states (similar to the Swap Test). This restriction arises from the specific nature of the protocol and the underlying mathematical framework used for its derivation, where we are sinking in Section 4.2. The protocol relies on specific properties and transformations of pure quantum states, which may not hold or produce meaningful results when applied to mixed states. It is crucial to highlight this limitation to avoid any misconceptions and misinterpretations when applying the protocol. If the initial state under investigation is known or suspected to be mixed, alternative approaches or techniques must be employed.

### 4.1 Methodology

To begin with, our first step on the protocol is to create the state $|\psi\rangle$ we want to focus on, which must be a bipartite pure state, living in a Hilbert Space composed of two subspaces $A$ and $B$, $\mathcal{H}_{AB}$. As we have already mentioned, the implementation of this code relies on several copies of the initial state $|\psi\rangle$ and an extra qubit. Specifically, to estimate the $n$-order Rényi entropy we need up to $n$ identical copies of $|\psi\rangle$ plus the auxiliary qubit. Thus, as analogous to the Swap Test, the Hadamard Test is performed by implementing several Control-Swap gates (also known as Fredkin gates) only to the qubits of subsystem $A$. The amount of Control-Swap gate will depend on the number of qubits that constitute the state and the number of $n$ copies, as $\#_{\text{C-Swap}} = (n-1) \cdot \#_{\text{qubits on A}}$. This is because we must apply the C-Swap gates to two consecutive copies, acting first on copies 1 and 2, then on 2 and 3, and so on. Thus, the general expression for $R_n$ is given by

$$R_n = \langle\psi|^{\otimes n} \, \mathrm{Perm}_A \, |\psi\rangle^{\otimes n} \,, \tag{3}$$

where $\mathrm{Perm}_A = \mathrm{Swap}_A^{n\leftrightarrow(n-1)} \, \mathrm{Swap}_A^{(n-1)\leftrightarrow(n-2)} \ldots \mathrm{Swap}_A^{3\leftrightarrow 2} \, \mathrm{Swap}_A^{2\leftrightarrow 1}$, and the $\mathrm{Swap}_A$ operator acts as the common Swap operator but only on subspace A of the system, as follows

$$\mathrm{Swap}_A \left|\psi\right\rangle \left|\psi\right\rangle = \mathrm{Swap}_A \sum_{ij} c_{ij} \left|a_i\right\rangle \left|b_j\right\rangle \sum_{i'j'} c_{i'j'} \left|a_{i'}\right\rangle \left|b_{j'}\right\rangle = \sum_{ij}\sum_{i'j'} c_{ij} c_{i'j'} \left|a_{i'}\right\rangle \left|b_j\right\rangle \left|a_i\right\rangle \left|b_{j'}\right\rangle . \tag{4}$$

where $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$ form a set of orthonormal bases of subspaces $A$ and $B$ respectively. Moreover, we would like to highlight that the special case of $R_2$ describes the purity of $\rho_A$, and is given just by

$$R_2 = \langle\psi| \langle\psi| \, \mathrm{Swap}_A \, |\psi\rangle |\psi\rangle = \mathrm{Tr}\{\mathrm{Swap}_A \, \rho_{AB} \otimes \rho_{AB}\}. \tag{5}$$

Figure 4 shows an example for the estimation of $R_2$, which is nothing more than the purity of subsystem $A$, performed in a quantum computer of 5 qubits.

Finally, $R_n$ is computed by the statistics obtained when measuring the auxiliary qubit multiple times as

$$R_n = P_{aux}(0) - P_{aux}(1). \tag{6}$$

## 4.2 Proofs and derivations

Let's prove Eq.(3). The protocol forces us to make $n$ copies of the initial state, of the form

$$|\psi\rangle^{\otimes n} = \sum_{\text{all indices}} c_{i_1 j_1} \dots c_{i_n j_n} \left|a_{i_1}\right\rangle \left|b_{j_1}\right\rangle \dots \left|a_{i_n}\right\rangle \left|b_{j_n}\right\rangle ,$$

where we are using the enumeration with numbers $i_l$ to simplify later on the calculations. First, let's get a general expression for the left-hand side of the equation.

$$\begin{aligned} R_n = \mathrm{Tr}\{\rho_A^n\} &= \mathrm{Tr}\left\{\left(\sum_{\text{all indices}} c_{i_1 j_1} c^*_{i'_1 j_1} \left|a_{i_1}\middle\rangle\middle\langle a_{i'_1}\right|\right)^n\right\} \\ &= \sum_{\text{all indices}} c_{i_1 j_1} c^*_{i_2 j_1} c_{i_2 j_2} c^*_{i_3 j_2} \dots c_{i_{n-1} j_{n-1}} c^*_{i_n j_{n-1}} c_{i_n j_n} c^*_{i_1 j_n} \\ &= \sum_{\text{all indices}} \gamma_{i_1 i_2} \gamma_{i_2 i_3} \dots \gamma_{i_{n-1} i_n} \gamma_{i_n i_1}, \end{aligned}$$

where $\gamma_{i_1 i_2} = \sum_{j_1} c_{i_1 j_1} c^*_{i_2 j_1}$. Now, let's compute the right-hand side of the equation,

$$\begin{aligned} &\langle\psi|^{\otimes n} \, \mathrm{Perm}_A \, |\psi\rangle^{\otimes n} = \langle\psi|^{\otimes n} \, \mathrm{Perm}_A \left(\sum_{\text{all indices}} c_{i_1 j_1} \dots c_{i_n j_n} \left|a_{i_1}\right\rangle \left|b_{j_1}\right\rangle \dots \left|a_{i_n}\right\rangle \left|b_{j_n}\right\rangle\right) \\ &= \left(\sum_{\text{all}} c^*_{i'_1 j'_1} \dots c^*_{i'_n j'_n} \left\langle a_{i'_1}\right| \left\langle b_{j'_1}\right| \dots \left\langle a_{i'_n}\right| \left\langle b_{j'_n}\right|\right) \cdot \left(\sum_{\text{all}} c_{i_1 j_1} \dots c_{i_n j_n} \left|a_{i_2}\right\rangle \left|b_{j_1}\right\rangle \dots \left|a_{i_1}\right\rangle \left|b_{j_n}\right\rangle\right) \\ &= \sum_{\text{all indices}} c_{i_1 j_1} c^*_{i'_1 j'_1} \dots c_{i_n j_n} c^*_{i'_n j'_n} \delta_{i'_1 i_2} \delta_{j'_1 j_1} \delta_{i'_2 i_3} \delta_{j'_2 j_2} \dots \delta_{i'_n i_1} \delta_{j'_n j_n} \\ &= \sum_{\text{all indices}} c_{i_1 j_1} c^*_{i_2 j_1} c_{i_2 j_2} c^*_{i_3 j_2} \dots c_{i_n j_n} c^*_{i_1 j_n} \\ &= \sum_{\text{all indices}} \gamma_{i_1 i_2} \gamma_{i_2 i_3} \dots \gamma_{i_{n-1} i_n} \gamma_{i_n i_1} . \end{aligned}$$

As we can see, both sides of the equation are the same, proving the general statement Eq.(3).

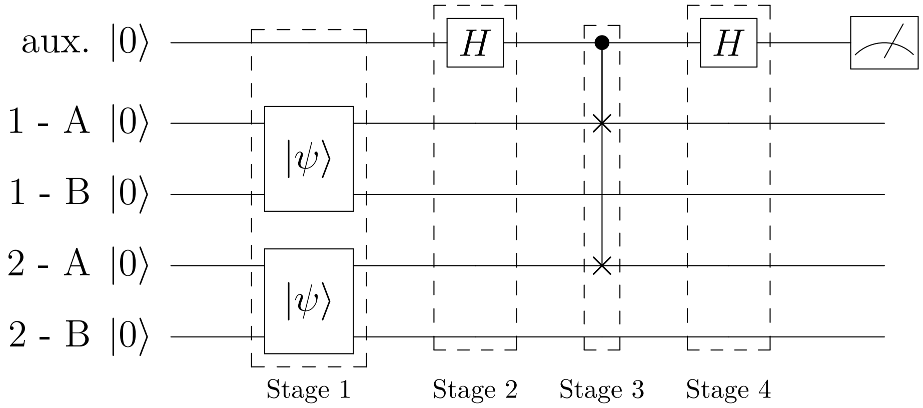

Figure 4: General circuit for the 2-Rényi entropy estimation using the Swap method. Numbers 1 and 2 represent the two different systems, while A and B represent their own subspaces. *aux.* indicates auxiliary qubit.

Now, let's focus on the derivation of Eq.(6) to show the usefulness of an auxiliary qubit. Let's have a look at the situation of all our systems at any stage of the circuit. Figure 4 is oriented in a 5-qubit system, but it can help us get a better notion of the overall process for the general case.

- **Stage 1**. Trivially, the state of the system at the beginning is
$$|\text{Stage 1}\rangle = |0\rangle\, |\psi\rangle^{\otimes n}\,.$$

- **Stage 2**. After applying the Hadamard gate on the auxiliary qubit we get
$$|\text{Stage 2}\rangle = \frac{|0\rangle\, |\psi\rangle^{\otimes n} + |1\rangle\, |\psi\rangle^{\otimes n}}{\sqrt{2}}.$$

- **Stage 3**. Control-Swap$_A$ gates are applied consecutively, affecting only subspace A, which is nothing more than the previous Perm$_A$ operator.
$$|\text{Stage 3}\rangle = \frac{|0\rangle\, |\psi\rangle^{\otimes n} + |1\rangle\, \text{Perm}_A\, |\psi\rangle^{\otimes n}}{\sqrt{2}}$$

- **Stage 4**. The final state of the system is then
$$\begin{aligned}|\text{Stage 3}\rangle &= \frac{|0\rangle\, |\psi\rangle^{\otimes n} + |1\rangle\, |\psi\rangle^{\otimes n} + |0\rangle\, \text{Perm}_A\, |\psi\rangle^{\otimes n} - |1\rangle\, \text{Perm}_A\, |\psi\rangle^{\otimes n}}{2} \\ &= \frac{1}{2}\, |0\rangle \left(|\psi\rangle^{\otimes n} + \text{Perm}_A\, |\psi\rangle^{\otimes n}\right) + \frac{1}{2}\, |1\rangle \left(|\psi\rangle^{\otimes n} - \text{Perm}_A\, |\psi\rangle^{\otimes n}\right).\end{aligned}$$

Assuming that $R_n$ must be a real value, as we have previously proven Eq.(3), then $\langle\psi|^{\otimes n}\, \text{Perm}_A\, |\psi\rangle^{\otimes n} = \langle\psi|^{\otimes n}\, \text{Perm}_A^{\dagger}\, |\psi\rangle^{\otimes n}$, and the probabilities of finding the auxiliary qubit at state 0 or 1 are

$$\begin{aligned}P_{aux}(0) &= \frac{1 + \langle\psi|^{\otimes n}\, \text{Perm}_A\, |\psi\rangle^{\otimes n}}{2} = \frac{1 + R_n}{2}, \\ P_{aux}(1) &= \frac{1 - \langle\psi|^{\otimes n}\, \text{Perm}_A\, |\psi\rangle^{\otimes n}}{2} = \frac{1 - R_n}{2}.\end{aligned}$$

Adding both previous relations, we get the final expression for the $n$-Rényi entropy as given by Eq.(6).

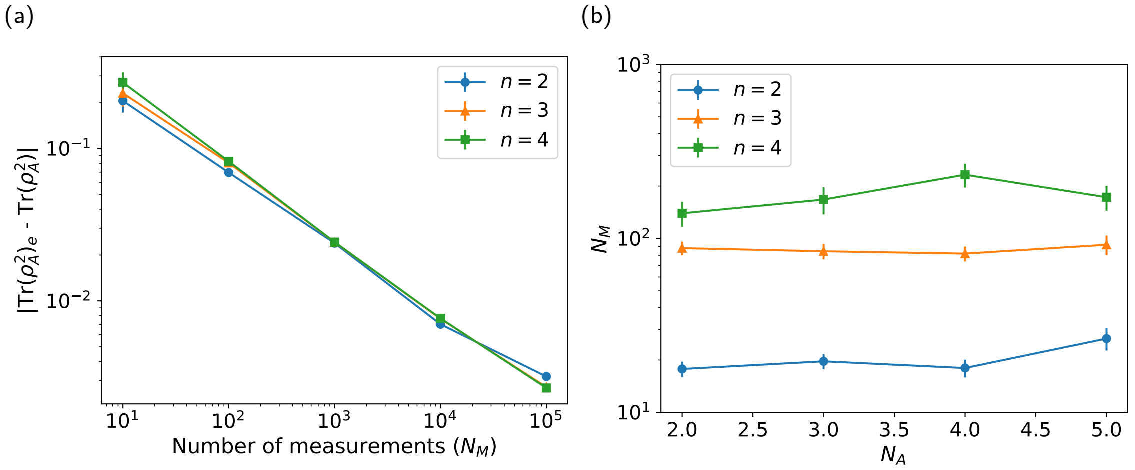

Figure 5: Error scaling of the Swap Test protocol to estimate the Rényi entropy of the GHZ state with 6 qubits. Plot (a) shows the difference of $\mathrm{Tr}\{\rho_A^n\}$ between the protocol and the expected value as a function of the number of measurements $N_M$ and the Rényi order $n$ and subsystem size $N_A = 5$; while plot (b) shows the scaling of the required number of measurements $N_M$ to determine $\mathrm{Tr}\{\rho_A^n\}$ up to an average relative error of 0.12 as a function of the subsystem size $N_A$ and the Rényi order $n$.

### 4.3 Results and error scaling

Here, we will take a look at how our implementation of the protocol on *Qibo* works for the maximally entangled Greenberger–Horne–Zeilinger state (GHZ), $|\psi\rangle_{GHZ} = \frac{|0\rangle^{\otimes N} + |1\rangle^{\otimes N}}{\sqrt{2}}$, with $N$ the number of qubits.

Figure 5a shows the difference on $\mathrm{Tr}\{\rho_A^n\}$ between the protocol and the expected value as a function of the number of measurements $N_M$ and the order $n$, for the maximally entangled GHZ state, with $N = 6$ qubits and subsystem $N_A = 5$. As the $N_M$ increases, the error decreases, because we get a better approximation to the real values of the probabilities $P_{aux}(0)$ and $P_{aux}(1)$. We can see how the influence of the Rényi order $n$ is minimum, as they scale similarly. In particular, it scales as $\frac{1}{\sqrt{N_M}}$.

Figure 5b shows how $N_M$ scales with the subsystem size $N_A$, where the error as been truncated to 0.12. As expected $N_M$ does not scale with $N_A$, as only measurements on the auxiliary qubits are made. However, it does scale with the order of the Rényi entropy. That is because the creation of multiple copies results in an expansion of the Hilbert space of the entire system, leading to an increased number of basis states and a greater variety of possible quantum states that can be observed upon the collapse. Consequently, a larger number of measurements is required to obtain a reliable approximation of the probabilities of the auxiliary qubit when the number of possible outcomes increases.

## 5 2-order Rényi entropy estimation by randomized measurements

In this section, we introduce a protocol for estimating the second-order Renyi entropy through statistical correlations between randomized measurements, as measurement outcomes performed on a random basis contain information about a system's purity $\mathrm{Tr}\{\rho_A^2\}$ of a reduced density matrix $\rho_A$ and, according to expression (2), the second-order Renyi entropy of the whole bipartite state $\rho_{AB}$. Our algorithm builds upon the proposals put forth in Ref. [68–70] and has been implemented on *Qibo*.

On the one hand, this method differs from the previous one (Section 4), as it requires only a single copy of the quantum system to compute its entropy. This represents a great advantage for the user, as all qubits in the quantum computer can be used for representing the quantum state of study. Besides, this protocol allows us to work with random mixed states, a consequence of the protocol's specific nature and its mathematical framework. This feature represents a major advantage with respect to the previous Swap method. On the other hand, similar to the previous approach, this protocol makes no *a priori* assumption regarding the final structure of the quantum state and it lowers the number of measurements needed compared to other protocols, providing an advantage over computationally expensive tomographic methods [71]. Furthermore, the inherent random nature of the algorithm allows us to mitigate the errors caused by NISQ devices, which translates into better data and statistics.

We must recall that this protocol only works as an entanglement witness if the initial bipartite state is pure. Thus, getting a second-order Rényi entropy value greater than zero does not guarantee the state is entangled without making the initial assumption that the state is completely pure. In Appendix C, we are delving deeper into the intricacies of this topic and mixed states, and we are discussing how it would be possible to implement an effective method to overcome this situation.

## 5.1 Methodology

Before we begin, it is worth mentioning that this procedure is not only feasible with qubits but it can be generalized into qudits. So, we are going to address the protocol in the more general path, only to focus later on the more specific part of qubits. The experimental procedure to estimate the purity of a reduced density matrix $\rho_A$ of a subsystem $A$ composed by $N_A$ qudits (each one inhabiting a Hilbert space $\mathcal{H}$ of dimension $d$) consists of several steps. First, a random unitary $U_A$ is applied to $\rho_A$. This can either be (i) a global random unitary sampled form the Circular Unitary Ensemble (CUE) defined on the entire Hilbert space $\mathcal{H}_A = \mathcal{H}^{\otimes N_A}$ of dimension $\mathcal{D}_A = d^{N_A}$, or (ii) local random unitaries of the form $U_A = \otimes_{i \in A} U_i$ where each $U_i$ is again sampled independently from the CUE defined on the local Hilbert space $\mathcal{H}$ of dimension $d$. We are discussing the CUE and the Haar measure in Appendix B, while the global and local approaches are discussed ahead in this section. Secondly, several measurements are performed in the computational basis, only over the subsystem $A$, with the same set of random unitaries $U_A$ in order to extract statistics and compute the occupation probabilities $P_U(s_A) = \mathrm{Tr}\left\{ U\rho U^\dagger \left|s_A\right\rangle\!\left\langle s_A\right| \right\}$, where $|s_A\rangle$ represents the possible outcome of the measurements over subsystem $A$, of the form $|s_A\rangle = |s_1, \ldots, s_{N_A}\rangle$ with $s_i = 1, \ldots, d$ for $i \in A$ (i.e. 0 or 1 for qubits). Finally, this procedure is repeated using different random unitaries, and the average probability over the ensemble of those random unitaries is estimated.

Thanks to the second-order cross-correlations across the random unitary ensemble given by the set of outcome probabilities $P_U(s_A)$, the purity of $\rho_A$ can be estimated following two expressions. For the global unitary case, we should follow

$$\mathrm{Tr}\left\{ \rho_A^2 \right\} = (\mathcal{D}_{\mathcal{A}} + 1) \sum_s \overline{P_U(s_A)^2} - 1. \tag{7}$$

However, for the local unitary case, the expression is

$$\mathrm{Tr}\left\{ \rho_A^2 \right\} = d^{N_A} \sum_{s_A, s'_A} (-d)^{D[s_A, s'_A]} \overline{P_U(s_A) P_U(s'_A)}, \tag{8}$$

where $D[s_A, s'_A]$ represents the Hamming distance, which given two states $|s_A\rangle = = |s_1, \ldots, s_{N_A}\rangle$ and $|s'_A\rangle = |s'_1, \ldots, s'_{N_A}\rangle$, it returns the number of constituents $i \in A$ where $s_i \neq s'_i$.

As mentioned before, the approach to this method regarding global and local unitaries differs in some aspects. In the realm of real quantum computers, the implementation of global random unitaries poses certain challenges due to the need for effective interactions between qubits. Therefore, the dynamic nature of these interactions often makes their implementation complex and subject to variation. As a result, local random unitaries emerge as a more viable solution. By focusing on single-qubit operations, local random unitaries offer several advantages. They can be implemented with higher fidelity, ensuring greater accuracy in the manipulation of individual qubits, and improving performance and reliability of the quantum system. Besides, local operations allow for increased repetition rates, an aspect particularly valuable for the protocol, as we must keep the same set of random unitaries to extract statistics. Therefore, the utilization of local random unitaries provides a practical and efficient approach to quantum computation, leveraging the strengths of single-qubit operations to overcome the challenges associated with implementing global random unitaries. Furthermore, if we already got the data regarding subsystem $A$, the local approach allows us to estimate the entropy of any other subsystem $A' \subseteq A$ without the need to run again the protocol. Finally, as we will see in future Section 5.3, both strategies manifest different error scaling and sensibility.

## 5.2 Proofs and derivations

Before beginning with the derivation, previous knowledge about the Haar measure and unitary $t$-designs is required. We recommend the reader take a closer look at Appendix B to get a better understanding of the topic. Our main objective here is to find an expression for the purity of a density matrix $\rho$ as a function of the ensemble average of second-order cross-correlation of the outcome probabilities, in order to prove Eqs. (7) and (8). A first approximation can be just a linear combination of the form

$$\mathrm{Tr}\left\{\rho^2\right\} = f\left(P_U(s)P_U(s')\right) = \sum_{s,s'} O_{s,s'} \overline{P_U(s)P_U(s')}, \tag{9}$$

where $O_{s,s'}$ are arbitrary coefficients depending on the outcome $s$. Remembering that $P_U(s) = \mathrm{Tr}\left\{U\rho U^\dagger |s\rangle\langle s|\right\}$, it is straightforward to check that $P_U(s)P_U(s') = = \mathrm{Tr}\left\{U^{\otimes 2}(\rho\otimes\rho)(U^\dagger)^{\otimes 2} |s\rangle\langle s| \otimes |s'\rangle\langle s'|\right\}$. We can then replace this expression with the previous one as

$$\begin{aligned}\sum_{s,s'} O_{s,s'} \overline{P_U(s)P_U(s')} &= \sum_{s,s'} O_{s,s'} \mathrm{Tr}\left\{\overline{U^{\otimes 2}(\rho\otimes\rho)(U^\dagger)^{\otimes 2}} |s\rangle\langle s| \otimes |s'\rangle\langle s'|\right\} \\ &= \mathrm{Tr}\left\{\sum_{s,s'} O_{s,s'} |s\rangle\langle s| \otimes |s'\rangle\langle s'| \, \overline{U^{\otimes 2}(\rho\otimes\rho)(U^\dagger)^{\otimes 2}}\right\},\end{aligned} \tag{10}$$

where we have used the cyclic properties of the trace. Now, it is possible to define an operator $O = \sum_{s,s'} O_{s,s'} |s\rangle\langle s| \otimes |s'\rangle\langle s'|$ and use the $k$-twirl channel, $\Phi^{(k)}_{Haar}(\cdot)$, to get

$$\mathrm{Tr}\left\{\rho^2\right\} = \mathrm{Tr}\left\{O \, \Phi^{(2)}_{Haar}(\rho\otimes\rho)\right\}. \tag{11}$$

However, implementing the $k$-twirl channel using the Haar measure becomes unfeasible (look Appendix B). That is why we can replace the $k$-twirl channel, $\Phi^{(k)}_{Haar}(\cdot)$, with the $k$-twirl channel over the well-limited ensemble $\mathcal{E}(N)$ (which depend on the number of qudits $N$ of the subsystem), $\Phi^{(k)}_{N}(\cdot) = \Phi^{(k)}_{Haar}(\cdot)$. Once again, using the cyclic properties of the trace and the previous definitions, it is possible to rewrite expression (11) as

$$\mathrm{Tr}\Big\{\rho^2\Big\} = \mathrm{Tr}\Big\{O\Phi^{(2)}_{N}(\rho\otimes\rho)\Big\} = \mathrm{Tr}\Big\{\Phi^{(2)}_{N}(O)\rho\otimes\rho\Big\}, \tag{12}$$

where the final expression is analogous to Eq.(5). This way, we can intuitively understand the ensemble average over the second cross-correlations as an effective construction of the swap operator on two virtual copies of the state $\rho$, making a connection with the Swap protocol. Thus, to get the final expression for the purity, we must find the operator $O$ such that

$$\Phi^{(2)}_{N}(O) = \mathrm{Swap}. \tag{13}$$

Before we continue, we must mention that the $k$-twirl channel, $\Phi^{(k)}_{Haar}(\cdot)$, may be spanned by the permutation operators $W_\pi$, for permutations $\pi = (\pi(1), \ldots, \pi(k)) \in \mathcal{S}_k$ in the symmetric group $\mathcal{S}_k$, as permutation operators are invariant under the projection $\Phi^{(k)}_{Haar}$. Proven by the Schur Weyl duality [72], we find

$$\Phi^{(k)}_{Haar}(O) = \sum_{\pi,\sigma\in\mathcal{S}_k} C_{\pi,\sigma}\,\mathrm{Tr}\{W_\sigma O\}W_\pi, \tag{14}$$

where the coefficients $C_{\pi,\sigma}$ are found in the so-called Weingarten matrix. It is worth mentioning that this matrix is invertible for $k \leq d$, with values $(C^{-1})_{\pi,\sigma} = d^{\#\mathrm{cycles}(\pi\sigma)}$, where #cycles$(\cdot)$ gives the number of cyclic permutations of a given permutation function.

To continue with, we are using an homogeneous ansatz $O = \otimes_{i=1}^{N} o$, where $o$ is a local operator on each qudit, $o = \sum_{s,s'=1}^{d} o_{s,s'} \left|s\middle\rangle\!\middle\langle s\right| \otimes \left|s'\middle\rangle\!\middle\langle s'\right|$. It can be checked that $\Phi^{(k)}_{N}(\otimes_{i=1}^{N} o) = \otimes_{i=1}^{N}\Phi^{(k)}_{1}(o) = \left(\Phi^{(k)}_{1}(o)\right)^{\otimes N}$. Thus, it is sufficient to find $o$ such that

$$\Phi^{(2)}_{1}(o) = W_{(2,1)} = \mathrm{Swap}. \tag{15}$$

Using Eq.(14), we need to find

$$\mathrm{Tr}\{W_\sigma o\} = (C^{-1})_{(2,1),\sigma} = d^{\#\mathrm{cycles}((2,1)\cdot\sigma)} \qquad \forall\sigma\in\mathcal{S}_2 \tag{16}$$

Therefore, replacing the local operator $o$, we find two equations that must be satisfied,

$$\mathrm{Tr}\Big\{W_{(1,2)}o\Big\} = \sum_{s,s'=1}^{d} o_{s,s'} = d, \tag{17}$$

$$\mathrm{Tr}\Big\{W_{(2,1)}o\Big\} = \sum_{s=1}^{d} o_{s,s} = d^2, \tag{18}$$

and the values that achieve this are

$$o_{s,s'} = (d+1)\delta_{s,s'} - 1 = d(-d)^{-D_G[s,s']} \tag{19}$$

with $D_G[s,s']$ the Hamming distance of the whole state $s$ and $s'$ (i.e. $D_G[s,s] = 0$ and $D_G[s,s'] = 1$ if $s \neq s'$). Finally, the operator $O$ remains as

$$O = o^{\otimes N} = d^N \sum_{s,s'} (-d)^{-D[s,s']} \left|s\middle\rangle\middle\langle s\right| \otimes \left|s'\middle\rangle\middle\langle s'\right| , \tag{20}$$

with $D[s,s'] = \sum_{i=1}^{N} D_G[s_i, s'_i]$. Replacing $O$ in Eq.(11) gives us the expression for the local random unitary case (8). Eq.(7) for the global approach can easily be proven departing from the local expression (8), by setting $N_A = 1$ and $d \to \mathcal{D}_A = d^{N_A}$, and replacing $\sum_{s_A, s'_A \neq s_A} P_U(s'_A) = 1 - P_U(s_A)$. Thus, they can be understood as a single qudit with dimension $\mathcal{D}_A$.

## 5.3 Results and error scaling

Let's take a look at how the error of the protocol scales with the number of measurements performed and the system's size. This will give us a better perspective of the protocol's performance. We are going to test it using two different states: a maximally entangled state GHZ, and a random mixed state (RMS). Those RMS are performed by creating 10 different random density matrices of the form $\rho^i_{AB} = \left|\psi_i\middle\rangle\middle\langle\psi_i\right|$ and adding them as $\rho_{RMS} = \sum_{i=1}^{10} p_i \rho^i_{AB}$, where $p_i$ are random probabilities that add to one, $\sum_{i=1}^{10} p_i = 1$.

For the case of global unitaries, Ref. [73] shows how the scaling of statistical error of the estimated purity would perform, following

$$|\operatorname{Tr}\left\{\rho_A^2\right\}_e - \operatorname{Tr}\left\{\rho_A^2\right\}| \sim \frac{1}{\sqrt{N_U \mathcal{D}_A}} \left( c_1 + c_2 \frac{\mathcal{D}_A}{N_M} \right) \tag{21}$$

where $c_1$ and $c_2$ are constants of order $\mathcal{O}(1)$. This expected tendency is represented in Figure 6a and 6c, showing a good agreement with the GHZ state. In the case of local random unitaries, Ref. [68] suggests an empirical scaling law of the form

$$|\operatorname{Tr}\left\{\rho_A^2\right\}_e - \operatorname{Tr}\left\{\rho_A^2\right\}| \sim \frac{1}{\sqrt{N_U}} \left( c_3 + \frac{2^{0.75 N_A}}{N_M} \right), \tag{22}$$

where $c_3$ is another constant of order $\mathcal{O}(N_A)$. Figure 6b and 6d shows how the obtained data follows the expected error.

Figure 6 shows the difference between the computed and expected purity for the GHZ state (a pure state) and RMS, with an average statistical error extracted from 100 numerical experiments. The scaling with $N_U$ an $N_M$ follow the expected behavior of Eq.(21) and Eq.(22) for the product state. Moreover, we can see how the statistical error of the mixed state is smaller. For a pure state, the density matrix represents a specific quantum state with no uncertainty, and fluctuations across the unitary ensemble can be significant. In contrast, mixed states represent ensembles of quantum states with varying probabilities, and this ensemble averaging leads to reduced fluctuations in the statistical properties. Besides, we acknowledge that the local protocol is more susceptible to statistical errors in comparison to the global protocol, probably because the estimator we are using on the global approach works better (see Appendix A for more information).

We can also foreshadow a serious problem with the protocol: when the total number of random unitaries $N_U$ is not big enough, the purity can be greater than 1. By definition, the purity of a reduced density matrix always satisfies $\operatorname{Tr}\{\rho_A^2\} \leq 1$, so getting a result over 1 does not have any physical sense. This problem can be explained by revisiting the mathematical derivation of the protocol. In the previous proof, we explicitly used the 2-fold twirl channel over a unitary ensemble $\mathcal{E}$, $\Phi^{(2)}_{\mathcal{E}}$, to efficiently substitute the 2-fold twirl channel over the Haar measure, $\Phi^{(2)}_{Haar}$. The problem is that we do not work with an ensemble $\mathcal{E}$ that spans a 2-design, but with a random ensemble that follows the distribution

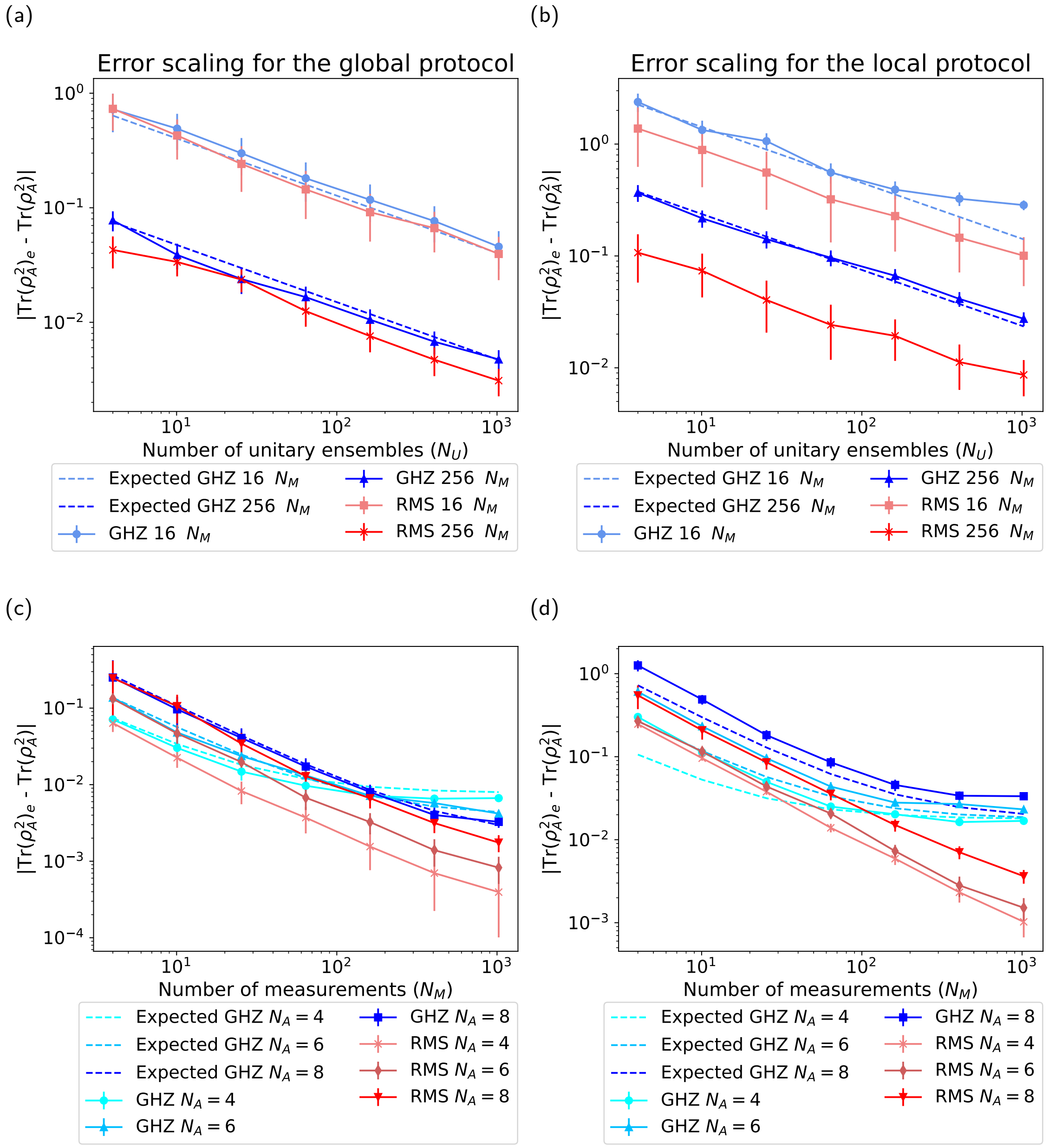

Figure 6: Error scaling of the randomized protocol. All plots show the difference in purity between the protocol and the expected value for systems of 10 qubits. Plots (a) global approach and (b) local approach as a function of the number of unitaries $N_U$ for a fixed set of measurements $N_M$ and different states, maximally entangled state GHZ (which is pure) and randomized mixed states RMS, in a subsystem size of $N_A = 8$ qubits. Plots (c) global approach and (d) local approach as a function of $N_M$ with fixed unitaries $N_U = 512$ and different subsystem size $N_A$. Dashed lines are estimated from the scaling laws. All data is extracted from 100 numerical experiments.

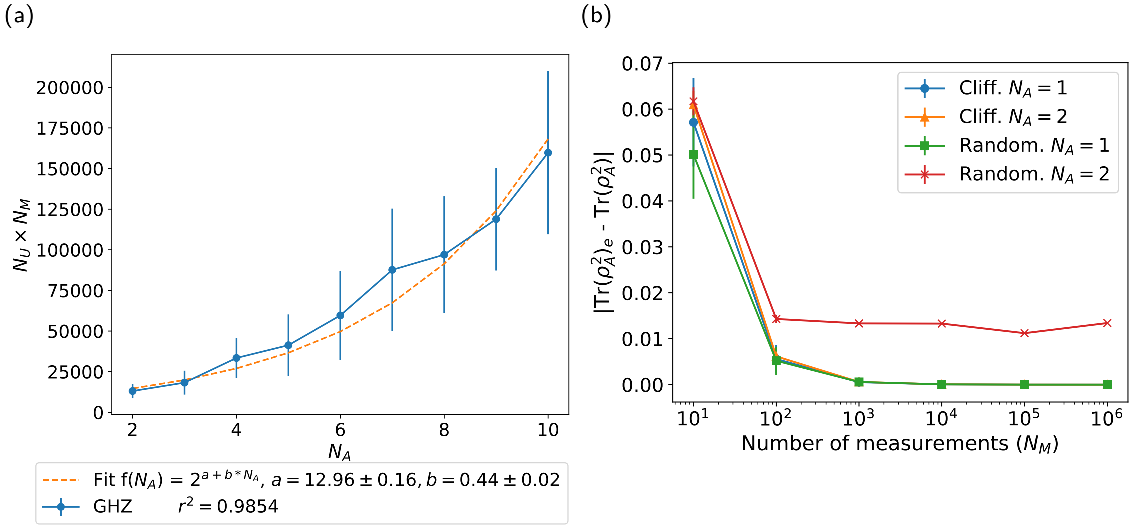

Figure 7: Error scaling of the randomized protocol. Plot (a) shows the required number of measurements $N_U N_M$ to determine the purity up to an average relative error inferior to 0.12 as a function of the subsystem size $N_A$. Plot (b) shows the difference of purity between the protocol and the expected value in the local approach as a function of the number of measurements $N_M$ for a fixed set of unitaries $N_U = 24^{N_A}$ in a system size of $N_A = \{1, 2\}$ qubits, for a GHZ state comparing random unitaries following the Haar measure (Random) and a set of fixed unitaries from the Clifford group (Cliff.).

given by the Haar measure and allows us to implement an $\epsilon$-approximate $t$-design. Thus, when $N_U$ is not big enough to reconstruct a good approximation of the entire Hilbert space, the equality does not hold and the obtained results should not be considered valid. We highlighted this flaw of the protocol, as it is not properly discussed on the original studies [68–70].

To continue with, Figure 7a shows the total number of shots $N_U N_M$ that must be performed to get an error lower than 12% as a function of the system size $N_A$, for a GHZ state. We have fitted the obtained data into a function of the form $2^{a+bN_A}$ so we could compare it with tomographic methods. We see that the required total number of measurements scales as $\sim 2^{(0.44\pm0.02)N_A}$, proving favorable compared to full state tomography [71].

Furthermore, we can see the effect of the chosen ensemble $\mathcal{E}$ in Figure 7b. Here, we are comparing the random ensemble extracted from the CUE $\mathcal{E}$ with a fixed ensemble given by the Clifford group $\mathcal{C}$ (composed of 24 elements). We are only focusing on subsystems of 1 and 2 qubits, as for the local unitary approach $N_U$ scales as $24^{N_A}$. A better discussion on the use of the Clifford group as a $t$-design can be consulted in Appendix B.2. We are also fixing $N_U = 24^{N_A}$ for both ensembles, otherwise the comparison would be futile. As we clearly see, using an exact unitary 2-design allows us to get better results for higher dimensions. In fact, for an infinite $N_M$, the estimation should be exact. This implementation is not usually practical, as we mentioned that $N_U$ scales exponentially as $24^{N_A}$, but may be of utility for QPUs of a few qubits, like the one available in the BSC with 5 qubits, where to estimate the entropy, the higher number of qubits we find in a subsystem is 2.

We would like to comment that all results simulated in *Qibo* are in good agreement with previous studies.

# 6 Conclusions

Throughout this research, we have presented cutting-edge techniques for estimating the $n$-order Rényi entropy of bipartite states on quantum computers, with the ultimate goal of developing quantum benchmarking protocols to assess entanglement properties. As a starting point, we provided a concise overview of various benchmarking protocols, serving as a gateway for new readers to this field. We ensured that the protocols were explained in a clear and accessible manner, accompanied by extensive references for those interested in delving deeper.

Next, we introduced a protocol utilizing the Swap Test to estimate the $n$-order Rényi entropy. This protocol is tailored for future fault-tolerant quantum computers and is restricted by the number of qubits available on the quantum device, as it relies on multiple identical copies of the quantum state. The error analysis we conducted on the protocol yields promising results, as with only a few thousand measurements on a single qubit, we achieved an error rate below 5% for pure states. With further research and improvements, and the future exploration and inclusion of mixed states, this protocol holds promise for enhancing the accuracy and scalability of entanglement estimation on future fault-tolerant quantum devices.

Lastly, we investigated a randomized protocol specifically designed for estimating the second-order Rényi entropy of bipartite states. This method was developed with NISQ devices in mind, as its randomized nature helps mitigate errors inherent in this kind of computers. We compared the local and global random unitary approaches and demonstrated that the global approach reduces statistical errors and improves performance. However, in the context of current NISQ devices, the local approach remains more practical. Additionally, our findings reveal that when the number of random unitaries $N_U$ is insufficiently large, the protocol yields non-physical outcomes. In such cases, it becomes challenging to achieve a reliable reconstruction of the complete Hilbert space. This limitation emphasizes the importance of employing a sufficiently high number of random unitaries in order to ensure accurate and meaningful results in the protocol. At last, we examined the scaling behavior of the protocol for the GHZ state, a maximally entangled pure state, and found it to exhibit favorable scaling compared to tomographic methods. However, further investigations are necessary to test this method against states with varying degrees of entanglement and different basis configurations.

Unfortunately, it is important to note that the protocols presented in this work primarily serve as entanglement detectors and quantifiers for pure states. In the case of mixed states, it only works as entanglement quantifiers but not as detectors. Certification would require the implementation of additional techniques, which can be achieved through randomized measurements. Although an initial study has been conducted and its findings can be consulted in Appendix C, further investigation in this area is planned for future research. By exploring the implementation of these techniques, we aim to enhance our understanding of entanglement properties in mixed states and develop more comprehensive benchmarking protocols. This ongoing exploration will contribute to the advancement of quantum benchmarking and provide valuable insights into the behavior of entangled states in practical scenarios.

Finally, we must mention that the initial main goal of this work was to run these protocols on the quantum computer of the BSC. Nevertheless, this was not possible due to some project delays, but it is planned to conduct the desired testing in the near future.

## Code repository

The current implementation of both quantum entanglement benchmarks programmed in *Qibo* is available on the github repository: `https://github.com/Juanfurk/Entanglement_measures/tree/main`.

## Acknowledgments

This work has been partially supported by the Ministry of Economic Affairs and Digital Transformation of the Spanish Government through the QUANTUM ENIA project call - Quantum Spain project, and by the European Union through the Recovery, Transformation and Resilience Plan - NextGenerationEU within the framework of the Digital Spain 2026 Agenda. I would like to express my deepest gratitude to the BSC for providing me with the opportunity to conduct my research and for hosting me throughout my master's thesis. I am immensely thankful to the BSC staff for their support, guidance, and assistance throughout my research. Their expertise and the resources provided by BSC have been invaluable to the completion of this work.

## References

[1] P. Shor, in Proceedings 35th annual symposium on foundations of computer science (1994), pp. 124–134.

[2] L. K. Grover, in Proceedings of the annual acm symposium on theory of computing (1996), pp. 212–219.

[3] J. Preskill, Quantum **2**, 79 (2018).

[4] M. Kjaergaard, M. E. Schwartz, J. Braumüller, et al., Annu. Rev. Condens. Matter Phys. **11**, 369–395 (2020).

[5] C. D. Bruzewicz, J. Chiaverini, R. McConnell, et al., Appl. Phys. Rev. **6**, 021314 (2019).

[6] V. T. Lahtinen and J. K. Pachos, SciPost Phys. **3**, 021 (2017).

[7] S. Slussarenko and G. J. Pryde, Appl. Phys. Rev. **6**, 41303 (2019).

[8] S. P. Harvey, *Quantum dots/spin qubits* (Oxford University Press, 2022).

[9] *Qibo | home page.* Available online: `https://qibo.science/` (accessed 06 July 2023).

[10] T. Lubinski, S. Johri, P. Varosy, et al., arXiv: 2110.03137 (2021).

[11] *Sri-international/qc-app-oriented-benchmarks: qed-c.* Available online: `https://github.com/SRI-International/QC-App-Oriented-Benchmarks` (accessed 20 March 2023).

[12] A. Li, S. Stein, S. Krishnamoorthy, et al., ACM Trans. Quantum Comput. **4**, 1–26 (2023).

[13] *Pnnl/qasmbench: a low-level openqasm benchmark suite for nisq evaluation and simulation. please see our paper for details.* Available online: `https://github.com/pnnl/QASMBench` (accessed 20 March 2023).

[14] T. Tomesh, P. Gokhale, V. Omole, et al., *Supermarq: a scalable quantum benchmark suite* (IEEE Computer Society, Feb. 2022), pp. 587–603.

[15] *Supertechlabs/superstaq-client.* Available online: `https://github.com/SupertechLabs/superstaq-client` (accessed 20 March 2023).

[16] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44-11**, 2455–2467 (1996).

[17] X.-D. Yu, J. Shang, and O. Gühne, Adv. Quantum Technol. **5**, 5 (2021).

[18] M. H. Devoret and R. J. Schoelkopf, Science **339**, 1169–1174 (2013).

[19] P. Stehle, Phys. Rep. **156**, 67–109 (1987).

[20] M. M. Amaral, L. V. Tarelho, M. A. De Souza, et al., J. Phys.: Conf. Ser. **733**, 012020 (2016).

[21] R. Blume-Kohout, J. K. Gamble, E. Nielsen, et al., arXiv: 1310.4492 (2013).

[22] S. T. Merkel, J. M. Gambetta, J. A. Smolin, et al., Phys. Rev. A **87** (2012).

[23] D. Greenbaum, arXiv: 1509.02921 (2015).

[24] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. A **85**, 042311 (2011).

[25] T. Proctor, K. Rudinger, K. Young, et al., Nat. Phys. **18**, 75–79 (2020).

[26] M. Sarovar, T. Proctor, K. Rudinger, et al., Quantum **4**, 321 (2020).

[27] D. C. McKay, A. W. Cross, C. J. Wood, et al., arXiv: 2003.02354 (2020).

[28] T. Proctor, S. Seritan, K. Rudinger, et al., Phys. Rev. Lett. **129**, 150502 (2022).

[29] J. Helsen, I. Roth, E. Onorati, et al., PRX Quantum **3**, 020357 (2022).

[30] A. Erhard, J. J. Wallman, L. Postler, et al., Nat. Commun. **10**, 5347 (2019).

[31] K. Mayer, A. Hall, T. Gatterman, et al., arXiv: 2108.10431 (2021).

[32] T. Proctor, S. Seritan, K. Rudinger, et al., Phys. Rev. Lett. **129**, 150502 (2021).

[33] A. W. Cross, L. S. Bishop, S. Sheldon, et al., Phys. Rev. A **100**, 032328 (2019).

[34] S. Aaronson and L. Chen, in Proceedings of the 32nd computational complexity conference (2017), pp. 1–67.

[35] R. Blume-Kohout and K. C. Young, Quantum **4**, 362 (2020).

[36] *Algorithmic qubits: a better single-number metric.* Available online: `https://ionq.com/resources/algorithmic-qubits-a-better-single-number-metric` (accessed 20 March 2023).

[37] *Ionq | trapped ion quantum computing.* Available online: `https://ionq.com/` (accessed 20 March 2023).

[38] F. Arute, K. Arya, R. Babbush, et al., Nature **574**, 505–510 (2019).

[39] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, et al., Nat. Phys. **14**, 595–600 (2018).

[40] J. Chen, D. Ding, C. Huang, et al., arXive: 2206.08293 (2022).

[41] D. Gottesman, arXiv: quant–ph/9807006 (1998).

[42] A. Wack, H. Paik, A. Javadi-Abhari, et al., arXiv: 2110.14108 (2021).

[43] A. Macaluso, L. Clissa, S. Lodi, et al., *A variational algorithm for quantum neural networks.* Vol. 12142 (Springer, Cham, 2020), pp. 591–604.

[44] S. Martiel, T. Ayral, and C. Allouche, IEEE Trans. Quantum Eng. **2**, 1–11 (2021).

[45] C. W. Commander, *Encyclopedia of optimization*, edited by C. A. Floudas and P. M. Pardalos (Springer US, Boston, MA, Aug. 2008), pp. 1991–1999.

[46] E. Farhi, J. Goldstone, and S. Gutmann, arXiv: 1411.4028 (2014).

[47] *Package for computing the atos q-score.* Available online: https://github.com/myQLM/qscore (accessed 20 March 2023).

[48] K. Mesman, H. Donkers, Z. Al-ars, et al., arXiv: 2103.17193 (2021).

[49] H. Donkers, K. Mesman, Z. Al-Ars, et al., arXiv: 2205.12142 (2022).

[50] R. B. Allan and R. Laskar, Discrete Math. **23**, 73–76 (1978).

[51] D. L. Applegate, R. E. Bixby, V. Chvátal, et al., *The traveling salesman problem: A computational study* (Princeton University Press, 2007), p. 608.

[52] *Qpack: application benchmark for quantum computing.* Available online: https://github.com/koenmesman/QPack (accessed 20 March 2023).

[53] *Benchmarks for multiple qaoa problems using the xacc platform for multiple quantum simulator backends.* Available online: https://github.com/huub-d96/xacc_qaoa_benchmarks (accessed 20 March 2023).

[54] P.-L. Dallaire-Demers, M. Stęchły, J. F. Gonthier, et al., arXiv: 2003.01862 (2020).

[55] V. E. Korepin, H. Frahm, F. Göhmann, et al., *The one-dimensional hubbard model* (Cambridge University Press, Jan. 2005), pp. 1–674.

[56] F. Levkovich-Maslyuk, J. Phys. A: Math. Theor. **49**, 323004 (2016).

[57] J. Tilly, H. Chen, S. Cao, et al., Phys. Rep. **986**, 1–128 (2022).

[58] A. Peruzzo, J. McClean, P. Shadbolt, et al., Nat. Commun. **5**, 4213 (2014).

[59] A. J. McCaskey, Z. P. Parks, J. Jakowski, et al., npj Quantum Inf **5**, 99 (2019).

[60] *Xacc - extreme-scale accelerator programming framework*, Available online: https://github.com/eclipse/xacc (accessed 20 March 2023).

[61] N. Friis, G. Vitagliano, M. Malik, et al., Nat. Rev. Phys. **1**, 72–87 (2019).

[62] K. E. Hamilton, N. Laanait, A. Francis, et al., arXiv: 2209.00678 (2022).

[63] C. Bravo-Prieto, C. Bravo-Prieto, D. García-Martín, et al., Phys. Rev. A **101**, 062310 (2020).

[64] X. Wang, Z. Song, and Y. Wang, Quantum **5**, 483 (2021).

[65] S. Johri, D. S. Steiger, and M. Troyer, Phys. Rev. B **96**, 195136 (2017).

[66] N. M. Linke, S. Johri, C. Figgatt, et al., Phys. Rev. A **98**, 052334 (2017).

[67] P. Horodecki and A. Ekert, Phys. Rev. Lett. **89**, 127902 (2001).

[68] A. Elben, B. Vermersch, C. F. Roos, et al., Phys. Rev. A **99**, 052323 (2018).

[69] T. Brydges, A. Elben, P. Jurcevic, et al., Science **364**, 260–263 (2018).

[70] A. Elben, S. T. Flammia, H.-Y. Huang, et al., Nat. Rev. Phys. **5**, 9–24 (2022).

[71] D. Gross, Y. K. Liu, S. T. Flammia, et al., Phys. Rev. Lett. **105**, 150401 (2010).

[72] D. A. Roberts and B. Yoshida, J. High Energ. Phys. **121**, 1–64 (2017).

[73] B. Vermersch, A. Elben, M. Dalmonte, et al., Phys. Rev. A **97**, 023604 (2018).

[74] *Qibo - random_ensambles (github).* Available online: https://github.com/qiboteam/qibo/blob/69665cfa94563ea8e5477dced9bd03b128226817/src/qibo/quantum_info/random_ensembles.py#L88 (accessed 20 March 2023).

[75] F. Mezzadri, arXiv: math–ph/0609050 (2006).

[76] Y. Nakata and M. Murao, Eur. Phys. J. Plus **129**, 152 (2014).

[77] Y. Nakata, C. Hirche, M. Koashi, et al., Phys. Rev. X **7**, 021006 (2017).

[78] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[79] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1–8 (1996).

[80] R. Horodecki, P. Horodecki, M. Horodecki, et al., Rev. Mod. Phys. **81**, 865 (2009).

[81] A. Elben, R. Kueng, H.-Y. Huang, et al., Phys. Rev. Lett. **125**, 200501 (2020).

[82] J. Gray, L. Banchi, A. Bayat, et al., Phys. Rev. Lett. **121**, 150503 (2017).

# A Error propagation

It is worth mentioning how the outcome probabilities $P(s_A)$ is computed. For the simple case of the probability $P$, a well-known estimator $\tilde{P}$ is found by performing $N_M$ measurements on the state and dividing the number of outcomes $s$ by the total $N_M$. Mathematically it is described by

$$\tilde{P} := \tilde{P}(s) = \frac{1}{N_M} \sum_{l=1}^{N_M} \delta_{s,l}, \tag{23}$$

where $l$ represents the variable describing the $l$-measurement. Going back to the randomized protocol, expressions (7) and (8) show the product of probabilities. This product can be performed between probabilities that will be the same and the rest that will be different. Thus, our interest is to find a better estimator to find the product of probabilities. It is not possible to find such an estimator for a product of different probabilities, but it is feasible to find one if those probabilities are the same, as we must estimate only $P^2$. Nevertheless, $\tilde{P}^2$ is a biased estimator of the probability squared. Thus, another unbiased estimator should be used, which we will denote by $\tilde{P}_2$ and is given by

$$\tilde{P}_2 = \frac{\tilde{P}(\tilde{P}N_M - 1)}{N_M - 1}. \tag{24}$$

We are using these unbiased estimators to simulate numerically the experiments with a finite number of measurements [73].

Regarding the confidence intervals of the data, we made use of one again of the Wald interval. The success probability $P$ is estimated as

$$P = \left( \tilde{P} \pm z\sqrt{\frac{\tilde{P}(1-\tilde{P})}{N_M}} \right), \tag{25}$$

where we use $z = 1$ for a standard error. In ther case of the Swap protocol, it is possible to study their error propagation through Eq.(6) using these error intervals, given by

$$\Delta R_n = \sqrt{(\Delta\tilde{P}_0)^2 + (\Delta\tilde{P}_1)^2}, \tag{26}$$

with $\Delta\tilde{P}_i = \sqrt{\frac{\tilde{P}_i(1-\tilde{P}_i)}{N_M}}$. Then, a mean of over 100 experiments is computed and the necessary statistical treatment through error propagation is performed.

Nevertheless, for the case where the product of probabilities differ $P_iP_j$, we use

$$P_iP_j = \left( \tilde{P}_i\tilde{P}_j \pm \sqrt{(\tilde{P}_i\Delta\tilde{P}_j)^2 + (\tilde{P}_j\Delta\tilde{P}_i)^2} \right). \tag{27}$$

However, for the case where the probabilities are the same $P_i^2$, we get

$$P_i^2 = \left( \tilde{P}_2 \pm \frac{2\tilde{P}N_M - 1}{N_M - 1}\Delta\tilde{P} \right). \tag{28}$$

From these error intervals, it is possible to study the error propagation through Eq.(7) and Eq.(8), using well-established statistical methods. Then, a mean of over 100 experiments is computed and the necessary statistical treatment is performed to compute the average and their error intervals.

# B Random unitaries

Before embarking on the formal demonstration of expressions (7) and (8), it is imperative to address pertinent topics that provide a comprehensive understanding of the procedure and shed light on the subsequent steps involved.

## B.1 The Haar measure

Quantum computer operations are mathematically described by unitary matrices, which can be parameterized using a set of specific coordinates. A single-qubit local unitary matrix can be conceptually understood as a rotation around the Bloch sphere, allowing any initial vector state to be transformed to any other point on the sphere without altering its dimensions. The most general rotation can be expressed according to three parameters as

$$U(\theta,\phi,\omega)=\begin{pmatrix} e^{-i\frac{\phi+\omega}{2}}\cos\frac{\theta}{2} & -e^{-i\frac{\phi-\omega}{2}}\sin\frac{\theta}{2} \\ e^{-i\frac{\phi-\omega}{2}}\sin\frac{\theta}{2} & e^{i\frac{\phi+\omega}{2}}\cos\frac{\theta}{2}. \end{pmatrix} \tag{29}$$

Furthermore, it is important to recall one of the general properties that define the set of unitary matrices: given a unitary $U$, its conjugate transpose (denoted by a dagger †) is its inverse,

$$UU^\dagger = U^\dagger U = \mathcal{I}. \tag{30}$$

The set of unitary matrices of size $N\times N$ constitute the so-called unitary group $U(N)$. As a group, it is possible to sample uniformly from it. When doing the sample, it is important to add correctly the measure. The measure describes the distribution of each parameter according to the space they live in, weighing points differently depending on which part of the space they inhabit. It allows to sample correctly, so the points are uniformly distributed. Therefore, the Haar measure emerges as the correct measure to work with when playing with the unitary group. It is worth mentioning that the distribution of the Circular Unitary Ensemble (CUE) follows this Haar measure on $U(N)$.

Suppose a function $f$ acting on the elements of the unitary group $U(N)$ and we would like to compute the integral over the whole group. Then, we would need the Haar measure, usually denoted by $\mu_N$, to tell us how the elements are distributed inside the group. This way, the integral could be computed as

$$\int_{V\in U(N)} f(V)d\mu_N(V). \tag{31}$$

*Qibo* offers the command 'qibo.quantum_info.random_unitary()' to compute random local unitaries (more details on its repository [74]). Besides, with the tool *measure* = 'haar', it is possible to compute a random local unitary following the Haar measure. The method used to perform the sample is by taking the QR decomposition of a complex matrix. The procedure is well explained in Ref. [75].

An important property of the Haar measure that is worth mentioning is the left and right invariance under unitary transformations. That is

$$\int_{V\in U(N)} f(WV)d\mu_N(V)=\int_{V\in U(N)} f(VW)d\mu_N(V)=\int_{V\in U(N)} f(V)d\mu_N(V), \tag{32}$$

which holds for any $W\in U(N)$, as the random nature of $V$ ensures that the product with any other unitary matrix remains random as well.

As we saw, our randomized protocol requires that we compute the so-called $k$-fold channel of a given operator $O$, of the form

$$\Phi^{(k)}_{Haar}(O) = \int_{Haar} dU (U^{\dagger})^{\otimes k} O U^{\otimes k}, \tag{33}$$

in order to compute the average ensemble over the cross-correlations of the probability $\overline{P_U(s)P_U(s')}$. However, computing this value requires an infinite number of unitaries drawn from the Haar distribution, which is not feasible in practice. Similarly, numerically evaluating the integral becomes impractical due to the overwhelming number of variables involved. Consequently, an alternative approach emerges: the use of unitary $t$-designs. These designs enable us to determine the exact value of the integral using a finite set of unitary matrices, circumventing the challenges posed by the infinite parameter space.

### B.2 Unitary t-design

Before diving into the unitary design, it may be better to have a look at the spherical design to gain some intuition. Suppose we have a polynomial in $d$ variables and we would like to compute its average over the surface of a $d$-dimensional sphere, $S(\mathcal{R}^d)$. One way to do it is by integrating over the sphere using the proper measure, but that would imply keeping track of a lot of parameters. Another simple possibility would be to sample random uniformly distributed points from the sphere, evaluate the function at those points and then compute the average value. This method is intuitive and will always take us close to the exact value, but it is just an approximation.

Therefore, a spherical $t$-design has been introduced to compute the exact result in a simpler and faster way. They can be understood as a set of points evenly distributed on the surface of the sphere. The only limitation of the $t$-design is that they can only be used if the terms in the polynomial have all the same degree of at most value $t$. Here we present a proper definition:

Let $p_t : S(\mathcal{R}^d) \to \mathcal{R}$ be a polynomial in $d$ variables, with all terms the same degree, at most $t$. A set $X = \{x : x \in S(\mathcal{R}^d)\}$ is a spherical $t$-design if

$$\frac{1}{|X|} \sum_{x \in X} p_t(x) = \int_{S(\mathcal{R}^d)} p_t(u) d\mu(u), \tag{34}$$

where $d\mu$ is the spherical measure. It is worth mentioning that a spherical $t$-design is also a $k$-design for all $k < t$.

With all this in mind, we can approach now the unitary design. Basically, unitary designs extend the previous concept from evenly-distributed points to evenly-distributed unitaries. Here we present a proper definition:

Let $P_t(U)$ be a polynomial with all degrees the same, at most $t$, in $d$ variables in the entries of a unitary matrix $U$. A unitary $t$-design is a set of $K$ unitaries $\{U_k\}$ such that

$$\frac{1}{K} \sum_{k=1}^{K} P_t(U_k) = \int_{V \in U(d)} P_t(V) d\mu(V), \tag{35}$$

where, one again, $d\mu(V)$ is the proper Haar measure.

Therefore, one can say that an ensemble $\mathcal{E} = \{p_i, U_i\}$ forms a unitary $k$-design if and only if $\Phi^{(k)}_{\mathcal{E}} = \Phi^{(k)}_{Haar}$. Once again, for our case, evaluating $\overline{P_U(s)P_U(s')}$ requires a unitary 2-design, as the unitary matrices $U$ and $U^{\dagger}$ of the $k$-fold channel (33) appear at the same degree of 2. A proper unitary 2-design to work with is the Clifford group.

The Clifford group $\mathcal{C}_n$ exhibits the property of being a unitary 3-design, thereby also serving as a 1-design and 2-design. This group consists of unitary matrices that, when applied to an $n$-qubit system, can transform any Pauli operator into another Pauli operator. In the case of the 1-qubit Clifford group $\mathcal{C}_1$, which is particularly relevant for the local unitary scenario of the randomized protocol, it comprises a mere 24 elements, which are primarily combinations of Hadamard and Phase gates. Consequently, when evaluating the purity in the randomized protocol, there is no longer a need for multiple random matrices following the Haar measure. Instead, only the 24 unitaries from $\mathcal{C}_1$ are required. However, a challenge arises with the size of the subsystem $N_A$. Since there are 24 elements, the number of unitaries to be applied $N_U$, scales exponentially, specifically as $24^{N_A}$. This exponential growth renders the "exact" evaluation of purity using unitary $k$-designs impractical, except for cases with a small number of qubits, such as 1, 2, or even 3. Furthermore, the impracticality intensifies when dealing with the global unitary measure, as the number of elements in $\mathcal{C}_n$ increases exponentially. To illustrate, the $\mathcal{C}_2$ group encompasses a staggering 11 520 elements, making the procedure highly unfeasible.

This way, when dealing with a large number of qubits, the randomized protocol must be performed using the so-called $\epsilon$-approximate $k$-design. Given by an ensemble $\mathcal{E}$, it allows to get and approximation of the form $\parallel \Phi_{\mathcal{E}}^{(k)} - \Phi_{Haar}^{(k)} \parallel < \epsilon$. On one hand, in the case of local unitaries, this ensemble will be nothing more than the needed number of random local unitaries following the Haar measure. If the number of unitaries drawn from the CUE is not enough, then our $\epsilon$-approximate $k$-design is not sufficient to obtain good data and nonphysical values may arise, just like we can observe in our simulations. On the other hand, the global unitary implementation is not straightforward. In our case, we decided to implement a global efficient unitary design using random diagonal circuits (RDC) [76], following the implementation given in [77].

While this procedure does not yield an "exact" result, it provides a reliable approximation. It allows us to obtain results that are sufficiently close to the outcome we would get using the complete Haar measure. Therefore, even though we do not achieve perfect accuracy, we can consistently obtain a good approximation of the desired results.

## C Mixed-state discussion of the protocols

All these presented protocols work assuming that the initial quantum state is pure. But we must also discuss the case where the initial bipartite state $\rho_{AB}$ is mixed. In this section, we will take a closer look and talk about it.

First, let's consider a pure bipartite state of the form $\rho_{AB} = |\psi\rangle\langle\psi|$. General properties of pure states tell us that the purity $\mathrm{Tr}\{\rho_{AB}^2\} = 1$ always. Now, when tracing over $B$, the reduced state $\rho_A$ will contain the information shared with system $B$. Thus, if the purity $\mathrm{Tr}\{\rho_A^2\} = 1$, that means that our original pure bipartite state could have been written as a product $\rho_{AB} = \rho_A \otimes \rho_B$. However, if $\mathrm{Tr}\{\rho_A^2\} < 1$ that means some correlation existed between subsystems $A$ and $B$ and, therefore, that some entanglement is manifested in the system. It could seem at first that purity could work as an entanglement checker, but we should continue.

Let's see what happens when working with bipartite mixed states. In this case, a bipartite state is said to be separable if $\rho_{AB} = \sum_i p_i \rho_{A,i} \otimes \rho_{B,i}$ and no entanglement is presented between both subsystems [78]. Nevertheless, it is straightforward to check that $\rho_A$ becomes a mixed state,

$$\rho_A = \mathrm{Tr}_B\, \rho_{AB} = \sum_i p_i \,\mathrm{Tr}_B\, \rho_{A,i} \,\mathrm{Tr}_B\, \rho_{B,i} = \sum_i p_i \rho_{A,i} \tag{36}$$

which means it fulfils the property $\mathrm{Tr}\{\rho_A^2\} < 1$ (presented on all mixed states). In this case, the purity does not seem to play the same role as before (for pure states) and now it does not work as an entanglement detector, as our initial hypothesis is that this bipartite state is separable and thus with no entanglement. A useful example to illustrate this dilemma is the one presented by the Werner state,

$$\rho_{AB}^W = p\,|\Psi^-\rangle\langle\Psi^-| + (1-p)\frac{\mathbf{I}}{4}, \tag{37}$$

This bipartite state is clearly mixed, but is it also separable? Finding the density matrix of subspace A, $\rho_A$, we see

$$\rho_A = \mathrm{Tr}_B\, \rho_{AB} = \frac{\mathbf{I}}{2}, \tag{38}$$

which is again a mixed state. That means its purity is $\mathrm{Tr}\{\rho_A^2\} < 1$, in particular $\mathrm{Tr}\{\rho_A^2\} = 1/2$. So, we could think that it may present some entanglement. However, by applying the Peres-Horodeki criterion [79] we find that it actually becomes an entangled mixed state for $p > 1/3$ (whenever there is a negative eigenvalue). This shows perfectly how purity does not work properly as a criterion to distinguish entanglement.

This way, the protocols discussed in this work only run under the assumption that the initial quantum state is pure. We can always run these methods to estimate the Rényi entropy, but that does not mean that actual entanglement is present in the system if the initial state is mixed. In this case, only once entanglement on the system is proven, we can accept the 2-order Rényi entropy estimation.

Furthermore, there is a straightforward method to guarantee entanglement certification of a state [80]: entanglement exists between two subsystems of a bipartite state if

$$\begin{aligned} S_2(\rho_A) &> S_2(\rho_{AB}) \\ S_2(\rho_B) &> S_2(\rho_{AB}) \end{aligned} \tag{39}$$

However, this criterion is sufficient but not necessary. Thus, entanglement may still be present when Eqs.(39) do not hold.

## C.1 $p_3$ - PPT Condition

As a matter of fact, a pure state only exists as an idea. The real world is full of actually highly mixed states, either due to decoherence or because they describe a subregion of a larger and globally entangled system. Thus, developing protocols to first detect entanglement in a mixed state is crucial. As we have previously seen, computing the Rényi entropy does not assure that the state presents entanglement; only if entanglement is present we can then compute its entropy to quantify it.

Therefore, the randomized protocol we previously presented can only be used assuming that the initial state is pure, so if we get any value on the entropy that is greater than zero it is sufficient to tell that the initial state exhibits any kind of entanglement. However, if the initial state is mixed, we cannot make the same statement. That is why the $p_3$ - PPT condition has been proposed and experimentally demonstrated to ensure entanglement certification [81]. As the name tells, the protocol is based on the Positive Partial Transpose Condition (PPT condition), which checks if the partially transpose density matrix $\rho_{AB}^{T_A}$

is positive semidefinite (i.e. all eigenvalues are non-negative). If the PPT condition is violated, A and B must be entangled. Furthermore, it is possible to turn this condition into a quantitative entanglement measure, the *negativity*, $\mathbf{N}(\rho_{AB}) = \sum_{\lambda<0} |\lambda|$ (where $\lambda$ are the eigenvalues), which is positive if and only if the state violates the PPT condition.

Unfortunately, computing the negativity requires estimating the spectrum of $\rho_{AB}^{T_A}$ accurately (i.e. use tomography techniques). This challenge is overcome by the moments of the partially transposed density matrix (PT-moments):

$$p_n = \mathrm{Tr}\Big\{(\rho_{AB}^{T_A})^n\Big\}, \qquad \text{for} \quad n = 1, 2, 3\ldots \tag{40}$$

Let's take a look at the first values. $p_1 = \mathrm{Tr}\{\rho_{AB}\} = 1$ always following the definition of density matrix; $p_2 = \mathrm{Tr}\{\rho_{AB}^2\}$ is the purity of the system, which we have already discussed that does not work as an entanglement certification measure. Hence, $p_3$ is the lowest PT-moment that can capture meaningful information about the partial transpose [82]. PT-moments can be used to define a simple and powerful test for bipartite entanglement certification:

$$\rho_{AB} \in PPT \implies p_3 \geq p_2^2. \tag{41}$$

We will name $p_3$-PPT condition to the contra-positive of the previous assertion. Thus, if $p_3 < p_2^2$, then $\rho_{AB}$ violates the PPT condition and $A$ and $B$ must be entangled.

It is important to emphasize that the randomized protocol can be implemented globally to estimate higher Rényi orders recursively, and thus estimate $p_3$. In Ref. [73] they showed that:

$$\overline{P_U(s)^k} = \frac{1}{\mathcal{D}_k} \sum_{\substack{b_1,\ldots,b_k \in \mathcal{N}_0 \\ \text{conditioned to} \\ \sum_{l=1}^k l b_l = k}} C_{b_1,\ldots,b_k} \prod_{l=1}^{k} \mathrm{Tr}\Big\{\rho^l\Big\}^{b_l}, \tag{42}$$

where $\mathcal{D}_k = \prod_{i=0}^{k-1}(\mathcal{D}+i)$ and $C_{b_1,\ldots,b_k}$ denotes the number of permutations $\sigma \in \mathcal{S}_k$ with typ($\sigma$)= $1^{b_1} 2^{b_2} \ldots k^{b_k}$ given by

$$C_{b_1,\ldots,b_k} = \frac{k!}{b_1! \cdot b_2! \cdot \ldots b_k! \cdot 1^{b_1} \cdot 2^{b_2} \cdot \ldots k^{b_k}}. \tag{43}$$

With all this, we can get an expression for $p_3$ as

$$\mathrm{Tr}\Big\{\rho^3\Big\} = \frac{1}{2}\left((\mathcal{D}+1)(\mathcal{D}+2)\sum_s \overline{P_U(s)^3} - 3\,\mathrm{Tr}\Big\{\rho^2\Big\}\right), \tag{44}$$

where $p_2$ can be estimated with the randomized method we addressed on this work. Nevertheless, the generalization of Eq.(44) using local unitaries is not simple and has not been studied yet.