



INSTYTUT INFORMATYKI TEORETYCZNEJ I STOSOWANEJ  
POLSKIEJ AKADEMII NAUK

---

METODY WALIDACJI WSPÓŁCZESNYCH ARCHITEKTUR  
KWANTOWYCH

---

ROZPRAWA DOKTORSKA

mgr Paulina LEWANDOWSKA

Promotor:

dr hab. Zbigniew PUCHAŁA, prof. IITiS PAN

Promotor pomocniczy:

dr hab. inż. Łukasz PAWELA, prof. IITiS PAN

Gliwice, 2023





INSTITUTE OF THEORETICAL AND APPLIED INFORMATICS, POLISH  
ACADEMY OF SCIENCES

---

METHODS OF VALIDATION OF MODERN QUANTUM  
ARCHITECTURES

---

DOCTORAL DISSERTATION

Paulina LEWANDOWSKA, MSc

Supervisor:

Zbigniew PUCHAŁA, PhD

Co-supervisor:

Łukasz PAWELA, PhD

Gliwice, 2023



# Contents

<b>1</b>	<b>Introduction</b>	<b>11</b>
<b>2</b>	<b>Mathematical preliminaries</b>	<b>17</b>
2.1	Complex Euclidean spaces and Dirac notation . . . . .	17
2.2	Linear operators . . . . .	20
2.2.1	Properties of linear operators . . . . .	21
2.2.2	Operator decompositions . . . . .	24
2.2.3	Operator norms . . . . .	27
2.2.4	Numerical range and $q$ -numerical range . . . . .	27
2.3	Linear superoperators . . . . .	29
2.3.1	Representations of linear superoperators . . . . .	31
2.3.2	Diamond norm of superoperators . . . . .	32
2.4	Semidefinite programming . . . . .	33
<b>3</b>	<b>Introduction to quantum information theory</b>	<b>35</b>
3.1	Quantum state . . . . .	35
3.2	State transformations . . . . .	36
3.3	Quantum channel representation . . . . .	37
3.4	Quantum measurement . . . . .	38
3.4.1	Quantum measurement and the Born rule . . . . .	38
3.4.2	Quantum measurement as a quantum channel . . . . .	38
3.4.3	Von Neumann measurements . . . . .	39
3.5	Fidelity measure . . . . .	40
3.5.1	Haar measure and the average fidelity function . . . . .	41
3.6	Quantum circuits . . . . .	42
3.6.1	One qubit gates . . . . .	42
3.7	Link product and quantum networks . . . . .	45
3.7.1	Link product . . . . .	45
3.7.2	Graphical representation of linear maps and quantum networks	46
3.7.3	Quantum network and tester . . . . .	47

<b>4</b>	<b>Storage and retrieval of von Neumann measurements</b>	<b>49</b>
4.1	Learning setup . . . . .	50
4.2	Main results . . . . .	51
4.2.1	Lower bound . . . . .	52
4.2.2	Upper bound . . . . .	55
4.3	Qubit cases ( $d = 2$ ) . . . . .	57
4.3.1	Pretty good learning scheme . . . . .	57
4.3.2	Learning scheme based on port-based teleportation . . . . .	60
4.3.3	Numerical investigation . . . . .	61
4.4	Conclusion and discussion . . . . .	61
<b>5</b>	<b>Indefinite causal structure theory in SAR of von Neumann measurements</b>	<b>65</b>
5.1	Process matrices – definition . . . . .	66
5.1.1	Bipartite process matrices . . . . .	66
5.1.2	N-partite process matrices . . . . .	68
5.2	Advantage of SAR of von Neumann measurements by using indefinite causal structure theory . . . . .	68
5.2.1	General approach of causal learning scheme . . . . .	69
5.2.2	$2 \rightarrow 1$ causal learning scheme of von Neumann measurements	70
5.2.3	$N \rightarrow 1$ causal learning scheme of qubit von Neumann measurements for $N \geq 3$ . . . . .	72
5.3	Conclusion and discussion . . . . .	73
<b>6</b>	<b>Discrimination of von Neumann measurements</b>	<b>75</b>
6.1	Optimal discrimination scheme . . . . .	76
6.1.1	Optimal probability . . . . .	77
6.1.2	Discriminator . . . . .	77
6.1.3	Final optimal measurement . . . . .	79
6.2	Discrimination scheme for parameterized family of Fourier measurements . . . . .	80
6.2.1	Optimal probability for parameterized family of Fourier measurements . . . . .	81
6.2.2	Optimal discrimination strategy for parameterized family of Fourier measurements . . . . .	84
6.3	Conclusion and discussion . . . . .	86
<b>7</b>	<b>PyQBench: a Python library for benchmarking gate-based quantum computers</b>	<b>89</b>
7.1	Limitations of NISQ devices and solutions . . . . .	91
7.2	Software Functionalities . . . . .	94

7.3	Software Architecture . . . . .	94
7.3.1	PyQBench as a CLI tool . . . . .	95
7.3.2	PyQBench as a library . . . . .	103
7.4	Conclusion and discussion . . . . .	103
<b>8</b>	<b>Certification of von Neumann measurements</b>	<b>105</b>
8.1	Certification of pure quantum states . . . . .	107
8.2	General concept of certification for quantum channels and measurements . . . . .	109
8.3	Certification of unitary channels . . . . .	110
8.4	Certification of von Neumann measurements . . . . .	112
8.5	Optimal certification protocol . . . . .	119
8.6	Certification scheme for parameterized family of Fourier measurements	119
8.7	Realization of the von Neumann measurement certification scheme .	124
8.8	Benchmarking based on certification scheme . . . . .	126
8.9	Conclusion and discussion . . . . .	128
<b>9</b>	<b>Summary</b>	<b>131</b>
	<b>Bibliography</b>	<b>143</b>
<b>A</b>	<b>Learning of von Neumann measurements</b>	<b>145</b>
A.1	Proof of Lemma 2 for qubit case . . . . .	145
A.1.1	Measurement learning via parallel storage of unitary transformations . . . . .	146
A.1.2	Objective function simplification . . . . .	149
A.1.3	Technical lemmas . . . . .	152
A.1.4	Proof of Lemma 2 for qubit case . . . . .	156
A.2	Proof of upper bound for any dimension $d$ . . . . .	157
A.3	Pretty good learning scheme . . . . .	158
<b>B</b>	<b>PyQBench: a Python library for benchmarking gate-based quantum computers</b>	<b>163</b>
B.1	PyQBench as a CLI . . . . .	163
B.2	PyQBench as a library . . . . .	168
<b>C</b>	<b>Certification of von Neumann measurements</b>	<b>177</b>
C.1	Distance of $q$ -numerical range to zero . . . . .	177





# List of publications related to the dissertation

1. Paulina Lewandowska, Aleksandra Krawiec, Ryszard Kukulski, Łukasz Pawela, and Zbigniew Puchała, *On the optimal certification of von Neumann measurements*, Scientific Reports, vol. 11, (2021). DOI: 10.1038/s41598-021-81325-1. arXiv: 2009.06776.
2. Paulina Lewandowska, Ryszard Kukulski, Łukasz Pawela, and Zbigniew Puchała, *Storage and retrieval of von Neumann measurements*, Physical Review A, vol. 106, issue 5, (2022). DOI: 10.1103/PhysRevA.106.052423. arXiv: 2204.03029.
3. Konrad Jałowiecki, Paulina Lewandowska and Łukasz Pawela, *PyQBench: a Python library for benchmarking gate-based quantum computers*, arXiv preprint, (2023). arXiv: 2304.00045. Under review SoftwareX.



# List of publications

## Published work

1. Przemysław Sadowski, Łukasz Pawela, Paulina Lewandowska, and Ryszard Kukulski *Quantum walks on hypergraphs*, International Journal of Theoretical Physics, vol. 58, (2019), DOI: 10.1038/s41598-018-27825-z. arXiv: 1809.04521.
2. Paulina Lewandowska, Ryszard Kukulski, and Łukasz Pawela, *Optimal representation of quantum channels*, Computational Science - ICCS 2020, (2020), DOI: 10.1007/978-3-030-50433-5\_47. arXiv: 2002.05507.
3. Ryszard Kukulski, Paulina Lewandowska, and Łukasz Pawela, *Perturbation of the numerical range of unitary matrices*, Computational Science - ICCS 2020, (2020), DOI: 10.1007/978-3-030-50433-5\_48. arXiv: 2002.05553.
4. Paulina Lewandowska, Aleksandra Krawiec, Ryszard Kukulski, Łukasz Pawela, and Zbigniew Puchała, *On the optimal certification of von Neumann measurements*, Scientific Reports, vol. 11, (2021), DOI: 10.1038/s41598-021-81325-1. arXiv: <https://arxiv.org/abs/2009.06776>
5. Paulina Lewandowska, Ryszard Kukulski, Łukasz Pawela, and Zbigniew Puchała, *Storage and retrieval of von Neumann measurements*, Physical Review A, vol. 106, issue 5, (2022), DOI: 10.1103/PhysRevA.106.052423. arXiv: 2204.03029.
6. Paulina Lewandowska, Łukasz Pawela, and Zbigniew Puchała, *Strategies for single-shot discrimination of process matrices*, Scientific Reports, vol. 13, (2023), DOI: 10.1038/s41598-023-30191-0. arXiv: 2210.14575.

## Preprints

1. Konrad Jałowiecki, Paulina Lewandowska and Łukasz Pawela, *PyQBench: a Python library for benchmarking gate-based quantum computers*, arXiv

preprint, (2023). arXiv: <https://arxiv.org/abs/2304.00045>. Under review SoftwareX.

# Streszczenie w języku polskim

W ostatniej dekadzie możemy obserwować, jak urzeczywistnia się idea budowy i praktycznego wykorzystania komputerów kwantowych. Producenci urządzeń typu NISQ (Noisy Intermediate-Scale Quantum) weszli szturmem na rynek techniki informatycznej, dzięki czemu mamy obecnie szeroki wybór komputerów kwantowych, opartych na różnych architekturach i odpowiadających im rodzajach oprogramowania. Wśród dostawców sprzętu, którzy oferują publiczny dostęp do urządzeń opartych na bramkowym modelu obliczeń, można wymienić takie firmy jak Rigetti, IBM, Oxford Quantum Group, IonQ czy Xanadu. Jeśli natomiast wziąć pod uwagę urządzenia wykorzystujące inne paradygmaty, to warto wspomnieć o D-Wave i ich kwantowych wyżarzaczach, czy urządzeniach QuEra opartych na atomach neutralnych. Jak już wspomniano wyżej, większość producentów zapewnia użytkownikom własne oprogramowanie czy interfejsy aplikacji służących uzyskiwaniu dostępu do urządzeń. Dzięki temu obecnie każdy jest w stanie wykonywać proste obliczenia na komputerach kwantowych.

Powszechnie wiadomo jednak, że urządzenia typu NISQ mają swoje ograniczenia. Powstaje więc naturalne pytanie: w jakim stopniu wiarygodne będą obliczenia przeprowadzane z ich wykorzystaniem? Odpowiedzi na to pytanie może dostarczyć nam proces walidacji. Przez pojęcie walidacji architektur kwantowych rozumiemy testowanie poprawności ich funkcjonowania i zdolności do wykonywania zadań, do których zostały zaprojektowane. Ze względu na licznie występujące błędy istnieje zatem istotna potrzeba opracowania procesów walidujących, które pozwolą dokładnie ocenić, na ile precyzyjnie działają pewne platformy obliczeniowe.

Celem niniejszej dysertacji jest opracowanie sposobów walidacji współczesnych architektur kwantowych typu NISQ opartych na bramkowym modelu obliczeń. W pracy analizujemy zarówno aspekty teoretyczne, jak i inżynierskie. Szczególną uwagę poświęcamy konstruowaniu metod walidacji, a następnie implementacji modeli na współczesnych architekturach kwantowych. Chcemy bowiem pokazać, że skonstruowane przez nas modele teoretyczne pozwolą również na uzyskanie praktycznych metod benchmarkowania współczesnych systemów kwantowych. Ostatnim, lecz nie mniej ważnym aspektem dysertacji jest wdrożenie i zaimplementowanie powstałych algorytmów na urządzeniach kwantowych.

Rozpoczynamy od rozpatrzenia metody walidacji opartej na schemacie uczenia się pomiarów von Neumanna. W tym przypadku rozważamy sytuację, w której używamy nieznanego pomiaru  $N$  razy eksperymentalnie, a następnie chcemy przybliżyć jego postać, gdy utracimy do niego dostęp. Strategię tę zwykle wykonuje się w dwóch etapach. Pierwszy z nich polega na przygotowaniu pewnego początkowego stanu kwantowego i zaimplementowaniu nieznanego pomiaru  $N$  razy, co pozwala nam przechowywać informację o tym pomiarze do późniejszego wykorzystania. Natomiast drugi etap polega na zaimplementowaniu operacji odzyskiwania, dzięki której otrzymujemy przybliżenie nieznanego pomiaru. Schemat jest optymalny, gdy osiąga najwyższą możliwą wierność przybliżenia. Zadanie to ma na celu wyznaczenie wierności przybliżenia oryginalnego pomiaru względem pomiaru nauczonego. Ponadto chcemy wyznaczyć najefektywniejszy algorytm, który realizuje to zadanie.

W standardowym schemacie uczenia się pomiarów von Neumanna wykorzystywane są sieci kwantowe, znane również jako kwantowe komby. W rozprawie rozszerzamy to podejście dzięki wykorzystaniu kwantowej teorii struktur przyczynowo-skutkowych. Pokazujemy potencjał i przewagę nad standardowym schematem użycia tej teorii w problemie uczenia się pomiarów von Neumanna. W tym celu opisujemy część strategii odpowiedzialną za przechowywanie operacji kwantowej za pomocą macierzy procesu, i wówczas liczymy wartość funkcji wierności. W efekcie funkcja wierności przybliżenia rośnie, a co za tym idzie – zwiększa się precyzja pomiaru nauczonego.

W naszej pracy wprowadzamy również metodę walidacji na podstawie schematu rozróżniania pomiarów von Neumanna. Aby zobrazować schemat rozróżniania, rozpatrzmy eksperyment, w którym wykorzystujemy nieznaną urządzenie pomiarowe. Wiemy jedynie, że wykonuje ono jeden z dwóch pomiarów kwantowych. Naszym celem jest wskazanie – możliwie z jak największym prawdopodobieństwem – który z pomiarów został wykonany w trakcie eksperymentu. Następnie chcielibyśmy opracować optymalną strategię rozróżniania, czyli taką, dla której uzyskujemy maksymalne prawdopodobieństwo poprawnego rozróżnienia. Będziemy zainteresowani głównie schematem rozróżniania w zakresie pewnej rodziny kubitowych pomiarów von Neumanna. Dla tej rodziny pomiarów obliczymy dokładną wartość prawdopodobieństwa poprawnego rozróżnienia i skonstruujemy optymalną strategię.

Jako inżynierski aspekt pracy wprowadzamy PyQBench – innowacyjną platformę open source przeznaczoną do testów porównawczych komputerów kwantowych opartych na bramkowym modelu obliczeń. PyQBench weryfikuje poprawność działania urządzeń typu NISQ dzięki wykorzystaniu schematu rozróżniania pomiarów von Neumanna. Oferujemy potencjalnym użytkownikom także interfejs wiersza poleceń (CLI) do uruchamiania testów porównawczych przy użyciu parametryzowanej rodziny pomiarów kubitowych w bazie Fouriera. Dodatkowo – w przypadku

bardziej zaawansowanych scenariuszy rozróżniania – PyQBench może być wykorzystywany jako biblioteka Python. Dzięki PyQBench pokażemy, że zaproponowane modele walidacji i uzyskane wyniki przyczyniły się do powstania nowego aspektu benchmarkingu urządzeń typu NISQ.

Rozważamy również schemat certyfikacji pomiarów von Neumanna rozumiany jako kwantowe testowanie hipotez. W tej rozprawie jesteśmy zainteresowani dwupunktowym schematem certyfikacji, w którym hipotezy – zerowa i alternatywna – są zbiorami jednoelementowymi. Naszym celem jest zminimalizowanie prawdopodobieństwa popełnienia błędu II rodzaju, przy założeniu pewnego ustalonego poziomu istotności. Podobnie jak w zadaniu rozróżniania chcemy wyznaczyć optymalną strategię certyfikacji, dla której prawdopodobieństwo popełnienia błędu II rodzaju będzie minimalne. Ponownie szczególną uwagę poświęcamy schematowi certyfikacji dla pewnej rodziny kubitowych pomiarów von Neumanna. Na podstawie otrzymanych wyników mamy zamiar rozszerzyć platformę PyQBench o kolejny schemat benchmarkowania komputerów kwantowych typu NISQ.

Wyniki przedstawione w niniejszej rozprawie można streścić w dwóch hipotezach:

1. Wykorzystanie kwantowej teorii struktur przyczynowo-skutkowych w problemie uczenia się pomiarów von Neumanna zapewnia wydajniejsze metody walidacji urządzeń kwantowych.

2. Metody walidacji, których podstawą jest zadanie rozróżniania i certyfikacji pomiarów von Neumanna, stanowią efektywne narzędzie benchmarkowania współczesnych komputerów kwantowych opartych na bramkowym modelu obliczeń.

Niniejsza praca składa się z dziewięciu rozdziałów oraz trzech dodatków. Pierwszy rozdział zawiera wstęp do teorii informacji kwantowej i opis motywacji prowadzonych badań. W Rozdziale 2 zamieszczono wprowadzenie do matematycznego języka informatyki kwantowej. Rozdział 3 poświęcony jest podstawowym pojęciom wykorzystywanym w teorii informacji kwantowej. Pozostała część dysertacji została napisana na podstawie trzech opublikowanych artykułów naukowych oraz jednego artykułu będącego w procesie recenzji.

Pierwsza z wymienionych wyżej prac [1], zaprezentowana w Rozdziale 4, dotyczy schematu uczenia się pomiarów von Neumanna. Następnie w Rozdziale 5 przedstawiono możliwość wykorzystania kwantowej teorii struktur przyczynowo-skutkowych w problemie uczenia się pomiarów von Neumanna. Rozdział ten stanowi mój autorski wkład w dysertację. W Rozdziale 6 omówiono kolejną metodę walidacji, jaką jest rozróżnianie pomiarów kwantowych. Należy nadmienić, że Rozdział 6 jest inspirowany pracą [2], w odniesieniu do której rozważany jest problem rozróżniania kubitowej rodziny pomiarów w bazie Fouriera. Ta część rozdziału stanowi mój autorski wkład w niniejszą dysertację. W rozdziale 7 zaprezentowano PyQBench, czyli innowacyjną platformę typu open source do testów porównawczych komputerów kwantowych wykorzystujących bramkowy model obliczeń. Rozdział ten został

napisany na podstawie pracy [3]. Natomiast w rozdziale 8 przedstawiono schemat certyfikacji pomiarów von Neumanna. W pierwszej części teoretycznych rozważań skupiono się na opracowaniu schematu certyfikacji dla dowolnie zdefiniowanych pomiarów von Neumanna, a następnie – na stworzeniu benchmarku komputerów kwantowych typu NISQ na podstawie schematu certyfikacji kubitowych pomiarów w bazie Fouriera. Rozdział ten został napisany z wykorzystaniem pracy [4] w części poświęconej omówieniu schematu certyfikacji dla stanów czystych, kanałów unitarych i pomiarów von Neumanna. Użycie schematu certyfikacji do benchmarkowania komputerów kwantowych stanowi mój autorski wkład w dysertację. Rozdział 9 zawiera wnioski z rozprawy i podsumowanie wyników przedstawionych badań. Do tekstu głównego załączono również trzy dodatki, które zawierają dowody twierdzeń, lematy i przykłady zbyt długie i techniczne, aby można je było umieścić w tekście głównym.



# Abstract in English

In the last decade the idea of quantum computing has become a reality. Noisy Intermediate-Scale Quantum (NISQ) devices are storming the market, with a wide selection of devices based on different architectures and accompanying software solutions. Among hardware providers offering public access to their gate-based devices, one could mention Rigetti, IBM, Oxford Quantum Group, IonQ or Xanadu. Other vendors offer devices operating in different paradigms. Especially, one could mention D-Wave and their quantum annealers or QuEra devices based on neutral atoms. Most vendors provide their own software stack and application programming interface for accessing their devices. Nowadays, everyone can make simple computations on these devices.

It is well known that NISQ devices have limitations. For that, a natural question arises: what extent can those devices perform meaningful computations? To answer this question, one has to devise a methodology for validating them. As a validation of quantum architectures, we refer to testing the correctness of their functioning and ability to perform the tasks they were designed for. The validation task has been highlighted as a significant challenge to scalable quantum computing technology. Due to the numerous errors, there is therefore a significant need to develop validation processes that allow for the best imaging of the accuracy and precision of the operation of computing platforms.

This dissertation aims to investigate a new validation methods for modern gate model NISQ devices. In the work, we analyze both theoretical and engineering aspects. We focus on the construction of validation methods and their adaptation to available NISQ quantum architectures. We want to show that the created theoretical models will also allow obtaining new concepts of benchmarking modern quantum systems. Another aspect of the dissertation is the implementation of obtained algorithms on quantum NISQ devices.

Initially, we consider a validation method based on the problem of learning von Neumann measurements known also as the storage and retrieval (SAR). In the general approach of SAR, we want to approximate an unknown von Neumann measurement which we were able to perform  $N$  times experimentally. This strategy usually consists of preparing some initial quantum state, applying the unknown

measurement  $N$  times, which allows us to store this operation for later use, and finally, a retrieval operation that returns an approximation of the unknown measurement. The scheme is optimal when it achieves the highest possible fidelity of the approximation. Our main goal is to estimate the asymptotic behaviour of the maximum value of the average fidelity function for storage and retrieval of von Neumann measurements and determine possibly the best approximation of the optimal scheme.

The primary tools used in storage and retrieval are quantum combs. In addition to quantum combs, we will also use the quantum causal structures theory – a completely new approach in quantum information theory. We will show the potential and advantage of this approach in the problem of learning quantum operations. For this purpose, we describe the part responsible for storing the quantum operation with a process matrix and then calculate the value of the fidelity function.

Next, we introduce a validation method based on the scheme of discrimination of von Neumann measurements. To illustrate the scheme of discrimination, let us consider an experiment in which we use an unknown measurement device. The only information we have is that it performs one of two measurements. Our goal is to indicate, with as high probability as possible, which of the measurements was used during the experiment. Next, we would like to construct an optimal discrimination strategy, for which we get the maximum probability of correct discrimination. We will be mainly interested in the scheme for distinguishing between a family of qubit von Neumann measurements. For this case, we will calculate the exact value of the probability of correct discrimination and construct the optimal strategy.

As an engineering aspect of the dissertation, we introduce PyQBench – an innovative open-source framework for benchmarking gate-based quantum computers. PyQBench benchmarks NISQ devices by verifying their capability based on the discrimination scheme. PyQBench offers a simplified, ready-to-use, command line interface (CLI) for running benchmarks using a predefined parametrized Fourier family of measurements. For more advanced scenarios, PyQBench offers a way of employing user-defined measurements instead of predefined ones. We will show that the proposed models and obtained results have led to a new aspect of benchmarking NISQ devices.

Finally, we consider the task of certification of von Neumann measurements. Here, we are interested in a binary certification scheme in which the null and alternative hypotheses are single-element sets. The goal of certification is to minimize the probability of the type II error given some fixed statistical significance. Next, we would like to construct an optimal certification strategy, which minimizes such a probability. Again, for the parameterized family of qubit measurements in the Fourier basis we calculate the exact value of the probability of the type II error and create the optimal certification scheme. Based on the obtained results,

we extend PyQBench creating benchmarks of NISQ devices using certification approach.

The results presented in this dissertation can be summarized in two hypotheses.

1. The usage of the quantum causal structure theory provides more efficient methods for the problem of learning von Neumann measurements.
2. The problem of distinguishing and certifying von Neumann measurements can be used to create a new aspect of benchmarking for modern gate model-inspired NISQ devices.

The work consists of nine chapters. The first Chapter presents an introduction to quantum information theory and the motivation for the research. Chapter 2 presents the necessary mathematical framework. Chapter 3 devotes the basic concepts used in quantum information theory. The rest of the dissertation was based on three published articles and one preprint.

The first paper [1], described in Chapter 4, concerns the von Neumann measurement learning scheme. Next, in Chapter 5, we explore the possibility of using the quantum causal structure theory in the task of von Neumann measurement learning. This Chapter is my contribution to the dissertation. In Chapter 6, we focus on the problem of distinguishing von Neumann measurements. This Chapter is inspired by the work [2]. Due to this work, we solve the problem of discrimination for the parameterized family of qubit von Neumann measurements in the Fourier basis, which is my contribution to the dissertation. In Chapter 7, we introduce PyQBench, an innovative open-source framework for benchmarking gate-based quantum computers. This chapter is written based on [3]. In Chapter 8 we will focus on a validation scheme by certifying two von Neumann measurements. This chapter is partly based on the work [4] describing the certification scheme of pure states, unitary channels, and von Neumann measurements. My contribution to the dissertation is using a certification scheme of von Neumann measurements for benchmarking quantum computers. Chapter 9 contains the conclusions of the dissertation and summarizes the results of the presented research. We also include three appendices that contain e.g. proofs of theorems, lemmas or examples too long and technical to be stated in the main text.



# Acknowledgment

For various reasons, I would like to show gratitude my mentor Karol Życzkowski, Ms. Anna Askaldowicz and my colleagues Ryszard Kukulski, Konrad Jałowiecki, Aleksandra Krawiec, Hanna Wojewódka-Ściążko, Bartłomiej Gardas and Piotr Gawron.

I acknowledge by the project „Near-term quantum computers Challenges, optimal implementations and applications” under Grant Number POIR.04.04.00-00-17C1/18-00, which is carried out within the Team-Net programme of the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund and European Unionscholarship through the European Social Fund, grant InterPOWER (POWR.03.05.00-00-Z305).

A special thanks go to Zbigniew Puchała and Łukasz Paweła for continuous support of my research.



# Chapter 1

## Introduction

In the last decade the idea of quantum computing has become a reality. Noisy Intermediate-Scale Quantum (NISQ) [5] devices are storming the market with a wide selection of devices based on different architectures and accompanying software solutions. Among hardware providers offering public access to their gate-based devices, one could mention IBM [6], Rigetti [7], Oxford Quantum Group [8], IonQ [9], or Xanadu [10]. Other vendors offer devices operating in different paradigms. Notably, one could mention D-Wave [11] and their quantum annealers, or QuEra devices [12] based on neutral atoms. Most vendors provide their own software stack and application programming interface for accessing their devices. To name a few, Rigetti’s computers are available through their Forest SDK [13] and PyQuil library [14] and IBM Q [6] computers can be accessed through Qiskit [15] or IBM Quantum Experience web interface [16]. Some cloud services, like Amazon Braket [17], offer access to several quantum devices under an unified API. On top of that, several libraries and frameworks can integrate with different hardware vendors. Examples of such frameworks include IBM Q’s Qiskit or Zapata Computing’s Orquestra [18]. Nowadays, everyone can make simple computations on these devices.

It is well known that NISQ devices have limitations [19]. For that, a natural question arises: what extent can those devices perform meaningful computations? To answer this question, one has to devise a methodology for validating them. As a validation of quantum architectures, we refer to testing the correctness of their functioning and ability to perform the tasks they were designed for. The validation task has been highlighted as a significant challenge to scalable quantum computing technology. Due to the numerous errors, there is therefore a significant need to develop validation processes that allow for the best imaging of the accuracy and precision of the operation of computing platforms. The search for practical and reliable tools for the validation of quantum architecture has attracted a lot of attention in recent years [20–25].

There are many approaches to testing NISQ devices. The method of choice may depend on, for example, the amount and type of information we have or want to obtain, the system size or structure, the computational resources, or obtained noise. Many times, the complexity of a protocol can be traded for the amount of information about the validated device (Fig. 1.1).

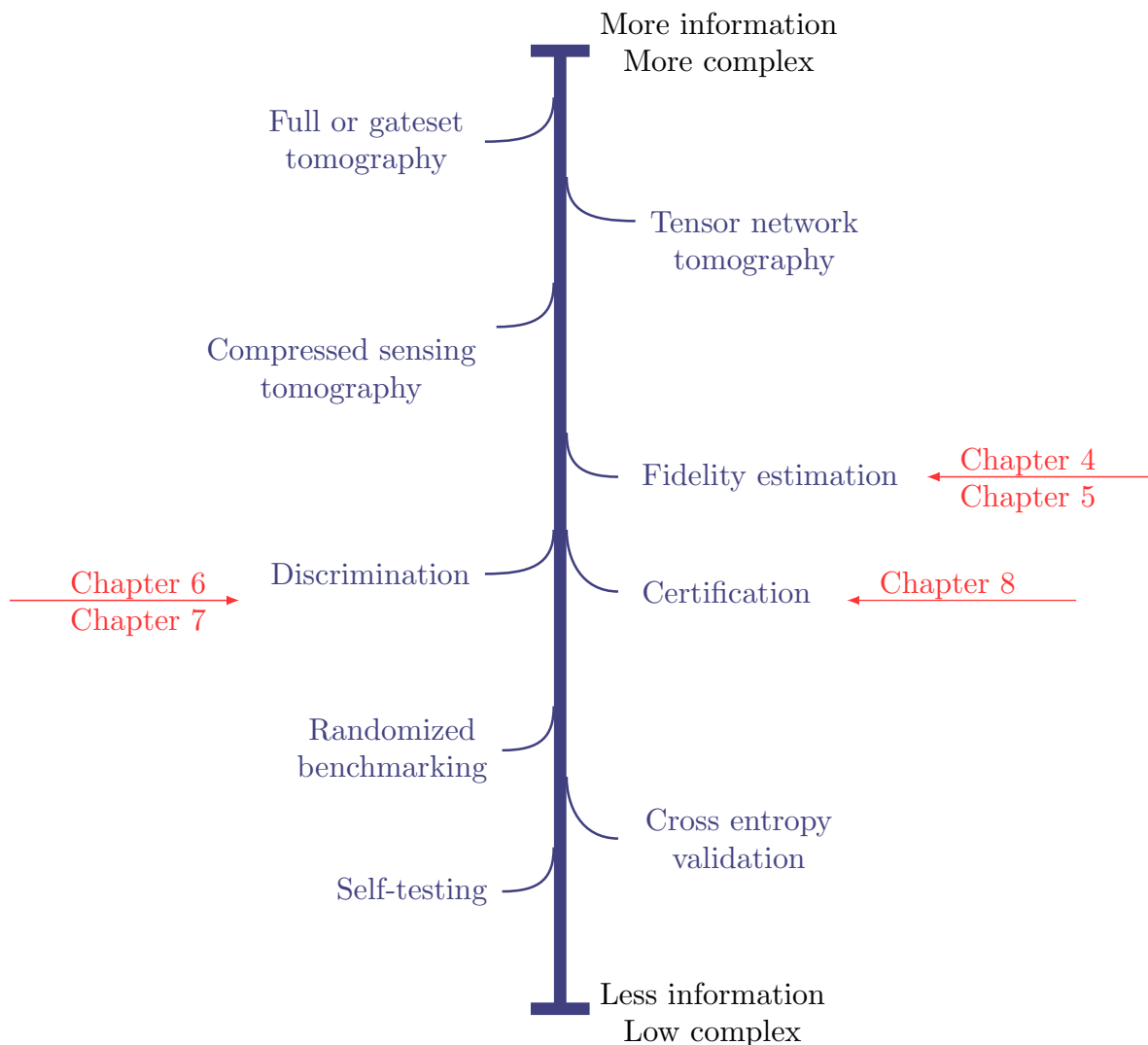


Figure 1.1: Validation methods of NISQ devices depending on the amount of information and complexity.

The most powerful but at the same time most resources demanded validating techniques are the full quantum tomography [26,27] and the gateset tomography [28,29]. The first idea is to obtain knowledge of the entire quantum state or



transformation by performing sufficiently many measurements. The second one, instead of focusing on a single component of the experiment, characterizes an entire set of quantum gates used during the experiment. However, as we mentioned before, the tomography process is excessively costly in the size of the quantum system. Fortunately, many quantum states and operations used in realistic experiments have strict structures. For example, quantum states are often close to being pure or having a fixed low rank. For such cases, one may use compressed sensing tomography [30]. Another approach is tensor network tomography which gives excellent approximation under the assumption reconstructing of a quantum state by product operators [31, 32].

In contrast to quantum tomography, fidelity estimation aims merely at determining the distance between the actual quantum state or operation and the theoretical one. While fidelity estimation yields much less information than full tomography, one saves tremendously in measurement, sample complexity and resources. The initial research estimated the fidelity of an imperfect preparation of pure quantum states [33]. This protocol is extended to estimating the fidelity of quantum channels [34] and, as we see in the dissertation, von Neumann measurements [1].

Another quantity used for validating NISQ devices is certification of quantum operations [35, 36], which can be viewed as the extension of quantum hypotheses testing. The standard certification scheme assumes that we have two hypotheses – the null and the alternative and there is possibly one of two outcomes: either we accept or reject the null hypothesis. Like in classical hypothesis testing, here, we also have two possible types of errors. The type I error happens if we reject the null hypothesis when it is actually true, whereas the type II error happens if we accept the null hypothesis when we should have rejected it. The main aim of certification is finding the optimal strategy that minimizes one type of error when the other is fixed.

Certification of quantum objects is closely related to the other well-known method of validation, which is the problem of discrimination of those objects [37]. Intuitively, in the discrimination problem we are given one of two quantum objects sampled according to a given a priori probability distribution. Hence, the probability of making an error in the discrimination task is equal to the average of the type I and type II errors over the assumed probability distribution. Therefore, the discrimination problem can be seen as symmetric distinguishability instead of certification, which is asymmetric. In other words, the main difference between both approaches is that the main task of discrimination is the minimization over the average of both types of possible errors, while the certification concerns the minimization over one type of error when the bound of the other one is assumed. Both symmetric and asymmetric discrimination schemes have been developed for quantum states [38], unitary channels [36, 38] and general quantum channels [39], SIC

POVM [40], or unknown quantum measurements [41]. This dissertation will extend these issues to the discrimination task of the von Neumann measurements [2, 42].

A still weaker method aims at the concept of randomized benchmarking [43]. In this approach, one sample circuits to be run from some predefined set of gates (e.g. from the Clifford group [44], or the random unitary gates [45]) and tests how much the output distribution obtained from the device running these circuits differs from the ideal one. It is also common to concatenate randomly chosen circuits with their inverses, which should yield the identity circuit, and run those concatenated circuits on the device.

We also mention validating methods characterizing an arbitrary quantum system based only on its classical input-output correlations. Examples include the cross-entropy benchmarking [46], or self-testing [47]. Building upon such notions of fidelities, specific quality measures have been introduced in different contexts. Examples include the cross-entropy function [48] strictly related to the maximum likelihood. Like the classical cross-entropy validation, this task aimed at verifying classical distributions and calculating the entropy function between the ideal and sampled distribution of the quantum system.

The work [49] has given birth to the field of self-testing. This work sets the terminology and formalism which was adopted by later works. In particular, a self-testing protocol can be seen as a device-independent validation of a quantum system, assuming that the system can be prepared many times in an independent, identically distributed manner.

In the scope of the dissertation, we aim to investigate new validation methods for modern gate model-inspired NISQ devices. In the work, we analyze both theoretical and engineering aspects. Initially, the key aspect of the dissertation is to constrain validation methods. Secondly, we focus on introducing new concepts for benchmarking current available NISQ devices.

The first method concerns the quantum learning of von Neumann measurements. This approach is also known in the literature as storage and retrieval (SAR). In the general approach of SAR, we want to approximate an unknown von Neumann measurement which we were able to perform  $N$  times experimentally. This strategy is usually divided into two parts. The first one consists of preparing some initial quantum state and applying the unknown measurement  $N$  times, which allows us to store this operation for later use. The second one, whereas, consists a retrieval operation that returns an approximation of the unknown measurement. The scheme is optimal when it achieves the highest possible fidelity of the approximation. Our main goal is to estimate the asymptotic behaviour of the maximum value of the average fidelity function for storage and retrieval of von Neumann measurements and determine possibly the best approximation of the optimal scheme.

The main tools used in SAR are quantum networks, also known as quantum

combs. Additionally, we will also use the quantum causal structures theory – a completely new approach in quantum information theory. We explore the possibility of using the quantum causal structure theory in the task of storage and retrieval of von Neumann measurements. For this purpose, we describe the part responsible for storing the quantum operation with a process matrix and then calculate the value of the fidelity function. This idea will be focused around **Hypothesis 1**.

*The usage of quantum causal structure theory can improve the value of fidelity function, providing a more efficient method of the storage and retrieval of von Neumann measurements.*

Next, we introduce a validation method based on the scheme of discrimination of von Neumann measurements. We have calculated the maximum value of the probability of correct discrimination between the parameterized family of qubit measurements in the Fourier basis and computational basis. Furthermore, we construct the optimal strategy, which maximizes the probability for this case. Next, we consider the certification task between two von Neumann measurements. Here, we are interested in a binary certification scheme in which the null and alternative hypotheses are single-element sets. The goal is to minimize the probability of the type II error given some fixed statistical significance. Again, for the parameterized family of qubit measurements in the Fourier basis we calculate the exact value of the probability of the type II error and create the optimal certification scheme.

As an engineering aspect of the dissertation, we introduce PyQBench – an innovative open-source framework for benchmarking gate-based quantum computers. PyQBench benchmarks NISQ devices by verifying their capability based on the discrimination and certification scheme. PyQBench offers a simplified, ready-to-use, command line interface (CLI) for running benchmarks using a predefined parametrized Fourier family of measurements. For more advanced scenarios, PyQBench offers a way of employing user-defined measurements instead of predefined ones. We will show that the proposed models and obtained results have led to a new aspect of benchmarking NISQ devices. Due to that, we formulate **Hypothesis 2**.

*Validating techniques based on discrimination and certification of von Neumann measurements can provide efficient methods for benchmarking current gate model-inspired NISQ devices.*

The work consists of nine chapters. The first Chapter presents an introduction to quantum information theory and the motivation for the research. Chapter 2 presents necessary mathematical framework. Chapter 3 devotes the basic concepts used in quantum information theory. The rest of the dissertation was based on three published articles and one preprint.

The first paper [1], described in Chapter 4, concerns the storage and retrieval of von Neumann measurements. In Chapter 5, we explore the possibility of using the quantum causal structure theory in the task of von Neumann measurement learning. This chapter is my contribution to the dissertation. In Chapter 6, we focus on the problem of discrimination von Neumann measurements. Due to obtained results, we introduce PyQBench, an innovative open-source framework for benchmarking gate-based quantum computers. We put the detailed description of PyQBench [3] in Chapter 7. Next, the work [4] presented in Chapter 8 focuses on another validation scheme based on the certification of von Neumann measurements. This chapter is partly based on the work [4] describing the certification scheme of pure states, unitary channels, and von Neumann measurements. My contribution to the dissertation is using a certification scheme of von Neumann measurements for benchmarking quantum computers. Chapter 9 contains the conclusions of the dissertation and summarizes the results of the presented research. We also include three appendices that contain e.g. proofs of theorems too long and technical to be stated in the main text.

# Chapter 2

## Mathematical preliminaries

Quantum information theory relies heavily on linear algebra and functional analysis in finite-dimensional Hilbert spaces. This section that follows presents an overview of the mathematical aspects that are most relevant to the theory of quantum information to be used throughout this dissertation. This chapter is written mostly based on the books [37, 50]. Unlike the other chapters in this dissertation, the present chapter does not include proofs.

### 2.1 Complex Euclidean spaces and Dirac notation

Let us consider complex Euclidean space denoted by scripted capital letters near the end of the alphabet, such as  $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ . The dimension of a complex Euclidean space  $\mathcal{X}$  will be denoted by  $\dim(\mathcal{X})$ .

In this dissertation we will use the Dirac notation. Its use in quantum mechanics is quite widespread to denote complex vectors and linear forms. This notation was introduced by Paul Dirac in 1939.

Let  $\mathcal{X}$  be a complex Euclidean space of dimension  $\dim(\mathcal{X}) = d$ . In Dirac notation, a column vector in  $\mathcal{X}$  is called a ket and denoted by

$$|\psi\rangle := (\psi_0, \dots, \psi_{d-1})^\top, \quad \psi_i \in \mathbb{C}, \quad i \in \{0, \dots, d-1\}, \quad (2.1)$$

whereas a row vector in  $\mathcal{X}^*$  is called a bra and denoted by

$$\langle\psi| := (\overline{\psi_0}, \dots, \overline{\psi_{d-1}}), \quad (2.2)$$

where  $\overline{\psi_i}$  denotes complex conjugate of  $\psi_i$ . From Riesz's representation theorem, each finite-dimensional complex Euclidean space  $\mathcal{X}$  and its dual space are isometrically isomorphic. Then, the isomorphism defining a one-to-one mapping between kets and bras are indicated as  $\dagger$ , that is  $|\psi\rangle^\dagger := \langle\psi|$ . Whereas, the inner product

between two vectors  $|\psi\rangle, |\phi\rangle$  is denoted by  $\langle\psi|\phi\rangle$ , more precisely

$$\langle\psi|\phi\rangle = \sum_{i=0}^{d-1} \overline{\psi_i} \phi_i. \quad (2.3)$$

The inner product satisfies the following conditions:

1. Linearity in the second argument:

$$\langle\psi|\alpha\phi_1 + \beta\phi_2\rangle = \alpha\langle\psi|\phi_1\rangle + \beta\langle\psi|\phi_2\rangle, \quad (2.4)$$

for all  $|\psi\rangle, |\phi_1\rangle, |\phi_2\rangle \in \mathcal{X}$  and scalars  $\alpha, \beta \in \mathbb{C}$ .

2. Conjugate symmetry

$$\langle\psi|\phi\rangle = \overline{\langle\phi|\psi\rangle}, \quad (2.5)$$

for all  $|\psi\rangle, |\phi\rangle \in \mathcal{X}$ .

3. Positive definiteness:

$$\langle\psi|\psi\rangle \geq 0, \quad (2.6)$$

for all  $|\psi\rangle \in \mathcal{X}$ , and  $\langle\psi|\psi\rangle = 0$  if and only if  $|\psi\rangle = 0$ .

The induced Euclidean norm of a vector  $|\psi\rangle \in \mathcal{X}$  is defined as

$$\| |\psi\rangle \| = \sqrt{\langle\psi|\psi\rangle}. \quad (2.7)$$

The Euclidean norm is the special case ( $p = 2$ ) of a wider class of  $p$ -norms defined for each  $|\psi\rangle \in \mathcal{X}$  as

$$\| |\psi\rangle \|_p = \left( \sum_{i=0}^{d-1} |\psi_i|^p \right)^{\frac{1}{p}}, \quad (2.8)$$

for  $p < \infty$ . Moreover, we define  $\infty$ -norm as

$$\| |\psi\rangle \|_\infty = \max\{|\psi_i| : i \in \{0, \dots, d-1\}\}. \quad (2.9)$$

Two vectors  $|\psi\rangle, |\phi\rangle \in \mathcal{X}$  are said to be orthogonal if  $\langle\psi|\phi\rangle = 0$ . We say that a collection of vectors  $\{|\psi_i\rangle\}_{i=0}^k$  creates an orthogonal set if for each pair of vectors holds  $\langle\psi_i|\psi_j\rangle = 0$  for all choices  $i, j$  such that  $i \neq j$ . Naturally, a collection of nonzero orthogonal vectors is necessarily linearly independent. An orthogonal set  $\{|\psi_i\rangle\}_{i=0}^k$  is called orthonormal if it additionally holds  $\langle\psi_i|\psi_i\rangle = 1$  for all  $i \in \{0, \dots, k\}$ . If  $k = \dim(\mathcal{X}) - 1$ , then the orthogonal set  $\{|\psi_i\rangle\}_{i=0}^{\dim(\mathcal{X})-1}$  forms a basis of  $\mathcal{X}$ . Nevertheless, if the basis is additionally orthonormal set, then it is called an orthonormal basis. We distinguish the orthonormal basis of the form

$\{|i\rangle\}_{i=0}^{d-1}$ , where  $|i\rangle$  is a vector with the  $i$ -th entry equal to one and others equal to zero. Such a basis is called the computational basis.

## Direct sum and tensor product of complex Euclidean spaces

Let  $|x\rangle = (x_0, \dots, x_k)^\top \in \mathcal{X}$  and  $|y\rangle = (y_0, \dots, y_l)^\top \in \mathcal{Y}$ . We define the direct sum of  $|x\rangle$  and  $|y\rangle$  as

$$|x\rangle \oplus |y\rangle = (x_0, \dots, x_k, y_0, \dots, y_l)^\top, \quad (2.10)$$

We define the direct sum of spaces  $\mathcal{X}$  and  $\mathcal{Y}$ , denoted by  $\mathcal{X} \oplus \mathcal{Y}$  as

$$\mathcal{X} \oplus \mathcal{Y} = \text{span} \{|x\rangle \oplus |y\rangle : |x\rangle \in \mathcal{X}, |y\rangle \in \mathcal{Y}\}. \quad (2.11)$$

Observe that  $\dim(\mathcal{X} \oplus \mathcal{Y}) = \dim(\mathcal{X}) + \dim(\mathcal{Y})$ . We can define the direct sum for more than two complex Euclidean spaces in an analogous way.

We further introduce the notion of the tensor product of complex Euclidean spaces. We define the tensor product of  $|x\rangle \in \mathcal{X}$  and  $|y\rangle \in \mathcal{Y}$  as

$$|x\rangle \otimes |y\rangle = (x_0 \cdot y_0, \dots, x_0 \cdot y_l, \dots, x_k \cdot y_0, \dots, x_k \cdot y_l)^\top, \quad (2.12)$$

By  $\mathcal{X} \otimes \mathcal{Y}$  we denote the tensor product of spaces  $\mathcal{X}$  and  $\mathcal{Y}$  given by

$$\mathcal{X} \otimes \mathcal{Y} = \text{span} \{|x\rangle \otimes |y\rangle : |x\rangle \in \mathcal{X}, |y\rangle \in \mathcal{Y}\}. \quad (2.13)$$

Observe,  $\dim(\mathcal{X} \otimes \mathcal{Y}) = \dim(\mathcal{X}) \cdot \dim(\mathcal{Y})$ . As previously we can generalize the tensor product for more than two complex Euclidean spaces.

Let  $\mathcal{X}_0, \dots, \mathcal{X}_{n-1}$  be Euclidean spaces. Then, we fix the notation

$$\bigoplus_{i=0}^{n-1} \mathcal{X}_i = \mathcal{X}_0 \oplus \dots \oplus \mathcal{X}_{n-1} \quad (2.14)$$

and

$$\bigotimes_{i=0}^{n-1} \mathcal{X}_i = \mathcal{X}_0 \otimes \dots \otimes \mathcal{X}_{n-1}, \quad (2.15)$$

respectively. It is often convenient to make the identification  $\mathcal{X}^{\oplus n}$  or  $\mathcal{X}^{\otimes n}$  under the assumption that  $\mathcal{X}_0, \dots, \mathcal{X}_{n-1}$  and  $\mathcal{X}$  refer to the same complex Euclidean space. Similarly, we will use a notation  $|x\rangle^{\oplus n}$  and  $|x\rangle^{\otimes n}$  refer to the  $n$ -fold direct sum and tensor product of  $|x\rangle$ .

## 2.2 Linear operators

For a given complex Euclidean spaces  $\mathcal{X}$  and  $\mathcal{Y}$ , one writes  $L(\mathcal{X}, \mathcal{Y})$  to refer to the set of all linear operators of the form

$$A : \mathcal{X} \rightarrow \mathcal{Y}. \quad (2.16)$$

For simplify notation, let  $L(\mathcal{X}) := L(\mathcal{X}, \mathcal{X})$ . The identity operator will be denoted by  $\mathbb{1}_{\mathcal{X}} \in L(\mathcal{X})$ .

The set  $L(\mathcal{X}, \mathcal{Y})$  forms a complex vector space when addition and scalar multiplication are defined as follows:

1. Addition: For operators  $A, B \in L(\mathcal{X}, \mathcal{Y})$ , the operator  $A + B \in L(\mathcal{X}, \mathcal{Y})$  is defined by the equation  $(A + B)|x\rangle = A|x\rangle + B|x\rangle$  for all  $|x\rangle \in \mathcal{X}$ .
2. Scalar multiplication: for an operator  $A \in L(\mathcal{X}, \mathcal{Y})$  and a scalar  $\alpha \in \mathbb{C}$  the operator  $\alpha A \in L(\mathcal{X}, \mathcal{Y})$  is defined by the equation  $(\alpha A)|x\rangle = \alpha A|x\rangle$  for all  $|x\rangle \in \mathcal{X}$ .

### Matrices and their correspondence with operators

For any choice of complex Euclidean spaces  $\mathcal{X}$  and  $\mathcal{Y}$ , there is a bijective linear correspondence between the set of operators  $L(\mathcal{X}, \mathcal{Y})$  and the collection of all matrices of size  $\dim(\mathcal{Y}) \times \dim(\mathcal{X})$  having the form

$$M = (m_{ij})_{\substack{i \in \{0, \dots, \dim(\mathcal{Y})-1\} \\ j \in \{0, \dots, \dim(\mathcal{X})-1\}}}. \quad (2.17)$$

For each operator  $X \in L(\mathcal{X}, \mathcal{Y})$  we associate the matrix  $M$  defined as  $m_{ij} = \langle i|X|j\rangle$  for  $i \in \{0, \dots, \dim(\mathcal{Y}) - 1\}$ ,  $j \in \{0, \dots, \dim(\mathcal{X}) - 1\}$ ,  $|i\rangle \in \mathcal{Y}$ , and  $|j\rangle \in \mathcal{X}$ . The operator  $A$  is uniquely determined by the matrix  $M$ . With respect to this correspondence, matrix multiplication is equivalent to operator composition. Hereafter in this dissertation, linear operators will be associated with matrices implicitly, and hence we will be using the words operator and matrix interchangeably.

### Direct sum and tensor product of linear operators

Let  $\mathcal{X}_0, \mathcal{X}_1, \mathcal{Y}_0, \mathcal{Y}_1$  be complex Euclidean spaces and let  $A_0 \in L(\mathcal{X}_0, \mathcal{Y}_0)$ ,  $A_1 \in L(\mathcal{X}_1, \mathcal{Y}_1)$  be linear operators. The direct product

$$A_0 \oplus A_1 \in L(\mathcal{X}_0 \oplus \mathcal{X}_1, \mathcal{Y}_0 \oplus \mathcal{Y}_1), \quad (2.18)$$



of these operators is uniquely defined by the equation

$$(A_0 \oplus A_1)(|x_0\rangle \oplus |x_1\rangle) = (A_0|x_0\rangle) \oplus (A_1|x_1\rangle), \quad (2.19)$$

for all choices  $|x_0\rangle \in \mathcal{X}_0$  and  $|x_1\rangle \in \mathcal{X}_1$ .

The tensor product

$$A_0 \otimes A_1 \in L(\mathcal{X}_0 \otimes \mathcal{X}_1, \mathcal{Y}_0 \otimes \mathcal{Y}_1), \quad (2.20)$$

of  $A_0 \in L(\mathcal{X}_0, \mathcal{Y}_0)$  and  $A_1 \in L(\mathcal{X}_1, \mathcal{Y}_1)$  to be the unique operator that satisfies

$$(A_0 \otimes A_1)(|x_0\rangle \otimes |x_1\rangle) = (A_0|x_0\rangle) \otimes (A_1|x_1\rangle), \quad (2.21)$$

for all choices  $|x_0\rangle \in \mathcal{X}_0$  and  $|x_1\rangle \in \mathcal{X}_1$ .

As previously, we can also generalize the notion of direct sum and tensor product for more than two linear operators. We will use a notation

$$A^{\oplus n} = \underbrace{A \oplus \dots \oplus A}_n, \quad (2.22)$$

and, for tensor product, we assume  $A^{\otimes n} = A \otimes \dots \otimes A$ .

## 2.2.1 Properties of linear operators

### Image and rank of an operator

The image of  $A \in L(\mathcal{X}, \mathcal{Y})$  is defined as a subspace of  $\mathcal{Y}$  given by

$$\text{im}(A) = \{A|x\rangle : |x\rangle \in \mathcal{X}\}. \quad (2.23)$$

The rank of an operator  $A \in L(\mathcal{X}, \mathcal{Y})$  is the dimension of the image of  $A$ , that means

$$\text{rank}(A) = \dim(\text{im}(A)). \quad (2.24)$$

### Trace of an operator

The trace of a square operator  $X \in L(\mathcal{X})$  is defined as the sum of its diagonal entries, that is  $\text{tr}(X) = \sum_{i=0}^{\dim(\mathcal{X})-1} x_{ii}$ . Equivalently, the trace is the unique linear function  $\text{tr} : L(\mathcal{X}) \rightarrow \mathbb{C}$  such that

$$\text{tr}(|x\rangle\langle y|) = \langle y|x\rangle, \quad (2.25)$$

for all vectors  $|x\rangle, |y\rangle \in \mathcal{X}$ . Using the trace, we define an inner product on the space  $L(\mathcal{X}, \mathcal{Y})$  as

$$\langle A, B \rangle = \text{tr}(A^\dagger B), \quad (2.26)$$

for all  $A, B \in L(\mathcal{X}, \mathcal{Y})$ , where  $A^\dagger$  denotes the Hermitian conjugate of  $A$ , that is  $A^\dagger = \overline{A^\top}$ .

The trace mapping possess a variety of properties, including the ones summarized in the following list:

1. The trace mapping is cyclic, that is

$$\text{tr}(AB) = \text{tr}(BA), \quad (2.27)$$

for any  $A \in L(\mathcal{X}, \mathcal{Y}), B \in L(\mathcal{Y}, \mathcal{X})$ .

2. For every  $X \in L(\mathcal{X})$  and  $Y \in L(\mathcal{Y})$  it holds that

$$\text{tr}(X \oplus Y) = \text{tr}(X) + \text{tr}(Y). \quad (2.28)$$

3. For every  $X \in L(\mathcal{X})$  and  $Y \in L(\mathcal{Y})$  it holds that

$$\text{tr}(X \otimes Y) = \text{tr}(X)\text{tr}(Y). \quad (2.29)$$

## Vectorization

There is a correspondence between the space  $L(\mathcal{X}, \mathcal{Y})$  and  $\mathcal{Y} \otimes \mathcal{X}$  for any choice of complex Euclidean spaces  $\mathcal{X}, \mathcal{Y}$ . This correspondence, called vectorization, is given by the linear mapping

$$|\cdot\rangle\rangle : L(\mathcal{X}, \mathcal{Y}) \rightarrow \mathcal{Y} \otimes \mathcal{X}, \quad (2.30)$$

defined for any  $A \in L(\mathcal{X}, \mathcal{Y})$  by

$$|A\rangle\rangle := \sum_{i=0}^{\dim(\mathcal{Y})-1} |i\rangle \otimes A^\top |i\rangle, \quad (2.31)$$

where  $|i\rangle \in \mathcal{Y}$ . This mapping is a linear bijection between spaces  $L(\mathcal{X}, \mathcal{Y})$  and  $\mathcal{Y} \otimes \mathcal{X}$ . It implies that every vector  $|x\rangle \in \mathcal{Y} \otimes \mathcal{X}$  can be uniquely determined by an operator  $A \in L(\mathcal{X}, \mathcal{Y})$  such that  $|A\rangle\rangle = |x\rangle$ . Moreover, it can be shown that the following holds

$$\langle A, B \rangle = \langle\langle A|B\rangle\rangle. \quad (2.32)$$

Another crucial property of the vectorization mapping will be useful throughout the dissertation, e.g.

$$(A \otimes B)|C\rangle\rangle = |ACB^\top\rangle\rangle, \quad (2.33)$$

for  $A \in L(\mathcal{X}, \mathcal{Y})$ ,  $B \in L(\mathcal{W}, \mathcal{Z})$  and  $C \in L(\mathcal{W}, \mathcal{X})$ . This property is sometimes called telegraphic notation.

## Types of linear operators

The following classes of operators have particular importance in the theory of quantum information.

The commutator  $AB - BA$  of operators  $A, B \in L(\mathcal{X})$  will be denoted by  $[A, B]$ . We call that an operator  $A \in L(\mathcal{X})$  is normal if it commutes with its Hermitian conjugate, that means  $[A, A^\dagger] = 0$ . A normal operator  $A \in L(\mathcal{X})$  is said to be Hermitian if satisfies the equation  $A = A^\dagger$ , while a Hermitian operator  $A \in L(\mathcal{X})$  is called positive semidefinite if it additionally satisfies  $\langle \psi | A | \psi \rangle \geq 0$  for every  $|\psi\rangle \in \mathcal{X}$ . The subset of  $L(\mathcal{X})$  consisting of Hermitian operators we will be denoted by  $\text{Herm}(\mathcal{X})$  whereas the set of all positive semidefinite by  $\text{Pos}(\mathcal{X})$ . A positive semidefinite operator  $\Pi \in \text{Pos}(\mathcal{X})$  satisfying  $\Pi^2 = \Pi$  is said to be a projection operator. We will denote the set of all projection operators by  $\text{Proj}(\mathcal{X})$ . For each subspace  $\mathcal{V} \subseteq \mathcal{X}$ , there is a uniquely defined projection operator  $\Pi \in \text{Proj}(\mathcal{X})$  satisfying  $\text{im}(\Pi) = \mathcal{V}$ . The notation  $\Pi_{\mathcal{V}}$  is used to refer to this projection operator. A positive semidefinite operator  $\rho \in \text{Pos}(\mathcal{X})$  is a density operator if  $\text{tr}(\rho) = 1$ . The set of all density operators will be denoted as  $\Omega(\mathcal{X})$ .

An operator  $V \in L(\mathcal{X}, \mathcal{Y})$ , is an isometry if it preserves the Euclidean norm of vectors that is  $\|V|x\rangle\| = \||x\rangle\|$  for all  $|x\rangle \in \mathcal{X}$ . We will consider also a unitary operator  $U \in L(\mathcal{X})$  that is an invertible operator satisfying the equation  $UU^\dagger = U^\dagger U = \mathbb{1}_{\mathcal{X}}$ . Naturally, each unitary operator also preserves the Euclidean norm of vectors. We will denote the set of all isometries as  $U(\mathcal{X}, \mathcal{Y})$  whereas the set of all unitary operators by  $U(\mathcal{X})$ .

An operator  $X = (x_{ij})_{i,j=0}^{\dim(\mathcal{X})-1} \in L(\mathcal{X})$  is diagonal if  $x_{ij} = 0$  for all  $i \neq j$ . The set of all diagonal operator will be denoted as  $\text{Diag}(\mathcal{X})$ . We will also consider unitary operators which are diagonal. Such a set of all diagonal unitary operators will be denoted by  $\text{DU}(\mathcal{X})$ . We will also introduce the operator  $\text{diag}(\cdot) : \mathcal{X} \rightarrow L(\mathcal{X})$  given by

$$\text{diag}(|x\rangle) := \begin{pmatrix} x_0 & 0 & \dots & 0 \\ 0 & x_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_{\dim(\mathcal{X})-1} \end{pmatrix}, \quad (2.34)$$

where  $|x\rangle = (x_0, \dots, x_{\dim(\mathcal{X})-1})^\top$ . and consider the adjoint operator  $\text{diag}^\dagger : L(\mathcal{X}) \rightarrow$

$\mathcal{X}$ , that is  $\text{diag}^\dagger(X) = (X_{0,0}, \dots, X_{\dim(\mathcal{X})-1, \dim(\mathcal{X})-1})^\top$ .

### Eigenvectors and eigenvalues

Let  $A \in L(\mathcal{X})$  be an operator and  $|x\rangle \in \mathcal{X}$  be a nonzero vector for which it holds that

$$A|x\rangle = \lambda|x\rangle, \quad (2.35)$$

for some  $\lambda \in \mathbb{C}$ . Then,  $|x\rangle$  is called an eigenvector of  $A$  and  $\lambda$  is its corresponding eigenvalue. Let us define the characteristic polynomial of  $A$  as

$$p(\alpha) = \det(\alpha \mathbb{1}_{\mathcal{X}} - A). \quad (2.36)$$

The spectrum of  $A$ , denoted  $\lambda(A)$ , is the multiset containing all roots of the polynomial  $p(\alpha)$ , that is all eigenvalues of  $A$ . Moreover, the trace may be expressed by using the term of spectrum in the following way

$$\text{tr}(A) = \sum_{\lambda \in \lambda(A)} \lambda, \quad (2.37)$$

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

For a given operator we have the following characterizations.

1. Each Hermitian operator  $H \in \text{Herm}(\mathcal{X})$  has real eigenvalues.
2. Each positive semidefinite operator  $P \in \text{Pos}(\mathcal{X})$  has non-negative eigenvalues.
3. Each projective operator  $\Pi \in \text{Pos}(\mathcal{X})$  of  $\text{rank}(\Pi) = k$  has  $k$  eigenvalues equal to one and  $(\dim(\mathcal{X}) - k)$  eigenvalues equal to zero.
4. Each density operator  $\rho \in \Omega(\mathcal{X})$  has non-negative eigenvalues which sum up to one.

### 2.2.2 Operator decompositions

Below we recall the theorems regarding decompositions of linear operators, which we use later in the dissertation.

**Theorem 1 (Spectral decomposition)** *Let  $X \in L(\mathcal{X})$  be a normal operator. There exists a positive integer  $m$ , distinct complex numbers  $\lambda_0, \dots, \lambda_{m-1} \in \mathbb{C}$  and nonzero projector operators  $\Pi_0, \dots, \Pi_{m-1} \in \text{Pos}(\mathcal{X})$  satisfying  $\Pi_0 + \dots + \Pi_{m-1} = \mathbb{1}_{\mathcal{X}}$ , such that*

$$X = \sum_{i=0}^{m-1} \lambda_i \Pi_i. \quad (2.38)$$

The scalars  $\lambda_0, \dots, \lambda_{m-1}$  and projection operators  $\Pi_0, \dots, \Pi_{m-1}$  are uniquely determined, that means each scalar  $\lambda_i$  is an eigenvalue of  $X$  with multiplicity equal to the rank of  $\Pi_i$ , and  $\Pi_i$  is the projection operator onto the space spanned by the eigenvectors of  $X$  corresponding to the eigenvalue  $\lambda_i$ .

**Corollary 1** *Let  $\dim(\mathcal{X}) = d$ . Let  $X \in \mathsf{L}(\mathcal{X})$  be a normal operator and let  $\lambda(X) = \{\lambda_0, \dots, \lambda_{d-1}\}$  be the spectrum of  $X$ . There exists an orthonormal basis  $\{|x_0\rangle, \dots, |x_{d-1}\rangle\} \in \mathcal{X}$  such that*

$$X = \sum_{i=0}^{d-1} \lambda_i |x_i\rangle \langle x_i|. \quad (2.39)$$

**Theorem 2 (Jordan–Hahn decompositions)** *Let us consider  $H \in \mathsf{Herm}(\mathcal{X})$  of the form*

$$H = \sum_{i=0}^{\dim(\mathcal{X})-1} \lambda_i \Pi_i, \quad (2.40)$$

where  $\Pi_i \in \mathsf{Proj}(\mathcal{X})$  are nonzero projection operators such that  $\sum_i \Pi_i = \mathbb{1}_{\mathcal{X}}$  and  $\lambda_i \in \mathbb{R}$ ,  $i \in \{0, \dots, \dim(\mathcal{X}) - 1\}$ . Let us define the operators  $P, Q \in \mathsf{Pos}(\mathcal{X})$  by

$$P = \sum_{i=0}^{\dim(\mathcal{X})-1} \max\{\lambda_i, 0\} \Pi_i, \quad (2.41)$$

and

$$Q = \sum_{i=0}^{\dim(\mathcal{X})-1} \max\{-\lambda_i, 0\} \Pi_i. \quad (2.42)$$

satisfying  $PQ = 0$ . Then, the operator  $H$  can be expressed as

$$H = P - Q. \quad (2.43)$$

The operators  $P$  and  $Q$  are uniquely defined for a given operator  $H$ . The expression given by Eq. (2.43) is called the Jordan–Hahn decomposition of  $H$ .

Now, we recall the singular value theorem. The singular value theorem has a close relationship to the spectral theorem. Unlike the spectral theorem, however, the singular value theorem holds for arbitrary (nonzero) operators, as opposed to just normal operators.

**Theorem 3 (Singular value decomposition)** *Let  $\mathcal{X}, \mathcal{Y}$  be complex Euclidean spaces. Let  $A \in \mathsf{L}(\mathcal{X}, \mathcal{Y})$  be a nonzero operator having rank equal to  $r \in \mathbb{N}$ . There*

exist sets of orthonormal vectors  $\{|x_0\rangle, \dots, |x_{r-1}\rangle\} \subset \mathcal{X}$  and  $\{|y_0\rangle, \dots, |y_{r-1}\rangle\} \subset \mathcal{Y}$  along with positive real numbers  $s_0, \dots, s_{r-1} > 0$  such that

$$A = \sum_{i=0}^{r-1} s_i |y_i\rangle \langle x_i|. \quad (2.44)$$

The expression of a given operator  $A$  in the form of Eq. (2.44) is said to be a singular value decomposition of  $A$ . The numbers  $s_0, \dots, s_{r-1}$  are called singular values of  $A$ , whereas the collection of vectors  $|x_0\rangle, \dots, |x_{r-1}\rangle$  and  $|y_0\rangle, \dots, |y_{r-1}\rangle$  are called right and left singular vectors of  $A$ , respectively. The set of all singular values of  $A$  will be denoted by  $\sigma(A)$ .

**Corollary 2** Let  $A \in L(\mathcal{X}, \mathcal{Y})$  be a nonzero operator with rank  $r$ , and let  $\mathcal{W}$  be a complex Euclidean space such that  $\dim(\mathcal{W}) = r$ . Then, there exists a diagonal, positive definite operator  $D \in \text{Pos}(\mathcal{W})$  of the form  $D = \text{diag}(\sigma(A))$  and isometries  $U \in U(\mathcal{W}, \mathcal{Y})$ ,  $V \in U(\mathcal{W}, \mathcal{X})$  such that

$$A = UDV^\dagger. \quad (2.45)$$

## Power of an operator

Now, we will recall the notion of power of an operator. Let  $A \in L(\mathcal{X})$  be a square operator. Naturally, we can define  $k$ -th power of the operator  $A$ , where  $k \in \mathbb{N}$  is a natural number, as the following multiplication

$$A^k = \underbrace{A \cdot \dots \cdot A}_k. \quad (2.46)$$

To extend this definition for any positive number  $k > 0$ , let us consider a positive semidefinite operator  $P \in \text{Pos}(\mathcal{X})$  of the form

$$P = \sum_{i=0}^m \lambda_i \Pi_i, \quad (2.47)$$

where  $\lambda_i \geq 0$  for all  $i \in \{0, \dots, m\}$ . The  $k$ -th power of the operator  $P$  is given by

$$P^k = \sum_{i=0}^m \lambda_i^k \Pi_i, \quad (2.48)$$

where  $k \in \mathbb{R}$ ,  $k > 0$ .

## Moore-Penrose pseudo inverse

**Definition 1** Let  $A \in L(\mathcal{X}, \mathcal{Y})$  be a nonzero operator written in the SVD form (see Theorem 3)

$$A = \sum_{i=0}^{m-1} s_i |y_i\rangle\langle x_i|. \quad (2.49)$$

We define an operator  $A^{-1} \in L(\mathcal{Y}, \mathcal{X})$ , known as the Moore-Penrose pseudo-inverse of  $A$ , as the unique operator given by

$$A^{-1} = \sum_{i=0}^{m-1} \frac{1}{s_i} |x_i\rangle\langle y_i|. \quad (2.50)$$

One may observe that  $AA^{-1}$  and  $A^{-1}A$  are projection operators, projecting onto the spaces spanned by the left singular vectors  $\{|y_i\rangle\}_{i=0}^{m-1}$  and right singular vectors  $\{|x_i\rangle\}_{i=0}^{m-1}$  of  $A$ , respectively.

### 2.2.3 Operator norms

Many useful norms can be defined on spaces of operators, but in quantum information theory we mostly use a single family of norms called Schatten  $p$ -norms.

**Definition 2** For any operator  $A \in L(\mathcal{X}, \mathcal{Y})$  and real number  $p \geq 1$ , we define the Schatten  $p$ -norm of  $A$  as

$$\|A\|_p = \left( \text{tr} \left( (A^\dagger A)^{\frac{p}{2}} \right) \right)^{\frac{1}{p}}. \quad (2.51)$$

The Schatten  $\infty$ -norm is defined as

$$\|A\|_\infty = \max\{\|A|\psi\rangle\| : |\psi\rangle \in \mathcal{X}, \langle\psi|\psi\rangle = 1\}, \quad (2.52)$$

This family of norms includes the three most commonly used norms in quantum information theory – the spectral norm ( $p = \infty$ ), the trace norm ( $p = 1$ ) and the Frobenius norm ( $p = 2$ ).

### 2.2.4 Numerical range and $q$ -numerical range

This dissertation will often make use of the terms of numerical range and  $q$ -numerical range.

**Definition 3** *The numerical range [51] is a subset of the complex plane defined for a matrix  $X \in L(\mathcal{X})$  as*

$$W(X) := \{\langle \psi | X | \psi \rangle : |\psi\rangle \in \mathcal{X}, \langle \psi | \psi \rangle = 1\}. \quad (2.53)$$

Observe that  $W(X)$  contains eigenvalues of  $X$ . It is, however, always the case that  $W(X)$  is convex and compact, which is the content of Toeplitz–Hausdorff theorem [52, 53]. Moreover, if we additionally assume that  $X$  is a normal operator that  $W(X)$  is a convex hull of its eigenvalues.

**Definition 4** *For a given operator  $X \in L(\mathcal{X})$ , the  $q$ -numerical range [54–57] is defined as*

$$W_q(X) := \{\langle \psi | X | \varphi \rangle : |\psi\rangle, |\varphi\rangle \in \mathcal{X}, \langle \psi | \psi \rangle = \langle \varphi | \varphi \rangle = 1, \langle \psi | \varphi \rangle = q, q \in \mathbb{C}\}. \quad (2.54)$$

Observe the standard numerical range makes up the special case of  $q$ -numerical range for  $q = 1$ . The  $q$ -numerical range, as a natural extension of standard numerical range preserves the properties of convexity and compactness (Tsing’s theorem [56]). Another key observation is the fact that  $W_q(X)$  contains each eigenvalues of  $X$  multiplied by  $q \in \mathbb{C}$  [57]. Other useful properties of  $q$ -numerical range [58] that will be used throughout this dissertation are

$$W_q(X \otimes \mathbb{1}_{\mathcal{Y}}) = W_q(X), \quad q \in \mathbb{C}, \quad (2.55)$$

for any dimension of  $\mathcal{Y}$  and

$$W_{q'} \subseteq \frac{q'}{q} W_q \quad \text{for } q \leq q', \quad q, q' \in \mathbb{R}. \quad (2.56)$$

The detailed properties of the numerical range and its generalizations we can also see on the website [59].

We will also use the notation

$$\nu_q(X) := \min\{|x| : x \in W_q(X)\}. \quad (2.57)$$

to denote the distance on a complex plane from  $q$ -numerical range to zero. In the case when  $q = 1$ , we will simply write  $\nu(X)$ . From the above properties of  $W_q(X)$  it is easy to see that

$$\nu_q(X \otimes \mathbb{1}_{\mathcal{Y}}) = \nu_q(X) \quad \text{for } q \in \mathbb{R}. \quad (2.58)$$



## 2.3 Linear superoperators

Now we will consider a mapping, called superoperator, acting on linear operator from  $L(\mathcal{X})$  and mapping them into operators on complex Euclidean space  $L(\mathcal{Y})$ . More precisely,

$$\mathcal{M} : L(\mathcal{X}) \rightarrow L(\mathcal{Y}). \quad (2.59)$$

The set of all such maps is denoted  $M(\mathcal{X}, \mathcal{Y})$ . According to the introduced convention, let  $M(\mathcal{X}) := M(\mathcal{X}, \mathcal{X})$ . The identity mapping on the space  $M(\mathcal{X})$  will be denoted by  $\mathcal{I}_{\mathcal{X}}$ .

**Definition 5** For  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  the adjoint of  $\mathcal{M}$  is defined to be the unique map  $\mathcal{M}^\dagger \in M(\mathcal{Y}, \mathcal{X})$  such that

$$\langle \mathcal{M}^\dagger(Y), X \rangle = \langle Y, \mathcal{M}(X) \rangle, \quad (2.60)$$

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

### Direct sum and tensor product of superoperators

The direct sum between the superoperators of the set  $M(\mathcal{X}, \mathcal{Y})$  are defined in a similar way to direct sum of operators. For the given superoperators  $\mathcal{M} \in M(\mathcal{X}_0, \mathcal{Y}_0)$  and  $\mathcal{N} \in M(\mathcal{X}_1, \mathcal{Y}_1)$  we define the direct sum of superoperators as

$$(\mathcal{M} \oplus \mathcal{N}) \in M(\mathcal{X}_0 \oplus \mathcal{X}_1, \mathcal{Y}_0 \oplus \mathcal{Y}_1), \quad (2.61)$$

to be the unique linear mapping that satisfies the following equation

$$(\mathcal{M} \oplus \mathcal{N})(X \oplus Y) = \mathcal{M}(X) \oplus \mathcal{N}(Y), \quad (2.62)$$

for all  $X \in L(\mathcal{X}_0)$  and  $Y \in L(\mathcal{X}_1)$ .

Similarly, let us define the tensor product between the superoperators  $\mathcal{M} \in M(\mathcal{X}_0, \mathcal{Y}_0)$  and  $\mathcal{N} \in M(\mathcal{X}_1, \mathcal{Y}_1)$  as the unique linear mapping

$$(\mathcal{M} \otimes \mathcal{N}) \in M(\mathcal{X}_0 \otimes \mathcal{X}_1, \mathcal{Y}_0 \otimes \mathcal{Y}_1), \quad (2.63)$$

which satisfies the following equation

$$(\mathcal{M} \otimes \mathcal{N})(X \otimes Y) = \mathcal{M}(X) \otimes \mathcal{N}(Y), \quad (2.64)$$

for all  $X \in L(\mathcal{X}_0)$  and  $Y \in L(\mathcal{X}_1)$ . As for vectors and operators, the notation  $\mathcal{M}^{\oplus n}$  and  $\mathcal{M}^{\otimes n}$  denotes the  $n$ -fold direct sum and tensor product of a map  $\mathcal{M}$  with itself, respectively.

## Partial trace

Let us consider the map

$$\text{tr} \otimes \mathcal{I}_{\mathcal{Y}} \in \text{M}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{Y}), \quad (2.65)$$

defined as the unique map

$$(\text{tr} \otimes \mathcal{I}_{\mathcal{Y}})(X \otimes Y) = \text{tr}(X)Y. \quad (2.66)$$

This map is called the partial trace, and is more commonly denoted  $\text{tr}_{\mathcal{X}}$ . Along similar ways, the map  $\text{tr}_{\mathcal{Y}} \in \text{M}(\mathcal{X} \otimes \mathcal{Y}, \mathcal{X})$  is defined as  $\text{tr}_{\mathcal{Y}} = \mathcal{I}_{\mathcal{X}} \otimes \text{tr}$ .

Now let us consider an operator  $A \in \text{L}(\mathcal{X} \otimes \mathcal{Y})$  not being a tensor product of two operators. Then we express the partial trace of  $A$  over space  $\mathcal{X}$  as

$$\text{tr}_{\mathcal{X}}(A) = \sum_{i=0}^{\dim(\mathcal{X})-1} (\langle i| \otimes \mathbb{1}_{\mathcal{Y}}) A (|i\rangle \otimes \mathbb{1}_{\mathcal{Y}}). \quad (2.67)$$

Along similar lines, the partial trace over space  $\mathcal{Y}$  is expressed as

$$\text{tr}_{\mathcal{Y}}(A) = \sum_{i=0}^{\dim(\mathcal{Y})-1} (\mathbb{1}_{\mathcal{X}} \otimes \langle i|) A (\mathbb{1}_{\mathcal{X}} \otimes |i\rangle). \quad (2.68)$$

We can generalize the concept of partial trace mappings for more than two linear operators and define them in an analogous way.

## Partial transposition

For a given operator  $X \otimes Y$ , where  $X \in \text{L}(\mathcal{X})$  and  $Y \in \text{L}(\mathcal{Y})$ , we define the map

$${}^T X \otimes \mathcal{I}_{\mathcal{Y}} \in \text{M}(\mathcal{X} \otimes \mathcal{Y}), \quad (2.69)$$

as the unique map satisfying

$$({}^T X \otimes \mathcal{I}_{\mathcal{Y}})(X \otimes Y) = X^{\top} \otimes Y. \quad (2.70)$$

Considering an operator which is not a tensor product of two operators  $A \in \text{L}(\mathcal{X} \otimes \mathcal{Y})$  of the form

$$A = \sum_{i,j=0}^{\dim(\mathcal{X})-1} \sum_{k,l=0}^{\dim(\mathcal{Y})-1} \alpha_{ijkl} |i\rangle\langle j| \otimes |k\rangle\langle l|, \quad (2.71)$$

then  $A^{T\mathcal{X}}$  has the following form

$$A^{T\mathcal{X}} = \sum_{i,j=0}^{\dim(\mathcal{X})-1} \sum_{k,l=0}^{\dim(\mathcal{Y})-1} \alpha_{ijkl} (|i\rangle\langle j|)^{\top} \otimes |k\rangle\langle l| = \sum_{i,j=0}^{\dim(\mathcal{X})-1} \sum_{k,l=0}^{\dim(\mathcal{Y})-1} \alpha_{ijkl} |j\rangle\langle i| \otimes |k\rangle\langle l|. \quad (2.72)$$

Analogously, we can define  $A^{T\mathcal{Y}}$ .

## Classes of superoperators

The following classes of linear maps  $M(\mathcal{X}, \mathcal{Y})$  will be used later in this dissertation. A linear map  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  is said to be Hermiticity-preserving if it holds  $\mathcal{M}(A) \in \text{Herm}(\mathcal{Y})$  for all  $A \in \text{Herm}(\mathcal{X})$ . We call that a Hermiticity-preserving map  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  is trace non-increasing map if satisfies the condition  $\text{tr}(\mathcal{M}(A)) \leq \text{tr}(A)$  for all  $A \in L(\mathcal{X})$ . The condition  $\text{tr}(\mathcal{M}(A)) = \text{tr}(A)$  defines trace-preserving maps. A linear mapping  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  is positive if it holds  $\mathcal{M}(A) \in \text{Pos}(\mathcal{Y})$  for every  $A \in \text{Pos}(\mathcal{X})$ . A linear map  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  is called completely positive if  $(\mathcal{M} \otimes \mathcal{I}_{\mathcal{Z}})(A) \in \text{Pos}(\mathcal{Y} \otimes \mathcal{Z})$  for every  $A \in \text{Pos}(\mathcal{X} \otimes \mathcal{Z})$  and complex Euclidean space  $\mathcal{Z}$ .

### 2.3.1 Representations of linear superoperators

A few specific representations of channels are known such the Choi, Stinespring or Kraus representation. These different representations reveal interesting properties of channels. Here, we will briefly describe each of them below.

#### Choi-Jamiołkowski representation

Let us consider a linear mapping  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$ . Let us define a mapping

$$J : M(\mathcal{X}, \mathcal{Y}) \rightarrow L(\mathcal{Y} \otimes \mathcal{X}) \quad (2.73)$$

as

$$J(\mathcal{M}) = (\mathcal{M} \otimes \mathcal{I}_{\mathcal{X}})(|\mathbb{1}_{\mathcal{X}}\rangle\rangle\langle\langle\mathbb{1}_{\mathcal{X}}|), \quad (2.74)$$

From the definition of vectorization, one may alternatively write

$$J(\mathcal{M}) = \sum_{i,j=0}^{\dim(\mathcal{X})-1} \mathcal{M}(|i\rangle\langle j|) \otimes |i\rangle\langle j|, \quad (2.75)$$

where  $|i\rangle, |j\rangle \in \mathcal{X}$ . The operator  $J(\mathcal{M})$  is called the Choi operator of  $\mathcal{M}$ . The action of a mapping  $\mathcal{M}$  is given by

$$\mathcal{M}(X) = \text{tr}_{\mathcal{X}} (J(\mathcal{M})(\mathbb{1}_{\mathcal{Y}} \otimes X^{\top})), \quad (2.76)$$

for all  $X \in L(\mathcal{X})$ . It is worth noting that  $J$  is a linear bijection between  $M(\mathcal{X}, \mathcal{Y})$  and operators  $L(\mathcal{Y} \otimes \mathcal{X})$ .

Most often, for clarity of notation, we will denote linear mappings with calligraphic font  $\mathcal{K}, \mathcal{L}, \mathcal{M}$  etc., whereas the corresponding Choi-Jamiołkowski matrices as plain symbols:  $K, L, M$  etc.

### Kraus representation

Let us define two sets of operators  $\{A_i\}_i$  and  $\{B_i\}_i \subset L(\mathcal{X}, \mathcal{Y})$ . We define a linear map  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  by

$$\mathcal{M}(X) = \sum_i A_i X B_i^{\dagger}. \quad (2.77)$$

Such an expression Eq. (2.77) is called a Kraus representation of  $\mathcal{M}$ . It can be shown that a Kraus representation exists for every map in  $M(\mathcal{X}, \mathcal{Y})$ . However, in contrast with Choi-Jamiołkowski representation, it is not unique.

### Stinespring representation

Let  $A, B \in L(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$  and  $X \in L(\mathcal{X})$ . Let us define a map  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  as

$$\mathcal{M}(X) = \text{tr}_{\mathcal{Z}}(AXB^{\dagger}). \quad (2.78)$$

The expression Eq. (2.78) is called a Stinespring representation of  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$ . Similar to Kraus representation, Stinespring representation is not unique.

The following fact shows the equivalence of these representations.

**Fact 1** *Let  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  be a linear map. Assume that the sets  $\{A_i\}_i, \{B_i\}_i \subset L(\mathcal{X}, \mathcal{Y})$  are Kraus representation of  $\mathcal{M}$ . Then, the Choi operator of  $\mathcal{M}$  has the form  $M = \sum_i |A_i\rangle\rangle\langle\langle B_i|$ . Moreover, for every  $\mathcal{Z}$  and operators  $A, B \in L(\mathcal{X}, \mathcal{Y} \otimes \mathcal{Z})$  defined as  $A = \sum_i A_i \otimes |i\rangle$  and  $B = \sum_i B_i \otimes |i\rangle$ , it holds that  $\mathcal{M}(X) = \text{tr}_{\mathcal{Z}}(AXB^{\dagger})$ .*

## 2.3.2 Diamond norm of superoperators

In this section, we will introduce the norm defined on the linear space  $M(\mathcal{X}, \mathcal{Y})$ , called diamond norm (or the completely bounded trace norm).

**Definition 6** Let  $\mathcal{X}, \mathcal{Y}$  be complex Euclidean spaces. The diamond norm  $\|\cdot\|_\diamond : \mathcal{M}(\mathcal{X}, \mathcal{Y}) \rightarrow \mathbb{R}$  of a linear mapping  $\mathcal{M} \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$  is defined as

$$\|\mathcal{M}\|_\diamond = \max \{ \|(\mathcal{M} \otimes \mathcal{I}_{\mathcal{X}})(X)\|_1 : X \in \mathcal{L}(\mathcal{X} \otimes \mathcal{X}), \|X\|_1 \leq 1 \}. \quad (2.79)$$

For Hermiticity-preserving maps  $\mathcal{M} \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$  with the Choi operator  $M$ , the following expressions of the diamond norm can be used equivalently

$$\|\mathcal{M}\|_\diamond = \max \{ \|(\mathcal{M} \otimes \mathcal{I}_{\mathcal{X}})(|x\rangle\langle x|)\|_1 : |x\rangle \in \mathcal{X} \otimes \mathcal{X} : \| |x\rangle \| = 1 \}, \quad (2.80)$$

or

$$\|\mathcal{M}\|_\diamond = \max \{ \|(\mathbb{1}_{\mathcal{Y}} \otimes \sqrt{\rho}) M (\mathbb{1}_{\mathcal{Y}} \otimes \sqrt{\rho})\|_1 : \rho \in \Omega(\mathcal{X}) \}. \quad (2.81)$$

## 2.4 Semidefinite programming

This section describes the formulation of semidefinite programming (SDP). Let  $\mathcal{X}$  and  $\mathcal{Y}$  be complex Euclidean spaces, and let  $\mathcal{M} \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$  be a Hermiticity-preserving map. Let  $A \in \text{Herm}(\mathcal{X})$  and  $B \in \text{Herm}(\mathcal{Y})$  be Hermitian operators. A semidefinite program is defined as a triple  $(\mathcal{M}, A, B)$  with which the following pair of optimization problems is associated.

<u>Primal problem</u>	<u>Dual problem</u>
maximize: $\langle A, X \rangle$	minimize: $\langle B, Y \rangle$
subject to: $\mathcal{M}(X) = B,$ $X \in \text{Pos}(\mathcal{X}).$	subject to: $\mathcal{M}^\dagger(Y) \geq A,$ $Y \in \text{Herm}(\mathcal{Y}).$

Table 2.1: Formulation of primal and dual problem.

Let us define the primal feasible set  $\mathbf{A}$  and the dual feasible set  $\mathbf{B}$  of  $(\mathcal{M}, A, B)$  as follows

$$\begin{aligned} \mathbf{A} &= \{X \in \text{Pos}(\mathcal{X}) : \mathcal{M}(X) = B\}, \\ \mathbf{B} &= \{Y \in \text{Herm}(\mathcal{Y}) : \mathcal{M}^\dagger(Y) \geq A\}. \end{aligned} \quad (2.82)$$

Operators  $X \in \mathbf{A}$  and  $Y \in \mathbf{B}$  are also said to be primal feasible and dual feasible, respectively. The optimum values associated with the primal and dual problems are defined as

$$\alpha = \sup \{ \langle A, X \rangle : X \in \mathbf{A} \}, \quad (2.83)$$

and

$$\beta = \inf\{\langle B, Y \rangle : B \in \mathbf{B}\}, \quad (2.84)$$

respectively.

Semidefinite programs have associated with them a notion of duality, which refers to the special relationship between the primal and dual problems. For many semidefinite programs, it happens that primary and dual problem values are equal. This situation is called strong duality. Slater's theorem provides one set of conditions under which strong duality is guaranteed.

**Theorem 4** (*Slater's theorem for semidefinite programs*) *Let  $\mathcal{X}$  and  $\mathcal{Y}$  be complex Euclidean spaces, let  $\mathcal{M} \in \mathbf{M}(\mathcal{X}, \mathcal{Y})$  be a Hermiticity-preserving map, and let  $A \in \text{Herm}(\mathcal{X})$  and  $B \in \text{Herm}(\mathcal{Y})$  be Hermitian operators. Let  $\mathbf{A}, \mathbf{B}, \alpha, \beta$  be as defined above for the semidefinite program  $(\mathcal{M}, A, B)$ . One has the following two implications:*

1. *If  $\alpha$  is finite and there exists a Hermitian operator  $Y \in \text{Herm}(\mathcal{Y})$  such that  $\mathcal{M}^\dagger(Y) > A$ , then  $\alpha = \beta$ , and moreover there exists a primal-feasible operator  $X \in \mathbf{A}$  such that  $\langle A, X \rangle = \alpha$ .*
2. *If  $\beta$  is finite and there exists a positive definite operator  $X > 0$  such that  $\mathcal{M}(X) = B$ , then  $\alpha = \beta$ , and moreover there exists a dual-feasible operator  $Y \in \mathbf{B}$  such that  $\langle B, Y \rangle = \beta$ .*

Here, we present the semidefinite program for calculating the diamond norm of the Hermiticity-preserving map  $\mathcal{M}$  with Choi operator  $M$  [60].

Primal problem	Dual problem
maximize: $\text{tr}(XM)$	minimize: $\ \text{tr}_{\mathcal{X}}(Y)\ _\infty$
subject to: $\begin{pmatrix} \mathbb{1}_{\mathcal{Y}} \otimes \rho & X \\ X^\dagger & \mathbb{1}_{\mathcal{Y}} \otimes \rho \end{pmatrix} \geq 0,$ $\rho \in \Omega(\mathcal{X}),$ $X \in \mathbf{L}(\mathcal{X} \otimes \mathcal{Y}).$	subject to: $\begin{pmatrix} Y & -M \\ -M & Y \end{pmatrix} \geq 0,$ $Y \in \text{Pos}(\mathcal{X} \otimes \mathcal{Y}).$

Table 2.2: Formulation of primal and dual problem for calculating the diamond norm of the Hermiticity-preserving map  $\mathcal{M}$ .

# Chapter 3

## Introduction to quantum information theory

Although the first chapter has been mathematical, at the moment we will introduce the elementary notions of quantum information theory, including definitions of quantum states, channels or measurements. Moreover, we will also remind the term of quantum networks. This chapter is mainly written based on [37, 61].

### 3.1 Quantum state

According to the convention adopted in books [37, 50], the quantum states are represented by density operators. The following definition makes this precise.

**Definition 7** *A quantum state, describing a quantum system  $X$ , is a density operator of the form  $\rho \in \Omega(\mathcal{X})$ .*

When we refer to a quantum state of a system  $X$ , it is to be understood that the state in question takes the form  $\rho \in \Omega(\mathcal{X})$  for  $\mathcal{X}$  being the complex Euclidean space associated with  $X$ . Note that if  $\mathcal{X}$  given by a tensor product  $\mathcal{X} = \mathcal{Y}_0 \otimes \dots \otimes \mathcal{Y}_k$  is associated with a compound system  $X=(Y_0, \dots, Y_k)$ .

We will distinguish few classes of quantum states describing below.

- Pure state: A quantum state  $\rho \in \Omega(\mathcal{X})$  is said to be a pure state, if it has rank equal to 1. Equivalently,  $\rho$  is a pure state if there exists a unit vector  $|\psi\rangle \in \mathcal{X}$  such that  $\rho = |\psi\rangle\langle\psi|$ .
- Mixed state: Let us assume an ensemble of pure states  $\{p_i, |\psi_i\rangle\}_{i=0}^k$ , where  $p_i > 0$  and  $\sum_{i=0}^k p_i = 1$ ,  $k \in \mathbb{N}$ . Therefore, it will be considered a convex

combination of pure states, called a mixed state, given by

$$\rho = \sum_{i=0}^k p_i |\psi_i\rangle\langle\psi_i|. \quad (3.1)$$

- **Product state:** Suppose  $X=(Y_0, Y_1)$  is a compound system. We say that a state  $\rho \in \Omega(\mathcal{X})$  is a product state of  $X$  if it has the form  $\rho = \sigma_0 \otimes \sigma_1$ , for  $\sigma_0 \in \Omega(\mathcal{Y}_0), \sigma_1 \in \Omega(\mathcal{Y}_1)$  being quantum states describing  $Y_0, Y_1$ , respectively.
- **Separable and entangled state:** A quantum state  $\rho \in D(\mathcal{X} \otimes \mathcal{Y})$  is said to be separable if there exist a positive number  $k \in \mathbb{N}$ , sets  $\{\sigma_0, \dots, \sigma_{k-1}\} \subset D(\mathcal{X}), \{\rho_0, \dots, \rho_{k-1}\} \subset D(\mathcal{Y})$  and a probability vector  $p = (p_0, \dots, p_{k-1})$ , such that

$$\rho = \sum_{i=0}^{k-1} p_i \rho_i \otimes \sigma_i. \quad (3.2)$$

Otherwise, if a state can not be written as a convex combination of product states, such a state is called an entangled state.

## 3.2 State transformations

Let us consider a quantum system described by a quantum state  $\rho \in \Omega(\mathcal{X})$ . The simplest states transformation may be given by a unitary operator  $U \in U(\mathcal{X})$ . As was mentioned before, each unitary operator preserve the norm of vectors that is for every  $|\psi\rangle$  we have  $\|U|\psi\rangle\| = \|\psi\|$ . From this, we can observe that a unitary state transformation maps pure state to pure state, mixed state to another mixed state, and this transformation is also reversible.

However, the unitary transformations are operations that map quantum systems of the same dimensions. A general linear transformation that maps a density operator to other one is known as a quantum channel and described by the following definition.

**Definition 8** *A superoperator  $\mathcal{M} \in M(\mathcal{X}, \mathcal{Y})$  is a quantum channel if satisfies the following restrictions:*

1.  $\mathcal{M}$  is trace-preserving;
2.  $\mathcal{M}$  is completely positive.

The set of all quantum channels we will denote by  $C(\mathcal{X}, \mathcal{Y})$ . According to convention, we take  $C(\mathcal{X}) := C(\mathcal{X}, \mathcal{X})$ . However, it is worth noting that the most general quantum operations are represented by quantum instruments [62, 63].



**Definition 9** A quantum instrument is a collection of completely positive trace non-increasing maps  $\{\mathcal{M}_i\}_i$  characterized by the property that  $\sum_i \mathcal{M}_i$  is a quantum channel.

### 3.3 Quantum channel representation

Due to Choi-Jamiołkowski isomorphism we obtain two criterions saying whether a linear map is a quantum channel.

**Proposition 1** A linear mapping  $\mathcal{M} \in \mathbf{M}(\mathcal{X}, \mathcal{Y})$  is completely positive if and only if  $M \in \mathbf{Pos}(\mathcal{Y} \otimes \mathcal{X})$ .

**Proposition 2** A linear mapping  $\mathcal{M} \in \mathbf{M}(\mathcal{X}, \mathcal{Y})$  is trace preserving if and only if  $\text{tr}_{\mathcal{Y}}(M) = \mathbb{1}_{\mathcal{X}}$ .

In this dissertation we will pay special attention to unitary channels

$$\Phi_U(X) = UXU^\dagger, \quad (3.3)$$

where  $X \in \mathbf{L}(\mathcal{X})$  and  $U \in \mathbf{U}(\mathcal{X})$ . Therefore, to emphasize their importance, we will denote them  $\Phi_U$  instead of  $\mathcal{U}$  or  $\mathcal{V}$ .

In this dissertation we will also consider a special class of quantum channels called non-signaling channels (or causal channels) [64, 65]. We say that  $\mathcal{N} : \mathbf{L}(\mathcal{X}_0 \otimes \mathcal{X}_1) \rightarrow \mathbf{L}(\mathcal{Y}_0 \otimes \mathcal{Y}_1)$  is non-signaling channel if it satisfies the following conditions

$$\begin{aligned} \text{tr}_{\mathcal{Y}_0}(N) &= \frac{\mathbb{1}_{\mathcal{X}_0}}{\dim(\mathcal{X}_0)} \otimes \text{tr}_{\mathcal{X}_0 \mathcal{Y}_0}(N), \\ \text{tr}_{\mathcal{Y}_1}(N) &= \frac{\mathbb{1}_{\mathcal{X}_1}}{\dim(\mathcal{X}_1)} \otimes \text{tr}_{\mathcal{X}_1 \mathcal{Y}_1}(N). \end{aligned} \quad (3.4)$$

It can be shown [66] that each non-signaling channel is an affine combination of product channels. More precisely, a non-signaling channel  $\mathcal{N} : \mathbf{L}(\mathcal{X}_0 \otimes \mathcal{X}_1) \rightarrow \mathbf{L}(\mathcal{Y}_0 \otimes \mathcal{Y}_1)$  can be written as

$$\mathcal{N} = \sum_i \lambda_i \mathcal{S}_i \otimes \mathcal{T}_i, \quad (3.5)$$

where  $\mathcal{S}_i \in \mathbf{M}(\mathcal{X}_0, \mathcal{Y}_0)$  and  $\mathcal{T}_i \in \mathbf{M}(\mathcal{X}_1, \mathcal{Y}_1)$  are quantum channels,  $\lambda_i \in \mathbb{R}$  such that  $\sum_i \lambda_i = 1$ . For the rest of the dissertation, by  $\mathbf{NS}(\mathcal{X}_0 \otimes \mathcal{X}_1 \otimes \mathcal{Y}_0 \otimes \mathcal{Y}_1)$  we will denote the set of Choi operators of non-signaling channels.

## 3.4 Quantum measurement

Quantum measurements provide the mechanism through which classical information may be extracted from quantum systems.

### 3.4.1 Quantum measurement and the Born rule

Let  $m \in \mathbb{N}$  be a finite positive number. A general quantum measurement  $\mathcal{E}$ , that is a positive operator valued measure (POVM) is a collection of positive semidefinite operators  $\{E_0, \dots, E_{m-1}\} \subset \text{Pos}(\mathcal{X})$ , called effects, which sum up to identity, that is  $\sum_{i=0}^{m-1} E_i = \mathbb{1}_{\mathcal{X}}$ .

The Born rule is a key postulate of quantum mechanics which gives the probability that a measurement of a quantum system will yield a given result. While performing a measurement on some quantum state  $\rho \in \Omega(\mathcal{X})$ , the probabilities of obtaining each of the outcomes  $i$  are given by

$$p_i = \text{tr}(\rho E_i). \quad (3.6)$$

### 3.4.2 Quantum measurement as a quantum channel

Quantum measurement can be defined by two equivalent definitions. The first one was defined in the previous subsection, as a set of effects associated with the outcome labels. Here, we will introduce the second definition of quantum measurement. The concept of quantum measurement will be described as a quantum-to-classical channel.

**Definition 10** *A mapping  $\Delta \in \mathcal{C}(\mathcal{X})$  of the form*

$$\Delta(X) = \sum_{i=0}^{\dim(\mathcal{X})-1} \langle i|X|i\rangle |i\rangle\langle i| = \text{diag}^\dagger(\text{diag}(X)), \quad (3.7)$$

*for all  $X \in \mathcal{L}(\mathcal{X})$  is said to be a completely dephasing channel.*

**Definition 11** *Let  $\mathcal{C} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$  be a quantum channel. The channel  $\mathcal{C} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$  is said to be a quantum-to-classical channels if the following holds*

$$\mathcal{C} = \Delta \circ \mathcal{C}. \quad (3.8)$$

The above theorem reveals the equivalence between quantum-to-classical channels and measurements.

**Theorem 5** *Let  $\mathcal{X}, \mathcal{Y}$  be two complex Euclidean spaces. The following complementary facts hold:*

1. For every quantum-to-classical channel  $\mathcal{C} \in \mathcal{C}(\mathcal{X}, \mathcal{Y})$ , there exists a unique measurement  $\mathcal{E} = \{E_0, \dots, E_{m-1}\}$  for which the equation

$$\mathcal{C}(X) = \sum_{i=0}^{m-1} \text{tr}(X E_i) |i\rangle\langle i|, \quad (3.9)$$

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

2. For every measurement  $\mathcal{E} = \{E_0, \dots, E_{m-1}\}$ , the mapping  $\mathcal{C} \in \mathcal{M}(\mathcal{X}, \mathcal{Y})$  defined in (3.9) is a quantum-to-classical channel.

Due to this theorem, each measurement can be considered as a measure-and-prepare channel. The action of the channel  $\mathcal{E}$  is given by

$$\mathcal{E}(\rho) = \sum_{i=0}^{m-1} \text{tr}(\rho E_i) |i\rangle\langle i|. \quad (3.10)$$

Note that in case of a quantum measurement  $\mathcal{E}$ , its Choi operator  $E$  has a block diagonal form with transposed effects on the diagonal, that means

$$E = \sum_{i=0}^{m-1} |i\rangle\langle i| \otimes E_i^\top. \quad (3.11)$$

### 3.4.3 Von Neumann measurements

In this dissertation we will be mostly interested in a special class of quantum measurements, that is von Neumann measurements.

**Definition 12** A quantum measurement  $\mathcal{E} = \{E_0, \dots, E_{m-1}\}$  is said to be a projective measurement if  $E_i \in \text{Proj}(\mathcal{X})$  for all  $i \in \{0, \dots, m-1\}$ .

If all the effects are rank-one projection operators, then such a measurement is called a von Neumann measurement. In this case we have  $m = \dim(\mathcal{X}) - 1$ . Every von Neumann measurement can be parameterized by a unitary operator  $U \in \mathcal{U}(\mathcal{X})$  in the sense that effects are of the form  $\{|u_i\rangle\langle u_i|\}_{i=0}^{\dim(\mathcal{X})-1}$ , where  $|u_i\rangle = U|i\rangle$  is the  $i$ -th column of the unitary operator  $U \in \mathcal{U}(\mathcal{X})$ . Due to this fact we will use the notation  $\mathcal{P}_U$ . The action of a von Neumann measurement  $\mathcal{P}_U$  is expressed as

$$\mathcal{P}_U(\rho) = \sum_{i=0}^{\dim(\mathcal{X})-1} \langle u_i | \rho | u_i \rangle |i\rangle\langle i| \quad (3.12)$$

for any choice of quantum state  $\rho \in \Omega(\mathcal{X})$ .

### 3.5 Fidelity measure

In this section we will define a distance between different quantum operations by using the notion of the fidelity functions.

#### Fidelity function for quantum states

Let  $P, Q \in \text{Pos}(\mathcal{X})$  be positive semidefinite operators. The fidelity function  $\mathcal{F}^+(P, Q)$  between  $P$  and  $Q$  is defined as

$$\mathcal{F}^+(P, Q) = \|\sqrt{P}\sqrt{Q}\|_1. \quad (3.13)$$

According to the definition of the trace norm, an alternative expression for the fidelity function is given by

$$\mathcal{F}^+(P, Q) = \text{tr} \left( \sqrt{\sqrt{Q}P\sqrt{Q}} \right). \quad (3.14)$$

#### Channel fidelity

The channel fidelity [67, 68], known also as the entanglement fidelity, poses a direct analogue to the fidelity measure for quantum states. The formal definition is as follows. Let  $\mathcal{M}$  and  $\mathcal{N}$  be two quantum channels in  $\mathcal{C}(\mathcal{X}, \mathcal{Y})$  with Choi matrices  $M, N \in \text{L}(\mathcal{Y} \otimes \mathcal{X})$ , respectively. The channel fidelity function  $\mathcal{F}^*(\mathcal{M}, \mathcal{N})$  is defined as

$$\mathcal{F}^*(\mathcal{M}, \mathcal{N}) = \left\| \sqrt{\frac{M}{\dim(\mathcal{X})}} \sqrt{\frac{N}{\dim(\mathcal{X})}} \right\|_1^2. \quad (3.15)$$

#### Fidelity measure for quantum measurements

Let  $\mathcal{X}$  be  $d$ -dimensional complex Euclidean space. Let us consider two quantum measurements  $\mathcal{P}$  and  $\mathcal{Q}$  with the effects  $\mathcal{P}_U = \{P_i\}_{i=0}^{d-1} \subset \text{Pos}(\mathcal{X})$  and  $\mathcal{Q}_U = \{Q_i\}_{i=0}^{d-1} \subset \text{Pos}(\mathcal{X})$ . The fidelity function for quantum measurements [69] is expressed as

$$\mathcal{F}_d(\mathcal{P}, \mathcal{Q}) = \frac{1}{d} \sum_{i=0}^{d-1} \text{tr}(P_i Q_i). \quad (3.16)$$

Note that in the case when  $\mathcal{P}$  or  $\mathcal{Q}$  is a von Neumann measurement we obtain the value of fidelity function  $\mathcal{F}_d$  belongs to the interval  $[0, 1]$  and equals one if and only if  $P_i = Q_i$  for all  $i \in \{0, \dots, d-1\}$ .

### 3.5.1 Haar measure and the average fidelity function

The probability measure having fundamental importance in quantum information theory is the Haar measure, defined on the set of unitary operators  $U(\mathcal{X})$ . This measure is closely analogous to the term of the standard Gaussian measure on the real line. Due to Haar measure, we define the average fidelity function used in the dissertation.

#### Haar measure

**Definition 13** A set  $\mathbf{B} \subseteq \mathbf{A}$  is said to be a Borel subset of  $\mathbf{A}$  if one or more of the following inductively defined properties holds

1.  $\mathbf{B}$  is an open set relative to  $\mathbf{A}$ .
2.  $\mathbf{B}$  is the complement of a Borel subset of  $\mathbf{A}$ .
3. For  $\{\mathbf{B}_0, \mathbf{B}_1, \dots\}$  being a countable collection of Borel subsets of  $\mathbf{A}$ , it holds that  $\mathbf{B} = \bigcup_{i=0}^{\infty} \mathbf{B}_i$ .

The collection of all Borel subsets of  $\mathbf{A}$  is denoted  $\text{Borel}(\mathbf{A})$ .

**Definition 14** A Borel measure (or simply a measure) defined on  $\text{Borel}(\mathbf{A})$  is a function

$$\mu : \text{Borel}(U(\mathcal{X})) \rightarrow [0, 1] \quad (3.17)$$

that possesses two properties:

1.  $\mu(\emptyset) = 0$ .
2. For any countable collection  $\{\mathbf{B}_0, \mathbf{B}_1, \dots\} \subseteq \text{Borel}(\mathbf{A})$  of pairwise disjoint Borel subsets of  $\mathbf{A}$ , it holds that

$$\mu \left( \bigcup_{i=0}^{\infty} \mathbf{B}_i \right) = \sum_{i=0}^{\infty} \mu(\mathbf{B}_i). \quad (3.18)$$

A measure  $\mu$  defined on  $\text{Borel}(\mathbf{A})$  is said to be probabilistic if it holds that  $\mu(\mathbf{A}) = 1$ .

**Definition 15** Let  $n \in \mathbb{N}$  and let  $\{X_{i,j}\}_{i,j=0}^n \cup \{Y_{i,j}\}_{i,j=0}^n$  be a collection of independent and identically distributed standard normal random variables. Define an operator-valued random variable  $Z$ , taking values in

$$Z = \sum_{i,j=0}^n (X_{i,j} + iY_{i,j}) |i\rangle\langle j|. \quad (3.19)$$

The Haar measure  $\mu$  on  $U(\mathcal{X})$  is the Borel probability measure

$$\mu : \text{Borel}(U(\mathcal{X})) \rightarrow [0, 1] \quad (3.20)$$

defined as

$$\mu(\mathbf{A}) = p(PZ \in \mathbf{A} \text{ for some } P > 0), \quad (3.21)$$

for every  $\mathbf{A} \in \text{Borel}(U(\mathcal{X}))$ .

One of the most used properties which we will use during the dissertation is that the Haar measure  $\mu$  is a unitarily invariant, that is

$$\mu(U\mathbf{A}) = \mu(\mathbf{A}) = \mu(\mathbf{A}U), \quad (3.22)$$

for every choice of  $\mathbf{A} \in \text{Borel}(U(\mathcal{X}))$  and  $U \in U(\mathcal{X})$ . Moreover, let us define a quantum channel  $\mathcal{C} \in \mathcal{C}(\mathcal{X}^{\otimes n})$  as

$$\mathcal{C}(X) = \int_U U^{\otimes n} X U^{\otimes n} d\mu(U), \quad (3.23)$$

for all  $X \in L(\mathcal{X}^{\otimes n})$ , where  $\mu$  denotes the Haar measure on  $U(\mathcal{X})$ . By the unitary invariance of Haar measure, we can observe that

$$[\mathcal{C}(X), U^{\otimes n}] = 0 \quad (3.24)$$

for every  $X \in L(\mathcal{X}^{\otimes n})$  and  $U \in U(\mathcal{X})$ .

## 3.6 Quantum circuits

In this section, we will introduce basic quantum gates and the methods of composing quantum circuits for algorithmic applications. First, we will introduce one qubit gates, next controlled two qubit gates, and finally we will show how to construct quantum circuit.

### 3.6.1 One qubit gates

We define four Pauli matrices are very important and useful in quantum computation. They are given by

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.25)$$

It is worth mentioning that an arbitrary qubit gate can be written as a linear combination of Pauli matrices. We would like to distinguish the rotation gates  $RX(\theta)$ ,  $RY(\theta)$  and  $RZ(\theta)$  of the forms respectively,

$$\begin{aligned}RX(\theta) &= \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) X, \\RY(\theta) &= \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) Y, \\RZ(\theta) &= \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) Z.\end{aligned}\tag{3.26}$$

Another important gate is Hadamard gate, which can also be obtained from Pauli matrices,

$$H = \frac{X+Z}{\sqrt{2}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.\tag{3.27}$$

In general, we will represent a single qubit gate  $U$  as in Figure 3.1.

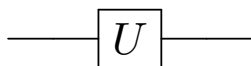


Figure 3.1: A schematic representation of the single qubit gate  $U$ .

## Two qubit gates

Controlled gates act on 2 or more qubits, where one or more qubits act as a control for some operation. For example, the controlled NOT gate (or CNOT or CX) acts on 2 qubits, and performs X operation on the second qubit only when the first qubit is  $|1\rangle$ , and otherwise leaves it unchanged. This gate is given by

$$CX = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X.\tag{3.28}$$

In similar way we can define an arbitrary controlled operator CU for any  $U \in U(\mathbb{C}^2)$  given by

$$CU = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U.\tag{3.29}$$

In general, we will represent a controlled unitary gate CU as in Figure 3.2. The only exception is SX gate which we will represent as in Figure 3.3.

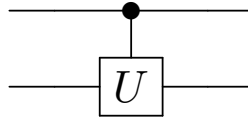


Figure 3.2: A schematic representation of the controlled unitary gate CX.

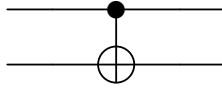


Figure 3.3: A schematic representation of the controlled unitary gate CU.

### Creating quantum circuit

In the standard approach in quantum computing, at the beginning of computation each qubit is in the initial state  $|0\rangle$ . The computation is completed by measuring each qubit in Z-basis, receiving a classical output label  $i$  on each qubit. The measurement in Z-basis with a classical output label  $i$  will be represented as in Figure 3.4.

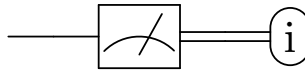


Figure 3.4: A schematic representation of measurement in Z-basis receiving a classical output label  $i$ .

After introducing the basic facts concerning the quantum gates, now we can create quantum circuits. For example, the following circuit in Fig. 3.5 can be

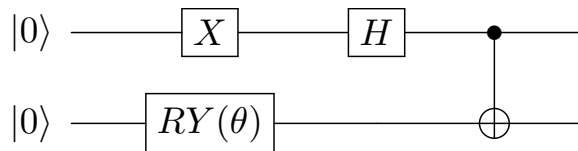


Figure 3.5: Example of circuit described by Eq. (3.30).



expressed as a matrix multiplication

$$CX(X \otimes RY(\theta))(H \otimes I)(|0\rangle \otimes |0\rangle), \quad (3.30)$$

for any angle  $\theta$ .

## 3.7 Link product and quantum networks

The notion of quantum networks is a generalization of the Choi isomorphism for the composition of any number of quantum operations. This approach allows us to efficiently represent a quantum circuit in terms of a single positive operator. For this purpose, we will introduce the definition of the link product. This section is inspired the work [61].

### 3.7.1 Link product

We will use the Choi-Jamiołkowski isomorphism for composition of two or more maps [70]. Let us consider a mapping  $\mathcal{R} = \mathcal{N} \circ \mathcal{M}$  for linear mappings  $\mathcal{N} : L(\mathcal{Z}) \rightarrow L(\mathcal{Y})$  and  $\mathcal{M} : L(\mathcal{X}) \rightarrow L(\mathcal{Z})$  with Choi operators  $N \in L(\mathcal{Z} \otimes \mathcal{Y})$  and  $M \in L(\mathcal{X} \otimes \mathcal{Z})$ , respectively. Therefore, the Choi operator of  $\mathcal{R}$  is given by

$$R = \text{tr}_{\mathcal{Z}} [(\mathbb{1}_{\mathcal{Y}} \otimes M^{T_{\mathcal{Z}}})(N \otimes \mathbb{1}_{\mathcal{X}})], \quad (3.31)$$

where  $M^{T_{\mathcal{Z}}}$  denotes the partial transposition of  $M$  on the space  $\mathcal{Z}$ . Above result can be expressed by introducing the notation

$$N * M := \text{tr}_{\mathcal{Z}} [(\mathbb{1}_{\mathcal{Y}} \otimes M^{T_{\mathcal{Z}}})(N \otimes \mathbb{1}_{\mathcal{X}})], \quad (3.32)$$

called the link product of the operators  $N$  and  $M$ .

If we consider maps such that their input and output spaces are tensor product of Hilbert spaces, it is possible to define the composition of maps only through some of these spaces. Let us consider the following example. Let  $\mathcal{M} \in M(\mathcal{X}_0 \otimes \mathcal{X}_2, \mathcal{X}_1 \otimes \mathcal{X}_3)$  and  $\mathcal{N} \in M(\mathcal{X}_3 \otimes \mathcal{X}_5, \mathcal{X}_4 \otimes \mathcal{X}_6)$  be two linear maps. We define the composition of  $\mathcal{N}$  and  $\mathcal{M}$  as

$$\mathcal{N} \star \mathcal{M} := (\mathcal{N} \otimes \mathcal{I}_{\mathcal{X}_1}) \circ (\mathcal{M} \otimes \mathcal{I}_{\mathcal{X}_5}). \quad (3.33)$$

Following in the same way as before we define the link product  $N * M$  by

$$N * M = \text{tr}_{\mathcal{X}_3} [(\mathbb{1}_{\mathcal{X}_{456}} \otimes M^{T_{\mathcal{X}_3}})(\mathbb{1}_{\mathcal{X}_{012}} \otimes N)], \quad (3.34)$$

where  $M \in L(\mathcal{X}_0 \otimes \mathcal{X}_2 \otimes \mathcal{X}_1 \otimes \mathcal{X}_3)$  and  $N \in L(\mathcal{X}_3 \otimes \mathcal{X}_5 \otimes \mathcal{X}_4 \otimes \mathcal{X}_6)$ . The generalization of the above definition for more than two maps is as follows.

**Definition 16** Let  $M$  be an operator in  $L\left(\bigotimes_{i \in I} \mathcal{X}_i\right)$  and  $N$  be an operator in  $L\left(\bigotimes_{j \in J} \mathcal{X}_j\right)$ , where  $I$  and  $J$  are two finite set of indexes. Then, the link product  $N * M$  is an operator in  $L\left(\bigotimes_{i \in I \setminus J} \mathcal{X}_i \otimes \bigotimes_{j \in J \setminus I} \mathcal{X}_j\right)$  defined as

$$N * M := \text{tr}_{\mathcal{X}_{I \cap J}} \left[ \left( \mathbb{1}_{\mathcal{X}_{J \setminus I}} \otimes M^{T_{\mathcal{X}_{I \cap J}}} \right) \left( \mathbb{1}_{\mathcal{X}_{I \setminus J}} \otimes N \right) \right]. \quad (3.35)$$

**Remark 1** It is worth noting that if  $I \cap J = \emptyset$  we have  $N * M = N \otimes M$  and if  $I = J$  we have  $N * M = \text{tr}(M^\top N)$ .

### 3.7.2 Graphical representation of linear maps and quantum networks

We will also consider the concept of quantum combs [71]. It is useful to provide a pictorial representation of linear maps and their composition. We will present this concept based on the following examples.

Let us consider a map  $\mathcal{M} : L(\mathcal{X}) \rightarrow L(\mathcal{Y})$ . For such a linear map, we will use convention presented in Fig. 3.6. We do not draw wires corresponding to one dimensional Hilbert spaces. We will present a map  $\mathcal{M} : \mathbb{C} \rightarrow L(\mathcal{X})$  in Figure 3.7. Whereas a lineal map  $\mathcal{M} : L(\mathcal{X}) \rightarrow \mathbb{C}$  we will represent in Figure 3.8. The composition of linear maps  $\mathcal{N} \star \mathcal{M}$  defined in Eq. (3.33), will be presented in Fig. 3.9.



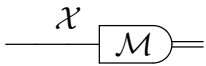
$$\begin{array}{c} \mathcal{X} \text{ --- } \boxed{\mathcal{M}} \text{ --- } \mathcal{Y} \end{array} \quad (3.36)$$

Figure 3.6: Schematic representation of a map  $\mathcal{M} : L(\mathcal{X}) \rightarrow L(\mathcal{Y})$ .



$$\begin{array}{c} \boxed{\mathcal{M}} \text{ --- } \mathcal{X} \end{array} \quad (3.37)$$

Figure 3.7: A schematic representation of the map  $\mathcal{M} : \mathbb{C} \rightarrow L(\mathcal{X})$ .



$$\begin{array}{c} \mathcal{X} \text{ --- } \boxed{\mathcal{M}} \text{ --- } \end{array} \quad (3.38)$$

Figure 3.8: A schematic representation of the map  $\mathcal{M} : L(\mathcal{X}) \rightarrow \mathbb{C}$ .

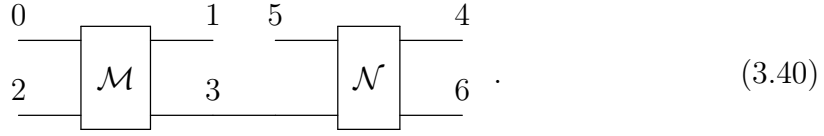
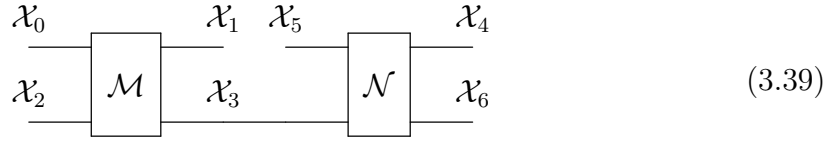


Figure 3.9: Two equivalent representation of the map  $\mathcal{N} \star \mathcal{M}$  defined in Eq. (3.33).

### 3.7.3 Quantum network and tester

**Definition 17** We say that  $\mathcal{S}^{(N)}$  is a probabilistic quantum network with  $N$  components if it is a concatenation of  $N$  completely positive trace non increasing maps  $\mathcal{S}^{(N)} = \mathcal{M}_1 \star \mathcal{M}_2 \star \dots \star \mathcal{M}_N$

We will represent a probabilistic quantum network  $\mathcal{S}^{(N)}$  as in Fig. 3.10.

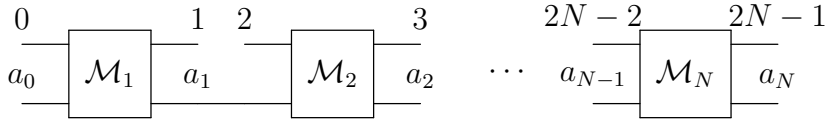


Figure 3.10: Schematic representation of the probabilistic quantum network  $\mathcal{S}^{(N)}$ . The mappings  $\mathcal{M}_i : \mathcal{L}(\mathcal{X}_{2i-2} \otimes \mathcal{X}_{a_{i-1}}) \rightarrow \mathcal{L}(\mathcal{X}_{2i-1} \otimes \mathcal{X}_{a_i})$  are completely positive trace non increasing linear maps,  $i \in \{1, \dots, N\}$ .

**Definition 18** We say that  $\mathcal{R}^{(N)}$  is a deterministic quantum network (or quantum comb) if it is a concatenation of  $N$  quantum channels and  $R^{(N)} \in \mathcal{L}\left(\bigotimes_{i=0}^{2N-1} \mathcal{X}_i\right)$  fulfills the following conditions

$$\begin{aligned} R^{(N)} &\geq 0, \\ \text{tr}_{\mathcal{X}_{2k-1}}(R^{(k)}) &= \mathbb{1}_{\mathcal{X}_{2k-2}} \otimes R^{(k-1)}, \end{aligned} \quad (3.41)$$

where  $R^{(0)} = 1$  and  $R^{(k-1)} \in \mathcal{L}\left(\bigotimes_{i=0}^{2k-3} \mathcal{X}_i\right)$  is the Choi matrix of the reduced quantum comb with concatenation of  $k-1$  quantum channels,  $k = 1, \dots, N$ .

**Remark 2** It can be shown that the Choi operator  $S^{(N)}$  satisfies  $0 \leq S^{(N)} \leq R^{(N)}$ , where  $R^{(N)}$  is Choi matrix of a deterministic quantum network  $\mathcal{R}^{(N)}$ .

**Definition 19** We say that a quantum network is a quantum tester if it is a collection of probabilistic quantum networks  $\{S_i^{(N)}\}_i$  whose sum is a quantum comb, that is  $\sum_i S_i^{(N)} = R^{(N)}$ , and additionally  $\dim(\mathcal{X}_0) = \dim(\mathcal{X}_{a_0}) = \dim(\mathcal{X}_{2N-1}) = \dim(\mathcal{X}_{a_N}) = 1$ .

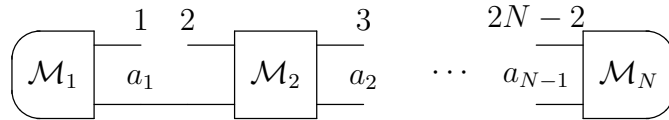


Figure 3.11: Schematic representation of a quantum tester defined in Definition 19.

## Chapter 4

# Storage and retrieval of von Neumann measurements

For the past researchers have discovered unusual properties of quantum systems, such that there is a lack of possibility of cloning quantum information [72, 73], or quantum processors cannot be universally programmed [74]. It states that no perfect universal quantum processor can exist. Hence, their approximate realizations are of common interest [75–78]. Therefore, a separate validation method was proposed based on the fidelity estimation.

One efficient fidelity estimation method is the storage and retrieval (SAR) known as a quantum learning. In the general approach of SAR, we want to approximate a given, unknown operation, we could perform  $N$  times experimentally. This scheme is called a  $N \rightarrow 1$  learning scheme. This strategy usually consists of preparing some initial quantum state, applying the unknown operation  $N$  times, which allows us to store the unknown operation for later use, and finally, a retrieval operation that gives us an approximation of a given operation. We assume that the scheme is optimal when it achieves the highest possible fidelity of the approximation [67, 79]. The task of SAR is two-fold. First, we would like to calculate the maximum value of the fidelity function, and next, we want to find an optimal strategy to realize it.

The seminal work in this field was the paper by Bisio and Chiribella [34]. It was devoted to learning an unknown unitary transformation. Therein, the authors focused on storing the unitary operation in a quantum memory while having limited resources. They proved that unitary operations could be learned optimally in the parallel scheme, meaning there is no additional processing after using the unknown unitary transformation. Hence, all the required uses of the black box can be performed in parallel. They also provide an upper bound on the fidelity of  $N \rightarrow 1$  learning scheme that is equal  $1 - \Theta\left(\frac{1}{N^2}\right)$ . A probabilistic version of SAR (PSAR) problem was also considered in [80, 81]. There, they showed the optimal success probability of  $N \rightarrow 1$  PSAR of unitary channels on  $d$ -dimensional quantum systems

is equal to  $N/(N - 1 + d^2)$ .

Subsequent works build upon these results but focus on different classes of operations, for example, the von Neumann measurements [69]. In contrast to previous works, they showed that, in general, the optimal algorithm for quantum measurement learning cannot be parallel and found the optimal learning algorithm for arbitrary von Neumann measurements for the case  $1 \rightarrow 1$  and  $2 \rightarrow 1$ . Nevertheless, a general optimal  $N \rightarrow 1$  scheme of measurement learning still remains an open problem, even for low-dimensional quantum systems. Hence, despite some partial results, the investigation of SAR for von Neumann measurements is still an open question.

In this chapter, we address the unsolved problem of  $N \rightarrow 1$  learning of unknown von Neumann measurement defined in [69]. We investigate the value of the average fidelity in the asymptotic regime. By using the deterministic port-based teleportation (PBT) protocol we state a lower bound which behaves as  $1 - \Theta\left(\frac{1}{N^2}\right)$ . Moreover, we provide an upper bound for the average fidelity function, which matches the lower bound and hence provides a solution to the problem. Additionally, we compare different learning schemes for the qubit case. Although, the learning scheme based on PBT is asymptotically optimal, it can be outperformed for low values of  $N$ . To show this, we introduce a new scheme, which we call pretty good learning scheme (PGLS). This scheme is a particular case of a parallel learning scheme which uses only two-qubit entangled memory states. The fidelity function calculated for the pretty good learning scheme is uniform over all qubit von Neumann measurements and behaves as  $1 - \Theta\left(\frac{1}{N}\right)$ .

This chapter is organized as follows. In Section 4.1 we formulate the problem of von Neumann measurement learning. In Section 3.5 we introduce necessary mathematical concepts. Our main result is then presented in Section 4.2. To prove this theorem, we first address the case of lower bound (Section 4.2.1), and subsequently, upper bound (Section 4.2.2). In Section 4.3 we compare the performance of different learning schemes for a qubit case. In particular, we introduce the pretty good learning scheme and present numerical results about the most efficient parallel and adaptive learning schemes. Finally, Section 4.4 concludes the article with a summary of the main results. In the Appendix A, we provide technical details of proofs. This chapter is written based on the paper [1].

## 4.1 Learning setup

Imagine we are given a black box with the promise that it contains some von Neumann measurement,  $\mathcal{P}_U$ , parameterized by a unitary matrix  $U$ . The exact value of  $U$  is unknown to us. We are allowed to use the black box  $N$  times. Our goal is to prepare some initial memory state  $\sigma$ , some intermediate processing channels

$\mathcal{C}_1, \dots, \mathcal{C}_{N-1}$  and a measurement  $\mathcal{R}$  such that we are able to approximate  $\mathcal{P}_U$  on an arbitrary state  $\rho$ . This approximation will be denoted throughout this work as  $\mathcal{Q}_U$ . We would like to point out that, generally,  $\mathcal{Q}_U$  will not be a von Neumann measurement. The schematic representation of this setup is presented in Fig. 4.1.

The initial memory state  $\sigma$  and the entire sequence of processing channels  $\{\mathcal{C}_i\}_{i=1}^{N-1}$  can be viewed as storing the unknown operation and will be denoted as  $\mathcal{S}$  whereas the measurement  $\mathcal{R}$  we will call a retrieval. During the storing stage, we apply  $\mathcal{S}$  on  $N$  copies of  $\mathcal{P}_U$ . As a result, the initial memory state  $\sigma$  is transferred to the memory state  $\sigma_{\mathcal{P}_U, \mathcal{S}}$ . After that, we measure an arbitrary quantum state  $\rho$  and the memory state  $\sigma_{\mathcal{P}_U, \mathcal{S}}$  by using  $\mathcal{R}$ . Equivalently, we can say that during retrieval stage, we apply the measurement  $\mathcal{Q}_U$  on the state  $\rho$ . The entire learning scheme will be denoted by  $\mathcal{L}$  and considered as a triple  $\mathcal{L} = (\sigma, \{\mathcal{C}_i\}_{i=1}^{N-1}, \mathcal{R})$ .

Considering a von Neumann measurement  $\mathcal{P}_U$  and its approximation  $\mathcal{Q}_U$  we introduce the average fidelity function [61] with respect to Haar measure as

$$\mathcal{F}_d^{\text{avg}}(\mathcal{L}) = \int_U \mathcal{F}_d(\mathcal{P}_U, \mathcal{Q}_U) d\mu(U). \quad (4.1)$$

Our main goal is to maximize  $\mathcal{F}_d^{\text{avg}}$  over all possible learning schemes  $\mathcal{L} = (\sigma, \{\mathcal{C}_i\}_{i=1}^{N-1}, \mathcal{R})$  which we will denote by  $F_d$ . More precisely,

$$F_d := \max_{\mathcal{L}} \mathcal{F}_d^{\text{avg}}(\mathcal{L}). \quad (4.2)$$

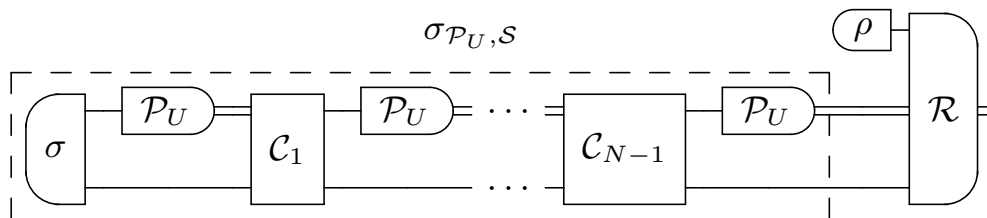


Figure 4.1: Schematic representation of the setup for  $N \rightarrow 1$  learning scheme of von Neumann measurements.

## 4.2 Main results

The main result can be summarized as the following theorem. This section provides just a sketch of the proof of Theorem 6, along with a general intuition behind our result. The full proofs are postponed to the Appendix A due to their technical nature.

**Theorem 6** *Let  $F_d$  be the maximum value of the average fidelity function, defined in Eq. (4.2) for  $N \rightarrow 1$  learning scheme of von Neumann measurements. Then, for arbitrary but fixed dimension  $d$  we obtain*

$$F_d = 1 - \Theta\left(\frac{1}{N^2}\right). \quad (4.3)$$

First, we will show the lower bound for  $F_d$  by constructing a learning scheme based on deterministic port-based teleportation protocol (DPBT). We will show that it achieves the scaling  $\mathcal{F}_d^{\text{avg}}(\mathcal{L}) = 1 - \Theta\left(\frac{1}{N^2}\right)$ . Next, we will provide a sketch of the upper bound. To prove it we use a learning scheme of unitary channels.

### 4.2.1 Lower bound

In the next section, we will see that the problem of SAR is closely related to the deterministic port-based teleportation (DPBT) [68, 82–85]. Before that happens, we recall the DPBT protocol.

#### Deterministic port-based teleportation (DPBT)

The DPBT protocol is described below and can be viewed in Fig. 4.2.

1. Consider two individuals – Alice and Bob. Alice and Bob share compound systems  $\mathcal{X}_A^{\otimes N}$  and  $\mathcal{X}_B^{\otimes N}$ , respectively. Alice system  $\mathcal{X}_A^{\otimes N}$  has the form  $\mathcal{X}_A^{\otimes N} = \mathcal{X}_A^{(0)} \otimes \cdots \otimes \mathcal{X}_A^{(N-1)}$ , where  $\dim(\mathcal{X}_A^{(i)}) = d$  for each  $i \in \{0, \dots, N-1\}$ . Analogously, we define Bob’s space  $\mathcal{X}_B^{\otimes N}$ .
2. They share an entangled state  $|\psi\rangle\langle\psi|_{A,B} \in \Omega(\mathcal{X}_A^{\otimes N} \otimes \mathcal{X}_B^{\otimes N})$ .
3. Their goal is to construct a scheme to teleport an unknown state  $\rho \in \Omega(\mathcal{X}^{(in)})$ , where  $\dim(\mathcal{X}^{(in)}) = d$ , from Alice to Bob in a way that this state appears in one of Bob’s systems. To achieve it, Alice performs a measurement  $\mathcal{Q} = \{Q_i\}_{i=0}^{N-1}$  on the state  $\rho \otimes |\psi\rangle\langle\psi|_A$ .
4. As a result, she receives one of the labels  $\{0, \dots, N-1\}$ .
5. Next, she communicates the outcome to Bob.
6. By using the outcome label  $i$ , Bob chooses  $i$ -th system of his state  $|\psi\rangle\langle\psi|_B$  as the one which contains the state  $\rho$ .



By  $\mathcal{T}(\rho)$  we will denote the output of the DPBT procedure, depending on the choice of  $|\psi\rangle\langle\psi|_{A,B}$  and  $\mathcal{Q}$ , given by

$$\mathcal{T}(\rho) = \sum_{i=0}^{N-1} \text{tr}_{\mathcal{X}^{(in)} \mathcal{X}_A^{\otimes N} \mathcal{X}_{B_i}^{\otimes N-1}} \left( (\mathcal{Q}_i \otimes \mathcal{I}_{\mathcal{X}_B^{\otimes N}}) (\rho \otimes |\psi\rangle\langle\psi|_{A,B}) \right), \quad (4.4)$$

where  $\mathcal{Q}_i(\sigma) = \sqrt{Q_i}\sigma\sqrt{Q_i}$  for every  $i \in \{0, \dots, N-1\}$  and  $\mathcal{X}_{B_i}^{\otimes N-1} := \mathcal{X}_B^{(0)} \otimes \dots \otimes \mathcal{X}_B^{(i-1)} \otimes \mathcal{X}_B^{(i+1)} \otimes \dots \otimes \mathcal{X}_B^{(N-1)}$ .

It is known [68], that for best teleportation procedure  $\mathcal{T}_0$  approximates the identity map  $\mathcal{I}_{\mathbb{C}^d}$  of dimension  $d$ , the entanglement fidelity is equal to

$$F^*(\mathcal{T}_0, \mathcal{I}_{\mathbb{C}^d}) = 1 - \Theta\left(\frac{1}{N^2}\right). \quad (4.5)$$

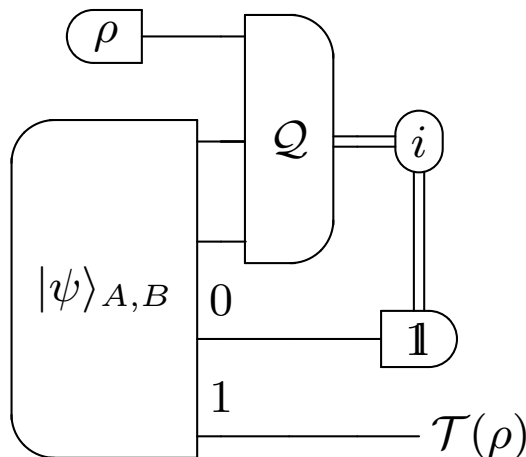


Figure 4.2: Schematic representation of DPBT for  $N = 2$ . In this case, the outcome label  $i = 1$  of the measurement  $\mathcal{Q}$  determines the partial trace of the system  $\mathcal{X}_B^{(0)}$  of the remaining quantum state.

### Learning scheme $N \rightarrow 1$ of von Neumann measurements based on DPBT protocol

Here, we use the DPBT protocol to construct  $N \rightarrow 1$  learning scheme of von Neumann measurements  $\mathcal{P}_U$ .

Assume  $|\psi_0\rangle\langle\psi_0|_{A,B}$  and  $\mathcal{Q}_0$  are the realization of optimal DPBT strategy  $\mathcal{T}_0$ . We take  $\sigma = |\psi_0\rangle\langle\psi_0|_{A,B}$  as an initial memory state and next, we implement  $N$  copies of von Neumann measurement  $\mathcal{P}_U$  on each system  $\mathcal{X}_B^{(i)}$  for  $i \in \{0, \dots, N-1\}$ .

Therefore, the result of the storage is a memory state  $\sigma_{\mathcal{P}_U, \mathcal{S}} = \left( \mathcal{I}_{\mathcal{X}_A}^{\otimes N} \otimes \mathcal{P}_U^{\otimes N} \right) (\sigma)$ , which consists of the remaining quantum state  $|\psi_0\rangle_A$ , and a tuple of measurement results  $(j_1, \dots, j_N) \in \{0, \dots, d-1\}^N$ . The retrieval part  $\mathcal{R}$  is a composition of a measurement  $\mathcal{Q}_0$  and classical postprocessing. In detail, first we apply  $\mathcal{Q}_0$  on  $\rho \otimes |\psi_0\rangle_A$  to obtain the label  $i \in \{0, \dots, N-1\}$ . Second, we return the result  $j_i$  as the output of  $\mathcal{R}$ . The schematic representation of this setup is presented in Fig. 4.3.

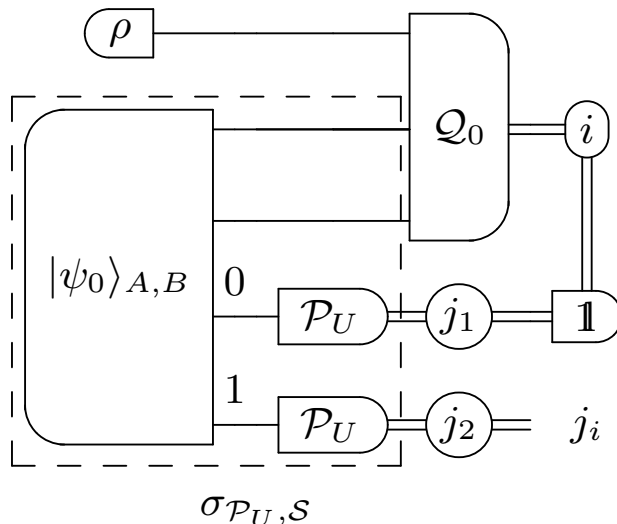


Figure 4.3: Schematic representation of the learning scheme for  $N = 2$  based on DPBT. In this case, the label  $i = 1$  of the measurement  $\mathcal{Q}_0$  indicates that the output of the learning procedure equals  $j_i = j_2$ .

**Lemma 1** *Let us fix  $d \in \mathbb{N}$  and let  $\mathcal{L}$  be  $N \rightarrow 1$  learning scheme of von Neumann measurements based on the DPBT protocol introduced in Section 4.2.1. It holds that*

$$\mathcal{F}_d^{avg}(\mathcal{L}) = 1 - \Theta\left(\frac{1}{N^2}\right). \quad (4.6)$$

**Proof.** Let  $|\psi_0\rangle\langle\psi_0|$  and  $\mathcal{Q}_0 = \{Q_i\}_{i=0}^{N-1}$  be a realization of the optimal teleportation protocol  $\mathcal{T}_0$ , which satisfies [68]

$$\mathcal{F}^*(\mathcal{T}_0, \mathcal{I}_{\mathcal{X}}) = 1 - \Theta\left(\frac{1}{N^2}\right). \quad (4.7)$$

Let us introduce the mappings  $\mathcal{Q}_{0,i}(\sigma) = \sqrt{Q_i}\sigma\sqrt{Q_i}$  for every  $i \in \{0, \dots, N-1\}$ .

Then, the approximation  $\mathcal{Q}_U$  acting on an arbitrary state  $\rho$  can be expressed as

$$\begin{aligned}
\mathcal{Q}_U(\rho) &= \sum_{i=0}^{N-1} \text{tr}_{\mathcal{X}^{(in)} \mathcal{X}_A^{\otimes N} \mathcal{X}_{B_i}^{\otimes N-1}} \left( (\mathcal{Q}_{0,i} \otimes \mathcal{P}_U^{\otimes N}) (\rho \otimes |\psi_0\rangle\langle\psi_0|) \right) \\
&= \mathcal{P}_U \left( \sum_{i=0}^{N-1} \text{tr}_{\mathcal{X}^{(in)} \mathcal{X}_A^{\otimes N} \mathcal{X}_{B_i}^{\otimes N-1}} \left( (\mathcal{Q}_{0,i} \otimes \mathcal{I}_{\mathcal{X}_B^{\otimes N}}) (\rho \otimes |\psi_0\rangle\langle\psi_0|) \right) \right) \\
&= \mathcal{P}_U(\mathcal{T}_0(\rho)),
\end{aligned} \tag{4.8}$$

where  $\mathcal{X}_{B_i}^{\otimes N-1} := \mathcal{X}_B^{(0)} \otimes \dots \otimes \mathcal{X}_B^{(i-1)} \otimes \mathcal{X}_B^{(i+1)} \otimes \dots \otimes \mathcal{X}_B^{(N-1)}$ .

Let  $J(\Delta) = \sum_{i=0}^{d-1} |i\rangle\langle i| \otimes |i\rangle\langle i|$  be the Choi matrix of the completely dephasing channel  $\Delta$ . By using the definition of  $\mathcal{F}_d(\mathcal{P}_U, \mathcal{Q}_U)$  given by

$$\mathcal{F}_d(\mathcal{P}_U, \mathcal{Q}_U) = \frac{1}{d} \text{tr}(P_U \mathcal{Q}_U), \tag{4.9}$$

we obtain

$$\begin{aligned}
\mathcal{F}_d^{\text{avg}}(\mathcal{L}) &= \frac{1}{d} \int_U d\mu(U) \text{tr}(P_U \mathcal{Q}_U) = \frac{1}{d} \int_U d\mu(U) \text{tr}((U \otimes \bar{U}) J(\Delta) (U \otimes \bar{U})^\dagger T_0) \\
&= \frac{1}{d} \text{tr} \left( \frac{1}{d+1} (\mathbb{1}_{\mathbb{C}^{d^2}} + |\mathbb{1}_{\mathbb{C}^d}\rangle\rangle\langle\langle \mathbb{1}_{\mathbb{C}^d}|) T_0 \right) = \frac{1}{d+1} + \frac{d}{d+1} \mathcal{F}^*(\mathcal{T}_0, \mathcal{I}_{\mathbb{C}^d}) \\
&= 1 - \Theta \left( \frac{1}{N^2} \right),
\end{aligned} \tag{4.10}$$

which completes the proof.  $\blacksquare$

**Corollary 3** *Due to the fact  $\mathcal{F}_d^{\text{avg}}(\mathcal{L}) \leq F_d$ , we immediately achieve the lower bound of  $F_d$ .*

## 4.2.2 Upper bound

To show the upper bound for  $F_d$ , we will construct a different learning scheme  $\tilde{\mathcal{L}}$  based on the learning of unitary channels. First, we will calculate the maximum value of the average fidelity function  $\widetilde{F}_d$  for  $\tilde{\mathcal{L}}$ . We will show the inequality  $\widetilde{F}_d \leq 1 - \Theta\left(\frac{1}{N^2}\right)$ . Next we will prove the inequality  $F_d \leq \widetilde{F}_d$ , which will complete the proof.

**Lemma 2** *For any dimension  $d$ , the maximum value of the average fidelity function*

$F_d$  defined in Eq. (4.2) is upper bounded by

$$F_d \leq 1 - \Theta\left(\frac{1}{N^2}\right). \quad (4.11)$$

To prove this lemma, first we will show that for  $d = 2$  this inequality holds. Next, we will present the sketch of key steps for any dimension  $d$ . The complete proof of this lemma for the qubit case we can see in Appendix A.1, whereas for any dimension  $d$  in Appendix A.2.

**Proof.** For the qubit case, let us consider a new learning scheme  $\tilde{\mathcal{L}}$  presented in Fig. 4.4. In this scheme, we are given  $N$  copies of a qubit unitary channel  $\Phi_{\bar{U}}$ , which we can use in parallel. We still want to approximate the measurement  $\mathcal{P}_U$ , but by using the unitary channel  $\Phi_{\bar{U}}$  instead. For this purpose, we will choose the appropriate initial memory state  $|\psi\rangle$  and a binary retrieval measurement  $\mathcal{R} = \{R_0, R_1\}$ .

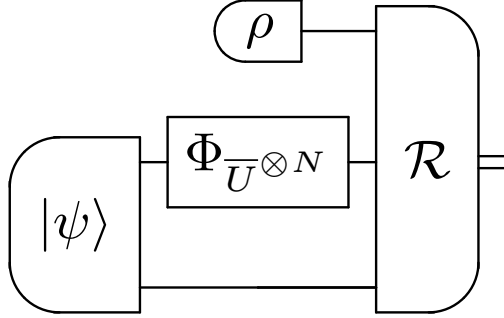


Figure 4.4: Schematic representation of the setup  $N \rightarrow 1$  qubit unitary channel  $\Phi_{\bar{U}}$  learning scheme in parallel way. The objective of this scheme is to approximate the von Neumann measurement  $\mathcal{P}_U$ .

Again, the goal is to maximize the value of the average fidelity function for  $\tilde{\mathcal{L}}$ , which is denoted as  $\widetilde{F}_2$ . In Appendix A.1.1 we derived the value of  $\widetilde{F}_2$  given by

$$\widetilde{F}_2 = \max_{\mathcal{R}, |\psi\rangle, \langle\psi|} \int_U d\mu(U) \sum_{i=0}^1 \frac{\text{tr} [R_i (P_{U,i} \otimes (\Phi_{\bar{U}}^{\otimes N} \otimes \mathcal{I}_{\mathbb{C}^{2^N}}) (|\psi\rangle\langle\psi|))]}{2}. \quad (4.12)$$

Calculating the value of  $\widetilde{F}_2$  is the crux of the proof, because next we showed that  $F_2 \leq \widetilde{F}_2$  (see Lemma 6 in Appendix A.1), which completes the upper bound for  $d = 2$ .

The proof of the upper bound for arbitrary  $d$  relies on the qubit case. Let us take the optimal learning scheme  $\mathcal{L}$  such that it achieves  $\mathcal{F}_d^{\text{avg}}(\mathcal{L}) = F_d$ . We use

$\mathcal{L}$  to construct a new learning scheme  $\mathcal{L}'$  of qubit von Neumann measurements that is we need to construct three parts of learning scheme  $\mathcal{L}'$  – measurement  $\mathcal{P}'_U$ , the initial state  $\sigma'$  and retrieval part  $\mathcal{R}'$ . The details about construction of  $\mathcal{L}'$  are described in Appendix A.2. Next, we calculate that for the learning scheme  $\mathcal{L}'$  of qubit von Neumann measurements, the following holds

$$dF_d - (d - 1) \leq \mathcal{F}_2^{\text{avg}}(\mathcal{L}') \leq F_2 \leq 1 - \Theta\left(\frac{1}{N^2}\right). \quad (4.13)$$

It directly implies that  $F_d \leq 1 - \Theta\left(\frac{1}{N^2}\right)$ , which completes the proof of upper bound for any dimension  $d$ . ■

We immediately obtain the following corollary.

**Corollary 4** *There is no perfect learning scheme of von Neumann measurements, that means  $F_d < 1$  for any  $N \in \mathbb{N}$ .*

### 4.3 Qubit cases ( $d = 2$ )

In this subsection we investigate more deeply the behavior of  $\mathcal{F}_2^{\text{avg}}$  for different examples of learning schemes  $\mathcal{L}$ . Although, the learning scheme based on deterministic port-based teleportation is asymptotically optimal, it can be outperformed for low values of  $N$ . To show this, we introduce a scheme, which we call pretty good learning scheme (PGLS).

#### 4.3.1 Pretty good learning scheme

The first scheme which we will analyze is the pretty good learning scheme. Despite its lack of optimality, it provides a relatively high value for the average fidelity function asymptotically behaving as  $\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PGLS}}) = 1 - \Theta\left(\frac{1}{N}\right)$ . However, this scheme employs a simple storage strategy, which uses only two-qubit entangled memory states and the learning process is done in parallel. Moreover, the achieved value of the fidelity function is uniform over all qubit von Neumann measurements.

Let us consider a parallel learning scheme with  $N$  copies of the von Neumann measurement  $\mathcal{P}_U$ . The pretty good learning scheme is presented in Fig. 4.5 and below we describe steps of the procedure:

1. Let  $|\omega\rangle$  be the maximally entangled state given by  $|\omega\rangle := \frac{1}{\sqrt{2}}|\mathbb{1}_{\mathbb{C}^2}\rangle\rangle$ . We prepare the initial memory  $\sigma$  state as a tensor product of  $N$  states  $|\omega\rangle$ , that is  $\sigma := |\omega\rangle^{\otimes N}$ .
2. We partially measure each state  $|\omega\rangle$  using  $\mathcal{P}_U$ , obtaining the state  $(\mathcal{P}_U \otimes \mathcal{I}_{\mathbb{C}^2})(|\omega\rangle\langle\omega|)$ .

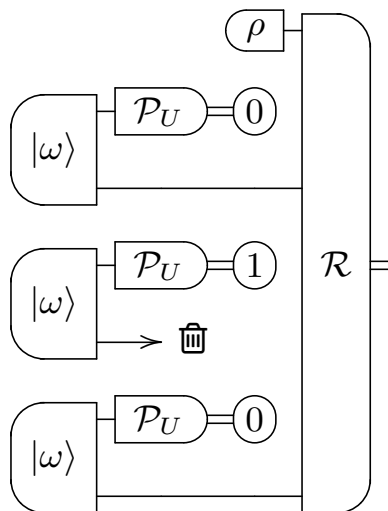


Figure 4.5: Schematic representation of the pretty good learning scheme for  $N = 3$ . In the learning process we obtained three labels: 0, 1, 0. As labels “0” are in majority, we reject the label “1” and the associated quantum part.

3. For each measurement  $\mathcal{P}_U$ , we obtain one of two possible measurement results: “0” or “1”. In consequence, we get  $N_0$  outcomes “0” and  $N_1$  outcomes “1”,  $N_0 + N_1 = N$ . The state of the remaining quantum part is equal to  $\overline{P_{U,0}}^{\otimes N_0} \otimes \overline{P_{U,1}}^{\otimes N_1}$  (up to permutation of subsystems). Without loss of a generality, we may assume that  $N_0 \geq N_1$ .
4. By majority vote we reject minority report, *i.e.* we reject all outcomes “1” and quantum states associated with them. As a result the memory state is given by  $\sigma_{\mathcal{P}_U, \mathcal{S}} = \overline{P_{U,0}}^{\otimes N_0}$ .
5. We prepare an arbitrary state  $\rho \in \Omega(\mathbb{C}^2)$ .
6. We perform a binary retrieval measurement  $\mathcal{R} = \{R, \mathbb{1}_{\mathbb{C}^{2^{n+1}}} - R\}$  on  $\rho \otimes \sigma_{\mathcal{P}_U, \mathcal{S}}$ .

To construct the effect  $R$ , let us fix  $N_0$  and let  $n = N_0 - 1$ . We introduce the family of Dicke states [86]. The Dicke state  $|D_k^n\rangle$  is the  $n$ -qubit state, which is equal to the superposition state of all  $\binom{n}{k}$  basis states of weight  $k$ . For example,  $|D_1^3\rangle = \frac{1}{\sqrt{3}}(|100\rangle + |010\rangle + |001\rangle)$ . Let us also define

$$s_n(k, m) := \sum_{i=0}^k \sum_{j=0}^{n-k} \delta_{i+j-m} \binom{k}{i} \binom{n-k}{j} (-1)^{n-k-j}, \quad (4.14)$$

being the convolution of binomial coefficients. Consider the effect  $R$  of the form

$$R = \sum_{k=0}^n |R_k\rangle\langle R_k|, \quad (4.15)$$

where  $|R_k\rangle := \frac{|M_k\rangle}{\|M_k\|_2}$  and matrices  $M_k \in L(\mathbb{C}^2, \mathbb{C}^{2^{n+1}})$  are given by

$$M_k = \sum_{m=0}^{n+1} \frac{s_n(k, n-m)|0\rangle + s_n(k, n+1-m)|1\rangle}{\sqrt{\binom{n+1}{m}}} \langle D_m^{n+1}|, \quad (4.16)$$

for  $k = 0, \dots, n$ . The proof that  $R$  is a valid effect is relegated to Lemma 14 in Appendix A.3.

In PGLS scheme the approximation  $\mathcal{Q}_U = \{Q_{U,0}, \mathbb{1}_{\mathbb{C}^2} - Q_{U,0}\}$  is determined by relation  $\text{tr}(\rho Q_{U,0}) = \text{tr}\left(\left(\rho \otimes \overline{P_{U,0}}^{\otimes N_0}\right) R\right)$ . Based on Lemma 15 in Appendix A.3, the effect  $Q_{U,0}$  has the form

$$Q_{U,0} = \frac{N_0}{N_0 + 1} P_{U,0}. \quad (4.17)$$

Provided we observed  $N_0$  outcomes “0”, we have that  $\mathcal{F}_2(\mathcal{P}_U, \mathcal{Q}_U) = \frac{2N_0+1}{2N_0+2}$ , where  $N_0$  satisfies  $N_0 \geq \lceil \frac{N}{2} \rceil$ . Note, that the value of  $\mathcal{F}_2(\mathcal{P}_U, \mathcal{Q}_U)$  does not depend on the choice of  $U$ . The average fidelity function  $\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PGLS}})$  defined for the pretty good learning scheme of qubit von Neumann measurements satisfies

$$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PGLS}}) = \begin{cases} \frac{1}{2^N} \sum_{l=k}^N 2 \binom{N}{l} \frac{2l+1}{2l+2}, & N = 2k - 1, \\ \frac{1}{2^N} \left( \binom{N}{k} \frac{2k+1}{2k+2} + \sum_{l=k+1}^N 2 \binom{N}{l} \frac{2l+1}{2l+2} \right), & N = 2k. \end{cases} \quad (4.18)$$

In the asymptotic regime, we may simplify the calculations to obtain

$$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PGLS}}) \geq \frac{2\lceil \frac{N}{2} \rceil + 1}{2\lceil \frac{N}{2} \rceil + 2} = 1 - \Theta\left(\frac{1}{N}\right). \quad (4.19)$$

**Corollary 5** *In the pretty good learning scheme  $\mathcal{L}_{\text{PGLS}} = (\sigma, \{\mathcal{C}_i\}_{i=1}^{N-1}, \mathcal{R})$  the initial state  $\sigma$  is defined as a product of  $N$  copies of maximally entangled state  $|\omega\rangle$ , processing channels  $\{\mathcal{C}_i\}_{i=1}^{N-1}$  are responsible for majority voting and the retrieving measurement  $\mathcal{R}$  with effect defined in Eq.(4.15).*

Finally, averaging the construction of  $\mathcal{Q}_U$  over all possible combinations of measurements' results  $\{0, 1\}^N$  leads to the following approximation of  $\mathcal{P}_U$ .

**Corollary 6** *The approximation  $\mathcal{Q}_U$  is a convex combination of the original measurement  $\mathcal{P}_U$  and the maximally depolarizing channel  $\Phi_*$ . More precisely,*

$$\mathcal{Q}_U = p_0 \mathcal{P}_U + (1 - p_0) \Phi_*, \quad (4.20)$$

where  $p_0 = 2 \cdot \mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PGLS}}) - 1$ .

### 4.3.2 Learning scheme based on port-based teleportation

#### Deterministic port-based teleportation

We have shown in Section 4.2.1 that  $N \rightarrow 1$  learning scheme of von Neumann measurements based on DPBT protocol achieves the average fidelity  $\mathcal{F}_d^{\text{avg}}(\mathcal{L}_{\text{DPBT}}) = 1 - \Theta\left(\frac{1}{N^2}\right)$ . Additionally, from the proof of lower bound in Lemma 1, we calculated that

$$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{DPBT}}) = \frac{1}{3} + \frac{2}{3} \mathcal{F}^*, \quad (4.21)$$

where  $\mathcal{F}^*$  is the entanglement fidelity of DPBT protocol. For qubit case, it is known [83] that  $\mathcal{F}^* = \cos^2\left(\frac{\pi}{N+2}\right)$ . Hence, we obtain

$$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{DPBT}}) = \frac{1}{3} + \frac{2}{3} \cos^2\left(\frac{\pi}{N+2}\right). \quad (4.22)$$

#### Probabilistic port-based teleportation

We can also construct a learning scheme  $N \rightarrow 1$  of von Neumann measurements  $\mathcal{P}_U$  using probabilistic version of port-based teleportation (PPBT) [68]. This protocol works similarly to DPBT presented in Fig. 4.3. The difference lies in the final measurement  $\mathcal{Q}_0$ , which returns a label  $i \in \{0, \dots, N-1\}$ , where the label  $i = 0$  indicates the protocol's failure. Otherwise, if the label  $i > 0$ , it indicates the success of the teleportation procedure. It means that the quantum state  $\rho$ , which we wanted to teleport, is in  $i$ -th system.

The result from [84] says that the corresponding optimal probability of success teleportation equals  $p_0 = \frac{N}{N+3}$ . It implies that the approximation  $\mathcal{Q}_U$  of  $\mathcal{P}_U$  achieved by using the learning scheme  $\mathcal{L}_{\text{PPBT}}$  is given by

$$\mathcal{Q}_U = p_0 \mathcal{P}_U + (1 - p_0) \Phi_*, \quad (4.23)$$



and thus the average fidelity function equals

$$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PPBT}}) = \frac{2N + 3}{2N + 6}. \quad (4.24)$$

### 4.3.3 Numerical investigation

It is generally difficult to find an optimal procedure for quantum operations learning. It is worth mentioning that the fidelity function for parallel learning scheme of von Neumann measurements matches adaptive ones for  $N = 1, 2$  but for  $N \geq 3$  adaptive strategy achieves slight advantage [69].

In the numerical analysis, we compare the average fidelity for the optimal parallel learning strategy  $\mathcal{L}_{\text{Parallel}}$  with the optimal adaptive strategy  $\mathcal{L}_{\text{Adaptive}}$  and other presented learning schemes. It is worth noting that  $\mathcal{L}_{\text{Adaptive}}$  is also the best possible learning scheme available by using quantum networks, that is  $\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{Adaptive}}) = F_2$ .

$N$	1	2	3	4	5
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{Adaptive}})$	0.74999	0.81141	0.86845	0.89684	0.91915
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{Parallel}})$	0.74999	0.81141	0.86764	0.89564	0.91880
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{DPBT}})$	0.5000	0.66667	0.76967	0.83333	0.87449
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PGLS}})$	0.7500	0.79167	0.84375	0.86250	0.88541
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PPBT}})$	0.6250	0.7000	0.7500	0.78571	0.81250

To optimize this problem we used the Julia programming language along with quantum package `QuantumInformation.jl` [87] and SDP optimization via SCS solver [88, 89] with a precision  $\epsilon_{\text{abs}} = 10^{-5}$  and  $\epsilon_{\text{rel}} = 10^{-5}$ . The code is available on GitHub [90]. We also compare the results obtained in this section in the Fig. 4.6.

## 4.4 Conclusion and discussion

In this chapter, we studied the problem of learning of unknown von Neumann measurement of dimension  $d$  from a finite number of copies  $N$ . The main goal was to find the asymptotic behavior of the maximum value for the average fidelity function  $F_d$ . This value was maximized over all possible learning schemes, and the average was taken over all von Neumann measurements. By using the deterministic port-based teleportation (DPBT) protocol, we were able to state the lower bound  $1 - \Theta\left(\frac{1}{N^2}\right)$ , which matched the obtained upper bound and hence, solved the given problem.

In the qubit case, we also introduced a new scheme called the pretty good learning scheme. This scheme was a particular case of a parallel learning protocol,

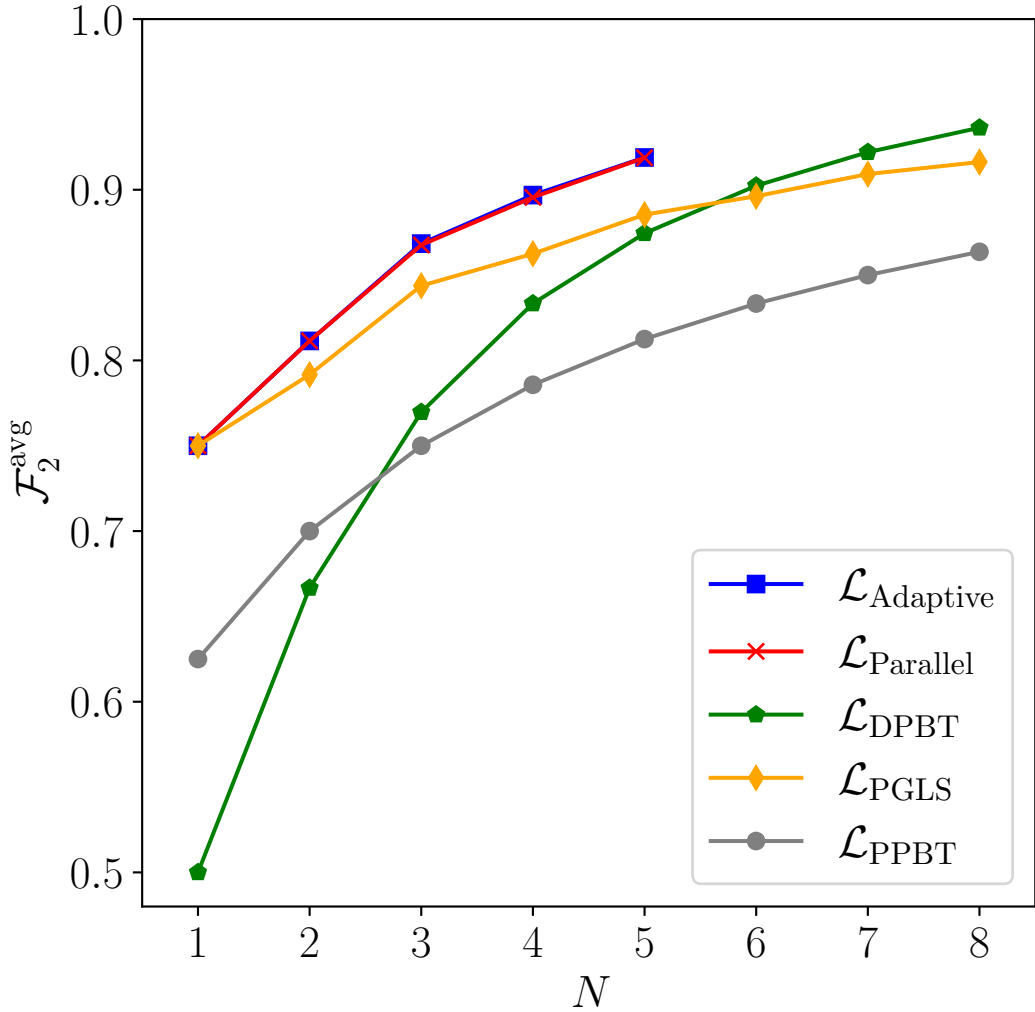


Figure 4.6: The average fidelity function  $\mathcal{F}_2^{\text{avg}}$  calculated for  $N \rightarrow 1$  learning scheme: optimal adaptive strategy,  $\mathcal{L}_{\text{Adaptive}}$  (numerical value, blue squares); optimal parallel scheme,  $\mathcal{L}_{\text{Parallel}}$  (numerical value, red crosses); learning scheme based on DPBT,  $\mathcal{L}_{\text{DPBT}}$  (green pentagons); pretty good learning scheme,  $\mathcal{L}_{\text{PGLS}}$  (orange diamonds); learning scheme based on PPBT,  $\mathcal{L}_{\text{PPBT}}$  (gray circles).

and it used only two-qubit entangled memory states. The average fidelity function calculated for the pretty good learning scheme behaved as  $1 - \Theta\left(\frac{1}{N}\right)$ . Moreover, we compared the performance of different learning schemes: adaptive, parallel, based on DPBT, based on PPBT and the pretty good learning scheme for the qubit case. Although, the learning scheme based on PBT were asymptotically optimal, we showed that the pretty good learning scheme outperforms it for low values of  $N$ .

This work paves the way toward a complete description of the capabilities of

von Neumann measurement learning schemes. One potential way forward is the probabilistic storage and retrieval approach, widely studied for unitary operations and phase rotations in [80, 81]. According to our numerical results, the probability of retrieval of a quantum measurement in a parallel scheme is exactly  $N/(N + 3)$ , which corresponds to the value obtained in [80] for unitary channels, while adaptive strategies for quantum measurements learning to provide slightly higher probability, starting from  $N \geq 3$ .

Another approach to improve the capability of von Neumann measurement learning maybe to use an emerging field of study in quantum information, that is, the indefinite causal structure theory [91]. Traditionally, the quantum theory assumes the existence of a fixed background causal structure. But what if we assume that such a background does not exist? In the next chapter, we will look closely at the concept of undefined causal structure and its potentially used in the problem of von Neumann measurement learning.



## Chapter 5

# Indefinite causal structure theory in SAR of von Neumann measurements

The topic of causality has remained a staple in quantum physics and quantum information theory for recent years. The idea of a causal influence in quantum physics is best illustrated by considering two characters, Alice and Bob, preparing experiments in two separate laboratories. Each of them receives a physical system and performs an operation on it. After that, they send their respective system out of the laboratory. In a causally ordered framework, there are three possibilities: Bob cannot signal to Alice, which means the choice of Bob's action cannot influence the statistics Alice records (denoted by  $A \prec B$ ), Alice cannot signal to Bob ( $B \prec A$ ), or neither party can influence the other ( $A||B$ ).

One may wonder if Alice's and Bob's action can influence each other. It might seem impossible, except in a world with closed time-like curves (CTCs) [92]. However, the existence of CTCs implies some logical paradoxes, such as the grandfather paradox. Possible solutions have been proposed in which quantum mechanics and CTCs can exist and such paradoxes are avoided, but modifying quantum theory into a nonlinear one [93]. A natural question arises: is it possible to keep the framework of linear quantum theory and still go beyond definite causal structures?

One such framework was proposed by Oreshkov, Costa and Brukner [94]. They introduced a new notion of a process matrix – a generalization of the notion of quantum state. This new approach has provided a consistent representation of correlations in causally and non-causally related experiments. Most interestingly, they have described a situation that two actions are neither causally ordered and one cannot say which action influences the second one. Thanks to that, the term of causally non-separable (CNS) structures started to correspond to superpositions of situations in which, roughly speaking, Alice can signal to Bob, and Bob can signal to Alice, jointly. While a causally neutral formulation of quantum theory is

described in terms of the quantum combs [69], the overview of indefinite causal connection theory we can see in [91].

The indefinite causal structures could make a new aspect of quantum information processing. This more general model of computation can outperform causal quantum computers in specific tasks, such as learning or discriminating between two quantum channels [95–97]. As the readers see later, it can be shown the advantage of using causal structure in the problem of the storage and retrieval of von Neumann measurements.

In this chapter we introduce a new concept of von Neumann measurement learning by using indefinite causal structure theory. To do so, we will introduce a notion of  $N$ -partite process matrices. We will call this scheme a causal learning scheme. As in the previous chapter, the main goal is to find the asymptotic behavior of the maximum value for the average fidelity function  $F_d$ . First, we consider a causal learning scheme for two copies of von Neumann measurements. For this case, we proved that using an indefinite causal structure does not improve the average fidelity function  $F_d$  for any dimension  $d$ . Next, however, we show numerical advantages of using causal structure theory in the  $N \rightarrow 1$  learning scheme of qubit von Neumann measurements for  $N \geq 3$ .

This chapter is organized as follows. In Section 5.1 we introduce the concept of  $N$ -partite process matrices. In Section 5.2, we present the advantage of the storage and retrieval of von Neumann measurement by using process matrices and, in general, the indefinite causal structure theory. To show this conjuncture, we first address the case  $2 \rightarrow 1$  causal learning scheme (Section 5.2.2), and subsequently,  $N \rightarrow 1$  causal learning scheme (Section 5.2.3). Finally, Section 5.3 concludes the chapter with a summary of the main results. This chapter is my contribution to the dissertation.

## 5.1 Process matrices – definition

This section introduces the formal definition of a process matrix with its characterization and intuition.

### 5.1.1 Bipartite process matrices

Let us define the operator  ${}_X Y$  as

$${}_X Y = \frac{\mathbb{1}_X}{\dim(X)} \otimes \text{tr}_X Y \quad (5.1)$$

for every  $Y \in \mathcal{L}(\mathcal{X} \otimes \mathcal{Z})$ , where  $\mathcal{Z}$  is an arbitrary complex Euclidean space. We will also need the following projection operator

$$L_V(W) = \mathcal{A}_O W + \mathcal{B}_O W - \mathcal{A}_O \mathcal{B}_O W - \mathcal{B}_I \mathcal{B}_O W + \mathcal{A}_O \mathcal{B}_I \mathcal{B}_O W - \mathcal{A}_I \mathcal{A}_O W + \mathcal{A}_O \mathcal{A}_I \mathcal{B}_O W. \quad (5.2)$$

where  $W \in \text{Herm}(\mathcal{A}_I \otimes \mathcal{A}_O \otimes \mathcal{B}_I \otimes \mathcal{B}_O)$ .

**Definition 20** *We say that  $W \in \text{Herm}(\mathcal{A}_I \otimes \mathcal{A}_O \otimes \mathcal{B}_I \otimes \mathcal{B}_O)$  is a process matrix if it fulfills the following conditions*

$$W \geq 0, \quad W = L_V(W), \quad \text{tr}(W) = \dim(\mathcal{A}_O) \cdot \dim(\mathcal{B}_O), \quad (5.3)$$

where the projection operator  $L_V$  is defined by Eq. (5.2).

The set of all process matrices will be denoted by  $\mathbf{W}^{\text{PROC}}$ . In the upcoming considerations, it will be more convenient to work with the equivalent characterization of process matrices which can be found in [98].

**Definition 21** *We say that  $W \in \mathbf{W}^{\text{PROC}}$  is a process matrix if it fulfills the following conditions*

$$\begin{aligned} W &\geq 0, \\ \mathcal{A}_I \mathcal{A}_O W &= \mathcal{A}_O \mathcal{A}_I \mathcal{B}_O W, \\ \mathcal{B}_I \mathcal{B}_O W &= \mathcal{A}_O \mathcal{B}_I \mathcal{B}_O W, \\ W + \mathcal{A}_O \mathcal{B}_O W &= \mathcal{B}_O W + \mathcal{A}_O W, \\ \text{tr}(W) &= \dim(\mathcal{A}_O) \cdot \dim(\mathcal{B}_O). \end{aligned} \quad (5.4)$$

The concept of a bipartite process matrix can be best illustrated by considering two characters, Alice and Bob, performing experiments in two separate laboratories. Each party acts in a local laboratory, which can be identified by an input space  $\mathcal{A}_I$  and an output space  $\mathcal{A}_O$  for Alice, and analogously  $\mathcal{B}_I$  and  $\mathcal{B}_O$  for Bob. In general, a label  $i$ , denoting Alice's measurement outcome, is associated with the CP map  $\Phi_{M_i^A}$  obtained from the instrument  $\{\Phi_{M_i^A}\}_i$ . Analogously, the Bob's measurement outcome  $j$  is associated with the map  $\Phi_{M_j^B}$  from the instrument  $\{\Phi_{M_j^B}\}_j$ . Finally, the joint probability for a pair of outcomes  $i$  and  $j$  can be expressed as

$$p_{ij} = \text{tr} [W (M_i^A \otimes M_j^B)], \quad (5.5)$$

where  $W \in \mathbf{W}^{\text{PROC}}$  is a process matrix that describes the causal structure outside of the laboratories. The valid process matrix is defined by the requirement that probabilities are well defined, that is, they must be non-negative and sum up

to one. These requirements give us the conditions present in Definition 20 and Definition 21.

In the general case, Alice's and Bob's strategies can be more complex than the strategy described by tensor product  $M_i^A \otimes M_j^B$  (which defines the probability  $p_{ij}$  given by Eq. (5.5)). If their action is somehow correlated, the strategy is written by an associated instrument of the form  $\{\mathcal{N}_{ij}^{AB}\}_{ij}$ . It was observed in [98] that the instrument  $\{\mathcal{N}_{ij}^{AB}\}_{ij}$  satisfies the following equation

$$\text{tr} \left( W \sum_{ij} N_{ij}^{AB} \right) = 1 \quad (5.6)$$

for all process matrix  $W \in \mathbf{W}^{\text{PROC}}$  if and only if

$$\sum_{ij} N_{ij}^{AB} \in \mathbf{NS}(\mathcal{A}_I \otimes \mathcal{A}_O \otimes \mathcal{B}_I \otimes \mathcal{B}_O). \quad (5.7)$$

### 5.1.2 N-partite process matrices

The generalization to the  $N$ -partite case is rather straightforward. Let us define a Hermitian operator  $W \in \text{Herm}(\mathcal{A}_I^1 \otimes \mathcal{A}_O^1 \otimes \dots \otimes \mathcal{A}_I^N \otimes \mathcal{A}_O^N)$  and the projection operator

$$L_V(W) = [1 - \prod_i (1 - \mathcal{A}_O^i + \mathcal{A}_I^i \mathcal{A}_O^i) + \prod_i \mathcal{A}_I^i \mathcal{A}_O^i] W. \quad (5.8)$$

**Definition 22** *We say that  $W \in \text{Herm}(\mathcal{A}_I^1 \otimes \mathcal{A}_O^1 \otimes \dots \otimes \mathcal{A}_I^N \otimes \mathcal{A}_O^N)$  is  $N$ -partite process matrix if it fulfills the following conditions*

$$W \geq 0, \quad W = L_V(W), \quad \text{tr}(W) = \dim(\mathcal{A}_O^1) \cdot \dots \cdot \dim(\mathcal{A}_O^N), \quad (5.9)$$

where the projection operator  $L_V$  is defined by Eq. (5.8).

A schematic representation of  $N$ -partite process matrix is described in Fig. 5.1.

## 5.2 Advantage of SAR of von Neumann measurements by using indefinite causal structure theory

In Chapter 4, we considered the storage and retrieval of a von Neumann measurement with fixed causal structure realized by quantum combs  $\mathcal{L}$  [70, 71]. In this section, we introduce a completely new approach of SAR using the theory



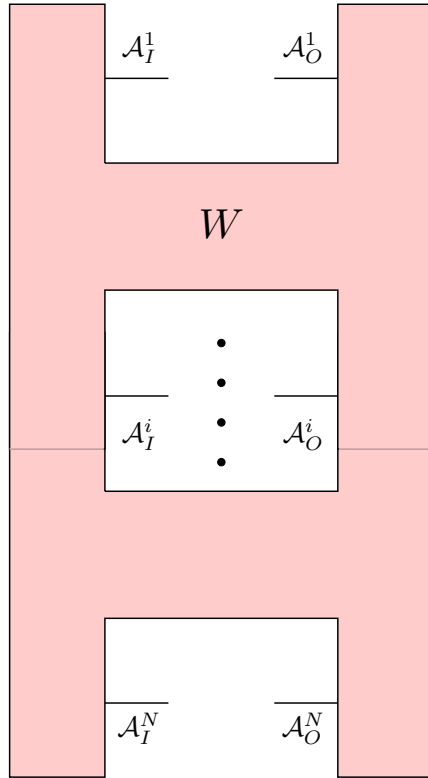


Figure 5.1: Schematic representation of  $N$ -partite process matrix  $W \in \text{Herm}(\mathcal{A}_I^1 \otimes \mathcal{A}_O^1 \otimes \dots \otimes \mathcal{A}_I^N \otimes \mathcal{A}_O^N)$ .

of causal connection. For this purpose we will consider a learning scheme such that its storage part is described by a process matrix. Although the usage of an indefinite causal structure does not improve the fidelity function in problem  $N \rightarrow 1$  unitary channels learning scheme [61] or  $2 \rightarrow 1$  von Neumann measurements learning scheme (see subsection 5.2.2), in subsection 5.2.3 we will show that the usage of indefinite causal structure theory can improve the capacity of  $N \rightarrow 1$  learning scheme of von Neumann measurements for  $N \geq 3$ .

### 5.2.1 General approach of causal learning scheme

Here, we characterize a quantum network describing  $N \rightarrow 1$  causal learning scheme. Let us consider a quantum network  $\mathcal{L}$  with Choi matrix  $L$ . As previous, we consider that  $L$  is a composition of storing and retrieving operations, that is  $L = R * S$ . Now, however, we assume that the storage  $S$  has an indefinite causal order. More precisely, the storage  $S$  is given by  $N$ -partite process matrix  $W$ , such that  $W = \text{tr}_a S$ . A schematic representation of the setup we can see in Fig. 5.2.

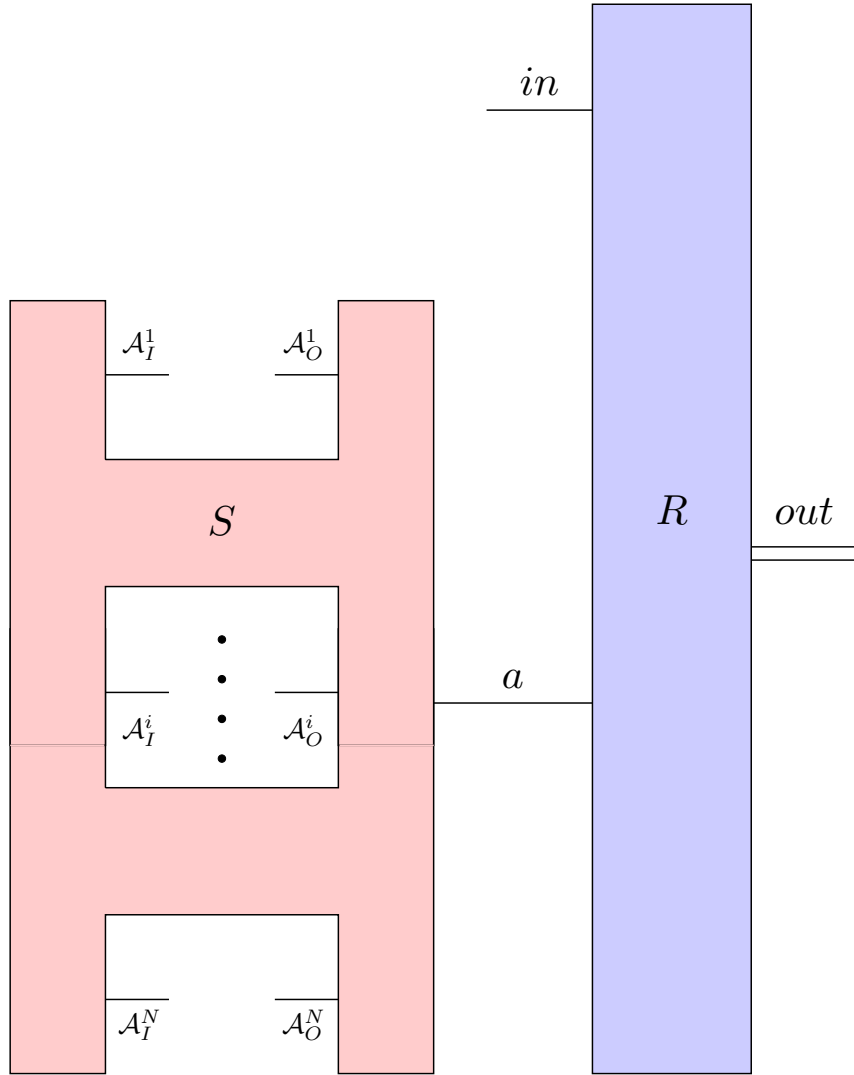


Figure 5.2: Schematic representation of a causal learning scheme  $N \rightarrow 1$  of von Neumann measurements.

### 5.2.2 $2 \rightarrow 1$ causal learning scheme of von Neumann measurements

Let us consider  $2 \rightarrow 1$  causal learning scheme  $\mathcal{L}$  in which we learn a von Neumann measurement  $\mathcal{P}_U$  by using two copies of it. A schematic representation of this setup is presented on Fig. 5.3. Here, the Choi operator  $L \in L(\mathcal{A}_I \otimes \mathcal{A}_O \otimes \mathcal{B}_I \otimes \mathcal{B}_O \otimes \mathcal{X}_{in} \otimes \mathcal{X}_{out})$  satisfies the condition

$$\mathrm{tr}_{\mathcal{X}_{out}} L = \mathbb{1}_{\mathcal{X}_{in}} \otimes \mathrm{tr}_{\mathcal{X}_a} W, \quad (5.10)$$

where  $W$  is a bipartite process matrix defined in Eq. (21).

The following proposition says that no matter which approach we use, the maximum value of the average fidelity for  $2 \rightarrow 1$  learning scheme of von Neumann measurements is the same.

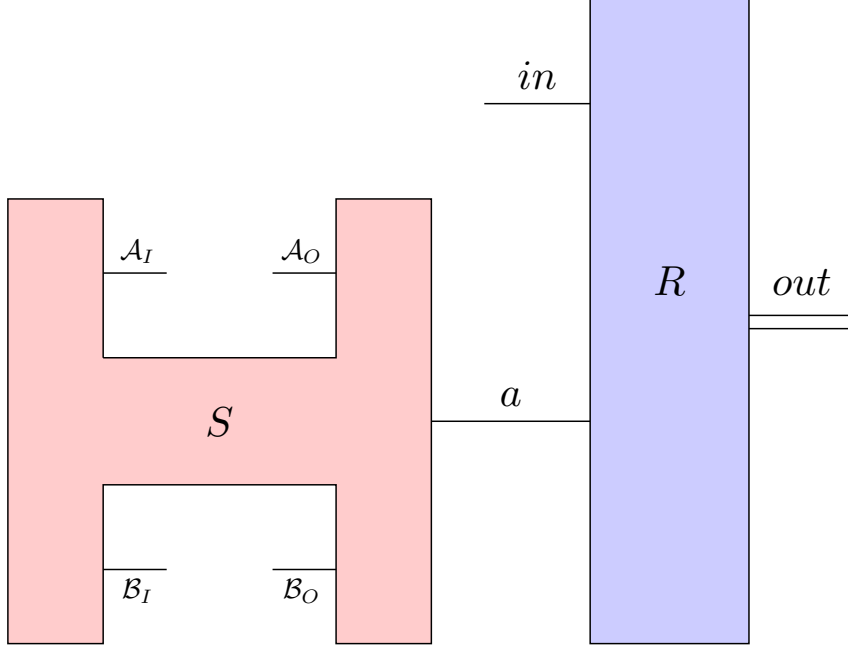


Figure 5.3: Causal learning scheme  $2 \rightarrow 1$  of von Neumann measurements.

**Theorem 7** *The usage of indefinite causal structure does not improve the fidelity function in  $2 \rightarrow 1$  learning scheme of von Neumann measurements.*

**Proof.** Due to the fact that the quantum network  $\mathcal{L}$  has classical labels on spaces  $\mathcal{A}_O, \mathcal{B}_O$  and  $\mathcal{X}_{out}$  then its Choi matrix  $L \in \text{Herm}(\mathcal{A}_I \otimes \mathcal{A}_O \otimes \mathcal{B}_I \otimes \mathcal{B}_O \otimes \mathcal{X}_{in} \otimes \mathcal{X}_{out})$  has the following form

$$L = \sum_i |i\rangle\langle i|_{\mathcal{X}_{out}} \otimes L_i, \quad (5.11)$$

where

$$L_i = \sum_{j,k} |j\rangle\langle j|_{\mathcal{B}_O} \otimes |k\rangle\langle k|_{\mathcal{A}_O} \otimes L_{ijk}. \quad (5.12)$$

Simultaneously, we know that

$$\sum_i L_i = \sum_{i,j,k} |j\rangle\langle j|_{\mathcal{B}_O} \otimes |k\rangle\langle k|_{\mathcal{A}_O} \otimes L_{ijk} = \mathbb{1}_{\mathcal{X}_{in}} \otimes W. \quad (5.13)$$

Without loss of generality we can assume that  $L_{ijk}$  satisfies the commutative relation (see Lemma 9.3 in [61])

$$[L_{ijk}, U_{\mathcal{A}} \otimes U_{\mathcal{B}} \otimes U^\dagger] = 0, \quad (5.14)$$

where  $U_{\mathcal{A}} \in L(\mathcal{A}_I \otimes \mathcal{A}_O)$ ,  $U_{\mathcal{B}} \in L(\mathcal{B}_I \otimes \mathcal{B}_O)$  and  $U^\dagger \in L(\mathcal{X}_{in} \otimes \mathcal{X}_{out})$ . So, we have

$$\forall j, k \quad \sum_i L_{ijk} = \mathbb{1}_{\mathcal{X}_{in}} \otimes \langle j |_{\mathcal{B}_O} \langle k |_{\mathcal{A}_O} W | j \rangle_{\mathcal{B}_O} | k \rangle_{\mathcal{A}_O}. \quad (5.15)$$

From relabeling symmetry property (see Lemma 9.4 in [61]) given by  $L_{ijk} = L_{\sigma(i)\sigma(j)\sigma(k)}$  we have

$$\sum_i L_{ijk} = \sum_i L_{\sigma(i)\sigma(j)\sigma(k)} = \mathbb{1}_{\mathcal{X}_{in}} \otimes \langle \sigma(j) | \langle \sigma(k) | W | \sigma(j) \rangle | \sigma(k) \rangle, \quad (5.16)$$

for any permutation  $\sigma$ . Therefore, we have

$$W = \sum_{j,k} |j\rangle \langle j|_{\mathcal{B}_O} \otimes |k\rangle \langle k|_{\mathcal{A}_O} \otimes \frac{1}{d} \sum_i \text{tr}_{\mathcal{X}_{in}} L_{ijk} = \sum_{j,k} |j\rangle \langle j|_{\mathcal{B}_O} \otimes |k\rangle \langle k|_{\mathcal{A}_O} \otimes W_{jk}. \quad (5.17)$$

Hence,

$$\forall j, k, \sigma \quad W_{jk} = W_{\sigma(j)\sigma(k)}. \quad (5.18)$$

It implies that  $W_{11} = \dots = W_{dd}$  and  $W_{12} = W_{ab}$  for all  $a \neq b$ . These properties together with Eq. (5.17) imply that

$$W = \mathbb{1}_{\mathcal{B}_O} \otimes \mathbb{1}_{\mathcal{A}_O} \otimes P + J(\Delta) \otimes (Q - P), \quad (5.19)$$

where  $P = W_{12}$  and  $Q = W_{11}$ . From the definition of then process matrix (more precisely from the condition  $W + {}_{\mathcal{A}_O \mathcal{B}_O} W = {}_{\mathcal{A}_O} W + {}_{\mathcal{B}_O} W$ ) we obtain that  $P = Q$ , which completes the proof.  $\blacksquare$

### 5.2.3 $N \rightarrow 1$ causal learning scheme of qubit von Neumann measurements for $N \geq 3$

**Conjecture 1** *The usage of indefinite causal structure improves the maximum value of the average fidelity function  $F_2$  defined in Eq. (4.2) in  $N \rightarrow 1$  learning scheme of qubit von Neumann measurements.*

To confirm Conjecture 1, we have computed the maximum value of the average fidelity  $F_2$  by using the semidefinite programming (SDP). To optimize this problem we used the Julia programming language along with quantum package

`QuantumInformation.jl` [87] and SDP optimization via SCS solver [88, 89] with a precision  $\epsilon_{\text{abs}} = 10^{-5}$  and  $\epsilon_{\text{rel}} = 10^{-5}$ . The code is available on GitHub [99].

### Numerical investigation

We have researched numerical advantage of SAR for qubit von Neumann measurements ( $d = 2$ ) by using indefinite causal structure approach. We wrote SDP program for  $N \rightarrow 1$  qubit von Neumann measurements learning schemes. For the first case, the numerical results confirm Proposition 7 and the average fidelity function  $F_c$  for indefinite causal structure approach is the same as for parallel and adaptive cases. However, for  $N \geq 3$  the causal learning scheme achieves slight advantage over the standard approach of SAR presented in Chapter 4. Below we summarize the results obtained in this section.

$N$	1	2	3	4	5
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{Causal}})$	0.75000	0.81142	0.86980	0.89816	0.92041
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{Adaptive}})$	0.74999	0.81141	0.86845	0.89684	0.91915
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{Parallel}})$	0.74999	0.81141	0.86764	0.89564	0.91880
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{DPBT}})$	0.50000	0.66667	0.76967	0.83333	0.87449
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PGLS}})$	0.75000	0.79167	0.84375	0.86250	0.88541
$\mathcal{F}_2^{\text{avg}}(\mathcal{L}_{\text{PPBT}})$	0.62500	0.70000	0.75000	0.78571	0.81250

We also present the advantage of causal learning scheme in Figure 5.4.

## 5.3 Conclusion and discussion

In this chapter, we studied the problem of learning an unknown von Neumann measurement from a finite number of copies  $N$  using indefinite causal connection theory. As in the previous chapter, the main goal was to find the asymptotic behavior of the maximum value for the average fidelity function  $F_d$ . First, we considered a causal learning scheme for two copies of  $d$ -dimensional von Neumann measurements. For this case, we proved that using an indefinite causal structure does not improve the average fidelity function  $F_d$ . Next, however, we show the numerical advantage of using causal structure theory for  $N \rightarrow 1$  learning scheme of qubit von Neumann measurements. For this purpose we wrote a semidefinite program calculating the average fidelity function  $F_2$ . This result paves the way toward a complete description of the capabilities of causal connection, not only in the problem of quantum operations learning but in others validation methods of quantum devices.

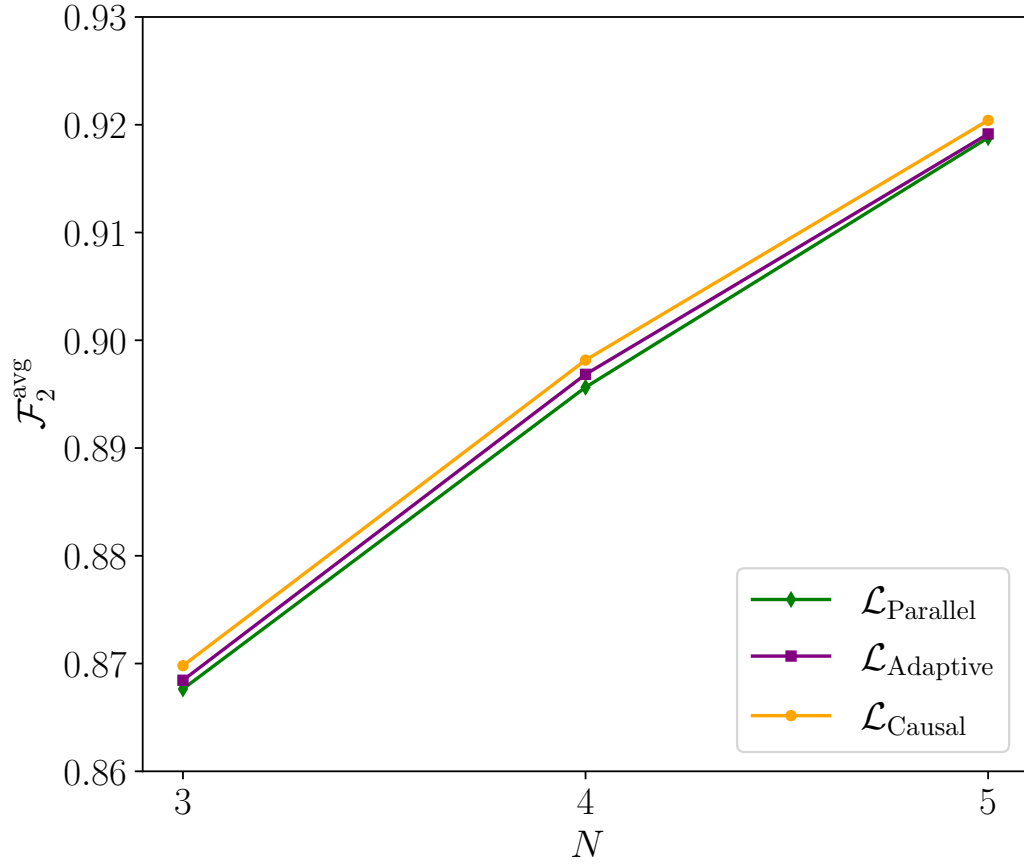


Figure 5.4: The average fidelity function  $\mathcal{F}_2^{\text{avg}}$  calculated for  $N \rightarrow 1$  learning scheme, where  $N = 3, 4, 5$  – optimal causal strategy,  $\mathcal{L}_{\text{Causal}}$  (numerical value, orange circles); optimal adaptive strategy,  $\mathcal{L}_{\text{Adaptive}}$  (numerical value, purple squares); optimal parallel scheme,  $\mathcal{L}_{\text{Parallel}}$  (numerical value, green triangles).

## Chapter 6

# Discrimination of von Neumann measurements

No possibility to perfect learning of von Neumann measurements has caused the rapid development of validation techniques for quantum devices. One of them is the discrimination of quantum operations. The task of discrimination can be best illustrated as follows. Suppose that Alice has access to a device performing one of two quantum operations. While such operations are known to Bob, he needs to find out which was performed by Alice. He would like to guess, with the highest possible probability, which of the operations was used by Alice. For this purpose, they need to construct an optimal strategy to maximize the probability of correct discrimination. What is the highest probability of making a correct guess? And what do you have to do to achieve it? We will answer these questions in this chapter.

The problem of discrimination of quantum states and channels was solved analytically by Helstrom a few decades ago in [35, 37]. For the case of a single-shot discrimination scenario, researchers have used different approaches, with the possibility of using entanglement. Although the use of entanglement improves the probability of correct discrimination, in [100]. Authors have shown that in the task of discrimination of unitary channels, the entanglement is not necessary. The following papers [2, 101–103] paved the way for studying the discrimination of quantum measurements.

Considering multiple-shot discrimination scenarios, researchers have utilized parallel or adaptive approaches. In the first case, they establish that the discrimination between operations does not require pre-processing and post-processing. One example of such an approach is distinguishing unitary channels [100], or von Neumann measurements [104]. The case when a quantum operation can be used multiple times in an adaptive way was investigated by the authors of [40, 105]. They proved that the use of adaptive strategy and a general notion of quantum

combs can improve discrimination. Examples of quantum channels which cannot be discriminated perfectly in the parallel scheme, but nonetheless can be possible using the adaptive approach, were discussed in [40, 106].

In the dissertation, we will mainly be interested in the discrimination task of quantum measurements. We restrict ourselves to discriminating von Neumann measurements. This is because, unlike other measurement types, they can be implemented on actual hardware. We will use the obtained results to create a benchmarking scheme presented in the further part of the dissertation.

This chapter is divided into two parts. First part, described in Section 6.1, presents a general overview about discrimination of quantum measurements. This section is written based on [2]. The second one, presented in Section 6.2, is dedicated to the discrimination task of single-qubit parameterized Fourier measurements, which is my contribution to the dissertation.

## 6.1 Optimal discrimination scheme

Without loss of generality<sup>1</sup>, we consider discrimination task between measurements  $\mathcal{P}_1$ , performed in the computational Z-basis, and a measurement  $\mathcal{P}_U$  performed in the basis  $U \in \mathcal{U}(\mathcal{X})$ . In general, the discrimination scheme presented in Fig. 6.1, requires an auxiliary system. First, the joint system is prepared in some state  $|\psi_0\rangle$ . Then, one of the measurements, either  $\mathcal{P}_U$  or  $\mathcal{P}_1$ , is performed on the first part of the system. Based on its outcome  $i$ , we choose another binary POVM  $\mathcal{Q}_i$  and perform it on the second system, obtaining the output in  $j$ . Finally, if  $j = 0$ , we say that the performed measurement is  $\mathcal{P}_U$ , otherwise we say that it was  $\mathcal{P}_1$ .

Unsurprisingly, both the  $|\psi_0\rangle$  and the final measurements  $\mathcal{Q}_i$  have to be chosen specifically for given  $U$  to maximize the probability of a correct guess. The detailed description how these choices are made in [37], and below we will recall the most important facts and theorems.

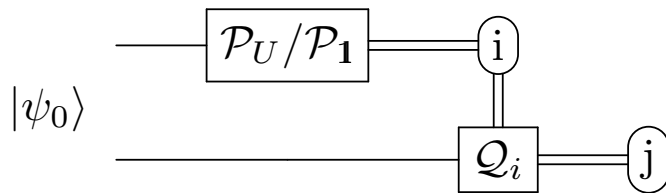


Figure 6.1: Theoretical scheme of discrimination between von Neumann measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$ .

<sup>1</sup>Explaining why we can consider only discrimination scheme between  $\mathcal{P}_1$  and  $\mathcal{P}_U$  is beyond the scope of this work. See [2] for a in depth explanation.



### 6.1.1 Optimal probability

The celebrated result by Helstrom [35] gives the optimal probability of correct discrimination between two quantum measurements,  $\mathcal{P}$  and  $\mathcal{Q}$ , in terms of the diamond norm

$$p_{\text{succ}}(\mathcal{P}, \mathcal{Q}) = \frac{1}{2} + \frac{1}{4} \|\mathcal{P} - \mathcal{Q}\|_{\diamond}, \quad (6.1)$$

where

$$\|\mathcal{P} - \mathcal{Q}\|_{\diamond} = \max_{\|\psi\rangle\|_1=1} \|((\mathcal{P} - \mathcal{Q}) \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)\|_1. \quad (6.2)$$

The quantum state  $|\psi_0\rangle$  which maximizes the diamond norm is called the discriminator.

Although the diamond norm can be efficiently computed via semidefinite programming, it is in general difficult to obtain analytical expressions and those are known only for a few particular cases. From [2, Theorem 1], we can rewrite the diamond norm between two von Neumann measurements in term of the distance between unitary channels.

**Theorem 8** *Let  $\mathcal{P}_U$  and  $\mathcal{P}_1$ ,  $U \in \mathcal{U}(\mathcal{X})$ , be two von Neumann measurements. The following equation holds*

$$\|\mathcal{P}_U - \mathcal{P}_1\|_{\diamond} = \min_{E \in \text{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_1\|_{\diamond}, \quad (6.3)$$

where  $\Phi_U$  is a unitary channel.

Furthermore, it can be shown [37] that the distance between two unitary channels  $\Phi_U$  and  $\Phi_1$  is given by

$$\|\Phi_U - \Phi_1\|_{\diamond} = 2\sqrt{1 - \nu^2(U^\dagger)}, \quad (6.4)$$

where  $\nu(U^\dagger) = \min_{x \in W(U^\dagger)} |x|$ . Due to above results we immediately obtain the following corollary.

**Corollary 7** *Let  $U \in \mathcal{U}(\mathcal{X})$ . Let  $\mathcal{P}_1$  and  $\mathcal{P}_U$  be two von Neumann measurements. The optimal success probability of correct discrimination between  $\mathcal{P}_1$  and  $\mathcal{P}_U$  is given by*

$$p_{\text{succ}}(\mathcal{P}_U, \mathcal{P}_1) = \frac{1}{2} + \frac{1}{4} \min_{E \in \text{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_1\|_{\diamond}. \quad (6.5)$$

### 6.1.2 Discriminator

Here, we present the facts about the discriminator and its properties. First, we recall the form of the discriminator when two von Neumann measurements can not

be perfectly distinguishable. Next, we recall the necessary and sufficient condition when two measurements are perfectly distinguishable.

**Lemma 3** (Lemma 5 from [2], direct implication) *Assume that  $E_0 \in \text{DU}(\mathcal{X})$  satisfies the condition*

$$\|\Phi_{UE_0} - \Phi_{\mathbf{1}}\|_{\diamond} = \|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_{\diamond} < 2. \quad (6.6)$$

*Let  $\lambda_0, \lambda_{d-1}$  be a pair of the most distant eigenvalues of  $UE_0$  and  $\Pi_0, \Pi_{d-1}$  be the projectors onto the subspaces spanned by the eigenvectors corresponding to  $\lambda_0$  and  $\lambda_{d-1}$ , respectively. Then, there exist states  $\rho_0, \rho_{d-1}$ , satisfying the following conditions*

$$\begin{aligned} \rho_0 &= \Pi_0 \rho_0 \Pi_0, \\ \rho_{d-1} &= \Pi_{d-1} \rho_{d-1} \Pi_{d-1}, \\ \text{diag}(\rho_0) &= \text{diag}(\rho_{d-1}). \end{aligned} \quad (6.7)$$

*Furthermore, the quantum state  $\left| \sqrt{\rho^\top} \right\rangle$ , where  $\rho$  is defined as*

$$\rho = \frac{1}{2} \rho_0 + \frac{1}{2} \rho_{d-1} \quad (6.8)$$

*is a discriminator in the problem of discrimination between von Neumann measurements.*

From the above lemma, we can observe the following properties of the discriminator  $\rho$  presented in Proposition 3.

**Proposition 3** *Let  $\rho$  be a quantum state defined in Lemma 3. Then, for each  $i \in \{0, \dots, d-1\}$  we have*

$$\text{tr}(\sqrt{\rho}|i\rangle\langle i|\sqrt{\rho}) = \text{tr}(\sqrt{\rho}U|i\rangle\langle i|U^\dagger\sqrt{\rho}). \quad (6.9)$$

*Moreover, for each  $i \in \{0, \dots, d-1\}$  such that  $\langle i|\rho|i\rangle \neq 0$ , we get*

$$\left| \frac{\langle i|\rho U|i\rangle}{\langle i|\rho|i\rangle} \right| = \left| \frac{\lambda_0 + \lambda_{d-1}}{2} \right|. \quad (6.10)$$

**Proof.** Let  $U = \sum_{i=0}^{d-1} \lambda_i \Pi_i$ , where  $\{\Pi_i\}_{i=0}^{d-1}$  is a set of orthogonal projectors. Then

$$\begin{aligned}
\text{tr}(\sqrt{\rho}U|i\rangle\langle i|U^\dagger\sqrt{\rho}) &= \langle i|U^\dagger\rho U|i\rangle = \langle i|U^\dagger\left(\frac{1}{2}\rho_0 + \frac{1}{2}\rho_{d-1}\right)U|i\rangle \\
&= \langle i|U^\dagger\left(\frac{1}{2}\Pi_0\rho_0\Pi_0 + \frac{1}{2}\Pi_{d-1}\rho_{d-1}\Pi_{d-1}\right)U|i\rangle \\
&= \langle i|\left(\sum_{i=0}^{d-1}\bar{\lambda}_i\Pi_i^\dagger\right)\left(\frac{1}{2}\Pi_0\rho_0\Pi_0 + \frac{1}{2}\Pi_{d-1}\rho_{d-1}\Pi_{d-1}\right)\left(\sum_{i=0}^{d-1}\lambda_i\Pi_i\right)|i\rangle \\
&= \langle i|\left(\frac{1}{2}\rho_0 + \frac{1}{2}\rho_{d-1}\right)|i\rangle = \text{tr}(\sqrt{\rho}|i\rangle\langle i|\sqrt{\rho}).
\end{aligned} \tag{6.11}$$

To prove the second part of the corollary we calculate

$$\begin{aligned}
\left|\frac{\langle i|\rho U|i\rangle}{\langle i|\rho|i\rangle}\right| &= \left|\frac{\langle i|\left(\frac{1}{2}\rho_0 + \frac{1}{2}\rho_{d-1}\right)\left(\sum_{i=0}^{d-1}\lambda_i\Pi_i\right)|i\rangle}{\langle i|\rho|i\rangle}\right| \\
&= \left|\frac{\langle i|\sum_{i=0}^{d-1}\lambda_i\left(\frac{1}{2}\Pi_0\rho_0\Pi_0 + \frac{1}{2}\Pi_{d-1}\rho_{d-1}\Pi_{d-1}\right)\Pi_i|i\rangle}{\langle i|\rho|i\rangle}\right| \\
&= \left|\frac{\langle i|\left(\frac{1}{2}\lambda_0\Pi_0\rho_0\Pi_0 + \frac{1}{2}\lambda_{d-1}\Pi_{d-1}\rho_{d-1}\Pi_{d-1}\right)|i\rangle}{\langle i|\rho|i\rangle}\right| \\
&= \left|\frac{\langle i|\left(\frac{1}{2}\lambda_0\rho_0 + \frac{1}{2}\lambda_{d-1}\rho_{d-1}\right)|i\rangle}{\langle i|\rho|i\rangle}\right| = \left|\frac{\lambda_0 + \lambda_{d-1}}{2}\right|.
\end{aligned} \tag{6.12}$$

■

Here, we recall the necessary and sufficient condition when two measurements are perfectly distinguishable.

**Proposition 4** (Proposition 3 from [2]) *Let  $U \in \mathcal{U}(\mathcal{X})$ . Then  $\mathcal{P}_U$  and  $\mathcal{P}_\mathbf{1}$  are perfectly distinguishable if and only if there exists  $\rho \in \Omega(\mathcal{X})$  such that*

$$\text{diag}(U^\dagger\rho) = 0. \tag{6.13}$$

Moreover, the quantum state  $\left|\sqrt{\rho^\top}\right\rangle$  is a discriminator in the problem of discrimination between von Neumann measurements  $\mathcal{P}_U$  and  $\mathcal{P}_\mathbf{1}$ .

### 6.1.3 Final optimal measurement

Furthermore, due to the proof of the Holevo-Helstrom theorem [37], it is possible to construct corresponding the optimal discrimination final measurements  $\mathcal{Q}_i$ ,

$i \in \{0, \dots, d-1\}$ . This procedure is described below.

1. Let us consider

$$X = (\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|) - (\mathcal{P}_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|), \quad (6.14)$$

where  $|\psi_0\rangle$  is the discriminator maximizing Eq. (6.2).

2. Notice that  $X \in \text{Herm}(\mathcal{X} \otimes \mathcal{X})$ . Then, from Hahn-Jordan decomposition, we express  $X$  as

$$X = P - Q, \quad (6.15)$$

where  $P, Q \in \text{Pos}(\mathcal{X} \otimes \mathcal{X})$  are defined in 2.

3. We create projectors  $\Pi_{\text{im}(P)}$  and  $\Pi_{\text{im}(Q)}$  onto the images of  $P$  and  $Q$ , respectively.
4. Due to that  $X$  is block-diagonal,  $P$  and  $Q$  are block-diagonal as well, so we may write  $\Pi_{\text{im}(P)}$  and  $\Pi_{\text{im}(Q)}$  as

$$\Pi_{\text{im}(P)} = \sum_{i=0}^{d-1} |i\rangle\langle i| \otimes \Pi_{\text{im}(P),i}, \quad (6.16)$$

and

$$\Pi_{\text{im}(Q)} = \sum_{i=0}^{d-1} |i\rangle\langle i| \otimes \Pi_{\text{im}(Q),i}, \quad (6.17)$$

where  $\Pi_{\text{im}(P),i}, \Pi_{\text{im}(Q),i} \in \text{Proj}(\mathcal{X})$  are orthogonal projectors.

5. For  $i \in \{0, \dots, d-1\}$ , we define  $\mathcal{Q}_i$  by

$$\mathcal{Q}_i = \{\Pi_{\text{im}(P),i}, \Pi_{\text{im}(Q),i}\}, \quad (6.18)$$

## 6.2 Discrimination scheme for parameterized family of Fourier measurements

So far, we only discussed how the discrimination scheme is performed, assuming that all needed components  $|\psi_0\rangle$ ,  $\mathcal{Q}_i$  are known. This section provides a specific example of a discrimination scheme using a parametrized family of Fourier qubit von Neumann measurements.

The parametrized family of Fourier measurements is defined as a set of the measurements  $\{\mathcal{P}_{U_\phi} : \phi \in [0, 2\pi]\}$ , where

$$U_\phi = H \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} H^\dagger, \quad (6.19)$$

and  $H$  is the Hadamard matrix of dimension two. For each element of this set, the discriminator is a Bell state

$$|\psi_0\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle). \quad (6.20)$$

The optimal final measurements  $\mathcal{Q}_0$  and  $\mathcal{Q}_1$  are von Neumann measurements defined as follow,  $\mathcal{Q}_i = \mathcal{P}_{V_i}$  and the unitaries  $V_0, V_1$ , which depend on  $\phi$ , have the following form:

$$V_0 = \begin{pmatrix} i \sin\left(\frac{\pi-\phi}{4}\right) & -i \cos\left(\frac{\pi-\phi}{4}\right) \\ \cos\left(\frac{\pi-\phi}{4}\right) & \sin\left(\frac{\pi-\phi}{4}\right) \end{pmatrix}, \quad (6.21)$$

$$V_1 = \begin{pmatrix} -i \cos\left(\frac{\pi-\phi}{4}\right) & i \sin\left(\frac{\pi-\phi}{4}\right) \\ \sin\left(\frac{\pi-\phi}{4}\right) & \cos\left(\frac{\pi-\phi}{4}\right) \end{pmatrix}. \quad (6.22)$$

Finally, the theoretical probability of correct discrimination between von Neumann measurements  $\mathcal{P}_{U_\phi}$  and  $\mathcal{P}_1$  is given by

$$p_{\text{succ}}(\mathcal{P}_{U_\phi}, \mathcal{P}_1) = \frac{1}{2} + \frac{|1 - e^{i\phi}|}{4}. \quad (6.23)$$

We prove the construction of  $|\psi_0\rangle$ ,  $V_0$  and  $V_1$  for parameterized Fourier family of measurements in Section 6.2.2 whereas the calculation of the optimal probability we can see in Section 6.2.1.

### 6.2.1 Optimal probability for parameterized family of Fourier measurements

In this section we present theoretical probability of correct discrimination scheme for parameterized Fourier family. To do that, we will present an auxiliary lemma.

**Lemma 4** *Let  $U = H \text{diag}(1, e^{i\phi})H^\dagger$ ,  $\phi \in [0, 2\pi)$ , and let  $\Phi_U$  and  $\Phi_1$  be two unitary channels. Then, the following equation holds*

$$\min_{E \in \text{DU}(\mathbb{C}^2)} \|\Phi_{UE} - \Phi_1\|_\diamond = \|\Phi_U - \Phi_1\|_\diamond. \quad (6.24)$$

**Proof.** Recall that the distance between two unitary channels is given by

$$\|\Phi_U - \Phi_{\mathbf{1}}\|_{\diamond} = 2\sqrt{1 - \nu^2(U^\dagger)}, \quad (6.25)$$

where  $\nu(U^\dagger) = \min_{x \in W(U^\dagger)} |x|$ . For  $U = H \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} H^\dagger$  we have

$$\nu^2(U^\dagger) = 1 - \frac{|1 - e^{-i\phi}|^2}{4} = 1 - \frac{|1 - e^{i\phi}|^2}{4}. \quad (6.26)$$

So, we obtain

$$\|\Phi_U - \Phi_{\mathbf{1}}\|_{\diamond} = |1 - e^{i\phi}|. \quad (6.27)$$

It implies that it is enough to prove

$$\min_{E \in \text{DU}(\mathbb{C}^2)} \|\Phi_{UE} - \Phi_{\mathbf{1}}\|_{\diamond} = |1 - e^{i\phi}|. \quad (6.28)$$

This condition is equivalent to show that for every  $E \in \text{DU}(\mathbb{C}^2)$

$$\nu(U^\dagger E) \leq \frac{|1 + e^{i\phi}|}{2}, \quad (6.29)$$

where  $\nu(U^\dagger E) = \min_{x \in W(U^\dagger E)} |x|$ . The celebrated Hausdorff-Töplitz theorem [52, 53] states that  $W(A)$  of any matrix  $A \in L(\mathcal{X})$  is a convex set, and therefore we have

$$W(A) = \{\text{tr}(A\rho) : \rho \in \Omega(\mathcal{X})\}. \quad (6.30)$$

So, we can assume that

$$\min_{|x \in \mathbb{C}^2: |x| = 1} |\langle x | U^\dagger | x \rangle| = \min_{\rho \in \Omega(\mathbb{C}^2)} |\text{tr}(U^\dagger \rho)|. \quad (6.31)$$

Then, we have

$$\nu(U^\dagger E) = \min_{\rho \in \Omega(\mathbb{C}^2)} |\text{tr}(\rho UE)|. \quad (6.32)$$

For that, our task is reduced to show that for every  $E \in \text{DU}(\mathbb{C}^2)$  there exists  $\rho \in \Omega(\mathbb{C}^2)$  such that

$$|\text{tr}(\rho UE)| \leq \frac{|1 + e^{i\phi}|}{2}. \quad (6.33)$$

Now, let us define  $E = \begin{pmatrix} E_0 & 0 \\ 0 & E_1 \end{pmatrix}$  and take  $\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ . From spectral

theorem, let us decompose  $U$  as

$$U = \lambda_0 |x_0\rangle\langle x_0| + \lambda_1 |x_1\rangle\langle x_1|, \quad (6.34)$$

where for eigenvalue  $\lambda_0 = 1$ , the corresponding eigenvector is of the form  $|x_0\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ , whereas for  $\lambda_1 = e^{i\phi}$  we have  $|x_1\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ . Then, for every  $E \in \text{DU}(\mathbb{C}^2)$  we have

$$\begin{aligned} |\text{tr}(\rho U E)| &= \frac{1}{2} |\text{tr}(H \text{diag}(1, e^{i\phi}) H^\dagger E)| = \frac{1}{2} |\text{tr}((|x_0\rangle\langle x_0| + e^{i\phi} |x_1\rangle\langle x_1|) E)| \\ &= \frac{1}{2} |\langle x_0| E |x_0\rangle + e^{i\phi} \langle x_1| E |x_1\rangle| = \frac{1}{2} \left| \frac{E_0 + E_1}{2} + e^{i\phi} \frac{E_0 + E_1}{2} \right| \\ &= \frac{|1 + e^{i\phi}|}{2} \left| \frac{E_0 + E_1}{2} \right| \leq \frac{|1 + e^{i\phi}|}{2}, \end{aligned} \quad (6.35)$$

which completes the proof.  $\blacksquare$

**Theorem 9** *The optimal probability of correct discrimination between von Neumann measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$  for  $U = H \text{diag}(1, e^{i\phi}) H^\dagger$ , where  $\phi \in [0, 2\pi)$  is given by*

$$p_{\text{succ}}(\mathcal{P}_U, \mathcal{P}_1) = \frac{1}{2} + \frac{|1 - e^{i\phi}|}{4}. \quad (6.36)$$

**Proof.** From Holevo-Helstrom theorem, we obtain

$$p_{\text{succ}}(\mathcal{P}_U, \mathcal{P}_1) = \frac{1}{2} + \frac{1}{4} \|\mathcal{P}_U - \mathcal{P}_1\|_\diamond. \quad (6.37)$$

From [2, Theorem 1], we have

$$\|\mathcal{P}_U - \mathcal{P}_1\|_\diamond = \min_{E \in \text{DU}(\mathbb{C}^2)} \|\Phi_{UE} - \Phi_1\|_\diamond. \quad (6.38)$$

From Lemma 4, we know that for  $U = H \text{diag}(1, e^{i\phi}) H^\dagger$ , it also holds that

$$\min_{E \in \text{DU}(\mathbb{C}^2)} \|\Phi_{UE} - \Phi_1\|_\diamond = \|\Phi_U - \Phi_1\|_\diamond, \quad (6.39)$$

which is exactly equal to

$$\|\Phi_U - \Phi_1\|_\diamond = 2\sqrt{1 - \nu^2(U^\dagger)} = |1 - e^{i\phi}|. \quad (6.40)$$

It implies that

$$p_{\text{succ}}(\mathcal{P}_U, \mathcal{P}_1) = \frac{1}{2} + \frac{|1 - e^{i\phi}|}{4}, \quad (6.41)$$

which completes the proof.  $\blacksquare$

## 6.2.2 Optimal discrimination strategy for parameterized family of Fourier measurements

In this section we create an optimal theoretical strategy of discrimination between  $\mathcal{P}_U$  and  $\mathcal{P}_1$ . To indicate the optimal strategy, we will present two propositions. The first one is concentrated around the discriminator  $|\psi_0\rangle$  (as the optimal input state), whereas the proposition describes the optimal final measurement  $\mathcal{P}_{V_i}$ .

### Discriminator

**Proposition 5** *Consider the problem of discrimination between von Neumann measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$ ,  $U = H \text{diag}(1, e^{i\phi})H^\dagger$  and  $\phi \in [0, 2\pi)$ . The discriminator has the form*

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}|\mathbb{1}_{\mathbb{C}^2}\rangle. \quad (6.42)$$

**Proof.** Observe that  $\mathcal{P}_U - \mathcal{P}_1$  is a Hermiticity-preserving map. Recall, for Hermiticity-preserving maps the diamond norm may be expressed as

$$\|\mathcal{M}\|_\diamond = \max_{\rho \in \Omega(\mathcal{X})} \|(\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho}) J(\mathcal{M}) (\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho})\|_1. \quad (6.43)$$

Let us calculate the Choi matrix of  $J(\mathcal{P}_U - \mathcal{P}_1)$ . We have

$$J(\mathcal{P}_U - \mathcal{P}_1) = \sum_{i=0}^1 |i\rangle\langle i| \otimes (|u_i\rangle\langle u_i| - |i\rangle\langle i|)^\top, \quad (6.44)$$

where the projector  $|u_i\rangle\langle u_i|$  is taken as  $(i+1)$ -th column of  $U$ . Assume that

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (6.45)$$



Hence, we have

$$\begin{aligned}
& \|(\mathbb{1}_{\mathbb{C}^2} \otimes \sqrt{\rho}) J(\mathcal{P}_U - \mathcal{P}_1) (\mathbb{1}_{\mathbb{C}^2} \otimes \sqrt{\rho})\|_1 = \\
&= \frac{1}{2} \left\| \sum_{i=0}^1 |i\rangle\langle i| \otimes (|u_i\rangle\langle u_i| - |i\rangle\langle i|)^\top \right\|_1 \\
&= \frac{1}{2} \sum_{i=0}^1 \| |u_i\rangle\langle u_i| - |i\rangle\langle i| \|_1.
\end{aligned} \tag{6.46}$$

One can prove that for all  $\alpha, \beta \geq 0$ , and unit vectors  $|x\rangle, |y\rangle \in \mathcal{X}$  the following equation holds [37]

$$\| \alpha |x\rangle\langle x| - \beta |y\rangle\langle y| \|_1 = \sqrt{(\alpha + \beta)^2 - 4\alpha\beta |\langle x|y\rangle|^2}. \tag{6.47}$$

By taking  $|x\rangle = |u_i\rangle$  and  $|y\rangle = |i\rangle$  we have

$$\frac{1}{2} \sum_{i=0}^1 \| |u_i\rangle\langle u_i| - |i\rangle\langle i| \|_1 = \frac{1}{2} \sum_{i=0}^1 \sqrt{4 - 4|\langle i|u_i\rangle|^2} = \sum_{i=0}^1 \sqrt{1 - |\langle i|u_i\rangle|^2}. \tag{6.48}$$

Direct calculation gives us the following equality

$$\begin{aligned}
& \|(\mathbb{1}_{\mathbb{C}^2} \otimes \sqrt{\rho}) J(\mathcal{P}_U - \mathcal{P}_1) (\mathbb{1}_{\mathbb{C}^2} \otimes \sqrt{\rho})\|_1 \\
&= \sum_{i=0}^1 \sqrt{1 - |\langle i|U|i\rangle|^2} \\
&= \sum_{i=0}^1 \sqrt{1 - \left| \frac{1 + e^{i\phi}}{2} \right|^2} = 2 \sqrt{1 - \left| \frac{1 + e^{i\phi}}{2} \right|^2} \\
&= |1 - e^{i\phi}|.
\end{aligned} \tag{6.49}$$

Due to Theorem 9 and the following equality

$$\| (\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho}) J(\mathcal{P}_U - \mathcal{P}_1) (\mathbb{1}_{\mathcal{X}} \otimes \sqrt{\rho}) \|_1 = \left\| ((\mathcal{P}_U - \mathcal{P}_1) \otimes \mathcal{I}_{\mathcal{X}}) \left( |\sqrt{\rho}^\top\rangle\rangle \langle\langle \sqrt{\rho}^\top| \right) \right\|_1, \tag{6.50}$$

the discriminator  $|\psi_0\rangle$  has the form

$$|\psi_0\rangle = \left| \sqrt{\rho}^\top \right\rangle\rangle = \frac{1}{\sqrt{2}} |\mathbb{1}_2\rangle\rangle, \tag{6.51}$$

which completes the proof. ■

## Final optimal measurement

To construct the final optimal measurements  $\mathcal{Q}_i$ ,  $i \in \{0, 1\}$ , we use the procedure presented in Section 6.1.3. Here,  $\mathcal{X} = \mathbb{C}^2$ , hence, it holds

$$\Pi_{\text{im}(P)} = \begin{pmatrix} |x_p\rangle\langle x_p| & 0 \\ 0 & |y_p\rangle\langle y_p| \end{pmatrix}, \quad (6.52)$$

and

$$\Pi_{\text{im}(Q)} = \begin{pmatrix} |x_q\rangle\langle x_q| & 0 \\ 0 & |y_q\rangle\langle y_q| \end{pmatrix}. \quad (6.53)$$

And finally, we define  $V_0$  as

$$\begin{cases} |x_p\rangle = V_0|0\rangle \\ |x_q\rangle = V_0|1\rangle \end{cases} \quad (6.54)$$

and  $V_1$  as

$$\begin{cases} |y_p\rangle = V_1|0\rangle \\ |y_q\rangle = V_1|1\rangle. \end{cases} \quad (6.55)$$

One can note that the optimal measurements  $\mathcal{Q}_0$  and  $\mathcal{Q}_1$  are von Neumann measurements and can be described by the relation  $\mathcal{Q}_i = \mathcal{P}_{V_i}$ . Next, we determine the values of  $V_i$ . We also include the `Mathematica` notebook [107] in `mathematics` file computing  $V_0$  and  $V_1$  of  $\mathcal{P}_{V_i}$ ,  $i \in \{0, 1\}$ . By the direct calculations, we immediately obtain the following proposition.

**Proposition 6** *Consider the problem of discrimination between von Neumann measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$ , where  $U = H \text{diag}(1, e^{i\phi})H^\dagger$ ,  $\phi \in [0, 2\pi)$ . The unitaries  $V_0$  and  $V_1$  have the following forms*

$$V_0 = \begin{pmatrix} i \sin\left(\frac{\pi-\phi}{4}\right) & -i \cos\left(\frac{\pi-\phi}{4}\right) \\ \cos\left(\frac{\pi-\phi}{4}\right) & \sin\left(\frac{\pi-\phi}{4}\right) \end{pmatrix}, \quad (6.56)$$

and

$$V_1 = \begin{pmatrix} -i \cos\left(\frac{\pi-\phi}{4}\right) & i \sin\left(\frac{\pi-\phi}{4}\right) \\ \sin\left(\frac{\pi-\phi}{4}\right) & \cos\left(\frac{\pi-\phi}{4}\right) \end{pmatrix}. \quad (6.57)$$

## 6.3 Conclusion and discussion

In this section, we studied the problem of discrimination of von Neumann measurements. In the beginning, we described the task of discrimination and the needed components to create the optimal strategy which maximizes the probability of correct discrimination. We also recalled the Holevo-Helstrom theorem, which gives

us the formula for the maximized value of correct discrimination. Due to the proof of the Holevo-Helstrom theorem, we obtain the necessary components to create the optimal discrimination strategy.

Although the theoretical problems of discrimination of quantum measurements are well-studied, an analytical form of probability and discrimination strategy is generally hard to determine. For this purpose, we focused on the discrimination task for single-qubit parameterized Fourier family of measurements. For this case, we calculated the exact value of the probability of correct discrimination and pointed out an optimal discrimination strategy for achieving such probability. Due to that, we will be able to use the discrimination task to create a new aspect of the benchmarking scheme, which we will describe in the next chapter.



# Chapter 7

## PyQBench: a Python library for benchmarking gate-based quantum computers

Suppose that we know a strategy that, for an ideal device, would yield a probability  $p_{\text{succ}}$  of successfully discriminating between two measurements. Will the probability be the same on an actual physical device? For current imperfect Noisy Intermediate Scale Quantum devices (NISQs) the answer is: not. It is well known that NISQ devices have their limitations [19]. The question is to what extent those devices can perform meaningful computations? To answer this question, one has to devise a methodology for benchmarking metric.

PyQBench helps us in organizing such experiments for a single-qubit parameterized Fourier family of measurements, executing them on real hardware or simulators, and computing probabilities based on the measured bitstrings. Initially, benchmarks in PyQBench work by experimentally determining the probability of correct discrimination between two qubit von Neumann measurements by the device under test and comparing the result with the ideal, theoretical predictions. Naturally, we need to repeat the same procedure multiple times for both measurements to obtain a reliable estimate of the underlying probability distribution. PyQBench can benchmark NISQ devices by verifying their capability of discriminating between two von Neumann measurements. PyQBench offers a simplified, ready-to-use, command line interface (CLI) for running benchmarks using a predefined parametrized Fourier family of measurements. For more advanced scenarios, PyQBench offers a way of employing user-defined measurements instead of predefined ones. Due to this possibility of extension, the users are able to test specific aspects of their architecture of interest. Finally, we provide the documentation [108] and source code of PyQBench on GitHub [107] under an open source license which will allow users to utilize and extend our package in their specific applications. Our package

allows the user to test various architectures, available through `qiskit` and Amazon BraKet [17] using problems with simple operational interpretation.

PyQBench is not the only software package for benchmarking gate-based devices. Probably the simplest benchmarking method one could devise is simply running known algorithms and comparing outputs with the expected ones. Analyzing the frequency of the correct outputs, or the deviation between actual and expected outputs distribution provides then a metric of the performance of a given device. Libraries such as Munich Quantum Toolkit (MQT) [109,110] or SupermarQ [111,112] contain benchmarks leveraging multiple algorithms, such as Shor’s algorithm or Grover’s algorithm. Despite being intuitive and easily interpretable, such benchmarks may have some problems. Most importantly, they assess the usefulness of a quantum device only for a very particular algorithm, and it might be hard to extrapolate their results to other algorithms and applications. For instance, the inability of a device to consistently find factorizations using Shor’s algorithms does not tell anything about its usefulness in Variational Quantum Algorithms. Another possible approach to benchmarking quantum computers is randomized benchmarking. In this approach, one samples circuits to be run from some predefined set of gates (e.g. from the Clifford group) and tests how much the output distribution obtained from the device running these circuits differs from the ideal one. It is also common to concatenate randomly chosen circuits with their inverses (which should yield the identity circuit) and run those concatenated circuits on the device. Libraries implementing this approach include Qiskit [113] or PyQuil [114]. Another quantity used for benchmarking NISQ devices is quantum volume. The quantum volume characterizes capacity of a device for solving computational problems. It takes into account multiple factors like number of qubits, connectivity and measurement errors. The Qiskit library allows one to measure quantum volume of a device by using its `qiskit.ignis.verification.quantum_volume`. Other implementations of Quantum Volume can be found as well, see e.g. [115]. However, we believe that our library has significant benefits over other benchmarking techniques such that simple operational interpretation, ability to control benchmarking process or easy-to-use CLI tool.

This chapter is organized as follows. In Section 7.1 we formulate limitations of NISQ devices and possible practical solutions to implement a discrimination scheme. In Section 7.2 we describe PyQBench software functionalities, whereas in Section 7.3 software architecture. In particular, we expose the CLI tool for parameterized Fourier family of measurements, as one of the possible functionalities of PyQBench. Next, we present the usage PyQBench as a Python library. We also include Appendix B with detailed description of usage PyQBench. Finally, Section 7.4 concludes this chapter with a summary of the main results. This chapter is based on [3].

## 7.1 Limitations of NISQ devices and solutions

Let us recall the scheme of discrimination of von Neumann measurements presented in Fig. 7.1

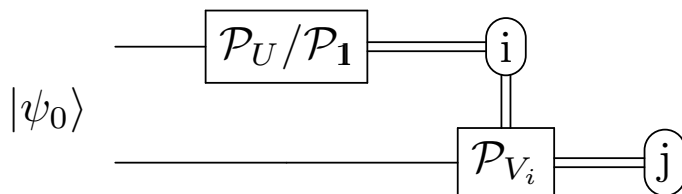


Figure 7.1: Theoretical scheme of discrimination between von Neumann measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$ .

It is a fact that current NISQ devices are unable to perform conditional measurements  $\mathcal{P}_{V_i}$ , which is the biggest obstacle to implementing our scheme on real hardware. However, we circumvent this problem by slightly adjusting our scheme so that it only uses components available on current devices. It can be achieved by the observation that each von Neumann measurement  $\mathcal{P}_U$  can be decomposed as in Figure 7.2. Due to that, we create two possible options: using a postselection or a direct sum.

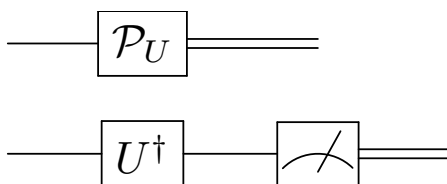


Figure 7.2: Implementation of a von Neumann measurement using measurement in computational basis. The upper circuit shows a symbolic representation of a von Neumann measurement  $\mathcal{P}_U$ . The bottom, equivalent circuit depicts its decomposition into a change of basis followed by measurement in the  $Z$ -basis.

### Postselection

The first idea uses a postselection scheme. In the original scheme, we measure the first qubit and only then determine which measurement should be performed on the second one. Instead of doing this choice, we can run two circuits, one with  $\mathcal{P}_{V_0}$  and one with  $\mathcal{P}_{V_1}$  and measure both qubits. We then discard the results of the circuit for which label  $i$  does not match measurement label  $k$ . Hence, the circuit for postselection looks as depicted in Fig. 7.3.

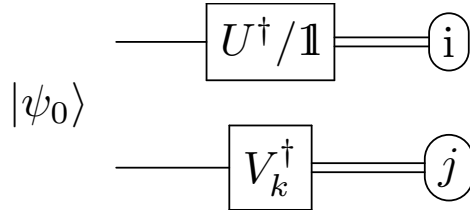


Figure 7.3: A schematic representation of the setup for distinguishing measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$  using postselection approach. In postselection scheme, one runs such circuits for both  $k = 0, 1$  and discards results for cases when there is a mismatch between  $k$  and  $i$ .

To perform the benchmark, one needs to run multiple copies of the postselection circuit, with both  $\mathcal{P}_U$  and  $\mathcal{P}_1$ . In PyQBench, we assume that the experiment is repeated the same number of times for both  $\mathcal{P}_U$  and  $\mathcal{P}_1$ . Each circuit has to be run in both variants, one with final measurement  $\mathcal{P}_{V_0}$  and the second with the final measurement  $\mathcal{P}_{V_1}$ . The experiments can thus be grouped into classes identified by tuples of the form  $(\mathcal{Q}, k, i, j)$ , where  $\mathcal{Q} \in \{\mathcal{P}_U, \mathcal{P}_1\}$  denotes the chosen measurement,  $k \in \{0, 1\}$  designates the final measurement used, and  $i \in \{0, 1\}$  and  $j \in \{0, 1\}$  being the labels of outcomes as presented in Fig. 7.3. We then discard all the experiments for which  $i \neq k$ . The total number of valid experiments is thus:

$$N_{\text{total}} = \#\{(\mathcal{Q}, k, i, j) : k = i\}. \quad (7.1)$$

Finally, we count the valid experiments resulting in successful discrimination. If we have chosen  $\mathcal{P}_U$ , then we guess correctly iff  $j = 0$ . Similarly, for  $\mathcal{P}_1$ , we guess correctly iff  $j = 1$ . If we define

$$N_{\mathcal{P}_U} = \#\{(\mathcal{Q}, k, i, j) : \mathcal{Q} = \mathcal{P}_U, k = i, j = 0\}, \quad (7.2)$$

$$N_{\mathcal{P}_1} = \#\{(\mathcal{Q}, k, i, j) : \mathcal{Q} = \mathcal{P}_1, k = i, j = 1\}, \quad (7.3)$$

then the empirical success probability can be computed as

$$p_{\text{succ}}(\mathcal{P}_U, \mathcal{P}_1) = \frac{N_{\mathcal{P}_U} + N_{\mathcal{P}_1}}{N_{\text{total}}}. \quad (7.4)$$

The  $p_{\text{succ}}$  is the quantity reported to the user as the result of the benchmark.

### Direct sum

The second idea uses the direct sum  $V_0^\dagger \oplus V_1^\dagger$  implementation. Here, instead of performing a conditional measurement  $\mathcal{P}_{V_k}$ , where  $k \in \{0, 1\}$ , we run circuits presented in Fig. 7.4.



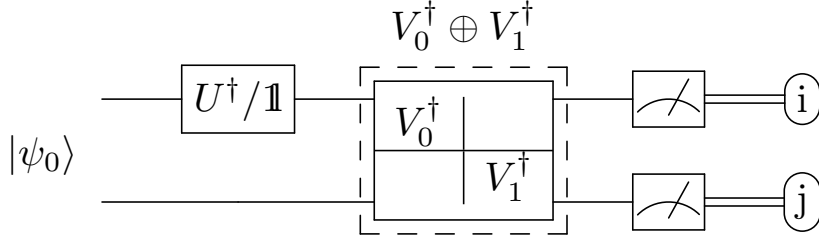


Figure 7.4: A schematic representation of the setup for distinguishing measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$  using the  $V_0^\dagger \oplus V_1^\dagger$  direct sum.

One can directly calculate that such a circuit is equivalent to the original discrimination scheme. Observe that, for any state  $\rho \in \Omega(\mathbb{C}^2)$ , it holds that

$$\begin{aligned}
 (\mathcal{P}_1 \otimes \mathcal{P}_1) \left( \Phi_{V_0^\dagger \oplus V_1^\dagger} \right) (\rho) &= (\mathcal{P}_1 \otimes \mathcal{P}_1) \left( \sum_{i=0}^1 |i\rangle\langle i| \otimes \left( \langle i| \otimes V_i^\dagger \right) \rho \left( |i\rangle \otimes V_i^\dagger \right) \right) \\
 &= \sum_{i=0}^1 |i\rangle\langle i| \otimes \mathcal{P}_{V_i} \left( \langle i| \otimes \mathbb{1} \right) \rho \left( |i\rangle \otimes \mathbb{1} \right),
 \end{aligned} \tag{7.5}$$

where we use the block-diagonal form of  $V_0^\dagger \oplus V_1^\dagger$ , that is

$$V_0^\dagger \oplus V_1^\dagger = |0\rangle\langle 0| \otimes V_0^\dagger + |1\rangle\langle 1| \otimes V_1^\dagger. \tag{7.6}$$

In direct sum scheme, the experiment can be characterized by a pair  $(\mathcal{Q}, i, j)$ , where  $\mathcal{Q} = \{\mathcal{P}_U, \mathcal{P}_1\}$  and  $i, j \in \{0, 1\}$  are the output labels. The number of successful trials for  $U$  and  $\mathbb{1}$ , respectively, can be written as

$$N_{\mathcal{P}_U} = \#\{(\mathcal{Q}, i, j) : \mathcal{Q} = \mathcal{P}_U, j = 0\}, \tag{7.7}$$

$$N_{\mathcal{P}_1} = \#\{(\mathcal{Q}, i, j) : \mathcal{Q} = \mathcal{P}_1, j = 1\}. \tag{7.8}$$

Then, the probability of correct discrimination between  $\mathcal{P}_U$  and  $\mathcal{P}_1$  is given by

$$p_{\text{succ}} = \frac{N_{\mathcal{P}_U} + N_{\mathcal{P}_1}}{N_{\text{total}}}, \tag{7.9}$$

where  $N_{\text{total}}$  is the number of trials.

## 7.2 Software Functionalities

The PyQBench can be used in two modes: as a Python library and as a CLI script. When used as a library, PyQBench allows the customization of discrimination scheme. The user provides a unitary matrix  $U$  defining the measurement to be discriminated, the discriminator  $|\psi_0\rangle$ , and unitaries  $V_0$  and  $V_1$  describing the final measurement. The PyQBench library provides then the following functionalities.

1. Assembling circuits for both postselection and direct sum-based discrimination schemes.
2. Executing the whole benchmarking scenario on specified backend (either real hardware or software simulator).
3. Interpreting the obtained outputs in terms of discrimination probabilities.

The PyQBench library also contains a readily available implementation of all necessary components needed to run discrimination experiments for parameterized Fourier family of measurements, defined previously in Chapter 6. If one only wishes to use this particular family of measurements in benchmarking process, then PyQBench offers a command line tool which might be more straightforward. PyQBench's command line interface (CLI) allows running the benchmarking process without writing Python code. The configuration of CLI is done by YAML [116] files describing the benchmark to be performed and the description of the backend on which the benchmark should be run. Notably, the YAML configuration files are reusable. The same benchmark can be used with different backends and vice versa.

## 7.3 Software Architecture

This section describes the most important architectural decisions taken when creating PyQBench, and how they affect the end-user experience.

### Overview of the software structure

As already described, PyQBench can be used both as a library and a CLI. Both functionalities are implemented as a part of `qbench` Python package. The exposed CLI tool is also named `qbench`. For brevity, we do not discuss the exact structure of the package here, and instead refer an interested reader to the source code available at GitHub [107] or at the reference manual [108].

PyQBench can be installed from official Python Package Index (PyPI) by running `pip install pyqbench`. In a properly configured Python environment the installation process should also make the `qbench` command available to the user without a need for further configuration.

## Integration with hardware providers and software simulators

PyQBench is built around the Qiskit [15] ecosystem. Hence, both the CLI tool and the `qbench` library can use any Qiskit-compatible backend. This includes, IBM Q backends (available by default in Qiskit) and Amazon Braket devices and simulators (available through `qiskit-braket-provider` package [117, 118]).

When using PyQBench as library, instances of Qiskit backends can be passed to functions that expect them as parameters. However, in CLI mode, the user has to provide a YAML file describing the backend. An example of using both approaches can be found in Section 7.3.2 and Section 7.3.1, and the detailed description is presented in PyQBench’s documentation.

### 7.3.1 PyQBench as a CLI tool

To introduce PyQBench as a CLI tool we use the parametrized family of Fourier measurements defined in Section 6.2.

#### Workflow

For the parametrized Fourier family of measurements, PyQBench offers a simplified way of conducting benchmarks using a command line interface (CLI). The workflow with PyQBench’s CLI can be summarized as the following list of steps:

1. Preparing configuration files describing the backend and the experiment scenario.
2. Submitting/running experiments. Depending on the experiment scenario, execution can be synchronous, or asynchronous.
3. (optional) Checking the status of the submitted jobs if the execution is asynchronous.
4. Resolving asynchronous jobs into the actual measurement outcomes.
5. Converting obtained measurement outcomes into tabulated form.

The general form of the CLI invocation is shown in the following Listing 7.1.

Listing 7.1: Invocation of `qbench` script

---

```
qbench <benchmark-type> <command> <parameters>
```

---

As we can see later, PyQBench’s CLI supports not only one type of benchmark – based on the discrimination between parametrized Fourier family of measurements.

For this purpose we decided on structuring the CLI in a hierarchical fashion to allow for future extensions. Here, the accepted value of `<benchmark-type>` is `disc-fourier`. The `qbench disc-fourier` command has four subcommands:

- **benchmark**: running benchmarks. This command creates either a result YAML file containing the measurements or an intermediate YAML file for asynchronous experiments.
- **status**: querying status of experiments submitted for given benchmark. This command is only valid for asynchronous experiments.
- **resolve**: querying the results of asynchronously submitted experiments and saving a result YAML file. The output of this command is almost identical to the result obtained from synchronous experiments.
- **tabulate**: interpreting the results of a benchmark and summarizing them as a table in the CSV file.

### **Asynchronous vs. synchronous execution**

PyQBench's CLI can be used in synchronous and asynchronous modes. The mode of execution is defined in the YAML file describing the backend. When running `qbench disc-fourier benchmark` in asynchronous mode, the PyQBench submits all the circuits needed to perform a benchmark and then writes an intermediate YAML file containing metadata of submitted experiments. The intermediate file can be used to query the status of the submitted jobs or to resolve them, i.e. to wait for their completion and get the measurement outcomes. In synchronous mode, PyQBench first submits all jobs required to run the benchmark and then immediately waits for their completion. The advantage of this approach is that no separate invocation of `qbench` command is needed to actually download the measurement outcomes. The downside, however, is that if the script is interrupted while the command is running, the intermediate results will be lost. Therefore, we recommend using asynchronous mode whenever possible.

### **Preparing configuration files – experiment file**

The configuration of PyQBench CLI is driven by YAML files. The first configuration file describes the experiment scenario to be executed. The second file describes the backend. Let us first describe the experiment configuration file, which might look as follow.

Listing 7.2: Defining the experiment file

---

```

type: discrimination-fourier
qubits:
  - target: 0
    ancilla: 1
  - target: 1
    ancilla: 2
  - target: 14
    ancilla: 16
angles:
  start: 0
  stop: 2 * pi
  num_steps: 32
gateset: ibmq
method: direct_sum
num_shots: 8192

```

---

The experiment file contains the following fields:

- **type**: a string describing the type of the experiment. Here, the option `type` is `discrimination-fourier`.
- **qubits**: a list enumerating pairs of qubits on which the experiment should be run. For experiment file defined in Listing 7.2, the benchmark will run on three pairs of qubits. The first pair consists of `target` 0 and `ancilla` 1, the second one is `target` 1 and `ancilla` 2, whereas the last pair is `target` 14 and `ancilla` 16. We describe a pair by using `target` and `ancilla` keys rather than using a plain list to emphasize that the role of qubits in the experiment is distinguishable.
- **angles**: an object describing the range of angles for Fourier parameterized family. The described range is always uniform, starts at the `start`, ends at `stop` and contains `num_steps` points, including both `start` and `stop`. The `start` and `stop` can be arithmetic expressions using `pi` literal. For instance, the experiment defined in Listing 7.2 contains 33 points:  $k \cdot \frac{\pi}{16}$ , where  $k = 0, \dots, 32$ .
- **gateset**: a string describing the set of gates used in the decomposition of circuits in the experiment. The PyQBench contains explicit implementations of circuits. The possible options are `[ibmq, lucy, rigetti]`, corresponding to decompositions compatible with IBM Q devices, OQC Lucy device, and Rigetti devices. Alternatively, one might wish to turn off the decomposition by using a special value `generic`.

- **method**: a string, either `postselection` or `direct_sum` determining which implementation of the conditional measurement is used.
- **num\_shots**: an integer defines how many shots are performed in the experiment for a particular angle, qubit pair and circuit. For direct sum we run  $2 \cdot \text{num\_steps} \cdot \text{num\_shots}$  circuits to achieved the empirical probability of correct discrimination. Note that the `postselection` method uses twice as many circuits as the `direct_sum` method.

## Preparing configuration files – backend file

The second configuration file describes the backend. Different Qiskit backends typically require different data for their initialization. Hence, there are multiple possible formats of the backend configuration files understood by PyQBench. We refer the interested reader to the PyQBench’s documentation [108].

Below we describe an example YAML file describing IBM Q backend named Kolkata using synchronous execution. Note, IBMQ backends typically require an access token to IBM Quantum Experience. Since it would be unsafe to store it in plain text, the token has to be configured separately in `IBMQ_TOKEN` environmental variable.

Listing 7.3: Defining IBMQ backend

---

```
name: ibmq_kolkata
asynchronous: false
provider:
  hub: ibm-q-psnc
  group: open
  project: main
```

---

Note that if we would like to run an experiment asynchronously, it is enough to fix the type `asynchronous` as `true`.

## Running the experiment and collecting measurements data

After preparing YAML files defining experiment and backend, running the benchmark can be launched by using the following command line invocation:

---

```
qbench disc-fourier benchmark experiment_file.yml backend_file.yml
```

---

The output file will be printed to stdout. Optionally, the `-output OUTPUT` parameter might be provided to write the output to the `OUTPUT` file instead.

---

```
qbench disc-fourier benchmark experiment_file.yml backend_file.yml
--output sync_results.yml
```

---

### (Optional) Getting status of asynchronous jobs

PyQBench provides also a helper command that will fetch the statuses of asynchronous jobs. The command is:

---

```
qbench disc-fourier status async_results.yml
```

---

and it will display dictionary with histogram of statuses.

### Resolving asynchronous jobs

If the status of jobs is `DONE` for asynchronous experiments, the stored intermediate data has to be resolved in actual measurements' outcomes. The following command will wait until all jobs are completed and then write a result file:

---

```
qbench disc-fourier resolve async-results.yml resolved.yml
```

---

The resolved results `resolved.yml` using asynchronous mode would look just like if the experiment was run synchronously.

### Tabulating results

As a last step in the processing workflow, no matter which method we choose (synchronous or asynchronous), the results file has to be passed to `tabulate` command:

---

```
qbench disc-fourier tabulate results.yml results.csv
```

---

A part of sample CSV file is provided in Table 7.1.

target	ancilla	phi	ideal_prob	disc_prob	mit_disc_prob
0	1	0	0.5	0.503	0.503
0	1	0.202	0.550	0.542	0.544
0	1	0.405	0.602	0.598	0.602
0	1	0.608	0.647	0.636	0.639
0	1	0.811	0.697	0.679	0.684
0	1	1.013	0.743	0.726	0.731
0	1	1.216	0.786	0.769	0.775
0	1	1.419	0.826	0.803	0.810
0	1	1.621	0.862	0.843	0.851
0	1	1.824	0.895	0.873	0.882

Table 7.1: The resulting CSV file contains table with columns `target`, `ancilla`, `phi`, `ideal_prob`, `disc_prob` and, optionally, `mit_disc_prob`. Each row in the table describes results for a tuple of `(target, ancilla, phi)`. The reference optimal value of discrimination probability is present in `ideal_prob` column, whereas the obtained, empirical discrimination probability can be found in the `disc_prob` column. The `mit_disc_prob` column contains empirical discrimination probability after applying the Mthree error mitigation [119, 120], if it was applied.

## Plotting results

For the experiment defined in Listing 7.2 and the backend defined in Listing 7.3, we also present graphical representations of the results in Fig. 7.5 and Fig. 7.6.

**Remark 3** *It is worth stressing why the benchmark using qubits (1, 2) returns poor results in comparison to other experiments. One may suppose a bug in code or implementation method. The benchmark, however, is compatible with error description provided by IBM Q vendor who indicates faulty implementation of CNOT gate. More precisely, the value CNOT error for pair (1, 2) which is equal to one in contrast to the CNOT error between qubits (0, 1) and (14, 16) fluctuates around  $10^{-3}$ .*



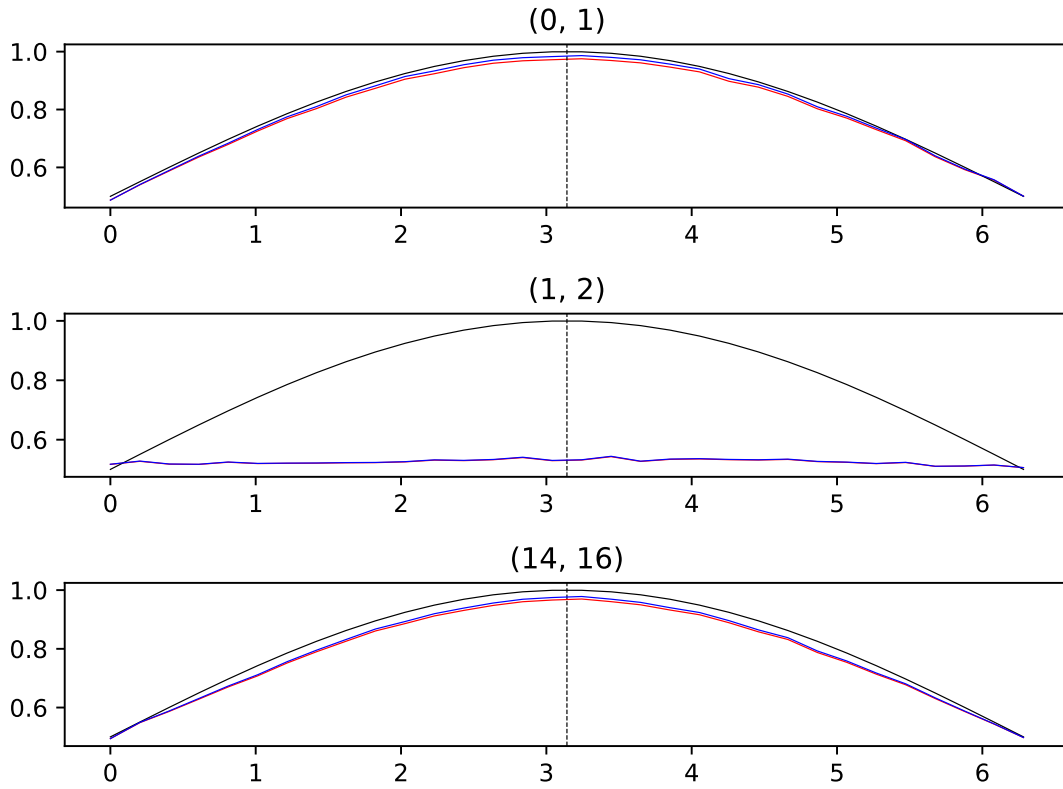


Figure 7.5: Discrimination experiment defined for the parameterized family of von Neumann measurement  $\mathcal{P}_U$  and  $\mathcal{P}_1$  using direct sum. We run experiment defined in Listing 7.2 on IBM Q device named Kolkata with 27 qubits using three pairs of qubits (0, 1), (1, 2) and (14, 16). The theoretical probability of correct discrimination is given by black line whereas the empirical probability is shown by the red line. The blue line represents the results after applying the Mthree error mitigation package.

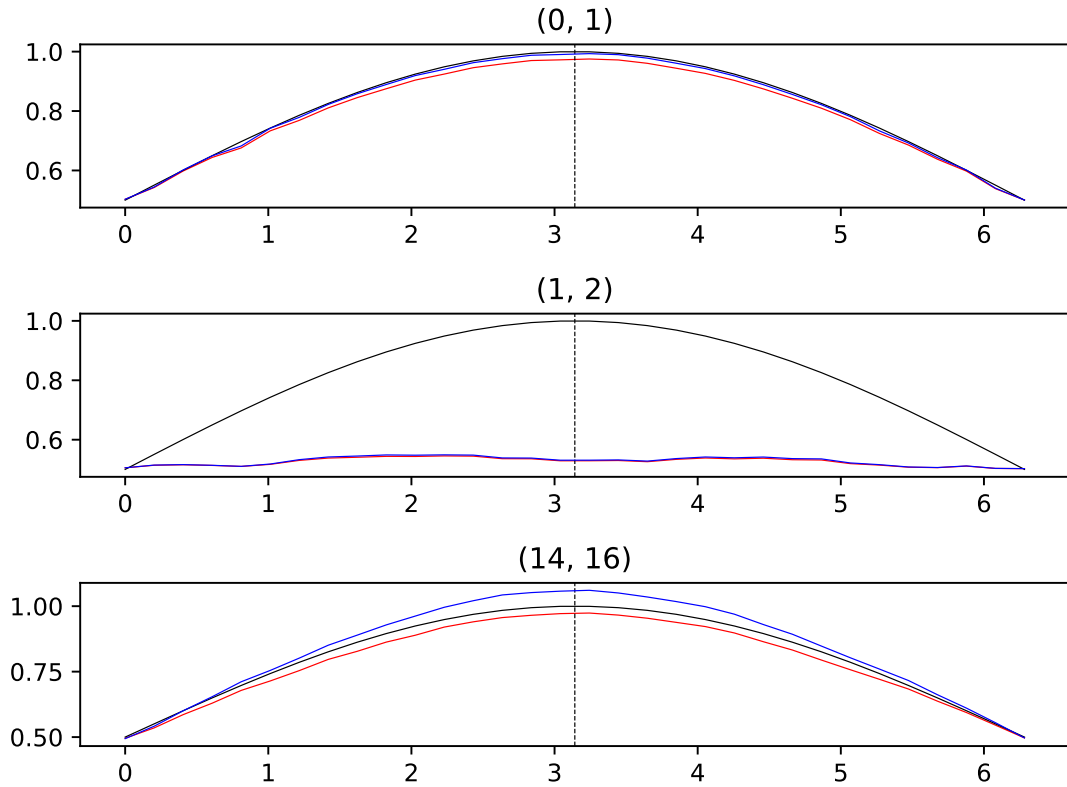


Figure 7.6: Discrimination experiment defined for the parameterized family of von Neumann measurement  $\mathcal{P}_U$  and  $\mathcal{P}_1$  using postselection. We run experiment defined in Listing 7.2 (with fixed `method` by `postselection`) on IBM Q device named Kolkata with 27 qubits using three pairs of qubits (0, 1), (1, 2) and (14, 16). The theoretical probability of correct discrimination is given by black line whereas the empirical probability is shown by the red line. The blue line represents the results after applying the Mthree error mitigation package.

### 7.3.2 PyQBench as a library

In the previous section we introduced benchmarks for specific family of measurements, but what if we want more control over some parts of this benchmark? Or what a user wants to prepare a benchmark for different von Neumann measurement? One possibility would be to add some additional parameters to `benchmark_using_xyz` functions, but this approach is not scalable. Moreover, anticipating all possible uses cases is impossible. Therefore, we decided on another approach. PyQBench provides functions performing:

1. Assembly of circuits needed for experiment, provided the components discussed above.
2. Interpretation of the obtained measurements.

The difference between the two approaches is illustrated on the diagrams in Fig. 7.7. The example of usage PyQBench as a Python library we can see in Appendix B.2.

## 7.4 Conclusion and discussion

In this section, we develop a Python library PyQBench [108], an innovative open-source framework for benchmarking gate-based quantum computers. Our package allows the user to test various architectures, available through `qiskit` and Amazon BraKet using problems with simple operational interpretation. PyQBench can benchmark NISQ devices by verifying their capability of discriminating between two qubit von Neumann measurements. PyQBench offers a simplified, ready-to-use, command line interface (CLI) for running benchmarks using a predefined parametrized Fourier family of measurements. Furthermore, we provide a powerful tool for the users to extend the range of available problems in a way that suits their needs. For more advanced scenarios, PyQBench offers a way of employing user-defined measurements instead of predefined ones. Due to this possibility of extension, the users are able to test specific aspects of their architecture of interest.

Finally, we provide the source code of PyQBench on GitHub [107] under an open source license which will allow users to utilize and extend our package in their specific applications.

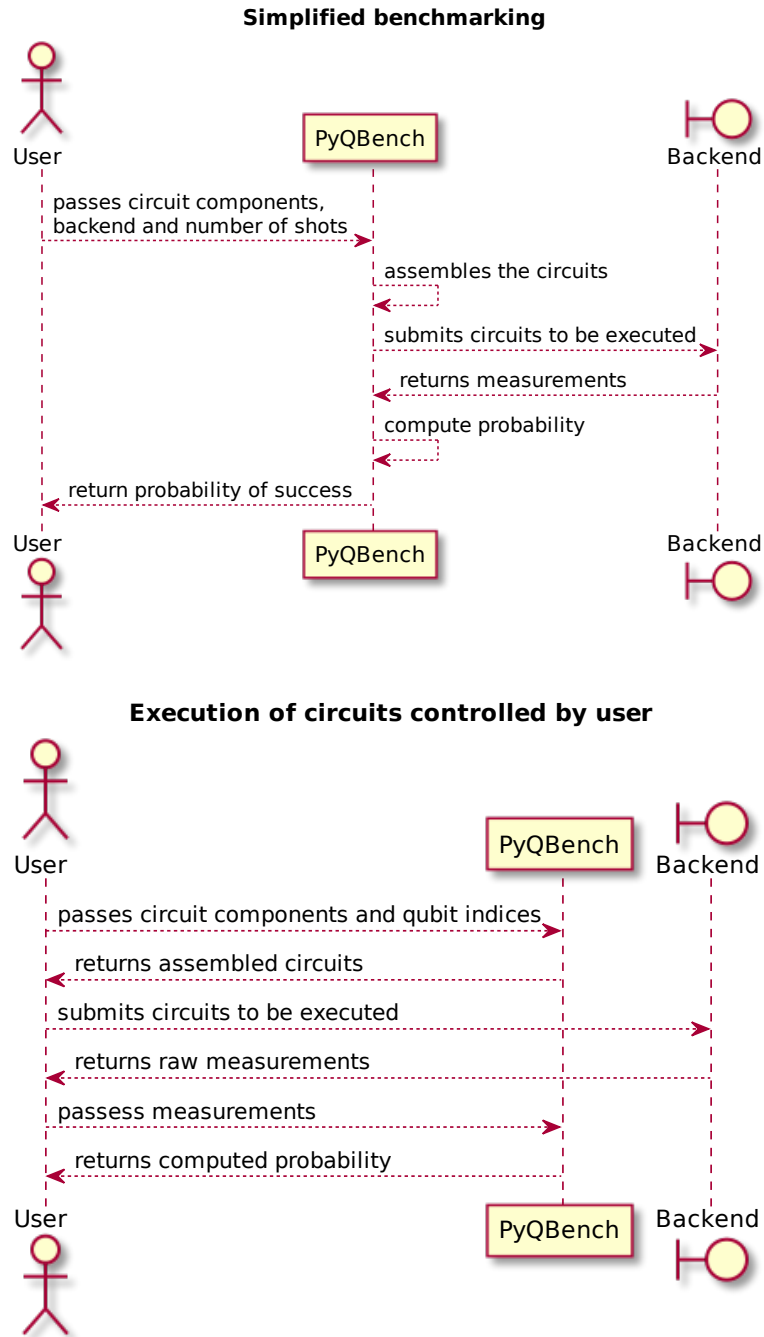


Figure 7.7: Differences between simplified (top) and user-controlled (bottom) execution of benchmarks in PyQBench. Compared to simplified benchmarking, in user-controlled benchmarks the user has direct access to the circuits being run, and hence can alter them (e.g. by adding noise) and/or choose the parameters used for executing them on the backend.

# Chapter 8

## Certification of von Neumann measurements

In this chapter we discuss another method of validation of quantum architectures, that is, certification of von Neumann measurements. To introduce the task of certification, let us imagine that Alice and Bob get a gift from Eve – a black-box with one of two known measurement device inside. The owner of the box, Eve, tells them which of the two possibilities is contained within the box. Yet, for some reason, they do not completely trust her and decide to perform some kind of hypothesis testing scheme on the black box. They decide to take Eve's promise as the null hypothesis,  $H_0$ , for this scheme and the second of the possibilities as the alternative hypothesis,  $H_1$ . Since now they own the box and are free to proceed as you want, they need to prepare some input into the box and perform a measurement on the output. A particular input state and final measurement will be called a certification strategy. Of course, just like in classical hypothesis testing, in the certification scheme there are two possible types of errors. The type I error happens if we reject the null hypothesis when it was in reality true. The type II error happens if we accept the null hypothesis when we should have rejected it. The main aim of certification task is finding the optimal strategy which minimizes type I error when the type II error is fixed.

Certification of quantum objects is closely related with the other well-know method of the validation, that is the problem of discrimination of those objects. Intuitively, in the discrimination problem we are given one of two quantum objects sampled according to a given a priori probability distribution (as we described in the previous section). Hence, the probability of making an error in the discrimination task is equal to the average of the type I and type II errors over the assumed probability distribution. In other words, the main difference between both approaches is that the main task of discrimination is the minimization over the average of both types of possible errors while the certification concerns the

minimization over one type of error when the bound of the other one is assumed. Therefore, this approach can be seen as a natural extension of the single-shot discrimination of von Neumann measurements [2].

The certification of quantum objects was first studied by Helstrom in [35], where the problem of pure state certification was considered. Further, certification schemes were established to the case of mixed states in [121]. A natural extension of quantum state certification is the certification of unitary operations. This problem was solved in [36]. For a more general overview of quantum certification we refer the reader to [38, 122].

This work will begin with recalling one-shot certification scenarios for quantum states and channels which will be used to constrain the certification setup for von Neumann measurements. For this purpose, we will often make use of the terms of numerical range and  $q$ -numerical range as essential tools in the proofs [54, 56, 57, 59]. More specifically, one of our results presented in this chapter is a geometric interpretation of the formula for minimized probability of the type II error in the problem of certification of unitary channels, which is strictly connected with the notion of  $q$ -numerical range. Later, basing on the results on the certification of unitary channels we will extend these considerations to the problem of certification of von Neumann measurements. It will show that the minimized probability of the type II error can also be connected with the notion of  $q$ -numerical range. On top of that, we will show that entanglement can significantly improve the certification of von Neumann measurements.

Based on the obtained results, we will create the algorithm realizing a protocol of optimal certification strategy. Due to the algorithm, we created an optimal certification strategy for the qubit parameterized family of Fourier measurements. For this case, we will also calculate the exact value of the probability of the type II error. Due to the obtained results, we will extend the platform PyQBench to another benchmark of NISQ devices based on the certification scheme. We will offer using PyQBench as a command line interface (CLI) for this family of measurements. However, PyQBench is constantly being developed and ultimately we would like to extend PyQBench based on the certification scheme as a Python library. The code is available in [123]. This part of chapter is my contribution to the dissertation.

This chapter is organized as follows. We begin with presentation a general aspect of certification in Section 8.2. Section 8.1 recalls about certification of quantum states, whereas Section 8.3 presents result for certification of unitary channels. In this section we also introduce geometrical interpretation of the problem of certification of unitary channels, expressed in terms of  $q$ -numerical range. Then, the certification of von Neumann measurements is studied in Section 8.4 and our main result is stated therein as Theorem 12. Based on the proof of this theorem, we create an protocol describing the optimal certification scheme. This protocol is

presented in Section 8.5. Next, in Section 8.6 we will discuss about certification scheme for parameterized Fourier family of qubit von Neumann measurements. In Section 5, we create a benchmark based on certification scheme using the previous defined parametrized Fourier family of qubit measurements whereas Section 8.8 presents the implementation of the certification scheme. Finally, concluding remarks are presented in the Section 8.9. In the Appendix C, we provide the technical calculations to determine a distance of  $q$ -numerical range to zero.

## 8.1 Certification of pure quantum states

The starting point towards the certification of quantum objects is the hypothesis testing of quantum states. Let  $H_0$  be a null hypothesis which states that the obtained state was  $|\psi\rangle \in \mathcal{X}$ , while the alternative hypothesis,  $H_1$ , states that the obtained state was  $|\varphi\rangle \in \mathcal{X}$ . The certification is performed by the use of a binary measurement  $\{\Omega, \mathbb{1}_{\mathcal{X}} - \Omega\}$ , where the effect  $\Omega$  corresponds to accepting the null hypothesis and  $\mathbb{1}_{\mathcal{X}} - \Omega$  accepts the alternative hypothesis. In this dissertation we will be considering only POVMs with two effects of this form. Therefore the effect  $\Omega$  uniquely determines the POVM and hence we will be using the words measurement and effect interchangeably.

Assume we have a fixed measurement  $\Omega$ . We introduce the probability of the type I error,  $p_{\text{I}}(\Omega)$ , that is the probability of rejecting the null hypothesis when in fact it was true, as

$$p_{\text{I}}(\Omega) = \langle \psi | (\mathbb{1}_{\mathcal{X}} - \Omega) | \psi \rangle = 1 - \langle \psi | \Omega | \psi \rangle. \quad (8.1)$$

The type II error,  $p_{\text{II}}(\Omega)$ , that is the probability of accepting the null hypothesis  $H_0$  when in reality  $H_1$  occurred, is defined as

$$p_{\text{II}}(\Omega) = \langle \varphi | \Omega | \varphi \rangle. \quad (8.2)$$

In the remainder of this work we will assume the statistical significance  $\delta \in [0, 1]$ , that is the probability of the type I error will be upper-bounded by  $\delta$ . Our goal will be to find a most powerful test, that is to minimize the probability of the type II error by finding the optimal measurement, which we will denote as  $\Omega_0$ . Such  $\Omega_0$ , which minimizes  $p_{\text{II}}(\Omega)$  while assuming the statistical significance  $\delta$ , will be called an optimal measurement. The minimized probability of type II error will be denoted by

$$p_{\text{II}} = \min_{\Omega: p_{\text{I}}(\Omega) \leq \delta} \langle \varphi | \Omega | \varphi \rangle =: \langle \varphi | \Omega_0 | \varphi \rangle, \quad (8.3)$$

The results of minimization of the probability of the type II error over measurements  $\Omega$  are summarized as the following theorem. A related study of this problem

can be found in [35].

**Theorem 10** *Consider the problem of two-point certification of unitary channels with hypotheses*

$$\begin{aligned} H_0 &: |\psi\rangle \\ H_1 &: |\phi\rangle. \end{aligned} \tag{8.4}$$

and statistical significance  $\delta \in [0, 1]$ . Then, for the most powerful test, the minimized probability of the type II error yields

$$p_{II} = \begin{cases} 0 & \text{if } |\langle\psi|\phi\rangle| \leq \sqrt{\delta}, \\ \left( |\langle\psi|\phi\rangle|\sqrt{1-\delta} - \sqrt{1-|\langle\psi|\phi\rangle|^2}\sqrt{\delta} \right)^2 & \text{if } |\langle\psi|\phi\rangle| > \sqrt{\delta}. \end{cases} \tag{8.5}$$

The proof of the above theorem is presented below. This proof gives us the method how to construct the optimal final measurement which minimizes the probability of the type II error. The exact form of the optimal measurement is stated as Corollary 8.

**Proof.** Without loss of generality we can assume that  $|\varphi\rangle = \alpha|\psi\rangle + \beta|\psi^\perp\rangle$ , for some  $\alpha, \beta \geq 0$  satisfying  $\alpha^2 + \beta^2 = 1$ . For any effect  $\tilde{\Omega}$  satisfying  $\langle\psi|\tilde{\Omega}|\psi\rangle \geq 1 - \delta$ , the effect  $\Omega$  defined as  $\Omega = \Pi\tilde{\Omega}\Pi$ , where  $\Pi = |\psi\rangle\langle\psi| + |\psi^\perp\rangle\langle\psi^\perp|$ , also satisfies the condition  $\langle\psi|\Omega|\psi\rangle \geq 1 - \delta$  and simultaneously returns the same value of probability of type II error. Hence, we can assume that rank-2 operator  $\Omega$  satisfies  $\Omega = \Pi\Omega\Pi$ . From the above, let  $\Omega = a\Pi + b|\omega\rangle\langle\omega|$ , where  $|\omega\rangle = c|\psi\rangle - d|\psi^\perp\rangle$ ,  $c \geq 0$ ,  $d \in \mathbb{C}$ , such that  $c^2 + |d|^2 = 1$  and  $a, b \in [0, 1]$ , such that  $a + b \leq 1$ . By the assumption on the value  $p_I$ , we have

$$1 - p_I(\Omega) = \langle\psi|\Omega|\psi\rangle = a + bc^2 \geq 1 - \delta. \tag{8.6}$$

Let us calculate the probability  $p_{II}$ :

$$p_{II} = \min_{\Omega: p_I(\Omega) \leq \delta} \langle\varphi|\Omega|\varphi\rangle = \min_{a, b, c, d \in \mathbf{A}} \left( \alpha^2(a + bc^2) + \beta^2(a + b|d|^2) - 2\alpha\beta bc\Re(d) \right) \tag{8.7}$$

where

$$\mathbf{A} := \{a, b, c, d : a + b \leq 1, a + bc^2 \geq 1 - \delta, c^2 + |d|^2 = 1, a, b, c \in [0, 1], d \in \mathbb{C}\}. \tag{8.8}$$

Note that the above formula is minimized when  $d \in \mathbb{R}$  is nonnegative. Hence, we have

$$\langle\varphi|\Omega|\varphi\rangle = a + b(\alpha c - \beta d)^2. \tag{8.9}$$



Thus, our task reduces to minimizing the formula

$$p_{\text{II}} = \min_{a,b,c \in \mathbf{B}} a + b \left( \alpha c - \beta \sqrt{1 - c^2} \right)^2 \quad (8.10)$$

where  $\mathbf{B} := \{a, b, c \in [0, 1], a + b \leq 1, a + bc^2 \geq 1 - \delta\}$ . We consider two cases.

1. If  $\alpha \leq \sqrt{\delta}$ , then we take  $a = 0, b = 1, c = \beta, d = \sqrt{1 - \beta^2}$ . In this case  $a, b, c \in \mathbf{B}$  and we obtain  $p_{\text{II}} = 0$ . The optimal strategy is represented by effect  $\Omega_0 = |\omega\rangle\langle\omega|$ , where  $|\omega\rangle = \beta|\psi\rangle - \alpha|\psi^\perp\rangle$ .
2. Let  $\alpha > \sqrt{\delta}$  and take  $a = 0, b = 1, c = \sqrt{1 - \delta}, d = \sqrt{\delta}$ . Again  $a, b, c \in \mathbf{B}$  and  $p_{\text{II}} = \left( \alpha\sqrt{1 - \delta} - \beta\sqrt{\delta} \right)^2$ . The optimal strategy is represented by effect  $\Omega_0 = |\omega\rangle\langle\omega|$  where  $|\omega\rangle = \sqrt{1 - \delta}|\psi\rangle - \sqrt{\delta}|\psi^\perp\rangle$ . The optimality of this value can be checked by using standard constrained optimization techniques. ■

**Corollary 8** *The optimal strategy for two-point certification of pure quantum states  $|\psi\rangle$  and  $|\varphi\rangle$ , with statistical significance  $\delta$  yields*

1. if  $|\langle\psi|\varphi\rangle| \leq \sqrt{\delta}$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$ , where  $|\omega\rangle = \frac{|\tilde{\omega}\rangle}{\|\tilde{\omega}\|}$ ,  $|\tilde{\omega}\rangle = |\psi\rangle - \langle\varphi|\psi\rangle|\varphi\rangle$ ;
2. if  $|\langle\psi|\varphi\rangle| > \sqrt{\delta}$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$  for  $|\omega\rangle = \sqrt{1 - \delta}|\psi\rangle - \sqrt{\delta}|\psi^\perp\rangle$ ,  $|\psi^\perp\rangle = \frac{|\tilde{\psi}^\perp\rangle}{\|\tilde{\psi}^\perp\|}$ , where  $|\tilde{\psi}^\perp\rangle = |\varphi\rangle - \langle\psi|\varphi\rangle|\psi\rangle$ .

## 8.2 General concept of certification for quantum channels and measurements

While certifying quantum channels or von Neumann measurements, we will also need to minimize over input states not only over the measurement  $\Omega$ . In the most general case, we will be allowed to use entanglement by adding an additional system. Hence, a quantum channel  $\Phi_0 \otimes \mathcal{I}_{\mathcal{X}}$  will be corresponded to hypothesis  $H_0$  and  $\Phi_1 \otimes \mathcal{I}_{\mathcal{X}}$  will be corresponded to hypothesis  $H_1$ . For a fixed state  $|\psi\rangle$  and  $\Omega$ , we define the conditional probability of type I error  $p_{\text{I}}^{|\psi\rangle}(\Omega)$  and type II error  $p_{\text{II}}^{|\psi\rangle}(\Omega)$  by

$$\begin{aligned} p_{\text{I}}^{|\psi\rangle}(\Omega) &= \text{tr}((\mathbb{1}_{\mathcal{X}} - \Omega)(\Phi_0 \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)) \\ p_{\text{II}}^{|\psi\rangle}(\Omega) &= \text{tr}(\Omega(\Phi_1 \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)). \end{aligned} \quad (8.11)$$

Naturally, for each input state we can consider minimized probability of type II error, that is

$$p_{\text{II}}^{|\psi\rangle} = \min_{\Omega: p_{\text{I}}^{|\psi\rangle}(\Omega) \leq \delta} p_{\text{II}}^{|\psi\rangle}(\Omega). \quad (8.12)$$

Finally, we will be interested in calculating optimized probability of type II error over all input states. This will be denoted as

$$p_{\text{II}} := \min_{|\psi\rangle} p_{\text{II}}^{|\psi\rangle}. \quad (8.13)$$

Note that the symbol  $p_{\text{II}}$  is used in two contexts. In the problem of certification of states the minimization is performed only over measurements  $\Omega$ , while in the problem of certification of unitary channels and von Neumann measurements the minimization is over both measurements  $\Omega$  and input states  $|\psi\rangle$ . In other words,  $p_{\text{II}}$  is equal to the optimized probability of the type II error in certain certification problem. The input state which minimizes  $p_{\text{II}}$  will be called an optimal state. We will use the term optimal strategy to denote both the optimal state and the optimal measurement.

### 8.3 Certification of unitary channels

In this section we will be interested in certification of two unitary channels  $\Phi_U$  and  $\Phi_{\mathbf{1}}$ . We will allow to use entanglement by adding an additional system. Hence, the null hypothesis  $H_0$  yields that the unknown channel is  $\Phi_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}}$  and the alternative hypothesis  $H_1$  yields that the unknown channel is  $\Phi_U \otimes \mathcal{I}_{\mathcal{X}}$ .

The idea behind the scheme of certification of unitary channels is to reduce this problem to certification of quantum states discussed in the previous section by preparing some (possibly entangled) input state  $|\psi\rangle$  and performing the unknown channel on it. The resulting state is either  $|\psi\rangle$  or  $(U \otimes \mathbb{1}_{\mathcal{X}})|\psi\rangle$ . Then, we perform the measurement  $\{\Omega, \mathbb{1}_{\mathcal{X}} - \Omega\}$  and make a decision whether the given channel was  $\Phi_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}}$  or  $\Phi_U \otimes \mathcal{I}_{\mathcal{X}}$ . The effect  $\Omega$  corresponds to accepting  $H_0$  hypothesis while  $\mathbb{1}_{\mathcal{X}} - \Omega$  corresponds to the alternative hypothesis  $H_1$ .

The results of minimization of the probability of the type II error over input states  $|\psi\rangle$  and measurements  $\Omega$  are summarized as the following theorem. This reasoning is based on the results from Theorem 10, while a related study of this problem can be found in [36], we connected the obtained results with  $q$ -numerical range.

**Theorem 11** *Consider the problem of two-point certification of unitary channels*

with hypotheses

$$\begin{aligned} H_0 &: \Phi_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}}, \\ H_1 &: \Phi_U \otimes \mathcal{I}_{\mathcal{X}}, \end{aligned} \quad (8.14)$$

and statistical significance  $\delta \in [0, 1]$ . Then, for the most powerful test, the probability of the type II error yields

$$p_{II} = \nu_{\sqrt{1-\delta}}^2(U). \quad (8.15)$$

**Proof.** Let us assume the optimal state  $|\psi\rangle$  and optimal measurement  $\Omega = |\omega\rangle\langle\omega|$ . We bound the probability of the type I error by  $\delta$ , that is  $p_1^{|\psi\rangle}(\Omega) = \text{tr}((\mathbb{1} - \Omega)|\psi\rangle\langle\psi|) \leq \delta$ . Hence, we have

$$\text{tr}(\Omega|\psi\rangle\langle\psi|) = |\langle\omega|\psi\rangle|^2 \geq 1 - \delta. \quad (8.16)$$

The probability of the type II error takes the form

$$\begin{aligned} p_{II} &= \min_{|\psi\rangle} \min_{\Omega: p_1^{|\psi\rangle}(\Omega) \leq \delta} \text{tr}(\Omega(U \otimes \mathbb{1})|\psi\rangle\langle\psi|(U^\dagger \otimes \mathbb{1})) \\ &= \min_{|\psi\rangle} \min_{|\omega\rangle: p_1^{|\psi\rangle}(|\omega\rangle\langle\omega|) \leq \delta} \langle\psi|(U^\dagger \otimes \mathbb{1})|\omega\rangle\langle\omega|(U \otimes \mathbb{1})|\psi\rangle \\ &= \min_{|\psi\rangle} \min_{|\omega\rangle: p_1^{|\psi\rangle}(|\omega\rangle\langle\omega|) \leq \delta} |\langle\psi|(U \otimes \mathbb{1})|\omega\rangle|^2. \end{aligned} \quad (8.17)$$

Let us recall that the  $q$ -numerical range is defined as

$$W_q(A) = \{\langle\xi_0|A|\xi_1\rangle : \langle\xi_0|\xi_1\rangle = q\}, \quad (8.18)$$

and we use the notation

$$\nu_q(X) = \min\{|x| : x \in W_q(X)\}. \quad (8.19)$$

Now from the definition of the  $q$ -numerical range for  $q = \sqrt{1-\delta}$  and its properties (see Subsection 2.2.4) we have

$$p_{II} = \nu_{\sqrt{1-\delta}}^2(U \otimes \mathbb{1}) = \nu_{\sqrt{1-\delta}}^2(U). \quad (8.20)$$

■

## Geometrical interpretation

Let  $\Theta$  be the angle between two most distant eigenvalues of a unitary matrix  $U$ . Then, from the above discussion we can draw a conclusion that for any statistical

significance  $\delta \in (0, 1]$ , if  $2 \arccos(\sqrt{\delta}) \leq \Theta < \pi$ , then although  $\Phi_U$  and  $\Phi_{\mathbf{1}}$  cannot be distinguished perfectly, they can be certified with  $p_{\text{II}} = 0$ . In other words, the numerical range  $W(U)$  does not contain zero but  $\sqrt{1-\delta}$ -numerical range,  $W_{\sqrt{1-\delta}}(U)$ , does contain zero. The situation changes when  $2 \arccos(\sqrt{\delta}) > \Theta$ . Then, both numerical range  $W(U)$  and  $\sqrt{1-\delta}$ -numerical range  $W_{\sqrt{1-\delta}}(U)$  do not contain zero. This situation is presented in Fig. 8.1 whereas the detailed calculation we can see in Appendix C.

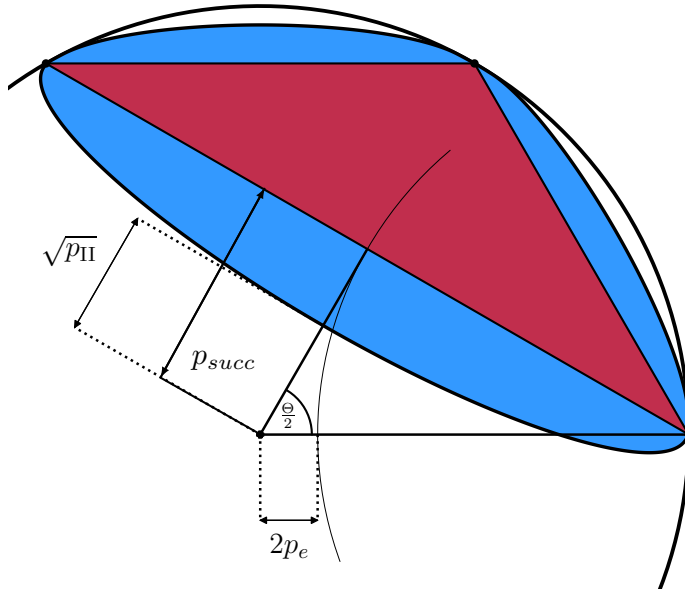


Figure 8.1: Numerical range  $W(U)$  (red triangle) and  $\sqrt{1-\delta}$ -numerical range  $W_{\sqrt{1-\delta}}(U)$  (blue oval) of  $U \in \text{U}(\mathbb{C}^3)$  with eigenvalues  $1, e^{\frac{\pi i}{3}}$  and  $e^{\frac{2\pi i}{3}}$  with statistical significance  $\delta = 0.05$ . The value  $p_{\text{succ}}$  denotes the probability of correct discrimination between unitary channels  $\Phi_U$  and  $\Phi_{\mathbf{1}}$  (see Subsection 6.1.1), whereas  $p_e$  is the probability of incorrect discrimination, that is  $p_e = 1 - p_{\text{succ}}$ . By  $p_{\text{II}}$  we denote the probability of the type II error for the task described in Theorem 11.

## 8.4 Certification of von Neumann measurements

Similar to the task of certification of unitary channels, we are also allowed to use entanglement hence, the null hypothesis  $H_0$  yields  $\mathcal{P}_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}}$  and the alternative hypothesis  $H_1$  is  $\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}}$ , where  $U \in \text{U}(\mathcal{X})$ . Similarly, like for unitary channels, the certification scheme is performed by finding the optimal input state  $|\psi\rangle$  and a binary measurement  $\{\Omega, \mathbb{1}_{\mathcal{X}} - \Omega\}$ , where the effect  $\Omega$  corresponds to accepting

the null hypothesis and  $\mathbb{1}_{\mathcal{X}} - \Omega$  accepts the alternative hypothesis. If we assume a fixed input state  $|\psi\rangle$  and measurement  $\Omega$ , we introduce the probability of the type I error,  $p_{\text{I}}(|\psi\rangle, \Omega)$ , that is the probability of rejecting the null hypothesis when in fact it was true, as

$$p_{\text{I}}(|\psi\rangle, \Omega) = \text{tr}((\mathbb{1}_{\mathcal{X}} - \Omega)(\mathcal{P}_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)). \quad (8.21)$$

The type II error,  $p_{\text{II}}(|\psi\rangle, \Omega)$ , is the probability of accepting the null hypothesis  $H_0$  when in reality  $H_1$  occurred, defined as

$$p_{\text{II}}(|\psi\rangle, \Omega) = \text{tr}(\Omega(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)). \quad (8.22)$$

Finally, the minimized probability of type II error will be denoted by

$$p_{\text{II}} := \min_{|\psi\rangle, \Omega: p_{\text{I}}(|\psi\rangle, \Omega) \leq \delta} p_{\text{II}}(|\psi\rangle, \Omega). \quad (8.23)$$

Again, our goal is to find an optimal input state  $|\psi\rangle$  and a final binary measurement for which the probability of the type II error is saturated, while the statistical significance  $\delta \in [0, 1]$  is assumed, that is the probability of the type I error will be upper-bounded by  $\delta$ .

**Theorem 12** *Consider the problem of two-point certification of von Neumann measurements with hypotheses*

$$\begin{aligned} H_0 &: \mathcal{P}_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}} \\ H_1 &: \mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}}. \end{aligned} \quad (8.24)$$

and statistical significance  $\delta \in [0, 1]$ . Then, for the most powerful test, the probability of the type II error yields

$$p_{\text{II}} = \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE), \quad (8.25)$$

where  $\nu_q(X)$  denotes the distance of  $q$ -numerical range of  $X$  to zero.

**Remark 4** *It is worth mentioning that we do not make any assumptions on the dimension of the auxiliary system, however its dimension is obviously upper-bounded by the dimension of the input states. Additionally, the dimension of the auxiliary system can be reduced to the Schmidt rank of the input state  $|\psi\rangle$  [ [2] Proposition 4]. It is worth mentioning here that in the certification of von Neumann measurements entanglement can significantly improve the outcome of the protocol while in the case of unitary channel certification it provides no benefit.*

The proof is divided into two parts. In the first part we will utilize data processing inequality presented in Lemma 5. Thanks to that, we will show the lower bound for  $p_{\text{II}}$ . In the second part we will utilize the results from the previous chapter about discrimination of von Neumann measurements to show the upper bound for  $p_{\text{II}}$ . Before we start, let us recall the data processing inequality.

**Lemma 5** (*Data processing inequality*) *Let  $\delta > 0$  and  $\Omega \in \text{Pos}(\mathcal{X})$  be a positive semidefinite operator such that  $\Omega \leq \mathbb{1}_{\mathcal{X}}$ . For any quantum channel  $\mathcal{C} \in \mathcal{C}(\mathcal{X})$  and quantum states  $\rho, \sigma \in \mathcal{D}(\mathcal{X})$  the following holds*

$$\min_{\Omega: \text{tr}(\Omega\rho) \geq 1-\delta} \text{tr}(\Omega\sigma) \leq \min_{\Omega: \text{tr}(\Omega\mathcal{C}(\rho)) \geq 1-\delta} \text{tr}(\Omega\mathcal{C}(\sigma)). \quad (8.26)$$

**Proof.** This inequality, along with its proof, can be found eg. in [124]. However, to keep this dissertation self-consistent we present our modified version of them. Let us consider two-point certification of two quantum states  $\rho$  and  $\sigma$  with statistical significance  $\delta \in [0, 1]$ . To calculate the probability of the type II error,  $p_{\text{II}}$ , we formulate the problem as

$$\min_{\Omega: \text{tr}(\Omega\rho) \geq 1-\delta} \text{tr}(\Omega\sigma). \quad (8.27)$$

Now, consider the scenario in which we use as processing the quantum channel  $\mathcal{C}$  on states  $\rho$  and  $\sigma$ . We want to calculate

$$\min_{\Omega: \text{tr}(\Omega\mathcal{C}(\rho)) \geq 1-\delta} \text{tr}(\Omega\mathcal{C}(\sigma)) \quad (8.28)$$

which is equivalent to

$$\min_{\Omega: \text{tr}(\mathcal{C}^\dagger(\Omega)\rho) \geq 1-\delta} \text{tr}(\mathcal{C}^\dagger(\Omega)\sigma). \quad (8.29)$$

Observe that  $\mathcal{C}^\dagger(\Omega)$  is also a measurement and

$$\{\mathcal{C}^\dagger(\Omega) : \text{tr}(\mathcal{C}^\dagger(\Omega)\rho) \geq 1 - \delta\} \subseteq \{\Omega : \text{tr}(\Omega\rho) \geq 1 - \delta\}. \quad (8.30)$$

Finally, we obtain the data processing inequality given by

$$\min_{\Omega: \text{tr}(\Omega\rho) \geq 1-\delta} \text{tr}(\Omega\sigma) \leq \min_{\Omega: \text{tr}(\Omega\mathcal{C}(\rho)) \geq 1-\delta} \text{tr}(\Omega\mathcal{C}(\sigma)). \quad (8.31)$$

■

**Proof of Theorem 12.** Recall, in the scheme of certification of von Neumann measurements the optimized probability of type II error can be expressed as

$$p_{\text{II}} := \min_{|\psi\rangle} \min_{\Omega: p_1^{|\psi\rangle}(\Omega) \leq \delta} \text{tr}(\Omega(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)). \quad (8.32)$$

Our goal is to prove that

$$p_{\text{II}} = \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE). \quad (8.33)$$

### The lower bound

This part of the proof mostly will be based on data processing inequality.

To show that

$$p_{\text{II}} \geq \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE), \quad (8.34)$$

let us recall that every quantum von Neumann measurement  $\mathcal{P}_U$  can be rewritten as  $\Delta \circ \Phi_{(UE)^\dagger}$ , where  $\Delta$  denotes the completely dephasing channel and  $E \in \text{DU}(\mathcal{X})$ . Therefore, utilizing data processing inequality in Lemma 5, along with the certification scheme of unitary channels in Theorem 11, the minimized probability of the type II error is lower-bounded by

$$p_{\text{II}} \geq \min_{|\psi\rangle, \Omega: p_{\text{I}}(|\psi\rangle, \Omega) \leq \delta} \text{tr}(\Omega(\Phi_{(UE)^\dagger} \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)) = \nu_{\sqrt{1-\delta}}^2((UE)^\dagger) = \nu_{\sqrt{1-\delta}}^2(UE) \quad (8.35)$$

which holds for each  $E \in \text{DU}(\mathcal{X})$ . Hence, maximizing the value of  $\nu_{\sqrt{1-\delta}}^2(UE)$  over  $E \in \text{DU}(\mathcal{X})$  leads to the lower bound of the form

$$p_{\text{II}} \geq \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE). \quad (8.36)$$

### The upper bound

Now we proceed to proving the upper bound. The proof of the inequality

$$p_{\text{II}} \leq \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE), \quad (8.37)$$

will be divided into two cases depending on diamond norm distance between considered measurements  $\mathcal{P}_U$  and  $\mathcal{P}_{\mathbf{1}}$ . In either case we will construct a strategy, that is choose a state  $|\psi_0\rangle$  and a measurement  $\Omega_0$ . As for every choice of  $|\psi\rangle$  and  $\Omega$  it holds that

$$p_{\text{II}} \leq \text{tr}(\Omega(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)), \quad (8.38)$$

we will show that for some fixed  $|\psi_0\rangle$  and  $\Omega_0$  it holds that

$$\text{tr}(\Omega_0(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)) = \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE). \quad (8.39)$$

First we focus on the case when  $\|\mathcal{P}_U - \mathcal{P}_{\mathbf{1}}\|_{\diamond} = 2$ . We take the discriminator

$|\psi_0\rangle$  for which it holds that

$$\|\mathcal{P}_U - \mathcal{P}_1\|_\diamond = \|((\mathcal{P}_U - \mathcal{P}_1) \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)\|_1. \quad (8.40)$$

Then, the output states  $(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)$  and  $(\mathcal{P}_1 \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)$  are orthogonal and by taking the measurement  $\Omega_0$  as the projection onto the support of  $(\mathcal{P}_1 \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)$  we obtain

$$\text{tr}(\Omega_0(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)) = 0. \quad (8.41)$$

Let us recall that

$$\|\Phi_U - \Phi_1\|_\diamond = 2\sqrt{1 - \nu^2(U)}, \quad (8.42)$$

where  $\nu(U) = \min\{|x| : x \in W(U)\}$  and it holds that [2]

$$\|\mathcal{P}_U - \mathcal{P}_1\|_\diamond = \min_{E \in \text{DU}(\mathcal{X})} \|\Phi_{UE} - \Phi_1\|_\diamond. \quad (8.43)$$

Then, utilizing Eq. (8.42) and (8.43) we obtain that

$$\max_{E \in \text{DU}(\mathcal{X})} \nu^2(UE) = 0. \quad (8.44)$$

Therefore, by the property that  $0 \in W_{\sqrt{1-\delta}}(UE)$  whenever  $0 \in W(UE)$  (see property Eq. (2.56)), we have that

$$\max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE) = 0. \quad (8.45)$$

Secondly, we consider the situation when  $\|\mathcal{P}_U - \mathcal{P}_1\|_\diamond < 2$ . Let

$$E_0 \in \arg \max_{E \in \text{DU}(\mathcal{X})} \nu(UE). \quad (8.46)$$

Again, by referring to Eq. (8.43) and (8.42) we obtain that  $\nu(UE_0) > 0$ . Without loss of generality assume that  $\dim(\mathcal{X}) = d$ . Let  $\lambda_0, \lambda_{d-1}$  be a pair of the most distant eigenvalues of  $UE_0$ . Note that the following relation holds

$$\nu(UE_0) = \frac{|\lambda_0 + \lambda_{d-1}|}{2}. \quad (8.47)$$

As the assumptions of Lemma 3 (see Chapter 6) are saturated for the defined  $E_0$ , we can consider the input state

$$|\psi_0\rangle = \sum_{i=0}^{d-1} \sqrt{\rho_0} |i\rangle \otimes |i\rangle, \quad (8.48)$$



where the existence of  $\rho_0$  together with its properties are described in Lemma 3 and Proposition 3. Let us define sets

$$\mathbf{C}_i := \left\{ \Omega : 0 \leq \Omega \leq \mathbb{1}, \operatorname{tr} \left( (\mathbb{1} - \Omega) \frac{\sqrt{\rho_0}|i\rangle\langle i|\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle} \right) \leq \delta \right\}, \quad (8.49)$$

for each  $i$  such that  $\langle i|\rho|i\rangle \neq 0$ . Now we take the measurement  $\Omega_0$  as

$$\Omega_0 = \sum_{i=0}^{d-1} |i\rangle\langle i| \otimes \Omega_i^\top, \quad (8.50)$$

where  $\Omega_i \in \mathbf{C}_i$  is defined as

$$\Omega_i \in \arg \min_{\tilde{\Omega} \in \mathbf{C}_i} \operatorname{tr} \left( \tilde{\Omega} \frac{\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle} \right), \quad (8.51)$$

for each  $i \in \{0, \dots, d-1\}$  such that  $\langle i|\rho_0|i\rangle \neq 0$  and  $\Omega_i = 0$  otherwise.

Now we check that the statistical significance is satisfied, that is for the described strategy we have

$$p_{\text{I}}(|\psi_0\rangle, \Omega_0) = 1 - \operatorname{tr}(\Omega_0(\mathcal{P}_{\mathbf{1}} \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)) = 1 - \sum_{i=0}^{d-1} \operatorname{tr}(\Omega_i\sqrt{\rho_0}|i\rangle\langle i|\sqrt{\rho_0}) \leq \delta. \quad (8.52)$$

Hence, it remains to show that for this setting

$$\operatorname{tr}(\Omega_0(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)) = \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE). \quad (8.53)$$

Direct calculations reveal that

$$\begin{aligned} \operatorname{tr}(\Omega_0(\mathcal{P}_U \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)) &= \sum_{i=0}^{d-1} \operatorname{tr}(\Omega_i\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}) \\ &= \sum_{i=0}^{d-1} \langle i|\rho_0|i\rangle \operatorname{tr} \left( \Omega_i \frac{\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle} \right). \end{aligned} \quad (8.54)$$

Let us define

$$p_{\text{II}}^i = \operatorname{tr} \left( \Omega_i \frac{\sqrt{\rho_0}U|i\rangle\langle i|U^\dagger\sqrt{\rho_0}}{\langle i|\rho_0|i\rangle} \right). \quad (8.55)$$

Note that due to Proposition 3, the absolute value of the inner product between pure states  $\frac{\sqrt{\rho_0}|i\rangle}{\|\sqrt{\rho_0}|i\rangle\|}$  and  $\frac{\sqrt{\rho_0}U|i\rangle}{\|\sqrt{\rho_0}U|i\rangle\|}$  is the same for every  $i \in \{0, \dots, d-1\} : \langle i|\rho|i\rangle \neq 0$ . Therefore we can consider the certification of pure states conditioned on the obtained

label  $i$  with statistical significance  $\delta$ . From Theorem 10 we know that  $p_{\text{II}}^i$  depends only on such an inner product between the certified states, hence  $p_{\text{II}}^i = p_{\text{II}}^j$  for each  $i, j : \langle i|\rho|i\rangle, \langle j|\rho|j\rangle \neq 0$ . Therefore, we have that the value of  $p_{\text{II}}^i$  will depend on  $\left|\frac{\lambda_1 + \lambda_d}{2}\right|$ . Thus, without loss of generality, we can assume that  $p_{\text{II}}^1 \neq 0$  and hence

$$\sum_{i=0}^{d-1} \langle i|\rho_0|i\rangle p_{\text{II}}^i = p_{\text{II}}^1 = \text{tr} \left( \Omega_1 \frac{\sqrt{\rho_0} U |1\rangle \langle 1| U^\dagger \sqrt{\rho_0}}{\langle 1|\rho_0|1\rangle} \right) \quad (8.56)$$

and in the remaining of the proof we will show that

$$p_{\text{II}}^1 = \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE). \quad (8.57)$$

It is sufficient to study two cases depending on the relation between  $\sqrt{\delta}$  and the inner product

$$\left| \frac{\langle 1|\rho_0 U|1\rangle}{\langle 1|\rho_0|1\rangle} \right| = \left| \frac{\lambda_1 + \lambda_d}{2} \right|. \quad (8.58)$$

In the case when  $\left|\frac{\lambda_1 + \lambda_d}{2}\right| \leq \sqrt{\delta}$ , then due to Theorem 10 we get  $p_{\text{II}}^1 = 0$ . On the other hand, we know that  $0 \in W_{\sqrt{1-\delta}}(UE_0)$ , and hence

$$\max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE) = 0. \quad (8.59)$$

In the case when  $\left|\frac{\lambda_1 + \lambda_d}{2}\right| > \sqrt{\delta}$ , then from Theorem 10 we know that

$$p_{\text{II}}^1 = \left( \left| \frac{\lambda_1 + \lambda_d}{2} \right| \sqrt{1-\delta} - \sqrt{1 - \left| \frac{\lambda_1 + \lambda_d}{2} \right|^2 \delta} \right)^2. \quad (8.60)$$

On the other hand, for  $E_0 \in \text{DU}(\mathcal{X})$  satisfying Eq. (8.46) we have

$$\nu_{\sqrt{1-\delta}}^2(UE_0) = \left( \left| \frac{\lambda_1 + \lambda_d}{2} \right| \sqrt{1-\delta} - \sqrt{1 - \left| \frac{\lambda_1 + \lambda_d}{2} \right|^2 \delta} \right)^2. \quad (8.61)$$

By the particular choice of  $E_0 \in \text{DU}(\mathcal{X})$ , this value is equal to

$$\max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE), \quad (8.62)$$

hence combining the above equations we finally obtain

$$p_{\text{II}}^1 = \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE). \quad (8.63)$$

To sum up, we indicated strategies  $\Omega_0$  and  $|\psi_0\rangle$  for which the optimized probability of type II error was equal to  $\max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE)$ . Combining this with the previously proven inequality

$$p_{\text{II}} \geq \max_{E \in \text{DU}(\mathcal{X})} \nu_{\sqrt{1-\delta}}^2(UE) \quad (8.64)$$

gives us Eq. (8.33) and proves that the proposed strategy  $|\psi_0\rangle, \Omega_0$  is optimal. ■

## 8.5 Optimal certification protocol

In the previous section, we have considered measurements certification scheme. We have certified between two von Neumann measurements,  $\mathcal{P}_U$  and  $\mathcal{P}_1$ . We have calculated the probability of type II error for the most powerful test. Due to the proof of Theorem 12, we introduce the protocol Algorithm 1 which describes the optimal certification strategy of von Neumann measurements.

## 8.6 Certification scheme for parameterized family of Fourier measurements

Based on the certification protocol presented in Algorithm 1, we have created the optimal certification strategy for parameterized family of Fourier measurements.

Let us consider the task of two-point certification between single-qubit von Neumann measurements  $\mathcal{P}_1$  and  $\mathcal{P}_U$ , where

$$U = H \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} H^\dagger \quad (8.65)$$

where  $H$  is the Hadamard matrix of dimension two and  $\phi \in [0, 2\pi)$ . We formulate the following hypotheses

$$\begin{aligned} H_0 &: \mathcal{P}_1 \otimes \mathcal{I}_{\mathbb{C}^2}, \\ H_1 &: \mathcal{P}_U \otimes \mathcal{I}_{\mathbb{C}^2}. \end{aligned} \quad (8.66)$$

with the statistical significance  $\delta \in [0, 1]$ .

---

**Algorithm 1:** Optimal strategy for the certification of von Neumann measurements.

---

**Input:** Measurement  $\mathcal{P}$ , which is either  $\mathcal{P}_U$  or  $\mathcal{P}_1$  and statistical significance  $\delta$

**Output:** Decision: “Accept  $H_0$ ” or “Reject  $H_0$ ”

- 1 Initialize input state  $|\psi_0\rangle\langle\psi_0| \in \Omega(\mathcal{X}^{\otimes 2})$ :
    - if**  $\|\mathcal{P}_U - \mathcal{P}_1\|_{\diamond} == 2$  **then**
      - |  $|\psi_0\rangle := \operatorname{argmax}_{|\psi\rangle \in \mathbb{C}^{d^2}: \langle\psi|\psi\rangle=1} \|((\mathcal{P}_U - \mathcal{P}_1) \otimes \mathcal{I}_{\mathcal{X}})(|\psi\rangle\langle\psi|)\|_1$ ;
    - else**
      - |  $|\psi_0\rangle := \sum_{i=0}^{d-1} \sqrt{\rho_0} |i\rangle \otimes |i\rangle$  for  $\rho_0$  defined in Lemma 3 and Proposition 3;
    - end**
  - 2 Perform  $\mathcal{P}$  on the first subsystem of  $|\psi_0\rangle\langle\psi_0|$ , that is  $(\mathcal{P} \otimes \mathcal{I}_{\mathcal{X}})(|\psi_0\rangle\langle\psi_0|)$ ;
  - 3 Read the measurement label  $i \in \{0, \dots, d-1\}$ ;
  - 4 Take the resulting quantum state  $|\psi_i\rangle\langle\psi_i|$  conditioned by label  $i$ :  
 $|\psi_i\rangle\langle\psi_i| \propto (\langle i| \otimes \mathbb{1}_{\mathcal{X}}) (\mathcal{P} \otimes \mathcal{I}_{\mathcal{X}}) (|\psi_0\rangle\langle\psi_0|) (|i\rangle \otimes \mathbb{1}_{\mathcal{X}})$ ;
  - 5 Prepare the measurement  $\{\Gamma_i, \mathbb{1}_{\mathcal{X}} - \Gamma_i\}$  conditioned on label  $i$ :
    - if**  $\|\mathcal{P}_U - \mathcal{P}_1\|_{\diamond} == 2$  **then**
      - |  $\Gamma_i$  is the projection onto the support of  $(\langle i| \otimes \mathbb{1}_{\mathcal{X}}) |\psi_0\rangle\langle\psi_0| (|i\rangle \otimes \mathbb{1}_{\mathcal{X}})$ ;
    - else**
      - |  $\Gamma_i = \Omega_i^{\top}$  where  $\Omega_i$  fullfills Eq. (8.51);
    - end**
  - 6 Perform the measurement  $\{\Gamma_i, \mathbb{1}_{\mathcal{X}} - \Gamma_i\}$  on  $|\psi_i\rangle\langle\psi_i|$ ;
  - 7 Read the measurement outcome  $j \in \{0, 1\}$ , where label  $j = 0$  is associated with effect  $\Gamma_i$  and the label  $j = 1$  with  $\mathbb{1}_{\mathcal{X}} - \Gamma_i$ ;
  - 8 **Result:**
    - if**  $j == 0$  **then**
      - | **return** “Accept  $H_0$ ”;
    - else**
      - | **return** “Reject  $H_0$ ”;
    - end**
-

### Optimal certification protocol

1. We prepare the optimal input state  $|\psi_0\rangle$ .
2. We perform one of the measurement  $\mathcal{P}$  on the first subsystem of  $|\psi_0\rangle\langle\psi_0|$ , that is  $(\mathcal{P} \otimes \mathcal{I}_{\mathbb{C}^2})(|\psi_0\rangle\langle\psi_0|)$ .
3. We read the measurement label either  $i = 0$  or  $i = 1$ .
4. We reduce the certification problem of the von Neumann measurements to the certification between quantum states depending on label  $i$ . Then, we obtain the following hypothesis.

$$\begin{aligned} H_0 &: |i\rangle, \\ H_1 &: H \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} H^\dagger |i\rangle. \end{aligned} \tag{8.67}$$

5. According to the optimal strategy for two-point certification of pure quantum states (Theorem 10 and Corollary 8), we prepare the optimal final measurement  $\Omega_0$  depending on the fixed statistical significance  $\delta$ .
6. We perform the measurement  $\Omega_0$  on  $|\psi_i\rangle\langle\psi_i|$ .
7. We read the measurement outcome  $k \in \{0, 1\}$ .
8. Finally, based on value of  $k$ , we make a decision whether we accept ( $k = 0$ ) or reject ( $k = 1$ ) the null hypothesis  $H_0$ .

### Optimal certification strategy

**Proposition 7** *For two-point certification of von Neumann measurements  $\mathcal{P}_1$  and  $\mathcal{P}_U$  defined in Eq. (8.65) with statistical significance  $\delta$ , the optimal input state  $|\psi_0\rangle$  is given by*

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}|\mathbb{1}_{\mathbb{C}^2}\rangle. \tag{8.68}$$

**Proof.** The proof of this proposition is straightforward. Observe that the input state  $|\psi_0\rangle$  need to maximize the diamond norm  $\|\mathcal{P}_U - \mathcal{P}_1\|_\diamond$ , where  $U$  is given by Eq. (8.65). So, the state  $|\psi_0\rangle$  can be determined in the same way as in Proposition 5

■

**Proposition 8** *For two-point certification of von Neumann measurements  $\mathcal{P}_1$  and  $\mathcal{P}_U$  defined in Eq. (8.65) with statistical significance  $\delta$ , the optimal measurement  $\Omega_0 = |\omega\rangle\langle\omega|$  is given by*

1. if  $i = 0$ , then we assume three cases:

(a) if  $\sqrt{1 + \cos \phi} \geq \sqrt{2\delta}$  and  $\phi \in [0, \pi)$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$ , where

$$|\omega\rangle = \sqrt{1 - \delta}|0\rangle - \sqrt{\delta}|1\rangle. \quad (8.69)$$

(b) if  $\sqrt{1 + \cos \phi} < \sqrt{2\delta}$  and  $\phi \in [0, 2\pi)$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$ , where

$$|\omega\rangle = \sin \frac{\phi}{2}|0\rangle - \cos \frac{\phi}{2}|1\rangle. \quad (8.70)$$

(c) if  $\sqrt{1 + \cos \phi} \geq \sqrt{2\delta}$  and  $\phi \in [\pi, 2\pi)$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$ , where

$$|\omega\rangle = \sqrt{1 - \delta}|0\rangle + \sqrt{\delta}|1\rangle. \quad (8.71)$$

2. if  $i = 1$ , then there are three cases:

(a) if  $\sqrt{1 + \cos \phi} \geq \sqrt{2\delta}$  and  $\phi \in [0, \phi)$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$ , where

$$|\omega\rangle = \sqrt{\delta}|0\rangle + \sqrt{1 - \delta}|1\rangle. \quad (8.72)$$

(b) if  $\sqrt{1 + \cos \phi} < \sqrt{2\delta}$  and  $\phi \in [0, 2\pi)$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$ , where

$$|\omega\rangle = \left| \cos \frac{\phi}{2} \right| |0\rangle + \frac{\sin \phi}{2 \left| \cos \frac{\phi}{2} \right|} |1\rangle. \quad (8.73)$$

(c) if  $\sqrt{1 + \cos \phi} \geq \sqrt{2\delta}$  and  $\phi \in [\pi, 2\pi)$ , then the optimal measurement is given by  $\Omega_0 = |\omega\rangle\langle\omega|$ , where

$$|\omega\rangle = -\sqrt{\delta}|0\rangle + \sqrt{1 - \delta}|1\rangle. \quad (8.74)$$

**Proof.** For the case  $i = 0$ , the certification scheme of von Neumann measurements is reduced to the certification scheme of quantum states:

$$\begin{aligned} H_0 &: |0\rangle \\ H_1 &: H \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} H^\dagger |0\rangle = \frac{1 + e^{-i\phi}}{2} |0\rangle + \frac{1 - e^{-i\phi}}{2} |1\rangle \end{aligned} \quad (8.75)$$

Next, we will use the the proof of Theorem 10. According to the assumption  $|\psi\rangle = \alpha|\phi\rangle + \beta|\phi^\perp\rangle$ , for some  $\alpha, \beta \geq 0$  satisfying  $\alpha^2 + \beta^2 = 1$ , we need to rewrite the state  $\frac{1}{2}(1 + e^{-i\phi})|0\rangle + \frac{1}{2}(1 - e^{-i\phi})|1\rangle$  to the form  $\frac{\sqrt{1+\cos\phi}}{\sqrt{2}}|0\rangle + \frac{i \sin\phi}{\sqrt{2+2\cos\phi}}|1\rangle$ . Hence, we obtain the following certification scheme

$$\begin{aligned} H_0 &: |0\rangle \\ H_1 &: \frac{\sqrt{1+\cos\phi}}{\sqrt{2}}|0\rangle + \frac{i \sin\phi}{\sqrt{2+2\cos\phi}}|1\rangle. \end{aligned} \quad (8.76)$$

To achieve the optimal final measurement for such defined certification scheme, we use Corollary 8, which complete first part of the proof.

For the case  $i = 1$ , the certification scheme of von Neumann measurements is reduced to the certification scheme of quantum states:

$$\begin{aligned} H_0 &: |1\rangle \\ H_1 &: H \begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix} H^\dagger |1\rangle = \frac{1 - e^{-i\phi}}{2}|0\rangle + \frac{1 + e^{-i\phi}}{2}|1\rangle \end{aligned} \quad (8.77)$$

In the same way, we rewrite the state  $\frac{1}{2}(1 - e^{-i\phi})|0\rangle + \frac{1}{2}(1 + e^{-i\phi})|1\rangle$  to the form  $\frac{\sin\phi}{\sqrt{2+2\cos\phi}}|0\rangle + \frac{\sqrt{1+\cos\phi}}{\sqrt{2}}|1\rangle$ . Then, we have the following hypothesis:

$$\begin{aligned} H_0 &: |1\rangle \\ H_1 &: \frac{\sin\phi}{\sqrt{2+2\cos\phi}}|0\rangle + \frac{\sqrt{1+\cos\phi}}{\sqrt{2}}|1\rangle \end{aligned} \quad (8.78)$$

Again, to achieve the optimal final measurement for such defined certification scheme, we use Corollary 8, which complete second part of the proof.  $\blacksquare$

**Remark 5** For two-point certification of von Neumann measurements  $\mathcal{P}_1$  and  $\mathcal{P}_U$  defined in Eq. (8.65) with statistical significance  $\delta$ , the optimal inputs state is given by

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}|\mathbb{1}_{\mathbb{C}^2}\rangle, \quad (8.79)$$

whereas the optimal measurement  $\mathcal{P}_{V_i}$  is given by

1. if  $\sqrt{1+\cos\phi} \geq \sqrt{2\delta}$  and  $\phi \in [0, \pi)$ , then

$$V_0 = \begin{pmatrix} \sqrt{1-\delta} & \sqrt{\delta} \\ -\sqrt{\delta} & \sqrt{1-\delta} \end{pmatrix}. \quad (8.80)$$

and

$$V_1 = \begin{pmatrix} \sqrt{\delta} & \sqrt{1-\delta} \\ \sqrt{1-\delta} & -\sqrt{\delta} \end{pmatrix}. \quad (8.81)$$

2. if  $\sqrt{1 + \cos \phi} < \sqrt{2\delta}$  and  $\phi \in [0, 2\pi)$ , then

$$V_0 = \begin{pmatrix} \sin \frac{\phi}{2} & |\cos \frac{\phi}{2}| \\ -\cos \frac{\phi}{2} & \frac{\sin \phi}{2|\cos \frac{\phi}{2}|} \end{pmatrix}. \quad (8.82)$$

and

$$V_1 = \begin{pmatrix} |\cos \frac{\phi}{2}| & \sin \frac{\phi}{2} \\ \frac{\sin \phi}{2|\cos \frac{\phi}{2}|} & -\cos \frac{\phi}{2} \end{pmatrix}. \quad (8.83)$$

3.  $\sqrt{1 + \cos \phi} \geq \sqrt{2\delta}$  and  $\phi \in [\pi, 2\pi)$ , then

$$V_0 = \begin{pmatrix} \sqrt{1-\delta} & -\sqrt{\delta} \\ \sqrt{\delta} & \sqrt{1-\delta} \end{pmatrix} \quad (8.84)$$

and

$$V_1 = \begin{pmatrix} -\sqrt{\delta} & \sqrt{1-\delta} \\ \sqrt{1-\delta} & \sqrt{\delta} \end{pmatrix}. \quad (8.85)$$

## 8.7 Realization of the von Neumann measurement certification scheme

In this section, we focus on the realization of certification scheme for parameterized family of qubit von Neumann measurements. A schematic representation of this setup is presented in 8.2. We will also use two methods to implement the controlled measurement  $\mathcal{P}_{V_i}$ , that is postselection and direct sum.

Note that the task of discrimination and certification the parameterized family of Fourier measurements are similar not only conceptually, but also implementationally. These two validation methods differ only in that each circuit in certification scheme consists of unitary gate  $U^\dagger$  whereas in discrimination scheme we implement  $U^\dagger$  or  $\mathbb{1}$ . It implies that number of circuits used in the certification scheme will be two times smaller than in discrimination one. Not surprisingly, the final conditional measurement has also different form, however the steps of postselection and direct sum will be the same. Below we describe the method of calculation the probability of type II error using postselection and direct sum approach.



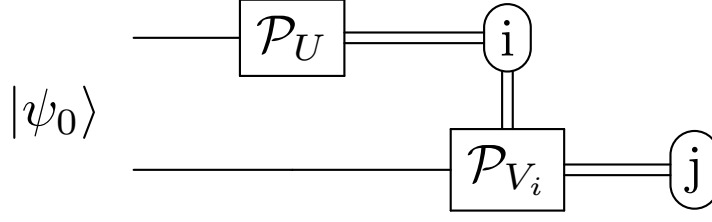


Figure 8.2: Schematic representation of the setup for von Neumann certification scheme with null hypothesis  $H_0 : \mathcal{P}_1 \otimes \mathcal{I}_X$  and alternative hypothesis  $H_1 : \mathcal{P}_U \otimes \mathcal{I}_X$  with statistical significance  $\delta$ . We implement the initial input state and the optimal final measurement  $\mathcal{P}_{V_i}$  based on Remark 5.

### Postselection

The postselection method is described in Fig. 8.3 and below we calculate the empirical probability of type II error.

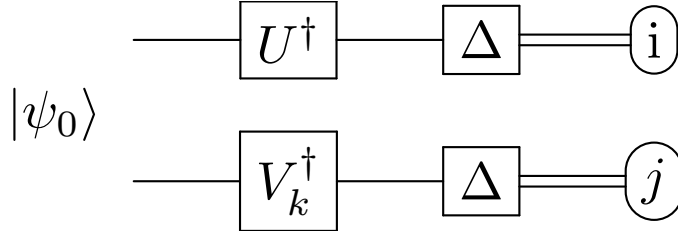


Figure 8.3: A schematic representation of the setup for certifying measurements using postselection. We prepare the Bell state  $|\psi_0\rangle$ , obtained from Proposition 7. Next, we apply the unitary  $U^\dagger$  on the first system. and  $V_k^\dagger$ , obtained from Proposition 8, on the second system. We measure both systems and reject cases when  $i \neq k$ . We make a decision based on the received label  $j$  on the second system. If  $j = 0$ , then we accept  $H_0$ . Otherwise, we decide to reject  $H_0$ .

Here, the experiments can be grouped into classes identified by tuples of the form  $(\mathcal{P}_U, k, i, j)$ . We discard all the experiments for which  $k \neq i$ . Hence, the total number of valid experiments is

$$N_{\text{total}} = \#\{(\mathcal{P}_U, k, i, j) : k = i\}. \quad (8.86)$$

Now we need to count the experiments (among the valid ones) including to calculation  $p_{\text{II}}$ . If we define

$$N_0 = \#\{(\mathcal{P}_U, k, i, j) : \mathcal{Q} = \mathcal{P}_U, k = i, j = 0\}, \quad (8.87)$$

then the empirical probability of type II error yields

$$p_{\text{II}} = \frac{N_0}{N_{\text{total}}}. \quad (8.88)$$

### Direct sum

The certification scheme by using the controlled unitary is described in Fig. 8.4 and below we calculate the empirical probability of type II error.

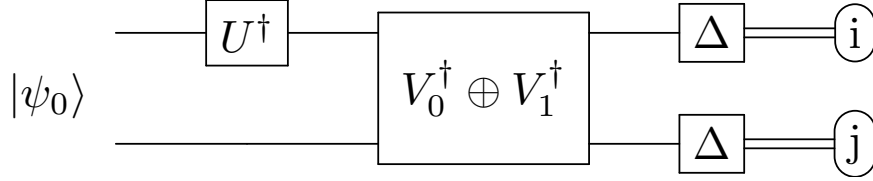


Figure 8.4: A schematic representation of the setup for certifying measurements using controlled unitary gate. We prepare the Bell state  $|\psi_0\rangle$ , obtained from Proposition 7. Next, we apply the unitary  $U^\dagger$  on the first system. We apply direct sum  $V_0^\dagger \oplus V_1^\dagger$ , obtained from Proposition 8, on the whole systems and measure them. We make a decision based on the received label  $j$  on the second system. If  $j = 0$ , then we accept  $H_0$ . Otherwise, we decide to reject  $H_0$ .

In this scheme, the experiment can be characterized by a pair  $(\mathcal{P}_U, i, j)$ . If we define

$$N_0 = \#\{(\mathcal{P}_U, i, j) : j = 0\}, \quad (8.89)$$

then the probability of type II error yields

$$p_{\text{succ}} = \frac{N_0}{N_{\text{total}}}, \quad (8.90)$$

where  $N_{\text{total}}$  is the number of trials.

## 8.8 Benchmarking based on certification scheme

Here, we extend the PyQBench package to certification experiments. Ultimately, we would like to have two approaches: using `qbench` as a library and as a CLI tool. Until now, however, PyQBench CLI is completed, whereas `qbench` as a library is a future work.

## PyQBench as a CLI tool

Recall, the CLI of PyQBench described in Chapter 7 has nested structure and the general form of the CLI is as below:

---

```
qbench <benchmark-type> <command> <parameters>
```

---

Here, for benchmarks using certification scheme, we fix the value of `<benchmark-type>` as `cert-fourier`, and next we obtain the same structure as previous consisting of four subcommands: `benchmark`, `status`, `resolve` and `tabulate`.

Below, we prepare the example of experiment YAML file and backend YAML file and the obtained results we present in Fig. 8.5 and Fig. 8.6. Observe that if we want to run benchmark on IBM Q device named Kolkata using certification scheme, it is enough to softly modify the experiment file from benchmarks based on discrimination scheme (see Listing 7.2), by fixing the `type` as `certification-fourier` and adding the component `delta`.

Listing 8.1: Defining certification experiment file

---

```
type: certification-fourier
qubits:
  - target: 0
    ancilla: 1
  - target: 1
    ancilla: 2
  - target: 14
    ancilla: 16
angles:
  start: 0
  stop: 2 * pi
  num_steps: 32
delta: 0.05
gateset: ibmq
method: postselection
num_shots: 8192
```

---

Listing 8.2: Defining certification backend file

---

```
name: ibmq_kolkata
asynchronous: true
provider:
  hub: ibm-q-psnc
  group: open
  project: main
```

---

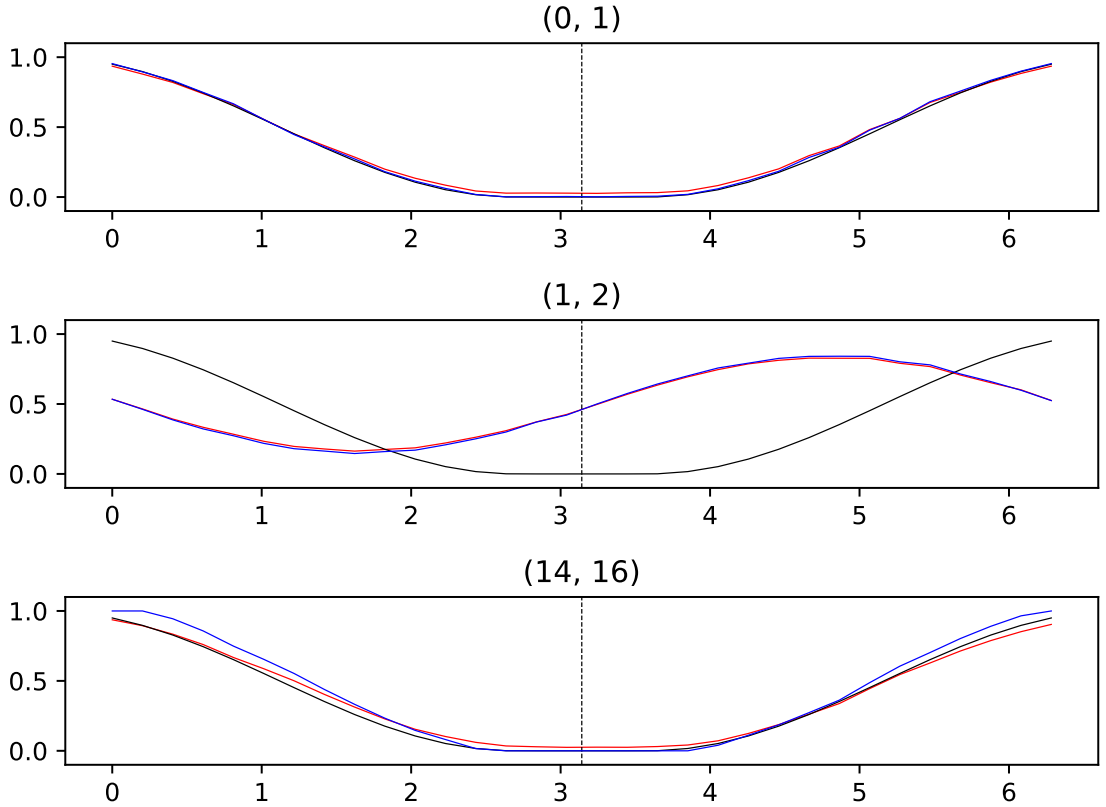


Figure 8.5: Certification experiment of the parameterized family of Fourier measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$  defined in Listing 8.1 using postselection with statistical significance  $\delta = 0.05$ . We run experiment on IBM Q device named Kolkata with 27 qubits using three pairs of qubits  $(0, 1)$ ,  $(1, 2)$  and  $(14, 16)$ . The theoretical probability of type II error is given by black line whereas the empirical probability is shown as the red line. The blue line represents the results after applying the Mthree error mitigation package.

## 8.9 Conclusion and discussion

This chapter studied the two-point certification of quantum states, unitary channels and von Neumann measurements. The problem of certification is inextricably related to quantum hypothesis testing. The aim of this task is to minimizing the probability of type II error with given the upper bound on the probability of type I error.

Although the problems of certification of quantum states and unitary channels

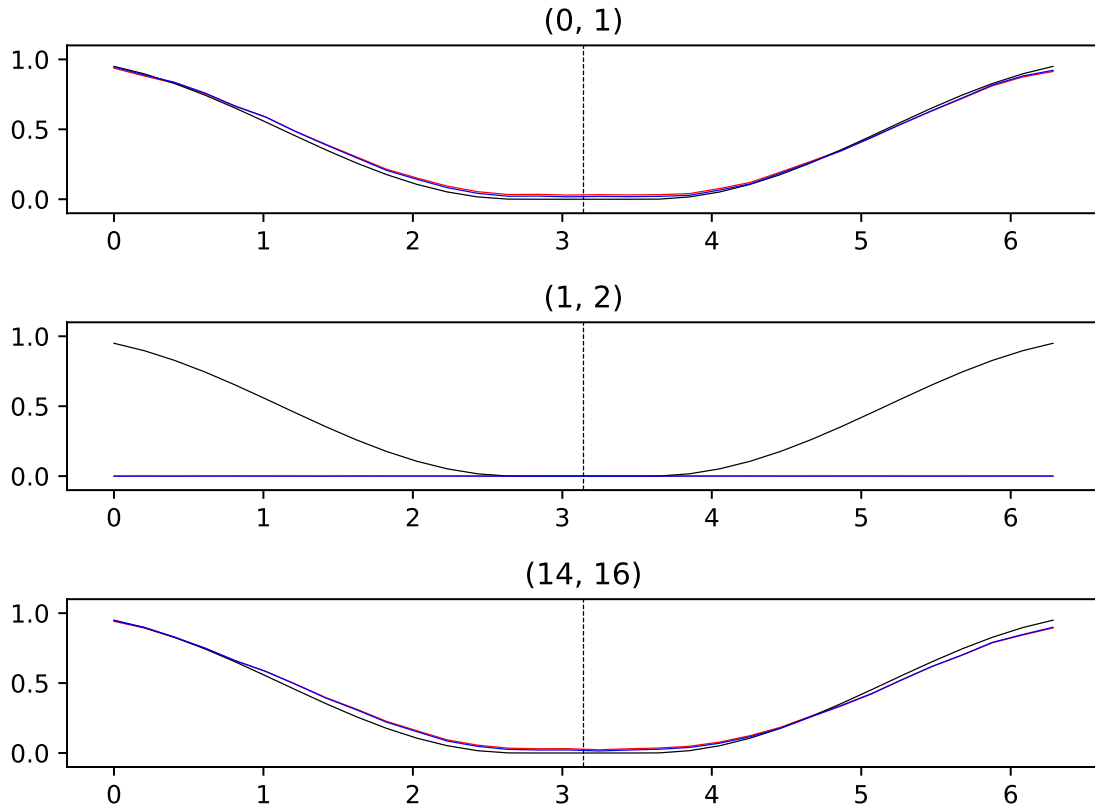


Figure 8.6: Certification experiment of the parameterized family of Fourier measurements  $\mathcal{P}_U$  and  $\mathcal{P}_1$  defined in Listing 8.1 using direct sum with statistical significance  $\delta = 0.05$ . We run experiment on IBM Q device named Kolkata with 27 qubits using three pairs of qubits  $(0, 1)$ ,  $(1, 2)$  and  $(14, 16)$ . The theoretical probability of type II error is given by black line whereas the empirical probability is shown by the red line. The blue line represents the results after applying the Mthree error mitigation package.

are well-studied, we pointed out the connection of certification of unitary channels with the notion of  $q$ -numerical range. Afterwards, we extended this approach to the certification of von Neumann measurements. We found a formula for the minimized probability of the type II error and the optimal certification strategy. It turned out that this formula can also be connected with the notion of  $q$ -numerical range.

Based on the obtained results, we have created the algorithm realizing a protocol of optimal certification strategy. Due to the algorithm, we created an optimal certification strategy for the parameterized family of qubit measurements on the Fourier basis. For this case, we also calculated the exact value of the probability of the type II error. Due to the obtained results, we extended PyQBench to another benchmark of NISQ devices based on the certification scheme. Until now, we offer using PyQBench as a command line interface (CLI) for this family of measurements. However, PyQBench is constantly being developed and ultimately we would like to extend PyQBench based on the certification scheme as a Python library. The code is available in [123]. This part of chapter is my contribution to the dissertation.

# Chapter 9

## Summary

This dissertation studied the problem of validation of modern quantum architectures. We investigated new validation methods for modern gate model-inspired NISQ devices. We analyzed both theoretical and engineering aspects. We studied various validation methods featuring three main approaches: learning, discrimination, and certification between von Neumann measurements. We showed that the created theoretical models allow for obtaining new concepts of benchmarking modern quantum systems. The following step of the dissertation was the implementation of obtained algorithms on quantum NISQ devices.

Initially, we considered a validation method based on the problem of learning von Neumann measurements. We estimated the asymptotic behaviour of the maximum value of the average fidelity function and determined possibly the best approximation of the optimal scheme. By using the deterministic port-based teleportation (DPBT) protocol, we were able to state the lower bound  $1 - \Theta\left(\frac{1}{N^2}\right)$ , which matched the obtained upper bound and hence, solved the given problem. In the qubit case, we also introduced a new scheme called the pretty good learning scheme. This scheme was a particular case of a parallel learning protocol using only two-qubit entangled memory states. Moreover, we compared the performance of different learning schemes: adaptive, parallel, based on DPBT, based on PPBT and the pretty good learning scheme for the qubit case.

We also used the quantum causal structures theory for the problem of quantum measurements learning. Such a scheme we called causal learning scheme. First, we considered causal learning scheme for two copies of von Neumann measurements. For this case, we proved that using an indefinite causal structure does not improve the average fidelity function  $F_d$ . Next, however, we showed the advantages of using causality in the problem of SAR of qubit von Neumann measurements. We wrote a semidefinite program to calculate  $F_2$ . As a result, we received a significant advantage between the standard approach of learning scheme and the causal learning scheme.

Next, we introduced a validation method based on the problem of discrimination

of von Neumann measurements. We were mainly interested in the discrimination scheme for the family of qubit measurements in the Fourier basis. For this case, we calculated the maximum value of the probability of correct discrimination and, additionally, we constructed the optimal strategy maximizing the probability of correct discrimination.

As an engineering aspect of the dissertation, we introduced PyQBench – an innovative open-source framework for benchmarking gate-based quantum computers. PyQBench benchmarks NISQ devices by verifying their capability based on the discrimination scheme. PyQBench provides a simplified, ready-to-use, command line interface (CLI) for running benchmarks using a predefined parametrized Fourier family of qubit measurements. For more advanced scenarios, PyQBench offers a way of employing user-defined measurements instead of predefined ones.

Finally, we considered the task of certification of von Neumann measurements. We analyzed a quantum hypothesis testing in which the null and alternative hypotheses are single-element sets. We calculated the minimized probability of the type II error and pointed out the connection of this result with the notion of  $q$ -numerical range. Based on obtained results, we created an algorithm describing the certification strategy, which minimizes the type II error and again, We determine such an optimal certification strategy for the qubit parameterized family of Fourier measurements. Furthermore, we extended PyQBench to benchmarking based on the certification scheme. Until now, we have offered a command line interface (CLI) for benchmarking NISQ devices using the pre-defined family of measurements. However, PyQBench is constantly being developed and ultimately we would like to extend PyQBench based on certification as a Python library.



# Bibliography

- [1] P. Lewandowska, R. Kukulski, Ł. Pawela, and Z. Puchała, “Storage and retrieval of von neumann measurements,” *Physical Review A*, vol. 106, no. 5, p. 052423, 2022.
- [2] Z. Puchała, Ł. Pawela, A. Krawiec, and R. Kukulski, “Strategies for optimal single-shot discrimination of quantum measurements,” *Physical Review A*, vol. 98, no. 4, p. 042103, 2018.
- [3] K. Jałowiecki, P. Lewandowska, and Ł. Pawela, “PyQBench: a Python library for benchmarking gate-based quantum computers,” *arXiv preprint arXiv:2304.00045*, 2023.
- [4] P. Lewandowska, A. Krawiec, R. Kukulski, Ł. Pawela, and Z. Puchała, “On the optimal certification of von neumann measurements,” *Scientific Reports*, vol. 11, no. 1, pp. 1–16, 2021.
- [5] J. Preskill, “Quantum computing in the nisq era and beyond,” *Quantum*, vol. 2, p. 79, 2018.
- [6] <https://www.ibm.com/quantum> The web resource at <https://www.ibm.com/quantum>. Accessed on 2023-02-18.
- [7] <https://www.rigetti.com/> The web resource at <https://www.rigetti.com/>. Accessed on 2023-02-18.
- [8] <http://oxfordquantum.org/> The web resource at <http://oxfordquantum.org/>. Accessed on 2023-02-18.
- [9] <https://ionq.com/> The web resource at <https://ionq.com/>. Accessed on 2023-02-18.
- [10] <https://www.xanadu.ai/> The web resource at <https://www.xanadu.ai/>. Accessed on 2023-02-18.

- [11] <https://www.dwavesys.com/> The web resource at <https://www.dwavesys.com/>. Accessed on 2023-02-18.
- [12] <https://www.quera.com/> The web resource at <https://www.quera.com/>. Accessed on 2023-02-18.
- [13] <https://docs.rigetti.com/qcs/> The web resource at <https://docs.rigetti.com/qcs/>. Accessed on 2023-02-18.
- [14] <https://pyquil-docs.rigetti.com/en/stable/> The web resource at <https://pyquil-docs.rigetti.com/en/stable/>. Accessed on 2023-02-18.
- [15] <https://qiskit.org/> The web resource at <https://qiskit.org/>. Accessed on 2023-02-18.
- [16] <https://quantum-computing.ibm.com/> The web resource at <https://quantum-computing.ibm.com/>. Accessed on 2023-02-18.
- [17] <https://aws.amazon.com/braket/> The web resource at <https://aws.amazon.com/braket/>. Accessed on 2023-02-18.
- [18] <https://www.zapatacomputing.com/orchestra-platform/> The web resource at <https://www.zapatacomputing.com/orchestra-platform/>. Accessed on 2023-02-18.
- [19] J. Preskill, “Quantum computing 40 years later,” *arXiv preprint arXiv:2106.10522*, 2021.
- [20] J. Carolan, J. D. Meinecke, P. J. Shadbolt, N. J. Russell, N. Ismail, K. Wörhoff, T. Rudolph, M. G. Thompson, J. L. O’Brien, J. C. Matthews, *et al.*, “On the experimental verification of quantum complexity in linear optics,” *Nature Photonics*, vol. 8, no. 8, pp. 621–626, 2014.
- [21] C. Chareton, S. Bardin, F. Bobot, V. Perrelle, and B. Valiron, “Toward certified quantum programming,” *arXiv preprint arXiv:2003.05841*, 2020.
- [22] Y.-D. Wu and B. C. Sanders, “Efficient verification of bosonic quantum channels via benchmarking,” *New Journal of Physics*, vol. 21, no. 7, p. 073026, 2019.
- [23] X. Jiang, K. Wang, K. Qian, Z. Chen, Z. Chen, L. Lu, L. Xia, F. Song, S. Zhu, and X. Ma, “Towards the standardization of quantum state verification using optimal strategies,” *npj Quantum Information*, vol. 6, no. 1, p. 90, 2020.

- [24] N. Spagnolo, C. Vitelli, M. Bentivegna, D. J. Brod, A. Crespi, F. Flamini, S. Giacomini, G. Milani, R. Ramponi, P. Mataloni, *et al.*, “Experimental validation of photonic boson sampling,” *Nature Photonics*, vol. 8, no. 8, pp. 615–620, 2014.
- [25] U. Chabaud, F. Grosshans, E. Kashefi, and D. Markham, “Efficient verification of boson sampling,” *Quantum*, vol. 5, p. 578, 2021.
- [26] Z. Hradil, “Quantum-state estimation,” *Physical Review A*, vol. 55, no. 3, p. R1561, 1997.
- [27] D. F. James, P. G. Kwiat, W. J. Munro, and A. G. White, “Measurement of qubits,” *Physical Review A*, vol. 64, no. 5, p. 052312, 2001.
- [28] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, “Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit,” *arXiv preprint arXiv:1310.4492*, 2013.
- [29] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, “Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography,” *Nature Communications*, vol. 8, no. 1, p. 14485, 2017.
- [30] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, “Quantum state tomography via compressed sensing,” *Physical Review Letters*, vol. 105, no. 15, p. 150401, 2010.
- [31] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, “Efficient quantum state tomography,” *Nature Communications*, vol. 1, no. 1, p. 149, 2010.
- [32] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, “Scalable reconstruction of density matrices,” *Physical review letters*, vol. 111, no. 2, p. 020401, 2013.
- [33] S. T. Flammia and Y.-K. Liu, “Direct fidelity estimation from few pauli measurements,” *Physical Review Letters*, vol. 106, no. 23, p. 230501, 2011.
- [34] A. Bisio, G. Chiribella, G. M. D’Ariano, S. Facchini, and P. Perinotti, “Optimal quantum learning of a unitary transformation,” *Physical Review A*, vol. 81, no. 3, p. 032324, 2010.
- [35] C. Helstrom, “Quantum detection and estimation theory, ser,” *Mathematics in Science and Engineering*. New York: Academic Press, vol. 123, 1976.

- [36] C. Lu, J. Chen, and R. Duan, “Optimal perfect distinguishability between unitaries and quantum operations,” *arXiv preprint arXiv:1010.2298*, 2010.
- [37] J. Watrous, *The theory of quantum information*. Cambridge university press, 2018.
- [38] C. W. Helstrom, “Quantum detection and estimation theory,” *Journal of Statistical Physics*, vol. 1, pp. 231–252, 1969.
- [39] A. Krawiec, Ł. Paweła, and Z. Puchała, “Excluding false negative error in certification of quantum channels,” *Scientific Reports*, vol. 11, no. 1, p. 21716, 2021.
- [40] A. Krawiec, Ł. Paweła, and Z. Puchała, “Discrimination of povms with rank-one effects,” *Quantum Information Processing*, vol. 19, pp. 1–12, 2020.
- [41] A. Krawiec, Ł. Paweła, and Z. Puchała, “Discrimination and certification of unknown quantum measurements,” *arXiv preprint arXiv:2301.04948*, 2023.
- [42] P. Lewandowska, R. Kukulski, and Ł. Paweła, “Optimal representation of quantum channels,” in *International Conference on Computational Science*, pp. 616–626, Springer, 2020.
- [43] J. Helsen, I. Roth, E. Onorati, A. H. Werner, and J. Eisert, “General framework for randomized benchmarking,” *PRX Quantum*, vol. 3, no. 2, p. 020357, 2022.
- [44] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, “Randomized benchmarking of quantum gates,” *Physical Review A*, vol. 77, no. 1, p. 012307, 2008.
- [45] J. Emerson, R. Alicki, and K. Życzkowski, “Scalable noise estimation with random unitary operators,” *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 7, no. 10, p. S347, 2005.
- [46] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell, *et al.*, “Quantum supremacy using a programmable superconducting processor,” *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.
- [47] I. Šupić and J. Bowles, “Self-testing of quantum systems: a review,” *Quantum*, vol. 4, p. 337, 2020.

- [48] Z. Shangnan and Y. Wang, “Quantum cross entropy and maximum likelihood principle,” *arXiv preprint arXiv:2102.11887*, 2021.
- [49] D. Mayers and A. Yao, “Self testing quantum apparatus,” *Quantum Information & Computation*, vol. 4, no. 4, pp. 273–286, 2004.
- [50] D. McMahon, *Quantum computing explained*. John Wiley & Sons, 2007.
- [51] F. D. Murnaghan, “On the field of values of a square matrix,” *Proceedings of the National Academy of Sciences*, vol. 18, no. 3, pp. 246–248, 1932.
- [52] O. Toeplitz, “Das algebraische analogon zu einem satze von fejér,” *Mathematische Zeitschrift*, vol. 2, no. 1, pp. 187–197, 1918.
- [53] F. Hausdorff, “Der wertvorrat einer bilinearform,” *Mathematische Zeitschrift*, vol. 3, no. 1, pp. 314–316, 1919.
- [54] C.-K. Li, “q-numerical ranges of normal and convex matrices,” *Linear and Multilinear Algebra*, vol. 43, no. 4, pp. 377–384, 1998.
- [55] M.-T. Chien and H. Nakazato, “Davis–wielandt shell and q-numerical range,” *Linear algebra and its applications*, vol. 340, no. 1-3, pp. 15–31, 2002.
- [56] N.-K. Tsing, “The constrained bilinear form and the c-numerical range,” *Linear Algebra and Its Applications*, vol. 56, pp. 195–206, 1984.
- [57] C.-K. Li and H. Nakazato, “Some results on the q-numerical,” *Linear and Multilinear Algebra*, vol. 43, no. 4, pp. 385–409, 1998.
- [58] R. Duan, Y. Feng, and M. Ying, “Perfect distinguishability of quantum operations,” *Physical Review Letters*, vol. 103, no. 21, p. 210501, 2009.
- [59] Ł. Pawela *et al.*, “Numerical shadow.” The web resource at <https://numericalshadow.org/>. Accessed on 2020-05-25.
- [60] D. Leung, B. Toner, and J. Watrous, “Coherent state exchange in multi-prover quantum interactive proof systems,” *Chicago Journal of Theoretical Computer Science*, vol. 11, pp. 1–18, 2013.
- [61] A. Bisio, G. D’Ariano, P. Perinotti, and G. Chiribella, “Quantum networks: General theory and applications,” *Acta Physica Slovaca*, vol. 61, no. 3, pp. 273–390, 2011.
- [62] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

- [63] E. B. Davies and J. T. Lewis, “An operational approach to quantum probability,” *Communications in Mathematical Physics*, vol. 17, no. 3, pp. 239–260, 1970.
- [64] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, “Causal and localizable quantum operations,” *Physical Review A*, vol. 64, no. 5, p. 052309, 2001.
- [65] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki, “Properties of quantum nonsignaling boxes,” *Physical Review A*, vol. 74, no. 1, p. 012305, 2006.
- [66] G. Chiribella, G. M. D’Ariano, P. Perinotti, and B. Valiron, “Quantum computations without definite causal structure,” *Physical Review A*, vol. 88, no. 2, p. 022318, 2013.
- [67] M. Raginsky, “A fidelity measure for quantum channels,” *Physics Letters A*, vol. 290, no. 1-2, pp. 11–18, 2001.
- [68] M. Christandl, F. Leditzky, C. Majenz, G. Smith, F. Spielman, and M. Walter, “Asymptotic performance of port-based teleportation,” *Communications in Mathematical Physics*, vol. 381, no. 1, pp. 379–451, 2021.
- [69] A. Bisio, G. M. D’Ariano, P. Perinotti, and M. Sedlák, “Quantum learning algorithms for quantum measurements,” *Physics Letters A*, vol. 375, no. 39, pp. 3425–3434, 2011.
- [70] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Theoretical framework for quantum networks,” *Physical Review A*, vol. 80, no. 2, p. 022339, 2009.
- [71] G. Chiribella, G. M. D’Ariano, and P. Perinotti, “Quantum circuit architecture,” *Physical Review Letters*, vol. 101, no. 6, p. 060401, 2008.
- [72] J. L. Park, “The concept of transition in quantum mechanics,” *Foundations of Physics*, vol. 23, no. 33, 2970.
- [73] W. K. Wootters and W. H. Zurek, “A single quantum cannot be cloned,” *Nature*, vol. 299, no. 5886, pp. 802–803, 1982.
- [74] M. A. Nielsen and I. L. Chuang, “Programmable quantum gate arrays,” *Physical Review Letters*, vol. 79, no. 2, p. 032324, 1997.
- [75] V. Bužek and M. Hillery, “Quantum copying: Beyond the no-cloning theorem,” *Physical Review A*, vol. 54, no. 3, p. 1844, 1996.

- [76] M. Hillery, V. Bužek, and M. Ziman, “Probabilistic implementation of universal quantum processors,” *Physical Review A*, vol. 65, no. 2, p. 022301, 2002.
- [77] Y. Yang, R. Renner, and G. Chiribella, “Optimal universal programming of unitary gates,” *Physical Review Letters*, vol. 125, no. 21, p. 210501, 2020.
- [78] M. Gschwendtner, A. Bluhm, and A. Winter, “Programmability of covariant quantum channels,” *Quantum*, vol. 5, p. 488, 2021.
- [79] V. P. Belavkin, G. M. D’Ariano, and M. Raginsky, “Operational distance and fidelity for quantum channels,” *Journal of Mathematical Physics*, vol. 46, no. 6, p. 062106, 2005.
- [80] M. Sedlák, A. Bisio, and M. Ziman, “Optimal probabilistic storage and retrieval of unitary channels,” *Physical Review Letters*, vol. 122, no. 17, p. 170502, 2019.
- [81] M. Sedlák and M. Ziman, “Probabilistic storage and retrieval of qubit phase gates,” *Physical Review A*, vol. 102, no. 3, p. 032618, 2020.
- [82] S. Ishizaka and T. Hiroshima, “Asymptotic teleportation scheme as a universal programmable quantum processor,” *Physical Review Letters*, vol. 101, no. 24, p. 240501, 2008.
- [83] S. Ishizaka and T. Hiroshima, “Quantum teleportation scheme by selecting one of multiple output ports,” *Physical Review A*, vol. 79, no. 4, p. 042306, 2009.
- [84] M. Studziński, S. Strelchuk, M. Mozrzyk, and M. Horodecki, “Port-based teleportation in arbitrary dimension,” *Scientific Reports*, vol. 7, no. 1, pp. 1–11, 2017.
- [85] M. Mozrzyk, M. Studziński, S. Strelchuk, and M. Horodecki, “Optimal port-based teleportation,” *New Journal of Physics*, vol. 20, no. 5, p. 053006, 2018.
- [86] C. S. Mukherjee, S. Maitra, V. Gaurav, and D. Roy, “Preparing dicke states on a quantum computer,” *IEEE Transactions on Quantum Engineering*, vol. 1, pp. 1–17, 2020.
- [87] P. Gawron, D. Kurzyk, and Ł. Paweła, “QuantumInformation.jl—a julia package for numerical computation in quantum information theory,” *PLOS ONE*, vol. 13, no. 12, p. e0209358, 2018.

- [88] B. O’Donoghue, E. Chu, N. Parikh, and S. Boyd, “Conic optimization via operator splitting and homogeneous self-dual embedding,” *Journal of Optimization Theory and Applications*, vol. 169, no. 3, pp. 1042–1068, 2016.
- [89] B. O’Donoghue, E. Chu, N. Parikh, and S. Boyd, “SCS: Splitting conic solver, version 3.2.1.” <https://github.com/cvxgrp/scs>, Nov. 2021.
- [90] <https://github.com/iitis/storage-and-retrieval-of-von-Neumann-measurements>. Permanent link to code/repository, Accessed: 2022-09-21.
- [91] Č. Brukner, “Quantum causality,” *Nature Physics*, vol. 10, no. 4, pp. 259–263, 2014.
- [92] K. Gödel, “An example of a new type of cosmological solutions of einstein’s field equations of gravitation,” *Reviews of Modern Physics*, vol. 21, no. 3, p. 447, 1949.
- [93] N. Gisin, “Weinberg’s non-linear quantum mechanics and supraluminal communications,” *Physics Letters A*, vol. 143, no. 1-2, pp. 1–2, 1990.
- [94] O. Oreshkov, F. Costa, and Č. Brukner, “Quantum correlations with no causal order,” *Nature Communications*, vol. 3, no. 1, pp. 1–8, 2012.
- [95] J. Bavaresco, M. Murao, and M. T. Quintino, “Strict hierarchy between parallel, sequential, and indefinite-causal-order strategies for channel discrimination,” *Physical Review Letters*, vol. 127, no. 20, p. 200504, 2021.
- [96] M. T. Quintino and D. Ebler, “Deterministic transformations between unitary operations: Exponential advantage with adaptive quantum circuits and the power of indefinite causality,” *Quantum*, vol. 6, p. 679, 2022.
- [97] J. Bavaresco, M. Murao, and M. T. Quintino, “Unitary channel discrimination beyond group structures: Advantages of sequential and indefinite-causal-order strategies,” *Journal of Mathematical Physics*, vol. 63, no. 4, p. 042203, 2022.
- [98] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, “Witnessing causal nonseparability,” *New Journal of Physics*, vol. 17, no. 10, p. 102001, 2015.
- [99] <https://github.com/iitis/causal-storage-and-retrieval-of-von-Neumann-measurements->. Permanent link to code/repository, Accessed: 2023-01-16.



- [100] R. Duan, Y. Feng, and M. Ying, “Entanglement is not necessary for perfect discrimination between unitary operations,” *Physical Review Letters*, vol. 98, no. 10, p. 100503, 2007.
- [101] Z. Ji, Y. Feng, R. Duan, and M. Ying, “Identification and distance measures of measurement apparatus,” *Physical Review Letters*, vol. 96, no. 20, p. 200401, 2006.
- [102] G. M. D’Ariano, P. L. Presti, and M. G. Paris, “Using entanglement improves the precision of quantum measurements,” *Physical Review Letters*, vol. 87, no. 27, p. 270404, 2001.
- [103] T.-Q. Cao, F. Gao, Z.-C. Zhang, Y.-H. Yang, and Q.-Y. Wen, “Perfect discrimination of projective measurements with the rank of all projectors being one,” *Quantum Information Processing*, vol. 14, no. 7, pp. 2645–2656, 2015.
- [104] Z. Puchała, Ł. Paweł, A. Krawiec, R. Kukulski, and M. Oszmaniec, “Multiple-shot and unambiguous discrimination of von neumann measurements,” *Quantum*, vol. 5, p. 425, 2021.
- [105] G. Wang and M. Ying, “Unambiguous discrimination among quantum operations,” *Physical Review A*, vol. 73, no. 4, p. 042301, 2006.
- [106] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, “Adaptive versus nonadaptive strategies for quantum channel discrimination,” *Physical Review A*, vol. 81, no. 3, p. 032339, 2010.
- [107] <https://github.com/iitis/PyQBench>. Permanent link to code/repository, Accessed: 2023-02-18.
- [108] <https://pyqbench.readthedocs.io/en/latest/>. Permanent link to code/repository, Accessed:2023-02-18.
- [109] N. Quetschlich, L. Burgholzer, and R. Wille, “Mqt bench: Benchmarking software and design automation tools for quantum computing,” *arXiv preprint arXiv:2204.13719*, 2022.
- [110] “MQTBench.” <https://github.com/cda-tum/MQTBench>. Accessed on 2023-02-18.
- [111] T. Tomesh, P. Gokhale, V. Omole, G. S. Ravi, K. N. Smith, J. Viszlai, X. Wu, N. Hardavellas, M. R. Martonosi, and F. T. Chong, “SupermarQ: A scalable quantum benchmark suite,” in *2022 IEEE International Symposium*

on *High-Performance Computer Architecture (HPCA)*, pp. 587–603, IEEE, 2022.

- [112] “SupermarQ.” <https://github.com/SupertechLabs/SupermarQ>. Accessed on 2023-02-18.
- [113] “Qiskit benchmarks.” <https://github.com/qiskit-community/qiskit-benchmarks>. Accessed on 2023-02-18.
- [114] “Forest Benchmarking: QCVV using PyQuil.” <https://github.com/rigetti/forest-benchmarking>. Accessed on 2023-02-18.
- [115] “Quantum volume in practice.” <https://github.com/lanl/Quantum-Volume-in-Practice>. Accessed on 2023-02-18.
- [116] <https://yaml.org/spec/1.2.2/>. Permanent link to code/repository, Accessed:2023-02-18.
- [117] <https://aws.amazon.com/blogs/quantum-computing/introducing-the-qiskit-provider-for-amazon-braket/> The web resource at <https://aws.amazon.com/blogs/quantum-computing/introducing-the-qiskit-provider-for-amazon-braket/>. Accessed on 2023-02-18.
- [118] <https://github.com/qiskit-community/qiskit-braket-provider> The web resource at <https://github.com/qiskit-community/qiskit-braket-provider>. Accessed on 2023-02-18.
- [119] <https://qiskit.org/documentation/partners/mthree/stubs/mthree.M3Mitigation.html> The web resource at <https://qiskit.org/documentation/partners/mthree/stubs/mthree.M3Mitigation.html>. Accessed on 2023-02-10.
- [120] P. D. Nation, H. Kang, N. Sundaresan, and J. M. Gambetta, “Scalable mitigation of measurement errors on quantum computers,” *PRX Quantum*, vol. 2, no. 4, p. 040326, 2021.
- [121] K. M. Audenaert, M. Mosonyi, and F. Verstraete, “Quantum state discrimination bounds for finite sample size,” *Journal of Mathematical Physics*, vol. 53, no. 12, p. 122205, 2012.
- [122] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, “Quantum certification and benchmarking,” *Nature Reviews Physics*, pp. 1–9, 2020.

- [123] <https://github.com/plewandowska777/PyQBench>. Permanent link to code/repository, Accessed: 2023-04-10.
- [124] L. Wang and R. Renner, "One-shot classical-quantum capacity and hypothesis testing," *Physical Review Letters*, vol. 108, no. 20, p. 200501, 2012.
- [125] R. Caron, X. Li, P. Mikusiński, H. Sherwood, and M. Taylor, "Nonsquare" doubly stochastic" matrices," *Lecture Notes-Monograph Series*, pp. 65–75, 1996.
- [126] <https://github.com/Qiskit/qiskit-aer> The web resource at <https://github.com/Qiskit/qiskit-aer>. Accessed on 2023-02-18.



# Appendix A

## Learning of von Neumann measurements

In this appendix we will proof the main theorem of Chapter 4 and next, we will describe in detail the pretty good learning scheme.

### A.1 Proof of Lemma 2 for qubit case

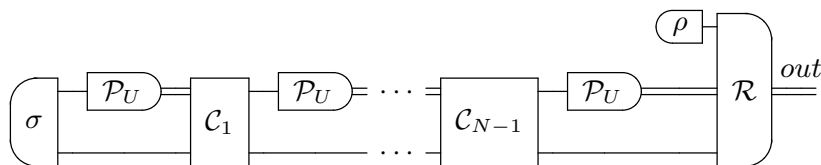


Figure A.1: The schematic representation of a learning scheme  $\mathcal{L} = (\sigma, \{\mathcal{C}_i\}_{i=1}^{N-1}, \mathcal{R})$ .

Let  $d = 2$  and let us fix  $N \in \mathbb{N}$ . By  $\mathcal{X}_i$  we denote the complex Euclidean space of dimension  $i$ , that is  $\mathcal{X}_i = \mathbb{C}^i$ . Let us consider a general single-qubit von Neumann measurement learning scheme  $\mathcal{L}$ , which is depicted in Fig. A.1. The Choi-Jamiołkowski representation of  $\mathcal{L}$  is given as  $L = \sum_{i=0}^1 |i\rangle\langle i| \otimes L_i$ , where  $|i\rangle \in \mathcal{X}_2^{(out)}$ . The result of composition of all copies of  $\mathcal{P}_U$  and the scheme  $\mathcal{L}$  is a measurement  $\mathcal{Q}_U = \{Q_{U,0}, Q_{U,1}\}$ , which is an approximation of  $\mathcal{P}_U$ . To define the effects  $Q_{U,i}$  we use the link product [61] in the following way  $\text{tr}(\rho Q_{U,i}) = \text{tr}(L_i^\top (\rho \otimes P_U^{\otimes N}))$  for  $\rho \in \Omega(\mathcal{X}_2)$  and  $i = 0, 1$ . Thus, we can calculate the fidelity defined in Eq. (3.16) between  $\mathcal{P}_U$  and  $\mathcal{Q}_U$

$$\mathcal{F}_2(\mathcal{P}_U, \mathcal{Q}_U) = \frac{1}{2} \sum_{i=0}^1 \text{tr}(P_{U,i} Q_{U,i}) = \frac{1}{2} \sum_{i=0}^1 \text{tr}[L_i^\top (P_{U,i} \otimes P_U^{\otimes N})]. \quad (\text{A.1})$$

Finally, we can express the maximum value of the average fidelity function  $F_2$  defined in Eq. (4.2) as

$$F_2 = \max_{\mathcal{L}} \int_U dU \frac{1}{2} \sum_{i=0}^1 \text{tr} [L_i^\top (P_{U,i} \otimes P_U^{\otimes N})]. \quad (\text{A.2})$$

In the following subsections we will upper bound  $F_2$  by using this simplified maximization formula.

### A.1.1 Measurement learning via parallel storage of unitary transformations

In this section we consider a new learning scheme, presented in Fig. A.2.

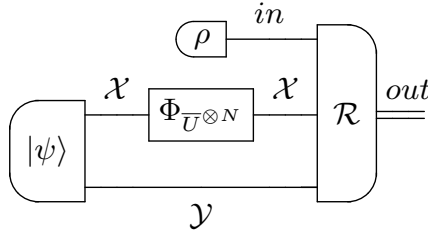


Figure A.2: Schematic representation of the setup, which we use to upper bound  $F_2$ . In this scenario, we are given  $N$  copies of unitary channel  $\Phi_{\bar{U}}$  in parallel. Our objective is to approximate the von Neumann measurement  $\mathcal{P}_U$ .

In this scheme, we are given  $N$  copies of a unitary channel,  $\Phi_{\bar{U}}$ , which we can use in parallel. Our goal is to approximate the measurement  $\mathcal{P}_U$  using the black box with the unitary channel  $\Phi_{\bar{U}}$  inside. To achieve this, we choose an initial memory state  $|\psi\rangle \in \mathcal{X} \otimes \mathcal{Y}$  and a retrieval binary measurement  $\mathcal{R} = \{R_0, R_1\}$ , such that  $R_i \in \mathcal{L}(\mathcal{Z} \otimes \mathcal{X} \otimes \mathcal{Y})$ , where  $\mathcal{Z} = \mathcal{X}_2^{(in)}$ ,  $\mathcal{X} = \mathcal{X}_{2N}$  and  $\mathcal{Y} = \mathcal{X}_{2N}$ . We maximize the value of the average fidelity function, which will be denoted as  $\widetilde{F}_2$ . To calculate  $\widetilde{F}_2$  we may observe that for a given  $\rho \in \Omega(\mathcal{Z})$ , the probability that outcome  $i$  occurs is equal to

$$p_i = \text{tr} \left( R_i \left( \rho \otimes (\bar{U}^{\otimes N} \otimes \mathbb{1}_{\mathcal{Y}}) |\psi\rangle\langle\psi| (U^{\top \otimes N} \otimes \mathbb{1}_{\mathcal{Y}}) \right) \right). \quad (\text{A.3})$$

Therefore, we obtain

$$\widetilde{F}_2 = \max_{\substack{\mathcal{R}=\{R_0, R_1\} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \int_U dU \frac{1}{2} \sum_{i=0}^1 \text{tr} \left[ R_i \left( (U \otimes \bar{U}^{\otimes N} \otimes \mathbb{1}_{\mathcal{Y}}) (|i\rangle\langle i| \otimes |\psi\rangle\langle\psi|) (U^\dagger \otimes U^{\top \otimes N} \otimes \mathbb{1}_{\mathcal{Y}}) \right) \right]. \quad (\text{A.4})$$

**Lemma 6** Let  $F_2$  and  $\widetilde{F}_2$  be the maximum value of the fidelity functions defined in Eq. (A.2) and Eq. (A.4), respectively. Then, it holds that  $F_2 \leq \widetilde{F}_2$ .

**Proof.** First, we observe that each von Neumann measurement  $\mathcal{P}_U$  can be written as a composition of the completely dephasing channel  $\Delta$  given by  $\Delta(X) = \sum_{i=0}^1 \langle i|X|i\rangle|i\rangle\langle i|$ , and a unitary channel  $\Phi_{U^\dagger}$ . Equivalently, that means  $P_U = (\Delta \otimes \mathcal{I}_{\mathbb{C}^2}) (|U^\dagger\rangle\langle U^\dagger|)$ . Due to the fact that the channel  $\Delta$  is self-adjoint, we obtain

$$\begin{aligned} \text{tr} [L_i^\top (P_{U,i} \otimes P_U^{\otimes N})] &= \\ &= \text{tr} \left[ \left( (\mathcal{I}_{\mathbb{C}^2} \otimes (\Delta \otimes \mathcal{I}_{\mathbb{C}^2})^{\otimes N}) (L_i) \right)^\top (P_{U,i} \otimes |U^\dagger\rangle\langle U^\dagger|^{\otimes N}) \right]. \end{aligned} \quad (\text{A.5})$$

Note that  $\sum_{i=0}^1 |i\rangle\langle i| \otimes (\mathcal{I}_{\mathbb{C}^2} \otimes (\Delta \otimes \mathcal{I}_{\mathbb{C}^2})^{\otimes N}) (L_i)$  represents the composition of the learning scheme  $\mathcal{L}$  and  $N$  copies of quantum channels  $\Delta$ . If we omit processing channels  $\Delta$ , we get the following upper bound of  $F_2$  defined in Eq. (A.2) given by

$$\begin{aligned} F_2 &\leq \max_{\mathcal{L}} \int_U dU \frac{1}{2} \sum_{i=0}^1 \text{tr} \left[ L_i^\top (P_{U,i} \otimes |U^\dagger\rangle\langle U^\dagger|^{\otimes N}) \right] \\ &= \frac{1}{2} \max_{\mathcal{L}} \int_U dU \text{tr} \left[ L^\top \left( (\mathbb{1}_{\mathbb{C}^2} \otimes U) J(\Delta) (\mathbb{1}_{\mathbb{C}^2} \otimes U^\dagger) \otimes |U^\dagger\rangle\langle U^\dagger|^{\otimes N} \right) \right], \end{aligned} \quad (\text{A.6})$$

where  $J(\Delta)$  is Choi matrix of  $\Delta$ . Observe that the maximal value of the integral in above equation is achievable by networks  $\mathcal{L}$  which satisfy the commutation relation

$$[L, \mathbb{1}_{\mathbb{C}^2} \otimes \bar{U} \otimes (\mathbb{1}_{\mathbb{C}^2} \otimes U)^{\otimes N}] = 0, \quad (\text{A.7})$$

for any unitary matrix  $U$ . To argue this fact, for any  $\mathcal{L}$  one can define a learning network  $\tilde{\mathcal{L}}$  given by

$$\tilde{L} = \int_U dU \left( (\mathbb{1}_{\mathbb{C}^2} \otimes \bar{U}) \otimes (\mathbb{1}_{\mathbb{C}^2} \otimes U)^{\otimes N} \right) L \left( (\mathbb{1}_{\mathbb{C}^2} \otimes U^\top) \otimes (\mathbb{1}_{\mathbb{C}^2} \otimes U^\dagger)^{\otimes N} \right). \quad (\text{A.8})$$

It is not difficult to see that  $\tilde{L}$  is a correctly defined Choi-Jamiołkowski representation of a quantum learning network [61, Theorem 2.5], which satisfies the relation Eq. (A.7). Moreover, for both  $L$  and  $\tilde{L}$  the value of the integral defined in Eq. (A.6) remains the same.

Let us divide  $\mathcal{L}$  into a storage network  $\mathcal{S}$  and a retrieval measurement  $\mathcal{R}$ , as shown in Fig. A.3. We introduce the input space  $\mathcal{X}_I := \bigotimes_{i=1}^N \mathcal{X}_2^{(2k)}$  (denoted with numbers  $2, 4, \dots, 2N$ ) and the output space  $\mathcal{X}_O := \bigotimes_{i=1}^N \mathcal{X}_2^{(2k-1)}$  (denoted with numbers  $1, 3, \dots, 2N-1$ ). Additionally, we define spaces  $\mathcal{X}_2^{(in)}$ ,  $\mathcal{X}_2^{(out)}$  and

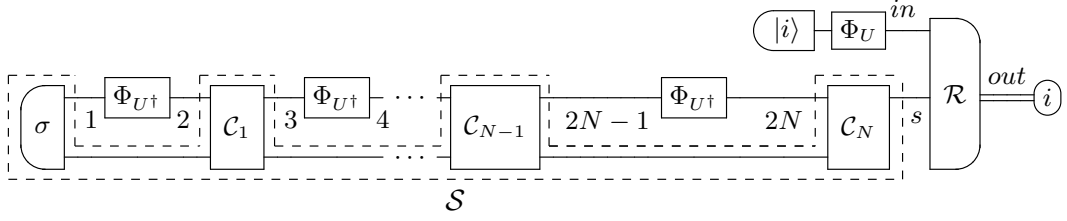


Figure A.3: Schematic representations of the right-hand side of Eq. (A.6). With probability  $1/2$  we prepare one of two states  $|0\rangle$  or  $|1\rangle$  and calculate the probability  $p_i$  given by Eq. (A.3) such that we obtain the output label  $i$ . So, Eq. (A.6) is the cumulative probability that provided the state  $|i\rangle\langle i|$  we measure  $i$ . The learning scheme  $\mathcal{L}$  is given as  $\mathcal{L} = (\sigma, \{\mathcal{C}_i\}_{i=1}^N, \mathcal{R})$  and the storage  $\mathcal{S}$  (marked with a dashed line) is defined as a composition of an initial memory state  $\sigma$  and processing channels  $\{\mathcal{C}_i\}_{i=1}^N$ .

$\mathcal{X}_s$  (see Fig. A.3). The space  $\mathcal{X}_s$  has arbitrary dimension  $s$ , but not smaller than the dimension of  $\mathcal{X}_I \otimes \mathcal{X}_O$ . The storage  $\mathcal{S}$  can be realized as a sequence of isometry channels followed by a partial trace operation [61, Theorem 2.6]. Therefore, by moving the partial trace operation to the retrieval part  $\mathcal{R}$ , we may assume that the storage  $\mathcal{S}$  consists of an initial pure state followed by a sequence of isometry channels  $\Phi_V$ . In consequence, the Choi-Jamiołkowski matrix of  $\mathcal{S}$  has the form  $S = |X\rangle\rangle\langle\langle X|$ . Observe that, there exists an isometry  $V \in \mathcal{U}(\mathcal{X}_s, \mathcal{X}_I \otimes \mathcal{X}_O)$ , such that  $X = \sqrt{\text{tr}_{\mathcal{X}_s} S} V^\top$ . In this notation,  $S$  is the solution of  $S = (\mathbb{1}_{\mathbb{C}^{4N}} \otimes V) |\sqrt{\text{tr}_{\mathcal{X}_s} S}\rangle\rangle\langle\langle \sqrt{\text{tr}_{\mathcal{X}_s} S}| (\mathbb{1}_{\mathbb{C}^{4N}} \otimes V)^\dagger$ . Hence, the isometry channel  $\Phi_V$  can be treated as a postprocessing of the storage  $\mathcal{S}$  and also viewed as a part of the retrieval  $\mathcal{R}$ . In summary, after all changes, the storage  $\mathcal{S}$  is of the form  $S = |\sqrt{\text{tr}_{\mathcal{X}_s} S}\rangle\rangle\langle\langle \sqrt{\text{tr}_{\mathcal{X}_s} S}|$ . By using the normalization condition [61, Theorem 2.5] for the network presented in Fig. A.3, we obtain  $\text{tr}_{\mathcal{X}_2^{(out)}} L = \mathbb{1}_{\mathbb{C}^2} \otimes \text{tr}_{\mathcal{X}_s} S$ . Therefore, using the property Eq. (A.7) we have

$$[\text{tr}_{\mathcal{X}_s} S, (\mathbb{1}_{\mathbb{C}^2} \otimes U)^{\otimes N}] = 0. \quad (\text{A.9})$$

Let us define the memory state  $\sigma_{\Phi_{U^\dagger}, \mathcal{S}}$  as an application of the storage  $\mathcal{S}$  on  $N$  copies of  $\Phi_{U^\dagger}$ . Then, we have

$$\begin{aligned} \sigma_{\Phi_{U^\dagger}, \mathcal{S}} &= \text{tr}_{\mathcal{X}_I \otimes \mathcal{X}_O} \left[ \left| \sqrt{\text{tr}_{\mathcal{X}_s} S} \right\rangle\rangle\langle\langle \sqrt{\text{tr}_{\mathcal{X}_s} S} \left| \left( |U^\top\rangle\rangle\langle\langle U^\top|^{\otimes N} \otimes \mathbb{1}_{\mathbb{C}^{4N}} \right) \right. \right] \\ &= \text{tr}_{\mathcal{X}_I \otimes \mathcal{X}_O} \left[ \left| (\mathbb{1}_{\mathbb{C}^2} \otimes U^\dagger)^{\otimes N} \sqrt{\text{tr}_{\mathcal{X}_s} S} \right\rangle\rangle\langle\langle (\mathbb{1}_{\mathbb{C}^2} \otimes U^\dagger)^{\otimes N} \sqrt{\text{tr}_{\mathcal{X}_s} S} \left| \right. \right. \\ &\quad \left. \left. \left( |\mathbb{1}_{\mathbb{C}^2}\rangle\rangle\langle\langle \mathbb{1}_{\mathbb{C}^2}|^{\otimes N} \otimes \mathbb{1}_{\mathbb{C}^{4N}} \right) \right] = (\mathbb{1}_{\mathbb{C}^2} \otimes \bar{U})^{\otimes N} |\psi\rangle\langle\psi| (\mathbb{1}_{\mathbb{C}^2} \otimes U^\top)^{\otimes N}, \end{aligned} \quad (\text{A.10})$$



where the last equality follows from the property Eq. (A.9) and taking

$$|\psi\rangle := \left( \langle\langle \mathbb{1}_{\mathbb{C}^2}^{\otimes N} \otimes \mathbb{1}_{4^N} \rangle\rangle \left| \sqrt{\text{tr}_{\mathcal{X}_s} S} \right\rangle \right). \quad (\text{A.11})$$

It means that an arbitrary storage strategy  $\mathcal{S}$ , which has access to  $N$  copies of a unitary channel  $\Phi_{U^\dagger}$  can be replaced with parallel storage strategy of  $N$  copies of a unitary channel  $\Phi_{\bar{U}}$ . By exploiting this property to Eq. (A.6) we finally obtain

$$\begin{aligned} F_2 &\leq \frac{1}{2} \max_{\mathcal{L}} \int_U dU \text{tr} \left[ L^\top \left( (\mathbb{1}_{\mathbb{C}^2} \otimes U) J(\Delta) (\mathbb{1}_{\mathbb{C}^2} \otimes U^\dagger) \otimes |U^\dagger\rangle\langle\langle U^\dagger|^{\otimes N} \right) \right] \\ &= \frac{1}{2} \max_{\mathcal{R}=\{R_0, R_1\}} \int_U dU \sum_{i=0}^1 \text{tr} \left[ R_i (U|i\rangle\langle i|U^\dagger \otimes \sigma_{\Phi_{U^\dagger}, \mathcal{S}}) \right] \\ &= \frac{1}{2} \max_{\substack{\mathcal{R}=\{R_0, R_1\} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X}_I \otimes \mathcal{X}_O)}} \int_U dU \sum_{i=0}^1 \text{tr} \left[ R_i (U|i\rangle\langle i|U^\dagger \otimes (\mathbb{1}_{\mathbb{C}^2} \otimes \bar{U})^{\otimes N} |\psi\rangle\langle\psi| \right. \\ &\quad \left. (\mathbb{1}_{\mathbb{C}^2} \otimes U^\top)^{\otimes N} \right] = \widetilde{F}_2, \end{aligned} \quad (\text{A.12})$$

which completes the proof. ■

## A.1.2 Objective function simplification

The aim of this section is to simplify the maximization of the fidelity function  $\widetilde{F}_2$  defined in Eq. (A.4). Recall,

$$\begin{aligned} \widetilde{F}_2 &= \max_{\substack{\mathcal{R}=\{R_0, R_1\} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \int_U dU \frac{1}{2} \sum_{i=0}^1 \text{tr} \left[ R_i \left( (U \otimes \bar{U}^{\otimes N} \otimes \mathbb{1}_Y) (|i\rangle\langle i| \otimes |\psi\rangle\langle\psi| \right. \right. \\ &\quad \left. \left. (U^\dagger \otimes U^{\top \otimes N} \otimes \mathbb{1}_Y) \right) \right). \end{aligned} \quad (\text{A.13})$$

Let us consider a binary measurement  $\mathcal{R} = \{R_0, R_1\}$  taken from the maximization domain in Eq. (A.4). It holds that  $R_0 + R_1 = \mathbb{1}_{\mathbb{C}^{2^{2N+1}}}$ , and hence we may write

$$\begin{aligned}
\widetilde{F}_2 &= \max_{\substack{\mathcal{R}=\{R_0, R_1\} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \int_U dU \frac{1}{2} \sum_{i=0}^1 \text{tr} [R_i ((U \otimes \bar{U}^{\otimes N} \otimes \mathbb{1}_Y)(|i\rangle\langle i| \otimes |\psi\rangle\langle\psi| \\
&\quad (U^\dagger \otimes U^{\top \otimes N} \otimes \mathbb{1}_Y))] = \frac{1}{2} + \frac{1}{2} \max_{\substack{\mathcal{R}=\{R_0, R_1\} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \int_U dU \text{tr} [R_0 ((U \otimes \bar{U}^{\otimes N} \otimes \mathbb{1}_Y) \\
&\quad (\sigma_z \otimes |\psi\rangle\langle\psi|)(U^\dagger \otimes U^{\top \otimes N} \otimes \mathbb{1}_Y))] = \\
&= \frac{1}{2} + \frac{1}{2} \max_{\substack{\mathcal{R}=\{R_0, R_1\} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \text{tr} \left[ \int_U dU (U^\dagger \otimes U^{\top \otimes N} \otimes \mathbb{1}_Y) R_0 (U \otimes \bar{U}^{\otimes N} \otimes \mathbb{1}_Y) \right. \\
&\quad \left. (\sigma_z \otimes |\psi\rangle\langle\psi|) \right], \tag{A.14}
\end{aligned}$$

where  $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ . Observe that, the integral of the matrix  $R_0$  over the unitary group  $\{U \otimes \bar{U}^{\otimes N} \otimes \mathbb{1}_Y\}_U$  is equivalent to taking  $R$  such that  $0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}}$  and

$$[R, U \otimes \bar{U}^{\otimes N} \otimes \mathbb{1}_Y] = 0, \tag{A.15}$$

for any qubit unitary matrix  $U$ . Equivalently, we may write

$$[R^{\top \mathcal{Z}}, U^{\otimes N+1} \otimes \mathbb{1}_Y] = 0, \tag{A.16}$$

where  $\cdot^{\top \mathcal{Z}}$  represents the partial transposition over subsystem  $\mathcal{Z}$ . According to [37, Theorem 7.15] the matrix  $R^{\top \mathcal{Z}}$  commutes with  $U^{\otimes N+1} \otimes \mathbb{1}_Y$  if and only if it is of the form

$$R^{\top \mathcal{Z}} = \sum_{\pi} W_{\pi} \otimes M_{\pi}, \tag{A.17}$$

where matrices  $W_{\pi} \in \text{L}(\mathcal{Z} \otimes \mathcal{X})$  represent subsystem permutation matrices acting on  $N + 1$  qubit systems, according to the equation

$$W_{\pi} |b_0, b_1, \dots, b_N\rangle = |b_{\pi(0)}, b_{\pi(1)}, \dots, b_{\pi(N)}\rangle, \quad b_k \in \{0, 1\}. \tag{A.18}$$

The matrices  $M_{\pi}$  belong to the set  $\text{L}(\mathcal{Y})$  and the index  $\pi$  goes over all permutations of the set  $\{0, \dots, N\}$ . Hence, we obtain the simplified formula  $\widetilde{F}_2$

$$\widetilde{F}_2 = \frac{1}{2} + \frac{1}{2} \max_{\substack{R: 0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}} \\ R = \sum_{\pi} W_{\pi}^{\top \mathcal{Z}} \otimes M_{\pi} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \text{tr} [R(\sigma_z \otimes |\psi\rangle\langle\psi|)]. \tag{A.19}$$

To simplify the form of  $\widetilde{F}_2$  even further, we introduce the following notation of basis states defined on  $N + 1$  qubit system with fixed weight. We enumerate qubit subsystems with numbers  $0, 1, \dots, N$ . For any subset  $A_k \subset \{1, \dots, N\}$ , such that  $|A_k| = k$ , we define:

$$\mathcal{X}_{2^N} \ni |A_k\rangle := \bigotimes_{i=1}^N (\delta(i \in A_k)|1\rangle + \delta(i \notin A_k)|0\rangle). \quad (\text{A.20})$$

Consider the following subspaces of the  $N + 1$  qubit space:

$$\mathcal{X}^{(k)} := \text{span}(|0\rangle|A_k\rangle, |1\rangle|A_{k+1}\rangle : A_k, A_{k+1} \subset \{1, \dots, N\}) \quad (\text{A.21})$$

for  $k = -1, \dots, N$ , where the vectors exist if and only if the expression is well-defined (for instance, the vectors  $|A_{-1}\rangle, |A_{N+1}\rangle$  do not exist). In this notation, subspaces  $\mathcal{X}^{(k)}$  constitute a decomposition of  $N + 1$  qubit space,  $\mathcal{X}_{2^{N+1}} = \bigoplus_{k=-1}^N \mathcal{X}^{(k)}$ . One may observe, that the matrix  $R$  appearing in the maximization domain of Eq. (A.19) is block diagonal in the introduced decomposition (in the partition  $\mathcal{Z} \otimes \mathcal{X}/\mathcal{Y}$ ). For such retrieval  $R$ , let us consider

$$H_R = \text{tr}_{\mathcal{Z}} (R(\sigma_z \otimes \mathbb{1}_{\mathbb{C}^{4^N}})). \quad (\text{A.22})$$

Observe that the matrix  $H_R$  is block diagonal in the decomposition

$$\mathcal{X}_{2^N} = \bigoplus_{k=0}^N \text{span}(|A_k\rangle : A_k \subset \{1, \dots, N\}). \quad (\text{A.23})$$

Hence, we will write  $H_R$  as

$$H_R = \bigoplus_{k=0}^N H_{R,k}. \quad (\text{A.24})$$

Utilizing the above observations, the maximization problem Eq. (A.19) can be written as

$$\begin{aligned} \widetilde{F}_2 &= \frac{1}{2} + \frac{1}{2} \max_{\substack{R: 0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}} \\ R = \sum_{\pi} W_{\pi}^{\top} \mathcal{Z} \otimes M_{\pi} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \text{tr} [R(\sigma_z \otimes |\psi\rangle\langle\psi|)] = \frac{1}{2} + \frac{1}{2} \max_{\substack{R: 0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}} \\ R = \sum_{\pi} W_{\pi}^{\top} \mathcal{Z} \otimes M_{\pi} \\ |\psi\rangle\langle\psi| \in \Omega(\mathcal{X} \otimes \mathcal{Y})}} \langle\psi|H_R|\psi\rangle \\ &= \frac{1}{2} + \frac{1}{2} \max_{k=0, \dots, N} \max_{\substack{R: 0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}} \\ R = \sum_{\pi} W_{\pi}^{\top} \mathcal{Z} \otimes M_{\pi}}} \lambda_1(H_{R,k}) \end{aligned} \quad (\text{A.25})$$

where  $\lambda_1(\cdot)$  stands for the largest eigenvalue. Finally, we observe that  $H_R = -(\sigma_x^{\otimes N} \otimes \mathbb{1}_y)H_R(\sigma_x^{\otimes N} \otimes \mathbb{1}_y)$ , where  $\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|$ . It implies that  $H_{R,k}$  is unitarily equivalent to  $-H_{R,N-k}$  for any  $k$ . We use this fact to write the final simplification of  $\widetilde{F}_2$ .

The following lemma sums up all the considerations presented in this section.

**Lemma 7** *For the fidelity function  $\widetilde{F}_2$  defined in Eq. (A.4) it holds that*

$$\widetilde{F}_2 = \frac{1}{2} + \frac{1}{2} \max_{k=0,\dots,\lfloor N/2 \rfloor} \max_{\substack{R: 0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}} \\ R = \sum_{\pi} W_{\pi}^{\top \mathcal{Z}} \otimes M_{\pi}}} \|H_{R,k}\|_{\infty}. \quad (\text{A.26})$$

### A.1.3 Technical lemmas

In the following lemma we will observe that optimization problem in Eq. (A.26) can be reduced to the case  $k \in \mathbb{N}$ , where  $N = 2k$ .

**Lemma 8** *Let  $N \in \mathbb{N}$  and let us take  $k \in \mathbb{N}$ , such that  $k \leq N/2$ . It holds that*

$$\max_{\substack{R: 0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}} \\ R = \sum_{\pi} W_{\pi}^{\top \mathcal{Z}} \otimes M_{\pi}}} \|H_{R,k}\|_{\infty} \leq \max_{\substack{\tilde{R}: 0 \leq \tilde{R} \leq \mathbb{1}_{\mathbb{C}^{2^{2\tilde{N}+1}}} \\ \tilde{R} = \sum_{\pi} \widetilde{W}_{\pi}^{\top \mathcal{Z}} \otimes \widetilde{M}_{\pi}}} \|\widetilde{H}_{\tilde{R},N-k}\|_{\infty}, \quad (\text{A.27})$$

and hence the number of systems on which the matrix  $\widetilde{W}_{\pi}$  acts is  $\tilde{N} + 1$ .

**Proof.** Let us fix  $R$  such that  $0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}}$  and  $R = \sum_{\pi} W_{\pi}^{\top \mathcal{Z}} \otimes M_{\pi}$ . Let us define

$$\tilde{R} := \sum_{\pi} (W_{\pi}^{\top \mathcal{Z}} \otimes \mathbb{1}_{\mathbb{C}^{2^{N-2k}}}) \otimes (M_{\pi} \otimes \mathbb{1}_{\mathbb{C}^{2^{N-2k}}}). \quad (\text{A.28})$$

Observe that the matrix  $\tilde{R}$  is in the maximization domain of the right-hand side of Eq. (A.27). Then, we have  $\widetilde{H}_{\tilde{R}} = \text{tr}_{\mathcal{Z}} \left( \tilde{R}(\sigma_{\mathcal{Z}} \otimes \mathbb{1}) \right) = \bigoplus_l \widetilde{H}_{\tilde{R},l}$ . The matrix  $\widetilde{H}_{\tilde{R},N-k}$  is defined on the space spanned by the vectors  $|A_{N-k}\rangle \in \mathcal{X}_{2^{\tilde{N}}}$  for  $A_{N-k} \subset \{1, \dots, \tilde{N}\}$ . These vectors can be expressed in the form  $|A_{N-k}\rangle = |B_i\rangle|B_{N-k-i}\rangle$ , where  $|B_i\rangle \in \mathcal{X}_{2^N}$  for  $B_i$  such that  $|B_i| = i$ ,  $B_i \subset \{1, \dots, N\}$ , and  $|B_{N-k-i}\rangle \in \mathcal{X}_{2^{N-2k}}$ ,  $B_{N-k-i} \subset \{N+1, \dots, \tilde{N}\}$ . Then, we have

$$\begin{aligned} & \left( \langle A_{N-k} | \otimes \mathbb{1}_{\mathbb{C}^{2^{\tilde{N}}}} \right) \widetilde{H}_{\tilde{R},N-k} \left( |A'_{N-k}\rangle \otimes \mathbb{1}_{\mathbb{C}^{2^{\tilde{N}}}} \right) = \\ & = \langle B_{N-k-i} | B'_{N-k-i'} \rangle \left( \langle B_i | \otimes \mathbb{1}_{\mathbb{C}^{2^N}} \right) H_R \left( |B'_{i'}\rangle \otimes \mathbb{1}_{\mathbb{C}^{2^N}} \right) \otimes \mathbb{1}_{\mathbb{C}^{4^{\tilde{N}-N}}}. \end{aligned} \quad (\text{A.29})$$

The non-zero blocks exist if and only if  $i = i'$  and  $B_{N-k-i} = B'_{N-k-i'}$ . Hence, we have

$$\tilde{H}_{\tilde{R}, N-k} = \bigoplus_{i=k}^{N-k} \bigoplus_{\substack{B_{N-k-i}: \\ B_{N-k-i} \subset \{N+1, \dots, \tilde{N}\}}} H_{R,i} \otimes \mathbb{1}_{\mathbb{C}^{4^{\tilde{N}-N}}}. \quad (\text{A.30})$$

That means

$$\|\tilde{H}_{\tilde{R}, N-k}\|_{\infty} = \max_{i=k, \dots, N-k} \|H_{R,i}\|_{\infty} \geq \|H_{R,k}\|_{\infty}. \quad (\text{A.31})$$

■

In the next lemma we will find the upper bound for Eq. (A.26) in the case  $N = 2k$  for  $k \in \mathbb{N}$ .

**Lemma 9** *Let  $k \in \mathbb{N}$  and  $N = 2k$ . For matrices  $R$  and  $H_{R,k}$  defined in Subsection A.1.2 we have*

$$\max_{\substack{R: 0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}} \\ R = \sum_{\pi} W_{\pi}^{\top} Z \otimes M_{\pi}}} \|H_{R,k}\|_{\infty} \leq 1 - \Theta\left(\frac{1}{k^2}\right). \quad (\text{A.32})$$

**Proof.** Let us fix  $R$  such that  $0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{2N+1}}}$  and  $R = \sum_{\pi} W_{\pi}^{\top} Z \otimes M_{\pi}$ . Through the rest of the proof, by  $B_l$  we denote subsets of  $\{1, \dots, 2k\}$ , such that  $|B_l| = l$ , for  $l = 0, \dots, 2k$ . Following the notation introduced in Subsection A.1.2, we define four types of vectors:

1.  $|+_{A_k}\rangle = x|0\rangle|A_k\rangle + \sum_{\substack{B_{k+1}: \\ |B_{k+1} \cap A_k| = k}} |1\rangle|B_{k+1}\rangle,$
2.  $|-_{A_k}\rangle = x|1\rangle|A_k\rangle + \sum_{\substack{B_{k-1}: \\ |B_{k-1} \cap A_k| = k-1}} |0\rangle|B_{k-1}\rangle,$
3.  $|\oplus_{A_k}\rangle = \sum_{\substack{B_{k+1}: \\ |B_{k+1} \cap A_k| = 1}} |1\rangle|B_{k+1}\rangle,$
4.  $|\ominus_{A_k}\rangle = \sum_{\substack{B_{k-1}: \\ |B_{k-1} \cap A_k| = 0}} |0\rangle|B_{k-1}\rangle,$

for each  $A_k \subset \{1, \dots, 2k\}$  and some  $x > 0$ . Now we define the following matrices:

1.  $I_+ = \sum_{A_k} |+_{A_k}\rangle\langle A_k|,$
2.  $I_- = \sum_{A_k} |-_{A_k}\rangle\langle A_k|,$
3.  $I_{\oplus} = \sum_{A_k} |\oplus_{A_k}\rangle\langle A_k|,$

$$4. I_{\ominus} = \sum_{A_k} |\ominus_{A_k}\rangle \langle A_k|.$$

For arbitrary  $A_k, A'_k \subset \{1, \dots, 2k\}$  we have

1.  $\langle +_{A_k} | +_{A'_k} \rangle = x^2 \delta(A_k = A'_k) + |\{B_{k+1} : |B_{k+1} \cap A_k| = k, |B_{k+1} \cap A'_k| = k\}|,$
2.  $\langle -_{A_k} | -_{A'_k} \rangle = x^2 \delta(A_k = A'_k) + |\{B_{k-1} : |B_{k-1} \cap A_k| = k-1, |B_{k-1} \cap A'_k| = k-1\}|,$
3.  $\langle \oplus_{A_k} | \oplus_{A'_k} \rangle = |\{B_{k+1} : |B_{k+1} \cap A_k| = 1, |B_{k+1} \cap A'_k| = 1\}|,$
4.  $\langle \ominus_{A_k} | \ominus_{A'_k} \rangle = |\{B_{k-1} : |B_{k-1} \cap A_k| = 0, |B_{k-1} \cap A'_k| = 0\}|.$

Observe that if  $A_k = A'_k$ , then the above inner products are  $x^2 + k, x^2 + k, k, k$ , respectively. If  $|A_k \cap A'_k| = k-1$ , then all the inner products are equal to one. Finally, if  $|A_k \cap A'_k| < k-1$ , then we obtain all the inner products are equal to zero. Moreover, we note two useful facts about matrices  $I_+, I_-, I_{\oplus}, I_{\ominus}$ . Firstly, we have

$$I_+^\dagger I_+ + I_{\oplus}^\dagger I_{\oplus} = I_-^\dagger I_- + I_{\ominus}^\dagger I_{\ominus}. \quad (\text{A.33})$$

Secondly, one can show that

$$\|I_+^\dagger I_+ + I_{\oplus}^\dagger I_{\oplus}\|_{\infty} = x^2 + 2k + 2k^2. \quad (\text{A.34})$$

As far as the first equality is straightforward, to show the second one, note that for each  $A_k$  there is exactly  $k^2$  sets  $A'_k$  such that  $|A_k \cap A'_k| = k-1$ . This means that, by the Birkhoff's theorem [125], we can express  $I_+^\dagger I_+ + I_{\oplus}^\dagger I_{\oplus}$  in the basis given by vectors  $|A_k\rangle$  as  $I_+^\dagger I_+ + I_{\oplus}^\dagger I_{\oplus} = (x^2 + 2k)\mathbb{1} + 2 \sum_{i=1}^{k^2} \Pi_i$ , where  $\Pi_i$  are permutation matrices. By the triangle inequality, we have that the spectral norm is no greater than  $x^2 + 2k + 2k^2$ . By taking the normalized vector  $|x\rangle \propto \sum_{A_k} |A_k\rangle$  we get  $\langle x | (I_+^\dagger I_+ + I_{\oplus}^\dagger I_{\oplus}) | x \rangle = x^2 + 2k + 2k^2$ .

To state the upper bound for  $\|H_{R,k}\|_{\infty}$  we will use the definition of  $H_R$  from Eq. (A.22) and the decomposition from Eq. (A.24). For a given  $A_k, A'_k \subset \{1, \dots, 2k\}$  we have that

$$\begin{aligned} & (\langle A_k | \otimes \mathbb{1}_y) H_{R,k} (|A'_k\rangle \otimes \mathbb{1}_y) = \\ & \sum_{\substack{\pi: \\ \pi(A_k)=A'_k}} M_{\pi} - \sum_{\substack{\pi: \\ \pi(0,A_k)=0,A'_k}} M_{\pi} = \sum_{\substack{\pi: \\ \pi(0)\neq 0, \\ \pi(A_k)=A'_k}} M_{\pi} - \sum_{\substack{\pi: \\ \pi(0)\neq 0, \\ \pi(0,A_k)=0,A'_k}} M_{\pi}. \end{aligned} \quad (\text{A.35})$$

Let us now define

$$\begin{aligned} G_{R,k} &= (I_+^\dagger \otimes \mathbb{1}_y) R (I_+ \otimes \mathbb{1}_y) + (I_{\oplus}^\dagger \otimes \mathbb{1}_y) R (I_{\oplus} \otimes \mathbb{1}_y) + \\ & - (I_-^\dagger \otimes \mathbb{1}_y) R (I_- \otimes \mathbb{1}_y) - (I_{\ominus}^\dagger \otimes \mathbb{1}_y) R (I_{\ominus} \otimes \mathbb{1}_y). \end{aligned} \quad (\text{A.36})$$

Taking  $A_k, A'_k \subset \{1, \dots, 2k\}$  we have:

$$\begin{aligned}
(\langle A_k | \otimes \mathbb{1}_y) G_{R,k}(|A'_k\rangle \otimes \mathbb{1}_y) &= \left( x^2 \sum_{\substack{\pi: \\ \pi(A_k)=A'_k}} M_\pi + x \sum_{\substack{B'_{k+1}, \pi: \\ |B'_{k+1} \cap A'_k|=k, \\ \pi(0, A_k)=B'_{k+1}}} M_\pi + x \sum_{\substack{B_{k+1}, \pi: \\ |B_{k+1} \cap A_k|=k, \\ \pi(B_{k+1})=0, A'_k}} M_\pi + \right. \\
&+ \left. \sum_{\substack{B_{k+1}, B'_{k+1}, \pi: \\ |B_{k+1} \cap A_k|=k, \\ |B'_{k+1} \cap A'_k|=k, \\ \pi(0, B_{k+1})=0, B'_{k+1}}} M_\pi + \sum_{\substack{B_{k+1}, B'_{k+1}, \pi: \\ |B_{k+1} \cap A_k|=1, \\ |B'_{k+1} \cap A'_k|=1, \\ \pi(0, B_{k+1})=0, B'_{k+1}}} M_\pi \right) - \left( x^2 \sum_{\substack{\pi: \\ \pi(0, A_k)=0, A'_k}} M_\pi + x \sum_{\substack{B'_{k-1}, \pi: \\ |B'_{k-1} \cap A'_k|=k-1, \\ \pi(A_k)=0, B'_{k-1}}} M_\pi + \right. \\
&+ x \sum_{\substack{B_{k-1}, \pi: \\ |B_{k-1} \cap A_k|=k-1, \\ \pi(0, B_{k-1})=A'_k}} M_\pi + \sum_{\substack{B_{k-1}, B'_{k-1}, \pi: \\ |B_{k-1} \cap A_k|=k-1, \\ |B'_{k-1} \cap A'_k|=k-1, \\ \pi(B_{k-1})=B'_{k-1}}} M_\pi + \left. \sum_{\substack{B_{k-1}, B'_{k-1}, \pi: \\ |B_{k-1} \cap A_k|=0, \\ |B'_{k-1} \cap A'_k|=0, \\ \pi(B_{k-1})=B'_{k-1}}} M_\pi \right). \tag{A.37}
\end{aligned}$$

Above formula can be simplified to

$$\begin{aligned}
(\langle A_k | \otimes \mathbb{1}_y) G_{R,k}(|A'_k\rangle \otimes \mathbb{1}_y) &= \\
&= \left( x^2 \sum_{\substack{\pi: \\ \pi(A_k)=A'_k}} M_\pi + x \sum_{\substack{B'_{k+1}, \pi: \\ |B'_{k+1} \cap A'_k|=k, \\ \pi(0, A_k)=B'_{k+1}}} M_\pi + x \sum_{\substack{B_{k+1}, \pi: \\ |B_{k+1} \cap A_k|=k, \\ \pi(B_{k+1})=0, A'_k}} M_\pi \right) \tag{A.38} \\
&- \left( x^2 \sum_{\substack{\pi: \\ \pi(0, A_k)=0, A'_k}} M_\pi + x \sum_{\substack{B'_{k-1}, \pi: \\ |B'_{k-1} \cap A'_k|=k-1, \\ \pi(A_k)=0, B'_{k-1}}} M_\pi + x \sum_{\substack{B_{k-1}, \pi: \\ |B_{k-1} \cap A_k|=k-1, \\ \pi(0, B_{k-1})=A'_k}} M_\pi \right).
\end{aligned}$$

Let us write the above as  $(\langle A_k | \otimes \mathbb{1}_y) G_{R,k}(|A'_k\rangle \otimes \mathbb{1}_y) = \sum_\pi c_\pi M_\pi$ , where  $c_\pi$  are some constants. For each  $\pi$ , let us determine the value of  $c_\pi$ :

- For  $\pi$  such that  $\pi(0) = 0, \pi(A_k) = A'_k$  we have  $c_\pi = x^2 - x^2 = 0$ .
- For  $\pi$  such that  $\pi(0) = 0, \pi(A_k) \neq A'_k$  we have  $c_\pi = 0$ .
- For  $\pi$  such that  $\pi(0) \neq 0, \pi(A_k) = A'_k$  we have  $c_\pi = x^2 + x + x = x^2 + 2x$ .
- For  $\pi$  such that  $\pi(0) \neq 0, \pi(A_k) \neq A'_k, \pi(0, A_k) \neq 0, A'_k$  there exists  $a_0 \notin \{0\} \cup A_k$ , such that  $\pi(a_0) \in \{0\} \cup A'_k$ . Therefore, we consider two subcases:
  - If for each  $a \notin \{0\} \cup A_k$  it holds  $\pi(a) \notin A'_k$ , then  $\pi(a_0) = 0, \pi(0) \in A'_k$  and  $A'_k \subset \pi(0, A_k)$ . Then,  $c_\pi = x - x = 0$ .
  - If  $\pi(a_0) \in A'_k$ , then we have two options:
    - \* If  $\pi(a_0, A_k) = 0, A'_k$ , then  $c_\pi = x - x = 0$ .
    - \* If  $\pi(a_0, A_k) \neq 0, A'_k$ , then  $c_\pi = 0$ .
- For  $\pi$  such that  $\pi(0) \neq 0$  and  $\pi(0, A_k) = 0, A'_k$ , we have  $c_\pi = -x^2 - x - x = -(x^2 + 2x)$ .

Therefore, we can see that  $G_{R,k} = (x^2 + 2x)H_{R,k}$ . Then, utilizing Eq. (A.33), Eq. (A.34) and Eq. (A.36) we get

$$-(x^2 + 2k + 2k^2)\mathbb{1}_{\mathbb{C}^{4N}} \leq G_{R,k} \leq (x^2 + 2k + 2k^2)\mathbb{1}_{\mathbb{C}^{4N}}, \quad (\text{A.39})$$

and finally we obtain

$$\|H_{R,k}\|_\infty \leq \frac{x^2 + 2k + 2k^2}{x^2 + 2x}. \quad (\text{A.40})$$

Minimizing over  $x > 0$ , we get for  $x \approx 2k^2$  that  $\|H_{R,k}\|_\infty \leq 1 - \Theta(1/k^2)$ , which completes this lemma.  $\blacksquare$

#### A.1.4 Proof of Lemma 2 for qubit case

**Proof of Lemma 2.** We have the following sequence of conclusions

$$\begin{aligned}
F_2 &\leq \widetilde{F}_2 && (\text{a}) \\
&= \frac{1}{2} + \frac{1}{2} \max_{k=0, \dots, \lfloor N/2 \rfloor} \max_{\substack{R: 0 \leq R \leq \mathbf{1}_{\mathbb{C}^{2N+1}} \\ R = \sum_{\pi} W_{\pi}^{\top} Z \otimes M_{\pi}}} \|H_{R,k}\|_\infty && (\text{b}) \\
&\leq \frac{1}{2} + \frac{1}{2} \max_{k=0, \dots, \lfloor N/2 \rfloor} \max_{\substack{\widetilde{R}: 0 \leq \widetilde{R} \leq \mathbf{1}_{\mathbb{C}^{2N+1}} \\ \widetilde{R} = \sum_{\pi} \widetilde{W}_{\pi}^{\top} Z \otimes \widetilde{M}_{\pi}}} \|\widetilde{H}_{\widetilde{R}, N-k}\|_\infty && (\text{c}) \\
&\leq \frac{1}{2} + \frac{1}{2} \max_{k=0, \dots, \lfloor N/2 \rfloor} 1 - \Theta\left(\frac{1}{(N-k)^2}\right) && (\text{d}) \\
&= 1 - \Theta\left(\frac{1}{N^2}\right),
\end{aligned}$$



where (a) follows by Eq. (A.2), Eq. (A.4) and Lemma 6, (b) by Lemma 7, (c) by Lemma 8 and (d) by Lemma 9.  $\blacksquare$

## A.2 Proof of upper bound for any dimension $d$

**Lemma 10** *Let  $d \in \mathbb{N}$ . The maximum value of the average fidelity function  $F_d$ , defined in Eq. (4.2) is upper bounded by*

$$F_d \leq 1 - \Theta\left(\frac{1}{N^2}\right). \quad (\text{A.41})$$

**Proof.** The thesis is true for  $d = 2$  which follows from Lemma 2. Now, let us fix  $d \in \mathbb{N}$ . Let us consider the optimal learning scheme  $\mathcal{L}$  such that  $\mathcal{F}_d^{\text{avg}}(\mathcal{L}) = F_d$ . Without loss of the generality, we assume that  $L$  satisfies the commutation condition, that is

$$[L, \mathbb{1}_{\mathbb{C}^d} \otimes U \otimes (\mathbb{1}_{\mathbb{C}^d} \otimes \bar{U})]^{\otimes N} = 0, \quad (\text{A.42})$$

for every unitary matrix  $U \in \text{U}(\mathbb{C}^d)$ . Then, for any  $U$ , we have

$$F_d = \mathcal{F}_d(\mathcal{P}_U, \mathcal{Q}_U), \quad (\text{A.43})$$

and

$$\mathcal{F}_d(\mathcal{P}_U, \mathcal{Q}_U) \leq \frac{d-1}{d} + \frac{1}{d} \min_i \text{tr}(P_{U,i} Q_{U,i}). \quad (\text{A.44})$$

Moreover, for  $j_1 \neq j_2$  it holds

$$\text{tr}(P_{U,j_1} Q_{U,j_2}) \leq \text{tr}(P_{U,j_1} (\mathbb{1}_d - Q_{U,j_1})) \leq 1 - \min_i \text{tr}(P_{U,i} Q_{U,i}). \quad (\text{A.45})$$

Now, we use  $\mathcal{L}$  to construct a new learning scheme  $\mathcal{L}'$  of qubit von Neumann measurements in the following way. To do so, we need to construct three parts of learning scheme – measurement  $\mathcal{P}'_U$ , the initial state  $\sigma'$  and retrieval part  $\mathcal{R}'$ .

To construct  $\mathcal{P}'_U$ , let  $\Pi \in \text{Proj}(\mathcal{X}_d)$  be a projector onto  $|0\rangle, |1\rangle$  and let us define an isometry matrix  $V = \sum_{i=0}^1 |i\rangle\langle i| \in \text{L}(\mathcal{X}_d, \mathcal{X}_2)$ . Having access to unknown qubit von Neumann measurement  $\mathcal{P}'_U$  we may use it to implement von Neumann measurement  $\mathcal{P}_{U \oplus \mathbb{1}_{d-2}}$  acting on  $\Omega(\mathcal{X}_d)$ . To do that, we take  $\sigma \in \Omega(\mathcal{X}_d)$  and measure it in the following way

$$|0\rangle\langle 0| \otimes \Pi \sigma \Pi + |1\rangle\langle 1| \otimes \sum_{i=2}^{d-1} \text{tr}(\sigma |i\rangle\langle i|) |i\rangle\langle i|. \quad (\text{A.46})$$

If on the first system we measure “0”, then we can project the initial state  $\sigma' \in \Omega(\mathcal{X}_2)$  into  $V^\dagger(\Pi \sigma \Pi)V$  and measure it by using  $\mathcal{P}_U$ . Otherwise, we do nothing. As a

result we implement the measurement of the form

$$\sum_{i=0}^1 \text{tr}(V^\dagger \sigma V U |i\rangle\langle i| U^\dagger) |i\rangle\langle i| + \sum_{i=2}^{d-1} \text{tr}(\sigma |i\rangle\langle i|) |i\rangle\langle i| = \mathcal{P}_{U \oplus \mathbf{1}_{d-2}}(\sigma). \quad (\text{A.47})$$

During the retrieval stage  $\mathcal{R}'$ , we project the input state  $\rho \in \Omega(\mathcal{X}_2)$  into  $V\rho V^\dagger$ . Moreover, as the output of  $\mathcal{L}$  is a classical label “0”, ..., “ $d-1$ ”, to finalize the construction of  $\mathcal{L}'$ , all the labels “1”, ..., “ $d-1$ ” are returned as “1”.

Finally, for a given unitary matrix  $U \in \text{U}(\mathcal{X}_2)$  and the learning network  $\mathcal{L}'$  we may calculate

$$\begin{aligned} \mathcal{F}_2(\mathcal{P}'_U, \mathcal{Q}'_U) &= \frac{1}{2} (\text{tr}(V P_{U,0} V^\dagger Q_{U \oplus \mathbf{1}_{d-2},0}) + \text{tr}(V P_{U,1} V^\dagger (\mathbb{1}_d - Q_{U \oplus \mathbf{1}_{d-2},0}))) \\ &\geq \frac{1}{2} \min_i \text{tr}(P_{U \oplus \mathbf{1}_{d-2},i} Q_{U \oplus \mathbf{1}_{d-2},i}) + \frac{1}{2} (1 - \text{tr}(P_{U \oplus \mathbf{1}_{d-2},1} Q_{U \oplus \mathbf{1}_{d-2},0})) \\ &\geq \min_i \text{tr}(P_{U \oplus \mathbf{1}_{d-2},i} Q_{U \oplus \mathbf{1}_{d-2},i}) \geq dF_d - (d-1). \end{aligned} \quad (\text{A.48})$$

Therefore, we get

$$dF_d - (d-1) \leq \mathcal{F}_2^{\text{avg}}(\mathcal{L}') \leq F_2 \leq 1 - \Theta\left(\frac{1}{N^2}\right), \quad (\text{A.49})$$

which completes the proof.  $\blacksquare$

### A.3 Pretty good learning scheme

The pretty good learning scheme  $\mathcal{L}_{\text{PGLS}} = (\sigma, \{\mathcal{C}_i\}_{i=1}^{N-1}, \mathcal{R})$  consists of the initial state  $\sigma$ , which is a tensor product of  $N$  copies of the maximally entangled state  $|\omega\rangle = \frac{1}{\sqrt{2}} |\mathbb{1}_{\mathbb{C}^2}\rangle\rangle$ , processing channels  $\{\mathcal{C}_i\}_{i=1}^{N-1}$  that are responsible for majority voting (see Section 4.3.1) and a binary measurement  $\mathcal{R} = \{R, \mathbb{1} - R\}$ . To construct the effect  $R$ , we fix  $N_0 \in \mathbb{N}$  and take  $n = N_0 - 1$ . Let us define

$$s_n(k, m) := \sum_{i=0}^k \sum_{j=0}^{n-k} \delta_{i+j-m} \binom{k}{i} \binom{n-k}{j} (-1)^{n-k-j}, \quad (\text{A.50})$$

being the convolution of binomial coefficients. We consider the effect  $R$  of the form

$$R = \sum_{k=0}^n |R_k\rangle\langle R_k| \quad (\text{A.51})$$

such that

$$|R_k\rangle = \frac{|M_k\rangle}{\|M_k\|_2}, \quad (\text{A.52})$$

$$\mathbb{L}(\mathbb{C}^2, \mathbb{C}^{2^{n+1}}) \ni M_k = \sum_{m=0}^{n+1} \frac{s_n(k, n-m)|0\rangle + s_n(k, n+1-m)|1\rangle}{\sqrt{\binom{n+1}{m}}} \langle D_m^{n+1}|, \quad (\text{A.53})$$

for  $k = 0, \dots, n$  and  $|D_m^{n+1}\rangle$  is the Dicke state.

**Lemma 11** *Let  $|x\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$ ,  $a, b \in \mathbb{C}$ . Then, we have  $M_k|x\rangle^{\otimes n+1} = (a+b)^k(a-b)^{n-k}|x\rangle$ .*

**Proof.** Direct calculations reveal

$$\begin{aligned} M_k|x\rangle^{\otimes n+1} &= \left[ \begin{array}{c} \sum_{m=0}^n \binom{n+1}{n-m} \cdot \frac{s_n(k, m)}{\binom{n+1}{n-m}} a^{m+1} b^{n-m} \\ \sum_{m=0}^n \binom{n+1}{n+1-m} \cdot \frac{s_n(k, m)}{\binom{n+1}{n+1-m}} a^m b^{n+1-m} \end{array} \right] = \sum_{m=0}^n s_n(k, m) a^m b^{n-m} |x\rangle \\ &= (a+b)^k (a-b)^{n-k} |x\rangle. \end{aligned} \quad (\text{A.54})$$

■

Moreover, to prove that  $R$  is a valid effect, let us now define

$$M := [s_n(k, m)]_{k, m=0}^n \quad (\text{A.55})$$

and a diagonal matrix

$$D := \sum_{m=0}^n \frac{1}{\binom{n}{m}} |m\rangle \langle m|. \quad (\text{A.56})$$

**Lemma 12** *With the notation given above, it holds that  $M^2 = 2^n \mathbb{1}_{\mathbb{C}^{n+1}}$ .*

**Proof.** First, observe that  $\mathbb{C}^{n+1} = \text{span}([x^k]_{k=0}^n : x \in \mathbb{C})$ . Let us take any vector of the form  $|x\rangle := [x^k]_{k=0}^n$ , where  $x \in \mathbb{C}$ . We have

$$\begin{aligned} M|x\rangle &= \left[ \sum_{m=0}^n s_n(k, m) x^m \right]_{k=0}^n = [(x+1)^k (x-1)^{n-k}]_{k=0}^n \\ &= (x-1)^n \left[ \left( \frac{x+1}{x-1} \right)^k \right]_{k=0}^n. \end{aligned} \quad (\text{A.57})$$

Finally, we calculate

$$M^2|x\rangle = (x-1)^n \left( \frac{x+1}{x-1} - 1 \right)^n \left[ \left( \frac{\frac{x+1}{x-1} + 1}{\frac{x+1}{x-1} - 1} \right)^k \right]_{k=0}^n = 2^n|x\rangle. \quad (\text{A.58})$$

■

**Lemma 13** *Using the notation presented above, we have the following equation  $MD = (MD)^\top$ .*

**Proof.** We will show that  $\langle k|MD|m\rangle = \langle m|MD|k\rangle$  for any  $m, k = 0, \dots, n$ . W.l.o.g. we can assume that  $k < m$ . On the one hand, it holds that

$$\begin{aligned} \langle k|MD|m\rangle &= \frac{s_n(k, m)}{\binom{n}{m}} = \sum_{\substack{i=0, \dots, k \\ j=0, \dots, n-k \\ i+j=m}} \frac{(-1)^{n-k-j} \binom{k}{i} \binom{n-k}{j}}{\binom{n}{m}} = \\ &= \sum_{i=\max(0, m+k-n)}^k \frac{(-1)^{n-k-m+i} \binom{k}{i} \binom{n-k}{m-i}}{\binom{n}{m}} = \\ &= (-1)^{n-k-m} \sum_{i=\max(0, m+k-n)}^k (-1)^i \frac{k!m!(n-k)!(n-m)!}{n!i!(k-i)!(m-i)!(n-k-m+i)!}. \end{aligned} \quad (\text{A.59})$$

On the other hand, we can calculate

$$\begin{aligned} \langle m|MD|k\rangle &= \frac{s_n(m, k)}{\binom{n}{k}} = \sum_{\substack{i=0, \dots, m \\ j=0, \dots, n-m \\ i+j=k}} \frac{(-1)^{n-m-j} \binom{m}{i} \binom{n-m}{j}}{\binom{n}{k}} \\ &= \sum_{i=\max(0, m+k-n)}^k \frac{(-1)^{n-k-m+i} \binom{m}{i} \binom{n-m}{k-i}}{\binom{n}{k}} \\ &= (-1)^{n-k-m} \sum_{i=\max(0, m+k-n)}^k (-1)^i \frac{k!m!(n-k)!(n-m)!}{n!i!(k-i)!(m-i)!(n-k-m+i)!}, \end{aligned} \quad (\text{A.60})$$

which gives us the desired equality and completes the proof. ■

**Lemma 14** *The operator  $R$  defined in Eq. (A.51) satisfies  $0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2n+2}}$ , and therefore  $\mathcal{R} = \{R, \mathbb{1} - R\}$  is a valid POVM.*

**Proof.** Let us fix  $N_0 \in \mathbb{N}$  and take  $n = N_0 - 1$ . Let us consider a matrix  $X := \frac{n+2}{n+1}MDM^\top$ . On the one hand, by using Lemma 12 and Lemma 13, we get

$$X = \frac{n+2}{n+1}(MD)^\top M^\top = \frac{n+2}{n+1}D(M^2)^\top = \frac{n+2}{n+1}2^n D. \quad (\text{A.61})$$

On the other hand, we have

$$\begin{aligned} \text{tr}\left(M_k^\dagger M_{k'}\right) &= \sum_{m=0}^n \frac{s_n(k, m)s_n(k', m)}{\binom{n+1}{n-m}} + \sum_{m=0}^n \frac{s_n(k, m)s_n(k', m)}{\binom{n+1}{n+1-m}} \\ &= \sum_{m=0}^n s_n(k, m)s_n(k', m) \left[ \frac{1}{\binom{n+1}{n-m}} + \frac{1}{\binom{n+1}{n-m+1}} \right] \\ &= \frac{n+2}{n+1} \sum_{m=0}^n \frac{s_n(k, m)s_n(k', m)}{\binom{n}{m}} = \langle k|X|k' \rangle. \end{aligned} \quad (\text{A.62})$$

Therefore, for all  $k \neq k'$  we get  $\text{tr}\left(M_k^\dagger M_{k'}\right) = 0$ . According to the definition Eq. (A.51), we get  $\langle R_k|R_{k'} \rangle = \delta_{k,k'}$ , which gives us  $0 \leq R \leq \mathbb{1}_{\mathbb{C}^{2^{n+2}}}$ .

■

**Lemma 15** *Let us fix  $N_0 \in \mathbb{N}$ . The approximation  $\mathcal{Q}_U = \{Q_{U,0}, \mathbb{1}_{\mathbb{C}^2} - Q_{U,0}\}$  of the von Neumann measurement  $\mathcal{P}_U$  obtained in the pretty good learning scheme is of the form*

$$Q_{U,0} = \frac{N_0}{N_0 + 1} P_{U,0}. \quad (\text{A.63})$$

**Proof.** Given a unitary matrix  $U \in U(\mathbb{C}^2)$  we take  $P_{U,0} = |x\rangle\langle x|$  for some unit vector  $|x\rangle \in \mathbb{C}^2$ . Let us decompose the  $(n+2)$ -qubit space in the following way  $\mathbb{C}^{2^{n+2}} = \mathcal{Z} \otimes \mathcal{X}$ , where  $\mathcal{Z} = \mathbb{C}^2$  and  $\mathcal{X} = \mathbb{C}^{2^{n+1}}$ . In the proof of Lemma 14 we defined the matrix  $X = \frac{n+2}{n+1}MDM^\top$  and showed that  $X = \frac{n+2}{n+1}2^n D$ , and  $\text{tr}\left(M_k^\dagger M_{k'}\right) = \langle k|X|k' \rangle$ . Therefore, for any  $k = 0, \dots, n$  we have  $\|M_k\|_2^2 = \frac{n+2}{n+1} \frac{2^n}{\binom{n}{k}}$ .

Due to this fact and by Lemma 11, we may express the effect  $Q_{U,0}$  as

$$\begin{aligned}
Q_{U,0} &= \text{tr}_{\mathcal{X}} \left( \left( \mathbb{1}_{\mathbb{C}^2} \otimes \overline{P_{U,0}}^{\otimes n+1} \right) R \right) = \left( \mathbb{1}_{\mathbb{C}^2} \otimes \overline{|x\rangle}^{\otimes n+1} \right) R \left( \mathbb{1}_{\mathbb{C}^2} \otimes \overline{|x\rangle}^{\otimes n+1} \right) \\
&= \sum_{k=0}^n \frac{1}{\|M_k\|_2^2} M_k |x\rangle \langle x|^{\otimes n+1} M_k^\dagger = \sum_{k=0}^n \frac{1}{\|M_k\|_2^2} |a + b|^{2k} |a - b|^{2(n-k)} |x\rangle \langle x| \\
&= \frac{n+1}{n+2} \sum_{k=0}^n \frac{\binom{n}{k}}{2^n} |a + b|^{2k} |a - b|^{2(n-k)} |x\rangle \langle x| = \frac{n+1}{n+2} \frac{(|a + b|^2 + |a - b|^2)^n}{2^n} |x\rangle \langle x| \\
&= \frac{n+1}{n+2} |x\rangle \langle x|,
\end{aligned} \tag{A.64}$$

which completes the proof. ■

# Appendix B

## PyQBench: a Python library for benchmarking gate-based quantum computers

In this Appendix we show examples of usage PyQBench. First, we will describe in detail how to use PyQBench's CLI after defining an experiment and a backend. Next, based on discrimination between measurement in Hadamard basis and computational one, we will show PyQBench as a library.

### B.1 PyQBench as a CLI

In this section we will present possible format of outputs which we can obtain using PyQBench's CLI.

#### Running the experiment and collecting measurements data

After preparing YAML files defining experiment and backend and running the benchmark, the result of running the above command can be two-fold:

- If the backend is synchronous, the output will contain measurement outcomes (bitstrings) for each of the circuits run.
- If backend is asynchronous, the output will contain intermediate data containing, among others, `job_ids` correlated with the circuit they correspond to.

For synchronous postselection experiment, the part of output looks similar to the one below.

Listing B.1: Output YML file `sync_results.yml`

---

```

metadata:
  experiments:
    type: certification-fourier
  qubits:
    - target: 0
      ancilla: 1
    - target: 1
      ancilla: 2
    - target: 14
      ancilla: 16
  angles:
    start: 0.0
    stop: 6.283185307179586
  num_steps: 32
  gateset: ibmq
  method: postselection
  num_shots: 8192
  backend_description:
    name: ibmq_kolkata
    asynchronous: false
  provider:
    group: open
    hub: ibm-q
    project: main
data:
  - target: 0
    ancilla: 1
    phi: 0.0
    results_per_circuit:
      - name: id
        histogram: {'00': 2108, '01': 2107, '10': 2064, '11': 1913}
        mitigation_info:
          target: {prob_meas0_prep1: 0.008399999999999963,
                  prob_meas1_prep0: 0.006}
          ancilla: {prob_meas0_prep1: 0.0126, prob_meas1_prep0:
                    0.0098000000000000031}
        mitigated_histogram: {'00': 0.25603047917671484, '01':
                              0.2573966558350486, '10': 0.2521648777124897, '11':
                              0.23440798727574674}
      - name: u
        histogram: {'00': 2108, '01': 2107, '10': 2064, '11': 1913}
        mitigation_info:

```



```
target: {prob_meas0_prep1: 0.008399999999999963, prob_meas1_prep0:
  0.006}
ancilla: {prob_meas0_prep1: 0.0126, prob_meas1_prep0:
  0.0098000000000000031}
mitigated_histogram: {'00': 0.25603047917671484, '01':
  0.2573966558350486, '10': 0.2521648777124897, '11':
  0.23440798727574674}
```

---

The data includes `target`, `ancilla`, `phi`, and `results_per_circuit`. The first three pieces of information have already been described. The last data `results_per_circuit` gives us the following additional information:

- **name**: the information which measurement is used during experiment, either string "u" for  $\mathcal{P}_U$  or string "id" for  $\mathcal{P}_1$ . In this example we consider  $\mathcal{P}_1$ .
- **histogram**: the dictionary with measurements' outcomes. The keys represent possible bitstrings, whereas the values are the number of occurrences.
- **mitigation\_info**: for backends corresponding to IBM Q devices, the object `backends.properties().qubits` contains information that might be used for error mitigation using the MThree method [119, 120]. If the mitigation info is not available this field will be absent.
- **mitigated\_histogram**: the histogram with measurements' outcomes after the error mitigation.

### Remarks on using the asynchronous flag

For backends supporting asynchronous execution, the `asynchronous` setting can be configured to toggle it. For asynchronous execution to work, the following conditions have to be met:

- Jobs returned by the backend have unique `job_id`.
- Jobs are retrievable from the backend using the `backend.retrieve_job` method, even from another process (e.g. if the original process running the experiment has finished).

Since PyQBench cannot determine if the job retrieval works for a given backend, it is the user's responsibility to ensure that this is the case before setting `asynchronous` to `true`. In this case, the output `async_results.yml` looks similar to the one below.

Listing B.2: Output YML file `async_results.yml`

---

```
metadata:
  experiments:
    type: discrimination-fourier
    qubits:
      - {target: 0, ancilla: 1}
      - {target: 1, ancilla: 2}
      - {target: 14, ancilla: 16}
    angles: {start: 0.0, stop: 6.283185307179586, num_steps: 32}
    gateset: ibmq
    method: postselection
    num_shots: 8192
  backend_description:
    name: ibmq_kolkata
    asynchronous: true
    provider: {group: open, hub: ibm-q-psnc, project: main}
data:
- job_id: 63e7f17a17b7ed49ca24e05b
  keys:
    - [0, 1, id_v0, 0.0]
    - [0, 1, id_v1, 0.0]
    - [0, 1, u_v0, 0.0]
    - [0, 1, u_v1, 0.0]
    - [0, 1, id_v0, 0.2026833970057931]
    - [0, 1, id_v1, 0.2026833970057931]
```

---

### (Optional) Getting status of asynchronous jobs

PyQBench provides also a helper command that will fetch the statuses of asynchronous jobs. The command is:

---

```
qbench disc-fourier status async_results.yml
```

---

and it will display dictionary with histogram of statuses.

### Resolving asynchronous jobs

Finally, if the status of jobs is `DONE`, for asynchronous experiments, the stored intermediate data has to be resolved in actual measurements' outcomes as follows:

---

```
qbench disc-fourier resolve async-results.yml resolved.yml
```

---

The resolved results, stored in `resolved.yml`, would look just like if the experiment was run synchronously. For this example the part of resolved output looks as below.

Listing B.3: Output YML file `resolved.yml`

---

```
metadata:
  experiments:
    type: discrimination-fourier
    qubits:
      - target: 0
        ancilla: 1
      - target: 1
        ancilla: 2
      - target: 14
        ancilla: 16
    angles:
      start: 0.0
      stop: 6.283185307179586
      num_steps: 32
    gateset: ibmq
    method: direct_sum
    num_shots: 8192
  backend_description:
    name: ibmq_kolkata
    asynchronous: true
    provider:
      group: open
      hub: ibm-q-psnc
      project: main
  data:
    - target: 0
      ancilla: 1
      phi: 0.0
      results_per_circuit:
        - name: id
      histogram:
        '00': 2108
        '01': 2107
        '10': 2064
        '11': 1913
  mitigation_info:
    target:
      prob_meas0_prep1: 0.008399999999999963
      prob_meas1_prep0: 0.006
```

```

    ancilla:
      prob_meas0_prep1: 0.0126
      prob_meas1_prep0: 0.0098000000000000031
  mitigated_histogram:
    '00': 0.25603047917671484
    '01': 0.2573966558350486
    '10': 0.2521648777124897
    '11': 0.23440798727574674
- name: u
  histogram:
    '00': 2015
    '01': 1989
    '10': 2116
    '11': 2072
  mitigation_info:
    target:
      prob_meas0_prep1: 0.0083999999999999963
      prob_meas1_prep0: 0.006
    ancilla:
      prob_meas0_prep1: 0.0126
      prob_meas1_prep0: 0.0098000000000000031
  mitigated_histogram:
    '00': 0.24453814824760045
    '01': 0.24254197782645848
    '10': 0.2585792025412591
    '11': 0.2543406713846819

```

---

Therefore, the final results will look the same no matter in which mode the benchmark was run, and hence in both cases the final output file is suitable for being an input for the command computing the discrimination probabilities.

### Tabulating results

As a last step in the processing workflow, the results file has to be passed to `tabulate` command saving the results in a CSV file presented in Table 7.1.

---

```
qbench disc-fourier tabulate results.yml results.csv
```

---

## B.2 PyQBench as a library

Here, we will demonstrate how `qbench` package can be used with user-defined measurement. For this purpose, we consider discrimination scheme between  $\mathcal{P}_H$  and

$\mathcal{P}_1$  for the Hadamard gate. Before we start let us determine the optimal strategy and the optimal discrimination probability for this scheme.

**Proposition 9** *Let us consider discrimination scheme between  $\mathcal{P}_H$  and  $\mathcal{P}_1$  where  $H$  is the Hadamard gate. The explicit formula for discriminator in this example reads:*

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad (\text{B.1})$$

with final measurements being equal to

$$V_0 = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \quad (\text{B.2})$$

and

$$V_1 = \begin{pmatrix} -\beta & \alpha \\ \alpha & \beta \end{pmatrix}, \quad (\text{B.3})$$

where

$$\alpha = \frac{\sqrt{2 - \sqrt{2}}}{2} = \cos\left(\frac{3}{8}\pi\right), \quad (\text{B.4})$$

and

$$\beta = \frac{\sqrt{2 + \sqrt{2}}}{2} = \sin\left(\frac{3}{8}\pi\right). \quad (\text{B.5})$$

Finally, the optimal probability of correct discrimination is equal to

$$p_{succ}(\mathcal{P}_H, \mathcal{P}_1) = \frac{1}{2} + \frac{\sqrt{2}}{4}. \quad (\text{B.6})$$

**Proof.** Let  $U = H$  and let  $\Phi_U$  and  $\Phi_1$  be two unitary channels. We will show that

$$\min_{E \in \text{DU}(\mathbb{C}^2)} \|\Phi_{UE} - \Phi_1\|_{\diamond} = \|\Phi_{UE_0} - \Phi_1\|_{\diamond}, \quad (\text{B.7})$$

where

$$E_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1+i & 0 \\ 0 & -1-i \end{pmatrix}. \quad (\text{B.8})$$

For  $HE_0$  we calculate that

$$\nu^2((HE_0)^\dagger) = \frac{1}{2} \quad (\text{B.9})$$

So, we obtain

$$\|\Phi_{HE_0} - \Phi_1\|_{\diamond} = \sqrt{2}. \quad (\text{B.10})$$

It implies that it is enough to prove

$$\min_{E \in \text{DU}(\mathbb{C}^2)} \|\Phi_{HE} - \Phi_{\mathbf{1}}\|_{\diamond} = \sqrt{2}. \quad (\text{B.11})$$

This condition is equivalent to show that for every  $E \in \text{DU}(\mathbb{C}^2)$

$$\nu(H^\dagger E) \leq \frac{1}{\sqrt{2}} \quad (\text{B.12})$$

which is equivalent to show that for every  $E \in \text{DU}(\mathbb{C}^2)$  there exists  $\rho \in \Omega(\mathbb{C}^2)$  such that

$$|\text{tr}(\rho HE)| \leq \frac{1}{\sqrt{2}}. \quad (\text{B.13})$$

Now, let us define  $E = \begin{pmatrix} E_0 & 0 \\ 0 & E_1 \end{pmatrix}$  and take  $\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ . Then, for every  $E \in \text{DU}(\mathbb{C}^2)$  we have

$$\begin{aligned} |\text{tr}(\rho HE)| &= \frac{1}{2\sqrt{2}} \left| \text{tr} \begin{pmatrix} E_0 & E_1 \\ E_0 & -E_1 \end{pmatrix} \right| = \frac{1}{2\sqrt{2}} |E_0 - E_1| \\ &\leq \frac{1}{\sqrt{2}}, \end{aligned} \quad (\text{B.14})$$

which completes the proof of Eq.(B.7). Now, let us take  $|\psi_0\rangle$  of the form

$$|\psi_0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle). \quad (\text{B.15})$$

We calculate the trace distance

$$\|(\mathcal{P}_H \otimes \mathbb{1})(|\psi_0\rangle\langle\psi_0|) - (\mathcal{P}_{\mathbf{1}} \otimes \mathbb{1})(|\psi_0\rangle\langle\psi_0|)\|_1 = \sqrt{2}. \quad (\text{B.16})$$

It implies  $|\psi_0\rangle$  is the discriminator, and moreover

$$p_{\text{succ}}(\mathcal{P}_H, \mathcal{P}_{\mathbf{1}}) = \frac{1}{2} + \frac{\sqrt{2}}{4}. \quad (\text{B.17})$$

Next, similarly as in Section 6.2.2, we determine the final measurements  $\mathcal{P}_{V_i}$ . For the discrimination task between  $\mathcal{P}_H$  and  $\mathcal{P}_{\mathbf{1}}$  the explicit form of  $V_0$  and  $V_1$  is given as follows:

$$V_0 = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}, \quad (\text{B.18})$$

and

$$V_1 = \begin{pmatrix} -\beta & \alpha \\ \alpha & \beta \end{pmatrix}, \quad (\text{B.19})$$

where

$$\alpha = \frac{\sqrt{2 - \sqrt{2}}}{2} = \cos\left(\frac{3}{8}\pi\right), \quad (\text{B.20})$$

and

$$\beta = \frac{\sqrt{2 + \sqrt{2}}}{2} = \sin\left(\frac{3}{8}\pi\right). \quad (\text{B.21})$$

We include the `Mathematica` notebook [107] in `mathematics` file computing  $V_0$  and  $V_1$  of the final optimal measurement  $\mathcal{P}_{V_i}$ ,  $i \in \{0, 1\}$ . ■

## Decomposition

To use the above benchmarking scheme in PyQBench, we first need to decompose the unitary matrices above into sequences of gates. In standard approach, a circuit taking  $|00\rangle$  to the Bell state  $|\psi_0\rangle$  comprises the Hadamard gate followed by CNOT gate on both qubits (see Fig. B.1). To decompose  $V_0$  observe that  $V_0 = \text{RY}\left(\frac{3}{4}\pi\right)$ ,

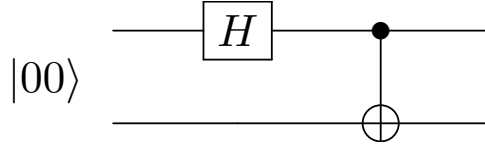


Figure B.1: Decomposition of the Bell state  $|\psi_0\rangle$ .

where RY is rotation gate around the  $Y$  axis. To obtain  $V_1$  we need only to swap the columns, i.e.

$$V_1 = \text{RY}\left(\frac{3}{4}\pi\right) X. \quad (\text{B.22})$$

**Remark 6** *It is worth mentioning that each NISQ device has own native build-in gates which can be implemented in quantum circuit. For example IBM Q devices have the following native gates: RZ, RX and CX, whereas Rigetti computers have: RZ, RX and CZ. By default, PyQBench uses native gates to perform benchmarks running through CLI. However, if a user creates their experiment by using the PyQBench library, we need to decompose a circuit with native gates on our own.*

## Implementation

We will now demonstrate how to implement this scheme in PyQBench. For this example we will use the Qiskit Aer simulator [126]. First, we import the necessary

functions and classes from PyQBench and Qiskit. We also import `numpy` for the definition of `np.pi` constant and the exponential function. The exact purpose of the imported functions will be described at the point of their usage.

Listing B.4: Imports needed for running benchmarking example

---

```
import numpy as np
from qiskit import QuantumCircuit, Aer
from qbench.schemes.postselection import benchmark_using_postselection
from qbench.schemes.direct_sum import benchmark_using_direct_sum
```

---

To implement the discrimination scheme in PyQBench, we need to define all the necessary components as Qiskit instructions. We can do so by constructing a circuit object `QuantumCircuit()` acting on two qubits 0 and 1 and then converting them using `to_instruction()` method.

Listing B.5: Defining components for Hadamard experiment

---

```
def state_prep():
    circuit = QuantumCircuit(2)
    circuit.h(0)
    circuit.cnot(0, 1)
    return circuit.to_instruction()

def u_dag():
    circuit = QuantumCircuit(1)
    circuit.h(0)
    return circuit.to_instruction()

def v0_dag():
    circuit = QuantumCircuit(1)
    circuit.ry(-np.pi * 3 / 4, 0)
    return circuit.to_instruction()

def v1_dag():
    circuit = QuantumCircuit(1)
    circuit.ry(-np.pi * 3 / 4, 0)
    circuit.x(0)
    return circuit.to_instruction()

def v0_v1_direct_sum_dag():
    circuit = QuantumCircuit(2)
    circuit.ry(-np.pi * 3 / 4, 0)
    circuit.cnot(0, 1)
    return circuit.to_instruction()
```



---

We now construct a backend object, which in this case is an instance of Aer simulator.

Listing B.6: Defining a backend

---

```
simulator = Aer.get_backend("aer_simulator")
```

---

In the simplest scenario, when we do not want to tweak execution details and simply wishes to run the experiment on a given backend, everything that is required is to run `benchmark_using_postselection` or `benchmark_using_direct_sum` function, depending on the user preference. It is worth mentioning that exactly here we can fix qubits (`target` and `ancilla`) which we want to use in the experiment and number of experiments sampled the empirical probability of correct discrimination. Recall, in the postselection scheme the total number of experiments equals  $4 \cdot \text{num\_shots\_per\_measurement}$ , whereas in the direct sum experiment the total number is  $2 \cdot \text{num\_shots\_per\_measurement}$ .

Listing B.7: Simulation benchmark by using postselection

---

```
postselection_result = benchmark_using_postselection(  
    backend=simulator,  
    target=0,  
    ancilla=1,  
    state_preparation=state_prep(),  
    u_dag=u_dag(),  
    v0_dag=v0_dag(),  
    v1_dag=v1_dag(),  
    num_shots_per_measurement=10000)
```

---

Listing B.8: Simulation benchmark by using direct sum

---

```
direct_sum_result = benchmark_using_direct_sum(  
    backend=simulator,  
    target=1,  
    ancilla=2,  
    state_preparation=state_prep(),  
    u_dag=u_dag(),  
    v0_v1_direct_sum_dag=v0_v1_direct_sum_dag(),  
    num_shots_per_measurement=10000)
```

---

The `postselection_result` and `direct_sum_result` variables contain now the empirical probabilities of correct discrimination. We can compare them to the theoretical value and compute the absolute error.

Listing B.9: Examining the benchmark results

---

```

p_succ = (2 + np.sqrt(2)) / 4
print(f"Analytical p_succ = {p_succ}")
print(f"Postselection: p_succ = {postselection_result}, abs. error =
      {p_succ - postselection_result}")
print(f"Direct sum: p_succ = {direct_sum_result}, abs. error =
      {p_succ - direct_sum_result}")

```

---

```

Analytical p_succ = 0.8535533905932737
Postselection: p_succ = 0.8559797193791593, abs. error =
      -0.0024263287858855564
Direct sum: p_succ = 0.85605, abs. error = -0.0024966094067262468

```

---

### Gaining more control of the experiment

In the example presented above we used functions that automate the whole process – from the circuit assembly, through running the simulations to interpreting the results. But what if we want more control over some parts of this process?

To show how to modify the basic example, for the rest of this example we focus only on the postselection case, as the direct sum case is analogous. We continue by importing two more functions from PyQBench.

Listing B.10: Assembling circuits

---

```

from qbench.schemes.postselection import (
    assemble_postselection_circuits,
    compute_probabilities_from_postselection_measurements)

circuits = assemble_postselection_circuits(
    target=0,
    ancilla=1,
    state_preparation=state_prep(),
    u_dag=u_dag(),
    v0_dag=v0_dag(),
    v1_dag=v1_dag())

```

---

Recall that for a postselection scheme we have two possible choices of the measurement,  $\mathcal{P}_U$  or  $\mathcal{P}_1$ , and two possible choices of a final measurement,  $\mathcal{P}_{V_0}$  or  $\mathcal{P}_{V_1}$ . It gives a total of four circuits needed to run the benchmark. The function `assemble_postselection_circuits` creates all four circuits and places them in a dictionary with keys "id\_v0", "id\_v1", "u\_v0", "u\_v1".

We will now run our circuits using noisy and noiseless simulation. We start by creating a noise model using Qiskit.

Listing B.11: Adding noise model

---

```
from qiskit.providers.aer import noise

error = noise.ReadoutError([[0.75, 0.25], [0.8, 0.2]])

noise_model = noise.NoiseModel()
noise_model.add_readout_error(error, [0])
noise_model.add_readout_error(error, [1])
```

---

Once we have our noise model ready, we can execute the circuits with and without noise. To this end, we will use Qiskit's `execute` function. One caveat is that we have to keep track which measurements correspond to which circuit. We do so by fixing an ordering on the keys in the `circuits` dictionary.

Listing B.12: Running circuits

---

```
from qiskit import execute

keys_ordering = ["id_v0", "id_v1", "u_v0", "u_v1"]
all_circuits = [circuits[key] for key in keys_ordering]

counts_noisy = execute(
    all_circuits,
    backend=simulator,
    noise_model=noise_model,
    shots=10000).result().get_counts()

counts_noiseless = execute(
    all_circuits,
    backend=simulator,
    shots=10000).result().get_counts()
```

---

Finally, we use the measurement counts to compute discrimination probabilities using `compute_probabilities_from_postselection_measurements` function.

Listing B.13: Computation probabilities

---

```
prob_succ_noiseless =
    compute_probabilities_from_postselection_measurements(
        id_v0_counts=counts_noiseless[0],
        id_v1_counts=counts_noiseless[1],
        u_v0_counts=counts_noiseless[2],
```

```
u_v1_counts=counts_noiseless[3])

prob_succ_noisy =
  compute_probabilities_from_postselection_measurements(
    id_v0_counts=counts_noisy[0],
    id_v1_counts=counts_noisy[1],
    u_v0_counts=counts_noisy[2],
    u_v1_counts=counts_noisy[3])
```

---

As a result, we obtained `prob_succ_noiseless = 0.8524401115559386` and `prob_succ_noisy = 0.5017958400693446`. As expected, for noisy simulations, the result lies further away from the target value of `0.8535533905932737`.

# Appendix C

## Certification of von Neumann measurements

In this Appendix, we will focus on calculating the distance from the  $q$ -numerical range to the origin of the coordinate system.

### C.1 Distance of $q$ -numerical range to zero

Let us begin with the two-dimensional case when the unitary matrix  $U$  has two eigenvalues  $\lambda_1$  and  $\lambda_2$ . Without loss of generality we can assume  $\lambda_1 = 1$ . From [57] we know that the  $q$ -numerical range is an elliptical disc with eccentricity equal to  $q$  and foci  $q\lambda_1$  and  $q\lambda_2$ , see Fig. C.1. Let  $c$  denote the distance from the center of the ellipse to the focus and  $a$  be the distance from the center of the ellipse to its vertex. Using this notation the eccentricity yields  $q = c/a$ . Let  $b$  denote the distance from the center of the ellipse to its co-vertex, which is the point which saturates the minimum.

First, we will calculate  $b$ . We note that

$$c = \frac{1}{2} \|q\lambda_1 - q\lambda_2\| = \frac{q}{2} \|\lambda_1 - \lambda_2\| = \frac{\sqrt{1-\delta}}{2} \|\lambda_1 - \lambda_2\|. \quad (\text{C.1})$$

From the properties of the ellipse and the form of the eccentricity  $q$  we have

$$b = \sqrt{a^2 - c^2} = \sqrt{\frac{c^2}{q^2} - c^2} = c\sqrt{\frac{1}{q^2} - 1} = c\sqrt{\frac{1}{1-\delta} - 1} = c\sqrt{\frac{\delta}{1-\delta}}. \quad (\text{C.2})$$

Hence

$$b = \frac{\sqrt{1-\delta}}{2} \|\lambda_1 - \lambda_2\| \sqrt{\frac{\delta}{1-\delta}} = \frac{\sqrt{\delta}}{2} \|\lambda_1 - \lambda_2\|. \quad (\text{C.3})$$

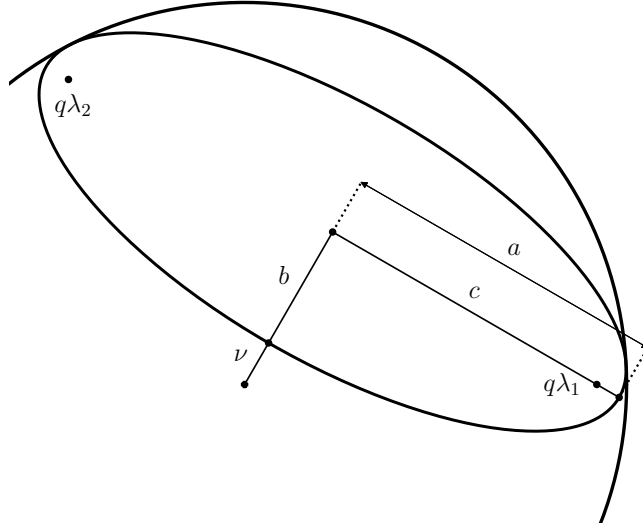


Figure C.1: Schematic illustration of an ellipse and notation used in Appendix, where we use shortcut notation  $\nu := \nu_q(U)$ .

On the other hand we have

$$\nu_q(U) + b = \left\| \frac{q\lambda_1 + q\lambda_2}{2} \right\| = \frac{q}{2} \|\lambda_1 + \lambda_2\| = \frac{\sqrt{1-\delta}}{2} \|\lambda_1 + \lambda_2\|, \quad (\text{C.4})$$

and therefore

$$\begin{aligned} \nu_q(U) &= \frac{\sqrt{1-\delta}}{2} \|\lambda_1 + \lambda_2\| - \frac{\sqrt{\delta}}{2} \|\lambda_1 - \lambda_2\| \\ &= \frac{1}{2} \left( \sqrt{1-\delta} \|\lambda_1 + \lambda_2\| - \sqrt{\delta} \|\lambda_1 - \lambda_2\| \right). \end{aligned} \quad (\text{C.5})$$

Now we need to show that the above expression for the distance  $\nu_q(U)$  is valid also for higher dimensions. The boundary of  $q$ -numerical ranges for larger matrices is described in [57]. It consists of parts of a few ellipses obtained in an analogous way. Let  $\lambda_1$  and  $\lambda_d$  be the pair of the most distant eigenvalues of  $U$ . Let  $\lambda_i$  and  $\lambda_j$  be some pair of eigenvalues such that  $i, j \neq 1, d$ . Let  $\tilde{\nu}_q(U)$  be the distance from zero to the ellipse built on  $\lambda_i$  and  $\lambda_j$  in the same way as above. Our goal is to prove that  $\tilde{\nu}_q(U) > \nu_q(U)$ .

We note that  $\|\lambda_1 - \lambda_2\| > \|\lambda_i - \lambda_j\|$ . Hence to prove that  $\tilde{\nu}_q(U) > \nu_q(U)$  it suffices to show that  $\|\lambda_1 + \lambda_2\| < \|\lambda_i + \lambda_j\|$ . As all the eigenvalues lie on the unit circle, then from the parallelogram law we have  $\|\lambda_1 + \lambda_2\|^2 = 4 - \|\lambda_1 - \lambda_2\|^2$ .

Therefore

$$\begin{aligned}\|\lambda_1 + \lambda_2\| &= \sqrt{4 - \|\lambda_1 - \lambda_2\|^2} < \sqrt{4 - \|\lambda_i - \lambda_j\|^2} \\ &= \sqrt{4 - (4 - \|\lambda_i + \lambda_j\|^2)} = \|\lambda_i + \lambda_j\|.\end{aligned}\tag{C.6}$$

and thus  $\tilde{\nu}_q(U) > \nu_q(U)$ , from which it follows that

$$\nu_{\sqrt{1-\delta}}(U) = \frac{1}{2} \left( \sqrt{1-\delta} \|\lambda_1 + \lambda_d\| - \sqrt{\delta} \|\lambda_1 - \lambda_d\| \right)\tag{C.7}$$

holds for any dimension  $d$ . The above formula can be easily translated into trigonometric functions where  $\Theta$  is the angle between  $\lambda_1$  and  $\lambda_d$ . Hence, we have

$$\nu_{\sqrt{1-\delta}}(U) = \sqrt{1-\delta} \cos\left(\frac{\Theta}{2}\right) - \sqrt{\delta} \sin\left(\frac{\Theta}{2}\right).\tag{C.8}$$

Therefore, we have

$$p_{\text{II}} = \nu_{\sqrt{1-\delta}}^2(U \otimes \mathbb{1}) = \nu_{\sqrt{1-\delta}}^2(U) = \left( \sqrt{1-\delta} \cos\left(\frac{\Theta}{2}\right) - \sqrt{\delta} \sin\left(\frac{\Theta}{2}\right) \right)^2.\tag{C.9}$$