

Multiple-shot discrimination of quantum unitary channels and its implementation

Vícenásobná diskriminace kvantových unitárních kanálů a její implementace

Jan Hlisnikovský

Diploma Thesis

Supervisor: Dr. Paulina Lewandowska

Ostrava, 2026

Diploma Thesis Assignment

Student: **Bc. Jan Hlisenikovsky**

Study Programme: N0541A170007 Computational and Applied Mathematics

Specialization: S01 Applied Mathematics

Title: Multiple-shot discrimination of quantum unitary channels and its implementation
Vícenásobná diskriminace kvantových unitárních kanálů a její implementace

The thesis language: English

Description:

Quantum information theory provides the theoretical foundation for quantum computing, a rapidly advancing field with the potential to revolutionize numerous applications.

A central challenge within quantum information theory is distinguishing between quantum channels. This task involves determining the maximum probability of correctly identifying a channel used in computation, given a finite number of channel copies and specific configurations, and developing strategies to achieve this optimal discrimination. However, limitations of Noisy Intermediate-Scale Quantum (NISQ) devices—such as restricted qubit coherence times and limited qubit connectivity—hinder the practical implementation of these optimal discrimination strategies, making it difficult to fully exploit quantum advantages in channel discrimination tasks.

This thesis is dedicated to designing a new approach for discriminating unitary channels within these existing constraints.

The tools for the thesis:

1. Comprehensive Literature Review: Conduct a thorough review of existing literature on quantum channel discrimination to identify current approaches, limitations, and gaps in the field.
2. Establishing Mathematical Foundations: Present essential mathematical preliminaries, along with the necessary concepts in quantum information theory, to provide a strong theoretical foundation.
3. Defining the Quantum Discrimination Task: Clearly define the quantum discrimination task, explaining its importance, challenges, and relevance to the field of quantum information.
4. Design and Implementation on Quantum Architecture: Develop and implement a practical quantum discrimination approach using a suitable quantum architecture.
5. Execution on Quantum Hardware or Simulator: Test the implemented discrimination strategy by running it on an available quantum computer or simulator to evaluate its performance.
6. Analysis of Results: Analyze the experimental or simulated results to assess the effectiveness and potential limitations of the discrimination approach.
7. Drawing Conclusions and Future Directions: Summarize findings, discuss implications, and suggest possible directions for further research in quantum channel discrimination.

References:

Watrous J. The Theory of Quantum Information. Cambridge University Press; 2018.
Nielsen, M. A., & Chuang, I. L. (2010). Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge: Cambridge University Press.
Liu, Shusen, Yinan Li, and Runyao Duan. "Distinguishing unitary gates on the IBM quantum processor." Science China Information Sciences 62 (2019): 1-7.
R. Duan, Y. Feng, and M. Ying, Physical review letters 98, 100503 (2007)
G. Chiribella, G. M. D'Ariano, and P. Perinotti, Physical review letters 101, 180501 (2008)
Qiskit documentation, available at <https://docs.quantum.ibm.com/>

Extent and terms of a thesis are specified in directions for its elaboration that are opened to the public on the web sites of the faculty.

Supervisor: **Dr. Paulina Lewandowska**

Date of issue: 01.09.2025

Date of submission: 30.04.2026

Study programme guarantor: prof. RNDr. Jiří Bouchala, Ph.D.

In IS EDISON assigned: 10.11.2025 08:48:43

Abstrakt

Tato práce zkoumá vícenásobné rozlišování unitárních kvantových kanálů na NISQ hardwaru, přičemž porovnává paralelní, sekvenční a hybridní konfigurace. V teoretické části představujeme geometrický algoritmus pro nalezení diskriminátoru pro dokonalé rozlišení. Experimentální data však ukazují, že teoreticky suboptimální obvody mohou na zašuměných zařízeních dosahovat lepších výsledků. Experimenty využívající 6 a 11 qubitů s metodou zmírňování chyb M3 navíc naznačují, že hardwarově přizpůsobená transpilace je žádoucí, jelikož konfigurace s velkou hloubkou a malou šířkou se ukazují jako spolehlivější než vysoce provázaná paralelní schémata. Při testování čistě paralelních schémat se na hardwaru IBM objevila anomálie.

Klíčová slova

kvantové počítání, diskriminace kvantových kanálů, NISQ zařízení, IBM Brisbane, unitární kanály

Abstract

This thesis investigates multiple-shot discrimination of unitary quantum channels on NISQ hardware, comparing parallel, sequential, and hybrid strategies. On the theoretical side, we introduce a hybrid geometric solver for constructing optimal input states for perfect discrimination tasks. Experimentally, we observe that theoretically suboptimal circuits can outperform optimal ones on noisy quantum devices. Results obtained on 6- and 11-qubit systems with M3 error mitigation highlight the importance of hardware-aware transpilation, with high-depth, low-width circuits proving more robust than highly entangled parallel schemes. Finally, we report an unexpected anomaly on IBM hardware when evaluating purely parallel strategies.

Keywords

quantum computing, discrimination of quantum channels, NISQ devices, IBM Brisbane, unitary channels

Acknowledgement

I would like to express my gratitude to my family for their support throughout my studies. I am also sincerely grateful to my supervisor, Dr. Paulina Lewandowska, for her invaluable guidance and crucial insights. Finally, I wish to thank AB, TB, RK, ML, TD, MB, and everyone else who supported me along the way.

Contents

List of symbols and abbreviations	8
List of Figures	10
List of Tables	13
1 Introduction	14
2 Mathematical preliminaries	17
2.1 Complex Euclidean spaces	17
2.2 Linear operators	18
2.3 Linear maps on square operators	26
3 Fundamentals of quantum information theory	31
3.1 Quantum states	31
3.2 Transformations of quantum states	34
3.3 Quantum computing	38
4 Discrimination of unitary quantum channels	42
4.1 Single-shot discrimination of unitary channels	42
4.2 Multiple-shot discrimination of unitary channels	43
4.3 Discriminator construction for perfect discrimination of unitary channels	46
4.4 Measurement construction via Jordan-Hahn decomposition	51
5 Components of the multiple-shot perfect discrimination task	52
5.1 Discrimination of unitary channels on IBM Brisbane without processing	53
5.2 Discrimination of unitary channels on IBM Brisbane with processing	54
6 Implementation of experiments on IBM Brisbane processor	56
6.1 Topology and gate set of IBM Brisbane	56
6.2 Experiments without processing	63

6.3 Experiments with processing	71
7 Conclusion	74
7.1 Future work	75
Bibliography	76
Appendices	80
A Publication	81

List of symbols and abbreviations

$\dim(\mathcal{X})$	– Dimension of $\mathcal{X}$
δ_{ij}	– Kronecker delta
$\mathbb{C}$	– Set of complex numbers
$\mathbb{R}$	– Set of real numbers
$\mathbb{R}_0^+$	– Set of non-negative real numbers
$\mathcal{X}^*$	– Dual space of $\mathcal{X}$
$\text{span}(S)$	– Linear span of set S
$\langle x y\rangle$	– Inner product of vectors $ x\rangle$ and $ y\rangle$
$ i\rangle$	– Basis vector where only the i -th position is one, rest is equal to zero
$\ x\rangle \ _p$	– Vector p -norm
$\ x\rangle \ $	– Euclidean vector norm ($p = 2$)
$\ x\rangle \ _\infty$	– Chebyshev vector norm
$\otimes$	– Tensor product
$ x\rangle^{\otimes k}$	– k -fold tensor product of vector $ x\rangle$
$\mathcal{L}(\mathcal{X}, \mathcal{Y})$	– Set of linear operators mapping $\mathcal{X}$ to $\mathcal{Y}$
$\mathcal{L}(\mathcal{X})$	– Set of linear operators mapping $\mathcal{X}$ to $\mathcal{X}$
$ A\rangle\rangle$	– Vectorization of operator A
$\overline{A}$	– Element-wise complex conjugate of operator A
A^T	– Transpose of operator A
$A^\dagger$	– Hermitian transpose (adjoint) of operator A
$\text{im}(A)$	– Image of operator A
$\text{rank}(A)$	– Rank of operator A
$ y\rangle\langle x $	– Outer product
$\text{Tr}(X)$	– Trace of operator X
$\langle A, B\rangle$	– Hilbert-Schmidt inner product of operators
$\text{Herm}(\mathcal{X})$	– Set of Hermitian operators on $\mathcal{X}$
$\text{Pos}(\mathcal{X})$	– Set of positive semidefinite operators on $\mathcal{X}$
$\text{Pd}(\mathcal{X})$	– Set of positive definite operators on $\mathcal{X}$

$\text{Proj}(\mathcal{X})$	– Set of projection operators on $\mathcal{X}$
$\Pi_{\mathcal{Y}}$	– Projection operator onto subspace $\mathcal{Y}$
$\mathbb{I}_{\mathcal{X}}$	– Identity operator on $\mathcal{X}$
$\text{U}(\mathcal{X})$	– Set of unitary operators on $\mathcal{X}$
$\text{D}(\mathcal{X})$	– Set of density operators on $\mathcal{X}$
$\ A\ _p$	– Schatten p -norm
$\ A\ _{\infty}$	– Spectral norm
$\lambda(X)$	– Spectrum (set of eigenvalues) of X
$\sigma(X)$	– Set of singular values of X
$W(X)$	– Numerical range of X
$\nu(X)$	– Distance of numerical range from the origin
$\theta(U)$	– Arc function of unitary U
$\text{T}(\mathcal{X}, \mathcal{Y})$	– Set of linear maps (superoperators) from $\text{L}(\mathcal{X})$ to $\text{L}(\mathcal{Y})$
Φ^*	– Adjoint of linear map Φ
$\mathbb{I}_{\text{L}(\mathcal{X})}$	– Identity superoperator
$\text{Tr}_{\mathcal{X}}$	– Partial trace over $\mathcal{X}$
$J(\Phi)$	– Choi representation of map Φ
$\ \Phi\ _1$	– Induced trace norm of a map
$\ \Phi\ _{\diamond}$	– Diamond norm (completely bounded trace norm)
SVD	– Singular Value Decomposition
CP	– Completely Positive
TP	– Trace-Preserving
SDP	– Semidefinite Programming
$P(i)$	– Probability of measurement outcome i
Φ_U	– Unitary quantum channel defined by operator U
$\mathbb{N}_0$	– Set of natural numbers with 0

List of Figures

2.1	Comparison of eigenvalues and numerical ranges in the complex plane for U in (a) and $U \otimes U$ in (b) where $U = \text{diag}([1, e^{\frac{2i\pi}{3}}]^T)$. In both cases the numerical range is a convex hull of eigenvalues. In (a) such hull creates line, in (b) the hull creates triangle.	25
2.2	Comparison of the smallest angle containing all eigenvalues of two unitaries U in (a) and $U \otimes U$ in (b) where $U = \text{diag}([1, e^{\frac{2i\pi}{3}}]^T)$	25
3.1	Schematic representation of a quantum circuit measurement performed in the computational basis.	39
3.2	Schematic representation of a von Neumann measurement realized by applying the inverse unitary $U^\dagger$ followed by a measurement in the computational basis.	40
3.3	Two-qubit quantum circuit demonstrating the sequence of X , H , and CNOT gate operations ended by measurement in computational basis.	40
4.1	Ancilla-assisted quantum discrimination scheme between the quantum channels Φ_U and $\Phi_{\mathbb{1}_{\mathbb{C}_2}}$. An input $ x\rangle$ (possibly entangled with an auxiliary system) interacts with an unknown quantum channel. A POVM measurement M is then performed on the system to identify the outcome.	43
4.2	Schematic representation of a sequential discrimination strategy. The unknown unitary W is applied to a single system three times, for $d = 3$ and $w = 1$. In the sequential strategy, the processing gates X_i might be used.	44
4.3	Schematic representation of a parallel discrimination strategy. Three copies of the transformed unitary W are applied in a parallel way. In this case, $w = 3$ and $d = 1$	45
4.4	Schematic representation of a hybrid discrimination strategy. The unknown unitary W is applied twelve times to a system of width $w = 3$ and depth $d = 4$. Intermediate processing gates $X_{k,i}$ ($X_{k,i}$ may be $\mathbb{1}$) are applied sequentially on each qubit.	46

6.1	The coupling map of the IBM Eagle R3 processor, representing a 127-qubit system. The architecture is defined by a heavy-hexagonal graph topology, where the maximum degree of any node is $d = 3$	57
6.2	A five qubit discriminator circuit designed to distinguish between RZ rotations and the identity operation $\mathbb{1}_X$. The implementation utilizes a hardware-native gate set consisting of $\sqrt{X}$, X , and ECR gates.	58
6.3	Comparison of qubit mapping topologies on the Eagle R3 heavy-hex mesh. In (a) , we utilize a one-way branching (linear) mapping, which represents the simplest path through the lattice. In (b) , a two-way branching strategy is shown for which the central node (41) serves as a hub for two outgoing ECR gates. Finally, (c) illustrates a three-way branching mapping, which fully saturates the maximum degree of a node within the heavy-hex topology.	59
6.4	A five-qubit discriminator circuit mapped using a two-way branching strategy (strategy b). Second qubit serves as the primary hub, initiating the entanglement chain that branches through third qubit to reach the peripheral qubits.	59
6.5	The XOR pre-measurement rotation circuit for 5 qubits. The circuit implements a mirrored ECR gate structure. Such circuit enables for post-processing using Hamming distance.	60
6.6	The short pre-measurement rotation circuit for 5 qubits. The circuit trades its length for potential for post-processing of results.	61
6.7	A scheme of circuit creating discriminator for 5 qubits using Hadamard gates and CNOT.	61
6.8	Implementation of quantum circuits for pre-measurement rotations for the 5-qubit branching scheme both using CNOT and H gates. In (a) we present equivalent to the short measurement. In (b) is circuit, which maximizes the Hamming distance for theoretical outcomes between identity and rotation.	62
6.9	Implementation of a 4-qubit discriminator state preparation using a traditional cascade pattern for the task of discrimination with processing involving $\sqrt{X}$. The discriminator is created using H , CNOT, $\sqrt{X}$ and X gates.	62
6.10	Implementation of a 4-qubit measurement for the discrimination task with processing. The measurement utilizes H and $RZ(\frac{\pi}{2})$ gates.	63
6.11	Schematic architectural mapping of a six-qubit discriminator across distinct basis gate configurations: (a) A standard decomposition utilizing Hadamard and CNOT gates; (b) A hardware-efficient implementation optimized for the IBM Eagle R3 processor, utilizing the native $\sqrt{X}$ and ECR gate set.	64

6.12	Schematic of six-qubit measurement circuits. Figs. (a) and (b) show the short and XOR schemes using non-native H and CNOT gates. Figs. (c) and (d) show the same schemes using native Eagle R3 gates, in this case the gates used are ECR and $\sqrt{X}$	65
6.13	Probability of successful discrimination between the identity operation and the $RZ(\pi/N)$ gate, where N is the number of copies of the unknown unitary. The dashed red line denotes the XOR measurement strategy, while the solid blue line represents the short measurement. (a) Performance of the purely sequential scheme, peaking at approximately 0.965. (b) Purely parallel scheme, displaying reduced performance from entanglement overhead. Each circuit ran with 10,000 shots. Ambiguous measurement outcomes were randomly assigned to either gate.	69
6.14	Successful discrimination probability between the identity operation and the $RZ(\pi/N)$ gate versus the hybrid scheme width. The dashed red line denotes the XOR measurement, and the solid blue line represents the short measurement. Panels show fixed total resources of (a) $N = 120$, (b) $N = 240$, and (c) $N = 1200$ copies. Performance degrades as scheme width increases, correlating with higher entanglement density. Circuits ran with 10,000 shots. Ambiguous outcomes were randomly assigned.	70
6.15	Successful discrimination probability between unitary operators $U = \sqrt{X}RZ(\frac{-\pi}{2N})\sqrt{X}$ and $V = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$ (N copies) versus the hybrid scheme width, using the short measurement. The blue line shows raw Brisbane results. The red line shows MThree error-mitigated results [49]. Panels display performance for total copies: (a) $N = 4$, (b) $N = 16$, and (c) $N = 32$. Each circuit ran with 10,000 shots.	72
6.16	Successful discrimination probability between unitary operators $U = \sqrt{X}RZ(\frac{-\pi}{2N})\sqrt{X}$ and $V = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$ (N copies) versus the hybrid scheme width, using the short measurement. The blue line shows raw IBM Brisbane results; the red line shows MThree error-mitigated results [49]. Panels display performance for total copies: (a) $N = 64$, (b) $N = 96$, and (c) $N = 1024$. Each circuit ran with 10,000 shots.	73

List of Tables

3.1	Single-qubit quantum gates with their circuit symbols and corresponding unitary matrix representations.	39
3.2	Two-qubit quantum gates with their circuit symbols and corresponding unitary matrix representations.	40
6.1	Comparison of 6-qubit parallel discrimination circuit metrics across different transpilation strategies and measurement types. Metrics include circuit depth and gate counts for ECR, RZ , $\sqrt{X}$, X , and $\mathbb{1}_X/RZ(\phi)$ gates.	66
6.2	Experimental accuracy results for different transpilation strategies. Comparison includes short and XOR measurement accuracies for CNOT and ECR-based implementations. Again, the setting is the parallel scheme.	66
6.3	Comparison of quantum circuit complexity across various transpilation strategies and measurement methods. The table tracks circuit depth and gate counts for ECR, RZ , $\sqrt{X}$, X , and $ID/RZ(\phi)$ basis gates.	67
6.4	Experimental accuracy results for the 11-qubit discrimination task. The table compares the performance of Short and XOR measurement schemes across selected transpilation strategies. Values represent the percentage of successful state discriminations, highlighting the impact of topology-aware and fixed mapping techniques on final readout fidelity.	68

Chapter 1

Introduction

In recent years, quantum computing has transitioned from small-scale experimental prototypes toward processors with more than 100 qubits [1]. We remain in the Noisy Intermediate-Scale Quantum (NISQ) era, a period defined by the availability of heterogeneous hardware platforms. These platforms include superconducting circuits provided by IBM and Rigetti [2, 3] and other options, such as trapped-ion systems from IonQ [4], both accessible via the cloud. While high-level programming frameworks abstract much of the hardware complexity, the fundamental limits of quantum information processing must be established to provide a baseline for assessing hardware-induced degradation.

One of the fundamental approaches to quantum hardware validation is quantum operator discrimination. The theoretical framework for quantum state discrimination relies on the density operator formalism introduced by von Neumann [5] to represent quantum states. Born’s rule [6] links this formalism to measurement probabilities, reflecting the inherently probabilistic nature of quantum measurements. This probabilistic nature gives rise to the problem of state discrimination. Holevo [7] and Helstrom [8] established the fundamental limits of state discrimination, deriving the optimal success probability for distinguishing between quantum states in a single-shot scenario. In the asymptotic regime of multiple copies, the focus shifts to the minimum error probability, governed by the quantum Chernoff bound [9]. This bound defines the optimal exponential decay rate of the discrimination error.

While the Holevo–Helstrom theorem defines the fundamental limits of quantum state discrimination, this paradigm must be extended to transformations of quantum states. Quantum channel discrimination generalizes these principles by quantifying the distinguishability of channels in both single- and multi-shot regimes. The diamond norm, introduced by Kitaev [10], defines an operationally meaningful metric on quantum channels. The diamond norm measures the distance between channels and directly characterizes the optimal success probability of distinguishing them. The discrimination of quantum measurements can also be described using the diamond norm by representing a measurement as a measure-and-prepare quantum channel [11]. In particular, the

discrimination of von Neumann measurements has been studied in a single-shot setting [11], in a multiple-shot setting [12], and has also been used for benchmarking gate-based quantum computers in PyQBench [13].

In this thesis, we will focus on a specific class of quantum channels, namely unitary channels. It turned out that the distance between two unitary channels U and V can be reduced to the analysis of the relative unitary $UV^\dagger$. In particular, the distance between the channels can be expressed in terms of the distance between zero and the numerical range of this operator [14, 15, 16, 17]. Furthermore, a fundamental result by Duan et al. [18] establishes that any two distinct unitary operations can be perfectly discriminated using a finite number of applications. The required number of uses can be further characterized through the numerical range and the arc function [19].

The efficiency of multiple-shot unitary discrimination strongly depends on the query architecture, such as parallel [20], sequential [18], or adaptive use of the unknown channel. These strategies can be described within higher-order frameworks for quantum transformations, including quantum combs [21]. Related higher-order discrimination problems have also been studied, for example in the context of process matrix discrimination [22]. The relative power of different query architectures has been studied in several settings: adaptive strategies can outperform nonadaptive ones for general channels [23], and strict separations between parallel, sequential, and indefinite-causal-order strategies have been compared for channel discrimination [24].

While theoretical frameworks establish the fundamental bounds for optimal discrimination, the performance of discrimination strategies under the constraints of contemporary NISQ hardware requires further empirical validation. This thesis evaluates and compares sequential, parallel, and hybrid discrimination schemes, specifically examining their robustness within a noisy experimental environment using the IBM quantum computer Brisbane. Moreover, we compare optimal and sub-optimal strategies to explore possible trade-offs. Lastly, we investigate practical improvements for the classical preprocessing of the discrimination task. The aim is to clarify how hardware topology, transpilation, entangling-gate overhead, and measurement noise affect the practical implementation of theoretically optimal discrimination protocols. This thesis is primarily based on a publication authored by the candidate, entitled *Experimental study of multiple-shot unitary channel discrimination using IBM quantum computers*, published in *Scientific Reports*.

Thesis organization

This thesis is organized into seven chapters, including this introduction. Chapter 2 introduces the mathematical concepts needed in the rest of the thesis. Chapter 3 presents the fundamental notions of quantum information theory, including quantum states, quantum channels, and measurements. Chapter 4 introduces the discrimination task of unitary quantum channels. It discusses single-shot and multiple-shot discrimination, compares parallel, sequential, and hybrid strategies, and presents the construction of discriminator states and final measurements. Chapter 5 applies this

theoretical framework to the specific discrimination tasks used in the experiments, both with and without intermediate processing. Chapter 6 describes the implementation of these tasks on the IBM Brisbane processor and presents the experimental results. Finally, Chapter 7 summarizes the results and outlines directions for future research.

Author contribution

Both the author and the supervisor of this thesis are authors of the scientific article [25] published in Scientific Reports. I designed and implemented quantum circuits realizing the discriminator and two types of measurements using native gate sets, as presented in Experiment 1 in [25]. These decompositions are essential for analyzing the trade-offs between circuit width and depth. In addition, I contributed to the statistical analysis by providing interval estimates for experiments without intermediate processing (Experiment 1), and I contributed to the development of the publication’s codebase. I also participated in the preparation of the manuscript and provided consultation on subsequent experiments (including mapping and related aspects) carried out by other authors.

Furthermore, building on the results of the publication, I developed a framework that employs Carathéodory’s theorem [26] in conjunction with sorting and searching algorithms to construct a discriminator for the multiple-shot scenario. These results are intended for submission to a peer-reviewed journal in the near future.

Statement on the usage of artificial intelligence

During the preparation of this thesis, Google’s Gemini models [27] were used for linguistic refinement, including grammar corrections and formatting.

Chapter 2

Mathematical preliminaries

Understanding quantum information theory requires a solid mathematical foundation. This chapter introduces the key concepts necessary for the remainder of the thesis. It is mainly based on [28], with the main difference being the use of Dirac notation [29].

2.1 Complex Euclidean spaces

In quantum information theory, vectors are represented using bra-ket notation (sometimes known as Dirac notation). Let $\mathcal{X}$ be a complex Euclidean space with $\dim(\mathcal{X}) = n + 1$. A column vector $|x\rangle \in \mathcal{X}$ is known as a *ket* and is expressed as:

$$|x\rangle = (x_0, \dots, x_n)^T, \quad (2.1)$$

where $x_i \in \mathbb{C}$ for all $i \in \{0, \dots, n\}$.

To define the corresponding row vector, consider the dual space $\mathcal{X}^*$. By the Riesz representation theorem [30, 31], there exists a unique anti-linear isometric isomorphism between $\mathcal{X}$ and $\mathcal{X}^*$. The unique linear functional corresponding to $|x\rangle$ is known as a bra, denoted as $\langle x|$. This functional is given by applying element-wise complex conjugation followed by transposition to $|x\rangle$ as:

$$\langle x| = (\bar{x}_0, \dots, \bar{x}_n), \quad (2.2)$$

where $\bar{x}_i$ denotes the complex conjugate of x_i . Consequently, for any $|x\rangle, |y\rangle \in \mathcal{X}$, the inner product is defined via the action of the linear functional $\langle x|$ on the vector $|y\rangle$:

$$\langle x|y\rangle = \sum_{i=0}^n \bar{x}_i y_i. \quad (2.3)$$

Definitions regarding vector norms, span, bases and potentially other elementary linear algebra definitions are adopted from [32, 33, 34, 28]. We define the vector $|i\rangle$ as a vector in which only the i -th position is one, and the rest are zero¹. The computational basis is then defined as $\{|i\rangle\}_{i=0}^n$.

2.1.1 Tensor product of complex Euclidean spaces

An essential concept for quantum information theory is the tensor product of two vectors, sometimes known as the Kronecker product. Consider two complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$ of dimension $\dim(\mathcal{X}) = n + 1, \dim(\mathcal{Y}) = m + 1$. Let $|x\rangle \in \mathcal{X}, |y\rangle \in \mathcal{Y}$, we define the tensor product of two vectors as:

$$|x\rangle \otimes |y\rangle = (x_0y_0, \dots, x_0y_m, \dots, x_ny_0, \dots, x_ny_m)^T. \quad (2.4)$$

This enables us to define the tensor product of the two complex Euclidean spaces as:

$$\mathcal{X} \otimes \mathcal{Y} = \text{span}(\{|x\rangle \otimes |y\rangle : |x\rangle \in \mathcal{X}, |y\rangle \in \mathcal{Y}\}). \quad (2.5)$$

We will use shorthand notation regarding the tensor product of one vector (or space) with itself. Let $\mathcal{X}$ be a complex Euclidean space and $|x\rangle \in \mathcal{X}$. We understand $\mathcal{X}^{\otimes k}$ as a k -fold tensor product between $\mathcal{X}$. Similar notation holds for $|x\rangle$, the $|x\rangle^{\otimes k}$ is k -fold tensor product of $|x\rangle$. Additionally, suppose each $\mathcal{X}_i$ is a complex Euclidean space for each $i \in \{0, \dots, n\}$. We define another shorthand as follows:

$$\bigotimes_{i=0}^n \mathcal{X}_i = \mathcal{X}_0 \otimes \dots \otimes \mathcal{X}_n, \quad (2.6)$$

moreover let $|x_i\rangle \in \mathcal{X}_i$ for all $i \in \{0, \dots, n\}$, then:

$$\bigotimes_{i=0}^n |x_i\rangle = |x_0\rangle \otimes \dots \otimes |x_n\rangle. \quad (2.7)$$

2.2 Linear operators

In this section, we introduce linear operators. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. By $L(\mathcal{X}, \mathcal{Y})$ we refer to the collection of all linear mappings of the form:

$$A : \mathcal{X} \rightarrow \mathcal{Y} \quad (2.8)$$

In case $\mathcal{Y} = \mathcal{X}$, the notation may be shortened to $L(\mathcal{X}, \mathcal{X}) = L(\mathcal{X})$. If not said otherwise, all operators considered in this thesis are linear. The set $L(\mathcal{X}, \mathcal{Y})$ constitutes a linear space over the field $\mathbb{C}$ where operator addition and scalar multiplication are defined as in [28].

¹We use indexing from zero as well as using binary representation of the index number.

There is a linear bijection from complex Euclidean spaces $L(\mathcal{X}, \mathcal{Y})$ to $\mathcal{Y} \otimes \mathcal{X}$ for any choice of complex Euclidean spaces $\mathcal{X}$ and $\mathcal{Y}$. Let $\dim(\mathcal{X}) = n+1$. Such a bijection, known as a vectorization, is then defined (with respect of computational basis) as:

$$|A\rangle\rangle = \sum_{i=0}^n A|i\rangle \otimes |i\rangle, \text{ for any } A \in L(\mathcal{X}, \mathcal{Y}). \quad (2.9)$$

2.2.1 Tensor product of linear operators

Similarly to the tensor product between vectors, one may consider defining a tensor product between linear operators. Let $\mathcal{X}_i, \mathcal{Y}_i$ be complex Euclidean spaces and $A_i \in L(\mathcal{X}_i, \mathcal{Y}_i)$ for $i \in \{0, \dots, n\}$, $n \in \mathbb{N}$. We define the tensor product of operators $A_i \in L(\mathcal{X}_i, \mathcal{Y}_i)$ as:

$$A_0 \otimes \dots \otimes A_n \in L\left(\bigotimes_{i=0}^n \mathcal{X}_i, \bigotimes_{i=0}^n \mathcal{Y}_i\right) \quad (2.10)$$

satisfying the condition:

$$(A_0 \otimes \dots \otimes A_n)(|x_0\rangle \otimes \dots \otimes |x_n\rangle) = \bigotimes_{i=0}^n A_i |x_i\rangle, \quad (2.11)$$

which holds for all $|x_i\rangle \in \mathcal{X}_i$, $i \in \{0, \dots, n\}$.

2.2.2 Basic properties of linear operators

Now, we discuss some basic properties and operations for linear operators. There exists a bijection between matrices and linear operators. Suppose that we are given an operator $M \in L(\mathcal{X}, \mathcal{Y})$, the corresponding matrix elements are given by $M_{i,j} = \langle i|M|j\rangle$.

For given complex Euclidean spaces $\mathcal{X}, \mathcal{Y}$ and operator $A \in L(\mathcal{X}, \mathcal{Y})$, we can define three more operators corresponding to A . Firstly, $\bar{A} \in L(\mathcal{X}, \mathcal{Y})$ is defined as the element-wise complex conjugate of the original matrix. This can be expressed as $\bar{A}_{i,j} = \overline{A_{i,j}}$. Secondly, $A^T \in L(\mathcal{Y}, \mathcal{X})$ can be understood as $A_{i,j}^T = A_{j,i}$. Thirdly, we define the Hermitian transpose (also known as the conjugate transpose or adjoint operator) of the original matrix A as $A^\dagger = \overline{A^T}$. $A^\dagger$ is then an element of $L(\mathcal{Y}, \mathcal{X})$.

Let $A \in L(\mathcal{X}, \mathcal{Y})$ be an operator, then the image of A is defined as:

$$\text{im}(A) = \{A|x\rangle : |x\rangle \in \mathcal{X}\}, \quad (2.12)$$

and the rank of A is defined as:

$$\text{rank}(A) = \dim(\text{im}(A)). \quad (2.13)$$

One can define an operator $A \in L(\mathcal{X}, \mathcal{Y})$ using the outer product of two vectors. Given $|x\rangle = (x_0, \dots, x_n)^T \in \mathcal{X}$ and $|y\rangle = (y_0, \dots, y_m)^T \in \mathcal{Y}$. We define the operator A as:

$$A = |y\rangle\langle x| = \begin{bmatrix} y_0\overline{x_0} & y_0\overline{x_1} & \cdots & y_0\overline{x_n} \\ y_1\overline{x_0} & y_1\overline{x_1} & \cdots & y_1\overline{x_n} \\ \vdots & \vdots & \ddots & \vdots \\ y_m\overline{x_0} & y_m\overline{x_1} & \cdots & y_m\overline{x_n} \end{bmatrix}. \quad (2.14)$$

The trace of a square operator is defined as the sum of the operator's diagonal elements. We denote the trace linear operator of a matrix $X \in L(\mathcal{X})$ as $\text{Tr}(X)$. We use the trace operator to define an inner product on the complex Euclidean space of linear operators $L(\mathcal{X}, \mathcal{Y})$. Let $\mathcal{X}, \mathcal{Y}$ be complex Euclidean spaces, $A, B \in L(\mathcal{X}, \mathcal{Y})$, then, the inner product is defined as:

$$\langle A, B \rangle = \text{Tr}(A^\dagger B). \quad (2.15)$$

One of the useful identities for the trace operator is for an operator defined as a tensor product of two other operators. Let $C = A \otimes B$ for arbitrary $A \in L(\mathcal{X})$ and $B \in L(\mathcal{Y})$, then:

$$\text{Tr}(C) = \text{Tr}(A) \text{Tr}(B). \quad (2.16)$$

The trace mapping has a useful property, called the cyclic property, which can be expressed as:

$$\text{Tr}(AB) = \text{Tr}(BA), \quad (2.17)$$

for all $A \in L(\mathcal{X}, \mathcal{Y})$ and $B \in L(\mathcal{Y}, \mathcal{X})$. For vectors, a similar identity can be shown. Let $|x\rangle, |y\rangle \in \mathcal{X}$, then:

$$\text{Tr}(|x\rangle\langle y|) = \text{Tr}(\langle y|x\rangle) = \langle y|x\rangle. \quad (2.18)$$

It holds that:

$$\langle\langle A|B\rangle\rangle = \langle A, B \rangle. \quad (2.19)$$

2.2.3 Schatten p-norms

For the discrimination task, it is crucial to define a set of p -norms on the vector space of linear operators.

Let $A \in L(\mathcal{X}, \mathcal{Y})$ be an operator and let $p \in \mathbb{R} : p \geq 1$, the p -norm is defined as:

$$\|A\|_p = \left(\text{Tr} \left((A^\dagger A)^{\frac{p}{2}} \right) \right)^{\frac{1}{p}}, \quad (2.20)$$

and the ∞ -norm is:

$$\|A\|_\infty = \max\{\|A|x\rangle\| : |x\rangle \in \mathcal{X}, \langle x|x\rangle = 1\}. \quad (2.21)$$

This family of norms includes the well-known trace norm and the Frobenius norm by setting $p = 1$ and $p = 2$, respectively. The $p = \infty$ norm is also known as the spectral norm.

2.2.4 Eigenvalues and spectrum

Let $X \in L(\mathcal{X})$, we say a nonzero $|x\rangle \in \mathcal{X}$ is an eigenvector and $\lambda \in \mathbb{C}$ is its corresponding eigenvalue of the operator X if:

$$X|x\rangle = \lambda|x\rangle. \quad (2.22)$$

The set of all the eigenvalues, the spectrum, of X is denoted as $\lambda(X)$.

2.2.5 Key categories of square linear operators

Here we present some of the well-known and important linear operators, which play an essential role during the thesis as they are essential to understanding the quantum information theory.

1. **Normal operators.** An operator $X \in L(\mathcal{X})$ is normal if the following holds:

$$XX^\dagger = X^\dagger X. \quad (2.23)$$

2. **Hermitian operators.** An operator $X \in L(\mathcal{X})$ is Hermitian if it is equal to its Hermitian conjugate. The set of Hermitian operators mapping complex Euclidean space $\mathcal{X}$ to itself is denoted by $\text{Herm}(\mathcal{X})$ and defined as:

$$\text{Herm}(\mathcal{X}) = \{X \in L(\mathcal{X}) : X = X^\dagger\}. \quad (2.24)$$

3. **Positive semidefinite operators.** A Hermitian operator $P \in L(\mathcal{X})$ is positive semidefinite if $\langle x|P|x\rangle \geq 0$ for all $|x\rangle \in \mathcal{X}$. Equivalently, $P \in L(\mathcal{X})$ is positive semidefinite if there exists $Y \in L(\mathcal{X})$ such that $P = Y^\dagger Y$. The set of positive semidefinite operators $\text{Pos}(\mathcal{X})$ is defined as:

$$\text{Pos}(\mathcal{X}) = \{P \in L(\mathcal{X}) : \forall |x\rangle \in \mathcal{X}, \langle x|P|x\rangle \geq 0\}. \quad (2.25)$$

4. **Projectors.** Another subset of positive semidefinite operators is the set of projection operators (projectors). A positive semidefinite operator $\Pi \in \text{Pos}(\mathcal{X})$ is a projection operator if $\Pi = \Pi^2$ is satisfied. The set of all projective operators is then denoted as $\text{Proj}(\mathcal{X})$ and is given by:

$$\text{Proj}(\mathcal{X}) = \{\Pi \in \text{Pos}(\mathcal{X}) : \Pi = \Pi^2\}. \quad (2.26)$$

5. **Unitary operators.** A crucial subset of square operators $L(\mathcal{X})$ for quantum information theory and quantum computing is the set of unitary operators. An operator $U \in L(\mathcal{X})$ is

unitary if: $UU^\dagger = U^\dagger U = \mathbb{1}_\mathcal{X}$. The subset of square operators that satisfy the condition is denoted as:

$$\mathcal{U}(\mathcal{X}) = \{U \in \mathcal{L}(\mathcal{X}) : UU^\dagger = U^\dagger U = \mathbb{1}_\mathcal{X}\}. \quad (2.27)$$

One important quality of unitary operators is that they preserve the Euclidean norm of vectors, in other words for any $|x\rangle \in \mathcal{X}$ holds:

$$\|U|x\rangle\| = \||x\rangle\|. \quad (2.28)$$

6. Density operators. A density operator (or density matrix) is a positive semidefinite operator with a unit trace. The set of all density operators is denoted as:

$$\mathcal{D}(\mathcal{X}) = \{\rho \in \text{Pos}(\mathcal{X}) : \text{Tr}(\rho) = 1\}. \quad (2.29)$$

7. Diagonal operators. An operator $D \in \mathcal{L}(\mathcal{X})$ is diagonal if its matrix representation has off-diagonal elements equal to zero. In other words let $|x\rangle = (x_0, \dots, x_n)^T$, then the matrix elements of this operator are given by $D_{ij} = \text{diag}(|x\rangle)_{ij} = x_i \delta_{ij}$, where δ_{ij} is the Kronecker delta. This also defines another mapping from vectors to matrices.

2.2.6 Operator decompositions

Theorem 2.1. (*Spectral theorem*) Let $\mathcal{X}$ be a complex Euclidean space and $X \in \mathcal{L}(\mathcal{X})$ be a normal operator. There exist distinct complex numbers $\lambda_0, \dots, \lambda_m$ and nonzero orthogonal projections $\Pi_0, \dots, \Pi_m$ satisfying $\Pi_0 + \dots + \Pi_m = \mathbb{1}$ and $\Pi_i \Pi_j = 0$ for $i \neq j$, such that

$$X = \sum_{i=0}^m \lambda_i \Pi_i. \quad (2.30)$$

The scalars $\lambda_0, \dots, \lambda_m$ and projection operators $\Pi_0, \dots, \Pi_m$ are unique (up to permutation). Specifically, each λ_k is an eigenvalue of X with multiplicity equal to the rank of Π_k , and Π_k is the orthogonal projection onto the eigenspace corresponding to λ_k .

It is also possible to rewrite Theorem 2.1 to a slightly different form, which is more useful for proofs or other uses.

Corollary 2.2 (Spectral decomposition for normal operators). Let $\mathcal{X}$ be a complex Euclidean space of dimension $n + 1$. Let $X \in \mathcal{L}(\mathcal{X})$ be a normal operator with spectrum

$$\lambda(X) = \{\lambda_0, \dots, \lambda_n\}. \quad (2.31)$$

There exists an orthonormal basis $\{|x_0\rangle, \dots, |x_n\rangle\} \subset \mathcal{X}$ such that

$$X = \sum_{k=0}^n \lambda_k |x_k\rangle\langle x_k|. \quad (2.32)$$

Equivalently, there exists a unitary operator $U \in \mathcal{U}(\mathcal{X})$ and a diagonal operator $D \in \mathcal{L}(\mathcal{X})$ such that

$$X = UDU^\dagger, \quad (2.33)$$

where $D = \text{diag}((\lambda_0, \dots, \lambda_n)^T)$ is the diagonal matrix of eigenvalues λ_k , and the columns of U are the corresponding eigenvectors $|x_k\rangle$.

We mention one more theorem, which can be seen as a generalization of Theorem 2.1, as we omit a requirement for the operator being normal.

Theorem 2.3 (Singular value decomposition (SVD)). *Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces. Let $X \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ be a nonzero operator and $\text{rank}(X) = r + 1$. Then, there exist two sets of orthonormal vectors $\{|x_0\rangle, \dots, |x_r\rangle\} \subset \mathcal{X}$ and $\{|y_0\rangle, \dots, |y_r\rangle\} \subset \mathcal{Y}$ along with positive real numbers $\sigma_0, \dots, \sigma_r > 0$ such that*

$$X = \sum_{i=0}^r \sigma_i |y_i\rangle\langle x_i|. \quad (2.34)$$

By $\sigma(X)$ we denote the set containing all σ_i of the operator $X \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ called the singular values as described in Eq. (2.34). Moreover we refer to $|y_i\rangle$ as a left singular vector and $\langle x_i|$ as a right singular vector of σ_i .

Corollary 2.4. *Let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ be a nonzero operator of rank $r + 1$, and let $\mathcal{W}$ be a complex Euclidean space for which $\dim(\mathcal{W}) = r + 1$. There exists a diagonal operator $D \in \text{Pos}(\mathcal{W})$ and isometries $U \in \mathcal{L}(\mathcal{W}, \mathcal{Y})$ and $V \in \mathcal{L}(\mathcal{W}, \mathcal{X})$ satisfying $U^\dagger U = V^\dagger V = \mathbb{1}_{\mathcal{W}}$ such that*

$$A = UDV^\dagger. \quad (2.35)$$

As per Corollary 2.2, the diagonal elements of D may be chosen from the $\sigma(A)$.

Theorem 2.5 (Jordan-Hahn decomposition). *Let us consider $H \in \text{Herm}(\mathcal{X})$ of dimension $\dim(\mathcal{X}) = n + 1$ of the form*

$$H = \sum_{i=0}^n \lambda_i \Pi_i \quad (2.36)$$

where $\Pi_i \in \text{Proj}(\mathcal{X})$ are nonzero projection operators such that $\sum_{i=0}^n \Pi_i = \mathbb{1}_{\mathcal{X}}$ and $\lambda_i \in \mathbb{R}$, $i \in \{0, \dots, n\}$. Let us define the operators $P, Q \in \text{Pos}(\mathcal{X})$ satisfying $PQ = 0$ by

$$P = \sum_{i=0}^n \max\{\lambda_i, 0\} \Pi_i; \quad Q = \sum_{i=0}^n \max\{-\lambda_i, 0\} \Pi_i. \quad (2.37)$$

Then, the operator H can be expressed as

$$H = P - Q. \quad (2.38)$$

The operators P and Q are uniquely defined for a given operator H . The expression given by Eq. (2.38) is therefore referred to as the Jordan-Hahn decomposition of H .

2.2.7 Numerical range and arc function

In the subsection, we define the notion of a numerical range and arc function. Let $X \in L(\mathcal{X})$. The numerical range of X is defined as the set:

$$W(X) = \{\langle x|X|x\rangle : |x\rangle \in \mathcal{X}, \langle x|x\rangle = 1\}. \quad (2.39)$$

We now discuss a few basic properties of the numerical range, beginning with the fact that it contains the eigenvalues.

Theorem 2.6. *Let $A \in L(\mathcal{X})$ be a linear operator on a complex Euclidean space $\mathcal{X}$. The numerical range of A contains the eigenvalues of A . That is, if λ is an eigenvalue of A , then $\lambda \in W(A)$.*

Proof. Let $|x\rangle \in \mathcal{X}$ be a non-zero eigenvector of $A \in L(\mathcal{X})$ and let λ be the corresponding eigenvalue. We define a normalized eigenvector $|y\rangle$ as:

$$|y\rangle = \frac{|x\rangle}{\| |x\rangle \|}. \quad (2.40)$$

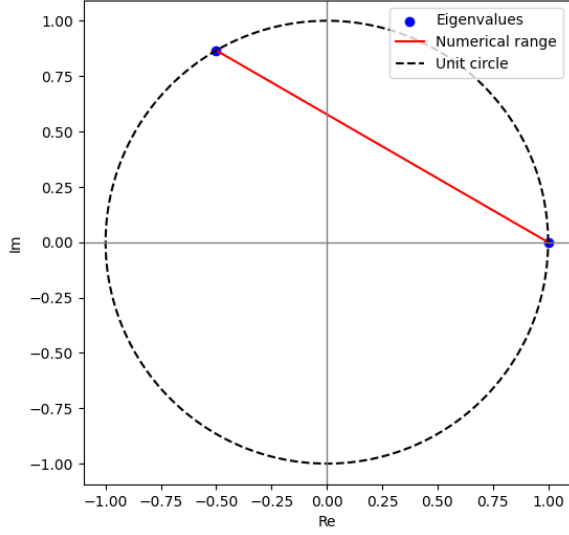
From the definition of the numerical range, the property of the induced Euclidean norm, and the eigenvalue equation $A|x\rangle = \lambda|x\rangle$, we find:

$$\langle y|A|y\rangle = \frac{\langle x|A|x\rangle}{\| |x\rangle \|^2} = \frac{\langle x|(\lambda|x\rangle)}{\langle x|x\rangle} = \lambda \frac{\langle x|x\rangle}{\langle x|x\rangle} = \lambda. \quad (2.41)$$

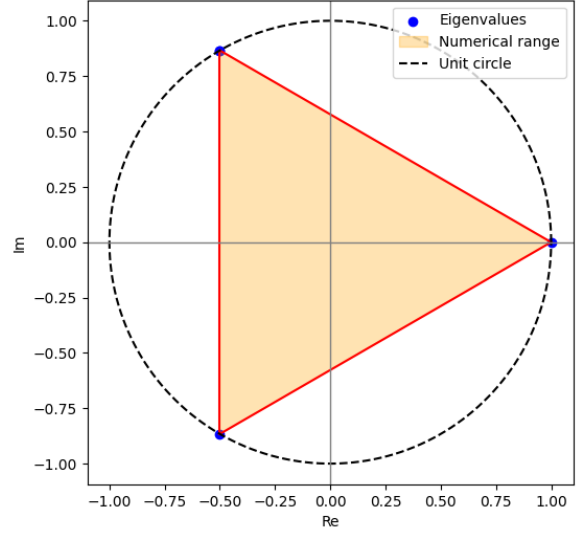
Since $|y\rangle$ is a unit vector, $\langle y|A|y\rangle$ is by definition an element of $W(A)$. Therefore, $\lambda \in W(A)$, which completes the proof. $\square$

The Hausdorff-Toeplitz [15, 14] theorem states that the numerical range is a convex set. Moreover, if the operator is normal, the numerical range is the convex hull of its eigenvalues. In the context of a unitary operator $U \in U(\mathcal{X})$, where the spectrum lies on the unit circle, the numerical range is linked to the angular spread of the eigenvalues. To formalize this, we introduce the *arc function* $\theta(U)$, defined as the minimal angle of an arc on the unit circle necessary to enclose the spectrum $\lambda(U)$. Formally, we define:

$$\theta(U) := \min \left\{ \Delta \in [0, 2\pi) : \exists \alpha \in [0, 2\pi) \text{ such that } \lambda(U) \subset \{e^{i\phi} : \phi \in [\alpha, \alpha + \Delta]\} \right\}. \quad (2.42)$$

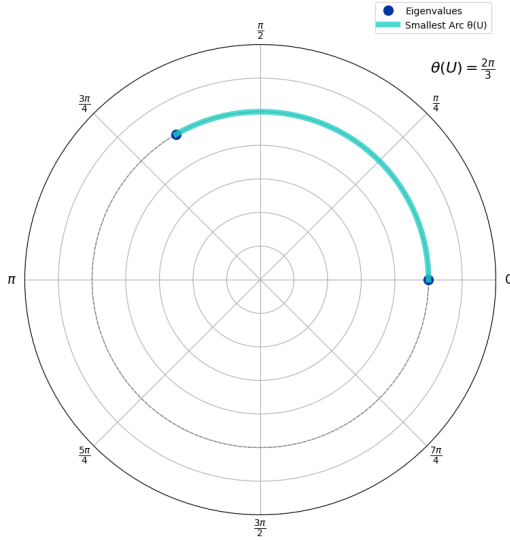


(a) Numerical range of operator U .

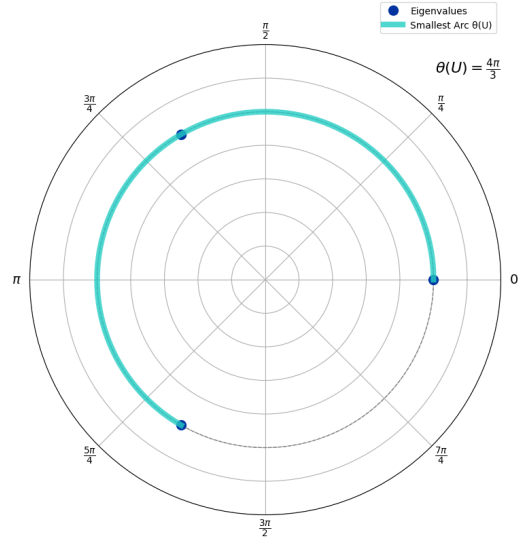


(b) Numerical range of operator $U \otimes U$.

Figure 2.1: Comparison of eigenvalues and numerical ranges in the complex plane for U in (a) and $U \otimes U$ in (b) where $U = \text{diag}([1, e^{\frac{2i\pi}{3}}]^T)$. In both cases the numerical range is a convex hull of eigenvalues. In (a) such hull creates line, in (b) the hull creates triangle.



(a) Arc function of operator U .



(b) Arc function of operator $U \otimes U$.

Figure 2.2: Comparison of the smallest angle containing all eigenvalues of two unitaries U in (a) and $U \otimes U$ in (b) where $U = \text{diag}([1, e^{\frac{2i\pi}{3}}]^T)$.

Observe the geometry of the numerical range, as illustrated in Fig. 2.1 and the arc function illustration in Fig. 2.2. A critical condition in our analysis is the containment of the origin within

the numerical range, which corresponds precisely to the condition $\theta(U) \geq \pi$. These conditions are met in the second example, given by $\text{diag}([1, e^{\frac{2i\pi}{3}}, e^{\frac{2i\pi}{3}}, e^{\frac{4i\pi}{3}}]T)$.

2.3 Linear maps on square operators

In this section, we define linear mappings between operators as well as some properties of them. Let $L(\mathcal{X})$ and $L(\mathcal{Y})$ be complex Euclidean spaces. A linear map Φ of the form:

$$\Phi : L(\mathcal{X}) \rightarrow L(\mathcal{Y}). \quad (2.43)$$

is referred to as a *superoperator* and plays a fundamental role in quantum information theory. By $T(\mathcal{X}, \mathcal{Y})$ we denote the set of all such mappings. Explicitly, the linear structure of the given $\Psi, \Phi \in T(\mathcal{X}, \mathcal{Y})$ is:

$$(\Psi + \Phi)(X) = \Psi(X) + \Phi(X) \quad (2.44)$$

for all $X \in L(\mathcal{X})$. Similarly, scalar multiplication for a map $\Phi \in T(\mathcal{X}, \mathcal{Y})$ and a complex scalar $\alpha \in \mathbb{C}$ is:

$$(\alpha\Phi)(X) = \alpha\Phi(X) \quad (2.45)$$

for all $X \in L(\mathcal{X})$. With these properties, $T(\mathcal{X}, \mathcal{Y})$ is a complex linear space.

We define the tensor product between the superoperators $\Psi \in T(\mathcal{W}, \mathcal{X})$ and $\Phi \in T(\mathcal{Y}, \mathcal{Z})$ as the unique linear mapping:

$$(\Psi \otimes \Phi) \in T(\mathcal{W} \otimes \mathcal{Y}, \mathcal{X} \otimes \mathcal{Z}), \quad (2.46)$$

which satisfies the following equation:

$$(\Psi \otimes \Phi)(A \otimes B) = \Psi(A) \otimes \Phi(B), \quad (2.47)$$

for all $A \in L(\mathcal{W})$ and $B \in L(\mathcal{Y})$. Similarly, as for vectors and operators, the notation $\Phi^{\otimes n}$ denotes the n -fold tensor product of a map Φ with itself. In a similar way to square linear operators, we will use $T(\mathcal{X})$ instead of $T(\mathcal{X}, \mathcal{X})$.

One can see a trace operator $\text{Tr} : L(\mathcal{X}) \rightarrow \mathbb{C}$ is an example of a linear superoperator (by considering $L(\mathbb{C}) = \mathbb{C}$). We also denote the identity superoperator as $\mathbb{1}_{L(\mathcal{X})}$ defining its action as $\mathbb{1}_{L(\mathcal{X})}(X) = X$ for all $X \in L(\mathcal{X})$.

The partial trace over $\mathcal{X}$, denoted as $\text{Tr}_{\mathcal{X}}$, is given by the superoperator $\text{Tr} \otimes \mathbb{1}_{L(\mathcal{Y})} \in T(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y})$. For product operator $X \otimes Y \in L(\mathcal{X} \otimes \mathcal{Y})$ it is defined as:

$$\text{Tr}_{\mathcal{X}}(X \otimes Y) = \text{Tr}(X)Y. \quad (2.48)$$

Let $\dim(\mathcal{X}) = n + 1$ and let $A \in \mathcal{L}(\mathcal{X} \otimes \mathcal{Y})$. One can compute the partial trace over $\mathcal{X}$ as:

$$\text{Tr}_{\mathcal{X}}(A) = \sum_{i=0}^n (\langle i| \otimes \mathbb{I}_{\mathcal{Y}}) A (|i\rangle \otimes \mathbb{I}_{\mathcal{Y}}). \quad (2.49)$$

In analogy, we can define $\text{Tr}_{\mathcal{Y}}(A)$.

We also define a partial transpose of linear operators. Let $\mathcal{X}$ and $\mathcal{Y}$ be complex Euclidean spaces of $\dim(\mathcal{X}) = n + 1$ and $\dim(\mathcal{Y}) = m + 1$. Let $X \in \mathcal{L}(\mathcal{X} \otimes \mathcal{Y})$ be an operator expressed in the computational basis as:

$$X = \sum_{i,k=0}^n \sum_{j,l=0}^m X_{ij,kl} (|i\rangle\langle k| \otimes |j\rangle\langle l|) \quad (2.50)$$

The partial transpose of X with respect to the first subsystem $\mathcal{X}$, denoted as $X^{\text{T}_{\mathcal{X}}}$, is defined by applying the transpose mapping only to the operator acting on $\mathcal{X}$:

$$X^{\text{T}_{\mathcal{X}}} = \sum_{i,k=0}^n \sum_{j,l=0}^m X_{ij,kl} (|i\rangle\langle k|)^{\text{T}} \otimes |j\rangle\langle l|. \quad (2.51)$$

Since $(|i\rangle\langle k|)^{\text{T}} = |k\rangle\langle i|$, this evaluates to:

$$X^{\text{T}_{\mathcal{X}}} = \sum_{i,k=0}^n \sum_{j,l=0}^m X_{ij,kl} (|k\rangle\langle i| \otimes |j\rangle\langle l|). \quad (2.52)$$

In analogy, we can define $X^{\text{T}_{\mathcal{Y}}}$.

2.3.1 Representations of linear maps on square operators

In this subsection, we present the representation of superoperators. The thesis primarily utilizes the Kraus and Choi formulations, which are detailed below.

Definition 2.7 (Kraus representation). Let $n \in \mathbb{N}$. Suppose that we are given two sets of operators $\{A_i\}_{i=0}^n$ and $\{B_i\}_{i=0}^n \subset \mathcal{L}(\mathcal{X}, \mathcal{Y})$. We define the superoperator $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ as:

$$\Phi(X) = \sum_{i=0}^n A_i X B_i^{\dagger}, \quad (2.53)$$

where $X \in \mathcal{L}(\mathcal{X})$. The existence of these sets demonstrates a Kraus representation of given Φ . The operators $\{A_i\}_{i=0}^n$ and $\{B_i\}_{i=0}^n$ are called the Kraus operators of the map. It should be noted that for a given superoperator, the Kraus representation is not unique.

Definition 2.8 (Choi representation). Let $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ be a superoperator with $\dim(\mathcal{X}) = n + 1$. The Choi representation of Φ is the operator $J(\Phi) \in \mathcal{L}(\mathcal{Y} \otimes \mathcal{X})$ defined as:

$$J(\Phi) = \sum_{i,j=0}^n \Phi(|i\rangle\langle j|) \otimes |i\rangle\langle j|. \quad (2.54)$$

Note that the Kraus representation explicitly characterizes the transformation of an operator X under the action of a superoperator. In contrast, the Choi representation encodes the map's properties within its associated Choi matrix. While the Choi representation establishes a bijection between completely positive linear maps and positive semidefinite matrices, the Kraus representation is not uniquely defined for a given map. We now present the relationship between these two representations.

Proposition 2.9. Let $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ be a linear map with $\dim(\mathcal{X}) = n + 1$. For any two collections of operators $\{A_i\}_{i=0}^m, \{B_i\}_{i=0}^m \subset \mathcal{L}(\mathcal{X}, \mathcal{Y})$, the following statements are equivalent:

1. The map Φ is defined by the operator-sum representation:

$$\Phi(X) = \sum_{i=0}^m A_i X B_i^\dagger, \quad \forall X \in \mathcal{L}(\mathcal{X}). \quad (2.55)$$

2. The map Φ is defined by the Choi representation:

$$J(\Phi) = \sum_{i=0}^m |A_i\rangle\langle B_i|. \quad (2.56)$$

Proof. By the definitions of the Choi representation and the operator-sum representation of Φ , we have:

$$J(\Phi) = \sum_{j,k=0}^n \Phi(|j\rangle\langle k|) \otimes |j\rangle\langle k| = \sum_{j,k=0}^n \left(\sum_{i=0}^m A_i |j\rangle\langle k| B_i^\dagger \right) \otimes |j\rangle\langle k|. \quad (2.57)$$

Adjusting the sums and applying the mixed-product property of the tensor product, specifically $(AC \otimes BD) = (A \otimes B)(C \otimes D)$, we can write:

$$\begin{aligned} J(\Phi) &= \sum_{i=0}^m \sum_{j,k=0}^n (A_i |j\rangle\langle k| B_i^\dagger) \otimes (|j\rangle\langle k|) \\ &= \sum_{i=0}^m \sum_{j,k=0}^n (A_i |j\rangle \otimes |j\rangle) (\langle k| B_i^\dagger \otimes \langle k|) \\ &= \sum_{i=0}^m \left(\sum_{j=0}^n A_i |j\rangle \otimes |j\rangle \right) \left(\sum_{k=0}^n \langle k| B_i^\dagger \otimes \langle k| \right). \end{aligned} \quad (2.58)$$

The term $\sum_{j=0}^n A_i |j\rangle \otimes |j\rangle$ corresponds exactly to the vectorization $|A_i\rangle\rangle$ of the operator A_i . Similarly, the second term is the Hermitian adjoint of $|B_i\rangle\rangle$, because:

$$\left(\sum_{k=0}^n B_i |k\rangle \otimes |k\rangle \right)^\dagger = \sum_{k=0}^n \langle k| B_i^\dagger \otimes \langle k| = \langle\langle B_i|. \quad (2.59)$$

Substituting these expressions simplifies the equation to:

$$J(\Phi) = \sum_{i=0}^m |A_i\rangle\rangle \langle\langle B_i|. \quad (2.60)$$

This confirms that the Choi matrix is the sum of the outer products of the vectorized operators. $\square$

2.3.2 Classes of linear maps on square operators

As with linear operators, we focus on several classes of superoperators defined below.

1. **Hermitian-preserving maps.** A linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ is called Hermitian-preserving if

$$\Phi(H) \in \text{Herm}(\mathcal{Y}) \quad (2.61)$$

for every $H \in \text{Herm}(\mathcal{X})$.

2. **Positive maps.** We define a linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ as a positive if for every positive semidefinite operator $P \in \text{Pos}(\mathcal{X})$ holds:

$$\Phi(P) \in \text{Pos}(\mathcal{Y}). \quad (2.62)$$

3. **Completely positive maps.** A linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ is called a completely positive (CP) if the map:

$$\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{Z})} \quad (2.63)$$

is positive for any choice of a complex Euclidean space $\mathcal{Z}$. The set of all such maps is denoted by $\text{CP}(\mathcal{X}, \mathcal{Y})$, equivalent statement is the map Φ is CP if:

$$J(\Phi) \in \text{Pos}(\mathcal{Y} \otimes \mathcal{X}). \quad (2.64)$$

4. **Trace-preserving maps.** A map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ is said to be trace-preserving (TP) if:

$$\text{Tr}(\Phi(X)) = \text{Tr}(X). \quad (2.65)$$

for all $X \in \mathcal{L}(\mathcal{X})$.

2.3.3 Superoperator norms

For a linear map $\Phi \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$, we define the induced trace norm $\|\Phi\|_1$ as:

$$\|\Phi\|_1 = \max \{ \|\Phi(X)\|_1 : X \in \mathcal{L}(\mathcal{X}), \|X\|_1 \leq 1 \}. \quad (2.66)$$

Finally, we introduce the diamond norm (also known as the completely bounded trace norm) on the space $\mathcal{T}(\mathcal{X}, \mathcal{Y})$. For Φ , the diamond norm is defined as:

$$\|\Phi\|_\diamond = \|\Phi \otimes \mathbb{1}_{\mathcal{L}(\mathcal{X})}\|_1. \quad (2.67)$$

This norm is used to quantify the distance $\|\Psi - \Phi\|_\diamond$ between two quantum channels Ψ and Φ . While calculating this distance is computationally expensive, it can be evaluated using Semidefinite Programming (SDP). This was famously shown by Watrous [35, 36].

Chapter 3

Fundamentals of quantum information theory

This chapter is dedicated to the establishment of a theoretical view of quantum information theory. To achieve this, we directly utilize and apply the fundamental mathematical definitions and terminology introduced in Chapter 2. The chapter is based on [28]. We begin with definition of a quantum state, which is followed by its transformation and measurement. A bridge connecting quantum computing and quantum information theory is established at the end of the chapter.

3.1 Quantum states

We associate each single quantum system X with the complex Euclidean space $\mathcal{X}$. A compound quantum system Z , consisting of subsystems $Y_0, \dots, Y_k$ (which may themselves be either single or compound), is associated with a complex Euclidean space $\mathcal{Z}$. This space is defined by the tensor product of the individual Euclidean spaces $\mathcal{Y}_0, \dots, \mathcal{Y}_k$ representing each subsystem $Y_0, \dots, Y_k$ given by:

$$\mathcal{Z} = \bigotimes_{i=0}^k \mathcal{Y}_i = \mathcal{Y}_0 \otimes \mathcal{Y}_1 \otimes \dots \otimes \mathcal{Y}_k. \quad (3.1)$$

Definition 3.1. Let $\mathcal{X}$ be a complex Euclidean space associated with a quantum system X . The quantum state of the system is described by a density operator $\rho \in \mathcal{D}(\mathcal{X})$.

Now, we are ready to present different classes of quantum states.

1. **Pure state.** A quantum state $\rho \in \mathcal{D}(\mathcal{X})$ is said to be a pure state if $\text{rank}(\rho) = 1$. Equivalently, ρ is a pure state if there exists a unit vector $|x\rangle \in \mathcal{X}$ such that $\rho = |x\rangle\langle x|$.

In this case, we identify $|x\rangle$ with ρ and treat them equivalently (while we may address $|x\rangle$ as a quantum state, it should be understood that ρ is meant and vice versa).

2. **Mixed state.** Suppose we have a finite set of pure states $\{\rho_i\}_{i=0}^j \subset D(\mathcal{X})$, $j \in \mathbb{N}$, and positive scalars $\{\alpha_i\}_{i=0}^j$ such that $\sum_{i=0}^j \alpha_i = 1$. Then, we define a mixed quantum state as follows:

$$\rho = \sum_{i=0}^j \alpha_i \rho_i. \quad (3.2)$$

3. **Product state.** Let $\mathcal{Z} = \mathcal{X} \otimes \mathcal{Y}$ be a complex Euclidean space. The quantum state $\rho \in D(\mathcal{Z})$ is a product state if it is of the form:

$$\rho = \tau \otimes \sigma, \quad (3.3)$$

where $\tau \in D(\mathcal{X})$, $\sigma \in D(\mathcal{Y})$.

4. **Separable state.** Let $j \in \mathbb{N}$. Suppose that we have a finite set of positive scalars $\{\alpha_i\}_{i=0}^{j-1}$ such that $\sum_{i=0}^{j-1} \alpha_i = 1$ and two collections of quantum states $\{\rho_0, \dots, \rho_{j-1}\} \subset D(\mathcal{X})$, $\{\sigma_0, \dots, \sigma_{j-1}\} \subset D(\mathcal{Y})$. The quantum state $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$ is separable if it can be written as:

$$\rho = \sum_{i=0}^{j-1} \alpha_i (\rho_i \otimes \sigma_i). \quad (3.4)$$

5. **Entangled state.** If the quantum state $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$ is not separable, it is called entangled.

To illustrate the definitions provided above, we will examine the following quantum states and determine their respective classifications.

Example 3.2. Let us define the following quantum states:

$$\chi = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \quad \xi = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}. \quad (3.5)$$

First, see that the rows in χ are linearly dependent, which results in $\text{rank}(\chi) = 1$, therefore the state is pure. Second, we compute the eigenvalues of ξ , which is $1/2$ with multiplicity of two. One can observe the respective eigenvectors are $|0\rangle$ and $|1\rangle$, which are normalized, so $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ are pure states. Therefore, by definition, ξ is a mixed state.

Example 3.3. For the next example we define $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. Now, suppose $\mathcal{Z} = \mathcal{X} \otimes \mathcal{Y}$ and the quantum state $\theta \in D(\mathcal{Z})$ given by:

$$\theta = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}. \quad (3.6)$$

Now, let us calculate the partial traces. We obtain $\text{Tr}_{\mathcal{X}}(\theta) = |+\rangle\langle+|$ and $\text{Tr}_{\mathcal{Y}}(\theta) = |1\rangle\langle 1|$. Finally, we have $\theta = |1\rangle\langle 1| \otimes |+\rangle\langle+|$, which means θ is a product state.

Furthermore, consider the following quantum states:

$$\sigma = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad \omega = \frac{1}{4} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad (3.7)$$

We begin by computing the eigenvalues of σ . These are 1 (with multiplicity one) and 0 (with multiplicity three). Since σ is a pure state, separability is equivalent to being a product state. Assuming that σ is a pure product state, it must satisfy $\sigma = \text{Tr}_{\mathcal{Y}}(\sigma) \otimes \text{Tr}_{\mathcal{X}}(\sigma)$. However, calculating the partial traces and taking their tensor product yields $\frac{1}{4} \text{diag}([1, 1, 1, 1])$, which does not equal σ . This contradiction shows that σ is not a product state and therefore must be entangled.

Lastly, we rewrite the state ω as a block matrix:

$$\omega = \frac{1}{4} \begin{bmatrix} \mathbb{1}_{\mathcal{X}} & X \\ X & \mathbb{1}_{\mathcal{X}} \end{bmatrix}, \quad (3.8)$$

for the operators:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbb{1}_{\mathcal{X}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.9)$$

The state ω can be rewritten in the form:

$$\omega = \frac{1}{4} (|0\rangle\langle 0| \otimes \mathbb{1}_{\mathcal{X}} + |1\rangle\langle 1| \otimes \mathbb{1}_{\mathcal{X}} + |0\rangle\langle 1| \otimes X + |1\rangle\langle 0| \otimes X). \quad (3.10)$$

Factoring out the identity and X , we get:

$$\begin{aligned} \omega &= \frac{1}{4} ((|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes \mathbb{1}_{\mathcal{X}} + (|0\rangle\langle 1| + |1\rangle\langle 0|) \otimes X) \\ &= \frac{1}{4} (\mathbb{1}_{\mathcal{X}} \otimes \mathbb{1}_{\mathcal{X}} + X \otimes X). \end{aligned} \quad (3.11)$$

Finally, observe $X = |+\rangle\langle+| - |-\rangle\langle-| = \Pi_+ - \Pi_-$ and $\mathbb{1}_{\mathcal{X}} = |+\rangle\langle+| + |-\rangle\langle-| = \Pi_+ + \Pi_-$. So ω can

be expressed as:

$$\begin{aligned}
\omega &= \frac{1}{4} ((\Pi_+ + \Pi_-) \otimes (\Pi_+ + \Pi_-) + (\Pi_+ - \Pi_-) \otimes (\Pi_+ - \Pi_-)) \\
&= \frac{1}{4} (\Pi_+ \otimes \Pi_+ + \Pi_+ \otimes \Pi_- + \Pi_- \otimes \Pi_+ + \Pi_- \otimes \Pi_- \\
&\quad + \Pi_+ \otimes \Pi_+ - \Pi_+ \otimes \Pi_- - \Pi_- \otimes \Pi_+ + \Pi_- \otimes \Pi_-) \\
&= \frac{1}{4} (2\Pi_+ \otimes \Pi_+ + 2\Pi_- \otimes \Pi_-) \\
&= \frac{1}{2} (\Pi_+ \otimes \Pi_+ + \Pi_- \otimes \Pi_-),
\end{aligned} \tag{3.12}$$

which indeed shows that the state ω is separable.

From the above analysis, we were able to decide the entanglement of σ and ω based on their specific algebraic properties. These findings are formally consistent with the positive partial transpose (PPT) or Peres-Horodecki criterion [37, 38]. For a state ρ , this criterion evaluates the positivity of its partial transpose ρ^{T_Y} (see Eq 2.52). While generally only a necessary condition for separability in arbitrary dimensions, it is both necessary and sufficient for low-dimensional systems like $\mathbb{C}^2 \otimes \mathbb{C}^2$ and $\mathbb{C}^2 \otimes \mathbb{C}^3$. Because our states $\sigma, \omega \in D(\mathbb{C}^2 \otimes \mathbb{C}^2)$, this sufficiency applies. Specifically, σ^{T_Y} yields a negative eigenvalue $(-1/2)$, confirming σ is entangled. Conversely, ω^{T_Y} is positive semidefinite, analytically verifying the separability demonstrated via decomposition.

3.2 Transformations of quantum states

In this section, we introduce quantum channels and the Born rule together with examples.

Definition 3.4. Let $\Phi \in T(\mathcal{X}, \mathcal{Y})$. The linear map Φ is a quantum channel if it is completely positive and trace preserving (CPTP). By $C(\mathcal{X}, \mathcal{Y})$ we denote the set of all such quantum channels.

These two constraints are physically motivated and essential for the mathematical consistency of the theory. The requirement of CP ensures that the output of the map remains a positive semidefinite operator, even when the channel is applied to a subsystem of a larger, entangled state. Simultaneously, the TP property guarantees the conservation of the trace. Collectively, these conditions imply that if $\Phi \in C(\mathcal{X}, \mathcal{Y})$, then the Φ maps the set of density operators $D(\mathcal{X})$ directly into the set of density operators $D(\mathcal{Y})$.

Recall the CP can be checked using Choi matrix, as it was provided as equivalent definition when discussing classes of channels. For proof see Theorem 2.22 in [28].

Example 3.5. Let $\mathcal{X}$ be a complex Euclidean space and let $U \in U(\mathcal{X})$ be a unitary operator. The linear map $\Phi_U : L(\mathcal{X}) \rightarrow L(\mathcal{X})$ defined by $\Phi_U(X) = UXU^\dagger$ is a CPTP map.

Proof. First, we check that Φ_U is linear. Let $A, B \in \mathcal{L}(\mathcal{X})$ and $\beta_1, \beta_2 \in \mathbb{C}$ be arbitrary complex scalars. By the distributive property of matrix multiplication:

$$\begin{aligned}\Phi_U(\beta_1 A + \beta_2 B) &= U(\beta_1 A + \beta_2 B)U^\dagger \\ &= \beta_1 UAU^\dagger + \beta_2 UBU^\dagger \\ &= \beta_1 \Phi_U(A) + \beta_2 \Phi_U(B).\end{aligned}\tag{3.13}$$

This shows the map is linear.

Then, we verify that Φ_U is trace preserving. For any $X \in \mathcal{L}(\mathcal{X})$, using the cyclic property of the trace and the unitary condition $U^\dagger U = \mathbb{1}_{\mathcal{X}}$, we obtain:

$$\text{Tr}(\Phi_U(X)) = \text{Tr}(UXU^\dagger) = \text{Tr}(U^\dagger UX) = \text{Tr}(\mathbb{1}_{\mathcal{X}}X) = \text{Tr}(X).\tag{3.14}$$

Thus, the map preserves the trace. Lastly, we verify that Φ_U is completely positive by examining its Choi representation, which is constructed using Proposition 2.9. The Choi matrix is then:

$$J(\Phi_U) = |U\rangle\langle U|.\tag{3.15}$$

For showing the map is CP we use the equivalent definition of CP map as shown in Eq. (2.64). We prove that $J(\Phi_U)$ is positive semidefinite. Let $|x\rangle \in \mathcal{X} \otimes \mathcal{X}$ be an arbitrary vector. We define a scalar $\alpha = \langle\langle U|x\rangle\rangle$, which implies $\bar{\alpha} = (\langle\langle U|x\rangle\rangle)^\dagger = \langle\langle x|U\rangle\rangle$. For any $|x\rangle \in \mathcal{X} \otimes \mathcal{X}$ we have:

$$\langle\langle x|U\rangle\rangle\langle\langle U|x\rangle\rangle = \alpha\bar{\alpha} = |\alpha|^2 \geq 0.\tag{3.16}$$

Since Φ_U is linear, trace preserving and completely positive, it is a correctly defined quantum channel.

□

We have established that a unitary channel $\Phi(\rho) = U\rho U^\dagger$ defined by some unitary $U \in \mathcal{U}(\mathcal{X})$ is a quantum channel. If we restrict ourselves to pure states only given by $\rho = |x\rangle\langle x|$ for a unit vector $|x\rangle \in \mathcal{X}$. Then, the action of the channel can be written as:

$$\Phi(\rho) = U|x\rangle\langle x|U^\dagger = (U|x\rangle)(U|x\rangle)^\dagger = |y\rangle\langle y|,\tag{3.17}$$

where we have defined the transformed vector $|y\rangle = U|x\rangle$. This relationship implies that for pure states, the evolution of the density operator is fully determined by the evolution of the vector. Therefore, it is sufficient to work directly with the vectors $|x\rangle$ and $|y\rangle$, as the density operator can be reconstructed using the outer product.

3.2.1 Quantum measurement and Born rule

In this subsection, we define the formal process of quantum measurement. To obtain information from a quantum system, a measurement must be performed. Such a procedure is inherently probabilistic, meaning that for a given density operator representing the state of a quantum system, the outcomes of repeated measurements may differ. This probabilistic nature is described by the Born rule [6].

Definition 3.6 (Positive Operator-Valued Measurement). Let $\mathcal{X}$ be a complex Euclidean space and $n \in \mathbb{N}$. A Positive Operator-Valued Measurement (POVM) is a finite set of positive semidefinite operators $\{E_0, \dots, E_n\} \subset \text{Pos}(\mathcal{X})$, referred to as effects, which satisfy the following relation:

$$\sum_{i=0}^n E_i = \mathbb{1}_{\mathcal{X}}, \quad (3.18)$$

where $\mathbb{1}_{\mathcal{X}} \in \text{L}(\mathcal{X})$ is the identity operator. A POVM provides the most general description of a quantum measurement, where the effects are not required to be mutually orthogonal or projective.

Definition 3.7 (The Born rule). The Born rule is the fundamental link between the mathematical formalism of quantum mechanics and physical observations. Consider a quantum state described by a density operator $\rho \in \text{D}(\mathcal{X})$ and a measurement described by $\{E_i\}_{i=0}^n$. The Born rule states that the probability $P(i)$ of obtaining the outcome (sometimes referred to as bitstring) i is given by the trace of the product of the corresponding effect and the state:

$$P(i) = \text{Tr}(E_i \rho). \quad (3.19)$$

Due to the definition of POVM (Eq. (3.18)), we ensure that the probabilities are normalized, $\sum_i P(i) = 1$, and the positive semidefiniteness of E_i and ρ ensures $P(i) \geq 0$.

In practice, we favor von Neumann measurements because these are currently the only projective measurements currently implementable on quantum architectures.

Definition 3.8 (Von Neumann measurement). Let $n \in \mathbb{N}$. Let $\mathcal{X}$ be a complex Euclidean space of $\dim(\mathcal{X}) = n + 1$. A POVM is called a von Neumann measurement if the effects E_i are rank-one orthogonal projectors $\Pi_i \in \text{Proj}(\mathcal{X})$. Such a measurement is associated with an orthonormal basis $\{U|i\rangle\}_{i=0}^n$ of $\mathcal{X}$, where each effect is defined as:

$$\Pi_i = U|i\rangle\langle i|U^\dagger. \quad (3.20)$$

The vector $U|i\rangle$ corresponds to the i -th column of a unitary matrix U , representing the transformation from the computational basis. These effects satisfy the following properties:

$$\Pi_i \Pi_j = \delta_{ij} \Pi_i, \quad \Pi_i^\dagger = \Pi_i, \quad \text{and} \quad \sum_{i=0}^n \Pi_i = \mathbb{1}_{\mathcal{X}}. \quad (3.21)$$

To illustrate concepts presented above, we will consider a few examples for $\mathcal{X} = \mathbb{C}^2$.

Example 3.9. General POVMs can have more effects than the dimension of the system. Consider POVM, which consists of three effects proportional to projectors onto vectors separated by $\frac{2\pi}{3}$. Let $|\psi_k\rangle = \cos(\theta_k)|0\rangle + \sin(\theta_k)|1\rangle$ with $\theta_k = \frac{2\pi k}{3}$ for $k = 0, 1, 2$. We define the effects as:

$$E_k = \frac{2}{3} |\psi_k\rangle\langle\psi_k|. \quad (3.22)$$

Observe that the effects are rank-one operators proportional to projectors, but they are not orthogonal. Evaluating for every possible k :

$$E_0 = \frac{2}{3} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad (3.23)$$

$$E_1 = \frac{2}{3} \begin{bmatrix} \frac{1}{4} & -\frac{\sqrt{3}}{4} \\ -\frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix}, \quad (3.24)$$

$$E_2 = \frac{2}{3} \begin{bmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} \\ \frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix}. \quad (3.25)$$

Summing these operators confirms the completeness relation: $\sum_{k=0}^2 E_k = \mathbb{1}_{\mathcal{X}}$. To illustrate the measurement outcomes, let the system be prepared in the state $\rho = |0\rangle\langle 0|$. The probabilities $P(k) = \text{Tr}(E_k \rho)$ for the three possible outcomes are:

$$P(0) = \text{Tr}(E_0 |0\rangle\langle 0|) = \frac{2}{3}, \quad (3.26)$$

$$P(1) = \text{Tr}(E_1 |0\rangle\langle 0|) = \frac{2}{3} \cdot \frac{1}{4} = \frac{1}{6}, \quad (3.27)$$

$$P(2) = \text{Tr}(E_2 |0\rangle\langle 0|) = \frac{2}{3} \cdot \frac{1}{4} = \frac{1}{6}. \quad (3.28)$$

Note that $\sum_k P(k) = \frac{2}{3} + \frac{1}{6} + \frac{1}{6} = 1$, as required.

Example 3.10. Consider the standard computational basis $\{|0\rangle, |1\rangle\}$. The von Neumann measurement corresponding to this basis consists of two projectors:

$$\Pi_0 = |0\rangle\langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Pi_1 = |1\rangle\langle 1| = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad (3.29)$$

Observe that $\Pi_0 + \Pi_1 = \mathbb{1}_{\mathcal{X}}$ and $\Pi_0\Pi_1 = 0$. For a general state ρ , the probability of obtaining outcome i after measurement is $P(i) = \text{Tr}(\Pi_i\rho) = \langle i|\rho|i\rangle$ for $i = 0, 1$.

Example 3.11. Similarly, consider the basis $\{|+\rangle, |-\rangle\}$. The von Neumann measurement corresponding with this basis is defined by the projectors:

$$\Pi_+ = |+\rangle\langle+| = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \Pi_- = |-\rangle\langle-| = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \quad (3.30)$$

Just as with the computational basis, these operators satisfy the completeness relation $\Pi_+ + \Pi_- = \mathbb{1}_{\mathcal{X}}$ and are orthogonal, $\Pi_+\Pi_- = 0$. The probability of obtaining the outcome $+$ for a state ρ is given by $\text{Tr}(\Pi_+\rho) = \langle +|\rho|+\rangle$.

3.3 Quantum computing

In this section, we present the concepts of quantum gates and quantum circuits, which are the fundamental components used for the implementation of quantum algorithms. Quantum gates are the building blocks of quantum circuits, analogous to logic gates in classical computers. Similarly to bits in classical computers, quantum ones use a quantum bit - qubit, which is represented by $|x\rangle \in \mathbb{C}^2$. A quantum gate acting on one qubit is then represented by a unitary transformation $U \in \text{U}(\mathbb{C}^2)$. The evolution of the state by the gate is then given by $|x\rangle \mapsto U|x\rangle$. In Table 3.1, we present a selection of the most frequently utilized single-qubit gates, detailing their circuit notation alongside their matrix representations. This collection includes the Pauli operators (X , Y , and Z) and the identity gate. Additionally, we include the Hadamard gate (H), usually used for creating superposition. Another point to note concerns the RZ gate. Observe that:

$$RZ(\phi) = \begin{bmatrix} e^{-\frac{i\phi}{2}} & 0 \\ 0 & e^{\frac{i\phi}{2}} \end{bmatrix} = e^{-\frac{i\phi}{2}} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} = e^{-\frac{i\phi}{2}} U(\phi). \quad (3.31)$$

This gate creates the same quantum channel as $\Phi_{U(\phi)}$ in Table 3.1. This can be seen in the following equation:

$$\Phi_{RZ(\phi)}(X) = e^{-\frac{i\phi}{2}} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} X e^{\frac{i\phi}{2}} \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\phi} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix} X \begin{bmatrix} 1 & 0 \\ 0 & e^{-i\phi} \end{bmatrix} = \Phi_{U(\phi)}(X). \quad (3.32)$$

Moreover, in Table 3.2 we introduce gates acting on two qubits. These are used to create entanglement, a phenomenon discussed later in the text. We show more known gates such as CNOT as well as some lesser-known gates, e.g. ECR, which is used in IBM devices based on the architecture EagleR3.

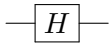
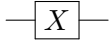
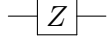
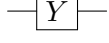
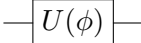
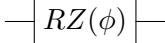
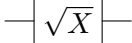
Name of the gate	Symbol	Matrix representation
Hadamard		$\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$
Pauli-X (X)		$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Pauli-Z		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Pauli-Y		$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$
Identity		$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Phase rotation		$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}$
Z-axis rotation		$\begin{bmatrix} e^{-\frac{i\phi}{2}} & 0 \\ 0 & e^{\frac{i\phi}{2}} \end{bmatrix}$
Square root of X		$\frac{1}{2} \begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}$

Table 3.1: Single-qubit quantum gates with their circuit symbols and corresponding unitary matrix representations.

It is a standard that during the start of the computation, each qubit is initialized in the quantum state $|0\rangle = (1, 0)^\top$. The overall quantum state is then determined by the number of qubits used (assume $N \in \mathbb{N}$), resulting in the quantum state $|0\rangle^{\otimes N}$. The computation ends with a quantum measurement, the gate of which is shown in Fig. 3.1.



Figure 3.1: Schematic representation of a quantum circuit measurement performed in the computational basis.

From a practical implementation standpoint on current quantum architectures, a von Neumann measurement is typically realized by applying the inverse unitary $U^\dagger$ followed by a standard basis measurement, as shown in Fig. 3.2.

The following example demonstrates the step-by-step evolution of a computation.

Example 3.12. Consider the quantum circuit illustrated in Fig. 3.3.

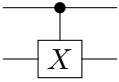
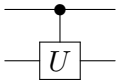
Name of the gate	Symbol	Matrix representation
CNOT		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
Controlled-U, CU		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & U_{1,1} & U_{1,2} \\ 0 & 0 & U_{2,1} & U_{2,2} \end{bmatrix}$
ECR		$\frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & 0 & i \\ 1 & 0 & -i & 0 \\ 0 & i & 0 & 1 \\ -i & 0 & 1 & 0 \end{bmatrix}$

Table 3.2: Two-qubit quantum gates with their circuit symbols and corresponding unitary matrix representations.

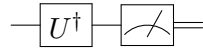


Figure 3.2: Schematic representation of a von Neumann measurement realized by applying the inverse unitary $U^\dagger$ followed by a measurement in the computational basis.

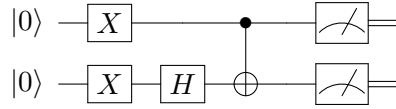


Figure 3.3: Two-qubit quantum circuit demonstrating the sequence of X , H , and CNOT gate operations ended by measurement in computational basis.

The circuit can be represented by the following equation:

$$|x\rangle = (\text{CNOT})(\mathbb{1}_{\mathcal{X}} \otimes H)(X \otimes X)(|0\rangle \otimes |0\rangle). \quad (3.33)$$

While the evolution is:

$$\begin{aligned}
|x\rangle &= (\text{CNOT})(\mathbb{1}_X \otimes H)(X \otimes X)(|0\rangle \otimes |0\rangle) \\
&= (\text{CNOT})(\mathbb{1}_X \otimes H)(|1\rangle \otimes |1\rangle) \\
&= (\text{CNOT})\left(|1\rangle \otimes \frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \\
&= (\text{CNOT})\left(\frac{|10\rangle - |11\rangle}{\sqrt{2}}\right) \\
&= \frac{|11\rangle - |10\rangle}{\sqrt{2}}.
\end{aligned} \tag{3.34}$$

The final density operator describing the system is then

$$\rho = |x\rangle\langle x| = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \tag{3.35}$$

which can be rewritten into a more convenient form for computing probabilities after measurement as:

$$\rho = \frac{1}{2}(|11\rangle\langle 11| - |11\rangle\langle 10| - |10\rangle\langle 11| + |10\rangle\langle 10|). \tag{3.36}$$

Since the computation concludes with a measurement in the computational basis, the corresponding measurement effects are given by the following operators:

$$\Pi_{00} = |00\rangle\langle 00|, \Pi_{01} = |01\rangle\langle 01|, \Pi_{10} = |10\rangle\langle 10|, \Pi_{11} = |11\rangle\langle 11|, \tag{3.37}$$

Now, we will compute the probabilities of obtaining outcomes ij where $i, j = 0, 1$. We obtain the outcomes with probabilities equal to:

- 00: $P(00) = \text{Tr}(\Pi_{00}\rho) = 0,$
- 01: $P(01) = \text{Tr}(\Pi_{01}\rho) = 0,$
- 10: $P(10) = \text{Tr}(\Pi_{10}\rho) = \frac{1}{2},$
- 11: $P(11) = \text{Tr}(\Pi_{11}\rho) = \frac{1}{2}.$

Chapter 4

Discrimination of unitary quantum channels

In this chapter, we introduce the discrimination of unitary channels and discuss the probability of correct discrimination. We consider a task involving the discrimination of quantum channels, illustrated with two individuals, Alice and Bob. Alice has access to several quantum channels. For simplicity, we restrict attention to two channels. She applies one of them but does not reveal which one she chose. First, let us assume that Alice can apply the chosen channel exactly once. Bob's goal is to determine which channel Alice used. To do this, he prepares an input state (possibly entangled), known as a discriminator, and then allows Alice to implement one of the quantum channels. Next, Bob performs his own sequence of quantum operations before making a measurement. From the Holevo-Helstrom theorem [8, 7], we know that the optimal probability of success of discrimination between two channels Φ_0, Φ_1 can be computed using the diamond norm as:

$$P_s = \frac{1}{2} + \frac{1}{4} \|\Phi_0 - \Phi_1\|_\diamond \quad (4.1)$$

The chapter is organized as follows. We first introduce single-shot and multiple-shot discrimination strategies, then detail the algorithmic construction of discriminators (states given as input to channels for discrimination), and conclude with measurement construction.

4.1 Single-shot discrimination of unitary channels

Recall the illustrative scenario from above. In the case of single-shot discrimination, we assume that Alice can only apply her channel once. Suppose that we have a complex Euclidean space $\mathcal{X}$ and two unitary channels $\Phi_U, \Phi_V \in \mathcal{C}(\mathcal{X})$. The diamond norm for such a case can be computed as [28, 39]:

$$\|\Phi_U - \Phi_V\|_\diamond = 2\sqrt{1 - \nu(V^\dagger U)^2}, \quad (4.2)$$

where we formally define $\nu : L(\mathcal{X}) \rightarrow \mathbb{R}_0^+$ as:

$$\nu(X) = \min\{|x| : x \in W(X)\}. \quad (4.3)$$

$\nu(V^\dagger U)$ quantifies the distance of the numerical range of operator $V^\dagger U$ to zero. One can see by combining Equations (4.2) and (4.1) that the two channels can be perfectly discriminated if and only if $0 \in W(V^\dagger U)$. As the numerical range and the arc function are closely related, the formulation of discrimination using the arc function requires $\theta(V^\dagger U) \geq \pi$. Using the results of [19, 18, 40], we can compute ν , which is used for the calculation of the diamond norm, $\nu = \cos(\frac{\theta}{2})$ for $\theta \in [0, \pi]$.

Example 4.1. Recall the numerical range and arc function for the operator $U = \text{diag}([1, e^{2i\pi/3}])$ (Fig 2.2 and Fig 2.1). Let us define the unitary channel $\Phi_U(X) = UXU^\dagger$, and assume an additional channel $\Phi_{\mathbf{1}_{\mathbb{C}^2}}(X) = X$ for $X \in L(\mathbb{C}^2)$. It follows that the two channels cannot be perfectly discriminated ($P_s \neq 1$). In fact, from Eq. (4.1) $P_s = \frac{1}{2} + \frac{\sqrt{3}}{4} \approx 0.933$. We illustrate this scenario in Fig. 4.1.

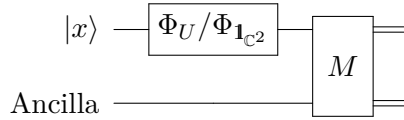


Figure 4.1: Ancilla-assisted quantum discrimination scheme between the quantum channels Φ_U and $\Phi_{\mathbf{1}_{\mathbb{C}^2}}$. An input $|x\rangle$ (possibly entangled with an auxiliary system) interacts with an unknown quantum channel. A POVM measurement M is then performed on the system to identify the outcome.

4.2 Multiple-shot discrimination of unitary channels

In contrast to the single-shot scenario, we now consider a multi-shot discrimination strategy where Alice applies one of the two quantum unitary channels N times. In this thesis, we consider three primary configurations. For all of these, assume we are given $N = dw$ copies of the channel, where d stands for the depth of the quantum circuit and w stands for its width. In all scenarios, we implicitly assume that only one of the unitaries $W \in \{U, V\}$ is applied in a single run of the discrimination task. We distinguish between the following strategies (sometimes referred to as schemes):

- **Sequential strategy:** This strategy applies the unitary W to a single qubit, allowing for intermediate quantum processing between applications, setting $N = d$ and $w = 1$.
- **Parallel strategy:** This strategy utilizes N simultaneous applications of W across distinct qubits, with exactly one operation per qubit, setting $N = w$ and $d = 1$.

- **Hybrid strategy:** This is a general framework where $N = dw$ applications of W are distributed across w qubits with depth d .

Sequential strategy

Let us begin with the sequential strategy. As depicted in Fig. 4.2, the unknown unitary W is applied serially to a single qubit system. Between each application of W , processing unitary gates X_i may be applied to adjust the state evolution.

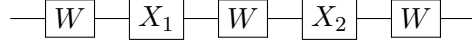


Figure 4.2: Schematic representation of a sequential discrimination strategy. The unknown unitary W is applied to a single system three times, for $d = 3$ and $w = 1$. In the sequential strategy, the processing gates X_i might be used.

To analyze this strategy, we define two unitaries A, B corresponding to the scenarios where the unknown unitary is V or U , respectively. Let

$$A = V \cdot X_1 \cdot V \cdot X_2 \cdot V \quad \text{and} \quad B = U \cdot X_1 \cdot U \cdot X_2 \cdot U. \quad (4.4)$$

By discriminating between the two new unitaries, we rewrite the multiple-shot sequential discrimination problem into an adjusted single-shot discrimination scheme.

Let us comment on the unitary operators X_i from the sequential scheme. Direct findings of [18] state the following theorem.

Theorem 4.2 (Duan's processing). *Let $\mathcal{X}$ be a complex Euclidean space, let U and V be two unitary operations, and let $N = \left\lceil \frac{\pi}{\theta(U^\dagger V)} \right\rceil$. Then there exist intermediate processing unitaries $X_1, \dots, X_{N-1} \in \mathcal{U}(\mathcal{X})$ and an initial quantum state $|\psi\rangle \in \mathcal{X}$ such that the output quantum states:*

$$|\rho\rangle = (UX_{N-1} \dots X_1 U)|\psi\rangle \quad \text{and} \quad |\phi\rangle = (VX_{N-1} \dots X_1 V)|\psi\rangle$$

are orthogonal, satisfying $\langle \rho | \phi \rangle = 0$.

This theorem establishes a relation between the arc function of $U^\dagger V$ and the minimum number of channel copies required for the perfect discrimination using processing unitaries X_i . Observe that the discriminator $|\psi\rangle$ in Theorem 4.2 is a specific quantum state, which can be entangled with some auxiliary system, that is optimized to maximize the distinguishability between the outputs of the possible channels. However, based on the results of R. Duan [18], one does not need an entangled system to achieve perfect discrimination between unitary transformations. We emphasize once more that the discriminator is in general a density operator. Since Duan's processing includes only unitaries, it can be represented entirely by a state vector.

Parallel strategy

Next, we examine the parallel strategy. In this setting, illustrated in Fig. 4.3, the unknown unitary W is applied three times (for $N = 3$) simultaneously across distinct subsystems without intermediate processing.

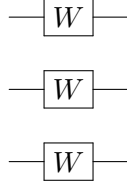


Figure 4.3: Schematic representation of a parallel discrimination strategy. Three copies of the transformed unitary W are applied in a parallel way. In this case, $w = 3$ and $d = 1$.

We define the unitary operators $A, B \in \text{U}(\mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \mathbb{C}^2)$ as:

$$A = U^{\otimes 3} \quad \text{and} \quad B = V^{\otimes 3}. \quad (4.5)$$

This view allows the parallel scheme to be adjusted into a single-shot framework in the larger complex Euclidean space $\mathcal{X} = (\mathbb{C}^2)^{\otimes 3}$.

Hybrid strategy

We now introduce the hybrid scheme, which combines the previous two approaches. Rather than restricting the circuit to a single qubit ($w = 1$) or a single layer of depth ($d = 1$), the hybrid strategy utilizes w parallel subsystems, each applying the sequential strategy with depth d . Consequently, the total number of applied copies of the discriminated channel is $N = wd$. As illustrated in Fig. 4.4 for a configuration with width $w = 3$ and depth $d = 4$ (totaling $N = 12$ applications), the unknown unitary W is applied serially d times across each of the w subsystems. Just as in the sequential strategy, intermediate processing unitaries $X_{k,i}$ can be inserted between the applications of W on the k -th qubit.

Since in both parallel and hybrid schemes only unitaries are employed (even the processing matrices $X_{i,j}$ are unitary), the condition for perfect discrimination reduces to the numerical range Eq. (2.39) or the arc function Eq. (2.42).

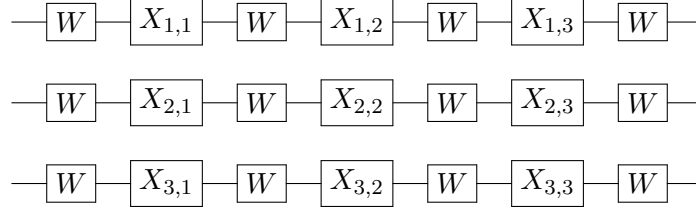


Figure 4.4: Schematic representation of a hybrid discrimination strategy. The unknown unitary W is applied twelve times to a system of width $w = 3$ and depth $d = 4$. Intermediate processing gates $X_{k,i}$ ($X_{k,i}$ may be $\mathbb{I}$) are applied sequentially on each qubit.

4.3 Discriminator construction for perfect discrimination of unitary channels

The subsequent sections of this chapter are devoted to the study of perfect distinguishability tasks ($P_s = 1$). To construct the discriminator for the perfect discrimination task, we derive an algorithm whose cornerstone is the triangle case of Carathéodory's Theorem [26]. This theorem states that any point in a convex subset of $\mathbb{R}^2$ can be expressed as a convex combination of at most three points.

Consider a unitary operator $U \in \mathcal{U}(\mathcal{X})$ and the identity $\mathbb{I}_{\mathcal{X}} \in \mathcal{U}(\mathcal{X})$ for a complex Euclidean space $\mathcal{X}$. Furthermore, suppose that their corresponding channels are perfectly distinguishable. The goal is to find a unit vector $|x\rangle \in \mathcal{X}$ such that $\langle x|U|x\rangle = 0$.

Let $I = \{0, \dots, \dim(\mathcal{X}) - 1\}$ and let us express U via its spectral decomposition

$$U = \sum_{j \in I} e^{i\phi_j} |u_j\rangle\langle u_j| \quad (4.6)$$

and expand $|x\rangle$ in this orthonormal eigenbasis as $|x\rangle = \sum_{j \in I} c_j |u_j\rangle$, where $c_j = \langle u_j|x\rangle$. We consolidate the normalization and orthogonality constraints into the following system:

$$\langle x|x\rangle = \sum_{j \in I} |c_j|^2 = \sum_{j \in I} p_j = 1, \quad (4.7)$$

$$\langle x|U|x\rangle = \langle x| \left(\sum_{j \in I} e^{i\phi_j} |u_j\rangle\langle u_j| \right) |x\rangle = \sum_{j \in I} e^{i\phi_j} |c_j|^2 = \sum_{j \in I} p_j e^{i\phi_j} = 0, \quad (4.8)$$

where we define the non-negative weights $p_j = |c_j|^2$. Now, we represent the scalar value $\lambda_j \in \mathbb{C}$ as a vector $|\lambda_j\rangle \in \mathbb{R}^2$ given by $|\lambda_j\rangle = (\text{Re}(\lambda_j), \text{Im}(\lambda_j))^T$. Perfect discrimination is possible if and only if the origin lies in the convex hull of the eigenvalues of U . By Carathéodory's theorem, this implies that the origin lies in the convex hull of at most three eigenvalues, i.e., $0 \in \text{Conv}(\lambda_a, \lambda_b, \lambda_c)$ [19],

where $\lambda_a, \lambda_b, \lambda_c \in \lambda(U)$ are associated with the respective eigenvectors $|u_a\rangle, |u_b\rangle, |u_c\rangle$.

We denote the eigenvalues as $\lambda_j = e^{i\phi_j}$ and parameterize the complex coefficients c_j using the real weights p_j and an arbitrary phase freedom $\theta_j \in [0, 2\pi)$. Since $|e^{i\theta_j}| = 1$, this definition inherently satisfies the requirement $|c_j|^2 = p_j$:

$$c_j = \sqrt{p_j} e^{i\theta_j} \implies |c_j|^2 = p_j. \quad (4.9)$$

Selecting specific values for the free parameters θ_j allows us to tailor the discriminator state to specific hardware-native gate decompositions. Restricting the solution to three eigenstates indexed by $S = \{a, b, c\}$, the orthogonality ($\sum_{j \in S} p_j \lambda_j = 0$) and normalization ($\sum_{j \in S} p_j = 1$) conditions form the linear system:

$$\begin{bmatrix} \text{Re}(\lambda_a) & \text{Re}(\lambda_b) & \text{Re}(\lambda_c) \\ \text{Im}(\lambda_a) & \text{Im}(\lambda_b) & \text{Im}(\lambda_c) \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} p_a \\ p_b \\ p_c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (4.10)$$

Solving for the weights p_j yields the exact discriminator state $|x\rangle = \sum_{j \in S} c_j |u_j\rangle$. In the specific case where the chosen subset contains antipodal eigenvalues ($\lambda_c = -\lambda_a$), the linear system reduces to two states, forcing equal weights $p_a = p_c = \frac{1}{2}$ and $p_b = 0$. The discriminator then simplifies to an equal superposition of two eigenvectors:

$$|x\rangle = \frac{1}{\sqrt{2}}(|u_a\rangle + |u_c\rangle). \quad (4.11)$$

The preceding discussion provides the main geometric idea behind the construction of the discriminator. We now summarize the procedure step by step in Algorithm 1.

4.3.1 Implementation and geometric optimizations

Quantum channel discrimination is traditionally formulated using semidefinite programming (SDP) [28]. While classical interior-point methods solve SDPs in polynomial time relative to the matrix dimension n (typically $\mathcal{O}(n^3)$ or higher-order polynomial complexity per iteration [41]), this scaling creates a fundamental bottleneck for multi-qubit systems. For parallel discrimination over w qubits, the density matrix dimension grows exponentially to $n = 2^w$, leading to a per-iteration complexity on the order of $\mathcal{O}(8^w)$ classical complexity [24]. For larger system (e.g. $w > 15$), constructing and solving the resulting optimization problems over such matrices becomes computationally prohibitive in practice. Furthermore, while previous studies utilized permutation symmetry to simplify parallel unitary discrimination, their reliance on algebraically resolving large linear systems effectively limited explicit solutions to small channel depths (e.g., $N \leq 4$ [42] and $N \leq 10$ [20]).

To avoid this exponential wall and bypass the algebraic bottlenecks of prior literature, we introduce a hybrid symbolic-geometric method. By translating the reduced combinatorial space into a geometric framework, our approach bypasses explicit representations of exponentially large matri-

Algorithm 1 Discriminator construction for unitary channels

Require: Two unitary operators $U, V \in \mathcal{U}(\mathcal{X})$ and arbitrary phase parameters $\theta_j \in [0, 2\pi)$.

Ensure: A discriminator state $|x\rangle$ such that $\langle x|V^\dagger U|x\rangle = 0$.

- 1: **Step 1: Transformation to discrimination between W and identity**
 - 2: Define unitary $W \leftarrow V^\dagger U$.
 - 3: **Step 2: Spectral Decomposition**
 - 4: Compute eigenvalues $\lambda(W) = \{\lambda_j = e^{i\phi_j}\}$ and eigenvectors $\{|u_j\rangle\}$ of W .
 - 5: **Step 3: Subset Selection** ▷ Find points containing origin in their convex hull
 - 6: **if** $\exists \lambda_a, \lambda_b \in \lambda(W)$ such that $\lambda_a = -\lambda_b$ (antipodal) **then**
 - 7: Select indices $S \leftarrow \{a, b\}$.
 - 8: Set weights $p_a \leftarrow 1/2, \quad p_b \leftarrow 1/2$.
 - 9: **else**
 - 10: Find indices $S \leftarrow \{a, b, c\}$ such that $0 \in \text{Conv}(\lambda_a, \lambda_b, \lambda_c)$.
 - 11: **Step 4: Solve linear system**
 - 12: Construct matrix M and vector $|b\rangle$:
$$M = \begin{bmatrix} \text{Re}(\lambda_a) & \text{Re}(\lambda_b) & \text{Re}(\lambda_c) \\ \text{Im}(\lambda_a) & \text{Im}(\lambda_b) & \text{Im}(\lambda_c) \\ 1 & 1 & 1 \end{bmatrix}, \quad |b\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$
 - 13: Solve $M|p\rangle = |b\rangle$ for $|p\rangle = [p_a, p_b, p_c]^T$.
 - 14: **end if**
 - 15: **Step 5: Construct vector**
 - 16: Calculate coefficients $c_j \leftarrow e^{i\theta_j} \sqrt{p_j}$ for $j \in S$.
 - 17: Construct the discriminator state:
$$|x\rangle \leftarrow \sum_{j \in S} c_j |u_j\rangle$$
 - 18: **return** $|x\rangle$
-

ces. Let L denote the number of distinct eigenvalues of the composite operator $W^{\otimes C}$, where C is the number of parallel channel instances and W has a constant number D of distinct eigenvalues (as is typical for hardware-native gates). In this case, $L = \binom{D+C-1}{C}$ grows polynomially in C , with degree at most $D - 1$.

This approach preserves exact analytical structure while restricting the computational complexity to $\mathcal{O}(L \log L)$, enabling the analytical determination of discriminators for high-dimensional systems (e.g., $C > 100$) that are otherwise inaccessible.

4.3.2 Eigenvalue evolution under tensor products

While permutation invariance has been applied to specific distinguishability problems [20], we generalize this reduction for an arbitrary number of distinct eigenvalues. For a composite operator W applied across w qubits, the transformation is governed by $W^{\otimes C}$, where $C = w / \log_2(\dim(W))$ is the number of parallel instances.

Given that W has distinct eigenvalues $\{\lambda_1, \dots, \lambda_D\}$, any eigenvalue Λ of $W^{\otimes C}$ is a product of exactly C base eigenvalues [43]. Because scalar multiplication commutes, the specific subsystem ordering is irrelevant. Consequently, Λ is uniquely defined purely by:

$$\Lambda = \prod_{i=1}^D (\lambda_i)^{c_i}, \quad \text{where } \sum_{i=1}^D c_i = C, \quad (4.12)$$

where $c_i \in \mathbb{N}_0$ for each $i \in \{1, \dots, D\}$ represents the frequency of occurrence of the base eigenvalue in the product. This symmetry allows us to bypass the diagonalization of the exponentially large $2^w \times 2^w$ matrix. Instead of computing 2^w eigenvalues, we only calculate the distinct partitions of the D base eigenvalues, vastly reducing the problem dimension to a purely combinatorial scalar space L .

4.3.3 Combinatorial state mappings

Generating the tensor product space for C parallel instances of a unitary with D distinct eigenvalues corresponds to combinations with repetition. By enforcing permutation symmetry via canonical lexicographical mapping, we reduce the computationally infeasible tensor space to exactly $L = \binom{D+C-1}{C}$ eigenvalue combinations.

We give an example for $W \otimes W$. Let $e^{i\phi_i}$ and $e^{i\phi_j}$ be eigenvalues of W , with corresponding eigenvectors $|e_i\rangle$ and $|e_j\rangle$, respectively. Since the complex scalars commute under multiplication ($e^{i\phi_j} e^{i\phi_i} = e^{i\phi_i} e^{i\phi_j}$), the distinct tensor product vectors $|e_j\rangle \otimes |e_i\rangle$ and $|e_i\rangle \otimes |e_j\rangle$ yield the same eigenvalue under $W \otimes W$, allowing us to group these permutations into a single equivalence class.

For instance, a single-qubit native gate ($D = 2, C = w$) results in $\mathcal{O}(w \log w)$. A two-qubit gate ($D = 4, C = w/2$) yields a cubically scaling space of $L = \binom{w/2+3}{3}$, which translates to $\mathcal{O}(w^3 \log w)$.

4.3.4 Hybrid solver execution and lazy evaluation

To avoid exponential symbolic or algebraic overhead, the algorithm projects symbolic eigenvalues into arbitrary-precision floats, applying geometric filtering and lazy evaluation:

1. **Preprocessing ($\mathcal{O}(\log d)$ for fixed D):** The operator $W = V^\dagger U$ is computed symbolically. Since physical gates typically have $D \leq 4$, the $\mathcal{O}(D^3)$ matrix operations are effectively bounded by a constant [44]. Base eigenvalues are then raised to channel depth d via $\mathcal{O}(\log d)$ binary exponentiation [45].
2. **Combinatorial generation:** The solver generates the purely combinatorial space of L integer partitions, bypassing explicit matrix instantiations.
3. **Phase sorting ($\mathcal{O}(L \log L)$):** The L combinations are mapped to their complex arguments $\theta \in [0, 2\pi)$ and sorted [46].
4. **Geometric sweeps ($\mathcal{O}(L \log L)$):** The solver restricts the search space geometrically using the sorted angles:
 - *Antipodal search:* Binary searches ($\mathcal{O}(\log L)$ per state) locate phase pairs intersecting the origin (θ and $\theta \pm \pi$).
 - *Convex triplet search:* A sliding window identifies spanning triangles enclosing the origin (all angular gaps $< \pi$), computing probability weights p_j via Cramer's rule.
5. **Lazy evaluation:** Heavy symbolic verification is restricted exclusively to the numeric candidates isolated by the geometric sweeps. Once confirmed, the exact discriminator $|x\rangle = \sum \sqrt{p_j} e^{i\theta_j} |u_j\rangle$ is provided.

4.3.5 Time complexity of the implementation of Algorithm 1

The derivation in the current section results in the following theorem.

Theorem 4.3 (Hybrid geometric solver). *Let $U, V \in \mathcal{U}(\mathcal{X})$ be native unitary operators acting on a small, constant number of qubits, such that the operator $W = V^\dagger U$ has constant $D \ll 2^w$ (e.g., $D \leq 4$ for single-qubit or two-qubit hardware-native gates). Suppose we evaluate C parallel instances of these channels, resulting in the composite operator $W^{\otimes C}$.*

Assuming the channels are perfectly distinguishable, an optimal pure discriminator state $|x\rangle$ satisfying the orthogonality condition $\langle x | W^{\otimes C} | x \rangle = 0$ can be constructed using the hybrid symbolic-geometric solver in $\mathcal{O}(L \log L)$ computational time, where:

$$L = \binom{D + C - 1}{C}$$

represents the polynomial combinatorial search space of eigenvalue combinations for fixed D .

4.4 Measurement construction via Jordan-Hahn decomposition

To construct the optimal measurement for discriminating between two density operators $\rho_0, \rho_1 \in \mathcal{D}(\mathcal{X})$, we consider the Hermitian operator of the form:

$$H = \rho_0 - \rho_1 \in \text{Herm}(\mathcal{X}). \quad (4.13)$$

By the Jordan-Hahn decomposition (Theorem 2.5), the operator H can be expressed as:

$$H = P - Q, \quad (4.14)$$

where $P, Q \in \text{Pos}(\mathcal{X})$ and $PQ = 0$.

To find the optimal final measurement $\{E_0, E_1\} \subset \text{Pos}(\mathcal{X})$ satisfying $E_0 + E_1 = \mathbb{1}_{\mathcal{X}}$. Let Π_P denote the orthogonal projector onto the support of the positive part P of H . Then, the optimal measurement operators are given by

$$E_0 = \Pi_P, \quad E_1 = \mathbb{1}_{\mathcal{X}} - \Pi_P. \quad (4.15)$$

This construction corresponds to the Helstrom measurement, which achieves the optimal success probability for state discrimination [28].

Chapter 5

Components of the multiple-shot perfect discrimination task

This chapter is dedicated to the multiple-shot perfect discrimination of unitary channels ($P_s = 1$). We present a discriminator and the corresponding optimal measurements for the introduced protocols. This chapter is divided into two major parts. The first section is dedicated to the discrimination task without processing. The second section describes the discrimination task in which processing operators are used.

For our experiments, we restrict $\mathcal{X} = \mathbb{C}^2$. We focus strictly on the discrimination of two unitary channels, generated by $U, V \in \mathcal{U}(\mathcal{X})$, under the distinguishability condition:

$$\theta(V^\dagger U) = \frac{\pi}{N}, \quad (5.1)$$

where $N = wd$ is the total number of channel queries, arranged in a hybrid layout of width w and depth d .

The general protocol applies N copies of the unknown unitary $W \in \{U, V\}$ across a w -qubit register. Incorporating intermediate processing layers $X_1, \dots, X_{d-1} \in \mathcal{U}((\mathbb{C}^2)^{\otimes w})$ as per Theorem 4.2, the global unitary evolution is given by:

$$W_{\text{global}} = W^{\otimes w} \cdot X_{d-1} \cdot W^{\otimes w} \cdots X_1 \cdot W^{\otimes w}. \quad (5.2)$$

These intermediate unitaries are responsible for the necessary processing between applications of the unknown channel, as established in Theorem 4.2. The arc function then results in:

$$\theta\left(\left((V^\dagger U)^d\right)^{\otimes w}\right) = w\theta\left((V^\dagger U)^d\right) = wd\theta(V^\dagger U) = N\frac{\pi}{N} = \pi. \quad (5.3)$$

5.1 Discrimination of unitary channels on IBM Brisbane without processing

Now, let us present the discrimination task without processing. In this configuration, we assume all intermediate processing matrices are identities ($X_i = \mathbb{1}_{\mathcal{X}}^{\otimes w}$). We aim to discriminate between the identity channel $\mathbb{1}_{\mathcal{X}}$ and the phase rotation $RZ(\phi)$ using $N = wd$ channel copies. The $RZ(\phi)$ gate is defined as:

$$RZ(\phi) = \begin{bmatrix} e^{-\frac{i\phi}{2}} & 0 \\ 0 & e^{\frac{i\phi}{2}} \end{bmatrix}. \quad (5.4)$$

The arc function for this pair evaluates to $\theta(\mathbb{1}_{\mathcal{X}}^\dagger \cdot RZ(\phi)) = \phi$. To satisfy the condition for perfect distinguishability (Eq. (5.1)) with N queries, we set the rotation angle to:

$$\phi = \frac{\pi}{N}. \quad (5.5)$$

5.1.1 Discriminator

To compute the discriminator, we explicitly expand the sequence of applied gates. The target rotation is:

$$RZ\left(\frac{\pi}{wd}\right) = \begin{bmatrix} e^{-\frac{i\pi}{2wd}} & 0 \\ 0 & e^{\frac{i\pi}{2wd}} \end{bmatrix}. \quad (5.6)$$

Applying this gate sequentially d times on a single qubit yields the depth-accumulated unitary:

$$RZ_{depth}\left(\frac{\pi}{w}\right) = \left(RZ\left(\frac{\pi}{wd}\right)\right)^d = \begin{bmatrix} e^{-\frac{i\pi}{2w}} & 0 \\ 0 & e^{\frac{i\pi}{2w}} \end{bmatrix}. \quad (5.7)$$

The global operation over the w -qubit register is the w -fold tensor product:

$$RZ_{global} = \bigotimes_{j=1}^w RZ_{depth}\left(\frac{\pi}{w}\right). \quad (5.8)$$

This composite unitary RZ_{global} is diagonal in the computational basis. Its extreme eigenvalues, which are critical for maximizing the arc function, correspond to the states $|0 \dots 0\rangle$ and $|1 \dots 1\rangle$. Specifically, $|0 \dots 0\rangle$ yields the eigenvalue $(e^{-i\pi/(2w)})^w = -i$, and $|1 \dots 1\rangle$ yields $(e^{i\pi/(2w)})^w = i$. All other computational basis states produce eigenvalues with phases strictly inside the interval $(-\pi/2, \pi/2)$.

Since the origin 0 lies precisely at the midpoint of the line segment connecting the extreme eigenvalues $-i$ and i , the optimal discriminator is an equal superposition:

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(e^{i\phi_1} |0 \dots 0\rangle + e^{i\phi_2} |1 \dots 1\rangle \right) \in \mathbb{C}^{2^w}. \quad (5.9)$$

Applying the corresponding channels to this discriminator yields the states $|\psi_{\mathbf{1}}\rangle = |\psi\rangle$ and $|\psi_{RZ}\rangle = \frac{-ie^{i\phi_1}}{\sqrt{2}}|0\dots 0\rangle + \frac{ie^{i\phi_2}}{\sqrt{2}}|1\dots 1\rangle$. It is straightforward to verify that these outputs are perfectly orthogonal, satisfying $\langle\psi_{\mathbf{1}}|\psi_{RZ}\rangle = 0$.

5.1.2 Measurement

To execute this measurement on quantum hardware, we must apply a pre-measurement unitary rotation U that maps the orthogonal discrimination states into the computational basis. We define the unitary U strictly by its action on our target states:

$$U|\psi_{\mathbf{1}}\rangle = |0\dots 0\rangle, \quad (5.10)$$

$$U|\psi_{RZ}\rangle = |1\dots 1\rangle. \quad (5.11)$$

The remainder of the operator U is constructed by completing the orthonormal basis. We then refer to the measurement as *measurement U* .

When executing this protocol on real quantum devices, one must consider the effect of hardware noise, such as bit-flip errors. Suppose a 5-qubit system yields an ideal state of $|00000\rangle$ for the identity channel. A single bit-flip error on the third qubit results in the measured state $|00100\rangle$. To identify the channels reliably despite such noise, we can simply evaluate the Hamming distance between the measured bitstring and our ideal targets. Because $|00100\rangle$ has a Hamming distance of one from $|00000\rangle$ and four from $|11111\rangle$, we can deduce the state was originally $|00000\rangle$, thereby identifying the identity channel.

5.2 Discrimination of unitary channels on IBM Brisbane with processing

As a second example, we consider a pair of unitary channels that require intermediate processing between applications to achieve perfect distinguishability. We define the unitary channels Φ_U and Φ_V via operators $U, V \in \text{U}(\mathbb{C}^2)$ as:

$$U = \sqrt{X}\text{RZ}\left(\frac{-\pi}{2N}\right)\sqrt{X} \quad \text{and} \quad V = \sqrt{X}\text{RZ}\left(\frac{\pi}{2N}\right)\sqrt{X}. \quad (5.12)$$

It is straightforward to verify that the condition $\theta(V^\dagger U) = \frac{\pi}{N}$ is satisfied, implying that perfect discrimination is theoretically achievable for $N = wd$ queries.

To align the evolution with hardware constraints, we employ a specific processing strategy. Instead of the theoretically optimal intermediate processing $X_i = (V^\dagger)^{\otimes w}$, we utilize a hardware-friendly processing layer $X_i = X^{\otimes w}$. Additionally, we introduce a pre-processing unitary $X_0 = (X\sqrt{X})^{\otimes w}$ and a post-processing operation $X_d = (\sqrt{X}X)^{\otimes w}$.

Combining these processing operators with the unknown channel Φ_U , the effective unitary operation U_{global} becomes:

$$U_{global} = X_d U^{\otimes w} X_{d-1} \cdots X_1 U^{\otimes w} X_0 = \left(\text{RZ} \left(\frac{-\pi}{2N} \right)^d \right)^{\otimes w}. \quad (5.13)$$

Analogously, for Φ_V , the combined unitary circuit yields:

$$V_{global} = \left(\text{RZ} \left(\frac{\pi}{2N} \right)^d \right)^{\otimes w}. \quad (5.14)$$

The angle between the processed unitaries is:

$$\theta(V_{global}^\dagger U_{global}) = \theta \left(\left(\text{RZ} \left(\frac{-\pi}{N} \right)^d \right)^{\otimes w} \right) = wd \cdot \theta \left(\text{RZ} \left(\frac{-\pi}{N} \right) \right) = N \frac{\pi}{N} = \pi. \quad (5.15)$$

Since the resulting angle is π , the channels are perfectly distinguishable. Note that X , $\sqrt{X}$, and RZ are native gates on the IBM quantum architecture, allowing this circuit to be implemented exactly as described. We omit the explicit derivation of the discriminator here, as it perfectly mirrors the procedure established in the previous section, again yielding output states $|\psi_{\mp i}\rangle$ corresponding to the accumulated phases.

5.2.1 Measurement

To distinguish between the resulting states $|\psi_{-i}\rangle$ and $|\psi_i\rangle$, we construct a hardware-efficient parity measurement circuit. The procedure involves applying an $\text{RZ}(\pi/2)$ rotation on the first qubit, followed by the parallel application of Hadamard gates $H^{\otimes w}$ on all qubits, and concluding with a standard Z -basis measurement.

Applying these pre-measurement operations to the output states yields:

$$H^{\otimes w} (\text{RZ}(\pi/2) \otimes \mathbb{1}_{\mathcal{X}}^{\otimes w-1}) |\psi_{\mp i}\rangle = |+\dots+\rangle \pm |-\dots-\rangle, \quad (5.16)$$

where $|+\rangle = H|0\rangle$ and $|-\rangle = H|1\rangle$.

Let $(b_1, \dots, b_w)$ denote the bit string obtained from the measurement. The discrimination logic relies entirely on the parity of this bit string:

- $\Phi = \Phi_U$ iff $b_1 \oplus \dots \oplus b_w \equiv 0$,
- $\Phi = \Phi_V$ iff $b_1 \oplus \dots \oplus b_w \equiv 1$,

where $\oplus$ denotes the XOR operation (sum modulo 2). Thus, the parity of the measurement outcome serves as the deterministic and decisive factor for identifying the unknown channel.

Chapter 6

Implementation of experiments on IBM Brisbane processor

This chapter presents the experimental implementation of unitary channel discrimination on the IBM Brisbane processor and evaluates the performance of the generated circuits. First, we introduce the topology of the Eagle R3 architecture. This is followed by a discussion about hardware-aware circuit mappings for quantum circuits. Subsequently, we detail the results of two experimental configurations: discrimination without intermediate processing, and discrimination with intermediate processing. This chapter is based on the results published in [25, 47].

6.1 Topology and gate set of IBM Brisbane

There are two critical hardware constraints to address regarding the IBM Eagle R3 architecture [1], which powers the IBM Brisbane processor. First, the native entangling operation is the Echoed Cross-Resonance (ECR) gate, previously introduced in Tab. 3.2, rather than the standard CNOT gate. Crucially, the ECR gate is asymmetric. The ordering of the control and target qubits matters. For two qubit states $|x\rangle$ and $|y\rangle$,

$$\text{ECR}|xy\rangle \neq \text{ECR}|yx\rangle, \quad (6.1)$$

which can be verified for $|x\rangle = |1\rangle$ and $|y\rangle = |0\rangle$ for example.

This asymmetry dictates that the processor's topology must be modeled as a directed graph, where edges represent the native, hardware-supported direction of ECR operations, as illustrated in Fig. 6.1 [48].

The processor features a heavy-hexagonal layout. This specific geometry restricts the connectivity such that any given qubit can have a maximum of three adjacent neighbors (a maximum node degree of three). Examples of such high-connectivity hub nodes include qubits 104 and 106.

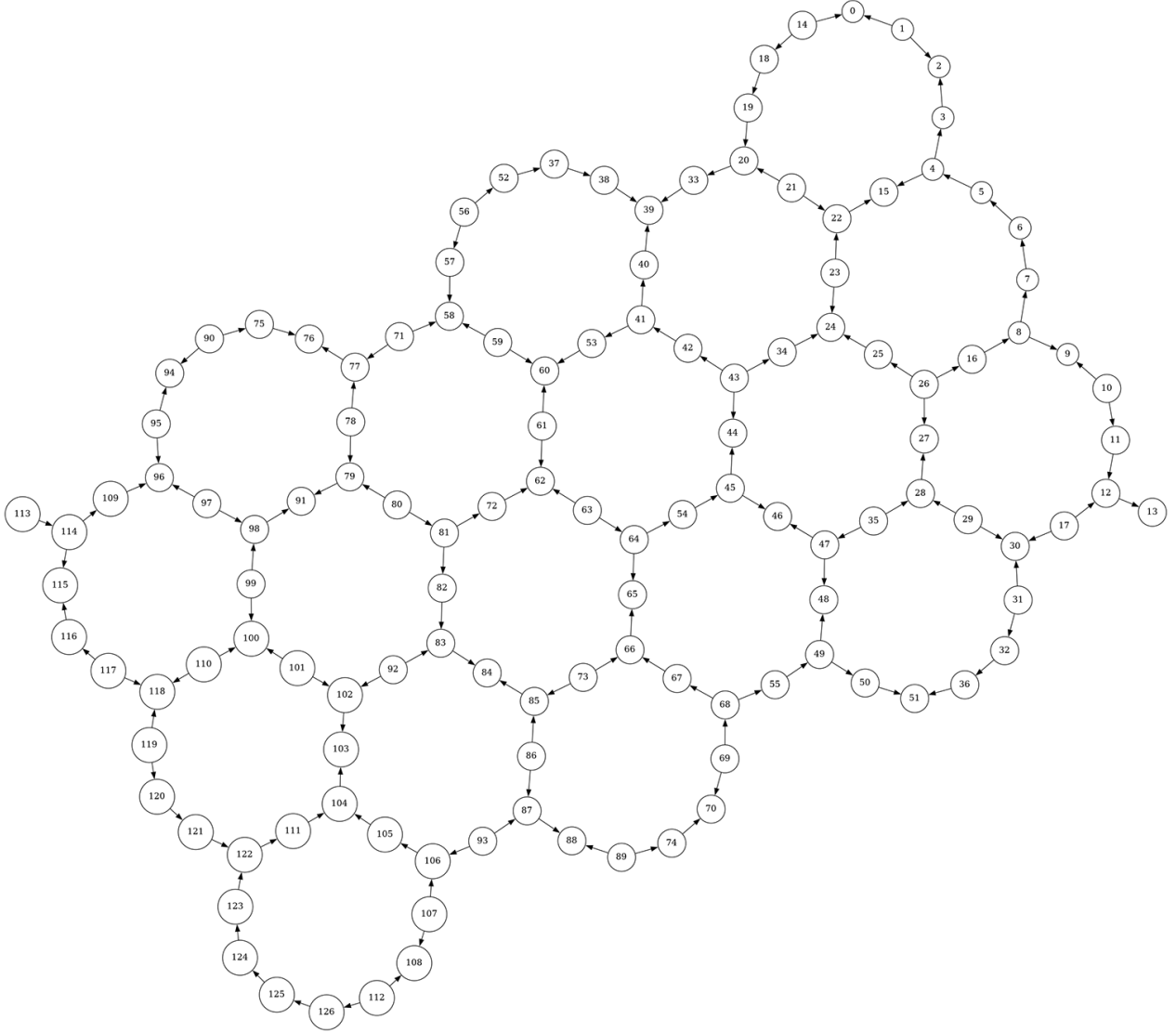


Figure 6.1: The coupling map of the IBM Eagle R3 processor, representing a 127-qubit system. The architecture is defined by a heavy-hexagonal graph topology, where the maximum degree of any node is $d = 3$.

6.1.1 Circuit construction

This subsection details the circuit construction strategies, divided into three parts. First, we discuss circuit design using strictly hardware-native gates. Second, we describe the implementation using standard, non-native quantum gates (such as CNOT and H). Finally, we detail the circuit design incorporating intermediate processing, as previously defined in Eq. (5.12).

6.1.1.1 Design for native gates without processing

Given the directed heavy-hexagonal mesh (Fig. 6.1) and the asymmetric nature of the ECR gate, routing the entanglement is needed. To implement the discriminator and the subsequent measurement, we consider three primary mapping strategies based on a tree topology. To illustrate this, consider the following 5-qubit circuit:

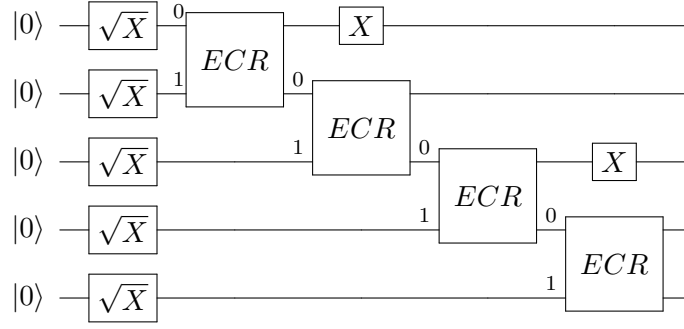


Figure 6.2: A five qubit discriminator circuit designed to distinguish between RZ rotations and the identity operation $\mathbb{1}_X$. The implementation utilizes a hardware-native gate set consisting of $\sqrt{X}$, X , and ECR gates.

The circuit in Fig. 6.2 prepares the target state $|\psi\rangle = \frac{1}{\sqrt{2}}e^{i\frac{\pi}{4}}|00000\rangle + \frac{1}{\sqrt{2}}e^{-i\frac{\pi}{4}}|11111\rangle$, which serves as the optimal discriminator according to Eq. (5.9). Translating this logical circuit to the physical hardware requires navigating two fundamental constraints of the coupling map (Fig. 6.1). First, the connectivity is *directed*. Applying an ECR gate against the native orientation necessitates the insertion of costly SWAP gates to invert the control and target qubits. Second, *the maximum degree of any node is strictly three*. These physical limitations naturally dictate three distinct topological mapping strategies. For each strategy discussed below, we provide a concrete example of physical nodes where the layout can be natively embedded within the heavy-hexagonal mesh.

Evaluating the mapping strategies from Fig. 6.3 within the context of the hardware coupling map (Fig. 6.1), we must select the optimal routing architecture. Strategy **(a)** provides a straightforward linear chain. Strategy **(b)** employs two-way branching, effectively splitting the control signal path into two distinct directions across the heavy-hex lattice. Strategy **(c)** utilizes three-way branching, where a single hub node (e.g., physical qubit 43) serves as the origin for three outgoing edges. These oriented edges are critical for compiler optimization, as any ECR gate applied in the reverse direction of the arrows would necessitate the insertion of additional swap gates to maintain logical consistency.

For the subsequent experiments, we adopt the two-way branching approach **(b)**. This configuration grants the transpiler greater flexibility to optimize routing and balance SWAP gate insertion overhead against individual qubit fidelities. Because a large majority of the processor’s nodes pos-

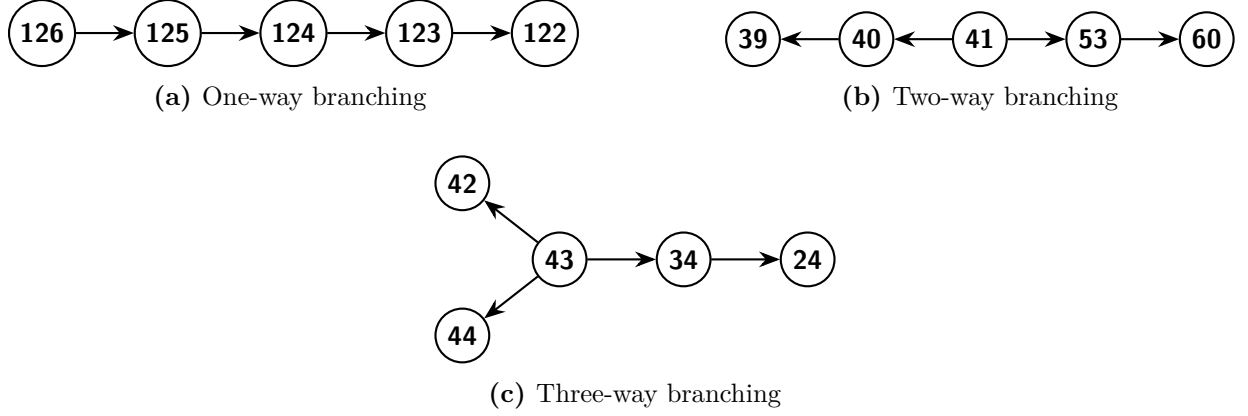


Figure 6.3: Comparison of qubit mapping topologies on the Eagle R3 heavy-hex mesh. In (a), we utilize a one-way branching (linear) mapping, which represents the simplest path through the lattice. In (b), a two-way branching strategy is shown for which the central node (41) serves as a hub for two outgoing ECR gates. Finally, (c) illustrates a three-way branching mapping, which fully saturates the maximum degree of a node within the heavy-hex topology.

sess at least two functional connections, the two-way strategy allows for high mobility across the chip. In contrast, nodes with a full degree of three are relatively sparse, severely restricting the physical placement options for strategy (c). Consequently, all following circuits are explicitly designed around this two-way branching architecture. The resulting hardware-adapted discriminator circuit is presented in Fig. 6.4.

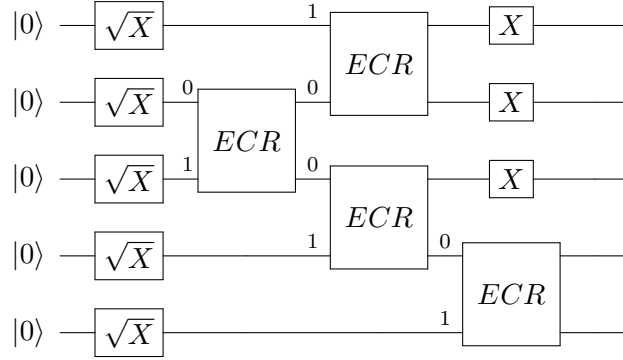


Figure 6.4: A five-qubit discriminator circuit mapped using a two-way branching strategy (strategy b). Second qubit serves as the primary hub, initiating the entanglement chain that branches through third qubit to reach the peripheral qubits.

Fig. 6.4 illustrates the adapted discriminator circuit utilizing the two-way branching strategy. The resulting state vector is given by $|\psi\rangle = \frac{1}{\sqrt{2}}e^{-i\frac{\pi}{4}}|00000\rangle + \frac{1}{\sqrt{2}}e^{i\frac{\pi}{4}}|11111\rangle$. Notably, the relative phases of the basis states are inverted compared to the original linear construction. Despite this

phase exchange, the state’s capacity to perfectly discriminate between the target channels remains mathematically equivalent.

Having established the hardware-adapted state preparation for the discriminator, we now detail the corresponding measurement protocols.

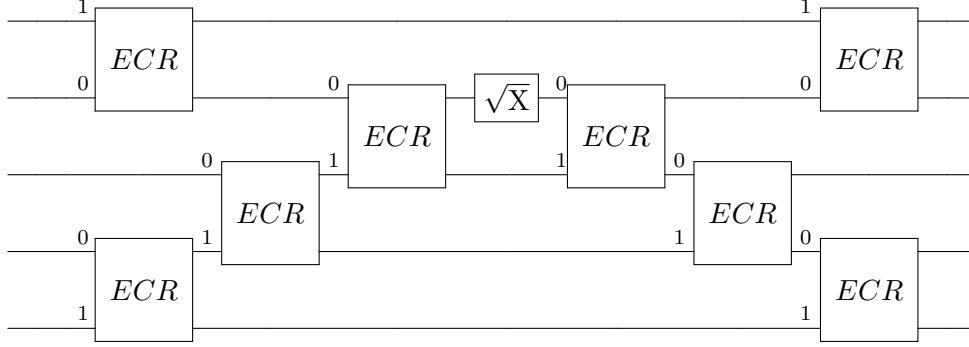


Figure 6.5: The XOR pre-measurement rotation circuit for 5 qubits. The circuit implements a mirrored ECR gate structure. Such circuit enables for post-processing using Hamming distance.

We first evaluate the XOR measurement scheme for the 5-qubit parallel configuration, as depicted in Fig. 6.5. The name derives from the property that the ideal output bitstrings corresponding to the two possible channels are exact bitwise negations of one another. This pre-measurement rotation is applied sequentially after the discriminator state preparation and the execution of the unknown quantum channel. If the applied channel is the RZ rotation, the final state prior to measurement evaluates to $|\psi\rangle = -i|00000\rangle$. Conversely, if the identity channel $\mathbb{1}_X$ is applied, the system evolves to $|\psi\rangle = i|11111\rangle$. These perfectly orthogonal computational basis states satisfy our theoretical requirements for exact discrimination. However, it is crucial to note that this specific measurement circuit requires eight ECR operations, which leads to a higher chance of errors.

Alternatively, we evaluate the Short measurement strategy for the 5-qubit parallel scheme, depicted in Fig. 6.6. Unlike the XOR measurement, which analytically maps the target states into distinct, maximal-Hamming-distance bitstrings, the Short measurement is a heuristic circuit aimed strictly at minimizing depth and the number of entangling gates. Consequently, the ideal output bitstrings do not follow a simple arithmetic parity pattern.

For the identity channel $\mathbb{1}_X$, the theoretical outcomes are the states $\{00010, 10110, 01011, 00111, 11111, 11010, 10011, 01110\}$. In contrast, the RZ channel yields $\{10111, 00011, 10010, 01010, 11011, 00110, 01111, 11110\}$. While these two outcome sets are strictly disjoint—satisfying the mathematical mandate for perfect discrimination—they must be determined numerically via a noiseless classical simulator prior to execution.

For practical post-processing, experimental hardware measurements are classified by comparing them against these simulated reference sets. However, the Hamming distance between the two sets

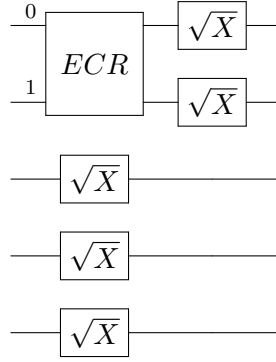


Figure 6.6: The short pre-measurement rotation circuit for 5 qubits. The circuit trades its length for potential for post-processing of results.

can be as low as one (e.g., observing 01110 for identity versus 11110 for rotation). This minimal distance severely limits the post-processing.

6.1.1.2 Design for CNOT and H gates without processing

While the previous subsection addressed both the physical connectivity and the strict orientation of the ECR gates (Fig. 6.1), in this section we omit the direction of the gates in the topology as we use CNOT and Hadamard gate sets. Because these gates form the conventional basis for logical quantum circuit design, their implementation is considerably more straightforward.

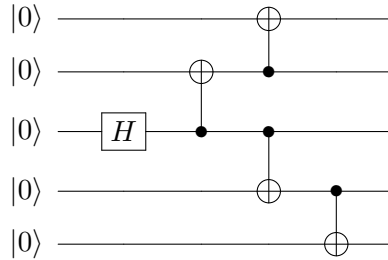
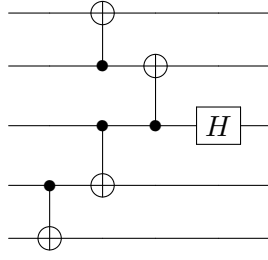
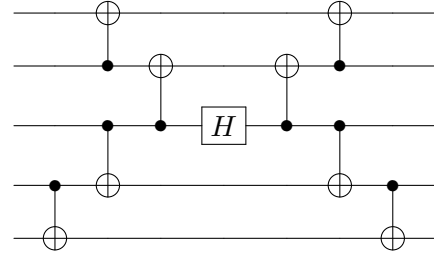


Figure 6.7: A scheme of circuit creating discriminator for 5 qubits using Hadamard gates and CNOT.

Fig. 6.7 presents a discriminator constructed from a different gate basis than the one shown in Fig. 6.4. This implementation utilizes H and CNOT gates. Although the central control qubit has been relocated, the underlying pattern remains easily recognizable. The resulting discriminator state is $|\psi\rangle = \frac{1}{\sqrt{2}}|00000\rangle + \frac{1}{\sqrt{2}}|11111\rangle$.



(a) Scheme of short measurement using CNOT and H .



(b) Scheme of XOR measurement using CNOT and H .

Figure 6.8: Implementation of quantum circuits for pre-measurement rotations for the 5-qubit branching scheme both using CNOT and H gates. In (a) we present equivalent to the short measurement. In (b) is circuit, which maximizes the Hamming distance for theoretical outcomes between identity and rotation.

Fig. 6.8 presents two measurements. These serve as the H and CNOT equivalents to the previously discussed ECR implementations. In (a), the results differ by a single qubit: for the identity channel, the output is the bitstring 00000, and for the rotation, it is 00100. In (b), the results are 00000 and 11111, respectively.

6.1.1.3 Design for discrimination with processing

This section introduces two circuits. The first prepares the discriminator for the task, and the second performs the measurement. The discriminator circuit in Fig. 6.9 uses a cascade pattern, while the measurement circuit in Fig. 6.10 utilizes Hadamard gates together with one $RZ(\frac{\pi}{2})$ gate.

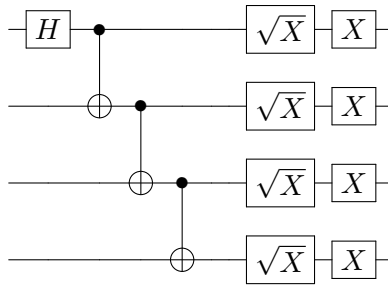


Figure 6.9: Implementation of a 4-qubit discriminator state preparation using a traditional cascade pattern for the task of discrimination with processing involving $\sqrt{X}$. The discriminator is created using H , CNOT, $\sqrt{X}$ and X gates.

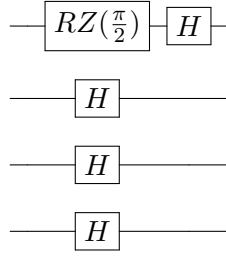


Figure 6.10: Implementation of a 4-qubit measurement for the discrimination task with processing. The measurement utilizes H and $RZ(\frac{\pi}{2})$ gates.

Measurement Error Mitigation

Physical measurements on current quantum hardware are inherently prone to readout assignment errors. To enhance readout fidelity, we employ the MThree (Matrix-free Measurement Mitigation) package [49]. MThree characterizes the readout noise of the active physical qubits and applies a scalable correction to the resulting distributions.

6.2 Experiments without processing

In this section, we present results in which the processing matrices are the identity. The presented results include a 6-qubit case and an 11-qubit case for parallel discrimination studies. The focus of these two studies is the influence of transpilation. Moreover, we also present a discussion about the sequential, parallel, and hybrid layouts, together with the probabilities of correct discrimination.

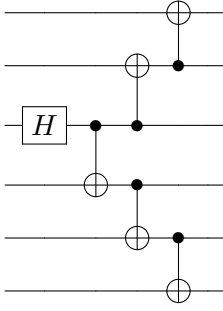
6.2.1 Experimental comparison for different techniques of transpilation

This subsection details our experimental results obtained from 6-qubit and 11-qubit systems. Having established the theoretical framework in the previous sections, we now shift our focus to the practical implementation and performance of these circuits.

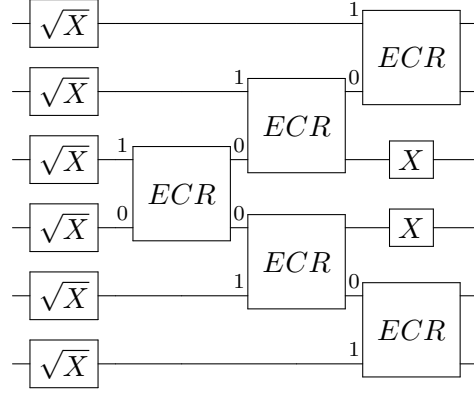
6.2.1.1 6-qubit system without processing

In this subsection, we present the circuits that we use for experiments and describe each setup, as well as the results from the experiments, including gate counts for each setup for 6-qubit system. We present both parts of the circuit in Fig. 6.11 and 6.12 for both the ECR and CNOT usage. Utilizing the insights provided in previous subsections.

Given the circuits, we now present circuit statistics (how many copies of each gate were used) for the following settings:



(a) Implementation of the discriminator using Hadamard and CNOT gates.

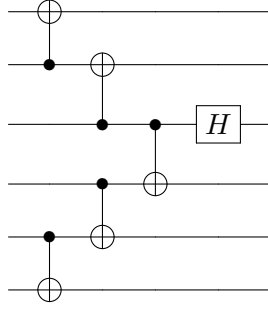


(b) Implementation of the discriminator using the native gates of the IBM Eagle R3 architecture.

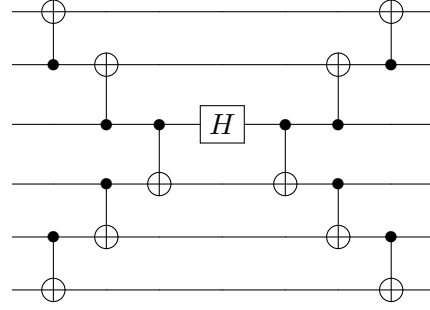
Figure 6.11: Schematic architectural mapping of a six-qubit discriminator across distinct basis gate configurations: (a) A standard decomposition utilizing Hadamard and CNOT gates; (b) A hardware-efficient implementation optimized for the IBM Eagle R3 processor, utilizing the native $\sqrt{X}$ and ECR gate set.

- **CNOT + Transpiler:** The circuit is implemented using CNOT basis gates and processed by the Qiskit transpiler with optimization level 3. This approach relies on the transpiler to decompose non-native CNOT operations into the hardware-native basis (ECR) and to handle qubit mapping and routing.
- **ECR + Transpiler:** The circuit is implemented directly using ECR basis gates, which are native to the IBM Eagle R3 architecture. It is then processed by the Qiskit transpiler with optimization level 3, allowing the software to perform optimizations without the overhead of decomposing entangling gates.
- **ECR + Transpiler + Fixed Mapping:** This strategy employs a topology-aware circuit design using ECR gates combined with a predetermined, fixed mapping of logical qubits to physical qubits. While the layout is manually fixed to align with device connectivity, the circuit is still processed by the transpiler with optimization level 3.
- **ECR + Fixed Mapping (No Opt.):** This method utilizes a topology-aware ECR circuit with a fixed logical-to-physical qubit mapping, explicitly bypassing subsequent transpiler optimization passes. This ensures that the executed circuit structure remains exactly as designed, avoiding the automatic insertion of SWAP gates or other alterations.

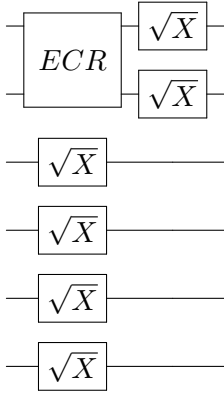
Tab. 6.1 details the circuit metrics for the 6-qubit parallel discrimination scheme, comparing four different transpilation strategies across two measurement settings: short and XOR. The presented



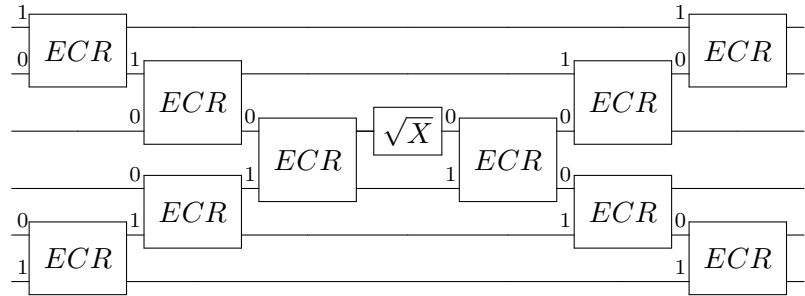
(a) CNOT implementation of the short measurement.



(b) CNOT implementation of the XOR measurement.



(c) ECR implementation of the short measurement.



(d) ECR implementation of the XOR measurement.

Figure 6.12: Schematic of six-qubit measurement circuits. Figs. (a) and (b) show the short and XOR schemes using non-native H and CNOT gates. Figs. (c) and (d) show the same schemes using native Eagle R3 gates, in this case the gates used are ECR and $\sqrt{X}$.

Transpilation Strategy	Measurement	Depth	ECR	RZ	$\sqrt{X}$	X	ID/RZ(ϕ)
CNOT + Transpiler	short	35	10	53	30	5	6
CNOT + Transpiler	XOR	44	14	62	36	9	6
ECR + Transpiler	short	20	6	18	19	1	6
ECR + Transpiler	XOR	40	14	59	34	2	6
ECR + Transpiler + Fixed Mapping	short	8	6	0	12	2	6
ECR + Transpiler + Fixed Mapping	XOR	21	14	7	12	2	6
ECR + Fixed Mapping (No Opt.)	short	8	6	0	12	2	6
ECR + Fixed Mapping (No Opt.)	XOR	13	15	0	7	2	6

Table 6.1: Comparison of 6-qubit parallel discrimination circuit metrics across different transpilation strategies and measurement types. Metrics include circuit depth and gate counts for ECR, RZ , $\sqrt{X}$, X , and $\mathbb{1}_X/RZ(\phi)$ gates.

Transpilation Strategy	Short accuracy	XOR Accuracy
CNOT + Transpiler	88.8%	86.4%
ECR + Transpiler	83.8%	90.0%
ECR + Transpiler + Fixed Mapping	84.4%	85.3%
ECR + Fixed Mapping (No Opt.)	83.3%	85.6%

Table 6.2: Experimental accuracy results for different transpilation strategies. Comparison includes short and XOR measurement accuracies for CNOT and ECR-based implementations. Again, the setting is the parallel scheme.

metrics in this table include the number of each gate used, as well as the depth of the circuit. Usually, the Short measurement scheme introduces less depth, moreover, such setting also lowers number of ECR gates.

Tab. 6.2 presents the corresponding experimental accuracies for these strategies. The table presents no noticeable variance when comparing measurement schemes, the schemes usually differs by a maximum of 2%. The maximum difference is found in ECR + Transpiler, where it is approx. 6.2%.

6.2.1.2 11-qubit system without processing

In this subsection, we present a flow similar to that of the 6-qubit, main difference lies in the schemes for circuits as they become large. However, the pattern is well described, and the reader may find the circuits in the github repository of the publication [47].

The settings for the 11-qubit parallel discrimination task are:

- **CNOT + Transpiler:** The circuit is constructed using CNOT basis gates and is processed by the transpiler using optimization level 3.
- **ECR + Transpiler:** This strategy implements the circuit using ECR basis gates and processes it via the transpiler with optimization level 3.

- **ECR + Transpiler (ID):** A specialized variation where the measurement or state discrimination is tuned specifically for the Identity (ID) channel. This setting directly influences the transpiler’s output, leading to distinct final circuit properties compared to the generic ECR strategy.
- **ECR + Transpiler ($RZ(\phi)$):** A specialized variation where the measurement or state discrimination is tuned specifically for the $RZ(\phi)$ channel. Similar to the ID case, this specialized setting impacts the transpilation process, resulting in different circuit characteristics.
- **ECR (Topology-Aware) + Transpiler:** The circuit is initially designed to account for device connectivity using ECR gates, followed by transpilation at optimization level 3.
- **ECR (Topology-Aware) + Transpiler + Fixed Mapping:** This method utilizes a topology-aware ECR circuit processed with optimization level 3, while enforcing a fixed logical-to-physical qubit mapping.
- **ECR (Topology-Aware) + Fixed Mapping (No Opt.):** This approach uses a topology-aware ECR circuit with a fixed mapping and explicitly bypasses subsequent transpiler optimization passes.

Transpilation Strategy	Measurement	Depth	ECR	RZ	$\sqrt{X}$	X	ID/ $RZ(\phi)$
CNOT + Transpiler	Short	63	20	99	58	10	11
CNOT + Transpiler	XOR	65	29	108	52	18	11
ECR + Transpiler	Short	31	11	30	32	2	11
ECR + Transpiler (ID)	XOR	81	32	108	65	3	11
ECR + Transpiler ($RZ(\phi)$)	XOR	85	35	115	66	4	11
ECR (Topology-Aware) + Transpiler	Short	26	11	23	30	2	11
ECR (Topology-Aware) + Transpiler	XOR	70	29	108	62	6	11
ECR (Topology-Aware) + Transpiler + Fixed Mapping	Short	11	11	0	22	2	11
ECR (Topology-Aware) + Transpiler + Fixed Mapping	XOR	30	29	7	17	2	11
ECR (Topology-Aware) + Fixed Mapping (No Opt.)	Short	11	11	0	22	2	11
ECR (Topology-Aware) + Fixed Mapping (No Opt.)	XOR	22	30	0	12	2	11

Table 6.3: Comparison of quantum circuit complexity across various transpilation strategies and measurement methods. The table tracks circuit depth and gate counts for ECR, RZ , $\sqrt{X}$, X , and $ID/RZ(\phi)$ basis gates.

Tab. 6.3 presents the circuit complexity metrics for the 11-qubit system, comparing different transpilation strategies and measurement types. The amount of additionally introduced gates and depth by XOR measurement compared to Short is evident. For depth, this is seen for example in ECR (Topology-Aware) + Transpiler, where the difference is 44 in absolute terms. In relative terms the XOR measurement utilizes around three times more depth than its Short counterpart.

Tab. 6.4 presents the success ratios of discrimination. While in the 6-qubit setting the differences were marginal, in this case, the results show a significant trend when comparing measurement schemes. The XOR measurement significantly increases the performance of the circuit. Only in one

Transpilation Strategy	Short Accuracy	XOR Accuracy
CNOT + Transpiler	43.3%	48.5%
ECR + Transpiler	55.0%	54.5%
ECR (Topology-Aware) + Transpiler	36.1%	47.2%
ECR (Topology-Aware) + Transpiler + Fixed Mapping	32.0%	71.5%
ECR (Topology-Aware) + Fixed Mapping (No Opt.)	33.4%	71.8%

Table 6.4: Experimental accuracy results for the 11-qubit discrimination task. The table compares the performance of Short and XOR measurement schemes across selected transpilation strategies. Values represent the percentage of successful state discriminations, highlighting the impact of topology-aware and fixed mapping techniques on final readout fidelity.

of the five experiments was the performance comparable for both measurements, in the rest, the XOR dominates, sometimes achieving twice the accuracy of the Short one.

However, one noticeable pattern emerges. While for lower number of gates (Short measurement) we introduce additional error when trying to optimize it for hardware in terms of topology. Such phenomena does not occur for XOR measurement, which had higher depth and more gates. This could mean the transpiler is hitting its limits and for larger circuits the help from user may significantly boost the results.

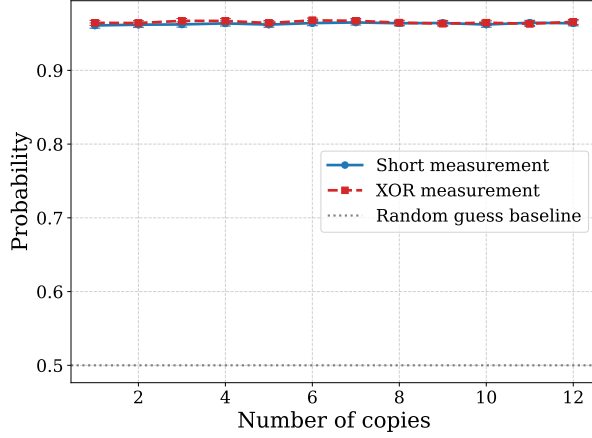
6.2.2 Experimental comparison of the different discrimination schemes

In this subsection, we present results from experiments run on IBM Brisbane. The results focus on comparing parallel, sequential, and hybrid schemes.

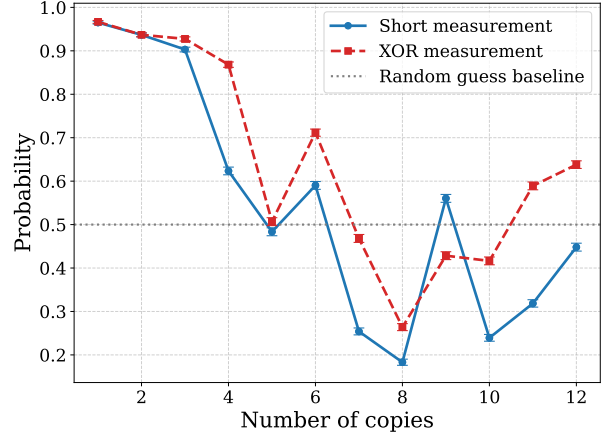
As illustrated in Fig. 6.13, the purely sequential scheme (Fig. 6.13 (a)) achieves a peak success probability of approximately 0.965. In contrast, the purely parallel scheme (Fig. 6.13b) exhibits significantly lower fidelity. The data indicates that the error overhead introduced by the entangling gates required for parallel distribution outweighs the decoherence accumulation of sequential protocols.

Expanding this analysis to hybrid schemes confirms the impact of circuit topology on error rates. We analyzed following number of copies: $N = 120$, $N = 240$, and $N = 1200$, which allowed for interpolation between sequential and parallel regimes (Fig. 6.14). The results demonstrate a consistent trend. Increasing the number of entangling gates leads to a systematic decrease in success probability.

Although the dashed red (XOR) and solid blue (Short) lines in Fig. 6.14 show distinct profiles, the probabilities of success show significant degradation from more width of the circuits. Consequently, optimization strategies yield the most substantial improvements when prioritizing entanglement reduction over readout modifications.



(a) Purely sequential scheme.

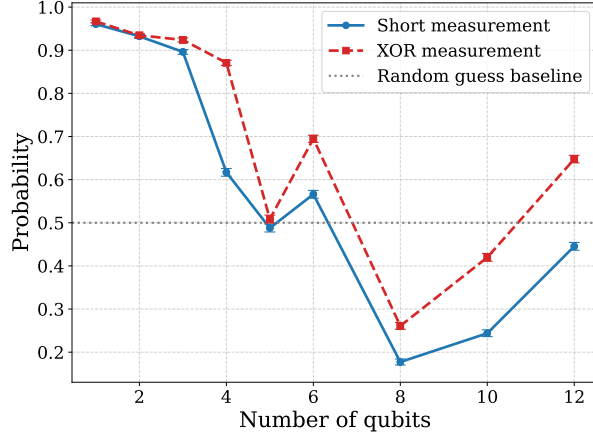


(b) Purely parallel scheme.

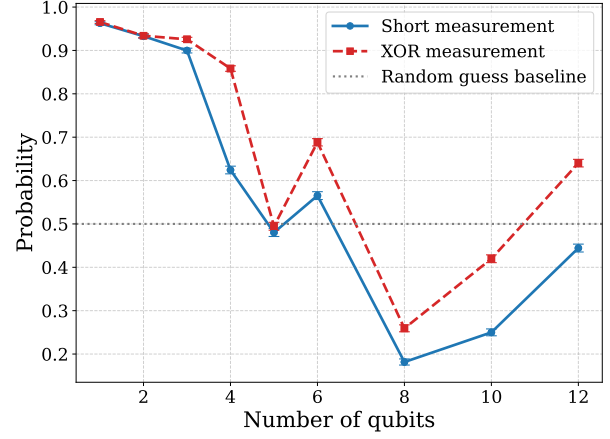
Figure 6.13: Probability of successful discrimination between the identity operation and the $RZ(\pi/N)$ gate, where N is the number of copies of the unknown unitary. The dashed red line denotes the XOR measurement strategy, while the solid blue line represents the short measurement. (a) Performance of the purely sequential scheme, peaking at approximately 0.965. (b) Purely parallel scheme, displaying reduced performance from entanglement overhead. Each circuit ran with 10,000 shots. Ambiguous measurement outcomes were randomly assigned to either gate.

Anomalous behavior discussion

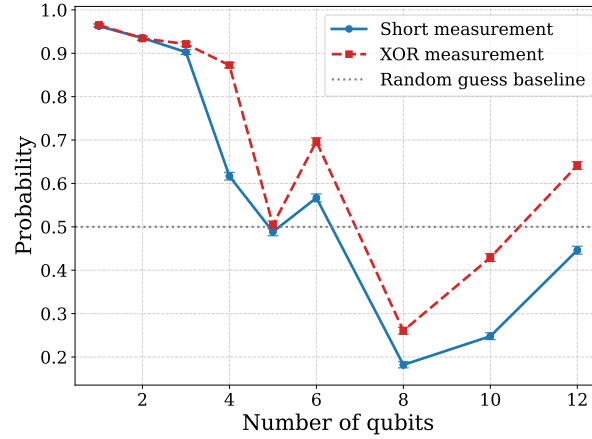
During our experiments on the Brisbane quantum device, we identified a consistent anomaly in measurements involving the entanglement of five or more qubits. This issue, characterized by random bit-flip errors appearing across all qubits simultaneously, is clearly visible in Fig. 6.13 and Fig. 6.14. Crucially, this behavior is independent of the measurement strategy employed. Unlike standard device noise, which typically pulls success probabilities toward the random baseline of 0.5, we observed a near-total inversion of results. This effectively corresponds to a swap of the anticipated answer sets, causing success rates to drop significantly below the 0.5 threshold. Although interchanging the target labels can restore the internal consistency of the data, this is not a valid mitigation strategy and is used here solely for illustrative purposes. We implemented several verification steps to ensure that the anomaly did not stem from our own software implementation. First, we analyzed experimental data collected across different dates. Second, the final transpiled circuits were verified using a noiseless simulator to confirm theoretical expectations. Third, we applied M3 error mitigation [49], which yielded no measurable improvement in the results. Because these errors persist regardless of the RZ gate angles or measurement methods, we hypothesize that they arise from a systematic hardware or software artifact specific to the Brisbane backend rather than from the algorithmic design.



(a) Hybrid scheme for $N = 120$.



(b) Hybrid scheme for $N = 240$.



(c) Hybrid scheme for $N = 1200$.

Figure 6.14: Successful discrimination probability between the identity operation and the $RZ(\pi/N)$ gate versus the hybrid scheme width. The dashed red line denotes the XOR measurement, and the solid blue line represents the short measurement. Panels show fixed total resources of (a) $N = 120$, (b) $N = 240$, and (c) $N = 1200$ copies. Performance degrades as scheme width increases, correlating with higher entanglement density. Circuits ran with 10,000 shots. Ambiguous outcomes were randomly assigned.

6.3 Experiments with processing

In this section, we present results for discrimination with processing. We do not include any circuits, as the pattern is simple and the description is enough. The experimental results for the discrimination between U and V on IBM Brisbane show clear trends. The results are compared with theoretical expectations and error-mitigated data using the MThree package [49]. We discuss two situations, the first is aimed at a lower number of copies, while the second focuses on a greater number of copies of the discriminated channels.

For a relatively small number of copies ($N = 4, 16, 32$), experiments utilizing purely sequential schemes (where width w is minimized and depth d is maximized) proved to be preferable, as shown in Fig. 6.15. The data suggest that the introduction of entangling gates required for parallel processing on real quantum hardware incurs a higher error overhead than the decoherence effects associated with the increased circuit depth of sequential protocols.

However, as the number of copies increases ($N = 64, 96, 1024$), the advantage shifts towards hybrid schemes. As illustrated in Fig. 6.16, utilizing a balanced w, d configuration allows for more precise results. This consistent trend indicates a threshold in gate composition: beyond a certain point, the error rates from decoherence and accumulative calibration imperfections in deep sequential circuits exceed the error rates introduced by creating entangled states in parallel schemes. In the extreme case of $N = 1024$, the standard schemes yielded effectively random results (see Fig. 6.16(c)). The purely sequential scheme suffered from excessive gate composition depth, while the fully parallel scheme required the preparation of a large GHZ state, both introducing too much noise to yield meaningful discrimination.

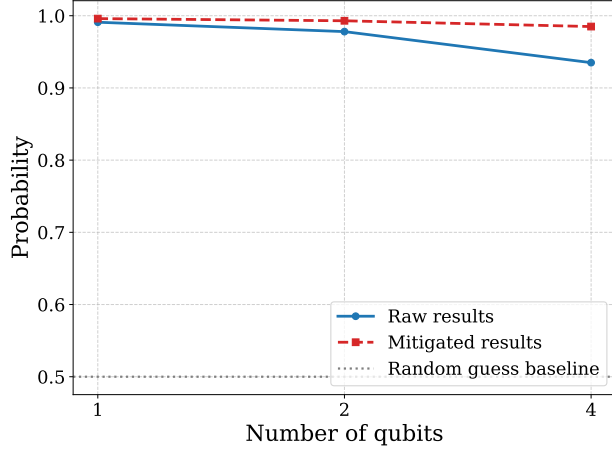
To address this, we explored a *suboptimal* circuit strategy. Instead of a fully entangled global scheme, we performed independent sequential experiments on w qubits, where each qubit processes a depth d of the unitary. Although theoretically suboptimal (as the total angle spread $\theta = \pi/w < \pi$), the probability of success can be boosted via majority voting.

- For $N = 1024$ using $w = 32$ and $d = 32$, this suboptimal procedure yielded a success probability of $P_s = 0.56765$, significantly outperforming the "optimal" schemes which failed due to noise.
- For $N = 96$ using $w = 3$ and $d = 32$, the suboptimal approach yielded $P_s = 0.74685$, surpassing the results shown in Fig. 6.16(b).

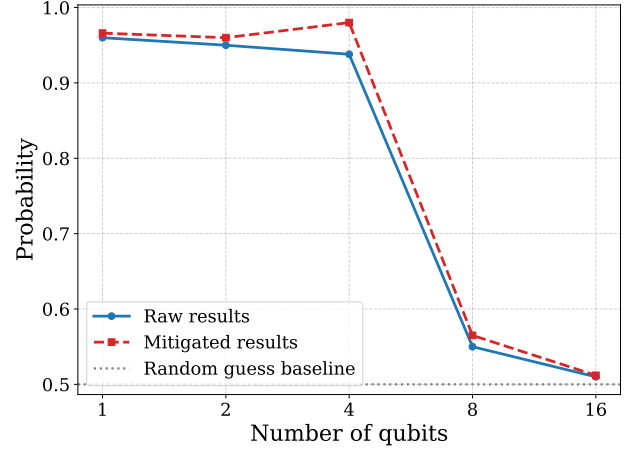
However, suboptimal circuits are not universally superior. For $N = 64$, the best hybrid result from Fig. 6.16(a) achieves approx. 0.85 and slightly outperforms the suboptimal procedure even in the absence of noise.

Finally, unlike previous experiments (see Experiment 1), no bit-flip anomalies were observed in these results despite comparable qubit counts and circuit depths. This stability suggests the

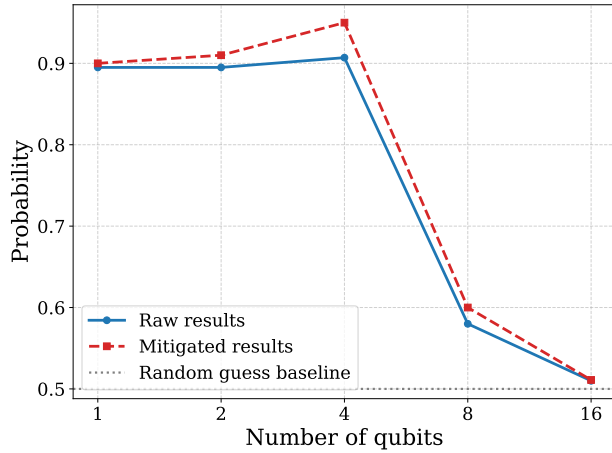
presence of undocumented backend-level optimizations or calibration updates to the IBM Brisbane execution pipeline or internal compilation procedures that perform low-level optimizations or calibration corrections.



(a) Hybrid scheme for $N = 4$.

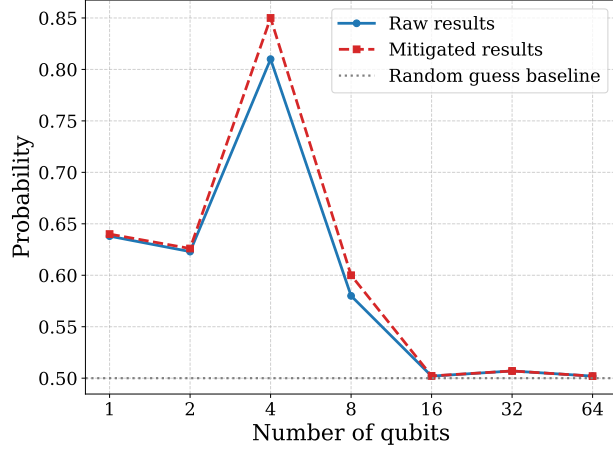


(b) Hybrid scheme for $N = 16$.

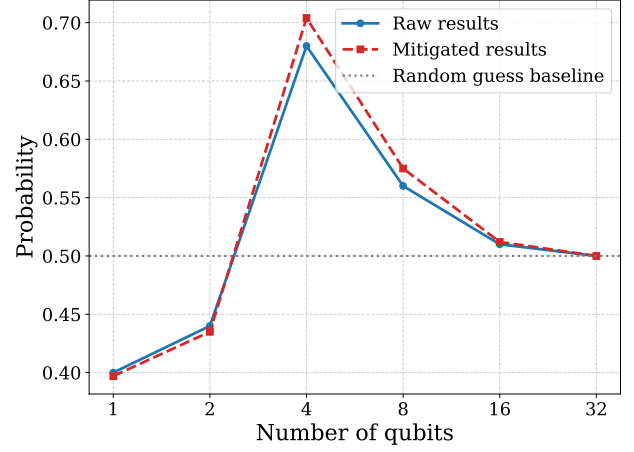


(c) Hybrid scheme for $N = 32$.

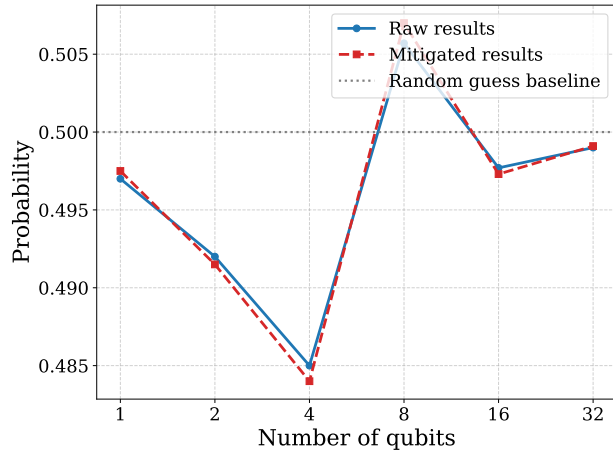
Figure 6.15: Successful discrimination probability between unitary operators $U = \sqrt{X}RZ(\frac{-\pi}{2N})\sqrt{X}$ and $V = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$ (N copies) versus the hybrid scheme width, using the short measurement. The blue line shows raw Brisbane results. The red line shows MThree error-mitigated results [49]. Panels display performance for total copies: (a) $N = 4$, (b) $N = 16$, and (c) $N = 32$. Each circuit ran with 10,000 shots.



(a) Hybrid scheme for $N = 64$.



(b) Hybrid scheme for $N = 96$.



(c) Hybrid scheme for $N = 1024$.

Figure 6.16: Successful discrimination probability between unitary operators $U = \sqrt{X}RZ(\frac{-\pi}{2N})\sqrt{X}$ and $V = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$ (N copies) versus the hybrid scheme width, using the short measurement. The blue line shows raw IBM Brisbane results; the red line shows MThree error-mitigated results [49]. Panels display performance for total copies: (a) $N = 64$, (b) $N = 96$, and (c) $N = 1024$. Each circuit ran with 10,000 shots.

Chapter 7

Conclusion

This thesis bridges the theoretical framework of multiple-shot unitary channel discrimination with its practical implementation on resource-constrained NISQ hardware. We investigate parallel, hybrid, and sequential strategies for perfect discrimination of unitary channels and map them onto the IBM Eagle R3 processor. The core of the theoretical contribution is Algorithm 1, a constructive framework to formulate optimal discriminators for unitary channels. By introducing an arbitrary phase angle into the state construction, this algorithm provides a crucial degree of freedom that allows the resulting quantum states to be analytically optimized for native hardware gate sets. For the restricted unitary setting considered in this thesis, the geometric solver avoids constructing the full SDP formulation and searches over a compressed set of eigenvalue classes.

Experimental results from 6-qubit and 11-qubit studies have demonstrated that the theoretical capabilities of current hardware are bottlenecked by the physical execution layer. The choice of transpilation technique - specifically, topology-aware manual mapping using XOR measurements versus automated transpilation, proved to be a decisive factor in success probability. Furthermore, our comparison of multiple-shot strategies revealed a clear hardware-specific complexity threshold with respect to the shape of the circuit. For a lower number of channel applications ($N \leq 32$), purely sequential, low-width circuits reliably outperformed their parallel counterparts due to the heavy error overhead introduced by ECR entangling gates. However, for deeper requirements ($N \geq 64$), the accumulative decoherence of deep sequential circuits forced a necessary shift toward hybrid rectangular configurations to balance depth and width.

Hardware evaluation also exposed critical systemic limitations. During the execution of multi-qubit circuits without processing, we observed a systematic anomaly on the IBM Brisbane device, characterized by simultaneous multi-qubit bit-flip errors. This resulted in a near-total inversion of the success probabilities, dropping the outcomes significantly below the 0.5 random baseline regardless of the measurement strategy. Additionally, scaling the total number of channel copies to $N = 1024$ in processed schemes identified a practical hardware limit where device noise completely overpowered the quantum state, reducing standard discrimination tasks to effectively random out-

comes. In these regimes, even M3 proved to be insufficient.

7.1 Future work

The findings and methodologies established in this thesis open three main avenues for future research. First, a more granular investigation of the IBM Brisbane processor and similar Eagle R3 architectures is required. Future studies should explore in depth the impact of hardware topology, for instance, by comparing three-way branching logic against the two-way mapping used in the oriented graph of Fig. 6.1. Second, the observed complexity threshold and the superiority of high-depth, low-width circuits must be validated across a wider range of hardware. Executing these experimental setups on different superconducting processors or trapped-ion systems will determine whether these findings are universal to NISQ devices or specific to the IBM transpilation pipeline. A third avenue involves further research of Algorithm 1, including optimized implementation, comparative performance against SDP, and exploration of alternative techniques for discriminator construction.

Bibliography

1. IBM QUANTUM. *Processor Types Guide* [online]. 2025. [visited on 2025-05-18]. Available from: <https://docs.quantum.ibm.com/guides/processor-types>.
2. IBM QUANTUM. *IBM Quantum Platform* [online]. 2026. [visited on 2026-04-25]. Available from: <https://quantum.cloud.ibm.com/>.
3. RIGETTI COMPUTING. *Rigetti Quantum Cloud Services (QCS)* [online]. 2026. [visited on 2026-04-25]. Available from: <https://qcs.rigetti.com/>.
4. IONQ. *IonQ Quantum Cloud* [online]. 2026. [visited on 2026-04-25]. Available from: <https://cloud.ionq.com/>.
5. VON NEUMANN, John. Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*. 1927, vol. 1927, pp. 245–272. Available also in English as: Probabilistic Structure of Quantum Mechanics.
6. BORN, Max. Zur Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik*. 1926, vol. 37, no. 12, pp. 863–867. Available from DOI: 10.1007/BF01397477.
7. HOLEVO, A. S. Statistical decision theory for quantum systems. *Journal of Multivariate Analysis*. 1973, vol. 3, no. 4, pp. 337–394. ISSN 0047-259X. Available from DOI: 10.1016/0047-259X(73)90028-6.
8. HELSTROM, C. W. Quantum detection and estimation theory. *Journal of Statistical Physics*. 1969, vol. 1, no. 2, pp. 231–252. ISSN 0022-4715. Available from DOI: 10.1007/BF01007479.
9. AUDENAERT, K. M. R.; CALSAMIGLIA, J.; MUÑOZ-TAPIA, R.; BAGAN, E.; MASANES, L.; ACIN, A. Discriminating States: The Quantum Chernoff Bound. *Physical Review Letters*. 2007, vol. 98, no. 16, p. 160501. Available from DOI: 10.1103/PhysRevLett.98.160501.
10. KITAEV, A. Y. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*. 1997, vol. 52, no. 6, p. 1191. Available from DOI: 10.1070/RM1997v052n06ABEH002155.
11. PUCHAŁA, Zbigniew; PAWELA, Łukasz; KRAWIEC, Aleksandra; KUKULSKI, Ryszard. Strategies for optimal single-shot discrimination of quantum measurements. *Physical Review A*. 2018, vol. 98, no. 4, p. 042103. Available from DOI: 10.1103/PhysRevA.98.042103.

12. PUCHAŁA, Zbigniew; PAWELA, Łukasz; KRAWIEC, Aleksandra; KUKULSKI, Ryszard; OSZMANIEC, Michał. Multiple-shot and unambiguous discrimination of von Neumann measurements. *Quantum*. 2021, vol. 5, p. 425. Available from DOI: 10.22331/q-2021-04-06-425.
13. JAŁOWIECKI, Konrad; LEWANDOWSKA, Paulina; PAWELA, Łukasz. PyQBench: A Python library for benchmarking gate-based quantum computers. *SoftwareX*. 2023, vol. 24, p. 101558. Available from DOI: 10.1016/j.softx.2023.101558.
14. TOEPLITZ, Otto. Das algebraische Analogon zu einem Satze von Fejér. *Mathematische Zeitschrift*. 1918, vol. 2, pp. 187–197. ISSN 0025-5874. Available from DOI: 10.1007/BF01212904.
15. HAUSDORFF, Felix. Der Wertvorrat einer Bilinearform. *Mathematische Zeitschrift*. 1919, vol. 3, pp. 314–316. ISSN 0025-5874. Available from DOI: 10.1007/BF01292610.
16. DUNKL, Charles F.; GAWRON, Piotr; HOLBROOK, John A.; MISZCZAK, Jarosław Adam; PUCHAŁA, Zbigniew; ŻYCZKOWSKI, Karol. Numerical shadow and geometry of quantum states. *Journal of Physics A: Mathematical and Theoretical*. 2011, vol. 44, no. 33, p. 335301. Available from DOI: 10.1088/1751-8113/44/33/335301.
17. PAWELA, Łukasz et al. *Numerical shadow* [online]. 2026. [visited on 2026-02-10]. Available from: <https://numericalshadow.org/>.
18. DUAN, Runyao; FENG, Yuan; YING, Mingsheng. Entanglement Is Not Necessary for Perfect Discrimination between Unitary Operations. *Physical Review Letters*. 2007, vol. 98, no. 10, p. 100503. ISSN 0031-9007. Available from DOI: 10.1103/PhysRevLett.98.100503.
19. ACIN, Antonio. Statistical distinguishability between unitary operations. *Physical Review Letters*. 2001, vol. 87, no. 17, p. 177901. ISSN 0031-9007. Available from DOI: 10.1103/PhysRevLett.87.177901.
20. LI, Chi-Kwong; LIU, Yue; MA, Chao; PELEJO, Diane Christine. *A Note on Parallel Distinguishability of Two Quantum Operations* [online]. 2020. [visited on 2026-03-06]. Available from arXiv: 1906.06209v4.
21. CHIRIBELLA, G.; D'ARIANO, G. M.; PERINOTTI, P. Quantum Circuit Architecture. *Physical Review Letters*. 2008, vol. 101, no. 6, p. 060401. Available from DOI: 10.1103/PhysRevLett.101.060401.
22. LEWANDOWSKA, Paulina; PAWELA, Łukasz; PUCHAŁA, Zbigniew. Strategies for single-shot discrimination of process matrices. *Scientific Reports*. 2023, vol. 13, p. 3046. Available from DOI: 10.1038/s41598-023-30191-0.
23. HARROW, Aram W; HASSIDIM, Avinatan; LEUNG, Debbie W; WATROUS, John. Adaptive versus nonadaptive strategies for quantum channel discrimination. *Physical Review A*. 2010, vol. 81, no. 3, p. 032339. Available from DOI: 10.1103/PhysRevA.81.032339.

24. BAVARESCO, Jessica; MURAO, Mio; QUINTINO, Marco Túlio. Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination. *Physical Review Letters*. 2021, vol. 127, no. 20, p. 200504. Available from DOI: 10.1103/PhysRevLett.127.200504.
25. BÍLEK, Adam; HLISNIKOVSKÝ, Jan; BEZDĚK, Tomáš; KUKULSKI, Ryszard; LEWANDOWSKA, Paulina. Experimental study of multiple-shot unitary channels discrimination using the IBM Q computers. *Scientific Reports*. 2026, vol. 16, p. 6142. Available from DOI: 10.1038/s41598-025-31665-z.
26. CARATHÉODORY, Constantin. Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen. *Rendiconti del Circolo Matematico di Palermo*. 1911, vol. 32, pp. 193–217. Available from DOI: 10.1007/BF03014795.
27. GOOGLE. *A new era of intelligence with Gemini 3* [online]. 2025. [visited on 2026-02-10]. Available from: <https://blog.google/products/gemini/gemini-3>.
28. WATROUS, John. *The Theory of Quantum Information*. Cambridge: Cambridge University Press, 2018. ISBN 9781107180567. Available from DOI: 10.1017/9781316848142.
29. DIRAC, Paul Adrien Maurice. *The Principles of Quantum Mechanics*. 4th ed. Oxford: Clarendon Press, 1958. ISBN 978-0-19-852011-5.
30. RIESZ, Frigyes. Sur les opérations fonctionnelles linéaires. *Comptes Rendus de l'Académie des Sciences*. 1909, vol. 149, pp. 974–977.
31. REED, Michael; SIMON, Barry. *Methods of Modern Mathematical Physics: Functional Analysis*. Revised and enlarged edition. San Diego: Academic Press, 1980. ISBN 978-0-12-585050-6.
32. DOSTÁL, Zdeněk; VONDRÁK, Vít. *Lineární algebra* [online]. Ostrava: Vysoká škola báňská – Technická univerzita Ostrava a Západočeská univerzita v Plzni, 2012 [visited on 2026-03-23]. Available from: https://mi21.vsb.cz/sites/mi21.vsb.cz/files/unit/linearni_algebra.pdf. Version 24.4.2012.
33. BOUCHALA, Jiří. *Úvod do funkcionální analýzy* [online]. Ostrava: VŠB – Technická univerzita Ostrava, 2012 [visited on 2026-03-23]. Available from: https://mi21.vsb.cz/sites/mi21.vsb.cz/files/unit/uvod_do_funkcionalni_analyzy.pdf.
34. NIELSEN, Michael A.; KNILL, Emanuel; LAFLAMME, Raymond. Complete quantum teleportation using nuclear magnetic resonance. *Nature*. 1998, vol. 396, pp. 52–55. Available from DOI: 10.1038/23891.
35. WATROUS, John. Semidefinite programs for completely bounded norms. *Theory of Computing*. 2009, vol. 5, no. 11, pp. 217–238. Available from DOI: 10.4086/toc.2009.v005a011.

36. WATROUS, John. Simpler semidefinite programs for completely bounded norms. *Chicago Journal of Theoretical Computer Science*. 2013, vol. 2013, no. 8, pp. 1–19. Available from DOI: 10.4086/cjtcs.2013.008.
37. PERES, Asher. Separability Criterion for Density Matrices. *Physical Review Letters*. 1996, vol. 77, no. 8, pp. 1413–1415. ISSN 0031-9007. Available from DOI: 10.1103/PhysRevLett.77.1413.
38. HORODECKI, Michał; HORODECKI, Paweł; HORODECKI, Ryszard. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*. 1996, vol. 223, no. 1-2, pp. 1–8. ISSN 0375-9601. Available from DOI: 10.1016/S0375-9601(96)00706-2.
39. WATROUS, John. *Channel Distinguishability and the Completely Bounded Trace Norm* [online]. University of Waterloo, 2011. [visited on 2026-01-05]. Available from: <https://cs.uwaterloo.ca/~watrous/TQI-notes/TQI-notes.20.pdf>. Lecture 20, CS 766 / QIC 820 – Theory of Quantum Information.
40. D’ARIANO, Giacomo Mauro; LO PRESTI, Paolillo; PARIS, Matteo G. A. Using entanglement improves the precision of quantum measurements. *Physical Review Letters*. 2001, vol. 87, no. 27, p. 270404. ISSN 0031-9007. Available from DOI: 10.1103/PhysRevLett.87.270404.
41. VANDENBERGHE, Lieven; BOYD, Stephen. Semidefinite Programming. *SIAM Review*. 1996, vol. 38, no. 1, pp. 49–95. Available from DOI: 10.1137/1038003.
42. DUAN, Runyao; GUO, Cheng; LI, Chi-Kwong; LI, Yinan. Parallel distinguishability of quantum operations. In: *2016 IEEE International Symposium on Information Theory (ISIT)*. Piscataway, NJ: IEEE, 2016, pp. 2259–2263. Available from DOI: 10.1109/ISIT.2016.7541701.
43. HORN, Roger A; JOHNSON, Charles R. *Matrix Analysis*. Cambridge: Cambridge University Press, 1985. ISBN 978-0-521-30586-0.
44. BURDEN, Richard L; FAIRES, J Douglas. *Numerical Analysis*. Boston, MA: Cengage Learning, 2011. ISBN 978-0-538-73564-3.
45. KNUTH, Donald E. *The Art of Computer Programming, Volume 2: Seminumerical Algorithms*. Reading, Mass.: Addison-Wesley Professional, 1997. ISBN 978-0-201-89684-8.
46. CORMEN, Thomas H; LEISERSON, Charles E; RIVEST, Ronald L; STEIN, Clifford. *Introduction to Algorithms*. Cambridge, Mass.: MIT press, 2022. ISBN 978-0-262-04630-5.
47. BÍLEK, Adam; HLISNIKOVSKÝ, Jan; BEZDĚK, Tomáš; KUKULSKI, Ryszard; LEWANDOWSKA, Paulina. *experimental_study_of_multiple_shot_channel_discrimination* [GitHub repository]. 2025. [visited on 2026-02-10]. Available from: https://github.com/Dotnester/experimental_study_of_multiple_shot_channel_discrimination.

48. IBM QUANTUM. *A new Eagle in the Poughkeepsie Quantum Datacenter: IBM Quantum's most performant system yet* [online]. 2022. [visited on 2026-01-20]. Available from: <https://www.ibm.com/quantum/blog/eagle-quantum-error-mitigation>.
49. MTHREE DEVELOPERS. *mthree.M3Mitigation Documentation* [online]. 2025. [visited on 2025-05-18]. Available from: <https://qiskit.github.io/qiskit-addon-mthree/stubs/mthree.M3Mitigation>.

Appendix A

Publication

The following article is included to provide the reader with complete documentation of the thesis evolution.



OPEN Experimental study of multiple-shot unitary channels discrimination using the IBM Q computers

Adam Bílek^{1,2}, Jan Hlisenikovsky^{1,2}✉, Tomáš Bezdek^{1,3}, Ryszard Kukulski² & Paulina Lewandowska²

Tasks involving black boxes appear frequently in the theory of quantum information, with quantum channel discrimination as a central example that has been deeply studied. In this work, we experimentally study the discrimination between two unitary quantum channels in the multiple-shot scenario. We challenge the theoretical results concerning the probability of correct discrimination with the results collected from experiments performed on the IBM Brisbane. Our analysis shows that neither too deep quantum circuits nor circuits that create too much entanglement are suitable for the discrimination task. We conclude that circuit architectures which minimize entanglement overhead while preserving discrimination power are significantly more resilient to hardware noise if their depth does not exceed a threshold value. Consequently, our findings necessitate a paradigm shift: for execution on noisy hardware, the theoretically suboptimal circuit is, counterintuitively, often the superior choice.

Keywords Quantum computing, Quantum networks, Quantum channel Discrimination, Benchmarking of NISQ devices

In the last decade, quantum computing has become a reality. Quantum algorithms of increasing complexity are being implemented on progressively more advanced quantum devices. In effect, high-quality solutions to some real-world problems are expected to arrive soon. This situation motivates the need for the certification and benchmarking of various quantum devices^{1–3}. The discrimination task of quantum operators constitutes one of the certification methods and quality metrics for benchmarking quantum architectures^{4,5}. The theoretical background of the quantum discrimination task has been widely developed. The primary task of discrimination involves a one-shot scenario of discrimination between quantum operators. We can imagine an unknown quantum device, a black box. The only information we have is that it performs one of two quantum operators, say $\mathcal{T}$ and $\mathcal{S}$. Our goal is two-fold. First, we want to determine the highest possible probability of correct guessing. Secondly, we need to devise an optimal strategy that maximizes the probability of success.

The problem of single-shot discrimination of quantum states was solved analytically by Helstrom a few decades ago in^{6,7}. The authors calculated the probability of correct discrimination between two quantum states using the notion of a trace norm. Next, there are many modifications of the origin problem for quantum states that were also considered^{8–12}. For the discrimination tasks of quantum channels⁷, the probability of success of the discrimination can be formulated by the diamond norm^{13,14}, which can be computed by semidefinite programming^{13,15}. However, as the dimension of quantum channels increases, computing the diamond norm becomes inefficient¹⁶. In general, for quantum channels, entanglement is necessary for optimal discrimination⁷. The exception is, for example, the discrimination task between unitary channels^{7,17}. Furthermore, the probability of correct discrimination between two unitary channels can be expressed in the notion of the numerical range¹⁸. Discrimination tasks for general quantum measurements, von Neumann measurements, or SIC POVMs were also considered in the literature^{19–24}. Lastly, the theory of an indefinite causal structure is one of the attractive

¹Department of Applied Mathematics, Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, 17. Listopadu 2172/15, 70833 Ostrava, Czech Republic. ²IT4Innovations, VSB-Technical University of Ostrava, 17. Listopadu 2172/15, 70833 Ostrava, Czech Republic. ³Department of Mathematics, TUM School of Computation, Information and Technology, Technical University of Munich, Boltzmannstr. 3/III, 85748 Garching bei München, Germany. ✉email: jan.hlisenikovsky@live.com

topics. This approach uses the notion of process matrices^{25,26}, which can be seen as a generalization of quantum combs. The single-shot discrimination of the process matrices was introduced²⁷.

What if we have multiple copies of quantum operators? The most general discrimination approach is known as an adaptive strategy²⁸. This strategy can be described by using the term quantum combs. Here, we assume freedom in choosing the setting of quantum operations and any processing between them. In addition, numerical investigation of quantum channel discrimination showed that using indefinite causal order, we can achieve greater success with discrimination^{29,30}. However, determining an optimal strategy to realize such tasks is a significant challenge in practice. This is why we often limit ourselves to parallel³¹ and sequential³² schemes. The parallel scheme assumes N copies of a given quantum operation distributed over N entangled systems and final measurement with no additional processing, whereas the sequential scheme assumes the implementation of N copies of the quantum operation on one system, allowing additional processing between them.

One of the first theoretical results was the study of discrimination of multipartite unitary operations²⁹. The authors showed that perfect discrimination can be achieved using a parallel scheme³¹. It is also possible to design a perfect discrimination scheme of unitary channels without introducing entanglement with an auxiliary system³³. It can be done by using the sequential approach and specific processing at the end of the circuit. On the other hand, for general quantum channels, neither the parallel nor the sequential strategy guarantees an optimal solution^{34,35}. The advantage of adaptive strategies became apparent for a task concerning entanglement-breaking channels³⁴. Adaptive discrimination scenarios were also investigated in^{22,32,36}. In Ref. ³⁷, the authors have formulated the necessary and sufficient conditions under which quantum channels can be perfectly discriminated, while in Ref. ²², the authors have formulated conditions for the perfect discrimination of two measurements. Furthermore, in³⁸, the authors have shown that any possible adaptive method does not offer any advantage over the parallel scheme for von Neumann measurements.

One of the first experimental results for the parallel and sequential discrimination scheme of unitary channels was presented in Ref. ³⁹. In the era of NISQ devices, we need to take into account certain limitations. Due to the existing decoherence, sequential schemes could not be practical for larger numbers of copies. At the same time, parallel schemes for NISQ architectures are also not possible to implement because of the limited number of qubits. Perhaps intermediate schemes, called sequentially-paralleled, may overcome both obstacles. This scheme assumes that we have $N = w \cdot d$ copies of a quantum operator, where w and d are natural numbers. The sequentially-paralleled scheme of width w and depth d consists of d applications of w copies of the quantum operator applied simultaneously to the quantum state. The initial state ρ evolves through these layers, combining parallelism within each layer and a sequential structure across layers. In general, one could consider incorporating additional intermediate processing operations between layers.

This raises the critical question of whether simpler, theoretically suboptimal circuits might prove more effective in practice. To address this, we utilize a simulator calibrated to the specific error profile of the IBM Brisbane device, providing a high-fidelity environment to test and compare the resilience of different strategies.

In this work, we present a comprehensive study of various scenarios of multiple-shot discrimination of quantum unitary channels. We consider parallel, sequential, and sequentially-paralleled quantum networks for discrimination between two qubit unitary channels. We challenge the theoretical results concerning the probability of correct discrimination with the results collected from experiments performed on the IBM Brisbane. Based on several examples, our analysis shows that neither excessively deep circuits nor those creating too much entanglement are suitable for the discrimination task.

Paper organization

This paper is organized as follows. **Section Mathematical Preliminaries** (page 2) recalls the notation and mathematical tools necessary for our research. We then formalize channel discrimination in single- and multiple-shot settings, including perfect-discrimination criteria as well as corresponding statistical methods for analysis. At page 4, we present **Experiment 1** of discrimination with no processing between Φ_1 and $\Phi_{RZ(\phi)}$: methodology (discriminator and two measurement schemes), device-aware transpilation (CNOT/ECR and mapping), and hardware results on IBM Brisbane. At page 8, we present **Experiment 2** of discrimination with processing between $U = \sqrt{X} RZ(-\pi/2N) \sqrt{X}$ and $V = \sqrt{X} RZ(\pi/2N) \sqrt{X}$, including hardware results on IBM Brisbane. **Section Noise Modeling and Ablation Analysis** (page 10) presents an analytic depolarizing model and calibrated Aer ablations (single-/two-qubit noise, T_1/T_2 , readout, and a small coherent RZ drift), and relates these to the observed hardware trends. Finally, the **conclusions** are presented in page 11.

Mathematical preliminaries

Let $\mathcal{X}$ be a complex Euclidean space. Then $L(\mathcal{X})$ denotes the collection of all linear mappings of the form $A : \mathcal{X} \rightarrow \mathcal{X}$. An operator $X \in L(\mathcal{X})$ is positive semi-definite if $\langle x | X | x \rangle \geq 0$ for all $|x\rangle \in \mathcal{X}$. The set of all such operators is written as $\text{Pos}(\mathcal{X})$. By $\Omega(\mathcal{X})$ we denote the set of quantum states $\rho \in \text{Pos}(\mathcal{X})$ such that $\text{Tr}(\rho) = 1$. Let $X \in L(\mathcal{X})$. We denote by $\text{spec}(X)$ the set of all eigenvalues of X .

The set of all linear mappings from $L(\mathcal{X})$ to $L(\mathcal{Y})$, $\Phi : L(\mathcal{X}) \rightarrow L(\mathcal{Y})$, will be denoted as $T(\mathcal{X}, \mathcal{Y})$. The tensor product of linear maps (or operators) Φ and Ψ will be denoted as $\Phi \otimes \Psi$. A linear map $\Phi \in T(\mathcal{X}, \mathcal{Y})$ is positive if it holds that $\Phi(P) \in \text{Pos}(\mathcal{Y})$ for all $P \in \text{Pos}(\mathcal{X})$, whereas Φ is a completely positive map (CP) if $\Phi \otimes \mathbb{1}_{L(\mathcal{Z})}$ is a positive map for every complex Euclidean space $\mathcal{Z}$. We say Φ is trace-preserving (TP) if it holds that $\text{Tr}(\Phi(X)) = \text{Tr}(X)$ for all $X \in L(\mathcal{X})$. A linear map $\Phi \in T(\mathcal{X}, \mathcal{Y})$, which is a CPTP map, is called a quantum channel. The collection of all quantum channels is denoted as $C(\mathcal{X}, \mathcal{Y})$ (with a shorthand $C(\mathcal{X}) := C(\mathcal{X}, \mathcal{X})$). Next, we distinguish a special subset of quantum channels known as unitary channels, $\Phi_U \in C(\mathcal{X})$, defined as $\Phi_U(X) = UXU^\dagger$ where $U \in L(\mathcal{X})$ is the unitary matrix.

A Positive Operator-Valued Measure (POVM) $\mathcal{P}$ is a collection of operators, the so-called effects $\{E_0, \dots, E_n\} \subset \text{Pos}(\mathcal{X})$ with the property of $\sum_{i=0}^n E_i = \mathbb{1}$. According to the Born rule for a given quantum state ρ the probability of obtaining the result E_i is given by $\text{Tr}(E_i \rho)$.

A useful tool for studying the discrimination of unitary channels is the concept of the numerical range of an operator. For $X \in \text{L}(\mathcal{X})$ we define the numerical range of X as the set

$$W(X) := \{\langle x | X | x \rangle : |x\rangle \in \mathcal{X}, \langle x | x \rangle = 1\}. \quad (1)$$

The Hausdorff-Töplitz Theorem^{40,41} states that $W(X)$ is a convex set. If X is normal, the numerical range is a convex hull of its eigenvalues. For unitary matrices U we define the arc function $\theta(U)$ as the length of the smallest arc on the unit circle that contains all the eigenvalues of the unitary operator U . Mathematically, this can be expressed as

$$\theta(U) := \min \left\{ \Delta \in [0, 2\pi) : \exists \alpha \in [0, 2\pi) \text{ such that } \text{spec}(U) \subset \{e^{i\theta} : \theta \in [\alpha, \alpha + \Delta]\} \right\}. \quad (2)$$

Lastly, let us introduce the diamond norm in the space $\text{T}(\mathcal{X}, \mathcal{Y})$. For $\Phi \in \text{T}(\mathcal{X}, \mathcal{Y})$ it is defined as

$$\|\Phi\|_{\diamond} = \|\Phi \otimes \mathbb{1}_{\text{L}(\mathcal{X})}\|_1, \quad (3)$$

where $\|\Phi\|_1 = \max\{\|\Phi(X)\|_1 : X \in \text{L}(\mathcal{X}), \|X\|_1 \leq 1\}$ and $\|Y\|_1$ is Schatten 1-norm of $Y \in \text{L}(\mathcal{Y})$.

Discrimination of quantum channels

We will consider the following scenario of quantum channel discrimination. Suppose that we have a classical description of two quantum channels $\Phi_0, \Phi_1 \in \text{C}(\mathcal{X}, \mathcal{Y})$ and a black box implementing an unknown channel Φ , which either is Φ_0 or Φ_1 . We would like to determine whether $\Phi = \Phi_0$ or $\Phi = \Phi_1$ have been hidden in the black box. To reveal the value of Φ , we construct a quantum experiment consisting of an initial state $\rho \in \Omega(\mathcal{X} \otimes \mathcal{Z})$ and a binary measurement $\{E_0, E_1\} \subset \text{Pos}(\mathcal{Y} \otimes \mathcal{Z})$. The channel $\Phi \otimes \mathbb{1}_{\text{L}(\mathcal{Z})}$ is applied to ρ and then the output state $(\Phi \otimes \mathbb{1}_{\text{L}(\mathcal{Z})})(\rho)$ is measured by $\{E_0, E_1\}$. The measurement label defines our guess about the hidden value of Φ (the label 0 associated with the effect E_0 indicates $\Phi = \Phi_0$). In such a setup, the probability of successful channel discrimination p_{succ} is given by

$$p_{\text{succ}} = \frac{1}{2} \text{Tr}(E_0(\Phi_0 \otimes \mathbb{1}_{\text{L}(\mathcal{Z})})(\rho)) + \frac{1}{2} \text{Tr}(E_1(\Phi_1 \otimes \mathbb{1}_{\text{L}(\mathcal{Z})})(\rho)). \quad (4)$$

The goal of our task is to construct ρ and $\{E_0, E_1\}$ that maximize p_{succ} . The auxiliary system $\mathcal{Z}$ is of arbitrary size and provides a resource to increase the probability of success. From the Holevo-Helstrom theorem^{6,42} we know the probability of successful channel discrimination can be expressed in terms of the diamond norm

$$p_{\text{succ}} = \frac{1}{2} + \frac{1}{4} \|\Phi_0 - \Phi_1\|_{\diamond}. \quad (5)$$

Statistical probability of discrimination

We report the probability of successful discrimination $p_{\text{succ}} = \frac{\# \text{correct}}{\text{shots}}$ for each circuit-backend pair under equal priors. Ambiguous outcomes (equal distance to both hypotheses) were resolved by a fair coin flip, as stated in the figure captions, so every run yielded a definite label.

To estimate confidence intervals for the probability of success in experiments, we use the method of Clopper and Pearson⁴³ with a 0.99 confidence level. Due to the large number of shots per discrimination task, the confidence intervals are very narrow and centered around the estimated p_{succ} values.

Single-shot discrimination of quantum unitary channels

It is worth emphasizing at this point that there is no need to use an auxiliary system $\mathcal{Z}$ for the discrimination of quantum unitary channels³⁰. Considering the task of single-shot discrimination between unitary channels Φ_U, Φ_V , the diamond norm between such channels can be expressed using the notion of the numerical range as^{7,44}

$$\|\Phi_U - \Phi_V\|_{\diamond} = 2\sqrt{1 - \nu^2}, \quad (6)$$

where $\nu = \min_{w \in W(V^\dagger U)} |w|$. From the above proposition, it follows that unitary channels Φ_U and Φ_V are perfectly distinguishable if and only if $0 \in W(V^\dagger U)$. The above can also be formulated as there exists a density matrix σ such that $\text{tr}(V^\dagger U \sigma) = 0$.

Using the results of^{33,45,46} we can also calculate ν using the arc function $\theta(V^\dagger U)$ defined in Eq. (2), determining the condition for perfect discrimination between U and V in the single-shot scenario. If we define $\theta := \theta(V^\dagger U) \leq \pi$ is the minimal covering arc, then $\nu = \cos(\theta/2)$ and therefore $\|\Phi_U - \Phi_V\|_{\diamond} = 2 \sin(\theta/2)$. Furthermore, we achieve perfect discrimination for unitary channels if and only if

$$\theta(V^\dagger U) \geq \pi. \quad (7)$$

Multiple-shot discrimination of unitary channels

When multiple uses of the unitary channel are available, perfect discrimination may be achieved even if the single-shot condition $\theta(V^\dagger U) \geq \pi$ does not hold⁴⁵. The theoretically best strategy in that case is a parallel scheme which consists of N copies of the unitary U applied simultaneously to N quantum registers. In that case, the problem is to discriminate identity $U^{\otimes N}$ and unitary matrix $V^{\otimes N}$. Since the arc function satisfies the scaling relation $\theta((V^{\otimes N})^\dagger U^{\otimes N}) = N\theta(V^\dagger U)$, for $N\theta(V^\dagger U) < 2\pi$, then the condition for perfect discrimination is $N\theta(V^\dagger U) \geq \pi$ ³³. This leads to the condition on the minimum number of copies required for perfect discrimination is given by

$$N \geq \left\lceil \frac{\pi}{\theta(V^\dagger U)} \right\rceil. \quad (8)$$

Thus, even when $\theta(V^\dagger U) < \pi$, as long as $\theta(V^\dagger U) > 0$, perfect discrimination can still be achieved by using a sufficient number of copies of U and V .

Experimental set-up

A wide range of multi-shot discrimination strategies has been explored in the literature. In this work, we restrict our attention to three representative classes: parallel schemes, sequential schemes, and rectangular hybrid schemes that interpolate between these two. Experimental comparison of the performance of these strategies constitutes the main contribution of our work. Our focus is on the discrimination of qubit unitary channels U and V . For simplicity we will consider only the cases that satisfy

$$\theta(V^\dagger U) = \frac{\pi}{N}. \quad (9)$$

Let $N = wd$ for $w, d \in \mathbb{N}$. In our set-up, N copies of the black box Φ are arranged in a rectangular layout over w qubits with circuit depth d . On each qubit, the unknown operation is applied d times, namely, we have

$$\Phi^{\otimes w} \Phi_{X_{d-1}}^{\otimes w} \Phi^{\otimes w} \cdots \Phi_{X_1}^{\otimes w} \Phi^{\otimes w}, \quad (10)$$

where $X_1, \dots, X_{d-1}$ are arbitrary unitary matrices defined on w qubits, which are responsible for additional processing. In special cases, when $w = N$ and $d = 1$, we obtain the parallel scheme $\Phi^{\otimes N}$ (no processing is needed) and when $w = 1$ and $d = N$, we obtain the sequential scheme $\Phi \Phi_{X_{d-1}} \cdots \Phi_{X_1} \Phi$.

$$\theta(((V^\dagger U)^d)^{\otimes w}) = w\theta((V^\dagger U)^d) = wd\theta(V^\dagger U) = N\frac{\pi}{N} = \pi. \quad (11)$$

The only remaining question is how to find optimal ρ and $\{E_0, E_1\}$. The precise form of these variables depends on U, V ,^{7,13,44} and will be calculated later for each pair of unitary channels considered.

Experiment 1—Discrimination of unitary channels on IBM Brisbane without processing

This section is divided into four main parts. At the beginning, we will discuss in detail the components of discrimination schemes and later on their decompositions into native gates. Next, we will talk about transpilation approaches, and finally we will present the results obtained from the IBM Quantum device. All experiments were executed on the IBM Brisbane. Throughout, we use IBM Quantum for the platform and IBM Brisbane for the device.

Methodology and setup

In this example, we will distinguish between identity Φ_1 and $\Phi_{\text{RZ}(\phi)}$ for $\text{RZ}(\phi) = \begin{pmatrix} e^{-i\frac{\phi}{2}} & 0 \\ 0 & e^{i\frac{\phi}{2}} \end{pmatrix}$ without processing between the particular application of the unitary channel ($X_i = \mathbb{1}$ for each $i = 1, \dots, d$). Let $N = wd$ for $w, d \in \mathbb{N}$. In our setup N copies of the black box Φ can be composed in the rectangular shape spread on w qubits with the depth of the circuit d . On each qubit the unknown operation is composed d times, namely, we have

$$\underbrace{\Phi^d \otimes \cdots \otimes \Phi^d}_w. \quad (12)$$

To determine a discriminator $|\psi\rangle$, we need to find a unit vector satisfying

$$\langle \psi | \text{RZ}(d\phi)^{\otimes w} | \psi \rangle = 0. \quad (13)$$

The most distant pair of eigenvalues of $\text{RZ}(d\phi)^{\otimes w}$ are $-i$ and i , corresponding to eigenvectors $|0 \cdots 0\rangle$ and $|1 \cdots 1\rangle$, respectively. Hence, we can show that

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0 \cdots 0\rangle + \lambda |1 \cdots 1\rangle) \in \mathbb{C}^{2^w}, \quad (14)$$

for any unit number $\lambda \in \mathbb{C}$. If $\Phi = \Phi_1$, the output state is equal to $|\psi_0\rangle = |\psi\rangle$. Otherwise, if $\Phi = \Phi_{\text{RZ}(\phi)}$, then the output states are equal to $|\psi_1\rangle = \text{RZ}(\phi)^{\otimes w} |\psi\rangle = \frac{-i}{\sqrt{2}} (|0 \cdots 0\rangle - \lambda |1 \cdots 1\rangle)$. As $|\psi_0\rangle$ and $|\psi_1\rangle$ are orthogonal, we can find $E_0 \geq |\psi_0\rangle\langle\psi_0|$ and $E_1 \geq |\psi_1\rangle\langle\psi_1|$ (in particular $E_0 = |\psi_0\rangle\langle\psi_0|$ and $E_1 = \mathbb{1} - E_0$), which guarantees $p_{\text{succ}} = \frac{1}{2} \text{Tr}(E_0(\Phi_1^{\otimes w})(|\psi\rangle\langle\psi|)) + \frac{1}{2} \text{Tr}(E_1(\Phi_{\text{RZ}(\phi)}^{\otimes w})(|\psi\rangle\langle\psi|)) = \frac{1}{2} \text{Tr}(E_0 |\psi_0\rangle\langle\psi_0|) + \frac{1}{2} \text{Tr}(E_1 |\psi_1\rangle\langle\psi_1|) = 1$

Components of the discrimination scheme and decompositions

We divide the implementation of our discrimination circuit into two distinct components: the discriminator and the measurement. The unknown unitary gates, representing the quantum channel to be identified, are inserted between these two parts. In practical experiments, we focus specifically on schemes that in theory achieve perfect discrimination between the identity operation and the $\text{RZ}(\phi)$ gate. To ensure that the condition previously discussed is met, we set the angle ϕ to π/N , where N denotes the total number of uses of the unknown gate in the circuit. This value is determined by the product of the width (the number of qubits used in parallel) and depth (the number of successive applications of the unknown gate).

After using the discriminator, the circuit must be in a maximally entangled state on N qubits in the form

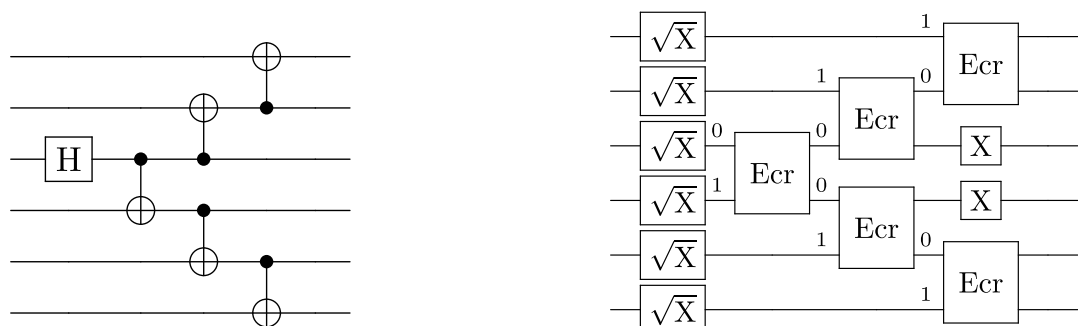
$$|\psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle^{\otimes N} + \alpha |1\rangle^{\otimes N}). \quad (15)$$

To achieve this, we used a cascade of CNOT gates, as can be seen for six qubits in Fig. 1a. This discriminator is created based on the standard pattern that is commonly used to create the GHZ state⁴⁷.

The experiments were carried out on IBM Brisbane, where the CNOT gate is not a native gate. To reduce the number of gates in the circuit, we prepared the discriminator specially designed for Eagle R3 architecture⁴⁸. Then, we use ECR gates to create entanglement between qubits, while omitting CNOT gates unrolling during transpilation at the same time. The single-qubit gate $\sqrt{X}$ (SX) has matrix form $\sqrt{X} = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}$. The first part of the discriminator consists of SX gates on all qubits and then a cascade of ECR gates of similar structure as the CNOT cascade in the first case. Unlike the discriminator based on CNOT gates, we also had to add several X gates to the end of the discriminator to get the desired quantum state. There is little to no pattern in the qubits on which the X gate has to be applied. Therefore, we prepared the discriminator for each number of qubits separately by hand. In Fig. 1b, we show the decomposition of six-qubit discriminator using ECR gates. As we could see, in this case the X gates are applied on second and third qubits.

For measurement, two distinct circuit implementations were used. The first method, referred to as the short measurement approach, is characterized by a reduced circuit depth and overall gate count. This method produces the disjoint sets of possible measurement outcomes, one for each unitary channel. However, this advantage is offset by an increased susceptibility to bit-flip errors. This circuit was implemented using CNOT gates, presented in Fig. 2a, or ECR gates, as in Fig. 2c. For the CNOT-based implementation, the measurement outcomes exhibit a regular structure: the scheme for the identity channel consistently yields the all-zero bitstring, while for the unitary channel $\Phi_{\text{RZ}(\theta)}$ produces bitstrings that are zero in all positions except for a single qubit set to one. The second measurement approach, referred to as the XOR measurement, uses a much deeper circuit. However, an added benefit is that the result for the identity channel consists of all zeros, and the result for the unitary channel consists of all ones. This approach helps mitigate bit-flip errors: if the majority of measured bits are zero, the result is taken as all zeros. In the case of a six-qubit system, and in the absence of noise, the observed bitstrings are 000000 for identity and 111111 for the unitary channel $\Phi_{\text{RZ}(\theta)}$. The implementation of this circuit is shown in Fig. 2b.

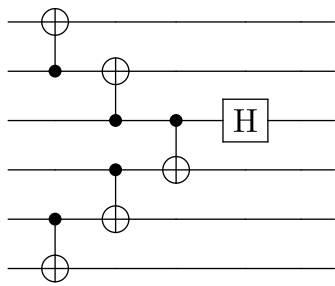
In contrast, the ECR-based implementation returns a more complex distribution of output bitstrings, lacking the clear structure observed in the CNOT-based case. With this implementation, we get two disjoint



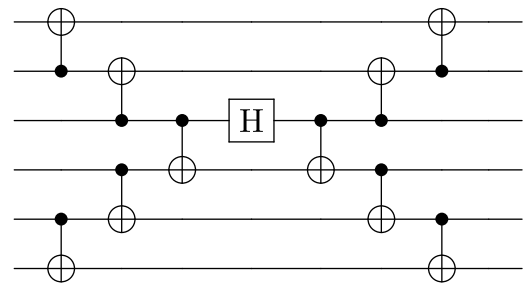
(a) Implementation of the discriminator using Hadamard and CNOT gates.

(b) Implementation of the discriminator using the native gates of the IBM Eagle R3 architecture.

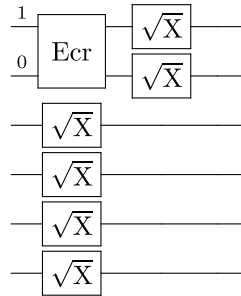
Fig. 1. Schematic implementation of a six-qubit discriminator using different gate sets: (a) Hadamard and CNOT; (b) IBM Eagle R3 native basis (single-qubit $\sqrt{X}$ (SX) and two-qubit ECR).



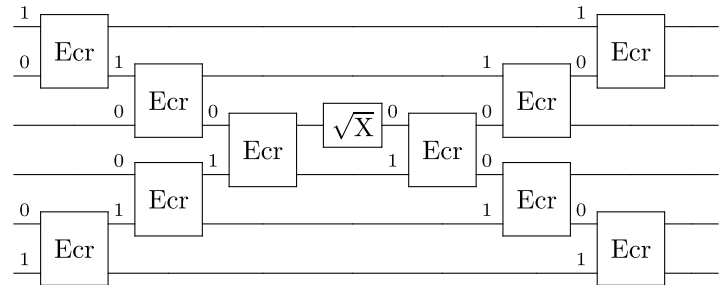
(a) CNOT implementation of the short measurement.



(b) CNOT implementation of the XOR measurement.



(c) ECR implementation of the short measurement.



(d) ECR implementation of the XOR measurement.

Fig. 2. Schematic implementations of six-qubit measurement circuits across gate sets. Panels (a) and (b) depict the circuits for the short and XOR measurement schemes, respectively, implemented using the non-native H and CNOT basis. Panels (c) and (d) illustrate the circuits for the short and XOR measurement schemes, respectively, utilizing the native IBM Eagle R3 basis gates ($\sqrt{X}$ and ECR).

subsets of all possible results, one subset for identity channel and the second subset for the unitary channel $\Phi_{RZ(\theta)}$. For example, on six-qubit system, we obtain the following two sets of bitstrings after measurement $\{001111, 010111, 101101, 110101, 100001, 100111, 011101, 110011, 000011, 111001, 011011, 111111, 001001, 000101, 010001, 101011\}$ for the identity channel and $\{110001, 001011, 011111, 100011, 001101, 110111, 000001, 010101, 000111, 101001, 111101, 010011, 101111, 111011, 011001, 100101\}$ for $\Phi_{RZ(\frac{\pi}{8})}$. Implementation of this ECR circuit could be seen in Fig. 2d.

The bitstring 001111 from the first set and 001101 from the second differ by only one bit, which may be problematic in a noisy environment. Even in this small example, processing these results is not straightforward, though this is compensated by the simplicity of the measurement circuit. In both measurement schemes, if it is not possible to determine which channel was applied during experiment, i.e., when the distance to both candidate channels is equal, then the output is assigned randomly via a coin flip. This approach ensures that each circuit run yields a definite outcome, thereby avoiding missing data points. Similarly, if the majority of bits have value one, we take this as the result of all of them.

Mapping on Eagle R3

- ECR gates are oriented, qubit order matters and simple swaps are constrained.
- For small widths w , hand-crafted ECR-native layouts avoid transpiler SWAP insertion.
- For larger w , either (i) size-specific hand layouts or (ii) transpiler-based layouts with SWAPs due to connectivity/orientation can be used.
- SWAPs increase depth and two-qubit gate count, impacting p_{succ} .

Comparative analysis of transpilation

To evaluate the efficacy of different quantum circuit transpilation approaches, particularly concerning the choice between CNOT and ECR gates and the impact of manual qubit mapping, experiments were conducted on the 6-qubit and the 11-qubit systems. The objective is to compare the performance, measured by result accuracy, for circuits utilizing either the short measurement or the XOR-based measurement scheme, while varying the transpilation method. Each circuit was executed with 100,000 shots.

The selection of CNOT and ECR gates for this study is based on their fundamental role in creating entanglement, which is a crucial component of the quantum circuits being investigated. A key consideration in this comparison is the fundamental difference in how these gates are implemented on the hardware, particularly on the Eagle R3 architecture. On this system, the ECR gate is a native gate, meaning it can be executed directly by the hardware. In contrast, the CNOT gate is not native and must be unrolled into a sequence of native gates.

Theoretically, a single CNOT gate requires one ECR gate plus additional single-qubit gates, already leading to a deeper circuit⁴⁹. Furthermore, the ECR gates on the Eagle R3 are oriented, meaning the qubit order matters and simple swaps are constrained. This necessitates careful mapping. For circuits with a small number of qubits, hand-crafted ECR-native layouts can be designed to avoid the insertion of additional SWAP gates by the transpiler. However, for larger circuits, either size-specific manual layouts or transpiler-based layouts are used. The latter often results in the automatic insertion of SWAP gates to accommodate qubit connectivity and gate orientation. The presence of these SWAP gates, which are themselves implemented as a sequence of native gates, significantly increases the circuit depth and the two-qubit gate count. This increased depth leads to a higher accumulation of errors, which directly impacts the result accuracy⁵⁰. Detail characteristics for different transpilation and measurement strategies for both 6-qubit and 11-qubit can be seen in Table 1.

6-Qubit system evaluation

For the 6-qubit configuration in a pure parallel discrimination scheme, four distinct transpilation strategies were evaluated using the short measurement and XOR-based measurement protocols:

1. **CNOT + Transpiler:** Circuit implemented using CNOT basis gates, processed by the Qiskit transpiler with optimization level 3. Short accuracy: 88.8%; XOR accuracy: 86.4%.
2. **ECR + Transpiler:** Circuit implemented using ECR basis gates, processed by the Qiskit transpiler with optimization level 3. Short accuracy: 83.8%; XOR accuracy: 90.0%.
3. **ECR + Transpiler + Fixed Mapping:** Circuit implemented using ECR basis gates, processed by the Qiskit transpiler with optimization level 3 and a predetermined, fixed mapping of logical qubits to physical qubits. Short accuracy: 84.4%, XOR accuracy: 85.3%.
4. **ECR + Fixed Mapping (No Opt.):** Circuit implemented using ECR gates, utilizing a fixed logical-to-physical qubit mapping, bypassing subsequent transpiler optimization passes. Short accuracy: 83.3%, XOR accuracy: 85.6%.

The experimental results indicate that for this system size, applying a fixed qubit mapping does not yield a discernible improvement in accuracy compared to relying solely on the transpiler's default ECR implementation. This observation is potentially attributable to the limited circuit depth and complexity inherent in six-qubit

(a) Circuit statistics for 6-qubits experiments							
Transpilation strategy	Measurement	Depth	ECR	RZ	$\sqrt{X}$	X	$RZ(\phi)$
CNOT + Transpiler	Short	35	10	53	30	5	6
CNOT + Transpiler	XOR	44	14	62	36	9	6
ECR + Transpiler	Short	20	6	18	19	1	6
ECR + Transpiler	XOR	40	14	59	34	2	6
ECR + Transpiler + Fixed Mapping	Short	8	6	0	12	2	6
ECR + Transpiler + Fixed Mapping	XOR	21	14	7	12	2	6
ECR + Fixed Mapping (No Opt.)	Short	8	6	0	12	2	6
ECR + Fixed Mapping (No Opt.)	XOR	13	15	0	7	2	6

(b) Circuit statistics for 11-qubits experiments.							
Transpilation strategy	Measurement	Depth	ECR	RZ	$\sqrt{X}$	X	$RZ(\phi)$
CNOT + Transpiler	Short	63	20	99	58	10	11
CNOT + Transpiler	XOR	65	29	108	52	18	11
ECR + Transpiler	Short	31	11	30	32	2	11
ECR + Transpiler (ID)	XOR	81	32	108	65	3	11
ECR + Transpiler ($RZ(\phi)$)	XOR	85	35	115	66	4	11
ECR (Topology-Aware) + Transpiler	Short	26	11	23	30	2	11
ECR (Topology-Aware) + Transpiler	XOR	70	29	108	62	6	11
ECR (Topology-Aware) + Transpiler + Fixed Mapping	Short	11	11	0	22	2	11
ECR (Topology-Aware) + Transpiler + Fixed Mapping	XOR	30	29	7	17	2	11
ECR (Topology-Aware) + Fixed Mapping (No Opt.)	Short	11	11	0	22	2	11
ECR (Topology-Aware) + Fixed Mapping (No Opt.)	XOR	22	30	0	12	2	11

Table 1. Circuit complexity metrics for the 6-qubit system presented in Table (a) and for the 11-qubit system presented in Table (b) across transpilation strategies. The table details the circuit *Depth*, two-qubit gate count (ECR), and single-qubit gate counts (RZ, $\sqrt{X}/SX$, X, ID/ $RZ(\phi)$) after transpilation to the native basis of the IBM Eagle architecture, comparing different initial circuits and measurement schemes. For the 11-qubit system the entries “ECR + Transpiler (ID)” and “ECR + Transpiler ($RZ(\phi)$)” reflect special cases where the measurement or state discrimination was tuned for the Identity (ID) and $RZ(\phi)$ channel, respectively. This specialized setting directly influenced the transpiler's output, leading to the distinct final circuit properties observed in these rows.

systems, where the transpiler's optimization might already find near-optimal solutions without explicit mapping constraints.

11-Qubit system evaluation

A subsequent set of experiments was performed using the 11-qubit system. Five transpilation strategies were assessed for each measurement type:

1. **CNOT + Transpiler:** Circuit using CNOT basis gates, transpiled with optimization level 3. Short accuracy: 43.3%, XOR accuracy: 48.5%.
2. **ECR + Transpiler:** Circuit using ECR basis gates, transpiled with optimization level 3. Short accuracy: 55.0%, XOR accuracy: 54.5%.
3. **ECR (Topology-Aware) + Transpiler:** Circuit initially designed considering device connectivity using ECR gates, then transpiled with optimization level 3. Short accuracy: 36.1%, XOR accuracy: 47.2%.
4. **ECR (Topology-Aware) + Transpiler + Fixed Mapping:** Topology-aware ECR circuit, transpiled with optimization level 3 and a fixed logical-to-physical qubit mapping. Short accuracy: 32.0%, XOR accuracy: 71.5%.
5. **ECR (Topology-Aware) + Fixed Mapping (No Opt.):** Topology-aware ECR circuit with fixed mapping, bypassing subsequent transpiler optimization. Short accuracy: 33.4%, XOR accuracy: 71.8%.

The results for the 11-qubit system demonstrate a general performance advantage for ECR-based implementations over CNOT-based ones for both measurement protocols. For the short measurement scheme, the standard ECR implementation processed by the transpiler (Method 2) yields the highest accuracy (55.0%). However, a more pronounced effect is observed for the XOR measurement protocol. Employing a topology-aware circuit design combined with fixed qubit mapping (Methods 4 and 5) results in a significant accuracy improvement, achieving approximately 71.5–71.8%, nearly a 20% absolute increase compared to the standard CNOT transpiled approach (48.5%).

To assess the effect of *Transpilation Strategy* and *Measurement Scheme* on accuracy in the hardware experiments summarized in the circuit-statistics tables, we performed a two-way ANOVA without replication⁵¹. For the 6-qubit system, neither factor was statistically significant at $\alpha = 0.05$ (*Transpilation Strategy*: $F(3, 3) = 0.745$, $p = 0.593$; *Measurement Scheme*: $F(1, 3) = 0.966$, $p = 0.398$). For the 11-qubit system *Transpilation Strategy* remained non-significant ($F(4, 4) = 0.327$, $p = 0.848$), while *Measurement Scheme* showed a marginal, non-significant trend ($F(1, 4) = 4.914$, $p = 0.091$). The modest sample size for the 11-qubit analysis limits statistical power, despite clear descriptive differences in some configurations.

The current experimental design, which utilized only a single accuracy result for each Transpilation Strategy and Measurement Scheme combination, precluded the testing of interaction effects between these categories for both the 6-qubit and 11-qubit systems, investigating these likely but untested relationships will be essential for future, more robust research. It is also worth observing that the substantial accuracy gain observed for the 11-qubit XOR measurement using hardware-aware design and fixed mapping highlights the potential benefits of tailoring circuits to specific device characteristics. However, this approach presents a significant practical challenge: it currently necessitates manual, device-specific, and qubit-count-specific circuit construction and mapping. This process lacks straightforward algorithmic automation and requires considerable expert intervention for each target configuration, limiting its scalability and general applicability.

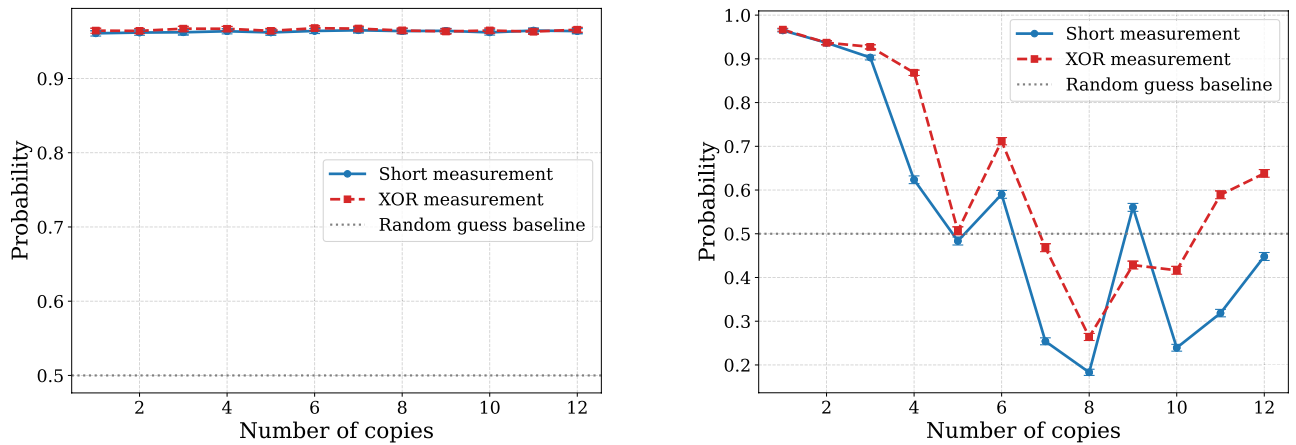
Preliminary results

From the initial experiments involving purely sequential and purely parallel discrimination schemes (see Fig. 3), we observe that the use of entangling gates on real quantum devices introduces a greater error overhead than the decoherence effects arising from increased gate depth in sequential protocols.

To further investigate this observation, we conducted a series of tests using hybrid rectangular (sequentially-parallel) schemes with a fixed number of unknown gate applications (see Fig. 4). The results consistently show a significant increase in error rates as more entangling gates are introduced. This trend reinforces the conclusion that the primary source of performance degradation in current devices stems from the imperfections of multi-qubit gate operations rather than decoherence from circuit depth alone. Another notable observation is that different measurement schemes have a visible but not dominant impact on the overall error rate, further suggesting that the circuit width and the number of entangled qubits are the primary contributors to performance degradation. This reinforces the conclusion that the main source of error arises from the entanglement rather than from the specifics of the measurement strategy.

Anomalous behavior discussion

During the experiments, we observed a specific issue with the IBM Quantum platform: measurements in experiments involving the entanglement of a certain number, typically five or more, qubits exhibited random bit-flip errors. This type of error is evident in Figs. 3 and 4 and appears to affect all qubits simultaneously, independent of the measurement strategy. Given that random guessing yields a baseline success probability of 0.5, device noise would normally drive the measured success rates toward this value. Instead, we observed something like inversion of outcomes, effectively corresponding to a swap of the expected answer sets. These anomalies can be reversed by interchanging the target labels, which restores internal consistency in the data. However, this procedure is purely for illustration, as the underlying cause of the bit-flip behavior remains unidentified and such swapping does not constitute a valid mitigation strategy. To ensure that the issue was not caused by errors in our own implementation, we performed several verification steps. First, multiple experimental results from different dates were analyzed. Second, the final transpiled circuits were simulated using a noiseless simulator to confirm



(a) Purely sequential scheme of discrimination.

(b) Purely parallel scheme of discrimination.

Fig. 3. Probability of successful discrimination between the identity operation and the $RZ(\pi/N)$ gate, where N is the number of copies of unknown unitary. The dashed red line corresponds to the XOR-based measurement strategy, while the solid blue line represents the short measurement scheme. Figure (a) illustrates the performance of the purely sequential scheme. In Figure (a), the probability is approximately 0.965. Figure (b) corresponds to the purely parallel scheme. Each circuit was executed with 10,000 shot. Measurement outcomes that could not be unambiguously associated with either gate were randomly assigned to one of the two possible answers.

the expected theoretical behavior. Third, we applied M3 error mitigation techniques⁵², which had no measurable effect on the observed phenomena. Given that the observed errors persist across all experiments, regardless of the RZ gate angle or the measurement strategy employed, we hypothesize the presence of a systematic hardware or software artifact specific to the IBM Quantum platform, particularly the Brisbane quantum device. Further investigation and hypothesis-driven testing are required to understand this behavior.

Experiment 2—Discrimination of unitary channels on IBM Brisbane with processing

In this section, we will present the second example of discrimination task with a different pair of unitary channels that require intermediate processing between applications to be perfectly distinguishable and run it on IBM Brisbane.

Methodology

In the second experiment, we take the following unitary channels to discriminate Φ_U for $U = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$ and Φ_V for $V = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$, where $\sqrt{X} = \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix}$. It is easy to check that the condition $\theta(V^\dagger U) = \frac{\pi}{N}$ is satisfied. It implies that for N -shot discrimination scenario, we will achieve perfect discrimination. Let us fix $N = wd$, where $w, d \in \mathbb{N}$. In this setup, instead of mid-processing $X_i = (V^\dagger)^{\otimes w}$ we use hardware-friendly processing $X_i = X^{\otimes w}$. For convenience we add also pre-processing unitary operation $X_0 = (X\sqrt{X})^{\otimes w}$ and post-processing operation $X_d = (\sqrt{X}X)^{\otimes w}$. Combining processing operators with the black-box Φ_U we get the following unitary operation

$$U_* = X_d U^{\otimes w} X_{d-1} \cdots X_1 U^{\otimes w} X_0 = \left(RZ\left(\frac{\pi}{2N}\right)^d \right)^{\otimes w}. \quad (16)$$

Similarly, for Φ_V the combined unitary circuit equals $V_* = \left(RZ\left(\frac{\pi}{2N}\right)^d \right)^{\otimes w}$. Note that $\theta\left(\left(RZ\left(\frac{\pi}{2N}\right)^d\right)^{\otimes w} \left(RZ\left(\frac{\pi}{2N}\right)^d\right)^{\otimes w}\right) = wd\theta\left(RZ\left(\frac{\pi}{2N}\right)\right) = N\frac{\pi}{N} = \pi$, so in theory, each shape w, d gives perfect discrimination. As $X, \sqrt{X}, RZ$ are native gates in IBM Quantum Brisbane this part of the circuit is implemented exactly as stated. To define the discriminator $|\psi\rangle$, we solve

$$\langle \psi | \left(RZ\left(\frac{\pi}{N}\right)^d \right)^{\otimes w} | \psi \rangle = 0. \quad (17)$$

We find that the input state can be taken as the GHZ state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\dots 0\rangle + |1\dots 1\rangle)$, which can be implemented by Hadamard gate followed by cascade of control-X gates. We let the IBM transpiler to optimize the input state circuit with the optimization level set to 3.

We distinguish two situations depending on the value of Φ . If $\Phi = \Phi_U$, then the input state $|\psi\rangle$ evolved under U_* is equal to $|\psi_{-i}\rangle = \frac{1}{\sqrt{2}}(|0\dots 0\rangle - i|1\dots 1\rangle)$. For $\Phi = \Phi_V$ we get after the evolution V_* that

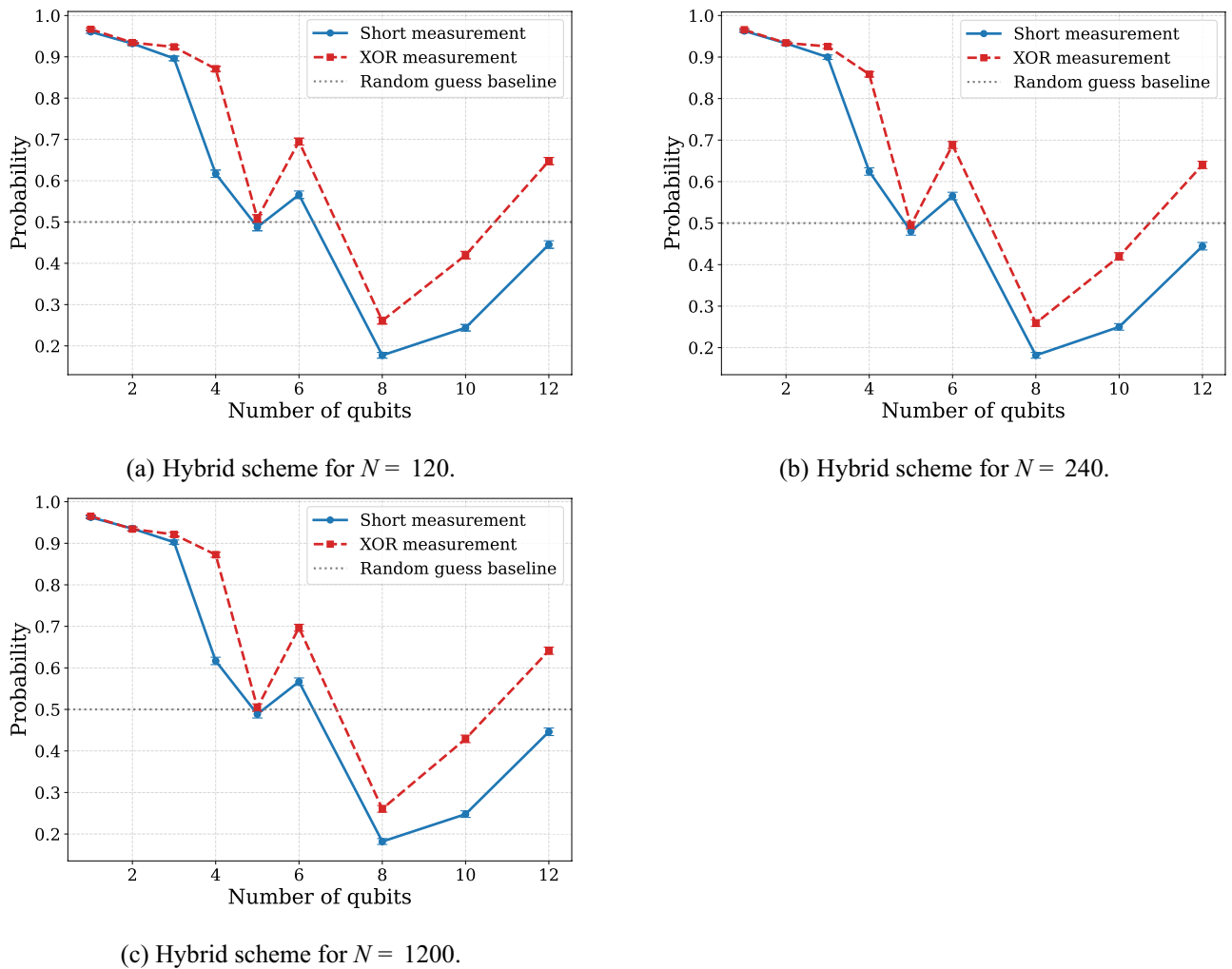


Fig. 4. Probability of successful discrimination between the identity operation and the $RZ(\pi/N)$ gate, where N is number of copies of unknown unitary, plotted as a function of the width of the hybrid rectangular scheme. The dashed red line corresponds to the XOR-based measurement strategy, while the solid blue line represents the short measurement scheme. Figure (a) illustrates the performance of hybrid scheme with 120 copies in total, Figure (b) corresponds to hybrid scheme with 240 copies in total, and Figure (c) corresponds to hybrid scheme with 1200 copies in total. Each circuit was executed with 10,000 shots. Measurement outcomes that could not be unambiguously associated with either gate were randomly assigned to one of the two possible answers.

$|\psi_i\rangle = \frac{1}{\sqrt{2}}(|0\dots 0\rangle + i|1\dots 1\rangle)$. At the measurement stage we implement $RZ(\pi/2)$ on the first qubit followed by parallel application of Hadamard gates $H^{\otimes w}$. Each qubit is then measured in the Z-basis. We let the IBM transpiler to optimize the measurement circuit with the optimization level set to 3. At the measurement stage, after applying $RZ(\pi/2)$ and $H^{\otimes w}$ we get

$$H^{\otimes w}(RZ(\pi/2) \otimes \mathbf{1})|\psi_{\mp i}\rangle = |+\dots+\rangle \pm |-\dots-\rangle, \quad (18)$$

where we used the notation of plus/minus states $|+\rangle = H|0\rangle$, $|-\rangle = H|1\rangle$. Observe that if $(b_1, \dots, b_w)$ is a bit string we received from measuring $|+\dots+\rangle \pm |-\dots-\rangle$, then in theory, $\Phi = \Phi_U$ if and only if $b_1 \oplus \dots \oplus b_w \equiv 0$, where $\oplus$ denotes XOR operation between bits. Similarly, $\Phi = \Phi_V$ if and only if $b_1 \oplus \dots \oplus b_w \equiv 1$.

Preliminary results

From relatively small number of copies ($N = 4, 16, 32$) of unitary channels to be discriminated, the experiments involving purely sequential schemes are preferable (see Fig. 5). We observe that the use of entangling gates on real quantum devices introduces a higher error overhead than the decoherence effects arising from increased gate depth in sequential protocols. For an increasing number of copies ($N = 64, 96, 1024$) of unitary channels, we can observe the advantage of usage sequentially-paralleled schemes to achieve more precise results (see Fig. 6). The results consistently show that after some threshold of gate composition, the existing decoherence or

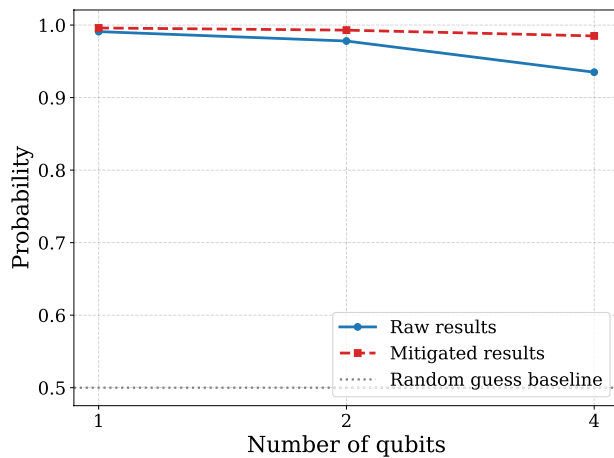
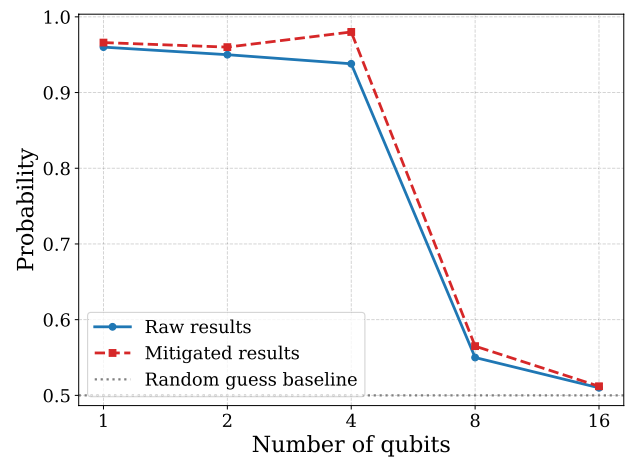
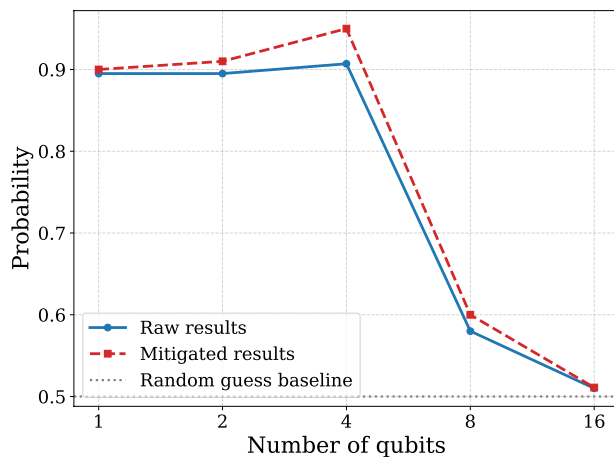
(a) Hybrid scheme for $N = 4$.(b) Hybrid scheme for $N = 16$.(c) Hybrid scheme for $N = 32$.

Fig. 5. Probability of successful discrimination between the unitary operator $U = \sqrt{X}RZ(\frac{-\pi}{2N})\sqrt{X}$ and $V = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$, where N is number of copies of unknown unitary, plotted as a function of the width of hybrid rectangular scheme, using the short measurement. The blue line corresponds to the results obtained directly from IBM Brisbane, while the red line represents the results after error mitigation using MThree package^{52,53}. Figure (a) illustrates the performance of hybrid schemes with 4 copies in total, Figure (b) corresponds to hybrid schemes with 16 copies in total, and Figure (c) corresponds to hybrid schemes with 32 copies in total. Each circuit was executed with 10,000 shots.

accumulative calibration imperfections gives higher error rates than the error rates used to create an entangled state.

Discussion

In the experiment with $N = 1024$ copies of black-box, we obtained effectively random results for all schemes considered (see Fig. 6c). The sequential scheme required excessive gate composition, while the parallel scheme required preparing a GHZ state that was too large. Both circuits introduced too many errors to yield any notable results. The final question that we can ask is if we can do better than that by exploring *suboptimal* circuits. A simple idea goes as follows. We perform independent sequential experiments on each of the w qubits. For each, we compose d quantum circuits and do independent measurements that indicate which black-box is preferable. As $\theta(V^\dagger U) = \frac{\pi}{N}$ for each qubit, we get the maximum angle spread of $\theta = \frac{\pi}{N}d = \frac{\pi}{w} < \pi$. Hence, the protocol is suboptimal. The probability of successful discrimination can be boosted by using majority voting on the results collected from the quantum computer. Let us assume that we collected k times label 0 indicating Φ_U and $w - k$ times label 1 standing for Φ_V . If $k > w - k$ we guess that $\Phi = \Phi_U$ (and for $k < w - k$ we guess $\Phi = \Phi_V$). In the case $k = w - k$, we make a random guess.

We applied the following suboptimal procedure for $N = 1024$. We used $w = 32$ qubits, and on each qubit we applied the operation Φ sequentially $d = 32$ times. According to Fig. 5, we should expect around 90% accuracy for each qubit. In combination, we obtained $p_{\text{succ}} = 0.56765$, which is better than any optimal

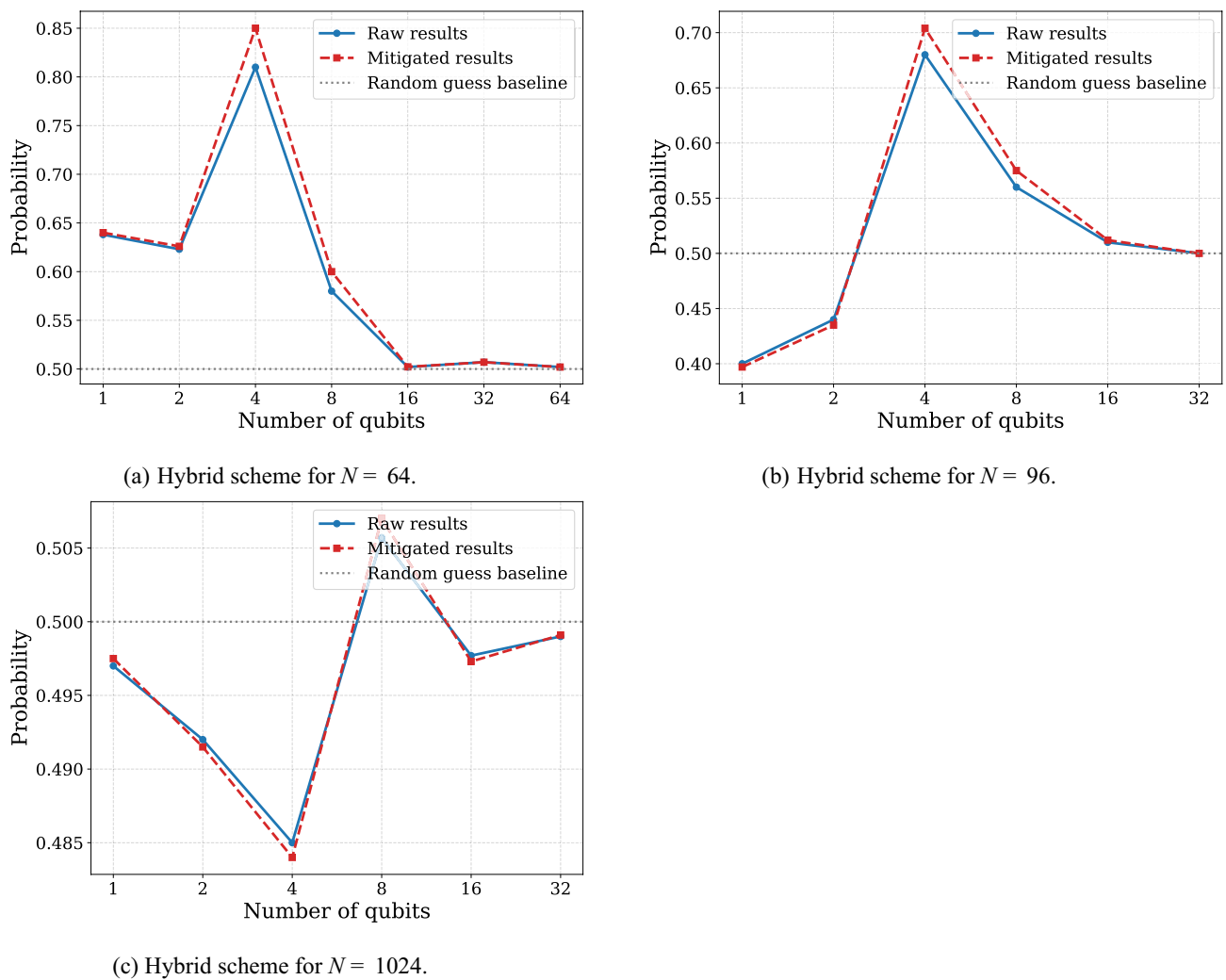


Fig. 6. Probability of successful discrimination between the unitary operator $U = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$ and $V = \sqrt{X}RZ(\frac{\pi}{2N})\sqrt{X}$, where N is number of copies of unknown unitary, plotted as a function of the width of hybrid rectangular scheme, using the short measurement. The blue line corresponds to the results obtained directly from IBM Brisbane, while the red line represents the results after error mitigation using MThree package^{52,53}. Figure (a) illustrates the performance of hybrid schemes with 64 copies in total, Figure (b) corresponds to hybrid schemes with 96 copies in total, and Figure (c) corresponds to hybrid schemes with 1024 copies in total. Each circuit was executed with 10,000 shots.

scheme. Similarly, we conducted the experiment for $N = 96$ and $w = 3, d = 32$ resulting in $p_{\text{succ}} = 0.74685$, which is a better result than indicated by Fig. 6b.

Can we say that suboptimal circuits are favorable for unitary channel discrimination? The answer is not straightforward. For example, for $N = 64$, the best result from Fig. 6a is slightly better than the suboptimal procedure presented above, even in the absence of noise, with $p_{\text{succ}} = 0.85295$. Interestingly, no analogous bit-flip anomalies were observed in these results compared to Experiment 1 (page 8), even though they employ comparable qubit counts and circuit depths. This may suggest that updates to the IBM Quantum execution pipeline or internal compilation procedures now perform additional low-level optimizations or calibration corrections before execution on hardware, which could explain the collective measurement flips seen in the simpler example. However, since these mechanisms are not publicly documented, this explanation remains speculative.

Noise modeling and ablation analysis Theoretical noise model

In this section we analyze the effect of symmetric depolarizing noise on the circuit presented in Experiment 1 for both optimal and suboptimal discrimination strategies. Let $U_{\pm} = RZ(\pm\pi/(2N))$ be two unitary operations to distinguish, where N is the number of copies given. We assume that $N = rwd$, where r is the number of independent circuits prepared, w is the number of qubits used per circuit, and d is the number of compositions of $U_{\pm}$ in each register. As a universal gate set, we choose single qubit operations and CNOT gates. Finally,

the depolarizing noise is modeled as follows. Let Ω_ϵ be a parametrized qubit depolarizing noise given by the equation $\Omega_\epsilon(\rho) = (1 - \epsilon)\rho + \epsilon\rho_*$, with the noise ratio $\epsilon \in [0, 1]$ and ρ_* indicating maximally mixed state. Each single qubit gate U is affected by Ω_ϵ , which we write as $\Omega_\epsilon(U\rho U^\dagger)$. The CNOT gate is affected by two independent noise sources acting on both qubits, $(\Omega_\epsilon \otimes \Omega_\epsilon)(CNOT\rho CNOT^\dagger)$. To implement suboptimal strategies, we use majority voting for r independent experiments. Hence, we will consider only odd values of r . If $r = 1$, we have the optimal strategy.

To implement the input state, we use single Hadamard gate and CNOT ladder containing $w - 1$ two-qubit operations. Hence, for the noiseless circuit ($\epsilon = 0$) we get a generalized GHZ state $\frac{1}{\sqrt{2}}(|0 \dots 0\rangle + |1 \dots 1\rangle)$. We further implement wd unitary $U_\pm$ gates as $(U_\pm^d)^{\otimes w}$ and obtain the state $\frac{1}{\sqrt{2}}(|0 \dots 0\rangle + e^{\pm(\pi wd)/(2N)i} |1 \dots 1\rangle)$. The measurement is performed using the short-measurement circuit,

where on the first qubit we put the gate $\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$ and on the remaining $w - 1$ qubits we put Hadamard gates. As a result, the state to measure is of the form $\frac{1}{\sqrt{2}}(|+\dots+\rangle + ie^{\pm(\pi wd)/(2N)i} |-\dots-\rangle)$. For any bitstring $b_1, \dots, b_w$ satisfying $b_1 \oplus \dots \oplus b_w = 0$, where $\oplus$ denotes XOR operation between bits, we get the probability $P(b_1, \dots, b_w) = \frac{1}{2^{w+1}} |1 + ie^{\pm(\pi wd)/(2N)i}|^2 = \frac{1}{2^w} (1 \mp \sin(\frac{\pi}{2r}))$. Therefore, each bitstring with the property $b_1 \oplus \dots \oplus b_w = 0$ will be connected with the guess U_- and $b_1 \oplus \dots \oplus b_w = 1$ with U_+ . In total, the success probability for correct discrimination in a single noiseless experiment reads $p_s = \frac{1}{2}(1 + \sin(\frac{\pi}{2r}))$.

In the presence of the noise $\epsilon > 0$ there is nonzero probability that at some point of the circuit, single qubit will decohere to ρ_* . If a single event like this happen, then the full coherence of the quantum state is lost. The information about $U_\pm$ is held in the phase of the GHZ state. Therefore, if at least one error occurs, the information about $U_\pm$ is lost and the probability of correct discrimination drops to $1/2$. Let us count the probability that there is no errors. Each error can occur with the same probability ϵ independently, so we need to count the number of places where the error can happen. To create the GHZ state, we use a single Hadamard gate and $w - 1$ CNOT gates. In total, there are $1 + 2(w - 1)$ possible positions where an error can occur. Then, we implement $U_\pm$, which adds d additional error positions per qubit, giving dw in total, where w is the circuit width. Finally, at the measurement stage we implement Hadamard gate for each qubit resulting in w error positions. To sum up, there are $1 + 2(w - 1) + wd + w = wd + 3w - 1$ error positions. The probability that there is no error is then $(1 - \epsilon)^{wd+3w-1}$. The success probability for a single experiment in the presence of noise is then given by

$$p_s(\epsilon) = (1 - \epsilon)^{wd+3w-1} p_s + (1 - (1 - \epsilon)^{wd+3w-1}) \frac{1}{2} = \frac{1}{2} + \frac{1}{2} (1 - \epsilon)^{wd+3w-1} \sin\left(\frac{\pi}{2r}\right). \quad (19)$$

Finally, if we add the majority vote procedure for r independent experiments, the total probability of success in the presence of noise Ω_ϵ reads as follows

$$P_{r,w,d}(\epsilon) = \sum_{k=\lceil r/2 \rceil}^r \binom{r}{k} p_s(\epsilon)^k (1 - p_s(\epsilon))^{r-k}. \quad (20)$$

If $r = 1$ and $\epsilon = 0$ we see that $P_{1,w,d}(0) = 1$ as intended for any shape (w, d) . However, in the presence of noise, we see that from all optimal strategies ($r = 1$), the formula $wd + 3w - 1 = N + 3w - 1$ is minimized for circuits with fewer qubits. Therefore, the circuits with a usage of a single qubit are preferable. They achieve a probability of success $P_{1,1,N}(\epsilon) = \frac{1}{2} + \frac{1}{2}(1 - \epsilon)^{N+2}$. The comparison with suboptimal strategies is more complex. As suboptimal strategies ($r > 1$) involve usage of majority voting they naturally constitute error correction method. Therefore, they could be preferable when the noise level exceeds some level for a given problem size. Indeed, let us take as an example $N = 150$ and compute $P_{1,1,150}(0.01) \simeq 0.6085$ and $P_{3,1,50}(0.01) \simeq 0.7158$. Three times repetition provides advantage for $\epsilon = 0.01$ even if for noiseless system we have $P_{1,1,150}(0) = 1$ and $P_{3,1,50}(0) = 0.84375$.

In conclusion, for a given quantum device with fixed error rate, large enough discrimination problems will benefit from utilizing suboptimal procedures that are more prone to errors.

Ablation study

To quantify the effect of different noise sources in Experiment 2, we performed an ablation study with the Aer Simulator using parameters calibrated to IBM Brisbane. Instead of a single model, we constructed configurations that enable only specific channels:

- 1q only: depolarization on single-qubit gates,
- 2q only: depolarization on two-qubit gates,
- T_1/T_2 only: thermal relaxation and dephasing,
- Readout only: asymmetric measurement errors,
- Full rebuilt: manual composition of average depolarization, T_1/T_2 , and readout noise,
- Full backend: calibration-based model from `NoiseModel.from_backend`.

The calibration data used for noise modeling were derived from the real IBM Brisbane device around the time of the experiments. The average single-qubit depolarizing error was set to 1.104×10^{-3} , and the two-qubit (ECR) gate error to 1.946×10^{-2} . The median relaxation times were taken as $T_1 = 210 \mu\text{s}$ and $T_2 = 155 \mu\text{s}$, with

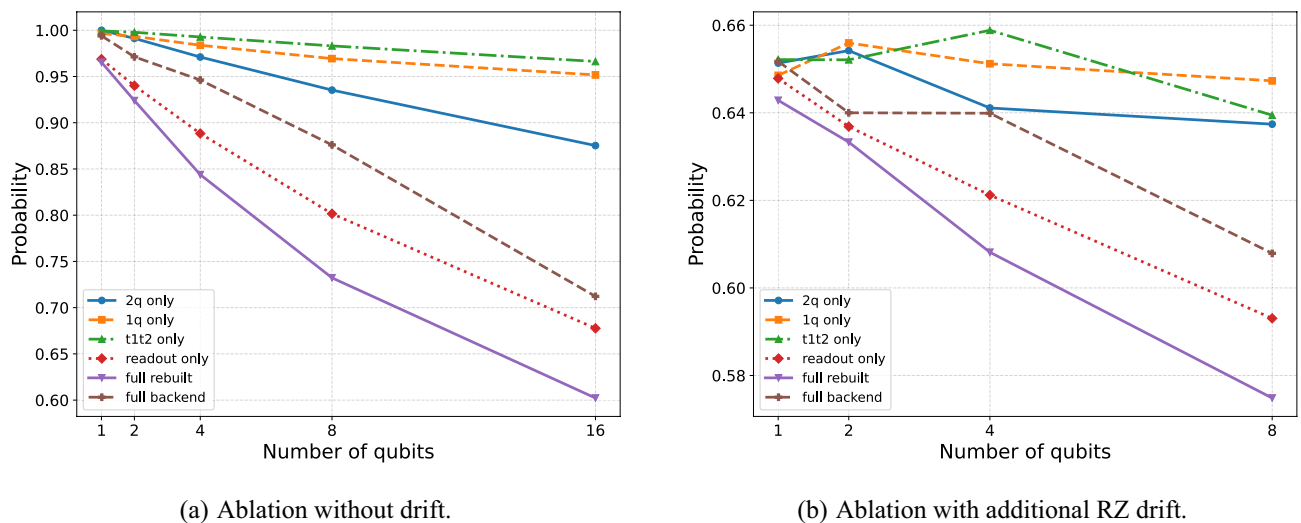


Fig. 7. Probability of successful discrimination between $U = \sqrt{X}, \text{RZ}(-\pi/128), \sqrt{X}$ and $V = \sqrt{X}, \text{RZ}(\pi/128), \sqrt{X}$ for $N = 64$ copies, plotted as a function of circuit width w (number of qubits). The dashed brown line corresponds to the full backend noise model of IBM Brisbane provided by Qiskit, while the solid purple line denotes the full rebuilt model that combines depolarization, thermal relaxation, and readout errors. The remaining curves represent ablation configurations with only single-qubit noise (dashed orange), two-qubit noise (solid blue), thermal relaxation (dash-dotted green), or readout noise (dotted red). The results confirm that two-qubit and readout noise are the dominant contributors to performance degradation. Figure (a) shows ablation results without coherent drift, whereas Figure (b) includes an added systematic phase offset $\theta \mapsto \theta + \pi/160$ on each RZ gate. Each circuit was executed with 10,000 shots.

corresponding gate durations of 35 ns for single-qubit gates and 248 ns for two-qubit gates. The readout errors were modeled with asymmetric probabilities of $p_{01} = 2.5 \times 10^{-2}$ and $p_{10} = 2.0 \times 10^{-2}$.

The results for $N = 64$ and widths $w \in \{1, 2, 4, 8, 16\}$ are shown in Fig. 7. The data confirm that two-qubit noise is the dominant source of degradation, while single-qubit errors and relaxation processes have a noticeably weaker but still increasing impact as the depth grows. Readout errors alone contribute greatly to the reduction in performance. However, this is likely an overestimation because in our noise model we apply uniform readout asymmetries rather than the full per-qubit readout map. The *full rebuilt* model consistently yields lower success probabilities than the *full backend*, as the reconstruction uses averaged T_1/T_2 values and applies depolarization and dephasing independently at each gate and also already mentioned readout errors. This slightly overestimates the combined effect of noise and neglects favorable device inhomogeneities, making the rebuilt model more pessimistic.

To further bridge the gap between simulations and hardware runs, we introduced a small coherent drift (biased gate calibration error) by replacing every $\text{RZ}(\theta)$ with $\text{RZ}(\theta + \pi/160)$. This modification yields the curves shown in Fig. 7b, which more closely reproduce the relative trends observed on real hardware (Fig. 6a), although the agreement is not exact.

Conclusion

In this work, we have studied the discrimination of two quantum unitary channels and benchmarked various schemes for perfect discrimination between them. The benchmarks were performed using the IBM Brisbane quantum device. All figures indicate the shot count (10,000 shots per circuit). Where multiple hardware runs were available, we aggregated by computing p_{succ} for each run and comparing across configurations as above. For noise-model simulations, we used the same success metric and classification rule as the hardware runs to ensure comparability.

As the first example, we chose the discrimination task between the identity operation and the $\text{RZ}(\phi)$ gate with no processing between them. To optimize circuit performance for this discrimination task, we also evaluated different quantum circuit transpilation approaches on 6- and 11-qubit subsystems, comparing the circuit implementation with CNOT and ECR gates and the impact of manual qubit mapping. Although fixed qubit mapping did not significantly improve accuracy in the smaller 6-qubit system, topology-aware circuit design combined with fixed mapping yielded substantial gains on the 11-qubit system when using XOR-based measurements, thereby underscoring the importance of hardware-aware optimization for larger circuits. In this experiment, we find the trend that deeper circuit architectures, which minimize entanglement overhead while preserving discrimination power, are significantly more resilient to hardware noise (see Fig. 3). Our results suggest that algorithm designers should prioritize circuit depth over width whenever possible. In the second example, we considered the discrimination between the unitaries $U = \sqrt{X}, \text{RZ}(-\pi/2N), \sqrt{X}$ and $V = \sqrt{X}, \text{RZ}(\pi/2N), \sqrt{X}$, where the processing consists of compositions of X and $\sqrt{X}$ gates. There, we have observed the advantage of using sequentially-paralleled schemes to achieve more precise results. The results

consistently show that beyond a certain threshold of gate composition, decoherence and cumulative calibration imperfections produce higher error rates than those required to generate an entangled state.

Our calibrated noise simulations identify two-qubit gate and readout errors as primary performance limitations. Consequently, for large-scale problems such as $N = 1024$, theoretically optimal schemes failed, producing random outcomes. In contrast, a suboptimal approach using majority voting proved more effective in some cases in accordance with the theoretical results. A final observation concerns the bit-flip anomaly reported in Experiment 1 (see Figs. 3 and 4), where correlated flips were observed across all measured qubits for certain qubit counts. In contrast, no such behavior was observed in Experiment 2 (see Figs. 5 and 6), despite its comparable width and depth combinations. This discrepancy may indicate that the IBM Quantum execution stack performs undisclosed optimizations or internal circuit simplifications for certain circuit classes, potentially suppressing these systematic artifacts in more complex layouts. However, without public access to the complete compilation and calibration pipeline, this hypothesis remains highly speculative. More research is needed to determine whether these effects arise from backend-level preprocessing, device-specific behavior, or a combination of both.

Circuit geometries beyond square layouts may offer a more accurate reflection of the capabilities of the device. These findings can be applied to various black-box tasks with many copies, such as quantum phase estimation⁵⁴. In that case, discrimination schemes that are theoretically suboptimal achieve good experimental performance⁵⁵. While comparing unitary channels leads to expected consistent and comparable results, the task of general quantum channel discrimination introduces complexity as its necessary implementation via the Stinespring representation requires adding ancillary systems and SWAP gates to enable entanglement. This operational overhead inherently subjects the experimental outcomes to a greater degree of error. Experimental studies of the discrimination of general quantum channels remain an open direction for future research.

Data availability

The data that support the findings of this study are openly available in Github repository at https://github.com/Dotnester/experimental_study_of_multiple_shot_channel_discrimination and on Zenodo⁵⁶.

Received: 18 June 2025; Accepted: 4 December 2025

Published online: 08 February 2026

References

- Eisert, J. et al. Quantum certification and benchmarking. *Nat. Rev. Phys.* **2**, 382–390 (2020).
- McKay, D. C. et al. Benchmarking quantum processor performance at scale. arXiv preprint [arXiv:2311.05933](https://arxiv.org/abs/2311.05933) (2023).
- Proctor, T., Young, K., Baczewski, A. D. & Blume-Kohout, R. Benchmarking quantum computers. *Nat. Rev. Phys.* **7**, 105–118 (2025).
- Jałowiecki, K., Lewandowska, P. & Pawela, Ł. PyQBench: A Python library for benchmarking gate-based quantum computers. *SoftwareX* **24**, 101558 (2023).
- Lewandowska, P. & Beseda, M. Benchmarking gate-based quantum devices via certification of qubit von neumann measurements. arXiv preprint [arXiv:2506.03514](https://arxiv.org/abs/2506.03514) (2025).
- Helstrom, C. W. Quantum detection and estimation theory. *J. Stat. Phys.* **1**, 231–252 (1969).
- Watrous, J. *The Theory of Quantum Information* (Cambridge University Press, 2018).
- Sen, K., Halder, S. & Sen, U. Incompatibility of local measurements providing an advantage in local quantum state discrimination. *Phys. Rev. A* **109**, 012415 (2024).
- Bandyopadhyay, S., Halder, S. & Nathanson, M. Optimal resource states for local state discrimination. *Phys. Rev. A* **97**, 022314 (2018).
- Calsamiglia, J., De Vicente, J., Muñoz-Tapia, R. & Bagan, E. Local discrimination of mixed states. *Phys. Rev. Lett.* **105**, 080504 (2010).
- Bergou, J. A., Herzog, U. & Hillery, M. 11 discrimination of quantum states. *Quantum state estimation* 417–465 (2004).
- Qiu, D. Minimum-error discrimination between mixed quantum states. *Phys. Rev. A*. <https://doi.org/10.1103/physreva.77.012328> (2008).
- Watrous, J. Simpler semidefinite programs for completely bounded norms. *Chicago J. Theor. Computer Sci.* **8**, 1–19 (2013).
- Benenti, G. & Strini, G. Computing the distance between quantum channels: Usefulness of the fano representation. *J. Phys. B Atomic Mol. Optical Phys.* **43**, 215508 (2010).
- Ramana, M. V. An exact duality theory for semidefinite programming and its complexity implications. *Math. Programming* **77**, 129–162 (1997).
- Gawron, P., Kurzyk, D. & Pawela, Ł. QuantumInformation.jl—A julia package for numerical computation in quantum information theory. *PLOS ONE* **13**, e0209358, <https://doi.org/10.1371/journal.pone.0209358> (2018).
- Ziman, M. & Sedláč, M. Single-shot discrimination of quantum unitary processes. *J. Modern Optics* **57**, 253–259 (2010).
- Pawela, Ł. et al. Numerical shadow. The web resource at <https://numericalshadow.org/>. Accessed on 2020-05-25.
- Sedláč, M. & Ziman, M. Optimal single-shot strategies for discrimination of quantum measurements. *Phys. Rev. A* **90**, 052312 (2014).
- Krawiec, A., Pawela, Ł. & Puchała, Z. Discrimination and certification of unknown quantum measurements. *Quantum* **8**, 1269 (2024).
- Puchała, Z., Pawela, Ł., Krawiec, A. & Kukulski, R. Strategies for optimal single-shot discrimination of quantum measurements. *Phys. Rev. A* **98**, 042103 (2018).
- Krawiec, A., Pawela, Ł. & Puchała, Z. Discrimination of povms with rank-one effects. *Quant. Information Process.* **19**, 1–12 (2020).
- Ji, Z., Feng, Y., Duan, R. & Ying, M. Identification and distance measures of measurement apparatus. *Phys. Rev. Lett.* **96**, 200401 (2006).
- Cao, T.-Q., Gao, F., Zhang, Z.-C., Yang, Y.-H. & Wen, Q.-Y. Perfect discrimination of projective measurements with the rank of all projectors being one. *Quant. Information Process.* **14**, 2645–2656 (2015).
- Oreshkov, O., Costa, F. & Brukner, Č. Quantum correlations with no causal order. *Nat. Commun.* **3**, 1–8 (2012).
- Araújo, M. et al. Witnessing causal nonseparability. *N. J. Phys.* **17**, 102001 (2015).
- Lewandowska, P., Pawela, Ł. & Puchała, Z. Strategies for single-shot discrimination of process matrices. *Sci. Rep.* **13**, 3046 (2023).
- Salek, F., Hayashi, M. & Winter, A. Usefulness of adaptive strategies in asymptotic quantum channel discrimination. *Phys. Rev. A* **105**, 022419 (2022).

29. Bavaresco, J., Murao, M. & Quintino, M. T. Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination. *Phys. Rev. Lett.* **127**, 200504 (2021).
30. Bavaresco, J., Murao, M. & Quintino, M. T. Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies. *J. Math. Phys.* **63**, 042203 (2022).
31. Duan, R., Guo, C., Li, C.-K. & Li, Y. Parallel distinguishability of quantum operations. In *2016 IEEE International Symposium on Information Theory (ISIT)*, 2259–2263 (IEEE, 2016).
32. Chiribella, G., D'Ariano, G. M. & Perinotti, P. Memory effects in quantum channel discrimination. *Phys. Rev. Lett.* **101**, 180501 (2008).
33. Duan, R., Feng, Y. & Ying, M. Entanglement is not necessary for perfect discrimination between unitary operations. *Phys. Rev. Lett.* **98**, 100503 (2007).
34. Harrow, A. W., Hassidim, A., Leung, D. W. & Watrous, J. Adaptive versus nonadaptive strategies for quantum channel discrimination. *Phys. Rev. A* **81**, 032339 (2010).
35. Hayashi, M. Discrimination of two channels by adaptive methods and its application to quantum system. *IEEE Trans. Inform. Theory* **55**, 3807–3820 (2009).
36. Wang, G. & Ying, M. Unambiguous discrimination among quantum operations. *Phys. Rev. A* **73**, 042301 (2006).
37. Duan, R., Feng, Y. & Ying, M. Perfect distinguishability of quantum operations. *Phys. Rev. Lett.* **103**, 210501 (2009).
38. Puchała, Z., Pawela, Ł., Krawiec, A., Kukulski, R. & Oszmaniec, M. Multiple-shot and unambiguous discrimination of von Neumann measurements. *Quantum* **5**, 425 (2021).
39. Liu, S., Li, Y. & Duan, R. Distinguishing unitary gates on the IBM quantum processor. *Sci. China Information Sci.* **62**, 1–7 (2019).
40. Hausdorff, F. Der Wertvorrat einer bilinearform. *Mathematische Zeitschrift* **3**, 314–316 (1919).
41. Toeplitz, O. Das algebraische analogon zu einem satze von fejér. *Mathematische Zeitschrift* **2**, 187–197 (1918).
42. Holevo, A. S. Statistical decision theory for quantum systems. *J. Multivariate Anal.* **3**, 337–394 (1973).
43. Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404–413 (1934).
44. Watrous, J. Channel distinguishability and the completely bounded trace norm. Tech. Rep. CS 766 / QIC 820 Lecture 20, University of Waterloo (2011). Available at <https://johnwatrous.com/wp-content/uploads/TQI-notes.20.pdf>.
45. Acín, A. Statistical distinguishability between unitary operations. *Phys. Rev. Lett.* **87**, 177901 (2001).
46. D'Ariano, G. M., Presti, P. L. & Paris, M. G. Using entanglement improves the precision of quantum measurements. *Phys. Rev. Lett.* **87**, 270404 (2001).
47. Eldredge, Z., Foss-Feig, M., Gross, J. A., Rolston, S. L. & Gorshkov, A. V. Optimal and secure measurement protocols for quantum sensor networks. *Phys. Rev. A* **97**, 042337 (2018).
48. IBM Quantum. Processor types guide. The web resource at <https://docs.quantum.ibm.com/guides/processor-types>. Accessed on 2025-05-18.
49. Tan, X. Quantum computing for phonon scattering effects on thermal conductivity, <https://doi.org/10.48550/arXiv.2407.15808> (2024).
50. Hardy, Y. & Steeb, W.-H. Decomposing the swap quantum gate. *J. Phys. A Math. General* **39**, 1463. <https://doi.org/10.1088/0305-4470/39/6/018> (2006).
51. Alin, A. & Kurt, S. Testing non-additivity (interaction) in two-way anova tables with no replication. *Stat. Methods Med. Res.* **15**, 63–85. <https://doi.org/10.1191/0962280206sm4260a> (2006).
52. Nation, P. D., Kang, H., Sundaresan, N. & Gambetta, J. M. Scalable mitigation of measurement errors on quantum computers. *PRX Quantum* **2**, 040326 (2021).
53. Qiskit and the mthree developers. mthree.m3mitigation documentation. The web resource at <https://qiskit.org/documentation/partners/mthree/stubs/mthree.M3Mitigation.html>. Accessed on 2025-05-18.
54. Dorner, U. et al. Optimal quantum phase estimation. *Phys. Rev. Lett.* **102**, 040403 (2009).
55. Svore, K. M., Hastings, M. B. & Freedman, M. Faster phase estimation. *Quant. Inform. Comput.* **14**, 306–328 (2014).
56. Bilek, A., Hlisnikovský, J., Bezdek, T., Kukulski, R. & Lewandowska, P. Zenodo: Experimental study of multiple-shot unitary channels discrimination using the IBM Q computers, <https://doi.org/10.5281/zenodo.15464711>, (2025).

Acknowledgements

It is a pleasure to thank Łukasz Pawela for numerous discussions concerning experiments on IBM quantum devices.

Author contributions

All authors have an equal contribution to the preparation of the paper. AB, JH, and TB contributed to the code base, setting up and running experiments, and prepared the figures. RK and PL set the research objective.

Funding

AB is supported by Grant of SGS No. SP2025/049, VŠB - Technical University of Ostrava, Czech Republic. JH is supported by the Ministry of Education, Youth and Sports of the Czech Republic through the e-INFRA CZ (ID:90254), with the financial support by Grant of SGS No. SP2025/049, VŠB - Technical University of Ostrava, Czech Republic. TB is supported by Grant of SGS No. SP2025/049, VŠB - Technical University of Ostrava, Czech Republic. PL is supported by the Ministry of Education, Youth and Sports of the Czech Republic through the e-INFRA CZ (ID:90254), with the financial support of the European Union under the REFRESH – Research Excellence For REgionSustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition. RK is supported by the European Union under the REFRESH – Research Excellence For REgionSustainability and High-tech Industries project number CZ.10.03.01/00/22_003/0000048 via the Operational Programme Just Transition.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025