Quantum Resources in Quantum Technologies

University of Warsaw, Faculty of Physics
Doctoral Thesis

Author: Moein Naseri
Supervisor: Dr. hab. Alexander Streltsov
Assistant Supervisor: Dr. Marco Fellous-Asiani

Abstract

This thesis delves into the fundamental quantum resources, say quantum coherence, entanglement and purity, that are essential for the development and optimization of quantum technologies. We study quantum speed limits, resource generation, and the impact of these resources on quantum algorithms and noisy quantum circuits.

As in many quantum technologies, one has to convert a collection of states through some unitary dynamics (in the lab), we introduce a novel notion of quantum speed limits which concerns transforming a collection of quantum states simultaneously, such as converting a basis into maximally coherent bases and basis permutation. The speed limits are particularly important when there is a limitation to implement arbitrarily many types of dynamic in the lab. We derive tight bounds for these transformations in systems of lower dimensions and establish general bounds for multi-qubit and higher-dimensional Hilbert spaces. Additionally, we present exact expressions for the speed limit of basis permutations in Hilbert spaces of arbitrary dimension. Using the tools we developed, we explore the minimal time needed to generate a specific amount of coherence under unitary evolution and building on it, we define the concept of the coherence-generating capacity of Hamiltonians, which measures the maximal rate at which coherence can be generated by a quantum system. Using the relative entropy of coherence as a figure of merit, we derive closed-form expressions for Hamiltonians and quantum states that achieve this maximum under the constraint of a bounded Hilbert-Schmidt norm for the Hilbert spaces of arbitrary dimension. These results provide valuable insights for optimizing quantum coherence in various quantum tasks. For the qubit systems, we find a complete characterization of the problem of coherence generation rate by a given Hamiltonian.

We further examine the role of quantum coherence and entanglement in quantum algorithms, focusing on the Bernstein-Vazirani algorithm and its probabilistic variant. We find an analytic relation connecting the performance of the algorithm with the coherence of the initial state. Our analysis further reveals that excessive entanglement can hinder optimal performance. In the context of quantum computation with mixed states, we demonstrate that pseudo-pure states can achieve optimal performance of this algorithm for a given level of purity.

We further study the resources in restricted model of computations in which a unitary is controlled by a single qubit and we show that the one clean qubit model can provide computational advantages even with minimal entanglement, coherence, and general quantum correlations. Motivated by the fact that the one clean qubit model provides computational advantages while working with a register of qubits in the fully mixed states, we then study if similar circuits could lead to some computational advantages in the case each gate in the computer happens to be noisy. For this, we consider a noise model inspired by superconducting cat qubits, which only introduces bit-flip after each gate of computation, and we show that using the asymmetry (which technically is called biasedness) of the noise (as only bit-flip errors are involved) and the coherence of the control qubit, one can design a class of noisy circuits which are highly resilient to the noise. However, we further show that this class of circuits is classically simulable. Using the simulability of these circuits, we establish a novel way, to benchmark the biasedness of the noise at the scale of the whole circuit. This benchmark protocol is particularly important as having this noise model is crucial for the scalability of the cat qubit circuits and it is not clear if they can maintain this property in large circuits due to crosstalks and correlated errors. The protocol is sensitive to these effects and is capable of validating the performance of circuits with up to 10^6 gates under realistic noise models, extending beyond traditional Pauli noise models.

Together, these results pave the way and show new directions for a more rigorous and comprehensive understanding of how various quantum resources can be quantified, optimized, and leveraged to enhance the performance and reliability of quantum technologies.

Streszczenie

Praca ta zagłębia się w tematykę podstawowych zasobów kwantowych, takich jak spójność kwantowa, splątanie i czystość, które są niezbędne do rozwoju i optymalizacji technologii kwantowych. Badamy kwantowe ograniczenia prędkości, generowanie zasobów oraz wpływ tych zasobów na algorytmy kwantowe i szumy w obwodach kwantowych.

Ponieważ w wielu technologiach kwantowych należy przekształcać zbiór stanów poprzez dynamikę unitarną (w laboratorium), wprowadzamy nowę koncepcję kwantowych ograniczeń prędkości, która dotyczy jednoczesnych transformacji zbioru stanów kwantowych, takich jak konwersja bazy na bazy maksymalnie spójne i permutacja baz. Ograniczenia prędkości są szczególnie ważne, gdy istnieje przeszkoda w implementacji dowolnie wielu typów dynamiki w laboratorium. Wyprowadzamy ścisłe ograniczenia dla tych transformacji w systemach o niższych wymiarach i ustalamy ogólne ograniczenia dla systemów wielokubitowych i przestrzeni Hilberta o wyższych wymiarach. Dodatkowo, przedstawiamy analityczne wyrażenia na ograniczenia dla permutacji baz w przestrzeniach Hilberta o dowolnym wymiarze. Korzystając z opracowanych narzędzi, badamy minimalny czas potrzebny do wygenerowania określonej ilości spójności poprzez ewolucję unitarną i na tej podstawie definiujemy pojęcie zdolności generowania spójności przez hamiltoniany, które mierzy maksymalną szybkość, z jaką spójność może być generowana przez system kwantowy. Używając względnej entropii spójności jako kryterium, wyprowadzamy ścisłe wyrażenia na hamiltoniany i stany kwantowe, które osiągają to maksimum pod warunkiem ograniczonej normy Hilberta-Schmidta, dla przestrzeni Hilberta o dowolnym wymiarze. Te wyniki dostarczają cennych informacji na temat optymalizacji spójności kwantowej w różnych zadaniach kwantowych. W systemach qubitowych znajdujemy pełne rozwiązanie problemu szybkości generowania spójności przez dany Hamiltonian.

Dalsze badania dotyczą roli spójności kwantowej i splątania w algorytmach kwantowych, skupiając się na algorytmie Bernsteina-Vaziraniego i jego odmianie probabilistycznej. Znajdujemy analityczną zależność łączącą wydajność algorytmu ze spójnością stanu początkowego. Nasza analiza ujawnia również, że nadmierne splątanie może obniżać optymalną wydajność. W kontekście obliczeń kwantowych z użyciem stanów mieszanych, wykazujemy, że stany pseudoczyste mogą osiągnąć optymalną wydajność

tego algorytmu dla pewnego poziomu czystości.

Kolejne rozważania dotyczą zasobów w szczególnym modelu obliczeń, w którym bramka unitarna jest kontrolowana przez jeden qubit, i wykazujemy, że model jednego czystego qubitu może zapewnić przewagę obliczeniową nawet przy minimalnym splątaniu, spójności i ogólnych korelacjach kwantowych. Motywowani faktem, że model jednego czystego qubitu zapewnia przewagi obliczeniowe, gdy dostępna jest pewna ilość qubitów znajdujących się w stanie w pełni mieszanym, badamy, czy podobne obwody mogą prowadzić do przewag obliczeniowych w przypadku, gdy wszystkie bramki w komputerze są zaszumione. W tym celu rozważamy model szumu inspirowany nadprzewodzącymi qubitami w stanie kota Schrödingera, który wprowadza jedynie błąd typu bit-flip po każdej bramce obliczeniowej, i wykazujemy, że dzięki asymetrii szumu (również zwanej stronniczością) oraz spójności qubitu kontrolnego, można zaprojektować klasę zaszumionych obwodów wysoce odpornych na szum. Pokazujemy jednak również, że ta klasa obwodów jest symulowalna klasycznie. Korzystając z symulowalności tych obwodów, wprowadzamy nowy sposób ustalania punktu odniesienia dla stronniczości szumu w skali całego obwodu. Taki protokół oceny jest szczególnie ważny, ponieważ model szumu jest kluczowy dla skalowalności obwodów qubitów w stanie kota Schrödingera i nie jest jasne, czy będą one w stanie utrzymać tę właściwość w dużych obwodach z powodu przesłuchu i skorelowanych błędów. Protokół jest zdolny do wykrywania tych efektów i weryfikacji wydajności obwodów zawierających do 10^6 bramek w realistycznych modelach szumu, wykraczających poza tradycyjne modele szumów Pauliego.

Wspólnie, te wyniki wskazują ścieżki i wyznaczają nowe kierunki dla bardziej ścisłego i kompleksowego zrozumienia, jak różne zasoby kwantowe mogą być kwantyfikowane, optymalizowane i wykorzystywane do poprawy wydajności i niezawodności technologii kwantowych.

Publications

The content of this dissertation is written mainly based on the results published in the following works by the author (of the dissertation), his supervisor and colleagues. These articles were published for the purpose of the dissemination of results of the works done by the author during his PhD studies.

- [NKG+22]: Entanglement and coherence in Bernstein-Vazirani algorithm

  Moein Naseri, Tulja Varun Kondra, Suchetana Goswami, Marco Fellous-Asiani, and Alexander Streltsov

  Phys. Rev. A **106 (6)** 062429 — Published 21 December 2022

- [NMB+24]: Quantum speed limits for change of basis

  Moein Naseri, Chiara Macchiavello, Dagmar Bruss, Paweł Horodecki, and Alexander Streltsov

  New J. Phys. **26 (2)** 023052 — Published 26 February 2024

- [SNS24]: Coherence generation with Hamiltonians

  Manfredi Scalici, Moein Naseri, and Alexander Streltsov

  Quantum Information and Computation, Vol. 24, No. 7 and 8 (2024) 0565–0575

- [FAND+23]: Scalable noisy quantum circuits with biased-noise qubits

  Marco Fellous-Asiani, Moein Naseri, Chandan Datta, Alexander Streltsov, and Michał Oszmaniec

  arXiv:2305.02045 [quant-ph]

My other research papers that are not included in this thesis are

- Local Purity Distillation in Quantum Systems: Exploring the Complementarity Between Purity and Entanglement

  Ray Ganardi, Piotr Masajada, Moein Naseri, Alexander Streltsov

  arXiv:2311.11820 [quant-ph]

- Coherence-based characterization of macroscopic quantumness
  Moein Naseri, and Sadegh Raeisi
  Phys. Rev. A **103 (3)** 032209 — Published 21 March 2021

Acknowledgements

I am deeply grateful to my parents and my brother, Soroush, whose unwavering support and encouragement have been a constant source of strength throughout my academic journey. I also extend heartfelt thanks to my grandfather, who patiently nurtured my critical and mathematical thinking from an early age. His wisdom and guidance have been a lifelong inspiration.

I wish to acknowledge Mr. M.H. Kheyrabadi, my physics teacher, who made me interested in the subject. His teachings laid the foundation for my understanding of problem-solving in the physical sciences. I also want to express my appreciation to my exceptional English teacher—Mr. Sadegh Tohidi, Mr. Mahmoudi, and Mr. Heyran. The knowledge and skills they taught me and their mentorship were instrumental in broadening my intellectual horizons, particularly in my chosen field of study.

I am grateful to all the teachers, professors, and collaborators, specifically Prof. S.J Akhtarshenas who have guided me at every stage of my academic life. In particular, I would like to thank my supervisor, Prof. Alexander Streltsov, for his firm support and patient guidance throughout the entirety of my PhD journey. His mentorship has been a source of great learning and inspiration for me.

Finally, I extend my thanks to Mr. Marcin Koźbiał for his assistance in translating the abstract of this thesis into Polish as well as reviewing the abstract.

A special thank goes to Agata Julia Szabłowska, who supported me to the full during my PhD years and has been the warmth and strength of my heart.

Contents

Chapter 1

Introduction

Quantum resource theories offer a structured method for examining the properties of quantum systems and their applications in quantum technologies [CG19]. Within this framework, the resource theories of entanglement [HHHH09a] and coherence [SAP17, WSR+21] are particularly notable examples.

The resource theory of entanglement studies the potential and limitations of agents working in separate quantum labs who can only communicate through classical means [HHHH09a]. In this theory, two parties, Alice and Bob, restricted to local operations and classical communications (LOCC), cannot create entanglement from a product state like $|0\rangle_{Alice} \otimes |0\rangle_{Bob}$ [CG19, Sha19, CLM+14]. This restriction shows that certain quantum state transformations cannot be achieved through LOCC alone. However, suppose Alice and Bob share a maximally entangled state. In that case, they can convert it into any other state deterministically [CG19] which demonstrates that entanglement is a valuable resource for state preparation [CG19, Sha19, CLM+14].

On the other hand, the resource theory of coherence investigates the difficulties and possibilities for an agent constrained in their capacity to generate and maintain quantum coherence [SAP17, WSR+21]. Additionally, the framework of quantum resource theories has been effectively utilized in the area of quantum thermodynamics [HO13, NW18]. This application has facilitated a deeper comprehension of how quantum systems can be controlled within the limits regarding the energy of the systems. In the following sections we aim to delve into a more rigorous discussion of these resources.

1.1 Quantum Resource Theories

A quantum resource theory (QRT) is a framework that outlines how one quantum state can be transformed into another using a specific set of quantum operations under defined conditions [CG19, Sha19, CLM+14]. Each quantum resource theory is built upon two essential concepts: free states and free operations [CG19]. Free states are those quantum states that can be easily created within a framework supported by physical principles. Regarding free operations, these ideally represent quantum manipulations that can be effortlessly carried out, founded on the physical principles underpinning the resource theory. For example, in the realm of entanglement resource theory, the combination of local operations and classical communication represents a set of free operations that can be easily interpreted in physical terms [BDSW96]. One may conclude, any other state outside the set of free states would be considered a resourceful state within the frame of our QRT. Hence, QRTs help one to identify and measure such resources, like entanglement, which facilitate specific quantum operations and transformations. In the following, we will present the formal definition of a quantum resource theory [CG19].

Definition 1.1. A quantum resource theory is defined by the pair (F, O_F), where F represents a set of quantum states known as free states, and O_F denotes a set of completely positive and trace-preserving maps (CPTP), referred to as free operations, such that:

$$\forall\ \Lambda \in O_F, \rho \in F\colon \Lambda(\rho) \in F. \tag{1.1}$$

In other words, the free operations map any free state to another free state. Note that this is also an important criterion for checking the consistency of the resource theory of interest.

Various quantum resource theories have been developed, including those for asymmetry, nonlocality, coherence, purity, and entanglement. For a comprehensive review of these quantum resource theories, one can refer to [CG19]. In this dissertation, we will focus on three key resource theories: purity, coherence, and entanglement. These resource theories are particularly important in the context of quantum technologies, such as quantum computation, quantum error correction, and quantum control [Wan23, DB07, MEKP16a, TV14]

In the framework of resource theory, quantifying the amount of the resource present in any state is a useful aspect. This quantification should be grounded in practical and meaningful tasks that highlight the utility of the resource [CG19]. For example, in the context of entanglement theory, the primary objective often involves distilling singlet states from a given quantum state. Singlet states are particularly valuable because they can be used to generate any other entangled state [Sha19]. Thus, a natural and meaningful way to measure entanglement is by assessing how many singlet states can be distilled from the state which ensures that the quantification aligns with the practical utility of entanglement.

Moreover, a well-defined resource quantification method must satisfy certain criteria to be meaningful: One expects that such a quantification must lead to zero for all the free states of the QRT. Also, it must not increase under the free operations otherwise it gives the possibility that the free operations do some valuable task for us and generate more amount of resources.

Definition 1.2. A resource measure is a map $M : L(\mathcal{H}) \longrightarrow R^+$ ($L(\mathcal{H})$ denotes the set of linear operations on the Hilbert space $\mathcal{H}$) such that:

- $M(\rho) = 0$ iff $\rho \in F$.
- (Monotonicity) $\forall \Lambda \in O_F$ we have $M(\Lambda(\rho)) \leq M(\rho)$.

In certain resource theories, there are special classes of states from which it is possible to derive any other state using the permitted free operations [CG19]. For instance, in the resource theory of bipartite entanglement, maximally entangled states are prime examples of such a special state. These can be transformed into any other state deterministically by using local operations and classical communication (LOCC) alone [CG19, Sha19]. These highly versatile states are referred to as maximally resourceful states [CG19]. In essence, maximally resourceful states are those that possess the highest possible level of resource value within a given theory, enabling them to serve for generating other states.

Definition 1.3. In the quantum resource theory (F, O_F), a maximally resourceful state (MRS) is a state that can be transformed to any other state by the usage of free operations [CG19].

It's worth mentioning that not all resource theories include a MRS. For instance, in the resource theory of multipartite entanglement with more than two parties ($N > 2$) there is not only one class of states that serves as an MRS for the entire system. Instead, there are various classes, each with its own distinct maximally resourceful states [OS06].

1.2 Quantum Resource Theory of Purity

A quantum state ρ is considered pure if and only if $\mathrm{Tr}(\rho^2) = 1$. However, to better understand and quantify the concept of purity, it is useful to establish a hierarchical structure within the space of quantum states. This hierarchy allows us to address the question: "To what extent is a quantum state pure?" . Resource theory of purity is rather a simple framework in which we can talk about the purity of quantum state as a resource [SKW+18].

Definition 1.4. The resource theory of purity is a quantum resource theory with the pair (F^p, O_F^p) (the pair of free states and free operations) such that:

$$F^p = \{\frac{\mathbb{I}}{d}\}, \tag{1.2}$$

$$O_F^p = \{\Lambda \,|\, \Lambda(\mathbb{I}) = \mathbb{I}\} \tag{1.3}$$

where $\mathbb{I}$ is the identity operator. We can see that the set of free operations in the resource theory of purity coincides with the set of unital channels on our Hilbert space.

Within the resource theory of purity, the state conversion follows from the classical theory of bistochastic maps [SKW+18].

Theorem 1.1. *the state ρ can be converted to the state σ using unital operations iff $\rho \succ \sigma$ (ρ majorizes σ) i.e. :*

$$\sum_{i=1}^{k} \lambda_i^-(\rho) \geq \sum_{i=1}^{k} \lambda_i^-(\sigma) \tag{1.4}$$

for all k where $\lambda_i^-(\rho)$ are the eigenvalues of ρ in the descending order.

The resource theory of purity admits a set of maximally resourceful states which coincides with the set of all pure states. Hence, starting from a pure state we can achieve any other state by unital channels [SKW+18]. Given n copy of the quantum state ρ, we are interested to know how many pure state one can achieve using only the unital channels. Considering the asymptotic limit of n, we have the following quantity of interest which is called the *purity of distillation* [SKW+18]:

$$P_d(\rho) = \sup\{R \,|\, \lim_{n\to\infty} \inf_{\Lambda\in O_F^p} \|\Lambda(\rho^{\otimes n}) - |\psi\rangle\langle\psi|^{\otimes nR}\|_1 = 0\}. \tag{1.5}$$

where $\|\sigma\|_1 \equiv \frac{1}{2}\,\mathrm{Tr}(|\sigma|)$ is the trace norm. It can be proved that in the asymptotic case, we have the following closed formula for the purity of distillation [SKW+18]:

$$P_d(\rho) = \log(d) - S(\rho) \tag{1.6}$$

where $S(\rho) \equiv -\,\mathrm{Tr}(\rho \ln \rho)$ is the von Neumann entropy function. Using the convexity property of $S(\rho)$ one can show that $P_d(\rho)$ is qualified for a purity monotone. Furthermore, due to the monotonicity of P_d and by the Eq. 1.6, we conclude the corollary below [SKW+18].

Corollary 1.1. *$S(\Lambda(\rho)) \geq S(\rho)$ iff Λ is a unital channel.*

1.3 Quantum Resource Theory of Coherence

The drive to explore the QRT of coherence originates from the concept of inevitable decoherence, implying that incoherent states remain unaffected in the presence of decoherence. Coherence resource theory is inherently dependent on the choice of basis, meaning that its framework varies according to the specific Hilbert space's basis chosen for analysis [SAP17]. This theory encompasses various approaches, each defined by different sets of free states and free operations. For example, coherence resource theories may differ based on whether they consider incoherent states or the states commuting with a Hamiltonian as free states[SAP17]. Similarly for free operations, various sets of maps have been considered such as incoherent operations (IO) and maximally incoherent operations (MIO) [BCP14a, WY16, Abe06, YMG+16, CG16, GS08, dVS16]. A shared characteristic among these groups is their incapacity to induce coherence from incoherent states. We refer to [SAP17] for an in-depth review of coherence resource theories. In the context of this thesis, we will focus on the following aspects of coherence resource theory.

Definition 1.5. The resource theory of coherence with respect to the basis $B = \{|i\rangle\}_{i=0}^{d-1}$ is the pair (F^c, O_F^c) such that:

$$F^c = \{\rho \,|\, \Delta(\rho) = \rho\} \tag{1.7}$$

and

$$O_F^c = \{\lambda_f \,|\, \lambda_f(\rho) \in F^c, \forall \rho \in F^c\} \tag{1.8}$$

where Δ is the dephasing channel in the basis B, defined as:

$$\Delta(\rho) \equiv \sum_{i=0}^{d-1} |i\rangle\langle i|\rho|i\rangle\langle i| \,. \tag{1.9}$$

In other words, in the resource theory defined above, the free states are those that are diagonal (i.e. incoherent) in our basis of interest and the free operations are the CPTP channels mapping each incoherent state to another incoherent one.

One can show that the QRT of coherence in the Hilbert space of dimension d, admits a set of maximally coherent states which can be written as [SAP17]:

$$|\psi_{\mathrm{MC}}\rangle_d = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} e^{\varphi_i} |i\rangle \tag{1.10}$$

where φ_i are some phases.

In the following, we discuss several coherence monotones that we will use in the next chapters.

1. l_1***-norm of Coherence*****:** is a coherence monotone defined as [SAP17]:

$$C_{l_1}(\rho) = \min_{\sigma \in F^c} |\rho - \sigma|_{l_1} \tag{1.11}$$

where $|\sigma|_{l_1} \equiv \sum_{i,j} |\sigma_{ij}|$ denotes the l_1-norm of σ and σ_{ij} are the components of σ in the basis $\{|k\rangle\}|_{k=1}^{d}$. It can be shown that C_{l_1} can be simplified as:

$$C_{l_1} = \sum_{i \neq j} |\rho_{ij}| \tag{1.12}$$

where ρ_{ij} are denoting the components of the density matrix ρ in the basis $\{|k\rangle\}|_{k=1}^{d}$.

2. Robustness of Coherence**:** Another coherence monotone is the robustness of coherence $R(\rho)$ which for a given state ρ quantifies the minimal mixing required to make ρ an incoherent state. Formally speaking we have [SAP17]:

$$R(\rho) = \min_{\sigma} \{s \geq 0 \,|\, \frac{\rho + s\sigma}{1+s} \in F^c\}. \tag{1.13}$$

It can be shown that the robustness of coherence coincides with the l_1-norm of coherence for pure states [SAP17].

3. Distillable Coherence: this quantity is defined by the task of distilling the maximally coherence single-qubit state $|\psi_{MC}^2\rangle$ as follow [SAP17]:

$$C_d(\rho) = \sup\{R \,|\, \lim_{n\to\infty} \inf_{\Lambda \in O_F^p} \|\Lambda(\rho^{\otimes n}) - |\psi_{MC}^2\rangle\langle\psi_{MC}^2|^{\otimes nR}\|_1 = 0\}. \tag{1.14}$$

Surprisingly, $C_d(\rho)$ can be expressed in a closed form as below [SAP17]:

$$C_d(\rho) = S(\Delta(\rho)) - S(\rho). \tag{1.15}$$

4. Distance Based Coherence: A distance-based approach to quantifying coherence, as described by [BCP14b], defines the coherence of a quantum state ρ as:

$$C_D(\rho) = \inf_{\sigma \in F^c} D(\rho, \sigma), \tag{1.16}$$

Here, the function D represents a chosen distance metric, and the infimum is calculated over the set of all incoherent states F^c. One can ascertain that the distance based quantifier of coherence is a convex function [BCP14b]. The practical meaning of this quantifier depends on the interpretation of the specific distance metric on the space of density matrices.

4. Relative Entropy of Coherence: The relative entropy of coherence for a state ρ is defined as [BCP14a]:

$$C_{\mathrm{r}}(\rho) = \min_{\sigma \in \mathcal{I}} S(\rho\|\sigma) = S(\Delta[\rho]) - S(\rho), \tag{1.17}$$

with the quantum relative entropy $S(\rho\|\sigma) = \mathrm{Tr}[\rho \log \rho] - \mathrm{Tr}[\rho \log \sigma]$. As we see, the relative entropy of coherence for the state ρ coincides with the distillable coherence of ρ.

1.4 Quantum Resource Theory of Entanglement

The resource theory of bipartite entanglement deals with scenarios where two spatially separated parties can only perform local quantum operations and communicate with each other classically [Sha19, CLM+14]. The operations permitted in this context are collectively referred to as local operations and classical communication (LOCC). Local operations involve quantum manipulations that each party can carry out independently on their subsystem [Sha19, CLM+14]. Classical communication allows the parties to exchange information through classical means, enabling them to coordinate their actions but not perform quantum operations remotely [Sha19, CLM+14].

A precise definition of LOCC operations, which we will adopt, is provided as follows [Sha19, CLM+14].

Definition 1.6. A one-round LOCC protocol, denoted as LOCC_1, is a quantum operation represented by an instrument $\mathcal{E}_x$, where the individual maps $\mathcal{E}_x$ are trace-non-increasing completely positive maps (CPMs) that remain local for all measurement outcomes x. In other words, each $\mathcal{E}_x$ decomposes into a tensor product $\bigotimes_j(\mathcal{E}_{xj})$, and there exists a site $j = K$ such that only at K the map $\mathcal{E}_x^K$ is not trace-preserving. This implies that the instrument can be executed by the party at site K applying the local instrument $\{\mathcal{E}_x^K\}$ and sharing the classical result x with all other parties, who then each perform (based on x) trace-preserving (deterministic) local quantum operations $\mathcal{T}_x^j$.

Note that LOCC_r are defined recursively as those operations that can be realized by following up an operation LOCC_{r-1} with a LOCC_1 operation. One can prove that the set of LOCC is a convex set [Sha19, CLM+14]. We refer to the thesis [CLM+14] for a detailed study of the set of LOCC operations. Another important element of the entanglement theory is the set of separable operations which is defined as follows [Sha19, CLM+14].

Definition 1.7. In quantum information theory, a quantum state ρ is called **separable** if it can be written as a mixture of product states. More precisely,

a quantum state ρ on a composite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ is said to be separable if there exists a probability distribution $\{p_i\}$ and a set of pure product states $\{|\psi_i\rangle_A \otimes |\varphi_i\rangle_B\}$ such that:

$$\rho = \sum_i p_i \left(|\psi_i\rangle_A\langle\psi_i| \otimes |\varphi_i\rangle_B\langle\varphi_i|\right), \tag{1.18}$$

We have the Horodecki criterion regarding the separability of multipartite quantum states [CLM+14].

Theorem 1.2. *Any separable state is positive under partial transposition. The pure state $|\psi\rangle$ is separable if and only if it is positive under partial transposition.*

Now we are prepared to introduce the resource theory of entanglement [Sha19, CLM+14].

Definition 1.8. The resource theory of multipartite entanglement with N number of parties is the pair (SEP, LOCC) where SEP denotes the convex set of separable states.

It has been proved that the resource theory of bipartite entanglement for the parties A and B with their Hilbert space's dimension d_A and d_B respectively, does admit a set of maximally entangled state as follow [CLM+14]:

$$|\psi_{\text{ME}}\rangle = \frac{1}{\sqrt{d}} \sum_{i}^{d} e^{i\varphi_i} |ii\rangle \tag{1.19}$$

where $d = \min\{d_A, d_B\}$ and $\{|i\rangle\}_{i=1}^{d}$ is a set of d orthogonal states. Moreover, regarding the state transformation in bipartite entanglement resource theory, we have the following theorem.

Theorem 1.3. *The bipartite pure state $|\psi\rangle_{AB}$ can be transformed to the state $|\varphi\rangle_{AB}$ by LOCC if and only if $|\psi\rangle_{AB}\langle\psi| \succ |\varphi\rangle_{AB}\langle\varphi|$.*

In the following, we also discuss some of the important and useful entanglement monotones.

1. Geometric Measure of Entanglement: For a pure n-partite state $\psi = |\psi\rangle\langle\psi|$, the multipartite geometric entanglement is defined as [Shi95, WG03, BL01, BNO02]

$$E_g(\psi) = 1 - \max_{\varphi \in \text{SEP}} |\langle\varphi|\psi\rangle|^2, \tag{1.20}$$

For mixed states, the geometric entanglement is defined as [WG03]

$$E_g(\rho) = \min_{\{p_i, |\psi_i\rangle\}} \sum_i p_i E_g(\psi_i), \tag{1.21}$$

where the minimum is taken over all pure state decompositions of ρ such that $\sum_i p_i \psi_i = \rho$. As has been shown in [SKB10], E_g can also be expressed as

$$E_g(\rho) = 1 - \max_{\sigma \in \text{SEP}} F(\rho, \sigma) \tag{1.22}$$

with the fidelity function $F(\rho, \sigma) = (\text{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})^2$. The geometric entanglement is zero for all separable states, and positive whenever the state is entangled. Moreover, E_g does not increase under LOCC [WG03].

2. Robustness of Entanglement: Another entanglement monotone is the robustness of entanglement $R_e(\rho)$ which for a given state ρ quantifies the minimal mixing required to make the state a separable state. Formally speaking we have [Sha19, CLM+14]:

$$R_e(\rho) = \min_{\sigma} \{s \geq 0 \,|\, \frac{\rho + s\sigma}{1+s} \in \text{SEP}\}. \tag{1.23}$$

3. Entanglement of Formation: Consider Alice and Bob possess m copies of the state ρ_{AB} in which $A(B)$ denotes the part of the system in the Alice's (Bob's) hand. Through a LOCC protocol, they aim to convert these m copies into n singlets $|\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$. This process is called entanglement distillation. Note that the process may result in some error, but the error should approach zero as $m \to \infty$. The distillable entanglement of ρ_{AB} is the highest possible value of n/m as m goes to infinity. The reverse process is called the entanglement dilution and the entanglement cost of ρ_{AB} is defined as the smallest ratio n/m when n approaches infinity. We have the following theorem in this regard [Sha19].

Theorem 1.4. *The distillable entanglement and the entanglement cost of a bipartite state $|\psi_{AB}\rangle$ are both equal to $S(\rho_A)$ in which $\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi|_{AB})$.*

Now, the entanglement of formation for the state ρ_{AB}, which is an entanglement monotone, is defined as

$$E_f(\rho_{AB}) = \min \sum_i p_i S(|\psi_i\rangle_{AB}) \tag{1.24}$$

and the minimum is taken over all decompositions $\{p_i, |\psi_i\rangle_{AB}\}$. This quantity is also convex [Sha19].

4. Distance Based Entanglement: A distance-based approach to quantifying entanglement, as described by [PV05], defines the entanglement of a quantum state ρ as:

$$C_D(\rho) = \inf_{\sigma \in \mathrm{SEP}} D(\rho, \sigma), \tag{1.25}$$

Here, the function D represents a chosen distance metric, and the infimum is calculated over the set of all separable states SEP. One can ascertain that the distance based quantifier of entanglement is a convex function. The practical meaning of this quantifier depends on the interpretation of the specific distance metric on the space of density matrices.

In the next section we will have a brief review of quantum speed limits for the state transformations in quantum systems.

1.5 Quantum Speed Limits

In the canonical quantization formalism of quantum mechanics, we attribute to the observables position x and momentum p, the corresponding operators $\hat{x}$ and $\hat{p}$ respectively, with the following commutation relation [DC17]:

$$[\hat{x}, \hat{p}] = i\hbar. \tag{1.26}$$

Due to this commutation relation, one can see that for any state in the position-momentum Hilbert space, we must have [DC17]:

$$\Delta x \Delta p \geq \hbar \tag{1.27}$$

which reflects the fact that the position and momentum of the corresponding particle cannot be measured simultaneously. Continuing this argument, one might conduct the following rough calculations:

$$\Delta t \approx \frac{\Delta x}{v} \tag{1.28}$$

and

$$\Delta E \approx \frac{\partial E}{\partial p} \Delta p = v \Delta p \tag{1.29}$$

in which v is the velocity of the particle. These calculations lead us to the following uncertainty relation [DC17]:

$$\Delta t \Delta E \geq \hbar. \tag{1.30}$$

However, the interpretation of the recent relation is not very clear. Mandelstam and Tamm realized that this relation expresses a bound on the speed of an evolution by Hamiltonian $\hat{H}$ with $\Delta \hat{H} = \Delta E$ [DC17]. They also, for the first time, proposed the notion of speed limit in a quantum evolution [DC17].

The standard approach for determining quantum speed limits is to consider a quantum state $|\psi\rangle$ that evolves unitarily as $U = e^{-iHt}$, transitioning into another state $|\varphi\rangle$. The goal is to find the minimal required time for the transition $|\psi\rangle \to |\varphi\rangle$, taking into account the energy scale of the Hamiltonian H. In recent years, more generalized concepts of quantum speed limit has been studied. The focus has expanded beyond unitary transitions between quantum states to include quantum speed limits for open system dynamics [dCEPH13, FSS19, TM21, TLM19, TSM+22]. Studies have also examined the speed limits for the evolution of observables in the Heisenberg picture [MP21].

In the following, we review some of the important fundamental bounds for the speed of evolution in quantum systems.

1. ***Mandelstam-Tamm Bound:*** The initial developments in this direction were focused on orthogonal states and are referred to as the Mandelstam-Tamm bound [MT45]. Considering a constrained variance of the the Hamiltonian $\hat{H}$ as $(\Delta E_\psi)^2 = \langle H^2 \rangle_\psi - \langle H \rangle_\psi^2$ with respect to the pure state $|\psi\rangle$, the time duration of transforming $|\psi\rangle$ to an orthogonal state is bounded below by:

$$T_\perp \geq \frac{\pi}{2\Delta E_\psi}\ , \tag{1.31}$$

However, this lower bound can be made arbitrarily small even if the mean energy $E_\psi = \langle H \rangle_\psi - E_0$ is bounded. To fix this issue Margolous and Levitin suggested another bound dependent on E_ψ [ML98].

2. *Margolous-Levitin Bound:* Considering a constrained mean energy as $E_\psi = \langle H\rangle_\psi - E_0$ with respect ot the pure state $|\psi\rangle$, Margolus and Levitin derived the following bound for the transformation of $|\psi\rangle$ to an orthogonal state [ML98],

$$T_\perp \geq \frac{\pi}{2E_\psi}\ , \tag{1.32}$$

with the mean energy $E_\psi = \langle H\rangle_\psi - E_0$, and E_0 is the ground state energy.

It is important to note that the speed limits given in equations (1.31) and (1.32) differ only in the choice of the energy scale. Comparing these bounds arises the apparent paradox that there exist two independent minimal time of orthogonalization of the pure state $|\psi\rangle$. The following theorem fixes this issue [LT09].

Theorem 1.5. *the unified bound*

$$T_\perp \geq \frac{\pi}{2min\{\Delta E_\psi, E_\psi\}} \tag{1.33}$$

is tight.

3. Generalized Bound: Generalized quantum speed limits have been presented for transitions between mixed states $\rho \to \sigma$ [LT09, PCC+16, CPBM18, SCMdC18] where

$$T(\rho \to \sigma) \geq \frac{\arccos F(\rho,\sigma)}{\min\{\Delta E_\rho, E_\rho\}.} \tag{1.34}$$

4. Geometric Bounds: The bounds we have mentioned so far, are regarding the unitary evolution of the system. However, there is a generalized geometric QSL that encompasses the minimal time of transformation given by any evolution [PCC+16]:

$$L^f(\rho_0,\rho_\tau) \leq l^f_\gamma(\rho_0,\rho_\tau) \tag{1.35}$$

where $L^f(\rho_0,\rho_\tau)$ is the geodesic distance between ρ_0 and ρ_τ with respect to the metric $\mathbf{g}^f$ and

$$l^f_\gamma(\rho_0,\rho_\tau) = \int_\gamma ds = \int_0^\tau dt \left(\frac{ds}{dt}\right) \tag{1.36}$$

is the length of the path γ from ρ_0 to ρ_τ with respect to the metric $\mathbf{g}^f$. If we parametrize the state $\rho(t)$ by the parameters $\{\lambda_\mu\}$, we then can generally write:

$$l^f_\gamma(\rho_0,\rho_\tau) = \int_0^\tau dt \sqrt{\sum_{\mu,\nu=1}^{r} g_{\mu\nu}\frac{d\lambda^\mu}{dt}\frac{d\lambda^\nu}{dt}} \tag{1.37}$$

However, the bound is in a very general form and it is expressed in terms of integral equations. Therefore it may be difficult to compute it for a given transformation.

In the next section, we present a brief introduction to the certification of quantum devices for the purpose of designing a class of quantum circuits in chapter 6, taking advantage of quantum resources to benchmark a specific class of noises.

1.6 Certifying and Benchmarking Quantum Devices

As quantum technologies advance, ensuring the reliability and performance of quantum devices becomes crucial. Certification involves verifying that these devices operate correctly and meet certain standards. This process is essential for both practical applications and the advancement of quantum research. Certifying quantum devices ensures that they function as intended and can be trusted in sensitive applications, such as quantum computing, quantum communication, and quantum sensing [EHW+20].

1.6.1 Certification and Benchmarking Methods

Quantum devices manipulate quantum states through operations that must be precisely controlled. Certification requires verifying that these states and operations adhere to theoretical expectations [EHW+20]. A quantum state is represented by a vector in a Hilbert space, and quantum operations are represented by unitary or non-unitary transformations on these vectors. Quantum tomography is a technique used to reconstruct the quantum state of a system based on measurement data. It provides a complete description of the quantum state but can be resource-intensive, especially for large systems. Fidelity measures quantify how close a quantum state or operation is to the desired state or operation. High fidelity indicates that the device performs accurately. For instance, the function $F(\rho, \sigma)$ (which has been defined previously after the equation 1.22) is a measure of fidelity between the two quantum states ρ and σ is. This measure is crucial for assessing the quality of quantum gates and circuits. We have two important ways of benchmarking the quantum devices.

1. Randomized Benchmarking: Randomized benchmarking is a method to assess the average error rate of quantum gates by applying sequences of random gate operations and comparing the output to the expected result. It provides a scalable way to evaluate device performance by averaging out specific errors and focusing on the overall error rate of the gates [KLR+08].

2. Cross-Entropy Benchmarking: Used primarily in the context of quantum computing, cross-entropy benchmarking involves running a quantum circuit on both a quantum device and a classical simulator, and then comparing the probability distributions of the outputs. This method is particularly useful for evaluating quantum processors designed to perform complex computations that are hard to simulate classically [BIS+18].

We also introduce another measure called the diamond norm [Wil11] which is useful in the context of quantum benchmark.

***Diamond Norm*:** Let $\mathcal{H}$ be a Hilbert space and $\Phi : L(\mathcal{H}) \to L(\mathcal{H})$ is a quantum channel on $L(\mathcal{H})$. The diamond norm of Φ, denoted by $\|\Phi\|_\diamond$ is defined as below:

$$\|\Phi\|_\diamond = \max_{\rho \in L(\mathcal{H} \otimes \mathcal{H})} \|\mathcal{I}_\mathcal{H} \otimes \Phi(\rho)\|_1 \tag{1.38}$$

where $\mathcal{I}_{\mathcal{H}}$ is the identity channel on $L(\mathcal{H})$.

The diamond norm can be used to quantify the performance for distinguishing two quantum channels in a single-shot scenario. If an agent is randomly given one of the two quantum channels Φ_1 and Φ_2 with the probabilities p and $1-p$ respectively and is allowed to apply it on only one state, then the maximal probability of success in discriminating the channels is given by [Wil11]

$$p_{success} = \frac{1}{2} + \frac{1}{2}\|p\varphi_1 + (1-p)\Phi_2\|_\diamond. \tag{1.39}$$

1.6.2 Challenges

As quantum systems grow, the complexity of certification increases exponentially. Techniques like quantum tomography become impractical for large systems due to the sheer number of measurements required. For example, the number of measurements needed for full-state tomography scales exponentially with the number of qubits, making it unfeasible for systems with more than a few qubits [CPF+10].

Furthermore, Quantum devices are susceptible to various types of noise and errors, which can complicate the certification process. Isolating and identifying these errors is a major challenge. Errors can arise from imperfect gate operations, decoherence, and environmental interactions, all of which need to be accounted for during certification [Pre18].

1.6.3 Recent Advances

Self-testing is an approach where the device's outputs are used to certify its performance, reducing the need for external references. This method leverages the concept of device independence, where the correctness of the device is inferred from the correlations in its outputs, assuming minimal assumptions about the internal workings of the device [MV21].

Machine learning algorithms are being explored to analyze measurement data and identify errors more efficiently. These techniques can potentially automate parts of the certification process. For example, neural networks can be trained to recognize patterns in the measurement data that indicate specific types of errors, thus streamlining the error identification process [DPM+23].

New protocols are being developed to certify large-scale quantum devices without exhaustive measurements. These protocols leverage statistical methods and innovative experimental designs. For instance, compressed sensing techniques can be used to reconstruct high-dimensional quantum states from a smaller number of measurements by exploiting the sparsity of the state in a certain basis [FGLE12].

While challenges remain, ongoing research and innovative approaches are paving the way for reliable and scalable certification methods. As quantum devices become more sophisticated, robust certification will be vital to harnessing the full potential of quantum technology. The development of efficient, scalable, and accurate certification techniques will play a pivotal role in the advancement and commercialization of quantum technologies.

Chapter 2

Quantum Speed Limits for Change of Basis and Resource Generation

In the pursuit of quantum advantages, such as faster computation, researchers have encountered a fundamental limitation imposed by nature. This limitation arises from the minimum time required for the unitary evolution of an initial quantum state to reach a final quantum state. The existence of this minimal time was first discussed in [MT45, ML98]. In a geometric perspective [JK10, Zwi12, PCC+16, CPBM18], the quantum speed limit is related to the shortest path length between the initial and final quantum states, quantified using an appropriate distance measure as discussed in the previous chapter. This approach offers valuable insights into the fundamental constraints of quantum processes and has been the focus of recent investigations [DC17]. The conventional approach to quantum speed limits involves considering a quantum state $|\psi\rangle$ that undergoes a unitary evolution $U = e^{-iHt}$ to transform into another state $|\varphi\rangle$. The objective is to find the optimal evolution time for the transition $|\psi\rangle \to |\varphi\rangle$ concerning the energy scale of the Hamiltonian H. For instance. in the early works by Mandelstam and Tamm [MT45] and Margolus and Levitin [ML98], the focus was on the speed limit for unitary transitions between two pure quantum states. However, in more recent years, researchers have developed more generalized versions of the speed limit. Indeed, the study of quantum speed limits has expanded beyond unitary transitions between quantum states. Researchers have explored quantum speed limits for open system dynamics [dCEPH13, FSS19, TM21, TLM19, TSM+22], as well as speed limits for the evolution of observables in the Heisenberg picture [MP21]. Additionally, there has been an investigation into speed limits for systems with a bounded energy spectrum [NAS22]. In a recent work [dC21], a theoretical approach has been introduced to measure quantum speed limits in an ultracold gas. Indeed, the exploration of

speed limits extends to various aspects of generating quantum resources. For instance, studies have been conducted to determine optimal rates for generating quantum entanglement [HHHH09b], quantum coherence [SAP17, MDP22], and quantum discord [MBC+12, Str15]. A recent study [BDLR21] utilized the concept of a speed limit to distinguish unitary channels by leveraging the characteristics of the diamond norm.

Earlier approaches to the concept of speed limits have primarily dealt with the transformation of one quantum state to another. In this chapter, we propose a new perspective by constructing a novel and well-defined notion of speed limit within the space of quantum state bases instead of focusing on the space of the states themselves. Additionally, we establish theorems and bounds for the minimal time needed to transform one basis into another.

2.1 Notion of Speed Limit for Change of Basis

In the previous studies, the focus was on the speed limit for transforming *one* quantum state into another. However, many quantum technological applications involve transformations of an entire collections of states. A prominent example is quantum computation, where operations like changing the basis are common, such as applying the Hadamard gate. The Hadamard gate transforms the computational qubit basis $|0\rangle, |1\rangle$ into the $|+\rangle, |-\rangle$ basis and $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. These collective transformations are crucial for performing various quantum algorithms and tasks. In this chapter, we aim to address the fundamental speed limits that apply to basis transformations. Specifically, we investigate bounds on the time required to transform an ordered set of quantum states to another ordered set of quantum states [NMB+24]. Our analysis considers the minimization of this transformation time over all possible Hamiltonians governing the dynamics of the quantum system [NMB+24]. Following the spirit of the Margolus-Levitin bound (1.32), our objective is to derive quantum speed limits in the form of [NMB+24]:

$$T(|\psi_j\rangle \to |\varphi_j\rangle) \geq \frac{g}{E}. \tag{2.1}$$

Here $\{|\psi_j\rangle\}, \{|\varphi_j\rangle\}$ are two ordered sets of orthonormal basis states, with $j = 1, ..., d$, where d is the dimension of the Hilbert space, and g can in general depend on the overlap of the two bases $\{|\psi_j\rangle\}$ and $\{|\varphi_j\rangle\}$ i.e. $|\langle\varphi_i|\psi_i\rangle|$. In Eq. (2.1), the term E refers to a notion of energy that is determined by the Hamiltonian. It should be noted that E must not depend on individual states (like E_ψ), as this would not make sense when discussing the simultaneous transition involving an entire set of basis states [NMB+24]. Additionally, it's essential to recognize that the concept of a speed limit, whether in classical or quantum context, becomes irrelevant without imposing energy constraints. With access to unlimited energy, state transformations could theoretically be completed in any arbitrarily small time span [NMB+24].

Since E represents some notion of energy, we expect it to satisfy the following properties [NMB+24]:

1. E is independent on the particular choice of basis $\{|\psi_j\rangle\}$.
2. E is additive for non-interactive Hamiltonians of the form $H^{AB} = H^A \otimes \mathbb{I}^B + \mathbb{I}^A \otimes H^B$:
$$E_{AB} = E_A + E_B, \tag{2.2}$$
where E_A and E_B represent the values of the function E associated with H^A and H^B, respectively.

Considering the introduced concepts, we present the following general theorem concerning the transformation of two arbitrary bases [NMB+24].

Theorem 2.1. *If we have a speed limit of the form*
$$T(|\psi_j\rangle \to |\varphi_j\rangle) \geq \frac{g}{E} \tag{2.3}$$
for two complete orthonormal bases $\{|\psi_j\rangle\}$ and $\{|\varphi_j\rangle\}$, we can immediately establish a speed limit for transforming $\{|\varphi_j\rangle\}$ to any basis that can be obtained from $\{|\varphi_j\rangle\}$ via the unitary transformation $V = \sum_j e^{i\alpha_j} |\psi_j\rangle\langle\psi_j|$ as follows:
$$T(|\psi_j\rangle \to V|\varphi_j\rangle) \geq \frac{g}{E} \tag{2.4}$$
where α_j are some phases. Additionally, the speed limit (2.4) is tight whenever Eq. (2.3) is tight.

Proof. [NMB+24] To prove this theorem, let's consider a Hamiltonian H such that
$$e^{-iHt}|\psi_j\rangle = |\varphi_j\rangle\,. \tag{2.5}$$
In this case, the Hamiltonian $H' = VHV^\dagger$ achieves the transformation $e^{-iH't}|\psi_j\rangle = e^{-i\alpha_j}V|\varphi_j\rangle$. Indeed, this can be seen by using the expression $e^{-iH't} = Ve^{-iHt}V^\dagger$ and substituting into $e^{-iHt}|\psi_j\rangle = |\varphi_j\rangle$ (we get $e^{-iH't}|\psi_j\rangle = Ve^{-iHt}V^\dagger|\psi_j\rangle = e^{-i\alpha_j}V|\varphi_j\rangle$). Since H and H' have the same value of the function E which is unitary invariant and g is only a function of the overlap $|\langle\psi_i|\varphi_i\rangle|$, the speed limit (2.3) for the transformation $e^{-iHt}|\psi_j\rangle = |\varphi_j\rangle$ implies the speed limit (2.4) for any unitary V that is diagonal in the $\{|\psi_j\rangle\}$ basis as $|\langle\psi_j|V\varphi_j\rangle| = |\langle\psi_j|\varphi_j\rangle|$. Moreover, the speed limit (2.4) is tight for all diagonal unitaries V whenever Eq. (2.3) is tight for the same reason. □

A natural selection for E that meets the requirements (i) and (ii) is [NMB+24]:
$$E = \frac{1}{d}\sum_j \langle\psi_j|H|\psi_j\rangle - E_0, \tag{2.6}$$
The choice of E is naturally analogous to the mean energy of Hamiltonian with respect to the evolving state, i.e. E_ψ, in the Margolus-Levitin bound (1.32), as both incorporate the concept of a mean, though they represent different quantities [NMB+24]. From

this point forward, we regard E as the quantity defined by Eq. 2.6 and refer to it as the "mean energy" of the corresponding Hamiltonian. It is also worth to note that the mean energy (2.6) is equivalent to $E = \mathrm{Tr}[H/d] - E_0$ [NMB+24].

In the following, we will soon explore interesting cases involving speed limits, focusing on the transformation of a basis to an unbiased one as well as the permutation of bases. In addition to examining the speed limits for basis change, we also investigate the maximum coherence that can be generated within a given time frame using a Hamiltonian with mean energy E. Specifically, we examine the maximum coherence that can be established within a certain time. These findings are of great significance in the resource theory of quantum coherence [BCP14b, WY16, SAP17]. They become even more relevant as recent studies suggest that quantum coherence is better suited than entanglement to characterize the efficiency of certain quantum algorithms [MEKP16b, ATE+22a, NKG+22].

2.2 Speed Limits for Pure States

Consider a Hamiltonian H of dimension d with eigenvalues E_i and corresponding eigenstates $|E_i\rangle$. Without compromising the generality, let us assume that the eigenvalues are arranged in the ascending order, such that $E_{\max} = E_{d-1}$ and $E_{\min} = E_0$.

Let's consider an initial state $|\psi\rangle$ that evolves over a time interval $0 \leq t \leq \pi/E_{\mathrm{gap}}$, where $E_{\mathrm{gap}} = E_{\max} - E_{\min}$ represents the energy gap of the Hamiltonian.

In the following our focus is to determine the smallest possible overlap between the initial state $|\psi\rangle$ and the state $|\psi_t\rangle$ evolved in time, given by $|\psi_t\rangle = e^{-iHt}|\psi\rangle$.

$$F_{\min} = \min_{|\psi\rangle} |\langle\psi|e^{-iHt}|\psi\rangle|, \tag{2.7}$$

minimized over all initial states $|\psi\rangle$ [NMB+24].

Theorem 2.2. *For a given Hamiltonian H and within the range of evolution time $0 \leq t \leq \pi/E_{\mathrm{gap}}$, we have the following:*

$$F_{\min} = |\langle\psi_{\min}|e^{-iHt}|\psi_{\min}\rangle| = \frac{1}{2}|e^{-iE_{\mathrm{gap}}t} + 1| \tag{2.8}$$

with $|\psi_{\min}\rangle = \frac{1}{\sqrt{2}}(|E_0\rangle + |E_{d-1}\rangle)$.

Proof. [NMB+24] By expressing the initial state in the eigenbasis of the Hamiltonian as $|\psi\rangle = \sum_j c_j |E_j\rangle$ with complex coefficients c_j, we can represent the overlap $|\langle\psi|e^{-iHt}|\psi\rangle|$ as:

$$|\langle\psi|e^{-iHt}|\psi\rangle| = |\sum_j |c_j|^2 e^{-iE_j t}|. \tag{2.9}$$

Considering that the coefficients c_j satisfy the normalization condition $\sum_j |c_j|^2 = 1$, we can express our figure of merit in the following manner:

$$F_{\min} = \min_{|\psi\rangle} |\langle\psi|e^{-iHt}|\psi\rangle| = \min_{\{p_j\}} |\sum_j p_j e^{-iE_j t}|. \tag{2.10}$$

Here, the minimum value on the right-hand side is determined by considering all potential probability distributions p_j. Emphasizing that $E_{\text{gap}}t \leq \pi$, it is straightforward to see that the minimum is realized by employing the following specific choice of p_j:

$$p_j = \begin{cases} \frac{1}{2} & \text{for } j = 0 \text{ and } j = d-1, \\ 0 & \text{for } 0 < j < d-1. \end{cases} \tag{2.11}$$

This implies that the optimal state $|\psi_{\min}\rangle$ for achieving the minimum overlap $|\langle\psi|e^{-iHt}|\psi\rangle|$ can be selected as:

$$|\psi_{\min}\rangle = \frac{1}{\sqrt{2}}(|E_0\rangle + |E_{d-1}\rangle), \tag{2.12}$$

as claimed. Lastly, we can straightforwardly confirm that:

$$|\langle\psi_{\min}|e^{-iHt}|\psi_{\min}\rangle| = \frac{1}{2}|e^{-iE_{\text{gap}}t} + 1| \tag{2.13}$$

which concludes the proof of the theorem. □

Interestingly, $F_{\min}$ is independent of the specific structure of the Hamiltonian, relying solely on the energy gap between its largest and smallest eigenvalues, E_{gap}.

Next, we will apply this result to establish a bound on the evolution time between pure states [NMB+24].

Theorem 2.3. *The evolution time required to transform a pure state $|\psi_0\rangle$ into another state $|\psi_1\rangle$ through unitary evolution $U = e^{-iHt}$ is constrained by the following bound:*

$$T(|\psi_0\rangle \to |\psi_1\rangle) \geq \frac{1}{E_{\text{gap}}} \arccos(2|\langle\psi_0|\psi_1\rangle|^2 - 1). \tag{2.14}$$

Proof. [NMB+24] If the states $|\psi_0\rangle$ and $|\psi_1\rangle$ satisfy the condition $|\psi_1\rangle = e^{-iHt}|\psi_0\rangle$ for $0 \leq t \leq \pi/E_{\text{gap}}$, then according to the theorem 2.2, we have:

$$|\langle\psi_0|\psi_1\rangle|^2 \geq \frac{1}{4}|e^{-iE_{\text{gap}}t} + 1|^2. \tag{2.15}$$

This inequality can be equivalently rewritten as:

$$t \geq \frac{1}{E_{\text{gap}}} \arccos(2|\langle\psi_0|\psi_1\rangle|^2 - 1). \tag{2.16}$$

Alternatively, if $|\psi_0\rangle$ and $|\psi_1\rangle$ satisfy $|\psi_1\rangle = e^{-iHt}|\psi_0\rangle$ with $t > \pi/E_{\text{gap}}$, the inequality in Eq. (2.14) is trivially fulfilled, as $\arccos(x) \leq \pi/2$ for all non-negative x. Thus, the theorem is proven. □

Considering that E_{gap} is bounded by dE, we readily deduce the subsequent theorem [NMB+24].

Theorem 2.4. *The lower bound for the time required to transform a pure state $|\psi_0\rangle$ into another state $|\psi_1\rangle$ through unitary evolution $U = e^{-iHt}$ can be expressed as follows:*

$$T(|\psi_0\rangle \to |\psi_1\rangle) \geq \frac{1}{dE} \arccos(2|\langle\psi_0|\psi_1\rangle|^2 - 1). \tag{2.17}$$

Furthermore, it is important to note that for any pair of pure states $|\psi_0\rangle$ and $|\psi_1\rangle$, there exists a Hamiltonian H that achieves the equality in the expression given by Eq. (2.17) [NMB+24]. To illustrate this point, it's worth noting that the inequality expressed in Eq.(2.17) is tight when $d = 2$, as also evidenced by Eq.(2.33) [NMB+24]. Now, consider a Hamiltonian denoted as $H = |\varphi\rangle\langle\varphi|$ which attains the lower bound for the case of $d = 2$. It's noteworthy that under this circumstance, the mean energy evaluates to $E = 1/2$ [NMB+24]. Consequently, this implies that the chosen Hamiltonian achieves the transformation $|\psi_0\rangle \to |\psi_1\rangle$ within a time interval of [NMB+24]

$$t = \arccos(2|\langle\psi_0|\psi_1\rangle|^2 - 1), \tag{2.18}$$

which is the shortest possible time for $E = 1/2$. In the case of $d > 2$, we can employ the same Hamiltonian $H = |\varphi\rangle\langle\varphi|$ to achieve the transformation within the same minimal time as specified in Eq.(2.18) [NMB+24]. For this scenario, the mean energy becomes $E = 1/d$, thereby fully saturating Eq.(2.17) [NMB+24].

2.3 Speed Limits for Unbiased Bases

In the following, we will derive the speed limits for transforming the computational basis $\{|n\rangle\}$ into an unbiased basis $\{|n_+\rangle\}$, where $|\langle n|n_+\rangle|^2 = 1/d$. Basis change refers to the simultaneous conversion of all vectors in the initial basis to their corresponding vectors in the target basis, as illustrated in Figure 2.1. For two arbitrary mutually unbiased bases, we establish the following lemma [NMB+24].

Lemma 2.1. *Suppose we have a unitary operator $U = e^{-iHt}$ that performs the transformation from the basis $\{|n\rangle\}$ to a maximally coherent basis $\{|n_+\rangle\}$ of dimension d. Then we have:*

$$-\sqrt{d} \leq \sum_i \cos(E_i t) \leq \sqrt{d}. \tag{2.19}$$

where E_i are the eigenvalues of the Hamiltonian.

Proof. [NMB+24] Any unitary operator that accomplishes the desired transformation must take the form:

$$U = \sum_{n=0}^{d-1} e^{i\varphi_n} |n_+\rangle\langle n| \tag{2.20}$$

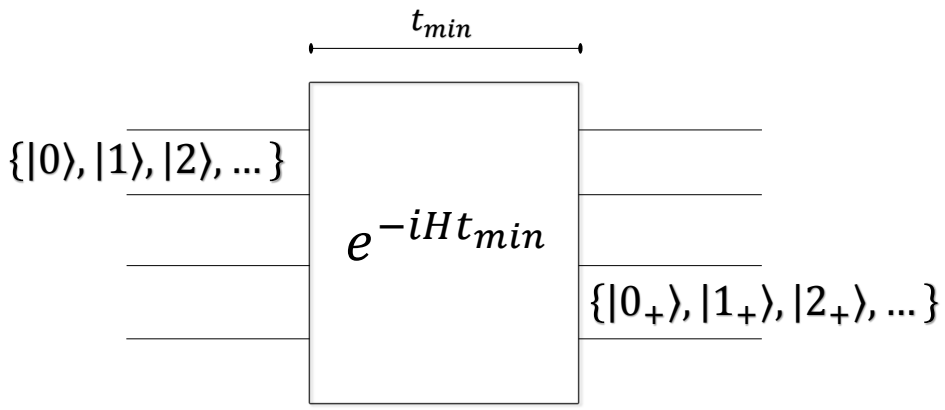

Figure 2.1: [NMB+24] Generating an unbiased basis $\{|n_+\rangle\}$ is achieved by evolving the computational basis $\{|n\rangle\}$ through a unitary process $e^{-iHt_{\min}}$. During this evolution, every vector in the computational basis transitions to its corresponding vector in the unbiased basis within a consistent time frame $t_{\min}$. It's important to highlight that $t_{\min}$ is the same for all transformations $|n\rangle \rightarrow |n_+\rangle$

with some phases φ_n. We also have:

$$\mathrm{Tr}[U + U^\dagger] = \sum_{n=0}^{d-1} (e^{i\varphi_n} \langle n|n_+\rangle + e^{-i\varphi_n} \langle n_+|n\rangle). \tag{2.21}$$

Noting that $\langle n|n_+\rangle = e^{i\gamma_n}/\sqrt{d}$ with some phases γ_n, we obtain the inequality:

$$-2\sqrt{d} \leq \mathrm{Tr}[U + U^\dagger] \leq 2\sqrt{d}. \tag{2.22}$$

Alternatively, if we recall that $U = e^{-iHt}$ with a corresponding Hamiltonian H, we can rewrite the inequality as follows:

$$\mathrm{Tr}[U + U^\dagger] = 2\sum_i \cos(E_i t), \tag{2.23}$$

where E_i are the eigenvalues of the Hamiltonian. In summary, in order for a unitary transformation $U = e^{-iHt}$ to achieve the transformation $|n\rangle \to |n_+\rangle$, it is necessary that the following condition holds:

$$-\sqrt{d} \leq \sum_i \cos(E_i t) \leq \sqrt{d}. \tag{2.24}$$

□

2.3.1 Systems with the Hilbert Space of Dimension 2, 3 and 4

The bound for single-qubit systems in the case of transformation to an unbiased basis is given by:

$$T_{\text{unbiased}} \geq \frac{\pi}{4E}, \tag{2.25}$$

and it is tight for any unbiased qubit basis [NMB+24]. We will see the proof in the following [NMB+24].

The general form of a single-qubit Hamiltonian is given by:

$$H = E_+ |E_+\rangle\langle E_+| + E_- |E_-\rangle\langle E_-|, \tag{2.26}$$

where the eigenvalues $E_\pm$ and eigenstates $|E_\pm\rangle$ can be parametrized as [NMB+24]:

$$E_\pm = G \pm E, , , , , , |E_\pm\rangle\langle E_\pm| = \frac{1}{2}(\mathbb{I} \pm \boldsymbol{n} \cdot \boldsymbol{\sigma}). \tag{2.27}$$

Here, G and $E \geq 0$ are real numbers, $\boldsymbol{n} = (n_x, n_y, n_z)$ is a normalized vector, and $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ contains the three Pauli operators. The Hamiltonian (2.26) can be rewritten in an equivalent form [NMB+24]:

$$H = (E\boldsymbol{n} \cdot \boldsymbol{\sigma} + G\mathbb{I}). \tag{2.28}$$

With these tools at our disposal, we can now introduce a bound for the evolution time between any two single-qubit states [NMB+24].

Theorem 2.5. *The transformation time for converting a single-qubit state ρ_0 into the state ρ_1 using unitary evolution $U = e^{-iHt}$ is subject to the following bound:*

$$T(\rho_0 \to \rho_1) \geq \frac{1}{2E} \arccos\left(\frac{\boldsymbol{r}_0 \cdot \boldsymbol{r}_1}{|\boldsymbol{r}_0||\boldsymbol{r}_1|}\right), \tag{2.29}$$

where $\boldsymbol{r}_i$ is the Bloch vector of the state ρ_i.

Proof. [NMB+24] It is important to observe that the unitary transformation $U(t) = e^{-iHt} = e^{-iGt}e^{-iEt\boldsymbol{n}\cdot\boldsymbol{\sigma}}$ can be understood as a rotation by an angle $2Et$ around the axis $\boldsymbol{n}$ of the Bloch sphere. The smallest value of Et is attained when selecting the rotation axis $\boldsymbol{n}$ to be perpendicular to both Bloch vectors $\boldsymbol{r}_0$ and $\boldsymbol{r}_1$.

$$\boldsymbol{n} = \frac{\boldsymbol{r}_0 \times \boldsymbol{r}_1}{|\boldsymbol{r}_0 \times \boldsymbol{r}_1|}, \tag{2.30}$$

$$Et = \frac{1}{2} \arccos\left(\frac{\boldsymbol{r}_0 \cdot \boldsymbol{r}_1}{|\boldsymbol{r}_0||\boldsymbol{r}_1|}\right). \tag{2.31}$$

This completes the proof. □

Noting that $\mathrm{Tr}[\rho_i\rho_j] = (1 + \boldsymbol{r}_i \cdot \boldsymbol{r}_j)/2$ we can reformulate Eq. (2.29) as follows [NMB+24]:

$$T(\rho_0 \to \rho_1) \geq \frac{1}{2E} \arccos\left(\frac{2\mathrm{Tr}[\rho_0\rho_1] - 1}{\sqrt{(2\mathrm{Tr}[\rho_0^2] - 1)(2\mathrm{Tr}[\rho_1^2] - 1)}}\right). \tag{2.32}$$

The proof of Theorem 2.5 demonstrates that this bound is optimal, meaning that for any pair of single qubit-states ρ_0 and ρ_1, there exists a Hamiltonian with a mean energy E that reaches the limit specified in Eq. (2.32) [NMB+24]. For pure qubit states, we obtain the tight bound:

$$T(|\psi_0\rangle \to |\psi_1\rangle) \geq \frac{1}{2E} \arccos(2|\langle\psi_0|\psi_1\rangle|^2 - 1). \tag{2.33}$$

For single-qubit systems, whenever a unitary transformation converts $|0\rangle$ into $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, it will also convert $|1\rangle$ into $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. Hence, for the transition from the computational basis $\{|0\rangle, |1\rangle\}$ to an unbiased qubit basis, we obtain [NMB+24]:

$$T_{\text{unbiased}} \geq \frac{\pi}{4E}, \tag{2.34}$$

as claimed.

As the dimension of the system increases ($d > 2$), one can intuitively expect that the evolution time needed to transform into an unbiased basis will also increase in comparison to the qubit setting. To further validate this intuition [NMB+24], let's consider a two-qubit system AB, and assume H^A and H^B are qubit Hamiltonians that

efficiently transform the computational basis $\{|0\rangle, |1\rangle\}$ to the unbiased basis $\{|+\rangle, |-\rangle\}$ within minimal times $\pi/(4E_A)$ and $\pi/(4E_B)$, respectively. If we set $E_A = E_B$, the Hamiltonian $H^{AB} = H^A \otimes \mathbb{I}^B + \mathbb{I}^A \otimes H^B$ achieves the transformation

$$\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\} \quad \rightarrow \quad \{|++\rangle, |+-\rangle, |-+\rangle, |--\rangle\} \tag{2.35}$$

within time $\pi/(4E_A) = \pi/(2E)$, where $E = 2E_A$ is the mean energy of the total Hamiltonian H^{AB}. Based on this argument [NMB+24], we observe that for a two-qubit system with $d = 4$, achieving an unbiased basis requires a time interval of $\pi/(2E)$, which is longer compared to the single-qubit setup.

However, as we proceed, we will discover that this intuition is not correct [NMB+24]. To explore this further, we will first shift our focus to qutrit systems [NMB+24].

Lemma 2.2. *A general unbiased qutrit basis can be achieved through a diagonal unitary transformation:*

$$V = \sum_j e^{i\alpha_j} |j\rangle\langle j| \tag{2.36}$$

from one of the following two bases (denoted by $\{|n_+\rangle\}$ and $\{|\tilde{n}_+\rangle\}$, respectively):

$$|0_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\frac{2}{3}\pi}|1\rangle + e^{i\frac{4}{3}\pi}|2\rangle), \tag{2.37a}$$

$$|1_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle), \tag{2.37b}$$

$$|2_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{-i\frac{2}{3}\pi}|1\rangle + e^{-i\frac{4}{3}\pi}|2\rangle), \tag{2.37c}$$

and

$$|\tilde{0}_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{-i\frac{2}{3}\pi}|1\rangle + e^{-i\frac{4}{3}\pi}|2\rangle), \tag{2.38a}$$

$$|\tilde{1}_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle), \tag{2.38b}$$

$$|\tilde{2}_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\frac{2}{3}\pi}|1\rangle + e^{i\frac{4}{3}\pi}|2\rangle). \tag{2.38c}$$

It is worth noting that these two class of basis states are odd permutations of each other.

Proof. [NMB+24] An arbitrary unbiased basis (with respect to the computational basis) for a qutrit can be expressed, up to an overall phase for each basis element, as follows:

$$|0_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\alpha_{0,1}}|1\rangle + e^{i\alpha_{0,2}}|2\rangle), \tag{2.39a}$$

$$|1_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\alpha_{1,1}}|1\rangle + e^{i\alpha_{1,2}}|2\rangle), \tag{2.39b}$$

$$|2_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\alpha_{2,1}}|1\rangle + e^{i\alpha_{2,2}}|2\rangle), \tag{2.39c}$$

where the phases $\alpha_{i,j}$ need to fulfill the condition

$$1 + e^{i(\alpha_{k,1}-\alpha_{l,1})} + e^{i(\alpha_{k,2}-\alpha_{l,2})} = 3\delta_{k,l}. \tag{2.40}$$

This condition specifies that the form of the basis can be either:

$$|0_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i(\alpha_{0,1}+\frac{2}{3}\pi)}|1\rangle + e^{i(\alpha_{0,2}+\frac{4}{3}\pi)}|2\rangle), \tag{2.41a}$$

$$|1_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\alpha_{0,1}}|1\rangle + e^{i\alpha_{0,2}}|2\rangle), \tag{2.41b}$$

$$|2_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i(\alpha_{0,1}-\frac{2}{3}\pi)}|1\rangle + e^{i(\alpha_{0,2}-\frac{4}{3}\pi)}|2\rangle), \tag{2.41c}$$

or

$$|0_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i(\alpha_{0,1}-\frac{2}{3}\pi)}|1\rangle + e^{i(\alpha_{0,2}-\frac{4}{3}\pi)}|2\rangle), \tag{2.42a}$$

$$|1_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\alpha_{0,1}}|1\rangle + e^{i\alpha_{0,2}}|2\rangle), \tag{2.42b}$$

$$|2_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i(\alpha_{0,1}+\frac{2}{3}\pi)}|1\rangle + e^{i(\alpha_{0,2}+\frac{4}{3}\pi)}|2\rangle). \tag{2.42c}$$

If we now introduce the unbiased bases

$$|0_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\frac{2}{3}\pi}|1\rangle + e^{i\frac{4}{3}\pi}|2\rangle), \tag{2.43a}$$

$$|1_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle), \tag{2.43b}$$

$$|2_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{-i\frac{2}{3}\pi}|1\rangle + e^{-i\frac{4}{3}\pi}|2\rangle), \tag{2.43c}$$

and

$$|\tilde{0}_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{-i\frac{2}{3}\pi}|1\rangle + e^{-i\frac{4}{3}\pi}|2\rangle), \tag{2.44a}$$

$$|\tilde{1}_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle), \tag{2.44b}$$

$$|\tilde{2}_+\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\frac{2}{3}\pi}|1\rangle + e^{i\frac{4}{3}\pi}|2\rangle), \tag{2.44c}$$

We observe that any basis of the form (2.41) or (2.42) can be obtained from the basis (2.43) or (2.44), respectively, by applying the diagonal unitary $V = |0\rangle\langle 0| + e^{i\alpha_{0,1}}|1\rangle\langle 1| + e^{i\alpha_{0,2}}|2\rangle\langle 2|$. □

As a consequence, speed limits for the transitions $|n\rangle \to |n_+\rangle$ and $|n\rangle \to |\tilde{n}+\rangle$ will also apply to general unbiased qutrit bases $|n\rangle \to V|n+\rangle$ and $|n\rangle \to V|\tilde{n}_+\rangle$ according to Theorem 2.1, where V is a diagonal unitary in the initial basis [NMB+24]. Now, we have the result stated below [NMB+24].

Theorem 2.6. *The time required to convert a qutrit basis to an unbiased basis is bounded below as follows:*

$$T_{\text{unbiased}} \geq \frac{2\pi}{9E}. \tag{2.45}$$

Proof. [NMB+24] Our aim is to establish that for any unitary transformation $U = e^{-iHt}$ resulting in the transition $|n\rangle \to |n_+\rangle$, the inequality holds:

$$Et \geq \frac{2}{9}\pi. \tag{2.46}$$

Given that E_i are in increasing order, we can observe that $E \geq \frac{E_2 - E_0}{3}$. Hence, to demonstrate Eq. (2.46), it is sufficient to prove:

$$(E_2 - E_0)t \geq \frac{2}{3}\pi. \tag{2.47}$$

To prove this, we will assume the opposite, i.e., that the transformation is possible with a unitary that violates Eq. (2.47). Violation of Eq. (2.47) implies that

$$(E_1 - E_0)t \leq \frac{\pi}{3} \quad \text{or} \quad (E_2 - E_1)t \leq \frac{\pi}{3}. \tag{2.48}$$

Let us consider the scenario where $(E_1 - E_0)t \leq \pi/3$. Without loss of generality, we can set $E_0 t = -\pi/6$, which leads to the following inequalities:

$$(|E_1 t)| \leq \frac{\pi}{6}, \quad E_2 t < \frac{\pi}{2}. \tag{2.49}$$

It follows that

$$\sum_i \cos(E_i t) > 2\cos(\frac{\pi}{6}), \tag{2.50}$$

which is a contradiction to Eq. (2.19). For the remaining case $(E_2 - E_1)t \leq \pi/3$, we can similarly choose $E_2 t = \pi/6$, which yields the following inequalities:

$$|E_1 t| \leq \frac{\pi}{6}, \quad E_0 t > -\frac{\pi}{2}. \tag{2.51}$$

Also in this case we obtain the inequality (2.50), in contradiction to Eq. (2.19). The proof of the bound (2.46) is now complete. Additionally, since the methods presented above apply to any qutrit basis that is unbiased with respect to the computational basis, this completes the proof of Theorem 2.6. □

After establishing a speed limit for basis change, a natural question arises: Is this bound tight? In other words, do there exist Hamiltonians H with mean energy E that saturate the bound (2.45) for any unbiased basis? To answer this question, we refer back to the definitions of the unbiased bases $\{|n_+\rangle\}$ and $\{|\tilde{n}_+\rangle\}$ in Eqs.(2.37) and(2.38). The following proposition addresses this matter [NMB+24].

Theorem 2.7. *The speed limit (2.45) is tight for the basis* $\{|n_+\rangle\}$*, but not tight for basis* $\{|\tilde{n}_+\rangle\}$.

Proof. [NMB+24] By the Theorem 2.6, the transition into the bases (2.43) and (2.44) can be described by the subsequent inequalities:

$$T(\{|n\rangle\} \to \{|n_+\rangle\}) \geq \frac{2\pi}{9E}, \tag{2.52a}$$

$$T(\{|n\rangle\} \to \{|\tilde{n}_+\rangle\}) \geq \frac{2\pi}{9E}. \tag{2.52b}$$

As can be checked by inspection, Eq. (2.52a) is saturated for the basis (2.43) by the Hamiltonian $H = |\alpha\rangle\langle\alpha|$ with

$$|\alpha\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{-i\frac{2}{3}\pi}\,|1\rangle + |2\rangle). \tag{2.53}$$

We shall demonstrate the strictness of the inequality (2.52b) for the basis (2.44). In other words, there does not exist an evolution e^{-iHt} that does the transformation $|n\rangle) \to |\tilde{n}_+\rangle$ within the time $t = 2\pi/(9E)$. Let us assume, through a proof by contradiction, that the bound is saturated for a certain unitary transformation denoted as $U = e^{-iHt}$:

$$|\tilde{n}_+\rangle = e^{-iHt}\,|n\rangle\,, \quad t = \frac{2\pi}{9E}. \tag{2.54}$$

Given that the energies E_i are arranged in the descending order, and based on the reasoning presented in the proof of Theorem 2.6, it necessarily follows that:

$$E_1 = E_0,\ (E_2 - E_0)t = \frac{2}{3}\pi. \tag{2.55}$$

Without loss of generality we can choose

$$E_0 t = E_1 t = -\frac{\pi}{6}, \quad E_2 t = \frac{\pi}{2}. \tag{2.56}$$

To summarize the presented arguments, there exists a unitary transformation $U = e^{-iHt}$ that satisfies Eq. (2.54) and possesses the following eigenvalues:

$$\lambda_0 = \lambda_1 = e^{i\frac{\pi}{6}}, \quad \lambda_2 = e^{-i\frac{\pi}{2}}, \tag{2.57}$$

Consequently, it follows that this unitary transformation fulfills the equation:

$$\mathrm{Tr}[U + U^\dagger] = 2\sqrt{3}. \tag{2.58}$$

Furthermore, the unitary transformation can also be expressed in the form:

$$U = \sum_{n=0}^{2} e^{i\varphi_n}\,|\tilde{n}_+\rangle\langle n|\,, \tag{2.59}$$

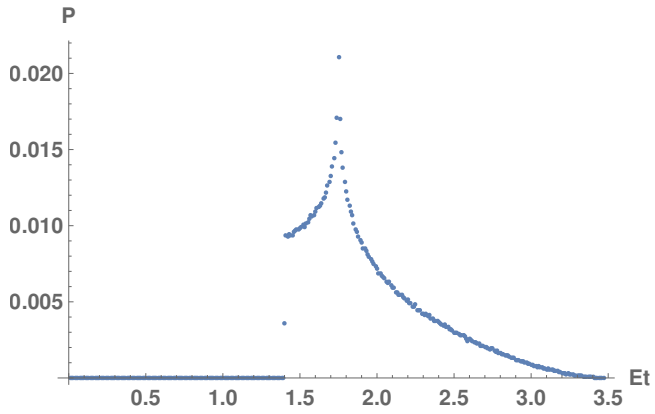

Figure 2.2: [NMB+24] We conduct a numerical assessment of Equation (2.64) by executing 10^6 samplings of unitaries in the form described by Equation (2.65), with random phases ranging from 0 to 2π. Subsequently, we compute Et utilizing Equation (2.6). The resulting plot illustrates the probability distribution as a function of Et. Our numerical analysis yields a lower bound of $Et \geq \frac{4}{9}\pi + \varepsilon$, where $\varepsilon \leq 10^{-5}$ in good agreement with the Equation (2.64).

where φ_n represents certain phases associated with each term. We find that

$$\begin{aligned} \mathrm{Tr}[U + U^\dagger] = \frac{2}{\sqrt{3}}(\cos\varphi_0 + \cos\varphi_1) \\ - \frac{1}{\sqrt{3}}\cos\varphi_2 + \sin\varphi_2. \end{aligned} \tag{2.60}$$

Together with Eq. (2.58) we obtain

$$\frac{2}{\sqrt{3}}(\cos\varphi_0 + \cos\varphi_1) - \frac{1}{\sqrt{3}}\cos\varphi_2 + \sin\varphi_2 = 2\sqrt{3}. \tag{2.61}$$

This equation has a unique solution within the range $0 \leq \varphi_i \leq 2\pi$, and it is given by:

$$\varphi_0 = \varphi_1 = 0, \quad \varphi_2 = \frac{2}{3}\pi. \tag{2.62}$$

This implies that the eigenvalues of U must be

$$\mu_0 = \mu_1 = e^{-i\frac{\pi}{6}}, \quad \mu_2 = e^{i\frac{\pi}{2}}, \tag{2.63}$$

which is a contradiction to Eq. (2.57). This completes the proof of the proposition. □

The findings presented indicate the existence of two separate classes of unbiased bases for qutrits [NMB+24]: Bases of the form $\{V|n_+\rangle\}$ can be acquired from the computational basis at time $T = 2\pi/9E$, whereas bases of the form $\{V|\tilde{n}_+\rangle\}$ require an

evolution time $T > 2\pi/9E$, where V is an arbitrary diagonal unitary in the basis $\{|n\rangle\}$. For the second class $\{V|\tilde{n}+\rangle\}$, there is numerical evidence suggesting that a tight speed limit is given by [NMB+24]:

$$T(\{|n\rangle\} \to \{|\tilde{n}+\rangle\}) \geq \frac{4\pi}{9E}. \tag{2.64}$$

To gain insight into this, consider that any unitary transformation enabling the transition $|n\rangle \to |\tilde{n}_+\rangle$ must exhibit the following structure:

$$U = \sum_{n=0}^{2} e^{i\varphi_n} |\tilde{n}_+\rangle\langle n| \tag{2.65}$$

with some phases φ_n. Consider $\lambda_j = e^{-i\alpha_j}$ as the eigenvalues of U, where the phases α_j are arranged in ascending order, satisfying the condition $-\pi \leq \alpha_j \leq \pi$. Given a specific collection of these phases α_j, there exists a corresponding Hamiltonian that realizes the unitary operator $U = e^{-iHt}$ with [NMB+24]:

$$E_j t = \alpha_j \quad \text{or} \quad E_j t = \alpha_j + 2\pi, \tag{2.66}$$

where E_j are the eigenvalues of H. Subsequently, the mean energy of the Hamiltonian obtained through numerical analysis satisfies the following [NMB+24]:

$$Et = \frac{1}{3}\sum_j E_j t - E_0 t. \tag{2.67}$$

Utilizing these findings, we are able to assess the validity of Eq.(2.64) by conducting numerical experiments. This involves randomly sampling phases within the range of $0 \leq \varphi_n \leq 2\pi$ and then calculating Et using Eq.(2.67) [NMB+24]. By selecting $E_j t$ as specified in Eq. (2.66), we ensure that the numerical Hamiltonians obtained through this procedure encompass those with the lowest possible value of Et [NMB+24]. In Fig. 2.2, the numerical probabilities are depicted, representing the likelihood of attaining specific values of Et over 10^6 samples. Notably, the minimum nonzero probability of Et is observed to be approximately 1.4 [NMB+24]. Based on the numerical outcomes, a lower limit for Et is suggested as follows [NMB+24]:

$$Et \geq \frac{4}{9}\pi + \varepsilon, \tag{2.68}$$

Here, ε is constrained by numerical analysis, yielding an upper bound of $\varepsilon \leq 10^{-5}$, which is notably consistent with the expression in Eq.(2.64). A Hamiltonian that saturates the bound specified in Eq.(2.64) is described by $\tilde{H} = -|\tilde{\alpha}\rangle\langle\tilde{\alpha}|$, where [NMB+24]:

$$|\tilde{\alpha}\rangle = \frac{1}{\sqrt{3}}(|0\rangle + e^{i\frac{2}{3}\pi}|1\rangle + |2\rangle). \tag{2.69}$$

A straightforward juxtaposition of Theorem 2.6 with the analogous qubit bound (2.25) reveals that achieving an unbiased qutrit basis demands a shorter duration, in comparison to realizing an unbiased qubit basis for the identical mean energy E [NMB+24].

In the subsequent discussion, we will delineate the primary distinctions between the qubit and qutrit scenarios.

When a single-qubit unitary transformation $U = e^{-iHt}$ is considered optimal for rotating the basis $\{|0\rangle, |1\rangle\}$ into an unbiased basis, it follows that the unitary $U^2 = e^{-2iHt}$ rearranges the constituent elements of the basis $\{|0\rangle, |1\rangle\}$ [NMB+24]. However, the situation differs when considering qutrits. Notably, it's worth recognizing that an optimal Hamiltonian for effecting the qutrit transition $|n_+\rangle = e^{-iHt}|n\rangle$ takes the form of $H = |\alpha\rangle\langle\alpha|$. Regarding this Hamiltonian, it is possible to calculate the fidelity between the initial state $|0\rangle$ and the state that evolves in time as $e^{-iHt}|0\rangle$ [NMB+24]:

$$|\langle 0|e^{-iHt}|0\rangle|^2 = \frac{1}{9}[5 + 4\cos(t)]. \tag{2.70}$$

It's important to highlight that the right-hand side of Eq. (2.70) is always nonzero. This implies that the evolution never causes the interchange of $|0\rangle$ with any other basis element, and analogous arguments can be applied to the states $|1\rangle$ and $|2\rangle$ [NMB+24].

In addition, when the basis $\{|0\rangle, |1\rangle\}$ are permuted by the single-qubit unitary U, an interesting observation is that the operation $\sqrt{U}$ invariably causes a rotation of the $\{|0\rangle, |1\rangle\}$ basis into an unbiased one [NMB+24]. However, this scenario changes in the qutrit context, as can be easily ascertained. A counterexample can be observed through the permutation $U = \sum_{n=0}^{2} |(n+1) \bmod 3\rangle\langle n|$. We further obtain

$$\sqrt{U} = \frac{1}{3}\begin{pmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{pmatrix}, \tag{2.71}$$

As a result, $\sqrt{U}|n\rangle$ fails to represent a maximally coherent state for any $0 \leq n \leq 2$. Furthermore, a direct examination reveals that $U^{1/3}$ similarly does not yield the transformation of any $|n\rangle$ state into a maximally coherent state [NMB+24].

Up to this point, our analysis has encompassed systems of dimensions 2 and 3. Moving ahead, we will take a further stride by determining the minimum evolution time required to establish an unbiased basis for two-qubit systems [NMB+24].

Theorem 2.8. *The lower bound on the time required to establish an unbiased basis for a two-qubit system is given by:*

$$T_{\text{unbiased}} \geq \frac{\pi}{4E}. \tag{2.72}$$

There exists a two-qubit Hamiltonian achieving this bound.

Notably, this constraint remains identical to the one observed in single-qubit systems, as indicated by Eq.(2.25). The specific Hamiltonian that reaches the limit outlined in Eq.(2.72) is formulated as follows [NMB+24]:

$$H = -\sigma_x \otimes \sigma_z + \sigma_y \otimes \sigma_y - \sigma_z \otimes \sigma_x. \tag{2.73}$$

The eigenvalues associated with this Hamiltonian are 3, −1, −1, −1, consequently yielding a mean energy of $E = 1$ for H.

Proof. [NMB+24] Let's establish the theorem by contradiction, supposing the existence of a unitary operator $U = e^{-iHt}$ that transforms the set $|n\rangle$ into a maximally coherent basis, such that:

$$Et < \frac{\pi}{4}. \tag{2.74}$$

We can make the assumption, without loss of generality, that $E_0 = 0$, leading to the conclusion that $E = (E_1 + E_2 + E_3)/4$. Let's proceed by defining $\alpha_i = E_i t$. It's worth noting that $\pi > \alpha_i \geq 0$. With reference to Eq. (2.74), we can deduce that $\alpha_3 < \pi - \alpha_1 - \alpha_2$, leading to the following implication:

$$\cos(\alpha_3) > \cos(\pi - \alpha_1 - \alpha_2) = -\cos(\alpha_1 + \alpha_2). \tag{2.75}$$

It follows that

$$\cos(\alpha_1) + \cos(\alpha_2) + \cos(\alpha_3) > \cos(\alpha_1) + \cos(\alpha_2) - \cos(\alpha_1 + \alpha_2). \tag{2.76}$$

Let's now delve into a more detailed examination of the right-hand side of Eq. (2.76), introducing the following definition:

$$f(\boldsymbol{\alpha}) = \cos(\alpha_1) + \cos(\alpha_2) - \cos(\alpha_1 + \alpha_2). \tag{2.77}$$

Particularly, our aim is to demonstrate the validity of $f(\boldsymbol{\alpha}) \geq 1$ in cases where:

$$\alpha_i \geq 0, \tag{2.78a}$$

$$\alpha_1 + \alpha_2 \leq \pi. \tag{2.78b}$$

To achieve this, we calculate the partial derivatives of f with respect to α_i:

$$\frac{\partial f}{\partial \alpha_1} = \sin(\alpha_1 + \alpha_2) - \sin(\alpha_1), \tag{2.79}$$

$$\frac{\partial f}{\partial \alpha_2} = \sin(\alpha_1 + \alpha_2) - \sin(\alpha_2). \tag{2.80}$$

To identify local extrema of f, we equate $\partial f/\partial \alpha_i$ to zero, resulting in $\sin(\alpha_1) = \sin(\alpha_2)$. This equation implies that either $\alpha_1 = \alpha_2$, or $\alpha_1 = \pi - \alpha_2$. When considering the condition $\alpha_1 = \alpha_2$, we derive $\sin(2\alpha_2) = \sin(\alpha_2)$, leading to the following solutions:

$$\alpha_1 = \alpha_2 = 0, \tag{2.81a}$$

$$\alpha_1 = \alpha_2 = \frac{\pi}{3}. \tag{2.81b}$$

Conversely, when examining the condition $\alpha_1 = \pi - \alpha_2$ along with $\partial f/\partial \alpha_i = 0$, we arrive at $\sin(\alpha_1) = \sin(\alpha_2) = 0$, yielding the following solutions:

$$\alpha_1 = 0, \alpha_2 = \pi, \tag{2.82a}$$

$$\alpha_1 = \pi, \alpha_2 = 0. \tag{2.82b}$$

To establish the inequality $f(\boldsymbol{\alpha}) \geq 1$, we assess $f(\boldsymbol{\alpha})$ at the extrema given by (2.81) and (2.82), as well as at the boundary of the region delineated by Eqs. (2.78). Upon substituting the solutions (2.81), we deduce $f(\boldsymbol{\alpha}) = 1$ and $f(\boldsymbol{\alpha}) = 3/2$, correspondingly. Furthermore, when we plug in the solutions (2.82), we find that $f(\boldsymbol{\alpha}) = 1$.

The task that remains is to establish that $f(\boldsymbol{\alpha}) \geq 1$ at the boundary of the region defined by Eqs. (2.78). When considering a particular value of $\alpha_1 \in [0, \pi]$, this boundary is realized either when $\alpha_2 = 0$ or $\alpha_2 = \pi - \alpha_1$.

Upon closer examination, it becomes evident that $f(\boldsymbol{\alpha}) = 1$ in both of these cases. In summation, this analysis conclusively demonstrates that the inequality $f(\boldsymbol{\alpha}) \geq 1$ remains valid throughout the region defined by (2.78).

Summing up the presented reasoning, we can deduce that Eq.(2.74) implies that there exists a unitary transformation $U = e^{-iHt}$ that accomplishes the transformation $|n\rangle \rightarrow |n_+\rangle$ while satisfying $\sum_i \cos(E_i t) > 2$. This directly contradicts the assertion made in Eq.(2.19). As a result, this concludes the proof of the lower bound (2.72). □

It's worth noting that even though the Hamiltonian H accomplishes the transformation with the least possible time, its interactive nature might introduce practical challenges and costs for its realization [NMB+24].

Taking $t = \pi/4$, we proceed to define the unitary operator $U = e^{-itH}$. The exertion of this unitary on the computational basis of two qubits unfolds as follows:

$$U(|0\rangle|0\rangle) = e^{i\pi/4}|+\rangle|+\rangle, \tag{2.83a}$$

$$U(|0\rangle|1\rangle) = e^{i\pi/4}|-\rangle|+\rangle, \tag{2.83b}$$

$$U(|1\rangle|0\rangle) = e^{i\pi/4}|+\rangle|-\rangle, \tag{2.83c}$$

$$U(|1\rangle|1\rangle) = e^{i\pi/4}|-\rangle|-\rangle. \tag{2.83d}$$

The Hamiltonian specified in Eq. (2.73) effectively facilitates the transformation of a two-qubit basis into an unbiased over a duration of $\pi/(4E)$.

The findings presented thus far underscore a noteworthy observation: the optimal time required for a transformation onto an unbiased basis remains the same for both single-qubit and two-qubit systems, standing at $\pi/(4E)$ in both instances. However, in the case of a qutrit system, this optimal time is further reduced to $2\pi/(9E)$ [NMB+24].

In the following sections, we investigate the speed limits pro regarding transfomation of a basis in the Hilbert space of arbitrary dimensions to an unbiased one.

2.3.2 Arbitrary Number of Qubits

We are about to broaden our analysis to encompass many-qubit systems. In doing so, we will show a universal threshold applicable to n-qubit systems, enabling the estab-

lishment of an unbiased basis within a finite duration [NMB+24].

Theorem 2.9. *In the systems comprised of n qubits, the minimum time required to establish an unbiased basis is constrained from above by:*

$$T_{\text{unbiased}} \leq \frac{\pi}{2E}. \tag{2.84}$$

Proof. [NMB+24] Consider the Hamiltonian H_n operating on n qubits:

$$H_n = V^{\otimes n}, \tag{2.85}$$

where V denotes the Hadamard gate. Notably, the mean energy associated with H_n is established at $E = 1$. Next, we introduce the unitary operator $U_n(t) = e^{-iH_n t}$. Utilizing the property $H_n^2 = \mathbb{I}$, we deduce that:

$$U_n(t) = \cos(t)\mathbb{I} - i\sin(t)H_n. \tag{2.86}$$

For $t = \pi/2$ we obtain

$$U_n(\pi/2) = -iV^{\otimes n}. \tag{2.87}$$

As a result of this unitary transformation, the computational basis of n qubits is converted into an unbiased basis, thereby concluding the proof. □

Theorem 2.9 provides evidence that an unbiased basis for n qubits can be established in a time frame of $\pi/(2E)$. This was concretely demonstrated by presenting a Hamiltonian that introduces interactions among all the qubits. In scenarios lacking interactions, that is, when each qubit evolves independently, the optimal time for evolution is determined as $n\pi/(4E)$ [NMB+24].

2.3.3 Systems with the Hilbert Space of Dimension 5 and 6

Upon contrasting the outcomes regarding the minimal transformation time T_{unbiased} required to transition a basis into an unbiased configuration within Hilbert spaces of dimensions $d = 2$, $d = 3$, and $d = 4$, one might be inclined to perceive a potential pattern. This pattern could suggest that for the subsequent prime number in the sequence, namely $d = 5$, the minimal transformation time might experience a reduction. However, we are now poised to demonstrate that this intuitive pattern is in fact erroneous. In the case of a dimension $d = 5$, we must have:

$$T_{\text{unbiased}} \geq \frac{\pi}{4E} \tag{2.88}$$

where $\frac{\pi}{4}$ serves as the minimal time T_{unbiased} for $d = 2$ and $d = 4$ [NMB+24]. To substantiate this, let's begin by supposing the existence of a Hamiltonian H characterized

by the eigenenergies $\{E_i\}_{i=1}^5$ and the mean energy E, such that the unitary e^{-iHT} can achieve the basis transformation in the time interval:

$$T \leq \frac{\pi}{4E}. \tag{2.89}$$

Then we must have $\sum_{i=1}^5 E_i T \equiv \sum_{i=1}^5 \alpha_i \leq \pi + \pi/4$. For the sake of simplicity, let's consider without loss of generality that the minimum eigenenergy of H is $E_{\min} = 0$. By minimizing the function

$$f(\boldsymbol{\alpha}) = \sum_{j=1}^{5} \cos \alpha_i, \tag{2.90}$$

while considering the constraint $0 \leq \sum_{i=1}^5 \alpha_i \leq \pi + \pi/4$, we determine that $\min f = 1 + 4\cos\frac{5\pi}{16} \approx 3.22$. This outcome contradicts the previously established result. This is because, as per Eq. 2.19, the upper limit of the function $f(\boldsymbol{\alpha})$ for a Hamiltonian that transforms a basis into an unbiased one within a Hilbert space of dimension 5 is $\sqrt{5} \approx 2.2$, which is smaller than the calculated value of 3.22.

Moving forward, we are poised to establish a lower limit for the speed constraint within a Hilbert space of dimension $d = 6$. Our objective is to demonstrate that the minimal time required to transform the basis $\{|i\rangle\}_{i=0}^5$ into an unbiased one through a Hamiltonian with a constant mean energy E is constrained from below by:

$$T \geq \frac{1}{3E} \arccos\left(-\frac{\sqrt{6}-4}{2}\right). \tag{2.91}$$

In order to establish this lower bound, let's assume the existence of a Hamiltonian for which

$$T < \frac{1}{3E} \arccos\left(-\frac{\sqrt{6}-4}{2}\right), \tag{2.92}$$

which implies that: $\sum_{i=0}^5 E_i T < 2\arccos\left(-\frac{\sqrt{6}-4}{2}\right)$. We define $E_i T = \alpha_i$ and without loss of generality, let's consider the minimum eigenenergy of the Hamiltonian as $E_{\min} = E_0 = 0$. According to Eq. (2.19), it is necessary to satisfy $-\sqrt{6} \leq \sum_j \cos\alpha_j \leq \sqrt{6}$. In the forthcoming analysis, we will demonstrate that the function $f(\boldsymbol{\alpha}) = \sum_j \cos\alpha_j$ surpasses the value of $\sqrt{d}$ within the region:

$$R = \{\sum_{i=0}^{5} \alpha_i < 2\arccos\left(-\frac{\sqrt{6}-4}{2}\right) \wedge \alpha_i > 0 \ \forall i\}. \tag{2.93}$$

Consequently, it would be evident that T cannot be reduced beyond $\frac{1}{3E}\arccos\left(-\frac{\sqrt{6}-4}{2}\right)$. We proceed to minimize the function $f(\boldsymbol{\alpha}) = \sum_{i=0}^5 \cos\alpha_i$ within the closure of R. Initially, our objective is to identify all critical points situated within this region. Upon differentiating the function $f(\boldsymbol{\alpha})$ and setting the resulting derivatives to zero, the critical points emerge as $\alpha_i = K_i\pi$, where $K_i \geq 0$ and K_i are integers. Because $\cos(K_i\pi) = \pm 1$, the minimum value of the function is attained among these critical points when we have

the maximum number of -1 (within the region R, the presence of only one -1 is permissible). Hence, the smallest value among these critical points is 4. Subsequently, we proceed to identify the minimum value along the boundaries $\sum_{i=0}^{5}\alpha_i = 2\arccos(-\frac{\sqrt{6}-4}{2})$. Let's make an assumption (without loss of generality) that we are considering the part of these boundaries where x number of the α_i values are zero. It's important to observe that the range of values for x is bounded by $0 \leq x \leq 3$, as exceeding this range would result in the function $f(\boldsymbol{\alpha})$ exceeding $\sqrt{6}$. This outcome aligns with the proof stipulated by Eq. (2.19). By employing the Lagrange multiplier method, we derive the subsequent system of equations:

$$\sin\alpha_i = k, \ \forall k, \tag{2.94}$$

where k represents the Lagrange multiplier. Solving these equations reveals that α_i must adhere to the either of following structure:

$$\alpha_i = \begin{cases} \lambda + 2K_i\pi \\ \pi - \lambda + 2K'i\pi, \end{cases} \tag{2.95}$$

where $0 \leq \lambda \leq \frac{\pi}{2}$ and K_i and K_i' are non-negative integers (they must be non-negative since α_i are non-negative). Additionally, we can assume (without loss of generality) that N instances of α_i correspond to the second form in Eq. (2.95). Guided by the boundary restriction within the enclosed region R, we arrive at:

$$(6-x-N)\lambda + N(\pi-\lambda) + 2(\sum_i K_i + \sum_j K_j') = 2\arccos -\frac{\sqrt{6}-4}{2}. \tag{2.96}$$

Solving this equation for λ we obtain:

$$\lambda = \frac{2\arccos(-\frac{\sqrt{6}-4}{2}) - (2K+N)\pi}{6-x-2N}. \tag{2.97}$$

where $K = \sum_i K_i + \sum_j K_j'$. Eq. (2.97) implies that $N < \frac{6-x}{2}$ otherwise $\lambda > \pi/2$ which is a contradiction (to the initial assumption that $0 \leq \lambda \leq \frac{\pi}{2}$). For critical points situated on the boundary, the function $f(\boldsymbol{\alpha})$ has the form $(6-x-2N)\cos(\frac{2\arccos(-\frac{\sqrt{6}-4}{2})-(2K+N)\pi}{6-x-2N})$. Given that $N < \frac{6-x}{2}$ and $0 \leq \lambda \leq \frac{\pi}{2}$, its minimum value is achieved for any x and N when $K = 0$. Consequently, the minimum of the function on the boundary must follow the pattern $(6-x-2N)\cos(\frac{2\arccos(-\frac{\sqrt{6}-4}{2})-N\pi}{6-x-2N})$, which consistently exceeds or equals $\sqrt{6}$ for all $1 \leq x \leq 3$ and $N < \frac{6-x}{2}$. As a result, the function $f(\boldsymbol{\alpha})$ attains a minimum value over the region R that surpasses $\sqrt{6}$, contradicting the condition in Eq. (2.19). Hence, the proof is concluded.

2.3.4 Hilbert Spaces with Arbitrary Dimension

In the subsequent discussion, we outline a comprehensive lower bound for the time needed to establish an unbiased basis within a system with the Hilbert space of dimension d [NMB+24].

Theorem 2.10. *The minimal time required to establish an unbiased basis in a system of dimension d is constrained from below by:*

$$T_{\text{unbiased}} > \frac{\pi(d-1)}{4Ed}. \tag{2.98}$$

It becomes apparent that as the dimension of the Hilbert space becomes larger, the lower bound converges towards $\pi/4E$ [NMB+24].

Proof. [NMB+24] We introduce the notation $T_{\text{low}} = \frac{d-1}{dE}\frac{\pi}{4}$, where $d \geq 2$. Let's consider that there exists a Hamiltonian and a unitary e^{-iHT} completing the task of basis transformation in a way that:

$$ET \leq \frac{d-1}{d}\frac{\pi}{4}. \tag{2.99}$$

Without loss of the generality, let's assume $E_0 = 0$ and ensure $E_j \geq 0$ for all j. Furthermore, we introduce the variables $\alpha_j = E_j T$, leading to:

$$\sum_j \alpha_j \leq (d-1)\frac{\pi}{4}. \tag{2.100}$$

By Eq. (2.19) we must have $-\sqrt{d} \leq \sum_j \cos\alpha_j \leq \sqrt{d}$. By minimizing the function $f(\boldsymbol{\alpha}) = \sum_j \cos\alpha_j$, we can demonstrate that $f(\boldsymbol{\alpha})$ always surpasses $\sqrt{d}$ within the defined region (2.100). Consequently, it follows that T cannot fall below the threshold of T_{low}. Let's begin by identifying the critical points of the function $f(\boldsymbol{\alpha})$ within the interior of the region (excluding the boundary). To achieve this, we differentiate the function with respect to α_i, resulting in the following set of equations:

$$\sin\alpha_i = 0 \quad \forall i \tag{2.101}$$

This shows that $\alpha_i = K_i\pi$ and $K_i \geq 0$. For these particular values, $\cos\alpha_i$ can only be either 1 or -1. As a result, the function's minimum (among these critical points) is attained when there is a maximal count of -1 values. Considering the constraint (2.100), this means that $\lfloor\frac{d-1}{4}\rfloor$ of the α_i variables should be equal to π, while the rest are set to zero. Consequently, the minimum value is given by $d - 2\lfloor\frac{d-1}{4}\rfloor$ if $\frac{d-1}{4}$ is not an integer. If $\frac{d-1}{4}$ is an integer, the critical point will lie on the boundary of the region, which we will delve into next.

Next, we proceed to identify the critical points on the boundary of the region (2.100), which satisfies $\sum_j \alpha_j = (d-1)\frac{\pi}{4}$ and $\alpha_j \geq 0$. In a more general context, let's assume that we are situated on a segment of the boundary where x of the $\{\alpha_i\}_{i=1}^{d-1}$ values are set to zero. Utilizing the Lagrange multipliers method, we arrive at the following set of equations:

$$\sin\alpha_i = k \quad \forall i, \tag{2.102}$$

where k is the Lagrange multiplier. The equations (2.102) reveal that either $\alpha_i = \lambda + 2K_i\pi$ or $\alpha_i = \pi - \lambda + 2K_i'\pi$, where $0 \leq \lambda \leq \frac{\pi}{2}$ and K_i, K_i' are non-negative integers (due

to the requirement that $\alpha_i \geq 0$). Considering the boundary segment where x of the α_i values are set to zero, and assuming that N of them follow the form $\alpha_i = \pi - \lambda + 2K_i^{'}\pi$, we are compelled to satisfy the equation (derived from $\sum_j \alpha_j = (d-1)\frac{\pi}{4}$):

$$(d - x - 2N)\lambda + (N + \sum_j K_j' + \sum_l K_l)\pi = (d-1)\frac{\pi}{4}. \tag{2.103}$$

We introduce the notation $K \equiv \sum_j K_j' + \sum_l K_l$. Expressing λ in terms of K and N, we derive:

$$\lambda = \frac{(d-1)/2 - 2(N+K)}{d - x - 2N}\frac{\pi}{2} \tag{2.104}$$

In this case, the function takes on the form $x + (d - x - 2N)\cos\lambda$. When we consider the range $N < \frac{d-x}{2}$, the function attains its minimum when λ is maximized, and this occurs when $K = 0$, applicable for any values of x and N. If we shift our focus to the interval $N > \frac{d-x}{2}$, we have:

$$\lambda = \frac{N - (d-1)/4}{N - (d-x)/2}\frac{\pi}{2} + \frac{K}{N - (d-x)/2}\frac{\pi}{2}. \tag{2.105}$$

Given that $x \leq \sqrt{d}$ (because if this were not the case, then $\sum_j \cos\alpha_j$ would exceed $\sqrt{d}$ and the proof would be complete), it can be demonstrated that the first term in Eq. (2.105) is greater than $\pi/2$. This is evident from the fact that the coefficient $\frac{N-(d-1)/4}{N-(d-x)/2}$ exceeds 1:

$$N - \frac{d-1}{4} \geq N - \frac{d-\sqrt{d}}{2} \geq N - \frac{d-x}{2} \rightarrow (\sqrt{d}-1)^2 \geq 0. \tag{2.106}$$

Furthermore, the second term in Eq. (2.105) is positive. Consequently, in the case where $N > \frac{d-x}{2}$, the value of λ would exceed $\frac{\pi}{2}$, which contradicts the initial assumption $\lambda \leq \frac{\pi}{2}$. Furthermore, if we consider the scenario where $N = \frac{d-x}{2}$, Eq. (2.105) leads to the equation $d + 1 + 2K = 2x$, which presents a contradiction as x is a positive integer and also satisfies the condition $x \leq \sqrt{d}$. Hence, we conclude that $N < \frac{d-x}{2}$, and in this case, the expression for λ takes the following form at the minimum of the function:

$$\lambda = \frac{(d-1)/4 - N}{(d-x)/2 - N}\frac{\pi}{2}. \tag{2.107}$$

Furthermore, based on Eq. (2.106), we have $\frac{d-1}{4} \leq \frac{d-x}{2}$, which implies $0 \leq N \leq \frac{d-1}{4}$ due to the non-negativity of λ. Thus, we need to determine which value of N within this domain minimizes the function. Our goal is to find the minimum of the following function as we vary the parameter N within the specified domain:

$$x + (d - x - 2N)\cos\left(\frac{(d-1)/4 - N}{(d-x)/2 - N}\frac{\pi}{2}\right). \tag{2.108}$$

Upon differentiating this function with respect to N, it becomes evident that the derivative is monotonically decreasing within the valid domain of N, hence the value $N_0 =$

$\lfloor (d-1)/4 \rfloor$ achieves the minimum of $f(\boldsymbol{\alpha})$ with the value of $x + (d - x - 2\lfloor (d-1)/4 \rfloor) \cos\left(\frac{(d-1)/4 - \lfloor (d-1)/4 \rfloor}{(d-x)/2 - \lfloor (d-1)/4 \rfloor} \frac{\pi}{2}\right)$ which is always greater than $\sqrt{d}$ for $d \geq 2$:

$$\begin{aligned}\sqrt{d} &\leq x\left(1 - \cos\left(\frac{(d-1)/4 - N_0}{(d-x)/2 - N_0} \frac{\pi}{2}\right)\right) \\ &\quad + \frac{d+1}{2} \cos\left(\frac{1}{\frac{d+1}{2} - \frac{\sqrt{d}}{2}} \frac{\pi}{2}\right) \\ &\leq x + (d - x - 2\lfloor (d-1)/4 \rfloor) \cos\left(\frac{(d-1)/4 - \lfloor (d-1)/4 \rfloor}{(d-x)/2 - \lfloor (d-1)/4 \rfloor} \frac{\pi}{2}\right) \end{aligned} \tag{2.109}$$

In deriving the second inequality, we have employed the conditions $1 \leq x \leq \sqrt{d}$ and $\frac{d-1}{4} - 1 \leq \lfloor \frac{d-1}{4} \rfloor \leq \frac{d-1}{4}$. Consequently, it is evident that the minimum of the function $f(\boldsymbol{\alpha})$ within the specified region (2.100) is consistently greater than $\sqrt{d}$. This discrepancy contradicts the condition posed by Eq. (2.19), leading to the completion of the proof.

□

By comparing the derived lower bound with the bound established in Theorem 2.9, a notable pattern becomes apparent. As the number of qubits, denoted by n, approaches infinity, it is clear that the minimal time T required for achieving an unbiased basis for this n-qubit system conforms to the inequality $\pi/4E \leq T \leq \pi/2E$ [NMB+24].

Note that from the proof, one can find the values of α_i and determine the Hamiltonian which is responsible for the task of transformation. However, the obtained Hamiltonian may not saturate the bound for the speed limits. One interesting direction (to find the saturable speed limits) would be checking how close the basis change by this Hamiltonian would be to the desired transformation in different dimensions.

In the next section, we study the minimal time of transformation of a basis to a permuted one.

2.4 Speed Limits for Basis Permutation

It would be instrumental to analyze the above findings concerning the speed limits associated with permuting the basis $\{|n\rangle\}$ as:

$$U|n\rangle = |(n+1) \bmod d\rangle \tag{2.110}$$

for all $0 \leq n \leq d-1$. First, we state and establish a lemma regarding the eigenvalues and eigenstates of the permutation unitary [NMB+24].

Lemma 2.3. *The eigenvalues of the permutation unitary are $\lambda_n = e^{i\frac{2\pi}{d}n}$, where n is an integer satisfying $0 \leq n \leq d-1$.*

Proof. [NMB+24] First, we will calculate the eigenvalues of the permutation unitary:

$$U|n\rangle = |(n+1) \bmod d\rangle . \tag{2.111}$$

Let $|\psi\rangle = \sum_n a_n |n\rangle$ be an eigenstate of U, i.e.,

$$U|\psi\rangle = e^{i\alpha}|\psi\rangle . \tag{2.112}$$

From Eq. (2.111) we obtain

$$a_n = e^{i\alpha} a_{(n+1) \bmod d}, \tag{2.113}$$

This indicates that the magnitudes of all coefficients a_n must be equal, specifically satisfying the condition $|a_n|^2 = 1/d$. As a result, any eigenstate $|\psi\rangle$ can be expressed in the form:

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} e^{i\varphi_j} |j\rangle . \tag{2.114}$$

This leads to the conclusion that U cannot have degenerate eigenvalues. To establish this, let's assume, for the sake of contradiction, that there exist two distinct eigenstates $|\psi_1\rangle$ and $|\psi_2\rangle$ sharing the same eigenvalue. Consequently, any linear combination of $|\psi_1\rangle$ and $|\psi_2\rangle$ would also be an eigenstate of U. Furthermore, through such a linear combination, we could generate an eigenstate that does not adhere to the form (2.114), leading to the sought contradiction.

Next, let us observe that any permutation unitary must satisfy the condition:

$$U^d = \mathbb{I}. \tag{2.115}$$

Furthermore, considering the non-degeneracy of U, we deduce that the eigenvalues of U take the specific form $\lambda_n = e^{i\frac{2\pi}{d}n}$, where n is an integer satisfying $0 \le n \le d-1$. □

Now, we are prepared to state the following theorem regarding the minimal time required for basis permutation [NMB+24].

Theorem 2.11. *The lower bound for the time required to permute a basis is given by:*

$$T_{\text{perm}} \ge \frac{\pi(d-1)}{dE}. \tag{2.116}$$

Proof. [NMB+24] As elaborated in the previous proof, the eigenvalues of the permutation unitary (2.111) are of the form:

$$\lambda_j = e^{-i\frac{2\pi j}{d}}, \tag{2.117}$$

where integer j is in the range $0 \le j \le d-1$. This leads to the conclusion that for any permutation unitary $U = e^{-iHt}$, the following must hold:

$$t \sum_j E_j = \sum_j \frac{2\pi j}{d} = \pi(d-1). \tag{2.118}$$

The proposition is fully established by noting that E can be expressed as the sum of E_j over d, leading to $E = \sum_j E_j/d$.

□

Remarkably, for a given Hamiltonian H, there exist only two possibilities [NMB+24]: Either the unitary operator $U = e^{-iHt}$ results in permutation with $t = \frac{\pi(d-1)}{dE}$, or the Hamiltonian does not lead to a basis permutation. It's important to emphasize that our analysis is specifically applicable to permutations of the form given in Eq. (2.110).

We conclude this chapter by providing speed limits for coherence generation utilizing our methods.

2.5 Speed of Evolution for Coherence Generation

Now, we will establish speed limits for generating quantum coherence through unitary evolution. Our focus is on determining the maximum attainable coherence $C_{\max}$ that can be generated from a given state ρ within a predetermined time interval t:

$$C_{\max}(\rho, t) = \max_H C(e^{-iHt}\rho e^{iHt}), \tag{2.119}$$

where the optimization is carried out over all possible Hamiltonians H with an mean energy $E = \mathrm{Tr}[H]/d - E_0$ [NMB+24]. We utilize the ℓ_1-norm of coherence as a measure of quantum coherence [BCP14b, SAP17].

$$C(\rho) = \sum_{i \neq j} |\rho_{ij}|, \tag{2.120}$$

This can be efficiently estimated in experimental setups through the use of collective measurements, as demonstrated in works such as [YHT+20, WSR+21].

Let's begin our exploration in the single-qubit context. In this scenario, the unitary operator $U(t) = e^{-iHt}$ can be understood as a rotation by an angle of $2Et$ around the axis $\boldsymbol{n}$ of the Bloch sphere. For single-qubit states, the measure of coherence C corresponds to the Euclidean distance from the incoherent axis. In this context, $C_{\max}(\rho, t)$ represents the maximum distance from the incoherent axis, achieved by optimizing over all possible rotations with a fixed angle of $2Et$. The optimal axis of rotation, denoted as $\boldsymbol{n}$, is orthogonal to both the Bloch vector $\boldsymbol{r}$ and the incoherent axis. The expression for the maximum coherence $C_{\max}$ is as follows [NMB+24]:

$$C_{\max}(\rho, t) = |\boldsymbol{r}| \cos(\arcsin([\frac{|r_z|}{|\boldsymbol{r}|}] - 2Et)). \tag{2.121}$$

It's important to note that $C_{\max}$ cannot exceed the magnitude of the Bloch vector $|\boldsymbol{r}|$, and this maximum value is achieved for the time:

$$T_{\mathrm{mc}} = \frac{1}{2E} \arcsin \frac{|r_z|}{|\boldsymbol{r}|}, \tag{2.122}$$

at which the final state lies in the maximally coherent plane [NMB+24]. When the initial state is pure, it can be represented using the parameterization:

$$|\psi\rangle = \cos(\theta/2)\,|0\rangle + e^{i\varphi}\sin(\theta/2)\,|1\rangle\ , \tag{2.123}$$

and the maximum attainable coherence within a time t is given by [NMB+24]:

$$C_{\max}(|\psi\rangle, t) = \cos\left(\arcsin([\cos\theta] - 2Et)\right). \tag{2.124}$$

Next, we extend our analysis to systems of arbitrary dimension $d \geq 2$ and determine the minimum time required to transform a pure state $|\psi\rangle$ into a maximally coherent state of the form:

$$|\psi_{\mathrm{MC}}\rangle_d = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1} e^{i\varphi_j}\,|j\rangle \tag{2.125}$$

with phases φ_j. The subsequent theorem provides a limit for the evolution time $T(|\psi\rangle \to |\psi_{\mathrm{MC}}\rangle_d)$ [NMB+24].

Theorem 2.12. *The duration required for transforming a state $|\psi\rangle$ into a maximally coherent state $|\psi_{MC}\rangle_d$ through unitary evolution $U = e^{-iHt}$ is bounded by:*

$$T(|\psi\rangle \to |\psi_{MC}\rangle_d) \geq \frac{1}{dE}\arccos[\frac{2}{d}\sum_j |\langle\psi|j\rangle|^2 - 1]. \tag{2.126}$$

Proof. [NMB+24] The time required to evolve a pure state into a maximally coherent state is bounded according to the result established in Lemma 2.4 as:

$$T(|\psi\rangle \to |\psi_{\mathrm{MC}}\rangle_d) \geq \frac{1}{dE}\arccos(2|\langle\psi|\psi_{\mathrm{MC}}\rangle_d|^2 - 1). \tag{2.127}$$

Therefore, to derive a universally applicable bound, we must evaluate the maximum overlap $|\langle\psi|\psi_{\mathrm{MC}}\rangle_d|$ across all states expressed in the form:

$$|\psi_{\mathrm{MC}}\rangle_d = \frac{1}{\sqrt{d}}\sum_{j=0}^{d-1} e^{i\varphi_j}\,|j\rangle \tag{2.128}$$

Expressing the initial state $|\psi\rangle$ in the basis of incoherent states $|i\rangle$, we have:

$$|\psi\rangle = \sum_{j=0}^{d-1} c_j e^{i\alpha_j}\,|j\rangle \tag{2.129}$$

with $c_j \geq 0$, it becomes evident that the overlap $|\langle\psi|\psi_{\mathrm{MC}}\rangle_d|^2$ is maximized when $\varphi_j = \alpha_j$. Thus we arrive at the expression:

$$\max_{|\psi_{\mathrm{MC}}\rangle_d} |\langle\psi|\psi_{\mathrm{MC}}\rangle_d|^2 = \frac{1}{d}(\sum_j |\langle\psi|j\rangle|)^2. \tag{2.130}$$

Alternatively, this result can be derived following the approach in [RFWA18, RLS18], where it is recognized that $\max_{|\psi_{\mathrm{MC}}\rangle_d} |\langle\psi|\psi_{\mathrm{MC}}\rangle_d|^2$ corresponds to the maximum fidelity between the state $\Lambda[|\psi\rangle\langle\psi|]$ and the specific maximally coherent state $|\psi_{\mathrm{MC}}\rangle_d = \sum_j |j\rangle / \sqrt{d}$, achieved by optimizing over all incoherent operations Λ. Using Eq. (2.130) in Eq. (2.127) completes the proof.

□

2.6 Discussion

We have explored the maximum speeds for basis transformations using unitary evolutions and have determined optimal time limits for these transformations in various relevant scenarios.

For dimensions $d \leq 4$, we identified the optimal time needed to convert the computational basis into an unbiased, or maximally coherent, basis. Notably, the shortest evolution times are the same for $d = 2$ and $d = 4$ when Hamiltonians with the same average energy E are considered. In the case of $d = 3$, achieving the speed limit requires a specific basis ordering that is unbiased relative to the computational basis. Additionally, we proved that an n-qubit Hadamard gate can be performed in a time of $\pi/2E$, indicating that in multi-qubit systems, a maximally coherent basis can be reached in a time that does not depend on the number of qubits. These results imply that interactive Hamiltonians can considerably shorten the evolution time compared to the time required if each qubit were evolved separately. We also demonstrated that for $d \to \infty$, the minimum time to establish an unbiased basis is at least $\pi/4E$. Additionally, we explored the speed limits associated with basis permutation.

We have also explored the limits on the speed for generating a given level of quantum coherence and for transforming a pure state into a maximally coherent state. We expect that these methods can be extended to find the minimal transformation times for various bases and other quantum resources, including quantum entanglement and imaginary states [HG18, WKR+21a, WKR+21b].

In the next chapter, we continue exploring the speed of coherence generation during an evolution governing by a Hamiltonian.

Chapter 3

Coherence Generation with Hamiltonian

As discussed in the introduction, central to any quantum resource theory is the fundamental investigation into the possibility of state transformations. Typically, the scenario involves having access to n copies of a specific initial state ρ, with the goal of converting them into m copies of a target state of interest σ. In this process, it's envisaged that there's an error margin that decreases as the number of initial state copies, n, grows. The success degree of this transformation is measured by the maximum achievable ratio of m/n, indicating the transformation rate. In the context of the coherence resource theory, the precise rate at which one quantum state can be transformed into another has been established while we focus on the set of maximally incoherent operations (MIO) [Abe06]. The highest attainable transformation rate within this framework is delineated by the formula [WY16]:

$$R(\rho \to \sigma) = \frac{C_{\mathrm{r}}(\rho)}{C_{\mathrm{r}}(\sigma)} = \frac{C_{\mathrm{d}}(\rho)}{C_{\mathrm{d}}(\sigma)}, \tag{3.1}$$

where $C_{\mathrm{r}}(.)$ and $C_{\mathrm{d}}(.)$ denote the relative entropy of coherence and distillable coherence respectively as defined in 1.17 and 1.14.

Given the importance of quantum coherence in both quantum information science and quantum technology [SAP17, WSR+21], it is imperative to delve into and grasp the most effective approaches for its creation. An approach for generating quantum coherence involves utilizing a fixed quantum channel denoted by Λ. This method can effectively induce coherence from an initially incoherent state, provided that Λ does not belong to the set of maximally incoherent operations (MIO). The endeavor to discover and enhance optimal approaches for generating coherence via stationary quantum channels has attracted significant interest and has been thoroughly investigated in numerous studies [BCP14a, MK15, DFW+18, TRS22, GDEP16].

This chapter centers on investigating the most effective techniques for inducing coherence through dynamic processes, with a particular emphasis on unitary evolutions $U_t = e^{-itH}$. By utilizing the relative entropy of coherence as a figure of merit, we analyze the highest possible rate of change of C_{r} within this framework, optimized over all initial states ρ and all Hamiltonians H with a finite Hilbert-Schmidt norm [SNS24]. This quantity holds a distinct operational significance as described in Eq. (3.1), representing the maximum achievable rate of coherence generation through Hamiltonians with a bounded Hilbert-Schmidt norm. We characterize the most advantageous initial states and Hamiltonians applicable to systems of any dimension d. Particularly, concerning qubit systems, we offer the optimal input state for any specified Hamiltonian [SNS24].

3.1 Coherence Generating Capacity of Hamiltonians

We establish the concept of the *coherence generating capacity* of a Hamiltonian H, representing the maximal increase rate in coherence achievable by a unitary evolution $U_t = e^{-itH}$ at time $t = 0$ [SNS24]:

$$C_{\mathrm{gen}}(H) = \max_{\rho} \left. \frac{C_{\mathrm{r}}(e^{-iHt}\rho e^{iHt})}{dt} \right|_{t=0}. \tag{3.2}$$

Similar concepts have been previously explored for the entanglement theory, particularly concerning the generation of entanglement through non-local Hamiltonians [DVC+01, Bra07].

We have the proposition 3.1 for an alternative expression for the coherence-generating capacity. Let's first state and prove the following lemma [SNS24].

Lemma 3.1. *For any Hermitian matrix A and any positive matrix B, we have:*

$$\mathrm{Tr}\left[\Delta(A)\log_2\Delta(B)\right] = \mathrm{Tr}\left[A\log_2\Delta(B)\right] \tag{3.3}$$

Proof. [SNS24] Consider A as a Hermitian matrix and B as a positive matrix. Then, we have:

$$\begin{aligned}
\mathrm{Tr}\left[A\log_2\Delta(B)\right] &= \mathrm{Tr}\left[A\log_2\left(\sum_i |i\rangle\langle i|\,B\,|i\rangle\langle i|\right)\right] = \sum_i \mathrm{Tr}\left[A\log_2\left(|i\rangle\langle i|\,B\,|i\rangle\langle i|\right)\right] \\
&= \sum_{i,j} \mathrm{Tr}\left[A\,|j\rangle\langle j|\log_2\left(|i\rangle\langle i|\,B\,|i\rangle\langle i|\right)|j\rangle\langle j|\right] \\
&= \sum_{i,j} \mathrm{Tr}\left[|j\rangle\langle j|\,A\,|j\rangle\langle j|\log_2\left(|i\rangle\langle i|\,B\,|i\rangle\langle i|\right)\right] \\
&= \mathrm{Tr}\left[\sum_j |j\rangle\langle j|\,A\,|j\rangle\langle j|\log_2\left(\sum_i |i\rangle\langle i|\,B\,|i\rangle\langle i|\right)\right] \\
&= \mathrm{Tr}\left[\Delta(A)\log_2\Delta(B)\right].
\end{aligned} \tag{3.4}$$

□

Therefore we have the following proposition [SNS24].

Proposition 3.1. *For any Hamiltonian H the following expression holds:*

$$C_{\text{gen}}(H) = \max_{\rho} i\text{Tr}(H[\rho, \log_2 \Delta(\rho)]). \tag{3.5}$$

Proof. [SNS24] Note that $\dot{x}$ represents the time derivative of the quantity x through the chapter. let define the state $\rho_t = e^{-iHt}\rho e^{iHt}$ and its time derivative as $\dot{\rho}_t = d\rho_t/dt$. We will demonstrate that the time derivative of the von Neumann entropy can be expressed as [DKSW18, MDP22]

$$\frac{d}{dt}S(\rho_t) = -\text{Tr}\left[\dot{\rho}_t \log_2 \rho_t\right]. \tag{3.6}$$

To begin, we express the density matrix ρ_t in terms of its eigenbasis: $\rho_t = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, where the eigenvalues λ_i and the eigenstates $|\psi_i\rangle$ depend on time. Consequently, we have:

$$\begin{aligned}\frac{d}{dt}(\rho_t \log_2 \rho_t) &= \sum_i \frac{d}{dt}\left(\lambda_i \frac{\ln \lambda_i}{\ln 2}\right)|\psi_i\rangle\langle\psi_i| + \sum_i (\lambda_i \log_2 \lambda_i)\frac{d}{dt}|\psi_i\rangle\langle\psi_i| \\ &= \sum_i \left(\dot{\lambda}_i \frac{\ln \lambda_i}{\ln 2} + \frac{\dot{\lambda}_i}{\ln 2}\right)|\psi_i\rangle\langle\psi_i| + \sum_i (\lambda_i \log_2 \lambda_i)\left[|\dot{\psi_i}\rangle\langle\psi_i| + |\psi_i\rangle\langle\dot{\psi_i}|\right].\end{aligned} \tag{3.7}$$

This implies that the von Neumann entropy's derivative can be represented as

$$\frac{d}{dt}S(\rho_t) = -\text{Tr}\left[\frac{d}{dt}(\rho_t \log_2 \rho_t)\right] = -\sum_i \left(\dot{\lambda}_i \frac{\ln \lambda_i}{\ln 2} + \frac{\dot{\lambda}_i}{\ln 2}\right), \tag{3.8}$$

regarding the fact that

$$\text{Tr}\left[|\dot{\psi_i}\rangle\langle\psi_i| + |\psi_i\rangle\langle\dot{\psi_i}|\right] = \text{Tr}\left[\frac{d}{dt}(|\psi_i\rangle\langle\psi_i|)\right] = \frac{d}{dt}\text{Tr}\left[|\psi_i\rangle\langle\psi_i|\right] = 0. \tag{3.9}$$

Noting that $\sum_i \dot{\lambda}_i = \frac{d}{dt}\sum_i \lambda_i = 0$, we obtain

$$\frac{d}{dt}S(\rho_t) = -\sum_i \dot{\lambda}_i \log_2 \lambda_i. \tag{3.10}$$

In the following step, we express $\dot{\rho}\log_2 \rho$ as

$$\dot{\rho}\log_2 \rho = \sum_{i,j}\left(\dot{\lambda}_i |\psi_i\rangle\langle\psi_i| + \lambda_i \frac{d}{dt}|\psi_i\rangle\langle\psi_i|\right)\log_2 \lambda_j |\psi_j\rangle\langle\psi_j|, \tag{3.11}$$

which implies

$$\mathrm{Tr}\left[\dot{\rho}\log_2\rho\right] = \sum_{i,j}\left(\dot{\lambda}_i\log_2\lambda_j\mathrm{Tr}\left[|\psi_i\rangle\langle\psi_i|\psi_j\rangle\langle\psi_j|\right] + \lambda_i\log_2\lambda_j\mathrm{Tr}\left[\frac{d}{dt}\left(|\psi_i\rangle\langle\psi_i|\right)|\psi_j\rangle\langle\psi_j|\right]\right) \tag{3.12}$$

$$= \sum_i \dot{\lambda}_i\log_2\lambda_i + \sum_{i,j}\lambda_i\log_2\lambda_j\mathrm{Tr}\left[\left(|\dot{\psi}_i\rangle\langle\psi_i| + |\psi_i\rangle\langle\dot{\psi}_i|\right)|\psi_j\rangle\langle\psi_j|\right]$$

$$= \sum_i \dot{\lambda}_i\log_2\lambda_i + \sum_i\lambda_i\log_2\lambda_i\left(\langle\psi_i|\dot{\psi}_i\rangle + \langle\dot{\psi}_i|\psi_i\rangle\right) = \sum_i\dot{\lambda}_i\log_2\lambda_i.$$

Therefore $\frac{d}{dt}S(\rho_t) = -\mathrm{Tr}\left[\dot{\rho}_t\log_2\rho_t\right]$.

Now, we obtain the time derivative of coherence as

$$\frac{dC_{\mathrm{r}}(\rho_t)}{dt} = \frac{dS(\Delta[\rho_t])}{dt} = -\mathrm{Tr}\left[\left(\frac{d}{dt}\Delta(\rho_t)\right)\log_2\Delta(\rho_t)\right]$$

$$= -\mathrm{Tr}\left[\Delta(\dot{\rho}_t)\log_2\Delta(\rho_t)\right], \tag{3.13}$$

where it is taken into account that the time derivative commutes with dephasing, i.e.,

$$\frac{d}{dt}\Delta(\rho_t) = \Delta(\dot{\rho}_t). \tag{3.14}$$

Using the von Neumann equation $\dot{\rho}_t = -i[H,\rho_t]$ Eq. (3.13) can be expressed as

$$\frac{dC_{\mathrm{r}}(\rho_t)}{dt} = i\mathrm{Tr}\left[\Delta([H,\rho_t])\log_2\Delta(\rho_t)\right]. \tag{3.15}$$

Using lemma 3.1 and choosing $A = i\Delta([H,\rho_t])$ and $B = \rho_t$ we further obtain:

$$\frac{dC_{\mathrm{r}}(\rho_t)}{dt} = i\mathrm{Tr}\left[[H,\rho_t]\log_2\Delta(\rho_t)\right]. \tag{3.16}$$

At $t = 0$:

$$\left.\frac{dC_{\mathrm{r}}(\rho_t)}{dt}\right|_{t=0} = i\mathrm{Tr}[[H,\rho]\log_2\Delta(\rho)], \tag{3.17}$$

where $\rho = \rho_{t=0}$. We can express this equation differently as

$$\left.\frac{dC_{\mathrm{r}}(\rho_t)}{dt}\right|_{t=0} = i\mathrm{Tr}(H[\rho,\log_2\Delta(\rho)]). \tag{3.18}$$

The proof is completed by maximizing over all states ρ.

□

3.2 Connection to the Surprisal of a Probability Distribution

Our aim here is to determine the maximum coherence-generating capacity across all Hamiltonians satisfying the constraint $\|H\|_2 \leq 1$, where $\|M\|_2 = \sqrt{\mathrm{Tr}[M^\dagger M]}$ denotes

the Hilbert-Schmidt norm of a matrix M. This problem demonstrates a close connection to the variance of surprisal, which was investigated in [RW15] (a similar approach was previously employed in [Bra07]).

Definition 3.1. For a probability distribution $\boldsymbol{p} = (p_0, \ldots, p_{d-1})$, the surprisal $-\log_2 p_i$ represents the measure of surprise to obtain the outcome i.

The variance of surprisal is expressed as

$$f(\boldsymbol{p}) = \sum_i p_i \left(-\log_2 p_i\right)^2 - \left[\sum_i p_i \left(-\log_2 p_i\right)\right]^2 . \tag{3.19}$$

We have the following lemma [SNS24].

Lemma 3.2. *Alternatively, $f(\boldsymbol{p})$ can be expressed as:*

$$f(\rho) = \frac{1}{2} \sum_{i,j} \rho_{ii}\rho_{jj} \left(\log_2 \rho_{jj} - \log_2 \rho_{ii}\right)^2 \tag{3.20}$$

Proof. [SNS24] This can be directly deduced from the following chain of equalities:

$$\sum_i \rho_{ii} \left(-\log_2 \rho_{ii}\right)^2 - \left[\sum_i \rho_{ii} \left(-\log_2 \rho_{ii}\right)\right]^2 \tag{3.21}$$

$$= \sum_i \rho_{ii} \left(\log_2 \rho_{ii}\right)^2 - \sum_i \rho_{ii}^2 \left(\log_2 \rho_{ii}\right)^2 - \sum_{i\neq j} \rho_{ii}\rho_{jj} \log_2 \rho_{ii} \log_2 \rho_{jj} \tag{3.22}$$

$$= \sum_i \rho_{ii} \left(1 - \rho_{ii}\right) \left(\log_2 \rho_{ii}\right)^2 - \sum_{i\neq j} \rho_{ii}\rho_{jj} \log_2 \rho_{ii} \log_2 \rho_{jj}$$

$$= \sum_i \rho_{ii} \left(\sum_{j\neq i} \rho_{jj}\right) \left(\log_2 \rho_{ii}\right)^2 - \sum_{i\neq j} \rho_{ii}\rho_{jj} \log_2 \rho_{ii} \log_2 \rho_{jj}$$

$$= \sum_{i\neq j} \rho_{ii}\rho_{jj} \left(\log_2 \rho_{ii}\right)^2 - \sum_{i\neq j} \rho_{ii}\rho_{jj} \log_2 \rho_{ii} \log_2 \rho_{jj}$$

$$= \frac{1}{2} \sum_{i\neq j} \rho_{ii}\rho_{jj} \left(\log_2 \rho_{ii}\right)^2 + \frac{1}{2} \sum_{i\neq j} \rho_{ii}\rho_{jj} \left(\log_2 \rho_{jj}\right)^2 - \sum_{i\neq j} \rho_{ii}\rho_{jj} \log_2 \rho_{ii} \log_2 \rho_{jj}$$

$$= \frac{1}{2} \sum_{i\neq j} \rho_{ii}\rho_{jj} \left[\left(\log_2 \rho_{ii}\right)^2 + \left(\log_2 \rho_{jj}\right)^2 - 2 \log_2 \rho_{ii} \log_2 \rho_{jj}\right]$$

$$= \frac{1}{2} \sum_{i\neq j} \rho_{ii}\rho_{jj} \left(\log_2 \rho_{ii} - \log_2 \rho_{ii}\right)^2 = \frac{1}{2} \sum_{i,j} \rho_{ii}\rho_{jj} \left(\log_2 \rho_{ii} - \log_2 \rho_{ii}\right)^2 .$$

Here, we have used the fact that $\sum_{j\neq i} \rho_{jj} = 1 - \rho_{ii}$. □

In the theorem that follows, we will illustrate how the maximum capacity to induce coherence in a system of dimension d is closely related to the highest variance of surprisal, denoted by f [SNS24].

Theorem 3.1. *It holds that*

$$\max_{\|H\|_2 \leq 1} C_{\text{gen}}(H) = \max_{\boldsymbol{p}} \sqrt{2f(\boldsymbol{p})}. \tag{3.23}$$

Proof. [SNS24] Let's define the following Hermitian matrix as

$$M = i[\rho, \log_2 \Delta(\rho)], \tag{3.24}$$

At time zero, the derivative of the relative entropy of coherence can be expressed as

$$\left.\frac{dC_{\text{r}}(\rho_t)}{dt}\right|_{t=0} = \text{Tr}(HM). \tag{3.25}$$

Now, we proceed with the maximization over all Hamiltonians H satisfying $\|H\|_2 \leq 1$. To accomplish this, we apply Holder's inequality, leading to the following result:

$$\left.\frac{dC_{\text{r}}(\rho_t)}{dt}\right|_{t=0} = \text{Tr}(HM) \leq \|H\|_2 \|M\|_2. \tag{3.26}$$

When considering a given M, this inequality achieves its maximum if H is selected as

$$H = \frac{M}{\|M\|_2}. \tag{3.27}$$

Upon maximizing across all Hamiltonians having bounded Hilbert-Schmidt norm, the result is

$$\max_{\|H\|_2 \leq 1} \left.\frac{dC_{\text{r}}(\rho_t)}{dt}\right|_{t=0} = \frac{\text{Tr}[M^2]}{\|M\|_2} = \|M\|_2 = \left\|[\rho, \log_2 \Delta(\rho)]\right\|_2. \tag{3.28}$$

To finalize the theorem's proof, we need to maximize $\left\|[\rho, \log_2 \Delta(\rho)]\right\|_2$ over all states ρ. If we denote the elements of ρ as ρ_{ij}, then we derive

$$\begin{aligned}
[\rho, \log_2 \Delta(\rho)] &= \rho \log_2 \Delta(\rho) - [\log_2 \Delta(\rho)]\,\rho \\
&= \sum_{i,j} \left(\rho_{ij} \log_2 \rho_{jj} - \rho_{ij} \log_2 \rho_{ii}\right) |i\rangle\langle j| \\
&= \sum_{i,j} \rho_{ij} \left(\log_2 \rho_{jj} - \log_2 \rho_{ii}\right) |i\rangle\langle j|.
\end{aligned} \tag{3.29}$$

With this, we obtain the following:

$$\left\|[\rho, \log_2 \Delta(\rho)]\right\|_2^2 = \sum_{i,j} |\rho_{ij}|^2 \left(\log_2 \rho_{jj} - \log_2 \rho_{ii}\right)^2. \tag{3.30}$$

Because ρ is a density matrix, we have:

$$\rho_{ii}\rho_{jj} \geq |\rho_{ij}|^2, \tag{3.31}$$

which implies the inequality

$$\left\|[\rho, \log_2 \Delta(\rho)]\right\|_2^2 \leq \sum_{i,j} \rho_{ii}\rho_{jj} \left(\log_2 \rho_{jj} - \log_2 \rho_{ii}\right)^2. \tag{3.32}$$

Now, define the following function:

$$f(\rho) = \frac{1}{2}\sum_{i,j}\rho_{ii}\rho_{jj}\left(\log_2\rho_{jj} - \log_2\rho_{ii}\right)^2. \tag{3.33}$$

We notice that the right side of Equation (3.32) is equivalent to $2f(\rho)$. It's worth noting that this function coincides with the variance of the surprisal function defined in Equation (3.19) when we set $p_i = \rho_{ii}$.

Let's now consider a pure state given by:

$$|\psi\rangle = \sum_{i=0}^{d-1}\sqrt{q_i}\,|i\rangle\,, \tag{3.34}$$

where the probabilities q_i are selected to maximize the variance of surprisal. Consider now the density matrix $\sigma = |\psi\rangle\langle\psi|$. Based on the aforementioned reasoning, it's evident that σ maximizes the function f, indicating $f(\sigma) = \max_\rho f(\rho)$. Furthermore, the matrix elements of σ satisfy $\sigma_{ii}\sigma_{jj} = |\sigma_{ij}|^2$, implying

$$\begin{aligned} 2f(\sigma) &= \sum_{i,j}\sigma_{ii}\sigma_{jj}\left(\log_2\sigma_{jj} - \log_2\sigma_{ii}\right)^2 \\ &= \sum_{i,j}|\sigma_{ij}|^2\left(\log_2\sigma_{jj} - \log_2\sigma_{ii}\right)^2 \\ &= \left\|[\sigma, \log_2\Delta(\sigma)]\right\|_2^2. \end{aligned} \tag{3.35}$$

Taking into account the arguments outlined earlier, we have the following conclusion:

$$\begin{aligned} \max_\rho\left\|[\rho, \log_2\Delta(\rho)]\right\|_2^2 &\le \max_\rho 2f(\rho) = 2f(\sigma) \\ &= \left\|[\sigma, \log_2\Delta(\sigma)]\right\|_2^2 \\ &\le \max_\rho\left\|[\rho, \log_2\Delta(\rho)]\right\|_2^2. \end{aligned} \tag{3.36}$$

This proves that

$$\max_\rho\left\|[\rho, \log_2\Delta(\rho)]\right\|_2^2 = 2f(\sigma), \tag{3.37}$$

and the proof of the theorem is complete. □

In the next section, we find the optimal initial state and Hamiltonian for the coherence generation rate.

3.3 Optimal Coherence Generation Rate

Following the proof presented in the previous section to establish the connection of the maximal coherence generation rate with the surprisal of the probability distribution of

the initial state, we provide an expression for the initial state and Hamiltonian to begin with to optimally generate the coherence resource[SNS24]. For the initial state, we have

$$|\psi\rangle = \sqrt{\gamma}\,|0\rangle + \sqrt{\frac{1-\gamma}{d-1}}\sum_{i=1}^{d-1}|i\rangle\,. \tag{3.38}$$

The value for γ is selected within the interval (0,1) in a manner that the probability distribution $(\gamma, \frac{1-\gamma}{d-1}, \ldots, \frac{1-\gamma}{d-1})$ to maximize the variance of the surprisal as mentioned in references [RW15, Bra07]. Also, the optimal Hamiltonian is described by Equation (3.27), along with [SNS24]:

$$\begin{aligned} M &= i[\psi, \log_2 \Delta(\psi)] = i\sum_{k,l}\psi_{kl}\left(\log_2\psi_{ll} - \log_2\psi_{kk}\right)|k\rangle\langle l| \\ &= i\sqrt{\gamma}\sqrt{\frac{1-\gamma}{d-1}}\left(\log_2\frac{1-\gamma}{d-1} - \log_2\gamma\right)\sum_{l=1}^{d-1}|0\rangle\langle l| \\ &+ i\sqrt{\gamma}\sqrt{\frac{1-\gamma}{d-1}}\left(\log_2\gamma - \log_2\frac{1-\gamma}{d-1}\right)\sum_{k=1}^{d-1}|k\rangle\langle 0| \\ &= i\alpha\left(|0\rangle\langle\varphi| - |\varphi\rangle\langle 0|\right) \end{aligned} \tag{3.39}$$

with a state $|\varphi\rangle = \sum_{i=1}^{d-1}|i\rangle / \sqrt{d-1}$ and some $\alpha \in \mathbb{R}$. Therefore, one can select an optimal Hamiltonian as

$$H = \frac{i}{\sqrt{2}}\left(|0\rangle\langle\varphi| - |\varphi\rangle\langle 0|\right). \tag{3.40}$$

In the next section, we study the problem for the case of a qubit.

3.4 Qubit Case

Let's now direct our attention specifically to the single-qubit scenario. Here, we'll assess $C_{\text{gen}}(H)$ for any given Hamiltonian H. Hereafter, we represent the components of the density matrix as ρ_{kl}, while H_{kl} denotes the elements of H in a similar fashion. Moreover, $\rho_{01} = |\rho_{01}|e^{i\alpha}$ and similarly $H_{01} = |H_{01}|e^{i\beta}$. Using Eq. (3.29) we obtain

[SNS24]

$$
\begin{aligned}
i\mathrm{Tr}\left(H[\rho, \log_2 \Delta(\rho)]\right) &= i \sum_{k,l} H_{lk}\rho_{kl}\left(\log_2 \rho_{ll} - \log_2 \rho_{kk}\right) \\
&= i\left[H_{10}\rho_{01}\left(\log_2 \rho_{11} - \log_2 \rho_{00}\right)\right] \\
&+ i\left[H_{01}\rho_{10}\left(\log_2 \rho_{00} - \log_2 \rho_{11}\right)\right] \\
&= i\left[H_{10}\rho_{01} - H_{01}\rho_{10}\right]\log_2 \frac{\rho_{11}}{\rho_{00}} \\
&= i\,|H_{10}|\,|\rho_{01}|\left[e^{i(\alpha-\beta)} - e^{-i(\alpha-\beta)}\right]\log_2 \frac{\rho_{11}}{\rho_{00}} \\
&= -2\,|H_{10}|\,|\rho_{01}|\sin(\alpha-\beta)\log_2 \frac{\rho_{11}}{\rho_{00}}.
\end{aligned}
\tag{3.41}
$$

Now, our objective is to optimize this expression across all values of α, $|\rho_{01}|$, ρ_{00}, and ρ_{11}, considering that ρ represents the density matrix of a single qubit. Maximizing with respect to α is straightforward; an optimal selection is $\alpha = \beta - \pi/2$. Hence, we obtain [SNS24]:

$$
C_{\mathrm{gen}}(H) = \max_{\rho_{ij}} 2\,|H_{10}|\,|\rho_{01}|\log_2 \frac{\rho_{11}}{\rho_{00}}. \tag{3.42}
$$

For any density matrix representing a qubit, it holds true that $|\rho_{01}| \leq \sqrt{\rho_{00}\rho_{11}}$, where equality occurs in the case of pure states. Utilizing this, we can conduct the maximization concerning $|\rho_{01}|$, resulting in [SNS24]:

$$
C_{\mathrm{gen}}(H) = \max_{\rho_{ij}} 2\,|H_{10}|\,\sqrt{\rho_{00}\rho_{11}}\log_2 \frac{\rho_{11}}{\rho_{00}}. \tag{3.43}
$$

This also implies that an optimal state can be selected to be pure. In the final step, we remind ourselves that $\rho_{11} = 1 - \rho_{00}$, thus:

$$
C_{\mathrm{gen}}(H) = \max_{\rho_{00}} 2\,|H_{10}|\,\sqrt{\rho_{00}(1-\rho_{00})}\log_2 \frac{1-\rho_{00}}{\rho_{00}}. \tag{3.44}
$$

This maximization can be carried out numerically, yielding $\rho_{00} \approx 0.083$ [SNS24].

Similar findings have been documented in entanglement theory previously. Specifically, optimal entanglement generation concerning two-qubit Hamiltonians was explored in [DVC+01], where optimal states for entanglement generation without ancillas were derived. Additionally, investigations into optimal entanglement generation with Hamiltonians of bounded operator norm were conducted in [Bra07].

3.5 Discussion

In summary, our findings in this chapter delivers a detailed examination of how Hamiltonians generate quantum coherence, offering a method to measure the maximum rate of coherence increase in quantum systems of any size, with Hamiltonians having limited Hilbert-Schmidt norms. Specifically, for qubit systems, we have fully resolved this

issue for any Hamiltonian, pinpointing the states that maximize the rate of change in the relative entropy of coherence.

This research opens several intriguing directions for future inquiry. One major area of interest is the potential for enhancing coherence through specific Hamiltonians in systems with dimensions larger than qubits. Although our approach offers a new framework for addressing this issue, it remains uncertain whether this optimization problem can be solved analytically or via semidefinite programming. Additionally, exploring whether our methods can be applied to other quantum resource theories, such as entanglement, is a compelling prospect. Given the similarities between coherence and entanglement resource theories, there is significant potential for our techniques to uncover optimal strategies for boosting entanglement in systems using particular Hamiltonian classes.

Contribution: This chapter is based on the work in the paper *Coherence Generation with Hamiltonian* (of which I am the second author). I contributed to the results of this work along with Dr. Manfredi Scalici. We found the results and established the proofs thinking together and through joint discussions under the supervision of Prof. Alexander Streltsov. However, the clean and final version of the results was written by Dr. Manfredi Scalici.

Chapter 4

Purity, Coherence and Entanglement in the Bernstein-Vazirani Algorithm

Quantum algorithms enable superior performance compared to classical counterparts in various tasks, with Shor's algorithm for efficient prime factorization on a quantum computer being a prominent example [Sho94]. The key factor contributing to this speedup is the superposition principle of quantum mechanics, allowing a quantum processor to exist in multiple states simultaneously. Although such superposition can lead to entanglement across different qubits, there are also quantum algorithms that leverage individual qubits' superpositions without entanglement to surpass classical algorithms. The Bernstein-Vazirani algorithm serves as an example by enabling the determination of a bit string encoded in an oracle [BV97]. While the classical version requires multiple calls to learn the bit string, a single query suffices in the quantum case. This chapter provides a detailed analysis of the quantum resources involved in the Bernstein-Vazirani algorithm, focusing on its probabilistic version, where the goal is to guess the bit string after a single oracle call. It demonstrates that the algorithm's performance directly links to the amount of quantum coherence in the initial state [NKG+22]. Additionally, it reveals that an excessive degree of entanglement in the initial state hinders the algorithm from achieving optimal performance [NKG+22]. The study also extends to quantum computation with mixed states, demonstrating that pseudo pure states attain optimal performance for a given purity in the Bernstein-Vazirani algorithm [NKG+22].

4.1 Probabilistic Bernstein-Vazirani Algorithm

The objective of the Bernstein-Vazirani algorithm (BV algorithm) [BV97] is to determine a concealed N-bit string $\boldsymbol{a} = a_1, \ldots, a_N$, where each a_i is either 0 or 1, and the string is encoded using a linear function:

$$f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x} \bmod 2 = \left(\sum_{k=0}^{N} a_k x_k \right) \bmod 2 \tag{4.1}$$

on the N-bit string $\boldsymbol{x} = x_1, \ldots, x_N$. The specific objective is to identify the string $\boldsymbol{a}$ using the fewest possible queries to the function f. In classical computing, the optimal approach involves evaluating f for every input $\boldsymbol{x}$ in which one bit is set to 1, while the remaining $N - 1$ bits are set to 0. This strategy results in a total of N queries to the function f [BV97].

In the quantum domain, however, only a single call of the function is required to acquire knowledge of the bit string $\boldsymbol{a}$ [BV97]. To achieve this, we employ the standard assumption that the bit string $\boldsymbol{x}$ is encoded in an N-qubit quantum state, denoted as $|\boldsymbol{x}\rangle = \otimes_{i=1}^{N} |x_i\rangle$. Furthermore, we assume that the function is encoded in an oracle, represented by a unitary operator $U_{\boldsymbol{a}}$ that acts on states of the form $|i\rangle |\boldsymbol{x}\rangle$, where $i \in 0, 1$, as described below:

$$U_{\boldsymbol{a}}(|i\rangle |\boldsymbol{x}\rangle) = |i \oplus f(\boldsymbol{x})\rangle |\boldsymbol{x}\rangle , \tag{4.2}$$

and $\oplus$ denotes addition modulo 2. In the subsequent description, the first qubit is referred to as the "oracle register," while the remaining N qubits are referred to as the "system qubits." When the oracle unitary $U_{\boldsymbol{a}}$ operates on the state $|-\rangle |+\rangle^{\otimes N}$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$, the resulting state can be expressed as $\sum_{\boldsymbol{x}} (-1)^{\boldsymbol{a}\cdot\boldsymbol{x}} |-\rangle |\boldsymbol{x}\rangle / 2^N$. By discarding the oracle register and applying a Hadamard gate to each qubit, the overall state is transformed into $|\boldsymbol{a}\rangle$. Finally, the bit string $\boldsymbol{a}$ can be obtained by measuring each qubit in the computational basis [BV97].

Thus far, we have observed that the Bernstein-Vazirani (BV) algorithm exhibits optimal performance when the initial state is $|-\rangle |+\rangle^{\otimes N}$. Now, let's examine the algorithm's performance for general input states. In the general case, we cannot expect the procedure to function optimally if the initial state differs from $|-\rangle |+\rangle^{\otimes N}$. To assess the performance in this broader scenario, we assume no prior knowledge regarding the bit string $\boldsymbol{a}$. In other words, each possible bit string is equally likely. Since there are 2^N potential bit strings, the probability of each bit string is $1/2^N$. To acquire information about $\boldsymbol{a}$, we permit the application of the oracle unitary $U_{\boldsymbol{a}}$ to a general quantum state ρ and execute a general quantum measurement on the final state $U_{\boldsymbol{a}} \rho U_{\boldsymbol{a}}^{\dagger}$. This protocol is referred to as the "probabilistic Bernstein-Vazirani algorithm" [NKG+22].

The performance of the probabilistic BV algorithm can be quantified by the optimal probability of correctly guessing the bit string $\boldsymbol{a}$, which corresponds to the highest probability of correctly identifying the oracle unitary $U_{\boldsymbol{a}}$. This definition is analogous to the notion of average guessing probability in channel discrimination tasks, as described in the literature such as [BK15]. In channel discrimination, a set of quantum

channels Λ_i is applied with probabilities p_i, and the goal is to discriminate between them by applying the channels to an initial quantum state ρ and performing a positive operator-valued measure (POVM) M_i on the final state. The average probability of correctly guessing the channel is given by $\sum_i p_i \operatorname{Tr}[\Lambda_i(\rho)M_i]$. Now, we define the performance of the probabilistic BV algorithm as the maximum probability of successfully guessing the bit string $\boldsymbol{a}$ over all possible POVM measurements [NKG+22]:

$$P(\rho) = \frac{1}{2^N} \max_{\{M_a\}} \sum_{\boldsymbol{a}} \operatorname{Tr}\left[U_{\boldsymbol{a}} \rho U_{\boldsymbol{a}}^{\dagger} M_{\boldsymbol{a}}\right]. \tag{4.3}$$

After establishing the performance definition of the BV algorithm for general input states, let us now derive a closed expression for its performance when considering all pure initial states. To do this, it is worth noting that any pure state of $N+1$ qubits can be expressed as follows:

$$|\mu\rangle = a\,|+\rangle\,|\varphi'\rangle + b\,|-\rangle\,|\varphi\rangle\,, \tag{4.4}$$

where $|\varphi\rangle$ and $|\varphi'\rangle$ are states of N qubits. In the following, $c_{\boldsymbol{x}}$ denote the coefficients of the state $|\varphi\rangle$ in the computational basis, i.e., $|\varphi\rangle = \sum_{\boldsymbol{x}} c_{\boldsymbol{x}} |\boldsymbol{x}\rangle$. We have the theorem stated below [NKG+22].

Theorem 4.1. *The performance of the probabilistic BV algorithm for a pure initial state is given as*

$$\begin{aligned} P(|\mu\rangle) \quad &= \frac{1}{2^N}\Bigg[1 + |b|^2 R(|\varphi\rangle\langle\varphi|) \\ &\quad + 2|b| \sum_{x\neq \mathbf{0}} |c_{\boldsymbol{x}}| \left(\sqrt{1 - |b|^2\left(1 - |c_{\mathbf{0}}^2|\right)} - |b||c_{\mathbf{0}}| \right) \Bigg]. \end{aligned} \tag{4.5}$$

where R is the robustness of coherence in the computational basis.

Proof. [NKG+22] To prove the result stated in Theorem 4.1, let us revisit the action of the oracle:

$$|\mu\rangle = a\,|+\rangle\,|\varphi'\rangle + b\,|-\rangle\,|\varphi\rangle \rightarrow |\mu_{\boldsymbol{a}}\rangle = a\,|+\rangle\,|\varphi'\rangle + b\,|-\rangle\,|\psi_{\boldsymbol{a}}\rangle\,. \tag{4.6}$$

To analyze the final state $|\mu_{\boldsymbol{a}}\rangle$, we will rearrange it in a more convenient manner. Let's consider the computational basis $\{|\boldsymbol{x}\rangle\}_{\boldsymbol{x}}$. We can express the final state as follows:

$$|\psi_{\boldsymbol{a}}\rangle = \sum_{\boldsymbol{x}} c_{\boldsymbol{x}} (-1)^{\boldsymbol{a}\cdot\boldsymbol{x}} |\boldsymbol{x}\rangle\,, \tag{4.7}$$

In the given expression, we represent the coefficient of the state $|\varphi\rangle = \sum_{\boldsymbol{x}} c_{\boldsymbol{x}} |\boldsymbol{x}\rangle$ as $c_{\boldsymbol{x}}$. To simplify the notation, we introduce the symbol $\mathbf{0}$, which represents the string of N zeros, denoted as 00...0. Hence,

$$|\mu_{\boldsymbol{a}}\rangle = a\,|+\rangle\,|\varphi'\rangle + b\,|-\rangle\,(c_{\mathbf{0}}\,|0\rangle^{\otimes N}) + b\,|-\rangle \left(\sum_{\boldsymbol{x}\neq\mathbf{0}} c_{\boldsymbol{x}} (-1)^{\boldsymbol{a}\cdot\boldsymbol{x}} |\boldsymbol{x}\rangle \right) \tag{4.8}$$

In this context, it is assumed that the dimension of the entire system, including the oracle register, is 2^{N+1}. Now,

$$|\mu_{\boldsymbol{a}}\rangle = \sqrt{|a|^2 + |b|^2|c_{\boldsymbol{0}}|^2}\,|\psi'_{\boldsymbol{0}}\rangle + b\sum_{\boldsymbol{x}\neq\boldsymbol{0}} c_{\boldsymbol{x}}(-1)^{\boldsymbol{a}\cdot\boldsymbol{x}}\,|\psi'_{\boldsymbol{x}}\rangle \tag{4.9}$$

with the states,

$$|\psi'_{\boldsymbol{0}}\rangle = \frac{a\,|+\rangle\,|\varphi'\rangle + bc_{\boldsymbol{0}}\,|-\rangle\,|0\rangle^{\otimes N}}{\sqrt{|a|^2 + |b|^2|c_{\boldsymbol{0}}|^2}} \tag{4.10}$$

and $|\psi'_{\boldsymbol{x}}\rangle = |-\rangle\,|\boldsymbol{x}\rangle$ for $\boldsymbol{x} \neq \boldsymbol{0}$. It is important to note that the state $|\psi'_{\boldsymbol{x}}\rangle$ is orthogonal to the state $|\psi'_{\boldsymbol{0}}\rangle$ for all $\boldsymbol{x} \neq \boldsymbol{0}$. By constructing a new basis using these orthogonal states, we can express the final state as follows:

$$|\mu_{\boldsymbol{a}}\rangle = \sum_{\boldsymbol{x}} c'_{\boldsymbol{x}}(-1)^{\boldsymbol{a}\cdot\boldsymbol{x}}\,|\psi'_{\boldsymbol{x}}\rangle \tag{4.11}$$

with $c'_{\boldsymbol{0}} = \sqrt{|a|^2 + |b|^2|c_{\boldsymbol{0}}|^2}$ and $c'_{\boldsymbol{x}} = bc_{\boldsymbol{x}}$ for $\boldsymbol{x} \neq \boldsymbol{0}$. Based on these observations, we can deduce that for an initial state $|\mu\rangle$, the action of the oracle unitary is equivalent to the unitary $U'_{\boldsymbol{a}} = \sum_{\boldsymbol{x}}(-1)^{\boldsymbol{a}\cdot\boldsymbol{x}}\,|\psi'_{\boldsymbol{x}}\rangle\langle\psi'_{\boldsymbol{x}}|$, which is diagonal in the basis $\{|\psi'_{\boldsymbol{x}}\rangle\}_{\boldsymbol{x}}$. Consequently, we can assess the optimal performance by determining the maximum probability of distinguishing between the unitaries $U'_{\boldsymbol{a}}$ when applied to the state $|\mu\rangle$. Analogous to Eq. (4.49), we obtain:

$$P(|\mu\rangle) = \frac{1 + R'(|\mu\rangle\langle\mu|)}{2^N}, \tag{4.12}$$

where R' is the robustness of coherence in the basis $\{|\psi'_{\boldsymbol{x}}\rangle\}_{\boldsymbol{x}}$. By utilizing the structure of the states $|\psi'_{\boldsymbol{x}}\rangle$ and leveraging the property of the robustness that it is equal to the l_1-norm of coherence for pure states, we can express this result in terms of the robustness of coherence R with respect to the computational basis.

$$\begin{aligned} P(|\mu\rangle) \quad &= \frac{1}{2^N}\Bigg[1 + |b|^2 R(|\varphi\rangle\langle\varphi|) \\ &\quad +2|b|\sum_{\boldsymbol{x}\neq\boldsymbol{0}} |c_{\boldsymbol{x}}|\Bigg(\sqrt{1 - |b|^2\Big(1 - |c_{\boldsymbol{0}}^2|\Big)} - |b||c_{\boldsymbol{0}}|\Bigg)\Bigg]. \end{aligned} \tag{4.13}$$

This completes the proof of Theorem 4.1. □

As we can observe from Eq. (4.12), in order to achieve the maximum performance $P = 1$, it is necessary for the robustness of coherence $R(|\mu\rangle\langle\mu|)$ to be equal to $2^N - 1$. Consequently, the state $|\mu\rangle$ must be a maximally coherent state in the basis $|\boldsymbol{x}\rangle_{\boldsymbol{x}}$ which implies [NKG+22]

$$|b|^2|c_{\boldsymbol{x}}|^2 = \frac{1}{2^N} \quad \forall \boldsymbol{x} \neq \boldsymbol{0}, \tag{4.14}$$

$$|a|^2 + |b|^2|c_{\boldsymbol{0}}|^2 = \frac{1}{2^N}. \tag{4.15}$$

Taking into account the relationships $|a|^2 = 1 - |b|^2$ and $\sum_{\boldsymbol{x}\neq\mathbf{0}} c_{\boldsymbol{x}} = 1 - |c_{\mathbf{0}}|^2$, we can rewrite the equations as follows:

$$|b|^2|c_{\mathbf{0}}|^2 = \frac{1}{2^N} \tag{4.16}$$

$$1 - |b|^2(1 - |c_{\mathbf{0}}|^2) = \frac{1}{2^N}. \tag{4.17}$$

By solving these two equations for $|b|$ and $|c_{\mathbf{0}}|$, we find that $|b| = 1$ and $c_{\mathbf{0}} = \frac{1}{\sqrt{2^N}}$ [NKG+22]. Moreover, from Eq. (4.14), we can deduce that $|c_{\boldsymbol{x}}| = \frac{1}{\sqrt{2^N}}$ for all $\boldsymbol{x}$. This implies that in order to achieve maximum performance, the initial state must be of the form $|\mu\rangle = |-\rangle|\psi_{\mathrm{MC}}\rangle$, where $|\psi_{\mathrm{MC}}\rangle$ represents a maximally coherent state in the computational basis [NKG+22].

We can extend the result of Theorem 4.1 to mixed states as well. If we initialize the BV algorithm in a state of the form $\rho = \sum_i p_i |\mu_i\rangle\langle\mu_i|$, where $|\mu_i\rangle = a_i|+\rangle|\varphi\rangle + b_i|-\rangle|\psi_i\rangle$, with $|a_i|^2 + |b_i|^2 = 1$ and $\langle 00...0|\psi_i\rangle = 0$, the action of the oracle unitary on this state is [NKG+22]:

$$U = \sum_{\boldsymbol{x}\neq\mathbf{0}} (-1)^{\boldsymbol{a}\cdot\boldsymbol{x}} |-\rangle\langle-| \otimes |\boldsymbol{x}\rangle\langle\boldsymbol{x}| + |+\rangle\langle+| \otimes |\varphi\rangle\langle\varphi|\,. \tag{4.18}$$

In this scenario, considering the basis $\{|+\rangle|\varphi\rangle\} \cup \{|-\rangle|\boldsymbol{x}\rangle\}_{\boldsymbol{x}\neq\mathbf{0}}$, the action of the unitary U is equivalent to the action of the oracle unitary on the system qubit in the computational basis when the oracle register is in the state $|-\rangle$. Thus, based on Eq. (4.49), we can express the performance as follows [NKG+22]:

$$P(\rho) = \frac{1 + R'(\rho)}{2^N} \tag{4.19}$$

in which $R'(\rho)$ is the robustness of coherence in the basis $\{|+\rangle|\varphi\rangle\} \cup \{|-\rangle|\boldsymbol{x}\rangle\}_{\boldsymbol{x}}$.

Let's compare the probabilistic version of the BV algorithm described earlier with its classical counterpart. To begin, we define the classical version of the probabilistic BV algorithm [NKG+22]. In the classical case, the BV algorithm maps the $N + 1$ bit string $(i, \boldsymbol{x})$ to $(i \oplus f(\boldsymbol{x}), \boldsymbol{x})$. Assuming that each possible function f is applied with an equal probability of $1/2^N$, the performance of the classical BV algorithm can be defined as the maximal probability of correctly guessing the bit string $\boldsymbol{a}$ when the algorithm is applied to the bit string $(i, \boldsymbol{x})$. We have the following Theorem [NKG+22].

Theorem 4.2. *The performance of the classical probabilistic BV algorithm is given by*

$$P_{\mathrm{c}}(\boldsymbol{x}) = \begin{cases} \frac{1}{2^N} & \text{if } \boldsymbol{x} = \mathbf{0}, \\ \frac{1}{2^{N-1}} & \text{otherwise.} \end{cases} \tag{4.20}$$

Proof. [NKG+22] To prove the theorem, we will provide the maximal probability for correctly guessing the bit string $\boldsymbol{a}$ in the classical probabilistic BV algorithm. For the

case where $\boldsymbol{x} = \mathbf{0}$, we have $f(\boldsymbol{x}) = 0$, indicating that the oracle does not provide any information about the bit string $(i, \boldsymbol{x})$. Therefore, the probability of success in this case is equal to 1.

Now, we will prove that for any $r \in 0, 1$ and any string $\boldsymbol{x} \neq \mathbf{0}$ with $x_i \in 0, 1$, there are 2^{N-1} different strings $\boldsymbol{a}$ with $a_i \in 0, 1$ such that $\boldsymbol{a} \cdot \boldsymbol{x} \bmod 2 = r$. Here, N represents the length of the strings $\boldsymbol{a}$ and $\boldsymbol{x}$. We have,

$$\boldsymbol{a} \cdot \boldsymbol{x} \bmod 2 = \left(\sum_i a_i x_i \bmod 2 \right) \bmod 2. \tag{4.21}$$

As $\boldsymbol{x} \neq \mathbf{0}$, there is at least one $x_l = 1$ ($l \in \{1, 2, .., N\}$). Therefore,

$$\boldsymbol{a} \cdot \boldsymbol{x} \bmod 2 = \left(\left[\sum_{i \neq l} a_i x_i \right] \bmod 2 + a_l \right) \bmod 2 = r. \tag{4.22}$$

Indeed, based on the result that for any $r \in 0, 1$ and any string $\boldsymbol{x} \neq \mathbf{0}$ with $x_i \in 0, 1$, there are 2^{N-1} different bit strings $\boldsymbol{a}$ satisfying $\boldsymbol{a} \cdot \boldsymbol{x} \bmod 2 = r$, we can conclude that knowing $(i \oplus f(\boldsymbol{x}), \boldsymbol{x})$ for any $\boldsymbol{x} \neq \mathbf{0}$, the success probability of correctly guessing $\boldsymbol{a}$ is equal to $1/2^{N-1}$. This means that the classical BV algorithm achieves a success probability of $1/2^{N-1}$ for correctly guessing $\boldsymbol{a}$ when applied to the bit string $(i, \boldsymbol{x})$. □

Indeed, upon comparing Eqs. (4.20) and (4.5), it is evident that in the classical version of the algorithm, the performance cannot exceed $1/2^{N-1}$ in a single oracle call. On the other hand, the quantum case allows for superior performance, as optimal performance $P = 1$ can be attained for certain initial states. This highlights the advantage of the quantum BV algorithm, as it enables a higher probability of successfully determining the target bit string in a single use of the oracle [NKG+22]. Next, we study how we can perform the BV algorithm without entanglement.

4.2 Probabilistic Bernstein-Vazirani Algorithm Without Entanglement.

It is worth noting the significance of quantum coherence in the performance of the BV algorithm, as stated in Theorem 4.1. The performance of the algorithm is explicitly given by Eq. (4.5), where the total initial state $|\mu\rangle$ can be entangled or separable. Considering that entanglement is generally considered a valuable and limited resource in quantum information theory [HHHH09b], it is indeed reasonable to explore the performance of the algorithm in scenarios where there is no entanglement between all $N + 1$ qubits, both before and after the action of the oracle. By investigating the algorithm's behavior in the absence of such entanglement, we can gain insights into its performance using more readily available resources.

Let's begin by focusing on pure initial states and then extend our discussion to mixed states later on. To demonstrate that the probabilistic BV algorithm can achieve performance above $1/2^N$ without entanglement in the initial and final state, we need to show that the total initial state must have the following form [NKG+22]:

$$|\mu\rangle = |-\rangle |\varphi\rangle \tag{4.23}$$

with an N-qubit product state $|\varphi\rangle$. It is worth noting that any pure initial state comprising $N+1$ qubits can be expressed as $|\mu\rangle = a|+\rangle|\varphi'\rangle + b|-\rangle|\varphi\rangle$. Following the application of the oracle unitary $U_{\boldsymbol{a}}$, the state adopts the following structure [NKG+22]:

$$U_{\boldsymbol{a}} |\mu\rangle = a|+\rangle|\varphi'\rangle + b|-\rangle|\psi_{\boldsymbol{a}}\rangle \tag{4.24}$$

with $|\psi_{\boldsymbol{a}}\rangle = \sum_{\boldsymbol{x}} c_{\boldsymbol{x}} (-1)^{\boldsymbol{a}\cdot\boldsymbol{x}} |\boldsymbol{x}\rangle$. When we trace out the oracle register, the resulting reduced state of the N-qubit system can be described by the following expression [NKG+22]:

$$\rho_{\boldsymbol{a}} = |a|^2 |\varphi'\rangle\langle\varphi'| + |b|^2 |\psi_{\boldsymbol{a}}\rangle\langle\psi_{\boldsymbol{a}}| . \tag{4.25}$$

To ensure that the state $U_{\boldsymbol{a}}|\mu\rangle$ remains separable for all bit strings $\boldsymbol{a}$, we can observe that it must satisfy one of the following conditions: either $|a| = 1$ or $|b| = 1$, or $|\psi_{\boldsymbol{a}}\rangle = |\varphi'\rangle$ for all $\boldsymbol{a}$ [NKG+22]. However, in the latter case, the final state $U_{\boldsymbol{a}}|\mu\rangle$ becomes independent of $\boldsymbol{a}$, leading to minimal performance. Similarly, if $|a| = 1$, the performance will also be minimal. Therefore, the only remaining scenario is when $|b| = 1$, which implies that the initial state must be in the form described by equation (4.23) [NKG+22].

Based on the arguments presented above, it becomes clear that the only viable option to avoid entanglement in both the initial and final states of the algorithm while maintaining nontrivial performance is to initialize the algorithm in a state of the form $|-\rangle|\varphi\rangle_p$, where $|\varphi\rangle_p$ represents a product state [NKG+22]. This choice of initialization ensures that the state remains separable throughout the algorithm's execution, while still allowing for significant performance.

The action of the unitary $U_{\boldsymbol{a}}$ on states of the form $|-\rangle|\varphi\rangle$ can be described as follows [NKG+22]:

$$U_{\boldsymbol{a}} (|-\rangle|\varphi\rangle) = |-\rangle \otimes (V_{\boldsymbol{a}} |\varphi\rangle) \tag{4.26}$$

where $V_{\boldsymbol{a}}$ is an N-qubit unitary operation that can be implemented by applying σ_z on the i-th qubit conditioned on the value of a_i, i.e., $V_{\boldsymbol{a}} = \otimes_{i=1}^{N} \sigma_{z,i}^{a_i}$. It is important to note that $V_{\boldsymbol{a}}$ does not introduce any entanglement in the N-qubit system (as it is local).

In order to ensure that entanglement does not have any influence on the algorithm, we will also investigate the feasibility of implementing the optimal positive operator-valued measure (POVM) $M_{\boldsymbol{a}}$ in Eq. (4.3) without requiring the use of entanglement. Although the density matrix immediately before applying the POVM consists of a combination of non-entangled states that we must differentiate in order to determine the bit-string $\boldsymbol{a}$, it is possible that the execution of the POVM to maximize Eq.(4.3) might necessitate the use of non-local operations[BDF+99, HBAB19]. In this case, we will

demonstrate that when qubits are initialized in $|\mu\rangle = |-\rangle|\varphi\rangle$, where $|\varphi\rangle = \otimes_{i=1}^{N}|\varphi^i\rangle$, it is feasible to attain the maximum value in Eq. (4.3) using non-entangling measurements [NKG+22]. To achieve this objective, we investigate a positive operator-valued measure (POVM) that consists of elements $M_{\boldsymbol{a}} = \bigotimes_{i=1}^{N} M^{(i)}a_i$, where $M^{(i)}a_i$ represents a single-qubit POVM applied to the i-th qubit. By utilizing the fact that the action of the oracle on $|-\rangle|\varphi\rangle$ corresponds to the implementation of the unitary $V_{\boldsymbol{a}} = \otimes_{i=1}^{N}\sigma_{z,i}^{a_i}$ on $|\varphi\rangle$, we can deduce the following relationship [NKG+22]:

$$\frac{1}{2^N}\sum_{\boldsymbol{a}}\mathrm{Tr}\left[U_{\boldsymbol{a}}|\mu\rangle\langle\mu|U_{\boldsymbol{a}}^{\dagger}M_{\boldsymbol{a}}\right] = \prod_{i=1}^{N}\frac{1}{2}\sum_{a_i}\mathrm{Tr}\left[\sigma_{z,i}^{a_i}|\varphi^i\rangle\langle\varphi^i|\sigma_{z,i}^{a_i}M_{a_i}^{(i)}\right]. \tag{4.27}$$

Considering that $P(|\mu\rangle\langle\mu|)$ corresponds to the maximum value attained by any POVM on the right-hand side of Eq. (4.3), we can state the following:

$$P(|\mu\rangle\langle\mu|) \geq \max_{\{M_{a_i}^{(i)}\}}\prod_{i=1}^{N}\frac{1}{2}\sum_{a_i}\mathrm{Tr}\left[\sigma_{z,i}^{a_i}|\varphi^i\rangle\langle\varphi^i|\sigma_{z,i}^{a_i}M_{a_i}^{(i)}\right]. \tag{4.28}$$

Our aim is now to establish that the inequality (4.28) is actually an equality. To demonstrate this, we can recall that $P(|\mu\rangle\langle\mu|) = [1 + R(|\varphi\rangle\langle\varphi|)]/2^N$ as indicated in Eq. (4.33). By utilizing the fact that for pure states, the robustness of coherence and the ℓ_1-norm of coherence are equivalent [PCB+16], and applying the properties of the ℓ_1-norm of coherence [BKZW17], we can establish the validity of the following equality [NKG+22]:

$$1 + R(|\varphi\rangle\langle\varphi|) = \prod_{i=1}^{N}[1 + R(|\varphi^i\rangle\langle\varphi^i|)]. \tag{4.29}$$

We can now choose $M_{a_i}^{(i)}$ in a manner that ensures the fulfillment of the following condition for all $i \in [1, N]$:

$$\begin{aligned}\frac{1}{2}\max_{\{M_{a_i}^{(i)}\}}\sum_{a_i}\mathrm{Tr}[\sigma_z^{a_i}|\varphi^i\rangle\langle\varphi^i|\sigma_z^{a_i}M_{a_i}^{(i)}] &= P(|-\rangle|\varphi^i\rangle) \\ &= \frac{1 + R(|\varphi^i\rangle\langle\varphi^i|)}{2}.\end{aligned} \tag{4.30}$$

This conclusion is derived from the observation that the BV algorithm, when applied to a single qubit in a pure state, attains its maximum performance [NKG+22]. This maximum performance corresponds to the $[1 + R(|\varphi^i\rangle\langle\varphi^i|)]/2$. Finally, as:

$$\begin{aligned}P(|\mu\rangle\langle\mu|) &\geq \max_{\{M_{a_i}^{(i)}\}}\prod_{i=1}^{N}\frac{1}{2}\sum_{a_i}\mathrm{Tr}\left[\sigma_{z,i}^{a_i}|\varphi^i\rangle\langle\varphi^i|\sigma_{z,i}^{a_i}M_{a_i}^{(i)}\right] \\ &\geq \prod_{i=1}^{N}\frac{1}{2}\max_{\{M_{a_i}^{(i)}\}}\sum_{a_i}\mathrm{Tr}\left[\sigma_{z,i}^{a_i}|\varphi^i\rangle\langle\varphi^i|\sigma_{z,i}^{a_i}M_{a_i}^{(i)}\right],\end{aligned} \tag{4.31}$$

we deduce [NKG+22]:

$$P(|\mu\rangle\langle\mu|) \geq \prod_{i=1}^{N}\frac{1 + R(|\varphi^i\rangle\langle\varphi^i|)}{2} = \frac{1 + R(|\varphi\rangle\langle\varphi|)}{2^N} \tag{4.32}$$

The equation $P(|\mu\rangle\langle\mu|) = (1 + R(|\varphi\rangle\langle\varphi|))/2^N$ reveals that the performance of the algorithm, as defined in Eq. (4.3), can be maximized when the system is initially in a pure product state of N qubits [NKG+22]. Moreover, we have provided a concrete example of a non-entangling positive operator-valued measure (POVM) that achieves this maximum performance. This demonstrates that it is possible to achieve the optimal result without the need for entanglement in the initial or final state of the algorithm [NKG+22].

Theorem 4.1 establishes that for states in the form (4.23), the performance of the algorithm can be given by $P(|-\rangle|\varphi\rangle) = [1 + R(|\varphi\rangle)]/2^N$. This expression remains valid regardless of whether the N-qubit state $|\varphi\rangle$ is a product state or exhibits entanglement. The results obtained so far can further be extended to mixed states, allowing for a comprehensive analysis of the algorithm's performance across a broader range of quantum states. In this regard, we have the following theorem [NKG+22].

Theorem 4.3. *For an arbitrary state σ,*

$$P(|-\rangle\langle-| \otimes \sigma) = \frac{1 + R(\sigma)}{2^N}, \tag{4.33}$$

Proof. [NKG+22] The robustness of asymmetry of a given density matrix ρ is a measure that quantifies the degree of asymmetry present in the matrix. It is the minimum amount of noise that needs to be added to ρ in order to make it symmetric which is define as below: [PCB+16]:

$$R_A(\rho) = \min_{\tau} \left\{ s \geq 0 : \frac{\rho + s\tau}{1 + s} \in \mathcal{F} \right\}, \tag{4.34}$$

In this context, the group F represents the set of symmetric states under a specific group action. The robustness of asymmetry, as discussed in [PCB+16], can be mathematically expressed through a semidefinite program (SDP). Interestingly, the resource theory of quantum coherence can be seen as a specific case of the resource theory of asymmetry when considering the symmetry associated with the $U(1)$ group [PCB+16]. This insight allows us to extend the SDP formulation originally developed for the robustness of asymmetry and adapt it to quantify the robustness of coherence. By formulating an SDP maximization problem over a variable X, as described in [PCB+16, NBC+16], we can effectively quantify the robustness of coherence and explore its properties within the framework of resource theories. The SDP can be expressed as follows:

$$R(\rho) = \max_{X} [\mathrm{Tr}(\rho X) - 1] \tag{4.35}$$

$$X \geq 0 \tag{4.36}$$

$$E(X) = \mathbb{I} \tag{4.37}$$

where $E(X) = \frac{1}{d}\sum_{\boldsymbol{a}} u_{\boldsymbol{a}} X u_{\boldsymbol{a}}^{\dagger}$ and $u_{\boldsymbol{a}} = \sum_{\boldsymbol{x}} e^{i\frac{2\pi}{d}\boldsymbol{a}\boldsymbol{x}} |\boldsymbol{x}\rangle\langle\boldsymbol{x}|$. To establish the validity of the expression in Eq. (4.33), we aim to demonstrate that $P(|-\rangle\langle-| \otimes \rho)$ is bounded from

below and above by $[1 + R(\rho)]/d$, where $d = 2^N$. As mentioned earlier, the oracle unitaries $U_{\boldsymbol{a}}$ act on states of the form $|-\rangle\langle-| \otimes \rho$ in the following manner:

$$U_{\boldsymbol{a}} |-\rangle\langle-| \otimes \rho U_{\boldsymbol{a}}^\dagger = |-\rangle\langle-| \otimes V_{\boldsymbol{a}} \rho V_{\boldsymbol{a}}^\dagger \tag{4.38}$$

with the N-qubit unitaries $V_{\boldsymbol{a}} = \otimes_{i=1}^N \sigma_{z,i}^{a_i}$. The performance of the probabilistic BV algorithm can be represented by the following expression:

$$P(|-\rangle\langle-| \otimes \rho) = \max_{\{M_{\boldsymbol{a}}\}} \frac{1}{d} \sum_{\boldsymbol{a}} \mathrm{Tr}\left(V_{\boldsymbol{a}} \rho V_{\boldsymbol{a}}^\dagger M_{\boldsymbol{a}}\right). \tag{4.39}$$

In order to verify the equation (4.33), we will first show that:

$$\max_{\{M_{\boldsymbol{a}}\}} \frac{1}{d} \sum_{\boldsymbol{a}} \mathrm{Tr}\left(V_{\boldsymbol{a}} \rho V_{\boldsymbol{a}}^\dagger M_{\boldsymbol{a}}\right) \geq \frac{1 + R(\rho)}{d}. \tag{4.40}$$

For this, we define the following operators:

$$M'_{\boldsymbol{a}} = \frac{1}{d} V_{\boldsymbol{a}} X V_{\boldsymbol{a}}^\dagger, \tag{4.41}$$

where X is the operator maximizing the SDP in Eqs. (4.37). The positivity of the operators $M'_{\boldsymbol{a}}$ follows from the positivity of X. Additionally, as $E(X) = \mathbb{I}$, we obtain:

$$\frac{1}{d} \sum_{k=1}^{d} X_{\boldsymbol{x},\boldsymbol{x}} = 1, \tag{4.42}$$

where $X_{\boldsymbol{x},\boldsymbol{y}}$ are the components of the matrix X. Therefore, the diagonal elements of X are all equal to 1. It is worth noting that $\frac{1}{d} \sum_{\boldsymbol{a}} e^{i \frac{2\pi}{D} \boldsymbol{a} \cdot (\boldsymbol{x} - \boldsymbol{x}')} = \delta_{\boldsymbol{x},\boldsymbol{x}'}$ due to the orthogonality of the exponentials, and since $V_{\boldsymbol{a}}$ is a diagonal unitary, we can write:

$$\begin{aligned} \sum_{\boldsymbol{a}} M'_{\boldsymbol{a}} &= \frac{1}{d} \sum_{f} V_{\boldsymbol{a}} X V_{\boldsymbol{a}}^\dagger \\ &= \sum_{\boldsymbol{x},\boldsymbol{y}} \frac{1}{d} \sum_{\boldsymbol{a}} e^{i \frac{2\pi}{D} \boldsymbol{a} \cdot (\boldsymbol{x} - \boldsymbol{y})} X_{\boldsymbol{x},\boldsymbol{y}} |\boldsymbol{x}\rangle\langle\boldsymbol{y}| = \mathbb{I}. \end{aligned} \tag{4.43}$$

Therefore $\{M'_{\boldsymbol{a}}\}$ forms a POVM. Thus, we must have

$$\begin{aligned} \frac{1}{d} \sum_{\boldsymbol{a}} \mathrm{Tr}\left(V_{\boldsymbol{a}} \rho V_{\boldsymbol{a}}^\dagger M'_{\boldsymbol{a}}\right) &= \frac{1}{d} \sum_{\boldsymbol{a}} \frac{1}{d} \mathrm{Tr}\left(V_{\boldsymbol{a}} \rho V_{\boldsymbol{a}}^\dagger V_{\boldsymbol{a}} X V_{\boldsymbol{a}}^\dagger\right) \\ &= \frac{1}{d} \mathrm{Tr}(\rho X) = \frac{1 + R(\rho)}{d} \\ &\leq \max_{\{M_{\boldsymbol{a}}\}} \frac{1}{d} \sum_{\boldsymbol{a}} \mathrm{Tr}\left(V_{\boldsymbol{a}} \rho V_{\boldsymbol{a}}^\dagger M_{\boldsymbol{a}}\right). \end{aligned} \tag{4.44}$$

Above we used the definition of the robustness of coherence to derive the expression $\mathrm{Tr}(\rho X) = 1 + R(\rho)$.

We will now indicate that the performance is also bounded above by $(1 + R(\rho))/d$. This can be inferred from the definition of the robustness of coherence, which implies that

$$\rho = [1 + R(\rho)]\sigma - R(\rho)\tau, \tag{4.45}$$

with some quantum state τ and some incoherent state σ. Thus, for an arbitrary POVM $\{M_a\}$ we have,

$$\begin{aligned}\sum_a \mathrm{Tr}\left(V_a \rho V_a^\dagger M_a\right) &= [1 + R(\rho)] \sum_a \mathrm{Tr}\left(V_a \sigma V_a^\dagger M_a\right) \\ &\quad - R(\rho) \sum_a \mathrm{Tr}\left(V_a \tau V_a^\dagger M_a\right) \\ &\leq [1 + R(\rho)] \sum_a \mathrm{Tr}\left(V_a \sigma V_a^\dagger M_a\right).\end{aligned} \tag{4.46}$$

Since σ is an incoherent state, we have $V_a \sigma V_a^\dagger = \sigma$, which implies that

$$\sum_a \mathrm{Tr}\left(V_a \sigma V_a^\dagger M_a\right) = \mathrm{Tr}\left(\sigma \sum_a M_a\right) = 1. \tag{4.47}$$

Hence, we can derive the inequality

$$\max_{\{M_a\}} \frac{1}{d} \sum_a \mathrm{Tr}\left(V_a \rho V_a^\dagger M_a\right) \leq \frac{1 + R(\rho)}{d}. \tag{4.48}$$

From Eqs. (4.44) and (4.48) we have

$$\max_{\{M_a\}} \frac{1}{d} \sum_a \mathrm{Tr}\left(V_a \rho V_a^\dagger M_a\right) = \frac{1 + R(\rho)}{d}. \tag{4.49}$$

This completes the proof of Eq. (4.33).

□

The relation given in Eq. (4.33) provides a meaningful operational interpretation for the robustness of coherence of a quantum state ρ in the context of quantum computation. It implies that the robustness of coherence directly impacts the performance of the probabilistic Bernstein-Vazirani algorithm. The higher the robustness of coherence, we will have the greater potential for successful computation using the algorithm. This observation highlights the significance of coherence as a valuable resource for quantum computational tasks and demonstrates the intimate connection between coherence and the computational capabilities of quantum systems [NKG+22].

It is worth noting that, in general, the POVM that achieves the optimal performance for mixed states, as expressed in Eq. (4.33), may require the implementation of entanglement [NKG+22].

In the following section we study the role of purity as a resource in the algorithm.

4.3 Purity in Probabilistic Bernstein-Vazirani Algorithm

In the subsequent analysis, we aim to determine the optimal performance of the Bernstein-Vazirani (BV) algorithm when subjected to noise [NKG+22]. To achieve this, we consider a scenario where the initial state of the algorithm exhibits a bounded purity. Specifically, we impose the condition $Tr(\rho^2) \leq \gamma$, which places a restriction on the purity of the state. In light of this constraint, we can characterize the optimal initial states through the theorem provided below [NKG+22].

Theorem 4.4. *Having the oracle register in the state $|-\rangle$, the optimal initial state of the N-qubit system maximizing the performance of the BV algorithm with bounded purity $\mathrm{Tr}[\rho^2] \leq \gamma$ is given by*

$$\rho_{\max,\gamma} = \frac{d}{2\lambda_1} |\psi_{MC}\rangle\langle\psi_{MC}| + \frac{\lambda_2}{2\lambda_1}\mathbb{I} \tag{4.50}$$

with

$$\lambda_1 = \frac{d\sqrt{1-\frac{1}{d}}}{2\sqrt{\gamma-\frac{1}{d}}}, \qquad \lambda_2 = \frac{\sqrt{1-\frac{1}{d}}}{\sqrt{\gamma-\frac{1}{d}}} - 1, \tag{4.51}$$

and $|\psi_{MC}\rangle$ being a maximally coherent state in the computational basis. The optimal performance in this case is given as

$$P(\rho_{\max,\gamma}) = \frac{1}{d} + \frac{d-1}{2\lambda_1} \tag{4.52}$$

where $d = 2^N$.

Proof. [NKG+22] In order to prove the results outlined in the Theorem 4.4, we begin by proving that for any pseudo-pure maximally coherent state, the robustness of coherence is equivalent to the ℓ_1-norm of coherence. The ℓ_1-norm of coherence, denoted as $C_{\ell_1}(\rho)$, can be expressed as the sum of absolute values of off-diagonal elements of the density matrix ρ [BCP14b]. Let us consider the pseudo-pure maximally coherent state denoted as ρ_s, which we define it as follows:

$$\rho_s = p\,|\psi_{\mathrm{MC}}\rangle\,\langle\psi_{\mathrm{MC}}| + (1-p)\frac{\mathbb{I}}{d}, \tag{4.53}$$

in which d is the dimension of the Hilbert space, $0 \leq p \leq 1$. For $C_{\ell_1}(\rho_s)$we have:

$$C_{\ell_1}(\rho_s) = \sum_{x,y,x\neq y} |\rho_{s,xy}| = p(d-1). \tag{4.54}$$

Here, $\rho_{s,xy}$ represents the elements of the density matrix of ρ_s in the basis $|x\rangle$. To evaluate the robustness of coherence $R(\rho_s)$, we utilize the semidefinite program (SDP)

formulation of the robustness of coherence. Therefore we have:

$$R(\rho_s) = \max_X \mathrm{Tr}(\rho_s X) - 1 \tag{4.55}$$
$$= \left[p \langle \psi_{\mathrm{MC}} | X^* | \psi_{\mathrm{MC}} \rangle + (1-p)\frac{1}{d}\mathrm{Tr}(X^*) \right] - 1,$$

where X^* is the matrix maximizing the SDP. Based on Equation (4.42), we observe that the diagonal elements of the matrix X are identical and equal to 1. Furthermore, since X is a positive matrix, we can represent it as $X^* \equiv d\rho_x$, where ρ_x denotes a quantum state. This enables us to derive the following expression:

$$\max_X \mathrm{Tr}(\rho_s X) - 1 = dp \langle \psi_{\mathrm{MC}} | \rho_x | \psi_{\mathrm{MC}} \rangle - p. \tag{4.56}$$

The constraint on X^* only requires that the diagonal elements of ρ_x are equal. In Equation (4.56), the term $\langle \psi_{\mathrm{MC}} | \rho_x | \psi_{\mathrm{MC}} \rangle$ is maximized when we consider ρ_x to be the maximally coherent state $|\psi_{\mathrm{MC}}\rangle\langle\psi_{\mathrm{MC}}|$, which satisfies the requirement for X^*. Therefore, we can conclude that $X^* = d\,|\psi_{\mathrm{MC}}\rangle\langle\psi_{\mathrm{MC}}|$ maximizes the semidefinite program (SDP) for ρ_s. Thus

$$R(\rho_s) = p(d-1) = C_{\ell_1}(\rho_s). \tag{4.57}$$

We proceed by applying the Lagrange multipliers method to maximize the ℓ_1-norm of coherence while considering a fixed amount of purity. The purity of the density matrix ρ can be expressed in terms of the absolute values of its components $|\rho_{i,j}|$ as shown below:

$$\mathrm{Tr}(\rho^2) = \sum_{i,j} |\rho_{i,j}|^2. \tag{4.58}$$

The objective is to maximize the ℓ_1-norm of coherence while maintaining a fixed purity of γ. Instead of directly maximizing C_{ℓ_1}, we maximize the function $g = \sum_{i,j} |\rho_{i,j}| = C_{\ell_1} + 1$ under the constraint $\sum_{i=j} \rho_{i,j} = 1$. This constraint ensures that we are effectively maximizing the C_{ℓ_1} function.

Therefore, our maximization problem is as follows:

- Constraint 1: $C_1 = \sum_{i,j} |\rho_{i,j}|^2 - \gamma = 0$.
- Constraint 2: $C_2 = \sum_{i=j} \rho_{i,j} - 1 = 0$.
- λ_1 and λ_2 are the Lagrange multipliers corresponding to C_1 and C_2 constraints respectively.
- The function $g = \sum_{i,j} |\rho_{i,j}|$ is the objective function that we aim to maximize with respect to the variables $|\rho_{i,j}|$.

It is important to note that we have two additional constraints to consider: the Hermiticity and positivity of the density matrix ρ. Although these constraints are not incorporated directly during the maximization process, it is essential to verify them for the final maximizing state to ensure its validity.

By applying the Lagrange multipliers method, we derive the following set of equations:

$$\frac{dg}{d|\rho_{i,j}|} - \lambda_1 \frac{dC_1}{d|\rho_{i,j}|} - \lambda_2 \frac{dC_2}{d|\rho_{i,j}|} = 0. \tag{4.59}$$

Simplifying these equations and taking into account the constraints, we obtain the following set of equations:

$$\text{for } i \neq j, 1 - 2\lambda_1 |\rho_{i,j}| = 0, \tag{4.60}$$

$$\text{for } i = j = k, 1 - 2\lambda_1 |\rho_{k,k}| + \lambda_2 = 0, \tag{4.61}$$

$$C_1 = \sum_{i,j} |\rho_{i,j}|^2 - \gamma = 0, \tag{4.62}$$

$$C_2 = \sum_{i=j} |\rho_{i,j}| - 1 = 0. \tag{4.63}$$

Solving these equations for $|\rho_{i,j}|$, λ_1 and λ_2 results in

$$i \neq j, |\rho_{i,j}| = \frac{1}{2\lambda_1}, \tag{4.64}$$

$$i = j = k, |\rho_{k,k}| = \frac{1 + \lambda_2}{2\lambda_1}. \tag{4.65}$$

Considering a d-dimensional system, we obtain,

$$\lambda_2 = \frac{\sqrt{1 - \frac{1}{d}}}{\sqrt{\gamma - \frac{1}{d}}} - 1, \tag{4.66}$$

$$\lambda_1 = \frac{d\sqrt{1 - \frac{1}{d}}}{2\sqrt{\gamma - \frac{1}{d}}}. \tag{4.67}$$

Since $|\rho_{i,j}| = |\rho_{j,i}|$ and we have the freedom to choose the phases in $\rho_{i,j} = |\rho_{i,j}|e^{i\varphi_{i,j}}$, we can select $\varphi_{i,j}$ such that $\rho_{\max,\gamma}$ becomes Hermitian.

Using the obtained values of $\rho_{i,j}$, we can represent the maximizing state $\rho_{\max,\gamma}$ in the following form:

$$\rho_{\max,\gamma} = \frac{d}{2\lambda_1} |\psi_{\text{MC}}\rangle\langle\psi_{\text{MC}}| + \frac{\lambda_2}{2\lambda_1}\mathbb{I}. \tag{4.68}$$

As $\lambda_1, \lambda_2 \geq 0$ and $\frac{d}{2\lambda_1} + \frac{d\lambda_2}{2\lambda_1} = 1$, the state $\rho_{\max,\gamma}$ is a valid density matrix that maximizes the ℓ_1-norm of coherence under the constraint of bounded purity γ. Additionally, the maximum amount of coherence achieved is

$$C_{\ell_1,\max} = \frac{d^2 - d}{2\lambda_1}. \tag{4.69}$$

Since we have previously shown that the ℓ_1-norm of coherence is equal to the robustness of coherence for any pseudo-pure maximally coherent state, and that $R(\rho) \leq C_{\ell_1}(\rho)$ for any ρ [PCB+16, NBC+16], we can conclude that the state $\rho_{\max,\gamma}$ given by Eq. (4.68) also maximizes the robustness of coherence with the same value as $C_{\ell_1,\max}$, subject to the constraint of purity $\mathrm{Tr}(\rho^2) = \gamma$:

$$R_{\max} = \frac{d(d-1)}{2\lambda_1}. \tag{4.70}$$

□

As we can see, the state that maximizes the robustness of coherence for a given bounded purity is a pseudo pure state, which can be prepared using NMR techniques [LP01, CFH97, SHC00]. Furthermore, if we choose $|\psi_{\mathrm{MC}}\rangle$ to be a product state, the resulting state $\rho_{\max,\gamma}$ will not exhibit entanglement. Since the performance $P(\rho)$ is monotonically related to the robustness of coherence, we can conclude that NMR quantum computing is a suitable platform for implementing the probabilistic BV algorithm.

In the next section we focus on the connection between multipartite entanglement and coherence and we will discuss how excessive entanglement can be detrimental to the algorithm's performance.

4.4 Multipartite Entanglement and Coherence

As we have observed, achieving optimal performance in the probabilistic BV algorithm does not necessitate the presence of entanglement. In fact, by using the initial state $|-\rangle|+\rangle^{\otimes N}$, it is possible to perfectly learn the bit string $\boldsymbol{a}$ with just a single application of the oracle unitary. However, we now aim to investigate more the role of multipartite entanglement among the N system qubits and its influence on the algorithm's performance. Specifically, we will investigate the relationship between the robustness of coherence and geometric entanglement for N-qubit systems. Surprisingly, our analysis reveals that a significant amount of geometric entanglement in the system state can actually be detrimental to the performance of the algorithm [NKG+22].

For this, we focus on N-qubit W-states [DVC00]:

$$|\Psi_W\rangle = \frac{1}{\sqrt{N}}(e^{i\varphi_1}|\psi_1\rangle|\psi_2\rangle\ldots|\psi_N^{\perp}\rangle + e^{i\varphi_2}|\psi_1\rangle\ldots|\psi_{N-1}^{\perp}\rangle|\psi_N\rangle + \ldots + e^{i\varphi_N}|\psi_1^{\perp}\rangle|\psi_2\rangle\ldots|\psi_N\rangle), \tag{4.71}$$

where $|\psi_i\rangle$ and $|\psi_i^{\perp}\rangle$ are orthogonal. It is worth noting that for $N = 3$, the W-states are the only type of states that can achieve maximal geometric entanglement [TWP09]. We have the following theorem regarding the connection of the W-state class of entanglement and coherence [NKG+22].

Theorem 4.5. *W-states with the number of particles $N \geq 3$ can never be a maximally coherent state.*

Proof. [NKG+22] Let us consider the most general form of a W-state in a given N-qubit system. [DVC00]:

$$\begin{aligned}|\Psi_W\rangle = \frac{1}{\sqrt{N}}(e^{i\varphi_1}|\psi_1\rangle|\psi_2\rangle..|\psi_N^{\perp}\rangle + e^{i\varphi_2}|\psi_1\rangle..|\psi_{N-1}^{\perp}\rangle|\psi_N\rangle \\ +... + e^{i\varphi_N}|\psi_1^{\perp}\rangle|\psi_2\rangle..|\psi_N\rangle) \\ = \frac{1}{\sqrt{N}}[|\psi_1\rangle(e^{i\varphi_1}|\psi_2\rangle..|\psi_N^{\perp}\rangle + .. + e^{i\varphi_{N-1}}|\psi_2^{\perp}\rangle..|\psi_N\rangle) \\ + |\psi_1^{\perp}\rangle e^{i\varphi_N}|\psi_2\rangle..|\psi_N\rangle]\end{aligned} \tag{4.72}$$

with $\{|\psi_i\rangle, |\psi_i^{\perp}\rangle\}_{i=1}^{N}$ forming a basis and φ_i are some phases. Now, we define

$$\begin{aligned}|\Phi\rangle_{N-1} &= \frac{1}{\sqrt{N-1}}\left(e^{i\varphi_1}|\psi_2\rangle\ldots|\psi_N^{\perp}\rangle + \ldots + e^{i\varphi_{N-1}}|\psi_2^{\perp}\rangle\ldots|\psi_N\rangle\right), \\ |\Phi^{\perp}\rangle_{N-1} &= e^{i\varphi_N}|\psi_2\rangle..|\psi_N\rangle.\end{aligned} \tag{4.73}$$

Note that the state $|\Phi^{\perp}\rangle_{N-1}$ is orthogonal to the state $|\Phi\rangle_{N-1}$. This implies that the state $|\Psi_W\rangle$ can be written as:

$$|\Psi_W\rangle = \frac{\sqrt{N-1}}{\sqrt{N}}|\psi_1\rangle|\Phi\rangle_{N-1} + \frac{1}{\sqrt{N}}|\psi_1^{\perp}\rangle|\Phi^{\perp}\rangle_{N-1}. \tag{4.74}$$

Substituting

$$|\psi_1\rangle = a|0\rangle + b|1\rangle, \tag{4.75}$$

$$|\psi_1^{\perp}\rangle = b^*|0\rangle - a^*|1\rangle \tag{4.76}$$

with $|a|^2 + |b|^2 = 1$ we further obtain

$$\begin{aligned}|\Psi_W\rangle = |0\rangle\left(\frac{\sqrt{N-1}}{\sqrt{N}}a|\Phi\rangle_{N-1} + \frac{1}{\sqrt{N}}b^*|\Phi^{\perp}\rangle_{N-1}\right) \\ + |1\rangle\left(\frac{\sqrt{N-1}}{\sqrt{N}}b|\Phi\rangle_{N-1} - \frac{1}{\sqrt{N}}a^*|\Phi^{\perp}\rangle_{N-1}\right).\end{aligned} \tag{4.77}$$

For any N-qubit maximally coherent state $|\psi_{\max,N}\rangle$, it can be represented as:

$$|\psi_{\mathrm{MC},N}\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|\psi_{\mathrm{MC},N-1}\rangle + |1\rangle|\psi'_{\mathrm{MC},N-1}\rangle\right), \tag{4.78}$$

where $|\psi_{\mathrm{MC},N-1}\rangle$ and $|\psi'_{\mathrm{MC},N-1}\rangle$ are $(N-1)$-qubit maximally coherent states. Comparing Eqs. (4.77) and (4.78), it can be observed that in order for a W-state to be maximally coherent, it is necessary for the (unnormalized) states $\frac{\sqrt{N-1}}{\sqrt{N}}a|\Phi\rangle_{N-1} + \frac{1}{\sqrt{N}}b^*|\Phi^{\perp}\rangle_{N-1}$

and $\frac{\sqrt{N-1}}{\sqrt{N}} b\,|\Phi\rangle_{N-1} - \frac{1}{\sqrt{N}} a^*\,|\Phi^{\perp}\rangle_{N-1}$ to have equal norms. By evaluating the norms of these states and setting them equal to each other, we obtain:

$$(N-1)|a|^2 + |b|^2 = (N-1)|b|^2 + |a|^2. \tag{4.79}$$

This equation implies that $|a|^2 = |b|^2$. Similar reasoning can be applied to other qubits in the W-state. Therefore, in order for $|\Psi_W\rangle$ to be a maximally coherent state, it is necessary to satisfy:

$$|\psi_i\rangle = \frac{|0\rangle + e^{i\theta_i}\,|1\rangle}{\sqrt{2}}, \tag{4.80}$$

$$|\psi_i^{\perp}\rangle = \frac{|0\rangle - e^{i\theta_i}\,|1\rangle}{\sqrt{2}} \tag{4.81}$$

with some phases θ_i. Now we consider an N-qubit W-state given by:

$$|W\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} e^{i\varphi_j}\,|(+)^{N-1}, (-)_j\rangle \tag{4.82}$$

with $|(+)^{N-1}, (-)_j\rangle = |+\rangle_1\,|+\rangle_2 \ldots |-\rangle_j \ldots |+\rangle_N$ and $|\pm\rangle_j = \frac{|0\rangle \pm e^{i\theta_j}|1\rangle}{\sqrt{2}}$. We also define $|1_k\rangle = e^{i\theta_k}\,|1\rangle$ and $|0_k\rangle = |0\rangle$. Our goal is to assess whether the state $|W\rangle$ qualifies as a maximally coherent state in the computational basis. If it meets the criteria of a maximally coherent state, we would expect all the states in the computational basis to have equal probabilities. Since the vectors $|(0)^{N-1}, 1_j\rangle = |0\rangle\,1\,|0\rangle\,2 \ldots |1\rangle\,j \ldots |0\rangle\,N$ represent the same state in the computational basis but with different phases, we expect that if $|W\rangle$ is a maximally coherent state, these vectors should have coefficients with the magnitude of $\frac{1}{\sqrt{2^N}}$ when $|W\rangle$ is expanded in the computational basis. Expanding $|W\rangle$ in the computational basis, we denote the coefficient of the state $|\boldsymbol{x}\rangle$ with $f(|\boldsymbol{x}\rangle)$ i.e. $|W\rangle = \sum_{\boldsymbol{x}} f(|\boldsymbol{x}\rangle)\,|x\rangle$. Let us first evaluate $f(|(0)^{N-1}, 1_k\rangle)$ and $f(|0\rangle^{\otimes N})$:

$$f(|0\rangle^{\otimes N}) = \frac{1}{\sqrt{N2^N}} \sum_{j=1}^{N} e^{i\varphi_j}, \tag{4.83}$$

$$f(|(0)^{N-1}, 1_k\rangle) = \frac{1}{\sqrt{N2^N}} \Big(\sum_{j=1}^{N} e^{i\varphi_j} - 2e^{i\varphi_k}\Big). \tag{4.84}$$

If $|W\rangle$ is a maximally coherent state, we expect the following condition to hold:

$$f(|0\rangle^{\otimes N}) = \frac{1}{\sqrt{N2^N}} \sum_{j=1}^{N} e^{i\varphi_j} = \frac{e^{i\alpha_0}}{\sqrt{2^N}}, \tag{4.85}$$

$$f(|(0)^{N-1}, 1_k\rangle) = \frac{1}{\sqrt{N2^N}} \left(\sum_{j=1}^{N} e^{i\varphi_j} - 2e^{i\varphi_k}\right) = \frac{e^{i\alpha_k}}{\sqrt{2^N}}. \tag{4.86}$$

For some α_k and α_0. Since only the relative phases are important, we can choose $\alpha_0 = 0$, resulting in the following equations:

$$\sum_{j=1}^{N} e^{i\varphi_j} = \sqrt{N}, \tag{4.87}$$

$$\sum_{j=1}^{N} e^{i\varphi_j} - 2e^{i\varphi_k} = \sqrt{N}e^{i\alpha_k}. \tag{4.88}$$

After solving this set of equations for all φ_k, we can determine the values of the phases. By substituting the first equation into the second equation and simplifying, we obtain:

$$\begin{aligned} \sqrt{N}\frac{1-e^{i\alpha_k}}{2} &= e^{i\varphi_k} \\ \iff \sqrt{\frac{N}{2}}\sqrt{1-\cos\alpha_k}e^{i\arctan\left(\frac{-\sin\alpha_k}{1-\cos\alpha_k}\right)} &= e^{i\varphi_k} \end{aligned} \tag{4.89}$$

These equations imply that $\cos\alpha_k = 1 - \frac{2}{N}$ and

$$\varphi_k = \pm\arctan\frac{|\sin\alpha_k|}{|1-\cos\alpha_k|} = \pm\arctan\sqrt{N-1}. \tag{4.90}$$

Now, let us calculate $f(|0\rangle^{\otimes N-2}\otimes|1\rangle\otimes|1\rangle)$:

$$f(|0\rangle^{\otimes N-2}\otimes|1\rangle\otimes|1\rangle) = \frac{1}{\sqrt{N2^N}}\left(\sum_{j=1}^{N} e^{i\varphi_j} - 2e^{i\varphi_{N-1}} - 2e^{i\varphi_N}\right). \tag{4.91}$$

If the $|W\rangle$ state is a maximally coherent state, then the coefficient $\frac{1}{\sqrt{2^{N-1}}}$ in front of each computational basis state $|(0)^{N-1}, 1_j\rangle$ should also be equal to $\frac{e^{i\theta}}{\sqrt{2^N}}$ for some phase θ. Substituting $\varphi_k = \pm\arctan\sqrt{N-1}$ and $\sum_{j=1}^{N} e^{i\varphi_j} = \sqrt{N}$ in Eq. (4.91) and equating the coefficient with $\frac{e^{i\theta}}{\sqrt{2^N}}$, we have:

$$\frac{1}{\sqrt{2^N}}\left[1 - \frac{2}{\sqrt{N}}\left(e^{\pm i\arctan\sqrt{N-1}} + e^{\pm i\arctan\sqrt{N-1}}\right)\right] = \frac{e^{i\theta}}{\sqrt{2^N}}. \tag{4.92}$$

Simplifying the above equation, we obtain:

$$1 - \frac{2}{\sqrt{N}}\left(e^{\pm i\arctan\sqrt{N-1}} + e^{\pm i\arctan\sqrt{N-1}}\right) = e^{i\theta}. \tag{4.93}$$

The last equation does not have any solutions for $N \in \mathbb{N}$ and $N > 2$. This implies that the magnitudes of the coefficients of the states $|0\rangle^{\otimes N-2}\otimes|1,1\rangle$ and $|(0)^{N-1}, 1_k\rangle$ when we expand the $|W\rangle$ state in the computational basis cannot be the same and equal to $\frac{1}{\sqrt{2^N}}$. Hence, it is clear that a W-state cannot be maximally coherent for $N > 2$.

□

This observation implies that when considering initial states of the form $|-\rangle|\psi\rangle$, there exists a threshold on the geometric entanglement of $|\psi\rangle$. Above this threshold, it is not possible to achieve the optimal performance $P = 1$ in the probabilistic BV algorithm [NKG+22]. This result shares a similar spirit with the findings presented in [GFE09], which demonstrate that quantum states can possess too much entanglement to be effectively utilized for quantum computation.

We conclude this chapter by generalizing the BV algorithm for an arbitrary number of qudits.

4.5 Probabilistic BV Algorithm for Qudits

In the BV algorithm using qudits, the objective is to determine the string $\boldsymbol{k}$ with $k_i \in 1, ..., D$ encoded in the linear function [NKG+22].

$$f(\boldsymbol{x}) = \boldsymbol{k} \cdot \boldsymbol{x} \bmod D = \sum_{i=1}^{N} k_i x_i \bmod D, \tag{4.94}$$

where $\boldsymbol{x} = (x_1, x_2, ..., x_N)$ with $x_i \in 1, ..., D$. Similarly to the qubit version of the algorithm, we assume that the function is encoded into an oracle unitary denoted as $U_{\boldsymbol{a}}$, which acts as [NKG+22]

$$U_{\boldsymbol{k}}|j\rangle|\boldsymbol{x}\rangle = |j + f(\boldsymbol{x}) \bmod D\rangle|\boldsymbol{x}\rangle \tag{4.95}$$

with $j \in \{0, ..., D-1\}$. As we will see shortly, there will be no entanglement between the oracle register and the qudit system when the oracle unitary is applied to a state of the form $|-_D\rangle|\varphi\rangle$, where

$$|-_D\rangle = \frac{1}{\sqrt{D}} \sum_{k=0}^{D-1} e^{-i\frac{2\pi}{D}k}|k\rangle, \tag{4.96}$$

and $|\varphi\rangle$ is a product state of N qudits. Furthermore, we can evaluate the performance of the protocol when it is applied to states of the form $|-_D\rangle\langle-_D| \otimes \rho$ as follows [NKG+22]:

$$P(|-_D\rangle\langle-_D| \otimes \rho) = \frac{1 + R(\rho)}{d}, \tag{4.97}$$

where $R(\rho)$ is the robustness of coherence in the computational basis and $d = D^N$.

By utilizing the relationship described in Eq. (4.97), it can be shown that when the state of the oracle register is $|-_D\rangle$, the optimal state of the system qudits that maximizes the performance $P(\rho)$ with a bounded purity constraint of $\mathrm{Tr}(\rho) = \gamma$ takes the following form [NKG+22]:

$$\rho = p|\psi_{\mathrm{MC}}\rangle_D\langle\psi_{\mathrm{MC}}| + (1-p)\frac{\mathbb{I}}{D^N} \tag{4.98}$$

with $0 \leq p \leq 1$ and $|\psi_{\mathrm{MC}}\rangle_D$ is an N-qudit maximally coherent state. The proofs for the aforementioned results in the case of qudits are very similar to the proofs presented earlier for qubits [NKG+22].

First, we will present a proof for the claim that in the probabilistic BV algorithm with qudits, if the oracle register is initially prepared in the state $|-_D\rangle$, there will be no entanglement between the oracle register and the system qudits after the action of the oracle. Additionally, we will show that the unitary $U_{\boldsymbol{k}}$ acts as a non-entangling gate [NKG+22].

Consider an arbitrary vector $|\boldsymbol{x}\rangle$ in the computational basis. We can write:

$$U_{\boldsymbol{k}} |-_D\rangle |\boldsymbol{x}\rangle = \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} e^{-i\frac{2\pi}{D}j} U_{\boldsymbol{k}} |j\rangle |\boldsymbol{x}\rangle . \tag{4.99}$$

The oracle unitary acts as $U_{\boldsymbol{k}} |j\rangle |\boldsymbol{x}\rangle = |j + f(\boldsymbol{x}) \bmod D\rangle |\boldsymbol{x}\rangle$ with $f(\boldsymbol{x}) = \boldsymbol{k} \cdot \boldsymbol{x} = \sum_{i=1}^{N} k_i x_i$. Thus, we can write:

$$\begin{aligned} U_{\boldsymbol{k}} |-_D\rangle |\boldsymbol{x}\rangle &= \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} e^{-i\frac{2\pi}{D}j} |j + f(\boldsymbol{x}) \bmod D\rangle |\boldsymbol{x}\rangle \\ &= e^{i\frac{2\pi}{D}f(\boldsymbol{x})} \frac{1}{\sqrt{D}} \sum_{j=0}^{D-1} e^{-i\frac{2\pi}{D}j} U_{\boldsymbol{k}} |j\rangle |\boldsymbol{x}\rangle \\ &= e^{i\frac{2\pi}{D}f(\boldsymbol{x})} |-_D\rangle |\boldsymbol{x}\rangle . \end{aligned} \tag{4.100}$$

Thus, we observe that if the state of the oracle register is $|-D\rangle$, the action of the oracle unitary $U\boldsymbol{k}$ corresponds to a non-entangling unitary transformation on the system qudits.

Now we prove the result stated in Eq. (4.97) [NKG+22]. Similar to the qubit case, we can evaluate the robustness of coherence using a semidefinite program, as shown in Eqs. (4.37). However, In contrast to the qubit setting, we have $E(X) = \frac{1}{d} \sum_{\boldsymbol{k}} u_{\boldsymbol{k}} X u_{\boldsymbol{k}}^{\dagger}$ and $u_{\boldsymbol{k}} = \sum_x e^{i\frac{2\pi}{d}\boldsymbol{k}\boldsymbol{x}} |\boldsymbol{x}\rangle\langle\boldsymbol{x}|$.

In order to establish the expression in Eq. (4.97), we will demonstrate that $P(|-_D\rangle\langle-_D| \otimes \rho)$ is bounded both from below and from above by $[1 + R(\rho)]/d$. As previously discussed, when the oracle register is in the state $|-D\rangle$, the action of the oracle unitaries $U\boldsymbol{k}$ on states of the form $|-_D\rangle\langle-_D| \otimes \rho$ can be described as follows:

$$U_{\boldsymbol{k}} |-_D\rangle\langle-_D| \otimes \rho U_{\boldsymbol{k}} = |-_D\rangle\langle-_D| \otimes V_{\boldsymbol{k}} \rho V_{\boldsymbol{k}}^{\dagger} \tag{4.101}$$

with the N-qudit unitaries

$$V_{\boldsymbol{k}} = \sum_{j=0}^{D-1} e^{i\frac{2\pi}{D}k_1 j} |j\rangle\langle j| \otimes \sum_{j=0}^{D-1} e^{i\frac{2\pi}{D}k_2 j} |j\rangle\langle j| \otimes \ldots \otimes \sum_{j=0}^{D-1} e^{i\frac{2\pi}{D}k_N j} |j\rangle\langle j| . \tag{4.102}$$

In case of qubits this unitary will become $V_{\boldsymbol{k}} = \otimes_{i=1}^{N} \sigma_{z,i}^{k_i}$ and $k_i \in \{0, 1\}$ [KMS16]. The performance of the probabilistic BV algorithm can be represented by the following expression:

$$P(|-_D\rangle\langle -_D| \otimes \rho) = \max_{\{M_k\}} \frac{1}{d} \sum_{\boldsymbol{k}} \mathrm{Tr}\left(V_{\boldsymbol{k}} \rho V_{\boldsymbol{k}}^{\dagger} M_{\boldsymbol{k}}\right). \tag{4.103}$$

Next, we will indicate that:

$$\max_{\{M_k\}} \frac{1}{d} \sum_{\boldsymbol{k}} \mathrm{Tr}\left(V_{\boldsymbol{k}} \rho V_{\boldsymbol{k}}^{\dagger} M_{\boldsymbol{k}}\right) \geq \frac{1 + R(\rho)}{d}. \tag{4.104}$$

To accomplish this, we introduce the following operators:

$$M'_{\boldsymbol{k}} = \frac{1}{d} V_{\boldsymbol{k}} X V_{\boldsymbol{k}}^{\dagger}. \tag{4.105}$$

It should be noted that the operator X is the one that maximizes the SDP in Eqs. (4.37). Additionally, the operators $M'_{\boldsymbol{k}}$ are positive because X is positive. Moreover, since $E(X) = \mathbb{I}$, we obtain the following relation:

$$\frac{1}{d} \sum_{k=1}^{d} X_{\boldsymbol{x},\boldsymbol{x}} = 1 \tag{4.106}$$

Here, $X_{\boldsymbol{x},\boldsymbol{y}}$ denotes the components of the matrix X. Consequently, the diagonal elements of X are identical and equal to 1, which can be expressed as $X_{\boldsymbol{x},\boldsymbol{x}} = 1$ for all $\boldsymbol{x}$. Note that as $\frac{1}{d} \sum_{\boldsymbol{k}} e^{i \frac{2\pi}{D} \boldsymbol{k} \cdot (\boldsymbol{x} - \boldsymbol{x}')} = \delta_{\boldsymbol{x},\boldsymbol{x}'}$ and $V_{\boldsymbol{k}}$ are diagonal unitaries, we have

$$\begin{aligned} \sum_{\boldsymbol{k}} M'_{\boldsymbol{k}} &= \frac{1}{d} \sum_{f} V_{\boldsymbol{k}} X V_{\boldsymbol{k}}^{\dagger} \\ &= \sum_{\boldsymbol{x},\boldsymbol{y}} \frac{1}{d} \sum_{\boldsymbol{k}} e^{i \frac{2\pi}{D} \boldsymbol{k} \cdot (\boldsymbol{x} - \boldsymbol{y})} X_{\boldsymbol{x},\boldsymbol{y}} |\boldsymbol{x}\rangle\langle \boldsymbol{y}| = \mathbb{I}. \end{aligned} \tag{4.107}$$

Therefore, the operators $M'_{\boldsymbol{k}}$ form a valid set of positive operator-valued measures (POVM). As a result, we must have:

$$\begin{aligned} \frac{1}{d} \sum_{\boldsymbol{k}} \mathrm{Tr}\left(V_{\boldsymbol{k}} \rho V_{\boldsymbol{k}}^{\dagger} M'_{\boldsymbol{k}}\right) &= \frac{1}{d} \sum_{\boldsymbol{k}} \frac{1}{d} \mathrm{Tr}\left(V_{\boldsymbol{k}} \rho V_{\boldsymbol{k}}^{\dagger} V_{\boldsymbol{k}} X V_{\boldsymbol{k}}^{\dagger}\right) \\ &= \frac{1}{d} \mathrm{Tr}(\rho X) = \frac{1 + R(\rho)}{d} \\ &\leq \max_{\{M_k\}} \frac{1}{d} \sum_{\boldsymbol{k}} \mathrm{Tr}\left(V_{\boldsymbol{k}} \rho V_{\boldsymbol{k}}^{\dagger} M_{\boldsymbol{k}}\right). \end{aligned} \tag{4.108}$$

Above we applied the definition of the robustness of coherence to derive the equation $\mathrm{Tr}(\rho X) = 1 + R(\rho)$.

Now, we will prove that the performance is also upper bounded by $(1 + R(\rho))/d$. Based on the definition of the robustness of coherence, we can deduce that

$$\rho = [1 + R(\rho)]\sigma - R(\rho)\tau, \tag{4.109}$$

with some quantum state τ and some incoherent state σ. Therefore, for any POVM $M_{\boldsymbol{a}}$, we have the following inequality:

$$\begin{aligned}\sum_{k} \mathrm{Tr}\left(V_k \rho V_k^\dagger M_k\right) &= [1 + R(\rho)] \sum_{k} \mathrm{Tr}\left(V_k \sigma V_k^\dagger M_k\right) \\ &\quad - R(\rho) \sum_{k} \mathrm{Tr}\left(V_k \tau V_k^\dagger M_k\right) \\ &\leq [1 + R(\rho)] \sum_{k} \mathrm{Tr}\left(V_k \sigma V_k^\dagger M_k\right).\end{aligned} \tag{4.110}$$

As σ is an incoherent state, it holds $V_k \sigma V_k^\dagger = \sigma$ and

$$\sum_{k} \mathrm{Tr}\left(V_k \sigma V_k^\dagger M_k\right) = \mathrm{Tr}\left(\sigma \sum_{k} M_k\right) = 1. \tag{4.111}$$

Thus, we arrive at the inequality which demonstrates that the performance of the probabilistic BV algorithm, when applied to states of the form $P(|-_D\rangle\langle -_D| \otimes \rho)$, is upper bounded by $\frac{1+R(\rho)}{d}$:

$$\max_{\{M_k\}} \frac{1}{d} \sum_{k} \mathrm{Tr}\left(V_k \rho V_k^\dagger M_k\right) \leq \frac{1 + R(\rho)}{d}. \tag{4.112}$$

From Eqs. (4.108) and (4.112) we have

$$\max_{\{M_k\}} \frac{1}{d} \sum_{k} \mathrm{Tr}\left(V_k \rho V_k^\dagger M_k\right) = \frac{1 + R(\rho)}{d}. \tag{4.113}$$

This completes the proof of Eq. (4.97).

4.6 Discussion

In this chapter, we have introduced and investigated a probabilistic version of the Bernstein-Vazirani algorithm where the goal is to accurately unveil a bit string $\boldsymbol{a}$ encoded within an oracle unitary. We have examined the algorithm's performance across all pure initial states, measured by its highest probability of correctly guessing the encoded bit string. We showed that there exists a direct relationship between performance and the amount of the coherence resource in the initial state. We further indicated that excessive multipartite entanglement can hinder the algorithm's ability to achieve peak performance.

Our approaches are applicable in quantum computation scenarios that involve mixed initial states. While analyzing the efficiency of the probabilistic Bernstein-Vazirani algorithm under noisy conditions, we observe that pseudo pure states offer superior performance while we have access to a constrained amount of purity. This observation carries significant implications for NMR-based quantum computation, suggesting that

NMR serves as an apt framework for implementing the probabilistic Bernstein-Vazirani algorithm.

In the next 2 chapters, we explore quantum resources in the circuit models with one control qubit, specifically for the task of computing the normalized trace of a unitary, we will illustrate that the model can achieve speedup over any known classical algorithm, even with arbitrary small multipartite entanglement, coherence, and general quantum correlations [NKG+22].

Chapter 5

Resources in Restricted Models of Quantum Computation based on Coherently Controlled Circuits

In this chapter, we examine quantum resources in restricted computational models. These models are based on circuits in which a single qubit coherently control unitary operation are applied on n qubits as shown in Fig. 5.1. One may consider two cases in these circuits: (i) when the circuit's gates are noiseless but the initial state of the qubits except the one controlling the unitaries are noisy, and the other case is (ii) when the initial state of the qubits is noiseless but the gates are noisy. Because in all our circuits in this chapter, the qubit controlling the gates is the only measured, we call it the "measured register". Apart when explicitly mentioned, we call the rest of the qubits the "data register".

For the first case, we study the quantum resources within the framework of deterministic quantum computation with one clean qubit (DQC1) [KL98]. This framework encompasses the circuits with noiseless unitary gates, one clean qubit (i.e. a qubit initialized in a pure state), and the state of the rest of the qubits being the maximally mixed state. Here we consider a class of DQC1 circuits where an n qubit unitary U is controlled by the clean qubit and by measuring this qubit, the goal is to find the normalized trace of U (i.e. $\mathrm{Tr}\, U/2^n$). To date, there is no known efficient classical algorithm for solving this problem [DFC05]. Several studies have attempted to understand the source of quantum speedup in this task by examining the properties of the quantum states used in the algorithm [DFC05, DSC08, DVB10, MEKP16a]. In a study by Datta et al. [DFC05], the level of bipartite entanglement in the final state of the DQC1 al-

gorithm was analyzed. The authors specifically examined the entanglement generated by the algorithm in various bipartitions, using the negativity as a measure of entanglement [idZHSL98, VW02]. The authors of [DFC05] concluded that negativity, as a measure of entanglement, is bounded by a constant that is independent of the number of qubits. This finding has led to the suggestion that other forms of quantum correlations, such as quantum discord [MBC+12, Str15], may be responsible for the observed speedup. Specifically, it has been observed that a typical instance of the algorithm exhibits non-zero quantum discord in a certain bipartition [DSC08]. Nevertheless, there is evidence indicating that exponential speedup can still occur even without the presence of quantum discord [DVB10]. Additionally, it has been proposed that the performance of trace estimation using DQC1 is more closely tied to the quantum coherence of the registers in the algorithm [MEKP16a]. We discover that the algorithm for finding the normalized trace of a unitary using these circuits, has the potential to achieve speedup compared to any known classical algorithm, even when there is only a minimal amount of multipartite entanglement, coherence, and general quantum correlations [NKG+22]. This finding highlights the robustness and power of the one clean qubit model as a quantum computing paradigm. It suggests that even with limited quantum resources, significant computational advantages can be attained.

The DQC1 model has the property that all but one of the qubits are initialized in a maximally mixed state while exhibiting a computational advantage. This feature made us wonder if by modifying the structure of these circuits, we could design circuits that are inherently resilient to noise while preserving a quantum advantage. By inherently resilient, we mean that the circuits provide reliable algorithm outcomes despite having noise everywhere: not only in the initial state but also in all the gates applied. Particularly, we ask that the noiseless outcome of the algorithm can be recovered at a cost (in the number of algorithm runs) that is not exponential in the circuit's characteristics (i.e. its total number of qubits, or algorithm depth which is the length of the longest sequence in the algorithm). It would lead to a way to fight the noise that is scalable, as its total cost in resources would not grow exponentially.

The motivation behind our question is that error correction is the only known approach to resist noise in arbitrary circuits that are scalable (for relatively general assumptions about the noise). However, it is challenging to implement as it requires gates to have error rates below very demanding threshold, making it impractical for the near term [TBG17, BMSSO18, MZO20, SQC+21, CBB+22, BMKT22, QCL21, Koc21]. Finding possible alternatives to error correction, ideally leading to less stringent requirements is then an interesting question. Note that other techniques than error correction do exist to resist the noise. They are usually called error mitigation. However for most noise models and circuit structures, these techniques suffer from scalability issues, as they require exponentially increasing resources with the size of the quantum circuit [CBB+22, QFK+22]. Research in this area suggests that, without the use of quantum error correction, achieving consistent and reliable algorithm outputs in scenarios that are classically intractable is not feasible under realistic noise conditions. However, recent studies have shown that even noisy quantum computers can surpass classical computers in certain oracle-based tasks [CCHL22]. A major challenge is that,

for most noise models, the fidelity of the output state declines exponentially as the number of gates increases [ZSW20]. This usually means that obtaining an accurate estimate of any expectation value would require an exponentially large number of algorithm executions, undermining the potential for exponential speedup. Other approaches may be more scalable but generally depend on specific noise models, not experimentally motivated [TLTC23], require information about the entanglement spectrum of quantum states [EMG+22], or involve potentially high algorithmic complexity [SMO22]. Considering these along with our motivation and some technical reasons (which we get into them in detail in this chapter) leads us to study the coherently controlled circuits with one control qubit initialized by pure states. These class of circuits obviously belong to the aforementioned case (ii) and they are called Hadmard tests.

There, we show that by assuming a noise model introducing only bit-flip which is inspired by existing superconducting cat qubits [PSJG+20], one can design a noisy restricted class of Hadamard tests, involving specific entangling operations and non-Clifford [1] gates that are resilient to noise in an asymptotic sense [FAND+23]. Loosely speaking, if the noise is only composed of bit-flips [NC10], the noise will corrupt the algorithm outcomes as a function of the algorithm size yet in a manner that a polynomial number of runs in the algorithm would recover the noiseless result. Our circuits avoid the usual no-go theorems in the literature regarding the possibility of preserving reliable outcomes for most noise models and circuits [FAND+23].

An important shortcoming of our finding is that we showed that our circuits cannot give rise to a computational advantage. We found an efficient classical algorithm able to produce samples that have the same probability distribution as measurement samples produced on the quantum computer [FAND+23]. This algorithm and the circuits we designed can nonetheless be used to benchmark the hardware. We can use our findings to see if the expected noise model of the qubits is still occurring in large-scale circuits [FAND+23]. This is a crucial requirement for superconducting cat qubits as the whole scalability of this platform relies on the fact that they only produce bit-flip errors [GRLR+23]. This is due to the fact that correcting only bit-flips would lead to lower overheads (number of qubits and gates dedicated to error correction). Hence our benchmarking is directly applicable to experimentalist working with superconducting cat qubits.

In this chapter, we focused on how these resilient circuits can be designed. In the next chapter, we will show how our circuits and simulations can be utilized as a protocol to benchmark the quality of these qubits.

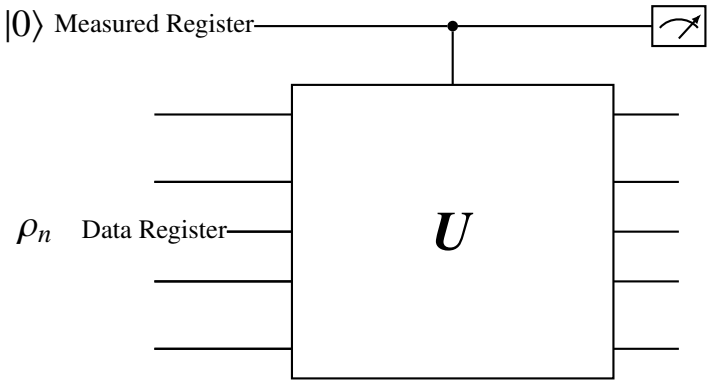

Figure 5.1: Schematic circuit for the Hadamard test or DQC1. The unitary matrix U acting on n number of qubits is controlled by the first one in the $\{+,-\}$ basis. In the case that the initial state $\rho_n = |\psi\rangle\langle\psi|$ is pure, the circuit is called a Hadammard test and is used to estimate $\langle\psi| U |\psi\rangle$ by only measuring the first qubit. When ρ_n is the maximally mixed state then we have a circuit belonging to the DQC1 class, and similar to the case of the Hadamard test, the circuit allows for the estimation of $\frac{1}{2^n}\operatorname{Tr} U$.

5.1 Hadamard Test and Power of One Qubit

The Hadamard test serves as the fundamental task upon which all of our examples are constructed. The Hadamard test is a procedure that enables the estimation of the expectation value $\langle\psi|U|\psi\rangle$ of a unitary operator U on a prepared n-qubit state $|\psi\rangle = B \otimes \bigotimes_{i=1}^{n} |\varphi_i\rangle$, where $|\varphi_i\rangle$ represents a single-qubit state and B is a unitary operation. A way to implement it is represented in Fig. 5.1. In the Hadamard test, the system consists of the measured register, initially prepared in the state $|0\rangle$, and the data register containing the prepared state $|\psi\rangle$. Just before measurement, the reduced state of the measured register can be written as:

$$\rho = \frac{1}{2}\left(I + \alpha_n \left(yY + zZ\right)\right) \tag{5.1}$$

where $y = -\Im(\langle\psi| U |\psi\rangle), z = \Re(\langle\psi| U |\psi\rangle)$ and $\alpha_n = 1$ for now on. Consequently, measuring the first register in the Y-basis enables us to estimate the imaginary part of $\langle\psi| U |\psi\rangle$, while measuring in the Z-basis allows us to estimate the real part of $\langle\psi| U |\psi\rangle$. According to Hoeffding's inequality [CBB+22], performing $N = 2\log(2/\delta)/\varepsilon^2$ experimental repetitions is adequate for estimating the values of y or z with an ε-precision and a probability of $1-\delta$. The problem of estimating $\langle\psi| U |\psi\rangle$, where U is generally a circuit composed of polynomially many gates as a function of n, to additive precision is known to be a BQP-complete problem [AJL09]. This implies that, in general,

[1] We mention the presence of the non-Clifford gates because they are required for an exponential quantum advantage as long as the qubits are initialized and measured in the eigenstates of a Pauli. However, if one allows for other state preparations or measurements, non-Clifford gates are not in general required for the advantage [AG04].

we do not expect an efficient classical algorithm that can efficiently solve this task. BQP-completeness signifies that this problem is computationally as hard as any other problem in the complexity class BQP (bounded-error quantum polynomial time). It is worth mentioning that BQP class is a complete class for quantum computation meaning that it exhibits the full computational power of a quantum computer. We notice that if the same model of computation as in the Hadamard test is initialized by the state $|0\rangle\langle 0| \otimes \frac{\mathbb{I}}{2^N}$ (the first qubit is the controlled qubit), we obtain a class of DQC1 circuits with the clean qubit to be the control qubit which is explained in more details in what follows.

As mentioned in the previous chapter, NMR has been identified as a viable platform for realizing the BV algorithm. Another prominent quantum computational model often examined in the context of NMR is DQC1 [KL98]. In the DQC1 model with one control qubit, the initial state is given by $\rho \otimes \mathbb{I}_n/2^n$, where $\rho = |0\rangle\langle 0|$, $\mathbb{I}_n/2^n$ is a maximally mixed n-qubit state. In the context of this model, a method proposed in [DFC05] allows for efficient estimation of the normalized trace of an n-qubit unitary U Given the efficient implementation of V_n using quantum gates. See also Fig. 5.1. This estimation is achieved by applying a controlled version of U to the initial state, where the first qubit serves as the control and the remaining n qubits act as the target:

$$V_n = |-\rangle\langle -| \otimes \mathbb{I} + |+\rangle\langle +| \otimes U. \tag{5.2}$$

Note that V_n is the control version of U. Just before measurement, the reduced state of the controlled qubit can be written as:

$$\rho = \frac{1}{2}\left(I + \alpha_n \left(yY + zZ\right)\right) \tag{5.3}$$

where $y = -\Im(\frac{\operatorname{Tr} U}{2^n})$, $z = \Re(\frac{\operatorname{Tr} U}{2^n})$ and $\alpha_n > 0$. As a result, measuring the first register in the Y-basis estimates the imaginary part of $\operatorname{Tr} U/2^n$, whereas measuring in the Z-basis provides an estimate of the real part of $\operatorname{Tr} U/2^n$. This estimation method is applicable when $\alpha_n > 0$ [DFC05].

In the next sections, we present the main results of this chapter.

5.2 Quantum Resources in DQC1

We will now demonstrate that the efficient implementation of normalized trace estimation with DQC1 is possible, even when considering a broad class of quantum resource and correlation quantifiers that can be arbitrarily small at each step of the algorithm. We will begin our analysis by considering general entanglement quantifiers and then extend it to other measures. We will prove that in the DQC1 protocol with one control qubit, any distance-based entanglement measure is bounded by a constant value [NKG+22]. We can observe that the maximally mixed state of $N + 1$ qubits, denoted as $\mathbb{I}_{N+1}/2^{N+1}$, is a fully separable state. Therefore, for any $N + 1$-qubit state σ, we

have the inequality $E(\sigma) \leq D(\sigma, \mathbb{I}_{N+1}/2^{N+1})$, where $E(\sigma)$ represents the distance based entanglement measure and $D(\sigma, \mathbb{I}_{N+1}/2^{N+1})$ denotes the distance between σ and the maximally mixed state [NKG+22]. We can recall that in the DQC1 protocol, the initial state is given by $\rho \otimes \mathbb{I}_n/2^n$. After the application of V_n, we obtain [NKG+22]

$$E\left(V_n\rho \otimes \frac{\mathbb{I}_n}{2^n}V_n^\dagger\right) \leq D\left(V_n\rho \otimes \frac{\mathbb{I}_n}{2^n}V_n^\dagger, \frac{\mathbb{I}_{n+1}}{2^{n+1}}\right) = D\left(\rho \otimes \frac{\mathbb{I}_n}{2^n}, \frac{\mathbb{I}_{n+1}}{2^{n+1}}\right) = D\left(\rho, \frac{\mathbb{I}_1}{2}\right). \tag{5.4}$$

Here, we utilized the property that any distance measure D satisfying the data processing inequality remains invariant under unitary operations and the addition or removal of ancillary systems, i.e. $D(\rho, \sigma) = D(U\rho U^\dagger, U\sigma U^\dagger)$ and $D(\rho, \sigma) = D(\rho \otimes \tau, \sigma \otimes \tau)$.

We observe that the amount of multipartite entanglement in the algorithm is bounded by a constant. This bound holds as long as the distance measure $D(\rho, \mathbb{I}_1/2)$ is limited to c for all qubit states ρ [NKG+22]. For any continuous distance measure D, it is possible to choose an arbitrarily small constant by selecting an appropriate value of α_n. This implies that the amount of multipartite entanglement in the DQC1 algorithm can be made arbitrarily small by adjusting the parameter α_n accordingly [NKG+22]. Note that the specific form of the unitary V_n is not relevant in Equation (5.4). The result holds true for any unitary operation acting on the total $n + 1$ qubit state in the protocol. As a consequence of this property, the conclusion also extends to the intermediate states of the algorithm ρ_i, which represent the states of the quantum processor after the application of i quantum gates [NKG+22]. This result indicates that normalized trace estimation with DQC1 can be accomplished with an extremely small amount of multipartite entanglement throughout the entire algorithm [NKG+22].

The methodology described above is not restricted to entanglement alone; it can be extended to encompass a wide range of quantum resource and correlation quantifiers that exhibit vanishing values on maximally mixed states. Therefore, the conclusion applies to a broad class of measures that capture various aspects of quantum resources and correlations [NKG+22]. To illustrate this, let's consider a general quantity of the form

$$\mathcal{M}(\rho) = \inf_{\sigma \in \mathcal{F}} D(\rho, \sigma), \tag{5.5}$$

where $\mathcal{F}$ is a set of $n + 1$-qubit states that includes the maximally mixed state, and D is a distance metric that satisfies the properties mentioned earlier. It is evident that the arguments presented in Eq. (5.4) can be applied to any quantity of this type. To see that the above results apply to the quantum mutual information [NKG+22]

$$I\left(\rho^{AB}\right) = S\left(\rho^A\right) + S\left(\rho^B\right) - S\left(\rho^{AB}\right), \tag{5.6}$$

recall that the mutual information can be expressed in the form given by Eq. (5.5), where the quantum relative entropy D is used as the distance measure and the set $\mathcal{F}$ consists of product states [BM18]. Here, the systems A and B represent arbitrary subsets of the $n+1$ qubits. The results discussed earlier extend to various measures of quantum correlations beyond entanglement, such as quantum discord [MBC+12, Str15],

when considering $\mathcal{F}$ as the set of classically correlated states $\rho_{\text{cc}} = \sum_{i,j} p_{ij} |a_i\rangle\langle a_i| \otimes |b_j\rangle\langle b_j|$ or classical-quantum states $\rho_{\text{cq}} = \sum_i p_i |a_i\rangle\langle a_i| \otimes \sigma_i$ with local orthonormal bases $\{|a_i\rangle\}$ and $\{|b_j\rangle\}$ and general local states σ_i. Our results also hold for the relative entropy of coherence, which is a measure of quantum coherence [BCP14b], by considering $\mathcal{F}$ as the set of incoherent states and utilizing the quantum relative entropy as the distance measure D [NKG+22]. Finally, In the case where $\mathcal{F}$ consists solely of the maximally mixed state of $N+1$ qubits, the quantifier $\mathcal{M}$ corresponds to a measure of purity [HHO03, GMN+15, SKW+18].

To summarize, our analysis demonstrates that the DQC1 protocol can efficiently estimate normalized traces even in the presence of minimal levels of multipartite entanglement, mutual information, general quantum correlations, coherence, or purity at each step of the algorithm [NKG+22].

5.3 Scalable Noisy Circuits under Bit-Flip Noise

In this section, we design a class of circuits that are noise-resilient in the asymptotic sense i.e. the noiseless answer can be recovered at the cost of running the algorithm polynomially many times as a function of the problem size (see Fig. 5.2). We will see that, roughly speaking, such scalability is the result of two interconnected resources which we are provided with; the coherence of the controlled qubit and the asymmetry of the noise (i.e. the fact that only bit-flips are produced after each gate in the algorithm) [FAND+23]. Let's first establish some notations and definitions for future convenience.

We denote the single-qubit Pauli matrices as $(\sigma_0, \sigma_1, \sigma_2, \sigma_3) \equiv (I, X, Y, Z)$. Let H denote the Hadamard gate. We define $\mathbb{P}_n^X$ as the set of X-Pauli operators acting on n qubits, denoted as $\mathbb{P}_n^X \equiv \{\bigotimes_{k=1}^n \sigma_{i_k} \mid \forall k, i_k \in 0, 1\}$.

Definition 5.1. We say that f_n belongs to the set poly(n) if there exist two positive real numbers C and a such that $\lim_{n\to\infty} \frac{f_n}{Cn^a} = 1$. Additionally, when we use poly(n) in an equation, it implies that the equation holds true if we replace poly(n) with any function f_n belonging to poly(n).

We say that $f_n = O(g_n)$ if there exists a positive constant C such that $\lim_{n\to\infty} \left|\frac{f_n}{g_n}\right| \leq C$. We also define the coherently controlled operation of a unitary operator A in the X-basis as $c_X A = |+\rangle\langle+| \otimes I + |-\rangle\langle-| \otimes A$. Consider a single-qubit unitary denoted as G. In a tensor product, the notation G_i signifies that G is applied specifically to the i-th qubit, while the identity operator $\mathbb{I}$ is applied to the remaining qubits.

Let's denote the unitary transformation corresponding to a unitary quantum gate G as $\mathcal{G}$, and the Completely Positive Trace Preserving (CPTP) operation representing the noisy implementation of this gate as $\mathcal{E}_{\mathcal{G}}$. The CPTP map $\mathcal{N}_{\mathcal{G}}$, defined as $\mathcal{N}_{\mathcal{G}} \equiv \mathcal{E}_{\mathcal{G}} \circ \mathcal{G}^\dagger$, is referred to as the "noise map of G" (or the noise map associated with $\mathcal{G}$). In the context of state preparation, the noise map is the CPTP operation that is applied

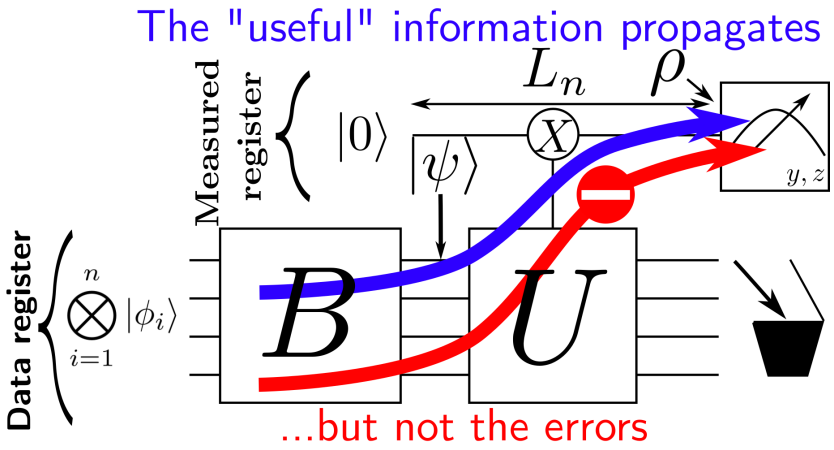

Figure 5.2: [FAND+23] The Hadamard test illustrated in this figure is used to estimate $\langle\psi| U |\psi\rangle$, where $|\psi\rangle$ is defined as $B \bigotimes_{i=1}^{n} |\varphi_i\rangle$ (with $|\varphi_i\rangle$ being individual qubit states and B a unitary operator). We define L_n as the count of gates that are applied to the measured register, including any potential noisy identity gates, after the unitary $c_X U$ has been broken down into a set of gates that can be implemented experimentally. In this study, we show that for a noise model characterized exclusively by bit-flips, and under specific conditions on the n-qubit unitaries U and B, the measurement process will only respond to bit-flips that occur in the measurement register, remaining entirely unaffected by bit-flips in the data register. Nevertheless, some of the useful information encoded in the entangled state $|\psi\rangle$ will still be transferred to the measured register. Our strategy involves encoding this important information in a 'Pauli Z (or Y) channel,' while keeping any errors confined to a 'Pauli X channel.' If a Pauli X error (such as a bit-flip) in the data register does not affect the measurement outcome, but Pauli Z information does, then the useful information will be successfully conveyed to the measurements. In contrast, the errors will not propagate to the measurement register, thereby isolating them from affecting the results. In summary, if L_n scales logarithmically with n (i.e., $L_n = O(\log(n))$), the presence of noise will only result in a polynomial increase in the number of algorithm repetitions required to ensure reliable results. This means that the algorithm can remain scalable even in the presence of noise. It is important to note that although $L_n = O(\log(n))$, the algorithm itself can still have polynomial depth.

following an ideal (noiseless) state preparation, ensuring that the sequence of ideal preparation followed by the CPTP map accounts for the noisy state preparation.

We recall that in the noiseless case, the outcome of the Hadamard test is prescribed by Eq. 5.1 with $\alpha_n = 1$. Then, one needs to run the algorithm $N = 2\log(2/\delta)/\varepsilon^2$ to fulfill the task with ε precision and the probability $1 - \delta$. For pedagogy, let's consider the case that there exists a noise whose effect is in such a way that $0 < \alpha_n < 1$. Then, we need $N_n = \frac{2\log(2/\delta)}{(\alpha_n \varepsilon)^2}$ repetitions of the algorithm to achieve an estimation accuracy

of ε with a desired confidence level δ [FAND+23]. If α_n decreases exponentially with n, the total number of algorithm calls required, as determined by N_n, would also grow exponentially with n [FAND+23]. In this case, the algorithm would not be considered scalable. The exponential growth in the number of algorithm repetitions would result in a significant increase in computational resources needed as the problem size n increases. In the case where $\alpha_n = 1/\text{poly}(n)$ and is efficiently computable, it is possible to achieve a reliable estimation of $\langle\psi| U |\psi\rangle$ with a polynomial overhead in the number of experiment repetitions [FAND+23]. The goal of everything that follows is to show that under some restrictions for the gates in the algorithm, everything will behave as in this example with $\alpha_n = \frac{1}{poly(n)}$, hence the algorithm will be scalable despite the presence of the noise.

5.3.1 Noise Model

In general, to implement both B and $c_X U$, they need to be decomposed into a gateset that is feasible to implement at the experimental level. Consider a unitary channel $\mathcal{G}$ representing a gate from the set of accessible gates, and let $\mathcal{E}_\mathcal{G}$ denote its noisy implementation in the laboratory. We will assume that every gates in the computation follows a local biased noise model: $\mathcal{E}_\mathcal{G} = \mathcal{N}_\mathcal{G} \circ \mathcal{G}$, where the "noise map" $\mathcal{N}_\mathcal{G}$ will only introduce (possibly correlated) bit-flip errors on the qubits on which $\mathcal{G}$ acts non-trivially (i.e. supp($\mathcal{G}$)),

$$\mathcal{N}_\mathcal{G}(\rho) = \sum_{\alpha \subset \text{supp}(\mathcal{G})} p_\alpha^\mathcal{G} X_\alpha \rho X_\alpha \,, \tag{5.7}$$

where we have $X_\alpha = \prod_{i\in\alpha} X_i$, and $\{p_\alpha^\mathcal{G}\}$ is a probability distribution that is defined on subsets of supp($\mathcal{G}$). As an illustration, let's consider the noise model of a two-qubit gate. In this case, the Kraus operators are proportional to $\sigma \otimes \sigma'$, where (σ, σ') can take values from the set $\{\mathbb{I}, X\}$. Additionally, a noisy measurement can be represented by conducting an ideal measurement followed by a probability p_{meas} of flipping the measurement outcome. Finally, in the case of single-qubit noisy state preparation, we consider a two-step process where a perfect state preparation is followed by the application of a Pauli X-error with probability p_{prep}. Our noise model is inspired by an idealization of cat qubits, which can effectively suppress noise channels other than bit-flip errors exponentially, at the expense of an increased bit-flip rate [GM21, LVP+20, CNAA+22]. Our noise model is an idealization for two reasons [FAND+23]. First, the Kraus operators are restricted to linear combinations of the Pauli X and $\mathbb{I}$ operators, which we refer to as perfect bias. Second, we disregard coherent errors, meaning that our noise model is entirely Pauli noise. However, the benchmarking protocol explained in the next chapter is valid for biased qubits that are designed such that their noise model can be represented by the first assumption alone [FAND+23] (see Theorem 6.3 and Definition 6.1). We now expound on what we mean by "an error".

Definition 5.2 (Error)**.** Let $|\Psi\rangle$ denote the desired state of the qubits at a particular timestep of the algorithm, assuming perfect gates. Due to the probabilistic nature of the noise model, the actual n-qubit quantum state will deviate from the ideal state

$|\Psi\rangle$. This deviation is described by the density operator ρ, which can be expressed as a weighted sum of terms involving the error operators E_i. Each error operator E_i represents a specific type of unitary operation that has affected the state $|\Psi\rangle$, and the corresponding weight p_i represents the probability of that particular error occurring ($p_i \geq 0$, $\sum_i p_i = 1$). In other words, the actual state ρ will be:

$$\rho = \sum_i p_i E_i |\Psi\rangle \langle\Psi| E_i^\dagger. \tag{5.8}$$

5.3.2 Characterization of the Gates Preserving the Biased Noises

The central concept underlying our work is to take advantage of the fact that only bit-flip errors occur, enabling us to design circuits that minimize the propagation of these errors to the measurement performed in the algorithm. Hence, we need to guarantee that (i) all along the algorithm we only have bit-flip (more precisely X-errors as defined in Def. 5.3) errors and (ii) most of these errors do not propagate toward the measured register [FAND+23]. In this section, we show how (i) can be satisfied. In the next section, we analyze how X-errors propagate through the gates that we will use to craft our circuits (once B and U as in Fig. 5.2 are decomposed in a sequence of primitive gates), so that in the section 5.3.4, we present how we can design the circuits satisfying (ii). For our purposes regarding (i), we introduce some technical definitions.

Definition 5.3 (X-type unitary operators and errors)**.** The set of unitary operators that can be expressed as a linear combination of Pauli X matrices is referred to as X-type unitary operators. Formally, for an n-qubit system, we define it as follows:

$$\mathbb{U}_n^X \equiv \{U = \sum_i c_i P_i,\ |P_i \in \mathbb{P}_n^X, c_i \in \mathbb{C}, U^\dagger = U^{-1}\}, \tag{5.9}$$

Another way to understand $\mathbb{U}_n^X$ is as the set of unitary operators that have a diagonal form when expressed in the product basis of the n-qubit system $|\mathbf{s}\rangle = |s_1\rangle |s_2\rangle \dots |s_n\rangle$, where $s_i = \pm$ and $|\pm\rangle$ are eigenstates of Pauli X matrix [FAND+23]. We will refer to an error as an "X-error" if it belongs to the set $\mathbb{U}_n^X$. If the error is a Pauli operator, we will specifically call it a "Pauli X error" or simply a "bit-flip".

Our objective is to ensure that any error that occurs at any step of the computation is an X-error. In other words, we want to restrict the errors to the set $\mathbb{U}_n^X$. To achieve this, we employ "bias-preserving" gates which we define as follow [FAND+23].

Definition 5.4 (Bias-preserving gates)**.** A unitary operator G on n qubits is said to preserve the X-errors (or X-bias) if it satisfies the following property:

$$\forall P \in \mathbb{P}_n^X, \exists A \in \mathbb{U}_n^X \text{ such that } GP = AG \tag{5.10}$$

We denote $\mathbb{B}_n$ the set of such gates.

Theorem 5.1 (Preservation of the bias)**.** *If a quantum circuit consists only of gates from $\mathbb{B}_n$, where each gate is subject to the local biased noise model described by Eq. (5.7), then any error that affects the state of the computation will be an X-error.*

Proof. [FAND+23] Following the initialization step and in the absence of errors, the state of the computation would be $|\psi_{\text{prep}}\rangle = \bigotimes_{i=1}^{n} |\varphi_i\rangle$. Due to the noisy initialization process, as described in the noise model around (5.7), the state after preparation will be in a mixed state:

$$\rho_i = \sum_{\alpha} p_{\alpha}^{\text{prep}} X_{\alpha} |\psi_{\text{prep}}\rangle\langle\psi_{\text{prep}}| X_{\alpha}^{\dagger} \tag{5.11}$$

for some probabilities p_{α}^{prep}. In this equation (and the subsequent equations), the sum over α ensures that all n-qubit Pauli-X operators will be reached exactly once, for some X_{α}. In more formal terms, the sum is defined such that $\alpha \subset \text{supp}(\mathbb{I}_n)$, where $\mathbb{I}_n$ represents the identity matrix applied on n qubits. We now consider a gate $G \in \mathbb{B}_n$. For any state $|\Psi\rangle$, we have, for some probability p_{α}^{G}:

$$\begin{aligned} \mathcal{E}(|\Psi\rangle\langle\Psi|) &= \mathcal{N}_{\mathcal{G}} \circ \mathcal{G}(|\Psi\rangle\langle\Psi|) \\ &= \sum_{\alpha} p_{\alpha}^{G} X_{\alpha} G |\Psi\rangle\langle\Psi| G^{\dagger} X_{\alpha}^{\dagger} \end{aligned} \tag{5.12}$$

Hence, we have:

$$\mathcal{E}(\rho_i) = \sum_{\alpha_1,\alpha_2} p_{\alpha_1}^{G} p_{\alpha_2}^{\text{prep}} X_{\alpha_1} G X_{\alpha_2} |\psi_{\text{prep}}\rangle\langle\psi_{\text{prep}}| X_{\alpha_2}^{\dagger} G^{\dagger} X_{\alpha_1}^{\dagger} \tag{5.13}$$

Using the fact $G \in \mathbb{B}_n$, we have: $GX_{\alpha_2} = E_{\alpha_2} G$ for some $E_{\alpha_2} \in \mathbb{U}_n^X$. Hence:

$$\mathcal{E}(\rho_i) = \sum_{\alpha_1,\alpha_2} p_{\alpha_1}^{G} p_{\alpha_2}^{\text{prep}} X_{\alpha_1} E_{\alpha_2} G |\psi_{\text{prep}}\rangle\langle\psi_{\text{prep}}| G^{\dagger} E_{\alpha_2}^{\dagger} X_{\alpha_1}^{\dagger} \tag{5.14}$$

The fact that $X_{\alpha_1} \times E_{\alpha_2} \in \mathbb{U}_n^X$ implies that any noisy gate G applied to ρ_i is only affected by errors belonging to $\mathbb{U}_n^X$. Indeed, the same reasoning can be applied recursively for any gate applied after this point, confirming the validity of this property throughout the circuit.

□

Examples of bias-preserving gates include all the unitaries in $\mathbb{U}_n^X$, the controlled-NOT (cNOT) gate, and a modified version of the Toffoli gate denoted as Toffoli' $\equiv H_1H_2H_3 \times \text{Toffoli} \times (H_1H_2H_3)^{\dagger}$, where H_i are Hadamard gates applied to the corresponding qubits [FAND+23]. To clarify, when we refer to the Toffoli' gate being implemented "natively," we mean that it can be directly applied as a single unitary operation without relying on separate Hadamard gates. This is crucial for preserving the X-bias property. An example of a gate that does not preserve the bias is the Hadamard gate [FAND+23].

Bias-preserving gates have an elegant interpretation: they can be seen as performing permutations (up to a phase) in the Pauli X-eigenstates basis [FAND+23].

Theorem 5.2 (Characterization of bias-preserving gates)**.** *A unitary operator V belongs to the set $\mathbb{B}_n$ if and only if, for any $\mathbf{s} \in \{+,-\}^n$, there exists a real phase $\varphi_{\mathbf{s},V}$ such that the action of V on the state $|\mathbf{s}\rangle$ can be expressed as $V|\mathbf{s}\rangle = e^{i\varphi_{\mathbf{s},V}}|\sigma_V(\mathbf{s})\rangle$, where σ_V represents a permutation of the elements in $\{+,-\}^n$ induced by the gate V.*

Proof. [FAND+23] We consider $\mathbf{s} \in \{+,-\}^n$ and $V \in \mathbb{B}_n^X$. We have $V|\mathbf{s}\rangle\langle\mathbf{s}|V^\dagger = |\Phi(\mathbf{s},V)\rangle\langle\Phi(\mathbf{s},V)|$ for some pure state $|\Phi(\mathbf{s},V)\rangle$. We also have $|\mathbf{s}\rangle\langle\mathbf{s}| = \sum_\alpha c_\alpha X_\alpha$ for some family of complex coefficients $\{c_\alpha\}$. Let us reiterate our notation, where α is a bit-string of n bits, and a bit value of 1 (resp. 0) at the i'th position indicates that a Pauli X (resp. $\mathbb{I}$) operator should be applied to the i'th tensor product, i.e., $X_\alpha = \prod_{i\in\alpha} X_i$. After observing that $VX_\alpha V^\dagger \in \mathbb{U}_n^X$, we can conclude that $|\Phi(\mathbf{s},V)\rangle\langle\Phi(\mathbf{s},V)|$ is diagonal in the local X-basis. In other words, it is composed of elements from the basis $|\mathbf{s}\rangle, \mathbf{s} \in \{+,-\}^n$. Being a pure state, it implies $|\Phi(s,V)\rangle \propto |\mathbf{s}'\rangle$ for some $\mathbf{s}' \in \{+,-\}^n$. Since unitary channels are invertible, for any $\mathbf{s} \in \{+,-\}^n$, there exists a permutation $\sigma_V : \{+,-\}^n \to \{+,-\}^n$ such that $V|\mathbf{s}\rangle\langle\mathbf{s}|V^\dagger = |\sigma_V(\mathbf{s})\rangle\langle\sigma_V(\mathbf{s})|$. This equation implies:

$$V|\mathbf{s}\rangle = \exp(i\varphi_{\mathbf{s},V})|\sigma_V(\mathbf{s})\rangle \tag{5.15}$$

where $\varphi_{\mathbf{s},V}$ is a real phase that can depend on $\mathbf{s}$ and V.

Reciprocally, if for any $\mathbf{s} \in \{+,-\}^n$, there exists a real phase $\varphi_{\mathbf{s},V}$ such that: $V|\mathbf{s}\rangle = e^{i\varphi_{\mathbf{s},V}}|\sigma_V(\mathbf{s})\rangle$, we have:

$$\sum_{\mathbf{s}} Vc_{\mathbf{s}}|\mathbf{s}\rangle\langle\mathbf{s}|V^\dagger = \sum_{\mathbf{s}} c_{\mathbf{s}}|\sigma_V(\mathbf{s})\rangle\langle\sigma_V(\mathbf{s})| \in \mathbb{U}_n^X. \tag{5.16}$$

Therefore, if an operator is diagonal in the local X-basis, it will remain diagonal in the same basis after applying the map $U \to VUV^\dagger$. This property implies that any unitary V satisfying equation (5.15) must be bias-preserving. In other words, the set of V's that fulfill the condition (5.15) corresponds to bias-preserving gates. □

5.3.3 Coherently Controlled Bias-Preserving Gates Limiting the Propagation of Errors

The theorem 5.1 of the previous part shows how we can guarantee that at any step of the computation, only X-errors are damaging the quantum state. In this section, we analyze how these errors propagate through the gates that we will use to design our circuits. Since in the circuits we will design (which rely on decomposing U and B from Fig. 5.2 into a set of primitive gates) the only gates interacting with the measured register will be controlled operations, we need to analyze how X-errors propagate through such gates. This is the goal of this section. Then, in the next section, we will use such gates in order to design circuits where most of the X-errors do not reach the measurement, leading to noise resilience of the circuits.

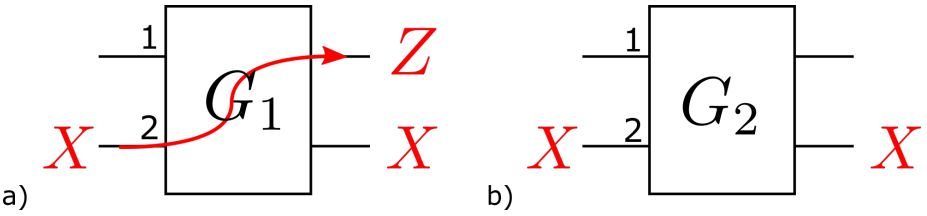

Figure 5.3: [FAND+23] **a)** Example of a unitary gate G_1 which propagates bit-flip error: when an X-Pauli error is present on qubit 2 before applying a two-qubit gate G_1, this error stays on qubit 2. However, it also induces a 'new' error on qubit 1 as a result of how the error propagates through G_1. In different terms, this diagram represents the fact that, for this gate G_1, we have: $G_1X_2 = Z_1X_2G_1$. **b)** For some other gates (for instance, if G_2 is a cNOT controlled by the qubit 1), X errors occurring at specific places do not propagate to other qubits but remain confined to the qubit where they originated, as demonstrated in the figure. Specific examples of how errors can propagate through different gates are detailed in Figure B.1.

.

We first formally define what we mean by the term "error propagation". It refers to the phenomenon where a pre-existing error occurring on some qubits before a gate can introduce errors on potentially additional qubits after the gate is applied [FAND+23]. In other words, the errors can spread or propagate to other qubits as the computation progresses through the gates. The "new" errors that arise after a gate application are a consequence of the gate dynamics, even if the gate itself is noiseless. This phenomenon is illustrated in Figure 5.3.

Ensuring the reliability of the measurement requires preventing X errors from propagating from the target (measured qubit) to the control side of the unitary [FAND+23]. Therefore, we are particularly interested in identifying conditions under which such X errors do not propagate from the target qubit to the control one. This is essential for maintaining the integrity of the computation and obtaining accurate results from the Hadamard test. In our notation, a coherently controlled unitary operation A is denoted as c_PA, where P represents a single-qubit matrix given by $\mathbf{n}.\boldsymbol{\sigma}$, with $\mathbf{n}$ being a unit vector, and $\boldsymbol{\sigma} = (X, Y, Z)$. This coherently controlled unitary is applied only if the control qubit is in the eigenstate -1 of the matrix P. The first notable observation is that A must commute with *every* element in $\mathbb{P}_n^X$. Otherwise, for a state $|\psi\rangle$ and a $P_X \in \mathbb{P}_n^X$, the following relation would hold:

$$\langle\psi| P_X^\dagger A P_X |\psi\rangle \neq \langle\psi| A |\psi\rangle \tag{5.17}$$

Such a scenario would imply that there exists a Hadamard test implemented on the unitary A and a state $|\psi\rangle$ that would be sensitive to X-errors occurring in the data register. This contradicts the noise-resilience property we aim to achieve. Hence, A must commute with all elements in $\mathbb{P}_n^X$. This remark can be summarized as follows [FAND+23].

Theorem 5.3 (Necessary conditions on a gate A such that c_PA does not propagate errors from the target to the control)**.** *If we define the controlled unitary c_PA that acts on $1+n$*

qubits as:

$$c_P A \equiv \frac{1}{2}(\mathbb{I} + P) \otimes \mathbb{I} + \frac{1}{2}(\mathbb{I} - P) \otimes A, \tag{5.18}$$

and does not propagate X errors from the target to the control, then A commutes with any element in $\mathbb{P}_n^X$, which implies $A \in \mathbb{U}_n^X$. Here, $P \equiv \mathbf{n}.\boldsymbol{\sigma}$, where $\mathbf{n}$ is a unit vector. The condition for the absence of propagation of X-errors from the target to the control is as follows:

$$\forall P_X \in \mathbb{P}_n^X, (c_P A)(\mathbb{I} \otimes P_X)(c_P A)^\dagger = \mathbb{I} \otimes B \tag{5.19}$$

for some unitary B.

We will now demonstrate the following two properties [FAND+23]:

Definition 5.5 (Bias-preserving controlled unitaries avoiding X-errors to propagate toward the control)**.** Let $c_P A \equiv \frac{1}{2}(\mathbb{I} + P) \otimes \mathbb{I} + \frac{1}{2}(\mathbb{I} - P) \otimes A$, where A acts on n qubits. We say that $c_P A$ is bias preserving and does not propagate X-errors to the control if it satisfies the following conditions:

$$\forall P_X \in \mathbb{P}_{n+1}^X, (c_P A) P_X (c_P A)^\dagger \in \mathbb{U}_n^X \tag{5.20}$$

$$\forall P_X \in \mathbb{P}_n^X, \exists U_X \in \mathbb{U}_n^X \, (c_P A) \, (\mathbb{I} \otimes P_X) \, (c_P A)^\dagger = \mathbb{I} \otimes U_X. \tag{5.21}$$

Theorem 5.4. *[Characterisation of controlled unitaries avoiding errors to propagate toward the control]*

The controlled unitary $c_X A$ is bias preserving and prevents X errors from propagating toward the control iff A belongs to $\mathbb{U}_n^X$.

The controlled unitary $c_P A$ with $P = yY + zZ$ (such that $(0, y, z)$ is a unit vector) is bias preserving and avoids X errors from propagating toward the control register iff A belongs to $\mathbb{U}_n^X$ and is Hermitian. None of the controlled unitaries $c_P A$ can have any other $\mathbf{n}.\boldsymbol{\sigma}$ than the ones previously discussed that is not trivial (i.e. $A \neq \mathbb{I}$) and that will satisfy the constraints on error propagation.

Proof. [FAND+23] For $c_P A$ to meet the requirements, it must satisfy conditions (5.20) and (5.21). We begin with (5.21):

$$c_P A(\mathbb{I} \otimes P_X)(c_P A)^\dagger = \mathbb{I} \otimes U_X \tag{5.22}$$

for some $U_X \in \mathbb{U}_n^X$ and for any $P_X \in \mathbb{P}_n^X$. This equation implies:

$$(\mathbb{I} + \mathbf{n} \cdot \boldsymbol{\sigma}) \otimes P_X + (\mathbb{I} - \mathbf{n} \cdot \boldsymbol{\sigma}) \otimes A P_X A^\dagger = 2\mathbb{I} \otimes U_X. \tag{5.23}$$

By expanding $c_P A$ in the Pauli basis and identifying the left and right hand sides, we can verify (5.23) implies that for any $P_X \in \mathbb{P}_n^X$, $P_X = A P_X A^\dagger$ and $P_X + A P_X A^\dagger = 2U_X$. The second equation is a consequence of the first one and is not independent. These

equations do not impose any additional constraints on A beyond the fact that $A \in \mathbb{U}_n^X$, as we have already established in property 5.3. Hence, we move on with the condition (5.20). The condition (5.20) implies that for all $P_X \in \mathbb{P}_n^X$, $(c_P A)\mathbb{I} \otimes P_X (c_P A)^\dagger \in \mathbb{U}_{n+1}^X$ and $(c_P A)X \otimes P_X (c_P A)^\dagger \in \mathbb{U}_{n+1}^X$. The second condition is the key aspect we need to focus on, as the first condition is already implied by the analysis of (5.21) that we have just completed. By considering $\mathbf{n} = (x, y, z)$, and using the fact $x^2 + y^2 + z^2 = 1$, we find that $(c_P A)X \otimes P_X (c_P A)^\dagger \in \mathbb{U}_{n+1}^X$ implies:

$$\begin{aligned} &\frac{x}{2}\mathbb{I} \otimes (P_X - AP_XA^\dagger) \\ +\frac{1}{2}X \otimes (x^2(P_X + AP_XA^\dagger) + (y^2 + z^2)(P_XA^\dagger + AP_X)) \\ &+\frac{1}{2}Y' \otimes (AP_X - P_XA^\dagger) \\ +\frac{1}{2}Z' \otimes (P_X - P_XA^\dagger - AP_X + AP_XA^\dagger) \in \mathbb{U}_{n+1}^X. \end{aligned} \tag{5.24}$$

where $Y' \equiv iyZ - izY$ and $Z' = x(yY + zZ)$. It can be easily verified that the matrices I, X, Y', Z' form a basis for the space of 2×2 complex matrices. Given this equation and the fact that $A \in \mathbb{U}_n^X$, we can deduce the following non-trivial implications:

$$Y' \otimes (AP_X - P_XA^\dagger) = 0 \tag{5.25}$$

$$Z' \otimes (P_X - P_XA^\dagger - AP_X + AP_XA^\dagger) = 0 \tag{5.26}$$

First case: either $y \neq 0$ or $z \neq 0$:
Knowing that $A \in \mathbb{U}_n^X$, (5.25) implies $A = A^\dagger$. Then, (5.26) implies that we must either have $x(yY+zZ) = 0$, or $P_X - P_XA^\dagger - AP_X + AP_XA^\dagger = 0$. If $P_X - P_XA^\dagger - AP_X + AP_XA^\dagger = 0$. By leveraging the properties of A being Hermitian and belonging to $\mathbb{U}_n^X$, we find that the only solution for A is $A = \mathbb{I}$, which corresponds to a trivial gate. As a result, to discover non-trivial gates, we must have $x = 0$. *In conclusion, if $y \neq 0$ or $z \neq 0$ then necessarily A is Hermitian and $x = 0$ in order to satisfy the constraints* (5.25)*,* (5.26) *with $A \neq I$.*
Second case: $y = z = 0 \Leftrightarrow x = 1$**:**
In such a scenario, no additional constraints are imposed, and any $A \in \mathbb{U}_n^X$ will satisfy the conditions (5.25) and (5.26).

Indeed, the condition imposed on the gate A and parameters (x, y, z), are both necessary and sufficient. By injecting any of the solutions into the constraints (5.25) and (5.26), we ensure that our conditions cover all possible cases. Thus, they represent both necessary and sufficient conditions.

□

In conclusion, we have characterized the controlled operations which preserve X-errors and do not propagate them from the target to the control unit. We now proceed to use these gates to design noise-resilient circuits in the next section.

5.3.4 Scalable Noise-Resillient Hadamard Test

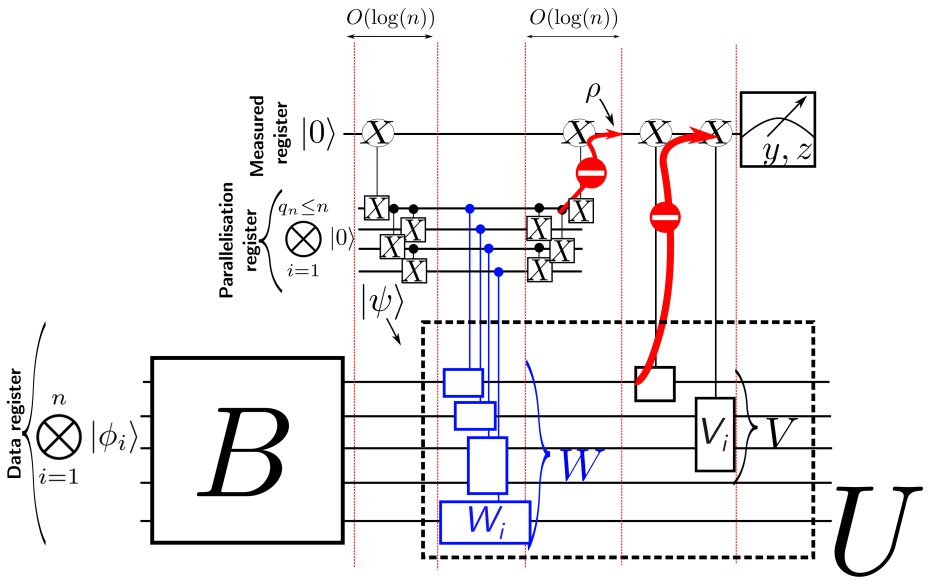

Figure 5.4: [FAND+23] This figure demonstrates how to implement $U = W \times V = \bigotimes_{i=1}^{N_W} W_i \times \bigotimes_{i=1}^{N_V} V_i$, as outlined in Theorem 5.5, in a manner that enhances resistance to noise. The unitary operations highlighted in blue represent the controlled Hermitian unitaries W_i, which collectively form the controlled unitary W. Similarly, the operations highlighted in black denote the controlled V_i unitaries, which together implement the controlled V. A central element of this construction is the parallelization register [MN01], which enables the implementation of the unitary W across all qubits in the data register while preserving noise resilience. Consequently, this setup allows the performance of a Hadamard test where U acts on every qubit in the data register. The main advantage of the parallelization register is that, despite introducing additional components that could lead to bit-flip errors, these errors are specifically designed to remain contained and do not affect the measurement register. They effectively commute with the last $c_X X$ gate applied inbetween the measured and parallelization register, thus avoiding any impact on the measurement results. This method illustrates the benefit of increasing the number of qubits to reduce the interaction of the measurement register from a polynomial number of gates to just $O(\log(n))$ gates, demonstrating a worthwhile trade-off regarding the implementation of the controlled W. While the parallelization register prevents bit-flip errors from reaching the measurement register, it does allow phase-flip errors to propagate. We exploit this property to effectively identify and evaluate whether there is an excessive rate of phase-flip errors in the benchmarking protocol of the next chapter.

We will now outline the essential elements that ensure the existence of noise-

resilient Hadamard tests. First, we assume the following conditions [FAND+23]:

1. Individual gate errors, as well as measurement and state-preparation errors, occur with a probability smaller than $p < 1/2$.
2. only X-errors occur in the algorithm (which can be satisfied with the assumptions of Property 5.1).
3. these errors cannot propagate from the data to the measured register
4. the number of interactions of the measured register with the data register satisfies $L_n = O(\log(n))$, implying that the measured register will only be impacted by X-errors introduced at $O(\log(n))$ locations.

Under the fulfillment of conditions (1-4), the reduced state ρ will satisfy Eq. (5.1), and the value of α_n can be efficiently computed classically. Moreover, α_n will satisfy the condition $\alpha_n \geq 1/\text{poly}(n)$. As described in the paragraph preceding 5.3.1, this property ensures the scalability of the algorithm for estimating $\langle\psi| U |\psi\rangle$. To be more precise, the following theorem is applicable [FAND+23].

Theorem 5.5 (Hadamard test resilient to biased noise)**.** *Let:*

$$|\psi\rangle = B \bigotimes_{i=1}^{N_B} |\varphi_i\rangle\,,\ U = W \cdot V\,,$$
$$W \equiv \prod_{i=1}^{N_W} W_i,\ V \equiv \prod_{i=1}^{N_V} V_i\,, \tag{5.27}$$

where B denotes a product of local bias-preserving gates, and the gates V_i and W_i are local gates belonging to $\mathbb{U}_n^X$. It is important to note that the gates W_i are further assumed to be Hermitian.

In our analysis, we incorporate the local bias noise model described in Eq. (5.7). We also consider the possibility that state preparation, measurements, and each non-trivial gate applied to the measurement register may introduce a bit-flip on the measured register, with a probability not exceeding $p < 1/2$.

Under these specified conditions, we can construct a quantum circuit that performs a noise-resilient Hadamard test. In the presence of noise, the reduced state ρ of this circuit will satisfy Eq. (5.1), and the parameter α_n will be bounded by $\alpha_n \geq (1-2p)^{O(N_V)}$. Moreover, the value of α_n can be efficiently computed classically. Thus, if the number of gate interactions N_V is on the order of $O(\log(n))$, it becomes feasible to implement the Hadamard test in a manner that running the algorithm a polynomial number of times would be enough to estimate the real and imaginary parts of $\langle\psi| U |\psi\rangle$ to a desired precision ε with high probability.

We construct the circuit through the scheme as shown in Fig. 5.4.

Proof. [FAND+23] Our main focus lies on Figure 5.4. In the diagram, the controlled unitaries outlined in blue correspond to the W_i gates, while the ones outlined in black correspond to the V_i gates. First, we will detail the implementation of the U gate as defined in Theorem 5.5. Then, we will demonstrate the noise-resilience of the circuit. We can assign indices m, p, and d to represent the measured, parallelization, and data registers, respectively. Let $q_n \leq n$ be the number of qubits in the parallelization register. At this stage, the full state of the quantum computer, just before the application of any controlled W_i or V_i, can be described as an entangled state:

$$|\Psi\rangle_0 \equiv \frac{|+\rangle_m |0\rangle_p^{\otimes q_n} + |-\rangle_m |1\rangle_p^{\otimes q_n}}{\sqrt{2}} \otimes |\psi\rangle_d \,. \tag{5.28}$$

This entanglement has been generated through a sequence of cNOT and $c_X X$ gates. (see Figure 5.4). Now, we apply the controlled W_i and the state transforms to:

$$\begin{aligned} &|\Psi\rangle_0 \rightarrow \\ &\frac{|+\rangle_m |0\rangle_p^{\otimes q_n} |\psi\rangle_d + |-\rangle_m |1\rangle_p^{\otimes q_n} \prod_{i=1}^{N_W} W_i |\psi\rangle_d}{\sqrt{2}}. \end{aligned} \tag{5.29}$$

The parallelisation register is then disentangled from the rest of the system by applying the reverse sequence of cNOT and $c_X X$ gates. After discarding this state, we apply the final sequence of coherently controlled V_i gates to obtain the required final state for the Hadamard test:

$$|\Psi\rangle_f = \frac{|+\rangle_m |\psi\rangle_d + |-\rangle_m U |\psi\rangle_d}{\sqrt{2}}, \tag{5.30}$$

where

$$U = \prod_{i=1}^{N_W} W_i \times \prod_{i=1}^{N_V} V_i. \tag{5.31}$$

It proves that the appropriate operation is implemented. The parallelisation technique we use is based on the work by Moore and Nilsson [MN01].

To establish the noise-resilience of this circuit, we define p_i as the probability that the i'th gate applied on the measured register introduces a bit-flip error at that particular point. For a multi-qubit gate $\mathcal{G}_i$, where the measured register is the first qubit in the tensor decomposition, we define the probability p_i as follows:

$$p_i \equiv \sum_{(1,\alpha') \subset \mathrm{supp}(\mathcal{G}_i)} p_{1,\alpha'}^{\mathcal{G}_i} \tag{5.32}$$

To clarify the notation with an example, let's consider a two-qubit gate with the following noise model:

$$\mathcal{N}(\rho) = \sum_{\substack{0 \leq i_1 \leq 1 \\ 0 \leq i_2 \leq 1}} p_{i_1,i_2} (\sigma_{i_1} \otimes \sigma_{i_2}) \rho (\sigma_{i_1} \otimes \sigma_{i_2}), \tag{5.33}$$

The probability of encountering a bit-flip error on the measured register (first tensor product) would be the sum of the probabilities $\sum_{0\leq i_2\leq 1} p_{1,i_2}$. In the subsequent analysis, we will also take into account noisy state preparation and measurements in our derivation.

It is worth noting that throughout the entire algorithm, only X-errors can occur. This is a consequence of utilizing the noise model described around (5.7), coupled with the fact that all of our gates are bias-preserving (as indicated by theorem 5.1 and 5.4). Another observation arising from the mentioned property is that the reason the W_i gates must be Hermitian is that they are coherently controlled on the Z basis. At this point, it is important to emphasize that the gates interacting with the measured register cannot propagate X-error from either the parallelization or the data register towards the measured register. This property is a direct consequence of theorem 5.4. This last point is of crucial importance: the introduction of the parallelization register allows us to trade space for time. While it introduces additional qubits in the algorithm, it enables the measured register to interact with only $O(N_V) = O(\log(n))$ gates instead of $O(N_W) \subset \text{poly}(n)$ gates. This property is the key to ensuring the noise-resilience of our circuit. The essential aspect of this tradeoff lies in the fact that errors arising in the parallelization register do not propagate to the measured register. This is due to the commutation of these errors with the only gate that interacts with the measured register, namely the $c_X X$ gate. Without the introduction of the parallelization register, the measured register would be susceptible to errors at approximately $O(N_W)$ more locations, which would compromise the scalability of the algorithm for $N_W = \text{poly}(n)$. Our observations remain valid in this context, with the addition of one further assumption that waiting locations (identity gates) can also introduce noise, as discussed in the subsequent section. Despite this, we have demonstrated that all the errors capable of affecting the measurements are directly generated within the measured register.

The final part of the proof involves computing the effect of noise on the measured register. After the i'th gate interacts with the measured register, the following bit-flip channel is applied to the measured register, for some p_i, (and σ denotes some density matrix):

$$\Lambda_i(\sigma) = (1 - p_i)\sigma + p_i X\sigma X. \tag{5.34}$$

This noise model naturally accounts for the errors introduced after state preparation. Furthermore, we can use this noise channel to model noisy measurements as well. In this model, noisy measurements are represented as perfect measurements followed by a probability of flipping the measured outcome. Since we measure Pauli Y and Z observables, we can equivalently represent them as perfect measurements followed by a bit-flip channel. This allows us to incorporate the measurement noise into our noise model consistently. Now, we exploit the fact that the bit-flip channel commutes with every gate applied to the measured register since these gates are coherently controlled in the X-basis. Therefore, denoting N_n as the number of gates applied to the measured register, (including state preparation and measurement), the full protocol is equivalent

as performing noiseless measurements in Y or Z bases on the state:

$$\rho = (\Lambda_{N_n} \circ ... \circ \Lambda_1)(\rho_{\text{ideal}}), \tag{5.35}$$

$$\rho_{\text{ideal}} = \frac{1}{2}(I + yY + zZ). \tag{5.36}$$

It is straightforward to show that:

$$\rho = \frac{1}{2}(I + \alpha_n(yY + zZ)),$$

$$\alpha_n = \prod_{i=1}^{N_n}(1 - 2p_i). \tag{5.37}$$

Under the assumption that identity gates are noiseless, we can determine that the total number of gates applied on the measured register denoted as N_n, is $N_n = N_V + 2 + 2 = O(N_V)$. Here, the first "+2" accounts for state preparation and measurement, while the last one corresponds to the implementation of the $c_X X$ gates using the parallelisation register, as depicted in Figure 5.4.

Under the assumption that identity gates are noisy, we explore this scenario in the following section of the proof. However, it's important to note that α_n remains efficiently computable classically, as long as N_n is polynomial in n. If we define p as the maximum probability of introducing a bit-flip error for any gate on the measured register ($p = \max_i p_i$), we can observe that $\alpha_n \geq (1-2p)^{O(N_V)}$. As a result, if $N_V = O(\log(n))$, α_n decreases at most at a polynomial rate, and we can still efficiently compute it.

Following the explanations provided after Eq. (5.1) in the main text, this implies that it is feasible to implement the Hadamard test in a manner that conducting the algorithm poly(n) times is sufficient to estimate the real and imaginary components of $\langle\psi| U |\psi\rangle$ with an accuracy of ε and a probability greater than $1 - \delta$.

Extension of Theorem 1 for noisy identity gates. In the context of Theorem 1, we assumed that any trivial gate (identity) applied on the measured register was noiseless. However, the inclusion of the parallelization register may introduce a considerable number of "waiting" locations for the measured qubit. These waiting locations are present during the initialization of the parallelization register or for the implementation of W gates, for instance. Thankfully, if we assume that the depth of W is in $O(\log(n))$, we can incorporate noisy identity gates into our reasoning without altering any of our conclusions. To simplify the explanation, we consider that all gates in the computation, including identity gates, have the same duration. Let p_I be the probability that an identity gate introduces a bit-flip, and let N_I be the total number of identity gates we apply on the measured register. In such a scenario, the only modification to the proof of Theorem 1 would be that:

$$\alpha_n = (1 - 2p_I)^{N_I} \prod_{i=1}^{N_n}(1 - 2p_i). \tag{5.38}$$

The term $\prod_{i=1}^{N_n}(1-2p_i)$ corresponds to the overall probability that the non-trivial gates, including state preparation and measurements, introduce bit-flip errors. The extra factor $(1-2p_I)^{N_I}$ represents the noise contribution from identity gates. In cases where W has a logarithmic depth ($O(\log(n))$), it is possible to execute the algorithm with $N_I = O(\log(n))$. (see figure 5.4): the algorithm would still be noise-resilient. Indeed, if the depth of W remains in $O(\log(n))$, α_n would decrease at a polynomial rate, and it can still be efficiently computable. However, if the depth of W exceeds $O(\log(n))$, the presence of identity gates could result in an exponential decay of α_n, which would undermine the scalability of the algorithm. For instance, if the depth of W grows linearly with n, then the number of identity gates N_I would also grow proportionally to n, and this could lead to an exponential decay of α_n, making the algorithm impractical for large n. Therefore, it is crucial to ensure that the depth of W remains in $O(\log(n))$ to maintain the noise-resilience and scalability of the algorithm. Indeed, the ability to create and disentangle the entanglement between the measured and parallelization registers in $O(\log(n))$ depth is a crucial factor in our discussion [MN01], as also illustrated on the figure 5.4.

□

It is important to emphasize that the Hadamard test is executed using the circuit depicted in Figure 5.4, which incorporates a parallelization register. This parallelization register enables the implementation of a unitary operation W (as mentioned in the theorem) that acts on all the qubits in the data register. Simultaneously, it maintains a logarithmic number of interactions with the measured register, which is necessary to maintain noise resilience [FAND+23].

Next, we will show that this class of biased noise resilient Hadamard tests can be efficiently simulated on a classical platform.

5.3.5 Classical Simulation

The previous observation shows that specific limited versions of Hadamard tests can withstand bit-flip errors occurring during the execution of quantum circuits. The upcoming theorem proves that the calculations done by such a restricted Hadamard test can unfortunately be efficiently simulated on a classical computer having access to polynomial resources [FAND+23].

Note: For the sake of clarity in the main text, our primary emphasis is on Pauli noise. However, generalization of the results and proofs for the most general perfectly biased noise (noises with their Kraus operators being a linear combination of the bit-flip channels) would be straightforward. We refer to the appendix A for more details regarding the general case.

Theorem 5.6 (Efficient classical simulation of restricted Hadamard test)**.** *Let $B \in \mathbb{B}_n, U \in \mathbb{U}_n^X$ be n qubit unitaries specified by R_B and R_U local qubit gates (belong-*

ing to respective classes $\mathbb{B}_n$ and $\mathbb{U}_n^X$). Let $|\psi_0\rangle = |\varphi_1\rangle|\varphi_2\rangle\cdots|\varphi_n\rangle$ be an initial product state. Then, there exists a randomized classical algorithm C that takes classical specifications of circuits defining B, U, and the initial state $|\psi_0\rangle$ as input. The algorithm efficiently and with high probability computes an additive approximation to the value of $\langle\psi_0| B^\dagger U B |\psi_0\rangle$. Specifically, we have

$$\Pr\left(|\langle\psi_0| B^\dagger U B |\psi_0\rangle - C| \leq \varepsilon\right) \geq 1 - \delta\,, \tag{5.39}$$

while the running time is $T = O\left(\frac{R_B+R_U+n}{\varepsilon^2}\log(1/\delta)\right)$.

Proof. [FAND+23] We observe that the operator $B^\dagger U B \in \mathbb{U}_n^X$ is diagonal in the local X basis. Therefore:

$$\langle\psi_0| B^\dagger U B |\psi_0\rangle = \mathrm{Tr}\left(B^\dagger U B\, \rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n\right)\,, \tag{5.40}$$

where the ρ_i are states that act on the i'th qubit and are defined as follows: $\rho_i = \frac{1}{2}|\varphi_i\rangle\langle\varphi_i| + \frac{1}{2}X|\varphi_i\rangle\langle\varphi_i|X$ (dephased version of $|\psi_i\rangle$ in the X basis). Each ρ_i can be expressed as $\rho_i = p_i^+|+\rangle\langle+| + p_i^-|-\rangle\langle-|$, where the computation of all the probabilities $p_i^\pm$ can be performed in $O(n)$ time. By exploiting the equation (5.40) and the specific structures of the unitaries B and U, we can develop an efficient sampling technique to estimate $\langle\psi_0| B^\dagger U B |\psi_0\rangle$. To this end, we decompose

$$\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n = \sum_{\mathbf{s}\in\{+,-\}^n} p_{\mathbf{s}} |\mathbf{s}\rangle\langle\mathbf{s}| \tag{5.41}$$

where $p_{\mathbf{s}} = \prod_{i=1}^n p_i^{s_i}$ is a product distribution. When we substitute equation (5.41) into the right-hand side of equation (5.40), we get the following expression:

$$\langle\psi_0| B^\dagger U B |\psi_0\rangle = \sum_{\mathbf{s}\in\{+,-\}^n} p_{\mathbf{s}}\, \mathrm{Tr}\left(UB|\mathbf{s}\rangle\langle\mathbf{s}|B^\dagger\right)\,. \tag{5.42}$$

The expression above can be further simplified by observing that $B|\mathbf{s}\rangle\langle\mathbf{s}|B^\dagger = |\sigma_B(\mathbf{s})\rangle\langle\sigma_B(\mathbf{s})|$ (as stated in Property 5.2), and by utilizing $U \in \mathbb{U}_n^X$, which implies $U|\mathbf{s}\rangle = \lambda_U(\mathbf{s})|\mathbf{s}\rangle$ for all $|\mathbf{s}\rangle$. By putting all these observations together, we arrive at the following expression:

$$\langle\psi_0| B^\dagger U B |\psi_0\rangle = \sum_{\mathbf{s}\in\{+,-\}^n} p_{\mathbf{s}} \lambda_U(\sigma_B(\mathbf{s}))\,. \tag{5.43}$$

The equation above reveals that the random variables $x_{\mathbf{s}} = \mathrm{Re}(\lambda_U(\sigma_B(\mathbf{s})))$ and $y_{\mathbf{s}} = \mathrm{Im}(\lambda_U(\sigma_B(\mathbf{s})))$ serve as unbiased estimators for the real and imaginary parts of $\langle\psi_0| B^\dagger U B |\psi_0\rangle$, respectively. We can summarize the approach in a simple three-step algorithm for constructing estimators of the quantities of interest:

1. generate $\mathbf{s} \sim \{p_{\mathbf{s}}\}$
2. compute $\mathbf{s}' = \sigma_B(\mathbf{s})$
3. evaluate $\lambda_U(s')$ in order to compute $x_{\mathbf{s}}, y_{\mathbf{s}}$.

Applying Hoeffding's inequality to the random variables $x_{\mathbf{s}}$ and $y_{\mathbf{s}}$, which take values in the interval $[-1, 1]$, allows us to conclude that by repeating the above procedure $O((1/\varepsilon^2)\log(1/\delta))$ times and calculating sample means, we can obtain an ε-accurate estimation of $\langle\psi_0| B^\dagger U B |\psi_0\rangle$ with a probability of at least $1-\delta$. What remains is to show the classical computability of the steps 1-3. Generating a single sample of $\mathbf{s}$ takes $O(n)$ time because $\mathbf{s}$ follows a simple product measure on n bits. In step (ii), we rely on Property 5.2 and break down B into a sequence of local gates R_B from $\mathbb{B}_n$. As each gate operates locally, it causes a manageable transformation on $|\mathbf{s}\rangle$ affecting only a constant-sized subset β of qubits. Accordingly, the computational complexity of step (ii) is $O(R_B)$. Similarly, for step (iii), we decompose the unitary U into a sequence of local unitaries g_j from $\mathbb{U}_n^X$. As each g_j acts locally on a subset of qubits, the overall computational cost of implementing (iii) is $O(R_U)$, where R_U represents the size of this local decomposition. Naturally, we have $U|\mathbf{s}'\rangle = \prod_{j=1}^{R_U} \lambda_j(\mathbf{s}')|\mathbf{s}'\rangle$, where $\lambda_j(\mathbf{s}')$ are defined by $g_j|\mathbf{s}'\rangle = \lambda_j(\mathbf{s}')|\mathbf{s}'\rangle$. The local nature of gates $g_j \in \mathbb{U}_n^X$ allows for the efficient computation of individual eigenvalues $\lambda_j(\mathbf{s}')$, taking only constant time. Moreover, this calculation depends solely on the few bits represented by $\mathbf{s}'$.

In conclusion, considering the runtimes estimated for each step of the protocol, we can combine them to obtain the desired result. The overall cost of the algorithm can be expressed as $O(n + R_B + R_U)$, which ensures efficient computation for estimating the quantities of interest. □

5.4 Discussion

We study quantum characteristics within the DQC1 model, which allows us to estimate the normalized trace of an n-qubit unitary that can be efficiently implemented using quantum gates. Our analysis shows that the DQC1 circuit can be realized efficiently with negligible quantum resources and correlations. This conclusion applies to a broad range of resource and correlation measures, including multipartite entanglement, mutual information, quantum coherence, and purity. We anticipate that the methods we have presented can be extended to other quantum algorithms that operate on mixed states, including the mixed-state version of Shor's algorithm [PP00, PP02]. The techniques and insights gained from our analysis can potentially be generalized to understand and characterize the behavior of these algorithms in the presence of noise and imperfections. The role of coherence in the implementation of Shor's algorithm has been the subject of the recent study, as highlighted in the work by Ahnefeld et al. [ATE+22b].

We further examine qubits with biased noise, which is relevant for superconducting cat qubits. This feature allows us to create a series of noisy Hadamard tests that include certain entangling and non-Clifford gates, which can be performed with high reliability and only a polynomial increase in the number of algorithm repetitions. On the other hand, we identified classical algorithms capable of efficiently simulating both

the noisy and noiseless versions of our specific Hadamard test variants. We recommend using these algorithms to benchmark the noise bias in large circuits. Although our circuits can be simulated efficiently, they exhibit strong noise resilience, which has been rigorously demonstrated even in the presence of Pauli bit-flip noise [2]. It raises the natural question of whether extensions of our work could result in noise-resilient circuits with computational significance. Our research also addresses this fundamental question and provides insights into it. To explore this, we note that the set of bias-preserving gates, $\mathbb{B}_n$, includes $c_X X$ gates. When these gates are combined with initial states $|0\rangle$ or $|1\rangle$, they can generate any graph state in the local X basis. It is known that typical n-qubit stabilizer states exhibit strong multipartite entanglement [SL06]. This, along with the fact that arbitrary graph states can be transformed into stabilizer states locally [VdNDDM04], indicates that bias-preserving circuits have the potential to generate a wide variety of highly entangled states. However, it is important to note that, for the specific computational problem we tackle in this work, these graph states do not offer practical utility. We can also include various non-Clifford gates in our design. For our objectives, any unitary $U \in \mathbb{U}_n^X$ can affect all qubits in the data register in a meaningful way. There are no restrictions on the number of gates in the preparation unitary B, and the circuit is designed to scale effectively even with noisy measurements.

In the next chapter, we demonstrate that the simulability of the discussed scalable Hadamard test proposes a simple algorithm to benchmark the biased noise model at the scale of large and complicated quantum circuits. By assessing the entire circuit, our benchmark can detect errors that are typically not visible in individual gate tomography, such as crosstalk or correlated errors suggesting it could be a valuable tool for scaling up cat qubit technologies.

Contribution: The results in this chapter regarding the DQC1 model of computation was originally found by Dr. Tulja Varun Kondra under supervision of Prof. Alexander Streltsov and I was involved in the technical discussions. The other results are based on the work in the paper *Scalable noisy quantum circuits with biased-noise qubits*. I contributed to all the technical discussions in this work. Moreover I did all the work on how the noise would propagate through controlled operations specifically the section 5.3.3.

[2] In cases where a non-Pauli noise model exactly meets the criteria outlined in definition 6.1, the results from the noisy algorithm typically won't align with those from the ideal noiseless algorithm. Despite this, such a mismatch does not affect the validity of our benchmarking approach.

Chapter 6

Benchmarking Biased Noise at the Scale of the Whole Circuit

The whole strategy of cat qubits relies on having phase-flip errors negligible (hence the consequence of imperfect bias is not noticeable), therefore we need a tool to check and ensure that the bias remains largely effective in big circuits[FAND+23]. There does not exist any protocol to do this today. In this chapter using the simulability of the circuits discussed in the previous chapter, we propose a simple protocol to benchmark the biased noise model in large-scale and complex quantum circuits, capable of detecting some violations of noise properties. This benchmark is particularly beneficial because it is scalable and capable of detecting collective noise effects that individual gate analysis cannot reveal, such as crosstalk and correlated errors. The core idea of the benchmark is to compare experimental results with a classical simulation that assumes each gate in the complete algorithm has the same noise model as determined from individual gate tomography. If there is a discrepancy between the simulation and the experiment, it suggests that collective effects usually invisible from individual gate's tomography, are reducing the hardware quality when running a full algorithm, posing a potential thread to the hardware's scalability.

6.1 Efficient Simulation of Noise-Resilient Hadamard Test

In general, the noise model for biased qubits can include Kraus operators that represent the noise effects of each gate, state preparation, and measurements, which do not precisely match a Pauli bit-flip channel. This means that the Kraus operators could be expressed as a linear combination of Pauli bit-flips, following the definition 6.1. Fur-

thermore, the bias might not be perfect. This implies that the Kraus operators representing the noise maps could have a non-zero Hilbert-Schmidt inner product with Pauli operators outside of $\mathbb{P}_n^X$. For instance, the Kraus operators might exhibit a non-zero overlap with a multi-qubit Pauli operator P that contains at least one Z or Y component in its tensor product, like $P = X \otimes Y$. If this happens, we will describe the noise model as also generating phase-flip errors. Our benchmarking protocol is capable of identifying deviations from the assumed noise model in such a broader scenario. The protocol relies on the ability to simulate the results of the Hadamard test in a noisy environment, assuming that the bias is perfect but not strictly Pauli, meaning it conforms to the criteria outlined in definition 6.1 [FAND+23]. This classical simulation uses the noise model of each individual gate, approximated to match that of a perfect bias, as its input. If the results of this classical simulation deviate from the experimental outcomes beyond some error budget that we will quantify, it would indicate that the noise model described by individual gate tomography does not accurately reflect the experimental conditions [FAND+23]. This error budget pertains to how closely the noise model derived from individual gate tomography approximates the perfect bias model. The formal details of the protocol are outlined in theorem 6.3. We will now present the definitions and theorems required for our discussion [FAND+23].

Definition 6.1. Perfect bias

Let G represent a quantum channel that describes either a noiseless unitary gate from the available gate set or a single-qubit state preparation, and let $\mathcal{E}_{\mathcal{G}}$ denote its noisy implementation in the laboratory. We say that the noisy implementation of the gate (or state preparation), $\mathcal{E}_{\mathcal{G}}$, follows a perfectly biased noise model if $\mathcal{E}_{\mathcal{G}} = \mathcal{N}_{\mathcal{G}} \circ \mathcal{G}$, where the noise map $\mathcal{N}_{\mathcal{G}}$ is a CPTP map that admits the following Kraus decomposition:

$$\begin{aligned} \mathcal{N}_{\mathcal{G}}(\rho) &= \sum_j K_j^{\mathcal{G}} \rho (K_j^{\mathcal{G}})^\dagger \,, \\ K_j^{\mathcal{G}} &= \sum_{\alpha \subset \mathrm{supp}(\mathcal{G})} c_\alpha^j X_\alpha, \end{aligned} \tag{6.1}$$

where $X_\alpha = \prod_{i \in \alpha} X_i$, $\forall j, c_\alpha^j \in \mathbb{C}$ and $\sum_j (K_j^{\mathcal{G}})^\dagger K_j^{\mathcal{G}} = \mathbb{I}_{\mathrm{supp}(\mathcal{G})}$, $\mathbb{I}_{\mathrm{supp}(\mathcal{G})}$ being the identity operator applied on the qubits where $\mathcal{G}$ acts non-trivially. A quantum map that meets the criteria specified in (6.1) will be referred to as perfectly biased [FAND+23].

To model a noisy single-qubit measurement, we consider it as an ideal measurement preceded by a perfectly biased single-qubit CPTP map applied to the qubit. This CPTP channel, known as the noise map of the measurement, accounts for the noise induced by the measurement [FAND+23].

To simplify the presentation of the upcoming theorems, we will assume that state preparation and measurements are noiseless. This assumption does not remove any generality under the condition that all gates are bias-preserving (which is assumed throughout this section) and the noise model is described by the criteria in Definition 6.1. This is because noise arising from state preparation (resp measurement) can

be effectively absorbed by updating the noise map of the following (resp preceding) gate . The revised noise map will remain consistent with the given definition in 6.1 [FAND+23]. To be more specific, let $\mathcal{N}_{\text{prep}}$ represent the noise map associated with state preparation, and $\mathcal{G}_{\text{prep}}$ denote the map for an ideal, noiseless state preparation. Then, the noisy state preparation is given by $\mathcal{N}_{\text{prep}} \circ \mathcal{G}_{\text{prep}}$. Similarly, if $\mathcal{N}_{\mathcal{G}_i}$ is the noise map for a quantum gate described by the unitary map $\mathcal{G}_i$, the noisy version of this quantum gate is represented by $\mathcal{N}_{\mathcal{G}_i} \circ \mathcal{G}_i$. Assuming this gate follows immediately after state preparation, the noise from state preparation can be accounted for by adjusting the gate's noise map to: $\mathcal{N}_{\mathcal{G}_i} \to \mathcal{N}'_{\mathcal{G}_i} \equiv \mathcal{N}_{\mathcal{G}_i} \circ \mathcal{G}_i \circ \mathcal{N}_{\text{prep}} \circ \mathcal{G}_i^{\dagger}$. Consequently, the total noise from state preparation to the first gate is given by: $\mathcal{N}'_{\mathcal{G}_i} \circ \mathcal{G}_i \circ \mathcal{G}_{\text{prep}} = \mathcal{N}_{\mathcal{G}_i} \circ \mathcal{G}_i \circ \mathcal{N}_{\text{prep}} \circ \mathcal{G}_{\text{prep}}$. Hence, if the bias is perfect and all gates are bias-preserving, $\mathcal{N}'_{\mathcal{G}_i}$ will still represent a perfectly biased noise model. The same reasoning applies to noisy measurements.

Theorem 6.1. *[FAND+23] Efficient classical simulation of a noisy Hadamard test under perfect bias.*

Let B, $U = W.V \in \mathbb{U}_n^X$, and N_W, N_V be such that they satisfy the conditions outlined in Theorem 5.5. Assume $|\psi_0\rangle = |\varphi_1\rangle|\varphi_2\rangle \cdots |\varphi_n\rangle$ represents the initial product state for the data register. Given that $N_V = O(\log(n))$, each gate's noise model adheres to the perfect bias definition 6.1, the total gate count is poly(n), and state preparation and measurements are considered noiseless (as stated earlier in the comments before Theorem 6.1), there exists a randomized classical algorithm C. This algorithm C requires:

1. *(I) the classical descriptions of the circuit performing the Hadamard test for the chosen $(B, U, |\psi_0\rangle)$,*
2. *(II) the quantum channel characterizing each gate's noise model, and*
3. *(III) the initial state $|\psi_0\rangle$.*

With these inputs, C can efficiently and with high probability produce an additive approximation to $Tr(P_1\rho_X)$, where ρ_X is the reduced density matrix of the measured register at the end of the noisy algorithm, and P_1 is a single-qubit Pauli matrix.

Specifically, we have:

$$\Pr\left(|Tr(P_1\rho_X) - C| \leq \varepsilon\right) \geq 1 - \delta\,, \tag{6.2}$$

while the running time is $T = O((1/\varepsilon^2)\log(1/\delta) \times poly(n))$.

We refer to the appendix A for the proof of the theorem. The theorem 6.1 establishes that the computation performed by the restricted Hadamard test we prescribed, even in the presence of a more general noise model from definition 6.1 can be efficiently simulated on a classical computer with polynomial effort.

In what follows, we will use this simulation algorithm to perform our benchmarking protocol. This protocol is designed to validate the assumption of bias noise across the

entire circuit, even for complex cases. It allows us to assess the impact of biased noise and gain valuable insights into the performance of the hardware.

6.2 Validity of the Perfect Bias Approximation

Consider $\mathcal{G}$ as the quantum map that represents a (noiseless) unitary gate, and let $\mathcal{E}_\mathcal{G}$ denote the quantum map of its noisy implementation. The noise map $\mathcal{N}_\mathcal{G}$ can be defined as:

$$\mathcal{E}_\mathcal{G} = \mathcal{N}_\mathcal{G} \circ \mathcal{G}. \tag{6.3}$$

In this study, we assume that each gate's noise map is perfectly biased. Nevertheless, this is an approximation because the gate's bias will not be ideal. Consequently, the Kraus operators of $\mathcal{N}_\mathcal{G}$ are not strictly linear combinations of elements in $\mathbb{P}_n^X$. We define $\mathcal{N}_{X,\mathcal{G}}$ a CPTP map having a perfect bias (according to definition 6.1) that approximates $\mathcal{N}_\mathcal{G}$. More precisely, we define $\Delta\mathcal{N}_\mathcal{G}$ and $\Delta\mathcal{E}_\mathcal{G}$ as:

$$\mathcal{N}_\mathcal{G} = \mathcal{N}_{X,\mathcal{G}} + \Delta\mathcal{N}_\mathcal{G} \tag{6.4}$$

$$\mathcal{E}_\mathcal{G} = \mathcal{E}_{X,\mathcal{G}} + \Delta\mathcal{E}_\mathcal{G}, \tag{6.5}$$

with:

$$\mathcal{E}_{X,\mathcal{G}} \equiv \mathcal{N}_{X,\mathcal{G}} \circ \mathcal{G}. \tag{6.6}$$

A smaller diamond norm of $\Delta\mathcal{E}_\mathcal{G}$ indicates a better approximation of a perfect bias. The theorem 6.2 helps quantify the maximum number of gates, N, that can be implemented without the assumption of a perfect bias causing the measurement outcomes of the Hadamard test to deviate beyond an error budget ε [FAND+23].

Theorem 6.2. *Quality of the perfect-bias assumption*

Let N represent the total number of gates in the algorithm. The noise map for the implementation of the i'th gate ($i \in [1, N]$) is referred to as $\mathcal{N}_{\mathcal{G}_i}$, and the unitary implementation of this gate is denoted by $\mathcal{G}_i$. We define $\mathcal{N}_{X,\mathcal{G}i}$ as a noise map that approximates $\mathcal{N}_{\mathcal{G}_i}$, so that $\mathcal{N}_{X,\mathcal{G}_i}$ has a perfect bias as per definition 6.1. We denote ρ_X as the final density matrix of the measured register when the noise model of each gate "i" is given by $\mathcal{N}_{X,\mathcal{G}_i}$. We refer to ρ as the final density matrix of the measured register when the noise model applied to the gate "i" is $\mathcal{N}_{\mathcal{G}_i}$. If:

$$\max_i \|\mathcal{N}_{X,\mathcal{G}_i} - \mathcal{N}_{\mathcal{G}_i}\|_\diamond \leq \frac{\varepsilon}{\sqrt{2}N}, \tag{6.7}$$

where $\|.\|_\diamond$ represents the diamond norm [Wil11],we can then establish a bound on the error in the measurement outcome probability of the Hadamard test based on our approximation of perfect bias. In particular, for any single-qubit Pauli P_1,

$$|Tr(\rho P_1) - Tr(\rho_X P_1)| \leq \varepsilon \tag{6.8}$$

Proof. [FAND+23] Let $\|\cdot\|_2$ denote the Hilbert-Schmidt norm. According to the Cauchy-Schwarz inequality, we have:

$$|Tr(\rho P_1) - Tr(\rho_X P_1)| \leq \|\rho - \rho_X\|_2 \|P_1\|_2. \tag{6.9}$$

Using the fact $\|\rho - \rho_X\|_2 \leq \|\rho - \rho_X\|_1$ [CCC19], and $\|P\|_2 = \sqrt{2}$, we also have:

$$|Tr(\rho P_1) - Tr(\rho_X P_1)| \leq \sqrt{2}\|\rho - \rho_X\|_1. \tag{6.10}$$

For every $i \in [1, N]$, we define $\mathcal{E}_{\mathcal{G}_i} \equiv \mathcal{N}_{\mathcal{G}_i} \circ \mathcal{G}_i$ and $\mathcal{E}_{X,\mathcal{G}_i} \equiv \mathcal{N}_{X,\mathcal{G}_i} \circ \mathcal{G}_i$. We now assume that the noiseless algorithm performs the unitary operation $\mathcal{G}_N \circ ... \circ \mathcal{G}_1$. In the noisy algorithm we then have: $\mathcal{E}_{\text{Algo}} = \mathcal{E}_{\mathcal{G}_N} \circ ... \circ \mathcal{E}_{\mathcal{G}_1}$. Under the perfect bias assumption, the noisy algorithm performs the following $\mathcal{E}_{X,\text{Algo}} = \mathcal{E}_{X,\mathcal{G}_N} \circ ... \circ \mathcal{E}_{X,\mathcal{G}_1}$. Let ρ_0 denote the initial state of the entire algorithm, including the data, the measured register, and any potential parallelization register as illustrated in figure 5.4. Given the definitions of ρ and ρ_X, we have:

$$\begin{aligned} \|\rho - \rho_X\|_1 &= \|Tr_{\neq \text{Measured}}[(\mathcal{E}_{\text{Algo}} - \mathcal{E}_{X,\text{Algo}})(\rho_0)]\|_1 \\ &\leq \|(\mathcal{E}_{\text{Algo}} - \mathcal{E}_{X,\text{Algo}})(\rho_0)\|_1 \leq \|\mathcal{E}_{\text{Algo}} - \mathcal{E}_{X,\text{Algo}}\|_\diamond \end{aligned} \tag{6.11}$$

In (6.11), $Tr_{\neq \text{Measured}}$ refers to taking the partial trace over all degrees of freedom except the measured register. This process relies on the fact that the trace distance is non-increasing when taking a partial trace, and the last inequality follows from the definition of the diamond norm. Lastly, we apply the chaining properties of the diamond norm [GLN05], we have:

$$\|\mathcal{E}_{\text{Algo}} - \mathcal{E}_{X,\text{Algo}}\|_\diamond \leq N \max_i \|\mathcal{E}_{\mathcal{G}_i} - \mathcal{E}_{X,\mathcal{G}_i}\|_\diamond \tag{6.12}$$

Combining (6.10), (6.11), (6.12), and using the property of unitary invariance of the diamond norm (i.e. $\|\mathcal{E}_i - \mathcal{E}_{X,\mathcal{G}_i}\|_\diamond = \|\mathcal{N}_{\mathcal{G}_i} - \mathcal{N}_{X,\mathcal{G}_i}\|_\diamond$), we deduce:

$$|Tr(\rho P) - Tr(\rho_X P)| \leq \sqrt{2}N \max_i \|\mathcal{N}_{\mathcal{G}_i} - \mathcal{N}_{X,\mathcal{G}_i}\|_\diamond \tag{6.13}$$

Hence, if $\max_i \|\mathcal{N}_{\mathcal{G}_i} - \mathcal{N}_{X,\mathcal{G}_i}\|_\diamond \leq \varepsilon/(\sqrt{2}N)$, then, $|Tr(\rho P) - Tr(\rho_X P)| \leq \varepsilon$ which proves the relation in 6.2. □

6.3 Benchmarking Protocol

Evaluating quantum gates individually poses a problem because it overlooks the cumulative interactions that arise when a complete circuit, consisting of numerous gates arranged both sequentially and in parallel, is executed. One instance is correlated errors, which are typically undetectable through individual gate tomography. Another

instance of collective effects is "scale-dependent noise," where the noise intensity per gate varies based on the number of qubits or the adjacent gates utilized in the algorithm [ZLL+22, KLR+20, SR20]. It's important to recognize that scale-dependent noise can occasionally be suggested by correlated noise models [FACW+21]. This kind of noise can be particularly harmful to biased qubits, as the initially low error rate may escalate with larger circuit scales, potentially undermining the benefits of employing these qubits. These noise behaviors pose a significant risk for both near term and the achievement of large-scale quantum computing [MMMJA20, ALL+23, ZLL+22]. Our benchmarking protocol, which we now sketch, allows us to detect some of these effects.

Theorem 6.3 (Benchmarking protocol)**.** *Consider a Hadamard test that satisfies the requirements established in Theorem 5.5. This test involves N unitary gates, each described by the unitary quantum channels $\{\mathcal{G}_i\}_{i=1}^N$. The noise associated with each gate $\mathcal{G}_i$, obtained through quantum tomography, is represented by $\mathcal{N}_{\mathcal{G}_i}$. We will assume that both state preparation and measurements are ideal and noiseless, which can be accommodated by appropriately redefining the noise maps of the quantum gates (as noted before Theorem 6.1).*

Let $\mathcal{N}_{X,\mathcal{G}_i}$ be an approximation to $\mathcal{N}_{\mathcal{G}_i}$, such that $\mathcal{N}_{X,\mathcal{G}_i}$ has a noise model satisfying definition 6.1.

Let ρ denote the density matrix of the measured register at the end of the algorithm, assuming that each gate $\mathcal{G}_i$ is affected by the noise map $\mathcal{N}_{\mathcal{G}_i}$. Define ρ_X as the reduced density matrix of the measured register under the noise map $\mathcal{N}_{X,\mathcal{G}_i}$ for each gate $\mathcal{G}_i$. Let ρ_{exp} be the reduced density matrix that precisely matches the experimental outcomes. In other words, performing the measurements used in the Hadamard test on ρ_{exp} would yield measurement results that exactly correspond to those observed experimentally.

Assume that there exists $\varepsilon > 0$ such that

$$\max_i \|\mathcal{N}_{X,\mathcal{G}_i} - \mathcal{N}_{\mathcal{G}_i}\| \leq \frac{\varepsilon}{\sqrt{2}N} \tag{6.14}$$

is satisfied. Then, for any single-qubit Pauli P_1:

$$|Tr(\rho P_1) - Tr(\rho_X P_1)| \leq \varepsilon \tag{6.15}$$

Principle of the benchmarking:

Here's how the benchmarking protocol functions: $Tr(\rho_X P_1)$ is the predicted outcome of the circuit when using a noise model for each gate that meets the criteria set out in Definition 6.1. This prediction can be calculated classically using Theorem 6.1. In contrast, $Tr(\rho_{exp} P_1)$ represents the result obtained from actual experiments. If the difference between $Tr(\rho_{exp} P_1)$ and $Tr(\rho_X P_1)$ exceeds ε, it indicates that ρ_{exp} and ρ are different, suggesting that the noise model derived from individual gate tomography (represented by the noise maps $\{\mathcal{N}_{\mathcal{G}_i}\}_{i=1}^N$) does not accurately reflect the experimental noise.

Noteworthy, $\Delta \equiv |Tr(\rho_{exp}P_1) - Tr(\rho_X P_1)| - \varepsilon$ *can quantify how strong the noise violation is at the scale of the whole circuit (if* $\Delta > 0$*). It is a consequence of the fact* $|Tr(\rho_{exp}P_1) - Tr(\rho P_1)| \geq \Delta$ *(the larger* Δ*, the larger the violation).*

Proof. The benchmarking protocol naturally follows from the theorem 6.1. If

$$\max_i \|\mathcal{N}_{X,\mathcal{G}_i} - \mathcal{N}_{\mathcal{G}_i}\| \leq \frac{\varepsilon}{\sqrt{2}N}, \tag{6.16}$$

then, from theorem 6.2, for any single-qubit Pauli P_1, we have:

$$|Tr(\rho P_1) - Tr(\rho_X P_1)| \leq \varepsilon \tag{6.17}$$

Therefore, if $|Tr(\rho_{\text{exp}}P_1) - Tr(\rho_X P_1)| > \varepsilon$, ρ_{exp} and ρ necessarily must be different. In conclusion, by utilizing a basic triangular inequality, we arrive at:

$$\begin{aligned} |Tr(\rho_{\text{exp}}P_1) - Tr(\rho_X P_1)| &\leq |Tr(\rho P_1) - Tr(\rho_X P_1)| \\ &\quad + |Tr(\rho P_1) - Tr(\rho_{\text{exp}}P_1)| \\ &\leq \varepsilon + |Tr(\rho P_1) - Tr(\rho_{\text{exp}}P_1)| \end{aligned} \tag{6.18}$$

Hence $|Tr(\rho_{\text{exp}}P_1) - Tr(\rho P_1)| \geq \Delta$. □

6.4 Estimating the Size of Implementable Circuits based on Literature

Our current aim is to estimate the largest circuit size for which our benchmarking protocol can be implemented. To do this, we need to assign specific values to $\max_i \|\mathcal{N}_{\mathcal{G}_i} - \mathcal{N}_{X,\mathcal{G}_i}\|_\diamond$ appearing in the benchmarking protocol (theorem 6.3) from the literature: This will help us calculate the maximum number of gates and state preparations, N, that can be accommodated in the circuit while ensuring that ε remains small enough to be effective [FAND+23]. We will represent quantum channels using their χ matrix. For an n-qubit quantum channel $\mathcal{N}$, the χ matrix is composed of elements χ_{ij}, such that $\mathcal{N}(\rho) = \sum_{ij}\chi_{ij}P_i\rho P_j$. In this expression, P_i denotes a basis of n-qubit Pauli matrices for the space of n-qubit operators. In [XIBJ22], numerical simulations provided the χ matrix for the noise map of a cNOT gate for superconducting cat qubits. We will base our quantitative estimates on this specific noise channel: Our estimate should be considered as a rough approximation for the maximum permissible circuit size. Since a cNOT gate is generally noisier than a single-qubit gate, it serves as a reasonable example for our calculations. However, a limitation is that [XIBJ22] (as well as other sources like [GM19, CNAA+22]) only provide absolute values for the real and imaginary components of the off-diagonal elements in the χ matrix. Consequently, we use the Pauli Twirling approximation, which involves ignoring these off-diagonal terms and approximating the noise channel as a perfectly biased Pauli channel. While this approach may underestimate the true diamond distance, it is the best we could do given

the best available method and the data that were openly accessible from the current literature. However, it's worth noting that this approximation may be quite reasonable in practice in certain situations, as the off-diagonal terms can often be much smaller compared to the dominant diagonal terms [CNAA+22]. Hence, we consider:

$$\mathcal{N}_{\text{cNOT}}(\rho) = \sum_{\substack{0\leq i\leq 3\\ 0\leq j\leq 3}} \chi^{\text{cNOT}}_{iji}(\sigma_i \otimes \sigma_j)\rho(\sigma_i \otimes \sigma_j), \tag{6.19}$$

where, from [XIBJ22] (we rounded the diagonal terms to the closest order of magnitude from their color plot), we have:

$$\begin{aligned}(&\chi_{0000},\chi_{0303},\chi_{0202},\chi_{0101},\chi_{3030},\chi_{3333},\chi_{3232},\chi_{3131},\\ &\chi_{2020},\chi_{2323},\chi_{2222},\chi_{2121},\chi_{1010},\chi_{1313},\chi_{1212},\chi_{1111})\\ &= \frac{1}{1.0012000066}(1, 10^{-9}, 10^{-10}, 10^{-4}, 10^{-9},\\ &10^{-9}, 10^{-10}, 10^{-10}, 10^{-9}, 10^{-9}, 10^{-10}, 10^{-10}, 10^{-3}, 10^{-9},\\ &10^{-10}, 10^{-4}).\end{aligned} \tag{6.20}$$

The term $1/1.0012000066$ is included to ensure that the map remains trace-preserving, as our rounding might otherwise affect this property. We note that in (6.20), while we used values from [XIBJ22], we swapped the roles of the Pauli matrices from $(I, X, Y, Z) \to (I, Z, -Y, X)$. This difference arises because [XIBJ22] considered the primary noise mechanism to be a phase-flip, whereas our work assumes it to be a bit-flip. We will approximate $\mathcal{N}_{\text{cNOT}}$ by the perfectly biased Pauli channel $\mathcal{N}_{X,\text{cNOT}}$:

$$\mathcal{N}_{X,\text{cNOT}}(\rho) = \sum_{\substack{0\leq i\leq 3\\ 0\leq j\leq 3}} \chi^{\text{cNOT}}_{X,iji}(\sigma_i \otimes \sigma_j)\rho(\sigma_i \otimes \sigma_j), \tag{6.21}$$

where:

$$\begin{aligned}(\chi^{\text{cNOT}}_{X,0000},\chi^{\text{cNOT}}_{X,0101},\chi^{\text{cNOT}}_{X,1010},\chi^{\text{cNOT}}_{X,1111}) = \frac{1}{1.0012}(1, 10^{-4},\\ 10^{-3}, 10^{-4}),\end{aligned} \tag{6.22}$$

All other coefficients for $\chi^{\text{cNOT}}_{X,iji}$ are zero. Using the formula for the diamond distance between Pauli channels [MGE12], we find that $\|\mathcal{N}_{\text{cNOT}} - \mathcal{N}_{X,\text{cNOT}}\|_\diamond \approx 1.31 \times 10^{-8}$. Since $|Tr(\rho P_1) - Tr(\rho_X P_1)| \leq 2$, a benchmark is considered useful if $\varepsilon < 2$. However, we adopt a more stringent criterion for a useful benchmark, requiring that when $Tr(\rho_X P_1)$ is approximately 1, $Tr(\rho P_1)$ should provide a reasonably accurate estimate of $Tr(\rho_X P_1)$. This implies we need $\varepsilon \ll 1$. In other words, we require our noise approximation to closely match the measurement outcomes that would be obtained with the expected noise model, which is based on individual gate tomography. Considering for instance $\varepsilon = 1/50$, the theorem 6.2 provides us:

$$N \leq 1.07 \times 10^6. \tag{6.23}$$

Note that it implicitly assumes that $\|\mathcal{N}_{\text{cNOT}} - \mathcal{N}_{X,\text{cNOT}}\|_\diamond$ gives a fair estimate to the left hand side of (6.7) (We consider this approach reasonable because a cNOT gate is generally noisier than single-qubit gates). Thus, our benchmark is suitable for circuits with up to 1.07×10^6 gates. This represents an increase of 3 to 4 orders of magnitude compared to what current hardware and circuits without noise bias can handle [PBE22, SDB+21]. Our benchmark is then in practice applicable for near and longer-term circuits.

6.5 Discussion

In this chapter, we evaluate the maximum number of quantum gates that our benchmark can handle effectively. According to existing research [PBE22], we find that, when applying the Pauli-Twirling approximation, the benchmark is suitable for circuits with up to 10^6 gates. This size is significantly larger—by three to four orders of magnitude—compared to the circuits used in contemporary experiments. This suggests that our benchmark is highly relevant for evaluating hardware reliability in the NISQ (noisy intermediate-scale quantum computation) regime and beyond, particularly for large circuits. In simple terms, our benchmark is intended for use in situations where the phase-flip error source is anticipated to be negligible which is precisely the regime where a benchmark is useful for this kind of qubits. Technically, this is due to its dependence on classical simulation, as outlined in theorem 6.1, which assumes perfect bias conditions with no phase-flips occurring in the algorithm.

Our protocol can identify some correlated errors that individual gate tomography might miss. This is because our classical simulation algorithm presumes that individual gate tomography accurately represents the noise behavior. If noise correlations are too strong, the simulation results might not align with experimental outcomes. However, we believe the most valuable aspect of our protocol is its capability to detect phase-flip errors occurring more frequently than anticipated (We note that increasing error rates with the size of the computer is a phenomenon observed experimentally in superconducting qubits [ZLL+22, KLR+20, SR20]). Indeed, such effects would typically result in discrepancies between the classical simulation and experimental results for algorithms with fewer gates than anticipated (i.e. a total number of gates, N, smaller than what (6.14) predicts). Specifically, an experimentalist should select the largest N such that the bound (6.14) is reached. If $\Delta \geq 0$, it suggests a violation of the assumptions underlying the noise model, which could signal a potential issue for the scalability of the platform. For example, when $\varepsilon = 1/50$, our quantitative analysis suggests that $N = 10^6$ would be effective.

We can also illustrate a specific circuit designed to efficiently detect phase-flip errors occurring at a higher rate than anticipated [FAND+23]. For example, consider a circuit implementing the controlled unitary $c_X U$, where $U = \otimes_{i=1}^n X_i$. This circuit satisfies the criteria specified in Theorem 5.5. (including its extension to noisy identity gates as discussed in the paragraph following Theorem 5.5): We can therefore utilize

this circuit for benchmarking purposes. It is particularly useful because it makes the measurement outcomes sensitive to phase-flip errors occurring on any qubit in the data register after the application of the preparation unitary B. This sensitivity arises because a phase-flip error on any qubit in the data register will initially propagate to the parallelization register (through the blue cNOT gates that implement $W = U = \otimes_{i=1}^{n} X_i$ as shown in Figure 5.4), and subsequently affect the measured register. Considering Figure 5.4, we use a parallelization register with n qubits, where each qubit interacts with the data register qubits through cNOT gates. For instance, if a Pauli Z error (or 'phase-flip') occurs on the second qubit (from the top) of the data register after applying B, this Z error will propagate to the second qubit (from the top) of the parallelization register, due to the cNOT gates transferring Z errors from the target qubit to the control qubit. The Z error will then spread to the top qubit of the parallelization register via the cNOT gate between the first and second qubits of this register. Eventually, the Z error will be transferred to the measured register through the final $c_X X$ gate connecting the parallelization and measured registers, where it will be converted into an X error. Consequently, the initial Z error will generally affect the measurement probability distribution. This explanation can be applied to a Z error on any qubit in the data register, making our protocol sensitive to such errors across the entire data register. Although we specifically discussed a Pauli-Z error on the second qubit of the data register, the same principles apply to other operators that have a non-zero Hilbert-Schmidt inner product with a Pauli operator involving a Pauli Z on the second qubit in their tensor product. In essence, the phase-flip error does not have to be a Pauli error due to the linearity of the process. As a result, phase-flip errors introduced during the application of B are likely to change the distribution of the measurement outcomes, thereby allowing for the efficient detection of an elevated rate of these errors [Pro]. An experimentalist could assess whether the circuit, consisting of bias-preserving gates encoded in the unitary B, produces more phase-flip errors than expected. Detecting such anomalies is critical for superconducting cat qubits, as their scalability strategy depends on maintaining minimal phase-flip error rates, even in large-scale circuits. [GRLR+23].

In summary, our benchmark protocol can identify deviations from the noise model that might be missed by individual gate tomography, as some noise effects are not detectable at that level, a point we elaborated on at the beginning of this chapter.

Contribution: The results in this chapter are based on the work in the paper *Scalable noisy quantum circuits with biased-noise qubits*. I contributed to all the technical discussions in this work. Moreover, I did the work on finding the original and preliminary proof of "Efficient Simulation of Noise-Resilient Hadamard Test" in section 6.1 and Appendix A. I also did some preliminary numerics on "Estimating the Size of Implementable Circuits based on Literature" in section 6.4.

Chapter 7

Conclusion

In this thesis, we have delved into various aspects of quantum computing, with a particular focus on the time constraints and resource limitations associated with basis transformations, quantum coherence, and algorithm performance under noise. Through this work, we have made several contributions to understanding how these factors influence the efficiency and reliability of quantum systems.

Firstly, we explored the maximum speeds for basis transformations using unitary evolutions, determining optimal time limits across different dimensionalities. Our analysis showed that for dimensions $d \leq 4$, the shortest evolution times are similar for $d = 2$ and $d = 4$ when Hamiltonians with equal average energy E are considered. For $d = 3$, achieving the speed limit requires a specific basis ordering that remains unbiased relative to the computational basis. We further demonstrated that an n-qubit Hadamard gate can be executed in a time independent of the number of qubits, highlighting how interactive Hamiltonians can accelerate the evolution process. These insights were extended to $d \to \infty$, where we established a lower bound on the transformation time for unbiased bases, $t_{\min} \geq \frac{\pi}{4E}$. Additionally, we examined the speed limits for generating quantum coherence and transforming pure states into maximally coherent ones, suggesting the potential to extend these methods to other quantum resources, such as entanglement.

Secondly, we provided a framework for quantifying the maximum rate of quantum coherence generation, specifically in systems where Hamiltonians possess limited Hilbert-Schmidt norms. Our research pinpointed the states that maximize the rate of change in coherence for qubit systems and suggested that similar approaches could be explored in higher-dimensional systems. This lays the groundwork for future research into optimizing Hamiltonians for enhancing quantum resources like entanglement. Given the theoretical similarities between quantum coherence and entanglement, our methods offer promising avenues for further exploration of resource theories in quantum computing.

In parallel, we investigated a probabilistic version of the Bernstein-Vazirani algorithm, focusing on how coherence in the initial state affects the algorithm's ability to decode a bit string $\boldsymbol{a}$. We demonstrated that multipartite entanglement, while beneficial up to a point, can diminish performance if present in excessive quantities. Our findings are especially relevant for NMR-based quantum computation, where pseudo-pure states may provide superior performance under noisy conditions. This opens up new possibilities for applying probabilistic algorithms in real-world quantum computing scenarios.

Our work also contributes to understanding how quantum circuits can be realized efficiently with minimal quantum resources and correlations. By analyzing the performance of the DQC1 circuit, we demonstrated that it can be implemented using only negligible quantum resources, suggesting that the observed quantum speedup is unlikely to be due to the intrinsic quantum states of the processor. This insight is applicable across a wide range of resource and correlation measures, such as multipartite entanglement, mutual information, quantum coherence, and purity. These findings, together with our extension to algorithms on mixed states, suggest that the methodology can be generalized to other quantum algorithms, including those operating under noisy conditions, such as the mixed-state version of Shor's algorithm.

A key part of our investigation involved examining qubits with biased noise, namely qubits mainly producing bit-flip errors such as superconducting cat qubits. This led to the development of a restricted class of noisy Hadamard tests that include certain entangling and non-Clifford gates. Despite these noise sources, the circuits we examined showed remarkable noise resilience in the presence of Pauli bit-flip errors. Our circuits allow to recover the noiseless's algorithm outcome with a polynomial increase in algorithm's repetition (as a function of the problem size), raising the question whether extensions of this work could lead to noise-resilient circuits with computational significance. Indeed, while the circuits we analyzed happened to be efficiently simulable, it is worth noticing that they allow for gates that exhibits interesting computational properties. In particular, they can generate highly entangled state and allow to implement non-Cllifordd gates which are usually required for a computational advantage.

Finally, we developed a benchmarking protocol to assess the impact of phase-flip errors in quantum circuits. We demonstrated that our benchmark can handle circuits with up to 10^6 gates when applying the Pauli-Twirling approximation, exceeding the scale of current experimental circuits by several orders of magnitude. This scalability makes our protocol useful in the NISQ regime and beyond. Furthermore, our protocol is designed to detect some errors that can go unnoticed in individual gate tomography such as correlated or crosstalk errors, providing a comprehensive approach to error detection in large-scale quantum systems. Our protocol also allows experimentalists to assess whether circuits with bias-preserving gates introduce more phase-flip errors than expected, providing valuable insights into the noise behavior in large quantum circuits (detecting such anomalies is critical for superconducting cat qubits, as their scalability strategy depends on maintaining minimal phase-flip error rates, even in large-scale circuits).

In conclusion, our work advances the understanding of the time and resource limitations in quantum computing and proposes new methods for optimizing both algorithmic performance and system scalability. These contributions not only enhance the theoretical foundation of quantum resource management but also offer practical insights for future quantum technologies. We anticipate that the methods and findings presented in this thesis will pave the way for further innovations in the optimization and implementation of quantum computing systems.

Appendix A

Proof of Theorem 6.1 and 6.3

We extend the idea from Theorem 5.6 to address the case of a noisy algorithm implementation. Specifically, we prove that for any single-qubit Pauli operator P_1 acting on the measured register, our classical algorithm can produce samples of measurement outcomes that match the probability distribution of those on a quantum computer, and does so with polynomial computational cost [FAND+23]. This proof assumes the circuit design described in Theorem 5.5. Therefore, the Hadamard test is implemented using the circuit shown in Figure 5.4, which allows for the most general form of the Hadamard test.

Proof. [FAND+23] To prove the result, we consider two specific conditions that simplify our analysis. First, assume that noise affects only the gates acting on the measured register (H.1). Second, assume that after initializing each qubit in the parallelization and data registers, a single-qubit "X-dephasing" map, defined as $\Delta_X(\rho) = \frac{1}{2}(\rho + X\rho X)$, is applied (H.2). Under these assumptions, we show that the measured register will still arrive at the final density matrix ρ_X. These conditions provide a basis for simplifying our derivations.

The reasoning behind (H.1) is based on several key points: (i) we assume a perfect bias as outlined in Definition 6.1 from the main text, (ii) the gates used in the algorithm are all part of $\mathbb{B}_n$,[1] and (iii) the gates interacting between the measured register and the data or parallelization registers commute with any X-Pauli operator, as shown in Figure 5.4.

To be more specific, when a noisy gate is applied locally to the parallelization or data register, it affects the qubit states through Kraus operators, which are linear combinations of X-Pauli operators, as described in (i). Since these gates act on localized

[1] Although Theorem 5.5 specifies that the unitary B is implemented with bias-preserving gates, it should be noted that all other gates also fall within $\mathbb{B}_n$ due to Property 5.4.

regions and we are employing a local noise model (see Definition 6.1), the Kraus operators will also be localized. Furthermore, if these Kraus operators are commuted with other gates in the circuit, they will remain linear combinations of X-Pauli operators, as indicated by (ii).

Additionally, if no gates have been placed between the measured register and the parallelization or data registers, the impact of these Kraus operators on the measured register will be minimal. Considering that these Kraus operators have been commuted in the circuit until a gate is applied between the measured register and either the parallelization or data register, and given that these Kraus operators are linear combinations of X-Pauli operators, they will commute with the applied gate, as noted in (iii). As a result, the Kraus operators will have no impact on the density matrix of the measured register (meaning their effect will be trivial on this register). Thus, any noise introduced by gates applied locally to the parallelization or data register will not alter the measured register's density matrix. This supports the validity of (H.1). The same rationale applies to (H.2), confirming its correctness.

We denote by ρ'_f the final density matrix of the algorithm under the conditions specified by (H.1) and (H.2). We have demonstrated that $\rho_X = \mathrm{Tr}_{\neq \mathrm{Measured}}(\rho'_f)$, where $\mathrm{Tr}_{\neq \mathrm{Measured}}$ refers to the partial trace over all degrees of freedom except those of the measured register.

We define the set of Kraus operators as

$$\{A^{\,i_1,i_2}_{j_1\cdots j_{N_V}}\}^{\,i_1,i_2}_{j_1\cdots j_{N_V}}$$

to represent the sequence of noisy operations applied in the algorithm, as depicted in Figure 5.4 and under the conditions specified by (H.1) and (H.2). Using this notation, we have:

$$\rho_X = Tr_{\neq \text{Measured}}(\rho'_f) \tag{A.1}$$

$$\rho'_f \equiv \sum_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}} \left(A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}}\right) \rho'_{\text{ini}} \left(A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}}\right)^\dagger \tag{A.2}$$

$$\rho'_{\text{ini}} \equiv |0\rangle\langle 0| \bigotimes_{i=1}^{q_n} \rho_i^{\text{par}} \bigotimes_{i=1}^{n} \rho_i^{\text{data}} \tag{A.3}$$

$$A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}} \equiv K_{j_{N_V}}^{c_X \mathcal{G}_{V_{N_V}}} c_X G_{V_{N_V}} \cdots K_{j_1}^{c_X \mathcal{G}_{V_1}} c_X G_{V_1} K_{i_2}^{c_X X} c_X X (\mathbb{I}_{\text{meas}} \otimes \widetilde{B}) K_{i_1}^{c_X X} c_X X (\mathbb{I}_{\text{meas}} \otimes \mathbb{I}_{\text{par}} \otimes B), \tag{A.4}$$

$$\rho_i^{\text{par}} \equiv \Delta_X(|0\rangle\langle 0|) \tag{A.5}$$

$$\rho_i^{\text{data}} \equiv \Delta_X(|\varphi_i\rangle\langle \varphi_i|) \tag{A.6}$$

In equation (A.3), $|0\rangle\langle 0|$ denotes the initial state of the measured register. Meanwhile, ρ_i^{par} and ρ_i^{data} represent the X-dephased initial states of the parallelized and data registers, respectively (as detailed in (A.5) and (A.6)). The term $\mathbb{I}_{\text{meas}} \otimes \mathbb{I}_{\text{par}} \otimes B$ in equation (A.4) indicates the ideal, noise-free implementation of the preparation unitary on the data register, with $\mathbb{I}_{\text{meas}}$ and $\mathbb{I}_{\text{par}}$ signifying identity operations applied to the measurement and parallelization registers. The expression $K_{i_1}^{c_X X} c_X X$ corresponds to the application of the first $c_X X$ gate between the measured and parallelization registers, followed by the Kraus operator associated with its noise map. The term $\mathbb{I}_{\text{meas}} \otimes \widetilde{B}$ represents the ideal implementation of both the cNOT gates applied locally on the parallelization register and the controlled operation W. The term $K_{i_2}^{c_X X} c_X X$ denotes the second noisy $c_X X$ gate applied between the measured and parallelization registers. Lastly, $K_{j_{N_V}}^{c_X \mathcal{G}_{V_{N_V}}} c_X G_{V_{N_V}} \cdots K_{j_1}^{c_X \mathcal{G}_{V_1}} c_X G_{V_1}$ describes the noisy execution of the controlled operation V as a series of N_V gates.

We will now demonstrate that a classical computer can efficiently produce samples of measurement results for the observable P_1 on ρ_X, which will have the same probability distribution as those generated by a quantum computer.

Initially, each ρ_i^{data} can be rewritten as: $\rho_i^{\text{data}} = p_i^{+|\text{data}} |+\rangle\langle +| + p_i^{-|\text{data}} |-\rangle\langle -|$, with the probabilities $p_i^{\pm|\text{data}}$ being computable in $O(n)$ time. The same approach applies to ρ_i^{par}, as the calculation can also be completed in $O(n)$ time, given that the number of qubits in the parallelization register, q_n, is at most n. We then proceed to the decomposition:

$$\bigotimes_{i=1}^{q_n} \rho_i^{\text{par}} \bigotimes_{i=1}^{n} \rho_i^{\text{data}} = \sum_{\substack{\mathbf{s}_{\text{par}} \in \{+,-\}^{q_n} \\ \mathbf{s}_{\text{data}} \in \{+,-\}^n}} p_{\mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}} |\mathbf{s}_{\text{par}}\rangle\langle \mathbf{s}_{\text{par}}| |\mathbf{s}_{\text{data}}\rangle\langle \mathbf{s}_{\text{data}}| \tag{A.7}$$

$$p_{\mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}} = p_{\mathbf{s}_{\text{par}}}^{\text{par}} \times p_{\mathbf{s}_{\text{data}}}^{\text{data}}. \tag{A.8}$$

Here, $p_{\mathbf{s}_{\text{data}}}^{\text{data}} = \prod_{i=1}^{n} p_i^{s_{\text{data},i}|\text{data}}$ represents a product distribution. To be precise, $\mathbf{s}_{\text{data}}$ is an

n-dimensional vector that specifies the state (+ or −) of each data qubit in the mixture given by (A.7) [2]. Since the parallelization register is made up of q_n qubits, all of which start in the state $|0\rangle$, the probability $p^{\text{par}}_{\mathbf{s}_{\text{par}}}$ is $1/2^{q_n}$. Consequently, we obtain:

$$Tr(P_1\rho_X) = Tr(P_1\rho'_f) = \sum_{\substack{\mathbf{s}_{\text{par}}\in\{+,-\}^{q_n}\\ \mathbf{s}_{\text{data}}\in\{+,-\}^{n}}} p_{\mathbf{s}_{\text{par}},\mathbf{s}_{\text{data}}} \times \lambda_{\mathbf{s}_{\text{par}},\mathbf{s}_{\text{data}}} \tag{A.9}$$

$$\lambda_{\mathbf{s}_{\text{par}},\mathbf{s}_{\text{data}}} = \sum_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}} Tr\left(P_1 A_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}} |0\rangle\langle 0|\mathbf{s}_{\text{par}}\rangle\langle\mathbf{s}_{\text{par}}|\mathbf{s}_{\text{data}}\rangle\langle\mathbf{s}_{\text{data}}|.(A_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}})^\dagger\right) \tag{A.10}$$

The classical algorithm operates as follows: (A) it generates samples $\mathbf{s} \equiv \{\mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\}$ according to the probability distribution $\{p_{\mathbf{s}_{\text{par}},\mathbf{s}_{\text{data}}}\}$ (i.e., samples $\{\mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\}$ are drawn based on this distribution), and (B) it computes $\lambda_{\mathbf{s}_{\text{par}},\mathbf{s}_{\text{data}}}$. By performing both steps (A) and (B), the algorithm produces measurement outcome samples for the observable P_1 that match the probability distribution of the quantum computer's measurements.

Step (A) is efficiently handled in $O(n)$ time classically because $\mathbf{s}$ is a product measure over $n + q_n = O(n)$ bits. We will also show that step (B) is computationally feasible. To do this, we first estimate the complexity of calculating one specific term in the summation (with fixed values of $(i_1, i_2, j_1, \cdots, j_{N_V})$) and then multiply by the total number of terms in the sum. The computation is performed by evaluating $A_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}} |0, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle$. To compute this, we start by expressing $|0, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle$ as the superposition $(|+, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle + |-, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle)/\sqrt{2}$. We then analyze the action of $A_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}}$ separately on each of these two states. Given that $A_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}}$ consists of elements that are either diagonal in the X-Pauli eigenbasis or belong to $\mathbb{B}_n$, and act on a finite number of qubits, the computation involves performing local operations in the X-Pauli basis and applying a global complex factor. We describe the steps for calculating $A_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}} |+, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle$ in detail, noting that the calculation for $A_{\substack{i_1,i_2\\ j_1\cdots j_{N_V}}} |-, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle$ follows a similar approach. This method requires a specific set of operations starting from $|+, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle$.

1. Applying $(\mathbb{I}_{\text{meas}} \otimes \mathbb{I}_{\text{par}} \otimes B)$ requires $O(R_B)$ operations. This is due to B being composed of R_B local permutations, which are performed within the eigenbasis of X-Pauli operators (see Property 5.2). These permutations are carried out on vectors that are already in this eigenbasis.
2. Applying $K^{c_X X}_{i_1} c_X X$ involves $O(1)$ operations. This is because the gate $K^{c_X X}_{i_1} c_X X$ (a) affects only a small, fixed number of qubits, (b) is represented as a diagonal matrix in the X-Pauli eigenbasis, and (c) operates on a state that is already expressed in this basis.
3. Applying $(\mathbb{I}_{\text{meas}} \otimes \widetilde{B})$ requires $O(R_{\widetilde{B}})$ operations, where $R_{\widetilde{B}}$ denotes the number of gates in $\widetilde{B}$. This is due to a similar rationale as explained in item 1.

[2] The i-th component of $\mathbf{s}_{\text{data}}$, denoted $s_{\text{data},i}$, is either + or −

4. Applying $K_{j_{N_V}}^{c_X \mathcal{G}_{V_{N_V}}} c_X G_{V_{N_V}} \cdots K_{j_1}^{c_X \mathcal{G}_{V_1}} c_X G_{V_1} K_{i_2}^{c_X X} c_X X$ requires $O(N_V)$ operations. This is because the product involves $O(N_V)$ individual operations, each of which takes $O(1)$ time to execute, for reasons similar to those in item 2.

Thus, after Step 4 the state has the following form:

$$\begin{aligned} & A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}} |+, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle \\ & = e^{j\varphi_+} |+, \sigma_{\widetilde{B}}(\mathbf{s}_{\text{par}}, \sigma_B(\mathbf{s}_{\text{data}}))\rangle \,, \end{aligned} \tag{A.11}$$

for some phase φ_+ introduced by each of the steps. The permutations $\sigma_{\widetilde{B}}$ and σ_B reflect the changes made to the qubits in the parallelization and data registers as outlined in Steps 1 and 3. To determine $A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}} |-, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle$, a parallel calculation is carried out, leading to a similar result but involving a different phase φ_-:

$$\begin{aligned} & A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}} |-, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle \\ & = e^{j\varphi_-} |-, \sigma_{\widetilde{B}}(\mathbf{s}_{\text{par}}, \sigma_B(\mathbf{s}_{\text{data}}))\rangle \,. \end{aligned} \tag{A.12}$$

Given that P_1 interacts solely with the first qubit and the final state, which results from summing (A.12) and (A.11), is a product state, the evaluation of

$$Tr(P_1 A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}} |0, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}\rangle\langle 0, \mathbf{s}_{\text{par}}, \mathbf{s}_{\text{data}}| (A_{\substack{i_1,i_2 \\ j_1 \cdots j_{N_V}}})^\dagger) \tag{A.13}$$

takes $O(1)$ number of operations, as it involves evaluating a single-qubit Pauli operator on a product state. Consequently, computing each term in the sum (A.10) takes $O(N_V + R_B + R_{\widetilde{B}})$ steps. Next, we need to determine the total number of terms in the sum (A.10).

Each index in the sum corresponds to the count of Kraus operators for a particular noise map. Let $j_{\max}$ denote the maximum number of qubits affected by any gate in the algorithm. As a result, the number of elements in the sum is bounded by $O(j_{\max}^{N_V})$. Therefore, computing (A.10) will require at most $O(j_{\max}^{N_V}(N_V + R_B + R_{\widetilde{B}}))$ operations. Consequently, the total computational cost for performing steps (A) and (B) is $O(n + j_{\max}^{N_V}(N_V + R_B + R_{\widetilde{B}}))$.

Our objective is to compute the expectation value $Tr(P_1 \rho_X)$ using samples generated by the classical algorithm. Hoeffding's inequality tells us that if P_1's expectation values are within $[-1, 1]$, then by repeating the process $O((1/\varepsilon^2)\log(1/\delta))$ times and averaging the results, we can achieve an ε-accurate estimate of $Tr(P_1 \rho_X)$ with probability at least $1 - \delta$. Therefore, the overall complexity of this method is at most $O((1/\varepsilon^2)\log(1/\delta) \times (n + j_{\max}^{N_V}(N_V + R_B + R_{\widetilde{B}})))$. Given that $N_V = O(\log(n))$ and R_B, $R_{\widetilde{B}}$ are polynomial functions of n, the total complexity is $O((1/\varepsilon^2)\log(1/\delta) \times \text{Poly}(n))$, which indicates that the algorithm is efficient. □

Appendix B

Example of Bias-Preserving Gates

In the following section, we will present specific examples of bias-preserving gates and demonstrate how these gates propagate errors in quantum circuits [FAND+23]. Indeed, while bias-preserving gates do have a theoretical existence, their practical implementation in a bias-preserving manner can be quite challenging. The challenge in implementing bias-preserving gates lies in the continuous Hamiltonian evolution used in the laboratory to realize these gates. When implementing gates through continuous Hamiltonian evolution, the presence of Pauli Z-terms in the Hamiltonian can lead to a phenomenon known as error transversal. In this context, a bit-flip (X-error) that occurs during the gate evolution might be converted into a phase flip (Z-error). In other words, the original type of error can change as the gate is being performed, which can introduce unexpected errors and impact the accuracy of the quantum computation. For example, while a cNOT gate preserves X-errors in principle, ensuring that this condition still holds during continuous-time evolution to implement the gate cannot be guaranteed. In practice, this issue can be overcome by utilizing cat-qubits [PSJG+20] and both the Toffoli' $\equiv H_1H_2H_3 \times \text{Toffoli} \times (H_1H_2H_3)^\dagger$ and cNOT gates have been demonstrated to preserve X-errors in the literature [GM21, PSJG+20]. In our work, it is important to note that we adopted a different convention for the dominant source of errors. (Indeed, in the references [GM21, PSJG+20], the main focus is on errors caused by phase-flips, i.e., Z errors, while in our work, we primarily consider errors resulting from bit-flips, i.e., X errors.). Hence, by "swapping" the role of X and Z errors, we can adapt the results from the references [GM21, PSJG+20] to our convention, which allows us to conclude that Toffoli and cNOT gates preserve the bit-flip bias, ensuring that errors introduced during their implementation remain in the X-error category. Figure B.1 illustrates the propagation of errors through cNOT and Toffoli gates, highlighting their bias-preserving property. In Figure B.1, we can observe that the occurrence of a pre-existing Pauli X error on a specific qubit can lead to the propagation of this error

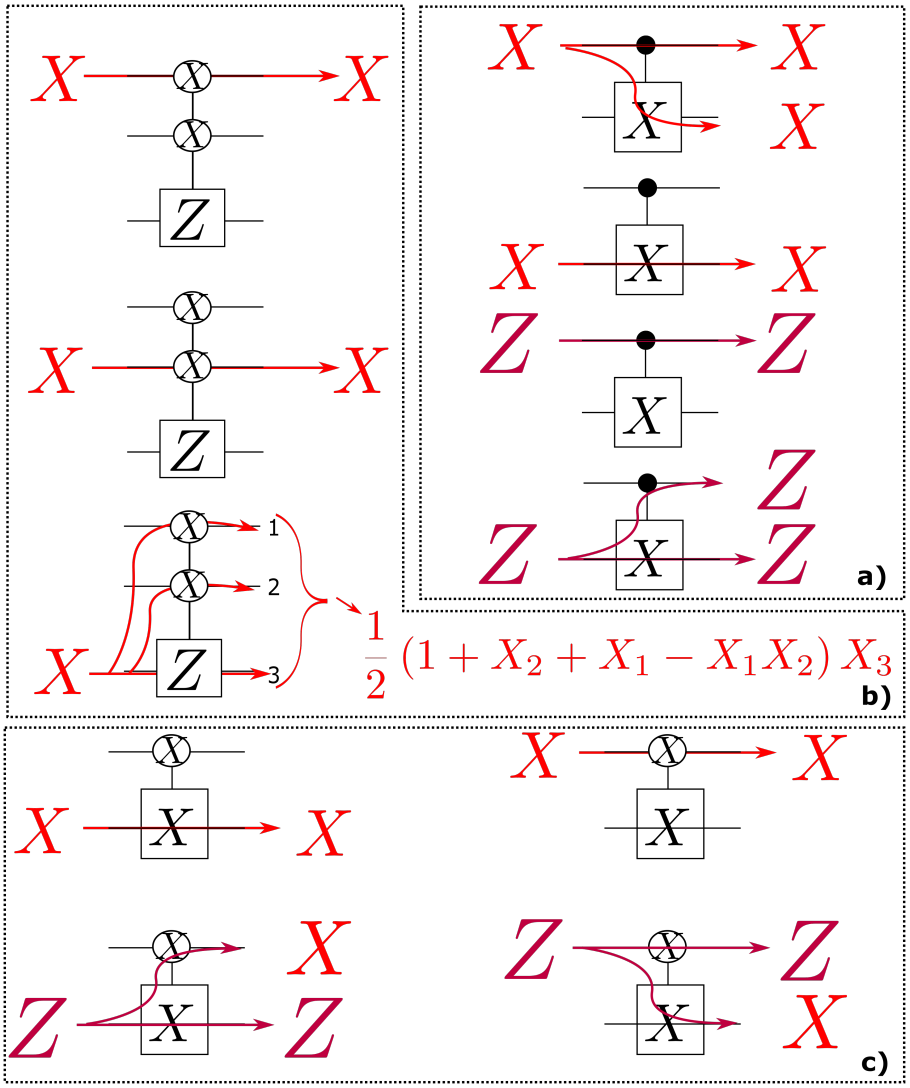

Figure B.1: [FAND+23] This image illustrates how initial Pauli errors (either X or Z) propagate through various bias-preserving gates. Specifically, we provide examples of bias-preserving gates that maintain X-errors, including the cNOT gate, the Toffoli' gate (which is a variant of the Toffoli gate where the roles of X and Z are swapped), and the c_XX gate (where c_XG denotes the application of a gate G controlled coherently in the X basis, meaning G is applied when the control qubit is in state $|-\rangle$ and not applied when it is in state $|+\rangle$). Coherent control in the X basis is represented by an X symbol within a black circle, as illustrated on the top qubit in Figure (c). In figure (a), it is shown that in a cNOT gate, an X error originating from the control qubit will propagate to the target qubit, but not in the reverse direction. Conversely, a Z error will move from the target qubit to the control qubit, but not from the control to the target. Figure (b) shows that the Toffoli gate allows X-errors to spread to multiple qubits, but this only occurs if there was already an error present on the target qubit beforehand. In this scenario, the resulting error transforms from being part of $\mathbb{P}_n^X$ to belonging to $\mathbb{U}_n^X$. The Toffoli gate does not maintain Z (or Y) errors. In part (c), it is shown that an X error commutes with the c_XX gate. However, Z errors are not preserved by this gate: a Z error occurring before the c_XX gate is equivalent to applying the c_XX gate followed by an operator that no longer qualifies as a Pauli Z operator, as demonstrated in the two examples at the bottom of the figure.

to multiple qubits, depending on the gate operations applied in the circuit. As depicted in Figure B.1 b), in certain cases, the error that propagates through the bias-preserving gates (cNOT and Toffoli') may no longer remain a simple Pauli X operator. However, due to the use of bias-preserving gates, we can ensure that any initial Pauli-X errors, which may be present before the gate operation, will always result in an error that belongs to the set $\mathbb{U}_n^X$ after the gate is applied. In Figure B.1, we also demonstrated the propagation of Z errors through a cNOT gate (similarly, the definitions can be extended to Y and Z operators). The way Y errors propagate can be inferred from how X and Z errors behave during the error propagation process. Toffoli' does not preserve Z errors, as a pre-existing Z error on any of the control qubits would no longer be a Z error after the gate.

Appendix C

Entanglement Properties of the States Produced by the Bias-Preserving Gates

In this section, we aim to discuss various characteristics of the entanglement produced by bias-preserving circuits and illustrate their application through specific examples using the Hadamard test [FAND+23]. One can show that highly entangled graph states can be generated by these circuits [VdNDDM04]. To accomplish this, the data register is first set to the state $|0\rangle^{\otimes n}$, and then specific gates $c_X X \in \mathbb{U}_n^X$ are applied within the preparation unitary B. This indicates that, in general, bias-preserving circuits can produce quantum states with intriguing computational properties, a point that is important to consider for potential future extension of the current research.

It is important to note that for the particular computational task examined in our study—executing a limited category of Hadamard tests, the $c_X X$ gates will not influence the measurement results [FAND+23]. Therefore, this implies that while bias-preserving circuits are capable of preparing highly entangled graph states, in our specific task, they would yield the same results as a Hadamard test, represented by $\langle 0^{\otimes n}| U |0^{\otimes n}\rangle$, for any U permitted by Theorem 5.5. It is noteworthy that U itself can serve as an entangling operation here.

However, this does not pose a problem for our objectives [FAND+23]: (i) it does not diminish the value of the benchmarking protocol for these states, and (ii) we can create other entangled states where the entangling operations used in their preparation influence the outcome of the Hadamard test. Regarding (i), the goal of the verification is to determine if the circuit designed to prepare these highly entangled states functions correctly. If there are no errors, our example shows that the expectation value for this Hadamard test will be identical to that for $|\psi\rangle = |0\rangle^{\otimes n}$[FAND+23]. When

there is an *imperfect* bias, the circuit, even if it only uses $c_X X$ gates in the preparation unitary B, will generally propagate Y and Z errors to the measured register. This results in substantial changes to its density matrix compared to the case with no noise [FAND+23]. Thus, the verification protocol can still assess whether the highly entangled states were prepared in a bias-preserving way, meaning they are only influenced by X-errors [FAND+23]. Verification is feasible for the unitaries permitted by Theorem 5.5, provided that $\langle 0^{\otimes n}| U |0^{\otimes n}\rangle \neq 0$, which can indeed be satisfied in our case [1]. Regarding (ii), it's straightforward to construct entangled states where the gates involved in their creation affect the expectation value of the Hadamard test. A basic example of this is an n-qubit GHZ state. You can create it by setting the first qubit to $|+\rangle$, initializing the remaining qubits to $|0\rangle^{\otimes n-1}$, and then applying a series of cNOT gates with the first qubit as the control and the others as targets [FAND+23]. Although the cNOT gates are part of $\mathbb{B}_n^X$, they are not included in $\mathbb{U}_n^X$. Consequently, the preparation unitary B will affect the measured outcome. In summary, providing a precise characterization of all entangled states achievable with our circuits is beyond the scope of the work.

Deploying the error mitigation techniques in these circuits along with studying their scalability could be a direction for future inspection, to evaluate the possibility and the degree of retrieving the entangled states that were supposed to be generated by the noiseless circuits.

[1] For example, if U is a single-qubit rotation around X with an angle different from $(\pi, 0)$, a perfect bias would result in a finite (non-zero) value for $\langle\psi| U |\psi\rangle$. However, with an imperfect bias, we generally expect $Tr(\rho Y)$ and $Tr(\rho Z)$ to be exponentially small allowing an experimentalist to detect this imperfection, as discussed in the benchmarking protocol in the following chapter.

Bibliography

[Abe06] Johan Aberg. Quantifying Superposition. *arXiv:quant-ph/0612146*, 2006.

[AG04] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70(5):052328, November 2004. doi:10.1103/PhysRevA.70.052328.

[AJL09] Dorit Aharonov, Vaughan Jones, and Zeph Landau. A Polynomial Quantum Algorithm for Approximating the Jones Polynomial. *Algorithmica*, 55(3):395–421, November 2009. doi:10.1007/s00453-008-9168-0.

[ALL+23] Abhishek Agarwal, Lachlan P Lindoy, Deep Lall, Francois Jamet, and Ivan Rungger. Modelling non-markovian noise in driven superconducting qubits. *arXiv:2306.13021*, 2023.

[ATE+22a] Felix Ahnefeld, Thomas Theurer, Dario Egloff, Juan Mauricio Matera, and Martin B. Plenio. Coherence as a Resource for Shor's Algorithm. *Phys. Rev. Lett.*, 129:120501, September 2022. doi:10.1103/PhysRevLett.129.120501.

[ATE+22b] Felix Ahnefeld, Thomas Theurer, Dario Egloff, Juan Mauricio Matera, and Martin B. Plenio. On the Role of Coherence in Shor's Algorithm. *arXiv:2203.10632*, 2022.

[BCP14a] T. Baumgratz, M. Cramer, and M. B. Plenio. Quantifying Coherence. *Phys. Rev. Lett.*, 113:140401, Sep 2014. doi:10.1103/PhysRevLett.113.140401.

[BCP14b] T. Baumgratz, M. Cramer, and M. B. Plenio. Quantifying coherence. *Phys. Rev. Lett.*, 113:140401, Sep 2014. doi:10.1103/PhysRevLett.113.140401.

[BDF+99] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, Feb 1999. doi:10.1103/PhysRevA.59.1070.

[BDLR21] Simon Becker, Nilanjana Datta, Ludovico Lami, and Cambyse Rouzé. Energy-constrained discrimination of unitaries, quantum speed limits, and a Gaussian Solovay-Kitaev theorem. *Physical Review Letters*, 126(19):190504, 2021. doi:10.1103/PhysRevLett.126.190504.

[BDSW96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, Nov 1996. doi:10.1103/PhysRevA.54.3824.

[BIS+18] Sergio Boixo, Sergei V Isakov, Vadim N Smelyanskiy, Ryan Babbush, Nan Ding, Zhang Jiang, Michael J Bremner, John M Martinis, and Hartmut Neven. Characterizing quantum supremacy in near-term devices. *Nature Physics*, 14(6):595–600, 2018. doi:10.1038/s41567-018-0124-x.

[BK15] Joonwoo Bae and Leong-Chuan Kwek. Quantum state discrimination and its applications. *Journal of Physics A: Mathematical and Theoretical*, 48(8):083001, 2015. doi:10.1088/1751-8113/48/8/083001.

[BKZW17] Kaifeng Bu, Asutosh Kumar, Lin Zhang, and Junde Wu. Cohering power of quantum operations. *Physics Letters A*, 381(19):1670–1676, 2017. doi:10.1016/j.physleta.2017.03.022.

[BL01] H Barnum and N Linden. Monotones and invariants for multi-particle quantum states. *J. Phys. A*, 34(35):6787, aug 2001. doi:10.1088/0305-4470/34/35/305.

[BM18] Mario Berta and Christian Majenz. Disentanglement cost of quantum states. *Phys. Rev. Lett.*, 121:190503, Nov 2018. doi:10.1103/PhysRevLett.121.190503.

[BMKT22] Ewout van den Berg, Zlatko K. Minev, Abhinav Kandala, and Kristan Temme. Probabilistic error cancellation with sparse Pauli-Lindblad models on noisy quantum processors. *arxiv:2201.09866*, January 2022.

[BMSSO18] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien. Low-cost error mitigation by symmetry verification. *Phys. Rev. A*, 98(6):062339, December 2018. doi:10.1103/PhysRevA.98.062339.

[BNO02] Ofer Biham, Michael A. Nielsen, and Tobias J. Osborne. Entanglement monotone derived from Grover's algorithm. *Phys. Rev. A*, 65:062312, Jun 2002. doi:10.1103/PhysRevA.65.062312.

[Bra07] Sergey Bravyi. Upper bounds on entangling rates of bipartite hamiltonians. *Phys. Rev. A*, 76:052319, Nov 2007. doi:10.1103/PhysRevA.76.052319.

[BV97] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM Journal on computing*, 26(5):1411–1473, 1997. doi:10.1137/S0097539796300921.

[CBB+22] Zhenyu Cai, Ryan Babbush, Simon C. Benjamin, Suguru Endo, William J. Huggins, Ying Li, Jarrod R. McClean, and Thomas E. O'Brien. Quantum Error Mitigation. *arXiv:2210.00921*, October 2022.

[CCC19] Patrick J Coles, M Cerezo, and Lukasz Cincio. Strong bound between trace distance and Hilbert-Schmidt distance for low-rank states. *Physical Review A*, 100(2):022103, 2019. doi:10.1103/PhysRevA.100.022103.

[CCHL22] Sitan Chen, Jordan Cotler, Hsin-Yuan Huang, and Jerry Li. The Complexity of NISQ. *arXiv:2210.07234*, October 2022.

[CFH97] David G. Cory, Amr F. Fahmy, and Timothy F. Havel. Ensemble quantum computing by nmr spectroscopy. *Proceedings of the National Academy of Sciences*, 94(5):1634–1639, 1997. doi:10.1073/pnas.94.5.1634.

[CG16] Eric Chitambar and Gilad Gour. Critical Examination of Incoherent Operations and a Physically Consistent Resource Theory of Quantum Coherence. *Phys. Rev. Lett.*, 117:030401, Jul 2016. doi:10.1103/PhysRevLett.117.030401.

[CG19] Eric Chitambar and Gilad Gour. Quantum resource theories. *Rev. Mod. Phys.*, 91:025001, Apr 2019. doi:10.1103/RevModPhys.91.025001.

[CLM+14] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. Everything you always wanted to know about locc (but were afraid to ask). *Communications in Mathematical Physics*, 328:303–326, 2014. doi:10.1007/s00220-014-1953-9.

[CNAA+22] Christopher Chamberland, Kyungjoo Noh, Patricio Arrangoiz-Arriola, Earl T. Campbell, Connor T. Hann, Joseph Iverson, Harald Putterman, Thomas C. Bohdanowicz, Steven T. Flammia, Andrew Keller, Gil Refael, John Preskill, Liang Jiang, Amir H. Safavi-Naeini, Oskar Painter, and Fernando G. S. L. Brandão. Building a Fault-Tolerant Quantum Computer Using Concatenated Cat Codes. *PRX Quantum*, 3(1):010329, February 2022. doi:10.1103/PRXQuantum.3.010329.

[CPBM18] Francesco Campaioli, Felix A. Pollock, Felix C. Binder, and Kavan Modi. Tightening quantum speed limits for almost all states. *Phys. Rev. Lett.*, 120:060409, February 2018. doi:10.1103/PhysRevLett.120.060409.

[CPF+10] Marcus Cramer, Martin B Plenio, Steven T Flammia, Rolando Somma, David Gross, Stephen D Bartlett, Olivier Landon-Cardinal, David Poulin, and Yi-Kai Liu. Efficient quantum state tomography. *Nature communications*, 1(1):149, 2010. doi:10.1038/ncomms1147.

[DB07] Wolfgang Dür and Hans J Briegel. Entanglement purification and quantum error correction. *Reports on Progress in Physics*, 70(8):1381, 2007. doi:10.1088/0034-4885/70/8/R03.

[DC17] Sebastian Deffner and Steve Campbell. Quantum speed limits: from Heisenberg's uncertainty principle to optimal quantum control. *J. Phys. A*, 50(45):453001, October 2017. doi:10.1088/1751-8121/aa86c6.

[dC21] Adolfo del Campo. Probing quantum speed limits with ultracold gases. *Phys. Rev. Lett.*, 126:180603, May 2021. doi:10.1103/PhysRevLett.126.180603.

[dCEPH13] A. del Campo, I. L. Egusquiza, M. B. Plenio, and S. F. Huelga. Quantum speed limits in open system dynamics. *Phys. Rev. Lett.*, 110:050403, January 2013. doi:10.1103/PhysRevLett.110.050403.

[DFC05] Animesh Datta, Steven T. Flammia, and Carlton M. Caves. Entanglement and the power of one qubit. *Phys. Rev. A*, 72:042316, Oct 2005. doi:10.1103/PhysRevA.72.042316.

[DFW+18] María García Díaz, Kun Fang, Xin Wang, Matteo Rosati, Michalis Skotiniotis, John Calsamiglia, and Andreas Winter. Using and reusing coherence to realize quantum processes. *Quantum*, 2:100, October 2018. doi:10.22331/q-2018-10-19-100.

[DKSW18] Siddhartha Das, Sumeet Khatri, George Siopsis, and Mark M. Wilde. Fundamental limits on quantum dynamics based on entropy change. *J. Math. Phys.*, 59(1):012205, 01 2018. doi:10.1063/1.4997044.

[DPM+23] Nicola D'Alessandro, Beatrice Polacchi, George Moreno, Emanuele Polino, Rafael Chaves, Iris Agresti, and Fabio Sciarrino. Machine-learning-based device-independent certification of quantum networks. *Physical Review Research*, 5(2):023016, 2023. doi:10.1103/PhysRevResearch.5.023016.

[DSC08] Animesh Datta, Anil Shaji, and Carlton M. Caves. Quantum discord and the power of one qubit. *Phys. Rev. Lett.*, 100:050502, Feb 2008. doi:10.1103/PhysRevLett.100.050502.

[DVB10] Borivoje Dakić, Vlatko Vedral, and Časlav Brukner. Necessary and sufficient condition for nonzero quantum discord. *Phys. Rev. Lett.*, 105:190502, Nov 2010. doi:10.1103/PhysRevLett.105.190502.

[DVC00] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62:062314, Nov 2000. doi:10.1103/PhysRevA.62.062314.

[DVC+01] W. Dür, G. Vidal, J. I. Cirac, N. Linden, and S. Popescu. Entanglement Capabilities of Nonlocal Hamiltonians. *Phys. Rev. Lett.*, 87:137901, Sep 2001. doi:10.1103/PhysRevLett.87.137901.

[dVS16] Julio I de Vicente and Alexander Streltsov. Genuine quantum coherence. *J. Phys. A*, 50(4):045301, dec 2016. doi:10.1088/1751-8121/50/4/045301.

[EHW+20] Jens Eisert, Dominik Hangleiter, Nathan Walk, Ingo Roth, Damian Markham, Rhea Parekh, Ulysse Chabaud, and Elham Kashefi. Quantum certification and benchmarking. *Nature Reviews Physics*, 2(7):382–390, 2020. doi:10.1038/s42254-020-0186-4.

[EMG+22] Andrew Eddins, Mario Motta, Tanvi P. Gujarati, Sergey Bravyi, Antonio Mezzacapo, Charles Hadfield, and Sarah Sheldon. Doubling the size of quantum simulators by entanglement forging. *PRX Quantum*, 3:010309, Jan 2022. doi:10.1103/PRXQuantum.3.010309.

[FACW+21] Marco Fellous-Asiani, Jing Hao Chai, Robert S. Whitney, Alexia Auffèves, and Hui Khoon Ng. Limitations in Quantum Computing from Resource Constraints. *PRX Quantum*, 2(4):040335, November 2021. doi:10.1103/PRXQuantum.2.040335.

[FAND+23] Marco Fellous-Asiani, Moein Naseri, Chandan Datta, Alexander Streltsov, and Michał Oszmaniec. Scalable noisy quantum circuits for biased-noise qubits. *arXiv:2305.02045*, 2023.

[FGLE12] Steven T Flammia, David Gross, Yi-Kai Liu, and Jens Eisert. Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators. *New Journal of Physics*, 14(9):095022, 2012.

[FSS19] Ken Funo, Naoto Shiraishi, and Keiji Saito. Speed limit for open quantum systems. *New Journal of Physics*, 21(1):013006, January 2019. doi:10.1088/1367-2630/aaf9f5.

[GDEP16] M. García-Díaz, D. Egloff, and M. B. Plenio. A note on coherence power of n-dimensional unitary operators. *Quant. Inf. Comp.*, 16:1282 – 1294, 2016. arXiv:1510.06683.

[GFE09] D. Gross, S. T. Flammia, and J. Eisert. Most quantum states are too entangled to be useful as computational resources. *Phys. Rev. Lett.*, 102:190501, May 2009. doi:10.1103/PhysRevLett.102.190501.

[GLN05] Alexei Gilchrist, Nathan K Langford, and Michael A Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, 2005. doi:10.1103/PhysRevA.71.062310.

[GM19] Jérémie Guillaud and Mazyar Mirrahimi. Repetition Cat Qubits for Fault-Tolerant Quantum Computation. *Phys. Rev. X*, 9(4):041053, December 2019. doi:10.1103/PhysRevX.9.041053.

[GM21] Jérémie Guillaud and Mazyar Mirrahimi. Error rates and resource overheads of repetition cat qubits. *Phys. Rev. A*, 103(4):042413, April 2021. doi:10.1103/PhysRevA.103.042413.

[GMN+15] Gilad Gour, Markus P. Müller, Varun Narasimhachar, Robert W. Spekkens, and Nicole Yunger Halpern. The resource theory of informational nonequilibrium in thermodynamics. *Physics Reports*, 583:1–58, 2015. doi:10.1016/j.physrep.2015.04.003.

[GRLR+23] Élie Gouzien, Diego Ruiz, Francois-Marie Le Régent, Jérémie Guillaud, and Nicolas Sangouard. Performance analysis of a repetition cat code architecture: Computing 256-bit elliptic curve logarithm in 9 hours with 126 133 cat qubits. *Physical Review Letters*, 131(4):040602, 2023. doi:10.1103/PhysRevLett.131.040602.

[GS08] Gilad Gour and Robert W Spekkens. The resource theory of quantum reference frames: manipulations and monotones. *New Journal of Physics*, 10(3):033023, mar 2008. doi:10.1088/1367-2630/10/3/033023.

[HBAB19] Saronath Halder, Manik Banik, Sristy Agrawal, and Somshubhro Bandyopadhyay. Strong quantum nonlocality without entanglement. *Phys. Rev. Lett.*, 122:040403, Feb 2019. doi:10.1103/PhysRevLett.122.040403.

[HG18] Alexander Hickey and Gilad Gour. Quantifying the imaginarity of quantum mechanics. *J. Phys. A*, 51(41):414009, September 2018. doi:10.1088/1751-8121/aabe9c.

[HHHH09a] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. doi:10.1103/RevModPhys.81.865.

[HHHH09b] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. doi:10.1103/RevModPhys.81.865.

[HHO03] Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. Reversible transformations from pure to mixed states and the unique measure of information. *Phys. Rev. A*, 67:062104, Jun 2003. doi:10.1103/PhysRevA.67.062104.

[HO13] Michał Horodecki and Jonathan Oppenheim. Fundamental limitations for quantum and nanoscale thermodynamics. *Nature Communications*, 4(1):2059, Jun 2013. doi:10.1038/ncomms3059.

[idZHSL98] Karol Życzkowski, Paweł Horodecki, Anna Sanpera, and Maciej Lewenstein. Volume of the set of separable states. *Phys. Rev. A*, 58:883–892, Aug 1998. doi:10.1103/PhysRevA.58.883.

[JK10] Philip J. Jones and Pieter Kok. Geometric derivation of the quantum speed limit. *Phys. Rev. A*, 82:022107, August 2010. doi:10.1103/PhysRevA.82.022107.

[KL98] E. Knill and R. Laflamme. Power of one bit of quantum information. *Phys. Rev. Lett.*, 81:5672–5675, Dec 1998. doi:10.1103/PhysRevLett.81.5672.

[KLR+08] Emanuel Knill, Dietrich Leibfried, Rolf Reichle, Joe Britton, R Brad Blakestad, John D Jost, Chris Langer, Roee Ozeri, Signe Seidelin, and David J Wineland. Randomized benchmarking of quantum gates. *Physical Review A—Atomic, Molecular, and Optical Physics*, 77(1):012307, 2008. doi:10.1103/PhysRevA.77.012307.

[KLR+20] Sebastian Krinner, Stefania Lazar, Ants Remm, Christian K Andersen, Nathan Lacroix, Graham J Norris, Christoph Hellings, Mihai Gabureac, Christopher Eichler, and Andreas Wallraff. Benchmarking coherent errors in controlled-phase gates due to spectator qubits. *Physical Review Applied*, 14(2):024042, 2020. doi:10.1103/PhysRevApplied.14.024042.

[KMS16] Rajath Krishna, Vishesh Makwana, and Ananda Padhmanabhan Suresh. A generalization of Bernstein-Vazirani algorithm to qudit systems. *arXiv:1609.03185*, 2016.

[Koc21] Bálint Koczor. Exponential error suppression for near-term quantum devices. *Phys. Rev. X*, 11:031057, Sep 2021. doi:10.1103/PhysRevX.11.031057.

[LP01] Noah Linden and Sandu Popescu. Good dynamics versus bad kinematics: Is entanglement needed for quantum computation? *Phys. Rev. Lett.*, 87:047901, Jul 2001. doi:10.1103/PhysRevLett.87.047901.

[LT09] Lev B Levitin and Tommaso Toffoli. Fundamental limit on the rate of quantum dynamics: the unified bound is tight. *Physical review letters*, 103(16):160502, 2009. doi:10.1103/PhysRevLett.103.160502.

[LVP+20] Raphaël Lescanne, Marius Villiers, Théau Peronnin, Alain Sarlette, Matthieu Delbecq, Benjamin Huard, Takis Kontos, Mazyar Mirrahimi, and Zaki Leghtas. Exponential suppression of bit-flips in a qubit encoded in an oscillator. *Nat. Phys.*, 16(5):509–513, May 2020. doi:10.1038/s41567-020-0824-x.

[MBC+12] Kavan Modi, Aharon Brodutch, Hugo Cable, Tomasz Paterek, and Vlatko Vedral. The classical-quantum boundary for correlations: Discord and related measures. *Rev. Mod. Phys.*, 84:1655–1707, Nov 2012. doi:10.1103/RevModPhys.84.1655.

[MDP22] Brij Mohan, Siddhartha Das, and Arun Kumar Pati. Quantum speed limits for information and coherence. *New Journal of Physics*, 24(6):065003, 2022. doi:10.1088/1367-2630/ac753c.

[MEKP16a] J M Matera, D Egloff, N Killoran, and M B Plenio. Coherent control of quantum systems as a resource theory. *Quantum Sci. Technol.*, 1(1):01LT01, 2016. doi:10.1088/2058-9565/1/1/01LT01.

[MEKP16b] J M Matera, D Egloff, N Killoran, and M B Plenio. Coherent control of quantum systems as a resource theory. *Quantum Science and Technology*, 1(1):01LT01, August 2016. doi:10.1088/2058-9565/1/1/01LT01.

[MGE12] Easwar Magesan, Jay M Gambetta, and Joseph Emerson. Characterizing quantum gates via randomized benchmarking. *Physical Review A*, 85(4):042311, 2012. doi:10.1103/PhysRevA.85.042311.

[MK15] Azam Mani and Vahid Karimipour. Cohering and decohering power of quantum channels. *Phys. Rev. A*, 92:032331, Sep 2015. doi:10.1103/PhysRevA.92.032331.

[ML98] Norman Margolus and Lev B. Levitin. The maximum speed of dynamical evolution. *Physica D*, 120(1):188 – 195, 1998. doi:10.1016/S0167-2789(98)00054-2.

[MMMJA20] Prakash Murali, David C McKay, Margaret Martonosi, and Ali Javadi-Abhari. Software mitigation of crosstalk on noisy intermediate-scale quantum computers. In *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, pages 1001–1016, 2020.

[MN01] Cristopher Moore and Martin Nilsson. Parallel quantum computation and quantum codes. *SIAM journal on computing*, 31(3):799–815, 2001. doi:10.1137/S0097539799355053.

[MP21] Brij Mohan and Arun Kumar Pati. Quantum speed limits for observable. *arXiv:2112.13789*, 2021.

[MT45] L. Mandelstam and Ig. Tamm. The Uncertainty Relation Between Energy and Time in Non-relativistic Quantum Mechanics. *J. Phys. USSR*, 9:249–254, 1945. doi:10.1007/978-3-642-74626-0_8.

[MV21] Tony Metger and Thomas Vidick. Self-testing of a single quantum device under computational assumptions. *Quantum*, 5:544, 2021. doi:10.22331/q-2021-09-16-544.

[MZO20] Filip B. Maciejewski, Zoltán Zimborás, and Michał Oszmaniec. Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography. *Quantum*, 4:257, April 2020. doi:10.22331/q-2020-04-24-257.

[NAS22] Gal Ness, Andrea Alberti, and Yoav Sagi. Quantum speed limit for states with a bounded energy spectrum. *Phys. Rev. Lett.*, 129:140403, September 2022. doi:10.1103/PhysRevLett.129.140403.

[NBC+16] Carmine Napoli, Thomas R. Bromley, Marco Cianciaruso, Marco Piani, Nathaniel Johnston, and Gerardo Adesso. Robustness of coherence: An operational and observable measure of quantum coherence. *Phys. Rev. Lett.*, 116:150502, Apr 2016. doi:10.1103/PhysRevLett.116.150502.

[NC10] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010. doi:10.1017/CBO9780511976667.

[NKG+22] Moein Naseri, Tulja Varun Kondra, Suchetana Goswami, Marco Fellous-Asiani, and Alexander Streltsov. Entanglement and coherence in the Bernstein-Vazirani algorithm. *Physical Review A*, 106(6):062429, 2022. doi:10.1103/PhysRevA.106.062429.

[NMB+24] Moein Naseri, Chiara Macchiavello, Dagmar Bruß, Paweł Horodecki, and Alexander Streltsov. Quantum speed limits for change of basis. *New Journal of Physics*, 26(2):023052, 2024. doi:10.1088/1367-2630/ad25a5.

[NW18] Nelly Huei Ying Ng and Mischa Prebin Woods. *Resource Theory of Quantum Thermodynamics: Thermal Operations and Second Laws*, pages 625–650. Springer International Publishing, Cham, 2018. doi:10.1007/978-3-319-99046-0_26.

[OS06] Andreas Osterloh and Jens Siewert. Entanglement monotones and maximally entangled states in multipartite qubit systems. *International journal of quantum information*, 4(03):531–540, 2006. doi:10.1142/S0219749906001980.

[PBE22] Elijah Pelofske, Andreas Bärtschi, and Stephan Eidenbenz. Quantum volume in practice: What users can expect from NISQ devices. *IEEE Transactions on Quantum Engineering*, 3:1–19, 2022. doi:10.1109/TQE.2022.3184764.

[PCB+16] Marco Piani, Marco Cianciaruso, Thomas R. Bromley, Carmine Napoli, Nathaniel Johnston, and Gerardo Adesso. Robustness of asymmetry and coherence of quantum states. *Phys. Rev. A*, 93:042107, Apr 2016. doi:10.1103/PhysRevA.93.042107.

[PCC+16] Diego Paiva Pires, Marco Cianciaruso, Lucas C. Céleri, Gerardo Adesso, and Diogo O. Soares-Pinto. Generalized geometric quantum speed limits. *Phys. Rev. X*, 6:021031, June 2016. doi:10.1103/PhysRevX.6.021031.

[PP00] S. Parker and M. B. Plenio. Efficient factorization with a single pure qubit and log$N$ mixed qubits. *Phys. Rev. Lett.*, 85:3049–3052, Oct 2000. doi:10.1103/PhysRevLett.85.3049.

[PP02] S. Parker and M. B. Plenio. Entanglement simulations of Shor's algorithm. *Journal of Modern Optics*, 49(8):1325–1353, 2002. doi:10.1080/09500340110107207.

[Pre18] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018. arXiv:1801.00862v3, doi:10.22331/q-2018-08-06-79.

[Pro] We use the term 'in general' because the detection of Z errors (i.e., phase-flips) can vary depending on the specific implementation of B. For example, if the overlap measured by the Hadamard test, $\langle\psi| U |\psi\rangle$, equals $\langle\psi| Z_i U Z_i |\psi\rangle$, a Z error occurring on the i-th data qubit just before applying U would not affect the measurement probability distribution.

[PSJG+20] Shruti Puri, Lucas St-Jean, Jonathan A. Gross, Alexander Grimm, Nicholas E. Frattini, Pavithran S. Iyer, Anirudh Krishna, Steven Touzard, Liang Jiang, Alexandre Blais, Steven T. Flammia, and S. M. Girvin. Bias-preserving gates with stabilized cat qubits. *Sci. Adv.*, 6(34):eaay5901, August 2020. doi:10.1126/sciadv.aay5901.

[PV05] Martin B Plenio and Shashank Virmani. An introduction to entanglement measures. *arXiv:0504163*, 2005.

[QCL21] Dayue Qin, Yanzhu Chen, and Ying Li. Error statistics and scalability of quantum error mitigation formulas. *arXiv*, December 2021. arXiv:2112.06255, doi:10.48550/arXiv.2112.06255.

[QFK+22] Yihui Quek, Daniel Stilck França, Sumeet Khatri, Johannes Jakob Meyer, and Jens Eisert. Exponentially tighter bounds on limitations of quantum error mitigation. *arXiv*, October 2022. arXiv:2210.11505, doi:10.48550/arXiv.2210.11505.

[RFWA18] Bartosz Regula, Kun Fang, Xin Wang, and Gerardo Adesso. One-shot coherence distillation. *Phys. Rev. Lett.*, 121:010401, July 2018. doi:10.1103/PhysRevLett.121.010401.

[RLS18] Bartosz Regula, Ludovico Lami, and Alexander Streltsov. Nonasymptotic assisted distillation of quantum coherence. *Phys. Rev. A*, 98:052329, November 2018. doi:10.1103/PhysRevA.98.052329.

[RW15] David Reeb and Michael M. Wolf. Tight Bound on Relative Entropy by Entropy Difference. *IEEE Transactions on Information Theory*, 61(3):1458–1473, 2015. doi:10.1109/TIT.2014.2387822.

[SAP17] Alexander Streltsov, Gerardo Adesso, and Martin B. Plenio. Colloquium: Quantum coherence as a resource. *Rev. Mod. Phys.*, 89:041003, Oct 2017. doi:10.1103/RevModPhys.89.041003.

[SCMdC18] B. Shanahan, A. Chenu, N. Margolus, and A. del Campo. Quantum speed limits across the quantum-to-classical transition. *Phys. Rev. Lett.*, 120:070401, February 2018. doi:10.1103/PhysRevLett.120.070401.

[SDB+21] Youngkyu Sung, Leon Ding, Jochen Braumüller, Antti Vepsäläinen, Bharath Kannan, Morten Kjaergaard, Ami Greene, Gabriel O. Samach, Chris McNally, David Kim, Alexander Melville, Bethany M. Niedzielski, Mollie E. Schwartz, Jonilyn L. Yoder, Terry P. Orlando, Simon Gustavsson, and William D. Oliver. Realization of High-Fidelity CZ and *ZZ*-Free iSWAP Gates with a Tunable Coupler. *Phys. Rev. X*, 11(2):021058, June 2021. doi:10.1103/PhysRevX.11.021058.

[Sha19] Farid Shahandeh. The resource theory of entanglement. *Quantum Correlations: A Modern Augmentation*, pages 61–109, 2019.

[SHC00] Yehuda Sharf, Timothy F. Havel, and David G. Cory. Spatially encoded pseudopure states for nmr quantum-information processing. *Phys. Rev. A*, 62:052314, Oct 2000. doi:10.1103/PhysRevA.62.052314.

[Shi95] Abner Shimony. Degree of entanglement. *Annals of the New York Academy of Sciences*, 755(1):675–679, 1995. doi:10.1111/j.1749-6632.1995.tb39008.x.

[Sho94] P.W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994. doi:10.1109/SFCS.1994.365700.

[SKB10] Alexander Streltsov, Hermann Kampermann, and Dagmar Bruß. Linking a distance measure of entanglement to its convex roof. *New Journal of Physics*, 12(12):123004, dec 2010. doi:10.1088/1367-2630/12/12/123004.

[SKW+18] Alexander Streltsov, Hermann Kampermann, Sabine Wölk, Manuel Gessner, and Dagmar Bruß. Maximal coherence and the resource theory of purity. *New Journal of Physics*, 20(5):053058, may 2018. doi:10.1088/1367-2630/aac484.

[SL06] Graeme Smith and Debbie Leung. Typical entanglement of stabilizer states. *Physical Review A*, 74(6):062314, 2006. doi:10.1103/PhysRevA.74.062314.

[SMO22] Tanmay Singal, Filip B. Maciejewski, and Michał Oszmaniec. Implementation of quantum measurements using classical resources and only a single ancillary qubit. *npj Quantum Information*, 8(1):82, 2022. doi:10.1038/s41534-022-00589-1.

[SNS24] Manfredi Scalici, Moein Naseri, and Alexander Streltsov. Coherence generation with hamiltonians. *arXiv:2402.17567*, 2024.

[SQC+21] Armands Strikis, Dayue Qin, Yanzhu Chen, Simon C. Benjamin, and Ying Li. Learning-Based Quantum Error Mitigation. *PRX Quantum*, 2(4):040330, November 2021. doi:10.1103/PRXQuantum.2.040330.

[SR20] Jaime Sevilla and C Jess Riedel. Forecasting timelines of quantum computing. *arXiv:2009.05045*, 2020.

[Str15] Alexander Streltsov. *Quantum Correlations Beyond Entanglement*. SpringerBriefs in Physics, 2015. doi:10.1007/978-3-319-09656-8.

[TBG17] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error Mitigation for Short-Depth Quantum Circuits. *Phys. Rev. Lett.*, 119(18):180509, November 2017. doi:10.1103/PhysRevLett.119.180509.

[TLM19] J Teittinen, H Lyyra, and S Maniscalco. There is no general connection between the quantum speed limit and non-Markovianity. *New Journal of Physics*, 21(12):123041, December 2019. doi:10.1088/1367-2630/ab59fe.

[TLTC23] Andrew K. Tan, Yuan Liu, Minh C. Tran, and Isaac L. Chuang. Error Correction of Quantum Algorithms: Arbitrarily Accurate Recovery Of Noisy Quantum Signal Processing. *arXiv*, January 2023. arXiv:2301.08542, doi:10.48550/arXiv.2301.08542.

[TM21] Jose Teittinen and Sabrina Maniscalco. Quantum speed limit and divisibility of the dynamical map. *Entropy*, 23(3):331, 2021. doi:10.3390/e23030331.

[TRS22] Masaya Takahashi, Swapan Rana, and Alexander Streltsov. Creating and destroying coherence with quantum channels. *Phys. Rev. A*, 105:L060401, Jun 2022. doi:10.1103/PhysRevA.105.L060401.

[TSM+22] Dimpi Thakuria, Abhay Srivastav, Brij Mohan, Asmita Kumari, and Arun Kumar Pati. Generalised quantum speed limit for arbitrary evolution. *arXiv:2207.04124*, 2022.

[TV14] Francesco Ticozzi and Lorenza Viola. Quantum resources for purification and cooling: fundamental limits and opportunities. *Scientific reports*, 4(1):5192, 2014. doi:10.1038/srep05192.

[TWP09] Sayatnova Tamaryan, Tzu-Chieh Wei, and Dae Kil Park. Maximally entangled three-qubit states via geometric measure of entanglement. *Phys. Rev. A*, 80:052315, Nov 2009. doi:10.1103/PhysRevA.80.052315.

[VdNDDM04] Maarten Van den Nest, Jeroen Dehaene, and Bart De Moor. Graphical description of the action of local clifford transformations on graph states. *Phys. Rev. A*, 69:022316, Feb 2004. doi:10.1103/PhysRevA.69.022316.

[VW02] G. Vidal and R. F. Werner. Computable measure of entanglement. *Phys. Rev. A*, 65:032314, Feb 2002. doi:10.1103/PhysRevA.65.032314.

[Wan23] Dong-Sheng Wang. Universal resources for quantum computing. *Communications in Theoretical Physics*, 75(12):125101, 2023. doi:10.1088/1572-9494/ad07d6.

[WG03] Tzu-Chieh Wei and Paul M. Goldbart. Geometric measure of entanglement and applications to bipartite and multipartite quantum states. *Phys. Rev. A*, 68:042307, Oct 2003. doi:10.1103/PhysRevA.68.042307.

[Wil11] Mark M Wilde. From classical to quantum shannon theory. *arXiv:1106.1445*, 2011.

[WKR+21a] Kang-Da Wu, Tulja Varun Kondra, Swapan Rana, Carlo Maria Scandolo, Guo-Yong Xiang, Chuan-Feng Li, Guang-Can Guo, and Alexander Streltsov. Operational resource theory of imaginarity. *Phys. Rev. Lett.*, 126:090401, March 2021. doi:10.1103/PhysRevLett.126.090401.

[WKR+21b] Kang-Da Wu, Tulja Varun Kondra, Swapan Rana, Carlo Maria Scandolo, Guo-Yong Xiang, Chuan-Feng Li, Guang-Can Guo, and Alexander Streltsov. Resource theory of imaginarity: Quantification and state conversion. *Phys. Rev. A*, 103:032401, March 2021. doi:10.1103/PhysRevA.103.032401.

[WSR+21] Kang-Da Wu, Alexander Streltsov, Bartosz Regula, Guo-Yong Xiang, Chuan-Feng Li, and Guang-Can Guo. Experimental progress on quantum coherence: Detection, quantification, and manipulation. *Advanced Quantum Technologies*, 4(9):2100040, 2021. doi:10.1002/qute.202100040.

[WY16] Andreas Winter and Dong Yang. Operational resource theory of coherence. *Phys. Rev. Lett.*, 116:120404, Mar 2016. doi:10.1103/PhysRevLett.116.120404.

[XIBJ22] Qian Xu, Joseph K. Iverson, Fernando G. S. L. Brandão, and Liang Jiang. Engineering fast bias-preserving gates on stabilized cat qubits. *Phys. Rev. Res.*, 4(1):013082, February 2022. doi:10.1103/PhysRevResearch.4.013082.

[YHT+20] Yuan Yuan, Zhibo Hou, Jun-Feng Tang, Alexander Streltsov, Guo-Yong Xiang, Chuan-Feng Li, and Guang-Can Guo. Direct estimation of quantum coherence by collective measurements. *npj Quantum Information*, 6(1):46, May 2020. doi:10.1038/s41534-020-0280-6.

[YMG+16] Benjamin Yadin, Jiajun Ma, Davide Girolami, Mile Gu, and Vlatko Vedral. Quantum Processes Which Do Not Use Coherence. *Phys. Rev. X*, 6:041028, Nov 2016. doi:10.1103/PhysRevX.6.041028.

[ZLL+22] Peng Zhao, Kehuan Linghu, Zhiyuan Li, Peng Xu, Ruixia Wang, Guangming Xue, Yirong Jin, and Haifeng Yu. Quantum Crosstalk Analysis for Simultaneous Gate Operations on Superconducting Qubits. *PRX Quantum*, 3(2):020301, April 2022. doi:10.1103/PRXQuantum.3.020301.

[ZSW20] Yiqing Zhou, E. Miles Stoudenmire, and Xavier Waintal. What Limits the Simulation of Quantum Computers? *Phys. Rev. X*, 10(4):041038, November 2020. doi:10.1103/PhysRevX.10.041038.

[Zwi12] Marcin Zwierz. Comment on "geometric derivation of the quantum speed limit". *Phys. Rev. A*, 86:016101, July 2012. doi:10.1103/PhysRevA.86.016101.